Diversity Order of Infinite-length Symbol-by-symbol Linear Equalization

Ali Tajer∗ Aria Nosratinia† Naofal Al-Dhahir†

Abstract

The processing complexity of optimal receivers in frequency selective channels grows exponentially with the channel memory length. Linear receivers, alternatively, offer significantly less complexity with the cost of performance degradation. Such degradation can be due to not fully capturing the inherent frequency diversity or degraded coding gain. In this paper we investigate how effectively the receivers with linear equalization can capture the frequency diversity gain. The analyses reveal that minimum mean-square error (MMSE) linear equalizers incur no diversity loss compared to maximum-likelihood sequence detection (MLSD). Specifically, we show that for a channel with memory length $\nu$, MMSE linear equalizers achieve full diversity gain of $(\nu + 1)$. On the other hand, zero-forcing (ZF) linear equalizers always achieve the diversity gain 1, irrespective of channel memory length.

Index Terms: Diversity gain, frequency diversity, linear equalization, minimum mean-square error, and zero-forcing.

∗Electrical Engineering Department, Columbia University, New York, NY 10027 (email: tajer@ee.columbia.edu).
†Electrical Engineering Department, The University of Texas at Dallas, Richardson, TX 75083 (email: {aria, aldahahir}@utdallas.edu).
1 Introduction

In broadband communication systems, the coherence bandwidth of the wireless fading channel is significantly less than the transmission bandwidth. This results in inter-symbol interference (ISI) and at the same time provides frequency diversity that can be exploited at the receiver to enhance transmission reliability [1]. It is well-understood that for Rayleigh flat-fading channels, the error rate decays only linearly with SNR [1]. For frequency-selective channels, however, proper exploitation of the available frequency diversity forces the error probability to decay at a possibly higher rate and thereof can potentially achieve higher diversity gains. The diversity gain heavily depends on the detection scheme employed at the receiver.

While maximum likelihood sequence detection (MLSD) [2] achieves optimum performance over ISI-inducing channels, its complexity (as measured by the number of MLSD trellis states) grows exponentially with the transmission spectral efficiency and the channel memory length. As a low-complexity alternative, filtering-based symbol-by-symbol equalizers (both linear and decision feedback) have been widely used over the past four decades (see [3] for an excellent tutorial). Despite their long history and successful commercial deployment, the performance of symbol-by-symbol linear equalizers over wireless fading channels is not fully characterized. More specifically, it is not known whether their observed sub-optimum performance is due to their inefficiency in fully exploiting channel frequency diversity, or it is due to a degraded coding gain. Therefore it is of paramount importance to investigate the diversity gain achieved by linear equalizers.

Among past work on symbol-by-symbol linear equalization we mention the empirical work in [4] and the analysis for decision feedback equalizers (DFE) [5], where it has been shown that both MMSE and ZF symbol-by-symbol DFE equalization achieves full diversity. Unlike symbol-by-symbol equalization, block transmission techniques with linear equalizations have been more investigated. The studies on the diversity order of linear receivers for block transmission techniques include the analysis of MMSE single-carrier frequency-domain equalization (SC-FDE) in [6,7] and diversity of MIMO-OFDM with linear equalizers [8].

We analyze the diversity gain achieved by symbol-by-symbol linear equalizers. Our analysis shows that while single-carrier infinite-length symbol-by-symbol MMSE linear equalization achieves
full diversity, ZF linear equalizers cannot exploit the frequency diversity provided by frequency-
selective channels.

The remainder of the paper is organized as follows. System descriptions and the equalizer-
related designs are provided in Section 2. We start the analysis by first looking into two-tap
channels (memory length 1) in Section 3, followed by the generalization to arbitrary channel lengths
in Sections 4 and 5 for MMSE and ZF equalizers, respectively. The concluding remarks are provided
in Section 6. To enhance the flow of the paper, most of the proofs are confined in the appendices.
Finally, we remark that whenever throughput the paper we mention linear equalization, it means
infinite-length symbol-by-symbol linear equalization.

2 System Descriptions

2.1 Transmission Model

Consider an inter-symbol interference (ISI)-inducing wireless fading channel with memory length
ν, denoted by \( h = [h_0, \ldots, h_\nu] \), and denote the input and the output of the channel at time
instance \( k \) by \( x_k \) and \( y_k \), respectively. Without loss of generality, we restrict our analyses to
channel realizations with \( h_0 \neq 0 \), as any other channels with \( h_0 = 0 \), can be fit into this model by
an appropriate time-shift. Transmission is contaminated by complex Gaussian additive white noise
\( \mathcal{CN}(0, N_0) \) represented by \( z_k \). Therefore, the basedband transmission model is given by

\[
y_k = \sum_{i=0}^{\nu} h_i x_{k-i} + z_k ,
\]

where the average transmission power is \( P_0 \), i.e., \( \mathbb{E}\{|x_k|^2\} \leq P_0 \). We assume quasi-static frequency
selective channels where the channel coefficients are assumed to be independent and the channel
coefficient \( \{h_i\} \), are distributed as zero-mean complex Gaussian random variable with variance \( \lambda_i \).
Channel coefficients are represented in the phasor form as

\[
h_i = |h_i|e^{j\theta_i} , \quad \text{for} \quad i = 0, \ldots, \nu ,
\]
where \( \theta_i \in [-\pi, \pi] \). By denoting the \( D \)-transform for a given sequence \( \{x_k\} \) by \( X(D) = \sum_k x_k D^k \), the baseband model can be cast in the \( D \)-domain as

\[
Y(D) = H(D) \cdot X(D) + Z(D) .
\]  

(3)

We denote the transmission signal-to-noise ratio by \( \text{SNR} \triangleq \frac{P_0}{N_0} \). Throughout the paper, the superscript \( * \) denotes complex conjugate and we use the shorthand \( D^{-*} \) for \( (D^{-1})^* \). Also, two functions \( f(\text{SNR}) \) and \( g(\text{SNR}) \) are said to be \textit{exponentially equal}, denoted by \( f(\text{SNR}) \approx g(\text{SNR}) \), when we have

\[
\lim_{\text{SNR} \to \infty} \frac{\log f(\text{SNR})}{\log \text{SNR}} = \lim_{\text{SNR} \to \infty} \frac{\log g(\text{SNR})}{\log \text{SNR}} .
\]  

(4)

We define the operators \( \leq \) and \( \geq \) accordingly. Also, we say that the \textit{exponential order} of \( f(\text{SNR}) \) is \( d \) if \( f(\text{SNR}) \approx \text{SNR}^d \). Finally we say that two functions \( f(\text{SNR}) \) and \( g(\text{SNR}) \) are \textit{asymptotically equal}, denoted by \( f(\text{SNR}) \approx g(\text{SNR}) \) if

\[
\lim_{\text{SNR} \to \infty} \frac{f(\text{SNR})}{g(\text{SNR})} = 1 .
\]

Throughout the paper \( \log \) denotes a base 2 logarithm and all rates are in bits/sec/Hz.

\[ \text{2.2 \ Linear Equalization} \]

As the simplest form of equalizers, zero-forcing (ZF) linear equalizers ignore the noise quality and process the output symbol \( y_k \) such that the equalizer’s output is free of ISI. By taking into account the combined effect of the ISI channel and its corresponding matched-filter at the receiver, the ZF linear equalizer for this ISI channel is given by [9, 3.87]

\[
W_{zf}(D) = \frac{\| \mathbf{h} \|}{H(D) \cdot H^*(D^{-1})} ,
\]  

(5)

where \( \| \mathbf{h} \| \) is the \( \ell_2 \)-norm of \( \mathbf{h} \), i.e., \( \| \mathbf{h} \|^2 = \sum_{i=0}^\nu |h_i|^2 \) and \( H^*(D^{-1}) \) is a shorthand for \( \sum_{i=0}^\nu h_i^* D^{-i} \).

The variance of the noise seen at the output of ZF equalizer is key in determining the performance of the equalizer. The variance of this noise which is produced by a linear filter acting on the Gaussian noise is given by

\[
\sigma_{zf}^2 \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{N_0}{|H(e^{-ju})|^2} \, du .
\]  

(6)
Therefore, by defining \( \text{SNR} \triangleq \frac{P_0}{N_0} \), the decision-point signal-to-noise ratio for any channel realization \( h \) and transmission power \( P_0 \) is given by

\[
\gamma_{zf}(\text{SNR}, h) \triangleq \text{SNR} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|H(e^{-ju})|^2} du \right]^{-1}.
\] (7)

The main shortcoming of ZF is that it ignores the effect of noise, which can lead to noise enhancement. To overcome this problem, minimum-mean square error (MMSE) linear equalizers strike a balance between ISI reduction and noise enhancement. MMSE linear equalizers are designed not to cancel the interference completely but rather to minimize the combined residual ISI and noise as always outperform ZF, while having the same implementation complexity. MMSE linear equalizers in the \( D \)-domain are given by [9, 3.148]

\[
W_{\text{mmse}}(D) = \frac{\|h\|}{H(D)H^*(D^{-1}) + \text{SNR}^{-1}}.
\] (8)

The variance of the combination of the residual ISI and channel noise as seen at the output of the equalizer is given by

\[
\sigma_{\text{mmse}}^2 \triangleq \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{N_0}{|H(e^{-ju})|^2 + \text{SNR}^{-1}} du.
\] (9)

As a result, the unbiased\(^{1,2} \) decision-point signal-to-noise ratio at the receiver for the channel realization \( h \) and the signal-to-noise ratio \( \text{SNR} \) is given by

\[
\gamma_{\text{mmse}}(\text{SNR}, h) \triangleq \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\text{SNR}|H(e^{-ju})|^2 + 1} du \right]^{-1} - 1.
\] (10)

2.3 Diversity Gain

We assume that data transmission rate by the source is \( R \) bits/sec/Hz. The system is said to be in outage if the channel is faded such that it cannot sustain an arbitrarily reliable communication at the intended data rate \( R \), or equivalently, the mutual information \( I(x_k, \tilde{y}_k) \) falls below the target rate \( R \), where \( \tilde{y}_k \) is the equalizer output. The probability of such outage for an average transmission power \( P_0 \) and receiver side signal-to-noise ratio \( \gamma(\text{SNR}, h) \) is

\[
P_{\text{out}}(\text{SNR}) \triangleq P_h \left( \log \left[ 1 + \gamma(\text{SNR}, h) \right] < R \right),
\] (11)

\(^{1}\)All MMSE equalizers are biased. Removing the bias decreases the decision-point signal-to-noise ratio by 1 (in linear scale) but improves the error probability [10].

\(^{2}\)All the results provided in this paper are valid for biased receivers as well.
where the probability is taken over all channel realizations $h$. The outage probability at high transmission powers ($\text{SNR} \to \infty$) is closely related to the average pairwise error probability, that we denote by $P_{\text{err}}(\text{SNR})$. $P_{\text{err}}(\text{SNR})$ which is the probability that a transmitted codeword $c_i$ is erroneously detected in favor of another codeword $c_j$, $j \neq i$, i.e.,

$$P_{\text{err}}(\text{SNR}) \triangleq \mathbb{E}_h \left[ P(c_i \to c_j \mid h) \right].$$  \hspace{1cm} (12)

When deploying channel codings with arbitrarily long code-length, the outage and error probabilities decay at the same rate with increasing $\text{SNR}$ and have the same exponential order [11]. In other words, by deploying sufficiently large codewords we have

$$P_{\text{out}}(\text{SNR}) = P_{\text{err}}(\text{SNR}).$$  \hspace{1cm} (13)

This is intuitively justified by noting that at high $\text{SNR}$ regimes, the effect of channel noise is diminishing and the dominant source of erroneous detection is channel fading which, as mentioned above, is the same source of outage events. As a result, in our setup diversity order which is the negative of the exponential order of the average pairwise error probability $P_{\text{err}}(\text{SNR})$ can be found as

$$d = - \lim_{\text{SNR} \to \infty} \frac{\log P_{\text{out}}(\text{SNR})}{\log \text{SNR}}.$$  \hspace{1cm} (14)

3 Two-Tap Channels

We start by analyzing the diversity order achieved for two-tap channels when its output is processed via linear MMSE and ZF linear equalization. We show that MMSE equalization fully captures the underlying frequency diversity gain, i.e., $P_{\text{err}}(\text{SNR}) \doteq \text{SNR}^{-2}$, and hence incurs no diversity loss compared to ML detection. In contrast, we show that through ZF equalization the diversity gain is 1, i.e., $P_{\text{err}}(\text{SNR}) \doteq \text{SNR}^{-1}$ which is inferior to that of MMSE and ML detection. We consider a two-tap channel $h = [h_0, h_1]$ and assume that $h_0, h_1 \neq 0$.

3.1 MMSE Linear Equalization

The aim of this section is to show that a second order diversity for two-tap channels is achievable when we deploy MMSE equalizers. The underlying idea of the approach to show such achievability is
that we first find a lower bound on $\gamma_{\text{mmse}}(\text{SNR}, h)$ given in (10). The outage probability associated with such lower bound provides an upper bound on the outage probability of the actual SNR. Then we show that the exponential order of this upper bound on the outage probability is $-2$, which in turn establishes that the diversity gain is 2.

**Theorem 1** For a non-zero two-tap ISI channel with symbol-by-symbol MMSE linear equalization we have

$$P_{\text{err}}^{\text{mmse}}(\text{SNR}) \doteq \text{SNR}^{-2}.$$  

**Proof:** By recalling that $H(e^{-ju}) = |h_0|e^{j\theta_0} + |h_1|e^{j(\theta_1-u)}$, the decision-point signal-to-noise ratio $\gamma_{\text{mmse}}(\text{SNR}, h)$ is lower bounded as follows.

$$\gamma_{\text{mmse}}(\text{SNR}, h) = \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\text{SNR}\left| H(e^{-ju}) \right|^2 + 1} \, du \right]^{-1} - 1$$

$$= \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\text{SNR}\left( |h_0|^2 + |h_1|^2 + 2|h_0h_1|\cos(u + \theta_0 - \theta_1) \right) + 1} \, du \right]^{-1} - 1$$

$$= \left[ \frac{1}{\text{SNR}^2 \left( |h_0|^2 + |h_1|^2 + \text{SNR}^{-1} \right)^2 - 4\text{SNR}^2|h_0h_1|^2} \right]^{-1/2} - 1$$

$$\geq \left[ 1 + 2\text{SNR} \left( |h_0|^2 + |h_1|^2 \right) \right]^{1/2} - 1 \doteq \gamma_{\text{mmse}}^L(\text{SNR}, h),$$

where (15) holds by noting that $\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{a+b\cos(u+\theta)} \, du = \frac{1}{\sqrt{a^2-b^2}}$. By recalling (11) and invoking the lower bound in (16) we find

$$P_{\text{out}}^{\text{mmse}}(\text{SNR}) = P_h\left( \gamma_{\text{mmse}}(\text{SNR}, h) + 1 < 2^R \right)$$

$$\leq P_h\left( \gamma_{\text{mmse}}^L(\text{SNR}, h) + 1 < 2^R \right)$$

$$= P_h\left( |h_0|^2 + |h_1|^2 < \frac{2^R - 1}{\text{SNR}} \right)$$

$$\leq P_{h_0}\left( |h_0|^2 < \frac{2^R - 1}{2\text{SNR}} \right) P_{h_1}\left( |h_1|^2 < \frac{2^R - 1}{2\text{SNR}} \right)$$

$$\doteq \text{SNR}^{-2},$$

7
where (17) holds due to statistical independence of $h_0$ and $h_1$, and the fact that for any $k > 0$, the event \( \{|h_0|^2 + |h_1|^2 < \frac{k}{\text{SNR}}\} \) is a subset of the event \( \{|h_0|^2 < \frac{k}{\text{SNR}} \text{ and } |h_1|^2 < \frac{k}{\text{SNR}}\} \). The last step in (18) is derived by noting that $|h_0|^2$ and $|h_1|^2$ are exponentially distributed and then by using the exponential equality of the Taylor expansion $1 - \exp(\frac{k}{\text{SNR}}) = \frac{\text{SNR}}{k}$. The result above along with the connection between $P_{\text{err}}(\text{SNR})$ and $P_{r,\text{SNR}}$ given in (13) and also invoking the fact that the diversity gain cannot exceed 2, the desired result is established.

3.2 ZF Linear Equalization

For analyzing the diversity gain of the two-tap channel with ZF equalization, we start off by finding a closed form representation for the decision-point signal-to-noise ratio $\gamma_{zf}(\text{SNR}, h)$ given (7) and show that the diversity gain achieved for it cannot exceed 1.

**Theorem 2** For a non-zero two-tap ISI channel with symbol-by-symbol ZF linear equalization we have

$$P_{\text{err}}^{zf}(\text{SNR}) = \frac{\text{SNR}}{k}.$$

*Proof:* We find $\gamma_{zf}(\text{SNR}, h)$ as follows

$$\gamma_{zf}(\text{SNR}, h) = \text{SNR} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|H(e^{-ju})|^2} \, du \right]^{-1}$$

$$= \text{SNR} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|h_0|^2 + |h_1|^2 + 2|h_0h_1|\cos(u + \theta_0 - \theta_1)} \, du \right]^{-1/2}$$

$$= \text{SNR} \left[ \frac{1}{(|h_0|^2 + |h_1|^2)^2 - 4|h_0h_1|^2} \right]^{-1/2}$$

$$= \text{SNR} \frac{1}{|h_0|^2 - |h_1|^2}.$$

By defining $t_i \triangleq |h_i|^2$, which is exponentially distributed with parameter $\lambda_i$ and also defining $A \triangleq \frac{2^R - 1}{\text{SNR}}$, we find the following lower bound on the outage probability ($t \triangleq [t_0, t_1]$)

$$P_{\text{out}}^{zf}(\text{SNR}) = P_h \left( \gamma_{zf}(\text{SNR}, h) + 1 < 2^R \right)$$

$$= P_h \left( \frac{|h_0|^2 - |h_1|^2}{\text{SNR}} < \frac{2^R - 1}{\text{SNR}} \right)$$

$$= P_t (|t_0 - t_1| < A)$$
\[
\begin{align*}
\int_0^\infty \int_{(t_0-A)^+}^{t_0+A} \frac{1}{\lambda_1} e^{-t_1/\lambda_1} \, dt_1 \frac{1}{\lambda_0} e^{-t_0/\lambda_0} \, dt_0 \\
\geq \int_0^\infty \int_{t_0-A}^{t_0+A} \frac{1}{\lambda_1} e^{-t_1/\lambda_1} \, dt_1 \frac{1}{\lambda_0} e^{-t_0/\lambda_0} \, dt_0 \\
= \frac{1}{\lambda_0} + \frac{1}{\lambda_1} \left( e^{A/\lambda_1} - e^{-A/\lambda_1} \right) \left( e^{-A/\lambda_0} + e^{-A/\lambda_1} \right) \\
= \frac{1}{\lambda_0} \sum_{i=1}^{\infty} \left( \frac{2R-1}{\lambda_1} \right)^i \frac{\text{SNR}^{-i}}{i!} \left( e^{-\frac{2R-1}{\lambda_0} \text{SNR}} + e^{-\frac{2R-1}{\lambda_1} \text{SNR}} \right) = \text{SNR}^{-1},
\end{align*}
\]

where in (19) we have used the notation \((x)^+ = \max(x, 0)\). Equation (20) is derived by using the exponential Taylor series and (21) holds as for high SNR regimes, the dominant term in (20) will be \(\text{SNR}^{-1}\). Therefore, (21) establishes that \(P_{\text{zf}}^\text{out}(\text{SNR}) \geq \text{SNR}^{-1}\) and therefore, the exponential order of \(P_{\text{out}}^\text{zf}(\text{SNR})\) cannot be greater than 1. On the other hand, we can also find the following upper bound on \(P_{\text{out}}^\text{zf}(\text{SNR})\)

\[
P_{\text{out}}^\text{zf}(\text{SNR}) = \int_0^\infty \int_{(t_0-A)^+}^{t_0+A} \frac{1}{\lambda_1} e^{-t_1/\lambda_1} \, dt_1 \frac{1}{\lambda_0} e^{-t_0/\lambda_0} \, dt_0 \\
\leq \int_0^\infty \int_{t_0-A}^{t_0+A} \frac{1}{\lambda_1} e^{-t_1/\lambda_1} \, dt_1 \frac{1}{\lambda_0} e^{-t_0/\lambda_0} \, dt_0 \\
= \frac{1}{\lambda_0} + \frac{1}{\lambda_1} \left( e^{A/\lambda_1} - e^{-A/\lambda_1} \right) \\
\leq \text{SNR}^{-1},
\]

Equations (21) and (22) together show that \(P_{\text{zf}}^\text{out}(\text{SNR}) \leq \text{SNR}^{-1}\). Further invoking (13) provides that \(P_{\text{err}}(\text{SNR}) \leq \text{SNR}^{-1}\), which confirms that the diversity gain achievable by ZF equalization is 1.

Different diversity gains achieved via MMSE and ZF liner equalization can be intuitively justified by considering that ZF equalization fully eliminates the interference terms, and as a result, the receiver observes only one version of each transmitted symbol. In contrast, MMSE equalization allows each received symbol be a linear combination of the last \(\nu\) transmitted symbols. Therefore, the receiver has access to multiple versions of each transmitted symbol.
4 Diversity Order of MMSE Linear Equalization

In this section, we generalize the result of two-tap channels with MMSE linear equalization to the channels with arbitrary memory lengths $\nu \geq 1$. The main result of this paper is provided in the following theorem.

**Theorem 3** For an ISI channel with channel memory length $\nu \geq 1$, and symbol-by-symbol MMSE linear equalization we have

$$P_{\text{err}}^{\text{mmse}}(\text{SNR}) = \text{SNR}^{-(\nu+1)}.$$ 

The scheme of the proof is as follows. First, we find a lower bound on the unbiased decision-point SNR. Then, we use this lower bound to show that for small enough data rates $R$, full diversity of $(\nu + 1)$ is achievable. Next, we show that increasing data rate $R$ to any arbitrary level does not incur a diversity loss, concluding that MMSE linear equalization are capable of fully collecting the frequency diversity gain of ISI channels.

4.1 Full Diversity for Low Rates

In this section we show that for arbitrarily small data transmission rates, $R$, full diversity is achievable. The following lemma is key in finding a lower bound on $\gamma_{\text{mmse}}(\text{SNR}, h)$.

**Lemma 1** For any channel realization $h$ with memory length $\nu$, ($h_0, h_\nu \neq 0$), there exists a non-zero interval $\mathcal{D}(h) \triangleq [\alpha(h), \beta(h)] \subseteq [-\pi, \pi]$ over which the function $f(h, u) \triangleq |H(e^{-ju})|^2 - \|h\|^2$ is strictly positive.

**Proof:** See Appendix A. ■

Now, for any channel realization $h$, we denote the interval over which $f(h, u)$ is strictly positive (provided by Lemma 1) by $\mathcal{D}(h) = \{ u \in [\alpha(h), \beta(h)] \}$. As follows, for any channel realization $h$ we establish a lower bound on $\gamma_{\text{mmse}}(\text{SNR}, h)$, given in (10) that depends on the interval $\mathcal{D}(h)$, as well as the channel realization $\|h\|$. From (16) and noting that $f(h, u) = |H(e^{-ju})|^2 - \|h\|^2$ we find that

$$\gamma_{\text{mmse}}(\text{SNR}, h)$$

10
\[
\begin{align*}
&= \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\text{SNR}(f(h, u) + \|h\|^2) + 1} \, du \right]^{-1} - 1 \\
&= \left[ \frac{1}{2\pi} \int_{D(h)} \frac{1}{\text{SNR}|H(e^{jw})|^2 + 1} \, du + \frac{1}{2\pi} \int_{D(h)} \frac{1}{\text{SNR}(f(h, u) + \|h\|^2) + 1} \, du \right]^{-1} - 1 \\
&\geq \left[ \frac{1}{2\pi} \int_{D(h)} 1 \, du + \frac{1}{2\pi} \int_{D(h)} \frac{1}{\text{SNR}\|h\|^2 + 1} \, du \right]^{-1} - 1 \\
&= \left[ \left( 1 - \frac{\beta(h) - \alpha(h)}{2\pi} \right) + \frac{\beta(h) - \alpha(h)}{2\pi} \cdot \frac{1}{\text{SNR}\|h\|^2 + 1} \right]^{-1} - 1.
\end{align*}
\] (23)

By using the above lower bound on the decision-point signal-to-noise ratio and using the outage probability provided in (11) we have

\[
P_{\text{mmse}}^{\text{out}}(\text{SNR}) = P_h \left( \gamma_{\text{mmse}}(\text{SNR}, h) + 1 < 2^R \right) \leq P_h \left\{ \left( 1 - \frac{\beta(h) - \alpha(h)}{2\pi} \right) + \frac{\beta(h) - \alpha(h)}{2\pi} \cdot \frac{1}{\text{SNR}\|h\|^2 + 1} > 2^{-R} \right\}. \] (24)

By some simple manipulations it can be readily verified the if the data rate \( R \) for all channel realizations \( h \) satisfies the constraint

\[
\forall h, \quad 2^{-R} > \left( 1 - \frac{\beta(h) - \alpha(h)}{2\pi} \right),
\] (26)

or equivalently

\[
2^{-R} > \max_h \left( 1 - \frac{\beta(h) - \alpha(h)}{2\pi} \right) = \left( 1 - \min_h \{ \beta(h) - \alpha(h) \} \right), \] (27)

then the probability term in (25) can be re-written as

\[
P_h \left\{ \text{SNR}\|h\|^2 < \frac{1 - 2^{-R}}{2^{-R} - (1 - \frac{\beta(h) - \alpha(h)}{2\pi})} \right\}. \] (28)

Now, we define

\[
R_{\text{max}} \triangleq - \log \left( 1 - \frac{\min_h \{ \beta(h) - \alpha(h) \}}{2\pi} \right), \] (29)

Note that according to Lemma 1, \( D(h) = [\alpha(h), \beta(h)] \) is a non-zero interval, indicating that \( \beta(h) > \alpha(h) \), or \( \min_h \{ \beta(h) - \alpha(h) \} > 0 \). As a result, \( R_{\text{max}} \) is guaranteed to bonded away from zero, i.e., \( R_{\text{max}} > 0 \). Therefore, based on equations (25)-(29) we have for all \( 0 < R < R_{\text{max}} \) we have

\[
P_{\text{mmse}}^{\text{out}}(\text{SNR}) \leq P_h \left\{ \text{SNR}\|h\|^2 < \frac{1 - 2^{-R}}{2^{-R} - (1 - \frac{\beta(h) - \alpha(h)}{2\pi})} \right\}. \] (30)

11
\[
\leq P_h \left\{ \text{SNR} \| h \|^2 < \max_h \left\{ \frac{1 - 2^{-R}}{2^{-R} - (1 - \frac{\beta(h) - \alpha(h)}{2\pi})} \right\} \right\} \\
= P_h \left\{ \text{SNR} \| h \|^2 < \frac{1 - 2^{-R}}{2^{-R} - (1 - \min_h \{\beta(h) - \alpha(h)\})} \right\} \\
= P_h \left\{ \text{SNR} \| h \|^2 < \frac{2^R - 1}{1 - 2^{-R - R_{max}}} \right\} .
\]

The transition from (30) to (31) holds by noting that

\[
\max_h \left\{ \frac{1 - 2^{-R}}{2^{-R} - (1 - \frac{\beta(h) - \alpha(h)}{2\pi})} \right\} \geq \frac{1 - 2^{-R}}{2^{-R} - (1 - \frac{\beta(h) - \alpha(h)}{2\pi})} .
\]

Next, by defining \( A \triangleq \frac{2^R - 1}{1 - 2^{-R - R_{max}}} \) and \( t_i \triangleq |h_i|^2 \) for \( i = 0, \ldots, \nu \), we further get (\( t \triangleq [t_1, \ldots, t_\nu] \)).

\[
P_{\text{out}}^{\text{mmse}}(\text{SNR}) \leq P_h \left\{ \sum_{m=0}^\nu |h_m|^2 < \frac{A}{\text{SNR}} \right\} \\
= P_t \left\{ \sum_{m=0}^\nu t_m < \frac{A}{\text{SNR}} \right\} \\
\leq \prod_{m=0}^\nu P_t \left\{ t_m < \frac{A}{\text{SNR}} \right\} \\
= \prod_{m=0}^\nu \left[ 1 - \exp \left( -\frac{A}{2\text{SNR}\lambda_m} \right) \right] \\
= \text{SNR}^{-(\nu+1)} ,
\]

where the inequality (32) holds due to statistical independence of \( h_i \) (and consequently \( t_i \)) and taking into account that

\[
\left\{ t_m : \sum_{m=0}^\nu t_m < \frac{A}{\text{SNR}} \right\} \subset \left\{ t_m : t_m < \frac{A}{\text{SNR}} \text{ for all } m \right\} .
\]

Therefore, from (25) and (33) it is concluded that for the rates lower than \( R_{max} \) as defined in (29) \( P_{\text{out}}^{\text{mmse}}(\text{SNR}) \leq \text{SNR}^{-(\nu+1)} \), which in conjunction with (13) proves that \( P_{\text{err}}^{\text{mmse}} \leq \text{SNR}^{-(\nu+1)} \).

Therefore, a diversity order of at least \((\nu + 1)\) is achievable. On the other hand, since the diversity order cannot exceed the number of channel taps, the achievable diversity order is exactly \((\nu + 1)\).

This establishes the proof of Theorem 3 for the range of the rates below \( R_{max} \) defined in (29).
4.2 Full Diversity for All Rates

In this section we intend to prove that increasing the rate beyond $R_{\text{max}}$ will incur no diversity loss. We use the two subsequent lemmas to show that the diversity gain achievable at the rates higher than $R_{\text{max}}$ is equal to those of smaller $R_{\text{max}}$.

**Lemma 2** For asymptotically large values of SNR, the decision-point signal-to-noise ratio $\gamma_{\text{mmse}}(\text{SNR}, h)$ will vary linearly with SNR, i.e.,

$$\lim_{\text{SNR} \to \infty} \frac{\partial \gamma_{\text{mmse}}(\text{SNR}, h)}{\partial \text{SNR}} = s(h), \quad \text{where } s(h): \mathbb{R}^{\nu+1} \to \mathbb{R}.$$  

*Proof:* See Appendix B.

**Lemma 3** For the continuous random variable $X$, variable $y \in \mathbb{R}$, constants $c_1, c_2 \in \mathbb{R}$ and function $G(X, y)$ continuous in $y$, we have

$$\lim_{y \to y_0} P_X\left(c_1 \leq G(X, y) \leq c_2 \right) = P_X\left(c_1 \leq \lim_{y \to y_0} G(X, y) \leq c_2 \right).$$

*Proof:* See Appendix C.

Now, we show that if for some rate $R^\dagger$ the achievable diversity order is $d$, then for all rates up to $R^\dagger + 1$, the same diversity order is achievable. By induction, we conclude that the diversity order remains unchanged by changing the data rate $R$. If for the rate $R^\dagger$, the negative of the exponential order of the outage probability is $d$, i.e.,

$$P_h\left(\log \left[1 + \gamma_{\text{mmse}}(\text{SNR}, h)\right] < R^\dagger\right) \simeq \text{SNR}^{-d}, \quad (34)$$

then by applying the results of lemmas 2 and 3 for the target rate $R^\dagger + 1$ we get

$$P_{\text{out}}^{\text{mmse}}(\text{SNR}) = P_h\left(\log \left[1 + \gamma_{\text{mmse}}(\text{SNR}, h)\right] < R^\dagger + 1\right)$$

$$= P_h\left(\gamma_{\text{mmse}}(\text{SNR}, h) + 1 < 2^{R^\dagger + 1}\right)$$

$$\simeq P_h\left(\text{SNR} \cdot s(h) < 2^{R^\dagger + 1}\right)$$

$$= P_h\left(\left(\frac{\text{SNR}}{2}\right) \cdot s(h) < 2^{R^\dagger}\right)$$

$$\simeq P_h\left(\gamma_{\text{mmse}}\left(\frac{\text{SNR}}{2}, h\right) + 1 < 2^{R^\dagger}\right) \quad (35)$$

13
Figure 1: Average detection error probability in two-tap and three-tap channels with MMSE linear equalization.

\[
\frac{1}{R} \left( \log \left[ 1 + \gamma_{\text{MMSE}} \left( \frac{\text{SNR}}{2}, h \right) \right] < R^\dagger \right)
\]

\[
\frac{1}{R} \left( \frac{\text{SNR}}{2} \right)^{-d}
\]

\[
\text{SNR}^{-d}
\]

Equations (35) and (36) are the immediate results of lemmas 2 and 3 which enable interchanging probability and limit and also permit replacing \( \gamma_{\text{MMSE}}(\text{SNR}, h) = \text{SNR} \cdot s(h) \) in the high SNR regimes.

Now, based on equations (34)-(37) it is concluded that the diversity orders achieved for rates up to \( R^\dagger \) and the rates up to \( R^\dagger + 1 \) are identical. Therefore, based on the results of Section 3.1 by induction we find that for any arbitrary rate \( R \), full diversity is achievable via MMSE linear equalization. This completes the proof of Theorem 3.

Figure 1 depicts the simulation results on the pairwise error probabilities for two different ISI channels with memory lengths \( \nu = 1 \) and 2 and MMSE equalizer. For each of these channels we consider signal transmissions with rates \( R = (1, 2, 3, 4) \) bits/sec/Hz. The simulation results confirm that for a channel with two taps the achievable diversity gain is two irrespective of the data rate.
Similarly it is observed that for a three-tap channel the achievable diversity gain is three.

5 Diversity Order of ZF linear Equalization

As briefly discussed earlier, ZF linear equalizers fully suppress all the inter-symbol interference terms. The major shortcoming of this type of filtering is that it boosts the noise terms at the channel nulls. Another deficiency of such filtering is that it refrains from delivering multiple versions of each transmitted symbol at the receiver. Such observation has also been made in other contexts, including pre-coded OFDM transmission [12] and MIMO-OFDM systems with linear equalization [8]. In this section, we show that the diversity order achieved by zero-forcing symbol-by-symbol linear equalization is always one and does not depend on the channel memory length.

Lemma 4 For any arbitrary set of normal complex Gaussian random variables \( \mathbf{\mu} \triangleq (\mu_1, \ldots, \mu_m) \) (possibly correlated) and for any \( B \in \mathbb{R}^+ \) we have

\[
P_{\mathbf{\mu}} \left( \sum_{k=1}^{m} \frac{1}{SNR|\mu_k|^2} > B \right) \geq SNR^{-1}.
\]

Proof: Define

\[
W_k \triangleq -\frac{\log |\mu_k|^2}{\log SNR}.
\]

Since \( |\mu_k|^2 \) has exponential distribution, it can be shown that for any \( k \) the probability density function (PDF) at the asymptote of high values of \( SNR \) is given by [13]

\[
p_{W_k}(w) = (\ln SNR) \exp(-w \log SNR),
\]

which indicates that when \( SNR \to \infty \), \( W_k \) has exponential distribution with parameter \( \frac{1}{\ln SNR} \).

Therefore, by substituting \( |\mu_k|^2 \triangleq SNR^{1-W_k} \), we find

\[
P_{\mathbf{\mu}} \left( \sum_{k=1}^{m} \frac{1}{SNR|\mu_k|^2} > B \right) \overset{\triangle}{=} P_W \left( \sum_{k=1}^{m} \frac{1}{SNR^{1-W_k}} > B \right)
\]

\[
= P_W \left( \sum_{k=1}^{m} SNR^{W_k-1} > B \right)
\]

\[
= P_W(\max_k W_k > 1)
\]
\[ P_{W_k}(W_k > 1) \]
\[ = \exp(-\ln \text{SNR}) = \text{SNR}^{-1}. \]

Equation (41) holds as the term \( \text{SNR}^{\max_{W_k}} - 1 \) is the dominant term in the summation \( \sum_{k=1}^{m} \text{SNR}^{W_k} - 1. \)

The transition from (41) to (42) is justified by noting that \( \max_{k} W_k \geq W_k \) and the last step is derived by noting that \( W_k \) has exponential distribution with parameter \( \frac{1}{\ln \text{SNR}}. \)

**Remark 1** For any \( u^\dagger \in [-\pi, \pi] \)
\[ H(e^{-ju}) \bigg|_{u=u^\dagger} = \sum_{k=0}^{\nu} h_k e^{-jk\pi}, \]
is a linear combination of complex Gaussian random variables \( h_k \) and therefore has Gaussian distribution.

**Theorem 4** The diversity order achieved by symbol-by-symbol zero-forcing linear equalization is always one, i.e.,
\[ P_{\text{zf}}^{\text{out}} = \text{SNR}^{-1}. \]

**Proof:** Using the decision-point signal-to-noise ratio of zero-forcing given in (7) we have
\[ P_{\text{ZF}}^{\text{out}}(\text{SNR}) = P_h \left( \gamma_{\text{zf}}(\text{SNR}, h) < 2^R - 1 \right) \]
\[ = P_h \left\{ \frac{1}{2\pi} \int_\pi^\pi \frac{1}{\text{SNR}|H(e^{-ju})|^2} \, du \right\}^{-1} < 2^R - 1 \}
\[ = \lim_{\Delta \to 0} \left\{ \frac{2\pi/\Delta}{\text{SNR}|H(e^{-j(-\pi+\Delta)})|^2} \right\}^{-1} < \frac{2^R - 1}{2\pi} \}
\[ = \lim_{\Delta \to 0} \left\{ \frac{2\pi/\Delta}{\text{SNR}|H(e^{-j(-\pi+\Delta)})|^2} \right\}^{-1} < \frac{2^R - 1}{2\pi} \}
\[ \geq \text{SNR}^{-1}. \]
Figure 2: Outage probability of a two-tap channel with ZF linear equalization

\[ H(e^{-j(-\pi+k\Delta)}) \] which as mentioned in Remark 1 has Gaussian distribution. Note that although \( \{\mu_k\} \) are not necessarily statistically uncorrelated, still (39) holds valid as Lemma 4 does not necessitate them to be uncorrelated. Therefore, the achievable diversity order is 1.

Figure 2 illustrates the simulation results on the pairwise error probability of communication over two ISI channels with memory lengths \( \nu = 1 \) and 2. The simulation results establish that the achievable diversity gain is constantly one, independently of the channel memory length or communication data rates.

6 Conclusion

We showed that infinite-length symbol-by-symbol MMSE linear equalization can fully capture the underlying frequency diversity of the ISI channel. Specifically the diversity order achieved is equal to that of maximum likelihood sequence detection (MLSD) which show that in the high-SNR regime, the performance of the MMSE linear equalization and MLSD only differ in coding gain. We also show that the diversity order achieved by zero-forcing linear equalizers is always one, regardless of
By using (47) and defining
\[ f(h,u) = H(e^{-ju})H^*(e^{-ju}) - \|h\|^2 = \left[ \sum_{m=0}^{\nu} |h_m|e^{j\theta_m}e^{-jmu} \right] \left[ \sum_{n=0}^{\nu} |h_n|e^{-j\theta_n}e^{jnu} \right] - \|h\|^2 \]

\[ = \sum_{m=0}^{\nu} \sum_{n=m+1}^{\nu} 2|h_mh_n| \cos ((n-m)u - (\theta_n - \theta_m)) \]

\[ = \sum_{m=1}^{\nu} \left( a_m \cos (mu) + b_m \sin (mu) \right) \text{,} \tag{47} \]

where \( a_i \) and \( b_i \) are real constants that depend on \( \{h_m\} \) and \( \{\theta_m\} \) and collect relevant terms.

Now, by utilizing the trigonometric identities \( \sin(mu) = (\sin u) \sum_{i=0}^{m-1} p_i \cos^i(u) \) and \( \cos(mu) = \sum_{i=0}^{m} q_i \cos^i(u) \) where \( p_i \) and \( q_i \) are real-valued constants, we get \( f(h,u) = \sum_{m=1}^{\nu} p'_m \cos^m(u) + (\sin u) \sum_{m=0}^{\nu-1} q'_m \cos^m(u) \), where \( p'_m \) and \( q'_m \) collect all relevant terms. Therefore, \( f(h,u) \) can be transformed to a polynomial equation where it can be readily verified that cannot have more than \( 4\nu \) solutions. Now we show that \( f(h,u) = 0 \) has at least one solution. For this purpose, for each \( k = 0, \ldots, \lfloor \log_2 \nu \rfloor \) we define the set \( A_k \) which consists of the natural numbers not greater than \( \nu \) that are dividable by \( 2^k \) but not \( 2^{k+1} \), i.e.,

\[ A_k \triangleq \{ m \mid m = (2t+1)2^k, 1 \leq m \leq \nu, t \in \mathbb{N} \} \text{ for } k = 0, \ldots, \lfloor \log_2 \nu \rfloor \text{.} \tag{48} \]

By using (47) and defining \( N \triangleq \lfloor \log_2 \nu \rfloor \) we get \( \forall u \in [-\pi, \pi] \)

\[ \sum_{n=0}^{2N+1-1} f(h,u+n\frac{\pi}{2^N}) = \sum_{n=0}^{2N+1-1} a_m \cos \left( mu + \frac{mn\pi}{2^N} \right) + b_m \sin \left( mu + \frac{mn\pi}{2^N} \right) \]

\[ = \sum_{k=0}^{N} \sum_{m \in A_k} 2^{k-1} \sum_{l=0}^{2^{N-k}+1} a_m \cos \left( mu + \frac{m(n+2^{N-k})\pi}{2^N} \right) + a_m \cos \left( mu + \frac{m(n+2^{N-k})\pi}{2^N} \right) \]

\[ + \sum_{k=0}^{N} \sum_{m \in A_k} 2^{k-1} \sum_{l=0}^{2^{N-k}+1} b_m \sin \left( mu + \frac{m(n+2^{N-k})\pi}{2^N} \right) + b_m \sin \left( mu + \frac{m(n+2^{N-k})\pi}{2^N} \right) = 0 \text{.} \]

\[ A \quad \text{Proof of Lemma 1} \]

If we show that \( f(h,u) \) has at least one zero and also the number of its zeros is finite, then by continuity of \( f(h,u) \) in \( u \), the lemma is concluded. We first show that \( f(h,u) \) has finite number of zeros.
Now since we showed that \( f(h, u) \) has a finite number of zeros, there exists \( \hat{u} \in [-\pi, -\frac{\pi}{2N}] \) such that \( f(h, \hat{u}) \neq 0 \). Therefore, there exists \( \hat{n} \) such that \( 0 \leq \hat{n} \leq 2N + 1 \) and the term \( f(h, \hat{u} + \frac{\hat{n}\pi}{2N}) \) should be greater than zero since otherwise the terms inside the summation in the LHS of (49) cannot sum up to zero. Recalling the range of \( \hat{u} \) we have \( -\pi \leq \hat{u} + \frac{\hat{n}\pi}{2N} \leq \pi \). Therefore in the interval \([-\pi, \pi]\) there exists a \( u^* = \hat{u} + \frac{\hat{m}\pi}{2N} \) for which \( f(h, u^*) > 0 \), and since for any channel realization \( h \), the function \( f(h, u) \) is continuous in \( u \), there is a neighborhood around \( u^* \) where \( f(h, u) \) is strictly positive. For each channel realization \( h \), we denote this neighborhood by \([\alpha(h), \beta(h)]\).

**B Proof of Lemma 2**

We define \( g(h, u) \triangleq |H(e^{-ju})|^2 \) where we can show that it has finite number of zero by following the same line of argument as for \( f(h, u) \) in the proof of Lemma 1. By using (10) we get

\[
\frac{\partial \gamma_{\text{mmse}}(\text{SNR}, h)}{\partial \text{SNR}} = \frac{\partial}{\partial \text{SNR}} \left( \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\text{SNR}g(h, u) + 1} \, du \right]^{-1} - 1 \right) = \left[ \frac{1}{2\pi} \int \frac{g(h, u)}{\left(\text{SNR}g(h, u) + 1\right)^2} \, du \right] \cdot \left[ \frac{1}{2\pi} \int \frac{1}{\text{SNR}g(h, u) + 1} \, du \right]^{-2} \tag{50}
\]

\[
= \left[ \frac{1}{2\pi} \int_{g(h, u) \neq 0} \frac{g(h, u)}{\left(\text{SNR}g(h, u) + 1\right)^2} \, du \right] \cdot \left[ \frac{1}{2\pi} \int_{g(h, u) \neq 0} \frac{1}{\text{SNR}g(h, u) + 1} \, du \right]^{-2} \tag{51}
\]

where (51) was obtained by removing a finite-number of points from the integral in (50).

**Theorem 5** Monotone Convergence \([14, \text{Thrm. 4.6}]\): if a function \( F(u, v) \) defined on \( U \times [a, b] \to \mathbb{R} \), is positive and monotonically increasing in \( v \), and there exists an integrable function \( \hat{F}(u) \), such that \( \lim_{v \to \infty} F(u, v) = \hat{F}(u) \), then

\[
\lim_{v \to \infty} \int_U F(u, v) \, du = \int_U \lim_{v \to \infty} F(u, v) \, du = \int_U \hat{F}(u) \, du . \tag{52}
\]

For further simplifying (51), we define \( F_1(u, \text{SNR}) \) and \( F_2(u, \text{SNR}) \) over,

\[
\left\{ u \mid u \in [-\pi, \pi], g(h, u) \neq 0 \right\} \times [1, +\infty] .
\]

as

\[
F_1(u, \text{SNR}) \triangleq \frac{1}{g(h, u)} - \frac{1}{\text{SNR}^2 g(h, u)} + \frac{g(h, u)}{(\text{SNR}g(h, u) + 1)^2} ,
\]

19
and $F_2(u, \text{SNR}) \triangleq \frac{1}{g(h, u)} - \frac{1}{\text{SNR}g(h, u)} + \frac{1}{\text{SNR}g(h, u) + 1}$.

It can be readily verified that $F_i(u, \text{SNR}) > 0$ and $F_i(u, \text{SNR})$ is increasing in SNR. Moreover, there exist $\hat{F}(u)$ such that

$$
\hat{F}(u) = \lim_{\text{SNR} \to \infty} F_1(u, \text{SNR}) = \lim_{\text{SNR} \to \infty} F_2(u, \text{SNR}) = \frac{1}{g(h, u)}.
$$

Therefore, by exploiting the result of Theorem 5 we find

$$
\lim_{\text{SNR} \to \infty} \int \left[ \frac{1}{g(h, u)} - \frac{1}{\text{SNR}^2 g(h, u)} + \frac{g(h, u)}{(\text{SNR}g(h, u) + 1)^2} \right] du = \int \frac{1}{g(h, u)} du,
$$

and

$$
\lim_{\text{SNR} \to \infty} \int \left[ \frac{1}{g(h, u)} - \frac{1}{\text{SNR}g(h, u)} + \frac{1}{\text{SNR}g(h, u) + 1} \right] du = \int \frac{1}{g(h, u)} du,
$$

or equivalently,

$$
\lim_{\text{SNR} \to \infty} \frac{1}{2\pi} \int \frac{g(h, u)}{\text{SNR}g(h, u) + 1} du = \lim_{\text{SNR} \to \infty} \frac{1}{2\pi} \int \frac{du}{\text{SNR}^2 g(h, u)}, \quad (53)
$$

and

$$
\lim_{\text{SNR} \to \infty} \frac{1}{2\pi} \int \frac{du}{\text{SNR}g(h, u) + 1} = \lim_{\text{SNR} \to \infty} \frac{1}{2\pi} \int \frac{du}{\text{SNR}g(h, u)} \quad (54)
$$

By using the equalities in (53)-(54) and proper replacement in (51) we get

$$
\frac{\partial \gamma_{\text{mmse}}(\text{SNR}, h)}{\partial \text{SNR}} = \left[ \frac{1}{2\pi} \int_{g(h, u) \neq 0} \frac{g(h, u)}{\text{SNR}g(h, u) + 1} du \right] \cdot \left[ \frac{1}{2\pi} \int_{g(h, u) \neq 0} \frac{1}{\text{SNR}g(h, u) + 1} du \right]^{-2}
$$

$$
= \left[ \frac{1}{2\pi} \int_{g(h, u) \neq 0} \frac{1}{\text{SNR}^2 g(h, u)} du \right] \cdot \left[ \frac{1}{2\pi} \int_{g(h, u) \neq 0} \frac{1}{\text{SNR}g(h, u)} du \right]^{-2}
$$

$$
= \left[ \frac{1}{2\pi} \int_{g(h, u) \neq 0} \frac{1}{g(h, u)} du \right]^{-1} = s(h),
$$

where $s(h)$ is independent of SNR and thus the proof is completed.

### C Proof of Lemma 3

For a given value of $y$ we define random variable $U_y \triangleq G(X, y)$ and denote its probability density function (PDF) by $p(u, y)$. We also define the indicator function

$$
1_A : U_y \to \{0, 1\}, \quad \text{and} \quad 1_A(u) = \begin{cases} 
1, & \text{if } u \in A \\
0, & \text{if } u \notin A
\end{cases}
$$
Theorem 6 Lebesgue’s Dominated Convergence: [14] Limit and integral are interchangeable if the integrand is dominated by an integrable function, i.e., for the function $F(u, v)$, if there exist $\hat{F}(u, v)$ such that $|F(u, v)| \leq \hat{F}(u, v)$ and $\hat{F}(u, v)$ is integrable for all $v$, then

$$\lim_{v \to v_0} \int F(u, v)du = \int \lim_{v \to v_0} F(u, v)du .$$

By defining $F(u, y) \triangleq 1_{c_1 \leq u \leq c_2}p(u, y)$ and $\hat{F}(u, y) \triangleq p(u, y)$ it is easily seen that $|F(u, v)| \leq \hat{F}(u, v)$ and $\int \hat{F}(u, v) = 1$. Hence, by applying the theorem above we get

$$\lim_{y \to y_0} P_X(c_1 \leq G(X, y) \leq c_2) = \lim_{y \to y_0} \int_{c_1}^{c_2} p(u, y) du$$

$$= \lim_{y \to y_0} \int_{-\infty}^{+\infty} 1_{c_1 \leq u \leq c_2} p(u, y) du$$

$$= \int_{-\infty}^{+\infty} \lim_{y \to y_0} 1_{c_1 \leq u \leq c_2} p(u, y) du$$

$$= \int_{c_1}^{c_2} \lim_{y \to y_0} p(u, y) du$$

$$= P_X(c_1 \leq \lim_{y \to y_0} G(X, y) \leq c_2) ,$$

where in transition from (55) to (56) we had used the result of Lebesgue’s dominated convergence theorem.

References


