ANALYZING SPATIAL DIVERSITY IN DISTRIBUTED RADAR NETWORKS

BY

RANI DAHER

A THESIS SUBMITTED IN CONFORMITY WITH THE REQUIREMENTS FOR THE DEGREE OF MASTER OF APPLIED SCIENCE, GRADUATE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING IN THE UNIVERSITY OF TORONTO.

Copyright © 2008 by Rani Daher. All Rights Reserved.
Analyzing Spatial Diversity in Distributed Radar Networks

Master of Applied Science Thesis
Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto

by Rani Daher
May 2008

Abstract

We introduce the notion of diversity order as a performance measure for distributed radar systems. We define the diversity order of a radar network as the slope of the probability of detection ($P_D$) versus signal-to-noise ratio (SNR) curve evaluated at $P_D = 0.5$. We prove that the diversity order of both joint detection and optimal binary detection grows as $\sqrt{K}$, $K$ being the number of widely distributed sensors. This result shows that the communication bandwidth between the sensors and the fusion center does not affect the asymptotic growth in diversity order. We also prove that, amongst fixed decision rules, the OR rule leads to the best performance and its corresponding diversity order only grows as $(\log K)$. Space-Time Adaptive Processing (STAP) systems depend to a great extend on geometry, and we introduce the notion of a random radar network in order to study the effect of geometry on overall system performance. We first approximate the distribution of the signal-to-interference-plus-noise ratio (SINR) at each sensor by an exponential distribution, and we derive the corresponding moments for a specific system model. We extend this analysis to multistatic systems and prove that each sensor should individually be large enough to cancel the interference so that the system exploits the available spatial diversity.
To Hani and Ihsane;

In the memory of Basil Fleihan and all the Martyrs who offered their lives to Beirut.
Acknowledgements

I would first like to express profound and sincere gratitude to Prof. Ravi Adve for offering me the chance to work on such an exciting project. Your guidance and support were indispensable for the completion of this work, I am forever in your debt.

My journey in Toronto could not have been more pleasant thanks to my ‘family’ at UofT. First I would like to mention the two friends to whom I owe my presence here: Nahi for all he has done for me and for all the coffee breaks, and Emad: my housemate, lab-mate and prof-mate. Then there is Chady our designated chef, Sari the king of \LaTeX, Elie my gym buddy and Hratch my flat-mate. I will also never forget the long political and existential discussions with Imad. Without any doubt, you all have made my decision to leave Toronto harder, but as we always say: I did what I had to do.

Of course, Khaled, I always considered you as my big brother. Thank you for all the support and help you have offered. I guess your name will be present in so many theses for so many years to come.

Many friends abroad also shaped my stay in Toronto. I would like to specifically mention Sany and Tarek for their constant availability on gtalk, Haytham for the long skype calls, and Liza and Arine for the long-distance care and support.

Last but not least, I owe all that I achieve to my family in Lebanon: my father Hani and mother Ihsane. Thank you for all the love, support, care, and for bearing my sky-high expenses. I guess no words can ever describe how much I am indebted to you.

Rani H. Daher
May 2008
Contents

1 Introduction

1.1 Motivation ........................................... 2
1.2 System Model and STAP .................................. 4
  1.2.1 System Model ...................................... 4
  1.2.2 Space-Time Adaptive Processing ......................... 6
  1.2.3 Distributed Detection and the Neyman-Pearson Test ....... 8
1.3 Literature Survey ....................................... 12
1.4 Outline ................................................ 13

2 The Diversity Order of a Distributed Radar System ............. 15

2.1 A Notion of Diversity Order in Radar Systems ................. 15
2.2 Diversity Order of a Single Sensor .......................... 18
2.3 Diversity Order of Joint Processing .......................... 19
2.4 Diversity Order of Joint Detection ........................... 21
  2.4.1 Gaussian Approximation ............................... 22
  2.4.2 Diversity Order of Optimal Joint Detection .......... 24
  2.4.3 Discussion ........................................ 25
2.5 Diversity Order of Fully Distributed Systems .................. 26
  2.5.1 Optimal Binary Detection ............................. 26
  2.5.2 Diversity Order of the OR Rule ....................... 29
  2.5.3 Diversity Order of the AND Rule ..................... 31
  2.5.4 Diversity Order of the MAJ Rules .................... 32
  2.5.5 Discussion ........................................ 32
List of Figures

1.1 Generic system model. .................................................. 5
1.2 STAP datacube. .......................................................... 6

2.1 Diversity in communication systems. ................................. 16
2.2 Probability of detection for increasing $K$, OR detection. ....... 17
2.3 $P_D$ for joint detection. .............................................. 23
2.4 Accuracy of the Gaussian approximation. ......................... 24
2.5 Diversity order of joint detection .................................... 25
2.6 $P_D$ versus SNR curves of optimal joint detection (MRC) and OR Rule. ......................................................... 30
2.7 Diversity Order of joint detection and the OR Rule as a function of $K$. .......................................................... 30
2.8 Performance when $n$ is varied, $K=4$. ............................. 33
2.9 Performance of quantization methods for $M = 2$ bits, $P_F = 10^{-2}$. ......................................................... 37

3.1 Accuracy of the approximation of Eqn. (3.9) ....................... 42
3.2 Mean of the SINR and $N$ and $J$ increase. ......................... 48
3.3 Variance of the SINR when $N$ and $J$ increase. .................. 49
3.4 Metric $\Gamma$ introduced in Eqn. (3.30). .............................. 52
3.5 SINR defined as the deflection ratio. ................................ 53
3.6 Plots for the average $P_D$ with $J = 1, 3$. .......................... 55
Chapter 1

Introduction

‘Detection with distributed sensors’ (the term being borrowed from [1]) has been the subject for sporadic research during the last three decades. This topic is receiving renewed attention after the significant advancements in STAP for weak target detection in the last decade [2–4]. Now that we have a better understanding of both the gains from and limitations of adaptive processing, distributed sensing is currently being investigated in order to address some of these limitations. Such systems enjoy better performance and enhanced reliability when compared to centralized detection with a single sensor.

Distributed detection systems can be divided into two classes: the first class includes systems where sensors do not perform any significant local processing; they transmit the raw data to a fusion center which subsequently performs signal detection. The second class - the subject of most of this work - are systems where individual sensors process the received data and transmit some information to the fusion center that makes the final decision. This information can range from binary decisions (in which case, we denote the system as fully distributed) to exact likelihood ratios (which theoretically require infinite bandwidth). Unless stated otherwise (for example, in Section 2.2), we will denote the latter class by ‘multistatic systems’ or more generally, ‘distributed systems’, terms which will be used interchangeably.
1.1 Motivation

The motivation to undertake this thesis arose from the fact that the literature on distributed radars has lacked any formal measure to predict the performance of these systems as a function of system parameters, such as the number of sensors, the antenna array size, the available bandwidth and geometry. Most of the available work has focused on empirical studies. In this regard our main objective is not the introduction of novel signal processing or detection schemes, but to provide pertinent analysis and design insights relating to existing detection schemes, in order to understand the tradeoffs entailed by distributed detection. We would also like to understand how does the availability of independent observations resulting from detection with multiple sensors that are far apart (in other words, spatial diversity) affect the overall system performance. Our first contribution in this work is the introduction of the notion of a diversity order in radar networks. Our definition not only enables the comparison of various distributed sensing schemes, but also provides insight into the design of these systems.

The notion of diversity played a major role in the development of Multiple-Input-Multiple-Output (MIMO) wireless communications where the diversity order is defined as the slope, in a log-log plot, of either the bit error rate or outage probability versus SNR at asymptotically high SNR [5]. It is crucial to note that even though diversity was defined as a high-SNR concept, wireless systems usually achieve this asymptotic behavior at practical SNR levels.

In the radar context, the analog of outage probability would be the miss probability for a fixed false alarm rate, which is the probability of declaring a target present when there is no target. In extending this notion of diversity to distributed detection we are faced by the fact that, since radar systems invariably deal with lower levels of SNR, an asymptotic definition may not to be as useful. For this specific reason, we were motivated to investigate useful definitions that truly reflect the behavior of the system.

In this work we define the diversity order as the slope of the probability of detection \((P_D)\)\(^1\) versus SNR curve at \(P_D = 0.5\). This \(P_D\) is derived for a fixed false alarm rate.

\(^1\)We will use uppercase subscripts \((P_D\) and \(P_F\)\) to denote the global probabilities of detection and false alarm respectively, and we will use the lowercase subscripts \((P_d^{(k)}\) and \(P_f^{(k)}\)\) for the individual probabilities of detection and false alarm at sensor \(k\).
1.1. MOTIVATION

The diversity order serves essentially as a proxy for the number of degrees of freedom in the radar system. This provides a consistent and useful definition and, as we will show, captures nicely the impact of varying system configurations and processing schemes. Furthermore, the SNR levels at which this $P_D$ is achieved are quite practical.

Note that, in the wireless communications context, diversity captures the number of independent signal paths and the diversity order is always an integer. However, as we will see, our definition of diversity order in a radar system captures the interaction of $P_D, P_F$ and the processing scheme and is rarely an integer. Each definition is useful in its own domain.

We first analyze systems that are limited by receiver noise. We prove that the diversity orders of both joint and optimal fully distributed detection grow as $\sqrt{K}$, $K$ being the number of sensors. This fact has two major implications. First, in noise-limited systems, due to coherent processing, the array gain outweighs spatial diversity gain, hence larger antenna arrays achieve better performance. The second major implication is that increasing the information flow to the fusion center does not increase diversity order, a result consistent with empirical results available in the literature.

Noise-limited systems provide some insight into the performance of distributed systems. Nonetheless, most practical systems also include jamming and clutter, the impact of which depends significantly on the geometry of the system. To our knowledge, all the work on geometry in distributed sensing assumes that sensor positions are fixed and known. However, radar engineers clearly do not influence where the target is or its velocity and consequently, the geometry relative to the target is random, and should be dealt with as such. In order to tackle this problem, we introduced the notion of a random radar network: the second concept introduced in this work.

We first study random unistatic radar networks and derive the characteristics of the output SNR at the individual sensors\(^2\). We extend this work to multistatic radars, and use our definition of diversity order in order to characterize a design tradeoff in distributed radar networks: each sensor should be large enough to null the effect of interference in order to exploit the spatial diversity available in the system. On the other hand, for

\(^2\)A random unistatic system is one where a single sensor, chosen randomly out of many available sensors, is used for target detection.
reasons of power and mobility, it is preferable to have smaller sensors. We provide the theory and the results of simulations to support this statement.

The analysis that will be presented throughout this work requires some knowledge in STAP and distributed detection; we dedicate the next section to providing this background.

\section*{1.2 System Model and STAP}

In this section we will present the system model and the necessary background relating to the topics that we discuss throughout this work. We briefly introduce STAP in order to grasp the importance of geometry in modern radar systems. We then turn our attention to distributed detection under a constant false alarm rate (CFAR) constraint and develop the Neyman-Pearson (NP) test for distributed STAP systems.

\subsection*{1.2.1 System Model}

The system comprises \( K \) distributed sensors detecting the presence of a target in a certain region in space. Each sensor possesses \( N \) collocated antennas. A uniform inter-element spacing of \( d = \lambda/2 \) is assumed throughout this work; hence the electrical distance is \( 2\pi d/\lambda = \pi \).

We assume throughout this work that the sensors are separated by a distance large enough to treat the observations at the individual sensors as statistically independent given the hypothesis \cite{6}. The \( k \)-th sensor receives a data vector of the form:

\begin{equation}
    z_k = \begin{cases}
        \alpha_k s_k + n_k, & \text{if target is present} \\
        n_k, & \text{if target is absent}
    \end{cases},
\end{equation}

where \( s_k \) is the normalized signature of the target, \( \alpha_k \) is the complex-valued amplitude, and \( n_k \) is the additive interference and noise vector. The target is modeled as a Swerling type-II and consequently, \( \{\alpha_k\}_{k=1}^{K} \) are independent and identically distributed (i.i.d.) drawn from a zero-mean complex Gaussian random variable. We do not model the source excitation that leads to the received signal.
1.2. **SYSTEM MODEL AND STAP**

We assume that the receiver noise statistics are known, and the corresponding Receiver Operating Characteristics (ROC)\(^3\) can be derived accordingly. We also assume that the input SNR, defined as the ratio of the signal received power (\(\mathcal{E}\{|\alpha|^2\}\)) to the receiver noise power (\(\sigma^2\)), is equal at all the sensors; we will denote this to be the ‘symmetry’ assumption. This assumption is prevalent in the literature since it allows a workable ground for any analytical attempt.

Our main concern in this work is to study the behavior of large networks. We will therefore limit our analysis to receiver noise and jammers, both of which are assumed statistically white in time. In other words, we will only deal with the spatial signature of the target and jammers. All the analysis pertaining to jamming can be easily extended to include clutter.

Finally, we introduce our model of a *random* distributed network. STAP detection, which will be discussed more thoroughly in Section 1.2.2, significantly depends on the geometry of the system. As its name suggests, a random radar network is a system where the relative positions and velocities of the sensors, target, jammers and clutter are all random. For reasons of practicality, we introduce a *generic* model (Figure 1.1) which assumes that the target is at the center of the area monitored by the sensors, which in turn are *randomly* (not necessarily uniformly) placed on a circle centered at the target; the jammers are randomly distributed inside this circle. We note that the generic model is a special case of a random network where the restriction to a circle allows for the

---

\(^3\)The ROC is a curve plotting \(P_D\) versus \(P_F\)
1.2. SYSTEM MODEL AND STAP

A random network model can be interpreted from two different perspectives. It can be first viewed as a set of $K$ sensors randomly dispersed over some area and the resulting distribution is being modeled. The second perspective is the one of a very large network out of which only $K$ sensors are randomly chosen for detection. These two perspectives are essentially equivalent and will lead to the same analysis.

1.2.2 Space-Time Adaptive Processing

We will now present a brief overview of space-time adaptive processing. The interested reader is referred to [2] for a more comprehensive discussion of STAP. A STAP detector is an array of $N$ antenna elements. The array transmits a pulse that reflects off the target (if present) in addition to other reflectors, such as the surface of the earth. The reflected signal is sampled $L$ times, each sample corresponds to a range. This process is repeated $M$ times within a coherent processing interval. This generates a datacube of size $M \times N \times L$ (Figure 1.2), a product that can be as large as several hundreds of thousands. The $N$ array elements generate the spatial signature vector $a(\theta_t)$ which corresponds to

![Figure 1.2: STAP datacube.](image)
the direction \( \theta_t \) of the target. Similarly, the \( M \) pulses generate the temporal signature vector \( \mathbf{b}(\varpi) \) corresponding to the normalized Doppler frequency \( \varpi \) of the target (hence its relative velocity). The reader is referred to [2] for a thorough description of the system model. We note that:

\[
\varpi = \frac{f_t}{f_r},
\]

(1.2)

where \( f_r \) is the Pulse Repetition Frequency (PRF), which is the inverse of the pulse repetition interval (PRI), or period of each of the \( M \) pulses, and \( f_t = \frac{2v_t}{\lambda_0} \) where \( \lambda_0 \) is the radar operating wavelength and \( v_t \) is the target radial velocity.

The space and time signatures are respectively of the form:

\[
\mathbf{a}(\theta_t) = [1, e^{i\pi \cos \theta_t}, e^{i2\pi \cos \theta_t}, \ldots, e^{i(N-1)\pi \cos \theta_t}]^T,
\]

(1.3)

\[
\mathbf{b}(\varpi) = [1, e^{i2\pi \varpi}, e^{i4\pi \varpi}, \ldots, e^{i(M-1)2\pi \varpi}]^T.
\]

(1.4)

The space-time steering vector \( \mathbf{s}(\theta_t, \varpi) \) of the target is the Kronecker product of \( \mathbf{a}(\theta_t) \) and \( \mathbf{b}(\varpi) \).

In a radar setting, there are three main contributors to the noise environment. First, the receiver generates additive white Gaussian noise. The second source is jamming, which appears as a localized source in the spatial dimension but as white in the time dimension. Finally, clutter is the third component and is considered to be the most challenging, especially for an airborne radar platform. Clutter is generated from echoes without tactical significance (as opposed to jamming). The surface of the earth is the main source of clutter, and the severity of the clutter is dependent on the reflectivity of the area illuminated. The three components are assumed mutually independent, thus the total \( (NM \times NM) \) noise covariance matrix \( \mathbf{R}_n \) is the sum of the covariance matrices of the individual components. Formally:

\[
\mathbf{R}_n = \mathbf{R}_\gamma + \mathbf{R}_j + \mathbf{R}_c,
\]

(1.5)

where \( \mathbf{R}_\gamma, \mathbf{R}_j \) and \( \mathbf{R}_c \) are the receiver noise, jamming and clutter covariance matrices respectively. In an airborne radar, the rank of the clutter covariance matrix is significantly
smaller than $NM$ [2].

In STAP the $NM$ data samples corresponding to each range cell are linearly combined using a vector $w$ of adaptive weights. The optimal weight vector that maximizes the output Signal-to-Noise-plus-Interference-Ratio (SINR) is [2]:

$$w = R_n^{-1}s.$$  \(1.6\)

The output statistic is then:

$$\eta = |w^H x|^2,$$  \(1.7\)

where $x$ is the length-$NM$ received data vector for a specific range cell. This statistic is then compared to a threshold $t_0$ in order to determine the presence (hypothesis $H_1$) or absence (hypothesis $H_0$) of a target in this cell. Formally, the STAP detector performs a test of the form:

$$f(\eta|H_1) \geq H_1 H_0 tf(\eta|H_0),$$  \(1.8\)

where $f(\eta|H_i)$ denotes the probability density function of $(\eta|H_i)$ evaluated at output statistic $\eta$, and $t$ is a threshold that determines the false alarm rate.

The STAP detector adapts to the interference environment in order to cancel the effect of jammers and clutter. Clearly, the performance greatly depends on the geometry of the system. For example, if the target is close to a strong jammer, the target will be nulled. Similarly, if the target moves in a direction parallel to the array, it appears stationary, like the ground clutter, and the sensor will not be able to detect the target. It is worth emphasizing that system designers have little control over this geometry.

The performance of a distributed system also relies on the signal processing scheme used to combine the information from different sensors. We will elaborate further on this topic in the following section.

### 1.2.3 Distributed Detection and the Neyman-Pearson Test

In distributed detection, each sensor $k$ transmits a decision $u_k$ to a fusion center, which makes the final decision $u_0$ indicating the presence or absence of a target in the region of space monitored by the sensors. We will adopt the convention that 1 symbolizes $H_1$,
0 symbolizes $H_0$, and $u$ is the length-$K$ vector of the decisions at the sensors.

In radar applications, we are particularly interested in preserving CFAR. In [7] the authors prove that the optimal detection rule under CFAR is an Neyman-Pearson (NP) test at both the fusion center and the local sensors. At the fusion center, the NP test is of the form:

$$u_0 = \begin{cases} 
1, & \text{if } \Pr(u|H_1) \geq t_0 \Pr(u|H_0) \\
0, & \text{if } \Pr(u|H_1) < t_0 \Pr(u|H_0) 
\end{cases}$$

where $t_0$ is a global threshold to be determined according to the required ($P_F$). By the monotonicity of the optimum fusion rule established in [8], and knowing that the NP test is the most powerful test [9], the local tests at the sensors are also NP tests, and in fully distributed systems, the tests are of the form:

$$u_k = \begin{cases} 
1, & \text{if } \Pr(y_k|H_1) \geq t_k \Pr(y_k|H_0) \\
0, & \text{if } \Pr(y_k|H_1) < t_k \Pr(y_k|H_0) 
\end{cases}$$

where $y_k$ and $u_k$ are the output of the STAP processor and the corresponding local decision at sensor $k$ respectively, and $\{t_k\}_{k=1}^K$ are the local thresholds to be determined in order to maintain a desired probability of false alarm at each sensor.

The problem of choosing the thresholds is reduced to maximizing the global probability of detection $P_D = \Pr(u_0 = 1|H_1)$ under the constraint that the global probability of false alarm $P_F = \Pr(u_0 = 1|H_0)$ is held constant. Finding the optimal thresholds is a non-convex problem in general and no global optima are guaranteed by the optimization process. We will not dwell on the optimization problem, and the reader is referred to [10] for more details. We also refer the reader to [9] and [11] for a comprehensive survey on distributed detection.

We will now derive the NP test for distributed STAP systems. The following derivation follows the same lines as in [6]. Given the Swerling type-II target model, and under both hypotheses, the received data vector, $z_k$, is complex Gaussian. The derivation below uses an arbitrary noise covariance matrix denoted by $R_n$ and the covariance matrix of
the signal is denoted by:

\[ S_k = A^2 s_k s_k^H, \]  

(1.11)

where \( A^2 = \mathcal{E}\{|\alpha|^2\} \) is the average signal power assumed equal at all sensors.

Under the null hypothesis \( H_0 \), and due to the independence assumption,

\[ f\{z_1, z_2, \ldots, z_K|H_0\} = \prod_{k=1}^K \frac{1}{\pi^N|R_n|} e^{-z_k^H R_n^{-1} z_k}. \]  

(1.12)

Similarly, under the target-present hypothesis, \( H_1 \),

\[ f\{z_1, \ldots, z_K|H_1\} = \prod_{k=1}^K \frac{1}{\pi^N|R_n + S_k|} e^{-z_k^H (R_n + S_k)^{-1} z_k}. \]  

(1.13)

The likelihood ratio corresponding to the NP test is defined as:

\[ \Lambda(z_1, \ldots, z_K) = \frac{f\{z_1, \ldots, z_K|H_1\}}{f\{z_1, \ldots, z_K|H_0\}}. \]  

(1.14)

We show in Appendix A.1 that the NP test statistic is of the form [6]:

\[ \zeta = \sum_{k=1}^K \frac{A^2 |s_k^H R_n^{-1} z_k|^2}{1 + A^2 s_k^H R_n^{-1} s_k}. \]  

(1.15)

Note that the numerator for each term of the summation is exponentially distributed under both hypotheses, and that the denominator is independent of the received vector.

We also note that the numerator is the amplitude of the output of the adaptive processor using the optimal fully adaptive weight vector \( w = R_n^{-1}s \) described in the previous section. Each term is proportional to the received signal power, which explains why this detection scheme is often denoted as Maximum-Ratio Combining (MRC) [12].

The NP test implies that each sensor should use the most powerful test, which is the NP test itself. Consequently, each sensor individually performs a test of the form of
1.2. SYSTEM MODEL AND STAP

Eqn. (1.15), which in the case of sensor $k$, reduces to:

$$
\zeta_k = \frac{A^2|s_k^H R_n^{-1} z_k|^2}{1 + A^2 s_k^H R_n^{-1}s_k} \geq H_0 \ T_h^{(k)},
$$

where $T_h^{(k)}$ is a threshold to be determined in order to maintain the desired probability of false alarm at the sensor.

For convenience, we drop the subscripts. Under the null hypothesis, the received vector $z$ is the zero-mean complex Gaussian noise vector, and the statistic is exponentially distributed with mean:

$$
\lambda_0 = \mathcal{E}\{\zeta|H_0\} = \frac{A^2 s^H R_n^{-1}s}{1 + A^2 s_k^H R_n^{-1}s_k}.
$$

Consequently, the probability of false alarm at sensor $k$ becomes:

$$
P_f^{(k)} = \Pr(\zeta > T_h^{(k)}|H_0) = e^{-T_h^{(k)}/\lambda_0}.
$$

Similarly, under the target-present hypothesis, the received vector is of the form:

$$
z = \alpha s + n,
$$

and the statistic for the Swerling type-II model is also exponentially distributed with mean:

$$
\lambda_1 = \mathcal{E}\{\zeta|H_1\} = \frac{A^2 s^H R_n^{-1}s + A^4 |s^H R_n^{-1}s|^2}{1 + A^2 s^H R_n^{-1}s},
$$

and the probability of detection at each sensor $k$ is:

$$
P_d^{(k)} = \Pr(\zeta > T_h^{(k)}|H_1) = e^{-T_h^{(k)}/\lambda_1}.
$$
1.3 Literature Survey

In this section we survey the available literature relating to distributed STAP detection, and we also refer the interested reader to references that are considered classics in the area.

In their landmark work [1], Tenney et al. introduced fully distributed detection by minimizing the Bayes cost function, a scheme that is not optimal under CFAR. When compared to their monostatic counterparts, such systems are said to be more reliable, survivable and they significantly reduce the communication bandwidth between the sensors and the fusion center. Distributed systems also enable the use of different sources of information. For example, the decisions of sonars can be combined with those of sensors from different platforms in order to generate the final decision regarding the presence or absence of a target. In this thesis, we provide a mathematical justification of these claims. In radar applications, we are usually interested in preserving a constant false alarm rate. Farina et al. analyzed detection under this criterion, and proved that detection is optimal when sensors transmit likelihood ratios that are proportional to the received power (hence the MRC nomenclature) [12]. They also use the fully distributed OR fusion rule wherein a target is declared present if any of the $K$ sensors detects it, without any proof regarding its optimality.

STAP resulted in significant advancements in combating jamming and clutter, and we refer the reader to [2] for a comprehensive survey on this topic. Distributed STAP was introduced often under the label of MIMO radar [13]. Such systems result in an extremely narrow main beam and lower grating lobes [3,14].

Geometry became a major concern with the introduction of STAP, which greatly depends on the relative positions and velocities of the sensors, target, jammers and clutter sources. In work that has been submitted after our proposal, Goodman et al. presented an empirical evaluation of the effect of geometry in [6]. In more recent work, they prove that the order of the decay of the miss probability when SNR approaches infinity is on the order of the total number of sensors involved in the detection procedure [15]. Pezeshki et al. also introduced an asymptotic definition for diversity in radar networks [16]. For reasons that will become apparent in Chapter 3, we believe that our definition constitutes
Distributed detection is optimal when the sensors transmit their exact, as opposed to quantized, likelihood ratios. However, optimal detection is not feasible in practice due to bandwidth limitations and radar designers are bound to using sub-optimal schemes. Most prevalent of these schemes is fully distributed detection where the sensors transmit a binary (hard) decision to the fusion center, which would again use the NP test to generate the system-wide decision. This NP test will be one of the “n out of K” fusion rules. Varying the fusion rule according to the sensors performance leads to significant gains in performance; however, the fusion rule is often fixed due to its simplicity in implementation. The fully distributed OR rule is prevalent in the literature, as is the case for the work by Farina et al. [12] and also in the work of Goodman [6], but the authors do not provide any reasoning backing the use of this rule. Again, we believe that the work in this thesis provides a mathematical justification for these claims.

Multi-bit detection occupies a middle ground between joint detection and fully distributed detection, wherein each sensor transmits an $M$-bit quantization of its likelihood ratio. In [17] and [18], the authors design optimum quantizers for distributed signal detection. Lee et al. applied quantization to the distributed detection problem under the CFAR criterion [19], and proved empirically that near-optimal performance can be achieved with a limited number of bits transmitted by each sensor. This topic will be briefly discussed in Section 2.6.

Finally, we assume that the fusion center receives the data from the local sensors without error. The reader is referred to [20] and the references therein for a summary on channel-aware distributed detection.

1.4 Outline

This document is organized as follows: In Chapter 2 we introduce the notion of diversity in radar systems and we analyze the diversity growth for joint detection and fully distributed detection for noise-limited systems. We also briefly describe two new quantization techniques for distributed detection. In Chapter 3 we introduce random radar
networks and we analyze the behavior of unistatic and multistatic systems, which enabled us to devise a design tradeoff between performance and diversity on the one hand, and power consumption and mobility on the other. We seal this document with some conclusions and projections for future work in Chapter 4.
Chapter 2

The Diversity Order of a Distributed Radar System

We now turn our attention to a central contribution of this work - the extension to radar systems of a notion that has been of critical importance in wireless communications: the diversity order. After we introduce our definition, we analyze systems that are limited by the internal noise generated at the receiver. We will prove that joint detection and ‘optimal’ binary detection lead to the same asymptotic growth in diversity order. We also analyze fully distributed systems where the fusion rule is assume fixed.

2.1 A Notion of Diversity Order in Radar Systems

In the context of wireless communications, diversity order is defined as the slope in a log-log scale of the probability of error ($P_e$) versus SNR curve at asymptotically high SNR [5]. The diversity order measures the number of independent paths over which the signal is received. This measure of performance has been particularly useful as a design and analysis tool because wireless systems generally achieve this asymptotic behavior at reasonable SNR levels.

Figure 2.1 illustrates the notion of diversity in communication systems. For a system with 1 antenna, when the SNR increases by 10 dB, the probability of error is reduced by a factor of 10, and the slope in log-log scale is 1. For a system with $M$ antennas, when
2.1. A NOTION OF DIVERSITY ORDER IN RADAR SYSTEMS

The SNR increases by 10 dB, the probability of error is reduced by $10^M$, thus the slope on a log-log scale is $M$.

A similar definition for diversity in distributed radar networks faces three major hindrances. First, radar systems invariably deal with SNR levels that are much lower than acceptable SNR ranges in wireless communications, hence an asymptotic definition might not be convenient. In addition, while $P_e$ ranges of $10^{-3}$ to $10^{-6}$ are usually required in wireless systems in order to guarantee a certain quality of service, such ranges (for the probability of miss) are rarely of interest in the radar context. Finally, and proceeding with the latter idea, radar engineers are mainly interested in the ‘rising’ portion of the $P_D$ curve as the steepness of this portion shows how fast the radar system switches from the low $P_D$ to the high $P_D$ regime. Therefore, there is a need for a revised definition for diversity relevant to distributed radar networks. $P_D$ being a function of SNR ($\gamma$), We choose the following definition:

**Definition 1** The **diversity order** of a distributed radar network is the derivative in a linear scale $\left( \frac{dP_D(\gamma)}{d\gamma} \right)$ evaluated at $P_D = 0.5$. 
Figure 2.2: Probability of detection for increasing $K$, OR detection.

Figure 2.2 illustrates our notion of diversity. It presents the $P_D$ versus SNR curve for a distributed network with $P_F = 10^{-4}$ (this value will be assumed throughout this thesis unless otherwise specified). The figure shows that the $P_D$ curve is steeper when the number of sensors increases. Note that the scale on the $x$-axis is in dB, and this plot is strictly for illustration purposes.

This figure underlines the relevance of our definition. First, the required probability of detection ($P_D = 0.5$) should be achievable (in fact expected) by most radar systems and for reasonable SNR levels. In addition, since we are mainly interested in the ‘rising’ portion of the $P_D$ curve, the slope at $P_D = 0.5$ is highly likely to best estimate an average slope over this portion of the curve. Thus, our definition is a simple, consistent and meaningful measure reflecting the effective performance of the system. Consequently, we will adopt it throughout this work as a performance measure for radar networks.

It is important to emphasize that our definition of diversity order does not completely characterize the radar detection system. As in wireless communications, our definition of diversity order does not address where, as a function of SNR, is the bend in the $P_D$
2.2. DIVERSITY ORDER OF A SINGLE SENSOR

curve. Various processing schemes might achieve the same diversity order but still have very different performance. We will draw attention to this fact when comparing ‘optimal’ binary detection with joint detection.

2.2 Diversity Order of a Single Sensor

When the system is noise-limited, the noise is assumed to follow a complex Gaussian distribution whose covariance matrix is of the form:

\[ \mathbf{R}_n = \sigma^2 \mathbf{I}_N \Rightarrow \mathbf{n}_k \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_N), \]  

(2.1)

hence:

\[ s^H \mathbf{R}_n^{-1} s = \frac{s^H s}{\sigma^2} = \frac{N}{\sigma^2}. \]  

(2.2)

In what follows, we assume that: \( A^2 = \mathbb{E}\{|\alpha|^2\} = 1 \). This is a simplifying assumption and does not entail any loss of generality. The input SNR is then defined as:

\[ \gamma = \frac{\mathbb{E}\{|\alpha|^2\}}{\sigma^2} = \frac{1}{\sigma^2}, \]  

(2.3)

and the output SNR is the array gain multiplied by the input SNR as follows:

\[ \gamma_o = \mathbb{E}\{|\alpha|^2\} s^H \mathbf{R}_n^{-1} s = N \gamma. \]  

(2.4)

Replacing Eqn. (2.4) in Eqns. (1.17) and (1.20) we get:

\[ \lambda_0 = \frac{N \gamma}{1+N \gamma}, \]  

(2.5)

\[ \lambda_1 = N \gamma. \]  

(2.6)

We now derive the diversity order of a system consisting of a single sensor with \( NK \) antennas, and we will use this result as a benchmark to evaluate the performance loss due to distributed detection.
2.3. DIVERSITY ORDER OF JOINT PROCESSING

Theorem 2.2.1 For a noise-limited system consisting of 1 sensor with NK collocated antennas, the diversity order is proportional to NK.

Proof: Since $P_F$ is a constant, applying Eqn. (1.21) at $P_D = 0.5$ we get:

\[
P_D = e^{\ln P_F} = 0.5,
\]

\[
1 + NK\gamma = -\frac{\ln P_F}{\ln 2}.
\]

Differentiating $P_D$ with respect to $\gamma$ yields:

\[
\frac{dP_D}{d\gamma} = -NK.e^{\ln P_F} \frac{\ln P_F}{(1 + NK\gamma)^2}.
\]

We finally combine Eqns. (2.8) and (2.9) to get:

\[
\frac{dP_D}{d\gamma}(P_D = 0.5) = -NK\ln^2 2
\]

\[
= \frac{-NK\ln^2 2}{2\ln P_F},
\]

i.e., the diversity order is proportional to $NK$. 


2.3 Diversity Order of Joint Processing

When the sensors transmit the raw data to the fusion center, where the space-time adaptive processing is later performed, we denote this scheme as Joint Processing. This is different than Distributed Detection (which is the subject of most of this work), where processing is locally performed at the sensors and the fusion center simply combines the received likelihood ratios.

The received vectors at the sensors follow the model of Eqn. (1.1). We follow the same notation as before, in which $s_i, z_i$ and $n_i$ denote the target steering vector, the received data vector and the additive noise vector at sensor $i$ respectively. The fusion center stacks the received vector into a length-$NK$ array of the form:

\[
z = \begin{cases}
S\alpha + n, & \text{if target is present} \\
n_i, & \text{if target is absent}
\end{cases}
\]
2.3. DIVERSITY ORDER OF JOINT PROCESSING

which can be expressed more elaborately as:

\[
\mathbf{z} = \begin{pmatrix}
\mathbf{z}_1 \\
\mathbf{z}_2 \\
\vdots \\
\mathbf{z}_k
\end{pmatrix}_{NK \times 1} = \begin{pmatrix}
\mathbf{s}_1 & 0 & \ldots & 0 \\
0 & \mathbf{s}_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \mathbf{s}_k
\end{pmatrix}_{NK \times K} \begin{pmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_k
\end{pmatrix}_{K \times 1} + \begin{pmatrix}
\mathbf{n}_1 \\
\mathbf{n}_2 \\
\vdots \\
\mathbf{n}_k
\end{pmatrix}_{NK \times 1}.
\] (2.12)

By the independent observations assumption, the noise covariance matrix is block-diagonal of the form:

\[
\mathbf{R}_n = \begin{pmatrix}
\mathbf{R}_{n,1} & 0 & \ldots & 0 \\
0 & \mathbf{R}_{n,2} & \ldots & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \mathbf{R}_{n,k}
\end{pmatrix}_{NK \times NK}.
\] (2.13)

where \(\mathbf{R}_{n,i}\) is the noise covariance matrix at sensor \(i\). We also note that under the Swerling type-II target model and the symmetry assumption, the \(\alpha_i\)'s are independent zero-mean complex Gaussian random variables of equal power, and thus the signal covariance matrix is:

\[
\Psi_{NK \times NK} = \mathbb{E}\{\mathbf{S} \mathbf{\alpha} \mathbf{\alpha}^H \mathbf{S}^H\} = \mathbf{S} \mathbb{E}\{\mathbf{\alpha} \mathbf{\alpha}^H\} \mathbf{S}^H = A^2 \mathbf{S} \mathbf{S}^H.
\] (2.14)

The received vector is complex Gaussian under both hypotheses, and we prove in Appendix A.2 that the NP test statistic for joint processing follows the same expression as Eqn (1.15). In fact, when the observations are independent, the information from the various sensors decouple, and joint processing leads to the same performance as joint detection, despite the fact that it requires a significantly larger communication bandwidth. For this reason, in what follows we focus on distributed detection, and we first analyze optimal joint detection before we study the more practical fully distributed systems.
2.4 Diversity Order of Joint Detection

Distributed detection is optimal when the exact (i.e., non-quantized) likelihood ratios at the sensors are jointly processed at the fusion center. This scheme assumes that the bandwidth available between each of the sensors and the fusion center is infinite, and hence, the following results are not practically achievable and only serve as a benchmark for sub-optimal schemes.

The NP test statistic for joint detection is described by Eqn. (1.15). Note that this approach is similar to Maximum Ratio Combining (MRC) in wireless communications because each sensor contributes to the statistic proportionately to its received power [12]. Under the Swerling type-II model and with the symmetry assumption, the statistic is a summation of $K$ i.i.d. exponential random variables. The sum follows a Gamma distribution of the form [21]:

$$f(x; K, \theta) = x^{K-1} \frac{e^{-x/\theta}}{\theta^K \Gamma(K)},$$

(2.15)

where $\theta$ is the common mean of the $K$ random variables and $\Gamma(K) = (K-1)!$ for integer $K$. Linear sums of exponential random variables is a well-studied problem (e.g., [21]), and it is straightforward to extend the current analysis to the non-symmetric case where each random variable possesses a different mean.

The probability of false alarm is the probability, given $H_0$, that the statistic surpasses a threshold $T_h$ to be determined. This is the complement of the cumulative distribution function (CDF) of $\zeta$ given the null hypothesis, and is given by the upper incomplete Gamma function $\Gamma(K, x)$ defined as:

$$P_F = \Pr(\zeta > T_h|H_0) = \Gamma(K, T_h/\lambda_0),$$

$$= \frac{1}{\Gamma(K)} \int_{T_h/\lambda_0}^{\infty} x^{K-1} e^{-x} dx.$$  

(2.16)
Similarly, the probability of detection is defined as follows:

\[
P_D = \Pr(\zeta > T_h|H_1) = \Gamma(K, T_h/\lambda_1),
\]
\[
= \frac{1}{\Gamma(K)} \int_{T_h/\lambda_1}^{\infty} x^{K-1}e^{-x}dx.
\]  

(2.17)

A closed form expression for the probability of false alarm, \(P_F\), does not exist, and we are reduced to numerical solutions. Given the desired \(P_F\), we perform a line search in order to find the corresponding threshold, \(T_h\) using Eqn. (2.16). We then replace the obtained value of \(T_h\) in Eqn. (2.17) to determine the probability of detection at a specific SNR. In turn, a numerical evaluation of the derivative at \(P_D = 0.5\) provides the diversity order.

The numerical approach described above is, however, hard to analyze. Figure 2.3 shows the \(P_D\) curve for joint detection when the number of sensors increases. It is clear that the curves are steeper when the number of sensor increases, however this growth is sub-linear. The following sections will be dedicated to quantifying this asymptotic growth in diversity order as a function of the number of sensors and the antenna array size.

2.4.1 Gaussian Approximation

Finding a closed-form solution for the slope and diversity order following the analysis above appears intractable. However, by the central limit theorem, when the number of sensors grows large, the test statistic \(\zeta\) asymptotically follows a Gaussian distribution. We use this approximation to analyze the diversity order of joint detection.

Under the null hypothesis and the symmetry assumption, each sensor will contribute an exponentially distributed term with mean \(\mu_k = \lambda_0\) and variance \(\sigma_k^2 = \lambda_0^2\). Hence, by the independence assumption, the statistic \(\zeta\) will be Gaussian with mean \(\mu_\zeta = K\lambda_0\) and standard deviation \(\sigma_\zeta = \sqrt{K}\lambda_0\). The probabilities of false alarm and detection are
respectively defined as follows:

\[
P_F = \Pr(\zeta > T_h | H_0) = Q\left(\frac{T_h - K\lambda_0}{\sqrt{K}\lambda_0}\right), \tag{2.18}\]

\[
P_D = \Pr(\zeta > T_h | H_1) = Q\left(\frac{T_h - K\lambda_1}{\sqrt{K}\lambda_1}\right), \tag{2.19}\]

where \(Q(x)\) is the upper tail of the standard normal distribution and is defined by:

\[
Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \tag{2.20}\]

Given the desired probability of false alarm \(P_F\), we invert Eqn. (2.18) to get the threshold:

\[
T_h = Q^{-1}(P_F)\sqrt{K}\lambda_0 + K\lambda_0 = \lambda_0\sqrt{K}\sqrt{K + Q^{-1}(P_F)}. \tag{2.21}\]

We combine Eqn. (2.19) and Eqn. (2.21) to get the probability of detection at a certain SNR. Figure 2.4 reflects the accuracy of this approximation. In the figure, the
2.4. DIVERSITY ORDER OF JOINT DETECTION

Figure 2.4: Accuracy of the Gaussian approximation.

lines marked ‘Q’ correspond to the Gaussian approximation while the lines marked ‘Γ’ correspond to the numerical line search approach. The figure shows that, as expected, the approximation become increasingly accurate as the number of sensors increases. Unfortunately, Fig. 2.4 also shows that a large number of sensors is required to achieve a reasonable accuracy.

2.4.2 Diversity Order of Optimal Joint Detection

Proceeding with the analysis presented in the previous section, we derive the diversity order for joint detection.

**Theorem 2.4.1** For a noise-limited system using optimal joint detection, and for $K$ large, the slope at $P_D = 0.5$ increases as $N\sqrt{K}$.

**Proof:** See Appendix A.3.

Figure 2.5 shows the diversity order for joint detection. We first note that the diversity order was calculated numerically, using Eqns. (2.16) and (2.17), and not from the
DIVERSITY ORDER OF JOINT DETECTION

2.4. Diversity Order of Joint Detection

Gaussian approximation. To illustrate the fact that the diversity order grows as $\sqrt{K}$, the figure also shows a curve of the form $\alpha\sqrt{K}$ with $\alpha$ chosen for a close fit. Figure 2.5 clearly shows that the diversity order for joint detection follows a square root behavior.

2.4.3 Discussion

Theorem 2.4.1 captures, in a single expression, the gains in performance due to two system parameters, the number of sensors ($K$) and the number of degrees of freedom per sensor (here the number of elements $N$ at each sensor). If we have a single array with $NK$ antennas, we have proved that the diversity order is on the order of $NK$. On the other hand, as proven in this section, the fact that joint detection for noise-limited systems achieves a diversity order proportional to $N\sqrt{K}$ suggests there is a significant performance loss due to distributed detection. This is true even when optimal detection
DIVERSITY ORDER OF FULLY DISTRIBUTED SYSTEMS

is performed. The loss in performance arises due the fact that, as shown in Eqn. (1.15), joint processing requires the sum of $K$ NP-terms, i.e., coherent processing is not possible across the $K$ platforms. On the other hand, with collocated antennas, coherent processing is possible across all $NK$ antennas. The $\sqrt{K}$ relationship accounts for the resulting interaction between $P_F$ and $P_D$.

We will now analyze fully distributed systems and prove that under NP fusion, binary systems can achieve the same growth in diversity order as optimal detection.

2.5 Diversity Order of Fully Distributed Systems

The performance characteristic of optimal joint detection, suggested by Theorem 2.4.1, raises the question of how to characterize more practical and sub-optimal schemes. Joint detection provides an upper bound on the performance of distributed systems, but is impractical due to bandwidth limitations. Therefore there is a need to analyze more practical systems in which each sensor transmits a finite message to the fusion center.

We call a system fully distributed when each sensor transmits a binary (one-bit) decision to the fusion center. We analyze the diversity order of such systems, and prove that in the case of ‘optimal’ binary detection, the system achieves full diversity. Nevertheless, many fully distributed systems assume a fixed fusion rule and we will also analyze such systems and prove that they incur a significant loss in performance even when compared to ‘optimal’ binary detection.

2.5.1 Optimal Binary Detection

For a fully distributed system, the optimum fusion rule is an NP test as described in Eqn. (1.9). This leads to a statistic $\Lambda(u)$ comprising a summation of $K$ log-likelihood
ratios of the form:

\[ \Lambda(u) = \log \frac{\Pr(u|H_1)}{\Pr(u|H_0)} = \log \prod_{k=1}^{K} \frac{\Pr(u_k|H_1)}{\Pr(u_k|H_0)} = \sum_{i=1}^{K} \mathcal{L}(u_i|H_j), \]

\[ = \sum_{i=1}^{n_{H_1}} \log \frac{\Pr(1_i|H_1)}{\Pr(1_i|H_0)} + \sum_{i=1}^{n_{H_0}} \log \frac{\Pr(0_i|H_1)}{\Pr(0_i|H_0)}, \]  

(2.22)

where \( \mathcal{L}(u_i|H_j) \) is defined below in Eqns. (2.23)-(2.24) and \( \Pr(j_i|H_\ell) \) is the probability that sensor \( i \) declares hypothesis \( H_j \) given hypothesis \( H_\ell \) is true, where \( j, \ell = 0, 1 \). For example, \( \Pr(1_i|H_1) \) is the probability of detection at sensor \( i \), \( \Pr(1_i|H_0) \) is the probability of false alarm. In addition, \( n_{H_j} \) is the number of sensors that declare that hypothesis \( H_j \) is true (clearly, \( n_{H_0} + n_{H_1} = K \)). The statistic is then compared to a global threshold \( t_0 \) and randomization is necessary due to the discrete nature of \( \Lambda(u) \). The procedure is summarized by the following test:

\[ \Lambda(u) = \begin{cases} > t_0 & \Rightarrow H_1 \text{ is declared} \\ < t_0 & \Rightarrow H_0 \text{ is declared} \\ = t_0 & \Rightarrow H_1 \text{ is declared with probability } \delta \end{cases}. \]

In what follows, we will adopt the following notation:

\[ \alpha = \Pr(1|H_0) \text{ is the probability of false alarm;} \]
\[ \beta = \Pr(1|H_1) \text{ is the probability of detection;} \]
\[ q = \Pr(0|H_1) = 1 - \beta \text{ is the probability of miss;} \]
\[ p = \Pr(0|H_0) = 1 - \alpha. \]

Each of the log-likelihood ratios above is a realization of a binary random variable given the hypothesis. Under \( H_1 \), the log-likelihood ratio corresponding to the binary decision at sensor \( i \) corresponds to a random variable of the form:

\[ \mathcal{L}(u_i|H_1) = \begin{cases} \log(\beta/\alpha), & \text{with probability } \beta \\ \log(q/p), & \text{with probability } q = 1 - \beta \end{cases}, \]  

(2.23)
2.5. **DIVERSITY ORDER OF FULLY DISTRIBUTED SYSTEMS**

and under the null hypothesis:

\[
\mathcal{L}(u_i|H_0) = \begin{cases} 
\log(\beta/\alpha), & \text{with probability } \alpha \\
\log(q/p), & \text{with probability } p = 1 - \alpha 
\end{cases}.
\]  

(2.24)

Let \( \mu_1 \) and \( \mu_0 \) denote the means of \( \mathcal{L}(u_i|H_1) \) and \( \mathcal{L}(u_i|H_0) \) respectively, \( \sigma_1 \) and \( \sigma_0 \) their corresponding standard deviations. Here again, we will invoke the central limit theorem to derive a closed-form expression for the diversity order of ‘optimal’ binary detection. First note that the statistic is again a summation of \( K \) i.i.d. random variables. Hence for large \( K \) it can be approximated by a Gaussian random variable. As in Section 2.4.1, the probabilities of false alarm and detection are respectively defined as:

\[
P_F = \Pr(\Lambda(u) > T_h|H_0) = Q\left(\frac{T_h - K \mu_0}{\sqrt{K} \sigma_0}\right),
\]

(2.25)

\[
P_D = \Pr(\Lambda(u) > T_h|H_1) = Q\left(\frac{T_h - K \mu_1}{\sqrt{K} \sigma_1}\right),
\]

(2.26)

where again, \( T_h \) is a global threshold to be determined.

**Theorem 2.5.1** For a noise-limited system using ‘optimal’ binary detection, and for \( K \) large, the slope at \( P_D = 0.5 \) increases as

\[
\frac{dP_D}{d\gamma}(P_D = 0.5) = \frac{\beta N \ln^2 \beta \sqrt{K}}{\sqrt{2\pi} \ln \alpha (1 - \beta) \sqrt{\beta (1 - \beta)}}.
\]

(2.27)

**Proof:** See Appendix A.4.

**Corollary 2.5.2** If \( \alpha \) is chosen independently of \( K \), \( \beta \) will converge to \( \alpha \) and the diversity order will behave as \( N\sqrt{K} \).

This result implies that there is no diversity gain achieved by the system when the sensors transmit additional bits to the fusion center. In other words, a 1-bit system essentially captures all the degrees of freedom in the system (achieves full diversity) when it adopts an ‘optimum’ NP fusion rule. This provides an explanation of the fact that a limited number of bits can achieve near-optimum performance as empirically proved in [19].
2.5. DIVERSITY ORDER OF FULLY DISTRIBUTED SYSTEMS

The optimal binary detection system studied in this section is also not of practical significance due to many considerations. Most notable of these is the fact that the fusion center is required to know the channel characteristics at each sensor, which, even if possible in real time, greatly increases the bandwidth requirement even if the decision itself is restricted to a single bit. Therefore, many fully distributed systems assume that the fusion rule is fixed. The rule is of the form “n out of K”, where $n = 1$ corresponds to the OR rule (a target is declared present if any one sensor declares $H_1$), $n = K$ corresponds to the AND rule (a target is declared present only if all the sensors declare $H_1$), and the MAJ rules lie in between. In the rest of this section we will analyze the diversity order achieved by these sub-optimal, but practical, decision rules.

2.5.2 Diversity Order of the OR Rule

If the fusion center adopts the “1 out of $K$” (OR) rule, the total miss probability, $P_M = 1 - P_D$, is the product of the individual miss probabilities at the local sensors. Applying Eqn. (1.21), the probability of detection at each sensor is:

$$P_d^{(k)} = e^{-\frac{1}{\lambda_1} \gamma h^{(k)}} = e^{\ln \frac{P_{f}^{(k)}}{1 + N \gamma}} = \left( P_{f}^{(k)} \right)^{\frac{1}{1 + N \gamma}}. \quad (2.28)$$

Note that the last expression is the most common form used in the literature. Proceeding with the analysis, the total probability of detection is:

$$P_D = 1 - \left( 1 - e^{\ln \frac{P_{f}^{(k)}}{1 + N \gamma}} \right)^K. \quad (2.29)$$

Differentiating with respect to $\gamma$, we get the following expression of the slope:

$$\frac{dP_D}{d\gamma} = \frac{-K \times N \times \ln(P_{f}^{(k)}) e^{\ln \frac{P_{f}^{(k)}}{1 + N \gamma}} \left( 1 - e^{\ln \frac{P_{f}^{(k)}}{1 + N \gamma}} \right)^{K-1}}{(1 + N \gamma)^2}. \quad (2.30)$$

Theorem 2.5.3 For a noise-limited system using the “1 out of $K$” (OR) fusion rule, and for large $K$, the slope at $P_D = 0.5$ increases as $N \log K$.

Proof: See Appendix A.5.
2.5. DIVERSITY ORDER OF FULLY DISTRIBUTED SYSTEMS

Figure 2.6: $P_D$ versus SNR curves of optimal joint detection (MRC) and OR Rule.

Figure 2.7: Diversity Order of joint detection and the OR Rule as a function of $K$. 
and $K = 8$ sensors. The gap for the SNR difference at $P_D = 0.5$ increases from 1 dB to 2.5 dB. This clearly shows that the OR rule incurs a significant loss when compared to joint detection. In addition, Fig. 2.7 plots the diversity order for $N = 2, 4$. Note that the range on the $x$-axis has been chosen to show the considerable difference between the diversity order achieved by joint detection and the OR rule. The logarithmic increase for the OR rule is illustrated by the fact that the diversity order doubles when the number of sensors is squared. Despite this significant loss in performance, the OR rule is widely used in the literature due to its simplicity. As we will see, the OR rule provides the best diversity order among all fixed rules.

### 2.5.3 Diversity Order of the AND Rule

At the other extreme from the OR rule, if the fusion center adopts the ‘$K$ out of $K$’ (AND) rule, the following expressions hold:

\[
P_f^{(k)} = \sqrt{P_F}, \tag{2.31}
\]

\[
P_m^{(k)} = 1 - e^{\frac{\ln P_f^{(k)}}{1 + N \gamma}}, \tag{2.32}
\]

\[
P_D = (1 - P_m^{(k)})^K. \tag{2.33}
\]

Combining Eqns. (2.31), (2.32) and (2.33) we get:

\[
P_D = e^{\frac{\ln P_F}{1 + N \gamma}} = \left( P_f^{(k)} \right)^{\frac{1}{1 + N \gamma}}, \tag{2.34}
\]

which is independent of $K$, and we conclude that for the AND case, we have no diversity gain when we add multiple sensors.
2.5.4 Diversity Order of the MAJ Rules

If “n out of K” sensors are required to declare detection, \( n = 2, \ldots, K-1 \), the probability of detection is a binomial sum of the form:

\[
P_D = \sum_{i=n}^{K} \binom{K}{i} (P_d^{(k)})^i (1 - P_d^{(k)})^{K-i}.
\]  

(2.35)

Let us denote each element of the summation in Eqn. (2.35) by \( P_{d,\ell} \). We replace \( P_d^{(k)} \) by its value from Eqn. (1.21) and differentiate with respect to \( \gamma \):

\[
\frac{dP_{d,\ell}}{d\gamma} = - \binom{K}{\ell} \frac{1}{(1 + N\gamma)^2} e^{\ln P_f^{(k)}} (1 - e^{\ln P_f^{(k)}})^{K-\ell-1} \left( (K - \ell) N \ln P_f^{(k)} + KN \ln P_f^{(k)} e^{\ln P_f^{(k)}} \right),
\]

which is strictly positive for all values of \( \gamma \). However, even if we assume that \( P_f^{(k)} \) is asymptotically increasing with \( n \), \( \frac{dP_{d,\ell}}{d\gamma} \) is not monotonic in \( P_f^{(k)} \), and no generalization can be made regarding the behavior of the system as the parameter \( n \) varies.

2.5.5 Discussion

Figure 2.8 plots the \( P_D \) curves for the case of \( K = 4 \) as \( n \) is varied. It is clear from the figure that, as \( n \) is increased, the detection performance worsens and the diversity order decreases. Not visible on the curve is the case of \( K = 1 \) since it overlaps exactly with the case of \( n = K = 4 \), i.e., the AND case, which is consistent with the results of Section 2.5.3. Note that, at low SNR, the \( n = 2 \) case provides the best performance; however, at more practical SNR values, it is the OR rule (\( n = 1 \)) that does.

We finally note that if we let the probability of false alarm \( \alpha \) at each sensor decay as \( \frac{1}{K} \) (for the OR rule, \( P_F^{(k)} \approx P_F/K \)), it is straightforward to prove by the means of Theorem 2.5.1 that an optimal sensor will achieve a diversity order of \( N \log K \), which shows that the OR rule achieves the maximum possible diversity given the restriction on the local probabilities of false alarm.
2.6 Quantization for Distributed Detection

In the previous section, we proved that fully distributed systems using the NP test as a fusion rule achieve the same growth in diversity order as joint detection. In this section we will investigate this result further. We will first present the work of Lee and Chao [19] on multibit inter-sensor communications and we will then propose two improved quantization methods. This section is not meant to be a survey on quantization in the radar context, it was only included to corroborate the results we achieved in the previous sections.

2.6.1 On Space Partitioning for Distributed Detection

We will briefly introduce the work of Lee and Chao, the interested reader is referred to the original reference [19] and to [17, 18] for more details.

We start with some notation. The decision space is divided into two hypotheses: $H_0$, the null hypothesis, and $H_1$, the target-present hypothesis. If sensor $i$ transmits
2.6. QUANTIZATION FOR DISTRIBUTED DETECTION

$M$ bits, the authors subdivide the decision space for each hypothesis into $2^{M-1}$ regions, $H_{0,i,1}, H_{0,i,2}, \ldots, H_{0,i,2M-1}$ and $H_{1,i,1}, H_{1,i,2}, \ldots, H_{1,i,2M-1}$. We adopt the convention that the higher the third index is, the more confident the sensor is about the decision. In which case,

$$\alpha_{i,j} = \Pr(u_i = H_{1,j} | H_0)$$

is the probability of false alarm with confidence $j$;

$$\beta_{i,j} = \Pr(u_i = H_{1,j} | H_1)$$

is the probability of detection with confidence $j$;

$$q_{i,j} = \Pr(u_i = H_{0,j} | H_1)$$

is the probability of miss with confidence $j$;

$$p_{i,j} = \Pr(u_i = H_{0,j} | H_0).$$

Following the lead of the authors, we will derive the partitions for $M = 2$ bits. The work was also extended to the generic case where $M$ can be an arbitrary number of bits. Following the same steps as in Section 2.5.1, we get the following log-likelihood ratio:

$$T(U) = \sum_{i=1}^{K} w_i,$$

(2.36)

where

$$w_i = \begin{cases} 
\log \beta_{i,j}, & \text{if } u_i = H_{1,j}, \\
\log \frac{q_{i,j}}{p_{i,j}}, & \text{if } u_i = H_{0,j}, \quad i = 1, 2, \ldots, N, \quad j = 1, 2. 
\end{cases}$$

(2.37)

Optimizing the probability of detection over the random variables $\alpha_{i,j}$ and $\beta_{i,j}$ is not tractable, and the authors recur to the J-divergence as an objective function [19]. The J-divergence is defined as:

$$J = \mathcal{E}_{H_1} \{ T(U) \} - \mathcal{E}_{H_0} \{ T(U) \}$$

(2.38)

$$= \sum_{i=1}^{K} \sum_{j=1}^{2M-1} \left[ (q_{i,j} - p_{i,j}) \log \left( \frac{q_{i,j}}{p_{i,j}} \right) + (\beta_{i,j} - \alpha_{i,j}) \log \left( \frac{\beta_{i,j}}{\alpha_{i,j}} \right) \right].$$

(2.39)

where the indices $i$ and $j$ correspond to the sensors and the partitions respectively.

Under the assumption of independent observations, the optimization of the J-Divergence
at the individual sensors decouples, and the partitioning for each sensor reduces to finding
the optimal $\alpha_{i,2}$ and $p_{i,2}$ that maximize its own terms in the J-Divergence. The authors
differentiated the J-Divergence with respect to these variables and set the resulting ex-
pression to zero in order to find the optimal partitioning. We note that in [19] the total
probability of false alarm $\alpha_i$ at each sensor $i$ is fixed \textit{a priori}. Thus, given $\alpha_{i,2}$ and $p_{i,2}$,
all other partitioning parameters can be calculated accordingly.

The authors suggest that the procedure they described is optimal. However, we
disagree with this statement for three reasons:

1. The J-Divergence is not proved to be optimal in the $P_D$ sense;

2. The authors assume that $\alpha_i$ is constant, and there is no reason to believe that
   a pre-determined $\alpha_i$ will give the best performance (the global false alarm rate
   probability is fixed, but the global threshold and the randomization parameter
   allow for flexibility in choosing $\alpha_i$);

3. There is no reason to believe that allocating the same number of partitions to $H_1$
   and $H_0$ is optimal.

Despite this list, the results by Lee and Chao come very close to optimum, even with a
small value of $M$.

2.6.2 Other Quantization Techniques

In this section we will introduce two additional quantization methods, both using the
J-Divergence as the objective function. We will first relax the assumption that the
probability of false alarm at each sensors is fixed, and we will formulate an optimization
problem which we will solve to get the corresponding partitioning thresholds. The second
method, which we call sequential quantization, optimizes the regions sequentially.
2.6. QUANTIZATION FOR DISTRIBUTED DETECTION

Optimum Quantization

We first note that for $M = 2$, the J-Divergence reduces to:

$$J_i = (q_{i,1} - p_{i,1}) \log \frac{q_{i,1}}{p_{i,1}} + (q_{i,2} - p_{i,2}) \log \frac{q_{i,2}}{p_{i,2}} + (\beta_{i,1} - \alpha_{i,1}) \log \frac{\beta_{i,1}}{\alpha_{i,1}} + (\beta_{i,2} - \alpha_{i,2}) \log \frac{\beta_{i,2}}{\alpha_{i,2}}.$$

(2.40)

When we relax the assumption on $\alpha_i$, we have the following optimization problem:

$$\max J_i(\alpha_i, \alpha_{i,2}, p_{i,1})$$

subject to

$$\alpha_i < 1;$$

$$\alpha_{i,2} < \alpha_i;$$

$$p_{i,1} < 1 - \alpha_i;$$

$$\alpha_i, \alpha_{i,2}, p_{i,1} > 0.$$

This optimization problem is non-concave and it is not possible to guarantee that any solution is the global optimum. The approach taken here is to approximate the problem with a sequence of quadratic problems, an approach called Sequential Quadratic Programming (SQP) [22].

Sequential Quantization

We now describe what we call sequential quantization. In this procedure, we first optimally partition the space for $M = 1$. In this case, the J-Divergence reduces to:

$$J_i = (\beta_i - \alpha_i) \log \frac{\beta_i(1 - \alpha_i)}{\alpha_i(1 - \beta_i)}.$$

(2.42)

We find the optimum $\alpha_i$, and use this value as a constant for the optimization problem for $M = 2$. This problem will be exactly similar to the one presented in the previous section, instead $\alpha_i$ is replaced by the value obtained from the optimization of $M = 1$. We repeat this procedure until we reach the required number of bits. This technique works

\[\text{MATLAB function} \ fmincon \ \text{is an efficient implementation of SQP}\]
because whenever a threshold is fixed (in this case $\alpha_i$), the optimization of the regions separated by this threshold decouples. Figure 2.9 illustrates the performance gains of multibit communication over the popular single bit approach (the OR rule). As can be seen the Lee-Chao approach, even with two bits, recovers approximately 1 dB of the 1.5 dB performance loss between joint detection and the OR rule. Note that the two schemes suggested here improve on the Lee-Chao approach if not by any significant amount. So while we disagree with the claim of optimality in [19], their work does provide an effective partitioning scheme.

Note that the system tested had 1 antenna, it is not a STAP system. Simulations will show that in STAP scenarios with harsh interference, even the OR rule can achieve performance close to optimal. The main conclusion that we draw out from this section is a simple corroboration of the results previously achieved in this chapter: bandwidth does not have a significant effect on performance, and no effect at all on the asymptotic growth in diversity order.
2.7 Summary

This chapter covers two main contributions. First, we proved that noise-limited systems favor larger arrays. As we will show in Chapter 3, the results of this chapter extend to all systems that are essentially (but not strictly) limited by noise. In other words, noise-limited systems do not allow for spatial diversity gains, in which case larger arrays best exploit array gains in order to mitigate the effect of receiver noise.

The second contribution is that near-optimal results can be achieved with limited communication bandwidth, and we proved that fully distributed systems achieve the same growth in diversity as joint detection. This result, though highly counter-intuitive, was confirmed through both theory and simulations.
The previous chapter dealt with understanding the impact of system size ($N$ and $K$) in noise-limited distributed systems. In some scenarios, receiver noise can indeed be the limiting factor for STAP systems. However, most often systems must deal with jamming and clutter, and this chapter will primarily deal with STAP detection under jamming. As reviewed in Chapter 1, there has been little work in random radar networks, and up to our knowledge, the entirety of the research in this area (e.g. [6,15]) assumed that the geometry is fixed and known \textit{a priori}. In this chapter we will:

- Use the notion of a \textit{random} radar network to address the effect of geometry on STAP systems;
- Analyze the behavior of each sensor individually; more specifically, we will determine the distribution, mean and variance of the SINR at each sensor;
- Extend the analysis pertaining to a single sensor to a multistatic scenario with $K$ sensors;
- Investigate the tradeoffs between spatial diversity and interference cancellation in distributed radar systems akin to the diversity multiplexing tradeoff in wireless communication systems [5].
3.1 Random Unistatic Systems

When a certain range is monitored by a single sensor that is randomly chosen from the set of available sensors, we refer to the system as a random unistatic system. In other words, the analysis presented in this section reflects the behavior of an individual sensor. We assume that the sensor is randomly chosen without resorting to any pre-processing scheme that might assist the network in choosing the sensor that is best-fitted for detection. The general setting appears not to be tractable for any significant analysis; we will therefore analyze a theoretical scenario where the system is limited by noise and a single strong jammer.

3.1.1 General Case

In the general case, the array receives returns from the target (if present), in addition to interference from an arbitrary number of jamming signals. Jammers are modelled as arising from a single point in space, while clutter can be modelled as the superposition of many weak jammers. As stated in Section 1.2.2, the optimum space-time adaptive filter that maximizes output SINR is given by the adaptive weight vector [2]:

\[ w = R_n^{-1}s, \]  
(3.1)

where \( R_n = R_\gamma + R_j \) is the noise-plus-interference covariance matrix, and \( s \) is the space-time steering vector corresponding to the assumed target location and velocity. Since our goal is to develop an understanding of the spatial diversity, we drop the time dimension and focus on the spatial dimension exclusively. The output SINR \( \gamma_o \) is given by [2]:

\[ \gamma_o = |\alpha|^2 s^H R_n^{-1}s. \]  
(3.2)

Under the Swerling type-II target model, the term \(|\alpha|^2\) is the power of a Rayleigh-distributed random variable and hence is exponentially distributed regardless of the geometry of the system. On the other hand, the scalar \( s^H R_n^{-1}s \) is dependent on the target direction \( \theta_t \) and on the noise covariance matrix, both of which are geometry-dependent.
We will now derive the distribution of the output SINR. The following section was inspired by [23], we will therefore adopt the notation therein. In particular, $\langle . \rangle_A$ will denote the statistical expectation with respect to random variable $A$. The characteristic function of the output SINR is:

$$
\Psi_{\gamma_o}(z) = \langle e^{-z\gamma_o} \rangle_{\theta_t, R_n}.
$$

We start by the innermost term. $\alpha$ is a complex Gaussian random variable with power $A^2 = \mathcal{E}\{|\alpha|^2\}$, thus:

$$
\langle e^{-z\gamma_o} \rangle_{\alpha} = \int \frac{1}{\pi A^2} e^{-\frac{|\alpha|^2}{\pi^2} - z|\alpha|^2 s^H R_n^{-1} s} d\alpha
$$

$$
= \frac{1}{\pi A^2} \int \alpha e^{\frac{-|\alpha|^2}{\pi^2} \left(A^{-2} + zs^H R_n^{-1} s\right)^{-1}} d\alpha
$$

$$
= \frac{1}{A^2 \left(A^{-2} + zs^H R_n^{-1} s\right)}
$$

$$
= \frac{1}{1 + zA^2 s^H R_n^{-1} s},
$$

and consequently,

$$
\Psi_{\gamma_o}(z) = \left\langle \frac{1}{1 + zA^2 s^H R_n^{-1} s} \right\rangle_{\theta_t, R_n}.
$$

We first note that given $\theta_t$ and $R_n$, i.e., when these random variables are fixed, the characteristic function is that of an exponential random variable with mean $A^2 s^H R_n^{-1} s$, a result in harmony with our previous analysis. Unfortunately, any further analysis appears intractable.

The characteristic function of the SINR $\Psi_{\gamma_o}(z)$ can be expressed as the weighted sum of its moments. As in [23], we replace the characteristic function by its first order approximation and get:

$$
\Psi_{\gamma_o}(z) \approx \frac{1}{1 + zA^2 \langle s^H R_n^{-1} s \rangle_{\theta_t, R_n}}.
$$

Again, this is the characteristic function of an exponential random variable with mean
3.1. RANDOM UNISTATIC SYSTEMS

\[ \lambda = A^2 \langle s^H R^{-1}_m s \rangle_{\theta_n, R_n} \]. Figure 3.1 compares the true distribution of the SINR with the approximation of Eqn. (3.9) for \( N = 8 \) and \( N = 25 \). The system is limited by noise and a single jammer with Jammer-to-Noise-Ratio (JNR) of 40 dB and for an input SNR of 0 dB, i.e. \( \mathbb{E}\{|\alpha|^2/\sigma^2\} = 1 \) (these values will be assumed throughout this chapter).

The figure shows that the approximation is reasonably accurate for practical values of \( N \), but becomes less so as the number of antenna elements per sensor increases.

No further simplification of the analysis above can be achieved; and we will consequently address a simple system model that is limited by noise and a single jammer. Although theoretical, this analysis will offer insight into the SINR behavior in STAP systems.

3.1.2 One Jammer Scenario

We will now analyze a system that is limited by noise and a single jammer. We will derive the mean and variance of the output SINR for a specific system model and we will finally provide the results of simulations pertaining to various system models in order to
3.1. RANDOM UNISTATIC SYSTEMS

support our analysis.

The jammer steering vector is given by:

\[ \mathbf{a}_j = [1, e^{i\pi \cos \theta_j}, e^{i2\pi \cos \theta_j}, \ldots, e^{i(N-1)\pi \cos \theta_j}]^T, \]  

(3.10)

where \( \theta_j \) is the jammer direction. The noise-plus-interference covariance matrix is then:

\[ \mathbf{R}_n = \sigma^2 (\mathbf{I} + \gamma_j \mathbf{a}_j \mathbf{a}_j^H), \]  

(3.11)

where \( \gamma_j \) is the JNR. Inverting \( \mathbf{R}_n \) using the matrix inversion lemma \(^1\) we get:

\[ \mathbf{R}_n^{-1} = \frac{1}{\sigma^2} \left( \mathbf{I} - \frac{\gamma_j \mathbf{a}_j \mathbf{a}_j^H}{1 + \gamma_j \mathbf{a}_j \mathbf{a}_j^H} \right). \]  

(3.12)

We note that \( \mathbf{a}_j^H \mathbf{a}_j = \mathbf{s}^H \mathbf{s} = N \) where \( N \) is the number of antenna elements. The SINR becomes:

\[ \gamma_o = |\alpha|^2 \mathbf{s}^H \mathbf{R}_n^{-1} \mathbf{s} \]
\[ = \frac{|\alpha|^2}{\sigma^2} \left( N - \frac{\gamma_j |\mathbf{s}^H \mathbf{a}_j|^2}{1 + N \gamma_j} \right) \]
\[ = \frac{|\alpha|^2}{\sigma^2} \left( N - \frac{\gamma_j \sum_{n=0}^{N-1} e^{in\pi(\cos(\theta_i) - \cos(\theta_j))} \sigma^2}{1 + N \gamma_j} \right). \]  

(3.13)

In this scenario, the SINR is a function of two independent random variables: \( \alpha \) and \( \mathbf{u} = [\cos(\theta_i) - \cos(\theta_j)] \). Under the Swerling type-II target assumption, \( \alpha \) is complex Gaussian with average power \( A^2 = \mathbb{E}\{|\alpha|^2\} \). We will further simplify the SINR expression

\(^1 (\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{UC}^{-1} + \mathbf{VA}^{-1} \mathbf{U}^{-1} \mathbf{VA}^{-1} \)
3.1. RANDOM UNISTATIC SYSTEMS

by using the fact that:

\[ G(u) = \left| \sum_{n=0}^{N-1} e^{i\pi u} \right|^2 = \left( \sum_{n=0}^{N-1} \cos(n\pi u) \right)^2 + \left( \sum_{n=0}^{N-1} \sin(n\pi u) \right)^2 \]

\[ = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} [\cos(n\pi u) \cos(m\pi u) + \sin(n\pi u) \sin(m\pi u)] \]

\[ = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \cos[(n - m)\pi u]. \quad (3.14) \]

In order to proceed, we now define the all-ones vector \( \mathbf{1}(N) \) of length \( N \). We then define the length-(2\( N \) − 1) vector \( \Delta \) as the linear convolution of \( \mathbf{1}(N) \) with itself. In the sum of Eqn. (3.14), each term \( j = (n - m) = -N + 1, \ldots, N - 1 \) is repeated \( \Delta(j + N) \) times. We can now combine the two summations into one as follows:

\[ G(u) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \cos[(n - m)\pi u] \]

\[ = \sum_{j=-N+1}^{N-1} \Delta(j + N) \cos(j\pi u) \]

\[ = N + 2 \sum_{j=1}^{N-1} \Delta(j) \cos(j\pi u), \quad (3.15) \]

where \( \Delta(N) = N \) by virtue of the linear convolution.

Mean and Variance of the SINR

We now derive the mean and variance of the SINR. We first assume that the random variable \( u \) is uniform over the range \([-2, 2]\), i.e. \( u \sim \mathcal{U}\{-2, 2\} \). The range is chosen so as to conform to the difference of 2 cosines. This is simply a tractable model to develop an understanding about the behavior of the mean and variance of the SINR. In the next section, we will provide simulation results for more practical models.
3.1. RANDOM UNISTATIC SYSTEMS

We start with the mean:

\[ E\{G(u)\} = \int_{-2}^{2} G(u)f_u(u)du \]

\[ = \frac{1}{4} \int_{-2}^{2} N + 2 \sum_{j=1}^{N-1} \Delta(j) \cos(j\pi u)du \]

\[ = N + \frac{1}{4} \sum_{j=1}^{N-1} \int_{-2}^{2} \Delta(j) \cos(j\pi u)du = N, \quad (3.16) \]

because the last integral evaluates to zero for all values of \( j \). Using Eqn. (3.16) in Eqn. (3.13) we get:

\[ E\{\gamma_o\} = \frac{A^2}{\sigma^2} \left( N - \frac{N\gamma_j}{1 + N\gamma_j} \right), \quad (3.17) \]

which for large JNR, reduces to:

\[ E\{\gamma_o\} \approx \frac{A^2}{\sigma^2} (N - 1), \quad (3.18) \]

thus providing a theoretical explanation for the intuitive result that nulling a single jammer costs the system on average one degree of freedom.

We now derive an expression for the variance of the SINR. We denote the second term of Eqn. (3.15) by:

\[ \mathcal{Y} = 2 \sum_{j=1}^{N-1} \Delta(j) \cos(j\pi u). \quad (3.19) \]

We first note that \( E\{\mathcal{Y}\} = 0 \) and \( \text{var}\{\mathcal{Y}\} = \text{var}\{G(u)\} \) because \( N \) is a constant. We now
derive the variance of $Y$.

\[
\text{var}\{Y\} = \mathcal{E}\{Y^2\} - \mathcal{E}\{Y\}^2 = \mathcal{E}\{Y^2\}
\]

\[
= \mathcal{E}\left\{4 \sum_{n=1}^{N-1} \sum_{m=1}^{N-1} \Delta(n) \Delta(m) \times \cos(n\pi u) \cos(m\pi u)\right\}
\]

\[
= 4 \sum_{n=1}^{N-1} \sum_{m=1}^{N-1} \Delta(n) \Delta(m) \times \mathcal{E}\left\{\cos(n\pi u) \cos(m\pi u)\right\}. \tag{3.20}
\]

Recalling that $u \sim \mathcal{U}\{-2, 2\}$ we get:

\[
\mathcal{E}\left\{\cos(n\pi u) \cos(m\pi u)\right\} = \frac{1}{4} \int_{-2}^{2} \cos(n\pi u) \cos(m\pi u) du
\]

\[
= \begin{cases} 
\frac{1}{2}, & \text{if } n = m \\
0, & \text{if } n \neq m
\end{cases}, \tag{3.21}
\]

from which we get:

\[
\text{var}\{Y\} = 2 \sum_{n=1}^{N-1} \Delta^2(n). \tag{3.22}
\]

In what follows, we will denote the summation above by:

\[
\mathcal{S}\{M\} = \sum_{n=1}^{M} \Delta^2(n) = 1^2 + 2^2 + \cdots + M^2, \quad M \leq N \tag{3.23}
\]

\[
= \frac{M(M+1)(2M+1)}{6}, \tag{3.24}
\]

where the last identity can be proved by induction over $M$. Let $\mathcal{Z} = s^H \mathbf{R}_n^{-1} s$ so that $\gamma_0 = |\alpha|^2 \mathcal{Z}$. We assume without loss of generality that $A^2 = \mathcal{E}\{|\alpha|^2\} = 1$ and consequently, $\mathcal{E}\{|\alpha|^4\} = 2$ because $|\alpha|^2$ is exponentially distributed. This means that the input SNR is
determined by the receiver noise variance $\sigma^2$. We first use Eqn. (3.13) to state that:

$$\text{var}\{Z\} = \frac{1}{\sigma^4} \left( \frac{\gamma_j}{1 + N\gamma_j} \right)^2 \text{var}\{Y\}$$

$$= \frac{2}{\sigma^4} \left( \frac{\gamma_j}{1 + N\gamma_j} \right)^2 S\{N - 1\}. \quad (3.25)$$

We can now proceed to compute the SINR variance:

$$\text{var}\{\gamma_o\} = \mathcal{E}\{|\alpha|^4 Z^2\} - \mathcal{E}^2\{|\alpha|^2\}\mathcal{E}\{Z\}$$

$$= \mathcal{E}\{|\alpha|^4\}\mathcal{E}\{Z^2\} - \mathcal{E}^2\{Z\}$$

$$= 2\mathcal{E}\{Z^2\} - \mathcal{E}^2\{Z\} = \mathcal{E}\{Z^2\} + \text{var}\{Z\} = 2\text{var}\{Z\} + \mathcal{E}^2\{Z\}.$$  

Using Eqns. (3.25), (3.17) and (3.26) we get:

$$\text{var}\{\gamma_o\} = \frac{1}{\sigma^4} \left[ \frac{4S\{N - 1\}\gamma_j^2}{(1 + N\gamma_j)^2} + \left( N - \frac{\gamma_j N}{1 + N\gamma_j} \right)^2 \right]. \quad (3.26)$$

For practical levels of JNR ($N\gamma_j \gg 1$) and the variance can be approximated as:

$$\text{var}\{\gamma_o\} \approx \frac{1}{\sigma^4} \left[ \frac{4S\{N - 1\}}{N^2} + (N - 1)^2 \right]. \quad (3.27)$$

Using the expression for $S(N)$, the SINR variance is:

$$\text{var}\{\gamma_o\} = \frac{1}{\sigma^4} \left[ \frac{2N(N - 1)(2N - 1)}{3N^2} + (N - 1)^2 \right]$$

$$= \frac{1}{\sigma^4} \left[ \frac{(N - 1)(N + 1)(3N - 2)}{3N} \right]. \quad (3.28)$$

We note that the variance increases on the order of $N^2$ while the expectation is on the order of $N$. 
3.1. RANDOM UNISTATIC SYSTEMS

(a) Mean of the SINR when $N$ increases.

(b) Mean of the SINR when $J$ increases.

Figure 3.2: Mean of the SINR and $N$ and $J$ increase.
3.1. RANDOM UNISTATIC SYSTEMS

(a) Variance of the SINR when $N$ increases.

(b) Variance of the SINR when $J$ increases.

Figure 3.3: Variance of the SINR when $N$ and $J$ increase.
3.2. RANDOM MULTISTATIC SYSTEMS

Simulation Results

Figure 3.2 presents the average SINR for the more practical generic model when the number of antenna elements $N$ and the number of jammers $J$ are varied. On the same graph, we also plot the straight line $y = N - J$. Figure 3.2(a) shows the mean value of SINR when the number of sensor increases and the number of jammers stays constant, for $J = 1, 3, 5$. We note that in this scenario, the interference is effectively cancelled and each jammer on average consumes a single degree of freedom in the system. On the other hand, Figure 3.2(b) shows the mean SINR value when the number of interfering sources grows large, for $N = 12, 24$. In fact, when the sensor is not able to null the interference, the SINR per sensor drops to zero, and the system turns futile. In extremely harsh interference scenarios, the mean SINR exceeds $y = N - J$, however the system is already working in extremely poor conditions and can be perceived as failing.

As for the variance of the SINR, the results are depicted in Figure 3.3. Figure 3.3(a) shows the parabolic behavior ($\sim N^2$) of the variance for $J = 1, 3, 5$. In addition, it shows that the variance is inversely proportional to the number of jammers. We note however that the mean also decreases and hence the reduction in the variance does not imply better performance. Figure 3.3(b) shows the behavior of the variance when the number of jammers increases for $N = 12, 24$, and the results therein corroborate those of the previous figure. On the one hand, we see that the variance decreases when $J$ increases; and as expected, when the mean approaches its minimum the variance does too. On the other hand, we notice a significant difference in the variance between $N = 12$ and $N = 24$; and the twofold difference in $N$ is translated faithfully into a fourfold difference in the variance for the lowest values of $J$.

3.2 Random Multistatic Systems

The previous section focused on a single sensor with $N$ collocated antennas. This sensor was picked randomly from a set of available sensors. We now analyze multistatic random radar networks. In such systems, $K$ sensors are randomly chosen to detect the presence of a target at a certain range. There are two main system design problems: 1) how many...
3.2. RANDOM MULTISTATIC SYSTEMS

3.2.1 An Upper Bound on SINR

In a system that adopts joint processing, and under the assumption that the observations at the sensors are independent, the output SINR is equal to the sum of the individual SINRs, and this is consequently an upper bound on performance (as we proved, this is similar to MRC). As a performance measure, we introduce the following metric, briefly alluded to earlier:

$$
\Gamma = \frac{\mathbb{E}\{\gamma_0\}}{\text{var}\{\gamma_0\}}.
$$

The numerator is a measure of the average ‘power’ available in the system; the denominator is the variance, which is a measure of the uncertainty in the system. The overall metric is, therefore, a measure of the relative reliability of the system.

Here again, the general case is not tractable, and we will give an expression for the noise-plus-1-jammer case described in Section 3.1.2. In this case, the metric for the overall system reduces to:

$$
\Gamma = K \frac{\mathbb{E}\{\gamma_0\}}{\text{var}\{\gamma_0\}} = \frac{3KN(N-1)}{(N+1)(3N-2)}.
$$

We would like to investigate the tradeoff of setting up a large number of smaller sensors versus having a small number of large sensors. To achieve this end, we assume that the total number of antenna elements in the system is fixed, and we denote it by $\eta$. 
Consequently, \( N = \frac{n}{K} \). The metric of Eqn. (3.30) reduces to:

\[
\Gamma = \frac{3K \frac{n}{K} \left( \frac{n}{K} - 1 \right)}{(\frac{2}{K} + 1)(\frac{\eta}{K} - 2)} = 3\eta \frac{K(\eta - K)}{(\eta + K)(3\eta - 2K)}. \tag{3.31}
\]

We first prove that \( \Gamma \) increases with increasing \( K \). In STAP \( N \geq 2 \) is assumed, hence \( K \leq \frac{\eta}{2} \). Differentiating \( \Gamma \) with respect to \( K \), we get:

\[
\frac{d\Gamma}{dK} = \frac{3\eta^2}{(\eta + K)^2(3\eta - 2K)^2} \times \left( K^2 - 6K\eta + 3\eta^2 \right). \tag{3.32}
\]

Term 1 is always positive. We now prove that for feasible values of \( K \), the second term is also positive. The roots of this polynomial are: \( k_1 = \eta(3 + \sqrt{6}) \) and \( k_2 = \eta(3 - \sqrt{6}) \approx 0.55\eta \). Knowing that the above polynomial describes a convex parabola and that \( K \leq \eta/2 \) which is in turn smaller than the smaller root \( k_2 \), this means that the polynomial evaluates to positive values for all feasible \( K \), hence \( d\Gamma/dK \) is positive and \( \Gamma \) is strictly increasing with \( K \). Figure 3.4 reflects this results. The product \( \eta = NK \) is set to 72, 96 and 120, and the metric is calculated when the number of sensors \( K \) is
3.2. RANDOM MULTISTATIC SYSTEMS

Figure 3.5: SINR defined as the deflection ratio.

increases, while $NK$ stays constant. The results consider cases when $N \geq 4$, meaning that interference is effectively canceled for all values shown on this plot.

This result shows that as the number of sensors increases, the system can potentially achieve better results. The following section presents the results of simulations that further validate the analysis we just presented.

3.2.2 Deflection Ratio

The statistic of Eqn. (1.15) cannot be separated into signal and noise terms, therefore there is no straightforward definition for SINR in a distributed system. This fact led Goodman et al. to use the deflection ratio as a proxy for SINR [6]:

$$
\gamma_o \triangleq \frac{|\mathcal{E}\{\zeta|H_1\} - \mathcal{E}\{\zeta|H_0\}|^2}{\frac{1}{2}[\text{var}\{\zeta|H_1\} + \text{var}\{\zeta|H_0\}]}.
$$

(3.33)
This metric is faithfully translated in terms of performance only when the metric is Gaussian-distributed. Nevertheless, the metric is presented in [6] as a proxy for performance. Note that the numerator is the difference between the means under the target present hypothesis and null hypothesis respectively, which is an indication of the certainty of detection. The denominator is the sum of the corresponding variances, which is in turn a measure of uncertainty. Figure 3.5 plots the deflection ratio for a constant product $NK = 120$ and for 1, 3 and 5 Jammers. The results mainly show that we have SINR gains as long as we are able to successfully null the interference caused by the jammers. However, for $J = 5$, the system is not able to null the jammer for $K = 30$ (which corresponds to $N = 4$) and the system SINR drops to zero.

### 3.2.3 Probability of Detection

In what follows, we develop the average $P_D$ results for joint detection (MRC) and the OR rule. Under both hypotheses, the test statistic (1.15) is exponentially distributed. For the OR rule, the total probability of detection is

$$P_D = 1 - \prod_{k=1}^{K} (1 - P_d^{(k)}),$$

(3.34)

where $P_d^{(k)}$ is the individual probability of detection at sensor $k$. The procedure is identical to that described in Sections 2.4 and 2.5. In joint detection, each sensor contributes an exponentially distributed term with mean $\lambda_0 = A^2 s_k^H R_n^{-1} s_k / (1 + A^2 s_k^H R_n^{-1} s_k)$ under $H_0$ and $\lambda_1 = A^2 s_k^H R_n^{-1} s_k$ under $H_1$. The statistic is a linear sum of exponential random variables and it follows a distribution of the form [21]:

$$f_{\sum x_j} (\zeta) = \left( \prod_{i=1}^{n} \lambda_i \right) \sum_{j=1}^{n} e^{-\lambda_j \zeta} \prod_{k=1, k \neq j}^{n} (\lambda_k - \lambda_j), \quad \zeta > 0.$$  

(3.35)

This distribution is used to derive the probability of detection: first we invert this expression with the corresponding $\lambda_i,0$ to get the global threshold, we then use this threshold into the same expression, but now with $\lambda_i,1$ instead of $\lambda_i,0$ in order to calculate the probability of detection.
3.2. RANDOM MULTISTATIC SYSTEMS

Figure 3.6: Plots for the average $P_D$ with $J = 1, 3$. 

(a) Average $P_D$ plots with 1 jammer

(b) Average $P_D$ plots with 3 jammers
Figure 3.6 shows the average $P_D$ results for a system with a total of $NK = 24$ antenna elements. Both MRC and the fully distributed OR rule are portrayed. Figure 3.6(a) shows the average $P_D$ for 1 jammer. We first note that in this case, all the sensors are able to null the interference. For low input SNR, the larger sensors achieve better performance thanks to the array gain. When channel conditions improve, interference cancellation becomes the bottleneck to performance, and a larger number of sensors achieve better results. It is interesting to note that in this case the system exploits the spatial diversity as clearly manifested in the $K = 3$ and $K = 6$ cases. In terms of diversity order [24], the more sensors we have, the steeper the $P_D$ curve is. This comes at the expense of poorer performance for low input SNR when the system is considered to be ‘noise-limited’.

On the other hand, Figure 3.6(b) shows the average $P_D$ for 3 jammers. The result is similar for $K = 3$ sensors. However, we see that for $K = 6$ which corresponds to $N = 4$, the sensors are not able to effectively null the 3 jammers, and the system achieves poor performance over all input SNR regimes. We note that the OR rule achieves results that are surprisingly close to optimum, especially when the interference environment become harsher.

### 3.3 Summary

In this chapter we analyzed distributed systems in the presence of interference. We first analyzed the behavior of each sensor individually and we derived the distribution of the SINR as well as its corresponding mean and variance. We extended this analysis to multistatic systems, and we proved that adding sensors to the system enables it to achieve higher spatial diversity gain. In fact, when the number of sensors increases, it becomes less likely for the target to be masked by the jammers and clutter, and also to fall outside of the field of ‘vision’ (direction and Doppler) of the overall system. We proved that the sensors should also be large enough to null the surrounding interference in order to exploit the available spatial diversity. In simpler terms: larger arrays enable better interference cancellation, more sensors lead to better spatial diversity gains; at the same time, power consumption is a main concern in a variety of applications, thus radar
engineers are required to find the proper tradeoff between the size of the arrays and their number in order to maximize performance and minimize power consumption.

Finally, we note that this chapter validates our claim that our definition for diversity is more suitable for radar networks than the communication-like definitions of [15, 16]. Specifically, if we follow any of the other definitions, the system with $K = 6$ sensors in Figure 3.6(b) would have higher diversity than the system with $K = 3$ sensors. Obviously, these definitions ignore the fact that interference should be effectively cancelled for the system to properly exploit spatial diversity; our definition tackles this problem properly.
Chapter 4

Conclusion and Future Work

We have made throughout this work several contributions relating to the analysis and design of distributed radar detection systems. Our first contribution was the notion of diversity order in radar networks. We first analyzed noise-limited systems and we derived the diversity order of joint detection and fully distributed detection, and we proved that both grow as $\sqrt{K}$. We also studied sub-optimal fully distributed systems, and we proved that the OR rule, despite achieving the best performance, only leads to a growth in diversity that is proportional to $(\log K)$. These results entail two main consequences: First, in noise-limited systems, array gain prevails over diversity gain, hence larger antenna arrays are preferable. The second contribution is the proof that increasing the communications bandwidth between sensors does not affect the growth in diversity order, which provides theoretical ground for many empirical evaluations available in the literature of fully and partially distributed systems.

We then turned our attention to more practical systems where jamming is taken into account. We introduced the notion of a random radar network in order to study the effect of geometry on distributed STAP detection. We analyzed the distribution of SINR at each individual sensor and we derived expressions for the corresponding mean and variance. We proved empirically that the mean of the SINR closely follows the difference between the degrees of freedom and the number of interferers. We then analyzed multistatic systems. We provided an upper bound on performance and proved that such systems can potentially achieve better performance when the number of sensors increases. We
then performed an empirical evaluation of multistatic networks. This evaluation lead to the characterization of a design tradeoff for distributed sensing systems, which was the last contribution of this work. In a few words: larger arrays cancel interference more effectively, and more sensors achieve higher diversity. However, for reasons of power and mobility, we need to reduce both the size and the number of sensors, so we need to carefully choose these parameters in order to meet the required performance with minimal power consumption and without sacrificing mobility.

4.1 Future Work

The extension of this work to clutter is straight forward and does not provide any particular insight to the problem at hand, thus we do not envisage it as a potential subject for future research. However, we have not had the chance to quantify power consumption and mobility. Clearly, the larger the array the more power it consumes, and this increase in power consumption per additional antenna is to be quantified. In addition, we also expect a significant overhead in power consumption from adding entire sensors, and we foresee another power-related tradeoff between adding more sensors or increasing the size of each of the sensors.

Mobility is also a major concern and has been the subject for recent research (e.g. [25]). But first, what is mobility? Is it the ability of the sensors to physically move? Or is it proportional to the level of coordination between the sensors? How large can a sensor be and still be considered mobile? Finally, how crucial is mobility? The answer to these questions is closely tied to the application. For example, on friendly grounds the sensors would be dispersed freely without obstacles, and the constraint on mobility tends to be less stringent. However, if the sensors are dispersed on enemy grounds, mobility becomes a primary concern.

An interesting extension to this work would be to formulate an optimization problem where power consumption is minimized given that the system achieves a certain diversity order and that a certain level of mobility is satisfied. Solving this problem will enable system designers to build efficient systems that both minimize power and maximize performance without sacrificing mobility.
List of Publications


1Finalist in the student paper competition
Appendix A

Derivations

A.1 NP Test Statistic for Distributed Detection

We derive the NP test for STAP distributed detection. The derivation follows the same lines as [6]. Under the null hypothesis $H_0$, and due to the independence assumption,

$$
\text{Pr}\{z_1, z_2, \ldots, z_K|H_0\} = \prod_{k=1}^{K} \frac{1}{\pi^N|\mathbf{R}_n|} e^{-z_k^H \mathbf{R}_n^{-1} z_k}. \quad (A.1)
$$

Similarly, under the target-present hypothesis, $H_1$,

$$
\text{Pr}\{z_1, \ldots, z_K|H_1\} = \prod_{k=1}^{K} \frac{1}{\pi^N|\mathbf{R}_n + \mathbf{S}_k|} e^{-z_k^H (\mathbf{R}_n + \mathbf{S}_k)^{-1} z_k}. \quad (A.2)
$$

The likelihood ratio corresponding to the NP test is of the form:

$$
\Lambda(z_1, \ldots, z_K) = \frac{\text{Pr}\{z_1, \ldots, z_K|H_1\}}{\text{Pr}\{z_1, \ldots, z_K|H_0\}} = \prod_{k=1}^{K} \frac{\text{Pr}(z_k|H_1)}{\text{Pr}(z_k|H_0)}. \quad (A.3)
$$
Substituting Eqns. (A.1) and (A.2) into Eqn. (A.3) and taking the natural logarithm we get the statistic:

\[ \ln \Lambda(z_1, z_2, \ldots, z_K) = \sum_{k=1}^{K} \ln \frac{|R_n|}{|R_n + S_k|} - \sum_{k=1}^{K} z_k^H (R_n^{-1} - (R_n + S_k)^{-1}) z_k. \]  (A.4)

Dropping the first term because it does not depend on the received vector, the test-statistic reduces to:

\[ \zeta = - \sum_{k=1}^{K} z_k^H (R_n^{-1} - (R_n + S_k)^{-1}) z_k. \]  (A.5)

From the matrix inversion lemma and Eqn. (1.11) we have:

\[ (R_n + S_k)^{-1} = (R_n + s_k A^2 s_k^H)^{-1}, \]

\[ = R_n^{-1} - R_n^{-1} s_k \left( \frac{1}{A^2 + s_k^H R_n s_k} \right)^{-1} s_k^H R_n^{-1}, \]

\[ = R_n^{-1} - \frac{A^2 R_n^{-1} s_k s_k^H R_n^{-1}}{1 + A^2 s_k^H R_n^{-1} s_k}. \]  (A.6)

Combining Eqn. (A.5) and Eqn. (A.6), we finally get the following test-statistic [6]:

\[ \zeta = \sum_{k=1}^{K} \frac{A^2 |s_k^H R_n^{-1} z_k|^2}{1 + A^2 s_k^H R_n^{-1} s_k}. \]  (A.7)

### A.2 NP Test Statistic for Joint Processing

We proceed from Eqn. (A.5), repeated for convenience:

\[ \zeta = z^H (R_n^{-1} - (R_n + \Psi)^{-1}) z. \]  (A.8)
From the Matrix Inversion Lemma and Eqn. (2.14) we have:

\[(R_n + \Psi)^{-1} = (R_n + A^2 S I_K S^H)^{-1},\]  
\[= R_n^{-1} - R_n^{-1} S \left( \frac{1}{A^2} I_K + S^H R_n^{-1} S \right)^{-1} S^H R_n^{-1}.\]  

Proceeding with the derivation:

\[
\left( \frac{1}{A^2} I_K + S^H R_n^{-1} S \right)^{-1} = \left( \frac{1}{A^2} I_K + \begin{pmatrix}
\frac{1}{A^2} R_n^{-1}_1 s_1 & 0 & 0 & 0 \\
0 & \frac{1}{A^2} R_n^{-1}_2 s_2 & 0 & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \frac{1}{A^2} R_n^{-1}_K s_K
\end{pmatrix} \right)^{-1},
\]  
\[= \begin{pmatrix}
\frac{1}{A^2} + \frac{1}{A^2} R_n^{-1}_1 s_1 & 0 & 0 & 0 \\
0 & \frac{1}{A^2} + \frac{1}{A^2} R_n^{-1}_2 s_2 & 0 & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \frac{1}{A^2} + \frac{1}{A^2} R_n^{-1}_K s_K
\end{pmatrix},
\]  
\[= \begin{pmatrix}
\frac{A^2}{1 + A^2 s_1^H R_n^{-1}_1 s_1} & 0 & 0 & 0 \\
0 & \frac{A^2}{1 + A^2 s_2^H R_n^{-1}_2 s_2} & 0 & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \frac{A^2}{1 + A^2 s_K^H R_n^{-1}_K s_K}
\end{pmatrix}.\]  

With simple matrix manipulations, it can be shown that:

\[\zeta = z^H (R_n^{-1} - (R_n + \Psi)^{-1}) z,\]  
\[= \sum_{k=1}^{K} \frac{A^2 |s_k^H R_n^{-1} z_k|^2}{1 + A^2 s_k^H R_n^{-1} s_k},\]  
which follows the same form as distributed detection.
A.3 Proof of Theorem 2.4.1

We derive a closed-form expression for the diversity order of joint detection. The following derivation assumes that the number of sensors is large and the statistic can be approximated by a Gaussian random variable.

We start by setting $P_D = 0.5$ and we get:

$$P_D = \Pr(\zeta > T_h | H_1) = Q\left(\frac{T_h - K\sigma_1}{\sqrt{K}\sigma_1}\right) = 0.5. \quad (A.16)$$

Inverting the $Q$-function, we have:

$$\frac{T_h - K\lambda_1}{\sqrt{K}\lambda_1} = 0. \quad (A.17)$$

We now replace $T_h$ by its value from Eqn. (2.18) and get:

$$\frac{\lambda_0 \sqrt{K} [\sqrt{K} + Q^{-1}(P_F)]}{\sqrt{K}\lambda_1} = \sqrt{K}. \quad (A.18)$$

Expanding Eqn. (A.18) and replacing $\lambda_0$ and $\lambda_1$ by their values in Eqn. (2.5) and Eqn. (2.6) respectively, we get:

$$1 + N\gamma = \frac{\sqrt{K} + Q^{-1}(P_F)}{\sqrt{K}}. \quad (A.19)$$

Let us now introduce a new variable $t$ such that:

$$P_D = Q\left(\frac{T_h - K\lambda_1}{\sqrt{K}\lambda_1}\right) = Q(t), \quad (A.20)$$

$$t = \frac{T_h - K\lambda_1}{\sqrt{K}\lambda_1} = \frac{\lambda_0 \sqrt{K} [\sqrt{K} + Q^{-1}(P_F)] - K\lambda_1}{\sqrt{K}\lambda_1}, \quad (A.21)$$

$$= \frac{[\sqrt{K} + Q^{-1}(P_F)]}{1 + N\gamma} - \sqrt{K}. \quad (A.22)$$
A.4. PROOF OF THEOREM 2.5.1

and differentiating with respect to $\gamma$ we get:

$$\frac{dt}{d\gamma} = -\frac{N.\left[\sqrt{K} + Q^{-1}(P_F)\right]}{(1 + N\gamma)^2}. \quad (A.23)$$

From the definition of the $Q$-function:

$$P_D = Q(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \quad (A.24)$$

Using the chain rule and the fundamental theorem of calculus:

$$\frac{dP_D}{d\gamma} = \frac{dP_D}{dt} \frac{dt}{d\gamma} = \frac{N.\left[\sqrt{K} + Q^{-1}(P_F)\right]}{\sqrt{2\pi}(1 + N\gamma)^2} e^{-\frac{1}{2}\left(\frac{\sqrt{K} + Q^{-1}(P_F)}{1 + N\gamma} - \sqrt{K}\right)^2}. \quad (A.25)$$

We first note that at $P_D = 0.5$, $t = 0$ and the exponent always reduces to zero. Hence, we plug Eqn. (A.19) into Eqn. (A.25) and the slope at $P_D = 0.5$ becomes:

$$\frac{dP_D}{d\gamma} = \frac{N.K}{\sqrt{2\pi}\left[\sqrt{K} + Q^{-1}(P_F)\right]}, \quad (A.26)$$

which behaves as

$$N\sqrt{K}, \quad (A.27)$$

for large $K$.

A.4 Proof of Theorem 2.5.1

We now derive a closed-form expression for the diversity order of a fully distributed system adopting the optimal fusion rule described in Eqn. (2.22), and repeated here for convenience:

$$\Lambda(u) = \log \frac{\Pr(u|H_1)}{\Pr(u|H_0)} = \log \prod_{k=1}^{K} \frac{\Pr(u_k|H_1)}{\Pr(u_k|H_0)} = \sum_{i=1}^{K} \mathcal{L}(u_i|H_j). \quad (A.28)$$
As we previously stated, the statistic is a summation of \( K \) binary random variables.

Under \( H_1 \):

\[
\mathcal{L}_1 \triangleq \mathcal{L}(u_i|H_1) = \begin{cases} 
\log(\beta/\alpha), & \text{with probability } \beta \\
\log(q/p), & \text{with probability } q = 1 - \beta 
\end{cases}.
\]  
(A.29)

\( \mathcal{L}_1 \) has mean and variance

\[
\mu_1 = \mathcal{E}\{\mathcal{L}_1\} = \beta \log(\beta/\alpha) + q \log(q/p), \quad \sigma_1^2 = \text{var}\{\mathcal{L}_1\} = \mathcal{E}\{\mathcal{L}_1^2\} - \mathcal{E}^2\{\mathcal{L}_1\}
\]

\[
= \beta \log^2(\beta/\alpha) + q \log^2(q/p) - \beta^2 \log^2(\beta/\alpha) - q^2 \log^2(q/p) - 2\beta \log(\beta/\alpha)q \log(q/p)
\]

\[
= \beta q[\log(\beta/\alpha) - \log(q/p)]^2.
\]  
(A.31)

respectively. Similarly, under \( H_0 \),

\[
\mathcal{L}_0 \triangleq \mathcal{L}(u_i|H_0) = \begin{cases} 
\log(\beta/\alpha), & \text{with probability } \alpha \\
\log(q/p), & \text{with probability } p = 1 - \alpha 
\end{cases}.
\]  
(A.32)

The corresponding moments are:

\[
\mu_0 = \mathcal{E}\{\mathcal{L}_0\} = \alpha \log(\beta/\alpha) + p \log(q/p),
\]

\[
\sigma_0^2 = \text{var}\{\mathcal{L}_0\} = \mathcal{E}\{\mathcal{L}_0^2\} - \mathcal{E}^2\{\mathcal{L}_0\}
\]

\[
= \alpha \log^2(\beta/\alpha) + p \log^2(q/p) - \alpha^2 \log^2(\beta/\alpha) - p^2 \log^2(q/p) - 2\alpha \log(\beta/\alpha)p \log(q/p)
\]

\[
= \alpha p[\log(\beta/\alpha) - \log(q/p)]^2.
\]  
(A.34)

When \( K \) is large, the statistic can be approximated by a Gaussian random variable. Hence, under \( H_0 \), the mean is \( \mu = K\mu_0 \) and the standard deviation \( \sigma = \sqrt{K}\sigma_0 \). The global probability of false alarm \( P_F \) is fixed, and the global threshold \( T_h \) can be calculated according to the following expression:

\[
P_F = Q\left( \frac{T_h - K\mu_0}{\sqrt{K}\sigma_0} \right),
\]  
(A.35)
from which we get:

\[ T_h = Q^{-1}(P_F)\sqrt{K\sigma_0} + K\mu_0. \] (A.36)

Under \( H_1 \), the statistic is also Gaussian distributed but now with mean \( \mu = K\mu_1 \) and standard deviation \( \sigma = \sqrt{K}\sigma_1 \). The probability of detection can be expressed as:

\[ P_D = Q\left(\frac{T_h - K\mu_1}{\sqrt{K}\sigma_1}\right) = Q(t), \] (A.37)

where:

\[ t = \frac{T_h - K\mu_1}{\sqrt{K}\sigma_1} = \frac{Q^{-1}(P_F)\sqrt{K}\sigma_0 + K\mu_0 - K\mu_1}{\sqrt{K}\sigma_1}, \] (A.38)

Combining Eqn. (A.31) and Eqn. (A.34) we get:

\[ \frac{\sigma_0}{\sigma_1} = \sqrt{\frac{\alpha(1 - \alpha)}{\beta(1 - \beta)}}. \] (A.39)

Combining Eqn. (A.30) and Eqn. (A.33) we get:

\[ \mu_0 - \mu_1 = (\alpha - \beta)[\log(\beta/\alpha) - \log(q/p)], \] (A.40)

which, combined with Eqn. (A.31) gives:

\[ \frac{\mu_0 - \mu_1}{\sigma_1} = \frac{\alpha - \beta}{\sqrt{\beta(1 - \beta)}}. \] (A.41)

We note that reversing the sign of this last equation which occurs when \( \log(q/p) > \log(\beta/\alpha) \) will lead to the same final result. Now combining Eqns. (A.38), (A.39), (A.40)
and (A.41) we get:

\[ t = Q^{-1}(P_F) \sqrt{\frac{\alpha(1-\alpha)}{\beta(1-\beta)}} + \sqrt{K} \frac{\alpha - \beta}{\sqrt{\beta(1-\beta)}}. \]  
(A.42)

\[ = \frac{Q^{-1}(P_F) \sqrt{\alpha(1-\alpha) + \sqrt{K} \alpha}}{\sqrt{\beta(1-\beta)}} + \frac{\sqrt{K} \sqrt{\beta}}{\sqrt{1 - \beta}}. \]  
(A.43)

\[ = \frac{Q^{-1}(P_F) \sqrt{\alpha(1-\alpha) + \sqrt{K} \alpha}}{\sqrt{\beta(1-\beta)}} - \frac{\sqrt{K} \sqrt{\beta}}{\sqrt{1 - \beta}}. \]  
(A.44)

We are given (Eqn. 2.28):

\[ \beta = \alpha^{\frac{1}{1+N\gamma}}, \]  
(A.45)

and we recall from the analysis of the diversity order of joint detection that:

\[ \frac{dP_D}{d\gamma} = -\frac{1}{\sqrt{2\pi}} \frac{dt}{d\gamma}, \]  
(A.46)

so the problem reduces to finding the behavior of \( \frac{dt}{d\gamma} \). We first note that \( \alpha \) and \( K \) are independent of \( \gamma \). So we begin by the first term of Eqn. (A.44):

\[ \frac{d}{d\gamma} \frac{1}{\sqrt{\beta(1-\beta)}} = \frac{-N(1 - 2\beta)(\ln \alpha)}{2(1 + N\gamma)^2(1-\beta)\sqrt{\beta(1-\beta)}}. \]  
(A.47)

As for the second term:

\[ \frac{d}{d\gamma} \frac{\sqrt{\beta}}{\sqrt{1-\beta}} = \frac{-N \ln \alpha \sqrt{\beta(1-\beta)}}{2(1 + N\gamma)^2(1-\beta)^2}. \]  
(A.48)
Combining Eqns. (A.44), (A.47) and (A.48) we get:

\[
\frac{dt}{d\gamma} = -\frac{[Q^{-1}(P_F)\sqrt{\alpha(1-\alpha)} + \sqrt{K}\alpha]N(1-2\beta)(\ln \alpha) + \sqrt{K}N(\ln \alpha)\beta(1-\beta)}{2(1+N\gamma^2)(1-\beta)^2},
\]

\[
= \frac{-[Q^{-1}(P_F)\sqrt{\alpha(1-\alpha)} + \sqrt{K}\alpha]N(1-2\beta)(\ln \alpha)(1-\beta) + \sqrt{K}N(\ln \alpha)\beta(1-\beta)}{2(1+N\gamma^2)(1-\beta)^2\sqrt{\beta(1-\beta)}},
\]

(A.49)

At \( P_D = 0.5 \), we have \( t = 0 \), so using Eqn. (A.43) we get:

\[
\beta = \frac{Q^{-1}(P_F)\sqrt{\alpha(1-\alpha)}}{\sqrt{K}} + \alpha,
\]

(A.51)

and from Eqn. (A.45) we have

\[
(1 + N\gamma) = \frac{\ln \alpha}{\ln \beta}.
\]

(A.52)

Using Eqns. (A.51) and (A.52) with Eqn. (A.50) and then Eqn. (A.46) we get:

\[
\frac{dP_D}{d\gamma}(P_D = 0.5) = \frac{-\beta N \ln \beta \sqrt{K}}{\sqrt{2\pi \ln \alpha(1-\beta)\sqrt{\beta(1-\beta)}}},
\]

(A.53)

thus concluding the proof.
A.5 Proof of Theorem 2.5.3

We finally prove that the diversity order for a distributed system using the OR rule grows logarithmically in the number of sensors. For $P_D = 0.5$,

$$P_D = 1 - \left(1 - e^{\ln P_f^{(k)}}\right)^K = 0.5,$$

(A.54)

$$e^{\ln P_f^{(k)}} = 1 - 0.5^{\frac{1}{K}};$$

(A.55)

$$\gamma = \frac{\ln \left(\frac{\ln P_f^{(k)}}{1-0.5^{\frac{1}{K}}}\right)}{N \ln \left(1 - 0.5^{\frac{1}{K}}\right)}.$$  

(A.56)

Consequently,

$$1 + N\gamma = \frac{\ln P_f^{(k)}}{\ln(1 - 0.5^{\frac{1}{K}})}.$$  

(A.57)

Using Eqns. (A.55) and (A.57) in conjunction with Eqn. (2.30) and noting that for the OR fusion rule:

$$P_f^{(k)} \approx \frac{P_F}{K},$$

(A.58)

we get the following expression for the slope of the $P_D$ curve at $P_D = 0.5$:

$$\frac{dP_D}{d\gamma} = \frac{-NK \ln^2 \left(1 - 0.5^{\frac{1}{K}}\right) \left(0.5^{\frac{1}{K}}\right)^{K-1} \left(1 - 0.5^{\frac{1}{K}}\right)}{\ln P_F - \ln K}$$

(A.59)

$$= \frac{-NK \ln^2 \left(1 - 0.5^{\frac{1}{K}}\right) \left(0.5^{1-\frac{1}{K}} - 0.5\right)}{\ln P_F - \ln K},$$

(A.60)

and the limit:

$$\lim_{K \to \infty} \frac{dP_D}{d\gamma} \frac{1}{\ln K} = N \frac{\ln 2}{2},$$

(A.61)

which is a constant, and the proof is concluded by noting that the limits of both the numerator and denominator exist at infinity.
Bibliography


