# Novel Approaches to Overloaded Array Processing

By

James Hicks

Dissertation submitted to the faculty of the

Virginia Polytechnic Institute & State University

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

in

Electrical Engineering

Approved:

______________________________

Dr. Jeffrey H. Reed, Chairman

______________________ ______________________

Dr. William H. Tranter Dr. Brian D. Woerner

______________________ ______________________

Dr. Warren L. Stutzman Dr. Christopher Beattie

August, 8, 2003

Blacksburg, Virginia

Keywords: Overloaded Array, Multi-user Detection, Antenna Array

# Novel Approaches to Overloaded Array Processing

James Hicks

(ABSTRACT)

An antenna array is overloaded when the number of cochannel signals in its operating environment exceeds the number of elements. Conventional space-time array processing for narrow-band signals fails in overloaded environments. Overloaded array processing (OLAP) is most difficult when signals impinging on the array are near equal power, have tight excess bandwidth, and are of identical signal type. Despite the failure of conventional beamforming in such environments, OLAP becomes possible when a receiver exploits additional signal properties such as the finite-alphabet property and signal excess-bandwidth. This thesis proposes three approaches to signal extraction in overloaded environments, each providing a different tradeoff in performance and complexity. The first receiver architecture extracts signals from an overloaded environment through the use of MMSE interference rejection filtering embedded in a successive interference cancellation (SIC) architecture. The second receiver architecture enhances signal extraction performance by embedding a stronger interference rejection receiver, the reduced-state maximum aposteriori probability (RS-MAP) algorithm in a similar SIC architecture. The third receiver fine-tunes the performance of spatially reduced search joint detection (SRSJD) with the application of an energy focusing transform (EFT), a complexity reducing front-end linear pre-processor. A new type of EFT, the Energy Focusing Unitary Relaxed Transform (EFURT) is developed. This transform facilitates a continuous tradeoff between noise-enhancement and error-propagation in an SRSJD framework. EFURT is used to study the role of this tradeoff for SRSJD receivers in a variety of signal environments. It is found that for the environments studied in this thesis, SRSJD enjoys an aggressive reduction in interference at the expense of possible noise-enhancement.

# ACKNOWLEDGMENTS

I would like to thank all those who have made this dissertation possible: my advisor, Dr. Jeffrey H. Reed for his support, encouragement, tutelage, and advice; his breadth and scope of knowledge about interference mitigation techniques was invaluable while scoping the research in this dissertation; Dr. Chris Beattie and Dr. William Tranter for numerous interesting hallway discussions, Brian Woerner for numerous insightful comments; Randall Nealy for infinite patience in the lab. I would like to provide special thanks to Louis Beex, who suggested using Gray and Markel's Spectral Flatness Measure (SFM) as a framework for interference rejection. Further, I would like to thank the graduate students at MPRG who have shared a legacy of outstanding research and whose stimulating discussions and ideas have contributed to this research including Tom Biedka, Max Robert, Bill Newhall, Jody Neal, Maruf Mohammad, Dr. Matt Valenti, and Dr. Neiyer Correal-Mendoza, Ran Gozali, Jay Tsai, Fakhrul Alam, Bruce Puckett, Bror Peterson, Saffet Bayram, and Rich Ertle. I would like to thank those in the industry who have provided valuable input into this dissertation: Ari Dimitriou and Daniel Sharp at Raytheon E-Systems; Keith Olds and Roger Peterson at Motorola. Most importantly, I would like to thank my wife, Rupashree Majali, whose love, support, and consideration have made this dissertation possible.

# CONTENTS

# FIGURES

# TABLES

# Chapter 1: INTRODUCTION

The use of linear Space-Time Processing (STP) for increased wireless system capacity has received considerable attention in the past decade, and significant advances have been made. However, the capacity increase achieved through the use of linear STP is bounded by the number of elements in the array [1]. An overloaded environment is defined as one in which the number of incident co-channel signals exceeds the number of antenna elements at the receiver. Overloaded Array Processing (OLAP) is usually most difficult when the interfering signals have nearly equal power and have little excess bandwidth, as is typical of narrow-band communication protocols. A similar situation exists in certain orthogonal frequency division multiplexed (OFDM) systems. In such environments, linear STAP fails as the array's degree of freedom is exceeded [1]. However, recently [9]-[18], multi-user detection techniques have evolved into OLAP techniques that spawn a myriad of surveillance and communications applications.

This thesis develops overloaded array processing algorithms with an emphasis on narrow-band signals. The question might arise: why focus on signal processing for narrowband signals? Firstly, narrow-band cellular standards are widely deployed and are expected to stay in existence beyond the next 10 years. In particular, developing countries have been expanding their existing cellular networks with second-hand narrow-band cellular infrastructure and 2.5G EDGE (Enhanced Data-rates for Global Evolution) and GPRS (General Packet Radio Service) upgrades. Secondly, it is possible that OLAP solutions for narrow-band signals can be extended to successful OLAP strategies for OFDM signals by applying a given OLAP algorithm on a per-sub-carrier basis. Since most communication standards apply some kind of multi-block interleaving, algorithms that do not exploit forward error correction coding (FEC) can help achieve latency requirements. Moreover, in surveillance applications, a ciphering scheme often inhibits FEC exploitation.

Multi-user detection applied to STAP for narrow-band signals has been investigated by many authors. The optimum multi-channel receiver was first investigated in [2]. In [4], [5], two different sub-optimal joint detection algorithms have been developed for detecting narrow-band signals with an antenna array. A sub-optimal algorithm that exploits error-correction coding is developed in and a variation is tailored to cellular systems in [43]. A block iterative, blind multi-user detection algorithm for the detection of synchronous PAM signals is proposed in [7]-[8]. The performance of this algorithm is analyzed for the overloaded case in [10]-[12].

All of the aforementioned joint detection algorithms either have complexities that increase exponentially with the number of jointly detected signals or require the exploitation of coding information. Hence, a need becomes apparent to develop signal extraction algorithms with a more manageable complexity. A search must be conducted for receiver architectures that can efficiently exploit signal properties such as the finite-alphabet property and excess bandwidth in an array processing setting. Such a study should address performance/complexity tradeoffs as they relate to the exploitation of various signal properties. Which signal properties provide the best performance? Which lead to the most computationally efficient implementations? These questions relate to the choice of receiver architecture, not the choice of adaptive implementation. Hence, in all circumstances we will assume that both the channel parameters and the noise statistics of every signal environment is known.

Since linear multi-user detectors and successive interference canceling (SIC) receivers have been well established in the CDMA literature, it is reasonable to consider these approaches first. In Chapter 2 of this thesis and [13], it is shown that for an environment where there are twice as many PAM signals as elements (an overloading factor of two) and each of these PAM signals have a sufficiently large excess bandwidth, even MMSE-based linear STAP can extract one or two signals from the environment. This surprising result motivates the use of a hybrid MMSE-SIC receiver. Simply put, the MMSE interference rejection receiver exploits signal excess bandwidth while SIC may be viewed as a crude way of exploiting the finite-alphabet property. The partial success of MMSE-SIC motivates an enhancement that is explored in Chapter 3 and [17]. This receiver replaces linear STAP with reduced state maximum apriori sequence estimation (RS-MAP) in an SIC framework. This so called MAP-SIC receiver has the potential to out-perform MMSE-SIC by exploiting not only excess bandwidth, but also the finite-alphabet property of the desired signal. Both SIC-based receivers leverage the ability of an interference rejection receiver to extract the first signal from an overloaded environment. The ability of an interference rejection receiver to do so is strongly dependent upon excess bandwidth. A need becomes apparent for an OLAP receiver that does not over-rely on the exploitation of excess bandwidth. Such a receiver must surely exploit differences between user's channels, but may also exploit the finite alphabet property. The later is a strong signal property that is central to almost all multi-user detection schemes [19].

The joint maximum likelihood receiver (JML) [19], is capable of exploiting the finite-alphabet property but, because its complexity increases exponentially with the number of received signals, it is impractical to implement. However, [16] describes a sub-optimal MUD algorithm called spatially reduced search joint detection (SRSJD) that attempts to approximate the JML receiver with a dramatic reduction in complexity. As in [7], it considers synchronous PAM signals with known channel parameters. A complexity reduction is achieved by exploiting the fact that when a phased array receives signals with low angle spread, signals closely spaced in Angle-of-Arrival (AOA) interfere with each other more than signals that are widely spaced. SRSJD forms a *spatial-trellis* which enables the use of reduced-state sequence estimation to extract all signals. It was found that with a uniform circular array, SRSJD could extract twice as many QPSK signals as elements with commercial voice quality.

The spatial-trellis concept enables novel Multi-user detection receivers that are ideally suited to OLAP. However, SRSJD by itself still has two major limitations that inhibit its practical application. The first limitation of SRSJD's original formulation is that it can only operate in environments with a select few array geometries; and of these geometries, the algorithm's performance tends to favor arrays with impracticably spaced elements (less than a quarter wavelength apart). This limitation stems from the use of an inflexible orthogonal beam-forming scheme. This beamforming scheme preserves the maximum-likelihood criterion by not coloring background noise.

Since SRSJD is an approximation of the Joint Maximum-Likelihood Estimator (JMLE), it is reasonable to consider beams that preserve the maximum likelihood criterion. However, using non-orthogonal beam-formers has the potential to suppress multi-access interference (MAI). Since SRSJD is a type of decision feedback receiver, suppressing MAI can potentially reduce its susceptibility to error-propagation. However, an overly aggressive reduction in MAI can introducing noise coloring, which in turn, can lead to higher symbol error rates. To understand the noise-coloring hazard, recall the most fundamental maximum likelihood estimator of all: a QPSK receiver in additive white Gaussian noise (AWGN). The signal space observed by a QPSK receiver is illustrated in Figure 1.1(a). Here, a QPSK alphabet is observed in AWGN. The effect of AWGN can be visualized as a cloud around each signal-point. Probability density function or PDF-contours are super-imposed on each cloud. As every communication engineer learns in an intro to digital communications class, the receiver that minimizes the symbol error rate, the maximum-likelihood (ML) receiver, is the receiver that guesses the signaling point that is closest to an observation. The decision boundaries that define the decision regions for each signaling point are illustrated as dashed lines in Figure 1.1. If the signal to noise ratio is large enough (i.e. if the cloud is small enough) received samples seldom cross over decision boundaries, resulting in an acceptable symbol-error rate. The "whiteness" of the noise can be recognized through the symmetry of the noise PDF-contours; they are perfect circles. However, when the receiver experiences noise that is colored across I/Q inputs, the noise PDF-contours become ellipsoidal as in Figure

1.1(b). If the ellipsoids become too elongated in the wrong direction, received samples will more frequently cross decision boundaries resulting in higher symbol error rates.

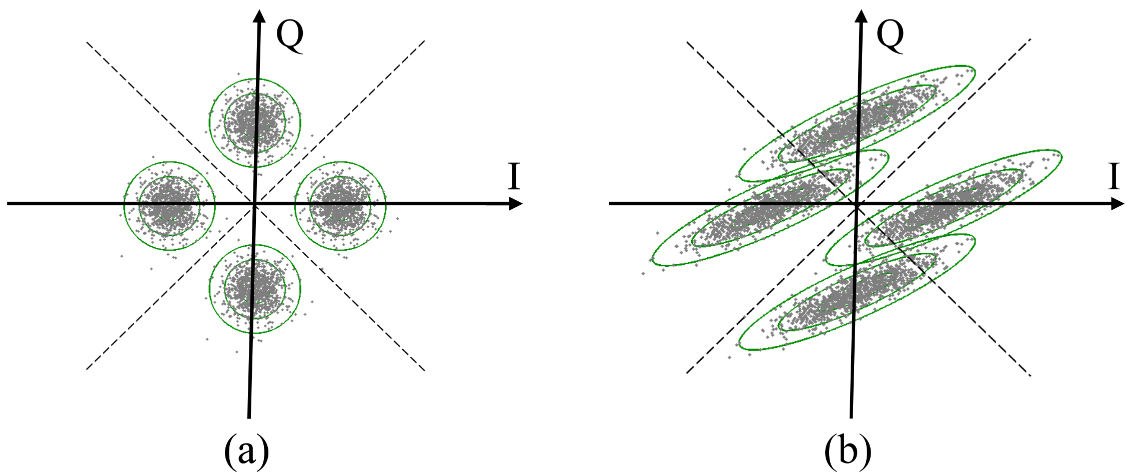


Figure 1.1: The effect of noise coloring on symbol error rate. (a) I/Q signal space for a QPSK signal in AWGN. (b) I/Q signal space for a QPSK signal in colored noise (noise correlated across in-phase and Quadrature components.

Chapter 4 investigates a family of beam-forming schemes that facilitate a tradeoff between noise enhancement (e.g. coloring) and multi-access interference rejection in an SRSJD framework. These so called Energy Focusing Transforms are chosen to accentuate a natural trellis-oriented structure in the spatial channel. Simulation results demonstrate that by relaxing noise-enhancement requirements, Energy Focusing Transforms enable SRSJD to achieve an order of magnitude improvement in symbol error rate over the approach in [18].

The second limitation of SRSJD is that it has only been developed for symbol-synchronous reception. That is, the formation of the spatial-trellis relies on the fact that all signals are received with symbols aligned in time and no signal experiences any ISI. However, it is expected that fine tuning SRSJD's performance will eventually inspire OLAP algorithms for asynchronous environments.

# Chapter 2: OVERLOADED ARRAY PROCESSING WITH MMSE-SIC

## 2.1 INTRODUCTION

In [17], it is shown that MMSE beam-forming can extract most signals from an overloaded environment if individual antenna elements have carefully designed, narrow beam-patterns. This study analyzes the case when narrow beam-patterns are not available at the receiver. Since linear multi-user detectors and successive interference canceling (SIC) receivers have been well established in the CDMA literature, it is reasonable to consider these approaches first. This chapter evaluates the ability of linear space-time processing to extract symbol-asynchronous linear pulse-amplitude modulated (PAM) signals with RRCOS pulse shaping (35%-75% rolloff) in an overloaded environment. The former rolloff value matches that of IS-136, while the 75% rolloff value was chosen because it is sufficiently high enough to show a dependence upon excess bandwidth. It is demonstrated that while linear space time processing may fail to extract most signals in overloaded environments, it can often extract at least one signal by exploiting both the spatial separation and the excess bandwidth of co-channel interferers. The performance of STP is quantified in terms of ranked output SINR statistics, compiled through the use of semi-analytic simulation. The effect of signal excess-bandwidth, adjacent channel interference, the number of users, array input SNR, and array type are considered. Since the channel parameters of all users are assumed random, the achievable output SINR is also a random variable (dependent on random time-delays, phases, and possibly fading amplitudes). Confidence intervals provide a framework for easily visualizing the range of achievable output SINR. Channel parameters of all signals are assumed known.

## 2.2 Signal Model

Consider $D_u$ linearly pulse shaped signals with pulse shape, $p(t)$, and symbol period, $T_s$, impinging on an $M$ element antenna array. Let the $d^{th}$ user arrive at the array with a time delay, $\tau_d$, and constant phase, $\phi_d$. Assume the delays of each user are independent and uniformly distributed in the range $\left[-T_s/2, T_s/2\right]$, and the phases are uniformly distributed in the range, $[0, 2\pi]$ but the carrier differences between co-channel signals are negligible. Further, assume that if any received signal is subject to multi-path, then it results in flat-fading.

It assumed that each element is front end filtered with a "brick-wall" anti-aliasing filter and over-sampled with over-sampling factor, $Q_0$. Then the received signal can be modeled as follows

$$\underline{r}_0[k] = \underline{r}_0\left(\frac{kT_s}{Q_o}\right) = \sum_{d=1}^{D_u} \underline{x}_d[k] + \underline{z}^{(0)}[k] \tag{2.1}$$

where

$$\underline{x}_d[k] = \underline{a}_d \sum_l s_d[l] g_d^{(0)}[k - lQ_0] \tag{2.2}$$

Here, $\underline{r}_0(t)$ is the $M \times 1$ vector whose $m^{th}$ element is the continuous time, complex envelope of the signal received at the $m^{th}$ antenna, $\underline{x}_d[k]$ is the $d^{th}$ user's signal component of $\underline{r}_0[k]$, $\underline{z}^{(0)}[k]$ is a vector of complex Gaussian noise. The $d^{th}$ user transmits a sequence of symbols, $s_d[l]$, over a channel with discrete equivalent impulse response, $g_d^{(0)}[k]$ and spatial signature, $\underline{a}_d$. The $d^{th}$ user's temporal impulse response is $g_d^{(0)}[k]$, which collectively models the phase, $\phi_d$, delay, $\tau_d$, and pulse-shape, $p(t)$, and can be expressed as

$$g_d^{(0)}[k] = e^{j\phi_d} p\left(\frac{kT_s}{Q_0} - \tau_d\right) \tag{2.3}$$

In this study, we will consider root-raised cosine (RRCOS) pulse shaping filters with 35%, and 75% roll-offs. A closed form expression for $p(t)$ can be found in [27]. For these roll-off factors, an over-sampling factor of $Q_0 = 2$ models all of the spectral features of interest. For the purpose of simulation, $g_d^{(0)}[k]$ can be truncated to $L_g - 1 = 6$ symbol periods and sufficiently delayed to assure causality [27]. Note that although we have assumed that each user's channel is not frequency-selective, the pulse shape itself introduces ISI. In the overloaded asynchronous case, this ISI cannot be easily removed with Nyquist matched filtering.

In this study, we will consider two spatial signatures. Firstly, we will consider the case when all co-channel signals impinge on an $M$ element uniform circular array (UCA) with AOAs, $\theta_d = (d-1)2\pi/D_u$, equally spaced in azimuth. The steering vector for an M element UCA is

$$\begin{aligned} \underline{a}(\theta_d) &= [a_1 \quad \cdots \quad a_M]^T \\ a_m &= \exp\left(-j\kappa_R \sin\left(\theta_d - \frac{2\pi(m-1)}{M}\right)\cos(\varepsilon_d)\right) \end{aligned}$$

where the constant $\kappa_R = 2\pi R_a / \lambda$, $\theta_d$ is the azimuthal angle of arrival (AOA); $\varepsilon_d$ is the elevation (or depression) angle; and $\lambda$ is the wavelength of the carrier frequency [26]. For simplicity, we will assume that $\varepsilon_d = 0$. Secondly, we will consider the case of quasi-static uncorrelated Rayleigh fading, where over the length of a detection window, the received amplitude and phase of the $d^{th}$ signal at the $m^{th}$ antenna is a complex Gaussian random variable. Hence, the $d^{th}$ user's spatial signature is modeled as a vector of independent complex Gaussian (ICG) random variables.

Background thermal noise, and adjacent channel interference (ACI) are modeled together with the $M \times 1$ complex Gaussian noise vector $\underline{z}_0[k]$. Figure 2.2 depicts the power spectral densities (PSD) of a single co-channel signal and a single adjacent channel interferer (ACI) for an over-sampling factor of $Q_0 = 2$. A symbol rate of 24.3 kHz and a channel spacing of 30kHz are assumed. Both signals have 35% RRCOS pulse shapes truncated to $L_g = 9$ symbols. Pulse truncation manifests itself in sidelobes of the PSD (ideal RRCOS pulse shapes have none). Note that because the assumed carrier-separation exceeds the Nyquist frequency, part of the ACI's PSD cut off by the front-end anti-alias filtering. However, for the purposes of performance assessment, it is only necessary to model ACI in the band edge of the desired signal. Each ACI is modeled as $25^{th}$ order moving average complex Gaussian noise process. The PSD of the modeled noise process is obtained with a Parks-McClellan fit to the visible portion of the RRCOS pulse shape. For the case of a UCA, the ACI spatial signatures are assumed to be the same as users in the desired channel. For the case of quasi-static fading, the ACI signatures are modeled as ICG vectors independent from the desired users with a power 15dB down from the desired channel.

For the given signal model, it may seem that space-time processing cannot achieve any additional interference rejection over space-only processing. However, by exploiting the excess bandwidth induced by the RRCOS pulse-shaping operation, temporal processing can increase the output signal to interference and noise ratio (SINR) [44]-[46]. This excess bandwidth is obscured by the implicit time-averaging of the PSD formulation [46]. The cyclo-stationary signal given in Equation (2.2) can be converted into a vector-wide-sense stationary random process by stacking sub-samples over a symbol interval as follows.

$$\begin{aligned}\underline{r}[n] &\triangleq \left[\underline{r}_0[nQ_0]^T \quad \cdots \quad \underline{r}_0[(n+1)Q_0-1]^T\right]^T \\ &= \sum_{d=1}^{D_u}\sum_{l=0}^{L_g-1} \underline{g}_d[l]s_d[n-l]+\underline{z}[n]\end{aligned} \tag{2.4}$$

where

$$\underline{g}_d[n] \triangleq \left[\underline{a}_d^T g_d^{(0)}[nQ_0] \quad \cdots \quad \underline{a}_d^T g_d^{(0)}[(n+1)Q_0-1]^T\right]^T \tag{2.5}$$

is the subsample-augmented vector impulse-response for the $d^{th}$ user.

## 2.3 SPACE-TIME PROCESSOR

Consider the linear space-time processor illustrated in Figure 2.1 where a delayed estimate of a symbol transmitted by the $d^{th}$ user, $s_d[n-\delta]$, is obtained from a linear combination of array signals collected over $L_w$ symbol intervals. The estimation lag, $\delta$, is a parameter left to the algorithm designer. A typical choice of estimation lag is $\delta = L_w/2$ and can improve the quality of estimates. An estimate of the delayed signal, $s_d[n-\delta]$, can be expressed as follows

$$\hat{s}_d[n-\delta] = y_d[n] = \underline{w}_d^H \underline{\tilde{r}}[n] \tag{2.6}$$

where

$$\underline{\tilde{r}}[n] \triangleq \left[\underline{r}[n]^T \quad \cdots \quad \underline{r}[n-L_w+1]^T\right]^T \tag{2.7}$$

is an $MQ_0L_w \times 1$ vector of time-reversed augmented array snapshots, and $\underline{w}_d$ is the space-time processor. In order to calculate a $\underline{w}_d$ that rejects co-channel interferers while enhancing a desired signal it is useful to construct the following linear channel model for $\underline{\tilde{r}}[n]$.

$$\underline{\tilde{r}}[n] = \tilde{\mathbf{G}}_{ST}\underline{\tilde{s}}[n] + \underline{\tilde{z}}[n] \tag{2.8}$$

where $\underline{\tilde{s}}[n]$ is the $D_u(L_g + L_w - 1)\times 1$ vector of transmitted signals such that $\left[\underline{\tilde{s}}[n]\right]_{lD+d} = s_d[n-l]$. For convenience let, $L_{gw} = L_g + L_w - 1$. The vector, $\underline{\tilde{z}}[n]$ is a vector constructed from $\underline{z}[n]$ according to (2.7). The $MQ_0L_w \times L_{gw}D_u$ matrix $\tilde{\mathbf{G}}_{ST}$, can be constructed as follows. Let $\mathbf{G}_0$ be a $MQ_0 \times L_g D_u$ matrix whose $(lD+d)^{th}$ column is $\underline{g}_d[l]$. Then, $\tilde{\mathbf{G}}_{ST}$ is the block Toeplitz matrix formed from $L_w$ shifts of $\mathbf{G}_0$, shifting $D_u$ columns for each block.

It is useful to re-express Equation (2.8) with the signal component corresponding to $s_d[n-\delta]$ separated from the interfering symbols. Let $\underline{\tilde{g}}_d$ be the $(\delta D_u + d)^{th}$ column of $\tilde{\mathbf{G}}_{ST}$ and let $\tilde{\mathbf{G}}_{ST}^{(\bar{d})}$ be the $MQ_0L_w \times (L_{gw}D_u - 1)$ matrix formed from $\tilde{\mathbf{G}}_{ST}$ with $\underline{\tilde{g}}_d$ removed. Then

$$\begin{aligned} \underline{\tilde{\tilde{g}}}_d &= \left[\underline{g}_d[L_g-1]^T \quad \cdots \quad \underline{g}_d[0]^T\right]^T \\ \underline{\tilde{g}}_d &= \left[\underline{0}^T_{(\delta-(L_g-1))MQ_0} \quad \underline{\tilde{\tilde{g}}}_d^T \quad \underline{0}^T_{(L_w-(\delta+1))MQ_0}\right]^T, \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} \underline{\tilde{r}}[n] &= \underline{\tilde{g}}_d s_d[n-\delta] + \underline{\tilde{i}}_{\bar{d}}[n], \\ \underline{\tilde{i}}_{\bar{d}}[n] &= \tilde{\mathbf{G}}_{ST}^{(\bar{d})}\underline{\tilde{s}}_{\bar{d}}[n] + \underline{\tilde{z}}[n], \end{aligned} \tag{2.10}$$

where $\underline{\tilde{s}}_{\bar{d}}[n]$ is the same as $\underline{\tilde{s}}[n]$ with $s_d[n-\delta]$.

With the linear signal models of Equations (2.8) and (2.10) in hand, elegant expressions can be obtained for the MMSE space-time processor and the output SINR that it achieves. Let $\mathbf{\Phi}_{\tilde{z}\tilde{z}} \triangleq E\left[\underline{\tilde{z}}[n]\underline{\tilde{z}}[n]^H\right]$ be the noise auto correlation matrix. Let

$$\begin{aligned}\mathbf{\Phi}_{\tilde{r}\tilde{r}} &\triangleq E\left[\underline{\tilde{r}}[n]\underline{\tilde{r}}[n]^H\right] \\ &= \sigma_s^2 \tilde{\mathbf{G}}_{ST} \tilde{\mathbf{G}}_{ST}{}^H + \mathbf{\Phi}_{\tilde{z}\tilde{z}},\end{aligned} \tag{2.11}$$

$$\begin{aligned}\mathbf{\Phi}_{\tilde{i}\tilde{i}}^{(d)} &\triangleq E\left[\underline{\tilde{i}}_{\bar{d}}[n]\underline{\tilde{i}}_{\bar{d}}[n]^H\right] \\ &= \sigma_s^2 \tilde{\mathbf{G}}_{ST}^{(\bar{d})} \tilde{\mathbf{G}}_{ST}^{(\bar{d})H} + \mathbf{\Phi}_{\tilde{z}\tilde{z}}\end{aligned} \tag{2.12}$$

and

$$\begin{aligned}\underline{\phi}_{\underline{\tilde{r}},s_d[n-\delta]} &\triangleq E\left[\underline{\tilde{r}}[n]s_d[n-\delta]^*\right] \\ &= \sigma_s^2 \underline{\tilde{g}}_d\end{aligned} \tag{2.13}$$

be the received signal auto correlation matrix, the interference auto correlation matrix and cross-correlation vector respectively. Then, the SINR achieved by the $d^{th}$ processor is

$$\begin{aligned}SINR_d &\triangleq \frac{P_d}{P_I^d} = \frac{E\left[\left|\underline{w}_d^H \underline{\tilde{g}}_d s_d[n-\delta]\right|^2\right]}{E\left[\left|\underline{w}_d^H \underline{\tilde{i}}_{\bar{d}}[n]\right|^2\right]} \\ &= \frac{\sigma_s^2 \left|\underline{w}_d^H \underline{\tilde{g}}_d\right|^2}{\underline{w}_d^H \mathbf{\Phi}_{\tilde{i}\tilde{i}}^{(d)} \underline{w}_d}\end{aligned} \tag{2.14}$$

The processor that maximizes the MMSE criterion

$$MMSE_d \triangleq E\left[\left|s_d[n-\delta] - \underline{w}_d^H \underline{\tilde{r}}[n]\right|^2\right] \tag{2.15}$$

is the so called MMSE space-time processor given by [47]

$$\underline{w}_d = \mathbf{\Phi}_{\tilde{r}\tilde{r}}^{-1} \underline{\phi}_{\tilde{r}, s_d[n-\delta]} . \tag{2.16}$$

If all interfering signals are uncorrelated, the MMSE STP achieves the highest attainable SINR. It should be noted, that because the vector impulse response, $\underline{g}_d[n]$, is a function of several random variables, the maximum obtainable output SINR, $SINR_d$, of each user is a random variable.

## 2.4 MMSE-SIC

Figure 2.6 illustrates the proposed minimum mean squared error successive interference cancellation (MMSE-SIC) architecture. Figure 2.6 (a) illustrates the multi-stage structure of MMSE-SIC. The raw received signal $\underline{r}^{(0)}[n] = \underline{r}[n]$ is input to the first stage. The $p^{th}$ stage accepts $\underline{r}^{(p-1)}[n]$ as input, and outputs a signal estimate, $\hat{s}_{d_p}[n]$, for the $d_p^{th}$ user as well as a refined received signal, $\underline{r}^{(p)}[n]$. Figure 2.6(c) illustrates the processing performed at each stage. First, consider the $p = 1$ stage. An MMSE space-time interference rejection filter is calculated for each user. Let $d_1$ denote the user with the largest output SINR obtained through the linear space-time processor of (2.16). Then, the signal estimate, $\hat{s}_{d_1}[n]$, is calculated with an MMSE space-time linear processor. The symbol sequence, $\hat{s}_{d_1}[n]$ is then re-modulated for successive cancellation. Two types of cancellation have been proposed for SIC, hard and soft cancellation [19]. Both attempt to cancel a demodulated signal by re-modulating a sequence of symbol estimates. However, in the case of soft-SIC, “soft” symbol estimates are obtained directly from the output of a linear filter, while in the case of hard-SIC, “hard” symbol estimates are obtained by “property-mapping” soft symbol estimates to a known alphabet. Although, property mapped modulation requires differentially-coherent detection, it has been observed to far outperform cancellation with no property mapping.

Property mapping for $\pi/4$-DQPSK signals must be handled with care. Figure 2.1(b) illustrates the signal constellation of $\pi/4$-DQPSK symbols. At any given time index, *n*, a transmitted symbol, $\hat{s}_{d_1}[n]$ takes on a value from one of two constellation sets. Hence, for even time-indices, $\hat{s}_{d_1}[n]$ is mapped to the nearest symbol in the even constellation set (open-circles). For odd time indices, $\hat{s}_{d_1}[n]$, is mapped to nearest symbol in the odd constellation set (closed-circles). This *signal property mapping* produces a sequence of symbols, $\breve{s}_{d_1}[n]$ drawn from a finite alphabet. Finally, an estimate of the signal component for the $d_1^{th}$ user, $\hat{\underline{x}}_{d_1}[n]$, is obtained as follows.

$$\hat{\underline{x}}_{d_1}[n] = \sum_{l} \breve{s}_{d_1}[l] \underline{g}_{d_1}[n-l]. \tag{2.17}$$

Then, the estimated sequence, $\hat{\underline{x}}_{d_1}[n]$ is subtracted from the received signal.

$$\underline{r}^{(1)}[n] = \underline{r}^{(0)}[n] - \hat{\underline{x}}_{d_1}[n]. \tag{2.18}$$

Continuing in this manner, at each stage, the symbol sequence of a new user is estimated, property mapped, re-modulated, and subtracted from the input passed from the previous stage. At each stage, the MMSE processor is calculated assuming that users on previous stages are cancelled perfectly. MMSE-SIC continues for $D_u$ stages until a symbol sequence is estimated for each user.

## 2.5 SIMULATION RESULTS

The probability distribution of $SINR_d$ is difficult to calculate because it is a highly non-linear function of random channel parameters. Instead, it is useful to characterize the symmetric confidence interval,

$$\Pr(SINR_d^1 \le SINR_d \le SINR_d^2) = \alpha. \tag{2.19}$$

where $SINR_d$ is equally likely to fall above $SINR_d^1$ or below $SINR_d^2$. For a given environment, $SINR_d^1$, $SINR_d^2$ is estimated through semi-analytic Monte-Carlo simulation.

Simulations were conducted as follows. For each simulation trial, channel parameters were randomly generated as described in a previous section and the vector-impulse responses, $\underline{g}_d[n]$, were calculated, assuming an over-sampling factor of $Q_0 = 2$. In each simulation, a detection window of $L_w = 20$ symbols is applied. The channel parameters, $\tilde{\mathbf{G}}_{ST}$, $\tilde{\mathbf{G}}_{ST}^{(\bar{d})}$, $\tilde{\underline{g}}_d$ of equations (2.8), and (2.10) were constructed, and the channel statistics, $\mathbf{\Phi}_{\tilde{z}\tilde{z}}$, $\mathbf{\Phi}_{\tilde{r}\tilde{r}}$, and $\mathbf{\Phi}_{\tilde{i}\tilde{i}}^{(d)}$, $\underline{\phi}_{\underline{\tilde{r}},s_d[n-\delta]}$ are calculated according to Equations (2.11), (2.12), and (2.13) respectively. Then, for each user, the MMSE processor, $\underline{w}_d$, and the output $SINR_d$ are calculated from Equations, (2.16) and (2.14), respectively. After each trial, the output SINRs from all users are sorted from greatest to least and recorded. For each simulation, 1000 trials are run and the 95% confidence intervals are obtained from histograms of the sorted $SINR_d$.

First, consider the case of $D_u = 6$ 35% RRCOS pulse shaped users impinging on an $M = 4$ element array. Figure 2.3 illustrates the estimated 95% confidence intervals for several users over a range of input SINR. Separate SINR confidence intervals are plotted for the best, second best, and worst users. The range of attainable SINRs for the best and second best users overlap. Both achieve output SINRs above 5dB for SINRs greater than 15dB. However, the worst user never attains acceptable SINRs, even for high input SNRs. It is evident that in overloaded environments, a linear STP may not be able to extract most of the users, but it can extract one or two. Secondly, consider the case where $D_u = 8$ signals impinge on an $M = 4$ element array. Again, ICG signatures are assumed. Figure 2.4 compares the attainable output SINR for 35% and 75% RRCOS pulse shaped signals. Again the attainable output SINR for both cases overlap; however, more excess bandwidth results in greater output SINR.

As a final case, consider the case of 35% RRCOS pulse shaped signals impinging on an $M = 4$ element UCA with diameter $\lambda / 2$. The AOAs of interfering users are assumed to be equally spaced in azimuth over $360^0$. The output SINR for the best user is plotted in Figure 2. 5 compares the output SINR of the best user for the cases when $D_u = 6,8$. First, observe a much smaller variance in output SINR than the ICG case. This is because the ICG spatial signature model assumes a much larger variation in received signal power than the UCA model. For the case of $D_u = 6$ users, marginally acceptable output SINRs can be obtained for large signal to input noise ratios. Increasing the number of users present reduces the SINR by more than 2dB.

The probability of error for MMSE-SIC is difficult to calculate because each symbol sequence estimate is a highly non-linear function of both the received signal and random channel parameters. Instead, bit error rates have been estimated through the use of Monte-Carlo Simulation.

Two array types were considered: a $\lambda / 2$ diameter uniform circular array (UCA), and independent complex Gaussian (ICG) spatial signatures. Both arrays have $M = 4$ elements. For the case of the UCA, all signals are assumed to be equally spaced in azimuth over $360^0$. Also both 35% and 75% RRCOS signals are considered. For each simulated frame (1000 symbols) a new set of channel parameters (e.g. phases, delays, and spatial signatures) are generated. For all simulation cases, a detection window of $L_w = 40$ symbols and an oversampling factor of $Q_0 = 2$ are assumed. It is expected that a much smaller detection window is required. Symbol error rate statistics are recorded according to *cancellation order*. A frame error is said to occur when all users have at least one error in a frame. Simulations were run until at least 30 frame errors and 50 symbol errors occurred on each frame.

Figure 2.7 shows the performance of MMSE-SIC for $D_u = 8$ 35% RRCOS pulse-shaped signals impinging on an $M = 4$ element $\lambda / 2$ diameter UCA. Figure 2.8 shows MMSE-SIC's performance for the same number of users and signal type but with ICG spatial signatures. In both cases symbol error rates are poor. However, for digital voice applications, SERs just less than 10% may result in intelligible voice to a trained operator. The increase in SER as SNR increases from 15dB to 20dB is attributed to simulation error.

Figure 2.9 and Figure 2.10 show MMSE-SIC's performance for the same number of signals and array types as Figure 2.7 and Figure 2.8, respectively. However, in these figures, 75% RRCOS pulse-shaping is considered. In both cases, greater excess bandwidth results in a marginal improvement in SER.

Figure 2.11 and Figure 2.12 illustrate MMSE-SIC's performance for the same array types and signal types as Figure 2.7 and Figure 2.8, respectively. However, in this case the number of signals considered is reduced to $D_u = 6$. Reducing the number of interfering users results in a moderate SER improvement for the UCA case, and an order of magnitude improvement for the ICG case. Most notably, in the ICG case, SERs fall below the 1% threshold for commercial grade voice quality. Also as SNR increases, the SERs of all users converge.

Note for all the cases above, the MMSE-SIC receiver becomes interference-limited for input SNRs greater than 20dB. Also, in all cases, the first user enjoys a better SER than the last cancelled user. This is expected, as symbol errors in the first re-modulated and cancelled signal can cause symbol errors for subsequently estimated signals. However, even in the case of ICG signatures, the variation in SER across users is not drastic.

## 2.6 Conclusions

The application of linear space-time processing (STP) applied to overloaded-array processing (OLAP) is investigated through semi-analytic simulation. The performance of a linear STP is quantified in terms of output SINR for the best user. Confidence intervals for output SINR are obtained through semi-analytic simulation. Although, linear space-time processing (STP) may not be able to extract most signals in an overloaded environment, it can often extract one or two signals by exploiting the excess bandwidth of interfering signals. More excess bandwidth results in greater average output SINR. Quasi-static fading results in a larger variation of output SINR than the assumed uniform circular array model. However, in both environments, linear STP can frequently extract at least one signal with better than 5dB output SINR for 6 signals impinging on a 4 element array. This suggests that MMSE STP embedded into a successive interference cancellation framework can result in low-cost overloaded array processing receivers.

The performance of MMSE-SIC is also evaluated through simulation. Both a uniform circular array (UCA) and independent complex Gaussian (ICG) vectors are considered. For the same number of signals and pulse-shape types, ICG performs better than UCA because the large variation in received power implied by the ICG model allows for greater separation among interfering signals. Also, an increase in excess bandwidth results in better symbol error rate performance; however, this trend is more pronounced in the ICG case. For the $M = 4$ element UCA case, MMSE-SIC can deliver commercial voice grade quality only for a high excess bandwidth of 75%, and a low number of users, $D_u = 6$. In contrast, for the $M = 4$ element ICG case, MMSE-SIC can deliver commercial voice grade quality for six 35% RRCOS pulse shaped users, and near commercial voice grade quality for eight 75% RRCOS pulse shaped users. In this study, the channel parameters of all signals are assumed known. Channel estimation is an area for future research. Finally for UCA, more powerful signal processing structures than MMSE-SIC are required.

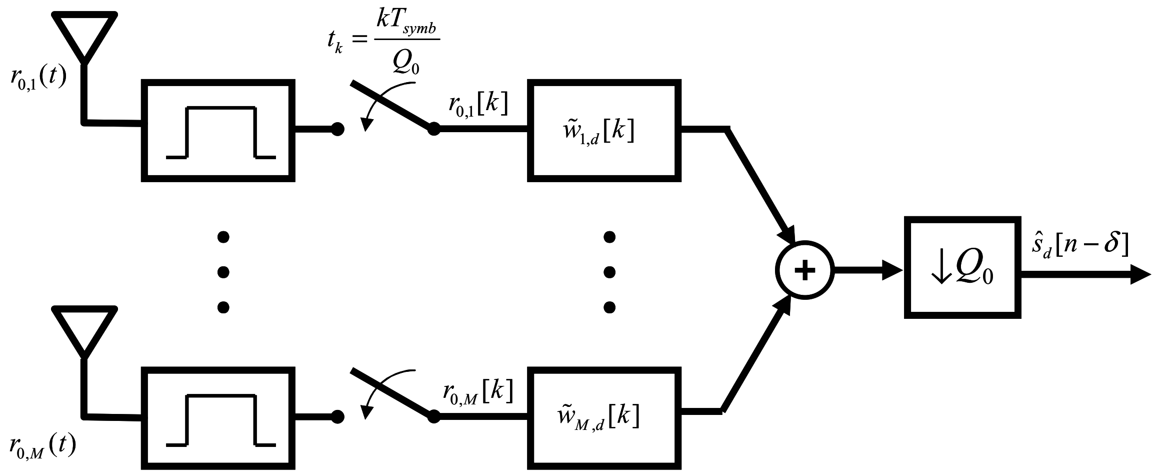


Figure 2.1: Linear Space-Time Processor

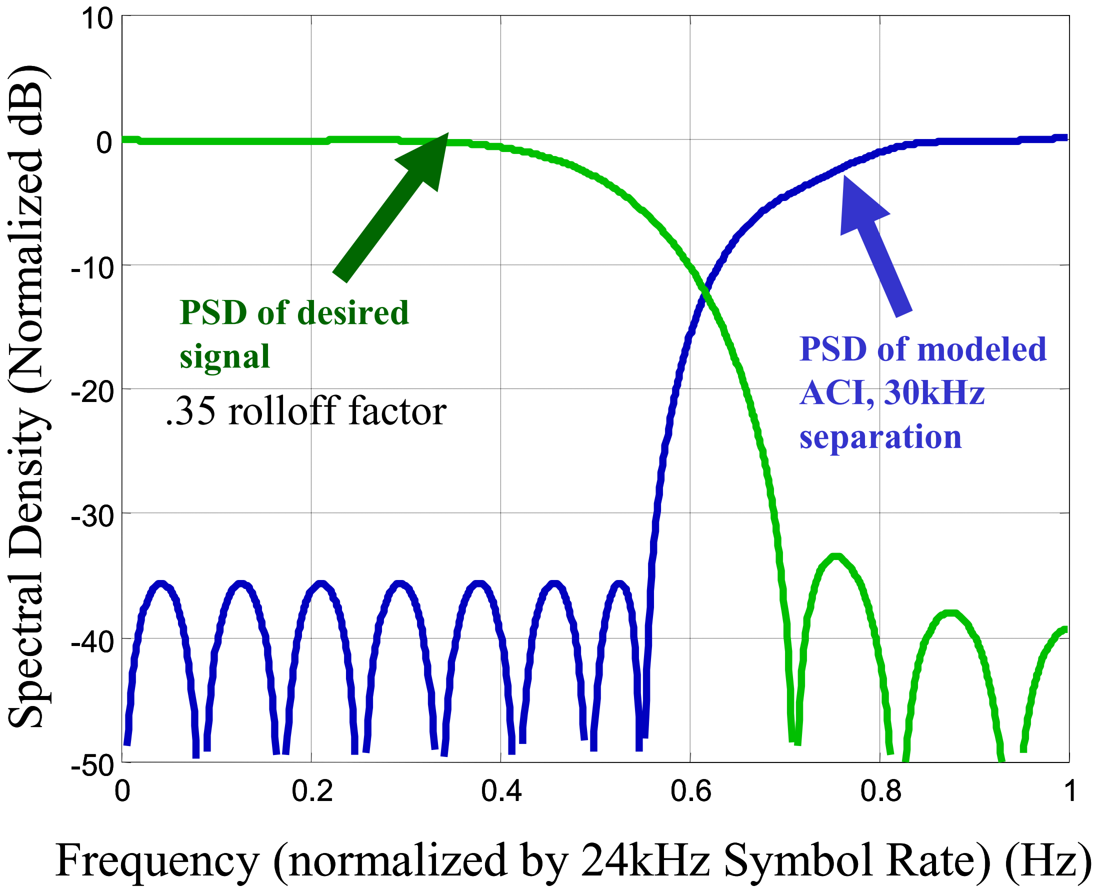


Figure 2.2: Spectral Representation of a signal on a desired channel and its adjacent channel interferer.

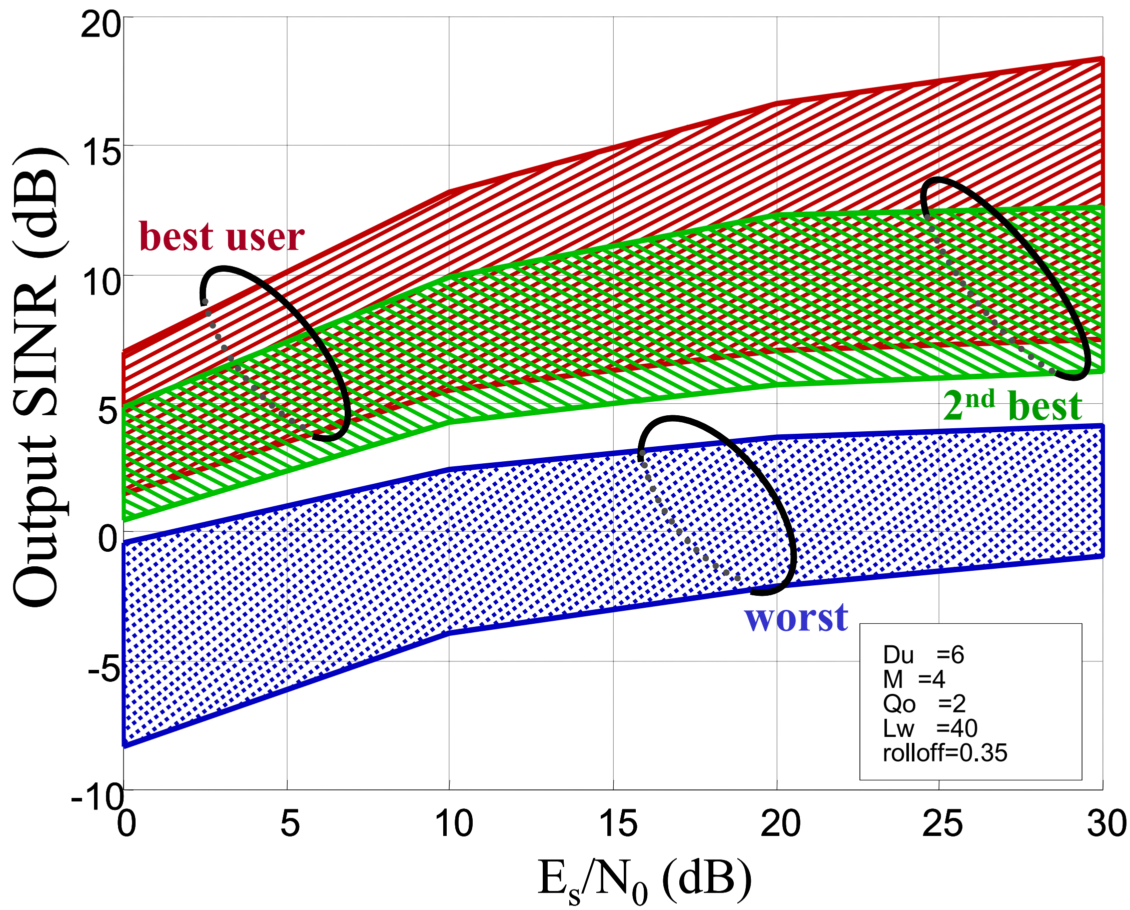


Figure 2.3: Variation of Output SINR over users, ICG signatures.

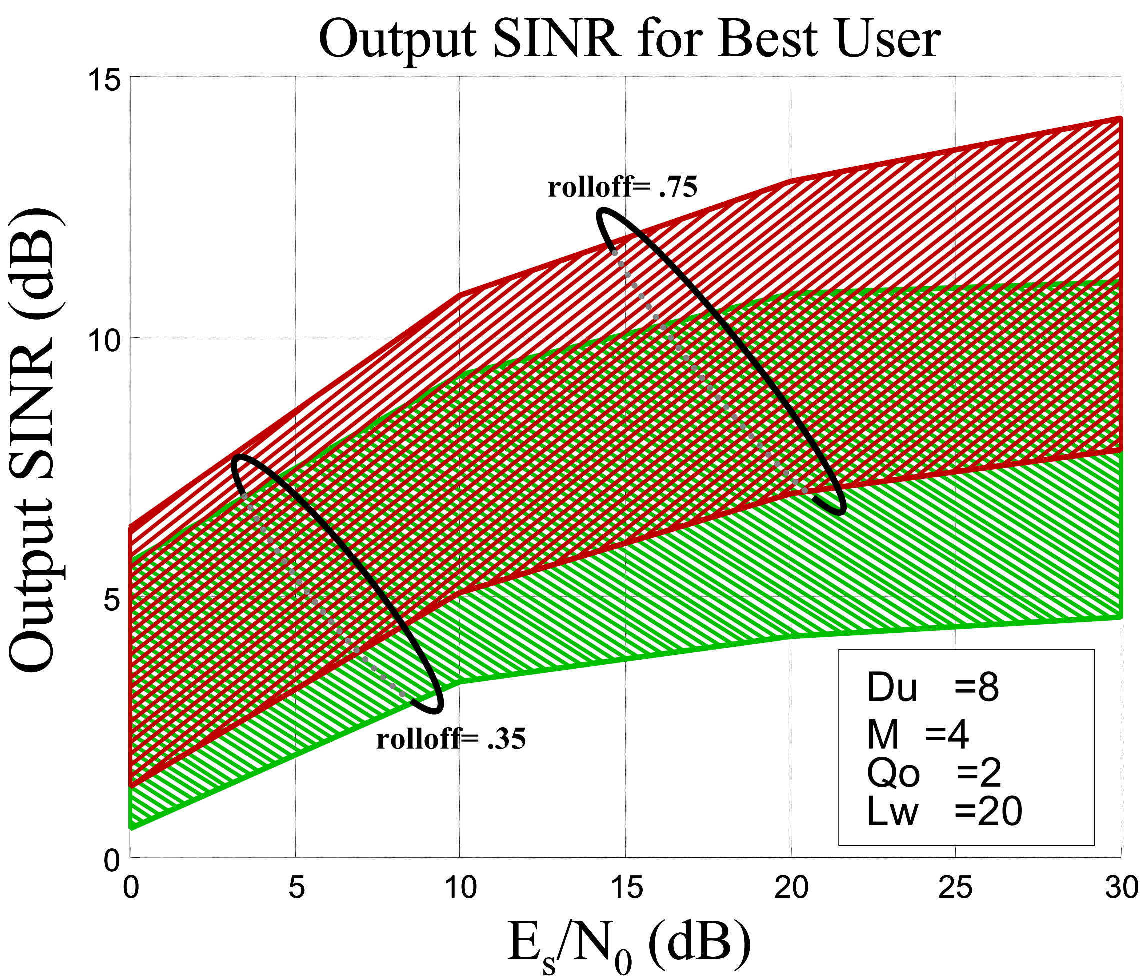


Figure 2.4: Effect of excess bandwidth on output SINR, ICG array.

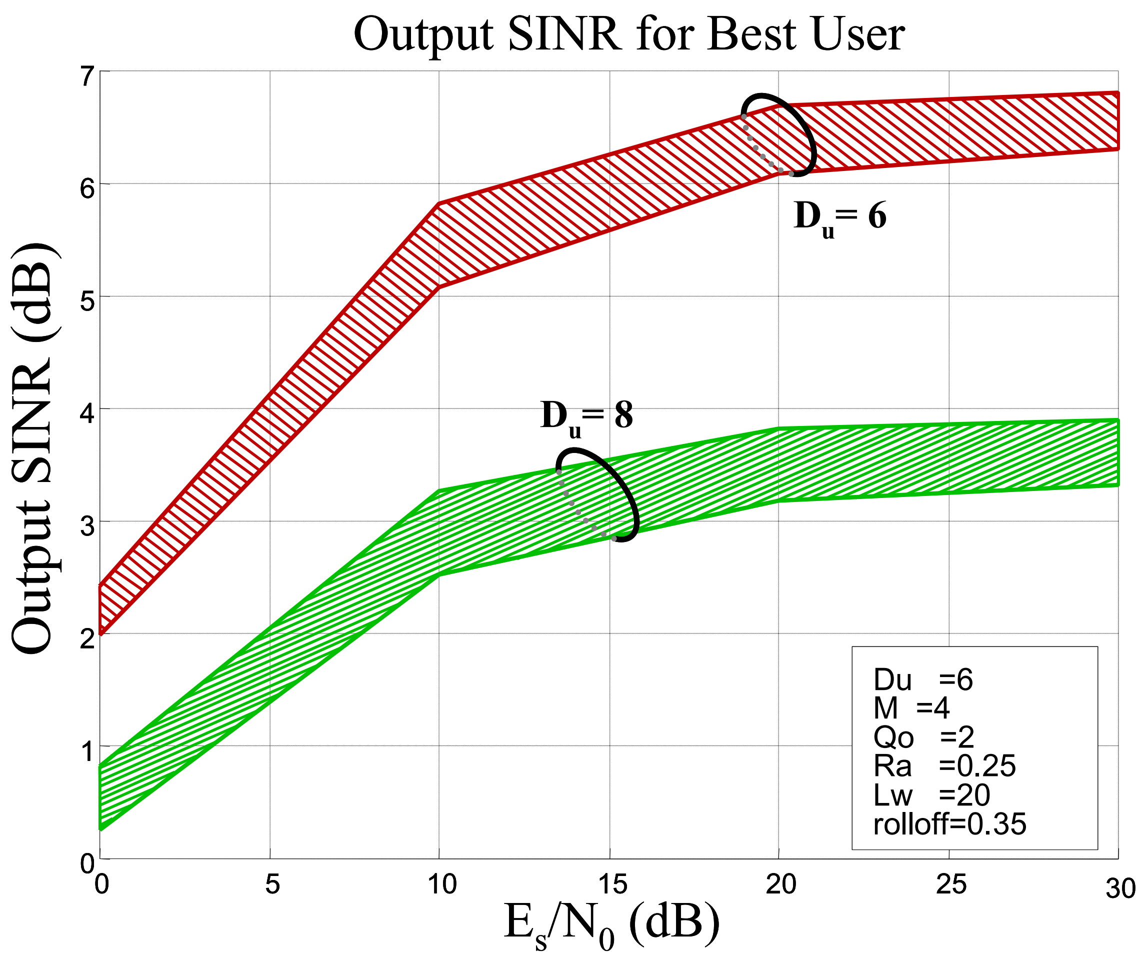


Figure 2.5: Effect of the number of users on output SINR, UCA.

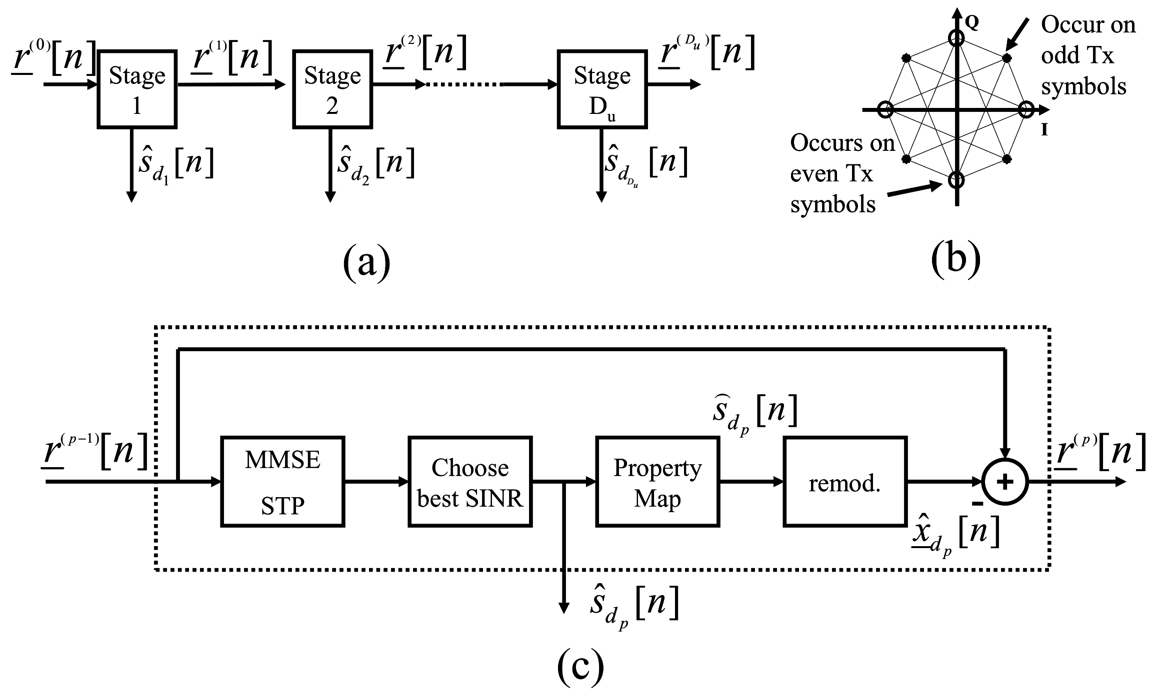


Figure 2.6: MMSE-SIC architecture. (a) multi-stage architecture. (b) signal constellation for property mapping operation (c) single-stage architecture.

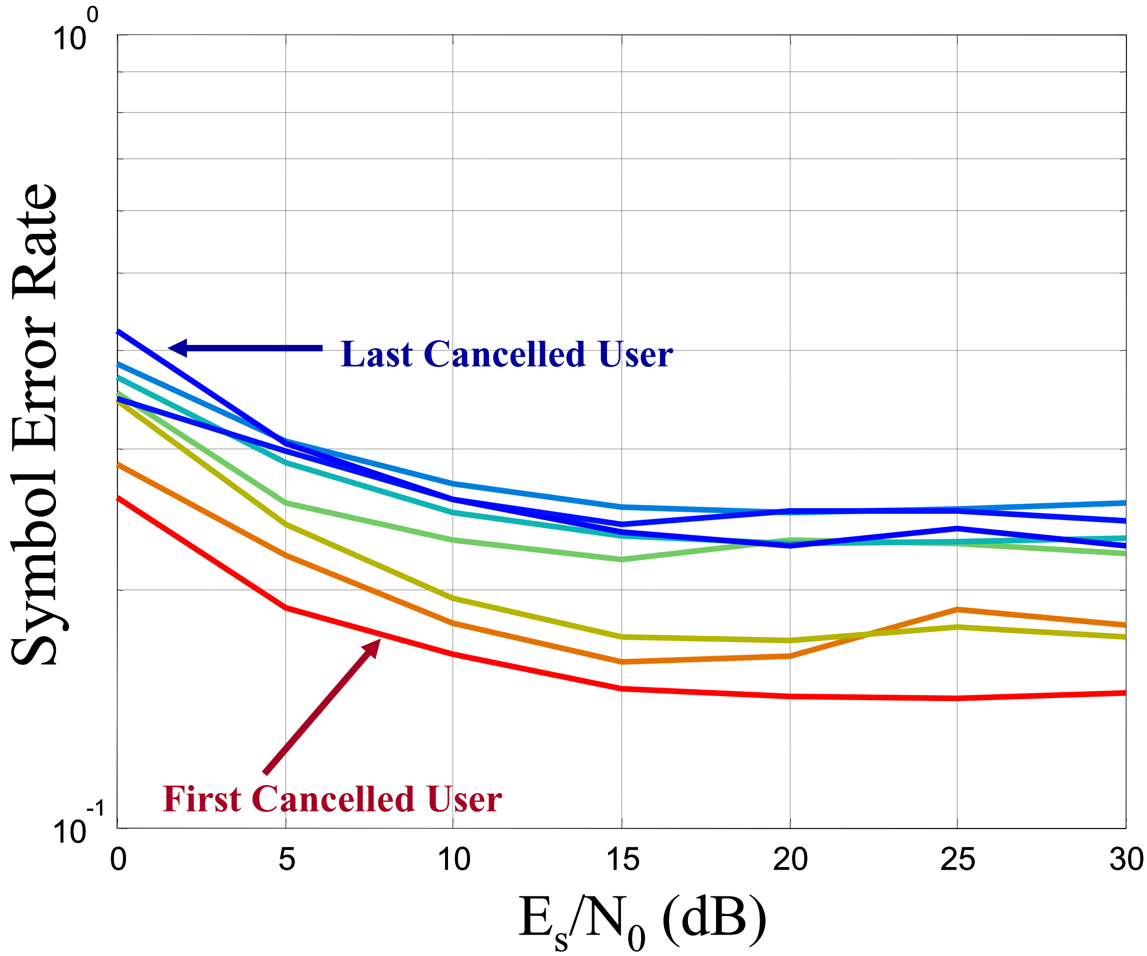


Figure 2.7: Symbol Error Rate for MMSE-SIC vs. Cancellation Order. $D_u = 8$, 35% rolloff, $M = 4$ element UCA.

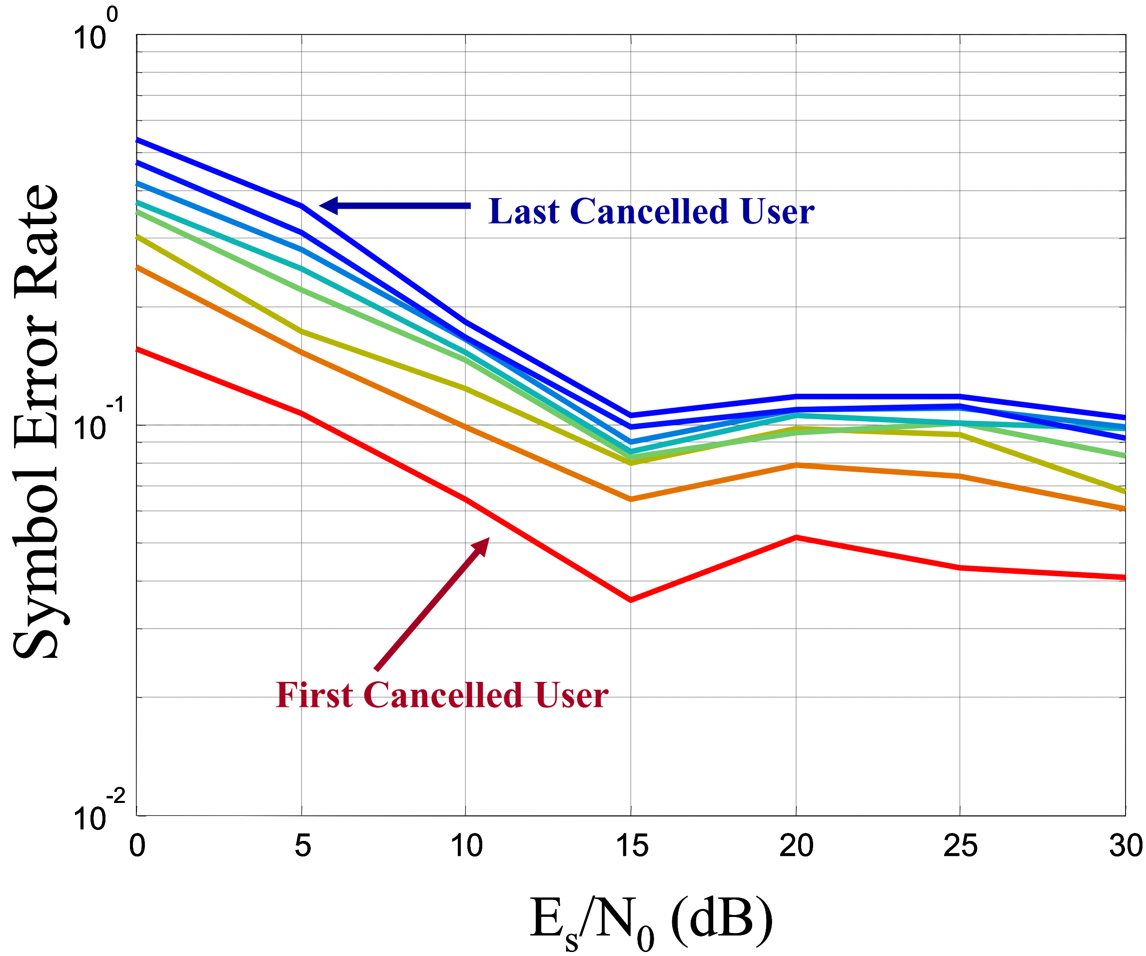


Figure 2.8: Symbol Error Rate for MMSE-SIC vs. Cancellation Order. $D_u = 8$, 35% rolloff, $M = 4$ element ICG spatial signatures

.

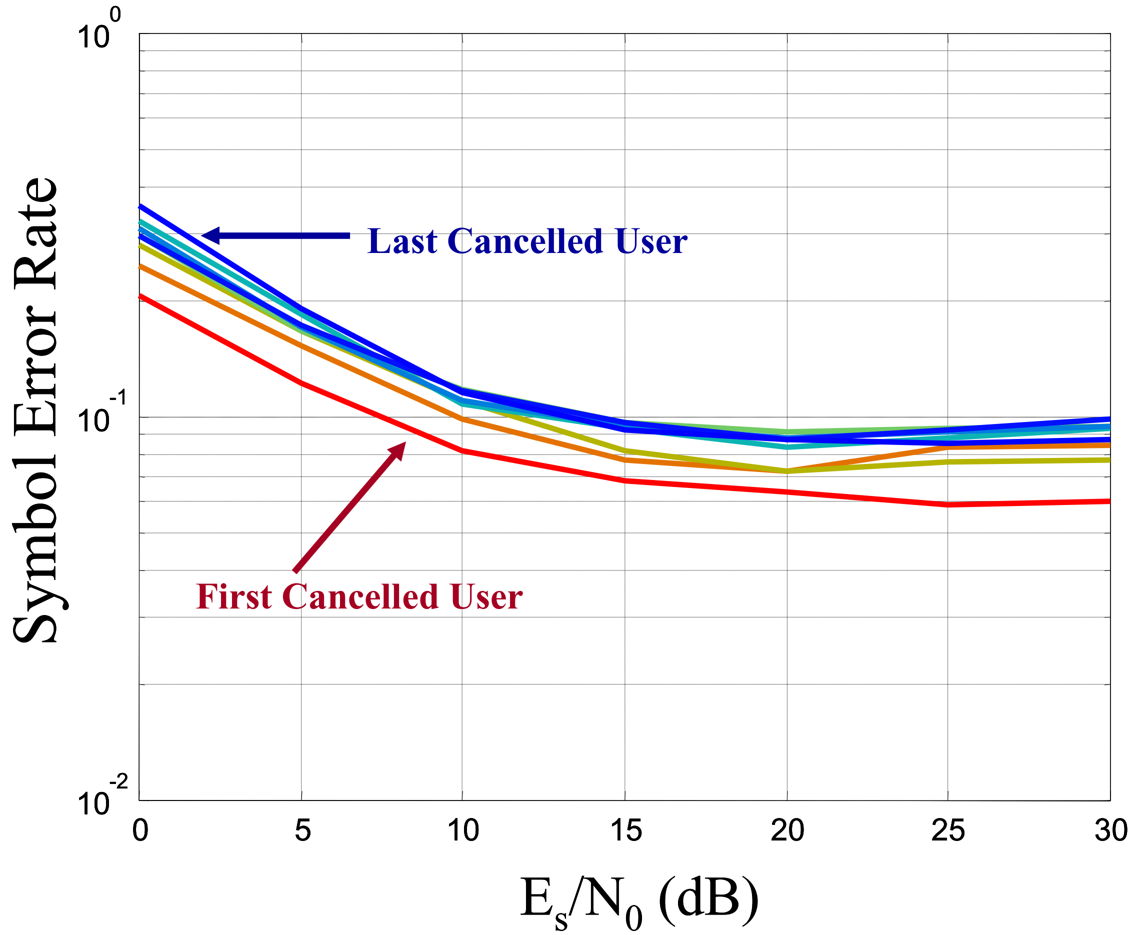


Figure 2.9: Symbol Error Rate for MMSE-SIC vs. Cancellation Order, . $D_u = 8$ , 75% rolloff, $M = 4$ element element UCA

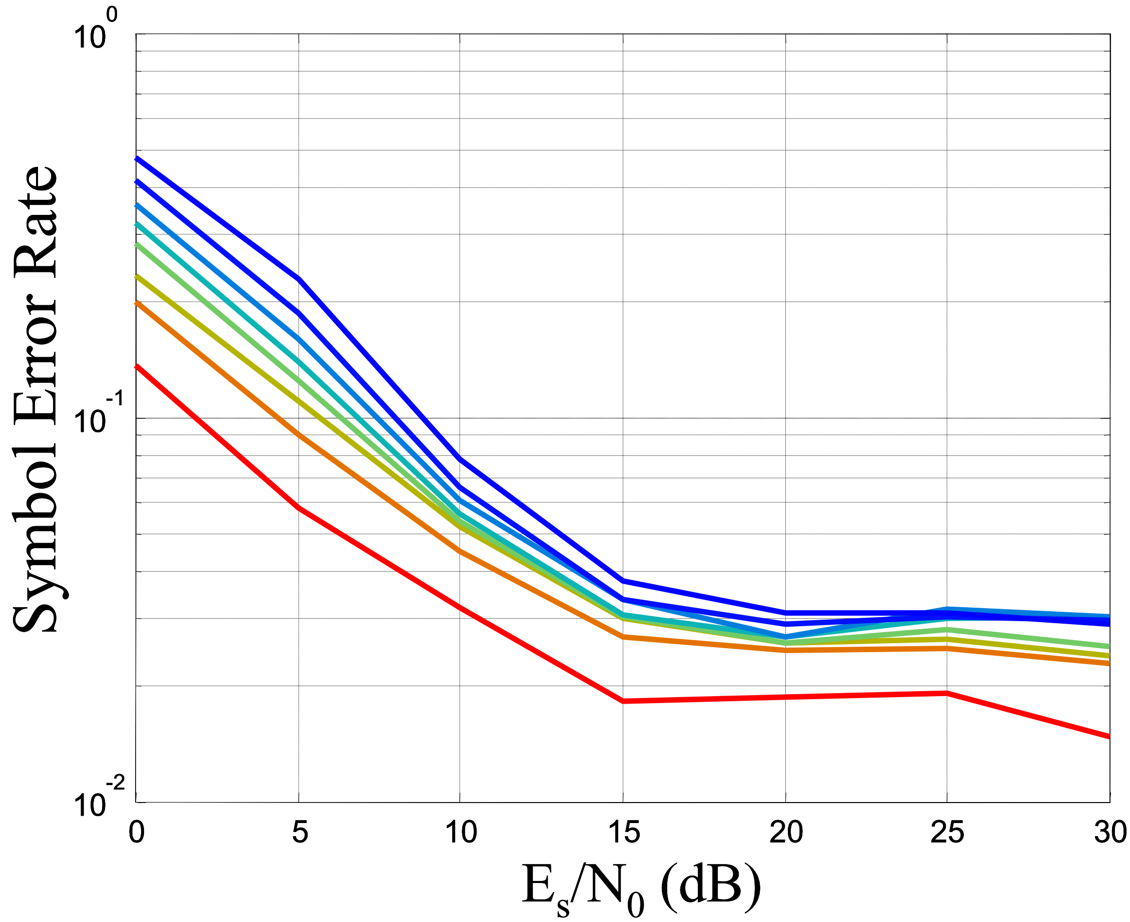


Figure 2.10: Symbol Error Rate for MMSE-SIC vs. Cancellation Order. $D_u = 8$, 75% rolloff, $M = 4$ element ICG spatial signatures

.

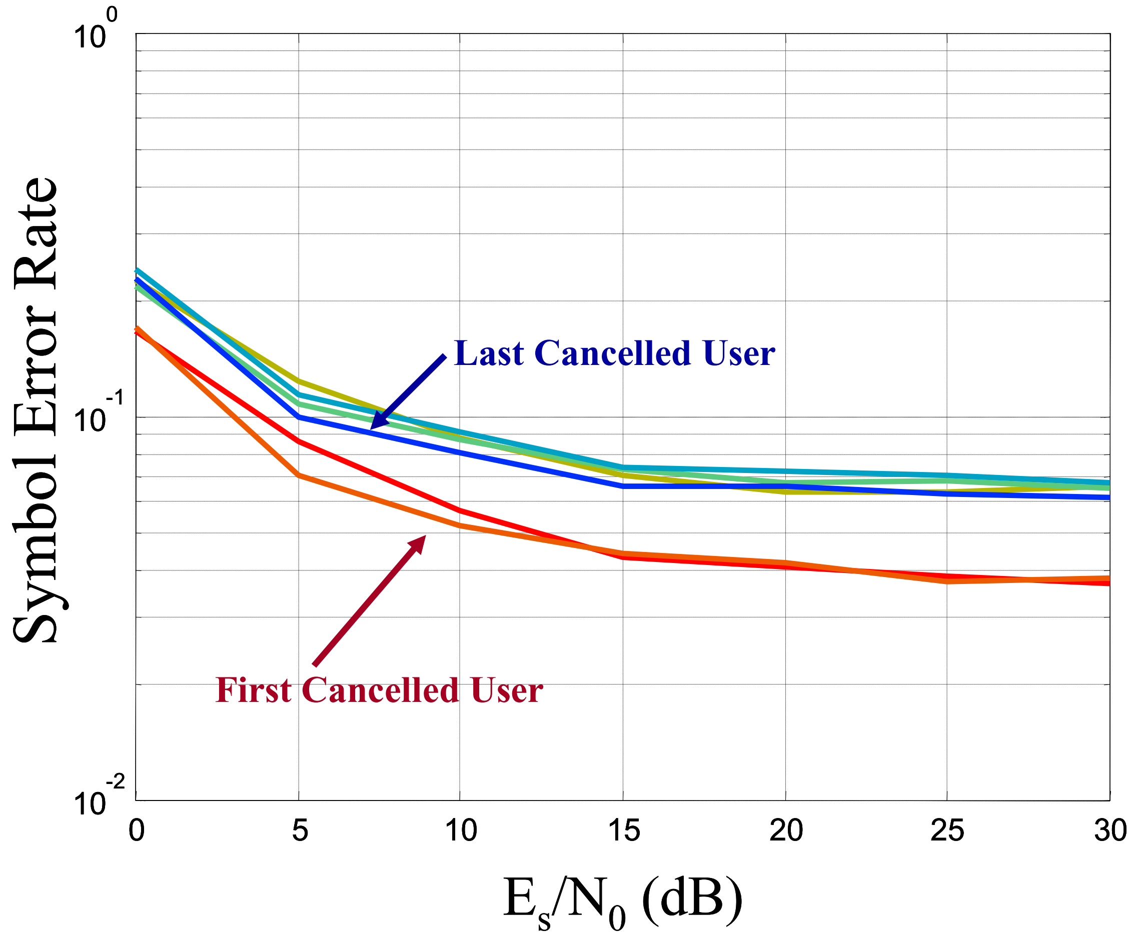


Figure 2.11: Symbol Error Rate for MMSE-SIC vs. Cancellation Order. $D_u = 6$ , 35% rolloff, $M = 4$ element UCA.

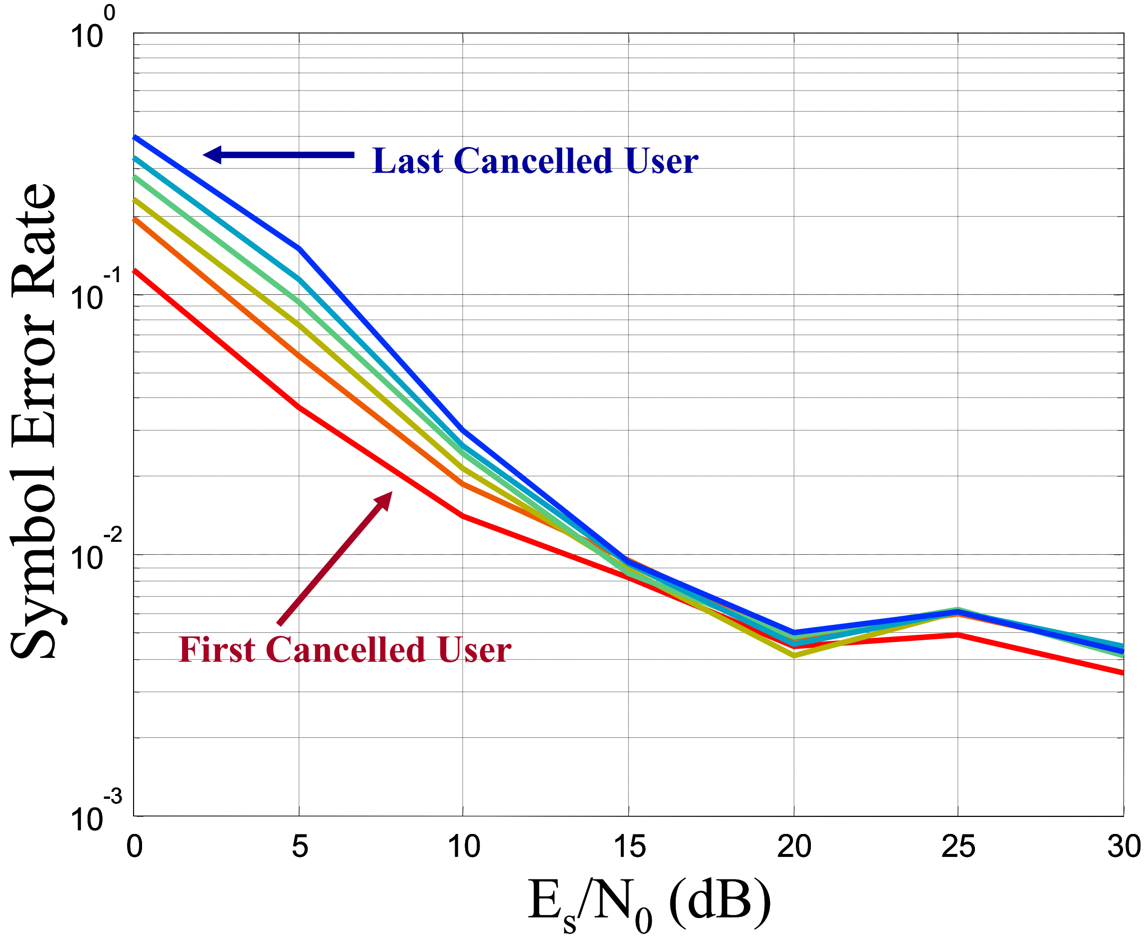


Figure 2.12: Symbol Error Rate for MMSE-SIC vs. Cancellation Order. $D_u = 6$ , 35% rolloff, $M = 4$ element ICG spatial signatures.

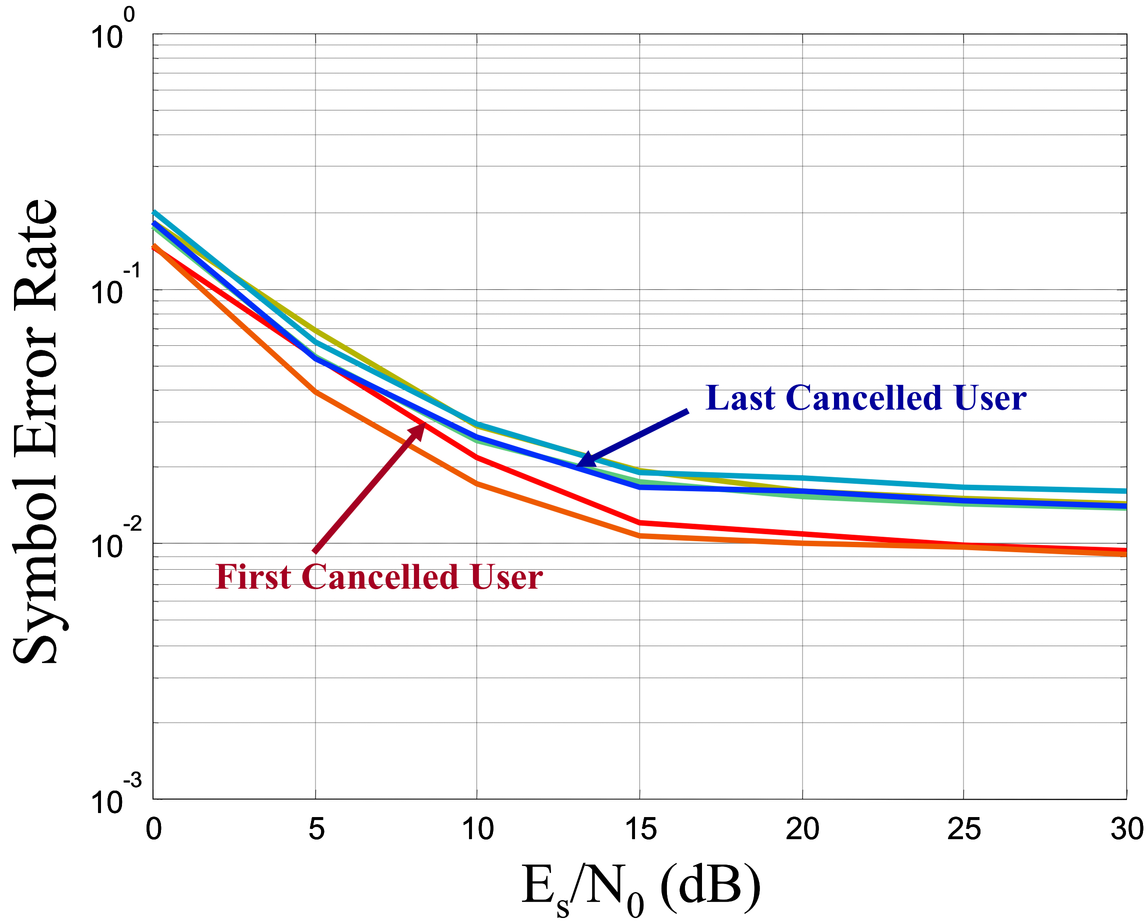


Figure 2.13: Error Rate for MMSE-SIC vs. Cancellation Order. $D_u = 6$ , 75% rolloff, $M = 4$ element UCA.

# Chapter 3: OVERLOADED ARRAY PROCESSING WITH MAP-SIC

## 3.1 INTRODUCTION

In the previous chapter, it is shown that MMSE space-time processing embedded in a successive interference cancellation framework (SIC) can extract most signals from a high SNR overloaded environment. This chapter analyzes the performance of more sophisticated receivers designed to extract signals with less excess bandwidth at lower SNRs. The application of Reduced-State Maximum Aposteriori Probability in tandem with Successive Interference Cancellation (MAP-SIC) is considered for the purpose of extracting symbol-asynchronous $\pi/4$-DQPSK signals with RRCOS pulse shaping (35%-75% rolloff) in overloaded environments. Algorithm performance is analyzed through simulation, and quantified in terms of symbol error rate ranked by cancellation order. The effect of signal excess-bandwidth, adjacent channel interference, the number of users, SNR, and array type are considered. The channel parameters of all users are assumed known.

The performance of MMSE-SIC pivots on the ability of a linear space-time processor (STP) to extract the first signal from an overloaded environment. This chapter investigates the replacement of linear space-time processing in MMSE-SIC with the use of non-linear processing. MMSE STP attempts to reject interfering signals with linear filters but must do so without significantly distorting the desired signal. The MMSE cost-function acts as a fulcrum in the tradeoff between interference rejection and distortion of the desired signal [13]. Alternatively, both Maximum Likelihood Sequence Estimation (MLSE) [21] and Maximum-Apriori Probability Sequence Estimation (MAPSE) [22] provide a non-linear interference rejection framework which divides interference rejection and desired signal estimation into separate blocks. This framework is attractive because it allows the receiver to simultaneously exploit both the finite alphabet and excess bandwidth properties of the desired signal without compromising the exploitation of interfering signal's excess bandwidth.

MLSE was originally developed for the detection of pulse-amplitude modulated signals in the presence of additive colored Gaussian noise. This so called "Forney-formulation" of MLSE accounts for noise-coloring with a pre-whitening approach. An alternative, "Ungerboeck-formulation" of MLSE was latter proposed in [23]. Its application to interference rejection and array processing was first investigated in [35] where interfering signals are treated as wide-sense stationary (WSS) colored Gaussian noise. In this approach, interference rejection is built into the cost-metric of the Viterbi Algorithm (VA). It has the benefit that channel parameters can be estimated with straight-forward averaging but has the drawback that it does not lend itself well to state-reduction techniques such as that of [20]. In contrast, Forney's formulation lends itself well to state-reduction, but in interference rejection applications, channel parameters are difficult to estimate. Nevertheless, Forney's formulation has the benefit over linear MMSE STP in that interference rejection and signal detection are broken into two stages of processing. In the first stage, interference is rejected through the use of a whitened matched filter (WMF). For example, if a desired signal is subject to AWGN and a single narrow-band interferer, a whitening filter will level the interference to the background noise level. Unlike MMSE filtering, the WMF is allowed to introduce frequency selective distortion to the desired signal. The relaxation of this constraint results in higher output signal to interference ratios (SINR) and hence, superior detection performance. Inter-symbol interference (ISI) introduced by the WMF is later accounted for in the VA. Alternatively, the soft-input/soft-output MAP algorithm may replace the

VA within the WMF interference rejection framework and has the following advantages. Firstly, the soft-input feature of the MAP algorithm allows for the potential exploitation of iterative decoding. Secondly, the soft-output feature of the MAP algorithm provides some measure of confidence in output symbol estimates that will aid its use in the SIC framework.

In [35], it was noted that treating interference as WSS Gaussian noise is a pessimistic model for the following reasons. Firstly, the Gaussian aspect fails to model the finite-alphabet property common to many digital signaling techniques. Finally, the WSS aspect fails to model the cyclo-stationarity induced by linear pulse shaping. Multi-user detection replaces the Guassian model with a more representative finite-alphabet model and can drastically improve signal extraction performance [16]. However, the complexity of these techniques grows exponentially with the number of users and hence, can be prohibitively complex. Often, the WSS assumption is justified in under-loaded environments, because it leads to simplified interference rejection structures without much performance penalty. However, in the case of overloaded arrays, excess bandwidth must be exploited in order to achieve adequate symbol error rates.

This study investigates the application of novel multiple-input, multiple output (MIMO) channel filtering techniques to exploit the excess bandwidth of interfering signals within a non-linear processing framework. In our treatment, the channels of all interfering signals are assumed known. However, this research motivates the investigation of adaptive implementations.

PAM digital signals (all assumed to have the same symbol rate) can be converted from cyclostationary random processes to vector wide-sense stationary processes simply by over-sampling the received signal and stacking sub-samples into a vector time series. Hence, single channel MLSE for cyclostationary Guassian noise can be converted to a multi-channel MLSE for wide-sense stationary noise. This characterization leads to the use of MIMO whitened matched filtering prior to non-linear detection.

MAPSE based interference rejection has many pitfalls. Firstly, strict adherence to the maximum likelihood criterion leads to very long WMFs [23]. For instance, a moving average noise process can be whitened only with an infinite impulse response filter (IIR). However, random processes can be approximately whitened, in some sense, with finite impulse response (FIR) filtering [39]. Causal, constrained length whitening filters are desirable in adaptive applications. For this reason, we have proposed a novel spectral flatness measure for multi-channel random processes that is an extension of the measure in [40]. Optimizing this criterion leads to a MIMO extension of the well known Yule-Walker Equations, for which there are efficient numerical solutions. Since the MAP algorithm's complexity increases exponentially with the length of the whitened matched filter, direct application of MAPSE can be prohibitively expensive. For this reason, the reduced-state MAP algorithm proposed in [22] based partially on DDFSE [20]. It was found that RS-MAP interference-rejection can extract one or two signals from an overloaded environment, warranting its application to successive interference cancellation (SIC). However, deciding which signal to extract at each cancellation stage is a difficult problem. For this reason, in this chapter, we propose an output-Entropy based cancellation decision.

## 3.2 SIGNAL MODEL

Consider $D_u$ linearly pulse shaped signals with pulse shape, $p(t)$ , and symbol period, $T_s$, impinging on an $M$ element antenna array. Let the $d^{th}$ user arrive at the array with a time delay, $\tau_d$, and constant phase, $\phi_d$. Assume the delays of each user are independent and uniformly distributed in the range $[-T_s/2, T_s/2]$, and the phases are uniformly distributed in the range, $[0, 2\pi]$ but the carrier differences between co-channel signals are negligible. Further, assume that if any received signal is subject to multi-path, then it results in flat-fading. Consider an antenna array where each element is front end filtered with a "brick-wall" anti-aliasing filter and over-sampled with over-sampling factor, $Q_0$. Then the received signal can be modeled as follows

$$\underline{r}_0[k] = \underline{r}_0\left(\frac{kT_s}{Q_o}\right) = \sum_{d=1}^{D_u} \underline{x}_d[k] + \underline{z}^{(0)}[k], \tag{3.1}$$

where

$$\underline{x}_d[k] = \underline{a}_d \sum_l s_d[l] g_d^{(0)}[k - lQ_0]. \tag{3.2}$$

Here, $\underline{r}_0(t)$ is the $M \times 1$ vector whose $m^{th}$ element is the continuous time, complex envelope of the signal received at the $m^{th}$ antenna; $\underline{x}_d[k]$ is the $d^{th}$ user's signal component of $\underline{r}_0[k]$; $\underline{z}^{(0)}[k]$ is a vector of complex Gaussian noise that models background thermal noise and adjacent channel interference. The adjacent channel interference model of [14] is used. The $d^{th}$ user transmits a sequence of symbols, $s_d[l]$, over a channel with discrete equivalent impulse response, $g_d^{(0)}[k]$, and spatial signature, $\underline{a}_d$. The $d^{th}$ user's temporal impulse response is $g_d^{(0)}[k]$, which collectively models the phase, $\phi_d$, delay, $\tau_d$, and pulse-shape, $p(t)$, and can be expressed as

$$g_d^{(0)}[k] = e^{j\phi_d} p\left(\frac{kT_s}{Q_0} - \tau_d\right). \tag{3.3}$$

In this study, we will consider root-raised cosine (RRCOS) pulse shaping filters with 35%, and 75% roll-offs. A closed form expression for $p(t)$ can be found in [25]. For these roll-off factors, an over-sampling factor of $Q_0 = 2$ models all of the spectral features of interest. For the purpose of simulation, $g_d^{(0)}[k]$ can be truncated to $L_g - 1 = 6$ symbol periods and sufficiently delayed to assure causality [25]. Note that although we have assumed that each user's channel is not frequency-selective, the pulse shape itself introduces ISI. In the overloaded asynchronous case, this ISI cannot be easily removed with matched filtering.

In this study, we will consider two spatial signatures. Firstly, we will consider the case when all co-channel signals impinge on an $M$ element uniform circular array (UCA) with AOAs, $\theta_d = (d-1)2\pi/D_u$, equally spaced in azimuth. The steering vector for an M element UCA is

$$\underline{a}(\theta_d) = [a_1 \quad \cdots \quad a_M]^T$$
$$a_m = \exp\left(-j\kappa_R \sin\left(\theta_d - \frac{2\pi(m-1)}{M}\right)\cos(\varepsilon_d)\right),$$

where the constant $\kappa_R = 2\pi R_a / \lambda$; $\theta_d$ is the azimuthal angle of arrival (AOA); $\varepsilon_d$ is the elevation (or depression) angle; and λ is the wavelength of the carrier frequency [24]. For simplicity, we will assume that $\varepsilon_d = 0$. Secondly, we will consider the case of quasi-static uncorrelated Rayleigh fading, where over a received frame, the received amplitude and phase of the $d^{th}$ signal at the $m^{th}$ antenna is a complex Gaussian random variable. Hence, the $d^{th}$ user's spatial signature is modeled as a vector of independent complex Gaussian (ICG) random variables.

The cyclo-stationary signal given in Equation 2.2 can be converted into a vector-wide-sense stationary random process by stacking subsamples over a symbol interval as follows.

$$\begin{aligned}\underline{r}[n] &\triangleq \left[\underline{r}_0[nQ_0]^T \quad \cdots \quad \underline{r}_0[(n+1)Q_0-1]^T\right]^T \\ &= \sum_{d=1}^{D_u}\sum_{l=0}^{L_g-1} \underline{g}_d[l]s_d[n-l] + \underline{z}[n],\end{aligned} \tag{3.4}$$

where

$$\underline{g}_d[n] \triangleq \left[\underline{a}_d^T g_d^{(0)}[nQ_0] \quad \cdots \quad \underline{a}_d^T g_d^{(0)}[(n+1)Q_0-1]^T\right]^T \tag{3.5}$$

is the subsample-augmented vector impulse-response for the $d^{th}$ user.

## 3.3 MULTI-CHANNEL RS-MAP

The Reduced State Maximum Aposteriori Probability (RS-MAP) algorithm was originally developed for estimating a sequence of symbols received with ISI and AWGN channel [22]. A non state-reduced version extended to the partially coherent reception of a $\pi/4-DQPSK$ symbol sequence transmitted over an inter-symbol interference (ISI) and received with AWGN is given in [35].

Consider a sequence, $\{\tilde{s}[k]\}$, of independent 4-QAM symbols, $\pi/4-DQPSK$ encoded into a sequence, $\{s[k]\}$, and transmitted over a number, $M$, of discrete time channels. Hence forth $\{\tilde{s}[k]\}$ will be called the message sequence, and $\{s[k]\}$ will be called the differentially encoded sequence. The sequence input to the base-band receiver is

$$\underline{y}[k]=\sum_{l=0}^{L_h-1}\underline{h}[l]s[k-l]+\underline{z}[k], \tag{3.6}$$

where $\underline{h}[k]$ is the $M\times 1$ vector channel impulse response of length $L_h$ and $\underline{z}[k]$ is assumed to be of complex additive white Gaussian noise vector sequence with a matrix auto-covariance sequence, $\mathbf{\Phi}_{\underline{zz}}[\Delta n]=E\left[\underline{z}[n]\underline{z}[n+\Delta n]^H\right]=\sigma_z^2\delta[\Delta n]\mathbf{I}$. A vector of all received array snapshots, $\underline{y}=\left[\underline{y}[0]^T \quad \cdots \quad \underline{y}[L_{rx}-1]^T\right]^T$. The input/output relationship of the MAP algorithm [35] is illustrated in Figure 3.1. Input to the algorithm is the noise variance per channel, $\sigma_z^2$, the vector channel coefficients, $\{\underline{h}[k]\}_{k=0}^{L_h-1}$, and a sequence of apriori probability distributions (APPi), $\{P(\tilde{s}[k])\}_{k=0}^{L_{tx}}$ which model possibly imperfect apriori information about the transmitted sequence. Note these so called “soft-input” parameters only describe information about message symbols. No apriori information about transmitted signals is needed. The MAP algorithm outputs sequences of aposteriori probability distributions for message, $\{P(\tilde{s}[k]\,|\,\underline{y})\}_{k=0}^{L_{tx}}$, (mAPPo) and differentially encoded symbols, $\{P(s[k]\,|\,\underline{y})\}_{k=0}^{L_{tx}}$, (eAPPo).

However, MAPSE has a complexity proportional to $4^{L_h}$, which prohibits its strict application to long channels or large alphabets. Nonetheless, in many applications, the channel may be very long but contains most of its energy at small lags[1]. RS-MAPSE [22] mitigates the complexity of sequence estimation by accounting for the dominant energy of $\underline{h}[k]$ with the MAPSE and the remaining ISI with decision feedback. Toward this end, the algorithm considers a portion of the vector impulse response, $\underline{h}[k]$, non-causal by advancing it by a pre-determined value, $\mu_b$. Hence, the vector impulse response is defined over the interval $k = -\mu_b, \ldots, L_h - \mu_b - 1$. The algorithm achieves a state reduction by decomposing the true channel state, $u[k] \triangleq \left(s_{k+\mu_b}, \ldots, s_{k-L_h+\mu_b+1}\right)$, into a reduced state, $\sigma[k] \triangleq \left(s_{k-1}, \ldots, s_{k-\mu}\right)$, a forward partial state, $\nu_f[k] \triangleq \left(s_{k-(\mu+1)}, \ldots, s_{k-L_h+1}\right)$ for the forward MAP recursion, and a backward partial state, $\nu_b[k] \triangleq \left(s_{k+\mu_b}, \ldots, s_{k+1}\right)$. The quantity, $\mu$, is an application specific design parameter called the reduced state size. RS-MAP approximates full-state MAPSE by constructing a trellis based on the reduced state sequence and applying a forward (or backward) partial state estimate (on a per-survivor basis) in tandem with the forward (or backwards, respectively) recursion of MAPSE. On the forward recursion a "forward partial state" signal component, $\underline{\hat{y}}_f^{(i)}[k] = \sum_{l=\mu+1}^{L_h-\mu_b-1} \underline{h}[l] s^{(i)}[k-l]$ is saved at each state, $i = 0, \ldots, 4^{\mu-1}$ of each stage for cancellation during the backwards recursion. Similarly, on the backwards recursion, a "backwards partial state" signal component, $\underline{\hat{y}}_b^{(j)}[k] = \sum_{l=-\mu_b}^{-1} \underline{h}[l] s^{(j)}[k-l]$ is saved at each state of each stage for cancellation during the forward recursion. Again, the choice of both $\mu$ and $\mu_b$ are application specific, but generally, a judicious choice captures the dominant energy of $\{\underline{h}[k]\}$ within $k = 0, \ldots, \mu$. In TDMA applications the trellis may be terminated to the all-zeros state at the beginning and end of a slot.

[1] For example, a non-obstructed multi-path channel has most of its energy in the first few echoes.

The extension of RS-MAP to π/4-DQPSK is straight-forward, but difficult to program. While only a single forward recursion and a single backwards recursion is required in MAPSE, RS-MAPSE's performance improves with multiple iterations. Henceforth, $N_{fb}$ describes the number of forward/backward iterations. In this chapter it was found that $N_{fb} = 2$ is adequate.

## 3.4 MIMO WHITENING FILTERS

Consider an $N^{th}$ order $M \times 1$ auto-regressive vector wide sense stationary (AR-VWSS) random process, $\underline{i}[n]$, with $M \times M$ auto-correlation matrix sequence $\mathbf{\Phi}_{\underline{i}\,\underline{i}}[\Delta n] = E\left[\underline{i}[k]\underline{i}[k+\Delta n]^H\right]$. Since $\underline{i}[k]$ is AR-VWSS, the Z-transform of the auto-correlation matrix sequence has a spectral factorization [39],

$$\mathbf{\Phi}_{\underline{i}\,\underline{i}}(z) = \tilde{\mathbf{W}}(z)^{-1}\mathbf{P}_0\mathbf{W}(z^{-*})^{-H}. \tag{3.7}$$

for some positive definite matrix, $\mathbf{P}_0$, and minimum-phase[2] matrix-polynomial

$$\tilde{\mathbf{W}}(z) = \mathbf{I} + \tilde{\mathbf{W}}[1]z^{-1} + \cdots + \tilde{\mathbf{W}}[N]z^{-N}. \tag{3.8}$$

The colored AR-VWSS process can be whitened with the FIR MIMO filtering operation

$$\underline{i}'[k] = \sum_{l=0}^{N} \tilde{\mathbf{W}}[l]\underline{i}[k-l]. \tag{3.9}$$

such that $\underline{i}'[k]$ has auto-correlation $\mathbf{\Phi}_{\underline{i}'\underline{i}'}[\Delta n] = \delta[\Delta n]\mathbf{P}_0$. However, when $N$ is very large, the application of shorter MIMO FIR filters for the purpose of *approximate whitening* become attractive. That is, consider filtering the sequence, $\underline{i}[k]$ with an $(L_w - 1)^{th}$ order MIMO filter, $\mathbf{W}[l]$, with $L_w < N+1$ such that

[2] A minimum phase matrix polynomial is a matrix polynomial with eigen-values within the unit circle [39].

$$\underline{v}[k] = \sum_{l=0}^{L_w-1} \mathbf{W}[l]\underline{i}[k-l]. \tag{3.10}$$

Since, the order of $\mathbf{W}[l]$ is less than $N$, the output sequence cannot be strictly white. In this case a filtering criterion for approximate whitening must be defined. For the single channel case, $M = 1$, the following measure of "spectral-flatness" or "white-ness" is proposed in [40].

$$\xi_v = \frac{\exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{vv}(f)df\right\}}{\int_{-\frac{1}{2}}^{\frac{1}{2}} P_{vv}(f)df}. \tag{3.11}$$

The criterion of Equation (3.11) is normalized, $0 \le \xi_v \le 1$, reaching its maximum value at $\xi_v = 1$. Although it places more emphasis on nulls than peaks, it has the benefit that its maximum is obtained by a solution to the well known Yule-Walker Equation (YWE).

In the multi-channel case, whitening must be handled in two stages, for reasons given below. Consider the vector sequence, $v^{(0)}[k]$ obtained by filtering $\underline{i}[k]$ with the length $L_w$, causal, MIMO FIR filter $\mathbf{W}^{(0)}[l]$. The single channel spectral flatness criterion of Equation (3.11) can be generalized to the following criterion.

$$\xi_{\underline{v}^{(0)}} = \frac{\exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} Tr\left[\ln \mathbf{\Phi}_{\underline{v}^{(0)}\underline{v}^{(0)}}(f)\right]df\right\}}{\int_{-\frac{1}{2}}^{\frac{1}{2}} Tr\left[\mathbf{\Phi}_{\underline{v}^{(0)}\underline{v}^{(0)}}(f)\right]df} \tag{3.12}$$

where, the natural log operation denotes the *matrix log* [29] and $\mathbf{\Phi}_{\underline{v}^{(0)}\underline{v}^{(0)}}(f)$ denotes the cross-spectral density matrix. As with the single channel case, Equation (3.12) is maximized by a MIMO extension of the Yule-Walker Equations.

$$\begin{bmatrix} \mathbf{\Phi}_{\underline{i}\underline{i}}[0] & \cdots & \mathbf{\Phi}_{\underline{i}\underline{i}}[L_w-2] \\ \vdots & \ddots & \vdots \\ \mathbf{\Phi}_{\underline{i}\underline{i}}[-(L_w-2)] & \cdots & \mathbf{\Phi}_{\underline{i}\underline{i}}[0] \end{bmatrix} \begin{bmatrix} \mathbf{W}^{(0)}[1]^H \\ \vdots \\ \mathbf{W}^{(0)}[L_w-1]^H \end{bmatrix} = \begin{bmatrix} -\mathbf{\Phi}_{\underline{i}\underline{i}}[-1] \\ \vdots \\ -\mathbf{\Phi}_{\underline{i}\underline{i}}[-(L_w-1)] \end{bmatrix} \tag{3.13}$$

The auto-correlation of the output sequence has a Z-transform

$$\begin{aligned} \mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}(z) &= \mathbf{W}^{(0)}(z)\mathbf{\Phi}_{\underline{i}\,\underline{i}}(z)\mathbf{W}^{(0)}(z^{-*})^H \\ &= \cdots + \mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}[-1]z^1 + \mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}[0] + \mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}[1]z^{-1} + \cdots \end{aligned} \tag{3.14}$$

The trace operation in both the numerator and denominator of (3.12) make this criterion *only* a measure of *net flatness* for the power spectral densities and cross power spectral densities in $\mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}(f)$. Hence, for sufficiently long $L_w$, the elements of $\mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}[\Delta n]$, $\Delta n \neq 0$, can be forced arbitrarily close to zero. However, Equation (3.12) is *not* a measure or inverse measure of cross correlation. This means that $\mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}[0]$ is not necessarily diagonal or even nearly diagonal. If $\mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}[0]$ is non-singular, it can be diagonalized as follows. Let, $\mathbf{W}_0 = \left(\mathbf{\Phi}_{\underline{\nu}^{(0)}\underline{\nu}^{(0)}}[0]\right)^{-1/2}$ and $\underline{\nu}[k] = \mathbf{W}_0\underline{\nu}^{(0)}[k]$ implying that $\mathbf{W}(z) = \mathbf{W}_0\mathbf{W}^{(0)}(z)$. Then, for sufficiently long $L_w$

$$\begin{aligned} \mathbf{\Phi}_{\underline{\nu}\underline{\nu}}(z) &= \mathbf{W}(z)\mathbf{\Phi}_{\underline{i}\,\underline{i}}(z)\mathbf{W}(z^{-*})^H \\ &\approx \mathbf{I}, \end{aligned} \tag{3.15}$$

as desired. The approximation improves with $L_w$ and is exact when $L_w = N+1$, where $N$ is the process order. Hence it is possible to whiten an AR-VWSS random process. In the case when $\underline{i}[k]$ is a moving-average (MA) or auto-regressive moving-average (ARMA) process, then equation (3.12) is still a valid criterion for approximate whitening. Any MA or ARMA process is equivalent to an AR random process of sufficiently high, possibly infinite order [47]. Examples of a MIMO whitening filter are given in a following section.

## 3.5 RS-MAPSE BASED INTERFERENCE REJECTION

Consider the receiver illustrated in Figure 3.2. An $M \times 1$ received vector signal, $\underline{r}_0(t)$, is oversampled with oversampling factor $Q_0$ and augmented into an $MQ_0 \times 1$ vector sequence, $\underline{r}[n]$, sampled at the symbol rate. The received sequence can be expressed as

$$
\begin{aligned}
\underline{r}[n] &= \sum_{d=1}^{D_u} \sum_{l=0}^{L_g-1} \underline{g}_d[l] s_d[n-l] + \underline{z}[n] \\
&= \sum_{l=0}^{L_g-1} \mathbf{G}_{MU}[l] \underline{s}[n-l] + \underline{z}[n] \\
&= \underline{x}_d[n] + \underline{i}_{\bar{d}}[n],
\end{aligned}
\tag{3.16}
$$

where $\underline{g}_d[l]$ is the vector-channel impulse response of the $d^{th}$ user. The quantity, $\underline{g}_d[l]$, is an aggregate model of the received phase, amplitude, time-delay, pulse-shaping function, and spatial signature. The $MQ_0 \times D_u$ matrix, $\mathbf{G}_{MU}[l] = \begin{bmatrix} \underline{g}_1[l] & \cdots & \underline{g}_{D_u}[l] \end{bmatrix}$ is the so called multi-user impulse response and $\underline{s}[n]$ is the multi-user symbol-vector, $\underline{s}[n] = \begin{bmatrix} s_1[n] & \cdots & s_{D_u}[n] \end{bmatrix}^T$. If the $d^{th}$ signal, $\underline{x}_d[n]$, is the desired signal component then it and its interference process, $\underline{i}_{\bar{d}}[n]$, are given by

$$
\begin{aligned}
\underline{x}_d[n] &= \sum_{l=0}^{L_g-1} \underline{g}_d[l] s_d[l] \\
\underline{i}_{\bar{d}}[n] &= \sum_{\substack{u=1 \\ u \neq d}}^{D_u} \sum_{l=0}^{L_g-1} \underline{g}_u[l] s_u[n-l] + \underline{z}[n] \\
&= \sum_{l=0}^{L_g-1} \mathbf{G}_{MU}^{(\bar{d})}[l] \underline{s}_{\bar{d}}[n-l] + \underline{z}[n],
\end{aligned}
\tag{3.17}
$$

where $\underline{s}_{\bar{d}}[n]$ is the $(D_u - 1) \times 1$ vector of symbols transmitted by all users excluding user $d$, obtained by deleting the $d^{th}$ element of $\underline{s}[n]$ and $\mathbf{G}_{MU}^{(\bar{d})}[l]$ is the $MQ_0 \times (D_u - 1)$ multi-user impulse response matrix obtained by deleting the $d^{th}$ column of $\mathbf{G}_{MU}[l], l = 0, \ldots, L_g - 1$. The sequence, $\underline{i}_{\bar{d}}[n]$, is VWSS with $MQ_0 \times MQ_0$ matrix auto-correlation sequence,

$$
\begin{aligned}
\mathbf{\Phi}_{\underline{i}\,\underline{i}}^{(\bar{d})}[\Delta n] &\triangleq E\left[\,\underline{i}[n]\underline{i}[n-\Delta n]^H\,\right] \\
&= \sum_{l=0}^{L_g-1} \sigma_s^2 \mathbf{G}_{MU}^{(\bar{d})}[l]\,\mathbf{G}_{MU}^{(\bar{d})}[l-\Delta n]^H + \mathbf{\Phi}_{\underline{z}\underline{z}}[\Delta n]
\end{aligned} \tag{3.18}
$$

Although, $\underline{i}_{\bar{d}}[n]$, is not strictly Gaussian, treating it as such leads to a processor with a complexity far less than the optimum multi-user detector [33].

The so-called Forney formulation [19] prescribes the use of a linear pre-whitening transform followed by non-linear sequence estimation. Although it is shown that a whitened matched filter (WMF) produces a scalar (as opposed to vector) sequence of sufficient statistics for MAPSE, this WMF is often non-causal and infinite in duration. Such filters can be difficult to estimate in practice. Instead of directly applying a WMF, we apply a constrained length approximate MIMO whitening filter to the input sequence, producing a baud-rate sampled vector output sequence for MAPSE. This study formally defines a filtering criterion for a causal, constrained length MIMO whitening filter.

Now, let $\underline{y}_d[n]$ be the output of the MIMO filtering operation

$$
\underline{y}_d[n] = \sum_{l=0}^{L_w-1} \mathbf{W}_d[l]\underline{r}[n-l]\,. \tag{3.19}
$$

Then, $\mathbf{W}_d[l]$ can be chosen to whiten the interference process for the $d^{th}$ user by applying Equation (3.12) with the auto-correlation sequence of Equation (3.18). The sequence $\underline{y}_d[n]$ can be input to the RS-MAP algorithm for an approximation of MAPSE. In practice, it was found that if $\underline{g}_d[n]$ includes a pulse-shaping model truncated to $2d_m$

symbol intervals, then $\mu_b = d_m$ is a good choice. For π/4-DQPSK, the reduced state size should be limited for complexity purposes to $\mu = 5$.

## 3.6 MAP-SUCCESSIVE INTERFERENCE CANCELLATION

While, RS-MAPSE cannot extract multiple signals, it can extract one or two. This motivates a receiver where RS-MAPSE is embedded in a successive interference cancellation architecture (MAP-SIC) as illustrated in Figure 3.3. At each stage, RS-MAPSE based interference rejection (RS-MAPSE-IR) extracts one user, re-modulates it, and subtracts it out. Let $\underline{r}^{(p-1)}[n]$ and $d_p$ be the input signal and user extracted at the $p^{th}$ stage, respectively. Then RS-MAPSE-IR outputs the eAPPo's for the symbol sequence $\left\{\hat{s}_{d_p}[k]\right\}$. A MAP estimate of the sequence $\left\{\hat{s}_{d_p}[k]\right\}$ is easily obtained by choosing the symbol value with the maximum eAPPo at each time instant, k. A new received signal is derived by canceling the remodulated $\hat{s}_{d_p}[n]$ from $\underline{r}^{(p-1)}[n]$ as follows.

$$\begin{aligned} \hat{\underline{x}}_{d_p}[n] &= \sum_l \hat{s}_{d_p}[l]\underline{g}_{d_p}[n-l] \\ \underline{r}^{(p)}[n] &= \underline{r}^{(p-1)}[n] - \hat{\underline{x}}_{d_p}[n]. \end{aligned} \tag{3.20}$$

The process begins with $\underline{r}^{(0)}[n] = \underline{r}[n]$. At each stage, MAP-IR calculates an interference rejection filter $\mathbf{W}_{d_p}[n]$ assuming perfect cancellation on previous stages. That is, cancelled users are excluded from Equation (3.18)

Even with knowledge of channel parameters, the choice of user, $d_p$, to extract at each stage is a difficult one. With MMSE-SIC, the user with the best output SINR can be chosen because closed-form formulas for output SINR are available. However, it is difficult to predict the performance of RS-MAPSE-IR based on channel parameters. However, eAPPo's output by RS-MAPSE provide a measure of confidence in the estimated sequence. To exploit this feature of this "soft-output" sequence, we propose the following measure of signal quality, for the $d^{th}$ signal at the $p^{th}$ stage.

$$\overline{\mathcal{H}}_d^{(p)} = \frac{1}{L_{tx}} \sum_{k=0}^{L_{tx}-1} \left( \sum_{\mathcal{S} \in \mathcal{A}[k]} \hat{P}_o(s_d[k] = \mathcal{S}) \log_{|\mathcal{A}|} \left( \hat{P}_o(s_d[k] = \mathcal{S}) \right) \right) \tag{3.21}$$

where $\mathcal{A}[k]$ is the π/4-DQPSK alphabet at the $k^{th}$ time instant. The inner most sum can be interpreted as the Entropy of the $k^{th}$ output eAPP. The outer-most sum can be interpreted as an average of this entropy over a received slot. Hence, equation (3.21) is called average aposteriori entropy (AAPE). If entropy is a measure of uncertainty, then at each stage, the user with the least uncertainty should be chosen, yielding the following min-APPE rule.

$$d_p = \min_{1 \le d \le D_u} \overline{\mathcal{H}}_d^{(p)} \tag{3.22}$$

In order to divorce the algorithm's performance from the cancellation selection criterion, we also consider a *genie aided* MAP-SIC. That is, all users are extracted and the number of symbol errors are counted by comparing the estimated symbol sequence with the true sequence. The user with the least number of errors is selected. Obviously, this selection criterion is not feasible in practical applications; however, simulating this selection criterion helps us understand the efficiency of the min-APPE rule.

As an example, consider the case of $D_u = 6$, 35% RRCOS pulse-shaped signals impinging on an $M = 4$ element diversity array. Run RS-MAP based interference rejection to extract each user in the environment. A sequence of sorted APPo's (differentially encoded symbols) is plotted for each user in Figure 3.4. Note, for all users except $d = 3, 4$, signal extraction quality is very poor. Demodulation of user 3 results in 1 error out of 100 and Demodulation of user 4 results in 2 errors out of 100. A genie aided MAP-SIC would proceed to re-modulate and cancel user 3's signal from the signal environment. However, the min-APPE rule will choose to cancel user 4's signal.

Although this example illustrates the fallibility of the min-APPE rule, in most cases, the min-APPE rule will make as good of a choice as the genie aided MAP-SIC.

## 3.7 Simulation Results

The probability of error for MAP-SIC is difficult to calculate because each symbol sequence estimate is a highly non-linear function of both the received signal and random channel parameters. Instead, symbol error rates have been estimated through the use of Monte-Carlo Simulation. SERs are reported for both π/4-DQPSK message symbols and differentially encoded symbols. Differentially encoded SERs are most important for understanding cancellation performance, since these symbols are remodulated for signal cancellation. Message SERs are most important for understanding signal extraction performance, since this quantity ultimately determines signal quality.

Two array types were considered: a $\lambda/2$ diameter uniform circular array (UCA) with, and independent complex Gaussian (ICG) spatial signatures. Both arrays have $M=4$ elements. For the case of the UCA, all signals are assumed to be equally spaced in azimuth over $360^0$. Root-Raised Cosine pulse shaping with 35% rolloff was considered. Simulations were run with 162 symbol frames at a time. For each simulated frame a new set of channel parameters (e.g. phases, delays, and spatial signatures) were generated. Symbol error rate statistics are recorded according to cancellation order. A frame error is said to occur when all users have at least one error in a frame. Simulations were run until at least 20 frame errors and 10 symbol errors occurred on each frame. For each simulated case, a detection window of $L_w = 15$ symbols and an oversampling factor of $Q_0 = 2$ is assumed. A reduced state size of $\mu = 5$ and a backward partial state size of $\mu_b = 4$. Very little benefit was found for increasing the number of forward-backward iterations of RS-MAPSE beyond two.

Figures 3.5-3.6 show the performance of MAP-SIC for $D_u = 6$, 35% RRCOS pulse-shaped signals impinging on an $M = 4$ element $\lambda/2$ diameter UCA. Note in both figures, the cancellation performance for the first and last signals are within 1 dB of one another. This is because the quality of signal extraction for the first SIC stage effects the extraction performance of successive SIC stages. The SER curve for message symbols resembles that of the differentially encoded symbols except it exhibits an approximate 3dB differential detection loss. In this environment, MAP-SIC supports commercial voice quality (SER < 1%) for all signals when the background SNR > 17dB. This is a high but achievable SNR in commercial cellular systems.

Note, that for Figures 3.5-3.6, the MAP-SIC algorithm employed the min-AAPE rule. To understand the efficiency of this rule, examine Figure 3.7, which plots the performance of a genie aided MAP-SIC. The genie-aided MAP-SIC's extraction performance on the first two stages is significantly better at high background SNRs, but within 0.5 dB of the min-APPE rule at moderate to low SNRs. Similar results were found all other signal environments considered in this chapter.

For higher overloading factors, the performance of MAP-SIC employing a UCA significantly degrades. Figures 3.8-3.9 show the performance of MAP-SIC for $D_u = 8$, 35% RRCOS pulse-shaped signals impinging on an $M = 4$ element $\lambda / 2$ diameter UCA. Observe that MAP-SIC is unable to guarantee an SER less than 10%, even at higher SNRs. However, if we consider the same scenario with users employing 75% rolloffs as shown in Figure 3.10, commercial voice quality is achieved with 15dB background SNRs. Hence, the performance of MAP-SIC, as with all signal extraction techniques that exploit cyclostationarity, is highly dependent on excess bandwidth. The higher excess bandwidth of a 75% RRCOS pulse shape affords better performance.

So far, we have considered only a phased array. Figures 3.11-3.12 consider the performance of MAP-SIC employing an M= 4 element diversity array with $D_u = 6$, 35% RRCOS pulse-shaped signals impinging on it. Note, that MAP-SIC employing diversity reception achieves a performance about 3dB better than MAP-SIC employing a UCA. This is because in a fading environment, MAP-SIC benefits from the fact that there is more variation between the powers of signals impinging on the array. This variation in received power yields better SINRs for the first cancelled user, and hence better cancellation performance. This is particularly beneficial at high background SNRs. Figure 3.13-3.14 illustrates the same receiver configuration with 8 signals impinging on it.

## 3.8 MEASUREMENT RESULTS

To demonstrate the feasibility of MAP-SIC, we have constructed a proof-of-concept demonstration system. Firstly, a wireless DSP programmable transmitter was constructed with a transmit power switchable between -4dBmW and 14dBmW and a local oscillator at 2051MHz. This transmitter was built with spare parts left over from previous MPRG wireless projects. The transmitter was programmed to transmit $\pi/4$-DQPSK signal in 162 symbol TDMA slots at a rate of 25ksps. The transmitter's data is generated by a M= 1048575 length M-sequence. This rate was chosen to be close as possible to the USDC symbol rate of 24.3ksps, as available hardware would allow. The transmitter's modular design allows for a variety of antenna designs constructed in the VA Tech antenna laboratory. In this experiment a vertically polarized coaxial dipole antenna with a 0dB Gain was used at the transmitter. At the receiver, a 4 element, vertically polarized UCA with an individual elemental gain of 0dB was connected to a bank of RF chains that down-convert the desired signal received at each individual element to an IF frequency of 1MHz. Note, to most accurately mimic a cost-effective commercial application, this array is not calibrated. Hence, AOA measurements are not possible. This IF signal is then sampled by a multi-channel deep memory oscilloscope with a sampling frequency of 5MHz. The resulting data-records are loadable in MATLAB for IF-down-conversion and base-band signal processing.

Base-band processing includes the following. For the purpose of channel estimation and symbol error-rate estimation, it is desired to know the exact sequence transmitted. Since the data sent by the transmitter is a known M-sequence, this is feasible, but since the M-sequence's duration is very long, (42 seconds, in fact), it is necessary and difficult to estimate the sequence's timing-phase. The first two recorded 2.5msec slots are collected, and the transmitted code phase is estimated with a differential correlator. The local oscillators of the transmitter and receiver, while exhibiting excellent phase-noise, were found on a spectrum analyzer to drift within 2kHz of one another. Also, since some frequency selective distortion is possible, due to multi-path and (more likely) non-ideal hardware, the channel must be estimated as an FIR filter. However, since the data-sent by the transmitter is known at the receiver, a joint channel frequency estimate is possible. In this case, we applied a Newton's block least-squares joint frequency/channel estimator. In simulation, this channel estimator has superior performance even in very low SNRs. In practice, however, we found signal to noise ratios to be un-expectedly low (24dB), even when the transmitter is directly connected to the receiver (with appropriate attenuation, not to exceed the maximum allowable power input to the receiver). We attribute these poor SNRs to signal distortion introduced by transmitter non-linearities. These non-linearities are untenable when the transmitter is operating with its maximum power setting.

The transmitter, transmitting at a power of -4dBm was placed at several random locations, all within 5m of the receiver, in the laboratory. A separate signal was recorded for each transmitter location. The resulting data-records were loaded into MATLAB, converted to complex base-band, and processed. The residual frequency error was estimated and the signal was mixed with a complex exponential to a carrier of 0Hz. The data records of all signals were then added in software and processed as if they were collected simultaneously. Note that since an enclosed laboratory with multiple obstructions is a very rich multi-path environment, AOA is not a meaningful description of the signal environment. Table 3.1 describes the SNR experienced by each signal. Average Received Power per element is just the total power received by the array divided by the number of elements. Average SNR/element is just the average signal power per element divided by the average noise power per element. Note, these numbers are very low for signal 2, 5, and 6. However, since most of the noise in this environment is generated by the transmitters, the received noise is spatially correlated and hence, a naïve application of one of the simulated SNR curves in this chapter (which assume spatially uncorrelated background noise) will yield very pessimistic predictions of performance. Note that since the min-AAPE rule is data dependent and the data received is random in nature, the resulting choice of cancellation order is dependent. Typical cancellation orders are (1, 3, 4, 5, 2, 6) and (3, 1, 4, 6, 2, 5). Hence, a particular signal may find itself cancelled at different stages for different slots. Table 3.2 provides a summary of measurement results. Note every user has a message SER less than 3%.

Table 3.1: Measurement Parameters

| Transmitter Number | Average SNR/element (dB) |
|---|---|
| 1 | 10.5 |
| 2 | 3.6 |
| 3 | 9.7 |
| 4 | 11.8 |
| 5 | 3.4 |
| 6 | 5.2 |

Table 3.2: Measurement results

| Cancellation Order | Number of errors in 24 slots (162 symbols/slot) | Average SNR/element (dB) |
|---|---|---|
| 1 | 0 | NA |
| 2 | 0 | NA |
| 3 | 0 | NA |
| 4 | 60 | 0.015 |
| 5 | 53 | 0.014 |
| 6 | 102 | 0.026 |

## 3.9 CONCLUSION

An algorithm for Overloaded Array Processing is proposed that embeds a RS-MAP based interference rejection within a successive interference cancellation framework. The performance of the array was evaluated through simulation for two different array types, assuming perfect channel estimates. MAP-SIC employing a UCA can provide commercial voice quality of service to six $\pi/4$-DQPSK users. MAP-SIC employing diversity reception provides commercial voice quality to eight $\pi/4$-DQPSK users. A key assumption that contributes to the algorithm's success is knowledge of channel parameters. Channel estimation is an area for future research.

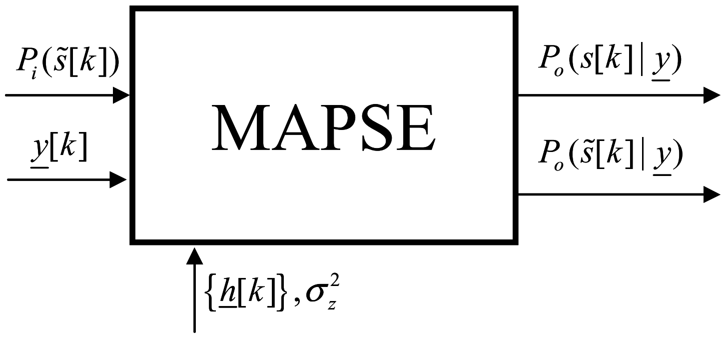


Figure 3.1: Input/Output Description of pi/4-DQPSK MAPSE.

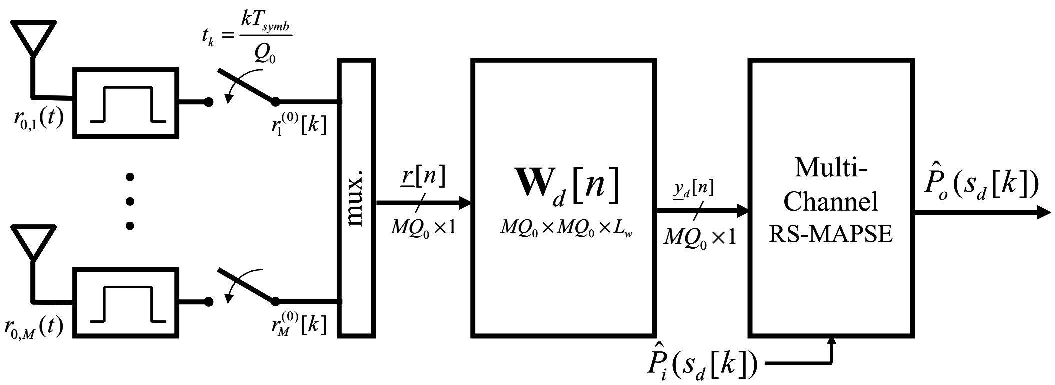


Figure 3.2: RS-MAPSE based interference rejection.

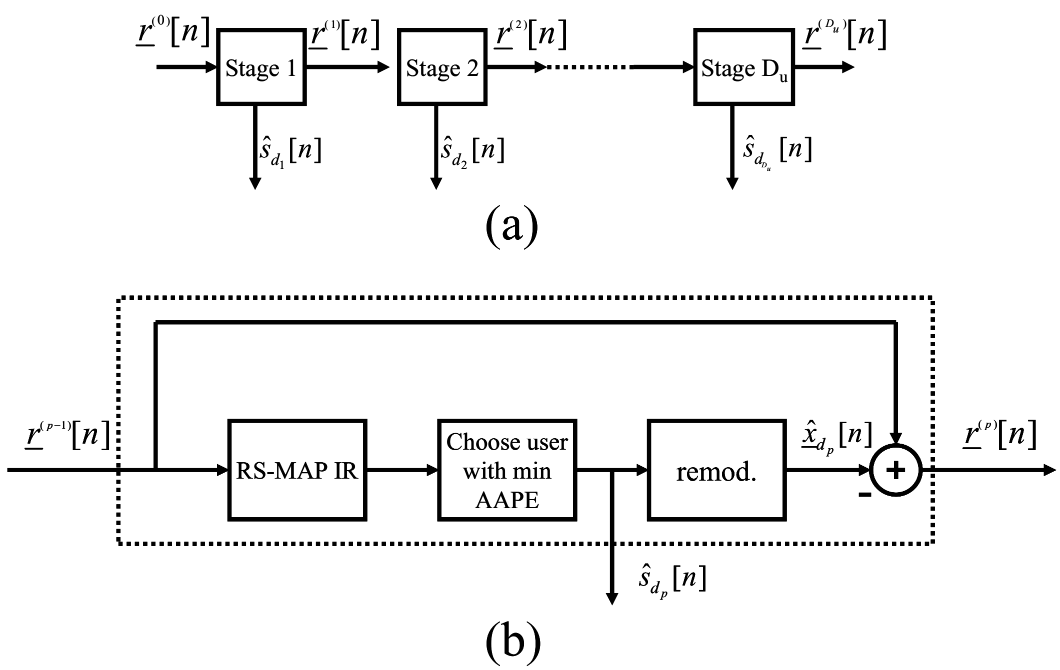


Figure 3.3: MAP-SIC architecture (a) Multi-Stage Structure (b) Structure of the $p^{th}$ stage.

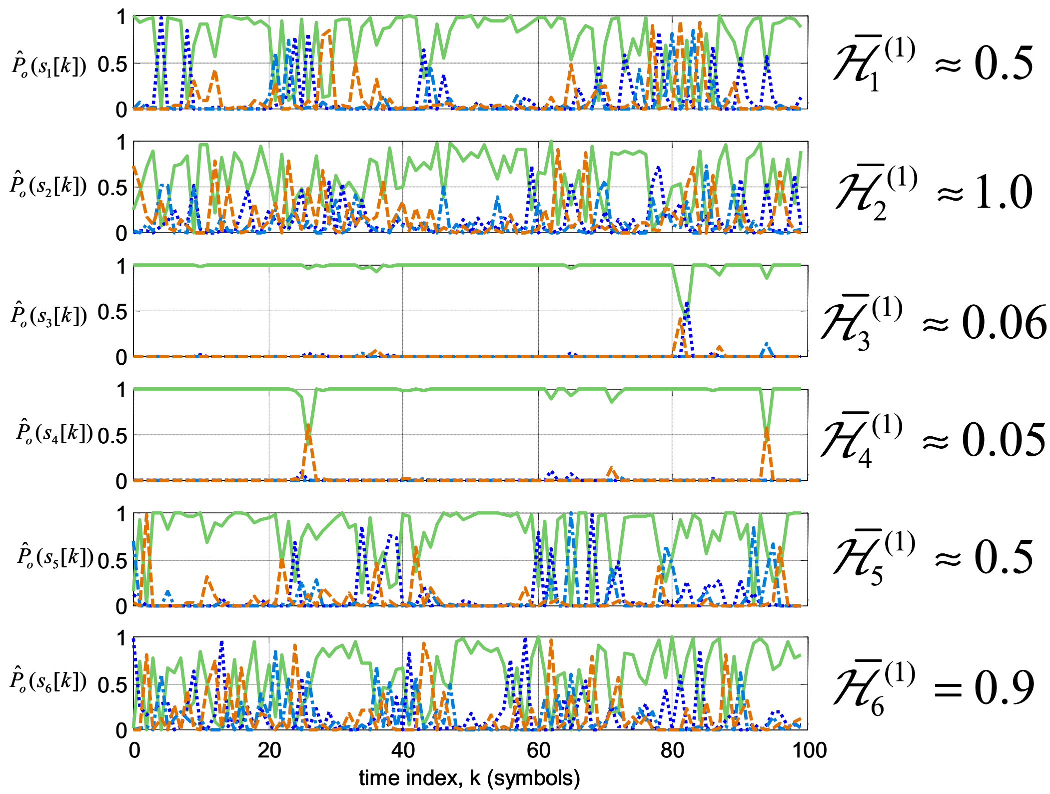


Figure 3.4: Differentially encoded symbol APPo's (sorted), output by first stage of MAP-SIC. The AAPE for each APPo sequence is labeled to the right. Signal Environment: $M$= 4 element UCA, $D_u$= 6 , 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5. Background SNR= 14dB.

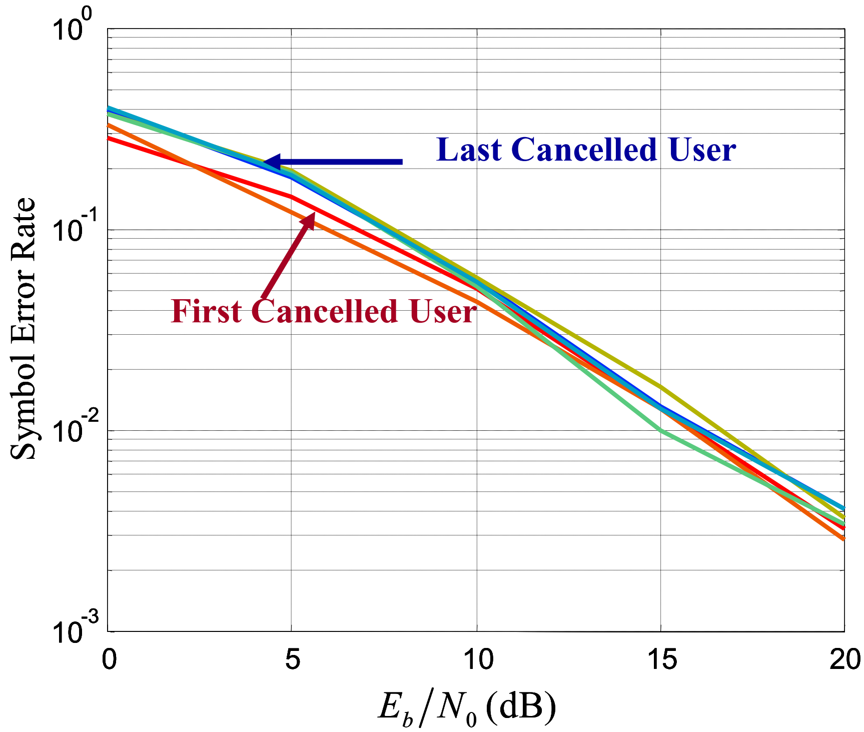


Figure 3.5: *Message* Symbol Error Rate for MAP-SIC, $M$= 4 element UCA, $D_u$= 6 , 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

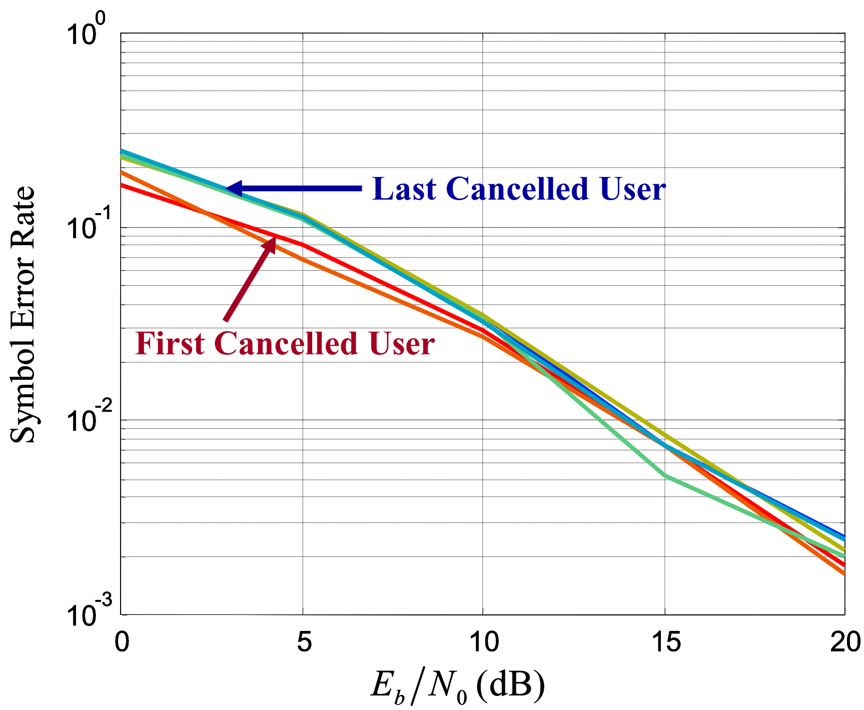


Figure 3.6: *Differentially Encoded* Symbol Error Rate for MAP-SIC, $M$= 4 element UCA, $D_u$= 6 , 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

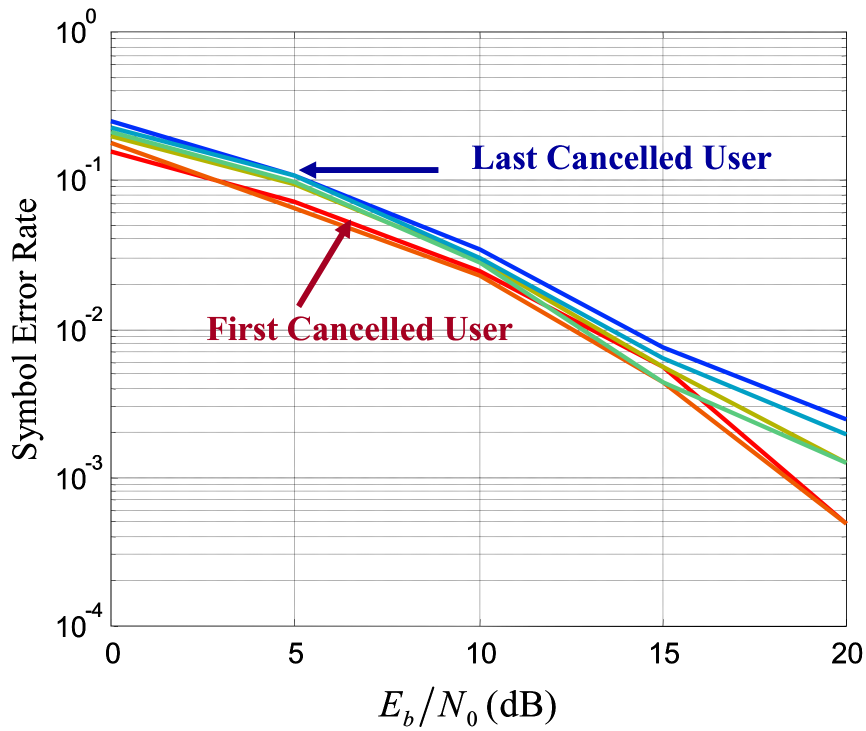


Figure 3.7: *Differentially Encoded* Symbol Error Rate for *Genie Aided* MAP-SIC, $M$= 4 element UCA, $D_u$= 6 , 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

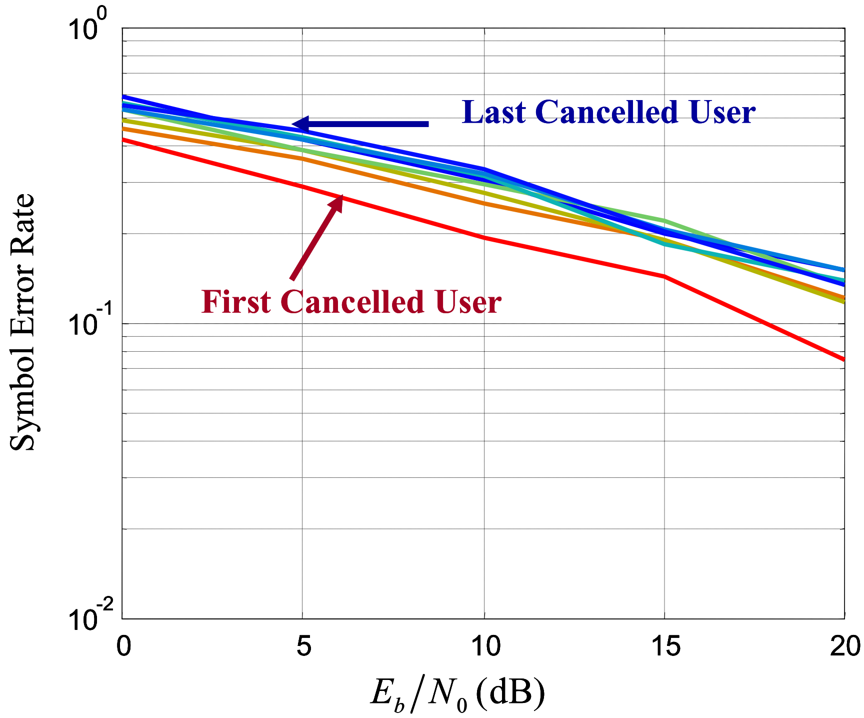


Figure 3.8: *Message* Symbol Error Rate for MAP-SIC, $M$= 4 element UCA, $D_u$= 8 , 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

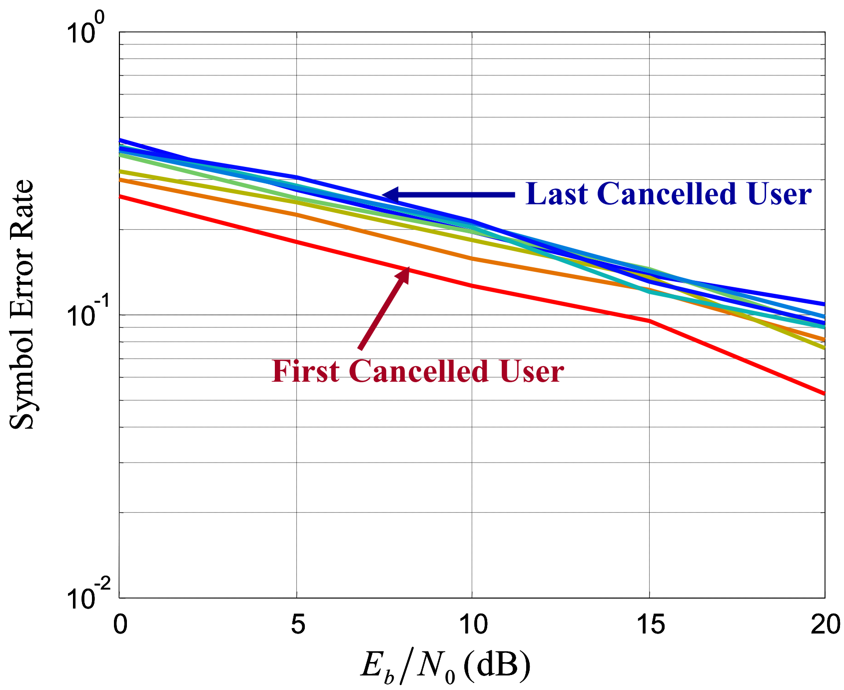


Figure 3.9: *Differentially Encoded* Symbol Error Rate for MAP-SIC, $M$= 4 element UCA, $D_u$= 8 , 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

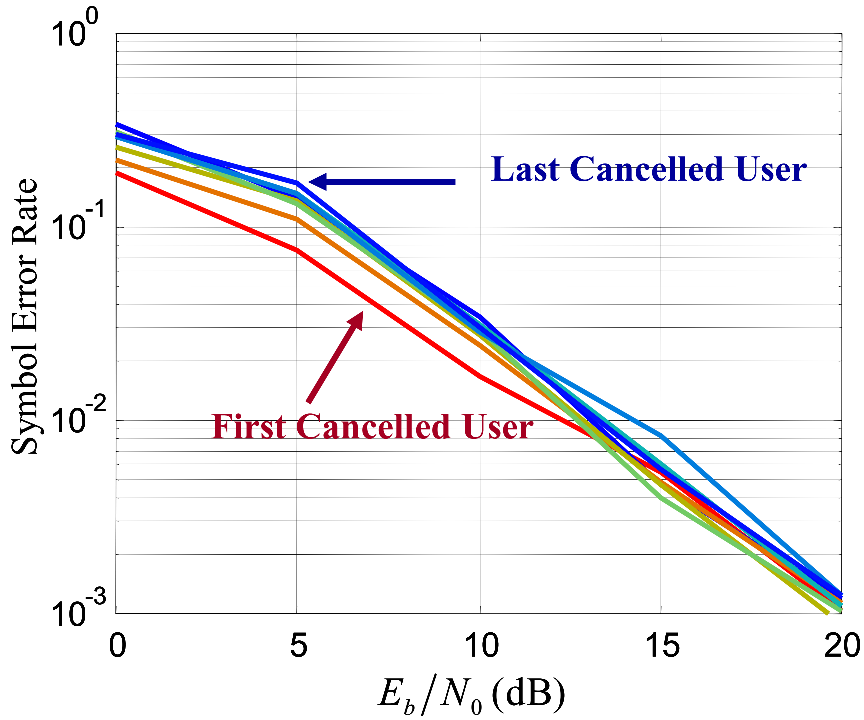


Figure 3.10: *Differentially Encoded* Symbol Error Rate for MAP-SIC, $M$= 4 element UCA, $D_u$= 8, 75% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

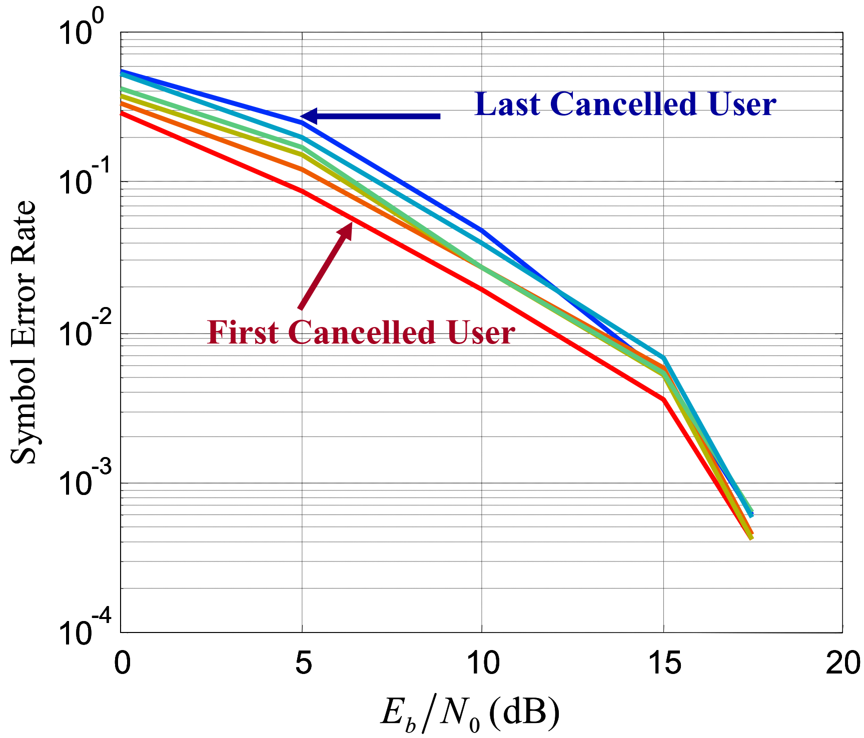


Figure 3.11: Message Symbol Error Rate for MAP-SIC, $M$= 4 element Diversity Array, $D_u$= 6, 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

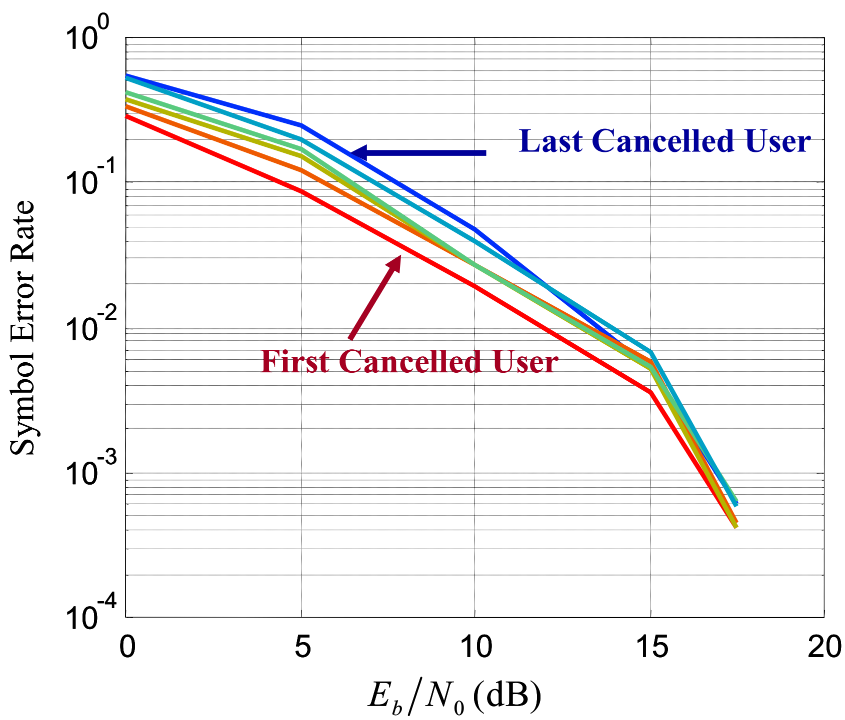


Figure 3.12: Diff. Encoded Symbol Error Rate for MAP-SIC, $M$= 4 element Diversity Array, $D_u$= 6, 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

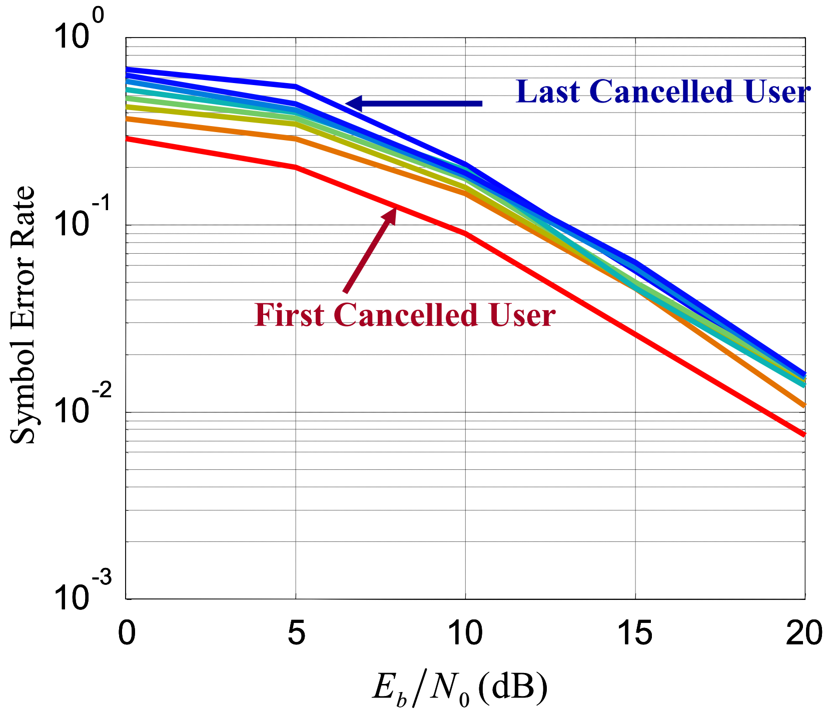


Figure 3.13: Message Symbol Error Rate for MAP-SIC, $M$= 4 element Diversity Array, $D_u$= 8, 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5.

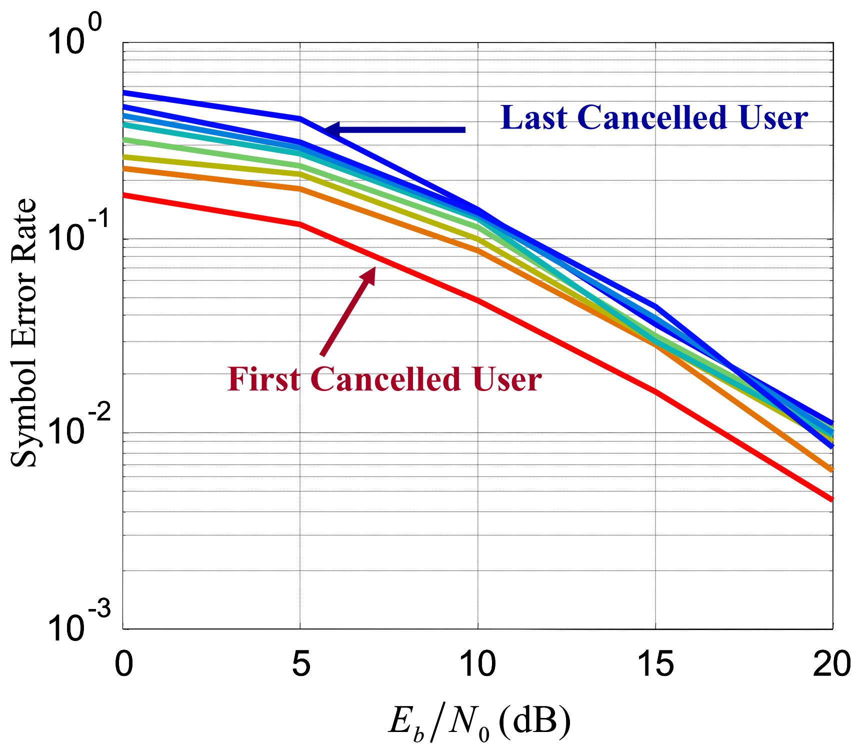


Figure 3.14: Diff. Encoded Symbol Error Rate for MAP-SIC, $M$= 4 element Diversity Array, $D_u$= 8, 35% rolloff, 15 symbol interference rejection filter with RS-MAP reduced state size of $\mu$= 5

# Chapter 4: Energy Focusing Transforms

## 4.1 Introduction

The OLAP approaches explored in previous chapters leveraged the ability of an interference rejection receiver to extract a signal from an overloaded environment. In this context, successive interference cancellation (SIC) may be viewed as a crude method of combining information from individual receivers in order to arrive at a joint estimate of all co-channel interfering signals. It is natural to ask if joint signal estimates can be achieved in a more integrated way. The joint maximum likelihood estimator (JMLE) provides a convenient starting point for answering this question. Although in most cases the JMLE is too expensive to implement in practice, attempting to approximate it can lead to novel signal extraction approaches. Both Successive Interference Cancellation (SIC) and Parallel Interference Cancellation (PIC) [19] may be thought of as JMLE approximations. This chapter will investigate the use of alternative JMLE approximations that leverage a synergistic approach to front-end linear pre-processing and back-end reduced complexity non-linear processing. We will build on the so-called spatially reduced search joint detection (SRSJD) approach of [16], which develops an OLAP algorithm for symbol-synchronous environments. Here, the receiver consists of a linear stage and a non-linear stage. The linear stage consists of a multi-input/multi-output (MIMO) beamformer that preserves the join-maximum likelihood criterion. This MIMO beamformer attempts to impose a structure on the channel matrix that can be exploited with an iterative-tail-biting variation of delayed decision feedback sequence estimation (ITB-DDFSE), a reduced state sequence estimation algorithm. We expect that refining this approach will eventually lead to space-time OLAP algorithms for asynchronous environments.

This chapter proposes an extension of SRSJD's original formulation that applies novel beam-forming criteria. These so called energy focusing transforms (EFT), linear transforms that focus the channel's energy to fit a specified pattern, are used to reduce the complexity of ITB-DDFSE. This chapter re-interprets the linear pre-processing used in [16] as a special type of energy focusing unitary transform (EFUT) and proposes new EFTs that improve symbol error rate (SER) performance. It is expected that future research will benefit from the energy-focusing transform concept. Hence, we will develop these EFTs in a general setting. However, in the final sections, we will return to the symbol-synchronous overloaded array problem to demonstrate the power of the approach.

## 4.2 JOINT MAXIMUM LIKELIHOOD ESTIMATION AND APPROXIMATIONS

Many communication systems may be reduced to the following discrete equivalent channel model.

$$\underline{y} = \mathbf{H}\underline{s} + \underline{z} \tag{4.1}$$

Here $\underline{s}$, is a Du$\times 1$ vector of transmitted symbols visible in the $D_i \times 1$ observation vector, $\underline{y}$. The symbols are assumed to be drawn independently with equal likelihood from a finite alphabet, $\mathcal{A}$. For instance, if the symbols are drawn from a QPSK alphabet, $\mathcal{A} = \{ 1, -1, j, -j\}$. The $D_i \times D_u$ matrix, $\mathbf{H}$, is the discrete equivalent channel model of the communications system. Finally, $\underline{z}$ is the $D_i \times 1$ complex noise vector with zero mean and auto–correlation matrix, $\Phi_{\underline{zz}} = E\left[\underline{zz}^H\right]$. For simplicity of discussion, we will assume that $\underline{z}$ is white in the range-space of $\mathbf{H}$ (see Appendix A). This linear model arises in many situations. For example, in array processing applications, $\underline{s}$ is a vector of symbols transmitted from independent signals impinging on the array with the same time of arrival; and the columns of $\mathbf{H}$ are the spatial signatures of their corresponding signals [7]. In equalization applications, the elements of $\underline{s}$ form a sequence of transmitted symbols and $\mathbf{H}$ is a Toeplitz matrix representation of convolution [34].

A joint maximum-likelihood (JML) estimate of the transmitted signal vector, $\underline{s}$, is any minimizing solution of the following [57].

$$\hat{\mathcal{S}} = \arg\min_{\underline{s} \in \mathcal{A}^{D_u}} \left\| \underline{y} - \mathbf{H}\underline{s} \right\|^2, \tag{4.2}$$

where $\hat{\mathcal{S}}$ is the set of minimizing solutions. Note, if the minimization were taken over the set $\{\underline{s} \in \mathbb{C}^{D_u}\}$, then (4.2) is a well known least-squares problem [31] efficiently solved with a linear estimator [57]. However, since the minimization is taken over a discrete alphabet, (4.2) can be very difficult to solve. If nothing is known about the structure of $\mathbf{H}$, then (4.2) may only be solved through an exhaustive search over all $|\mathcal{A}|^{D_u}$ possible values of $\underline{s}$. For instance, if $|\mathcal{A}| = 4$ and $D_u = 23$, then this would result in $4^{23} \approx 10^{14}$ evaluations of the JML cost function per received sample, which is prohibitively complex. However, if $\mathbf{H}$ is Toeplitz or, more generally, banded, then (4.2) is solved by the well known Viterbi Algorithm [21]. More recently [15]-,[16], it has been found that if $\mathbf{H}$ has energy focused on the diagonal, then it is often possible to very closely approximate the JML estimator with a variation of Delayed Decision Feedback Sequence Estimation DDFSE [22].

This chapter will investigate the application of one-sided linear transformations of (4.2) that can improve the complexity of the JML search. That is, given $\underline{y} \in \mathbb{C}^{D_i \times 1}$ and a number of outputs, $D_o$, find a matrix, $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$ such that

$$\begin{aligned}\underline{\tilde{y}} &= \mathbf{Q}^H(\mathbf{H}\underline{s} + \underline{z}) \\ &= \tilde{\mathbf{H}}\underline{s} + \underline{\tilde{z}}\end{aligned} \tag{4.3}$$

results in a least squares estimate that can be approximated with less complexity. Note, the estimator

$$\hat{\mathcal{S}} = \underset{\underline{s}\in\mathcal{A}^{D_u}}{\arg\min} \left\| \underline{\tilde{y}} - \tilde{\mathbf{H}}\underline{s} \right\|^2 \tag{4.4}$$

is not necessarily maximum likelihood, and hence could have a performance, measured in symbol error rate (SER), significantly better than the true maximum likelihood estimator. To distinguish it from the true ML estimator, we will call (4.4) the Transformed Enumerated Least Squares Estimate (TELSE).

If $\mathbf{Q}$ can produce an advantageous (e.g. banded) structure in the composite channel, $\tilde{\mathbf{H}} = \mathbf{Q}^H\mathbf{H}$, similar to that achieved in [16] then the TELSE estimator might be approximated with a reduced state sequence estimator without sacrificing SER performance. In fact, it is demonstrated in [16] that this structure need only be approximate. The desired structure of $\tilde{\mathbf{H}}$ is most easily described with a so-called sparsity pattern, $\mathbf{P}$, a $D_o \times D_u$ matrix of ones and zeros, where $p_{i,j} = 0$ indicates that $\tilde{h}_{i,j}$ should be as small as possible. For instance, the pattern

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \tag{4.5}$$

indicates that $\tilde{\mathbf{H}}$ is desired to be as close to a circulant-banded matrix as possible. Also of interest is the complement pattern $\overline{\mathbf{P}} = [\overline{p}_{i,j}]$, for-which $\overline{p}_{i,j} = 0$ if $p_{i,j} = 1$, and $\overline{p}_{i,j} = 1$, otherwise. An element, $\tilde{h}_{i,j}$ is said to be *in-pattern* if $p_{i,j} = 1$ and *off-pattern* if $p_{i,j} = 0$. In subsequent developments, it will be useful to define the following notation. For each output, $k = 1, \ldots, D_o$, define the ordered set $\mathcal{U}_e[k] = \{d = 1, \ldots, D_u : p_{k,d} = 1\}$, that lists the column indices of the ones in the sparsity pattern's $k^{th}$ row in the order that they occur. Also, form the ordered set $\overline{\mathcal{U}_e}[k] = \{d = 1, \ldots, D_u : p_{k,d} = 0\}$, that lists the zeros in the sparsity pattern's $k^{th}$ row. Let $N_e[k] = |\mathcal{U}_e[k]|$ be the number of ones that occur on the $k^{th}$ row of $\mathbf{P}$. For the example of Equation (4.5), $\mathbf{P}$ is a regular pattern with $\mathcal{U}_e[k] = \{10, 1, 2, 3\}$, $\overline{\mathcal{U}_e}[k] = \{4, 5, 6, 7, 8, 9\}$, and $N_e[k] = N_e = 4, \forall k$. It will be convenient to reference individual elements of $\mathcal{U}_e[k]$ with an additional subscript. For instance, $\mathcal{U}_e[k](1)$ denotes the first element of the set, $\mathcal{U}_e[k]$. Of special interest is when $\mathcal{U}_e[k]$ is contiguous subset of $\{1, \ldots, D_u\}$. That is, its elements are either consecutive integers in the range $\{1, \ldots, D_u\}$, or they wrap around the endpoints of $\{1, \ldots, D_u\}$. Finally, two sets, $\mathcal{U}_e[k_1]$ and $\mathcal{U}_e[k_2]$ are *well offset if* $\mathcal{U}_e[k_1] \cap \mathcal{U}_e[k_2]$ is either empty or contiguous.

A pattern is *tail-biting* if $\mathcal{U}_e[1] \cap \mathcal{U}_e[D_o] \neq \emptyset$. A *pattern of constant-width* has $N_e[k] = N_e$ for every $k$. A pattern is square if $D_u = D_o$. When $D_o = D_u$, a *uniform pattern* is a pattern for which a cyclic permutation of its rows yields a pattern for which, $\mathcal{U}_e[k] = \{k, \cdots, k + N_e - 1\}$, (wrapping indices back in the set $\{1, \ldots, D_u\}$, when appropriate). When $D_o < D_u$, a *uniform pattern* is any pattern that can be obtained by extracting every $\lceil D_u / D_o \rceil^{th}$ row from a square uniform pattern. A *well offset* sparsity pattern is one for which every consecutive pair of rows are well-offset. Finally, a sparsity pattern is *trellis-oriented*, if it is well-offset, and all its rows and columns are contiguous.

In [16], it is demonstrated that if the energy of $\tilde{\mathbf{H}}$ exactly fits a *trellis-oriented* sparsity pattern, then a TELSE can be efficiently implemented with the Viterbi Algorithm. Moreover, if the energy of $\tilde{\mathbf{H}}$ exactly fits a *tail-biting* sparsity pattern, then a TELSE can be efficiently implemented with the Tail-biting Viterbi Algorithm. Usually it is not possible to find a linear transform, $\mathbf{Q}$, such that the off-pattern elements of $\tilde{\mathbf{H}}$ are zero, but as this chapter will show, it is often possible to drive them to a acceptably small value. As demonstrated in [16], a carefully chosen pattern results in an $\tilde{\mathbf{H}}$ for which the TELSE can be closely approximated with a combination of sequence estimation and decision feedback. In such a setting, $\| \overline{\mathbf{P}} \circ (\mathbf{Q}^H \mathbf{H}) \|_F^2$ (where "$\circ$" denotes the element-wise or Hadamard product, and $\| \circ \|_F^2$ is the Frobinius matrix norm) is a measure of residual-interference.

JMLE approximations will be developed in two steps. First, we will apply linear transforms for which the resulting TELSE is either equivalent to the true ML estimate or a close approximation. Second, we will try to approximate the TELSE with a combination of sequence estimation and decision feedback. While the SER performance of such an approach is too difficult to analyze, a fundamental tradeoff exists (as with any decision feedback approach) between noise-enhancement and error-propagation. Noise-enhancement refers to an increase in output noise power along any given direction in signal space. Error-propagation refers to the adverse effect of incorrect symbol estimates upon subsequent symbol estimates in a decision feedback framework. An aggressive approach to reducing error-propagation, is to reduce the off-pattern elements of $\tilde{\mathbf{H}}$ without regards to the effect that such a transformation will have on the noise. However, if the resulting transformation, $\mathbf{Q}$, has large singular values, then the resulting noise auto-correlation, $\mathbf{\Phi}_{\underline{\tilde{z}\tilde{z}}} = E\left[ \underline{\tilde{z}\tilde{z}}^H \right] = \mathbf{Q}^H \mathbf{\Phi}_{\underline{zz}} \mathbf{Q}$ may have large singular values. When noise is uncorrelated, e.g. $\mathbf{\Phi}_{\underline{zz}} = \sigma_z^2 \mathbf{I}$, these singular values can be interpreted as the relative noise powers along selected directions in the signal space [47]. Excessive noise power along any direction in signal space can result in a TELSE with unacceptable SER degradation. On the other hand, a conservative approach to noise-enhancement results in an $\tilde{\mathbf{H}}$ with severe residual interference, increasing the estimator's susceptibility to error-propagation. A prudent choice of linear pre-processing should make the proper tradeoff between noise-enhancement and error-propagation. Such a choice is application specific, dependent on the input noise statistics and the geometry of the signal space. In cases of practical interest, this signal space is possible to analyze because it consists of a very large number of signaling points in possibly a large number of dimensions. Hence, this thesis proposes that this tradeoff be explored through simulation.

This chapter will develop approaches to investigate this tradeoff for weakly structured problems (i.e. when $\mathbf{H}$ has little structure). To accommodate issues relating to the number of outputs, we will consider $\mathbf{Q} \in \mathbb{C}^{D_o \times D_i}$ with arbitrary $D_o$. The first section provides a solution that maximizes the Signal to Interference and Noise Ratio output by each column of $\mathbf{Q}$ individually. This so called maximum individual Group SINR (MIGSINR) weakly controls noise enhancement by implicitly controlling the diagonal entries of $\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}$. It is by now well known that the MIGSINR linear transform can be found by solving a sequence of generalized eigen-vector problems (in other applications MIGSINR will go by another name). However, for clarity of development, we will re-derive this result in terms of a specified sparsity-pattern. The second section will investigate the application of linear transforms that do not enhance noise at all. While, the third section investigates the application of linear transforms that indirectly control noise- enhancement through $\mathbf{Q}$'s output noise variance. Numerical methods for approximating these approaches and numerical results are presented in the last two sections. While these numerical methods are not computationally efficient, they are suitable for investigating the noise-enhancement/error-propagation tradeoff.

## 4.3 Maximum Individual Group SINR Transforms (MIGSINR)

Upon close inspection, equation (4.3) can be decomposed into three parts,

$$\underline{\tilde{y}} = (\mathbf{P} \circ \tilde{\mathbf{H}})\underline{s} + (\overline{\mathbf{P}} \circ \tilde{\mathbf{H}})\underline{s} + \underline{\tilde{z}} \tag{4.6}$$

A reduced search joint detection scheme accounts for the first term, $\underline{\hat{\tilde{y}}} = (\mathbf{P} \circ \tilde{\mathbf{H}})\underline{s}$, with trellis-based detection and the second term or the interference term, $\underline{\tilde{y}}^{(i)} = (\bar{\mathbf{P}} \circ \tilde{\mathbf{H}})\underline{s}$, with decision feedback. A crude way to achieve a prudent tradeoff between noise-enhancement and error propagation is to consider both noise and residual interference as equally harsh impairments to the trellis-based detection process. Defining the error vector $\underline{\tilde{\varepsilon}} = \underline{\tilde{y}}^{(i)} + \underline{\tilde{z}}$, a metric for signal quality in the $k^{th}$ output is the signal to interference and noise ratio (SINR)

$$SINR_k\left(\underline{q}_k\right) = \frac{E\left[|\,\hat{\tilde{y}}_k\,|^2\right]}{E\left[|\,\tilde{\varepsilon}_k\,|^2\right]} = \frac{\sigma^2_{\hat{\tilde{y}}_k}}{\sigma^2_{\tilde{y}_k^{(i)}} + \sigma^2_{\tilde{z}_k}} \tag{4.7}$$

We seek a linear transform, $\mathbf{Q} = \left[\underline{q}_1 \quad \cdots \quad \underline{q}_{D_o}\right]$, that maximizes each SINR independently. Toward this end, form the matrix $\breve{\mathbf{H}}_k \in \mathbb{C}^{D_i \times N_e[k]}$ and $\bar{\mathbf{H}}_k \in \mathbb{C}^{D_i \times (D_u - N_e[k])}$ from the columns of $\mathbf{H}$ indexed by the elements of $\mathcal{U}_e[k]$ and $\bar{\mathcal{U}}_e[k]$ respectively. That is, $\breve{\mathbf{H}}_k = [p_{k,d}\underline{h}_d]_{d\in\mathcal{U}_e[k]}$ and $\bar{\mathbf{H}}_k = [\bar{p}_{k,d}\underline{h}_d]_{d\in\bar{\mathcal{U}}_e[k]}$. Then,
$\hat{\tilde{y}}_k = \underline{q}_k^H \breve{\mathbf{H}}_k \underline{\breve{s}}_k$ and $\varepsilon_k = \underline{q}_k^H\left(\bar{\mathbf{H}}_k \underline{\bar{s}}_k + \underline{z}_k\right)$. So, if

$$\breve{\mathbf{\Psi}}_k \triangleq E\left[\left(\breve{\mathbf{H}}_k \underline{\breve{s}}_k\right)\left(\breve{\mathbf{H}}_k \underline{\breve{s}}_k\right)^H\right] = \sigma_s^2 \breve{\mathbf{H}}_k \breve{\mathbf{H}}_k^H$$

and

$$\begin{aligned}\bar{\mathbf{\Psi}}_k &= E\left[\left(\bar{\mathbf{H}}_k \underline{\bar{s}}_k + \underline{z}_k\right)\left(\bar{\mathbf{H}}_k \underline{\bar{s}}_k + \underline{z}_k\right)^H\right] \\ &= \sigma_s^2 \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H + \mathbf{\Phi}_{\underline{zz}}\end{aligned} \tag{4.8}$$

then

$$SINR_k\left(\underline{q}_k\right) = \frac{\underline{q}_k^H \breve{\mathbf{\Psi}}_k \underline{q}_k}{\underline{q}_k^H \bar{\mathbf{\Psi}}_k \underline{q}_k}. \tag{4.9}$$

Note $\breve{\mathbf{\Psi}}_k$ and $\bar{\mathbf{\Psi}}_k$ are positive semi-definite by construction. It is by now well known, that maximizing the Generalized Rayleigh Ratio in Equation (4.9) is achieved by finding the generalized eigen-vector associated with the largest generalized eigen-value of the following system.

$$\breve{\mathbf{\Psi}}_k \underline{q}_k = \lambda \bar{\mathbf{\Psi}}_k \underline{q}_k. \tag{4.10}$$

The complexity of the algorithm is dominated by the calculation of the eigen-vectors. Note that the MIGSINR solution indirectly controls noise by controlling the error variance. However, as revealed by (4.7), only the diagonal elements of $\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}$ are controlled. It is possible that all of the diagonal elements of $\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}$ are reasonable, but one singular value of $\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}$ is excessively large. Thus, the severity of noise-enhancement is hidden by the short-sighted method of optimizing each individual output without considering the effect of all outputs to the entire receiver. This is not a concern in conventional array processing where decisions are made in each output individually.

## 4.4 Energy Focusing Unitary Transforms (EFUT)

An aggressive approach to control noise enhancement is to restrict the receiver to linear transforms that strictly preserve the maximum likelihood criterion. That is, restrict the estimator to only apply linear transforms for which the TELSE of Equation (4.4) is equivalent to the JMLE defined by Equation(4.2). This is assured if there exist positive real constants, $\kappa$, and $\beta$ such that

$$\left\|\underline{y} - \mathbf{H}\underline{s}\right\|^2 = \kappa \left\|\underline{\tilde{y}} - \tilde{\mathbf{H}}\underline{s}\right\|^2 + \beta \quad \forall \underline{\mathrm{s}} \in \mathbb{C}^{D_u}. \tag{4.11}$$

in which case, the two least squares cost functions are said to be *equivalent.* They are said to be equivalent without scaling if $\kappa = 1$. It is well known that unitary transforms preserve the Euclidean norm. However, there is a larger class of linear transforms that preserve the ML criterion as well: scaled, subspace-unitary transforms. The following theorem guarantees that this is the largest such class.

**Theorem:** Two least-squares cost functions, $\left\|\underline{y} - \mathbf{H}\underline{s}\right\|^2$, and $\left\|\underline{\tilde{y}} - \tilde{\mathbf{H}}\underline{s}\right\|^2$ are equivalent if and only if there exists a matrix, $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$, unitary with respect to $\mathcal{R}(\mathbf{H})$, and a $\kappa > 0$ such that

(i) $\mathbf{P}_{\mathcal{R}(\tilde{\mathbf{H}})}\underline{\tilde{y}} = \kappa \mathbf{Q}^H (\mathbf{P}_{\mathcal{R}(\mathbf{H})}\underline{y})$

(ii) $\tilde{\mathbf{H}} = \kappa \mathbf{Q}^H \mathbf{H}$

**Proof:**

Let $\kappa, \beta$ be defined as in (4.11). With $\beta_2 = \beta + \|\underline{\tilde{y}}\|^2 - \|\underline{y}\|^2$, it can be seen that (4.11) holds with $\kappa = 1$ if and only if

$$\underline{s}^H \mathbf{H}^H \mathbf{H}\underline{s} - 2\,\mathrm{Re}\{\underline{s}^H \mathbf{H}^H \underline{y}\} = \underline{s}^H \tilde{\mathbf{H}}^H \tilde{\mathbf{H}}\underline{s} - 2\,\mathrm{Re}\{\underline{s}^H \tilde{\mathbf{H}}^H \underline{\tilde{y}}\} + \beta_2$$

for every $\underline{s} \in \mathbb{C}^{D_u}$. This, in turn, can occur if and only if both $\mathbf{H}^H\mathbf{H} = \tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$ and $\mathbf{H}^H\underline{y} = \tilde{\mathbf{H}}^H\underline{\tilde{y}}$ or equivalently

$$\mathbf{H}^H\mathbf{H} = \tilde{\mathbf{H}}^H\tilde{\mathbf{H}} \tag{4.12}$$

and

$$\mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})}\underline{y} = \tilde{\mathbf{H}}^H \mathbf{P}_{\mathcal{R}(\tilde{\mathbf{H}})}\underline{\tilde{y}}\,. \tag{4.13}$$

We must then establish that (4.12) and (4.13) are equivalent to (i) and (ii). Now, suppose both (4.12) and (4.13) hold. Then, it must be the case that $\mathcal{R}(\mathbf{H}^H) = \mathcal{R}(\tilde{\mathbf{H}}^H)$ and moreover, by the matrix factorization lemma [56], there exists a $\mathcal{R}(\mathbf{H})$-unitary $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$ such that $\tilde{\mathbf{H}} = \mathbf{Q}^H \mathbf{H}$. The former guarantees that (4.13) has a solution, and in particular,

$$\begin{aligned} \mathbf{P}_{\mathcal{R}(\tilde{\mathbf{H}})} \underline{\tilde{y}} &= \tilde{\mathbf{H}}^{H\dagger} \mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \underline{y} \\ &= \mathbf{Q}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \underline{y} \ . \end{aligned}$$

Conversely, suppose (i) and (ii) hold. Then

$$\begin{aligned} \tilde{\mathbf{H}}^H \tilde{\mathbf{H}} &= \mathbf{H}^H \mathbf{Q} \mathbf{Q}^H \mathbf{H} \\ &= \mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{Q} \mathbf{Q}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{H} \\ &= \mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{H} \\ &= \mathbf{H}^H \mathbf{H} \end{aligned}$$

and multiplying both sides of (ii) by $\mathbf{H}^H \mathbf{Q} = \mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{Q}$, we have

$$\begin{aligned} (\mathbf{H}^H \mathbf{Q}) \mathbf{P}_{\mathcal{R}(\tilde{\mathbf{H}})} \underline{\tilde{y}} &= (\mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{Q}) \mathbf{Q}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \underline{y} \\ (\mathbf{Q}^H \mathbf{H})^H \mathbf{P}_{\mathcal{R}(\tilde{\mathbf{H}})} \underline{\tilde{y}} &= \mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \underline{y} \\ \tilde{\mathbf{H}}^H \mathbf{P}_{\mathcal{R}(\tilde{\mathbf{H}})} \underline{\tilde{y}} &= \mathbf{H}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \underline{y} \ , \end{aligned}$$

which establishes (4.12). In the above argument, if $\kappa \neq 1$, then the same logic applies with $\mathbf{Q}$ replaced by $\kappa \mathbf{Q}$. This concludes the proof.

Although the following corollary is easy to establish by itself, it is a convenient consequence of the preceding theorem.

**Corollary:** Given $\underline{y} \in \mathbb{C}^{D_i}$ and $\mathbf{H} \in \mathbb{C}^{D_i \times D_o}$, then the least squares cost functions $\left\| \underline{y} - \mathbf{H}\underline{s} \right\|^2$, and $\left\| \mathbf{P}_{\mathcal{R}(\mathbf{H})}\underline{y} - \mathbf{H}\underline{s} \right\|^2$ are equivalent.

Appendix A and the previous theorem reveal that the linear transforms preserving the maximum likelihood criterion are exactly those that do not color noise in the subspace of $\mathbf{H}$. This fact relates to Wozencraft's theorem of irrelevance [57], which, succinctly stated, says the portion of an observed signal outside the signal space, $\mathcal{R}(\mathbf{H})$, neither contributes to nor detracts from the performance of the maximum likelihood estimate. For the remainder of this chapter, we will assume that $\underline{y} \in \mathcal{R}(\mathbf{H})$. Together with the assumption that $\underline{z}$ is white in $\mathcal{R}(\mathbf{H})$, this implies that there exists a number, $\sigma_z^2$, such that $\mathbf{\Phi}_{\underline{zz}} = \sigma_z^2 \mathbf{P}_{\mathcal{R}(\mathbf{H})}$.

We are now ready to define the energy focusing problem, the problem of finding a linear transform that focuses the energy of $\mathbf{H}$ within a specified pattern, $\mathbf{P}$. This intuitive goal should be formally defined through the use of a cost function. First, such a cost function should both penalize the residual, $\overline{\mathbf{P}} \circ (\mathbf{Q}^H \mathbf{H})$, and reward focused energy, $\mathbf{P} \circ (\mathbf{Q}^H \mathbf{H})$. Secondly, since the maximum likelihood cost function is invariant to scaling factors (i.e. if $\mathbf{Q}$ preserves the maximum likelihood metric then so does $\alpha\mathbf{Q}$ for every $\alpha \neq 0$), the cost function should also be invariant to scaling factors. Along these lines, the following EFUT cost function is proposed

$$J(\mathbf{Q}) = \frac{\| \mathbf{P} \circ (\mathbf{Q}^H \mathbf{H}) \|_F^2}{\| \overline{\mathbf{P}} \circ (\mathbf{Q}^H \mathbf{H}) \|_F^2} \tag{4.14}$$

which measures the ratio of the in-pattern energy of $\tilde{\mathbf{H}} = \mathbf{Q}^H \mathbf{H}$ to the off-pattern energy. Equation (4.14) may also be interpreted as a total signal to total interference ratio (TSTIR). Thus, the best energy focusing unitary transform is any solution to the following optimization problem.

$$\max\left\{\frac{\|\mathbf{P}\circ(\mathbf{Q}^H\mathbf{H})\|_F^2}{\|\overline{\mathbf{P}}\circ(\mathbf{Q}^H\mathbf{H})\|_F^2}:\mathbf{Q}\in U_{\mathcal{R}(\mathbf{H})}^{D_i\times D_o}\right\} \tag{4.15}$$

This is a non-convex function with a set of quadratic (non-convex) constraints. Equation (4.15) is too difficult to optimize but lends insight to other approaches and (4.14) stands as a figure of merit for competing energy focusing solutions.

Equation (4.14) has a number of properties that will prove insightful later. First, it is invariant to phase rotations of the columns of $\mathbf{Q}$ and constant scaling factors. That is, if $\alpha\neq 0,\ \phi_k\in[0,2\pi]$, $k=1,\ldots,D_o$, and

$$\mathbf{\Lambda}_{\underline{\phi}}=\begin{bmatrix} e^{j\phi_1} & & 0 \\ & \ddots & \\ 0 & & e^{j\phi_{D_o}}\end{bmatrix}, \tag{4.16}$$

then

$$\begin{aligned} J(\alpha\mathbf{Q}\mathbf{\Lambda}_{\underline{\phi}}) &= \frac{\sum_{k=1}^{D_o}\sum_{d=1}^{D_u}|\,p_{k,d}(\alpha e^{j\phi_k}\underline{q}_k^H)\underline{h}_d\,|^2}{\sum_{k=1}^{D_o}\sum_{d=1}^{D_u}|\,\overline{p}_{k,d}(\alpha e^{j\phi_k}\underline{q}_k^H)\underline{h}_d\,|^2} \\ &= \frac{\sum_{k=1}^{D_o}\sum_{d=1}^{D_u}|\,p_{k,d}\underline{q}_k^H\underline{h}_d\,|^2}{\sum_{k=1}^{D_o}\sum_{d=1}^{D_u}|\,\overline{p}_{k,d}\underline{q}_k^H\underline{h}_d\,|^2} \\ &= J(\mathbf{Q}). \end{aligned}$$

Secondly, both the numerator and denominator have a quadratic expansion. Defining

$$\mathbf{\Phi}_k=\mathbf{H}\begin{bmatrix} p_{k,1} & & 0 \\ & \ddots & \\ 0 & & p_{k,D_u}\end{bmatrix}\mathbf{H}^H=\breve{\mathbf{H}}_k\breve{\mathbf{H}}_k^H$$

and

$$\bar{\mathbf{\Phi}}_k = \mathbf{H}\begin{bmatrix} \bar{p}_{k,1} & & 0 \\ & \ddots & \\ 0 & & \bar{p}_{k,D_u} \end{bmatrix}\mathbf{H}^H = \bar{\mathbf{H}}_k \bar{\mathbf{H}}_k^H \ ,$$

we have

$$\begin{aligned} J(\mathbf{Q}) &= \frac{\sum_{k=1}^{D_o}\sum_{d=1}^{D_u} | p_{k,d} \underline{q}_k^H \underline{h}_d |^2}{\sum_{k=1}^{D_o}\sum_{d=1}^{D_u} | \bar{p}_{k,d} \underline{q}_k^H \underline{h}_d |^2} \\ &= \frac{\sum_{k=1}^{D_o} \underline{q}_k^H \left( \sum_{d=1}^{D_u} p_{k,d} \underline{h}_d \underline{h}_d^H \right) \underline{q}_k}{\sum_{k=1}^{D_o} \underline{q}_k^H \left( \sum_{d=1}^{D_u} \bar{p}_{k,d} \underline{h}_d \underline{h}_d^H \right) \underline{q}_k} \\ &= \frac{\sum_{k=1}^{D_o} \underline{q}_k^H \mathbf{\Phi}_k \underline{q}_k}{\sum_{k=1}^{D_o} \underline{q}_k^H \bar{\mathbf{\Phi}}_k \underline{q}_k}. \end{aligned} \tag{4.17}$$

Equation (4.17) is a special form of a Rayleigh Ratio whose numerator and denominator are separable in the columns of $\mathbf{Q}$ .

### 4.4.1 Example

It is interesting to investigate the consequences of omitting the subspace unitary constraints from (4.15). Consider two cases. First, if $Ker(\bar{\mathbf{\Phi}}_k) \cap \mathcal{R}(\mathbf{\Phi}_k) \neq \emptyset$ for any k, then choosing $\underline{q}_k \in Ker(\bar{\mathbf{\Phi}}_k) \cap \mathcal{R}(\mathbf{\Phi}_k)$ and $\underline{q}_j = \underline{0}, \ \forall j \neq k$ , makes the numerator of (4.17) non-zero but the denominator zero, making $J(\mathbf{Q}) = \infty$. Secondly, consider the case when $Ker(\bar{\mathbf{\Phi}}_k) \cap \mathcal{R}(\mathbf{\Phi}_k) = \emptyset$ for every k, then maximizing the separable Rayleigh ratio in (4.17) can be achieved in three steps. First, find the maximum generalized eigen-vector of each system $\mathbf{\Phi}_k \underline{v} = \lambda_k \overline{\mathbf{\Phi}}_k \underline{v}$, $k = 1,\ldots,D_o$. Second, find any index, $k_o$, such that $\lambda_{k_0} \le \lambda_k$, $\forall k$. Third, set $\underline{q}_{k_0}$ to be the eigen-vector associated with $\lambda_{k_0}$ and $\underline{q}_j = \underline{0}$, $\forall j \ne k_0$. In both cases, there is a trivial maximizing solution of (4.17), $\mathbf{Q}$, that has only one non-zero column! This example will later help illustrate the extreme behavior of energy focusing solutions with relaxed unitary constraints. It is believed that $J(\mathbf{Q})$ is bounded on the constrained set, $U^{D_i \times D_o}_{\mathcal{R}(\mathbf{H})} \subseteq \mathbb{C}^{D_i \times D_o}$, but this is difficult to prove.

### 4.4.2 Simplified EFUT

The best EFUT problem in (4.15) can be simplified as follows. Since $\mathbf{P} + \overline{\mathbf{P}} = \underline{1}\,\underline{1}^T$ (i.e. the all ones matrix) and $\mathbf{P} \circ \overline{\mathbf{P}} = \mathbf{0}$ (i.e. the patterns are complementary), then $\|\mathbf{Q}^H\mathbf{H}\|_F^2 = \|(\mathbf{P}+\overline{\mathbf{P}}) \circ (\mathbf{Q}^H\mathbf{H})\|_F^2 = \|\mathbf{P} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2 + \|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2$. Further, if $\mathbf{Q} \in U^{D_i \times D_o}_{\mathcal{R}(\mathbf{H})}$, then $\|\mathbf{Q}^H\mathbf{H}\|_F^2 = \|\mathbf{H}\|_F^2$ and hence, we have

$$0 \le \|\mathbf{H}\|_F^2 = \|\mathbf{P} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2 + \|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2. \tag{4.18}$$

Substituting $\|\mathbf{P} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2 = \|\mathbf{H}\|_F^2 - \|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2$ into (4.14), we have

$$\begin{aligned} J(\mathbf{Q}) &= \frac{\|\mathbf{H}\|_F^2 - \|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2}{\|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2} \\ &= \frac{\|\mathbf{H}\|_F^2}{\|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2} - 1 \quad , \end{aligned} \tag{4.19}$$

which is easily recognized as a strictly decreasing function of $\|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2$. Hence, the maximization problem of (4.15) is equivalent to solving the following optimization problem.

$$\min\left\{\|\overline{\mathbf{P}} \circ (\mathbf{Q}^H\mathbf{H})\|_F^2 : \mathbf{Q} \in U^{D_i \times D_o}_{\mathcal{R}(\mathbf{H})}\right\} \tag{4.20}$$

In other words, because sub-space unitary transforms preserve energy, the best EFURT problem can be solved by minimizing the off-pattern energy. Similarly, the best EFUT problem may be solved by maximizing the in-pattern energy.

$$\max\left\{\|\mathbf{P}\circ(\mathbf{Q}^H\mathbf{H})\|_F^2 : \mathbf{Q}\in U_{\mathcal{R}(\mathbf{H})}^{D_i\times D_o}\right\} \tag{4.21}$$

As with (4.15), the unitary constraint in (4.20) renders it difficult to solve analytically; however, because the cost function of (4.20) is convex, it is significantly easier to solve numerically than (4.15). Numerical EFUT solutions are discussed in a subsequent section.

On a final note, it is often the case that $\mathbf{H}$, itself, has no apparent structure, but $\mathbf{H}^H\mathbf{H}$ reveals an underlying physical structure. For example, when $\mathbf{H}$ is a matrix of CDMA spreading codes, $\mathbf{H}^H\mathbf{H}$ is the so called code cross-correlation matrix [19]; or in array processing when $\mathbf{H}$ is a matrix of spatial signatures, $\mathbf{H}^H\mathbf{H}$ reveals which signals are closely spaced in angle-of-arrival [16]. In both cases, $\mathbf{H}^H\mathbf{H}$ is a weakly diagonally dominant matrix whose spectral square root, $\tilde{\mathbf{H}}=(\mathbf{H}^H\mathbf{H})^{1/2}$, preserves the diagonal dominance. In [16], this property is exploited to design a reduced-complexity overloaded array processing algorithms. As the next theorem will show, this approach is equivalent to solving an alternative EFUT formulation closely resembling (4.21), with the special choice of $D_o = D_u$ and pattern, $\mathbf{P}=\mathbf{I}$; namely,

$$\max\left\{\|\mathbf{I}\circ(\mathbf{Q}^H\mathbf{H})\|_1 : \mathbf{Q}\in U_{\mathcal{R}(\mathbf{H})}^{D_i\times D_o}\right\}, \tag{4.22}$$

where $\|\circ\|_1$ is the $l_1$-matrix norm defined as $\|\mathbf{A}\|_1=\sum_{i,j}|a_{i,j}|$ [31].

**Theorem**[1]**:** Let $\mathbf{H} = \mathbf{U\Sigma V}^H$ be the small singular value decomposition of $\mathbf{H}$. Then $\mathbf{Q} = \mathbf{UV}^H \in U_{D_i \times D_o}^{\mathcal{R}(\mathbf{H})}$ is a minimizing solution of (4.22).

**Proof:**

Expanding $\| \mathbf{I} \circ (\mathbf{Q}^H \mathbf{H}) \|_1 = \sum_{k=1}^{D_o} | \underline{q}_k^H \underline{h}_k |$ we can see that $\| \mathbf{I} \circ (\mathbf{Q}^H \mathbf{H}) \|_1$ is invariant to phase rotations of the columns of $\mathbf{Q}$. Hence, for every maximizing solution of (4.22), there is another maximizing solution such that $\| \mathbf{I} \circ (\mathbf{Q}^H \mathbf{H}) \|_1 = \sum_{k=1}^{D_o} | \underline{q}_k^H \underline{h}_k | = \sum_{k=1}^{D_o} \mathrm{Re}\left\{ \underline{q}_k^H \underline{h}_k \right\} = \mathrm{Re}\, Tr\left\{ \mathbf{Q}^H \mathbf{H} \right\}$. Therefore, a maximizing solution of

$$\max\left\{ \mathrm{Re}\, Tr\{\mathbf{Q}^H \mathbf{H}\} : \mathbf{Q} \in U_{\mathcal{R}(\mathbf{H})}^{D_i \times D_o}, \right\} \tag{4.23}$$

is a maximizing solution of (4.22) as well. Applying the identity $Tr\{\mathbf{AB}\} = Tr\{\mathbf{BA}\}$, we have $\mathrm{Re}\, Tr\{\mathbf{Q}^H \mathbf{H}\} = \mathrm{Re}\, Tr\left\{ \mathbf{Q}^H \mathbf{U\Sigma V}^H \right\} = \mathrm{Re}\, Tr\left\{ \mathbf{V}^H \mathbf{Q}^H \mathbf{U\Sigma} \right\}$. Letting $\mathbf{Z}_{\mathbf{QH}} = \mathbf{V}^H \mathbf{Q}^H \mathbf{U}$, we have

$$\begin{aligned}
&\mathrm{Re}\, Tr\{\mathbf{Q}^H \mathbf{H}\} = \mathrm{Re}\, Tr\left\{ \mathbf{Z}_{\mathbf{QH}} \mathbf{\Sigma} \right\} \\
&= \sum_{i=1}^{rank(\mathbf{H})} \mathrm{Re}\{[\mathbf{Z}_{\mathbf{QH}}]_{i,i}\} \sigma_i \\
&\leq \sum_{i=1}^{rank(\mathbf{H})} \sigma_i
\end{aligned}$$

with equality if and only if $\mathbf{I} = \mathbf{Z}_{\mathbf{QH}} = \mathbf{V}^H \mathbf{Q}^H \mathbf{U}$ which is achieved when $\mathbf{Q} = \mathbf{UV}^H$. Moreover, $\mathbf{U}^H \mathbf{Q} \mathbf{Q}^H \mathbf{U} = \mathbf{V}^H \mathbf{V} = \mathbf{I}$ establishing that $\mathbf{Q}$ is unitary with respect to $\mathcal{R}(\mathbf{H})$. Finally, $\tilde{\mathbf{H}} = \mathbf{Q}^H \mathbf{H} = (\mathbf{UV}^H)^H \mathbf{U\Sigma V}^H = \mathbf{V\Sigma V}^H \triangleq (\mathbf{H}^H \mathbf{H})^{1/2}$. This completes the proof.

[1] This theorem and its proof is attributed to Dr. Chris Beattie.

For convenience, we will call the choice of $\mathbf{Q}$ in the previous theorem the Procrustean EFUT solution because its proof closely resembles the development of Procrustean orthogonalization [31]. Numerical experiments suggest that for square, uniform patterns (i.e. $D_o = D_u$ and $N_e[k]$ constant for every $k$), the best EFUT solution yields very little improvement over the Procrustean EFUT.

## 4.5 Energy Focusing Unitary (Relaxed) Transforms (EFURT)

The previous two sections prescribed energy focusing solutions that achieved their goal with two extreme approaches to noise enhancement. The MIGSINR solution controls noise only by controlling the noise power observed in each individual output. This liberal treatment of noise allows sufficient room to achieve a superior reduction of residual interference but at the expense of potentially detrimental noise enhancement. In contrast, EFUTs allow no noise enhancement to occur, but by doing so possibly have poor residual interference suppression. This tradeoff begs the question: is it possible to make a more prudent tradeoff between noise enhancement, and residual interference suppression? Such a question hinges on a meaningful definition of noise coloring. The dual of a coloring measure, a spectral flatness measure, has been considered by researchers in the spectral estimation literature [40]. Motivated by the spirit of these attempts, this section attempts to define a meaningful measure of noise coloring that easily lends itself to the energy focusing problem.

First of all, a coloring measure should certainly be zero when a random vector is not colored. Secondly, since noise coloring is essentially a feature of the shape of a random vector's spectrum, a coloring measure should only describe the shape of the random variable's spectrum, and should not be a function of the spectrum's absolute magnitude. That is to say, it is meaningful to speak of two random vectors; one with large spectrum values, the other with low spectrum values, but both have the same coloring measure. Put another way, a coloring measure should be invariant to scaling factors. Thirdly, a coloring measure should only measure coloring in the estimator's signal space. For example, it is easy to conceive of two different energy focusing transforms. One introduces a severe amount of noise coloring outside the estimator's signal space but only moderately colors noise in the receiver's signal space, while the other does completely the opposite. A coloring measure should favor the former over the later. Fourthly, it is desired but not necessary that a coloring measure be normalized in the range [0, 1]. Finally, a coloring measure should reflect only meaningful distortion of the maximum likelihood function. We will find that this final coloring measure requirement is not unrelated to the previous three.

Let us assess the discrepancy between the TELSE cost function and the ML cost function. A previous theorem tells us that we may consider only observations, $\underline{y} \in \mathcal{R}(\mathbf{H})$, without a loss of generality. Recall that the TELSE cost function can be expanded as follows.

$$\begin{aligned}
&\| \underline{\tilde{y}} - \tilde{\mathbf{H}}\underline{s} \|^2 = \| \mathbf{Q}^H (\underline{y} - \mathbf{H}\underline{s}) \|^2 \\
&= \| \mathbf{Q}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} (\underline{y} - \mathbf{H}\underline{s}) \|^2 \\
&= (\underline{y} - \mathbf{H}\underline{s})^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{Q}\mathbf{Q}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} (\underline{y} - \mathbf{H}\underline{s})
\end{aligned}$$

Further, since a least squares cost function is equivalent to scaled versions of itself, we may scale it by a real positive scaling factor, say $\sqrt{\kappa}$. And for a meaningful measure of discrepancy between the TELSE and ML cost functions, we may minimize their difference over the scaling factor, $\kappa$.

$$\begin{aligned}
&\min_{\kappa>0} \left( \| \sqrt{\kappa} (\underline{\tilde{y}} - \tilde{\mathbf{H}}\underline{s}) \|^2 - \| \underline{y} - \mathbf{H}\underline{s} \|^2 \right) \\
&= \min_{\kappa>0} (\underline{y} - \mathbf{H}\underline{s})^H \left( \kappa \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{Q}\mathbf{Q}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} - \mathbf{P}_{\mathcal{R}(\mathbf{H})} \right) (\underline{y} - \mathbf{H}\underline{s}),
\end{aligned}$$

which tends to zero as

$$\xi(\mathbf{Q}) = \min_{\kappa>0} \frac{1}{r_H} \| \kappa \mathbf{P}_{\mathcal{R}(\mathbf{H})} \mathbf{Q}\mathbf{Q}^H \mathbf{P}_{\mathcal{R}(\mathbf{H})} - \mathbf{P}_{\mathcal{R}(\mathbf{H})} \|^2 \tag{4.24}$$

tends to zero. Moreover, if $\mathbf{V}_H$ is a matrix whose columns form an orthonormal basis for $\mathcal{R}(\mathbf{H})$, then (4.24) is equivalent[2] to

$$\xi(\mathbf{Q}) = \min_{\kappa>0} \frac{1}{r_H} \| \kappa \mathbf{V}_\mathbf{H}^H \mathbf{Q}\mathbf{Q}^H \mathbf{V}_\mathbf{H} - \mathbf{I} \|^2 \tag{4.25}$$

While (4.25) will provide an easy interpretation of what $\xi(\mathbf{Q})$ measures, we can easily solve for the minimum of (4.25) to obtain an expression that is easy to compute. Solving for the scaling constant that minimizes (4.25),

$$\kappa_{\min} = \frac{\| \mathbf{V}_\mathbf{H}^H \mathbf{Q} \|_F^2}{\| \mathbf{V}_\mathbf{H}^H \mathbf{Q}\mathbf{Q}^H \mathbf{V}_\mathbf{H} \|_F^2}, \tag{4.26}$$

and substituting into (4.25), we obtain

$$\xi(\mathbf{Q}) = 1 - \frac{\| \mathbf{V}_\mathbf{H}^H \mathbf{Q} \|_F^4}{r_H \| \mathbf{V}_\mathbf{H}^H \mathbf{Q}\mathbf{Q}^H \mathbf{V}_\mathbf{H} \|_F^2}. \tag{4.27}$$

First it is easy to see from (4.25) that $\xi(\mathbf{Q}) \geq 0$ and from (4.27) that $\xi(\mathbf{Q}) \leq 1$. To understand what (4.25) measures, define $\tilde{\mathbf{Q}} = \mathbf{V}^H \mathbf{Q}$ and let $\tilde{\mathbf{Q}} = \tilde{\mathbf{U}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{V}}^H$ be a small SVD for $\tilde{\mathbf{Q}}$. $\tilde{\mathbf{Q}}$ may be interpreted as a representation of $\mathbf{P}_{\mathcal{R}(\mathbf{H})}\mathbf{Q}$ operating on points in a coordinate system local to $\mathcal{R}(\mathbf{H})$. Then,

---

[2] Recall that $r_H \triangleq rank(\mathbf{H})$

$$\begin{aligned}\xi(\mathbf{Q}) &= \min_{\kappa>0}\frac{1}{r_H}\|\kappa\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^H-\mathbf{I}\|^2\\ &= \min_{\kappa>0}\frac{1}{r_H}\|\kappa\tilde{\mathbf{V}}\tilde{\mathbf{\Sigma}}^2\tilde{\mathbf{V}}^H-\mathbf{I}\|^2\\ &= \min_{\kappa>0}\frac{1}{r_H}\|\kappa\tilde{\mathbf{\Sigma}}^2-\mathbf{I}\|^2\\ &= \min_{\kappa>0}\frac{1}{r_H}\sum_{i=1}^{r_H}(\kappa\tilde{\sigma}_i^2-1)^2\end{aligned} \tag{4.28}$$

Hence, (4.25) measures the average deviation of the linear transform $\mathbf{Q}$'s singular values local to $\mathcal{R}(\mathbf{H})$. Now, we claim that (4.25) is a coloring measure for $\underline{\tilde{z}}$ in $\mathcal{R}(\mathbf{Q}^H\mathbf{H})$. First note that $\tilde{\mathbf{V}}$ is an orthonormal basis for $\mathcal{R}(\tilde{\mathbf{Q}}^H)=\mathcal{R}(\mathbf{Q}^H\mathbf{V})$ $=\mathcal{R}(\mathbf{Q}^H\mathbf{H})$. Hence if we define $\underline{\tilde{v}}=\tilde{\mathbf{V}}^H\underline{\tilde{z}}$, then $\underline{\tilde{v}}$ may be interpreted as the coordinates of $\underline{\tilde{z}}$ in a coordinate system local to $\mathcal{R}(\mathbf{Q}^H\mathbf{H})$. Moreover,
$\mathbf{\Phi}_{\underline{\tilde{v}}\underline{\tilde{v}}}=\tilde{\mathbf{V}}^H\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}\tilde{\mathbf{V}}=\tilde{\mathbf{V}}^H\mathbf{Q}^H\mathbf{\Phi}_{\underline{z}\underline{z}}\mathbf{Q}\tilde{\mathbf{V}}=\tilde{\mathbf{V}}^H\mathbf{Q}^H(\sigma_z^2\mathbf{V}\mathbf{V}^H)\mathbf{Q}\tilde{\mathbf{V}}$
$=\sigma_z^2\tilde{\mathbf{V}}^H(\mathbf{Q}^H\mathbf{V})(\mathbf{V}^H\mathbf{Q})\tilde{\mathbf{V}}=\sigma_z^2\tilde{\mathbf{V}}^H\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^H\tilde{\mathbf{V}}=\sigma_z^2\tilde{\mathbf{\Sigma}}^2$. So,

$$\begin{aligned}&\min_{\kappa>0}\frac{1}{r_H}\|\kappa\mathbf{\Phi}_{\underline{\tilde{v}}\underline{\tilde{v}}}-\mathbf{I}\|_F^2\\ &\quad= \min_{\kappa>0}\frac{1}{r_H}\|\kappa\sigma_z^2\tilde{\mathbf{\Sigma}}^2-\mathbf{I}\|_F^2\\ &\quad= \min_{\kappa>0}\frac{1}{r_H}\|\kappa\tilde{\mathbf{\Sigma}}^2-\mathbf{I}\|_F^2\\ &\quad= \xi(\mathbf{Q})\end{aligned} \tag{4.29}$$

Equation (4.29) measures the deviation of the spectra of $\underline{\tilde{z}}$ in $\mathcal{R}(\mathbf{Q}^H\mathbf{H})$ from a purely white spectra. It is obviously zero when $\underline{\tilde{z}}$ is white, scale invariant, and measures coloring only in the estimator's signal space. Hence, the proposed coloring measure of (4.25) has all the desirable properties of a coloring measure. However, the form of (4.27) is preferred since it is differentiable with respect to $\mathbf{Q}$. Note, as mentioned earlier, implicit in this discussion is the assumption that $\underline{z} \in \mathcal{R}(\mathbf{H})$. If this is not true, the relationship in (4.29) breaks down.

Armed with a coloring measure, we are now ready to define the energy focusing unitary (relaxed) transform (EFURT). The best EFURT problem is to find an energy focusing transform that maximizes (4.14) with constrained noise coloring, as defined by (4.25). To avoid behavior similar to that of Example 4.4.1, it will help to add additional constraints to assure that a specified set of $\tilde{\mathbf{H}}$ elements, indexed by $I_r = \{(k, d_k)\}$, each achieve a minimum acceptable amplitude, $\psi_k$. The best EFURT problem is to find any maximizing solution to the following problem.

$$
\begin{aligned}
&\max_{\mathbf{Q} \in \mathbb{C}^{D_o \times D_i}} \frac{\| \mathbf{P} \circ (\mathbf{Q}^H \mathbf{H}) \|_F^2}{\| \bar{\mathbf{P}} \circ (\mathbf{Q}^H \mathbf{H}) \|_F^2} \\
&\text{s.t.} \quad \xi(\mathbf{Q}) \le \xi_0 \\
&\qquad | \underline{q}_k^H \underline{h}_{d_k} | \ge \psi_k, \quad \forall (k, d_k) \in I_r
\end{aligned} \tag{4.30}
$$

Here, $\xi_0$ is a specified coloring bound. The last constraints, the so called regularization constraints, can be used to prevent one row of $\tilde{\mathbf{H}}$ from dominating others. As demonstrated in Example 4.4.1, if $\xi_o = \infty$, and $\psi_k = 0, \ \forall k$ are chosen (i.e. there is no constraint on coloring and no regularization), then maximizing the unregularized EFURT cost function can be achieved with a trivial solution, where $\mathbf{Q}$ has only one non-zero column. If $\xi_o$ is chosen to be large, then without the regularization constraints, a single column of $\mathbf{Q}$ can strongly dominate others. This is prevented in (4.30) with the last set of constraints. Both the regularization targets, $d_k$, and the regularization thresholds, $\xi_k$, may be obtained from the Procrustean solution or by some other method.

## 4.6 Numerical Approximations

While (4.30) is arguably the best-EFURT formation that is easiest to interpret, it is very difficult to solve. Hence, in this section, we will develop alternative cost functions that attempt to approximate (4.30) but are also reasonable to optimize. Motivated by Equation (4.20), if the coloring constraint of $\mathbf{Q}$ is kept sufficiently small, the following optimization problem provides a reasonable approximation of (4.30).

$$\begin{aligned} &\min_{\mathbf{Q}\in\mathbb{C}^{D_O\times D_I}} \|\overline{\mathbf{P}}\circ(\mathbf{Q}^H\mathbf{H})\|_F^2 \\ &\text{s.t.}\quad \xi(\mathbf{Q})\le\xi_0 \\ &\qquad |\underline{q}_k^H\underline{h}_{d_k}|\ge\psi_k,\quad \forall(k,d_k)\in I_r \end{aligned} \tag{4.31}$$

A common way to numerically solve such optimization problems is to apply *penalty-function methods* [59], in which the solution of a constrained optimization problem is approximated with the solution of an unconstrained optimization problem by adding terms to the cost function that penalize a violation of the constraints. In such methods, the following function is useful for defining penalties applicable to inequality constraints.

$$|x|_+=\begin{cases} x & x\ge 0 \\ 0 & x<0 \end{cases} \tag{4.32}$$

Staying true to the spirit of applying penalty function methods, it is reasonable to minimize

$$\|\overline{\mathbf{P}}\circ(\mathbf{Q}^H\mathbf{H})\|_F^2+\rho\left|\xi(\mathbf{Q})-\xi_0\right|_+^2+\delta\sum_{(k,d_k)\in I_r}\left|\psi_k-|\underline{q}_k^H\underline{h}_{d_k}|\right|_+^2 \tag{4.33}$$

Here $\rho$ and $\delta$ are the so called penalty function gains. By choosing $\rho$ and $\delta$ large, (4.33) places a great deal of emphasis on the constraint penalties. Hence, its minimization results in solutions either within or close to the constrained solution set. Unfortunately, (4.33) is non-differentiable and the coloring term is eighth-order, rendering it difficult to optimize. Hence, we propose one final modification of the penalty function in (4.33). Consider,

$$J_{\rho,\delta}(\mathbf{Q}) = \| \bar{\mathbf{P}} \circ (\mathbf{Q}^H \mathbf{H}) \|_F^2 + \rho \| \mathbf{V}_{\mathbf{H}}^H \mathbf{Q} \mathbf{Q}^H \mathbf{V}_{\mathbf{H}} - \mathbf{I} \|_F^2 + \delta \sum_{(k,d_k) \in I_r} \left| \psi_k^2 - | \underline{q}_k^H \underline{h}_{d_k} |^2 \right|_+^2 \quad (4.34)$$

This function is differentiable and fourth-order but has the dis-advantage that the coloring penalty is neither scale-invariant nor is it strongly linked to a coloring bound. However, by varying the coloring penalty gain, $\rho$, one may obtain a continuous tradeoff between residual interference reduction and coloring penalty. Since (4.34) is differentiable, when $\rho, \delta$ are fixed, $J_{\rho,\delta}(\mathbf{Q})$ can be minimized with a simple gradient descent algorithm. In this case, it is convenient to define the gradient of $J_{\rho,\delta}(\mathbf{Q})$ with respect to a matrix, $\nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}) \in \mathbb{C}^{D_i \times D_o}$ as

$$\nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}) = \left[ \nabla_{\underline{q}_1} J_{\rho,\delta}(\mathbf{Q}) \quad \cdots \quad \nabla_{\underline{q}_{D_o}} J_{\rho,\delta}(\mathbf{Q}) \right]. \quad (4.35)$$

It is tedious but straight forward to calculate

$$\nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}) = \mathbf{\Psi}(\mathbf{Q}) + 4\rho \mathbf{V}_{\mathbf{H}} (\mathbf{V}_{\mathbf{H}}^H \mathbf{Q} \mathbf{Q}^H \mathbf{V}_{\mathbf{H}} - \mathbf{I}) \mathbf{V}_{\mathbf{H}}^H \mathbf{Q} + \delta \mathbf{\Psi}_{reg}(\mathbf{Q}), \quad (4.36)$$

where

$$\mathbf{\Psi}(\mathbf{Q}) = 2 \left[ \bar{\mathbf{\Phi}}_1 \underline{q}_1 \quad \cdots \quad \bar{\mathbf{\Phi}}_{D_o} \underline{q}_{D_o} \right],$$

$$\mathbf{\Psi}_{reg}(\mathbf{Q}) = -4 \left[ \varepsilon_{reg,1} \left( \underline{q}_1^H \underline{h}_{d_1} \right)^* \underline{h}_{d_1} \quad \cdots \quad \varepsilon_{reg,D_o} \left( \underline{q}_1^H \underline{h}_{D_o} \right)^* \underline{h}_{D_o} \right],$$

and

$$\varepsilon_{reg,k} = \left| \psi_k^2 - | \underline{q}_k^H \underline{h}_{d_k} |^2 \right|_+, \quad k = 1, \ldots, D_o.$$

An iterative gradient descent algorithm operates as follows. On each iteration, the algorithm calculates a step direction and finds the functions minimum along that direction. The step direction on the $l^{th}$ iteration is

$$\Delta \mathbf{Q}_l = -\frac{\nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}_l)}{\| \nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}_l) \|_F} \tag{4.37}$$

A step along this direction is found by a so called *line-search* operation which minimizes the function

$$f_l(\lambda) = J_{\rho,\delta}(\mathbf{Q}_l + \lambda \Delta \mathbf{Q}_l) \tag{4.38}$$

If the minimum of (4.38) is given by $\lambda_l$. Then the next iteration is given by $\mathbf{Q}_{l+1} = \mathbf{Q}_l + \lambda_l \Delta \mathbf{Q}_l$. Linear searches are commonly executed in two ways:

1. Exact Line Search: (4.38) can expanded as a piecewise-defined, fourth order polynomial in $\lambda$ which can be easily minimized with a root-finding method.

2. Armijo Line Search: (4.38) can be approximately minimized by executing a series of quadratic fits.

Since the number of "pieces" of a piece-wise polynomial expansion of (4.38) increases as $D_o^2$ and root-finding methods are notoriously complex for higher order polynomials, the Armijo line search procedure is considerably less complex. For a large class of functions one or two quadratic fits are enough. For details see [59]. In practice, it has been found that for any given $\mathbf{Q}_l \in \mathbb{C}^{D_i \times D_o}$, (4.38) has exactly two local minima: one occurring at some $\lambda_+ > 0$, and another occurring for some $\lambda_- < 0$. Numerical results suggest that both step-sizes are equally advantageous. In the results reported in the next section, only $\lambda_+ > 0$ is chosen. This is because the fourth order coloring penalty term in (4.34) dominates the expansion in (4.38) for large $\rho$. A summary of the algorithm is given in Table 4.1.

Table 4.1: Steepest Penalty EFURT

**Input:** $\mathbf{H} \in \mathbb{C}^{D_i \times D_u}$, the discrete equivalent channel matrix

$\mathbf{V}_H \in \mathbb{C}^{D_i \times r_H}$, an orthonormal basis for $\mathcal{R}(\mathbf{H})$

$D_o$, the number of desired outputs

$N_{iter}$, the number of iteratations

$\psi_k$, $\forall k = 1,\ldots,D_o$, regularization thresholds

$d_k$, $\forall k = 1,\ldots,D_o$, regularization targets

**Output:** $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$, the approximate EFURT solution

$\tilde{\mathbf{H}} = \mathbf{Q}^H \mathbf{H}$

**Algorithm:**

1) Initialize $\mathbf{Q}_0 \in \mathbb{C}^{D_i \times D_o}$ to any $\mathcal{R}(\mathbf{H})$-unitary matrix.
2) Pre-compute $\mathbf{\Phi}_k = \breve{\mathbf{H}}_k \breve{\mathbf{H}}_k^H$, $\forall k = 1,\ldots,D_o$
3) for $l = 1,\ldots,N_{iter}$
   a) Compute the gradient,
   $$\mathbf{Q} = \mathbf{Q}_{l-1}$$
   $$\tilde{\mathbf{H}} = \mathbf{Q}^H \mathbf{H}$$
   $$\tilde{\mathbf{Q}} = \mathbf{V}_{\mathbf{H}}^H \mathbf{Q}$$
   $$\mathbf{\Psi}(\mathbf{Q}) = 2\begin{bmatrix} \mathbf{\Phi}_1 \underline{q}_1 & \cdots & \mathbf{\Phi}_{D_o} \underline{q}_{D_o} \end{bmatrix}$$
   $$\varepsilon_{reg,k} = \left| \psi_k^2 - |\tilde{h}_{k,d_k}|^2 \right|_+, \ \forall k = 1,\ldots,D_o$$
   $$\mathbf{\Psi}_{reg}(\mathbf{Q}) = -4\begin{bmatrix} \varepsilon_{reg,1} \tilde{h}_{k,d_k}^* \underline{h}_{d_1} & \cdots & \varepsilon_{reg,D_o} \tilde{h}_{k,d_k}^* \underline{h}_{D_o} \end{bmatrix}$$
   $$\nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}) = \mathbf{\Psi}(\mathbf{Q}) + 4\rho \mathbf{V}_{\mathbf{H}} (\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^H - \mathbf{I})\tilde{\mathbf{Q}} + \delta \mathbf{\Psi}_{reg}(\mathbf{Q})$$
   b) Compute the search direction.
   $$\Delta \mathbf{Q}_l = -\nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}_{l-1}) / \| \nabla_{\mathbf{Q}} J_{\rho,\delta}(\mathbf{Q}_{l-1}) \|_F$$
   c) Perform Line Search to obtain $\lambda_l$
   d) Update iterate: $\mathbf{Q}_l = \mathbf{Q}_{l-1} + \lambda_l \Delta \mathbf{Q}_l$

## 4.7 Numerical Examples

To illustrate the operation of steepest penalty EFURT, we will construct a numerical example. Consider $D_u$ QPSK signals impinging on an $M$ element uniform circular antenna array (UCA). For simplicity, we will assume that all signals are symbol-synchronous. We seek a MIMO beamformer that aids the ITB-DDFSE algorithm in an

SRSJD receiver. This can be achieved by focusing channels energy into a trellis-oriented pattern.

The channel model for one received array snapshot is [7]:

$$\underline{x} = \mathbf{A}\underline{s} + \underline{z}\,. \tag{4.39}$$

The quantity $\underline{x}$ is an $M\times 1$ vector of the arrays received snap-shot. The matrix, $\mathbf{A}$, is the composite array response for all signals, and for all practical OLAP applications is rank $M$. The $d^{th}$ column of the $\mathbf{A}$ matrix is the steering vector for the $d^{th}$ signal. The steering vector for an $M$ element UCA of radius $R_a$ is given as

$$\begin{aligned}
&\underline{a}(\varepsilon_d,\theta_d) = [a_1 \quad \cdots \quad a_M]^T \\
&a_m = \exp\left(-j\kappa_R \sin\left(\theta_d - \frac{2\pi(m-1)}{M}\right)\cos(\varepsilon_d)\right),
\end{aligned}$$

where the constant, $\kappa_R = 2\pi R_a/\lambda$, $\theta_d$ is the azimuthal angle of arrival (AOA); $\varepsilon_d$ is the elevation (or depression) angle; and λ is the wavelength of the carrier frequency [26]. As evident in the cosine relationship between $\varepsilon_d$ and the steering vector, a UCA has little resolution in elevation. Motivated by applications where a receiver is demodulating signals on the horizon, we will assume $\varepsilon_d = 0$. Finally, $\underline{z}[n]$ is the $M\times 1$ temporally uncorrelated noise vector with zero mean and autocorrelation, $\mathbf{\Phi}_{\underline{zz}} = E\left[\underline{z}[n]\underline{z}[n]^H\right]$. For simplicity, we will assume that the noise is spatially uncorrelated, or $\mathbf{\Phi}_{\underline{zz}} = \sigma_z^2\mathbf{I}$.

As an example, we will consider $M = 5$ element UCA with $R_a = 0.43\lambda$, and $D_u = 10$ signals with angles of arrival (AOA), $\theta_d = 13^0$, $33^0$, $92^0$, $186^0$, $207^0$, $268^0$, $279^0$, $291^0$, $323^0$, $351^0$, as illustrated in Figure 4.1(a). For all signals, the received SNR per element, $10\log_{10}(\sigma_s^2/\sigma_z^2) = 20$ dB. Note that (4.39) is equivalent to (4.1) with $D_i = M$, $\mathbf{H} = \mathbf{A}$, and $\mathbf{V}_H = \mathbf{I}$. A checkerboard plot of the composite channel, $\tilde{\mathbf{H}} = \mathbf{Q}^H\mathbf{A}$, for the

Procrustean EFUT is illustrated in Figure 4.2(a). Each checkerboard plot in Figure 4.2(a) is normalized by the largest element. Qualitatively, Figure 4.2 (a) shows that the Procrustean EFUT, to some limited degree, focuses energy along the diagonal. The resulting Procrustean EFUT solution may also be viewed as a set of beam-formers. The $k^{th}$ beam-former, $\underline{q}_k$, is the $k^{th}$ column of $\mathbf{Q}$, and focuses the in-pattern energy on the $k^{th}$ row of $\tilde{\mathbf{H}}$. Figure 4.1(a) plots the gain patterns of all beam-formers as a function of azimuth. Again, each beam is normalized by its maximum gain. In this figure, it can be seen that the $k^{th}$ beam focuses its gain in the direction of a desired user. However, undesired back-lobes in some beams fail to reject undesired interference. In this interpretation, $\tilde{\mathbf{H}}$, may be viewed as a table whose $(k,d)^{th}$ element lists the amplitude and phase of the $d^{th}$ signal in the $k^{th}$ beam's output. By examining the composite channel of the Procrustean EFUT, Figure 4.2(a), a sparsity pattern of uniform width, $N_e = 4$, can be chosen that captures as much energy as possible in a tail-biting form. This results in the sparsity pattern (4.5), which is illustrated in Figure 4.2(a) by highlighting the boundary of pattern elements in a light color. While there is some noticeable focusing of energy on the diagonal, the quality of the focusing, as measured by $J(\mathbf{Q}) = 5.6 \text{ dB}$ of (4.14), is poor.

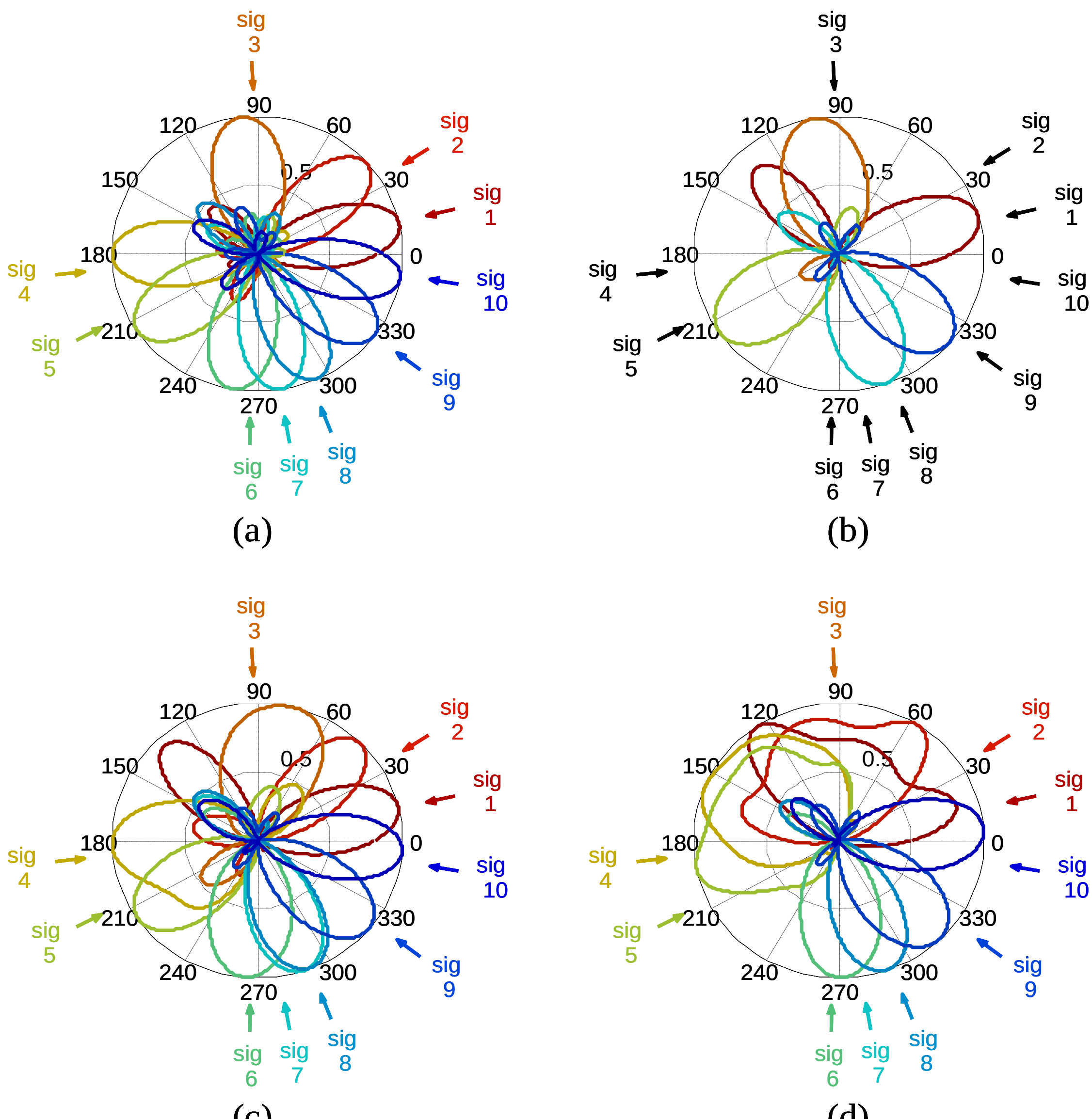


Figure 4.1: Gain Patterns of different energy focusing transforms (a) Procrustean EFUT (b) $D_o = 5$ Steepest Penalty EFURT (c) $D_o = 10$ Steepest Penalty EFURT (d) MIGSINR Transform

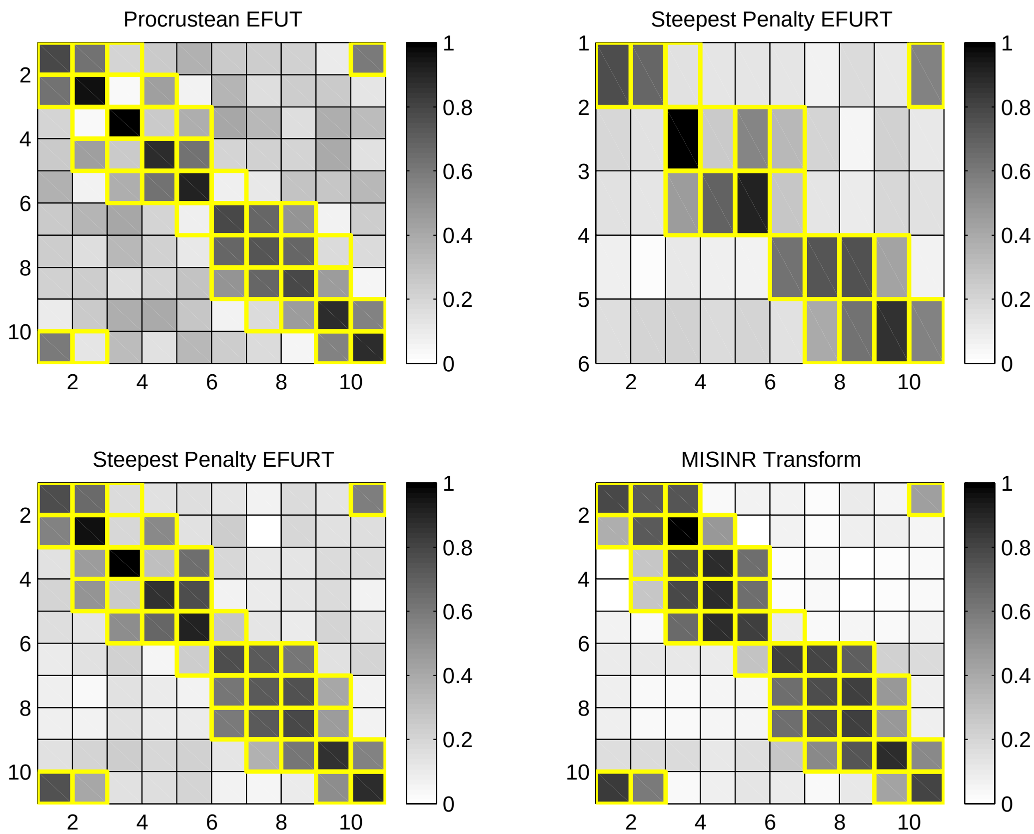


Figure 4.2: Comparison of composite channel, $\tilde{\mathbf{H}}$, for different energy focusing transforms.

Now consider the steepest penalty EFURT solution with $\rho = 0.07$, and $\delta = 1$. Unless specified otherwise, we will run the steepest penalty EFURT algorithm employing exact line searches for 100 iterations. We will compare results for two choices of $D_o : D_o = M = 5$, and $D_o = D_u = 10$. For $D_o = 5$, we use the pattern,

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

and initialize the steepest penalty EFURT algorithm with $\mathbf{Q}_0 \in \mathbb{C}^{4\times 4}$ chosen from the odd columns of the Procrustean EFUT. For $D_o = 10$, we apply the pattern of (4.5) and initialize the steepest penalty EFURT algorithm with the Procrustean EFUT, itself. The resulting composite channel for both the $D_o = 5$, and $D_o = 10$ cases are shown in Figure 4.2(b) and (c), respectively. The resulting beam-patterns are shown in Figure 4.1(b) and (c). Note for the $D_o = 10$ case there is one beam per signal, while for the $D_o = 5$ case there is not.

The evolution of the steepest penalty EFURT algorithm for the $D_o = 10$ case is shown in Figure 4.3-Figure 4.4. Figure 4.3 (a) shows the evolution of the steepest penalty cost function. Here the algorithm exhibits a slow convergence marked by a rapid initial improvement followed by diminishing returns. Figure 4.3 (b) plots the step-size used at each iteration. As iterates approach the minimizing solution, the step-size tends to decrease, although not monotonically. The dampened oscillation of the step-size curve is often indicative of either poor problem scaling or a cost-function similar to that of the infamous *Banana Function* [59]. Figure 4.4 (a) and (b) illustrate the evolution of the EFURT cost function and coloring measure. This figure shows a rapid initial improvement followed by diminishing returns. The coloring measure, initialized to the unitary Procrustean solution, starts at zero and steadily increases as the EFURT cost function improves. After 100 iterations a 5 dB improvement in the EFURT cost function is achieved at the expense of noise enhancement. However, a 4dB improvement may be achieved after just 18 iterations. The steepest penalty EFURT algorithm for the $D_o = 5$ case exhibits similar characteristics which are not shown here. In both cases, the choice of algorithm initialization, $\mathbf{Q}_0$, did not significantly effect the final values of $J(\mathbf{Q})$ or $\xi(\mathbf{Q})$. In fact even random initializations (non-unitary) yielded iterates with measures similar to the Procrustean initialization after just 2 or 3 iterations. Moreover, this fact holds true for a wide range of overloaded array processing problems and a wide range of $\rho, \delta$ values. The independence of the final EFURT and coloring measures to initial values suggests that all local minima of $J_{\rho,\delta}(\mathbf{Q})$ are equally effective.

Finally, the composite channel and beam-patterns of the MIGSINR energy focusing transform are shown in Figure 4.2(c) and Figure 4.1(d) respectively. These figures show an even better quality of energy focusing than the steepest penalty EFURT at the cost of even more noise-coloring. The resulting EFURT cost function, $J(\mathbf{Q})$, and coloring measure, $\xi(\mathbf{Q})$, for all solutions are summarized in Table 4.2 and Table 4.3. The results for the steepest penalty EFURT algorithm are reported for the 100th iterate. We offer no explanation why the reduced output $D_o = 5$ EFURT has a greater coloring measure than the $D_0 = 10$ EFURT.

Table 4.2: Comparison of the EFURT cost function and coloring measure for different energy focusing techniques.

| | Procrustean | Steepest Penalty EFURT $D_o = 5$ | Steepest Penalty EFURT $D_o = 10$ | MIGSINR Beamformer |
|---|---|---|---|---|
| $J(\mathbf{Q})$ (dB) | 5.6 | 10.7 | 10.9 | 15.7 |
| $\xi(\mathbf{Q})$ | 0 | 0.52 | 0.39 | 0.76 |

Table 4.3: Comparison of the singular values for different energy focusing techniques.

| Singular Value | Procrustean | Steepest Penalty EFURT $D_o = 5$ | Steepest Penalty EFURT $D_o = 10$ | MIGSINR Beamformer |
|---|---|---|---|---|
| 1 | 1 | 1.4315 | 1.5013 | 1.0969 |
| 2 | 1 | 0.4701 | 0.7030 | 0.0603 |
| 3 | 1 | 0.3398 | 0.4730 | 0.0289 |
| 4 | 1 | 0.1062 | 0.2151 | 0.0152 |
| 5 | 1 | 0.0520 | 0.1476 | 0.0067 |

We will now investigate the dependence of the EFURT measure and coloring measure upon the choice of coloring penalty gain. The steepest penalty EFURT algorithm was run for $D_o = 10$ and 50 values of $\rho$ evenly spaced in the log-domain between $\rho = 10^{-3}$ to $\rho = 100$. For each value of $\rho$, the algorithm was run for 1000 iterations. Such a large number of iterations assures that the steepest penalty EFURT algorithm is well within two significant digits of its final value. The resulting EFURT measure and coloring measure is plotted in Figure 4.5. For reference, the EFURT measure and coloring measure of both the Procrustean EFUT and the MIGSINR transform are also shown. As expected, the coloring measure approaches zero for large $\rho$ and approaches the coloring measure of the MIGSINR solution for small $\rho$. Although not shown in this plot, both EFURT curves approach the MIGSINR results as $\rho \to 0$. However, in general, the steepest-penalty EFURT with $\rho = 0$ may yield an EFURT measure and coloring measure that slightly exceeds those for MIGSINR.

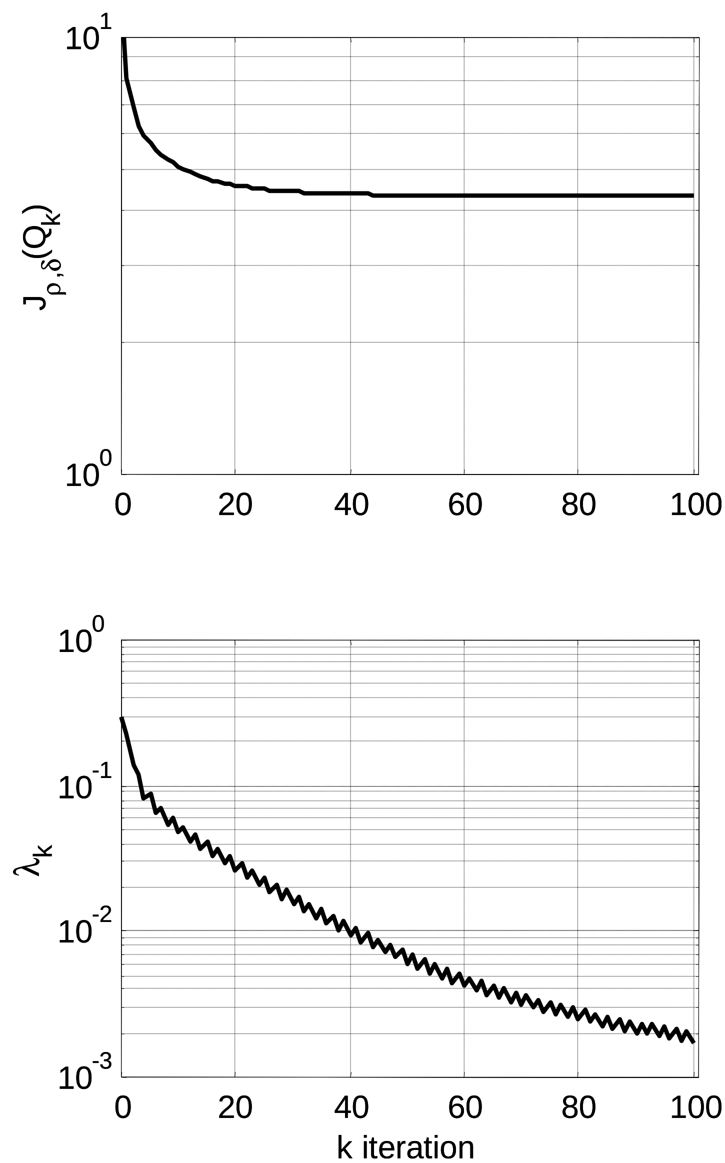


Figure 4.3: Evolution of Steepest Penalty EFURT, $\rho = 0.07$

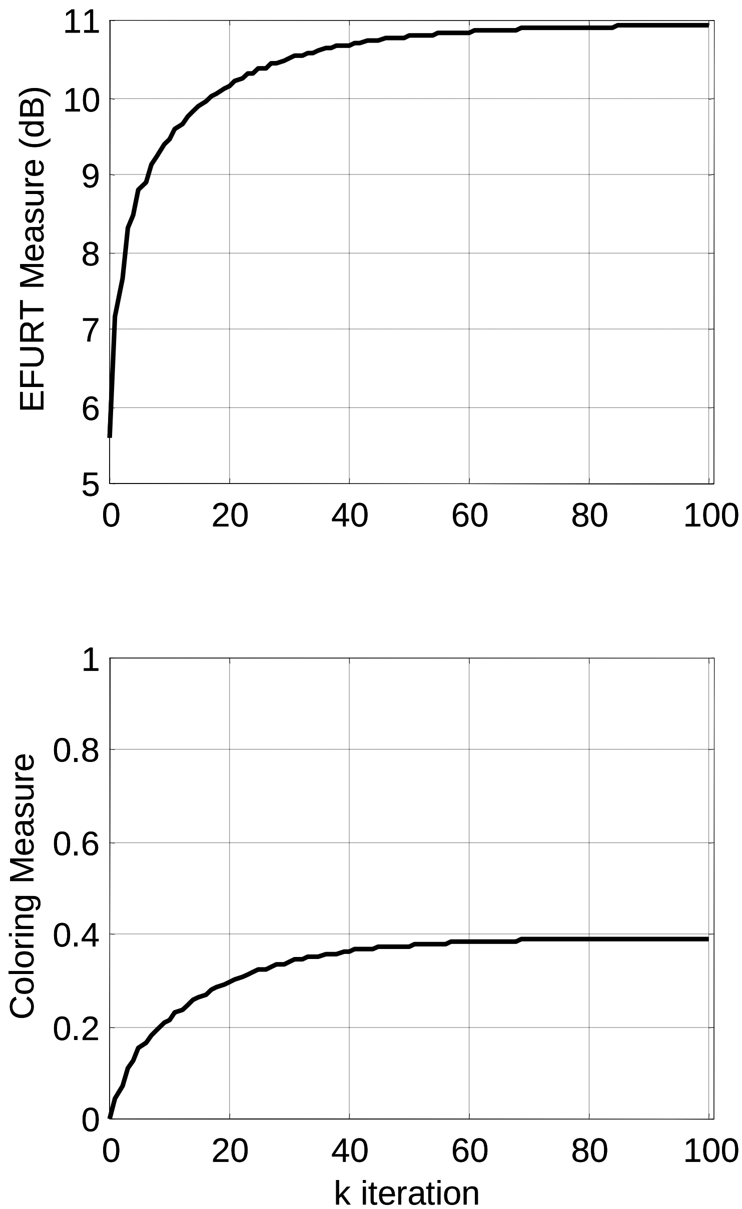


Figure 4.4: Evolution of the EFURT and coloring measures for the Steepest Penalty Algorithm

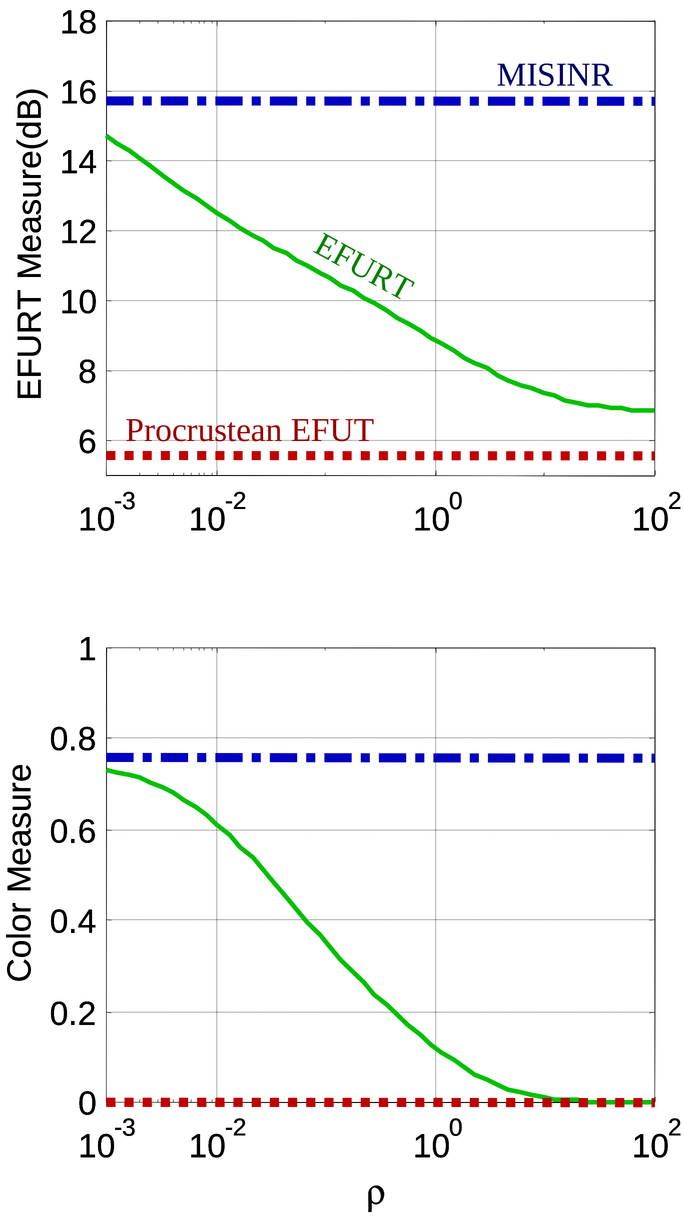


Figure 4.5: Relationship between coloring measure and coloring penalty gain

## 4.8 Sparsity Pattern Selection

Thus far, we have only discussed the computation of EFTs for a particular sparsity pattern without saying much about how a pattern should be chosen. Recall that when SRSJD uses a Procrustean EFT, the sparsity pattern only affects the ITB-DDFSE's operation. Meanwhile, for all other proposed EFT types, the sparsity-pattern affects the EFT-solution, itself. The pattern choice is application specific but there are some general guidelines. An obvious first choice would be a uniform pattern. For instance, for an overloaded UCA, if AOAs are reasonably evenly distributed around the array, an obvious candidate for a successful sparsity pattern is a circulant matrix of ones and zeros. However, the use of a non-uniform sparsity-pattern can conceivably improve SRSJD's performance if it results in significantly higher EFT measures. The question is, what is the best criteria for choosing such a pattern? Given a particular EFT-type, a pattern search algorithm guaranteed to yield the best EFT-measure is as follows: enumerate through every possible pattern choice and choose the pattern with the maximum EFT. If the sought-after pattern width is fixed on each row, (i.e. $N_e[k] = N_e$), and if we ignore superficial permutations of the pattern's rows, this would require the computation of $D_u^{D_o}$ EFTs. Indeed, such a search would be more complex than even JMLE! Thus, a need for a heuristic pattern search algorithm arises.

Another candidate pattern-search criterion would be AOA. However, considering AOA alone does not account for the fact that a strong interferer can interfere with other signals more than signals closely spaced in AOA. An approach that would account for the effect of both AOA and received signal power would be to use a *pre-beamformer*, $\mathbf{Q}'$. Choosing the so-called *max signal-to-noise ratio (MSNR) pre-beamformer,* $\mathbf{Q}' = \mathbf{A}$, has the following interpretation: the $d^{th}$ column of $\mathbf{Q}'$ is the beam-former that would maximize signal $d$ 's output SNR if the other signals were not present [60]. Examining the energy-distribution of the channel after pre-beamforming, $\tilde{\mathbf{H}}' = \mathbf{Q}'^H \mathbf{A}$ tells which signals tend to interfere with each other the most. One may think of many other conceivable choices of a pre-beamformer, but none stands out as optimal. Hence, unless otherwise specified, we will consider only the MSNR pre-beamformer for the remainder of this thesis. Table 4.4 outlines the so-called *Greedy pattern-search* algorithm, which attempts to find the square, non-uniform pattern of fixed width, $N_e$, that captures the most energy in the pre-beamformed channel. It is called greedy because each row of the pattern is chosen without regard to the other rows. Henceforth, a *Greedy pattern* is a sparsity pattern produced by this algorithm. Non-square patterns are obtained from a square greedy pattern by down-selecting rows according to the following criterion: choose $D_o$ rows with the best SIR such that every signal is represented in the pattern.

If $N_e$ is chosen less than $D_u/2$, the Greedy pattern search is guaranteed to produce a trellis-oriented pattern. Step one of the algorithm assures that every row is contiguous. The choice of $N_e$ assures that the pattern is well offset. Finally the sorting operation in the second set assures that every column of the pattern is contiguous.

Table 4.4: Greedy Sparsity Pattern Search Algorithm

**Input:** $\tilde{\mathbf{H}}' \in \mathbb{C}^{D_o \times D_u}$, the channel matrix after beamforming

$N_e$, the pattern width

**Output:** $\mathbf{P} \in \mathbb{C}^{D_u \times D_u}$, a sparsity pattern

$\mathcal{U}_e[k],\ 1 \le k \le D_u,$ a sequence of ordered sets listing the ones on the $k^{th}$ row of $\mathbf{P}$ as they occur in a stretch

**Algorithm:**

1) For each $k,\ 1 \le k \le D_u$
   a) Consider all of the length $N_e$ stretches of elements on the $k^{th}$ row of $\mathbf{H}'$ containing $[\mathbf{H}']_{k,k}$.
   b) Of all such stretches, let $\mathcal{U}_e'[k]$ be the column indices of the stretch with maximum energy.
2) Let $\mathcal{U}_e[k]$ be $\mathcal{U}_e'[k]$, sorted by $\mathcal{U}_e'[k](1)$
3) Construct $\mathbf{P}$ directly from the sequence $\mathcal{U}_e[k]$

## 4.9 SIMULATION RESULTS

This section compares the use of various pattern-search and MIMO beamformers for use in SRSJD. The ITB-DDFSE symbol error rate (SER) performance is difficult to evaluate analytically, because each symbol estimate is a highly non-linear function of both the received signal and random channel parameters. Instead, symbol error rates have been estimated through the use of Monte-Carlo simulation. It is instructive to revisit the previous example before the discussion of some more comprehensive simulations.

In order to investigate the relationship between AOAs, the choice of beamformer, and SER in the previous example, a simulation of all four receiver configurations was run. The channel and the previously calculated energy focusing transforms were fixed for 4735 simulation trials over which all users experienced at least 10 errors. On each simulation trial, the ITB-DDFSE algorithm starts on the tail-biting trellis stage with the best SINR and proceeds around the tail-biting trellis in the direction for which the second stage has the best SINR. The symbol error rate is reported for all receivers and signals in Table 4.5. As expected, for all receivers, the signal with the greatest inter-AOA spacing between neighboring signals enjoy the best SER. Also note that the MIGSINR transform out-performs all other receivers. This suggests that in some cases, no noise control beyond that already loosely provided by the MIGSINR transform is required.

Now let us consider a more comprehensive set of simulations. To compare results with those of [16], a uniform circular array (UCA) is used for all experiments. A number of receiver attributes are varied during simulation: the type of MIMO beamformer, the type of sparsity pattern, the pattern width, $N_e$, the number of elements, $M$, the number of beamformer outputs, $D_o$. Procrustean, EFURT, and MIGSINR beamformers are simulated. Two types of patterns are used: a uniform tail-biting pattern and a greedy pattern employing a MSNR pre-beamformer. To demonstrate SRSJD's scalability, two array sizes have been considered, a four element array, and an eight element array. Finally, recall that the number of SRSJD beamformer outputs, $D_o$, becomes the number of stages for a tail-biting trellis. There are two seemingly valid but conflicting rationales for choosing this number. One line of reasoning says that the received signal space is of dimension $M$, the number of antenna elements; and since a MIMO beamformer is just a linear transform, it cannot increase this dimension. Moreover, since SRSJD is an approximation of JMLE, whose performance cannot be improved by increasing $D_o$. Hence any choice other than $D_o = M$ could be wasteful. However, another line of reasoning could interpret SRSJD as a multi-user detector designed to extract a desired signal from each beam in the presence of strong interferers. Hence, allocating one beam for each user (i.e. choosing $D_o = D_u$) could treat all users more fairly. To resolve this dilemma, both $D_o = M$ and $D_o = D_u$ have been considered.

Table 4.5: Resulting Symbol Error Rate for all signals in previous example

| Signal Number | Procrustean EFUT | Steepest Penalty EFURT $D_o = 5$ | Steepest Penalty EFURT $D_o = 10$ | MIGSINR Transform |
|---|---|---|---|---|
| 1 | 0.2106 | 0.0944 | 0.0543 | 0.0070 |
| 2 | 0.2021 | 0.0936 | 0.0439 | 0.0063 |
| 3 | 0.1143 | 0.0530 | 0.0327 | 0.0021 |
| 4 | 0.2030 | 0.0834 | 0.0598 | 0.0040 |
| 5 | 0.1814 | 0.0822 | 0.0642 | 0.0023 |
| 6 | 0.2636 | 0.0874 | 0.0583 | 0.0051 |
| 7 | 0.2781 | 0.0743 | 0.0536 | 0.0072 |
| 8 | 0.2739 | 0.0819 | 0.0513 | 0.0065 |
| 9 | 0.1937 | 0.0841 | 0.0604 | 0.0055 |
| 10 | 0.1962 | 0.0661 | 0.0560 | 0.0070 |

A meaningful comparison of beam-forming criteria must demonstrate the effect of inevitable variations in the operating environment. To demonstrate a robustness to (AOA) variations, we have used the following so-called *random-sectored* angle-of-arrival model. For a simulation with $D_u$ signals, the azimuthal plane is divided up into $D_u$ equally spaced sectors. Signal's AOAs are random, independent, and uniformly distributed in their sectors. Note that this signal model makes the most sense when the receiver has the ability to allocate radio-resources. To demonstrate SRSJD's robustness to fluctuations in received power, the following Ricean fading model with angular spread has been used. The AOA, $\theta_d$, of the $d^{th}$ signal's specular component is generated according to the random-sectored model. The diffuse component is modeled with the sum of $L_p = 10$ paths. The paths' AOAs are random, independent, and uniformly distributed in the range $[\theta_d - \Delta/2, \theta_d + \Delta/2]$, where the parameter $\Delta$ is the azimuthal angular spread. Normally, a UCA is assumed to have isotropic elements, but it is instructive to investigate the effect of this assumption on SRSJD's performance. A model for radially directed element patterns is given in Appendix C. In all of the subsequent results, ITB-DDFSE's symbol error-rate (SER) performance is averaged over all signals.

Figure 4.6 depicts the case of a four element array in a random-sectored environment with eight signals. Here the receiver employs a square ($D_o = 8$) and uniform tail-biting sparsity pattern of width $N_e = 4$. Three types of beamformers: Procrustean, EFURT (of varying coloring penalty gain), and MIGSINR have been used. First, the Procrustean pattern's poor performance belies its typically low EFT measures and hence susceptibility to error propagation. The beamformer that aids SRSJD the most is the MIGSINR, which introduces the most noise coloring but achieves the best EFT measure. Meanwhile the EFURT beamformer family exhibits a decrease in SER with coloring penalty gain. These results suggest that, for this signal environment, SRSJD benefits from an aggressive reduction in interference at the expense of noise-enhancement. In fact, we will find that this is a common characteristic of many random-sectored signal environments.

Figure 4.7 depicts SRSJD's performance for the same type of signal environment as Figure 4.6, except a greedy sparsity pattern of width $N_e = 4$ was used. Again the Procrustean beamformer fairs poorly, as do EFURT beam-formers with high coloring-penalty gain values. However, with this type of sparsity pattern, there is a slight benefit to using EFURT beamformers over MIGSINR beamformers at moderate SNRs. Nonetheless, a quick comparison of Figures 4.6 and Figures 4.7 reveals that the greedy sparsity pattern fairs *worse* than a uniform one. There are two explanations. First, nothing prevents the algorithm that constructs the greedy pattern from duplicating the enumeration set, $\mathcal{U}_e[k]$ for multiple rows. Not only does this deprive ITB-DDFSE of multiple viewpoints of the signal space pertinent to a particular signal, but it can also lead to duplicate or near-duplicate observations input to ITB-DDFSE. The later phenomenon is most readily apparent in Figure 4.2 (b), where rows three and seven are duplicated. These duplicate rows are produced by duplicate beams, which in turn, produce duplicate inputs to ITB-DDFSE. Subsequently, during its search for the least-cost trellis path, ITB-DDFSE can over-emphasize errors incurred on the trellis stages associated with these duplicate inputs. Duplicate beams can be discarded, but this does not solve the more pressing problem of the greedy pattern-search: when encountering a cluster of AOAs, the greedy algorithm chooses the entire cluster for many beams, thereby depriving signals on the cluster edge of other signal-space views provided by alternative beam patterns.

Figure 4.8 illustrates ITB-DDFSE's performance for the same signal environment, but this time the pattern-width, $N_e$ is varied. Again $D_o = D_u = 8$ was chosen. This figure shows that for an overloading factor of two, it is not possible to reduce SRSJD's complexity to $N_e = 3$ without a large penalty in SER. Figure 4.9 illustrates the effect of the number of beam-former outputs on SRSJD's performance. Clearly, the choice $D_o = D_u = 8$ exhibits a superior performance. The conclusions of both figures were found to apply to all the random-sectored environments that have been simulated.

Figure 4.10 illustrates SRSJD's performance for the same signal environment, but for different Ricean K-factors. In this figure, ITB-DDFSE employing a uniform pattern and a MIGSINR beamformer is considered, since these universally result in the best performance. Here, SRSJD exhibits a reasonable robustness to amplitude variations with angle-spreads as high as $10^0$. Little variation in performance was found for angle-spreads less than $10^0$.

Figure 4.11 depicts the case of an eight element array in a random-sectored environment with sixteen signals employing a square ( $D_o = 16$ ) and uniform sparsity pattern of width $N_e = 4$. This time no fading has been simulated ( $K = \infty$ ). Again, here we find a reduction in SER with coloring penalty gain similar to that previously encountered. Also note that even when SRSJD employs its best beamformer, MIGSINR, it achieves 1% error probability only at high SNRs. This poor performance might be improved by employing wider sparsity patterns, however, even for $N_e = 5$, no significant performance benefit was found; and choosing wider patterns results in impractical computational complexities. An alternative approach would be to replace the UCA's isotropic elements with radially directive ones. The proposed antenna pattern model is discussed in Appendix C. Figure 4.12 illustrates the performance of SRSJD in the same environment as Figure 4.11, but here the performance is characterized in terms of the element-pattern's half-power beam-width (HPBW). For all cases an SNR= 20dB was used. Again, SRSJD tends to favor an aggressive reduction in interference at the possible expense of noise-enhancement. SRSJD, employing MIGSINR beam-formers and high-gain antenna elements, enjoys an order of magnitude improvement in SER performance. Beam patterns with a null at $180^0$ from bore-sight enjoy better SER performance than other patterns with a similar HPBW. This corresponds to HPBW values of $34^0$, $51^0$, $105^0$. Again, little benefit was found for increasing the pattern width to $N_e = 5$.

## 4.10 CONCLUSION

To transcend the over-dependence of previous OLAP approaches on excess bandwidth, this chapter explores a new approach to Overloaded Array Processing based on reduced-complexity JMLE approximation. This approximation is most easily understood in two steps: the JML estimator is approximated by a TELSE, for which a careful choice of linear transform is chosen to assure that TELSE both well approximates JMLE and can be well approximated with a reduced complexity search. The performance of such a scheme pivots on a prudent choice of linear pre-processing. A choice of front-end linear processing should strike a prudent tradeoff between noise-enhancement and error-propagation within an SRJSD setting. On one side of this tradeoff are EFUTs which do not introduce any noise-enhancement but often exhibit poor interference rejection performance. On the other side of this tradeoff are MIGSINR transforms which weakly control noise enhancement by indirectly controlling the diagonal of the output noise processes' auto-correlation matrix. To explore the possibility that a better tradeoff between the two transforms exists, this chapter has developed a noise-coloring measure whose minimization also minimizes the discrepancy between the TELSE and JMLE cost functions. Armed with this coloring measure it is possible to explore the noise-enhancement/error-propagation tradeoff.

Simulation results suggest for the random-sectored signal environment, a uniform pattern far out-performs a greedy pattern. Simulation results also suggest that the SER performance of the MIGSINR enabled SRSJD outperforms other linear transforms, even when it introduces a severe degree of noise-enhancement. This is a surprising but advantageous result. It is a surprising result because, generally speaking, one would not expect that the linear transform that distorts the maximum-likelihood criterion the most would result in the best performance of its approximation. However, it is advantageous since symmetric generalized eigen-value problems are by now well understood, with numerous numerical methods available for finding them. Nevertheless, the tradeoff between noise-enhancement and error-propagation is problem specific, and hence the optimum tradeoff can never be completely resolved. However, the framework developed in this chapter for exploring this tradeoff can be applied to gain understanding about both TELSE and SRSJD's performance in numerous signal environments.

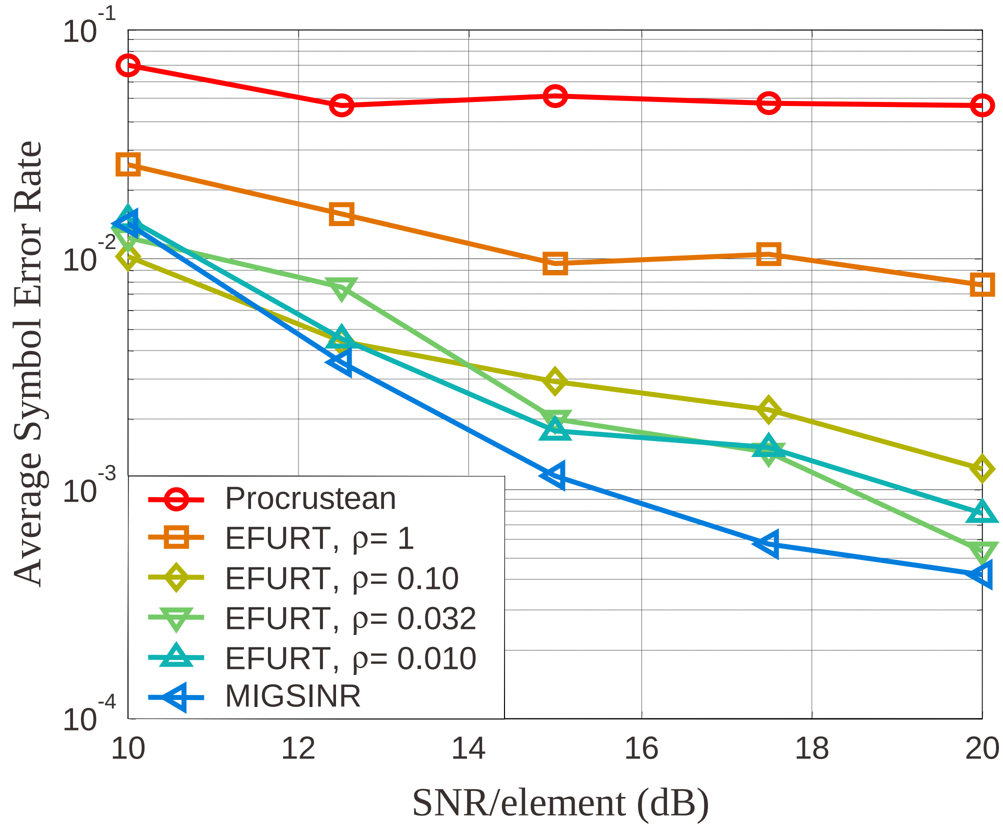


Figure 4.6: ITB-DDFSE's performance for various MIMO beamformers, $M = 4$, $D_u = 8$, $K = \infty$, random-sectored AOA model, uniform tail-biting sparsity pattern.

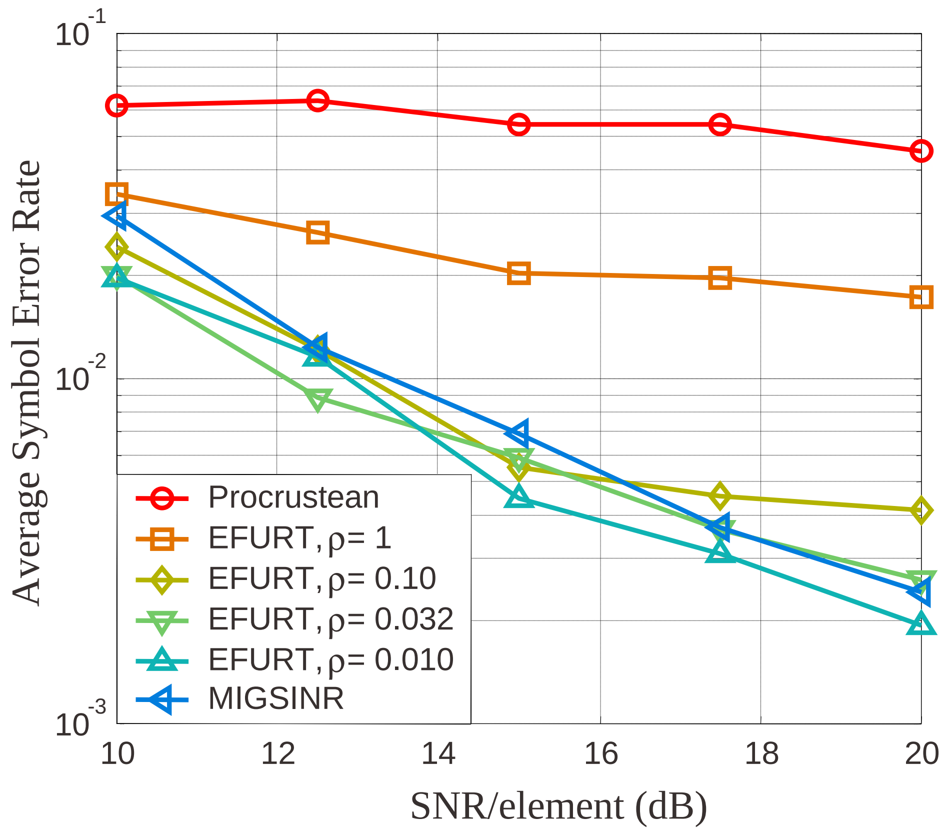


Figure 4.7: ITB-DDFSE's performance for various MIMO beamformers, $M = 4$, $D_u = 8$, $K = \infty$, random-sectored AOA model, greedy sparsity pattern.

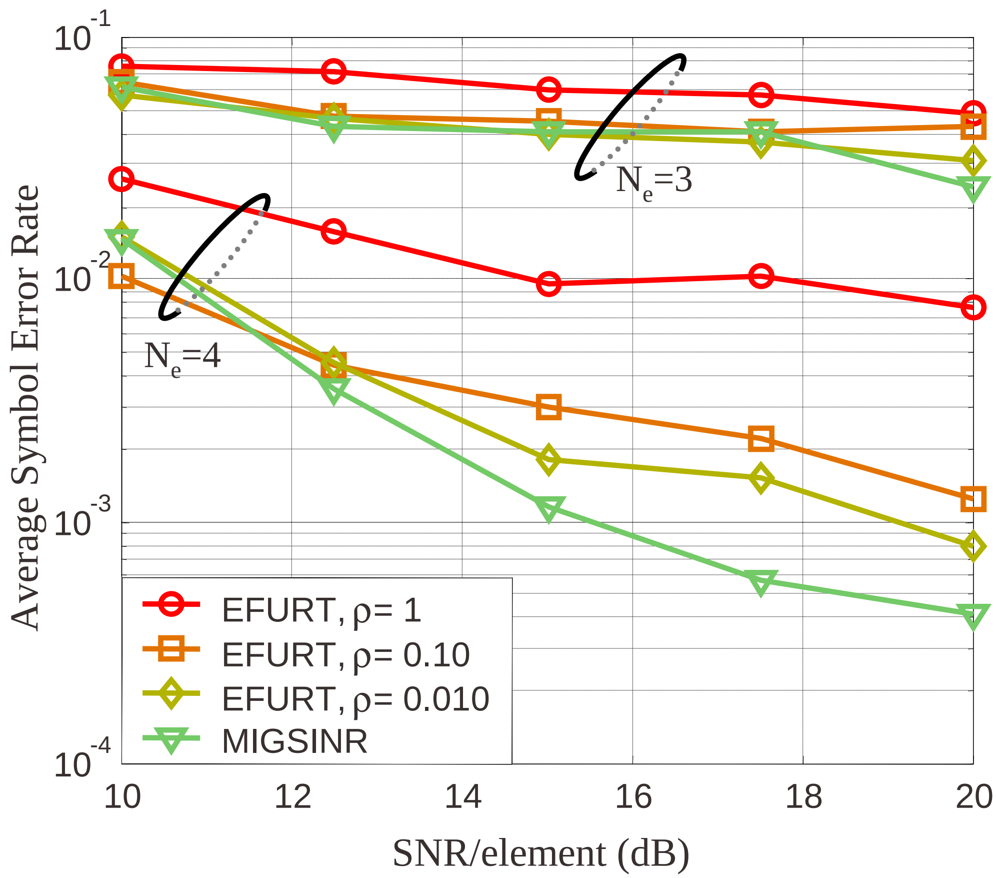


Figure 4.8: The effect of pattern width, $N_e$, on ITB-DDFSE's performance. $M = 4$, $D_u = 8$, $K = \infty$, random-sectored AOA model, uniform tail-biting sparsity pattern.

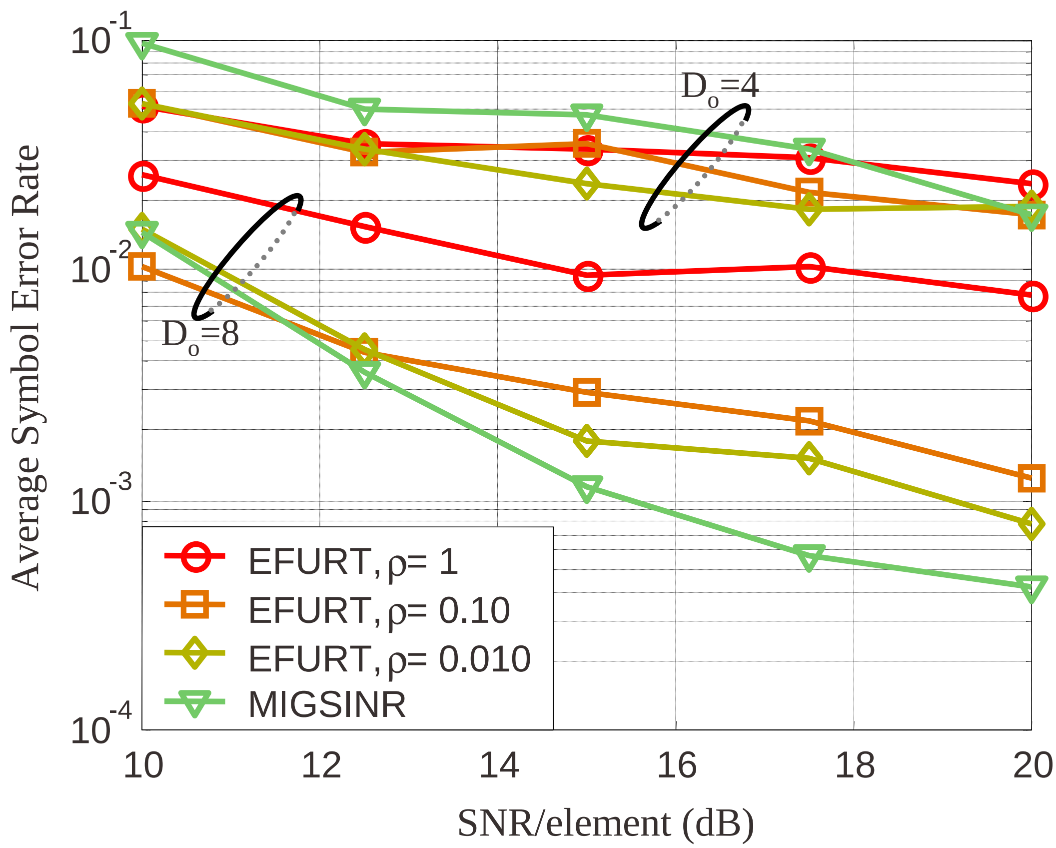


Figure 4.9: The effect of the number of MIMO beamformer outputs, $D_o$, on ITB-DDFSE's performance. $M=4$, $D_u=8$, $K=\infty$, random-sectored AOA model, uniform tail-biting sparsity pattern.

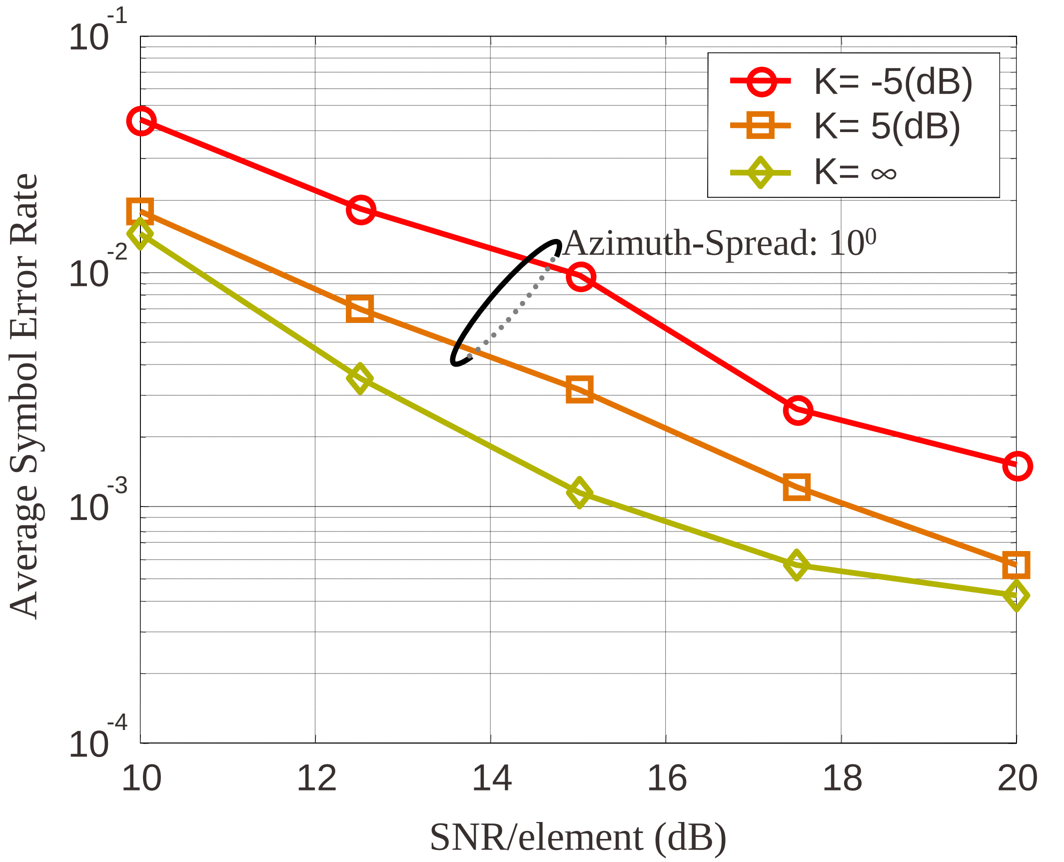


Figure 4.10: The effect of Ricean K-factor on ITB-DDFSE's performance. $M = 4$, $D_u = 8$, $K = \infty$, random-sectored AOA model, uniform tail-biting sparsity pattern

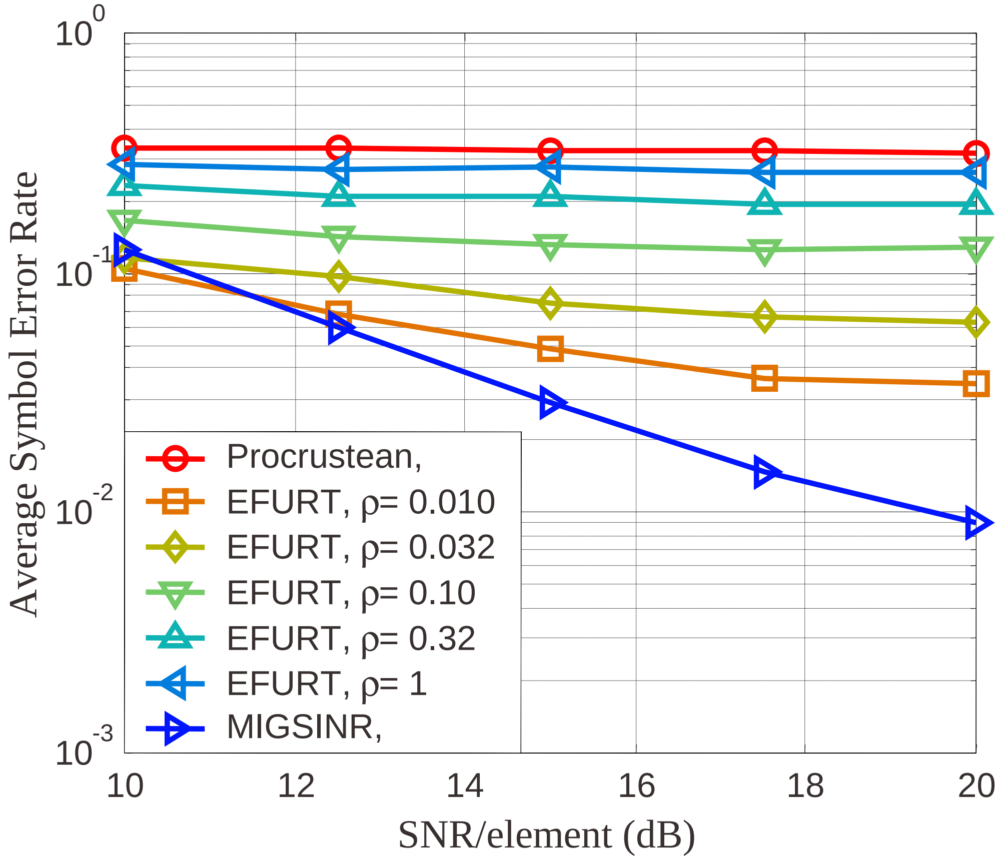


Figure 4.11: ITB-DDFSE's performance for various MIMO beamformers, $M = 8$, $D_u = 16$, $K = \infty$, random-sectored AOA model, uniform tail-biting sparsity pattern.

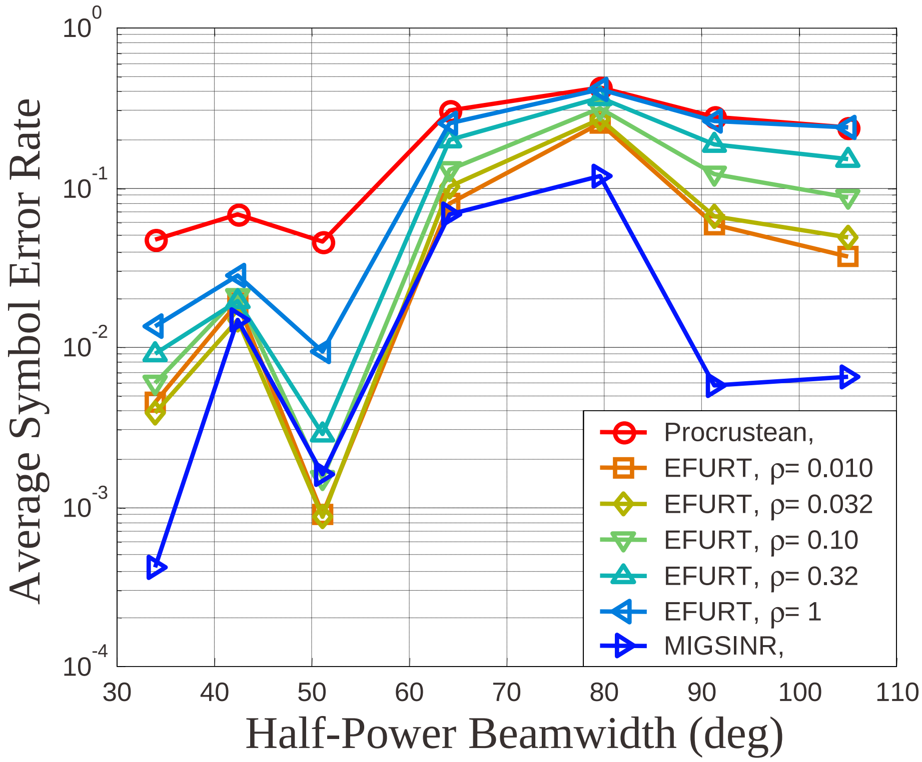


Figure 4.12: The Benefit of using radially directed element patterns for ITB-DDFSE's performance, SNR/element= 20 dB

# Chapter 5: CONCLUSION AND FUTURE WORK

This thesis has explored three OLAP approaches that do not exploit FEC. Each approach offers a different performance/complexity tradeoff by exploiting signal properties in different ways. The easiest signal property to take advantage of is excess-bandwidth which may be exploited with linear space-time processing. In Chapter 2 of this thesis it is shown that for an environment where there are twice as many PAM signals as elements (an overloading factor of two) and each of these PAM signals have a sufficiently large excess bandwidth, even MMSE-based linear STAP can extract one or two signals from the environment. This surprising result motivated a hybrid MMSE-SIC receiver. It was found that a phased array subject to an overloading factor of 1.5 could extract 75% excess bandwidth signals with *intelligible*[1] voice quality. An enhancement to this scheme, considered in Chapter 3, replaces linear STAP with reduced state maximum apriori sequence estimation (RS-MAP) in an SIC framework. This so called MAP-SIC receiver employing a phased array can extract 75% excess bandwidth signals with commercial voice quality in environments with an overloading factor of two. While the performance of such schemes is enticing, a need becomes apparent for an OLAP receiver that does not over-rely on the exploitation of excess bandwidth.

This need motivates a fresh approach to OLAP based upon the SRSJD receiver in [16]. Chapter 4 investigates the use of linear pre-processing that reduces the complexity of back-end non-linear signal processing while minimizing the ensuing SER performance

---

[1] For linear-predictive vocoders, an uncoded bit error rate of 10% is often considered to be intelligible to a trained operator.

penalty. Since SRSJD is essentially a reduced complexity JMLE approximation, and noise-enhancement (e.g. coloring) is known to degrade JMLE's performance, this is also a concern for SRSJD. A coloring measure is proposed that quantifies the degree of noise-enhancement introduced by linear preprocessing. This coloring measure is used to investigate the noise-enhancement/ error-propagation tradeoff inherent to all reduced complexity JMLE schemes, including SRSJD. Simulations suggest that when signals are near-equal power and have AOAs that are reasonably evenly distributed over $360^0$ of azimuth, ITB-DDFSE benefits from an aggressive reduction in interference at the expense of noise enhancement. Moreover, in these environments ITB-DDFSE was found to favor the use of uniform sparsity patterns. While the issue of noise-enhancement can never be resolved for all situations, the framework developed in this thesis for exploring this tradeoff can yield insight about both TELSE's and SRSJD's performance in numerous signal environments.

To enjoy practical application, SRSJD must be extended to handle the processing of asynchronous signals. Perhaps the most obvious extension is to the application of OFDM signals. In multi-user OFDM systems employing tight frequency control and moderate frame-synchronization, users appear to be symbol synchronous on each sub-carrier. However, inter-carrier interference (ICI) (induced by both Doppler frequencies and the use of imprecise local oscillators at the transmitters) prohibit the direct application of SRSJD on a per-subcarrier basis. Perhaps a more difficult task is to extend SRSJD's application to asynchronous PAM signal environments like those considered in Chapters 2 and 3. In both situations, the chief difficulty arises from the fact that SRSJD's operation rests upon the formation of a spatial-trellis. Since trellises are by definition, directed graphs, they can only be formed over one signal dimension. In the case of OFDM signal environments interference is spread over both the spatial domain (multi-access interference) and frequency domains (ICI). In the case of asynchronous PAM signals, interference is spread over both the spatial domain and the time domain. A similar situation exists in the maximum-likelihood decoding of linear block codes [48], or the half-toning of an image [51], both of which have been successfully solved with iterative decoding or "TURBO" principles. Future research will investigate the use of the

TURBO principle to asynchronous OLAP problems. This approach is very similar to that taken in [6], where array processing is viewed from an iterative decoding perspective.

# Appendix A: Subspace Unitary Transforms and Subspace White Noise

A matrix, $\mathbf{Q} \in \mathbb{C}^{D \times D}$, is said to be unitary if $\mathbf{Q}^H \mathbf{Q} = \mathbf{I}$. Unitary matrices have many well known properties: they preserve length (e.g. are invariant to the Euclidian norm), their eigen-values and singular values have unit magnitude, and they may be thought of as a combination of rigid rotations and reflections. A natural question to ask is how one might characterize matrices that have similar properties in a particular linear subspace, $\mathcal{S}$. In particular, what linear transforms preserve the length of a vector in $\mathcal{S}$? Which linear transforms preserve the length of a vectors $\mathcal{S}$ component? We will find that these are two different questions and the latter leads to the development of subspace-unitary matrices.

**Definition:** A matrix $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$ is said to be *outer unitary* if $D_i \le D_o$ and $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$.

**Definition:** A matrix $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$ is said to be *inner unitary* if $D_i \ge D_o$ and $\mathbf{Q}^H \mathbf{Q} = \mathbf{I}$.

**Definition:** $U_o^{D_i \times D_o} = \{\mathbf{Q} \in \mathbb{C}^{D_o \times D_i} : \mathbf{Q}\mathbf{Q}^H = \mathbf{I}\}$, $D_i \le D_o$ is the set of all $D_i \times D_o$ outer unitary matrices.

**Definition:** $U_I^{D_i \times D_o} = \{\mathbf{Q} \in \mathbb{C}^{D_o \times D_i} : \mathbf{Q}^H \mathbf{Q} = \mathbf{I}\}$, $D_i \le D_o$ is the set of all $D_i \times D_o$ inner unitary matrices.

The subsequent development will deal with a linear subspace $\mathcal{S} \subseteq \mathbb{C}^{D_i}$, with $\dim(\mathcal{S}) = r$. For convenience, let $\mathbf{V} \in \mathbb{C}^{D_i \times r}$ be any matrix whose columns form an ortho-normal basis for $\mathcal{S}$, and $\mathbf{V}_\perp \in \mathbb{C}^{D_i \times (D_i - r)}$ be any matrix whose columns form an ortho-normal basis for this subspace's orthogonal complement, $\mathcal{S}^\perp$. Finally, let $\mathbf{P}_S = \mathbf{V}\mathbf{V}^H$ be the projection matrix onto $\mathcal{S}$ and $\mathbf{P}_S^\perp = \mathbf{I} - \mathbf{V}\mathbf{V}^H = \mathbf{V}_\perp \mathbf{V}_\perp^H$ be the projection onto $\mathcal{S}^\perp$.

**Definition:** A matrix $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$ is said to be *isometric on* $\mathcal{S}$, or $\mathcal{S}$*-isometric* if any one of the following equivalent statements are true.

(i) $\mathbf{P}_{\mathcal{S}}\mathbf{Q}\mathbf{Q}^H\mathbf{P}_S = \mathbf{P}_S$
(ii) $\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V} = \mathbf{I}$
(iii) $\mathbf{P}_S\mathbf{Q} = \mathbf{V}\tilde{\mathbf{Q}}$ for some $\tilde{\mathbf{Q}} \in U_o^{r \times D_o}$
(iv) $\| \mathbf{Q}^H \underline{x} \| = \| \underline{x} \|$, $\forall \underline{x} \in \mathcal{S}$

**Proof:**

We will prove that the aforementioned definitions are indeed equivalent. First note that $\mathbf{P}_{\mathcal{S}}\mathbf{Q}\mathbf{Q}^H\mathbf{P}_S = \mathbf{P}_S$, if and only if $\mathbf{V}\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V}\mathbf{V}^H = \mathbf{V}\mathbf{V}^H$, if and only if $\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V} = \mathbf{I}$, since $\mathbf{V}^H\mathbf{V} = \mathbf{I}$. This establishes that (i) and (ii) are equivalent. To establish that (ii) implies (iii) and vice-versa, consider $\mathbf{Q}_1 = \begin{bmatrix} \underline{q}_{11} & \underline{q}_{12} & \cdots & \underline{q}_{1D_o} \end{bmatrix} = \mathbf{P}_{\mathcal{S}}\mathbf{Q}$. Since for every $i = 1, \ldots, D_o$, $\underline{q}_{1i} \in \mathcal{S}$, there exists a $\underline{\tilde{q}}_{1i} \in \mathbb{C}^{r \times 1}$ such that $\underline{q}_{1i} = \mathbf{V}\underline{\tilde{q}}_i$. Let $\tilde{\mathbf{Q}} = \begin{bmatrix} \underline{\tilde{q}}_1 & \underline{\tilde{q}}_2 & \cdots & \underline{\tilde{q}}_{D_o} \end{bmatrix}$, so $\mathbf{Q}_1 = \mathbf{V}\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{Q}} = \mathbf{V}^H\mathbf{Q}_1$. Moreover, $\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V} = \mathbf{V}^H\mathbf{V}\tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^H\mathbf{V}\mathbf{V} = \tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^H$, so $\tilde{\mathbf{Q}} \in U_{r \times D_o}$ if and only if (ii) holds. Finally, to show that (iii) implies (iv), notice that by letting $\tilde{\mathbf{Q}} = \mathbf{V}^H\mathbf{Q}$, we have $\mathbf{V}\tilde{\mathbf{Q}} = \mathbf{V}\mathbf{V}^H\mathbf{Q} = \mathbf{P}_{\mathcal{S}}\mathbf{Q}$ and for every $\underline{x} \in \mathcal{S}$, there exists $\tilde{x} \in \mathbb{C}^r$ such that $\underline{x} = \mathbf{V}\underline{\tilde{x}}$. Thus, $\mathbf{Q}^H\underline{x} = \mathbf{Q}^H\mathbf{P}_{\mathcal{S}}\underline{x} = (\mathbf{P}_{\mathcal{S}}\mathbf{Q})^H\underline{x} = (\tilde{\mathbf{Q}}\mathbf{V}^H)\mathbf{V}\underline{\tilde{x}} = \tilde{\mathbf{Q}}^H\underline{\tilde{x}}$ and hence, $\| \mathbf{Q}^H\underline{x} \| = \| \tilde{\mathbf{Q}}^H\underline{\tilde{x}} \|$. It is then enough to show that (iv) is equivalent to $\tilde{\mathbf{Q}} \in U_{r \times D_o}$. Indeed, $\| \mathbf{Q}^H\underline{x} \| = \| \underline{x} \|$ for every $\underline{x} \in \mathcal{S}$ if and only if $\| \tilde{\mathbf{Q}}^H\underline{\tilde{x}} \| = \| \underline{\tilde{x}} \|$ for every $\underline{\tilde{x}} \in \mathbb{C}^r$, which, in turn, occurs if and only if $\tilde{\mathbf{Q}} \in U_{r \times D_o}$. This concludes the proof.

The previous theorem provides four convenient characterizations of subspace isometry. Perhaps the most intuitive definition, (iv), says that $\mathbf{Q}$ is unitary with respect to $\mathcal{S}$ if it doesn't change the length of any vector in $\mathcal{S}$. (iii) simply says that an $\mathcal{S}$-unitary transformation can be represented in a coordinate system local to $\mathcal{S}$ with an outer unitary transformation. (i) and (ii) are simply convenient restatements of (iii) that will prove useful in subsequent discussions. For the remainder of the discussion, we will use the notation of the previous proof without re-introduction.

**Definition:** A matrix $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$ preserves orthogonality with respect to $\mathcal{S}$ if any one of the following equivalent statements are true.

(i) $\mathbf{P}_{\mathcal{S}}\mathbf{Q}\mathbf{Q}^H\mathbf{P}_{\mathcal{S}}^{\perp} = \mathbf{0}$
(ii) $\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V}_{\perp} = \mathbf{0}$
(iii) $\mathbf{Q}^H\mathcal{S}^{\perp} \subseteq (\mathbf{Q}^H\mathcal{S})^{\perp}$

**Proof:**

To show (i) is equivalent to (ii), note that $\mathbf{0} = \mathbf{P}_{\mathcal{S}}\mathbf{Q}\mathbf{Q}^H\mathbf{P}_{\mathcal{S}}^{\perp} = \mathbf{V}\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V}_{\perp}\mathbf{V}_{\perp}^H$ if and only if $\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V}_{\perp} = \mathbf{0}$. Now let us show that (ii) is equivalent to (iii). Given $\underline{y}_{\perp} \in \mathbf{Q}^H\mathcal{S}^{\perp}$, there exists a $\underline{\tilde{x}}_{\perp} \in \mathbb{C}^{(D_i - r)\times 1}$ such that $\underline{y}_{\perp} = \mathbf{Q}^H\mathbf{V}_{\perp}\underline{\tilde{x}}_{\perp}$. Given $\underline{y} \in \mathbf{Q}^H\mathcal{S}$ there exists a $\underline{\tilde{x}} \in \mathbb{C}^{r\times 1}$ such that $\underline{y} = \mathbf{Q}^H\mathbf{V}\underline{\tilde{x}}$. Now, (iii) is true if and only if $\forall \underline{y}_{\perp} \in \mathbf{Q}^H\mathcal{S}^{\perp}, \underline{y} \in \mathbf{Q}^H\mathcal{S}$, $\underline{y}^H\underline{y}_{\perp} = 0$. This is true if and only if $\forall \underline{\tilde{x}}_{\perp} \in \mathbb{C}^{(D_i - r)\times 1}, \underline{\tilde{x}} \in \mathbb{C}^{r\times 1}$, $0 = (\mathbf{Q}^H\mathbf{V}\underline{\tilde{x}})^H(\mathbf{Q}^H\mathbf{V}_{\perp}\underline{\tilde{x}}_{\perp}) = \underline{\tilde{x}}\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V}_{\perp}\underline{\tilde{x}}_{\perp}$, which, in turn, is true if and only if $\mathbf{V}^H\mathbf{Q}\mathbf{Q}^H\mathbf{V}_{\perp} = \mathbf{0}$. This concludes the proof.

Recall that every $\underline{x} \in \mathbb{C}^{D_i\times 1}$ can be decomposed into a component, $\underline{x}_{\mathcal{S}} \in \mathcal{S}$ and a component, $\underline{x}_{\perp} \in \mathcal{S}^{\perp}$ as follows: $\underline{x} = \left(\mathbf{P}_{\mathcal{S}} + \mathbf{P}_{\mathcal{S}}^{\perp}\right)\underline{x} = \mathbf{P}_{\mathcal{S}}\underline{x} + \mathbf{P}_{\mathcal{S}}^{\perp}\underline{x} = \underline{x}_{\mathcal{S}} + \underline{x}_{\perp}$. (iii) of the previous definition says that when $\mathbf{Q}^H$ maps a point $\underline{x}$ to a new space, according to the mapping, $\underline{y} = \mathbf{Q}^H\underline{x} = \mathbf{Q}^H\underline{x}_{\mathcal{S}} + \mathbf{Q}^H\underline{x}_{\perp} = \underline{y}_{\mathbf{Q}^H\mathcal{S}} + \underline{y}_{\perp}$, the orthogonality of $\underline{x}_{\mathcal{S}}$ and $\underline{x}_{\perp}$ are preserved in $\underline{y}_{\mathbf{Q}^H\mathcal{S}}$ and $\underline{y}_{\perp}$. That is, $\underline{y}_{\perp}^H\underline{y}_{\mathbf{Q}^H\mathcal{S}} = 0$.

If a matrix is unitary with respect to a subspace, it has the useful property that the linear map it represents and projections can be considered in any order.

**Lemma:** If $\mathbf{Q}$ preserves orthogonality with respect to $\mathcal{S}$, then $\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}\mathbf{Q}^H = \mathbf{Q}^H\mathbf{P}_{\mathcal{S}}$

**Proof:**

First note that $\mathbf{P}_{\mathbf{Q}^H\mathcal{S}}\mathbf{Q}^H\mathbf{P}_{\mathcal{S}}^{\perp} = \mathbf{0}$ since $\mathcal{R}\left(\mathbf{Q}^H\mathbf{P}_{\mathcal{S}}^{\perp}\right) = \mathbf{Q}^H\mathcal{S}^{\perp} \subseteq \left(\mathbf{Q}^H\mathcal{S}\right)^{\perp}$. Also, $\mathbf{P}_{\mathbf{Q}^H\mathcal{S}}\mathbf{Q}^H\mathbf{P}_{\mathcal{S}} = \mathbf{Q}^H\mathbf{P}_{\mathcal{S}}$, since $\mathcal{R}\left(\mathbf{Q}^H\mathbf{P}_{\mathcal{S}}\right) = \mathbf{Q}^H\mathcal{S}$. Thus,

$$\begin{aligned}
\mathbf{P}_{\mathbf{Q}^H\mathcal{S}}\mathbf{Q}^H &= \mathbf{P}_{\mathbf{Q}^H\mathcal{S}}\mathbf{Q}^H\left(\mathbf{P}_{\mathcal{S}} + \mathbf{P}_{\mathcal{S}}^{\perp}\right)\\
&= \mathbf{P}_{\mathbf{Q}^H\mathcal{S}}\mathbf{Q}^H\mathbf{P}_{\mathcal{S}} + \mathbf{P}_{\mathbf{Q}^H\mathcal{S}}\mathbf{Q}^H\mathbf{P}_{\mathcal{S}}^{\perp}\\
&= \mathbf{P}_{\mathbf{Q}^H\mathcal{S}}\mathbf{Q}^H\mathbf{P}_{\mathcal{S}}\\
&= \mathbf{Q}^H\mathbf{P}_{\mathcal{S}}
\end{aligned}$$

This concludes the proof.

**Definition:** A matrix $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$ is *unitary with respect to* $\mathcal{S}$*,* or $\mathcal{S}$*-unitary* if it is both $\mathcal{S}$-isometric and preserves orthogonality with respect to $\mathcal{S}$.

Subspace-unitary matrices are a relaxation of unitary matrices. All unitary matrices are also subspace-unitary matrices for two reasons. Firstly, a unitary matrix is isometric on every linear subspace. Secondly, the rigid-rotation that a unitary matrix represents preserves orthogonality of *all* orthogonal vectors, and in particular, it preserves the orthogonality of $\mathcal{S}$ and $\mathcal{S}^{\perp}$. However, subspace-unitary matrices are not necessarily unitary; in fact, they do not even need to be square! But when a subspace-unitary matrix is not strictly unitary, it is not isometric on $\mathcal{S}^{\perp}$.

**Notation:** $U^{D_i \times D_o}(\mathcal{S})$ is the set of all $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$, *unitary with respect to* $\mathcal{S}$.

On a final note, it might be argued that what we have defined to be "unitary with respect to $\mathcal{S}$", might more aptly be called "outer unitary with respect to $\mathcal{S}$". However, for brevity we will adhere to the former. It is easy to conceive of a similar definition of *inner unitary with respect to* $\mathcal{S}$ but this will not prove useful for us.

There is an intimate relationship between white vector random variables (e.g. uncorrelated) and unitary matrices. The terms white and colored are borrowed from the terminology used to describe white and colored stochastic processes. Namely, if $\underline{z} \in \mathbb{C}^D$ is a white vector random variable (i.e. has an auto-correlation matrix $\mathbf{\Phi}_{\underline{zz}} = E\left[\underline{zz}^H\right] = \sigma_z^2 \mathbf{I}$), then for every $\mathbf{Q} \in U_{D\times D}$, $\underline{\tilde{z}} = \mathbf{Q}^H \underline{z}$ is also white. One might expect a similar relationship to hold for subspace unitary matrices. It is natural to say that a random vector, $\underline{z}$, is *white in* $\mathcal{S}$ if $\mathbf{V}^H \underline{z}$ is white. That is, $E\left[(\mathbf{V}^H \underline{z})(\mathbf{V}^H \underline{z})^H\right] = \mathbf{V}^H \mathbf{\Phi}_{\underline{zz}} \mathbf{V} = \sigma_z^2 \mathbf{I}$ for some $\sigma_z^2$. The following equivalent definition is easier to work with and does not depend upon a particular choice of basis for $\mathcal{S}$.

**Definition:** A vector random variable, $\underline{z}$, with auto-correlation matrix $\mathbf{\Phi}_{\underline{zz}} = E\left[\underline{zz}^H\right]$ is said to be *white in* $\mathcal{S}$ if $\mathbf{P}_{\mathcal{S}} \mathbf{\Phi}_{\underline{zz}} \mathbf{P}_{\mathcal{S}} = \sigma_z^2 \mathbf{P}_{\mathcal{S}}$ for some real, positive $\sigma_z$.

**Definition:** A vector random variable is said to be *colored in* $\mathcal{S}$ if it is not white respect to $\mathcal{S}$.

The following theorem provides the link between subspace unitary matrices and subspace coloring. Namely, subspace unitary matrices are exactly those that preserve subspace coloring.

**Theorem:** Let $\underline{z} \in \mathbb{C}^{D_i}$ be a random vector that is white in $\mathcal{S}$, and $\mathbf{Q} \in \mathbb{C}^{D_i \times D_o}$. Then $\mathbf{Q}$ is unitary with respect to $\mathcal{S}$ if and only if $\dim(\mathbf{Q}^H(\mathcal{S})) = \dim(\mathcal{S})$ and $\underline{\tilde{z}} \triangleq \mathbf{Q}^H \underline{z}$ is white in $\mathbf{Q}^H(\mathcal{S})$.

**Proof:**

Let $\tilde{\mathbf{Q}} = \mathbf{V}^H\mathbf{Q}$, $\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}} = E\left[(\mathbf{Q}^H\underline{z})(\mathbf{Q}^H\underline{z})^H\right] = \mathbf{Q}^H\mathbf{\Phi}_{\underline{zz}}\mathbf{Q}$, and $\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})} = (\mathbf{Q}^H\mathbf{V})(\mathbf{Q}^H\mathbf{V})^\dagger = \tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}^{H\dagger} = \mathbf{P}_{\mathcal{R}(\tilde{Q}^H)}$.

First, suppose $\mathbf{Q}$ is unitary with respect to $\mathcal{S}$ and $\mathbf{P}_{\mathcal{S}}\mathbf{\Phi}_{\underline{zz}}\mathbf{P}_{\mathcal{S}} = \mathbf{P}_{\mathcal{S}}$. Then, $\tilde{\mathbf{Q}} = \mathbf{V}^H\mathbf{Q} \in U_o^{r\times D_o}$, $\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})} = \tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}^{H\dagger} = \tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}$, and $\dim(\mathbf{Q}^H(\mathcal{S})) = \dim(\mathcal{R}(\mathbf{Q}^H\mathbf{V})) = rank(\tilde{\mathbf{Q}}) = r$. We must show that $\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})} = \mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}$. By the previous lemma, we have

$$\begin{aligned}
\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})} &= \mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}\mathbf{Q}^H\mathbf{\Phi}_{\underline{zz}}\mathbf{Q}\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})} \\
&= \mathbf{Q}^H\mathbf{P}_{\mathcal{S}}\mathbf{\Phi}_{\underline{zz}}\mathbf{P}_{\mathcal{S}}\mathbf{Q} \\
&= \mathbf{Q}^H\mathbf{P}_{\mathcal{S}}\mathbf{Q} \\
&= \mathbf{Q}^H\mathbf{V}\mathbf{V}^H\mathbf{Q} \\
&= \tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}} \\
&= \mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}
\end{aligned}$$

Conversely, suppose $\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})} = \mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}$ and $\dim(\mathbf{Q}^H(\mathcal{S})) = \dim(\mathcal{S}) = r$. Since, $\dim(\mathbf{Q}^H(\mathcal{S})) = \dim(\mathcal{R}(\mathbf{Q}^H\mathbf{V})) = rank(\tilde{\mathbf{Q}})$, we have $rank(\tilde{\mathbf{Q}}) = r$. Moreover, since $\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}}\mathbf{P}_{\mathbf{Q}^H(\mathcal{S})} = \mathbf{P}_{\mathbf{Q}^H(\mathcal{S})}$ and $\mathbf{\Phi}_{\underline{\tilde{z}}\underline{\tilde{z}}} = \mathbf{Q}^H\mathbf{\Phi}_{\underline{zz}}\mathbf{Q} = (\tilde{\mathbf{Q}}^H\mathbf{V}^H)\mathbf{\Phi}_{\underline{zz}}(\mathbf{V}\tilde{\mathbf{Q}}) = \tilde{\mathbf{Q}}^H(\mathbf{V}^H\mathbf{\Phi}_{\underline{zz}}\mathbf{V})\tilde{\mathbf{Q}} = \tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}$, $\mathbf{P}_{\mathcal{R}(\tilde{Q}^H)}\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\mathbf{P}_{\mathcal{R}(\tilde{Q}^H)} = (\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}^{H\dagger})$. But $\mathbf{P}_{\mathcal{R}(\tilde{Q}^H)}\tilde{\mathbf{Q}}^H = \tilde{\mathbf{Q}}^H$ and $\tilde{\mathbf{Q}}\mathbf{P}_{\mathcal{R}(\tilde{Q}^H)} = \tilde{\mathbf{Q}}$, so $\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}} = \tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}^{H\dagger}$, which, since $\tilde{\mathbf{Q}}$ is rank $r$, can only happen if $\tilde{\mathbf{Q}} \in U_o^{r\times D_o}$. Thus, by the third characterization of subspace unitary matrices, $\mathbf{Q}$ is unitary with respect to $\mathcal{S}$. This completes the proof.

We will conclude the discussion of subspace-unitary matrices with a theorem that establishes a connection with strict unitary matrices.

**Theorem:** Consider a linear subspace $\mathcal{S} \subseteq \mathbb{C}^{D_i}$. For every matrix, $\mathbf{Q}_1 \in \mathbb{C}^{D_i\times D_o}$, that is unitary with respect to $\mathcal{S}$, there exists a $\mathbf{Q}_2 \in \mathbb{C}^{D_i\times D_o}$ with $\mathcal{R}(\mathbf{Q}_2) \subseteq \mathcal{S}^\perp$, such that if $\mathbf{Q} = \mathbf{Q}_1 + \mathbf{Q}_2$, then

$$\begin{aligned}
&\mathbf{Q}\mathbf{Q}^H = I,\ D_i \le D_o \\
&\mathbf{Q}^H\mathbf{Q} = I,\ D_i \ge D_o
\end{aligned}$$

Here, we call $\mathbf{Q}_2$ the $\mathcal{S}$*-unitary completion of* $\mathbf{Q}_1$.

**Proof:**

Since $\mathbf{Q}_1$ is unitary with respect to $\mathcal{S}$, $\mathcal{R}(\mathbf{Q}_1) \supseteq \mathcal{S}$. It is enough to consider $\mathcal{R}(\mathbf{Q}_1) = \mathcal{S}$. Otherwise, if $\mathbf{P}_{\mathcal{S}}^{\perp}\mathbf{Q}_1 \neq \mathbf{0}$, and we find the completion of $\mathbf{P}_{\mathcal{S}}\mathbf{Q}_1$, $\mathbf{Q}_2'$, and then we can let $\mathbf{Q}_2 = \mathbf{Q}_2' - \mathbf{P}_{\mathcal{S}}^{\perp}\mathbf{Q}_1$.

Now, with $\dim(\mathcal{S}) = r$, two trivial cases occur: when $r = D_i$ or $r = D_o$. When $r = D_i$, then $\mathcal{S} = \mathbb{C}^{D_i}$ and so by hypothesis, $\mathbf{Q}_1\mathbf{Q}_1{}^H = \mathbf{I}$ and $\mathbf{Q}_2 = \mathbf{0}$ satisfies the theorem. When $r = D_o$, $\mathbf{Q}_1{}^H\mathbf{Q}_1 = \mathbf{I}$ by hypothesis, and again, it is sufficient to choose $\mathbf{Q}_2 = \mathbf{0}$. Hence, we may now consider the case when both $r < D_i$ and $r < D_o$. In the remaining discussion, let $\mathbf{V} \in \mathbb{C}^{D_i \times r}$ be matrix whose columns for an orthonormal basis for $\mathcal{S}$ and $\tilde{\mathbf{Q}}_1 \in \mathbb{C}^{r \times D_o}$ be such that $\mathbf{Q}_1 = \mathbf{V}_1\tilde{\mathbf{Q}}_1$ and $\tilde{\mathbf{Q}}_1\tilde{\mathbf{Q}}_1^H = \mathbf{I}$.

Now consider the case when $r < D_i \leq D_o$. Then $\mathbf{Q}_1\mathbf{Q}_1^H = \mathbf{V}_1\tilde{\mathbf{Q}}_1\tilde{\mathbf{Q}}_1^H\mathbf{V}_1^H = \mathbf{V}_1\mathbf{V}_1^H$. Let $\mathbf{V}_2 \in \mathbb{C}^{D_i \times (D_i - r)}$ be an orthonormal basis for $\mathcal{S}^{\perp}$. Let $\tilde{\mathbf{Q}}_2 \in \mathbb{C}^{(D_i - r) \times D_o}$ be a matrix whose rows are any $D_i - r$ size subset of an orthonormal basis for $\mathcal{R}(\tilde{\mathbf{Q}}_1^H)^{\perp}$. By choice, $\tilde{\mathbf{Q}}_1\tilde{\mathbf{Q}}_2^H = \mathbf{0}$. Finally, let $\mathbf{Q}_2 = \mathbf{V}_2\tilde{\mathbf{Q}}_2$. Then $\mathbf{Q}_1\mathbf{Q}_2^H = \mathbf{V}_1\tilde{\mathbf{Q}}_1\tilde{\mathbf{Q}}_2^H\mathbf{V}_2^H = \mathbf{0}$ and $\mathbf{Q}_2\mathbf{Q}_2^H = \mathbf{V}_2\tilde{\mathbf{Q}}_2\tilde{\mathbf{Q}}_2^H\mathbf{V}_2^H = \mathbf{V}_2\mathbf{V}_2^H$. Let $\mathbf{Q} = \mathbf{Q}_1 + \mathbf{Q}_2$ and $\mathbf{V} = [\mathbf{V}_1 \quad \mathbf{V}_2]$. Then

$$
\begin{aligned}
\mathbf{Q}\mathbf{Q}^H &= \mathbf{Q}_1\mathbf{Q}_1^H + \mathbf{Q}_1\mathbf{Q}_2^H + \mathbf{Q}_2\mathbf{Q}_1^H + \mathbf{Q}_2\mathbf{Q}_2^H \\
&= \mathbf{Q}_1\mathbf{Q}_1^H + \mathbf{Q}_2\mathbf{Q}_2^H \\
&= \mathbf{V}_1\mathbf{V}_1^H + \mathbf{V}_2\mathbf{V}_2^H \\
&= \mathbf{V}\mathbf{V}^H \\
&= \mathbf{I}
\end{aligned}
$$

Now consider the case when $r < D_o \leq D_i$. We will construct a $\mathbf{Q}_2 \in \mathbb{C}^{D_i \times D_o}$ in a manner similar to the previous case. Then, $\mathbf{Q}_1^H\mathbf{Q}_1 = \tilde{\mathbf{Q}}_1\mathbf{V}_1^H\mathbf{V}_1\tilde{\mathbf{Q}}_1^H = \tilde{\mathbf{Q}}_1\tilde{\mathbf{Q}}_1^H$. Let $\mathbf{V}_2 \in \mathbb{C}^{D_i \times r}$ be a matrix whose columns are chosen from an orthonormal basis for $\mathcal{S}^{\perp}$. (There may be many possible choices). Let $\tilde{\mathbf{Q}}_2 \in \mathbb{C}^{(D_o - r)\times D_o}$ be any orthonormal basis for $\mathcal{R}(\tilde{\mathbf{Q}}_1^H)^{\perp}$. Let $\mathbf{Q}_2 = \mathbf{V}_2\tilde{\mathbf{Q}}_2$. Then $\mathbf{Q}_2^H\mathbf{Q}_1 = \tilde{\mathbf{Q}}_2^H\mathbf{V}_2^H\mathbf{V}_1\tilde{\mathbf{Q}}_1 = \mathbf{0}$ and $\mathbf{Q}_2^H\mathbf{Q}_2 = \tilde{\mathbf{Q}}_2^H\mathbf{V}_2^H\mathbf{V}_2\tilde{\mathbf{Q}}_2 = \tilde{\mathbf{Q}}_2^H\tilde{\mathbf{Q}}_2$. Similar to before, letting $\mathbf{Q} = \mathbf{Q}_1 + \mathbf{Q}_2$ and $\mathbf{V} = [\mathbf{V}_1 \quad \mathbf{V}_2]$, we have $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$. This completes the proof.

# Appendix B: SPECTRAL FLATNESS MEASURES FOR MULTI-CHANNEL RANDOM PROCESSES

It is well known that a single channel wide-sense stationary (WSS) $N^{th}$ order auto-regressive (AR) random process can be whitened with an $N^{th}$ order monic finite-impulse-response (FIR) filter, whose coefficients are the solution to the Yule-Walker Equations (YWE). Moreover, a vector wide-sense stationary (VWSS) $N^{th}$ order AR process can be whitened and de-correlated with an $N^{th}$ order multiple input multiple output (MIMO) FIR filter, whose coefficients are the solution to the MIMO YWE [39-40]. However, an FIR filter cannot whiten an AR process whose order is greater than the filter order, nor can it whiten moving average processes. Nevertheless, in many cases an FIR filter of an appropriate order can approximately whiten such processes. For single channel random processes, the notion of "approximate-whitening" has been formalized in terms of the following spectral flatness measure. If $P_{vv}(f)$ is the power spectral density of a random process, $v[n]$, then the spectral flatness measure (SFM) of $v[n]$ is then,

$$\xi_v = \frac{\exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln P_{vv}(f)df\right\}}{\int_{-\frac{1}{2}}^{\frac{1}{2}} P_{vv}(f)df}. \tag{B.1}$$

The criterion of Equation (B.1) is normalized, $0 \le \xi_v \le 1$, reaching its maximum value at $\xi_v = 1$. Although it places more emphasis on peaks than nulls, it has the benefit that its maximum is obtained by a solution to the well known Yule-Walker Equation

(YWE). This appendix generalizes this SFM to VWSS random processes. It is shown that a solution to the MIMO YWEs yields the filter coefficients of the monic MIMO filter that maximizes spectral flatness.

In the discussion we will use the following definitions. Given a VWSS random process, $\underline{v}[n] \in \mathbb{C}^{n\times 1}$, we will denote its matrix auto-correlation function as $\mathbf{\Phi}_{\underline{vv}}[\Delta n] = E\left[\underline{v}[n]\underline{v}[n+\Delta n]^H\right]$, whose z-transform is $\mathbf{\Phi}_{\underline{vv}}(z)$.

**Definition:** The spectral flatness measure (SFM) of a full-rank VWSS random process, $\underline{v}[n] \in \mathbb{C}^{n\times 1}$, is

$$\xi_{\underline{v}} = \frac{\exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} Tr\left[\ln \mathbf{\Phi}_{\underline{vv}}(e^{j2\pi f})\right] df\right\}}{\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{n} Tr\left[\mathbf{\Phi}_{\underline{vv}}(e^{j2\pi f})\right] df\right)^n} \tag{B.2}$$

The integral of (B.2) has a well defined indefinite form when the points on the unit circle at which $\mathbf{\Phi}_{\underline{vv}}(z)$ becomes singular are isolated. This is the case when $\mathbf{\Phi}_{\underline{vv}}(z)$ is analytic on a neighborhood of the unit circle. Secondly, like the single-channel SFM, we will follow the convention that $\xi_{\underline{v}} = 0$ if $\mathbf{\Phi}_{\underline{vv}}(z)$ is singular on a measurable subset of the unit circle. This convention is in accordance with the following intuitive reasoning. In this case, the argument of the numerator's integral becomes $-\infty$ on a measurable set. Hence, the argument of the exponent is $-\infty$, while the denominator is finite. Finally, we claim that $0 \le \xi_{\underline{v}} \le 1$ , reaching its upper bound only when the random process has auto-correlation function of the form $\mathbf{\Phi}_{\underline{vv}}(z) = \sigma_v^2 \mathbf{I}$. The first inequality follows from the fact that the denominator of (B.2) is positive, since symmetric positive definite matrices have a positive trace. The second inequality will be established shortly.

To understand what this SFM measures, consider, for each value of $f \in [-1/2, 1/2]$, an eigen-value decomposition of $\mathbf{\Phi}_{\underline{vv}}(e^{j2\pi f}) = \mathbf{U}(f)\mathbf{\Lambda}(f)\mathbf{U}(f)^H$. This eigen-value decomposition exists because, by definition, $\mathbf{\Phi}_{\underline{vv}}(e^{j2\pi f})$ is a symmetric matrix for every $f \in [-1/2, 1/2]$. The eigenvector function, $\mathbf{U}(f)$ describes the geometric structure of $\mathbf{\Phi}_{\underline{vv}}(z)$ while the elements of the diagonal matrix $\mathbf{\Lambda}(f)$, $\{\lambda_i(f)\}_{i=1}^{n}$, describe the *principle spectra* of $\mathbf{\Phi}_{\underline{vv}}(z)$. Hence, the MIMO-SFM can be re-expressed as

$$\begin{aligned}
\xi_{\underline{v}} &= \frac{\exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} Tr\left[\ln\left(\mathbf{U}(f)\mathbf{\Lambda}(f)\mathbf{U}(f)^H\right)\right]df\right\}}{\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{n} Tr\left[\mathbf{U}(f)\mathbf{\Lambda}(f)\mathbf{U}(f)^H\right]df\right)^n} \\
&= \frac{\exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} Tr\left[\ln\left(\mathbf{\Lambda}(f)\right)\right]df\right\}}{\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{n} Tr\left[\mathbf{\Lambda}(f)\right]df\right)^n} \\
&= \frac{\exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{i=1}^{n} \ln \lambda_i(f) df\right\}}{\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{n} \sum_{i=1}^{n} \lambda_i(f) df\right)^n}
\end{aligned} \tag{B.3}$$

That is to say, from the point of view of spectral flatness, the geometric structure of the process' cross-spectra is superficial. Only, the principle spectra affect flatness. Letting $P_i = \int_{-1/2}^{1/2} \lambda_i(f) df$ and $\xi_i$ denote the power and SFM of the $i^{th}$ principle spectrum, respectively, we have,

$$
\begin{aligned}
\xi_{\underline{v}} &= \frac{\prod_{i=1}^{n} \exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \lambda_i(f)\,df\right\}}{\left(\frac{1}{n}\sum_{i=1}^{n} P_i\right)^n} \\
&= \frac{\prod_{i=1}^{n} \exp\left\{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln \lambda_i(f)\,df\right\}}{\prod_{i=1}^{n} P_i} \frac{\prod_{i=1}^{n} P_i}{\left(\frac{1}{n}\sum_{i=1}^{n} P_i\right)^n} \\
&= \left(\prod_{i=1}^{n} \xi_i\right) \frac{\prod_{i=1}^{n} P_i}{\left(\frac{1}{n}\sum_{i=1}^{n} P_i\right)^n}
\end{aligned}
\tag{B.4}
$$

This final equality splits the MIMO-SFM into the product of the principle SFMs and a geometric factor which describes the effect that cross-channel correlation has upon the SFM. Also, Equation (B.4) makes the promised inequality $\xi_{\underline{v}} \le 1$ immediately apparent. Obviously, for each value of $i$, $\xi_i \le 1$, since it is a single-channel SFM; but the geometric factor is also less than one by the following well known *geometric-mean inequality* [59]:

$$
\sqrt[n]{\prod_{i=1}^{n} P_i} \le \frac{1}{n}\sum_{i=1}^{n} P_i
\tag{B.5}
$$

with equality if and only if, $P_i = P$, a real-non-negative constant, for every $i$. Hence, the MIMO-SFM reaches its maximum value of unity only if all principle spectra are identical and flat; that is, $\lambda_i(f) = P, \forall i$, for some real positive constant, $P$. This can happen only if $\mathbf{\Phi}_{\underline{vv}}(z) = P\mathbf{I}$.

We have thus far shown that the multi-channel SFM is a normalized spectral flatness measure that only measures the flatness of principle spectra and geometric asymmetries across channels. Moreover, the multi-channel SFM is one when the process is decorrelated across channels and across time-samples. We will now move toward proving the central result of this Appendix by first establishing some pre-requisite Lemmas.

**Proposition:** For upper triangular matrix, $\mathbf{T} \in \mathbb{C}^{nxn}$, and non-negative integer, $k$, $\mathbf{T}^k$ has the form

$$\mathbf{T}^k = \begin{bmatrix} t_{11}^k & & * \\ & \ddots & \\ 0 & & t_{nn}^k \end{bmatrix}$$

**Proof:**

Let $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ be upper triangular. Then,

$$\begin{aligned} [\mathbf{AB}]_{i,i} &= \sum_{j=1}^{n} a_{ij} b_{ji} \\ &= \sum_{j=1}^{i-1} a_{ij} b_{ji} + a_{ii} b_{ii} + \sum_{j=i+1}^{n} a_{ij} b_{ji} \\ &= a_{ii} b_{ii} \end{aligned} \tag{B.6}$$

where the last line follows from the following fact. For every term in the first sum of line two, $a_{ij} = 0$, and every term of the last sum of line two, $b_{ji} = 0$. Applying $\mathbf{T} = \mathbf{A}$, $\mathbf{T} = \mathbf{B}$ establishes that $\left[\mathbf{T}^2\right]_{i,i} = t_{ii}^2$. Moreover, an induction argument with $\mathbf{T}^{k-1} = \mathbf{A}$, $\mathbf{T} = \mathbf{B}$, establishes the desired result. This concludes the proof.

**Lemma:** For every $\mathbf{M} \in \mathbb{C}^{n\times n}$, $\det(e^{\mathbf{M}}) = e^{Tr\mathbf{M}}$

**Proof:**

Let $\mathbf{M} = \mathbf{UTU}^H$ a Schur decomposition of $\mathbf{M}$. Recalling the definition of a matrix-exponent,

$$e^{\mathbf{T}} = \sum_{k=0}^{\infty} \frac{\mathbf{T}^k}{k!}. \tag{B.7}$$

By the previous proposition,

$$\begin{aligned} \left[e^{\mathbf{T}}\right]_{i,i} &= \sum_{k=0}^{\infty} \frac{\left[\mathbf{T}^k\right]_{i,i}}{k!} \\ &= \sum_{k=0}^{\infty} \frac{t_{ii}^k}{k!} \\ &= e^{t_{ii}}. \end{aligned} \tag{B.8}$$

where the second line follows from the fact that $\mathbf{T}$ is upper triangular. But,

$$\begin{aligned} e^{\mathbf{M}} &= \sum_{k=0}^{\infty} \frac{\mathbf{M}^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{\left(\mathbf{U}\mathbf{T}\mathbf{U}^H\right)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{\mathbf{U}\mathbf{T}^k\mathbf{U}^H}{k!} \\ &= \mathbf{U}\left(\sum_{k=0}^{\infty} \frac{\mathbf{T}^k}{k!}\right)\mathbf{U}^H \\ &= \mathbf{U}e^{\mathbf{T}}\mathbf{U}^H \end{aligned} \tag{B.9}$$

where the third line follows from the fact that $\mathbf{U}$ is a unitary matrix. Thus,

$$\begin{aligned} \det(e^{\mathbf{M}}) &= \det(\mathbf{U})\det(e^{\mathbf{T}})\det(\mathbf{U}^H) \\ &= \det(e^{\mathbf{T}}) \\ &= \prod_{i=1}^{n} \left[e^{\mathbf{T}}\right]_{i,i} \\ &= \prod_{i=1}^{n} e^{t_{ii}} \\ &= \exp \sum_{i=1}^{n} t_{ii} \\ &= e^{Tr\mathbf{T}} \\ &= e^{Tr\mathbf{M}}. \end{aligned}$$

Here, the second line follows from the fact that $\mathbf{U}$ is unitary; the third follows from the fact that $e^{\mathbf{T}}$ is upper-triangular; the fourth from a previous lemma; and the last from the fact that $\mathbf{T}$ and $\mathbf{M}$ are similar matrices. This concludes the proof.

Of course, for any two real positive numbers, $\log(ab) = \log(a) + \log(b)$. However, for matrices, this is not usually the case. However, as the following Lemma shows, this identity is preserved under the "real-trace" operation.

**Lemma** (Beattie)**:** For non-singular, $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$, $\operatorname{Re} Tr \log \mathbf{AB} = \operatorname{Re} Tr(\log \mathbf{A} + \log \mathbf{B})$, where log(x) is any branch of the natural log for which these quantities are defined.

**Proof:**

By a previous proposition,

$$\exp\{Tr \log(\mathbf{AB})\} = \det(\exp(\log(\mathbf{AB}))) \,. \tag{B.10}$$

Now, recall that $\forall z \in \mathbb{C},\ |e^z| = \exp(\operatorname{Re} z)$. Hence,

$$\begin{aligned} \exp\{\operatorname{Re} Tr \log(\mathbf{AB})\} &= |\det(\exp(\log(\mathbf{AB})))| \\ &= |\det(\mathbf{AB})| \\ &= |\det(\mathbf{A})\det(\mathbf{B})| \\ &= |\det(\exp(\log(\mathbf{A})))\det(\exp(\log(\mathbf{B})))| \\ &= \exp\{\operatorname{Re} Tr(\log \mathbf{A})\}\exp\{\operatorname{Re} Tr(\log \mathbf{B})\} \\ &= \exp\{\operatorname{Re} Tr(\log \mathbf{A}) + \operatorname{Re} Tr(\log \mathbf{B})\} \end{aligned} \tag{B.11}$$

Since $e^x$ is one-to-one on the real axis, taking the log of both sides establishes the desired result. This concludes the proof.

**Definition:** A Multiple Input Multiple Output (MIMO) FIR filter with Z-transform, $\mathbf{W}(z) = \mathbf{W}_o + \mathbf{W}_1 z^{-1} + \cdots + \mathbf{W}_N z^{-N}$ is said to be *minimum phase* if

$$\int_{-1/2}^{1/2} Tr \ln\left\{ \mathbf{W}(e^{j2\pi f})\mathbf{W}(e^{j2\pi f})^H \right\} df = Tr \ln\left( \mathbf{W}_0 \mathbf{W}_0^H \right) \tag{B.12}$$

To shed light on this definition, consider the definition of a single input single output (SISO) polynomial system, $W(z) = w_o + w_1 z^{-1} + \cdots + w_N z^{-N}$ : all its roots are in the unit circle. Recall that, although every polynomial system can be factored into monic terms, this is not always the case for matrix polynomial systems [39]. However, when a matrix polynomial system *is* monic-factorable, the concept of *minimum-phase* has an intuitively appealing link to SISO systems.

**Lemma:** A monic-factorable matrix polynomial system, $\mathbf{W}(z) = \mathbf{W}_0 \prod_{i=1}^{N} \left( \mathbf{I} - z^{-1}\mathbf{W}_i \right)$ is minimum phase if and only if the eigen-values of its matrix-roots, $\{\mathbf{W}_i\}_{i=1}^{N}$, lie within the unit circle.

**Proof:**

Note for the proof of this theorem, the definitions of $\{\mathbf{W}_i\}_{i=1}^{N}$ have changed from that of the previous definition. First consider the case when $\mathbf{W}(z)$ is monic (i.e. $\mathbf{W}_0 = \mathbf{I}$ ). The proof will be based on the fact that a monic SISO polynomial system, $\tilde{W}(z)$, is minimum phase if and only if

$$\int_{-1/2}^{1/2} \ln |\tilde{W}(e^{j2\pi f})| \, df = 0 \tag{B.13}$$

Further, let $\mathbf{W}_i = \mathbf{U}_i \mathbf{T}_i \mathbf{U}_i^H$ be the Schur decomposition of the $i^{th}$ monic root and let $\lambda_{i,j}$, $1 \le i \le N,\ 1 \le j \le n$ be the $j^{th}$ eigen-value (including multiplicity) of the $i^{th}$ monic root. Then,

$$\int_{-1/2}^{1/2} Tr \ln\left\{ \mathbf{W}(e^{j2\pi f})\mathbf{W}(e^{j2\pi f})^H \right\} df = \int_{-1/2}^{1/2} Tr \ln\left\{ \left( \prod_{i=1}^{N} (\mathbf{I} - e^{-j2\pi f}\mathbf{W}_i) \right) \left( \prod_{i=1}^{N} (\mathbf{I} - e^{-j2\pi f}\mathbf{W}_i) \right)^H \right\} df \tag{B.14}$$

$$= \int_{-1/2}^{1/2} \sum_{i=1}^{N} Tr \ln\left\{(\mathbf{I} - e^{-j2\pi f}\mathbf{U}_i\mathbf{T}_i\mathbf{U}_i^{\ H}) + (\mathbf{I} - e^{j2\pi f}\mathbf{U}_i\mathbf{T}_i^H\mathbf{U}_i^{\ H})\right\} df$$
$$= \int_{-1/2}^{1/2} \sum_{i=1}^{N} Tr \ln\left\{(\mathbf{I} - e^{-j2\pi f}\mathbf{T}_i) + (\mathbf{I} - e^{j2\pi f}\mathbf{T}_i^H)\right\} df$$
$$= \int_{-1/2}^{1/2} \sum_{i=1}^{N}\sum_{j=1}^{n} Tr \ln\left\{(1 - e^{-j2\pi f}\lambda_{ij}) + (1 - e^{-j2\pi f}\lambda_{ij})^*\right\} df$$
$$= \int_{-1/2}^{1/2} \ln\left\{\prod_{i=1}^{N}\prod_{j=1}^{n} | 1 - e^{-j2\pi f}\lambda_{ij} |^2\right\} df$$

Letting $\tilde{W}(z) = \prod_{i=1}^{N}\prod_{j=1}^{n}(1 - z^{-1}\lambda_{ij})$, then by (B.13), (B.14) is zero if and only if all the terms, $\lambda_{ij}$, lie within the unit circle. This establishes the lemma's hypothesis for monic systems.

Now suppose that $\mathbf{W}(z) = \mathbf{W}_0\mathbf{W}_1(z)$ is not monic, where $\mathbf{W}_1(z) = \prod_{i=1}^{N}\left(\mathbf{I} - z^{-1}\mathbf{W}_i\right)$. Then,

$$\int_{-1/2}^{1/2} Tr \ln \mathbf{W}(e^{j2\pi f})\mathbf{W}(e^{j2\pi f})^H\, df = Tr \ln \mathbf{W}_0\mathbf{W}_0^H + \int_{-1/2}^{1/2} Tr \ln \mathbf{W}_1(e^{j2\pi f})\mathbf{W}_1(e^{j2\pi f})^H\, df$$

This left hand side evaluates to $Tr \ln \mathbf{W}_0\mathbf{W}_0^H$ if and only if the latter integral is zero, which, by the preceding argument, can only happen if all matrix-roots have eigen-values within the unit circle. This concludes the proof.

**Theorem:** Consider a VWSS random process, $\underline{i}[n]$, that is to be approximately whitened with an $N^{th}$ order $n \times n$ minimum-phase, monic, MIMO FIR filter, $\mathbf{W}(z) = \mathbf{I} + \mathbf{W}_1 z^{-1} + \cdots + \mathbf{W}_N z^{-N}$. Let $\underline{v}[n] = \mathbf{W}[n] * \underline{i}[n]$. Then a solution to the Yule-Walker Equations yields a set of filter coefficients that maximize $\xi_{\underline{v}}$.

**Proof:**

A key integral in the output SFM expression evaluates as

$$\begin{aligned}
&\int_{-1/2}^{1/2} Tr\ln[\mathbf{W}(e^{j2\pi f})\mathbf{\Phi}_{\underline{i}\,\underline{i}}(e^{j2\pi f})\mathbf{W}(e^{j2\pi f})^H]df \\
&= \int_{-1/2}^{1/2} Tr\ln[\mathbf{W}(e^{j2\pi f})\mathbf{W}(e^{j2\pi f})^H]df + \int_{-1/2}^{1/2} Tr\ln[\mathbf{\Phi}_{\underline{i}\,\underline{i}}(e^{j2\pi f})]df \\
&= \int_{-1/2}^{1/2} Tr\ln[\mathbf{\Phi}_{\underline{i}\,\underline{i}}(e^{j2\pi f})]df
\end{aligned} \tag{B.15}$$

where in the second line, we have used a previous Lemma, and the last line, we have used the fact that $\mathbf{W}(z)$ is minimum phase. Thus, the SFM of the final output process, $\underline{v}[n]$, is

$$\begin{aligned}
\xi_{\underline{v}} &= \frac{\exp\left\{\int_{-1/2}^{1/2} Tr\ln\left[\mathbf{\Phi}_{\underline{v}\underline{v}}\left(e^{j2\pi f}\right)\right]df\right\}}{\left(\frac{1}{n}\int_{-1/2}^{1/2} Tr\left[\mathbf{\Phi}_{\underline{v}\underline{v}}\left(e^{j2\pi f}\right)\right]df\right)^n} \\
&= \frac{\exp\left\{\int_{-1/2}^{1/2} Tr\ln[\mathbf{W}(e^{j2\pi f})\mathbf{\Phi}_{\underline{i}\,\underline{i}}(e^{j2\pi f})\mathbf{W}(e^{j2\pi f})^H]df\right\}}{\left(\frac{1}{n}\int_{-1/2}^{1/2} Tr\left[\mathbf{\Phi}_{\underline{v}\underline{v}}\left(e^{j2\pi f}\right)\right]df\right)^n} \\
&= \frac{\exp\left\{\int_{-1/2}^{1/2} Tr\ln\left[\mathbf{\Phi}_{\underline{i}\,\underline{i}}\left(e^{j2\pi f}\right)\right]df\right\}}{\left(\frac{1}{n}\int_{-1/2}^{1/2} Tr\left[\mathbf{\Phi}_{\underline{v}\underline{v}}\left(e^{j2\pi f}\right)\right]df\right)^n} \\
&= \xi_{\underline{i}}\left(\frac{P_{\underline{i}}}{P_{\underline{v}}}\right)^n
\end{aligned} \tag{B.16}$$

Since both the spectral flatness, $\xi_{\underline{i}}$ and power, $P_{\underline{i}}$ of the input process are fixed, the only parameter of Equation (B.16) that varies with the filter coefficients is the process output power, $P_{\underline{v}}$. Thus, a solution to the YWE, which minimizes $P_{\underline{v}}$, also maximizes spectral flatness. This concludes the proof.

# Appendix C: MODELING OF ELEMENTAL BEAM PATTERNS FOR AN ANTENNA ARRAY

The steering vector equation that describes an antenna array's response does not account for the directivity of individual elements. A more complete model must account for the patterns of individual elements. This Appendix describes how to model elemental antenna patterns in a simulation of an antenna system. The discussion proposes a generic elemental pattern model that is easy to apply, but compatible with arrays that have $360^0$ of azimuthal view, such as a uniform circular array (UCA). The Appendix concludes with a discussion of noise modeling.

In [60], [61], it is shown that if $\underline{\tilde{a}}(\varepsilon,\theta) \in \mathbb{C}^{M\times 1}$ denotes the steering vector of an M-element array with isotropic elements, and if the $m^{th}$ element of a non-isotropic antenna array has an response, $p_m(\varepsilon,\theta)$, then the steering vector's $m^{th}$ element is given as

$$a_m(\varepsilon,\theta) = \tilde{a}_m(\varepsilon,\theta)\, p_m(\varepsilon,\theta),\ 1 \le m \le M \tag{C.1}$$

Here $\varepsilon$ denotes the elevation angle of an impinging plane-wave, while $\theta$ denotes the azimuth angle. A uniform circular array (UCA) is said to have *radially directed elemental patterns* if all elements employ identical patterns with radially directed bore-sights. If each element has a beam-pattern, $p(\varepsilon,\phi)$, where $\phi$ is the azimuthal angle referenced from bore-sight, $p_m(\varepsilon,\theta) = p(\varepsilon,\theta-\theta_m)$, for each element's position angle, $\theta_m = 2\pi(m-1)/M$. It is often convenient to consider antenna gain,

$$G(\varepsilon,\phi) = |p(\varepsilon,\phi)|^2 \tag{C.2}$$

which describes the power received from a particular direction relative to an isotropic antenna; or it is convenient to specify a normalized antenna response,

$$F(\varepsilon,\phi) = \frac{p(\varepsilon,\phi)}{p_{\max}} \tag{C.3}$$

For the purpose of studying tradeoffs in antenna array systems without considering the specific details of physical antenna design, we propose the following normalized antenna response.

$$F(\varepsilon,\phi) = \mathrm{sinc}(\alpha\sin(\phi/2)) = \frac{\sin(\pi\alpha\sin(\phi/2))}{\pi\alpha\sin(\phi/2)} \tag{C.4}$$

Here $\alpha$ is a parameter that affects the pattern's beam-width. For simplicity, we have assumed no directivity in elevation-angle. Because this beam-pattern has period $2\pi$, it is useful when employing arrays with $360^0$ of azimuthal view (for instance a UCA). If 100% efficient antennas are used, the antenna's gain pattern is then

$$G(\varepsilon,\phi) = G\,|F(\varepsilon,\phi)|^2 \tag{C.5}$$

where the antenna gain, $G = 4\pi/\Omega_A$, and

$$\begin{aligned}\Omega_A &= \int\int |F(\varepsilon,\phi)|^2\, d\Omega \\ &= \int_{-\pi}^{\pi}\int_{-\pi/2}^{\pi/2} |F(\varepsilon,\phi)|^2 \cos\varepsilon\; d\varepsilon\; d\phi \\ &= 4\int_0^{\pi} |F(\varepsilon,\phi)|^2\, d\phi\end{aligned} \tag{C.6}$$

where the last line uses the fact that the pattern is symmetric about $\phi = 0$. This last integral can be computed easily with a numerical integration package. For instance MATLAB's "quad8" command can compute (C.6) to 3 significant digits. A final expression for the antenna's response becomes

$$p(\varepsilon,\phi) = \sqrt{G}F(\varepsilon,\phi) \tag{C.7}$$

The $\alpha$ parameter of the model is best interpreted in terms of the beam-width that it produces. The half-power angle is the closest positive angle to bore-sight for which the antenna gain is one half of its peak. This occurs when $|F(\varepsilon,\phi_{HP})|^2 = 1/2$. Hence the half-power beam-width, $HPBW = 2\phi_{HP}$, is given as

$$HPBW \approx 4\sin^{-1}\left(\frac{1.39}{\pi\alpha}\right),\ \alpha \gtrsim 1.39/\pi \tag{C.8}$$

The normalized gain pattern, $G(\varepsilon,\phi)/G$, for several values of HPBW is plotted in Figure C.1.

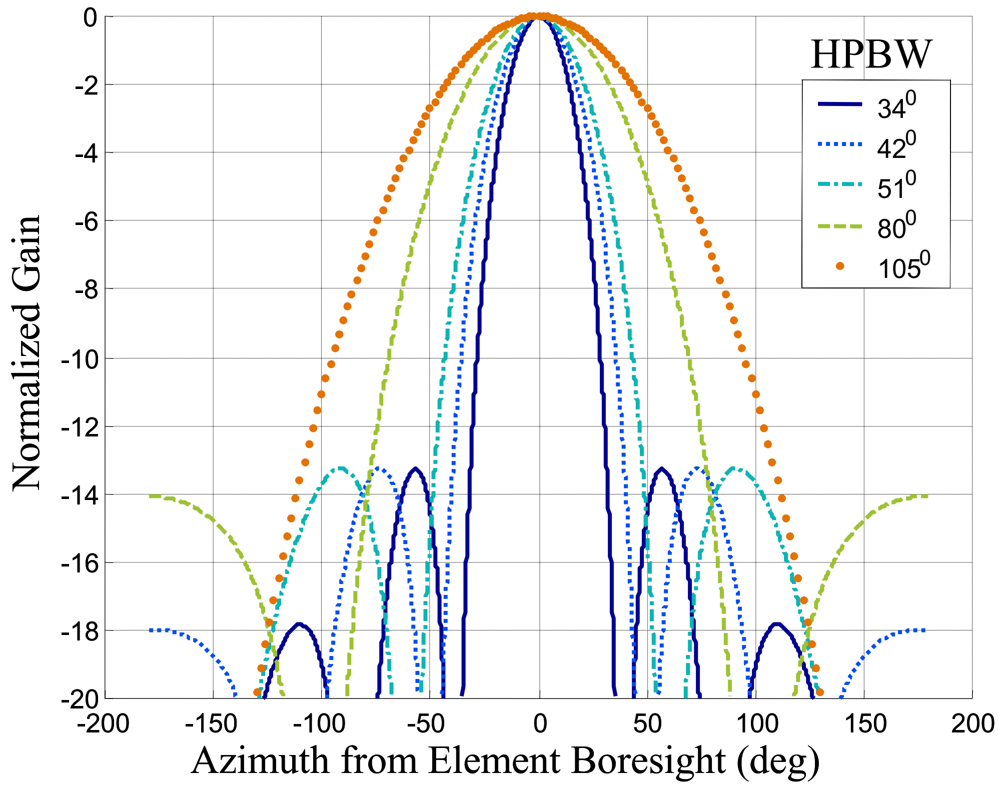


Figure C.1: Normalized Gain patterns for several values of HPBW.

While (C.7) and (C.1) together describe the effect of an element's pattern on signal amplitude, a model of noise that is consistent across multiple gain-patterns is also needed. Toward this end, define the *noise power angular density*, $P_z(\varepsilon,\phi)$, as the function that describes the spatial distribution of noise power measured in Watts/steradian as a function of azimuth and elevation. The total power received at an antenna with gain pattern $G(\varepsilon,\phi)$ is then

$$\sigma_z^2 = \int\int P_z(\varepsilon,\phi)G(\varepsilon,\phi)d\Omega \tag{C.9}$$

It is often the case that noise doesn't appear to come from any particular direction, in which case $P_z(\varepsilon,\phi) = \sigma_z^2 / 4\pi$ (W/steradian). Surprisingly, in the presence of non-directive noise, the received power at an antenna *does not change* with gain pattern. This becomes readily apparent from (C.9).

$$\begin{aligned}\int\int \frac{\sigma_z^2}{4\pi} G(\varepsilon,\phi) d\Omega &= \frac{\sigma_z^2}{4\pi}\int\int \frac{4\pi}{\Omega_A} |F(\varepsilon,\phi)|^2\, d\Omega \\ &= \frac{\sigma_z^2}{4\pi}\frac{4\pi}{\Omega_A}\int\int |F(\varepsilon,\phi)|^2\, d\Omega \\ &= \sigma_z^2\end{aligned} \tag{C.10}$$

This result is independent of Gain pattern.

When simulating an antenna array system employing elemental patterns, it is useful to define noise power based on the signal-to-noise ratio of an isotropic system. For instance, if $\underline{\tilde{a}}$ is the steering vector for a signal, $s(t)$ with variance $\sigma_s^2$, *signal-to-noise ratio per element with respect to isotropic* is

$$(SNR)_{iso} = (\|\underline{\tilde{a}}\|^2 / M)\frac{\sigma_s^2}{\sigma_z^2} \tag{C.11}$$

For phased arrays, it is usually the case that $|\tilde{a}_m| = 1$, in which case,

$$(SNR)_{iso} = \frac{\sigma_s^2}{\sigma_z^2} \tag{C.12}$$

Equations (C.1) and (C.10) say that in fixed signal environment, if the isotropic antenna patterns are replaced by directive ones, received signal for any given element will change, but noise power does not. This means that in a simulation of an array employing elemental gain-patterns, signal amplitudes should be computed using Equation (C.1), while noise variances are computed from (C.12).

# Appendix D: BIBLIOGRAPHY


[1] A. Paulraj and C. Papadias, "Space-Time Processing for Wireless Communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49-83, November 1997.

[2] Van Etten, "Maximum Likelihood Receiver for Multiple Channel Transmission Systems," *IEEE Transactions on Communications*, vol. 24, pp. 276- 283, February 1976.

[3] B. Agee, S. Bruzzone, M. Bromberg, "Exploitation of Signal Structure in Array-Based Blind Copy and Copy-Aided DF Systems," Vehicular Technology Conference, June, 1998, pp. 2489-2492.

[4] H. Yoshino, K. Fukawa, and H. Suzuki, "Interference Canceling Equalizer (ICE) for Mobile Radio Communication," *IEEE Transactions on Vehicular Technology,* vol.46, no.4, pp. 849-861, November 1997.

[5] K. Fukawa, "A Suboptimal Receiver for Joint Spatial-Temporal Filtering and Multi-User Detection in Mobile Radio Communications," *IEEE Transactions on Vehicular Technology,* vol. 49, no. 4, pp. 1196-1206, July 2000.

[6] M. Moher, "Multi-User Decoding for Multi-Beam Systems," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 4, pp. 1226-1234, July 2000.

[7] S. Talwar, M. Viberg, and A. Paulraj, "Blind Separation of Synchronous Co-Channel Digital Signals Using an Antenna Array – Part I: Algorithms," *IEEE Transactions on Signal Processing,* vol. 44, no.5, pp. 1184-1197, May 1996.

[8] S. Talwar, M. Viberg, and A. Paulraj, "Blind Separation of Synchronous Co-Channel Digital Signals Using an Antenna Array – Part II: Performance Analysis," *IEEE Transactions on Signal Processing,* vol. 45, no.3, pp. 706-718, March 1997.

[9] S. Grant and J. Cavers, "Performance Enhancement Through Joint Detection of Cochannel Signals Using Diversity Arrays," *IEEE Transactions on Communications*, vol. 46, no. 8, pp. 1038-1049, August 1998.

[10] S. Bayram, J. Hicks, J.H. Reed, and B. Boyle, " Joint ML Approach in Overloaded Array Processing," *Vehicular Technology Conference*, Boston, MA, Sept. 2000, pp. 394-400.

[11] S. Bayram, J. Hicks, J.H. Reed, and B. Boyle, “Overloaded Array Processing: Non-Linear vs. Linear Signal Extraction Techniques,” *Wireless 2000 Conference*, Calgary, Canada, July 2000, pp. 492-498.

[12] S. Bayram, J. Hicks, J.H. Reed, and B. Boyle, “Overloaded Array Processing in Wireless Airborne Communication Systems,” *MILCOM 2000*, Los Angeles, CA, October 2000, CD-ROM.

[13] J. E. Hicks, J. Tsai, J. H. Reed, W. H. Tranter, B. D. Woerner, “Overloaded Array Processing with MMSE-SIC,” *VTC, Spring, 2002.*

[14] J. E. Hicks, J. Tsai, J. H. Reed, W. H. Tranter, B. D. Woerner, “The Performance of Linear Space Time Processing in Overloaded Environments,” Wireless 2002, July, 8, 2002,Calgary, AB.

[15] J. Hicks, *Overloaded Array Processing with Spatially Reduced Search Joint Detection*, M.Sc. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, May 2000.

[16] J. E. Hicks, S. Bayram, W. H. Tranter, B. Boyle, J. H. Reed,“Overloaded Array Processing with Spatially Reduced Search Joint Detection,” *IEEE Journal on Selected Areas of Communication*, Aug. 2001.

[17] J. Tsai, J. E. Hicks, et. al., “Overloaded Array Processing with Multi-user Detection,” MILCOM 2001.

[18] J. E. Hicks, J. H. Reed, “Overloaded Array Processing with MAP-SIC,” *Motorola Science Advisory Board Association Conference*, October 19, 2002.

[19] S. Verdu, *Multi-user Detection*, Cambridge, UK, © 1998, Cambridge University Press.

[20] K. Giridhar, J.J. Shynk, A. Mathur, S. Chari, and R.P. Gooch, “Nonlinear Techniques for the Joint Estimation of Cochannel Signals,” *IEEE Transactions on Communications,* vol. 45, no. 4, pp. 473-483, April 1997.

[21] G. Forney, “Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference,” *IEEE Transactions on Information Theory*, vol. 18, pp. 363-378, May 1972.

[22] A. Duel-Hallen and C. Heegard, “Delayed Decision-Feedback sequence Estimation,” *IEEE Transactions on Communications*, vol. 37, no. 5, pp. 428-436, May 1989.

[23] G. Ungerboeck, “Adaptive Maximum-Likelihood Receiver for Carrier-Modulated Data-Transmission Systems,” *IEEE Transactions on Communications*, vol 22., no. 5, pp. 624-636, May 1974.

[24] G. Colavolpe, G. Ferrari, R. Raheli, “Reduced-State BCJR-Type Algorithms,” *IEEE Journ. On Selected Areas in Communications*, vol. 19, no. 5, pp. 848-859, May 2001.

[25] N. Al-Dhahir, “FIR Channel-Shortening Equalizers for MIMO ISI Channels,” IEEE Transactions on Communications, vol. 49, no. 2, pp. 214-218, Feb. 2001.

[26] R. Monzingo, T. Miller, *Introduction to Adaptive Arrays*, © 1980, John Wiley & Sons, Inc., NY, NY.

[27] G. L. Stuber, *Principles of Mobile Communications*, © 1996, Kluwer Press.

[28] J. Goldstein, I. Reed, and L. Scharf, “A Multistage Representation of the Wiener Filter Based on Orthogonal Projections,” *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2943-2959, November 1998.

[29] A.R. Calderbank, G. D, and A. Vardy, “Minimal Tail-Biting Trellises: The Golay Code and More,” *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1435-1455, July 1999.

[30] J. Anderson and S. Hladik, “Tailbiting MAP Decoders,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 297-302, Feb. 1998.

[31] R., C. Johnson, *Matrix Analysis*, © 1985, Camgridge University Press, NY, NY.

[32] F. Kschischang, V. Sorokine, “On the Trellis Structure of Block Codes,” *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1924-1934, November, 1995.

[33] M. J. Colella, J.N. Martin, F. Akyildiz, “The Halo Network,” *IEEE Communications Magazine*, vol. 38, no. 6, pp. 142-148, June, 2000.

[34] Proakis, J.G, *Digital Communications*, © 2001, McGraw Hill, Boston, MA.

[35] Bottomley, G.E, Molnar, K. J., Chennakeshu, S. C., “Interference Cancellation with an Array Processing Receiver”, *IEEE Transactions on Vehicular Technology*, vol. 48, no. 5, Sept.., 1999.

[36] Bottomley, G.E, Molnar, K. J., “Adaptive Channel Estimation for Multichannel MLSE Receivers”, *IEEE Communications Letters*, vol. 3, no. 2, Feb., 1999.

[37] G. E. Bottomley, H. Arslan, R. Ramésh, G. Brismark, “Coherent MAP Detection of DQPSK Bits,” *IEEE Comm. Letters*, vol. 4, no. 11, pp. 354-356, Nov. 2000.

[38] S. N. Diggavi, A. Paulraj, “Performance of multisensor adaptive MLSE in fading channels”, in *Proc. VTC ’97,* Pheonix, AZ, May 4-7, 1997, pp. 2148-2152.

[39] S. Kay, *Modern Spectral Estimation*, © 1988, Prentice Hall, Englewood Cliffs, NJ.

[40] A. H. Gray, J. D. Markel, "A Spectral Flatness Measure for Studying the Autocorrelation Method of Linear Prediction of Speech, " *IEEE Transactions on Acoustics and Speech Signal Processing*, vol. ASSP22, pp. 207-217, 1974.

[41] W. A. Gardner, "Computationally Efficient MIMO Demodulation", 11th Annual *Virginia Tech/MPRG Symposium on Wireless Personal Communications,* June 6-8, 2001, Blacksburg, VA.

[42] E. A. Robinson, *Multichannel Time Series Analysis*, © 1983, Goose Pond Press, Houston, TX.

[43] J. Thomas, E. Geraniotis, "Space-Time Iterative Receivers for Narrowband Multichannel Networks," *IEEE Transactions on Communications*, vol. 50, no. 7, July 2002.

[44] Reed, Hsia, "The Performance of Time-Dependent Adaptive Filters for Interference Rejection," *IEEE Trans. On Acoustics, Speech, and Signal Processing*, Vol. 38, No. 8, August 1990.

[45] J.D. Laster and J.H. Reed, "Interference Rejection in Digital Wireless Communications," *IEEE Signal Processing Magazine*, pp. 37-62, May 1997.

[46] W.A. Gardner, "Cyclic Wiener Filtering: Theory and Method," *IEEE Transactions on Communications*, Vol. 41, No. 1, January 1993, pp. 151-163.

[47] Therrin, Statistical Signal Processing, © 2000, Prentice Hall, Englewood Cliffs, NJ.

[48] J. Hagenauer, E. Offer, L. Papke, "Iterative Decoding of Binary Block and Convolutional Codes," *IEEE Transactions on Information Theory*, pp. 429-445, vol. 42, no. 2, March 1996.

[49] P. Thiennviboon, A. Ortega, K. M. Chugg , "Simplified grid message-passing algorithm with application to digital image halftoning," International Conference on Image Processing, 2001, 7-10 Oct 2001, p.p 1061 -1064, vol.2.

[50] N. Al-Dhahir, J. M. Cioffi, "Efficiently Computed Reduced-Parameter Input-Aided MMSE Equalizers for ML Detection: a unified approach," *IEEE Transactions on Information Theory*, vol. 42, no. 3, May 1996.

[51] X. Chen, K Chugg, "Reduced-State Soft-Input/Soft-Output Algorithms for Complexity Reduction in Iterative and Non-Iterative," *ICC 2000*, vol. 1, pp. 6-10, no. 1.

[52] S. Choi, D. Shim, "A Novel Adaptive Beamforming Algorithm for a Smart Antenna System in a CDMA Mobile Communication Environment", *IEEE Transactions on Vehicular Technology*, vol. 49, no. 5, Sept. 2000.

[53] W. Younis, N. Al-Dhahir, "Joint Prefiltering and MLSE Equalization of Space-Time-Coded Transmissions Over Frequency-Selective Channels," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 1, January 2002.

[54] C. Schlegel, S. Roy, P. Alexander, Z. Xiang, "Multiuser Projection Receivers," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 8, Oct. 1996.

[55] I. Medvedev, V. Tarokh, "A Channel-Shortening Multiuser Detector for DS-CDMA Systems," *VTC 2002,* pp. 1834-1838.

[56] S. Haykin, *Adaptive Filter Theory*, fourth edition, Prentice Hall, Upper Saddle River, NJ, © 2002.

[57] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, John Wiley, NY, © 1968.

[58] A. V. Fiacco, G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley, NY, © 1968.

[59] M. S. Bazaraa, H. D. Sherali, C. M. Shetty, *Nonlinear Programming, Theory and Algorithms*, second edition, John Wiley, NY, © 1993.

[60] J. Litva, T. Lo, *Digital Beamforming in Wireless Communications,* © 1996, Artech House, Boston, MA.

[61] W. L. Stutzman, G. A. Thiele, *Antenna Theory and Design*, © 1998, John Wiley and Sons, NY, NY.

# VITA

James Hicks was born in Fairfax, VA on September 13, 1974. He received his Bachelor of Science degree in Electrical Engineering from George Mason University, Fairfax, VA (GMU) in May 1997. While pursuing his undergraduate studies at GMU, he completed two Cooperative Education programs. From June 1994 to June 1996, James worked at the United States Naval Research Laboratory as a graphics programmer for real-time tactical warfare simulation. There he wrote a graphics library for the real-time culling of drawing objects in a virtual reality environment optimized for tactical warfare simulation. From June 1996 to June 1997, he worked at Stanford Telecommunications where he helped develop several satellite propagation model tools. From 1997-2002 James has been consulting part time for Information Systems Laboratories (ISL) in Vienna, VA. While at ISL, he has developed algorithms and system analysis for single satellite position location systems. James is currently pursuing a Ph.D. in electrical engineering and a M. S. in Mathematics at Virginia Tech, Blacksburg, VA as a Bradley Fellow. His research interests are digital signal processing and system modeling for wireless communications with a special interest in multi-user detection, array processing, and Markov modeling.