RESEARCH ARTICLE

New approach to equalizing antenna elements: analysis and performance evaluation

Pierre-Richard Cornely

Eastern Nazarene College

Abstract:The antenna face of a phased array radar typically consists of several hundred of antenna elements, and they degrade independently. This poses a challenging problem to radar target detection, discrimination, and classification, which rely on adaptive beamforming and assume that the channels are matched to each other. In this research, a channel equalization algorithm is developed compensating for the mismatch between the reference and testing channels using the least-squares error (LSE) criterion. The equalized output is precisely the projection of the reference channel data onto the columns of the equalization matrix, which is solely a function of the testing channel output. Through the analysis of the equalization matrix, the performance metrics including the squares error, instantaneous correlation coefficient, and cancellation ratio (CR) of the proposed equalizer are expressed in closed forms. The analysis also allows us to postulate on the effect of system parameters including: window size, equalizer length, and input signal-to-noise ratio (SNR) on the performance metrics. Extensive Monte Carlo simulations show that higher values of the equalizer length, input SNR, or window size improves the CR; however, once a system parameter approaches a certain threshold, further incrementing the size of these parameters has a diminishing return on system performance. Simulations also reveal that an equalizer with good CR or correlation coefficient results into the equalized testing channel output being almost a replica of the reference channel's output. Correspondingly, degradation in the CR or correlation coefficient affects the equalized testing channel output. The simulation results agree closely with known theoretical analyses. The research in this paper demonstrates the importance of channel equalization and system parameter selection in obtaining a satisfactory antenna elements/subarrays output.

Keywords:channel equalization; least-squares error criterion; singular value decomposition; matrix eigen-analysis; subspace projection; correlation coefficient; jamming cancellation ratio; complex Gaussian noise; adaptive beam-forming.

1. Introduction

The antenna array in a pulse-Doppler radar is comprised of a multitude of sensor elements. Incoming data sensed by these elements determine whether a target is present or not[1–2], and if it is present, its dynamic behavior and physical attributes. For self-protection, the target dynamics sometimes includes electronic counter-measures to jam the radar. Electronic counter-measures have required that modern radar deploy adaptive digital beamforming (ABF) techniques to electronically steer the sensor elements away from the direction of jamming[3–7]. The effectiveness of ABF hinges on sensor elements with identical temporal/spectral characteristics. illustrates two antenna elements/subarrays that receive a broadband intermediate frequency (IF) signal from the far sight. Note that for the purpose of simplicity, the hardware devices used for frequency down conversion of the incoming radio frequency (RF) signal and phase compensation for the path difference between the source and the antenna elements/subarrays preceding are not shown. The IF signal is corrupted by IF noise and jamming on the way to the rest of the antenna elements, which are lumped

doi: 10.18063/wct.v2i1.629

together as a channel. Typically, a channel is made of a cascade of cables and an analog Band-Pass Filter (BPF); thelatter entails a lattice of inductors and capacitors actively/passively accepting signals within a certain frequency spectrum and suppressing out-of-band signals. The resulting signal emerging from the channel is sampled by a high speed Analog-to-Digital Converter (ADC) whose output yields a set of complex data. Due to cable mismatch and components of the BPFs, which are vulnerable to natural and operational wear, the spectral characteristics of the individual channels are hardly the same. Without loss of generality, we assume that the channel at the bottom of Figure 1, called the reference channel, encounters miniscule degradation. On the other hand, the channel at the top, called the testing channel, contains all the discrepancies of the channel characteristics as compared to the reference channel. Correspondingly, the outputs from the digital demodulators (BPF plus ADC) of the testing and reference channels, known as the testing $\{x_i \mid i = 1, 2, \cdots\}$, and reference data $\{x_{\text{ref};i} \mid i = 1, 2, \cdots\}$ respectively, do not have identical characteristics and as such, the ABF technique cannot be executed efficiently. To overcome this difficulty, the testing channel is equalized, which is the main thrust of the research in this paper. The paper is organized as follows. Section 2 provides a mathematical formulation of antenna elements/subarrays channel equalization, and the derivation of a channel equalization algorithm. Section 3 is dedicated to the analysis of the equalizer components, and an introduction to performance metrics. Section 4 describes simulation results and the effects of system parameters on the performance metrics. Further remarks and conclusions are provided in Section 5.

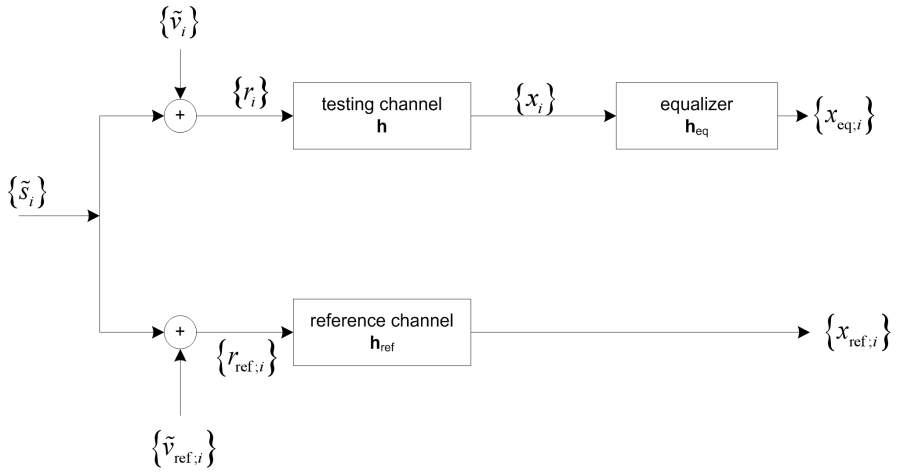

Figure 1; Schematic of a channel equalizer.

2. Channel Equalization Algorithm

2.1 Preliminaries

A sequence of n_s complex information waveforms, $\{\tilde{s}_i\}_{i=1}^{n_s}$, together with noise plus jamming is injected into a time invariant, and mismatched testing channel with Finite Impulse Response (FIR) function $\{h_i\}_{i=1}^{n_h}$. n_h is the length of support of the channel, and h_i is a discrete-time sample of the "composite" analog channel impulse response function $h(t)$ lumped with the pulse shaping filter, the channel, the receiving filters, mismatches in cables, and imperfection in frequency conversion data sampled at the sampling interval T_s: $h_i = h\big((i-1)T_s\big)$.

The complex output signal x_i at the testing channel is the convolution of the noise-corrupted information symbols and the system FIR

$$x_i = \sum_{k=-\infty}^{\infty} r_{i-k+1}\, h_k \qquad (2.1)$$

$$r_i = \tilde{s}_i + \tilde{v}_i, \qquad i = 1, 2, \cdots, n_s \tag{2.2}$$

where $\tilde{v}_i$ is an additive, wide-sense stationary channel plus jamming noise that is complex, normally distributed with mean zero and variance $\sigma_{\tilde{v}}^2$: $\tilde{v}_i \sim \mathrm{CN}\left(0, \sigma_{\tilde{v}}^2\right)$, $0.5E\left\{\tilde{v}_i^* \tilde{v}_i\right\} = \sigma_{\tilde{v}}^2 \delta_{ik}$. The Kronecker delta δ_{ik} equals one for $i = k$, and zero otherwise; and the asterisk denotes complex conjugation.

The same information sequence is injected into the reference channel to yield the output $x_{\mathrm{ref};i}$

$$x_{\mathrm{ref};i} = \sum_{k=-\infty}^{\infty} r_{\mathrm{ref};i-k+1}\, h_{\mathrm{ref};k} \tag{2.3}$$

$$r_{\mathrm{ref};i} = \tilde{s}_i + \tilde{v}_{\mathrm{ref};i}, \qquad i = 1, 2, \cdots, n_s \tag{2.4}$$

and $\tilde{v}_{\mathrm{ref};i}$ has the same statistical distribution as $\tilde{v}_i$.

The idea of channel equalization is to process the measurements $\{x_i\}_{i=1}^{n_s}$ of the testing channel incorporating the reference channel data $\{x_{\mathrm{ref};i}\}_{i=1}^{n_s}$ with the objective that the processed data will resemble the referenced data in some way. One way to achieve this is to filter the mismatched channel with an equalizer with FIR $\{h_{\mathrm{eq};i}\}_{i=1}^{n_{\mathrm{eq}}}$ to produce the equalized output

$$x_{\mathrm{eq};i} = \sum_{l=-\infty}^{\infty} x_{i-l+1}\, h_{\mathrm{eq};l} = \sum_{l=1}^{n_{\mathrm{eq}}} x_{i-l+1}\, h_{\mathrm{eq};l}, \qquad i = 1, 2, \cdots, n_s. \tag{2.5}$$

The formulation in (2.5) while seemingly simple is not practical, because equalization has to be applied to the entire data set acquired in a time snapshot interval. In order to accomplish this, we first note that (2.1) is equivalent to

$$x_i = \begin{bmatrix} r_i & r_{i-1} & \cdots & r_{i-n_h+1} \end{bmatrix} \begin{bmatrix} h_1 & r_2 & \cdots & r_{n_h} \end{bmatrix}'.$$

As a result, the entire observation set over one-time snapshot can be represented by

$$\begin{aligned} \mathbf{x} &= \tilde{\boldsymbol{\Psi}}(\mathbf{r})\mathbf{h} \\ &= \tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{s}} + \tilde{\mathbf{v}})\mathbf{h} = \tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{s}})\mathbf{h} + \tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{v}})\mathbf{h} \\ &= \mathbf{s} + \mathbf{v}, \end{aligned} \tag{2.6}$$

where prime denotes matrix/vector transpose. The measurements $\{x_i | i = 1, \cdots, n_s\}$ are stacked into a column vector $\mathbf{x}$; the same for the discretized FIR filter $\{h_i | i = 1, \cdots, n_h\}$. The measurements $\{r_i\}$ are then mapped into an $n_s \times n_h$ matrix $\tilde{\boldsymbol{\Psi}}(\mathbf{r})$, which is a superposition of the noise samples $\{\tilde{v}_i\}$ and the excitation signals $\{\tilde{s}_i\}$:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_s} \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{n_h} \end{bmatrix}, \quad \tilde{\boldsymbol{\Psi}}(\boldsymbol{\tau}) = \begin{bmatrix} \tau_1 & \tau_0 & \cdots & \tau_{2-n_h} \\ \tau_2 & \tau_1 & \cdots & \tau_{3-n_h} \\ \vdots & \vdots & \cdots & \vdots \\ \tau_{n_s} & \tau_{n_s-1} & \cdots & \tau_{n_s-n_h+1} \end{bmatrix}, \boldsymbol{\tau} \in \mathbf{r}, \tilde{\mathbf{s}}, \tilde{\mathbf{v}}. \tag{2.7}$$

Similarly, the equalized output is

$$x_{\mathrm{eq};i} = \begin{bmatrix} x_i & x_{i-1} & \cdots & x_{i-n_{\mathrm{eq}}+1} \end{bmatrix} \begin{bmatrix} h_{\mathrm{eq};1} & r_2 & \cdots & r_{\mathrm{eq};n_{\mathrm{eq}}} \end{bmatrix}'$$

and by the linearity property, it is related to the excitation signal

$$\begin{aligned}\mathbf{x}_{\text{eq}} &= \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}} \\ &= \boldsymbol{\Psi}\left(\tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{s}}+\tilde{\mathbf{v}})\mathbf{h}\right)\mathbf{h}_{\text{eq}} \\ &= \boldsymbol{\Psi}\left(\tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{s}})\mathbf{h}\right)\mathbf{h}_{\text{eq}} + \boldsymbol{\Psi}\left(\tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{v}})\mathbf{h}\right)\mathbf{h}_{\text{eq}} \\ &= \boldsymbol{\Psi}(\mathbf{s})\mathbf{h}_{\text{eq}} + \boldsymbol{\Psi}(\mathbf{v})\mathbf{h}_{\text{eq}}.\end{aligned} \tag{2.8}$$

$\mathbf{x}_{\text{eq}}$ is a n_{eq} vector of $\{x_{\text{eq};i}\}$; the mismatched noise-free data $\{s_i\}$ is mapped into a $n_s \times n_{\text{eq}}$ matrix $\boldsymbol{\Psi}(\mathbf{s})$; $\mathbf{h}_{\text{eq}}$ is a n_{eq} vector of $\{h_{\text{eq};i}\}$; and $\mathbf{v}_{eq}$ is a n_{eq} vector $\{v_i\}$ mapped into $\boldsymbol{\Psi}(\bullet)$ followed by the linear operation of the equalizer:

$$\mathbf{x}_{\text{eq}} = \begin{bmatrix} x_{\text{eq};1} \\ x_{\text{eq};2} \\ \vdots \\ x_{\text{eq};n_s} \end{bmatrix}, \quad \boldsymbol{\Psi}(\boldsymbol{\tau}) = \begin{bmatrix} \tau_1 & \tau_0 & \cdots & \tau_{2-n_{\text{eq}}} \\ \tau_2 & \tau_1 & \cdots & \tau_{3-n_{\text{eq}}} \\ \vdots & \vdots & \cdots & \vdots \\ \tau_{n_s} & \tau_{n_s-1} & \cdots & \tau_{n_s-n_{\text{eq}}+1} \end{bmatrix}, \ \boldsymbol{\tau} \in \mathbf{x}, \mathbf{s}, \mathbf{v}, \tag{2.9}$$

$$\mathbf{h}_{\text{eq}} = \begin{bmatrix} h_{\text{eq};1} \\ h_{\text{eq};2} \\ \vdots \\ h_{\text{eq};n_s} \end{bmatrix}, \quad \mathbf{v}_{\text{eq}} = \begin{bmatrix} v_{\text{eq};1} \\ v_{\text{eq};2} \\ \vdots \\ v_{\text{eq};n_s} \end{bmatrix} = \boldsymbol{\Psi}(\mathbf{v})\mathbf{h}_{\text{eq}} \tag{2.10}$$

The exact form of the equalizer will be derived pending on a suitable maximization criterion, as described in the next Section.

Least Squares Error Criterion

According to **Figure 1**, output snapshots from the equalizer and reference channels are given by $\mathbf{x}_{\text{eq}}$, $\mathbf{x}_{\text{ref}}$ respectively, where $\mathbf{x}_{\text{eq}} = \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}$. One approach is to minimize the square of the Euclidean distance between these two data sets[9] (superscript H denotes complex conjugation)

$$\begin{aligned}\varepsilon &= \left\|\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right\|^2 \\ &= \left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)^H \left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right).\end{aligned} \tag{2.11}$$

Invoking the identities $\frac{d}{d\boldsymbol{\beta}}\boldsymbol{\alpha}'\boldsymbol{\beta} = \boldsymbol{\alpha}$, $\frac{d}{d\boldsymbol{\beta}}\boldsymbol{\beta}'\boldsymbol{\alpha} = \boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ are vectors, the derivative of ε with respect to $\mathbf{h}_{\text{eq}}$ is

$$\begin{aligned}\frac{d\varepsilon}{d\mathbf{h}_{\text{eq}}} &= -\boldsymbol{\Psi}^H(\mathbf{x})\left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right) - \left[\boldsymbol{\Psi}^H(\mathbf{x})\left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)\right]^* \\ &= -2\,\text{Re}\left[\boldsymbol{\Psi}^H(\mathbf{x})\left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)\right].\end{aligned} \tag{2.12}$$

Setting it to zero, a necessary condition for minimizing ε is

$$\mathbf{h}_{\text{eq}} = \left(\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{x})\mathbf{x}_{\text{ref}}. \tag{2.13}$$

The second derivative is obtained from the first line of (2.12)

$$\frac{d^2\varepsilon}{d\mathbf{h}_{\text{eq}}^2} = \left[\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right]' + \left[\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right]^H$$
$$= 2\,\text{Re}\left[\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right]. \quad (2.14)$$

For any nonzero vector $\boldsymbol{\alpha}$, $\boldsymbol{\Psi}(\mathbf{x})\boldsymbol{\alpha}$ equals $\boldsymbol{\beta} = \begin{bmatrix}\beta_1 & \cdots & \beta_{n_s}\end{bmatrix}'$, and $\boldsymbol{\alpha}^H\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\boldsymbol{\alpha} = \sum_{i=1}^{n_s}|\beta_i|^2$ is a positive quantity. Therefore, $\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})$ is a positive-definite matrix. As $\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})$ is Hermitian, it contains only real eigenvalues. Consequently, all eigenvalues in $\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})$ are positive. This means that $d^2\varepsilon / d\mathbf{h}_{\text{eq}}^2$ is positive-definite and $\mathbf{h}_{\text{eq}}$ (2.13) is a necessary and sufficient condition for Least-Squares Error (LSE) optimization[10]. On the other hand, expand (2.11) as

$$\varepsilon = \left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)^H \left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)$$
$$= -\mathbf{h}_{\text{eq}}^H\left[\boldsymbol{\Psi}^H(\mathbf{x})\left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)\right] + \mathbf{x}_{\text{ref}}^H\left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)$$

then, by selecting $\mathbf{n}_{\text{eq}}$ such that it is orthogonal to the expression in brackets, the same solution (2.13) is obtained. The knowledge of the "error being always orthogonal to the estimate" is in fact, the basic principle of optimization. Bearing in mind that the first term vanishes and, when coupled with (2.13), the resulting minimum error is

$$\varepsilon = \mathbf{x}_{\text{ref}}^H\left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}}\right)$$
$$= \mathbf{x}_{\text{ref}}^H\left(\mathbf{x}_{\text{ref}} - \boldsymbol{\Psi}(\mathbf{x})\left(\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{x})\mathbf{x}_{\text{ref}}\right)$$
$$= \mathbf{x}_{\text{ref}}^H\left\{\mathbf{I} - \boldsymbol{\Psi}(\mathbf{x})\left(\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{x})\right\}\mathbf{x}_{\text{ref}}. \quad (2.15)$$

By definition ε (2.11) is strictly positive, so $\mathbf{I} - \boldsymbol{\Psi}(\mathbf{x})\left(\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{x})$ is a positive-definite matrix.

3. Properties of the Equlaizer

3.1 Equalizer Output Decomposition

From the results in Section 2.2, the equalizer output is the convolution of a sequence of measurements with the equalizer. With the data $\mathbf{x}$ stacked as a 2-D matrix $\boldsymbol{\Psi}(\mathbf{x})$, convolution is the product of $\boldsymbol{\Psi}(\mathbf{x})$ and the equalizer in (2.13). The equalizer output in (2.8) is

$$\mathbf{x}_{\text{eq}} = \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}} = \mathbf{A}^H(\mathbf{x})\mathbf{x}_{\text{ref}} \quad (3.1)$$

Where: $\mathbf{A}(\mathbf{x})$ is a $n_s \times n_s$ matrix

$$\mathbf{A}(\mathbf{x}) = \boldsymbol{\Psi}(\mathbf{x})\left(\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{x}). \quad (3.2)$$

In (3.1), the equalization matrix $\mathbf{A}$ is composed of n_s columns, with each column being a vector of length n_s. The equalizer output is the projection of the reference channel output onto the columns of the equalization matrix[11]. In particular, the projection of $\mathbf{x}_{\text{ref}}$ onto the kth column of $\mathbf{A}$ produces the kth element output, $\mathbf{x}_{\text{eq};k}$. This indicates that the matrix $\mathbf{A}$ plays a prominent role in channel equalization. The following theorem asserts that $\mathbf{A}$ can be expressed in a simpler form.

Proposition 3.1. The matrix $\mathbf{A}^{n_s \times n_s}$ defined in (3.2) can be decomposed into

$$\mathbf{A}(\mathbf{x}) = \boldsymbol{\xi}\boldsymbol{\Sigma}\boldsymbol{\xi}^H$$

with $\boldsymbol{\xi}^{n_s \times n_s}$ a unitary matrix such that $\boldsymbol{\xi}^H\boldsymbol{\xi} = \boldsymbol{\xi}\boldsymbol{\xi}^H = \mathbf{I}^{n_s}$, and $\boldsymbol{\Sigma}^{n_s \times n_s}$ is a diagonal matrix.

Proof: Since $\mathbf{A}^{n_s \times n_s}$ is Hermitian it has the eigenvectors $\left\{\xi_k^{n_s \times 1}, k = 1, \cdots, n_s\right\}$ and the positive eigenvalues $\left\{\sigma_k, k = 1, \cdots, n_s\right\}$ satisfying[11]

$$\mathbf{A}\xi_k = \sigma_k \xi_k, \quad k = 1, \cdots, n_s.$$

Without loss of generality, the eigenvectors are assumed orthonomal, i.e., $\xi_k^H \xi_l = \delta_{kl}$, and δ_{kl} is the Kronecker delta. Incorporating all eigenvectors, then

$$\mathbf{A}\left[\xi_1 \quad \cdots \quad \xi_{n_s}\right] = \left[\xi_1 \quad \cdots \quad \xi_{n_s}\right]\boldsymbol{\Sigma}$$

or

$$\mathbf{A}\xi = \xi\boldsymbol{\Sigma}$$

where $\xi = \left[\xi_1 \quad \cdots \quad \xi_{n_s}\right]$, $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \cdots, \sigma_{n_s})$.

Post-multiplying the equation by ξ^H and, by viture of the orthogonality of ξ, the result follows.Q.E.D

From the proof of Proposition 3.1, the equalization matrix $\mathbf{A}(\mathbf{x})$ is comprised of matrices ξ, $\boldsymbol{\Sigma}$. Columns of ξ are eigenvectors of $\mathbf{A}(\mathbf{x})$, and $\boldsymbol{\Sigma}$ is a diagonal matrix with part of its diagonal elements populated by the eigenvalues of $\mathbf{A}(\mathbf{x})$. The following theorem states that $\mathbf{A}^{n_s \times n_s}$ is not of full rank.

Proposition 3.2. The eigenvalues of $\mathbf{A}^{n_s \times n_s}$ (3.2) are either ones or zeros. More specifically, $\boldsymbol{\Sigma}$ contains r $(r \le n_{\mathrm{eq}})$ ones and the remaining $n_s - r$ diagonal elements are zero.

Proof: Refer to the Appendix in Section 6.

We introduce the $n_s \times 1$ vectors $\left\{\boldsymbol{\varphi}_k, k = 1, \cdots, r\right\}$, where $\boldsymbol{\varphi}_k$ is the kth column vector of $\xi\boldsymbol{\Sigma}^{0.5}$, i.e., $\boldsymbol{\varphi}_k = \sigma_k^{0.5}\xi_k$, and σ_k is the kth diagonal element of $\boldsymbol{\Sigma}$. Following the proof of Proposition 3.2 then the vectors $\left\{\boldsymbol{\varphi}_k, k = 1, \cdots, r\right\}$ span an $n_s \times 1$ vector space V_r. In particular, the testing channel output $\mathbf{x}$ lies in V_r. The goal of channel equalization is to project $\mathbf{x}_{\mathrm{ref}}$ onto the V_r space to yield $\mathbf{x}_{\mathrm{eq}}$ such that the error vector $\mathbf{x}_{\mathrm{ref}} - \mathbf{x}_{\mathrm{eq}}$ incurs the least squares error, as in **Figure 2.**

Similarly, we define the $n_s \times 1$ dimensional basis vectors $\left\{\hat{\boldsymbol{\varphi}}_k, k = 1, \cdots, n_s\right\}$ that spans the vector space $\hat{V}_{n_s}$. These basis vectors can be generated using for example, the Gram-Schmidt orthogonalization procedure on the data $\mathbf{x}_{\mathrm{ref}}$. In particular, the referenced channel output $\mathbf{x}_{\mathrm{ref}}$ lies in $\hat{V}_{n_s}$. Normally, V_r, $\hat{V}_{n_s}$ have no relationship with each other, because they are derived from two sets of data contaminated by two independent sets of noise. On the other hand, in high input signal-to-noise ratio environments, it is not difficult to see that V_r is a subspace $\hat{V}_{n_s}$. Further, if n_{eq} is selected sufficiently high, then the spaces V_r, $\hat{V}_{n_s}$ almost overlap, with the impact that, under such conditions, $\mathbf{x}_{\mathrm{eq}}$ is almost equal to $\mathbf{x}_{\mathrm{ref}}$, and accounts for almost zero error (2.15) [12].

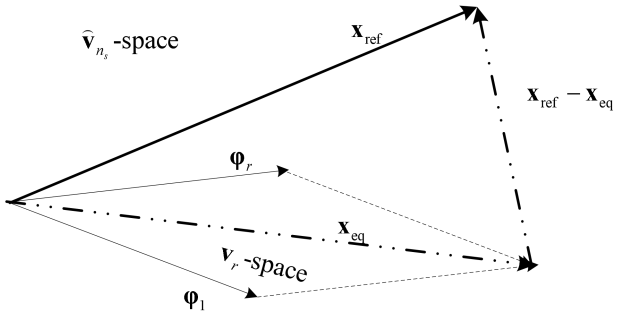

Figure 2; Geometrical interpretation of channel equalization

in the n_s-dimensional hyperspace.

3.2 Some Performance Measures

Proposition (3.1) simplifies the resulting error (2.15) as

$$\begin{aligned}\varepsilon &= \mathbf{x}_{\text{ref}}^H(\mathbf{I}-\boldsymbol{\xi}\boldsymbol{\Sigma}\boldsymbol{\xi}^H)\mathbf{x}_{\text{ref}} \\ &= (\boldsymbol{\xi}^H\mathbf{x}_{\text{ref}})^H(\mathbf{I}-\boldsymbol{\Sigma})\boldsymbol{\xi}^H\mathbf{x}_{\text{ref}}.\end{aligned} \quad (3.3)$$

Introduce an $n_s \times 1$ vector $\mathbf{y}$

$$\mathbf{y} = \boldsymbol{\xi}^H\mathbf{x}_{\text{ref}} \quad (3.4)$$

which is the projection of $\mathbf{x}_{\text{ref}}$ onto the columns of $\boldsymbol{\xi}$, where $\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_{n_s} \end{bmatrix}'$.

Correspondingly, y_k is the projection of $\mathbf{x}_{\text{ref}}$ onto the kth column of $\boldsymbol{\xi}$

$$y_k = \boldsymbol{\xi}_k^H\mathbf{x}_{\text{ref}}. \quad (3.5)$$

Therefore, (3.3) becomes

$$\begin{aligned}\varepsilon &= \mathbf{y}^H(\mathbf{I}-\boldsymbol{\Sigma})\mathbf{y} \\ &= \sum_{k=1}^{n_s}(1-\sigma_k)|y_k|^2.\end{aligned} \quad (3.6)$$

(3.6) is a succinct expression which can be used to emphasize the merits of channel equalization. For example, let us suppose that the equalizer output is a zero vector, i.e., a completely dysfunctional device, then the resulting error is: $\varepsilon = \mathbf{y}^H\mathbf{y} = \mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}}$. Channel equalization reduces the error by the quantity $\mathbf{y}^H\boldsymbol{\Sigma}\mathbf{y}$ to yield an error as seen in (3.6). Due to the distribution of the eigenvalues (see Proposition 3.2) the error is simplified as $\varepsilon = \sum_{k=r+1}^{n_s}|y_k|^2$. The error, originally defined in (2.11) and rewritten in (3.3), measures the distance square between the reference channel output and the equalized output in the Euclidean space. In fact, ε is the energy of the error vector $\mathbf{x}_{\text{ref}} - \mathbf{x}_{\text{eq}}$ over a snapshot time interval. Another suitable metric, for assessing the goodness of the proposed equalization scheme, is the instantaneous Correlation Coefficient (CC) between the equalized output and the reference channel output, given by

$$\rho = \frac{\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{eq}}}{\sqrt{\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}}}\sqrt{\mathbf{x}_{\text{eq}}^H\mathbf{x}_{\text{eq}}}} \quad (3.7)$$

Using the Cauchy-Schwarz inequality

$$\left|\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{eq}}\right| \le \sqrt{\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}}}\sqrt{\mathbf{x}_{\text{eq}}^H\mathbf{x}_{\text{eq}}},$$

where the equality is true when $\mathbf{x}_{\text{ref}}$, $\mathbf{x}_{\text{eq}}$ differ by a scaling factor. Thus for $0 \le |\rho| \le 1$, the instantaneous correlation coefficient measures how close the equalized output is with respect to the referenced channel output, leading to $\rho = 1$ as a perfect equalization. The following theorem provides a precise computation of ρ.

Proposition 3.3. Given the equalizer output $\mathbf{x}_{\text{eq}} = \boldsymbol{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}} = \mathbf{A}^H(\mathbf{x})\mathbf{x}_{\text{ref}}$, then the instantaneous CC is a real positive scalar equal to

$$\rho = \sqrt{\frac{\sum_{k=1}^{n_s}\sigma_k|y_k|^2}{\sum_{k=1}^{n_s}|y_k|^2}}.$$

where y_k is the projection of $\mathbf{x}_{\text{ref}}$ onto ξ_k (see (3.5)), and σ_k is one or zero.

Proof: Because $\mathbf{A}(\mathbf{x})$ is Hermitian and idempotent (consult proof of Proposition 3.2), then $\mathbf{x}_{\text{eq}}^H\mathbf{x}_{\text{eq}} = (\mathbf{A}^H(\mathbf{x})\mathbf{x}_{\text{ref}})^H\mathbf{A}^H(\mathbf{x})\mathbf{x}_{\text{ref}} = \mathbf{x}_{\text{ref}}^H\mathbf{A}(\mathbf{x})\mathbf{x}_{\text{ref}}$. Also, $\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{eq}} = \mathbf{x}_{\text{ref}}^H\mathbf{A}^H(\mathbf{x})\mathbf{x}_{\text{ref}} = \sum_{k=1}^{n_s}\sigma_k|y_k|^2$, and $\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}} = \sum_{k=1}^{n_s}|y_k|^2$ (see (3.6)). Upon replacing (3.7) by these three components, the result is obtained.

Lastly, we introduce the (jamming) Cancellation Ratio (CR):

$$J = 1 - |\rho|^2. \tag{3.8}$$

When the reference and the equalized channel outputs are in close proximity, ρ is nearly 1 or in logarithmic scale, i.e., close to 0 dB. Clearly, while ρ is useful in measuring channel equalization performance, it is not quite as sensitive in differentiating the differences in the data sets $\mathbf{x}_{\text{ref}}$, $\mathbf{x}_{\text{eq}}$ being close to each other. In contrast, ρ approaching one translates to J approaching zero with its logarithm a very negative number. Based on this argument, J can be readily used to access the degree of channel equalization achieved at any point in the process and is therefore a more useful tool in channel equalization evaluation.

Proposition 3.3 facilitates the computation of the cancellation ratio

$$J = 1 - \frac{\sum_{k=1}^{n_s}\sigma_k|y_k|^2}{\sum_{k=1}^{n_s}|y_k|^2} = \frac{\sum_{k=1}^{n_s}(1-\sigma_k)|y_k|^2}{\sum_{k=1}^{n_s}|y_k|^2}. \tag{3.9}$$

The error in (3.6) is next expressed as

$$\varepsilon = J\sum_{k=1}^{n_s}|y_k|^2 = J\|\mathbf{x}_{\text{ref}}\|^2 \tag{3.10}$$

and shows explicitly that the CR is the error normalized by the energy of the reference channel output. In the above, $\|\mathbf{x}_{\text{ref}}\|$ denotes the norm of $\mathbf{x}_{\text{ref}}$, i.e., $\|\mathbf{x}_{\text{ref}}\|^2 = \mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}}$.

3.3 Effects of System Parameters

In Section 2.2, we show that the equalizer output is

$$\mathbf{x}_{\text{eq}} = \mathbf{\Psi}(\mathbf{x})\mathbf{h}_{\text{eq}} = \mathbf{A}^H(\mathbf{x})\mathbf{x}_{\text{ref}} \tag{3.11}$$

where $\mathbf{x} = \mathbf{s} + \mathbf{v}$, $\mathbf{\Psi}(\mathbf{x})$ is $n_s \times n_{\text{eq}}$, $\mathbf{A}(\mathbf{x})$ is given by (3.2), and $\mathbf{x}$ is the output from the testing channel $\mathbf{h}$ in response to the input $\tilde{\mathbf{r}}$: $\tilde{\mathbf{r}} = \tilde{\mathbf{s}} + \tilde{\mathbf{v}}$: $\mathbf{x} = \tilde{\mathbf{\Psi}}(\tilde{\mathbf{r}})\mathbf{h}$.

In this setting, the equalizer output is a function of system parameters, including window size n_s, equalizer length n_{eq}, and input signal-to-noise ratio (ISNR). Their impact on channel equalization is of prime concern and will be studied in details in later sections. As discussed in Section 2, the LSE ε, the instantaneous correlation coefficient ρ, and the cancellation ratio J can be used to provide reasonable performance metrics for the proposed channel equalization algorithm. Nevertheless, due to its goodness in differentiating between two data sets, the cancellation ratio will be the metric we primarily use.

3.3.1 Effects of equalizer length

Proposition 3.2 indicates that the first r $(r \le n_{\text{eq}})$ eigenvalues are one and the last $n_s - r$ eigenvalues are zero. As a result, the cancellation ratio (3.9) reduces to

$$J = \frac{\sum_{k=r+1}^{n_s} |y_k|^2}{\sum_{k=1}^{n_s} |y_k|^2}. \quad (3.12)$$

Suppose n_{eq} increases, then the numerator becomes smaller because it encapsulates fewer number of terms. With n_s fixed, the set of eigenvectors $\{\boldsymbol{\xi}_k, k = 1, \cdots, n_s\}$ remains the same; this property extends to y_k also (see (3.5)). As the numerator holds fewer terms while the denominator (3.12) remains constant, J decreases (a highly desirable property) and, in fact, is a strictly decreasing function of the parameter r $(r \le n_{\text{eq}})$.

In particular, when n_{eq} reaches n_s then the number of terms in the numerator in (3.12) is near zero. As a result, J drops to zero at which point, we reach a perfect channel equalization.

Alternatively, suppose n_{eq} equals n_s, then $\boldsymbol{\Psi}(\mathbf{x})$ is a square matrix and $\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})$ is invertible, leading to

$$\begin{aligned}\mathbf{x}_{\text{eq}} &= \mathbf{A}^H(\mathbf{x})\mathbf{x}_{\text{ref}} \\ &= \boldsymbol{\Psi}(\mathbf{x})\left(\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{x})\mathbf{x}_{\text{ref}} = \mathbf{x}_{\text{ref}}\end{aligned}$$

which implies that $J = 0$ and once more, we reach perfect equalization.

3.3.2 Effects of window size

Suppose the window size n_s is raised to $\dot{n}_s$, where $\dot{n}_s = n_s + \kappa$, and κ a positive integer, with all other system parameters unchanged.

Following the two paragraphs preceding Section 0, the equalization matrix $\mathbf{A}(\mathbf{x})$, that is the source of the eigenvectors $\{\boldsymbol{\xi}_k, k = 1, \cdots, n_s\}$, becomes $\dot{\mathbf{A}}(\dot{\mathbf{x}})$ with the eigenvectors $\{\dot{\boldsymbol{\xi}}_k, k = 1, \cdots, \dot{n}_s\}$; correspondingly, $\mathbf{x}_{\text{ref}}$ switches to $\dot{\mathbf{x}}_{\text{ref}}$. Note also that, for example, $\boldsymbol{\xi}_1$ is an $n_s \times 1$ vector while $\dot{\boldsymbol{\xi}}_1$ is an $\dot{n}_s \times 1$ vector.

Applying (3.4) and, bearing in mind that $\boldsymbol{\xi}$ is a unitary matrix, then

$$\dot{\mathbf{y}}^H\dot{\mathbf{y}} = \dot{\mathbf{x}}_{\text{ref}}^H\dot{\mathbf{x}}_{\text{ref}} = \mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}} + \sum_{k=n_s+1}^{\dot{n}_s} \left|x_{\text{ref};k}\right|^2.$$

Therefore, the cancellation ratio (3.12) at window size $\dot{n}_s$ is

$$J(\dot{n}_s) = \frac{\sum_{k=r+1}^{\dot{n}_s} |\dot{y}_k|^2}{\sum_{k=1}^{\dot{n}_s} |\dot{y}_k|^2} = \frac{\sum_{k=r+1}^{n_s} |\dot{y}_k|^2 + \sum_{k=n_s+1}^{\dot{n}_s} |\dot{y}_k|^2}{\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}} + \sum_{k=n_s+1}^{\dot{n}_s} \left|x_{\text{ref};k}\right|^2}. \quad (3.14)$$

For convenience, the cancellation ratio (3.12) based on window size n_s is rewritten here

$$J(n_s) = \frac{\sum_{k=r+1}^{n_s} |y_k|^2}{\sum_{k=1}^{n_s} |y_k|^2} = \frac{\sum_{k=r+1}^{n_s} |y_k|^2}{\mathbf{x}_{\text{ref}}^H\mathbf{x}_{\text{ref}}}. \quad (3.15)$$

Clearly, the denominator of $J(\dot{n}_s)$ is greater than $J(n_s)$. The disparity between the two formulations of the cancellation ratio widens when the number of extra terms κ increases. Since $y_k = \boldsymbol{\xi}_k^H\mathbf{x}_{\text{ref}}$,

$\dot{y}_k = \dot{\boldsymbol{\xi}}_k^H \left[\mathbf{x}_{\text{ref}} \quad x_{\text{ref;n}_s+1} \quad \cdots \quad x_{\text{ref;}\dot{n}_s} \right]$, where $\boldsymbol{\xi}_k^{n_s \times 1}$, $\dot{\boldsymbol{\xi}}_k^{\dot{n}_s \times 1}$ are unitary vectors, it is unclear whether $|\dot{y}_k|$ exceeds $|y_k|$ or vice versa, $n_{\text{eq}} \le r \le k \le n_s$, unless the actual data $\mathbf{x}_{\text{ref}}$, $\mathbf{x}$, $\dot{\mathbf{x}}_{\text{ref}}$, $\dot{\mathbf{x}}$ are available. Thus, there is no concrete evidence to support the fact that larger window size improves the cancellation ratio.

On the other hand, larger $\dot{n}_s$ has some effects in both the numerator and denominator of the cancellation ratio $J(\dot{n}_s)$ which become dependent on an ever increasing summation series $\sum_{k=n_s+1}^{\dot{n}_s} |\dot{y}_k|^2$ (due to window size increment). Therefore, provided the components $\sum_{k=r+1}^{n_s} |\dot{y}_k|^2$, $\sum_{k=r+1}^{n_s} |y_k|^2$ are close to each other, it would seem that $J(\dot{n}_s)$ will be greater than $J(n_s)$. The above ad hoc explanation suggests that higher window size will degrade system's performance.

3.3.3 Effects of input signal-to-noise ratio

Define the input SNR γ the ratio of the power of excitation signals $\{\tilde{s}_i\}$ to the power of noise $\{\tilde{v}_i\}$ at the channel front end:

$$\gamma = \frac{\frac{0.5}{n_s}\sum_{i=1}^{n_s} |S_i|^2}{0.5E\left\{|\tilde{v}_i|^2\right\}} = \frac{\frac{0.5}{n_s}\tilde{\mathbf{s}}^H\tilde{\mathbf{s}}}{0.5E\left\{|\tilde{v}_i|^2\right\}}, \tag{3.16}$$

where

$$\tilde{\mathbf{s}} = \left[\tilde{s}_1 \quad \cdots \quad \tilde{s}_{n_s} \right]', \quad E\left\{|\tilde{v}_i|^2\right\} = E\left\{|\tilde{v}_{\text{ref};i}|^2\right\}.$$

Combine (2.15), (3.10) for the cancellation ratio

$$\begin{aligned} J &= \frac{\mathbf{x}_{\text{ref}}^H \left\{ \mathbf{I} - \boldsymbol{\Psi}(\mathbf{x})\left(\boldsymbol{\Psi}^H(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{x}) \right\} \mathbf{x}_{\text{ref}}}{\mathbf{x}_{\text{ref}}^H \mathbf{x}_{\text{ref}}} \\ &= \frac{(\mathbf{s}_{\text{ref}} + \mathbf{v}_{\text{ref}})^H \left\{ \mathbf{I} - \boldsymbol{\Psi}(\mathbf{s}+\mathbf{v})\left(\boldsymbol{\Psi}^H(\mathbf{s}+\mathbf{v})\boldsymbol{\Psi}(\mathbf{s}+\mathbf{v})\right)^{-1}\boldsymbol{\Psi}^H(\mathbf{s}+\mathbf{v}) \right\} (\mathbf{s}_{\text{ref}} + \mathbf{v}_{\text{ref}})}{(\mathbf{s}_{\text{ref}} + \mathbf{v}_{\text{ref}})^H (\mathbf{s}_{\text{ref}} + \mathbf{v}_{\text{ref}})}, \end{aligned} \tag{3.17}$$

where $\mathbf{s} = \tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{s}})\mathbf{h}$, $\mathbf{v} = \tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{v}})\mathbf{h}$, $\mathbf{s}_{\text{ref}} = \tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{s}})\mathbf{h}_{\text{ref}}$, and $\mathbf{v}_{\text{ref}} = \tilde{\boldsymbol{\Psi}}(\tilde{\mathbf{v}}_{\text{ref}})\mathbf{h}_{\text{ref}}$. Due to its computational complexity, evaluation of the CR is left to be implemented with Monte Carlo simulations.

4. Simulation Results

Figure 3 shows the magnitude of the frequency response of the testing (solid line) and reference (dashed line) channels, with its zoomed-in version shown in **Figure 4.** They clearly show that these two channels are not matched. A sequence of excitation signals utilizing eight-level pseudo-random phase-shift keying (PSK) corrupted by complex Gaussian noise and an input SNR γ (see (2.2), (2.4), (3.16)), is injected in parallel onto the channels.

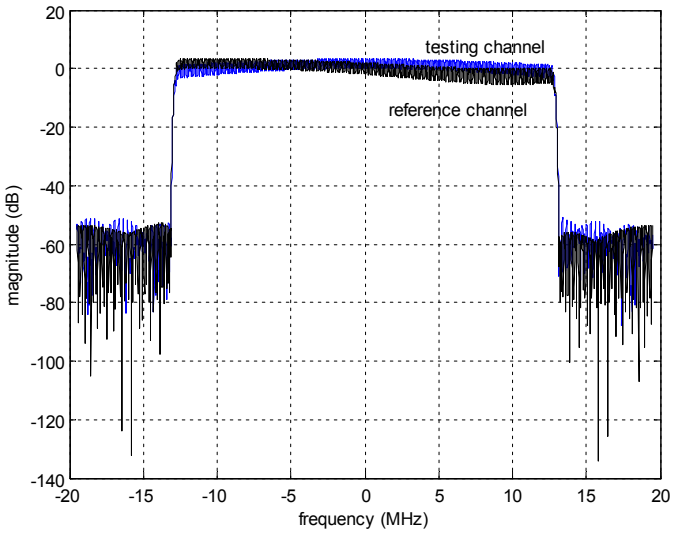

Figure 3; Magnitude plot of the frequency response of the testing and reference channels.

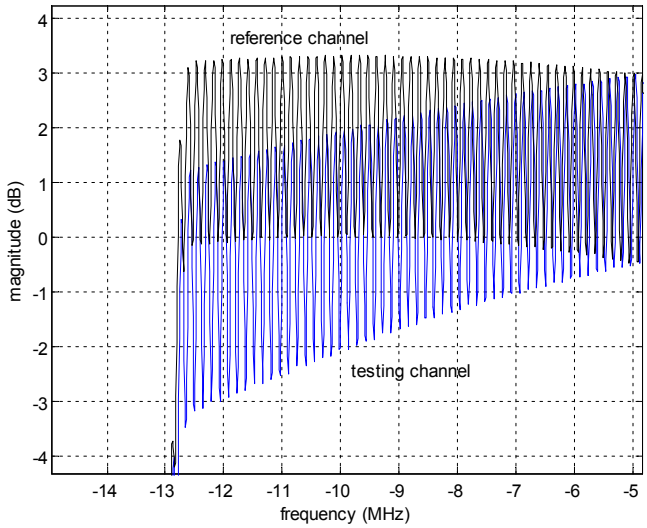

Figure 4; Zoomed-in portions of Figure 1.

This drives the difference in magnitude of the direct outputs at the testing (solid) and reference (dashed) channels, as shown in **Figure 5**, top panel. The corresponding phase plots are depicted in **Figure 6**, top panel. The testing channel output $\mathbf{x}$ is equalized using (3.1), to yield the output $\mathbf{X}_{eq}$.

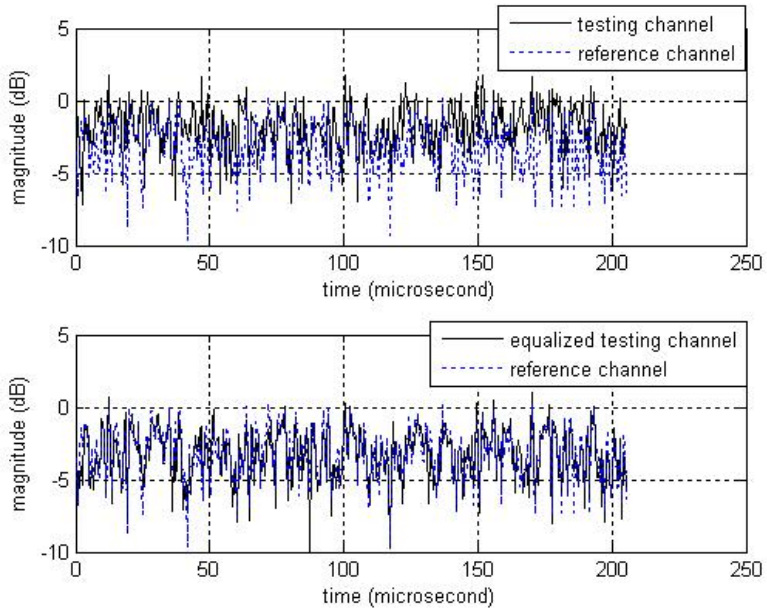

Figure 5; Magnitude plot (at input SNR=5 dB) of

Top: reference and direct testing channel outputs

Bottom: reference and equalized testing channel outputs.

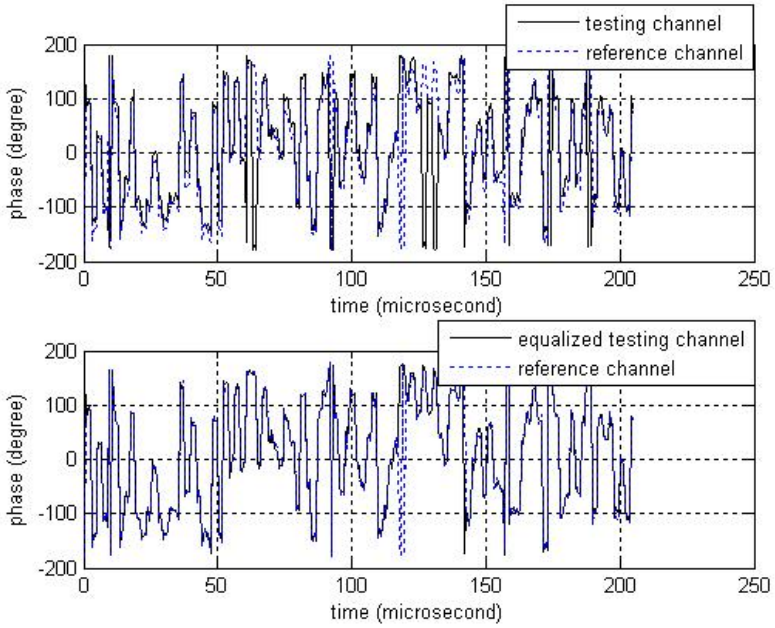

Figure 6; Phase plot (at input SNR=5 dB) of

Top: reference and direct testing channel outputs

Bottom: reference and equalized testing channel outputs.

Depending on the setting of the system parameters, including equalizer length, window size, and input SNR (ISNR), it is found that the equalized testing channel output and the reference channel output resemble each other to a certain extent, with the correlation coefficient and cancellation ratio used as evaluators. For comparison purpose, one would like to assess the differences in the metrics between the direct testing channel, i.e., un-equalized output and the reference channel output. Unless stated otherwise, the equalizer length and window size are set to 51, 400 respectively. In the Monte Carlo simulations being used here, an experiment that evaluates correlation coefficient (3.7) is repeated for a specified number of runs. In each trial, the same signal sequence $\tilde{\mathbf{s}}$ is employed, and is subsequently contaminated by independent noise sequences $\tilde{\mathbf{v}}$, $\tilde{\mathbf{v}}_{\text{ref}}$ and finally injected into the testing, and reference channel

respectively. This produces a set of correlation coefficients, for identical system-parameters and independent, identically distributed noise environments. The set up allows for a straight computation of the average correlation coefficient. The same procedure is applied to evaluate the average cancellation ratio on the basis of a constellation of cancellation ratios (3.8). However, in this section the average correlation coefficient, and average cancellation ratio are used. Also, the number of runs for each experiment is set equal to 3.

4.1 Variation of System Parameters

4.1.1 Equalizer length variation

Figure 7 shows that at 10 dB ISNR, the equalizer CR fares at least 10 dB better than the direct testing channel output over the course of equalizer length varying from 25 to 270, which corresponds to 6.3% to 67.5% of the window size (which is fixed).

With the ISNR raised to 60 dB, the landscape changes dramatically. It is found that the CR of the equalizer continues to improve over the course of increasing the equalizer length. According to **Figure 7**, the CR curve of the equalizer can be subdivided into two regions: over

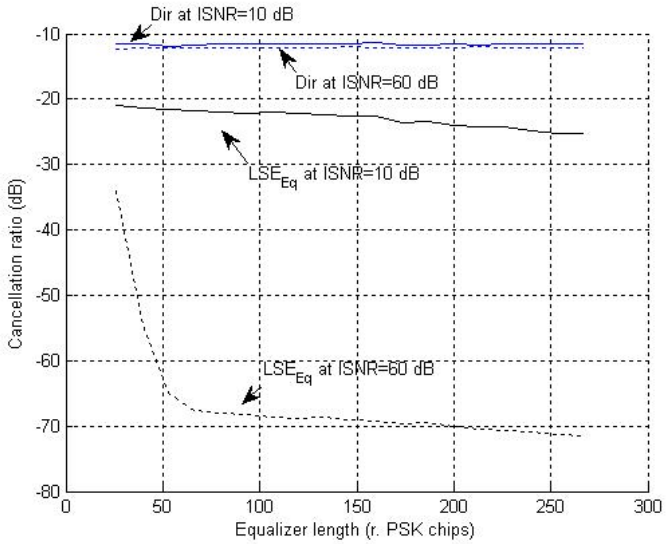

Figure 7; Cancellation ratio for the equalized and direct testing channel outputs at various equalizer lengths with input SNR as the parameter.

The portion of the equalizer length extending from 25 to 75 (6.3% to 17.5% of the window size), the equalizer CR hypothetically improves from -33 dB down to -67 dB. Further increase of the equalizer length from 17.5% to 67.5% of the window size improves the CR consistently and, in fact the CR in dB scale decreases linearly at the rate 0.025 dB per unit of increase in equalizer length. In contrast, the CR of the direct testing channel output is invariant with respect to equalizer length and stays at -12 dB. This behavior was expected, because the correlation coefficient in (3.7) with $\mathbf{x}_{eq}$ replaced by $\mathbf{x}$ is independent to variations in the equalizer. Also, the CRs of the direct testing channel output at 60 dB and 10 dB ISNR coincide almost perfectly.

The CC has also been studied as a function of equalizer length for the same ISNR parameters. When comparing **Figure 7** and **Figure 8**, it is apparent that the CR provides superior information on the correlation coefficient with varying equalizer length. This supports our earlier observation about the fidelity of the CR parameter, as described in previous sections.

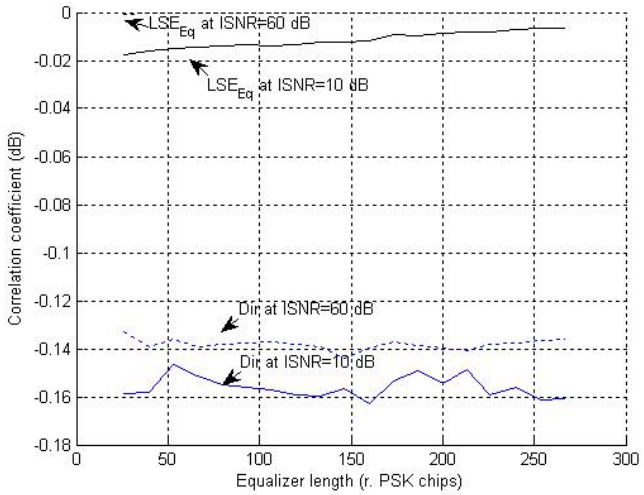

Figure 8; Correlation coefficient for the equalized and direct testing channel outputs at various equalizer lengths with ISNR as the parameter.

4.1.2 Window size variation

Suppose the ISNR is set equal to 60 dB and the window size is varied from 20 to 1000, i.e., 40% to 2000% of the nominal equalizer length (which is kept fixed). **Figure 9** illustrates that the LSE equalizer begins with a CR of 0 dB and falls asymptotically when the window size climbs uphill starting at 40% of the equalizer length. The CR values stop falling at the point where the window size reaches 100% of the nominal equalizer length, which coincides to the point where the CR reaches $-125\,\mathrm{dB}$ (or perfect equalization). With the window size increasing from 51 to 100 (100% to 200% of the equalizer length), the CR increases asymptotically until it attains a value of -58 dB. Increasing window size beyond 200 (400% of the nominal equalizer length) has minimal effect on the CR, which seems to stabilize at $-58\,\mathrm{dB}$. On the other hand, the direct testing channel output has a CR equal to a value of -12 dB, which is almost insensitive to window size variations.

When the input SNR is lowered to 10 dB exactly the same behavior, as described in the previous paragraph, is observed for the equalized output, except at the point where the window size reaches 100% of the equalizer length. Then, the equalizer achieves -115 dB CR, and subsequent window size increments stabilize the CR at -22 dB, as shown in **Figure 9.** Once again, the direct testing channel at 10 dB ISNR is almost the same as that when the ISNR is set at 60 dB. The same study is done for the CC as shown in Figure 10.

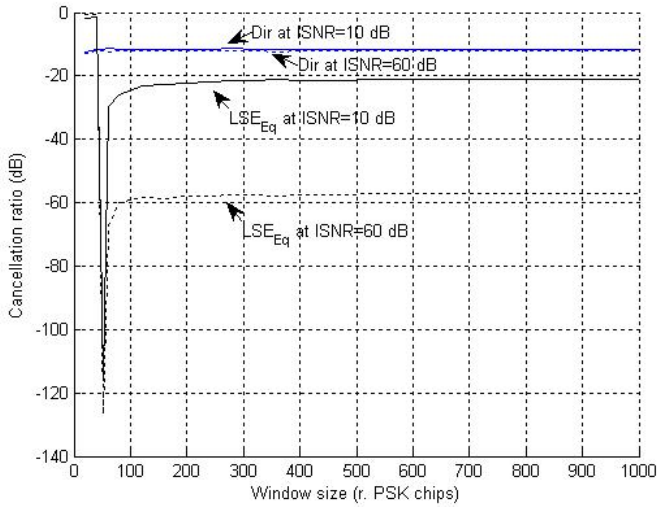

Figure 9; Cancellation ratio for the equalized and direct testing channel outputs at various window size with input SNR as the parameter.

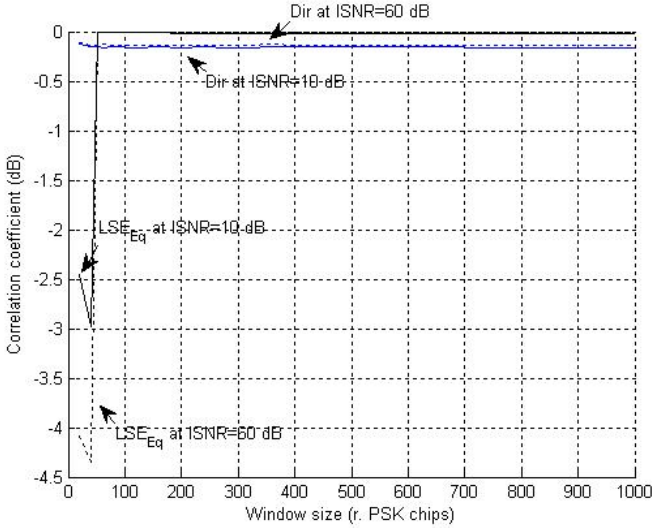

Figure 10; Correlation coefficient for the equalized and direct testing

channel outputs at various window size with input SNR as the parameter.

Figure 9 and **Figure 10** again confirm that the CR is a more efficient metric than the correlation coefficient specifically when trying to assess the equalizer performance as a function of window size variation.

4.1.3 Input signal-to-noise ratio variation

Figure 11 illustrates the variations of CR as a function of the equalizer as the ISNR is increased from 5 dB to 90 dB, while all other system parameters remain unchanged. We observe that over the course of increasing the ISNR from 5 dB to 40 dB, the CR in dB scale improves linearly and is 10 dB lower than the input SNR. Increasing the ISNR from 40 dB to 60 dB causes the CR (in dB scale) to behave as a convex function of ISNR, such that the CR drops from -50 to -57 dB. Raising the ISNR beyond 60 dB has practically no effect on the system, and provides no insignificant improvement in the CR, which converges to -57 dB. Once again, except for the limited range from 5 dB to 10 dB of the ISNR, the CR curve of the direct testing channel output is -12 dB and is almost invariant to changes in ISNR.

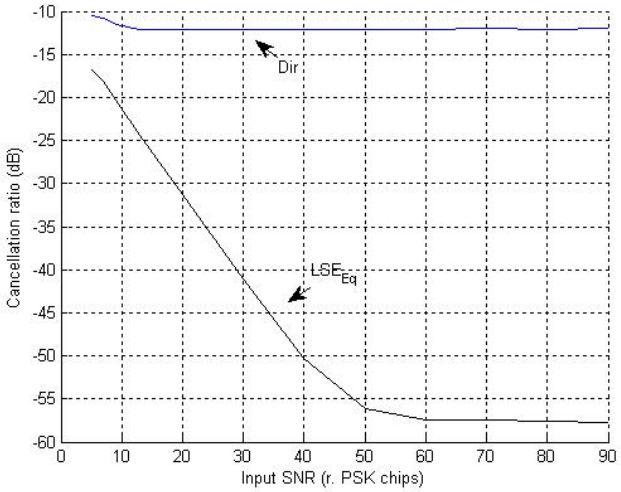

Figure 11; Cancellation ratio for the equalized and direct testing

channel outputs at various input SNR.

4.2 Equalized Output and Performance Metrics Relationship

Figure 12, bottom panel, shows the magnitude of the equalizer output (solid line) at 5 dB ISNR, and **Figure 13** is the zoomed-in version. For comparison purpose, the reference channel output (dash line) is overlaid on the same diagram. The presence of glitches between these two sets of outputs is a good indication that, in low ISNR

environments, the equalized output does not match the reference channel output. This is expected because, according to **Figure 7**, the corresponding CR (approximately -17 dB) is quite far from its stabilized value (approximately -57dB).

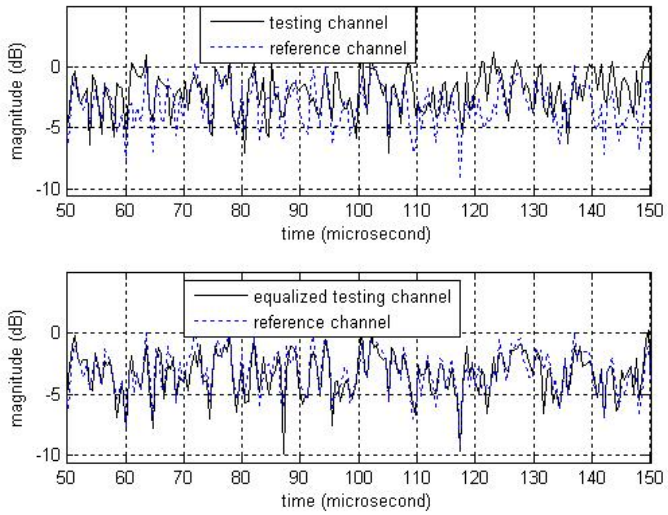

Figure 12; Zoomed-in magnitude plot (at input SNR=5 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

Nevertheless, by comparing the bottom and top panels of **Figure 11** or **Figure 12**, it seems that the magnitude of the equalizer output is a better match to the magnitude of the reference channel output than to the testing channel.

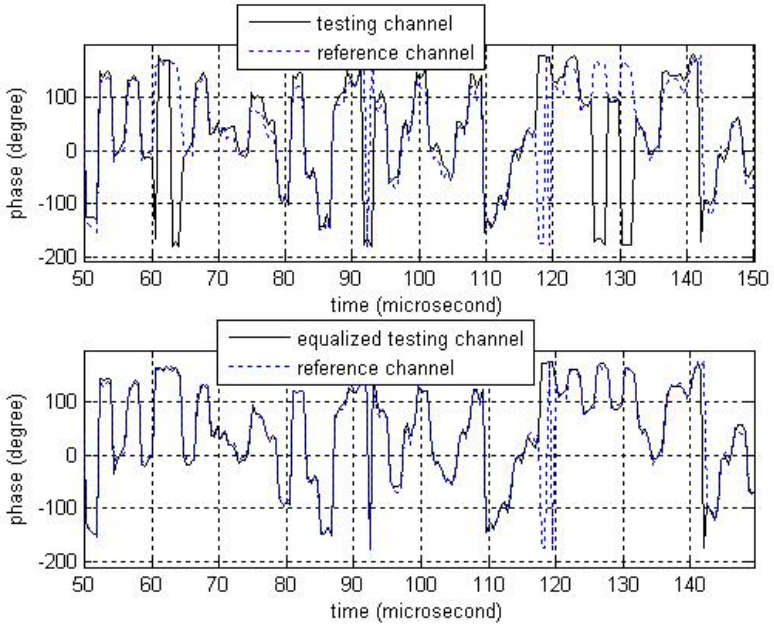

Figure 13; Zoomed-in magnitude plot (at input SNR=5 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

With the ISNR raised to 20 dB, **Figure 14** and **Figure 15** show the magnitude, and phase plots respectively, of the direct testing, equalized testing, and reference channel outputs, while **Figure 16** and **Figure 17** are their zoomed-in versions. The bottom panels of **Figure 14** and **Figure 15** (or **Figure 16** and **Figure 17**) show that at 20 dB ISNR, the magnitude (or phase) of the equalized channel, reference channel outputs are undistinguishable from each other over one-time snapshot interval.

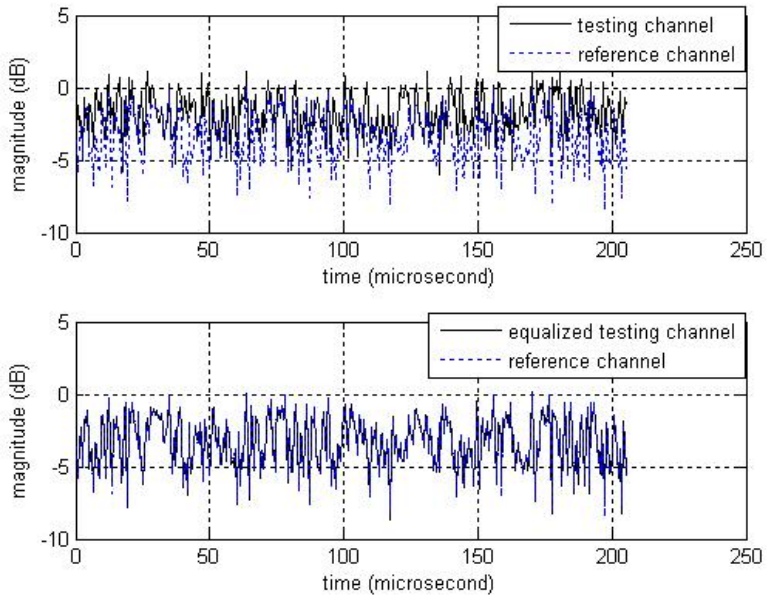

Figure 14; Magnitude plot (at input SNR=20 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

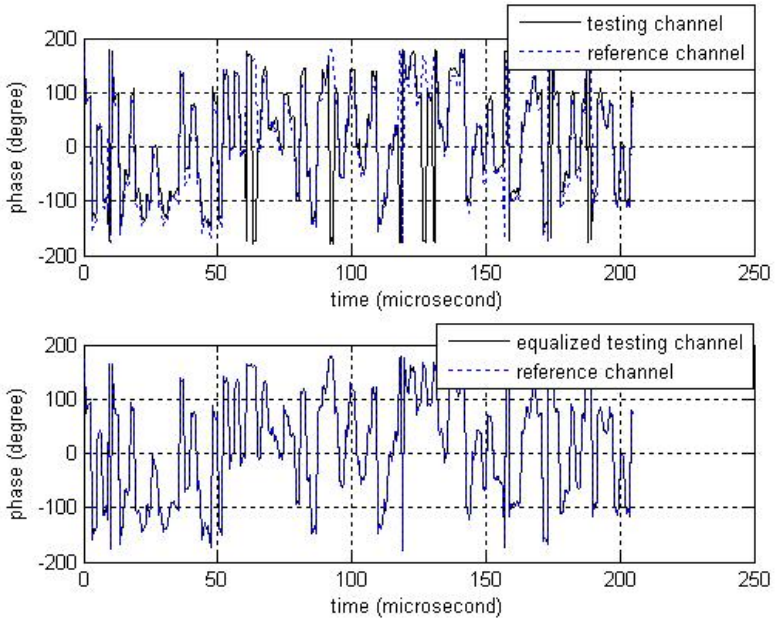

Figure 15; Phase plot (at input SNR=20 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

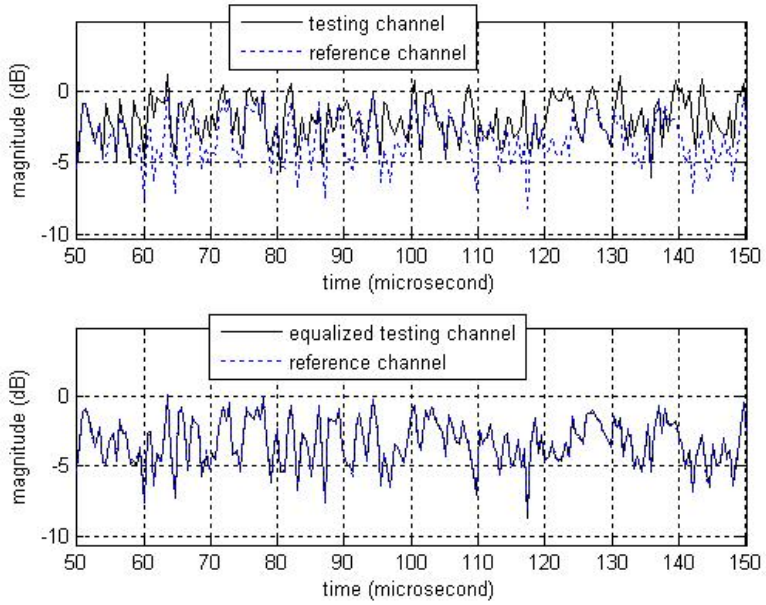

Figure 16; Zoomed-in magnitude plot (at input SNR=20 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

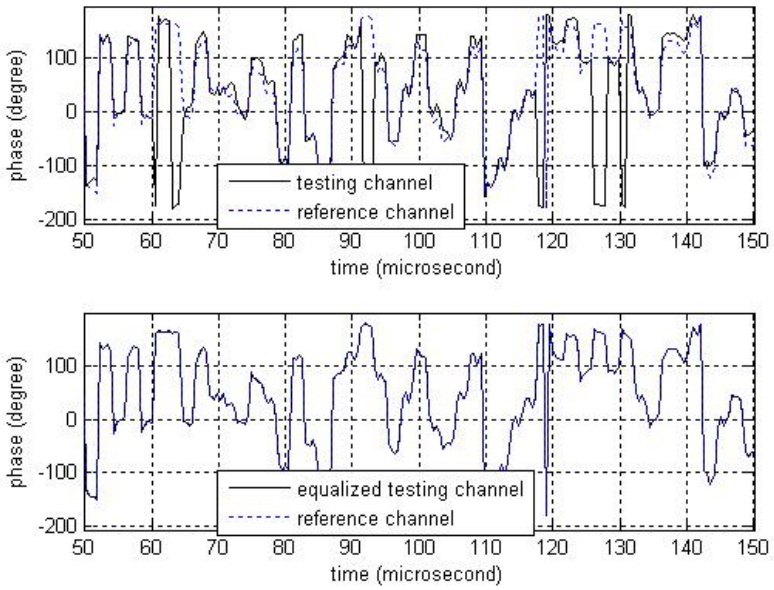

Figure 17; Zoomed-in phase plot (at input SNR=20 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

The concept of exact matching between the equalized testing and reference channel outputs can finally be illustrated in the bottom panels of **Figure 18** and its zoomed-in **Figure 20** (or **Figure 19** and its zoomed-in **Figure 21**) with regards to the way in which the magnitude (or phase) of the equalized testing and reference channel outputs matched at 40 dB ISNR.

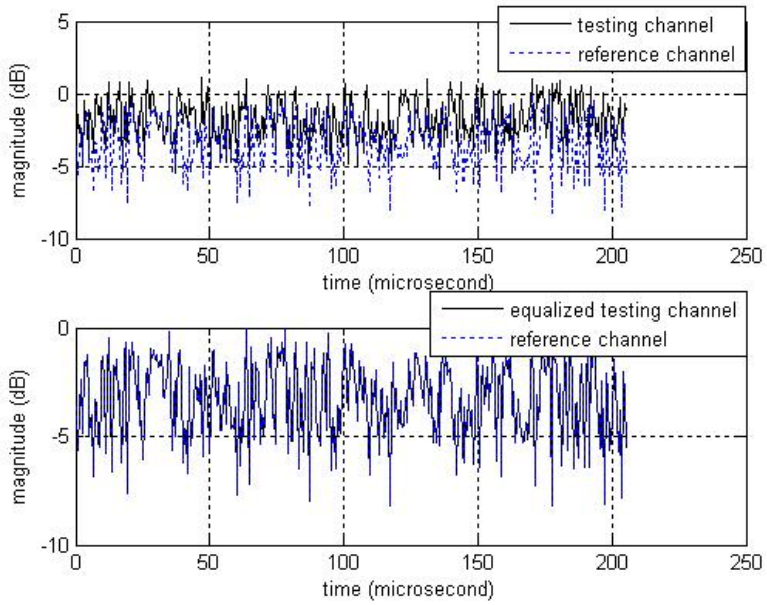

Figure 18; Magnitude plot (at input SNR=40 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

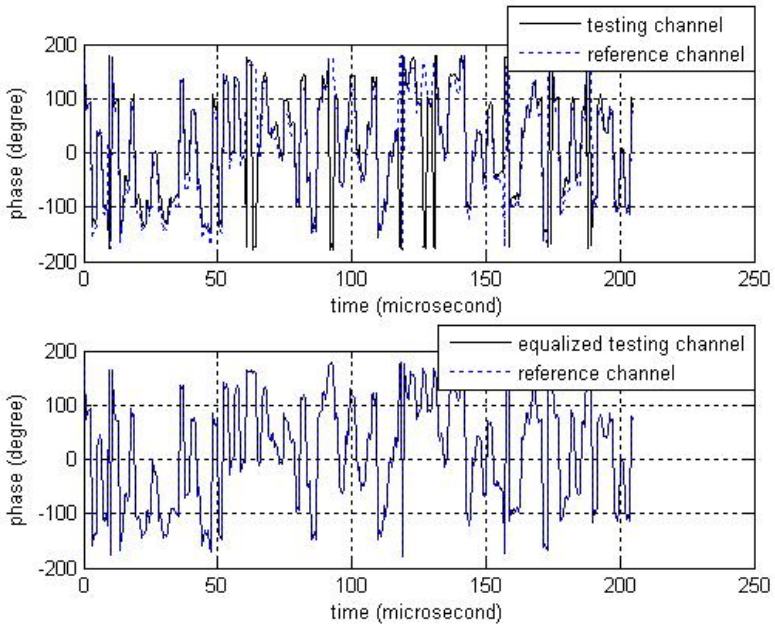

Figure 19; Phase plot (at input SNR=40 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

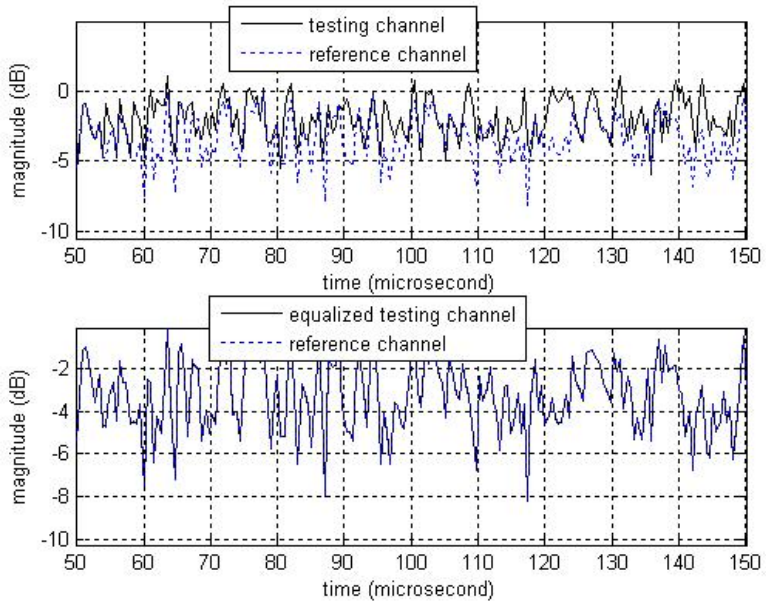

Figure 20; Zoomed-in magnitude plot (at input SNR=40 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

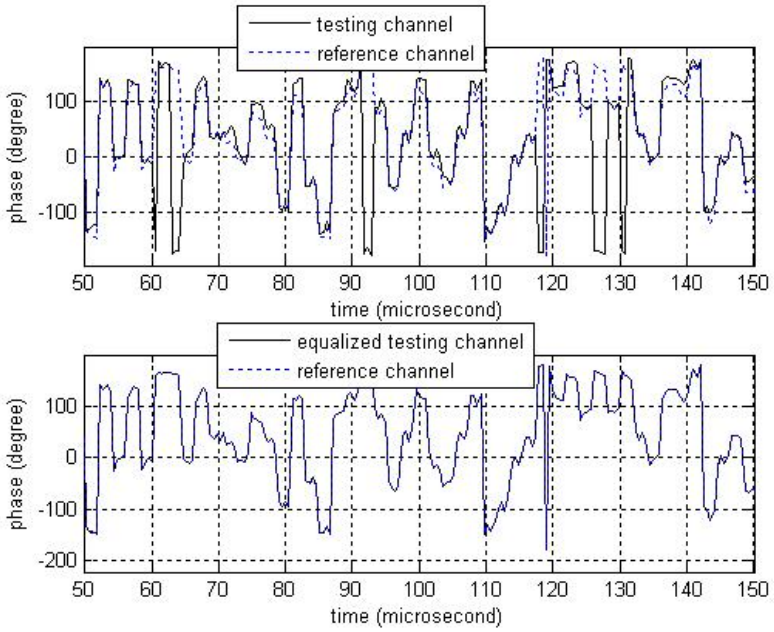

Figure 21; Zoomed-in phase plot (at input SNR=40 dB) of

Top: reference and testing channel outputs

Bottom: reference and equalized testing channel outputs.

5. Discussion and Conclusions

In this paper, a channel equalization algorithm was developed compensating for the mismatches between the reference and testing channels of a phased array radar. By stacking the testing channel output as a Toeplitz matrix, the equalizer output is the product of the Toeplitz matrix and the impulse response function of the equalizer, with no need for the convolution operation. The equalizer is determined using the least-squares error (LSE) criterion. Geometrically, the equalizer output is the projection of the reference channel data onto the columns of the equalization matrix, with the latter being a nonlinear function of the testing channel data. Using singular value decomposition, the equalization matrix is the product of a unitary matrix, a diagonal matrix, and the Hermitian transpose of the unitary matrix. The

columns of the unitary matrix are eigenvectors of the equalization matrix patched by some linearly independent orthonormal vectors, due to the rank deficiency and diagonal matrix elements are either ones or zeros. The nature of the channel equalization formulation results in the squares error, instantaneous correlation coefficient, and cancellation ratio (CR) metrics expressed in closed form. Extensive Monte Carlo simulations were used to evaluate the performance of the proposed equalizer, using CR as the main assessment metric. In particular, it was found that at higher input signal-to-noise ratio (ISNR) $(\geq 50\,\text{dB})$ the CR improves consistently with increasing equalizer length while the window size is fixed. The equalizer attains almost perfect equalization when it reaches 17.5 % of the window size. At 10 dB ISNR even though the CR improves consistently with respect to increasing equalizer length, the equalizer cannot reach perfect equalization. Analogously, at higher ISNR increasing the window size to 80% of the equalizer length (which is fixed) asymptotically improves the CR. In addition, the CR attains perfect equalization when the window size is exactly equal to the equalizer length. Once the window size goes beyond the equalizer length, the CR degrades noticeably. A similar behavior is exhibited at low ISNR. However, the CR never reaches perfect equalization. It was also shown that over the course of increasing ISNR from 5 dB to 40 dB, while all other system parameters remain fixed, the CR (in dB scale) improves linearly and is always -10 dB lower (or better) than the ISNR. No significant improvement (in CR) occurs beyond ISNR levels of 40 dB. When both the amplitude and phase of the simulated data are used, it is evident that good values of the CR ($\leq -40\,\text{dB}$) produce equalized output responses, which are almost a perfect replica of the reference channel output, despite the evident disparities between the direct and reference channel outputs. The simulation results demonstrate that an equalizer with good values of CR or Correlation Coefficient CC produces an equalized testing channel output that matches the reference channel output. Conversely, progressive degradation in CR or CC affects the resulting equalized testing channel output when compared to the reference channel output. The simulation results agree closely with the theoretical work as described in Section 3. The work in this paper show the importance of channel equalization, as well as the effect of system parameters, in our ability to produce matched equalized testing channel outputs. The paper also points us to potential new directions for investigation. Firstly, since larger window size (with respect to fixed equalizer length) degrades the CR, it seems that data in a snapshot may provide more information if partitioned into sub-snapshots. Then, the equalization process can be implemented on shorter sub-snapshots rather than a single long snapshot. In this case, the tradeoff between CR improvement and computational complexity will need to be carefully studied. Secondly, the proposed algorithm, in its current form, uses batch processing only, i.e., it processes one frame of data at a time and then discards the processed data prior to moving to the next frame. A recursive technique, which can make use of data in previous snapshots may potentially provide better solutions. The computational complexity of such approach as well as its potential benefit in improving the CR will need to be carefully studied.

6. Appendix

A Proof of Proposition 3.2

According to (3.2), $\mathbf{A}^2(\mathbf{x}) = \mathbf{A}(\mathbf{x})$ and so $\mathbf{A}(\mathbf{x})$ is idempotent, with its eigenvalues either 1 or 0. The next interest is the precise number of ones amid the eigenvalues. The singular value decomposition is deployed to factor the $n_{\text{eq}} \times n_s$ matrix $\boldsymbol{\Psi}^H(\mathbf{x})$

$$\boldsymbol{\Psi}^H(\mathbf{x}) = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H \tag{a1}$$

where $\mathbf{U}^{n_{\text{eq}} \times r}$ and $\mathbf{V}^{n_s \times r}$ are unitary matrices: $\mathbf{U}^H\mathbf{U} = \mathbf{V}^H\mathbf{V} = \mathbf{I}^r$; $\mathbf{I}^r$ denotes an $r \times r$ identity matrix.

This changes the equalization matrix (3.2) as

$$\begin{aligned}\mathbf{A}(\mathbf{x}) &= \mathbf{V}\boldsymbol{\Lambda}\mathbf{U}^H\left(\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H\mathbf{V}\boldsymbol{\Lambda}\mathbf{U}^H\right)^{-1}\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H \\ &= \mathbf{V}\boldsymbol{\Lambda}\mathbf{U}^H\left(\mathbf{U}\boldsymbol{\Lambda}^2\mathbf{U}^H\right)^{-1}\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H.\end{aligned} \tag{a2}$$

In view of $\mathbf{U}$ not being a square matrix, its inverse cannot be readily found. This hurdle is circumvented by inserting $(n_{\text{eq}} - r)$ additional arbitrary, orthonormal, vectors $\left\{\boldsymbol{\varsigma}_k^{n_{\text{eq}}\times 1}, k = 1, \cdots, n_{\text{eq}} - r\right\}$. The augmented matrix is $\tilde{\mathbf{U}} = \left[\mathbf{U} \quad \boldsymbol{\varsigma}\right]$, where $\boldsymbol{\varsigma} = \left[\boldsymbol{\varsigma}_1 \quad \cdots \quad \boldsymbol{\varsigma}_{n_{\text{eq}}-r}\right]$. Clearly, $\tilde{\mathbf{U}}^H\tilde{\mathbf{U}} = \tilde{\mathbf{U}}\tilde{\mathbf{U}}^H = \mathbf{I}^{n_{\text{eq}}}$. The same procedure enlarges $\mathbf{V}$ to become $\tilde{\mathbf{V}}$: $\tilde{\mathbf{V}} = \left[\mathbf{V} \quad \boldsymbol{\psi}\right]$, where $\boldsymbol{\psi} = \left[\boldsymbol{\psi}_1^{n_s\times 1} \quad \cdots \quad \boldsymbol{\psi}_{n_s-r}^{n_s\times 1}\right]$.

Also, $\boldsymbol{\Lambda}$ is reconstructed as $\tilde{\boldsymbol{\Lambda}} = \begin{bmatrix} \boldsymbol{\Lambda} & \mathbf{0}_a^{r\times(n_s-r)} \\ \mathbf{0}_b^{(n_{eq}-r)\times r} & \mathbf{0}_c^{(n_{eq}-r)\times(n_s-r)} \end{bmatrix}$.

With the aid of the identity

$$\begin{aligned} \tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{V}}^H &= \begin{bmatrix} \mathbf{U} & \boldsymbol{\varsigma} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda} & \mathbf{0}_a \\ \mathbf{0}_b & \mathbf{0}_c \end{bmatrix} \begin{bmatrix} \mathbf{V} & \boldsymbol{\psi} \end{bmatrix}^H \\ &= \begin{bmatrix} \mathbf{U}\boldsymbol{\Lambda} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V} & \boldsymbol{\psi} \end{bmatrix}^H = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H, \end{aligned} \quad \text{(a3)}$$

an alternative representation of $\boldsymbol{\Psi}^H(\mathbf{x})$ is

$$\boldsymbol{\Psi}^H(\mathbf{x}) = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^H = \tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{V}}^H \quad \text{(a4)}$$

This is applied to the first line of (a2)

$$\begin{aligned} \mathbf{A}(\mathbf{x}) &= \tilde{\mathbf{V}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{U}}^H(\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{V}}^H\tilde{\mathbf{V}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{U}}^H)^{-1}\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{V}}^H \\ &= \tilde{\mathbf{V}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{U}}^H(\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}^{-2}\tilde{\mathbf{U}}^H)\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{V}}^H \\ &= \tilde{\mathbf{V}}(\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Lambda}}^{-2}\tilde{\boldsymbol{\Lambda}})\tilde{\mathbf{V}}^H \end{aligned} \quad \text{(a5)}$$

In the equation above, $\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}^2\tilde{\mathbf{U}}^H$ is invertible because, unlike before $\tilde{\mathbf{U}}$ is a square matrix. Also, in $\tilde{\boldsymbol{\Lambda}}^{-1}$ the diagonal elements are equal to the reciprocals of the corresponding non-zero elements in $\tilde{\boldsymbol{\Lambda}}$, and zero otherwise. The same interpretation leads to $\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Lambda}}^{-2}\tilde{\boldsymbol{\Lambda}}$ a diagonal matrix whose leading r diagonal elements are one and the remaining $n_s - r$ diagonal elements are zero.

Equating (a5) and Proposition 3.1, yields $\boldsymbol{\xi} = \tilde{\mathbf{V}}$, $\boldsymbol{\Sigma} = \tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{\Lambda}}^{-2}\tilde{\boldsymbol{\Lambda}}$. The distribution of the diagonal elements of $\boldsymbol{\Sigma}$ (hence the eigenvalues of $\mathbf{A}(\mathbf{x})$) is as stated in the previous paragraph.

References

1. Reed IS, Mallet JD, Brennan LE. Rapid convergence rate in adaptive arrays. IEEE Trans. Aeros.Elect. Syst. 1974; AES-10 (6): 853-863.
2. Kelly EJ. An adaptive detection algorithm. IEEE Trans. Aeros. Elect. Syst. 1986; AES-22(1): 115-127.
3. Godara LC , Cantoni A. Analysis of constrained LMS algorithm with application to adaptive beamforming using perturbation sequences. IEEE Trans. Ant. Propag 1986; AP-34(3): 368-379.
4. Chen YH, Chen CH. A new structure for adaptive broad-band beamforming. IEEE Trans. Ant. Propag 1991; AP-39: 551-555.
5. Griffiths LJ, Jim CW. An alternative approach to linearly constrained adaptive beamforming. IEEE Trans. Ant. Propag 1982; AP-30(1): 27-34.
6. Jablon NK. Steady state analysis of the generalized sidelobe canceller by adaptive noise cancelling techniques.IEEE Trans. Ant. Propag 1986; 34 (3): 330-337.
7. Yu KB, Murrow DJ. Adaptive digital beamforming for angle estimation in jamming. IEEE Trans. AES 2001. 2002; 37 (2): 508-523.
8. Oppenheim AV, Schafer RW. Discrete-time signal processing, Prentice Hall 1999.
9. TüChler M, Singer AC , Koetter R. Minimum mean squared error equalization using a priori information. IEEE Trans. Signal Processing 2002; 50 (3): 673-683.
10. Pierre DA. Optimization theory with applications. John Wiley 1969; 21 (3): 379-379.
11. Bellman R. Introduction to matrix analysis, second edition, McGraw-Hill. 1960; 56 (3): 225-226.
12. Scheffe H. The analysis of variance, John Wiley 1999; 43 (2): 233-234.
13. Monzingo RA, Haupt RL, Miller TW. Optimum array processing. IET Digital Library 2011 :81-149.