

OPTIMAL POWER ALLOCATION TECHNIQUES FOR VECTOR PARAMETER ESTIMATION WITH FISHER INFORMATION BASED OBJECTIVES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

By
Doğa Gürgünoğlu
June 2021

Optimal Power Allocation Techniques for Vector Parameter Estimation
with Fisher Information based Objectives

By Doğa Gürgünoğlu

June 2021

We certify that we have read this thesis and that in our opinion it is fully adequate,
in scope and in quality, as a thesis for the degree of Master of Science.

Sinan Gezici (Advisor)

Berkan Dölek

Tolga Mete Duman

Approved for the Graduate School of Engineering and Science:

Ezhan Karahan
Director of the Graduate School

ABSTRACT

OPTIMAL POWER ALLOCATION TECHNIQUES FOR VECTOR PARAMETER ESTIMATION WITH FISHER INFORMATION BASED OBJECTIVES

Doğa Gürgünoğlu

M.S. in Electrical and Electronics Engineering

Advisor: Sinan Gezici

June 2021

In this thesis, optimal power allocation problems are investigated for vector parameter estimation according to various Fisher information based optimality criteria. By considering a generic observation model involving a linear/nonlinear transformation of the parameter vector and an additive noise component with an arbitrary joint probability distribution, six different optimal power allocation problems are formulated based on Fisher information based objective functions. Various closed-form solutions are derived for the proposed problems using optimization theoretic approaches for the cases in which the transformation acting on the parameter vector is linear. Also, the results are extended to cases in which nuisance parameters exist in the system model, and to the cases when the transformation acting on the parameter vector is nonlinear. It is shown that the proposed methods are also valid for the provided extensions under certain conditions. Numerical examples are presented to investigate performance of the proposed power allocation strategies, and it is shown that they provide significant performance gains over the equal power allocation strategy.

Keywords: Cramér-Rao lower bound, estimation, Fisher information, power adaptation.

ÖZET

FISHER BİLGİSİ TABANLI OBJEKTİFLERLE VEKTÖR PARAMETRE KESTİRİMİ İÇİN OPTİMAL GÜÇ DAĞITIM TEKNİKLERİ

Doğa Gürgünoğlu

Elektrik Elektronik Mühendisliği, Yüksek Lisans

Tez Danışmanı: Sinan Gezici

Haziran 2021

Bu tezde vektör parametre kestirimi için optimal güç adaptasyonu problemleri, farklı Fisher bilgisi temelli optimallik kriterlerine göre incelenmektedir. Parametre vektörünün doğrusal/doğrusal olmayan bir dönüşümünü ve herhangi bir olasılık dağılımına sahip olabilecek toplanır bir gürültü bileşenini içeren genel bir gözlem modeli benimsenip, farklı Fisher bilgisi temelli amaç fonksiyonları kullanılarak altı farklı optimal güç paylaşırma problemi formüle edilmektedir. Optimizasyon kuramsal yaklaşımlar ile formüle edilen problemler için sistem dönüşümünün doğrusal olduğu durumlar için çeşitli kapalı yapıda çözümler elde edilmektedir. Aynı zamanda çözümler, doğrusal olmayan dönüşümlere ve önemsiz parametrelerin var olduğu durumlara uyarlanmaktadır ve önerilen yöntemlerin belirli koşullar altında bu durumlarda da uygulanabileceği gösterilmektedir. Önerilen güç paylaşım stratejilerinin performansının incelenmesi için sayısal örnekler sunulmakta ve güç uyarlamasının eşit güç dağılımına kıyasla önemli performans kazançları getirdiği gösterilmektedir.

Anahtar sözcükler: Cramér-Rao alt sınırı, kestirim, Fisher bilgisi, güç uyarlaması.

Acknowledgement

I would like to express my sincere gratitude to my supervisor Prof. Sinan Gezici for his continuous support even at the most difficult times. Thanks to him, I learned how to construct a research question, formulate problems, use rigorous approaches to solve the formulated problems and how to document academic work systematically, clearly and concisely. His systematic approach and communication skills made him an idol for me.

Secondly, I would like to express my sincere thanks to my colleagues Ayşen Balbaş Bulu, Taner Kolçak, Mehmet Akif Ayas, Ege Engin, Mustafa Levent Eksert, Yiğit Uğur, Canan Göbekli, Serkan Buhurcu, Anıl Şeker, Tamer GÜDÜ and Cansu Şen at Aselsan for their continuous support and expertise. I could not undertake such an intensive process without them.

I also would like to thank Alper Umut Uçar, Süleyman Taylan Topaloğlu, Can Soygür and Onur Özder, Ünal Ege Gaznepoğlu and Emir Ceyani for their friendship and support. The last couple of years would be very dark and unbearable without them.

I would like to express my endless gratitude for my parents Serap and Feridun Gürgünoğlu for their support throughout my life. It takes a lot of effort and patience to raise a child. They have done everything for me, they guided and supported me through each and every step. In every step, there lies a meaning. In every failure, a seed for success is planted. I learned my first principles from them, and I could not have done anything without those principles.

This work is supported by TÜBİTAK within the scope of 2210/A scholarship.

Contents

1	Introduction	1
1.1	Related Work	2
1.2	Our Contribution	4
1.3	Organization of the Thesis	5
2	System Model	6
3	Optimal Power Allocation for Vector Parameter Estimation	8
3.1	Average Mean Squared Error Criterion	8
3.2	Shannon Information Criterion	12
3.3	Worst-Case Error Variance Criterion	15
3.4	Worst-Case Coordinate Error Variance Criterion	18
3.5	Average Fisher Information Criterion	19
3.6	Worst-Case Coordinate Fisher Information Criterion	22

4	Extensions	24
4.1	Presence of Nuisance Parameters	24
4.2	Extension to Nonlinear Model	26
5	Numerical Results	27
5.1	Results for Average MSE Criterion	29
5.2	Results for Shannon Information Criterion	30
5.3	Results for Worst-Case Error Variance Criterion	31
5.4	Results for Worst-Case Coordinate Error Variance Criterion	33
5.5	Results for Average Fisher Information Criterion	34
5.6	Results for Worst Case Coordinate Fisher Information	35
6	Conclusion and Future Work	37

List of Figures

5.1	CRLB on the average MSE versus k for the equal and optimal power allocation strategies.	30
5.2	Shannon information versus k for the optimal (equal) power allocation strategy.	31
5.3	Maximum eigenvalue of the CRLB (inverse FIM) versus k for the optimal and equal power allocation strategies.	32
5.4	The largest diagonal entry of the CRLB (worst-case coordinate CRLB) versus k for the optimal and equal power allocation strategies.	33
5.5	Average Fisher information versus k for the optimal and equal power allocation strategies.	34
5.6	The minimum diagonal entry of the FIM (worst-case coordinate Fisher information) versus k for the optimal and equal power allocation strategies.	35

Chapter 1

Introduction

In vector parameter estimation, the aim is to design an optimal estimator for a number of unknown parameters based on a set of observations. The design of an optimal estimator commonly involves the calculation of posterior distributions or likelihood functions based on the statistical relation between the observation and the parameter vector. If the prior distribution of parameters is known, the Bayesian approach can be adopted and estimators such as the minimum mean squared error (MMSE) estimator, the minimum mean absolute error (MMAE) estimator, or the maximum *a posteriori* probability (MAP) estimator can be derived based on the posterior distribution, i.e., the probability distribution of the parameter vector given the observation [1]. On the other hand, in the absence of prior information, parameters can be modeled as a deterministic unknown vector and estimators such as the minimum variance unbiased estimator (MVUE), the maximum likelihood (ML) estimator, or the best linear unbiased estimator (BLUE) can be employed for vector parameter estimation [2].

Performance of the aforementioned estimators depends on system parameters such as noise variance and linear/nonlinear transformations acting on the parameter vector, and it is usually challenging to find exact and closed-form expressions for estimation errors of the corresponding estimators. Therefore, in order to assess estimation performance, various theoretical bounds such as the Cramer-Rao

lower bound (CRLB), Ziv-Zakai lower bound (ZZLB), and Barankin-type bounds are used as gold standards [3]. Such bounds are mainly determined by the statistics of the observation, which depends on system parameters. This means that for a given system model, estimation performance can be improved only to a certain extent by using an optimal estimator. In order to realize further improvements in estimation performance, the effects of the system on the parameter vector should be adapted. One common way of achieving such an improvement is to perform power adaptation, i.e., transmitting different components of the parameter vector with different power levels [4]. Since the total available power is usually limited [5], the problem of power adaptation arises as a constrained optimization problem. In this manuscript, the aim is to develop optimal power allocation strategies for vector parameter estimation in the absence of prior information by using Fisher information based optimality criteria [2, Section IV.E.1], [6], [7, Section 9.2.1].

1.1 Related Work

Power adaptation and in general resource allocation have been considered for various estimation problems in the literature. For example, in wireless sensor networks (WSNs), the problem of optimal resource allocation for vector parameter estimation with respect to various performance metrics is the main focus in many studies. In [8], the optimal transmit power allocation and quantization rate allocation schemes are investigated to minimize the average mean squared error (MSE). In [9], the optimal power allocation strategy that minimizes the ℓ_2 -norm of the transmit power vector is derived under a maximum variance constraint for the best linear unbiased estimator. In addition, the optimal codebook is computed via the Lloyd algorithm when the channel state information (CSI) is limited, which is usually the case for large WSNs. In [4], estimation of an unknown Gaussian random vector with known mean vector and covariance matrix is considered in a WSN setting, where the fusion center uses the linear MMSE (LMMSE) estimator to estimate the parameter vector based on sensor observations, which are fading channel impaired and noise corrupted versions of the

transmitted parameter vector. An upper bound on the MSE is minimized by first computing the optimal bit allocation to minimize the MSE distortion. Then, the optimal power allocation strategy is computed to minimize the channel errors. In [10], optimal power allocation for vector parameter estimation is investigated with the aim of maximizing the average Bayesian Fisher information between the random parameter vector and the observation vector. In [11], optimal power allocation schemes for LMMSE estimation are derived by taking channel estimation errors into account. In [12–21], the optimal power allocation problem is considered for position estimation in wireless localization and radar systems. In [14], the transmit power allocation problem is formulated as a semidefinite program by using the squared position error bound as the objective function. In [18], the total transmit power is minimized by imposing a constraint on the CRLB for target localization in a distributed multiple-radar system. In addition, the dual problem of CRLB minimization for a predefined total power budget is considered.

It is noted that theoretical lower bounds for estimation error are commonly used in the literature to define optimality criteria for developing power adaptation strategies in estimation problems [10], [12–22]. In the absence of prior information, lower bounds generated from the Fisher information matrix (FIM) are usually adopted due to their practicality. As the most widely used bound, the CRLB is obtained as the inverse of the FIM and specifies a lower limit on the covariance matrix of any unbiased estimator with respect to the positive semidefinite cone. Various scalarizations of the FIM are employed in the literature [18, 22]. In particular, the log-determinant of the FIM, the maximum (minimum) eigenvalue of the CRLB (FIM), the maximum diagonal entry of the CRLB, the trace of the FIM, and the minimum diagonal entry of the FIM are utilized for quantifying estimation performance from various perspectives such as estimation robustness and probabilistic confinement of estimator error [7, Section 9.2.1], [23–28]. In this manuscript, the power adaptation problem for vector parameter estimation is considered according to such Fisher information based optimality criteria and the corresponding optimal strategies are characterized.

1.2 Our Contribution

Although there exist a multitude of studies on power allocation for various estimation problems in the literature, a general investigation of the optimal power allocation problem for vector parameter estimation according to various Fisher information based criteria has not been available. In particular, we consider a generic additive noise model, where the observation vector is a linear or nonlinear function of the parameter vector corrupted by additive noise with an arbitrary probability distribution [29]. Based on this generic model, we first present the FIM in terms of the system parameters, including the power allocation parameters. Then, we formulate optimal power allocation problems according to six different estimation performance criteria based on the FIM, and derive various closed-form solutions [29]. We also extend our results to cases in which nuisance parameters exist in the problem [29]. The main contributions and novelty of this thesis can be summarized as follows:

- According to various Fisher information based optimality criteria, we propose optimal power allocation problems for vector parameter estimation by considering a generic system model, where the parameter vector is processed by any linear/nonlinear transformation and corrupted by additive noise with a generic probability distribution.
- Based on optimization theoretic approaches, we provide various closed-form solutions for the proposed power allocation problems.
- We show that the proposed optimal power allocation strategies are also valid for nonlinear system models under certain conditions and in the presence of nuisance parameters.

In addition, we provide numerical examples to illustrate the performance of the proposed strategies and compare them with the equal power allocation strategy. It should be noted that providing closed-form solutions for optimal power allocation is important for real-time applications due to delay and computational complexity requirements.

1.3 Organization of the Thesis

The rest of the thesis is organized as follows: The system model is presented in Section 2 and optimal power allocation strategies are derived in Section 3. In Section 4, extensions to nonlinear models and presence of nuisance parameters are considered. Numerical results are provided in Section 5 followed by the concluding remarks in Section 6.



Chapter 2

System Model

Consider the following linear¹ model relating a vector of unknown deterministic parameters $\boldsymbol{\theta} = [\theta_1, \dots, \theta_k]^T \in \mathbb{R}^k$ with their measurements $\mathbf{X} \in \mathbb{R}^n$:

$$\mathbf{X} = \mathbf{F}^T \mathbf{P} \boldsymbol{\theta} + \mathbf{N} \quad (2.1)$$

In (2.1), $\mathbf{F}$ is a $k \times n$ real matrix with full row rank ($k \leq n$) that is assumed to be known, $\mathbf{N} \in \mathbb{R}^n$ is the additive noise vector with a joint probability density function $f_{\mathbf{N}}(\cdot)$, which is independent of $\boldsymbol{\theta}$, and $\mathbf{P}$ is a $k \times k$ diagonal power allocation matrix (to be optimized) expressed as

$$\mathbf{P} = \begin{bmatrix} \sqrt{p_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{p_k} \end{bmatrix} \quad (2.2)$$

subject to the total power constraint

$$\sum_{i=1}^k p_i \leq P_{\Sigma} \quad (2.3)$$

where p_i denotes the power allocated to the parameter θ_i and P_{Σ} denotes the (available) total power. For the linear model given in (2.1), the FIM of the

¹Extensions to nonlinear models are presented in Section 4.

measurement vector $\mathbf{X}$ with respect to the parameter vector $\boldsymbol{\theta}$ is obtained as [30, Lemma 5].

$$\mathbf{I}(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{P} \mathbf{I}(\mathbf{N}) \mathbf{F}^T \mathbf{P}, \quad (2.4)$$

where $\mathbf{P} = \mathbf{P}^T$ is employed, and $\mathbf{I}(\mathbf{N})$ is a special form of the FIM, namely the FIM of the random vector $\mathbf{N}$ with respect to a translation parameter $\boldsymbol{\phi}$ [30, Equation 8], defined as

$$\begin{aligned} \mathbf{I}(\mathbf{N}) &= \mathbf{I}(\boldsymbol{\phi} + \mathbf{N}; \boldsymbol{\phi}) \\ &= \int \frac{1}{f(\mathbf{n})} \left(\frac{\partial f(\mathbf{n})}{\partial \mathbf{n}} \right) \left(\frac{\partial f(\mathbf{n})}{\partial \mathbf{n}} \right)^T d\mathbf{n} \end{aligned} \quad (2.5)$$

It is noted that the FIM under translation is a function of only the probability density function (pdf) of the random vector $\mathbf{N}$, and consequently, $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})$ in (2.4) does not depend on the parameter vector $\boldsymbol{\theta}$.

In the following, we provide closed-form solutions for optimal power allocation problems by considering various estimation accuracy criteria based on the FIM in (2.4).

Chapter 3

Optimal Power Allocation for Vector Parameter Estimation

3.1 Average Mean Squared Error Criterion

The inverse of the FIM, known as the Cramer-Rao lower bound (CRLB) provides a lower bound on the MSE of any unbiased estimator $\hat{\boldsymbol{\theta}}(\mathbf{X})$ via the following matrix inequality [2]:

$$\text{Cov}(\hat{\boldsymbol{\theta}}(\mathbf{X})) \geq \mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta}) \quad (3.1)$$

where $\text{Cov}(\hat{\boldsymbol{\theta}}(\mathbf{X})) = \text{E}[(\hat{\boldsymbol{\theta}}(\mathbf{X}) - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}(\mathbf{X}) - \boldsymbol{\theta})^T]$ due to the unbiasedness and the expectation is taken with respect to the pdf of $\mathbf{X}$ given $\boldsymbol{\theta}$. Consequently, the lower bound on the average MSE¹ of the vector parameter can be stated as

$$\text{E}[\|\hat{\boldsymbol{\theta}}(\mathbf{X}) - \boldsymbol{\theta}\|^2] \geq \text{tr}\{\mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta})\} \quad (3.2)$$

Consideration of the lower bound in (3.2) as a performance metric in optimal design problems is referred as the A-optimality criterion in the literature [7, 23].

The optimal power allocation problem that minimizes the lower bound on the

¹Dividing both sides of the inequality in (3.2) by k , a lower bound on the average MSE is obtained. Therefore, minimizing the trace expression in (3.2) is equivalent to minimizing the lower bound on the average MSE.

average MSE subject to a sum-power constraint can be formulated as

$$\begin{aligned}
& \min_{\{p_i\}_{i=1}^k} \quad \text{tr}\{\mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta})\} \\
& \text{s.t.} \quad \sum_{i=1}^k p_i \leq P_\Sigma \\
& \quad p_i \geq 0, \quad i = 1, 2, \dots, k
\end{aligned} \tag{3.3}$$

For the convenience of notation, two system dependent matrices can be defined as

$$\begin{aligned}
\mathbf{J} &\triangleq \mathbf{F}\mathbf{I}(\mathbf{N})\mathbf{F}^T \\
\mathbf{A} &\triangleq \mathbf{J}^{-1}
\end{aligned} \tag{3.4}$$

From (2.4) and (3.4), the FIM with respect to the parameter vector $\boldsymbol{\theta}$ and the corresponding CRLB are expressed respectively as $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{P}\mathbf{J}\mathbf{P}$ and $\mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-1}$. Then, the objective function in (3.3) can be written in terms of the power allocation coefficients and the diagonal entries of $\mathbf{A}$ as

$$\begin{aligned}
\text{tr}\{\mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta})\} &= \text{tr}\{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-1}\} \\
&= \text{tr}\{(\mathbf{P}^{-1})^2\mathbf{A}\} \\
&= \sum_{i=1}^k \frac{a_{ii}}{p_i}
\end{aligned} \tag{3.5}$$

where a_{ii} denotes the i th diagonal entry of $\mathbf{A}$. From (2.5) and (3.4), it is noted that $\mathbf{A}$ is positive semi-definite; hence, $a_{ii} \geq 0$ for $i = 1, \dots, k$. To have a non-trivial problem, it is assumed that at least one element of $\{a_{ii}\}_{i=1}^k$ is strictly positive.

As the objective function is convex (see (3.5)) and the constraints are linear, the problem in (3.3) is a convex optimization problem. In addition, Slater's condition holds [24]. Therefore, Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for optimality. From (3.5), the Lagrangian for (3.3) is expressed as

$$\mathcal{L}(\{p_i\}_{i=1}^k, \{v_i\}_{i=1}^{k+1}) = \sum_{i=1}^k \frac{a_{ii}}{p_i} + v_1 \left(\sum_{i=1}^k p_i - P_\Sigma \right) - \sum_{i=1}^k v_{i+1} p_i \tag{3.6}$$

where $v_1, \dots, v_{k+1}$ are the dual variables. Then, KKT conditions for optimality are obtained as follows [24]:

- *Primal Feasibility:* The optimal power allocation strategy $\{p_i^*\}_{i=1}^k$ must satisfy the constraints $\sum_{i=1}^k p_i^* \leq P_\Sigma$ and $p_i^* \geq 0, \forall i \in \{1, \dots, k\}$.
- *Dual Feasibility:* The dual variables must be non-negative, i.e., $v_i^* \geq 0$ for $i = 1, \dots, k+1$.
- *Stationarity:* The derivatives of the Lagrangian in (3.6) with respect to p_i must be equal to zero at $p_i = p_i^*$ for $i = 1, \dots, k$. That is,

$$\left. \frac{\partial \mathcal{L}}{\partial p_i} \right|_{p_i=p_i^*} = -\frac{a_{ii}}{(p_i^*)^2} + v_1^* - v_{i+1}^* = 0 \quad (3.7)$$

for $i = 1, \dots, k$.

- *Complementary Slackness:* At the optimal solution, the following conditions hold:

$$v_1^* \left(\sum_{i=1}^k p_i^* - P_\Sigma \right) = 0 \quad (3.8)$$

$$v_{i+1}^* p_i^* = 0, \quad i = 1, \dots, k \quad (3.9)$$

For the condition in (3.8), the case of $v_1^* = 0$ is not possible since the derivative in (3.7) could be set to zero only for $p_i^* \rightarrow \infty$ in that case (for some positive a_{ii}), which would violate the primal feasibility condition. Therefore, (3.8) implies that $v_1^* > 0$ and

$$\sum_{i=1}^k p_i^* = P_\Sigma \quad (3.10)$$

That is, full-power utilization is required for optimality². Then, two cases are investigated depending on the values of a_{ii} 's. Let $\mathcal{A}_z$ and $\mathcal{A}_p$ denote the sets of indices i for which a_{ii} 's are zero and positive, respectively. That is, $\mathcal{A}_z = \{i \in \{1, \dots, k\} \mid a_{ii} = 0\}$ and $\mathcal{A}_p = \{i \in \{1, \dots, k\} \mid a_{ii} > 0\}$.

Case 1: Consider an index i such that $i \in \mathcal{A}_z$. Suppose that $p_i^* > 0$. Then, (3.9) implies that $v_{i+1}^* = 0$ and the expression in (3.7) becomes equal to v_1^* .

²This fact can also be seen by noting that the objective function in (3.5) is a decreasing function of p_i 's.

However, $v_1^* > 0$ as discussed before, which leads to a contradiction (i.e., the stationary condition could not be satisfied). Hence, it is concluded that

$$p_i^* = 0 \text{ for } i \in \mathcal{A}_z. \quad (3.11)$$

Case 2: Consider an index i such that $i \in \mathcal{A}_p$. In that case, it can be concluded from (3.7)–(3.9) that $p_i^* > 0$ and $v_{i+1}^* = 0$ for $i \in \mathcal{A}_p$. Then, (3.7) leads to

$$p_i^* = \sqrt{\frac{a_{ii}}{v_1^*}} \text{ for } i \in \mathcal{A}_p \quad (3.12)$$

From (3.10), a relation for v_1^* can be obtained as

$$\sum_{j=1}^k p_j^* = \sum_{i \in \mathcal{A}_p} p_i^* = \sum_{j \in \mathcal{A}_p} \sqrt{\frac{a_{jj}}{v_1^*}} = P_\Sigma, \quad (3.13)$$

which yields

$$\frac{1}{\sqrt{v_1^*}} = \frac{P_\Sigma}{\sum_{j \in \mathcal{A}_p} \sqrt{a_{jj}}} = \frac{P_\Sigma}{\sum_{j=1}^k \sqrt{a_{jj}}} \quad (3.14)$$

Based on (3.11), (3.12) and (3.14), the optimal power allocation strategy to minimize the average MSE in (3.3) is specified as follows:

$$p_i^* = \frac{P_\Sigma \sqrt{a_{ii}}}{\sum_{j=1}^k \sqrt{a_{jj}}}, \quad i = 1, \dots, k \quad (3.15)$$

Hence, a closed-form solution to the problem in (3.3) is obtained.

3.2 Shannon Information Criterion

An alternative criterion for estimation accuracy is to maximize the log-determinant of the FIM, i.e.,

$$\mathbf{I}_S(\mathbf{X}; \boldsymbol{\theta}) = \log \det \mathbf{I}(\mathbf{X}; \boldsymbol{\theta}) \quad (3.16)$$

which is associated with volume of confidence ellipsoid containing the estimation error [24, Section 7.5.2]. This criterion is known as the Shannon information criterion, and also the D-optimal design in the literature [23, 25, 26].

The optimal power allocation problem with the objective of maximizing the Shannon information under the sum-power constraint can be expressed as

$$\begin{aligned} \max_{\{p_i\}_{i=1}^k} \quad & \log \det \mathbf{I}(\mathbf{X}; \boldsymbol{\theta}) \\ \text{s.t.} \quad & \sum_{i=1}^k p_i \leq P_\Sigma \\ & p_i \geq 0, \quad i = 1, \dots, k \end{aligned} \quad (3.17)$$

The problem in (3.17) involves the maximization of a concave function, and the feasible region has an interior point; hence, Slater's condition is satisfied. Consequently, KKT conditions are necessary and sufficient for optimality. In order to find the optimal solution, the Shannon information can be expressed in terms of the known matrices. From (2.4) and (3.4), the Shannon information can be written as

$$\begin{aligned} \log \det \mathbf{I}(\mathbf{X}; \boldsymbol{\theta}) &= \log \det \mathbf{PJP} \\ &= 2 \log \det \mathbf{P} + \log \det \mathbf{J} \\ &= 2 \sum_{i=1}^k \log p_i + \log \det \mathbf{J} \end{aligned} \quad (3.18)$$

As seen in (3.18), the Shannon information separates into a power allocation dependent component and a system dependent component, the latter being constant for a fixed $\mathbf{F}$ and $\mathbf{I}(\mathbf{N})$. Therefore, it suffices to consider $\sum_{i=1}^k \log p_i$ in order to maximize the Shannon information. Therefore, the problem in (3.17) reduces

to

$$\begin{aligned}
& \max_{\{p_i\}_{i=1}^k} \quad \sum_{i=1}^k \log p_i \\
& \text{s.t.} \quad \sum_{i=1}^k p_i \leq P_\Sigma \\
& \quad p_i \geq 0, \quad i = 1, \dots, k
\end{aligned} \tag{3.19}$$

which is a convex optimization problem.³

Considering the minimization of the negative of the objective function in (3.19), the Lagrangian can be expressed as

$$\mathcal{L}(\{p_i\}_{i=1}^k, \{v_i\}_{i=1}^{k+1}) = - \sum_{i=1}^k \log p_i + v_1 \left(\sum_{i=1}^k p_i - P_\Sigma \right) - \sum_{i=2}^{k+1} v_i p_{i-1} \tag{3.20}$$

The KKT conditions are the same as those in Section 3.1 except that the stationarity condition becomes

$$\left. \frac{\partial \mathcal{L}}{\partial p_i} \right|_{p_i=p_i^*} = -\frac{1}{p_i^*} + v_1^* - v_{i+1}^* = 0 \tag{3.21}$$

for $i = 1, \dots, k$. From the complementary slackness condition in (3.8) and the stationarity condition in (3.21), it is concluded that $v_1^* > 0$ and $\sum_{i=1}^k p_i^* = P_\Sigma$ must hold for not violating the primal feasibility. In addition, the fact that $v_1^* > 0$ together with the conditions in (3.9) and (3.21) imply that $v_{i+1}^* = 0$ for $i = 1, \dots, k$. Hence, (3.21) becomes $p_i^* = 1/v_1^*$ for $i = 1, \dots, k$. Together with $\sum_{i=1}^k p_i^* = P_\Sigma$, the optimal power allocation strategy according to the Shannon information criterion is calculated as

$$p_i^* = \frac{P_\Sigma}{k}, \quad i = 1, \dots, k \tag{3.22}$$

Corresponding to the optimal strategy, the maximum Shannon information is achieved as

$$\mathbf{I}_S^*(\mathbf{X}; \boldsymbol{\theta}) = 2k \log \left(\frac{P_\Sigma}{k} \right) + \log \det \mathbf{J} \tag{3.23}$$

It is noted that the optimal power allocation strategy in (3.22) corresponds to allocating equal power to all the parameters at the sum-power limit. Since the

³This problem is equivalent to maximizing the product of nonnegative numbers whose sum is constant

optimal power allocation strategy does not depend on $\mathbf{J}$ or $\mathbf{A}$, the system parameters do not affect the optimal solution according to the Shannon information criterion.



3.3 Worst-Case Error Variance Criterion

The worst-case error variance criterion is a measure of robustness rather than average estimation accuracy and is associated with the maximum eigenvalue of the CRLB [24], [27]. In order to reduce the worst-case error variance, the maximum eigenvalue of the CRLB can be minimized. Optimality according to this criterion is also known as E-optimality, where the minimum diameter of the FIM is maximized [23, 28]. When variances vary significantly, the confidence ellipsoid can have very different diameters along different dimensions; hence, the log-volume minimization approach in the D-optimal design can be misleading [23, 28].

The optimal power allocation strategy that minimizes the maximum eigenvalue of the CRLB corresponds to maximizing the minimum eigenvalue of the FIM. Hence, the following problem can be considered:

$$\begin{aligned}
 & \max_{\{p_i\}_{i=1}^k} \quad \lambda_{\min}\{\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})\} \\
 & \text{s.t.} \quad \sum_{i=1}^k p_i \leq P_{\Sigma} \\
 & \quad p_i \geq 0, \quad i = 1, \dots, k
 \end{aligned} \tag{3.24}$$

From (2.4) and (3.4), $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})$ in (3.24) can be expressed as $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{P}\mathbf{J}\mathbf{P}$, where $\mathbf{J}$ is positive semi-definite and $\mathbf{P}$ is diagonal (see (2.2)). It can be shown that the eigenvalues of $\mathbf{P}\mathbf{J}\mathbf{P}$ are the same as those of $\mathbf{P}^2\mathbf{J}$ based on their characteristic equations. However in general, there is not a closed-form relationship between the eigenvalues of $\mathbf{P}^2$ and $\mathbf{J}$ and the eigenvalues of their product. Therefore, it is challenging to obtain a closed-form solution to (3.24). One way to solve (3.24) is to apply global optimization tools such as particle swarm optimization (PSO) or the multistart algorithm [31]. This approach is adopted in Section 5 to obtain the solution of (3.24).

To perform further investigations on the problem in (3.24), we can derive a bound on the objective function in (3.24). To that aim, the following lemma can be utilized to provide bounds for the eigenvalues of the FIM, $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})$.

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{A}^T = \mathbf{A}$, $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T) = n$ and $\mathbf{B} \in \mathbb{R}^{n \times n}$. Let $\{\nu_i\}_{i=1}^n$, $\{\mu_i\}_{i=1}^n$ and $\{\gamma_i\}_{i=1}^n$ denote, respectively, the eigenvalues of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{A} + \mathbf{B}$ increasing in the absolute value. Then,

$$|\nu_1||\mu_1| \leq |\gamma_1| \leq \dots \leq |\gamma_n| \leq |\nu_n||\mu_n| \quad (3.25)$$

For the Euclidean matrix norm, it is known that $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F$. In addition, the norm of a symmetric matrix is equal to its spectral radius, i.e., $\|\mathbf{A}\|_2 = \max_j |\nu_j|$ and $\|\mathbf{B}\|_2 = \max_j |\mu_j|$. Therefore, $\|\mathbf{A} + \mathbf{B}\|_2 = \|\mathbf{A} + \mathbf{B}\|_F \leq \|\mathbf{A}\|_F + \|\mathbf{B}\|_F = |\nu_n| + |\mu_n|$. Since $\mathbf{A} + \mathbf{B}$ is the product of two symmetric matrices, it is also symmetric. Therefore, $\|\mathbf{A} + \mathbf{B}\|_2 = |\gamma_n|$. Hence, the upper bound on the absolute value of the maximum eigenvalue of $\mathbf{A} + \mathbf{B}$ is obtained as $|\gamma_n| \leq |\nu_n| + |\mu_n|$. Since $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{B}) = n$, the lower bound can be derived via inversion; that is, $(\mathbf{A} + \mathbf{B})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$. Therefore, $\|(\mathbf{A} + \mathbf{B})^{-1}\|_2 = \|\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\|_2 \leq \frac{1}{|\nu_1|} + \frac{1}{|\mu_1|}$. Through the same reasoning, $\|\mathbf{A}^{-1}\|_2 = \frac{1}{|\nu_1|}$. Hence, the relation of $\frac{1}{|\gamma_1|} \leq \frac{1}{|\nu_1|} + \frac{1}{|\mu_1|}$ is obtained, which yields the lower bound in (3.25).

Lemma 1 can be used to derive a lower bound on the objective function in (3.24) as follows:

$$\lambda_{\min}\{\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})\} = \lambda_{\min}\{\mathbf{P}^2 \mathbf{J}\} \geq \lambda_{\min}\{\mathbf{J}\} \min_{i \in \{1, \dots, k\}} p_i \quad (3.26)$$

where $\lambda_{\min}\{\mathbf{J}\}$ denotes the minimum eigenvalue of $\mathbf{J}$. In (3.26), the absolute value operators in (3.25) are not used since all the eigenvalues are non-negative and the eigenvalues of $\mathbf{P}^2$ are taken as $\{p_1, \dots, p_k\}$ based on (2.2).

Instead of maximizing the minimum eigenvalue of $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})$ in (3.24), consider the maximization of the lower bound on it. As noted from (3.26), the lower bound on the minimum eigenvalue of $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})$ depends on the minimum power allocated to an individual parameter. Therefore, instead of (3.24), we get the following convex optimization problem:

$$\begin{aligned} & \max_{\{p_i\}_{i=1}^k} \min_{j \in \{1, \dots, k\}} p_j \\ & \text{s.t.} \quad \sum_{i=1}^k p_i \leq P_{\Sigma} \\ & \quad p_i \geq 0, \quad i = 1, \dots, k \end{aligned} \quad (3.27)$$

The Lagrangian for (3.27) can be expressed as

$$\mathcal{L}(\{p_i\}_{i=1}^k, \{v_i\}_{i=1}^{k+1}) = - \min_{j \in \{1, \dots, k\}} p_j + v_1 \left(\sum_{i=1}^k p_i - P_\Sigma \right) - \sum_{i=1}^k v_{i+1} p_i \quad (3.28)$$

Accordingly, the KKT conditions are again the same as those in Section 3.1 except that the stationarity condition becomes

$$\left. \frac{\partial \mathcal{L}}{\partial p_i} \right|_{p_i=p_i^*} = -\mathbb{I}_{\{p_i \leq p_j, \forall j \neq i\}} + v_1^* - v_{i+1}^* = 0 \quad (3.29)$$

for $i = 1, \dots, k$, where $\mathbb{I}$ denotes the indicator function, which is equal one if the condition is satisfied and zero otherwise. From the complementary slackness condition in (3.8) and the stationarity condition in (3.29), it is concluded that $v_1^* > 0$ and $\sum_{i=1}^k p_i^* = P_\Sigma$ must hold. (Otherwise, $\mathbb{I}_{\{p_i \leq p_j, \forall j \neq i\}} = 0$ for all $i = 1, \dots, k$, which is impossible.) In addition, the fact that $v_1^* > 0$ together with the conditions in (3.9) and (3.29) imply that $v_{i+1}^* = 0$ for $i = 1, \dots, k$. Hence, (3.29) reduces to

$$\mathbb{I}_{\{p_i \leq p_j, \forall j \neq i\}} = v_1^*, \quad i = 1, \dots, k \quad (3.30)$$

The only possibility to satisfy the k equations in (3.30) is to have $p_1^* = p_2^* = \dots = p_k^*$. Therefore, $\sum_{i=1}^k p_i^* = P_\Sigma$ implies that $p_i^* = P_\Sigma/k$ for $i = 1, \dots, k$; that is, the optimal power allocation strategy to maximize (minimize) the lower (upper) *bound* on the minimum (maximum) eigenvalue of the FIM (CRLB) is the equal power allocation strategy. Consequently, the lower bound on the minimum eigenvalue of $\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})$ becomes $\lambda_{\min}\{\mathbf{J}\}P_\Sigma/k$.

Remark 1: The preceding analysis indicates that the equal power allocation strategy solves the problem of maximizing a **lower bound** on the objective function in (3.24). Hence, it does not necessarily yield the optimal power allocation strategy. (The numerical example in Section 5.3 illustrates this fact.)

3.4 Worst-Case Coordinate Error Variance Criterion

As an alternative measure of robustness, one can consider the worst-case coordinate error variance, which is bounded by the largest diagonal entry of the CRLB; i.e., $\max_{j \in \{1, \dots, K\}} [\mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta})]_{j,j}$. This criterion is referred as G-optimality [7, 23], and it has the effect of reducing the worst-case error variance as well.

From (3.5), the j th diagonal entry of the CRLB can be expressed as

$$[\mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta})]_{j,j} = \frac{a_{jj}}{p_j} \quad (3.31)$$

Therefore, the problem of minimizing the maximum diagonal entry of the CRLB can be formulated as

$$\begin{aligned} \min_{\{p_i\}_{i=1}^k} \quad & \max_{j \in \{1, \dots, k\}} \frac{a_{jj}}{p_j} \\ \text{s.t.} \quad & \sum_{i=1}^k p_i \leq P_\Sigma \\ & p_i \geq 0, \quad i = 1, \dots, k \end{aligned} \quad (3.32)$$

The problem in (3.32) is a convex optimization problem, it can be shown that the solution of (3.32) satisfies $\sum_{i=1}^k p_i^* = P_\Sigma$ and

$$\frac{a_{ii}}{p_i^*} = \alpha, \quad \forall i \in \{1, \dots, k\} \quad (3.33)$$

where α is a constant (i.e., an equalizer solution [32]). Then, parameter α in (3.33) obtained from

$$\sum_{i=1}^k p_i^* = \frac{1}{\alpha} \sum_{i=1}^k a_{ii} = P_\Sigma \quad (3.34)$$

which yields

$$\alpha = \frac{\text{tr}\{\mathbf{A}\}}{P_\Sigma}. \quad (3.35)$$

Hence, the optimal power allocation strategy is given by

$$p_i^* = \frac{P_\Sigma a_{ii}}{\text{tr}\{\mathbf{A}\}}, \quad i = 1, \dots, k \quad (3.36)$$

When the optimal power allocation strategy is employed, the diagonal entries of the CRLB are the same, and the worst-case coordinate error variance becomes $\alpha = \text{tr}\{\mathbf{A}\}/P_\Sigma$.

3.5 Average Fisher Information Criterion

When the aim is to estimate a vector of parameters, the average Fisher information indicates the overall usefulness of the observation vector to estimate the parameter vector. The informativeness of the observation vector to estimate the i th parameter corresponds to the i th diagonal entry of the FIM. Therefore, the average Fisher information is related to the trace of the FIM. Accordingly, the optimal power allocation problem for maximizing the trace of the FIM is formulated as follows:

$$\begin{aligned} \max_{\{p_i\}_{i=1}^k} \quad & \text{tr}\{\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})\} \\ \text{s.t.} \quad & \sum_{i=1}^k p_i \leq P_\Sigma \\ & p_i \geq 0, \quad i = 1, \dots, k \end{aligned} \tag{3.37}$$

From (2.4) and (3.4), the objective function in (3.37) can be rewritten in terms of the known matrices as

$$\text{tr}\{\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})\} = \text{tr}\{\mathbf{PJP}\} = \sum_{i=1}^k p_i j_{ii} \tag{3.38}$$

where $j_{ii} \triangleq [\mathbf{J}]_{i,i}$. Based on (3.38), the problem in (3.37) can be converted to a linear program (LP) by defining

$$\mathbf{p} \triangleq \text{diag}(\mathbf{PP}^T) = [p_1 \dots p_k]^T \tag{3.39}$$

$$\mathbf{j} \triangleq \text{diag}(\mathbf{J}) = [j_{11} \dots j_{kk}]^T, \tag{3.40}$$

and expressing (3.37) as

$$\begin{aligned} \max_{\mathbf{p}} \quad & \mathbf{j}^T \mathbf{p} \\ \text{s.t.} \quad & \mathbf{1}^T \mathbf{p} \leq P_\Sigma \\ & \mathbf{p} \geq 0 \end{aligned} \tag{3.41}$$

The solution of (3.41) is provided in the following proposition.

Let i^* denote the index of the maximum element of $\mathbf{j}$ in (3.40); i.e., $i^* = \arg \max_{l \in \{1, \dots, k\}} j_{ll}$. Then, the optimal power allocation strategy that maximizes

the average Fisher information under the sum-power constraint is given by $\mathbf{p}^* = [p_1^* \cdots p_k^*]^T$, where

$$p_i^* = \begin{cases} P_\Sigma, & i = i^* \\ 0, & \text{otherwise} \end{cases} \quad (3.42)$$

for $i = 1, \dots, k$. (In case of multiple maxima, the indices of any non-empty subset can be selected.) Although optimization problems in the form of (3.41) have been solved in various contexts [22, 33, 34], a proof is provided below for completeness.

Consider any feasible power allocation strategy represented by $\mathbf{p} = [p_1 \cdots p_k]^T$, which achieves the objective value of

$$\mathbf{j}^T \mathbf{p} = \sum_{i=1}^k j_{ii} p_i \quad (3.43)$$

Suppose that this strategy is changed by using a modification vector $\Delta = [\delta_1 \cdots \delta_k]^T$ with the following properties:

$$\begin{aligned} \delta_i &= \begin{cases} -\sum_{l \neq i} \delta_l, & i = i^* \\ \text{a nonpositive real number,} & \text{otherwise} \end{cases} \\ -\delta_i &\leq p_i, \quad \forall i \neq i^* \end{aligned} \quad (3.44)$$

where $i^* = \arg \max_{l \in \{1, \dots, k\}} j_{ll}$. Accordingly, the following new strategy is obtained:

$$\tilde{\mathbf{p}} = \begin{bmatrix} p_1 + \delta_1 \\ p_2 + \delta_2 \\ \vdots \\ p_{i^*} - \sum_{l \neq i^*} \delta_l \\ \vdots \\ p_k + \delta_k \end{bmatrix} \quad (3.45)$$

It can be seen that $\sum_{i=1}^k \tilde{p}_i = \sum_{i=1}^k p_i$ and $\tilde{p}_i \geq 0$ for $i = 1, \dots, k$. Hence, the new strategy is also feasible. (Essentially, the new strategy accumulates some of the power distributed among indices into a single index.) The improvement in the objective function of (3.41) achieved by the new strategy can be calculated

as follows:

$$\mathbf{j}^T(\tilde{\mathbf{p}} - \mathbf{p}) = \sum_{i \neq i^*} j_{ii} \delta_i - j_{i^* i^*} \sum_{i \neq i^*} \delta_i \quad (3.46)$$

$$= - \sum_{i \neq i^*} (j_{i^* i^*} - j_{ii}) \delta_i \geq 0 \quad (3.47)$$

where the inequality is obtained due to the facts that δ_i 's are nonpositive and $j_{i^* i^*}$ is the maximum element of $\mathbf{j}$. The maximum improvement is obtained if and only if $\delta_i = -p_i$, $\forall i \neq i^*$. Then, the optimal strategy becomes $p_i^* = 0$, $\forall i \neq i^*$, and $p_i^* = \sum_{l=1}^k p_l^*$ for $i = i^*$ (see (3.45)). Moreover, it is straightforward to show that $\sum_{l=1}^k p_l^* = P_\Sigma$ must hold via similar arguments to those in the previous sections. Therefore, $p_i^* = P_\Sigma$ for $i = i^*$. That is, the optimal power allocation strategy is to allocate all available power to the parameter corresponding to the maximum element of $\mathbf{j}$.

Proposition 1 states that in order to maximize the average FI, the whole power must be allocated to the parameter corresponding to the maximum diagonal entry of $\mathbf{J}$. When the optimal power adaptation strategy is used, the average Fisher information achieves a maximum value of $P_\Sigma \max_i j_{ii}$.

3.6 Worst-Case Coordinate Fisher Information Criterion

Depending on the system properties, the observation vector can be more informative about some parameters and less informative about the others. Consequently, the estimation performance of individual parameters can vary to some extent. Such variations can be undesirable as certain performance requirements should be satisfied for estimation of all parameters. To alleviate this effect, one approach is to maximize the minimum Fisher information contained the observation vector w.r.t. individual parameters. Such an optimization increases the robustness of estimation against accuracy variations.

The minimum Fisher information contained in the observation vector w.r.t. individual parameters is called the worst-case coordinate FI, which corresponds to the minimum diagonal entry of the FIM, that is, $\min_{i \in \{1, \dots, k\}} [\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})]_{i,i}$. Based on this objective function, the following maximization problem is defined under the sum-power constraint:

$$\begin{aligned} & \max_{\{p_i\}_{i=1}^k} \quad \min_{i \in \{1, \dots, k\}} [\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})]_{i,i} \\ & \text{s.t.} \quad \sum_{i=1}^k p_i \leq P_\Sigma \\ & \quad p_i \geq 0, \quad i = 1, \dots, k \end{aligned} \tag{3.48}$$

Based on (3.38), the objective function in (3.48) can be written as

$$\min_i [\mathbf{I}(\mathbf{X}; \boldsymbol{\theta})]_{i,i} = \min_i p_i j_{ii} \tag{3.49}$$

Then, (3.48) is observed to have a very similar form to the problem in (3.27). Hence, the same steps can be followed and it can be shown that the solution of (3.48) satisfies $\sum_{i=1}^k p_i^* = P_\Sigma$ and

$$p_i^* j_{ii} = \tilde{\alpha}, \quad \forall i \in \{1, \dots, k\} \tag{3.50}$$

where $\tilde{\alpha}$ is a constant that is specified by

$$\sum_{i=1}^k p_i^* = \sum_{i=1}^k \frac{\tilde{\alpha}}{j_{ii}} = P_\Sigma \tag{3.51}$$

From (3.51), $\tilde{\alpha}$ is obtained as

$$\tilde{\alpha} = \frac{P_{\Sigma}}{\sum_{i=1}^k \frac{1}{j_{ii}}} . \quad (3.52)$$

Therefore, the optimal power allocation strategy becomes

$$p_i^* = \frac{P_{\Sigma}}{j_{ii} \sum_{l=1}^k \frac{1}{j_{ll}}} , \quad i = 1, \dots, k \quad (3.53)$$

When the optimal power allocation strategy is used, the worst case Fisher information achieves a maximum value of $P_{\Sigma} / \sum_{l=1}^k \frac{1}{j_{ll}}$. It is noted that the optimal power allocation strategy in (3.53) equalizes the Fisher information contained in $\mathbf{X}$ w.r.t. the i th element of $\boldsymbol{\theta}$ for all $i \in \{1, \dots, k\}$; that is, $p_i^* j_{ii} = P_{\Sigma} / \sum_{l=1}^k \frac{1}{j_{ll}}$ for all $i \in \{1, \dots, k\}$.

Chapter 4

Extensions

4.1 Presence of Nuisance Parameters

In some vector parameter estimation scenarios, only a subset of parameters can be of interest for estimation purposes. Let only r out of the k parameters be relevant and the remaining $k - r$ parameters be nuisance parameters. We assume that the nuisance parameters must be transmitted with unit power and power adaptation is not available for them. Without loss of generality, we can arrange the vector of parameters to be transmitted as

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_\gamma \\ \boldsymbol{\theta}_\sigma \end{bmatrix} \quad (4.1)$$

where $\boldsymbol{\theta}_\gamma \in \mathbb{R}^r$ denotes the vector of relevant parameters, and $\boldsymbol{\theta}_\sigma \in \mathbb{R}^{k-r}$ represents the vector of nuisance parameters. Then, the power allocation matrix becomes

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_\gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k-r} \end{bmatrix} \quad (4.2)$$

where $\mathbf{I}_{k-r}$ denotes the $(k - r) \times (k - r)$ identity matrix. Under the same system model, matrix $\mathbf{J}$ defined in (3.4) can be expressed as

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_\gamma & \mathbf{B} \\ \mathbf{B}^T & \mathbf{J}_\sigma \end{bmatrix} \quad (4.3)$$

where $\mathbf{J}_\gamma \in \mathbb{R}^{r \times r}$ and $\mathbf{J}_\sigma \in \mathbb{R}^{(k-r) \times (k-r)}$ are the components of $\mathbf{J}$ corresponding to the parameters of interest and the nuisance parameters, respectively, and matrix $\mathbf{B} \in \mathbb{R}^{r \times (k-r)}$ and its transpose are the cross-terms. Similarly, matrix $\mathbf{A}$ in (3.4), i.e., $\mathbf{A} \triangleq \mathbf{J}^{-1}$, can also be arranged as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_\gamma & \mathbf{C} \\ \mathbf{C}^T & \mathbf{A}_\sigma \end{bmatrix} \quad (4.4)$$

where $\mathbf{C} \in \mathbb{R}^{r \times (k-r)}$ and its transpose are the cross-terms. Based on (4.2), (4.3), and (4.4), the FIM and the CRLB can be expressed as

$$\mathbf{I}(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{PJP} = \begin{bmatrix} \mathbf{P}_\gamma \mathbf{J}_\gamma \mathbf{P}_\gamma & \mathbf{P}_\gamma \mathbf{B} \\ \mathbf{B}^T \mathbf{P}_\gamma & \mathbf{J}_\sigma \end{bmatrix} \quad (4.5)$$

$$\mathbf{I}^{-1}(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}^{-1} = \begin{bmatrix} \mathbf{P}_\gamma^{-1} \mathbf{A}_\gamma \mathbf{P}_\gamma^{-1} & \mathbf{P}_\gamma^{-1} \mathbf{C} \\ \mathbf{C}^T \mathbf{P}_\gamma^{-1} & \mathbf{A}_\sigma \end{bmatrix} \quad (4.6)$$

The related terms of the FIM and the CRLB are the ones involving only the parameters of interest. In this setting, only the first r rows and the first r columns are taken into account; that is,

$$\mathbf{I}_\gamma(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{P}_\gamma \mathbf{J}_\gamma \mathbf{P}_\gamma \quad (4.7)$$

$$\mathbf{I}_\gamma^{-1}(\mathbf{X}; \boldsymbol{\theta}) = \mathbf{P}_\gamma^{-1} \mathbf{A}_\gamma \mathbf{P}_\gamma^{-1} \quad (4.8)$$

where $\mathbf{A}_\gamma = (\mathbf{J}_\gamma - \mathbf{B} \mathbf{J}_\sigma^{-1} \mathbf{B}^T)^{-1}$ [35]. As seen from above, the power allocation strategies developed in Section 3 (which are developed in the absence of nuisance parameters) can also be used in this case.

4.2 Extension to Nonlinear Model

In some practical applications, the linear system model in (2.1) may not be valid, and the parameter vector, after power adaptation, can be processed by a nonlinear transformation $\mathbf{f}(\cdot)$ as follows:

$$\mathbf{X} = \mathbf{f}(\mathbf{P}\boldsymbol{\theta}) + \mathbf{N} \quad (4.9)$$

In this case, the FIM w.r.t. parameter $\boldsymbol{\theta}$ can be expressed in the same form as (2.4) after replacing $\mathbf{F}$ with the Jacobian of the vector valued function $\mathbf{f}(\cdot)$ [30, Lemma 4]. More explicitly, let $\boldsymbol{\phi} \triangleq \mathbf{P}\boldsymbol{\theta}$ in (4.9), and the Jacobian of $\mathbf{f}(\boldsymbol{\phi})$ w.r.t. its argument $\boldsymbol{\phi}$ is given as

$$\mathbf{F} \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial \phi_1} & \frac{\partial f_2}{\partial \phi_1} & \cdots & \frac{\partial f_n}{\partial \phi_1} \\ \frac{\partial f_1}{\partial \phi_2} & \frac{\partial f_2}{\partial \phi_2} & \cdots & \frac{\partial f_n}{\partial \phi_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial \phi_k} & \frac{\partial f_2}{\partial \phi_k} & \cdots & \frac{\partial f_n}{\partial \phi_k} \end{bmatrix} \quad (4.10)$$

If $\mathbf{f}(\cdot)$ is continuously differentiable w.r.t. $\boldsymbol{\phi}$ and $\mathbf{F}$ in (4.10) does not depend on p_i 's for $i = 1, \dots, k$, $\mathbf{F}$ in (4.10) can be substituted into (2.4), and the developed techniques can be employed without further modification for power adaptation in the presence of a nonlinear system model, as well. If $\mathbf{F}$ depends on p_i 's, (4.10) is still valid; however, the objective functions should be modified accordingly, leading to possibly nonconvex optimization problems. In that case, numerical methods can be employed. On the other hand, if $\mathbf{f}(\cdot)$ is not continuously differentiable w.r.t. $\boldsymbol{\phi}$, further analysis is required and new techniques should be developed.

Chapter 5

Numerical Results

In this section, we provide numerical examples for the optimal power allocation strategies in Section 3. In all cases, the equal power allocation strategy is also implemented for comparison purposes. The noise is modeled as a zero-mean Gaussian random vector with independent components; that is, $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, where $\Sigma = \text{Diag}(\sigma_1^2, \dots, \sigma_n^2)$. For this noise model, the FIM of $\mathbf{N}$ in (2.5) is obtained as

$$\mathbf{I}(\mathbf{N}) = \Sigma^{-1} = \text{Diag}(1/\sigma_1^2, \dots, 1/\sigma_n^2) \quad (5.1)$$

In the simulations, σ_i^2 's are set to $\sigma_i^2 = 10^{-7+3(i-1)/(n-1)}$ for $i = 1, \dots, n$. The dimension of the parameter vector, k , is varied between 2 and 30, and the dimension of the observation vector, n , is taken to be equal to the number of parameters, i.e., $k = n$. Also, for matrix $\mathbf{F}$ in (2.1), two different scenarios are considered. In the first scenario, $\mathbf{F} = \mathbf{F}_1$, where $\mathbf{F}_1$ is the $k \times k$ identity matrix ($k = n$), that is, $\mathbf{F}_1 = \mathbf{I}_{k \times k}$. In this scenario, we can observe the effects of power adaptation on the estimation performance when the main source of error is additive noise. In the second scenario, $\mathbf{F} = \mathbf{F}_2$, which is specified as

$$\mathbf{F}_2 = \mathbf{I}_{k \times k} + \kappa \mathbf{V}^T \quad (5.2)$$

with

$$\begin{aligned}\kappa &= \frac{\|\mathbf{I}_{k \times k}\|_F}{\|\mathbf{V}\|_F} \\ \mathbf{V} &= \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 + \epsilon & (1 + \epsilon)^2 & \dots & (1 + \epsilon)^{k-1} \\ 1 & 1 + 2\epsilon & (1 + 2\epsilon)^2 & \dots & (1 + 2\epsilon)^{k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1.5 & 1.5^2 & \dots & 1.5^{k-1} \end{bmatrix} \\ \epsilon &= \frac{0.5}{k-1}\end{aligned}\tag{5.3}$$

That is, $\mathbf{F}_2$ is the sum of the $k \times k$ identity matrix and the transpose of a normalized $k \times k$ Vandermonde matrix, where the normalization factor κ makes sure that the Frobenius norms of the added matrices are equal. In this scenario, the entries of the system matrix $\mathbf{F}$ differ from each other significantly, which implies that the system affects the individual parameters differently.

In the following, the optimal power allocation strategies are obtained according to the Fisher information based criteria in Section 3 for the considered simulation setup, and the performance metrics are plotted against the dimension of the parameter vector, k , under a unit sum-power constraint, that is, $P_\Sigma = 0$ dB.

5.1 Results for Average MSE Criterion

In this case, the problem in (3.3) is considered, and the CRLBs achieved by the optimal power allocation strategy in (3.15) and by the equal power allocation strategy (i.e., $p_i^* = P_\Sigma/k$, $i = 1, \dots, k$) are plotted versus k in Fig. 5.1. It is noted that as the dimension of the parameter vector increases, the CRLB on the average MSE increases for both optimal and equal power allocation strategies except for the slight initial decrease in the optimal strategy for $\mathbf{F} = \mathbf{F}_1$. It is also observed that the optimal power allocation strategy consistently outperforms the equal power allocation strategy for both system matrices. As an example, for $\mathbf{F} = \mathbf{F}_2$, the CRLB is around 10^{-4} when $k = n = 7$ for the equal power allocation strategy, and the same level of CRLB is attained when $k = n = 14$ for the optimal power allocation strategy. Hence, significant improvements can be achieved by the optimal power allocation strategy.

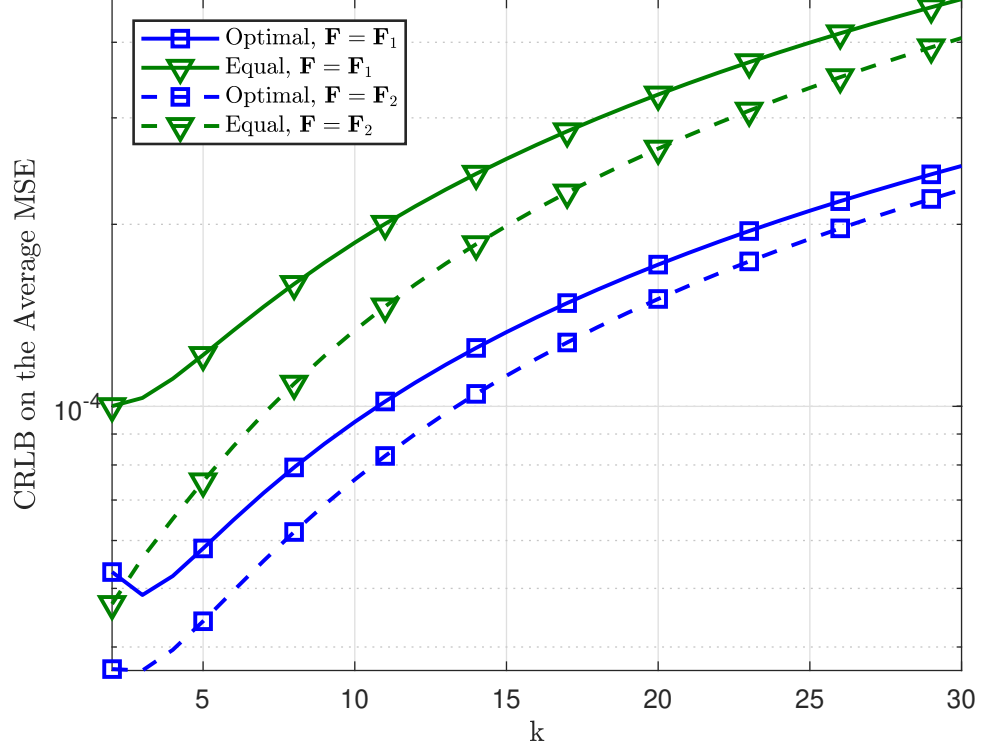


Figure 5.1: CRLB on the average MSE versus k for the equal and optimal power allocation strategies.

5.2 Results for Shannon Information Criterion

For this criterion, the problem in (3.17) (equivalently, (3.19)) is considered, which leads to the solution in (3.22). That is, the optimal and equal power allocation strategies yield the same solution in this case. The Shannon information achieved by the optimal (equal) power allocation strategy is plotted versus k in Fig. 5.2. It is observed that the Shannon information increases as the dimension of the parameter vector, k , increases. The increase in Shannon information is linear for both system matrices, and the achieved Shannon information scores are nearly the same for $\mathbf{F} = \mathbf{F}_1$ and $\mathbf{F} = \mathbf{F}_2$. This means that increased Shannon information can be achieved while estimating more parameters instead of fewer parameters when the number of observations is equal to the number of parameters.

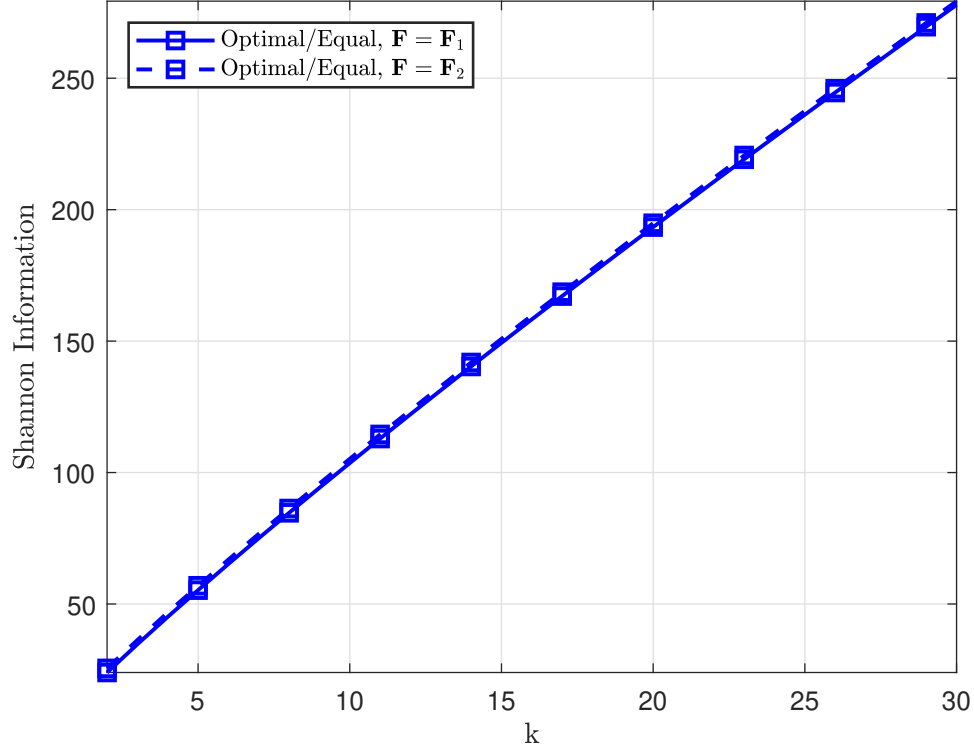


Figure 5.2: Shannon information versus k for the optimal (equal) power allocation strategy.

5.3 Results for Worst-Case Error Variance Criterion

In this case, the problem in (3.24) and the alternative problem in (3.27) are solved. The solution of (3.24) is obtained via the multistart global optimization algorithm in MATLAB. On the other hand, the equal power allocation strategy is the solution of (3.27), as shown in Section 3.3. In Fig. 5.3, the maximum eigenvalues of the CRLBs achieved by the optimal and equal power allocation strategies are plotted versus k . It can be seen in Fig. 5.3 that the optimal power allocation strategy can significantly outperform the equal power allocation strategy, and the difference between the two power allocation strategies increases as the number of parameters increases. One implication of this result is that power adaptation can

get more effective when there exist more parameters to estimate. In addition, it is noted that maximizing the lower bound on the eigenvalues of the CRLB is not sufficient to obtain the optimal power allocation strategy, as stated in Remark 1.

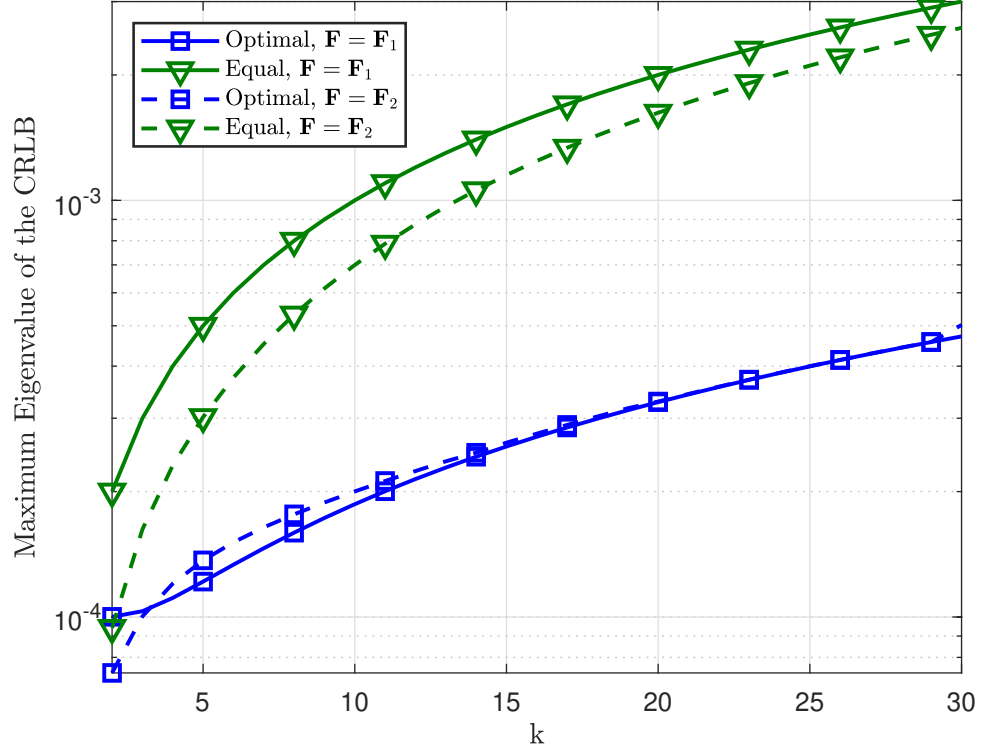


Figure 5.3: Maximum eigenvalue of the CRLB (inverse FIM) versus k for the optimal and equal power allocation strategies.

5.4 Results for Worst-Case Coordinate Error Variance Criterion

For this criterion, the problem in (3.32) is considered, which leads to the optimal solution in (3.36). The largest diagonal entry of the CRLB is plotted versus k for both the optimal solution and the equal power allocation strategy in Fig. 5.4. It is noted that the trend is similar to that in Fig. 5.3. Namely, the benefits of optimal power adaptation are observed for the worst-case coordinate error variance criterion, as well.

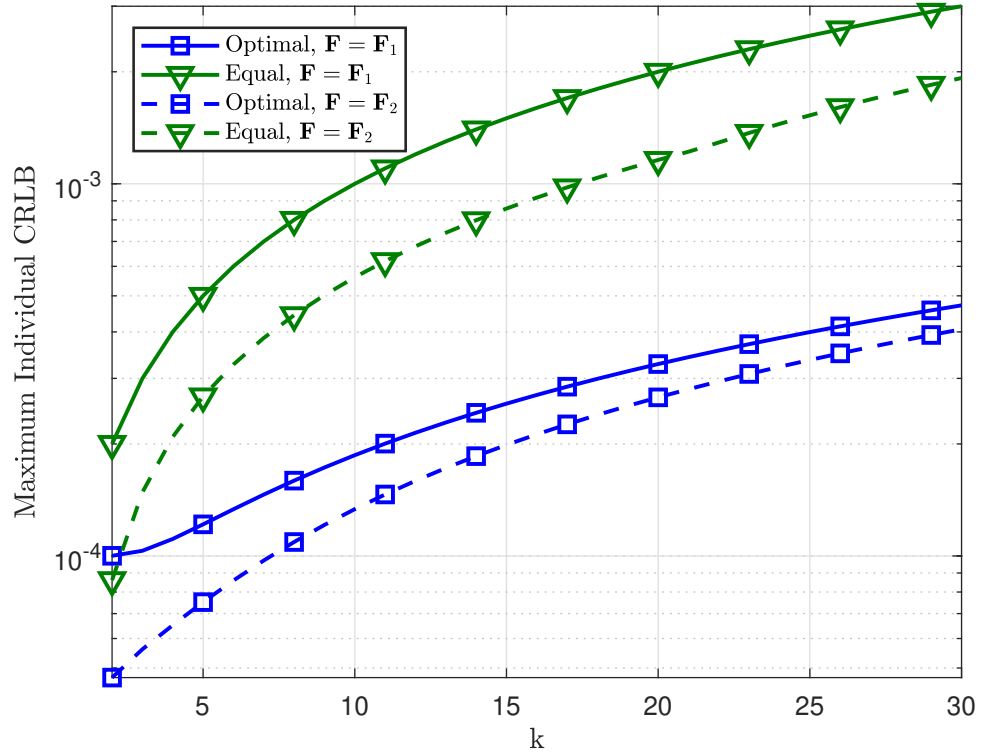


Figure 5.4: The largest diagonal entry of the CRLB (worst-case coordinate CRLB) versus k for the optimal and equal power allocation strategies.

5.5 Results for Average Fisher Information Criterion

In this case, we focus on the problem in (3.37), the solution of which is provided by (3.42) in Proposition 1. The impact of the dimension of the parameter vector, k , on the average Fisher information is shown in Fig. 5.5 for both the optimal solution in (3.42) and the equal power allocation strategy. It is observed that the average Fisher information rapidly decreases when $k \leq 10$ for both the optimal and equal power allocation strategies. However for $k > 10$, power adaptation significantly mitigates the adverse effects of the increased parameter dimension. While the optimal power allocation strategy is superior to the equal power allocation for all values of k , significant enhancements are observed for large values of k .

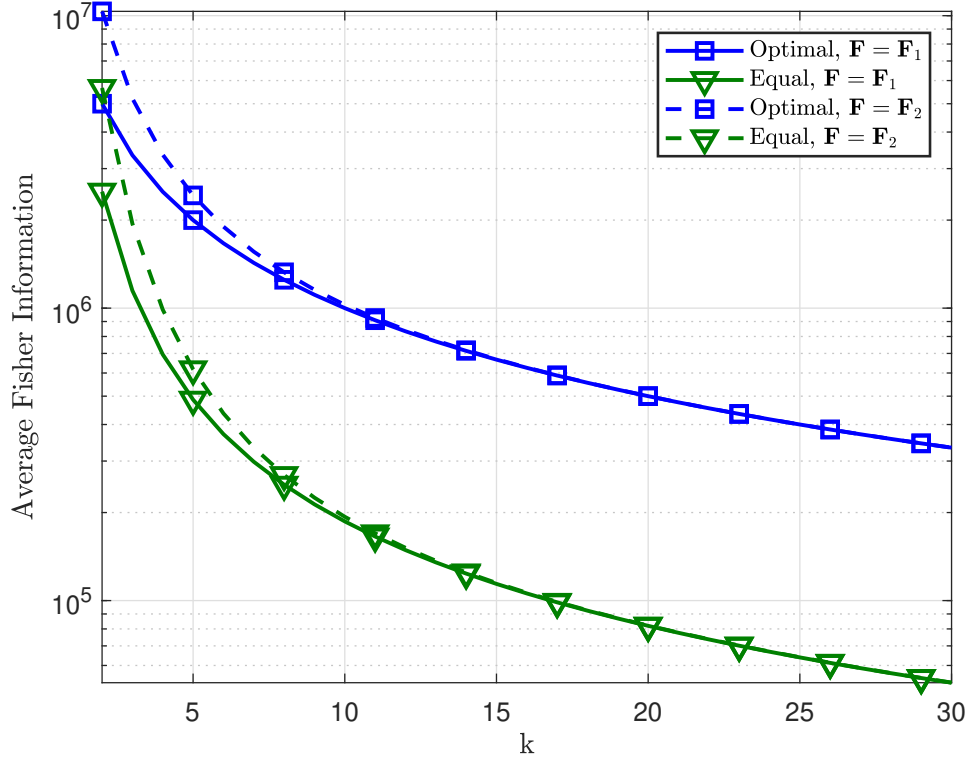


Figure 5.5: Average Fisher information versus k for the optimal and equal power allocation strategies.

5.6 Results for Worst Case Coordinate Fisher Information

In this scenario, the minimum diagonal entry of the FIM is maximized as in (3.48), leading to the optimal power allocation strategy in (3.53). The minimum diagonal entry of the FIM is plotted versus k for both the optimal and equal power allocation strategies in Fig. 5.6. When $\mathbf{F} = \mathbf{F}_2$, the worst-case coordinate Fisher information rapidly decreases for small k , while the trend is more steady when $\mathbf{F} = \mathbf{F}_1$. The decrease in worst-case Fisher information slows down for large values of k . Overall, the impact of power adaptation can be observed more clearly when k is large.

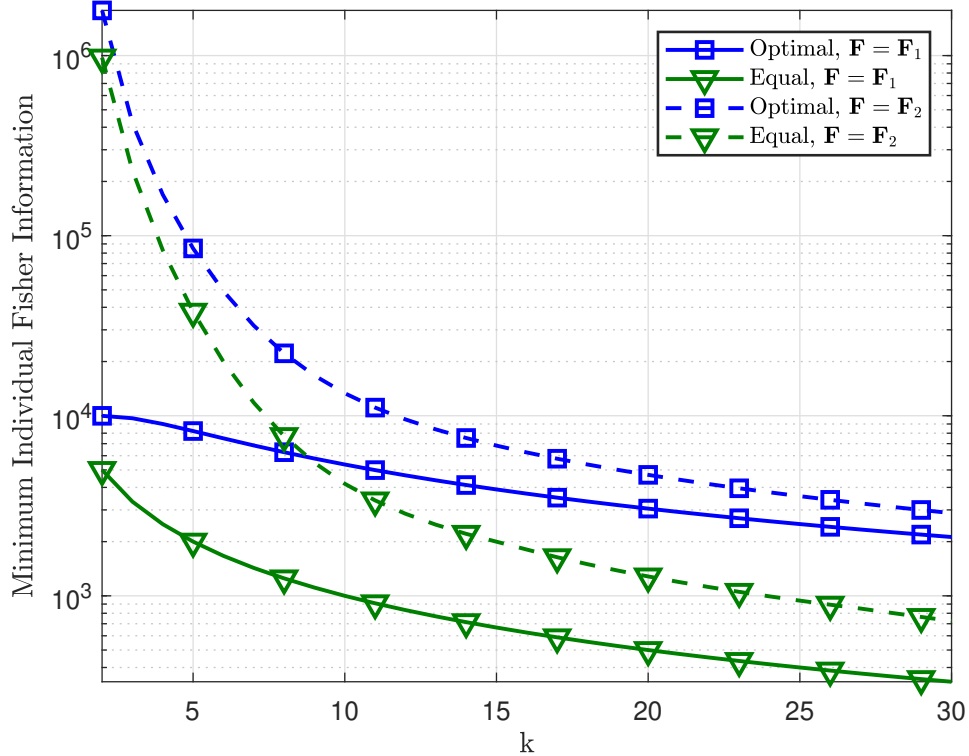


Figure 5.6: The minimum diagonal entry of the FIM (worst-case coordinate Fisher information) versus k for the optimal and equal power allocation strategies.

It is noted from the simulation results that when the dimensions of the parameter and observation vectors are large, power adaptation becomes more critical and the optimal power allocation strategies can provide more significant improvements over the equal power allocation strategy. In addition, the trends show that power adaptation can mitigate the adverse effects of increases in the dimension of the parameter vector when the observation vector has the same dimension as the parameter vector.



Chapter 6

Conclusion and Future Work

The optimal power allocation problem has been investigated for vector parameter estimation in the absence of prior information according to various Fisher information based optimality criteria. After deriving the FIM for a generic observation model, six different optimal power allocation problems have been formulated. Then, some closed-form solutions have been provided based on optimization theoretic approaches. It has been shown that the proposed power allocation strategies are also valid for nonlinear system models under certain conditions and in the presence of nuisance parameters. Numerical examples have shown that the use of the optimal power allocation strategies can provide significant improvements in estimation performance over the equal power allocation strategy.

One of the possible extensions to the investigated problems is to derive closed-form solutions to the problems with non-linear models. The fundamentals for this extension are discussed in Chapter 4. In addition, optimizing a precoding matrix that is not necessarily diagonal or square instead of a diagonal power allocation matrix is also a possible extension. Solving such a problem, however, would be more challenging due to the contributions of the off-diagonal entries. Lastly, introduction of secrecy constraints for the optimal secure transmission of parameters would be an interesting extension. In this case, however, the secrecy constraints can be non-convex; therefore, KKT conditions are no longer sufficient

for optimality. Hence, algorithmic approaches in conjunction with KKT-based characterization can lead to the solution of such a problem. Another possible extension is to consider a physically relevant nonlinear model. If the nonlinear function is continuously differentiable and the Jacobian of the nonlinear function does not depend on p_i 's for $i = 1, \dots, k$, then the proposed techniques would be applicable. On the other hand, the case where the nonlinear function is not differentiable w.r.t. its argument or a system involving a power parameter dependent nonlinear transformation still remain as an open problem.

Bibliography

- [1] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, 1993.
- [2] H. V. Poor, *An Introduction to Signal Detection and Estimation (2nd Ed.)*. Berlin, Heidelberg: Springer-Verlag, 1994.
- [3] A. Nehorai, G. Tang, and P. M. Green, *Bounds on Estimation*, pp. 1–8. London: Springer London, 2013.
- [4] A. Sani and A. Vosoughi, “Resource allocation optimization for distributed vector estimation with digital transmission,” in *48th Asilomar Conference on Signals, Systems and Computers*, pp. 1463–1467, 2014.
- [5] J. Li and G. AlRegib, “Distributed estimation in energy-constrained wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3746–3758, 2009.
- [6] N. H. Tran, L. Jiménez Rodríguez, and T. Le-Ngoc, “Optimal power control and error performance for full-duplex dual-hop af relaying under residual self-interference,” *IEEE Communications Letters*, vol. 19, no. 2, pp. 291–294, 2015.
- [7] R. Kalaba and K. Spingarn, *Control, Identification, and Input Optimization*. Mathematical Concepts and Methods in Science and Engineering, Springer US, 1 ed., 1982.

- [8] A. Sani and A. Vosoughi, “Distributed vector estimation for power- and bandwidth-constrained wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3879–3894, 2016.
- [9] M. Fanaei, M. C. Valenti, and N. A. Schmid, “Power allocation for distributed BLUE estimation with full and limited feedback of CSI,” in *IEEE Military Communications Conference (MILCOM 2013)*, pp. 418–423, 2013.
- [10] M. Shirazi and A. Vosoughi, “On Bayesian Fisher information maximization for distributed vector estimation,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 4, pp. 628–645, 2019.
- [11] X. Liu, Z. Li, X. Liu, and J. Wang, “Power allocation in vector estimation systems with the impact of wireless channel uncertainty: Power allocation in vector estimation systems,” *International Journal of Communication Systems*, vol. 30, 03 2015.
- [12] Y. Shen and M. Z. Win, “Energy efficient location-aware networks,” in *2008 IEEE International Conference on Communications*, pp. 2995–3001, 2008.
- [13] W. W. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, “Efficient anchor power allocation for location-aware networks,” in *2011 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2011.
- [14] W. W. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, “Robust power allocation via semidefinite programming for wireless localization,” in *2012 IEEE International Conference on Communications (ICC)*, pp. 3595–3599, 2012.
- [15] W. W. Li, Y. Shen, Y. J. Zhang, and M. Z. Win, “Robust power allocation for energy-efficient location-aware networks,” *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1918–1930, 2013.
- [16] T. Wang, G. Leus, and L. Huang, “Ranging energy optimization for robust sensor positioning based on semidefinite programming,” *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4777–4787, 2009.

- [17] T. Wang and G. Leus, “Ranging energy optimization for robust sensor positioning with collaborative anchors,” in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2714–2717, 2010.
- [18] H. Godrich, A. P. Petropulu, and H. V. Poor, “Power allocation strategies for target localization in distributed multiple-radar architectures,” *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3226–3240, 2011.
- [19] S. Bayram, N. D. Vanli, B. Dulek, I. Sezer, and S. Gezici, “Optimum power allocation for average power constrained jammers in the presence of non-Gaussian noise,” *IEEE Communications Letters*, vol. 16, no. 8, pp. 1153–1156, 2012.
- [20] Y. Shen, W. Dai, and M. Z. Win, “Optimal power allocation for active and passive localization,” in *2012 IEEE Global Communications Conference (GLOBECOM)*, pp. 3713–3718, 2012.
- [21] Y. Shen, W. Dai, and M. Z. Win, “Power optimization for network localization,” *IEEE/ACM Transactions on Networking*, vol. 22, no. 4, pp. 1337–1350, 2014.
- [22] S. Gezici, M. R. Gholami, S. Bayram, and M. Jansson, “Jamming of wireless localization systems,” *IEEE Transactions on Communications*, vol. 64, pp. 2660–2676, June 2016.
- [23] A. F. Emery and A. V. Nenarokomov, “Optimal experiment design,” *Measurement Science and Technology*, vol. 9, pp. 864–876, jun 1998.
- [24] S. Boyd and L. Vandenberghe, *Convex Optimization*. USA: Cambridge University Press, 2004.
- [25] K. Dogancay, “Online optimization of receiver trajectories for scan-based emitter localization,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 1117–1125, 2007.
- [26] K. Dogancay, “UAV path planning for passive emitter localization,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1150–1166, 2012.

- [27] B. Dulek and S. Gezici, “Cost minimization of measurement devices under estimation accuracy constraints in the presence of Gaussian noise,” *Digital Signal Processing*, vol. 22, p. 828–840, Sep. 2012.
- [28] E. Tzoreff and A. J. Weiss, “Single sensor path design for best emitter localization via convex optimization,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 939–951, 2017.
- [29] D. Gürgünoğlu, B. Dulek, and S. Gezici, “Power adaptation for vector parameter estimation according to Fisher information based optimality criteria,” 2020.
- [30] R. Zamir, “A proof of the Fisher information inequality via a data processing argument,” *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 1246–1250, 1998.
- [31] K. E. Parsopoulos and M. N. Vrahatis, *Particle swarm optimization method for constrained optimization problems*, pp. 214–220. IOS Press, 2002. in *Intelligent Technologies—Theory and Applications: New Trends in Intelligent Technologies*.
- [32] I. Jóźwiak and J. Legut, “Minimax decision rules for identifying an unknown distribution of a random variable,” in *Proceedings of 39th International Conference on Information Systems Architecture and Technology (ISAT 2018)* (J. Świątek, L. Borzemski, and Z. Wilimowska, eds.), pp. 308–317, 2019.
- [33] Y. Karisan, D. Dardari, S. Gezici, A. D’Amico, and U. Mengali, “Range estimation in multicarrier systems in the presence of interference: Performance limits and optimal signal design,” *IEEE Transactions on Wireless Communications*, vol. 10, pp. 3321–3331, Oct. 2011.
- [34] A. Goldsmith, *Wireless Communications*. Cambridge, UK: Cambridge University Press, 2005.
- [35] Y. Shen and M. Win, “Fundamental limits of wideband localization— part i: A general framework,” *Information Theory, IEEE Transactions on*, vol. 56, pp. 4956 – 4980, 11 2010.