Diplomarbeit

Channel Aware Inference Based on the Fisher Information

ausgeführt am
Institute of Telecommunications (E 389)
der Technischen Universität Wien

unter der Leitung von

Dipl.-Ing. Dr. techn. Alexander Jung
Ao. Univ.-Prof. Dipl.-Ing. Dr. Franz Hlawatsch

von

Bernhard Kausl
Traisenbachstraße 2
A-3184 Türnitz

Wien, Oktober 2012

Acknowledgements

I would like to express my very great appreciation to my supervisor Dr. Alexander Jung for his valuable and constructive suggestions. His willingness to give his time so generously has been very much appreciated.

Thanks also to Prof. Franz Hlawatsch for his support and that he gave me the opportunity, as part of his research group, to write a thesis.

I would like to express my thanks to my parents. Their support made my study of electrical engineering possible.

Abstract

We consider a resource constrained wireless sensor network, where a set of distributed sensors and a fusion center (FC) collaborate to estimate an unknown vector source. The basic question is, how should a sensor encode and/or compress the locally observed data before transmitting it over an imperfect channel to the FC. This encoding should be such that the FC can estimate the unknown vector source most accurately under the given bandwidth and power constraints for the data transmission. In this thesis, we focus our discussion on linear systems, where each sensor linearly encodes its local observed data and also the FC applies a linear mapping in order to estimate the unknown vector source, based on its received data. We adopt the Fisher information as our performance metric, which is motivated by their relation to the Cramér–Rao lower bound. We investigate two types of channel usage between sensors and FC, an orthogonal (i.e., non–interfering) and a coherent multiple access channel (MAC). For the case when the source is scalar–valued, we derive the optimal local sensor rule, when the channels between sensors and FC are orthogonal. We also derive an optimal power scheduling strategy, when a given total power is optimally scheduled among sensors. Simulations show that the proposed power scheduling performs much better than that for the uniform power scheduling. For a scalar–valued source, we also study the coherent MAC under a total power constraint and derive optimal local sensor rules in closed form for certain assumptions on the channel states. We also show in simulation that the asymptotic performance, when the number of sensors increases, critically depends on the different multiple access schemes. For the general case, when the source is vector–valued, we consider only the case of an orthogonal MAC. We derive optimal local sensor rules for certain assumptions on the channel states in closed form.

Zusammenfassung

Wir betrachten ein drahtloses Sensornetz mit begrenzten Ressourcen, in dem die Sensoren ihre lokalen Beobachtungen an einer unbekannten, im allgemeinen vektorwertigen Quelle, einem so genannten Fusion Center (FC) übermitteln. Vor allem in drahtlosen Sensornetzen ist die Bandbreite limitiert und Energieeffizienz von großer Bedeutung. Aus diesem Grund sollte jeder Sensor seine Beobachtungen (Messdaten) komprimiert und/oder codiert zum FC übertragen. Das Codieren soll dabei in einer Art und Weise geschehen, damit das FC die unbekannte Quelle möglichst genau (optimal) schätzen kann, für eine vorgegebene maximale Bandbreite und Sendeleistung. Im Rahmen dieser Diplomarbeit beschränken wir uns auf lineare Systeme, wo die Codiervorschrift am lokalen Sensor (Sensor–Regel) als auch die Schätzfunktion am FC mit einer linearen Transformationen beschrieben werden. Als Performancekriterium verwenden wir die Fisher Information, motiviert durch ihre Beziehung zur Cramér–Rao-Schranke. Wir betrachten einerseits einen orthogonalen (d.h. ohne Nachbarkanal–Interferenzen), andererseits einen koheränten Mehrfachzugriffskanal (MZK) zwischen den Sensoren und dem FC. Für den Spezialfall einer skalarwertigen Quelle und der Annahme eines orthogonalen MZKs geben wir die optimale Codiervorschrift am lokalen Sensor an; eine optimale Leistungs–Verteilungsstrategie, wenn eine vorgegebene maximale Gesamtleistung im Sensornetz auf die einzelnen Sensoren optimal aufgeteilt werden soll, sodass die maximale Systemperformance resultiert. Durch Simulationen wird gezeigt, dass dadurch ein signifikanter Performance–Gewinn resultiert, gegenüber der einer gleichverteilten Verteilungsstrategie. Für einen koheränten MZK und einer skalarwertigen Quelle werden optimale Codiervorschriften unter einer Gesamtleistungbegrenzung und für gewisse Spezialfälle an das Kanalmodell gezeigt. Für den allgemeinen Fall einer vektorwertigen Quelle wird im Rahmen dieser Diplomarbeit nur der orthogonale MZK Fall studiert, wobei unter bestimmten Annahmen an den Kanalzuständen optimale Codiervorschriften in geschlossener Form gezeigt werden.

Contents

Acronyms

CRLB Cramér–Rao lower bound

FIM Fisher information matrix

FI Fisher information

FC fusion center

MVU minimum variance unbiased

MSE mean square error

BLUE best linear unbiased estimator

pdf probability density function

ML maximum likelihood

LGM linear Gaussian model

MAC multiple access channel

SDP semi–definite programming

WSN wireless sensor network

iid independent and identically distributed

SVD singular value decomposition

EVD eigenvalue decomposition

KKT Karush–Kuhn–Tucker

SNR signal to noise ratio

MMSE minimum mean square error

Chapter 1

Introduction

1.1 Motivation – Wireless Sensor Networks

Consider a distributed wireless sensor network (WSN), where sensors observe data from a vector source and transmit it, possibly after performing some preprocessing, to a fusion center (FC) over an imperfect channel. An example could be a target tracking scenario, where several sensors track the movement of a target object and transmit their observations to a central unit. Cooperative communications between sensors would cost much more local energy and in addition increase the system complexity. Therefore, distributed schemes (i.e., non–cooperative locally) are of more practical importance. The FC receives the transmitted data set from the different sensors, which are in general affected by the channel more or less, and based on the received data, it generates a final estimate on the unknown source for a specific signal processing task. We investigate two types of channel usage between sensors an FC: an orthogonal and a coherent multiple access channel (MAC). For the case of an orthogonal MAC, the sensors have their independent non–interfering channels to the FC. As for a coherent MAC, we allow all sensors transmit simultaneously, by assuming that all transmit messages reaches the FC in a coherent sum.

However, communications between sensors and FC is costly, as is the case in WSNs. Especially in such networks is an important fact, energy efficiency of their operation. E.g., battery capacities may be small and their replacement unfeasible. There can be significant power savings, if less information is transmitted to the FC, without degrading the overall performance. The basic question is, how to encode and/or compress the locally observed data before transmitting it over the channel to the FC. This encoding should be such that the FC can estimate the parameter of interest most accurately under the given bandwidth and power constraints for the data transmission. There are at least two approaches in which the finite bandwidth constraint can be modeled. On the one hand, we can limit the number of binary bits that each sensor transmit to the FC per observation period (source coding). This bandwidth measure is from a digital communications point of view. On the other hand, we can limit the number of real–valued messages that each sensor transmit to the FC per observation period, which is directly proportional to the physical frequency bandwidth in the system. This approach is suited for analog transmission schemes. Throughout this thesis, we adopt the second bandwidth measure, i.e., we consider analog transmission of real–valued sensor messages (cf. intro of [1]). The power constraint in contrast limits the strength of the transmit signals.

In this thesis, we discuss the joint estimation of a vector parameter by a sensor network with a FC. The transmission between sensors and FC is subject to bandwidth and power constraints. We focus our discussion on linear systems, where encoding functions at the local sensors and the fusion function at the FC are all linear. The reason behind restricting to linear models is tractability. For non–linear models, there are often numerical/iterative techniques necessarily, which in particular often converges only to local optimum solutions. As a consequence of considering linear models, we can describe our local sensor rule by some sensor matrix and additive systematic noise. The power constraint limits the strength of each transmitted data, while the bandwidth constraint limits the number of real–valued transmitted symbols (messages) per observation period. Under a Cramér–Rao lower bound (CRLB) criterion, we design the optimal local sensor rule, based on the channel states and the second order statistics of the local observation.

1.2 State of the Art

Similar questions were addressed by the authors of [1] and [2]. They designed optimal local sensor rules under the mean square error (MSE) criterion and considered a Bayesian setting in contrast (minimum mean square error (MMSE) estimator).

In [1], they designed optimal local sensor rules by assuming non–orthogonal channel usage (the orthogonal channel usage has been studied in [3] before), subject to bandwidth and/or power constraints, for cases where the parameter of interest and local sensor observations are scalars or vectors.

For the scalar case, they used solutions for the optimal local sensor rules from [3] and derived an optimal power scheduling strategy, i.e., where a given total transmit power is optimally scheduled among all sensors such that the achieved MSE is maximized. Simulations show that the proposed power scheduling strategy significantly improves the MSE performance when compared to an uniform power scheduling (i.e., all sensors use the same transmit power). They have also shown that the MSE performance critically depends on the different multiple access schemes (orthogonal and coherent MAC), which has in particular significantly different asymptotic behaviours (in the sense when the number of sensors increases). When the parameter of interest and local sensor observations are vectors, they derived a closed–form solution of an optimal local sensor rule for a noiseless channel (i.e., neglecting the additive channel noise). For a noisy channel, the problem can be efficiently solved by a numerical method (semi–definite programming (SDP)).

In [2], they differentiate between uncorrelated and correlated local sensor observations, i.e., whether the local sensor observations are uncorrelated among different sensors or not. They considered the case of estimating a vector parameter and analyzed the MSE performance for a system setup with an orthogonal MAC. For correlated sensor observations, they derived a closed–form MSE optimal local sensor rule and showed an optimal power scheduling strategy in a waterfilling–like manner so as to balance channel strength and additive channel noise variance. For correlated sensor observations, they further developed an iterative algorithm with guaranteed convergence to at least a stationary point of the MSE–cost.

By contrast, we consider a classical estimation problem, where the parameter vector is modeled as unknown deterministic and use the Fisher information (FI) as the performance metric. The motivation for using the FI is based on its relation to the CRLB.

1.3 Organization of this Thesis

The rest of this thesis is organized as follows. In Chapter 2, we review some elementary concepts of classical estimation theory. In particular, we introduce the concepts of the CRLB, the Fisher information matrix (FIM) and discuss their properties. We then specialize to the linear Gaussian model (LGM) which will be used throughout the thesis. In Chapter 3, we give a general problem formulation for our system model. The FI performance metric and power constraints are derived in terms of the local sensor rules. We also discuss some fundamental notions of the optimal experiment design. In particular, we introduce various optimality criteria which can be used in the case of a vector–valued parameter. Then, we formulate the basic design problem in the most general form, in order to obtain the optimal local sensor rule. In Chapter 4, we show the main results of this thesis. We solve the basic design problem for certain special cases, first, for the special case of a scalar parameter, afterwards, for the general case of a vector parameter, where in the latter, we are particularly interested in two optimality criteria. For the scalar parameter, we also show an optimal power scheduling strategy. Finally, in Chapter 5, we show numerical experiments, first, for the scalar parameter, where we are interested in the optimal power scheduling performance gain. For a vector parameter, we compare the two cases of optimal designs with regard to the MSE performance in a single sensor setup.

1.4 Symbols and Notations

Throughout this thesis we adopt the following notations: A lower/upper case letter a/A denotes a real scalar, a boldface/lowercase letter $\mathbf{a}$ denotes a vector and a boldface/uppercase letter $\mathbf{A}$ denotes a matrix; $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}^+$ denotes the set of positive real numbers including 0, $\mathbb{R}\backslash\{0\}$ denotes the set of real numbers excluding 0, $\mathbb{R}^+\backslash\{0\}$ denotes the set of positive real numbers excluding 0, $\mathbb{R}^m$ denotes the set of all real vectors of dimension m, $\mathbb{R}^{+m}$ denotes the set of all real vectors with positive elements; $\sqrt[+]{\cdot}$ denotes the positive square root; $|\cdot|$ denotes the absolute value; $\operatorname{sign}(\cdot)$ denotes the signum function (returns 1 or -1 depending on the sign of the argument); $(\cdot)^*$ denotes an optimum; $\min\{a_1, a_2, \ldots, a_K\}$ and $\max\{a_1, a_2, \ldots, a_K\}$ denote the minimum and maximum of the set $\{a_1, a_2, \ldots, a_K\}$.

Matrix and Vector Analysis: The notations $\mathbf{A}^T$, $\mathbf{A}^{-1}$, $\mathbf{A}^{-T}$, $\mathbf{A}^\dagger$ mean the transpose, the matrix–inverse, the matrix–inverse–transpose and the pseudo–inverse of a matrix $\mathbf{A}$; $\mathbf{I}$ denotes the identity matrix; $\mathbf{1}$ denotes a vector of ones; $\mathbf{0}$ denotes a vector of zeros. The ith element of a vector $\mathbf{a}$ is denoted by a_i, the element of the ith row and jth column of $\mathbf{A}$ is denoted by $(\mathbf{A})_{i,j}$, the ith column vector of $\mathbf{A}$ is denoted by $\mathbf{a}_i$, the ith row vector of $\mathbf{A}$ is denoted by $\mathbf{a}_i^r$. We denote the set of all real $(k \times k)$–symmetric matrices by $\operatorname{Sym}(k)$, –positive semi–definite matrices by $\operatorname{NND}(k)$ and –positive definite matrices by $\operatorname{PD}(k)$. Let $\mathbf{A}, \mathbf{B} \in \operatorname{Sym}(k)$, then the relations $\mathbf{A} \geq \mathbf{B}$ or $\mathbf{B} \leq \mathbf{A}$ means that $\mathbf{A} - \mathbf{B} \in \operatorname{NND}(k)$, similarly, $\mathbf{A} > \mathbf{B}$ or $\mathbf{B} < \mathbf{A}$ means that $\mathbf{A} - \mathbf{B} \in \operatorname{PD}(k)$ (Loewner ordering among symmetric matrices). The relation '$\succeq$', '$\preceq$', '$\succ$' and '$\prec$' denote the corresponding element–wise inequalities for vectors. The notation $\operatorname{R}(\mathbf{A}) \triangleq \{\mathbf{Av} \in \mathbb{R}^m : \mathbf{v} = \mathbf{0}\}$ and $\operatorname{N}(\mathbf{A}) \triangleq \{\mathbf{v} \in \mathbb{R}^n : \mathbf{Av} = \mathbf{0}\}$ mean the range and the nullspace of the matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$. The notation $\operatorname{rank}(\mathbf{A})$ means the rank of the matrix $\mathbf{A}$. The vector $\mathbf{e}_i$ denotes the ith unit vector. The notation $1 \leq i, j \leq N$ means that $i, j \in \{1, 2 \ldots, N\}$. $\operatorname{diag}\{x_1, x_2, \ldots, x_K\}$ denotes a diagonal matrix with entries x_i for $1 \leq i \leq K$. $\operatorname{tr}\{\mathbf{A}\}$ and $\det\{\mathbf{A}\}$ denote the trace ad the determinant of a matrix $\mathbf{A}$, $\|\cdot\|$ denotes the Euclidean norm (l_2–norm).

Statistical Signal Processing: The notations $\mathbf{C}_a$ and $\boldsymbol{\mu}_a$ mean the auto–covariance matrix and the mean of the random vector $\mathbf{a}$; $\mathbf{a} \sim \mathcal{N}(\boldsymbol{\mu}_a, \mathbf{C}_a)$ means that $\mathbf{a}$ is Gaussian distributed with mean $\boldsymbol{\mu}_a$ and covariance matrix $\mathbf{C}_a$; $\operatorname{var}\{\mathbf{a}\}$ and $\operatorname{cov}\{\mathbf{a}\}$ mean the variance and the auto–covariance matrix (equivalent to the notation $\mathbf{C}_a$) of $\mathbf{a}$; $\operatorname{cov}\{\mathbf{a}, \mathbf{b}\}$ means the cross–covariance matrix between $\mathbf{a}$ and $\mathbf{b}$; $\operatorname{E}\{\cdot\}$ denotes the expectation operator.

Abbreviations: "Fig.", "w.r.t.", "w.l.o.g." denote "Figure", "with respect to", "without log of generality"; "iff" means "if and only if"; "cf." means "confer"; "ev.", "p.", "ftn." stand for "evaluated", "page", "footnote".

Chapter 2

Basic Concepts of Classical Estimation Theory

In this chapter, we will introduce the *Cramér–Rao lower bound* (CRLB) and the *Fisher information matrix* (FIM) for the general case of a vector parameter, which is a fundamental result in classical estimation theory. The CRLB is a lower bound on the variance of any unbiased estimator and is practically useful since it provides a benchmark against which we can compare the performance of any unbiased estimator. In certain cases, it even allows us to find the minimum variance unbiased (MVU) estimator. Before we go into details of the CRLB and the FIM, we review some basic concepts of classical estimation theory.

2.1 The Estimation Problem

Let us consider an unknown, deterministic parameter vector $\boldsymbol{\theta} \in \mathbb{R}^n$. A so called *fusion center* (FC) receives the data $\mathbf{z}$ and estimates the parameter vector $\boldsymbol{\theta}$, based on the observed data $\mathbf{z}$. It should be noted that the FC has no prior information about the parameter vector $\boldsymbol{\theta}$, i.e., we consider the classical estimation setting in contrast to the Bayesian setting, where the parameter vector is modeled random with a known prior probability density function (pdf). The dependence of the observed data $\mathbf{z}$ and $\boldsymbol{\theta}$ is described by the family of pdfs

$$f\left(\mathbf{z};\boldsymbol{\theta}\right), \tag{2.1}$$

i.e., the notation in (2.1) means, that the pdf of $\mathbf{z}$ is parameterized (indexed) by $\boldsymbol{\theta}$. For an estimator $\hat{\boldsymbol{\theta}}\left(\mathbf{z}\right)$ the estimation error $\mathbf{e}$ is defined as

$$\mathbf{e} = \hat{\boldsymbol{\theta}}\left(\mathbf{z}\right) - \boldsymbol{\theta}. \tag{2.2}$$

The *mean square error* (MSE) of an estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is given by

$$\begin{aligned} \mathrm{MSE}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \frac{1}{n}\mathrm{E}\left\{\|\mathbf{e}\|^2\right\} &\overset{(2.2)}{=} \frac{1}{n}\mathrm{E}\left\{\left\|\hat{\boldsymbol{\theta}}(\mathbf{z}) - \boldsymbol{\theta}\right\|^2\right\} \\ &= \frac{1}{n}\int_{\mathbf{z}}\left\|\hat{\boldsymbol{\theta}}(\mathbf{z}) - \boldsymbol{\theta}\right\|^2 f(\mathbf{z};\boldsymbol{\theta}). \end{aligned} \tag{2.3}$$

It is important to note that the expectation in (2.3) is only with respect to $\mathbf{z}$, since $\boldsymbol{\theta}$ is non-random. As the notation in (2.3) suggests, the MSE depends on the parameter vector $\boldsymbol{\theta}$ in general. The MSE can be decomposed into two terms[1]:

$$\mathrm{MSE}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \frac{1}{n}\left\|\mathrm{bias}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}\right\|^2 + \mathrm{var}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}, \tag{2.4}$$

where the bias of $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is defined as the expectation of the estimation error $\mathbf{e}$, i.e.,

$$\mathrm{bias}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} \triangleq \mathrm{E}_{\boldsymbol{\theta}}\{\mathbf{e}\} \overset{(2.2)}{=} \mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z}) - \boldsymbol{\theta}\right\} = \mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} - \boldsymbol{\theta}, \tag{2.5}$$

and the variance of the estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is given by

$$\mathrm{var}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \frac{1}{n}\mathrm{E}_{\boldsymbol{\theta}}\left\{\left\|\hat{\boldsymbol{\theta}}(\mathbf{z}) - \mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}\right\|^2\right\}. \tag{2.6}$$

As the MSE, also the bias and the variance of an estimator depend on $\boldsymbol{\theta}$ in general.

Definition 2.1.1 *An estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is said to be unbiased iff*

$$\mathit{bias}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = E_{\boldsymbol{\theta}}\{\mathbf{e}\} = \mathbf{0} \quad \textit{for all } \boldsymbol{\theta}. \tag{2.7}$$

As can be verified easily, for an unbiased estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$ it holds that

$$\mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \boldsymbol{\theta} \quad \text{for all } \boldsymbol{\theta}$$

(cf. (2.5)) and moreover, by (2.4), we have

$$\mathrm{MSE}_{\boldsymbol{\theta}} = \mathrm{var}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}.$$

We also define the covariance matrix of $\hat{\boldsymbol{\theta}}(\mathbf{z})$ by

$$\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \mathrm{E}_{\boldsymbol{\theta}}\left\{\left(\hat{\boldsymbol{\theta}}(\mathbf{z}) - \mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{z}) - \mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}\right)^T\right\}. \tag{2.8}$$

---

[1] The decomposition is only valid in the classical context, i.e., if $\boldsymbol{\theta}$ is modeled as deterministic.

If $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is unbiased, i.e., $\mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, the covariance matrix of $\boldsymbol{\theta}(\mathbf{z})$ from (2.8) equals the "MSE-matrix" $\mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{e}\mathbf{e}^T\right\}$:

$$\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \mathrm{E}_{\boldsymbol{\theta}}\left\{\left(\hat{\boldsymbol{\theta}}(\mathbf{z}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{z}) - \boldsymbol{\theta}\right)^T\right\} \overset{(2.2)}{=} \mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{e}\mathbf{e}^T\right\}.$$

Furthermore, $\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z}) - \boldsymbol{\theta}\right\} = \mathrm{cov}_{\boldsymbol{\theta}}\left\{\mathbf{e}\right\}$ since $\boldsymbol{\theta}$ is deterministic. Note that the kth diagonal element of $\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}$ equals the variance of the kth estimator component $\hat{\theta}_k$, i.e., $\mathrm{var}_{\boldsymbol{\theta}}\left\{\hat{\theta}_k\right\} = \left(\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}\right)_{k,k}$, and thus it equals the MSE of $\hat{\theta}_k$, if $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is unbiased. In particular, the MSE of $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is obtained as the arithmetic mean of all individual MSEs of $\hat{\theta}_k$ for $1 \leq k \leq n$. Hence, the MSE is also given by the trace of the "MSE-matrix"/error covariance matrix/covariance matrix of an unbiased estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$, divided by n:

$$\mathrm{MSE}_{\boldsymbol{\theta}} = \frac{1}{n}\mathrm{tr}\left\{\mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{e}\mathbf{e}^T\right\}\right\} = \frac{1}{n}\mathrm{tr}\left\{\mathrm{cov}_{\boldsymbol{\theta}}\left\{\mathbf{e}\right\}\right\} = \frac{1}{n}\mathrm{tr}\left\{\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}\right\}. \quad (2.9)$$

2.2 The Cramér–Rao Lower Bound

Given an observation $\mathbf{z}$ and an estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$, it is desirable to quantify how good the estimator performs, e.g., by comparing it against some benchmark. We now introduce our central performance benchmark for the set of all unbiased estimators for a classical estimation problem, which is related to the *Cramér–Rao lower bound* (CRLB) and the *Fisher information matrix* (FIM).

2.2.1 The Fisher Information Matrix

In the following, we assume that $\frac{\partial}{\partial\theta_k}\ln f(\mathbf{z};\boldsymbol{\theta})$ and $\frac{\partial^2}{\partial\theta_k\partial\theta_l}\ln f(\mathbf{z};\boldsymbol{\theta})$ exist and are absolutely integrable with respect to $\mathbf{z}$. Consider an estimation problem based on the observation vector $\mathbf{z}$, whose pdf $f(\mathbf{z};\boldsymbol{\theta})$ is parametrized by the parameter vector $\boldsymbol{\theta}$, which we would like to estimate. We can then define the corresponding FIM as

$$\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) \triangleq \mathrm{E}_{\boldsymbol{\theta}}\left\{\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f(\mathbf{z};\boldsymbol{\theta})\right]\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f(\mathbf{z};\boldsymbol{\theta})\right]^T\right\}. \quad (2.10)$$

The FIM is a square matrix of size $n \times n$, where n is the dimension of the parameter vector $\boldsymbol{\theta}$. It should be noted that $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ depends on the parameter $\boldsymbol{\theta}$ in general. The elements of $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ are thus given by

$$\left(\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})\right)_{k,l} = \int_{\mathbf{z}}\left[\frac{\partial}{\partial\theta_k}\ln f(\mathbf{z};\boldsymbol{\theta})\right]\left[\frac{\partial}{\partial\theta_l}\ln f(\mathbf{z};\boldsymbol{\theta})\right] f(\mathbf{z};\boldsymbol{\theta})\, d\mathbf{z}. \quad (2.11)$$

We can write (2.10) and (2.11) also in the more compact form as

$$\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = -\mathrm{E}_{\boldsymbol{\theta}}\left\{\frac{\partial}{\partial\boldsymbol{\theta}}\frac{\partial}{\partial\boldsymbol{\theta}}\ln f(\mathbf{z};\boldsymbol{\theta})\right\} \tag{2.12}$$

and

$$(\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}))_{k,l} = -\mathrm{E}_{\boldsymbol{\theta}}\left\{\frac{\partial^2}{\partial\theta_k\partial\theta_l}\ln f(\mathbf{z};\boldsymbol{\theta})\right\}. \tag{2.13}$$

The FIM is symmetric, i.e., $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = \mathbf{J}_{\mathbf{z}}^T(\boldsymbol{\theta}) \in \mathrm{Sym}(n)$ and positive semi-definite, i.e., $\mathbf{J}_{\mathbf{z}} \in \mathrm{NND}(n)$. For the set of symmetric matrices, we can define a partial ordering to be able to compare two or more FIMs [4].

Definition 2.2.2 *The partial ordering $\geq$, is defined on* $\mathrm{Sym}(s)$*, by*

$$\mathbf{A} \geq \mathbf{B} \quad \Longleftrightarrow \quad \mathbf{A}-\mathbf{B} \geq 0 \quad \Longleftrightarrow \quad \mathbf{A}-\mathbf{B} \in \mathrm{NND}(s),$$

which is known as the Loewner ordering of symmetric matrices. Note that the notation $\mathbf{B} \leq \mathbf{A}$ is equivalent to $\mathbf{A} \geq \mathbf{B}$. We also define the closely related variant $>$, by

$$\mathbf{A} > \mathbf{B} \quad \Longleftrightarrow \quad \mathbf{A}-\mathbf{B} > 0 \quad \Longleftrightarrow \quad \mathbf{A}-\mathbf{B} \in \mathrm{PD}(s).$$

In the scalar case, i.e., for $s = 1$, the Loewner ordering reduces to the familiar total ordering on the real line $\mathbb{R}$. Or, the other way around, the total ordering of the real line $\mathbb{R}$ is extended to the partial ordering of the matrix spaces $\mathrm{Sym}(s)$, with $s > 1$. In Chapter 3, we will define our basic design problem, which is based on the Loewner ordering among FIMs. Another important property holds, if the data z_k are statistically independent for all $1 \leq k \leq n$. Then, the FIM can be written as

$$\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = \sum_{k=1}^{n}\mathbf{J}_{z_k}(\boldsymbol{\theta}), \tag{2.14}$$

where $\mathbf{J}_{z_k}(\boldsymbol{\theta})$ is the FIM for the z_kth data. This property can be easily verified since for independent data z_k, the pdf $f(\mathbf{z};\boldsymbol{\theta})$ can be factored into the form

$$f(\mathbf{z};\boldsymbol{\theta}) = \prod_{k=1}^{n} f(z_k;\boldsymbol{\theta}).$$

2.2.2 The Cramér–Rao Lower Bound

If the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ is non-singular, i.e., the inverse $\mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta})$ exists for all $\boldsymbol{\theta}$, it can be shown that the MSE matrix/error covariance matrix/covariance matrix of any unbiased estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$ is bounded below by the inverse FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ [5],

$$\mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{e}\mathbf{e}^T\right\} = \mathrm{cov}_{\boldsymbol{\theta}}\left\{\mathbf{e}\right\} = \mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} \geq \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta}). \tag{2.15}$$

The inequality (2.15) is referred to as the *Cramér–Rao lower bound* (CRLB). Throughout this thesis we only consider estimation problems where the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ is non-singular. However, there are also generalizations of the CRLB to situations where the FIM is singular [6].

2.2.3 Efficient Estimators

If the covariance matrix $\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\}$ of an unbiased estimator $\hat{\boldsymbol{\theta}}(\mathbf{z})$, i.e., $\mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, attains the CRLB, i.e.,

$$\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta}),$$

then such an estimator is called *efficient*, denoted by $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z})$. An efficient estimator exists if and only if $\frac{\partial}{\partial\boldsymbol{\theta}}\ln f(\mathbf{z};\boldsymbol{\theta})$ can be written as

$$\frac{\partial}{\partial\boldsymbol{\theta}}\ln f(\mathbf{z};\boldsymbol{\theta}) = \mathbf{K}(\boldsymbol{\theta})\left[\mathbf{g}(\boldsymbol{\theta}) - \boldsymbol{\theta}\right], \tag{2.16}$$

with some $n \times n$ matrix $\mathbf{K}(\boldsymbol{\theta})$ and some function $\mathbf{g}(\boldsymbol{\theta})$ [5]. This estimator is then given by

$$\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z}) = \mathbf{g}(\mathbf{z}), \tag{2.17}$$

and its covariance matrix is given by

$$\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}_{\text{eff}}(\boldsymbol{\theta})\right\} = \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta}) = \mathbf{K}^{-1}(\boldsymbol{\theta}), \tag{2.18}$$

i.e., the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = \mathbf{K}(\boldsymbol{\theta})$. If an efficient estimator exists, it coincides with the MVU estimator and the maximum likelihood (ML) estimator.

In the following, we will define the CRLB and the FIM for the special case of a Gaussian distributed observation $\mathbf{z}$. Furthermore, we specialize it to a linear observation model, i.e., when observation $\mathbf{z}(\boldsymbol{\theta})$ is linear in $\boldsymbol{\theta}$, because we will consider only that system model exclusively throughout this thesis.

2.2.4 The Gaussian Case

For the case of a Gaussian distributed observation, i.e., we assume $\mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu}_z(\boldsymbol{\theta}), \mathbf{C}_z(\boldsymbol{\theta}))$, where $\mathbf{C}_z$ is non-singular, it can be shown [5] that

$$\begin{aligned}(\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}))_{k,l} = \left[\frac{\partial\boldsymbol{\mu}_z(\boldsymbol{\theta})}{\partial\theta_k}\right]^T \mathbf{C}_z^{-1}(\boldsymbol{\theta})\left[\frac{\partial\boldsymbol{\mu}_z(\boldsymbol{\theta})}{\partial\theta_l}\right] + \\ \frac{1}{2}\mathrm{tr}\left\{\mathbf{C}_z^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}_z(\boldsymbol{\theta})}{\partial\theta_k}\mathbf{C}_z^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}_z(\boldsymbol{\theta})}{\partial\theta_l}\right\}.\end{aligned} \tag{2.19}$$

2.2.5 The Linear Gaussian Model

We now consider the LGM, i.e., the observation $\mathbf{z}$ can be written in the form

$$\mathbf{z} = \mathbf{H}\boldsymbol{\theta} + \mathbf{v},$$

where $\mathbf{H}$ is a deterministic matrix of size $p \times q$, with $p \geq q$ and full column-rank, i.e., $\operatorname{rank}(\mathbf{H}) = q$. The random vector $\mathbf{v}$ is Gaussian distributed with mean $\boldsymbol{\mu}_v$ and the non-singular covariance matrix $\mathbf{C}_v$, i.e., $\mathbf{v} \sim \mathcal{N}(\boldsymbol{\mu}_v, \mathbf{C}_v)$. The LGM is a special case of the general Gaussian model with mean $\boldsymbol{\mu}_v$ and covariance matrix $\mathbf{C}_v$. Therefore, we obtain the FIM of the LGM by specializing (2.19) as

$$\mathbf{J}_{\mathbf{z}} = \left(\frac{\partial \boldsymbol{\mu}_z}{\partial \boldsymbol{\theta}}\right)^T \mathbf{C}_z^{-1} \frac{\partial \boldsymbol{\mu}_z}{\partial \boldsymbol{\theta}} = \mathbf{H}^T \mathbf{C}_v^{-1} \mathbf{H}. \tag{2.20}$$

An important property of the LGM is that the FIM $\mathbf{J}_{\mathbf{z}}$ from (2.20) does not depend on $\boldsymbol{\theta}$. Furthermore, it can be shown [5] that for a LGM,

- there always exist an efficient estimator given by

$$\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z}) = \left(\mathbf{H}^T \mathbf{C}_v^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^T \mathbf{C}_v^{-1} (\mathbf{z} - \boldsymbol{\mu}_v), \tag{2.21}$$

- the estimator is unbiased, $\mathrm{E}\left\{\hat{\boldsymbol{\theta}}(\mathbf{z})\right\} = \boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, and its covariance matrix is given by

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z})} = \mathbf{J}_{\mathbf{z}}^{-1} = \left(\mathbf{H}^T \mathbf{C}_{\mathbf{v}}^{-1} \mathbf{H}\right)^{-1}$$

  and does not depend on $\boldsymbol{\theta}$,

- the efficient estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z})$ is Gaussian distributed, i.e.,

$$\hat{\boldsymbol{\theta}}_{\text{eff}} \sim \mathcal{N}\left(\boldsymbol{\theta}, \left(\mathbf{H}^T \mathbf{C}_{\mathbf{v}}^{-1} \mathbf{H}\right)^{-1}\right),$$

- the estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z})$ coincides with the MVU, the ML and the best linear unbiased estimator (BLUE).

Chapter 3

Problem Formulation and System Model

In the last chapter, we introduced the CRLB and the FIM for a classical estimation problem. Since the CRLB is a lower bound on the MSE matrix/error covariance matrix/covariance matrix of any unbiased estimator, it can be used as a performance benchmark for this class of estimators. As already indicated in Chapter 1, our goal is to design optimal local sensor rules, in order to obtain maximal overall performance for estimating the unknown deterministic parameter at the FC, subject to bandwidth and/or power constraints of the transmit signals. It is close therefore to use the FIM, based on the final observation at the FC, as a performance indicator, due to their relation to the CRLB (cf. (2.15)). In what follows, we setup the system model and problem statement that will be considered.

3.1 System Model

Suppose, there are $L \geq 1$ sensors, each making an observation $\mathbf{y}_i \in \mathbb{R}^{m_i}$ about an unknown source, which is described by a parameter vector $\boldsymbol{\theta} \in \mathbb{R}^n$. We assume that $\boldsymbol{\theta}$ is deterministic, i.e., we have no prior information available. The relation between the sensor observation $\mathbf{y}_i$ and the parameter vector $\boldsymbol{\theta}$ is fully described by the parametrized pdf $f(\mathbf{y}_i; \boldsymbol{\theta})$. The local sensors communicate to a FC, which computes a final estimate on $\boldsymbol{\theta}$.

In most WSNs, sensors only have limited battery power and limited communication capability. For this reason, local data encoding/compression at each sensor is of importance, to reduce communication requirement between sensors and the FC. Therefore, we introduce as discussed in Chapter 1, bandwidth and power constraints on each transmit signal. We assume, that the distributed sensors have no inter–sensor communication. The role of each sensor is to encode/compress the observed local data $\mathbf{y}_i$ to a transmit data $\mathbf{s}_i$ by a mapping

$$\mathrm{LO}_i : \mathbf{y}_i \rightarrow \mathbf{s}_i \quad \text{for } 1 \leq i \leq L.$$

In what follows, we denote such an local sensor rule by writing LO. The transmit data $\mathbf{s}_i$ for $1 \leq i \leq L$ are then transmitted over a MAC to the FC, which is due to the bandwidth constraint, limited by a finite dimension. The FC produces a final estimate $\hat{\boldsymbol{\theta}}(\mathbf{z})$ of the true parameter vector by applying some fusion rule, which is a deterministic estimator function to the received vector $\mathbf{z}$ (cf. Fig. 3.1). As already mentioned in Chapter 1, we consider

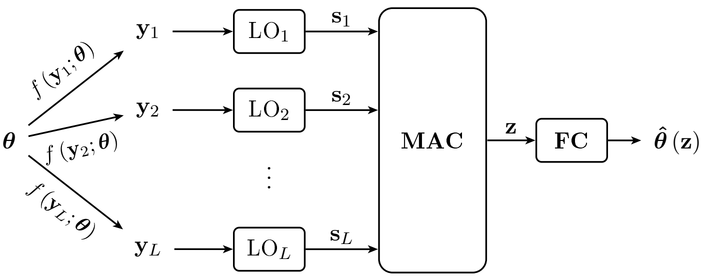


Figure 3.1: System model for a sensor network with FC and MAC.

throughout this thesis a linear Gaussian setting, i.e., every block in Fig. 3.1 corresponds to a matrix multiplication and addition of a Gaussian noise vector.

Specifically, we assume the sensor observation $\mathbf{y}_i \in \mathbb{R}^{m_i}$ are the linear combination of $\boldsymbol{\theta}$ corrupted by additive noise and can be described as

$$\mathbf{y}_i = \mathbf{G}_i\boldsymbol{\theta} + \mathbf{n}_i, \tag{3.1}$$

where $\mathbf{G}_i \in \mathbb{R}^{m_i \times n}$ is the known, deterministic observation matrix of sensor i. The additive observation noise $\mathbf{n}_i \in \mathbb{R}^{m_i}$ is assumed to be zero–mean and Gaussian distributed with fixed and known covariance matrix $\mathbf{C}_{n_i}$, i.e., $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{n_i})$. We assume that the observation noise vectors $\mathbf{n}_i$ for all i are uncorrelated across different sensors, i.e., the cross–covariance matrix

$$\operatorname{cov}\{\mathbf{n}_i, \mathbf{n}_j\} = \mathrm{E}\left\{\mathbf{n}_i\mathbf{n}_j^T\right\} = \mathbf{0} \quad \text{for } 1 \leq i, j \leq L, \quad i \neq j.$$

The main task of the ith local sensor is to map the local observed data $\mathbf{y}_i$ to a trasmit data vector $\mathbf{s}_i$ before transmitting over the channel to the FC, in order to maximize the overall performance. As already mentioned, our performance indicator is based on the CRLB or the FIM for the final observation at the FC (already introduced in Chapter 2). Since we have assumed a linear and Gaussian setup, we describe the ith local sensor rule LO_i by a deterministic matrix $\mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}$ and some additive systematic noise $\mathbf{n}_{l_i}$, which we restrict to be zero–mean and Gaussian distributed with covariance matrix $\mathbf{C}_{l_i}$, i.e., $\mathbf{n}_{l_i} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{l_i})$. Both matrices $\mathbf{A}_i$ and $\mathbf{C}_{l_i}$, fully describe our local sensor rule LO of the ith local sensor, i.e., $\mathrm{LO}_i \triangleq (\mathbf{A}_i, \mathbf{C}_{li})$. The LO_i performs a linear transformation of $\mathbf{y}_i$ and adds systematic noise $\mathbf{n}_{l_i}$, to generate the transmit data $\mathbf{s}_i$, which is given by

$$\mathbf{s}_i = \mathbf{A}_i \mathbf{G}_i \boldsymbol{\theta} + \mathbf{A}_i \mathbf{n}_i + \mathbf{n}_{li} \quad \text{for } 1 \leq i \leq L, \tag{3.2}$$

where we assume that $\mathbf{n}_i$ is uncorrelated with $\mathbf{n}_{l_j}$ (also orthogonal, since both are zero–mean) for all i and j, i.e.,

$$\mathrm{cov}\left\{\mathbf{n}_i, \mathbf{n}_{l_j}\right\} = \mathrm{E}\left\{\mathbf{n}_i \mathbf{n}_{l_j}^T\right\} = \mathbf{0} \quad \text{for } 1 \leq i, j \leq L.$$

Moreover, we also request that all systematic noise vectors $\mathbf{n}_{l_i}$ for all i are uncorrelated across different sensors, i.e.,

$$\mathrm{cov}\left\{\mathbf{n}_{l_i}, \mathbf{n}_{l_j}\right\} = \mathrm{E}\left\{\mathbf{n}_{l_i} \mathbf{n}_{l_j}^T\right\} = \mathbf{0} \quad \text{for } 1 \leq i, j \leq L, \quad i \neq j.$$

The bandwidth constraint on $\mathbf{s}_i$ leads to dimensionality condition on $\mathbf{A}_i$, i.e., $\mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}$. I.e., the ith local sensor can transmit q_i messages (real–valued symbols) to the FC, which is determined by the degrees if freedom (dimension) of the channel from sensor i to the FC and is potentially decided by the channel bandwidth [1]. The power constraint on $\mathbf{s}_i$ will be defined in the next section.

Each sensor thus transmits their encoded and/or compressed data $\mathbf{s}_i$ over a channel to the FC. Depending on the different multiple access schemes, we investigate two cases for the MAC between sensors and FC, an *orthogonal* and a *coherent* MAC [1]. For the case of an orthogonal MAC, we assume that the sensors have their own separate non–interfering channel to the FC. This can be realized, e.g., by a time–division, code–division or frequency–division multiple access scheme (TDMA/CDMA/FDMA). As for a coherent MAC, we allow all sensors transmitting simultaneously by using for example the same frequency band or time slot. Here, we assume perfect synchronization between sensors and FC, i.e., the transmitted data from all sensors reaches the FC in a coherent sum. In the following we complete our model system for both multiple access schemes and derive the corresponding expressions for the FIM.

3.1.1 Orthogonal MAC

The orthogonal MAC consists of L separate and non–interfering channels between each local sensor and the FC. The received vector $\mathbf{z}$ at the FC is given by the concatenation of L individual receive vectors $\mathbf{z}_i$ corresponding to the local sensors (cf. Fig. 3.2). The signal $\mathbf{z}_i$ received at the FC from the

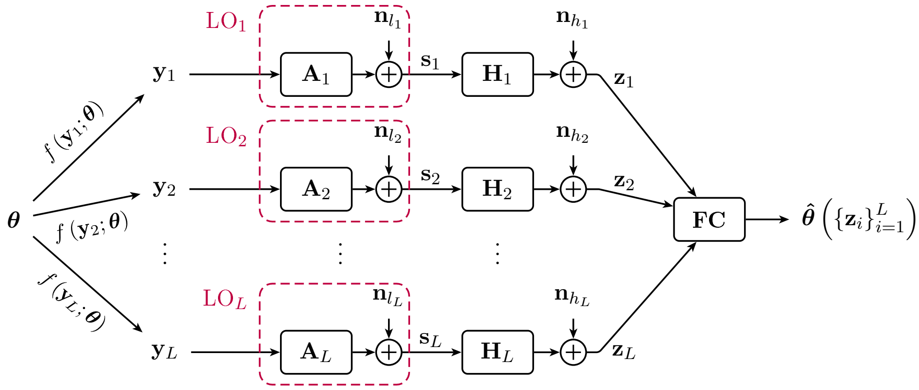


Figure 3.2: Linear decentralized estimation with orthogonal MAC.

ith local sensor, can be written as

$$\mathbf{z}_i = \mathbf{H}_i\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{H}_i\mathbf{A}_i\mathbf{n}_i + \mathbf{H}_i\mathbf{n}_{li} + \mathbf{n}_{hi} \quad \text{for } 1 \le i \le L, \tag{3.3}$$

where $\mathbf{H}_i \in \mathbb{R}^{p_i \times q_i}$ is the known, deterministic channel matrix from sensor i to the FC and $\mathbf{n}_{h_i} \in \mathbb{R}^{p_i}$ is the additive channel noise, which is again assumed to be zero–mean and Gaussian distributed with the known covariance matrix $\mathbf{C}_{h_i}$, i.e., $\mathbf{n}_{h_i} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{h_i})$. Here, we assume that the covariance matrix $\mathbf{C}_{h_i}$ is non–singular. Moreover, we assume that, firstly, $\mathbf{n}_{h_i}$ is uncorrelated with $\mathbf{n}_j$ for all i and j, i.e.,

$$\operatorname{cov}\{\mathbf{n}_{h_i}, \mathbf{n}_j\} = \mathrm{E}\left\{\mathbf{n}_{h_i}\mathbf{n}_j^T\right\} = \mathbf{0} \quad \text{for } 1 \le i, j \le L,$$

and secondly, $\mathbf{n}_{h_i}$ is uncorrelated with $\mathbf{n}_{l_j}$ for all i and j, i.e.,

$$\operatorname{cov}\left\{\mathbf{n}_{h_i}, \mathbf{n}_{l_j}\right\} = \mathrm{E}\left\{\mathbf{n}_{h_i}\mathbf{n}_{l_j}^T\right\} = \mathbf{0} \quad \text{for } 1 \le i, j \le L.$$

Additionally, we again request that all channel noise vectors $\mathbf{n}_{h_i}$ for all i are uncorrelated across different sensors, i.e.,

$$\operatorname{cov}\left\{\mathbf{n}_{h_i}, \mathbf{n}_{h_j}\right\} = \mathrm{E}\left\{\mathbf{n}_{h_i}\mathbf{n}_{h_j}^T\right\} = \mathbf{0} \quad \text{for } 1 \le i, j \le L, \quad i \ne j.$$

Further, we assume that $\mathbf{z}_i$ from (3.3) for all i are jointly Gaussian. Note that the signal model in (3.3) is an instance of the linear Gaussian model (cf. Subsection 2.2.5) with system matrix $\mathbf{H} = \mathbf{H}_i\mathbf{A}_i\mathbf{G}_i$ and noise covariance matrix $\mathbf{C}_n = \mathbf{C}_{z_i}$, where

$$\mathbf{C}_{z_i} = \mathbf{H}_i\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\mathbf{H}_i^T + \mathbf{H}_i\mathbf{C}_{l_i}\mathbf{H}_i^T + \mathbf{C}_{h_i}. \tag{3.4}$$

Therefore, we can use the expression (2.20) for the FIM of a LGM, to obtain

$$\mathbf{J}_{\mathbf{z}_i}(\boldsymbol{\theta}) = \mathbf{G}_i^T\mathbf{A}_i^T\mathbf{H}_i^T\left(\mathbf{C}_{h_i} + \mathbf{H}_i\mathbf{C}_{l_i}\mathbf{H}_i^T + \mathbf{H}_i\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\mathbf{H}_i^T\right)^{-1}\mathbf{H}_i\mathbf{A}_i\mathbf{G}_i, \tag{3.5}$$

since $\mathbf{C}_{z_i}$ (cf. (3.4)) is non–singular (i.e., invertible), due to the assumption that $\mathbf{C}_{h_i}$ for $1 \leq i \leq L$ is non–singular. The estimation problem is fully characterized by the joint pdf $f(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L; \boldsymbol{\theta})$ parametrized by $\boldsymbol{\theta}$. Note that the joint pdf can be factored as

$$f(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L; \boldsymbol{\theta}) = f(\mathbf{z}_1; \boldsymbol{\theta}) \cdot f(\mathbf{z}_2; \boldsymbol{\theta}) \ldots f(\mathbf{z}_L; \boldsymbol{\theta}) = \prod_{i=1}^{L} f(\mathbf{z}_i; \boldsymbol{\theta}), \quad (3.6)$$

since all data vectors $\mathbf{z}_i$ are statistically independent. This follows from our assumption that all in the system occuring noise vectors are uncorrelated to each other (thus $\mathrm{cov}\{\mathbf{z}_i, \mathbf{z}_j\} = \mathbf{0}$ for $1 \leq i, j \leq L$ and $i \neq j$) and the assumed joint Gaussianity of the data vectors $\mathbf{z}_i$. Hence, the FIM for the final observation $\mathbf{z} \triangleq \{\mathbf{z}_i\}_{i=1}^{L}$ at the FC, according to (2.10), can be obtained as

$$\begin{aligned}
\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = & \; \mathrm{E}\left\{\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L; \boldsymbol{\theta})\right]\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L; \boldsymbol{\theta})\right]^T\right\} \\
\overset{(3.6)}{=} & \; \mathrm{E}\left\{\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln \prod_{i=1}^{L} f(\mathbf{z}_i; \boldsymbol{\theta})\right]\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln \prod_{i=1}^{L} f(\mathbf{z}_i; \boldsymbol{\theta})\right]^T\right\} \\
= & \; \mathrm{E}\left\{\left[\frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^{L} \ln f(\mathbf{z}_i; \boldsymbol{\theta})\right]\left[\frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^{L} \ln f(\mathbf{z}_i; \boldsymbol{\theta})\right]^T\right\} \\
\overset{(a)}{=} & \; \sum_{i=1}^{L} \mathrm{E}\left\{\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{z}_i; \boldsymbol{\theta})\right]\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{z}_i; \boldsymbol{\theta})\right]^T\right\} \\
= & \; \sum_{i=1}^{L} \mathbf{J}_{\mathbf{z}_i}(\boldsymbol{\theta}),
\end{aligned} \quad (3.7)$$

where $\mathbf{J}_{\mathbf{z}_i}(\boldsymbol{\theta})$ has been already derived in (3.5). The derivation in (3.7) verifies the general composition property (2.14) of the FIM for independent data. In step (a) of (3.7), we used the linearity property of the operators $\frac{\partial}{\partial \boldsymbol{\theta}}(\cdot)$ and $\mathrm{E}\{\cdot\}$, respectively. Combining (3.5) with (3.7) yields

$$\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = \sum_{i=1}^{L} \mathbf{G}_i^T \mathbf{A}_i^T \mathbf{H}_i^T \left(\mathbf{C}_{h_i} + \mathbf{H}_i \mathbf{C}_{l_i} \mathbf{H}_i^T + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{H}_i \mathbf{A}_i \mathbf{G}_i. \quad (3.8)$$

It is important to note that the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$, according to (3.8), does not depend on the (unknown) parameter $\boldsymbol{\theta}$, which is conceptually appealing.

One immediate question that arises here is, for which conditions on the system model with an orthogonal MAC exists at least one efficient unbiased estimator. The condition for the existence of an efficient estimator is given in (2.16). Invoking [5, p.89], we can strictly follow the derivation of the efficient estimator for a LGM model. With $\mathbf{X}_i \triangleq \mathbf{H}_i \mathbf{A}_i \mathbf{G}_i$, the first derivative of $\ln f(\mathbf{z}_i; \boldsymbol{\theta})$ can thus be written as

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{z}_i; \boldsymbol{\theta}) = \mathbf{X}_i^T \mathbf{C}_{z_i}^{-1} \mathbf{z}_i - \mathbf{J}_{\mathbf{z}_i}(\boldsymbol{\theta}) \boldsymbol{\theta},$$

where $\mathbf{J}_{\mathbf{z}_i}(\boldsymbol{\theta})$ is given in (3.5) and since all $\mathbf{z}_i$ are statistically independent to each other (cf. (3.6) and the derivation in (3.7)), we obtain

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L; \boldsymbol{\theta}) &= \sum_{i=1}^{L} \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\mathbf{z}_i; \boldsymbol{\theta}) \\
&= \sum_{i=1}^{L} \mathbf{X}_i^T \mathbf{C}_{z_i}^{-1} \mathbf{z}_i - \mathbf{J}_{\mathbf{z}_i}(\boldsymbol{\theta}) \boldsymbol{\theta} \\
&\overset{(a)}{=} \mathbf{Y}\mathbf{z} - \mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) \boldsymbol{\theta} \\
&\overset{(b)}{=} \mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) \left[(\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}))^{-1} \mathbf{Y}\mathbf{z} - \boldsymbol{\theta} \right] \\
&= \mathbf{K} [\mathbf{g}(\mathbf{z}) - \boldsymbol{\theta}],
\end{aligned} \tag{3.9}$$

where in step (a), we used (3.7) and introduced

$$\begin{aligned}
\mathbf{Y} &\triangleq \begin{bmatrix} \mathbf{X}_1^T \mathbf{C}_{z_1}^{-1} & \mathbf{X}_2^T \mathbf{C}_{z_2}^{-1} & \ldots & \mathbf{X}_L^T \mathbf{C}_{z_L}^{-1} \end{bmatrix} \quad \text{and} \\
\mathbf{z} &\triangleq \begin{bmatrix} \mathbf{z}_1^T & \mathbf{z}_2^T & \ldots & \mathbf{z}_L^T \end{bmatrix}^T .
\end{aligned} \tag{3.10}$$

In step (b), we assumed that $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ is non-singular. In the last equation of (3.9) we introduced

$$\mathbf{K} \triangleq \mathbf{Y}\mathbf{X} \overset{(3.10)}{=} \sum_{i=1}^{L} \mathbf{X}_i^T \mathbf{C}_{z_1}^{-1} \mathbf{X}_i \overset{(3.8)}{=} \mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$$

and

$$\begin{aligned}
\mathbf{g}(\mathbf{z}) \triangleq (\mathbf{Y}\mathbf{X})^{-1} \mathbf{Y}\mathbf{z} &\overset{(3.10)}{=} \left(\sum_{i=1}^{L} \mathbf{X}_i^T \mathbf{C}_{z_1}^{-1} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{L} \mathbf{X}_i^T \mathbf{C}_{z_i}^{-1} \mathbf{z}_i \right) \\
&= \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta}) \left(\sum_{i=1}^{L} \mathbf{G}_i^T \mathbf{A}_i^T \mathbf{H}_i^T \mathbf{C}_{z_i}^{-1} \mathbf{z}_i \right).
\end{aligned}$$

Comparing with (2.16), we conclude the following:

- It exists an efficient estimator (cf. (2.17)), which is given by

$$\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L) = \mathbf{g}(\mathbf{z}) = \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta}) \left(\sum_{i=1}^{L} \mathbf{G}_i^T \mathbf{A}_i^T \mathbf{H}_i^T \mathbf{C}_{z_i}^{-1} \mathbf{z}_i \right), \tag{3.11}$$

and is simultaneously the MVU estimator. It exists iff $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ and $\mathbf{C}_{z_i}$ for all i are non–singular.

- The FIM is $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = \mathbf{K}$; it does not depend on $\boldsymbol{\theta}$.
- The estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L)$ from (3.11) is obviously unbiased, $\mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L)\right\} = \boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, and its covariance matrix (cf. (2.18)) is given by

$$\operatorname{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_L)\right\} = \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta}). \tag{3.12}$$

This is the efficient estimator (MVU estimator) for our system model with an orthogonal MAC. It exists if

$$\begin{aligned} &\mathbf{J}_{\mathbf{z}} \text{ is non–singular and} \\ &\mathbf{C}_{z_i} \text{ is non–singular} \quad \text{for all } 1 \leq i \leq L, \end{aligned} \tag{3.13}$$

where again, the last condition is guaranteed, to due our assumption that $\mathbf{C}_{h_i}$ is non–singular for all i.

3.1.2 Coherent MAC

As a second model for the link between local sensors and FC, we consider the case of coherent MAC. Here, the individual transmit signals $\mathbf{s}_i$ of the local sensors, add up at the FC in a coherent sum (signals are perfectly synchronized between sensors and FC[1]). We also assume that all by the channel corrupted trasmitted data vectors have the same length $p = p_i$ for $1 \leq i \leq L$. Then, we can use the following observation model at the FC (see Fig. 3.3),

$$\mathbf{z} = \sum_{i=1}^{L} \mathbf{H}_i \mathbf{s}_i + \mathbf{n}_h = \sum_{i=1}^{L} \left(\mathbf{H}_i \mathbf{A}_i \mathbf{G}_i \boldsymbol{\theta} + \mathbf{H}_i \mathbf{A}_i \mathbf{n}_i + \mathbf{H}_i \mathbf{n}_{l_i}\right) + \mathbf{n}_h, \tag{3.14}$$

where again $\mathbf{H}_i \in \mathbb{R}^{p \times q_i}$ is the known, deterministic channel matrix from sensor i to the FC and $\mathbf{n}_h \in \mathbb{R}^p$ is the additive channel noise, which is again assumed to be zero–mean and Gaussian distributed with the known covariance matrix $\mathbf{C}_h$, i.e., $\mathbf{n}_h \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_h)$. As in the orthogonal MAC case, we assume that $\mathbf{n}_h$ is uncorrelated with $\mathbf{n}_j$ for all j, i.e.,

$$\operatorname{cov}\{\mathbf{n}_h, \mathbf{n}_j\} = \mathrm{E}\left\{\mathbf{n}_h \mathbf{n}_j^T\right\} = \mathbf{0} \quad \text{for } 1 \leq j \leq L,$$

[1] In the orthogonal MAC case, we only need to assume pair–wise synchronization between each sensor and the FC, where synchronization among different sensors is not required.

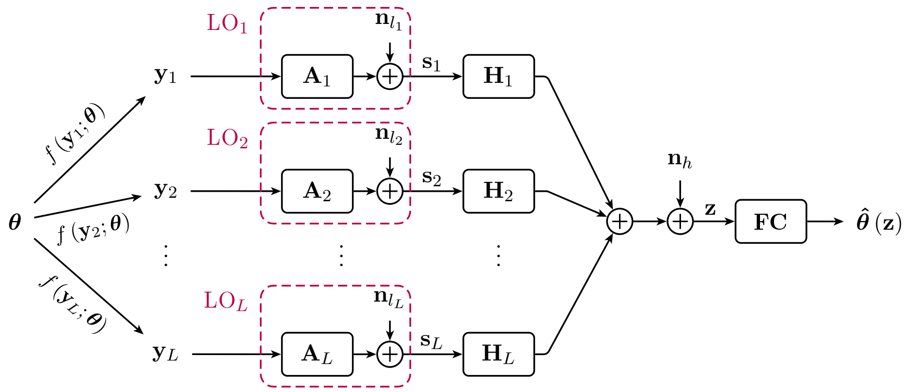


Figure 3.3: Linear decentralized estimation with coherent MAC

and further that $\mathbf{n}_h$ is uncorrelated with $\mathbf{n}_{l_j}$ for all j, i.e.,

$$\operatorname{cov}\left\{\mathbf{n}_h, \mathbf{n}_{l_j}\right\} = \mathrm{E}\left\{\mathbf{n}_h \mathbf{n}_{l_j}^T\right\} = \mathbf{0} \quad \text{for } 1 \leq j \leq L.$$

Unless otherwise stated, we assume that the channel matrices $\mathbf{H}_i$ for $1 \leq i \leq L$ are of full column–rank and the noise covariance matrix $\mathbf{C}_h$ is non–singular.

Let us introduce the shorthand

$$\widetilde{\mathbf{A}}_i \triangleq \mathbf{H}_i \mathbf{A}_i \in \mathbb{R}^{p \times m_i}. \tag{3.15}$$

Due to our assumption, that $\mathbf{H}_i$ has full column–rank, implying $p \geq q_i$, we can reobtain $\mathbf{A}_i$ from $\widetilde{\mathbf{A}}_i$ via $\mathbf{A}_i = \mathbf{H}_i^{\dagger} \widetilde{\mathbf{A}}_i$. Here, $\mathbf{H}_i^{\dagger}$ denotes the pseudo–inverse of $\mathbf{H}_i$ and is given by $\mathbf{H}_i^{\dagger} = \left(\mathbf{H}^T \mathbf{H}\right)^{-1} \mathbf{H}^T$. We refer to $\widetilde{\mathbf{A}}_i$ as the sensor–channel matrix of sensor i. Let us furthmore define

$$\begin{aligned} \widetilde{\mathbf{A}} &\triangleq \begin{bmatrix} \widetilde{\mathbf{A}}_1 & \widetilde{\mathbf{A}}_2 & \dots & \widetilde{\mathbf{A}}_L \end{bmatrix}, \quad \widetilde{\mathbf{A}} \in \mathbb{R}^{p \times k}, \\ \mathbf{G} &\triangleq \begin{bmatrix} \mathbf{G}_1^T & \mathbf{G}_2^T & \dots & \mathbf{G}_L^T \end{bmatrix}^T, \quad \mathbf{G} \in \mathbb{R}^{k \times n}, \\ \mathbf{H} &\triangleq \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \dots & \mathbf{H}_L \end{bmatrix}, \quad \mathbf{H} \in \mathbb{R}^{p \times q}, \end{aligned} \tag{3.16}$$

where $k = m_1 + m_2 + \cdots + m_L$ and $q = q_1 + q_2 + \cdots + q_L$. Let us refer to $\widetilde{\mathbf{A}}$ and $\mathbf{G}$ as the total sensor–channel matrix and the total observation matrix, respectively. Analog, we define the total observation noise and the total systematic noise by

$$\begin{aligned} \mathbf{n} &\triangleq \begin{bmatrix} \mathbf{n}_1^T & \mathbf{n}_2^T & \dots & \mathbf{n}_L^T \end{bmatrix}^T, \quad \mathbf{n} \in \mathbb{R}^{k}, \\ \mathbf{n}_l &\triangleq \begin{bmatrix} \mathbf{n}_{l1}^T & \mathbf{n}_{l2}^T & \dots & \mathbf{n}_{lL}^T \end{bmatrix}^T, \quad \mathbf{n}_l \in \mathbb{R}^{q}. \end{aligned} \tag{3.17}$$

Using the notations in (3.16) and (3.17), we can write the observation $\mathbf{z}$ at the FC, given in (3.14), in the form

$$\mathbf{z} = \widetilde{\mathbf{A}} \mathbf{G} \boldsymbol{\theta} + \widetilde{\mathbf{A}} \mathbf{n} + \mathbf{H} \mathbf{n}_l + \mathbf{n}_h. \tag{3.18}$$

Note that the signal model in (3.18) is again an instance of the linear Gaussian model (cf. Subsection 2.2.5) with system matrix $\mathbf{H} = \widetilde{\mathbf{A}}\mathbf{G}$ and noise covariance matrix $\mathbf{C}_n = \mathbf{C}_z$, where

$$\mathbf{C}_z = \mathbf{C}_h + \mathbf{H}\mathbf{C}_l\mathbf{H}^T + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T. \tag{3.19}$$

Therefore, we can use the expression (2.20) for the FIM of a LGM, to obtain

$$\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta}) = \mathbf{G}^T\widetilde{\mathbf{A}}^T\left(\mathbf{C}_h + \mathbf{H}\mathbf{C}_l\mathbf{H}^T + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T\right)^{-1}\widetilde{\mathbf{A}}\mathbf{G}. \tag{3.20}$$

We can invoke (2.20), since the covariance matrix $\mathbf{C}_z$ (cf. (3.19)) is non–singular. This is guaranteed by the assumption that $\mathbf{C}_h$ is non–singular. Note that the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ in (3.20) does not depend on the parameter $\boldsymbol{\theta}$.

As discussed at the end of Subsection 3.1.1, we will now analyze conditions on our system model with a coherent MAC, such that it exists at least one efficient unbiased estimator. Now, it is much easier to find an efficient estimator as in the orthogonal MAC, since we can directly use the derivation for a simple LGM in [5, p.89], and conclude that it exists an efficient unbiased estimator (MVU estimator) iff the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ from (3.20) and $\mathbf{C}_z$ from (3.19) are both non-singular. Then, we conclude the following:

- It exists an efficient estimator, which is given in (2.21) for system matrix $\mathbf{H} = \widetilde{\mathbf{A}}\mathbf{G}$ (not be confused with $\mathbf{H}$ from (3.16)) and noise covariance matrix $\mathbf{C}_n = \mathbf{C}_z$, where $\mathbf{H}$ has full column–rank n and $\mathbf{C}_n$ is non–singular, i.e.,

$$\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z}) = \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta})\,\mathbf{G}^T\widetilde{\mathbf{A}}^T\mathbf{C}_z^{-1}\mathbf{z}, \tag{3.21}$$

  and is simultaneously the MVU estimator.

- The estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z})$ from (3.21) is obviously unbiased, $\mathrm{E}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z})\right\} = \boldsymbol{\theta}$ for all $\boldsymbol{\theta}$, and its covariance matrix is given by

$$\mathrm{cov}_{\boldsymbol{\theta}}\left\{\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{z})\right\} = \mathbf{J}_{\mathbf{z}}^{-1}(\boldsymbol{\theta}). \tag{3.22}$$

This is the efficient estimator (MVU estimator) for our system model with a coherent MAC. It exists iff $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ and $\mathbf{C}_z$ are both non–singular. Note that we already assumed that the channel noise covariance matrix $\mathbf{C}_h$ is non–singular and thus $\mathbf{C}_z$ is non–singular (cf. (3.19)).

Lemma 3.1.3 *Consider matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$, then*

$$\mathrm{rank}(\mathbf{A}) + \mathrm{rank}(\mathbf{B}) - n \leq \mathrm{rank}(\mathbf{AB}) \leq \min\{\mathrm{rank}(\mathbf{A}), \mathrm{rank}(\mathbf{B})\}.$$

Proof. see [7, Lemma 2.1, p.16]. □

Let us finally derive conditions on the system matrices $\widetilde{\mathbf{A}}$ and $\mathbf{G}$, such that it exists an efficient estimator (MVU estimator), which is given in (3.21). This occurs iff the matrix product $\mathbf{H} = \widetilde{\mathbf{A}}\mathbf{G}$ has full column–rank. According to Lemma 3.1.3, we conclude that $\operatorname{rank}\left(\widetilde{\mathbf{A}}\mathbf{G}\right) = n$ when $\operatorname{rank}(\mathbf{G}) = n$ and $n \leq \operatorname{rank}\left(\widetilde{\mathbf{A}}\right) \leq k$. Hence, the conditions for the existence of at least one efficient, unbiased estimator on our system model with a coherent MAC can be summarized as follows:

$$\begin{aligned}
&n \leq \operatorname{rank}\left(\widetilde{\mathbf{A}}\right) \leq k, \quad \operatorname{rank}(\mathbf{G}) = n, \\
&p \geq n, \quad k = m_1 + m_2 + \cdots + m_L \geq n, \\
&\text{and} \\
&\mathbf{C}_z \text{ is non–singular,}
\end{aligned} \tag{3.23}$$

where again, the last condition is guaranteed, to due our assumption that $\mathbf{C}_h$ is non–singular.

So far, we have derived the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ for both multiple access schemes. In both cases, it is important to note that, due to the assumption on a LGM, the FIM $\mathbf{J}_{\mathbf{z}}(\boldsymbol{\theta})$ does not depend on the parameter $\boldsymbol{\theta}$. Thus, we simply write $\mathbf{J}_{\mathbf{z}}$ in what follows. Before, we define our optimization problem in detail, we will now introduce the power constraint on the transmit data $\mathbf{s}_i \in \mathbb{R}^{q_i}$ for $1 \leq i \leq L$, which seems in addition to the already mentioned bandwidth constraint.

3.2 Power Constraint

Remember that our goal is to determine each LO_i for $1 \leq i \leq L$, such that the FIM or, equivalently, the CRLB for the observation at the FC is optimized. In WSN, energy efficiency is highly desirable, e.g., due to using battery powered devices and changing battery is not possible easily. Hence, each local sensor has only limited power available for transmitting the prepared data $\mathbf{s}_i$ to the FC over the channel. Therefore, we have to introduce an appropriate power constraint for the transmitted data $\mathbf{s}_i$. On the other hand, without considering such a power constraint, we can always ensure ideal links between sensors and the FC, by scaling the sensor matrices $\mathbf{A}_i$ for $1 \leq i \leq L$, with an arbitrarily large factor. Throughout this thesis, we consider two types of power constraints. The first, more natural power constraint, is given by

$$E_{\boldsymbol{\theta}}\left\{\|\mathbf{s}_i\|^2\right\} \leq P_{0,i} \quad \text{for } 1 \leq i \leq L, \tag{C1}$$

which is the mean power of the transmit data $\mathbf{s}_i$. The constant $P_{0,i}$ denotes the known, maximum power for $\mathbf{s}_i$, which we allow for sensor i. The second power constraint reads

$$\mathrm{var}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\} = E_{\boldsymbol{\theta}}\left\{\left\|\mathbf{s}_i - E_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\}\right\|^2\right\} \leq P'_{0,i} \quad \text{for } 1 \leq i \leq L, \tag{C2}$$

which consider in contrast to (C1), the variance of the transmit data $\mathbf{s}_i$. The constant $P'_{0,i}$ denotes the known, maximum (variance) power for $\mathbf{s}_i$, correspondingly. It is important to note that the expectation in (C1)) and (C2) is only with respect to $\mathbf{s}_i$, since $\boldsymbol{\theta}$ is modeled as deterministic. The subscript $\boldsymbol{\theta}$ in (C1) and (C2) indicates that the expectation of $\mathbf{s}_i$ depends on $\boldsymbol{\theta}$ in general. Both power constraints have their justification. Note that the constraint (C2) is equivalent to (C1), if we choose

$$P'_{0,i} = P_{0,i} - \left\|\mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\}\right\|^2, \tag{3.24}$$

which follows directly from the identity

$$\mathrm{var}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\} = E_{\boldsymbol{\theta}}\left\{\mathbf{s}_i^T\mathbf{s}_i\right\} - \left\|\mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\}\right\|^2. \tag{3.25}$$

The optimum design for each local sensor with consideration to constraint (C1), will in all likelihood depend on the unknown parameter vector $\boldsymbol{\theta}$, which makes an implementation not practicable. However, we can estimate the parameter $\boldsymbol{\theta}$, first locally, at each local sensor, i.e., for sensor i, we compute an estimate $\hat{\boldsymbol{\theta}}_{\mathrm{LO}_i}$. With the estimate $\hat{\boldsymbol{\theta}}_{\mathrm{LO}_i}$, we are then able to design a, of course, sub–optimum LO_i. Note that the ith local sensor has to redesign itself dynamically, according to the value of the estimate $\hat{\boldsymbol{\theta}}_{\mathrm{LO}_i}$.

Let us now specialize the constraints (C1) and (C2) to our specific system model. The expected power of the transmit data $\mathbf{s}_i$, i.e., $\mathrm{E}_{\boldsymbol{\theta}}\left\{\left\|\mathbf{s}_i\right\|^2\right\}$, where $\mathbf{s}_i$ is given in (3.2), can be expressed as

$$\begin{aligned}
\mathrm{E}_{\boldsymbol{\theta}}\left\{\left\|\mathbf{s}_i\right\|^2\right\} &= E_{\boldsymbol{\theta}}\left\{\mathbf{s}_i^T\mathbf{s}_i\right\} \\
&= \mathrm{E}_{\boldsymbol{\theta}}\left\{\left(\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{A}_i\mathbf{n}_i + \mathbf{n}_{l_i}\right)^T\left(\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{A}_i\mathbf{n}_i + \mathbf{n}_{l_i}\right)\right\} \\
&\overset{(a)}{=} \mathrm{E}_{\boldsymbol{\theta}}\left\{\boldsymbol{\theta}^T\mathbf{G}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{n}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{n}_i + \mathbf{n}_{l_i}^T\mathbf{n}_{l_i}\right\} \\
&\quad + \mathrm{E}_{\boldsymbol{\theta}}\left\{\boldsymbol{\theta}^T\mathbf{G}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{n}_i + \mathbf{n}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{n}_{l_i}^T\mathbf{A}_i\mathbf{n}_i\right. \\
&\qquad \left. + \boldsymbol{\theta}^T\mathbf{G}_i^T\mathbf{A}_i^T\mathbf{n}_{l_i} + \mathbf{n}_i^T\mathbf{A}_i^T\mathbf{n}_{l_i} + \mathbf{n}_{l_i}^T\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\right\} \\
&\overset{(b)}{=} \mathrm{E}_{\boldsymbol{\theta}}\left\{\boldsymbol{\theta}^T\mathbf{G}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\right\} + \mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{n}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{n}_i\right\} + \mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{n}_{l_i}^T\mathbf{n}_{l_i}\right\} \\
&\overset{(c)}{=} \left\|\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\right\|^2 + \mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\}.
\end{aligned} \tag{3.26}$$

In step (a) and (b), we used the linearity of the expectation operator $\mathrm{E}\left\{\cdot\right\}$. In step (b), we used the fact that observation noise $\mathbf{n}_i$ and systematic noise $\mathbf{n}_{l_i}$ have been accepted as zero–mean and uncorrelated to each other for all

i, i.e., $\mathrm{cov}\left\{\mathbf{n}_i, \mathbf{n}_{l_i}\right\} = \mathrm{E}\left\{\mathbf{n}_i\mathbf{n}_{l_i}^T\right\} = \mathbf{0}$ for $1 \leq i \leq L$. The second expectation term in (a) is thus zero. In the last step (c), we obtain the final expression for the expected power on $\mathbf{s}_i$. If we write $\|\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\|^2 = \boldsymbol{\theta}^T\mathbf{G}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}$, we can reformulate the last equation of (3.26) with the trace operator property $a = \mathrm{tr}\left\{a\right\}$ for $a \in \mathbb{R}$ [8], also in the more compact form as

$$\begin{aligned}
\mathrm{E}_{\boldsymbol{\theta}}\left\{\|\mathbf{s}_i\|^2\right\} &= \mathrm{tr}\left\{\boldsymbol{\theta}^T\mathbf{G}_i^T\mathbf{A}_i^T\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\right\} + \mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} \\
&\overset{(a)}{=} \mathrm{tr}\left\{\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\boldsymbol{\theta}^T\mathbf{G}_i^T\mathbf{A}_i^T + \mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} \\
&= \mathrm{tr}\left\{\mathbf{A}_i\left(\mathbf{G}_i\boldsymbol{\theta}\boldsymbol{\theta}^T\mathbf{G}_i^T + \mathbf{C}_{n_i}\right)\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} \\
&\overset{(b)}{=} \mathrm{tr}\left\{\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\},
\end{aligned} \tag{3.27}$$

where in step (a) we used the cyclic property and the linearity of the trace operator $\mathrm{tr}\left\{\cdot\right\}$ [8] and in step (b) we introduced the matrix

$$\mathbf{M}_i \triangleq \mathbf{G}_i\boldsymbol{\theta}\boldsymbol{\theta}^T\mathbf{G}_i^T + \mathbf{C}_{n_i}. \tag{3.28}$$

The variance of $\mathbf{s}_i$, i.e., $\mathrm{var}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\}$, can be directly obtained by inserting the last equation of (3.26) into (3.25), i.e.,

$$\begin{aligned}
\mathrm{var}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\} &= \|\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\|^2 + \mathrm{tr}\left(\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right) + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} - \|\mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\}\|^2 \\
&\overset{(a)}{=} \|\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\|^2 + \mathrm{tr}\left(\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right) + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} - \|\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}\|^2 \\
&= \mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\},
\end{aligned} \tag{3.29}$$

where in step (a) we insert the mean of $\mathbf{s}_i$, as can be verified easily by computing the expectation of $\mathbf{s}_i$, i.e., $\mathrm{E}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\} = \mathbf{G}_i\mathbf{A}_i\boldsymbol{\theta}$. Note that from (3.29), we conclude that the second power constraint (C2) does not depend on the parameter $\boldsymbol{\theta}$, which is indeed unknown.

Let us summarize both power constraints, (C1) and (C2), by

$$\mathrm{E}_{\boldsymbol{\theta}}\left\{\|\mathbf{s}_i\|^2\right\} = \mathrm{tr}\left\{\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} \leq P_{0,i} \quad \text{for } 1 \leq i \leq L, \tag{C1}$$

where $\mathbf{M}_i$ is given in (3.28), and

$$\mathrm{var}_{\boldsymbol{\theta}}\left\{\mathbf{s}_i\right\} = \mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} \leq P'_{0,i} \quad \text{for } 1 \leq i \leq L, \tag{C2}$$

obtained from the last equations in (3.27) and (3.29), respectively.

3.3 Problem Formulation

We are now able to define our basic design problem in a general form. Inspired by the CRLB for the MSE of the MVU, as discussed in the previous chapter, we choose the local sensor rules LO_i for $1 \leq i \leq L$, such that the CRLB is minimized or equivalently the FIM is maximized, w.r.t. the Loewner ordering.

3.3.1 The Loewner Optimality

Remember from Subsection 2.2.2, that the CRLB is the inverse of the FIM (cf. (2.15)).

Corollary 3.3.4 *Let* $\mathbf{A}, \mathbf{B} \in \mathrm{PD}(k)$*, then*

$$\mathbf{A} \geq \mathbf{B} \quad \Leftrightarrow \quad \mathbf{B}^{-1} \geq \mathbf{A}^{-1}.$$

Proof. see [9, p.471, corollary 7.7.4] □

According to Corollary 3.3.4, we conclude that minimizing the CRLB is equivalent to maximizing the FIM, w.r.t. the Loewner ordering. Hence, we can define our optimal LO_i, in the sense of maximizing the FIM $\mathbf{J_z}$ or minimizing the CRLB $\mathbf{J_z}^{-1}$.

Definition 3.3.5 *A local sensor rule* LO_i*, given by* $(\mathbf{A}_i, \mathbf{C}_{l_i})$*, is called Loewner optimal, when the FIM* $\mathbf{J_z}(\mathbf{A}_i, \mathbf{C}_{l_i})$ *satisfies*

$$\mathbf{J_z}\left(\mathbf{A}_i^*, \mathbf{C}_{l_i}^*\right) \geq \mathbf{J_z}\left(\mathbf{A}_i, \mathbf{C}_{l_i}\right),$$

or equivalently,

$$\mathbf{J_z}^{-1}\left(\mathbf{A}_i^*, \mathbf{C}_{l_i}^*\right) \leq \mathbf{J_z}^{-1}\left(\mathbf{A}_i, \mathbf{C}_{l_i}\right),$$

for all $(\mathbf{A}_i, \mathbf{C}_{l_i})$*. The pair* $(\mathbf{A}_i^*, \mathbf{C}_{l_i}^*)$*, which satisfies the power constraint* (C1) *or* (C2)*, respectively, denotes the (Loewner) optimal* LO_i *for sensor* i*, denoted by* $\mathrm{LO}_i^* \triangleq (\mathbf{A}_i^*, \mathbf{C}_{l_i}^*)$.

The question arises how to maximize a matrix valued FIM in the sense of Loewner optimality. For a scalar parameter, i.e., $\theta \in \mathbb{R}$ and $n = 1$, the FIM reduces to a scalar function on $\mathrm{LO}_i = (\mathbf{A}_i, \mathbf{C}_{l_i})$. In that case we have to maximize a scalar real-valued function under the constraint (C1) or (C2). For the general case, i.e., for $n \geq 1$, the focus includes all parameters to be estimated. Such optimal designs are considered in [4], where they introduced real-valued optimality criteria.

3.3.2 Real–Valued Optimality Criteria

In this subsection, we introduce real–valued optimality criteria, i.e., real–valued functions, which measure (in some sense) the "largeness" of an information matrix. Thus, an optimality criteria is a real–valued function ϕ from the domain of positive semi–definite matrices (i.e., on the closed cone $\mathrm{NND}(s)$) into the real line,

$$\phi : \mathrm{NND}(s) \to \mathbb{R}. \tag{3.30}$$

The function ϕ should capture the idea of whether an information matrix (an information matrix includes the class of FIMs as special cases, cf. [4, Chapter

3]) is large or small. It is important to note, that such a transformation from the high dimensional matrix cone to the one dimensional real line, can only retain partial aspects. Let $\mathbf{C}$ and $\mathbf{D}$ be two information matrices of size $s \times s$. The main properties, which have to be satisfied by those class of functions are [4, Chapter 4]:

- *Isotonic*: The main aspect of an optimality criterion ϕ is the ordering among information matrices. They are isotonic relative to the Loewner ordering, i.e.,

$$\mathbf{C} \geq \mathbf{D} \geq \mathbf{0} \Leftrightarrow \phi\left\{\mathbf{C}\right\} \geq \phi\left\{\mathbf{D}\right\}. \tag{3.31}$$

- *Concativity*, i.e.,

$$\phi\left\{(1-\alpha)\,\mathbf{C} + \alpha\mathbf{D}\right\} \geq (1-\alpha)\,\phi\left\{\mathbf{C}\right\} + \alpha\phi\left\{\mathbf{D}\right\}, \tag{3.32}$$

  for all $\alpha \in [0;1]$, $\mathbf{C}, \mathbf{D} \geq 0$.

- *Positive homogeneity*, i.e.,

$$\phi\left(\delta\mathbf{C}\right) = \delta\phi\left(\mathbf{C}\right) \quad \text{for all } \delta > 0,\ \mathbf{C} > \mathbf{0}. \tag{3.33}$$

- *Super–additive*, i.e.,

$$\phi\left\{\mathbf{C} + \mathbf{D}\right\} \geq \phi\left\{\mathbf{C}\right\} + \phi\left\{\mathbf{D}\right\} \quad \text{for all } \mathbf{C}, \mathbf{D} \geq \mathbf{0}. \tag{3.34}$$

- *Non–negative*, i.e.,

$$\phi\left(\mathbf{C}\right) \geq 0 \quad \text{for all } \mathbf{C} \geq \mathbf{0}, \tag{3.35}$$

  and *positive*, iff

$$\phi\left(\mathbf{C}\right) > 0 \quad \text{for all } \mathbf{C} > \mathbf{0}. \tag{3.36}$$

  Thus, information can never be negative. Notice, that positive homogeneity (3.33) implies that ϕ vanishes for the null matrix, $\phi\left(\mathbf{0}\right) = 0$, because $\phi\left(\mathbf{0}\right) = \phi\left(2 \cdot \mathbf{0}\right) = 2\phi\left(\mathbf{0}\right)$.

- *Non–constant*, i.e.,

$$\phi\left\{\mathbf{C}\right\} = \phi\left\{\mathbf{D}\right\} \quad \Leftrightarrow \quad \mathbf{C} = \mathbf{D}. \tag{3.37}$$

- *Upper–semicontinuity*, i.e.,

$$\left\{\phi \geq \alpha\right\} = \left\{\mathbf{C} \in \mathrm{NND}\left(s\right) : \phi\left\{\mathbf{C}\right\} \geq \alpha\right\} \tag{3.38}$$

  are closed, for all $\alpha \in \mathbb{R}$.

The most prominent optimality criteria are the so-called *matrix means* (cf. [4]), which are denoted by ϕ_p for $p \in [-\infty; 1]$ (throughout this subsection, p denotes the paramater for matrix means; not be confused with the dimension p of the observation vector $\mathbf{z}$ in our system model).

Definition 3.3.6 *For positive semi–definite matrices, $\mathbf{C} \in \mathrm{NND}(s)$, the matrix mean ϕ_p is represented by [4, Sec. 6.7]*

$$\phi_p(\mathbf{C}) = \begin{cases} \lambda_{\max}(\mathbf{C}) & p = \infty, \\ \left(\frac{1}{s}\mathrm{tr}\left\{\mathbf{C}^p\right\}\right)^{1/p} & p \in (-\infty, 0) \cap (0, \infty), \\ \left(\det\left\{\mathbf{C}\right\}\right)^{1/s} & p = 0, \\ \lambda_{\min}(\mathbf{C}) & p = -\infty, \end{cases}$$

where $\lambda_{\max}(\mathbf{C})$ and $\lambda_{\min}(\mathbf{C})$ denote the largest- and smallest eigenvalue of $\mathbf{C}$, respectively.

For the parameter $p \in [-\infty; 1]$, all stated properties (3.31)-(3.38) for an optimality criterion function ϕ are satisfied [4, Sec. 6.7]. This family of optimality criteria functions contain the well–known criteria termed *D–, A–, E– and T–optimality* as special cases.

Let us now consider the A– and T–optimality criteria in more detail, since throughout this thesis, we will consider only these two particular examples of the ϕ_p–family, with parameter $p = \{1, -1\}$. In what follows, let $\mathbf{C}$ be a FIM of size $s \times s$, i.e., $\mathbf{C} \in \mathrm{NND}(s)$ and furthermore $\mathbf{C} \in \mathrm{Sym}(s)$.

A–Criterion: The A–criterion (also known under the name *average-variance* criterion) can be obtained for $p = -1$, i.e., $\phi_{-1}(\mathbf{C})$ is defined by

$$\phi_{-1}(\mathbf{C}) \triangleq \left(\frac{1}{s}\mathrm{tr}\left\{\mathbf{C}^{-1}\right\}\right)^{-1}, \tag{3.39}$$

for a non–singular $\mathbf{C}$. An A-optimal design minimizes the MSE of an efficient unbiased estimator (cf. (2.9), since $\mathbf{C}^{-1}$ is the CRLB). Note that maximizing the average–variance criterion ϕ_{-1} among information matrices is the same as minimizing

$$\frac{1}{\phi_{-1}(\mathbf{C})} = \frac{1}{s}\mathrm{tr}\left\{\mathbf{C}^{-1}\right\}. \tag{3.40}$$

T–Criterion: The T–criterion (also known under the name *trace* criterion) can be obtained for $p = 1$, i.e., $\phi_1(\mathbf{C})$ is defined by,

$$\phi_1(\mathbf{C}) \triangleq \frac{1}{s}\mathrm{tr}\left\{\mathbf{C}\right\}. \tag{3.41}$$

The usage of the T–optimality criterion has no direct practical justification. However, the T–criterion has the appealing property of being linear, i.e.,

$$\begin{aligned}\phi_1\left(k_1\mathbf{C}_1+k_2\mathbf{C}_2\right)&=\frac{1}{s}\mathrm{tr}\left\{k_1\mathbf{C}_1+k_2\mathbf{C}_2\right\}\\&=k_1\frac{1}{s}\mathrm{tr}\left\{\mathbf{C}_1\right\}+k_2\frac{1}{s}\mathrm{tr}\left\{\mathbf{C}_2\right\}\\&=k_1\phi_1\left(\mathbf{C}_1\right)+k_2\phi\left(\mathbf{C}_2\right),\end{aligned}\tag{3.42}$$

where $\mathbf{C}_1$ and $\mathbf{C}_2$ are two arbitrary FIMs and $k_1, k_2 \in \mathbb{R}^+$. Deriving the T–optimal solution might give some intuition about the structure of the optimal rules for different criteria, e.g., the A–optimality criterion.

With the introduction on optimality criteria, we are now able to define our basic design problem.

3.3.3 The Basic Optimization Problem

In the previous subsection, we have introduced real-valued optimality criteria, which enables us to measure information of FIMs. Given such an optimality criterion ϕ, which is defined on $\mathrm{NND}\,(n)$, the basic optimization problem then reads

$$\begin{array}{ll}\underset{\substack{\mathrm{LO}_i=\left(\mathbf{A}_i,\mathbf{C}_{l_i}\right)\\1\le i\le L}}{\text{maximize}} & \phi\left\{\mathbf{J}_\mathbf{z}\right\}\\ \text{subject to} & \mathbf{J}_\mathbf{z}\text{ satisfies (3.8) or (3.20)}\\ & \mathbf{A}_i\in\mathbb{R}^{q_i\times m_i}\quad\text{for }1\le i\le L,\\ & \mathbf{C}_{l_i}\in\mathbb{R}^{q_i\times q_i}:\mathbf{C}_{l_i}\ge\mathbf{0}\quad\text{for }1\le i\le L,\\ & \text{(C1) or (C2) is satisfied, i.e.,}\\ & \text{(C1)}:\ \mathrm{tr}\left\{\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^T\right\}+\mathrm{tr}\left\{\mathbf{C}_{l_i}\right\}\le P_{0,i}\quad\text{for }1\le i\le L,\\ & \text{(C2)}:\ \mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\}+\mathrm{tr}\left\{\mathbf{C}_{l_i}\right\}\le P'_{0,i}\quad\text{for }1\le i\le L.\end{array}\tag{P-I}$$

Definition 3.3.7 *A local sensor rule, given by* $(\mathbf{A}_i, \mathbf{C}_{li})$*, which solves* (P-I)*, is said to be formally* ϕ*-optimal and is denoted by* $\mathrm{LO}^*_{i\,\phi} \triangleq \left(\mathbf{A}^*_i, \mathbf{C}^*_{l_i}\right)_\phi$.

The FIM $\mathbf{J}_\mathbf{z}$ is given in (3.8) for a system with an orthogonal MAC and in (3.20) for a system with a coherent MAC, respectively. This calls for maximizing information as measured by an optimality criterion ϕ. For solving (P-I), we assume that the FC has perfect knowledge of the observation model and the channel states, i.e., the matrices $\{\mathbf{G}_i, \mathbf{C}_{n_i}, \mathbf{H}_i\}_{i=1}^L$, moreover, the channel noise covariance matrices $\{\mathbf{C}_{h_i}\}_{i=1}^L$ (for the orthogonal MAC case) and $\mathbf{C}_h$ (for the coherent MAC case), are assumed to be known. This assumption is reasonable for a quasi-static scenario, i.e., when $\{\mathbf{G}_i, \mathbf{C}_{n_i}, \mathbf{H}_i\}_{i=1}^L$ and $\{\mathbf{C}_{h_i}\}_{i=1}^L$ or $\mathbf{C}_h$, respectively, change slowly in a quasi-static manner. The optimization problem in the form (P-I), with an optimality criteria ϕ

is the most general one and is also called ϕ-optimization problem, since the solution, i.e., $\mathrm{LO}^*_{i_\phi} = \left(\mathbf{A}^*_i, \mathbf{C}^*_{l_i}\right)_\phi$ for $1 \leq i \leq L$ depends on the choice of the optimal design criterion ϕ. Note that we do not necessarily require the existence of an efficient estimator (cf. conditions in (3.13) or (3.23), respectively) - we also study cases, when the FIM is singular, unless permitted by the specific optimality criterion ϕ, e.g., the T-optimality criterion.

After having formulated our basic optimization problem (P-I), we will now present the solutions for (P-I), separately for the case of a coherent and an orthogonal MAC. Within this work, we focus on solving (P-I) for the orthogonal MAC case, first for a scalar parameter and afterwards for a vector valued parameter, where we are interested on T- and the A-optimal designs. For the coherent MAC we consider only the case of a scalar parameter.

Chapter 4

Optimal Local Sensor Rules

In this chapter we will solve (P-I) for certain special cases. First we solve (P-I) with respect to the systematic noise covariance matrix $\mathbf{C}_{l_i}$ and show that the optimal $\mathbf{C}_{l_i}$ occurs with $\mathbf{C}_{l_i}^* = \mathbf{0}$, without any restriction on our system model. From then on, we restrict our system model with $\mathbf{C}_{l_i} = \mathbf{0}$ for $1 \leq i \leq L$. Furthermore, we will show that for an orthogonal MAC and in particular for the class of linear optimality criterion functions ϕ, we can determine each optimal LO_i independently of each other, whereby we may consider a single sensor setup. Afterwards, we will show that we can reduce our original system model to an equivalent model, in which observation and channel noise $\left\{\mathbf{n}_i', \mathbf{n}_{h_i}', \mathbf{n}_h',\ 1 \leq i \leq L\right\}$ are beeing independent and identically distributed (iid) and the channel matrices $\{\mathbf{H}_i',\ 1 \leq i \leq L\}$ (only for an orthogonal MAC) have diagonal structure. This model reduction concerning iid observation and channel noise, only occurs, when we consider positive definite observation and channel noise covariance matrices in the original system model, i.e., for $\{\mathbf{C}_{n_i} > 0, \mathbf{C}_{h_i} > 0, \mathbf{C}_h > 0,\ 1 \leq i \leq L\}$. Using this foundation, we first consider the special case of a scalar parameter and afterwards the general case of a vector parameter.

4.1 Systematic Noise

Let us now consider the additive systematice noise $\mathbf{n}_{l_i} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}_{l_i} \geq \mathbf{0}\right)$ at sensor i, in particular. As a designer of $\mathbf{C}_{l_i}$ for LO_i, we have to choose $\mathbf{C}_{l_i}$ optimally for (P-I). The next theorem shows that the optimal $\mathbf{C}_{l_i}$ occurs with $\mathbf{C}_{l_i}^* = \mathbf{0}$, i.e., neglecting the systematic noise $\mathbf{n}_{l_i}$ at sensors i.

Lemma 4.1.8 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *are real symmetric matrices of size* $s \times s$*, then*

$$\mathbf{A} \geq \mathbf{B} \quad \Rightarrow \mathbf{T}^T\mathbf{A}\mathbf{T} \geq \mathbf{T}^T\mathbf{B}\mathbf{T},$$

for all $\mathbf{T}$ *of size* $s \times k$*. If* $k \leq s$*, we also have*

$$\mathbf{A} > \mathbf{B} \quad \Rightarrow \mathbf{T}^T\mathbf{A}\mathbf{T} > \mathbf{T}^T\mathbf{B}\mathbf{T},$$

whenever $\mathbf{T}$ *has rank* k.

Proof. see [10, p.470, Observation 7.7.2]. □

Theorem 4.1.9 *The optimal* $\mathbf{C}_{l_i}$ *for a local sensor rule* $\mathrm{LO}_i = (\mathbf{A}_i, \mathbf{C}_{l_i})$ *with fixed sensors matrix* $\mathbf{A}_i$ *is optimal for* (P-I) *if and only if* $\mathbf{C}^*_{l_i} = \mathbf{0}$, *for an arbitrary optimality criterion function* ϕ.

Proof. We consider a given design for the ith LO, i.e., a sensor matrix $\mathbf{A}_i$ and a covariance matrix $\mathbf{C}_{l_i}$, which satisfies the power constraint (C1) or (C2), respectively. Then, we will show that the LO_i with same sensor matrix $\mathbf{A}_i$, but covariance matrix $\mathbf{C}'_{l_i} = \mathbf{0}$, will never result in a decrease of the FIM $\mathbf{J_z}$ from (P-I) (w.r.t. to the Loewner ordering) or, equivalently, of $\phi\{\mathbf{J_z}\}$ (cf. (3.31)) from (P-I). Finally, we will show that $\mathbf{C}'_{l_i} = \mathbf{0}$ also satisfies the power constraints (C1) or (C2), respectively.

To that end, we recall the general expression of the FIM $\mathbf{J_z}$, first, for the orthogonal MAC case, given in (3.8), i.e.,

$$\mathbf{J_z} = \sum_{j=1}^{L} \mathbf{G}_j^T \mathbf{A}_j^T \mathbf{H}_j^T \mathbf{C}_{z_j}^{-1} \mathbf{H}_j \mathbf{A}_j \mathbf{G}_j, \tag{4.1}$$

where the covariance matrix $\mathbf{C}_{z_j}$ is given in (3.4), i.e.,

$$\mathbf{C}_{z_j} = \mathbf{C}_{h_j} + \mathbf{H}_j \mathbf{C}_{l_j} \mathbf{H}_j^T + \mathbf{H}_j \mathbf{A}_j \mathbf{C}_{n_j} \mathbf{A}_j^T \mathbf{H}_j^T \quad \text{for } 1 \le j \le L. \tag{4.2}$$

Let us first consider the covariance matrix $\mathbf{C}_{z_i}$ (sensor i) from (4.2) for $i = j$. It is evident that

$$\mathbf{C}_{h_i} + \mathbf{H}_i \mathbf{C}_{l_i} \mathbf{H}_i^T + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T \ge \mathbf{C}_{h_i} + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T$$

for all $\mathbf{C}_{l_i} \ge \mathbf{0}$, since $\mathbf{H}_i \mathbf{C}_{l_i} \mathbf{H}_i^T \ge \mathbf{0}$ (positive semi-definite) for all $\mathbf{H}_i$, since $\mathbf{x}^T \mathbf{H}_i \mathbf{C}_{l_i} \mathbf{H}_i^T \mathbf{x} = \mathbf{y}^T \mathbf{C}_{l_i} \mathbf{y} \ge 0$ for all $\mathbf{x} \in \mathbb{R}^{p_i}$, $\mathbf{y} = \mathbf{H}_i^T \mathbf{x} \in \mathbb{R}^{q_i}$. By Corollary 3.3.4, we have that

$$\left(\mathbf{C}_{h_i} + \mathbf{H}_i \mathbf{C}_{l_i} \mathbf{H}_i^T + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \le \left(\mathbf{C}_{h_i} + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \tag{4.3}$$

for all $\mathbf{C}_{l_i} \ge \mathbf{0}$. We now introduce $\mathbf{T}_i \triangleq \mathbf{H}_i \mathbf{A}_i \mathbf{G}_i \in \mathbb{R}^{p_i \times n}$. With (4.3) and Lemma 4.1.8, we conclude that

$$\begin{aligned} &\mathbf{T}_i^T \left(\mathbf{C}_{h_i} + \mathbf{H}_i \mathbf{C}_{l_i} \mathbf{H}_i^T + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{T}_i \\ &\qquad \le \mathbf{T}_i^T \left(\mathbf{C}_{h_i} + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{T}_i \quad \text{for all } \mathbf{C}_{l_i} \ge \mathbf{0}. \end{aligned} \tag{4.4}$$

From (4.4), it is obvious that the FIM $\mathbf{J_z}$ from (4.1) for any $\mathbf{C}_{l_i} \ge \mathbf{0}$ (at sensor i), is bounded above (for a variable $\mathbf{C}_{l_i}$, but otherwise fixed system setup) by

$$\begin{aligned}\mathbf{J_z} &= \sum_{j=1}^{L} \mathbf{T}_j^T \left(\mathbf{C}_{h_j} + \mathbf{H}_j \mathbf{C}_{l_j} \mathbf{H}_j^T + \mathbf{H}_j \mathbf{A}_j \mathbf{C}_{n_j} \mathbf{A}_j^T \mathbf{H}_j^T\right)^{-1} \mathbf{T}_j \\ &\leq \sum_{j=1, j\neq i}^{L} \mathbf{T}_j^T \left(\mathbf{C}_{h_j} + \mathbf{H}_j \mathbf{C}_{l_j} \mathbf{H}_j^T + \mathbf{H}_j \mathbf{A}_j \mathbf{C}_{n_j} \mathbf{A}_j^T \mathbf{H}_j^T\right)^{-1} \mathbf{T}_j \\ &\qquad + \mathbf{T}_i^T \left(\mathbf{C}_{h_i} \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{T}_i \triangleq \mathbf{J}_\mathbf{z}^i \quad \text{for all } \mathbf{C}_{l_i} \geq \mathbf{0}.\end{aligned} \tag{4.5}$$

Note that $\mathbf{J}_\mathbf{z}^i$ is equivalent to $\mathbf{J_z}$ evaluated for $\mathbf{C}'_{l_i} = \mathbf{0}$.

Let us now recall the system model with a coherent MAC. The general expression for the FIM $\mathbf{J_z}$ is then given in (3.20), i.e.,

$$\mathbf{J_z} = \mathbf{G}^T \widetilde{\mathbf{A}}^T \mathbf{C}_z^{-1} \widetilde{\mathbf{A}} \mathbf{G}, \tag{4.6}$$

where the covariance matrix $\mathbf{C}_z$ is given in (3.19), i.e.,

$$\mathbf{C}_z = \mathbf{C}_h + \mathbf{H}\mathbf{C}_l\mathbf{H}^T + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T, \tag{4.7}$$

and $\mathbf{G}$, $\widetilde{\mathbf{A}}$, $\mathbf{C}_h$, $\mathbf{C}_n$, $\mathbf{C}_l$ are defined in (3.16) and (3.17), respectively. Note that $\mathbf{C}_l$ has block-diagonal structure, whose block-diagonal entries are $\mathbf{C}_{l_j}$ for $1 \leq j \leq L$ (cf. (3.17) and the assumption that all systematic noise vectors $\mathbf{n}_j$ are uncorrelated among different sensors), i.e., $\mathbf{C}_l = \mathrm{diag}\{\mathbf{C}_{l_1}, \mathbf{C}_{l_2}, \ldots, \mathbf{C}_{l_L}\}$. For a given (i.e., fixed) set $\left\{\mathbf{C}_{l_j}\right\}_{j=1, j\neq i}^{L}$ and variable covariance matrix $\mathbf{C}_{l_i} \geq \mathbf{0}$ for sensor i, we have that

$$\mathbf{C}_l = \mathrm{diag}\{\mathbf{C}_{l_1}, \ldots, \mathbf{C}_{l_i}, \ldots, \mathbf{C}_{l_L}\} \geq \mathrm{diag}\left\{\mathbf{C}_{l_1}, \ldots, \mathbf{C}'_{l_i} = \mathbf{0}, \ldots, \mathbf{C}_{l_L}\right\} \triangleq \mathbf{C}_l^i \tag{4.8}$$

for all $\mathbf{C}_{l_i} \geq \mathbf{0}$, since[1]

$$\begin{aligned}\mathbf{x}^T\mathbf{C}_l\mathbf{x} &= \mathbf{x}^T \mathrm{diag}\{\mathbf{C}_{l_1}, \mathbf{0}, \ldots, \mathbf{0}\}\mathbf{x} + \cdots + \mathbf{x}^T \mathrm{diag}\{\mathbf{0}, \ldots, \mathbf{C}_{l_i}, \ldots, \mathbf{0}\}\mathbf{x} \\ &\qquad + \cdots + \mathbf{x}^T \mathrm{diag}\{\mathbf{0}, \ldots, \mathbf{C}_{l_L}\}\mathbf{x} \\ &\geq \mathbf{x}^T \mathrm{diag}\{\mathbf{C}_{l_1}, \mathbf{0}, \ldots, \mathbf{0}\}\mathbf{x} + \cdots + \mathbf{x}^T \mathrm{diag}\left\{\mathbf{0}, \ldots, \mathbf{C}'_{l_i} = \mathbf{0}, \ldots, \mathbf{0}\right\}\mathbf{x} \\ &\qquad + \cdots + \mathbf{x}^T \mathrm{diag}\{\mathbf{0}, \ldots, \mathbf{C}_{l_L}\}\mathbf{x} = \mathbf{x}^T\mathbf{C}_l^i\mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^q, \mathbf{C}_{l_i} \geq \mathbf{0}.\end{aligned}$$

Thus from (4.8), we conclude that

$$\left(\mathbf{C}_h + \mathbf{H}\mathbf{C}_l\mathbf{H}^T + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T\right)^{-1} \leq \left(\mathbf{C}_h + \mathbf{H}\mathbf{C}_l^i\mathbf{H}^T + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T\right)^{-1} \tag{4.9}$$

for all $\mathbf{C}_{l_i} \geq \mathbf{0}$, since $\mathbf{H}\mathbf{C}_l\mathbf{H}^T \geq \mathbf{H}\mathbf{C}_l^i\mathbf{H}^T$ for all $\mathbf{H}$ (cf. Lemma 4.1.8). We now introduce $\mathbf{T} \triangleq \widetilde{\mathbf{A}}\mathbf{G} \in \mathbb{R}^{p\times n}$. With (4.9) and Lemma 4.1.8, it is obvious

[1] The Loewner ordering $\mathbf{A} \geq \mathbf{B}$, where $\mathbf{A}, \mathbf{B} \in \mathrm{Sym}(k)$, is equivalent to $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq \mathbf{x}^T\mathbf{B}\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^k$ (cf. [4]).

that the FIM $\mathbf{J_z}$ from (4.6) for any $\mathbf{C}_{l_i} \geq \mathbf{0}$, is again bounded above by

$$\begin{aligned}\mathbf{J_z} &= \mathbf{T}^T \left(\mathbf{C}_h + \mathbf{H}\mathbf{C}_l\mathbf{H}^T + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T\right)^{-1} \mathbf{T} \\ &\leq \mathbf{T}^T \left(\mathbf{C}_h + \mathbf{H}\mathbf{C}_l^i\mathbf{H}^T + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T\right)^{-1} \mathbf{T} \triangleq \mathbf{J}_\mathbf{z}^i \quad \text{for all } \mathbf{C}_{l_i} \geq \mathbf{0}.\end{aligned} \tag{4.10}$$

Hence, for a variable covariance matrix $\mathbf{C}_{l_i} \geq \mathbf{0}$ at sensor i, but otherwise fixed system setup (orthogonal or coherent MAC), we have showed that the FIM $\mathbf{J_z}$ is upper bounded by the FIM $\mathbf{J}_\mathbf{z}^i$ for the same system setup, but covariance matrix $\mathbf{C}'_{l_i} = \mathbf{0}$, in the sense of the Loewner ordering. Since any optimality criterion ϕ is isotonic, relative to the Loewner ordering (cf. subsection 3.3.2), it holds:

$$\mathbf{J_z} \leq \mathbf{J}_\mathbf{z}^i \quad \Leftrightarrow \quad \phi\left(\mathbf{J_z}\right) \leq \phi\left(\mathbf{J}_\mathbf{z}^i\right) \quad \text{for all } \mathbf{C}_{l_i} \geq \mathbf{0},$$

i.e., the objective function of (P-I) is bounded above accordingly.

It remains to verify that the power constraint, i.e., (C1) or (C2), which is assumed to be satisfied for a given LO_i with $\mathbf{C}_{l_i} \geq \mathbf{0}$, is also be satisfied when $\mathbf{C}'_{l_i} = \mathbf{0}$. Let us recall the ith constraint of (C1). We conclude that (C1) is also satisfied for $\mathbf{C}'_{l_i} = \mathbf{0}$, since

$$\mathrm{tr}\left\{\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}'_{l_i} = \mathbf{0}\right\} \leq \mathrm{tr}\left\{\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} \leq P_{0_i},$$

because $\mathrm{tr}\{\mathbf{0}\} = 0$ and $\mathrm{tr}\{\mathbf{C}_{l_i}\} \geq 0$ for all $\mathbf{C}_{l_i} \geq \mathbf{0}$ [8]. Analog, we recall the ith constraint of (C2). We conclude that (C2) is also satisfied for $\mathbf{C}'_{l_i} = \mathbf{0}$, since

$$\mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}'_{l_i} = \mathbf{0}\right\} \leq \mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} + \mathrm{tr}\left\{\mathbf{C}_{l_i}\right\} \leq P'_{0_i},$$

again, because $\mathrm{tr}\{\mathbf{0}\} = 0$ and $\mathrm{tr}\{\mathbf{C}_{l_i}\} \geq 0$ for all $\mathbf{C}_{l_i} \geq \mathbf{0}$.

Thus, we have showed that $\mathbf{C}'_{l_i} = \mathbf{0}$ is the optimum for (P-I), i.e., $\mathbf{C}^*_{l_i} = \mathbf{C}'_{l_i} = \mathbf{0}$. □

We conclude from Theorem 4.1.9 that optimal $\mathbf{C}_{l_i}$ for $1 \leq i \leq L$ for (P-I), is given by $\mathbf{C}^*_{l_i} = \mathbf{0}$. Thus, we can reformulate the basic optimization problem (P-I) as

$$\begin{aligned} \underset{\mathbf{A}_i,\ 1\leq i\leq L}{\text{maximize}} \quad & \phi\left\{\mathbf{J_z}\right\} \\ \text{subject to} \quad & \mathbf{J_z} \text{ satisfies (3.8) or (3.20)} \\ & \qquad \text{ev. for } \mathbf{C}_{l_i} = \mathbf{0} \quad \text{for } 1 \leq i \leq L, \\ & \mathbf{A}_i \in \mathbb{R}^{q_i \times m_i} \quad \text{for } 1 \leq i \leq L, \\ & \text{(C1) or (C2) is satisfied, i.e.,} \\ & \text{(C1)}: \quad \mathrm{tr}\left\{\mathbf{A}_i\mathbf{M}_i\mathbf{A}_i^T\right\} \leq P_{0,i} \quad \text{for } 1 \leq i \leq L, \text{ or} \\ & \text{(C2)}: \quad \mathrm{tr}\left\{\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\right\} \leq P'_{0,i} \quad \text{for } 1 \leq i \leq L. \end{aligned} \tag{P-II}$$

4.2 The Basic Optimization Problem for an Orthogonal MAC Reformulated

Consider the basic optimization problem (P-II) with an optimality criterion ϕ and for the case of an orthogonal MAC, in particular. The FIM $\mathbf{J_z}$ is then given in (3.8) for $\mathbf{C}_{l_i} = \mathbf{0}$, i.e., it consists of the sum of all individual FIMs $\mathbf{J}_{\mathbf{z}_i}$ from (3.5) for $\mathbf{C}_{l_i} = \mathbf{0}$ and for all i, i.e.,

$$\mathbf{J_z} = \sum_{i=1}^{L} \mathbf{J}_{\mathbf{z}_i}.$$

The next theorem shows that for the case of an orthogonal MAC, we can split the joint problem (P-II) into L individual and independent problems, if we suppose ϕ to be linear.

Theorem 4.2.10 *Consider an optimality criterion ϕ, which also is still linear*[1]*. Then, an optimal $\mathbf{A}_i$ for* (P-II)*, is also optimal for*

$$\begin{aligned} &\underset{\mathbf{A}_i}{\text{maximize}} && \phi\left\{\mathbf{J}_{\mathbf{z}_i}\right\} \\ &\text{subject to} && \mathbf{J}_{\mathbf{z}_i} \text{ satisfies (3.5) ev. for } \mathbf{C}_{l_i} = \mathbf{0}, \\ &&& \mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}, \\ &&& (\mathrm{C1})_i \text{ or } (\mathrm{C2})_i \text{ is satisfied, i.e.,} \\ &&& (\mathrm{C1})_i: \quad \mathrm{tr}\left\{\mathbf{A}_i \mathbf{M}_i \mathbf{A}_i^T\right\} \leq P_{0,i} \text{ or} \\ &&& (\mathrm{C2})_i: \quad \mathrm{tr}\left\{\mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T\right\} \leq P'_{0,i}. \end{aligned} \tag{P-III}$$

The notation $(\mathrm{C1})_i$ *and* $(\mathrm{C2})_i$ *mean the* i*th constraint of* (C1) *and* (C2)*. Conversely, an optimal $\mathbf{A}_i$ for* (P-III)*, is also optimal for* (P-II).

Proof. Let us first interpret $\mathbf{J}_{\mathbf{z}_k}$ from (3.5) as a function on $\mathbf{A}_k$ for all k, i.e., $\mathbf{J}_{\mathbf{z}_k} = \mathbf{J}_{\mathbf{z}_k}(\mathbf{A}_k)$.

Let us start with the proof so that an optimal $\mathbf{A}_i$ for (P-II) implies optimality for (P-III). We consider a given (i.e., fixed) set $\left\{\mathbf{A}'_j \in \mathbb{R}^{q_i \times m_i}\right\}_{j=1, j\neq i}^{L}$, which satisfies the constraint in (P-II). A sensor matrix $\mathbf{A}_i^*$ is optimal for (P-II) iff

$$\phi\left\{\mathbf{J}_{\mathbf{z}_i}(\mathbf{A}_i) + \sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}'_j\right)\right\} \leq \phi\left\{\mathbf{J}_{\mathbf{z}_i}(\mathbf{A}_i^*) + \sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}'_j\right)\right\} \tag{4.11}$$

for all $\mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}$, which satisfies the constraint in (P-II) (i.e., the ith power constraint (C1) or (C2)). Once we have accepted ϕ to be linear, (4.11) yields

$$\phi\left\{\mathbf{J}_{\mathbf{z}_i}(\mathbf{A}_i)\right\} + \phi\left\{\sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}'_j\right)\right\} \leq \phi\left\{\mathbf{J}_{\mathbf{z}_i}(\mathbf{A}_i^*)\right\} + \phi\left\{\sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}'_j\right)\right\}$$

[1] An optimality criterion ϕ on $\mathrm{NND}(s)$ is linear iff $\phi\{k_1\mathbf{J}_1 + k_2\mathbf{J}_2\} = k_1\phi\{\mathbf{J}_1\} + k_2\phi\{\mathbf{J}_2\}$, where $\mathbf{J}_1, \mathbf{J}_2 \in \mathrm{NND}(s)$ and $k_1, k_2 \in \mathbb{R}^+$.

for all $\mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}$, again which satisfies the constraint in (P-II), implying

$$\phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i\right)\right\} \leq \phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i^*\right)\right\} \quad \text{for all } \mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}, \tag{4.12}$$

which satisfies the constraint in (P-II). Since both contraints from (P-II) and (P-III) are identical, (4.12) implies optimality of $\mathbf{A}_i^*$ for (P-III).

Conversely, if $\mathbf{A}_i^*$ is optimal for (P-III), then it is evident that

$$\phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i\right)\right\} \leq \phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i^*\right)\right\} \quad \text{for all } \mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}, \tag{4.13}$$

which satisfies the constraint in (P-III). Without violating (4.13), we can add on both sides of (4.13)

$$\phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i\right)\right\}+\phi\left\{\sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}_j''\right)\right\} \leq \phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i^*\right)\right\}+\phi\left\{\sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}_j''\right)\right\} \tag{4.14}$$

for all $\mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}$, which satisfies the constraint in (P-III) and, for an arbitrary, but fixed set $\left\{\mathbf{A}_j'' \in \mathbb{R}^{q_i \times m_i}\right\}_{j=1, j\neq i}^{L}$, which satisfies the constraint in (P-II), in particular. With the assumption that ϕ is linear, (4.14) can also be written as

$$\phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i\right)+\sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}_j''\right)\right\} \leq \phi\left\{\mathbf{J}_{\mathbf{z}_i}\left(\mathbf{A}_i^*\right)+\sum_{j=1, j\neq i}^{L} \mathbf{J}_{\mathbf{z}_j}\left(\mathbf{A}_j''\right)\right\} \tag{4.15}$$

for all $\mathbf{A}_i \in \mathbb{R}^{q_i \times m_i}$, which satisfies the constraint in (P-III). Again, since both ith contraints from (P-II) and (P-III) are identical, (4.15) implies optimality of $\mathbf{A}_i^*$ for (P-II). □

Supposing, an optimality criterion ϕ is linear and the system setup includes an orthogonal MAC for solving (P-II). Then, according to Theorem 4.2.10 we can solve (P-II) or, equivalently, (P-III) for solving an optimal LO_i, i.e., the optimal sensor matrix $\mathbf{A}_i^*$. Note that this result clearly holds for a scalar paramater, as a special case. Any optimality criterion ϕ' would be equivalent and especially linear, since $\phi'\{k_1+k_2\} = k_1+k_2 = \phi'\{k_1\}+\phi'\{k_2\}$ for $k_1, k_2 \in \mathbb{R}$. Thus, for the remaining part of the thesis, we will consider problem (P-III) when we treat the special case of a scalar parameter and furthermore, when we determine T-optimal local sensors for a vector parameter, both, for a system with an orthogonal MAC.

4.3 Reduction to Standard Model

In Section 4.1, we have already solved (P-I), w.r.t. $\mathbf{C}_{l_i}$ for all i, where we have determined $\mathbf{C}_{l_i}^* = \mathbf{0}$ for $1 \leq i \leq L$. In what follows, we restrict our original system model (for both multiple access schemes) with $\mathbf{C}_{l_i} = \mathbf{0}$ for all i. Let us rewrite both LGMs, first, for the orthogonal MAC case, i.e., the final observation $\mathbf{z}_i$ for all i from (3.3) as

$$\mathbf{z}_i = \mathbf{H}_i\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{H}_i\mathbf{A}_i\mathbf{n}_i + \mathbf{n}_{h_i} \quad \text{for } 1 \leq i \leq L, \tag{4.16}$$

where, due to the LGM, $\mathbf{z}_i$ is Gaussian distributed with mean

$$\boldsymbol{\mu}_{z_i} = \mathbf{H}_i\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}, \tag{4.17}$$

and covariance matrix

$$\mathbf{C}_{z_i} = \mathbf{C}_{h_i} + \mathbf{H}_i\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{H}_i^T\mathbf{A}_i^T. \tag{4.18}$$

For the coherent MAC case, we rewrite the LGM, i.e., the final observation $\mathbf{z}$ from (3.18) as

$$\mathbf{z} = \widetilde{\mathbf{A}}\mathbf{G}\boldsymbol{\theta} + \widetilde{\mathbf{A}}\mathbf{n} + \mathbf{n}_h, \tag{4.19}$$

where again, $\mathbf{z}$ is Gaussian distributed with mean

$$\boldsymbol{\mu}_z = \widetilde{\mathbf{A}}\mathbf{G}\boldsymbol{\theta}, \tag{4.20}$$

and covariance matrix

$$\mathbf{C}_z = \mathbf{C}_h + \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T. \tag{4.21}$$

Note that the model parameters for the coherent MAC case, i.e., $\widetilde{\mathbf{A}}$, $\mathbf{G}$ and $\mathbf{n}$ can be obtained from (3.16) and (3.17), respectively. The covariance matrix $\mathbf{C}_n = \operatorname{diag}\{\mathbf{C}_{n_i}\}_{i=1}^{L}$, due to the assumption that the observation noise vectors $\mathbf{n}_i$ for all i, are uncorrelated among different sensors.

Definition 4.3.11 *Two observations $\mathbf{z}_1$ and $\mathbf{z}_2$ are said to be equivalent, iff $\mathbf{z}_1$ and $\mathbf{z}_2$ have the same pdf for every $\boldsymbol{\theta}$, i.e., $f_{\mathbf{z}_1}(\mathbf{z};\boldsymbol{\theta}) = f_{\mathbf{z}_2}(\mathbf{z};\boldsymbol{\theta})$.*

The next theorem shows that we can reduce the original model, i.e., the LGMs for both multiple access schemes from (4.16) and (4.19), into an equivalent model by the observation- and channel noise vectors are beeing iid w.l.o.g. Note that this applies only if we suppose non-singular observation- and channel noise covariance matrices in the original system model. All other assumptions we have made so far to our original system model (cf. Section 3.1) are also valid for the equivalent model.

Theorem 4.3.12 (Noise Whitening) *Consider the original system model, shown in* (4.16) *and* (4.19) *for both multiple access schemes. Assuming non-singular coariance matrices $\mathbf{C}_{n_i}$, $\mathbf{C}_{h_i}$ for $1 \leq i \leq L$ (thus $\mathbf{C}_n$) and $\mathbf{C}_h$, respectively. Then we can define an equivalent system model, according to Definition 4.3.11, for the orthogonal MAC case as*

$$\mathbf{z}'_i = \frac{1}{\sigma_{h'_i}}\mathbf{C}_{h_i}^{1/2}\left(\mathbf{H}'_i\mathbf{A}'_i\mathbf{G}'_i\boldsymbol{\theta} + \mathbf{H}'_i\mathbf{A}'_i\mathbf{n}'_i + \mathbf{n}'_{h_i}\right) \quad 1 \leq i \leq L, \tag{4.22}$$

where $\mathbf{n}_i' \sim \mathcal{N}\left(\mathbf{0}, \sigma_{n_i'}^2 \mathbf{I}\right)$, $\mathbf{n}_{h_i}' \sim \mathcal{N}\left(\mathbf{0}, \sigma_{h_i'}^2 \mathbf{I}\right)$, $\mathbf{H}_i' \triangleq \sigma_{h_i'} \mathbf{C}_{h_i}^{-1/2} \mathbf{H}_i$, $\mathbf{A}_i' \triangleq \frac{1}{\sigma_{n_i'}} \mathbf{A}_i \mathbf{C}_{n_i}^{1/2}$ *and* $\mathbf{G}_i' \triangleq \sigma_{n_i'} \mathbf{C}_{\mathbf{n}_i}^{-1/2} \mathbf{G}_i$ *with* $\sigma_{n_i'}^2 > 0$ *and* $\sigma_{h_i'}^2 > 0$. *Here, we assume that* $\mathbf{n}_i'$ *is uncorrelated with* $\mathbf{n}_{h_j}'$ *for all* i *and* j, *i.e.,* $cov\left\{\mathbf{n}_i, \mathbf{n}_{h_j}\right\} = \mathbf{0}$ *for all* i *and* j. *For the coherent MAC case, we can define*

$$\mathbf{z}' = \frac{1}{\sigma_{h'}} \mathbf{C}_h^{1/2} \left(\widetilde{\mathbf{A}}' \mathbf{G}' \boldsymbol{\theta} + \widetilde{\mathbf{A}}' \mathbf{n}' + \mathbf{n}_h' \right), \tag{4.23}$$

where $\mathbf{n}' \sim \mathcal{N}\left(\mathbf{0}, \sigma_{n'}^2 \mathbf{I}\right)$, $\mathbf{n}_h' \sim \mathcal{N}\left(\mathbf{0}, \sigma_{h'}^2 \mathbf{I}\right)$, $\widetilde{\mathbf{A}}' \triangleq \frac{\sigma_{h'}}{\sigma_{n_i'}} \mathbf{C}_h^{-1/2} \widetilde{\mathbf{A}} \mathbf{C}_n^{1/2}$ *and* $\mathbf{G}' \triangleq \sigma_{n'} \mathbf{C}_n^{-1/2} \mathbf{G}$ *with* $\sigma_{n'}^2 > 0$ *and* $\sigma_{h'}^2 > 0$. *Again, we assume that* $\mathbf{n}'$ *is uncorrelated with* $\mathbf{n}_h'$.

Proof. Let us start with the proof for the orthogonal MAC case. To that end, we insert $\mathbf{H}_i'$, $\mathbf{A}_i'$ and $\mathbf{G}_i'$ into (4.22) yields

$$\begin{aligned}
\mathbf{z}_i' &= \frac{1}{\sigma_{h_i'}} \mathbf{C}_{h_i}^{1/2} \left(\mathbf{H}_i' \mathbf{A}_i' \mathbf{G}_i' \boldsymbol{\theta} + \mathbf{H}_i' \mathbf{A}_i' \mathbf{n}_i' + \mathbf{n}_{h_i}' \right) \\
&= \frac{\sigma_{h_i'} \sigma_{n_i'}}{\sigma_{h_i'} \sigma_{n_i'}} \mathbf{C}_{h_i}^{1/2} \mathbf{C}_{h_i}^{-1/2} \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i}^{1/2} \mathbf{C}_{n_i}^{-1/2} \mathbf{G}_i \boldsymbol{\theta} \\
&\quad + \frac{\sigma_{h_i'}}{\sigma_{h_i'} \sigma_{n_i'}} \mathbf{C}_{h_i}^{1/2} \mathbf{C}_{h_i}^{-1/2} \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i}^{1/2} \mathbf{n}_i' + \frac{1}{\sigma_{h_i'}} \mathbf{C}_{h_i}^{1/2} \mathbf{n}_{h_i}' \\
&= \mathbf{H}_i \mathbf{A}_i \mathbf{G}_i \boldsymbol{\theta} + \frac{1}{\sigma_{n_i'}} \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i}^{1/2} \mathbf{n}_i' + \frac{1}{\sigma_{h_i'}} \mathbf{C}_{h_i}^{1/2} \mathbf{n}_{h_i}',
\end{aligned}$$

where $\mathbf{z}_i'$ is Gaussian distributed, with mean

$$\boldsymbol{\mu}_{z_i'} = \mathbf{H}_i \mathbf{A}_i \mathbf{G}_i \boldsymbol{\theta}, \tag{4.24}$$

and covariance matrix

$$\begin{aligned}
\mathbf{C}_{z_i'} &= \frac{\sigma_{n_i'}^2}{\sigma_{n_i'} \sigma_{n_i'}} \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i}^{1/2} \mathbf{C}_{n_i}^{1/2} \mathbf{A}_i^T \mathbf{H}_i^T + \frac{\sigma_{h_i'}^2}{\sigma_{h_i'} \sigma_{h_i'}} \mathbf{C}_{h_i}^{1/2} \mathbf{C}_{h_i}^{1/2} \\
&= \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T + \mathbf{C}_{h_i},
\end{aligned} \tag{4.25}$$

due to our assumption that $\mathbf{n}_i'$ is uncorrelated with $\mathbf{n}_{h_j}'$ for all i and j. Comparing (4.24) with (4.17) and (4.25) with (4.18), we conclude that $\mathbf{z}_i'$ from (4.22) is equivalent to $\mathbf{z}_i$ from (4.16) for $1 \leq i \leq L$, according to Definition 4.3.11, since both are Gaussian distributed random variables with same mean and same covariance matrix.

Let us now consider the coherent MAC. To that end, we insert $\widetilde{\mathbf{A}}'$ and

$\mathbf{G}'$ into (4.23) yields

$$\begin{aligned}\mathbf{z}' &= \frac{1}{\sigma_{h'}}\mathbf{C}_h^{1/2}\left(\widetilde{\mathbf{A}}'\mathbf{G}'\boldsymbol{\theta} + \widetilde{\mathbf{A}}'\mathbf{n}' + \mathbf{n}_h'\right) \\ &= \frac{\sigma_{h'}\sigma_{n'}}{\sigma_{h'}\sigma_{n'}}\mathbf{C}_h^{1/2}\mathbf{C}_h^{-1/2}\widetilde{\mathbf{A}}\mathbf{C}_n^{1/2}\mathbf{C}_n^{-1/2}\mathbf{G}\boldsymbol{\theta} \\ &\qquad + \frac{\sigma_{h'}}{\sigma_{h'}\sigma_{n'}}\mathbf{C}_h^{1/2}\mathbf{C}_h^{-1/2}\widetilde{\mathbf{A}}\mathbf{C}_n^{1/2}\mathbf{n}' + \frac{1}{\sigma_{h'}}\mathbf{C}_h^{1/2}\mathbf{n}_h' \\ &= \widetilde{\mathbf{A}}\mathbf{G}\boldsymbol{\theta} + \frac{1}{\sigma_{n'}}\widetilde{\mathbf{A}}\mathbf{C}_n^{1/2}\mathbf{n}' + \frac{1}{\sigma_{h'}}\mathbf{C}_h^{1/2}\mathbf{n}_h',\end{aligned}$$

where $\mathbf{z}'$ is Gaussian distributed, with mean

$$\boldsymbol{\mu}_{z'} = \widetilde{\mathbf{A}}\mathbf{G}\boldsymbol{\theta}, \tag{4.26}$$

and covariance matrix

$$\begin{aligned}\mathbf{C}_{z'} &= \frac{\sigma_{n'}^2}{\sigma_{n'}\sigma_{n'}}\widetilde{\mathbf{A}}\mathbf{C}_n^{1/2}\mathbf{C}_n^{1/2}\widetilde{\mathbf{A}}^T + \frac{\sigma_{h'}^2}{\sigma_{h'}\sigma_{h'}}\mathbf{C}_h^{1/2}\mathbf{C}_h^{1/2} \\ &= \widetilde{\mathbf{A}}\mathbf{C}_n\widetilde{\mathbf{A}}^T + \mathbf{C}_h,\end{aligned} \tag{4.27}$$

again, due to the assumption that $\mathbf{n}'$ is uncorrelated with $\mathbf{n}_h'$. Comparing (4.26) with (4.20) and (4.27) with (4.21), we conclude that $\mathbf{z}'$ from (4.23) is equivalent to $\mathbf{z}$ from (4.19), according to Definition 4.3.11, since both are Gaussian distributed random variables with same mean and same covariance matrix. □

According to Theorem 4.3.12, we can assume $\mathbf{C}_{n_i} = \sigma_{n_i}^2\mathbf{I}$ with $= \sigma_{n_i}^2 > 0$ and/or $\mathbf{C}_{h_i} = \sigma_{h_i}^2\mathbf{I}$ with $\sigma_{h_i}^2 > 0$ for our system model with an orthogonal MAC. Otherwise, we can absorb $\sigma_{n_i'}\mathbf{C}_{n_i}^{-1/2}$ and $\sigma_{n_i'}^{-1}\mathbf{C}_{n_i}^{1/2}$ into the observation matrix $\mathbf{G}_i$ and the sensor matrix $\mathbf{A}_i$, respectively, and/or $\sigma_{h_i'}\mathbf{C}_h^{-1/2}$ and $\sigma_{h_i'}^{-1}\mathbf{C}_{h_i}^{1/2}$ into the channel matrix $\mathbf{H}_i$ and the FC, respectively, such that we obtain an equivalent model in which $\mathbf{C}_{n_i'} = \sigma_{n_i'}^2\mathbf{I}$ and/or $\mathbf{C}_{h_i'} = \sigma_{h_i'}^2\mathbf{I}$. Similarly, we can assumme $\mathbf{C}_n = \sigma_n^2\mathbf{I}$ with $\sigma_n^2 > 0$ and $\mathbf{C}_h = \sigma_h^2\mathbf{I}$ with $\sigma_h^2 > 0$ for the system model with a coherent MAC, w.l.o.g. Otherwise, we again can absorb $\sigma_{n'}\mathbf{C}_n^{-1/2}$ and $\sigma_{n'}^{-1}\mathbf{C}_n^{1/2}$ into the total observation matrix $\mathbf{G}$ and the total sensor-channel matrix $\widetilde{\mathbf{A}}$, respectively, and/or $\sigma_{h'}\mathbf{C}_h^{-1/2}$ and $\sigma_{h'}^{-1}\mathbf{C}_h^{1/2}$ into the total sensor-channel matrix $\widetilde{\mathbf{A}}$ and the FC, respectively, such that we obtain an equivalent model in which $\mathbf{C}_{n'} = \sigma_{n'}^2\mathbf{I}$ and/or $\mathbf{C}_{h'} = \sigma_{h'}^2\mathbf{I}$ (cf. Fig. 4.1).

Unless otherwise stated, we assume throughout the remaining part of this thesis iid channel noise for both multiple access schemes w.l.o.g. The next theorem shows that for the orthogonal MAC case, we can assume diagonal channel matrices $\mathbf{H}_i$ for $1 \leq i \leq L$, w.l.o.g.

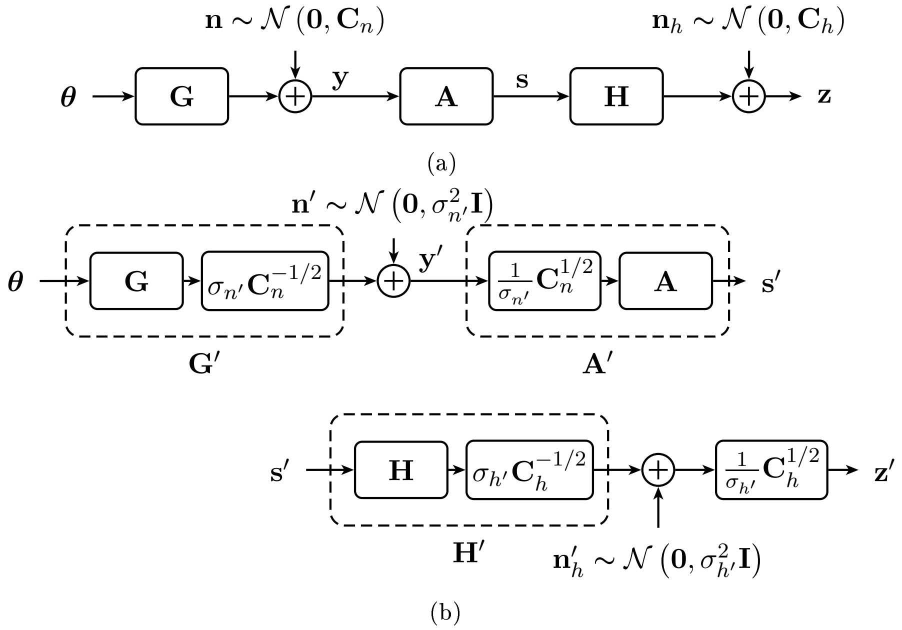


Figure 4.1: Equivalence of two system models - (a) original model (b) equivalent, noise whitened model. For the case of an orthogonal MAC, we have to index all model parameters (i.e., all occuring vectors and matrices, except $\boldsymbol{\theta}$) by an index i. Then, the figure illustrates sensor i with an appropriate observation- and channel model from sensor to the FC. For the case of a coherent MAC, figure shows the whole system model, if we merge $\mathbf{A}$ with $\mathbf{H}$ to obtain $\widetilde{\mathbf{A}} = \mathbf{H}\mathbf{A}$.

Theorem 4.3.13 *Let us consider a system model with an orthogonal MAC, shown in* (4.16)*. According to Theorem 4.3.12, we assume iid channel noise (zero-mean, Gaussian distributed)* $\mathbf{n}_{h_i}$ *with covariance matrix* $\mathbf{C}_{h_i} = \sigma_{h_i}^2 \mathbf{I}$ *(*$\sigma_{h_i}^2 > 0$*), w.l.o.g. Then we can define an equivalent system model, according to Definition 4.3.11 as*

$$\mathbf{z}_i' = \mathbf{U}_{h_i} \left(\mathbf{H}_i' \mathbf{A}_i' \mathbf{G}_i \boldsymbol{\theta} + \mathbf{H}_i' \mathbf{A}_i' \mathbf{n}_i + \mathbf{n}_{h_i} \right), \tag{4.28}$$

where $\mathbf{H}_i' = \mathbf{\Sigma}_{h_i}$ *and* $\mathbf{A}_i' \triangleq \mathbf{V}_{h_i}^T \mathbf{A}_i$*.* $\mathbf{U}_{h_i}$ $\mathbf{\Sigma}_{h_i}$ *and* $\mathbf{V}_{h_i}$ *follows from the singular value decomposition (SVD) of the original channel matrix* $\mathbf{H}_i = \mathbf{U}_{h_i} \mathbf{\Sigma}_{h_i} \mathbf{V}_{h_i}^T$*, i.e.,* $\mathbf{U}_{h_i}$ *and* $\mathbf{V}_{h_i}$ *are both unitary and the diagonal matrix* $\mathbf{\Sigma}_{h_i}$ *of size* $p_i \times q_i$ *contains the singular values of* $\mathbf{H}_i$ *on the main diagonal. Note that* $\mathbf{n}_i'$ *is uncorrelated with* $\mathbf{n}_{h_j}'$ *for all* i *and* j *as already assumed.*

Proof. Inserting $\mathbf{H}_i'$ and $\mathbf{A}_i'$ into (4.28) yields

$$\begin{aligned}\mathbf{z}_i' &= \mathbf{U}_{h_i}\left(\mathbf{\Sigma}_{h_i}\mathbf{A}_i'\mathbf{G}_i\boldsymbol{\theta} + \mathbf{\Sigma}_{h_i}\mathbf{A}_i'\mathbf{n}_i + \mathbf{n}_{h_i}\right)\\ &= \mathbf{U}_{h_i}\mathbf{\Sigma}_{h_i}\mathbf{V}_{h_i}^T\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{U}_{h_i}\mathbf{\Sigma}_{h_i}\mathbf{V}_{h_i}^T\mathbf{A}_i\mathbf{n}_i + \mathbf{U}_{h_i}\mathbf{n}_{h_i}\\ &= \mathbf{H}_i\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta} + \mathbf{H}_i\mathbf{A}_i\mathbf{n}_i + \mathbf{U}_{h_i}\mathbf{n}_{h_i},\end{aligned} \tag{4.29}$$

where $\mathbf{z}_i'$ is again Gaussian distributed, with mean

$$\boldsymbol{\mu}_{z_i'} = \mathbf{H}_i\mathbf{A}_i\mathbf{G}_i\boldsymbol{\theta}, \tag{4.30}$$

and covariance matrix

$$\begin{aligned}\mathbf{C}_{z_i'} &= \mathbf{H}_i\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\mathbf{H}_i^T + \sigma_{h_i}^2\mathbf{U}_h\mathbf{U}_h^T\\ &= \mathbf{H}_i\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\mathbf{H}_i^T + \sigma_{h_i}^2\mathbf{I} = \mathbf{H}_i\mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T\mathbf{H}_i^T + \mathbf{C}_{h_i}.\end{aligned} \tag{4.31}$$

Comparing (4.30) with (4.16) and (4.31) with (4.18) for $\mathbf{C}_{h_i} = \sigma_{h_i}^2\mathbf{I}$, we conclude that $\mathbf{z}_i'$ from (4.28) is equivalent to $\mathbf{z}_i$ from (4.16) for $\mathbf{n}_{h_i} = \sigma_{h_i}^2\mathbf{I}$ and for $1 \leq i \leq L$, according to Definition 4.3.11, since both are Gaussian distributed random variables, with same mean and same covariance matrix. □

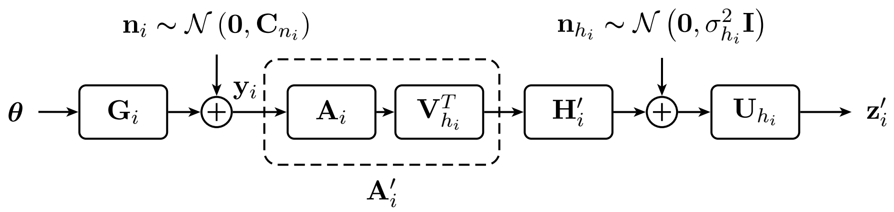


Figure 4.2: Equivalent system model - channel diagonalization between sensor i and FC.

According to Theorem 4.3.13, we thus can assume a diagonal channel matrix $\mathbf{H}_i$ in the case of an orthogonal MAC w.l.o.g. Otherwise, we can absorb $\mathbf{U}_{h_i}$ and $\mathbf{V}_{h_i}^T$, which follows from the SVD $\mathbf{H}_i = \mathbf{U}_{h_i}\mathbf{\Sigma}_{h_i}\mathbf{V}_{h_i}^T$ into the FC and the sensor matrix $\mathbf{A}_i$, respectively, such that we obtain an equivalent model in which $\mathbf{H}_i'$ is diagonal, where the diagonal entries are the singular values of $\mathbf{H}_i$. Note that the order of the diagonal entries, i.e., the singular values of $\mathbf{H}_i$ can be assumed arbitrarily. An illustration of the equivalent model for a diagonalized channel matrix can be seen in Fig. 4.2. The original model is shown in Fig. 4.1(a) for the orthogonal MAC.

Definition 4.3.14 (Standard Model) *Assuming a system model with an orthogonal MAC, that has iid observation noise vectors* $\mathbf{n}_i$ *and iid channel noise vectors* $\mathbf{n}_{h_i}$ *(zero-mean, Gaussian distributed) with covariance matrices* $\mathbf{C}_{n_i} = \sigma_{n_i}^2\mathbf{I}$ *and* $\mathbf{C}_{h_i} = \sigma_{h_i}^2\mathbf{I}$ *for* $1 \leq i \leq L$*, and especially diagonal channel matrices* $\mathbf{H}_i$ *for* $1 \leq i \leq L$ *- is called the standard model.*

4.4 Scalar Parameter Case

In this section, we consider the special case of estimating a scalar parameter, i.e., $\theta \in \mathbb{R}$. In that case, the FIMs $\mathbf{J}_\mathbf{z}$ and $\mathbf{J}_{\mathbf{z}_i}$ are both scalar-valued, since $n = 1$ - thus we us the notations $J_\mathbf{z}$ and $J_{\mathbf{z}_i}$ in what follows and set $\mathbf{J}_\mathbf{z} = J_\mathbf{z}$ and $\mathbf{J}_{\mathbf{z}_i} = J_{\mathbf{z}_i}$. In Theorem 4.1.9, we have already showed that each LO$_i$, can be restrict to the sensor matrix $\mathbf{A}_i$, since $\mathbf{C}_{l_i}^* = \mathbf{0}$ for $1 \leq i \leq L$, i.e., we consider (P-II) for solving optimal LO$_i$ - even for this special case. Note that for a scalar parameter, we do not need any optimality criterion ϕ for solving problem (P-II), since $\phi\{J_\mathbf{z}\} = J_\mathbf{z}$ for $J_\mathbf{z} \in \mathbb{R}$ - in that case all optimality criteria are equivalent.

Unless otherwise stated, we do not especially assume iid observation noise $\mathbf{n}_i$ for sensor i for the case of a scalar parameter. We take only the assumption - which has already been made in Section 4.3 - that the channel noise $\mathbf{n}_{h_i}$, for the orthogonal MAC case and, $\mathbf{n}_h$, for the coherent MAC case are both assumed to be iid w.l.o.g., with covariance matrices $\mathbf{C}_{h_i} = \sigma_{h_i}^2 \mathbf{I}$ with $\sigma_{h_i}^2 > 0$ for all i and $\mathbf{C}_h = \sigma_h^2 \mathbf{I}$ with $\sigma_h^2 > 0$, respectively - that the channel matrices $\mathbf{H}_i$ are assumed to be diagonal w.l.o.g., for the case of an orthogonal MAC. Furthermore, we assume only for the coherent MAC that $\mathbf{g}_i \in \mathrm{R}(\mathbf{C}_{n_i}) \setminus \{\mathbf{0}\}$ for all i. The reason for this particular assumption is explained later by itself, when we consider the coherent MAC case in Subsubsection 4.4.3.2.

Let us now customize the notations for our system model, especially for the scalar parameter case. Since $n = 1$, the observation matrix $\mathbf{G}_i \in \mathbb{R}^{m_i \times 1}$ for sensor i reduces to vector - thus we use the notation $\mathbf{g}_i \in \mathbb{R}^{m_i}$ and set $\mathbf{G}_i = \mathbf{g}_i$ in what follows. Note that $m_i \geq 1$. As a consequence, the notation $\mathbf{G} \in \mathbb{R}^{k \times 1}$ from (3.16) reduces to a vector too - thus we use $\mathbf{g} \in \mathbb{R}^k$ an set $\mathbf{G} = \mathbf{g}$, accordingly. Thus, $\mathbf{g}$ is in the form $\mathbf{g} = \begin{bmatrix} \mathbf{g}_1^T & \mathbf{g}_2^T & \ldots & \mathbf{g}_L^T \end{bmatrix}^T$. Both deterministic vectors $\mathbf{g}_i$ and $\mathbf{g}$ stands for the observation vector for sensor i and for the total observation vector, respectively.

Let us first recall the FIMs $\mathbf{J}_{\mathbf{z}_i}$ from (3.5) and $\mathbf{J}_\mathbf{z}$ from (3.8), both evaluated for $\mathbf{C}_{l_i} = \mathbf{0}$ and $\mathbf{C}_{h_i} = \sigma_{h_i}^2 \mathbf{I}$ - for the orthogonal MAC case. Adapted to our notations for the scalar case, this means:

$$J_{\mathbf{z}_i} = \mathbf{g}_i^T \mathbf{A}_i^T \mathbf{H}_i^T \left(\sigma_{h_i}^2 \mathbf{I} + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{H}_i \mathbf{A}_i \mathbf{g}_i \tag{4.32}$$

and thus

$$J_\mathbf{z} = \sum_{i=1}^{L} \mathbf{g}_i^T \mathbf{A}_i^T \mathbf{H}_i^T \left(\sigma_{h_i}^2 \mathbf{I} + \mathbf{H}_i \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{H}_i \mathbf{A}_i \mathbf{g}_i. \tag{4.33}$$

For the coherent MAC case, we recall the FIM $\mathbf{J}_\mathbf{z}$ from (3.20) for $\mathbf{C}_l = \mathbf{0}$ and $\mathbf{C}_h = \sigma_h^2 \mathbf{I}$, thus

$$J_\mathbf{z} = \mathbf{g}^T \widetilde{\mathbf{A}}^T \left(\sigma_h^2 \mathbf{I} + \widetilde{\mathbf{A}} \mathbf{C}_n \widetilde{\mathbf{A}}^T\right)^{-1} \widetilde{\mathbf{A}} \mathbf{g}. \tag{4.34}$$

Note that the model parameters for (4.34), are given in (3.16) and (3.17) for $\mathbf{G}_i = \mathbf{g}_i$ and $\mathbf{G} = \mathbf{g}$.

The remaining part of this section is organized as follows: Before we solve the optimization problem (P-II) or (P-III), especially for the scalar parameter case, i.e., we consider problem (P-II) for $\mathbf{J}_{\mathbf{z}} = J_{\mathbf{z}}$ from (4.33) or (4.34), and problem (P-III) for $\mathbf{J}_{\mathbf{z}_i} = J_{\mathbf{z}_i}$ from (4.32) - we first show that for the special case of a scalar parameter $\theta \in \mathbb{R}$, a local sensor rule LO_i can be reduced to an equivalent scalar observation model w.l.o.g. Subsequently, we can solve a simplified, but equivalent optimization problem. Then we give an optimal power scheduling strategie, where a given total power is optimally scheduled among all sensors. Finally, we will show how we can implement optimal local sensors.

4.4.1 Reduction to Scalar Observation

Let us consider the original observation model for a local sensor i, which spezializes for a scalar parameter θ to

$$\mathbf{y}_i = \mathbf{g}_i \theta + \mathbf{n}_i, \tag{4.35}$$

where $\mathbf{g}_i \in \mathbb{R}^{m_i}$ is the known, deterministic observation vector os sensor i and $\mathbf{n}_i$ is again the observation noise vector, i.e., $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{n_i})$. The observation model for sensor i is illustrated in Fig. 4.3, where the ith local sensor rule LO_i is also shown, even here with the additive systematic noise $\mathbf{n}_{l_i}$. Note that the following theorem considers $\mathbf{n}_{l_i}$ with some $\mathbf{C}_{l_i} \geq \mathbf{0}$, even though we already know how it is to be chosen optimally for (P-II). However,

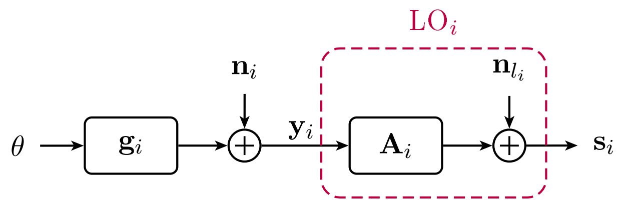


Figure 4.3: Observation model for the ith local sensor for a scalar parameter.

$\mathbf{y}_i$ from (4.35) can be equivalently modeled as a Gaussian distributed random variable, i.e.,

$$\mathbf{y}_i \sim \mathcal{N}(\mathbf{g}_i \theta, \mathbf{C}_{n_i}). \tag{4.36}$$

The next theorem shows an equivalence of two local sensor rules LOs, where the equivalence is based on Definition 4.4.15. As a result, we conclude that we can reduce our original observation model from (4.35) - with a vector observation $\mathbf{y}_i$ - to an equivalent observation model, with an appropriate scalar observation y_i. To that end, we first decompose the observation vector $\mathbf{g}_i$ for sensor i into

$$\mathbf{g}_i = \mathbf{g}'_i + \mathbf{g}'_{i\perp}, \tag{4.37}$$

where $\mathbf{g}'_i$ denotes those components of $\mathbf{g}$, which lies in the range of $\mathbf{C}_{n_i}$, i.e., $\mathbf{g}'_i \in \mathrm{R}(\mathbf{C}_{n_i})$ and $\mathbf{g}'_{i\perp}$ denotes consequently those components of $\mathbf{g}$, which lies in the nullspace of $\mathbf{C}_{n_i}$, i.e., $\mathbf{g}'_{i\perp} \in \mathrm{N}(\mathbf{C}_{n_i})$[1]. Let us also define even the equivalence of two LOs for an observation $\mathbf{y} = \mathbf{y}_i$.

Definition 4.4.15 *Two local sensor rules,* $\mathrm{LO}' = (\mathbf{A}', \mathbf{C}'_l)$ *and* $\mathrm{LO}'' = (\mathbf{A}'', \mathbf{C}''_l)$ *are equivalent, if and only if the corresponding outputs* $\mathbf{s}'$ *and* $\mathbf{s}''$*, i.e.,* $\mathbf{s}' = \mathbf{A}'\mathbf{y} + \mathbf{n}'$ *and* $\mathbf{s}'' = \mathbf{A}''\mathbf{y} + \mathbf{n}''$ *have the same pdf for every* θ*, i.e.,* $f_{\mathbf{s}'}(\mathbf{s};\theta) = f_{\mathbf{s}''}(\mathbf{s};\theta)$.

Exclusively for the next theorem we use the specific observation model parameters (cf. (4.35), (4.36) and (4.37)): $\mathbf{y} = \mathbf{y}_i$, $\mathbf{g} = \mathbf{g}_i$ (thus $\mathbf{g}' = \mathbf{g}'_i$ and $\mathbf{g}'_\perp = \mathbf{g}'_{i\perp}$), $\mathbf{n} = \mathbf{n}_i$ and finally $\mathbf{C}_n = \mathbf{C}_{n_i}$.

Theorem 4.4.16 *The set of local sensor rules* $\mathrm{LO}s$*, given by*

$$\{\mathbf{A}, \mathbf{C}_l\}_{\mathbf{A}\in\mathbb{R}^{n\times m}, \mathbf{C}_l \geq \mathbf{0}}, \tag{4.38}$$

is equivalent to the set of $\mathrm{LO}s$ *given by*

$$\left\{\mathbf{A}' = \mathbf{A}\mathbf{a}_2\mathbf{a}_1^T, \mathbf{C}''_l = \mathbf{C}_l + \mathbf{A}\mathbf{C}'_l\mathbf{A}^T\right\}_{\mathbf{A}\in\mathbb{R}^{n\times m}, \mathbf{C}_l \geq \mathbf{0}}, \tag{4.39}$$

where $\mathbf{a}_1$, $\mathbf{a}_2$ *and* $\mathbf{C}'_l$ *are to be chosen as follows, depending on the observation vector* $\mathbf{g}$ *and the observation noise covariance matrix* $\mathbf{C}_n$*: For the first case (case1), when* $\mathbf{g} \in \mathrm{R}(\mathbf{C}_n) \setminus \{\mathbf{0}\}$*, then*

$$\mathbf{a}_1 \triangleq \mathbf{C}_n^\dagger \mathbf{g}, \quad \mathbf{a}_2 \triangleq \frac{1}{\mathbf{g}^T\mathbf{C}_n^\dagger\mathbf{g}}\mathbf{g} \quad \textit{and} \quad \mathbf{C}'_l \triangleq \frac{\mathbf{g}\mathbf{g}^T}{\mathbf{g}^T\mathbf{C}_n^\dagger\mathbf{g}},$$

where $\mathbf{C}_n^\dagger$ *is the pseudo-inverse of* $\mathbf{C}_n$*. For the second case (case2), when* $\mathbf{g}$ *has at least one component in* $\mathrm{N}(\mathbf{C}_n)$ *or, equivalently,* $\mathbf{g}'_\perp \neq \mathbf{0}$*, then*

$$\mathbf{a}_1 \triangleq \mathbf{g}'_\perp, \quad \mathbf{a}_2 \triangleq \frac{\mathbf{g}}{\mathbf{g}'^T_\perp \mathbf{g}'_\perp} \quad \textit{and} \quad \mathbf{C}'_l \triangleq \mathbf{C}_n.$$

The last case (case3) is the trivial one, when $\mathbf{g} = \mathbf{0}$*. Then, we can define (where* $\mathbf{a}_2$ *could be chosen arbitrarily)*

$$\mathbf{a}_1 \triangleq \mathbf{0}, \quad \mathbf{a}_2 \triangleq \mathbf{0} \quad \textit{and} \quad \mathbf{C}'_l \triangleq \mathbf{C}_n.$$

The equivalence of the sets (4.38) *and* (4.39) *is taken according to Definition 4.4.15. Note that* $\mathbf{C}'_l$ *has to be simulated/generated by some additive Gaussian and zero-mean systematic noise* $\mathbf{n}'_l$*. The structure (setup) of the original- and the equivalent* $\mathrm{LO}s$ *are illustrated in Fig. 4.4.*

[1] Note that for a symmetric matrix $\mathbf{C}$: $\{\mathrm{R}(\mathbf{C})\}^\perp = \mathrm{N}(\mathbf{C})$.

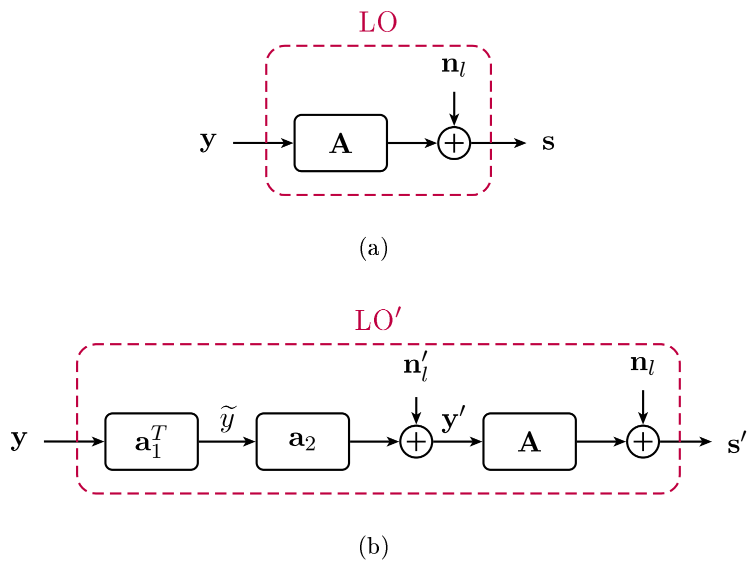


Figure 4.4: Equivalence of LOs (a) original LO model (b) equivalent LO model.

Proof. Let us now consider the equivalent LO′ model from Fig. 4.4(b), where we assume that $\mathbf{n}_l'$ is uncorrelated with the observation noise vector $\mathbf{n}$ and, furthermore, with the systematic noise vector $\mathbf{n}_l$. As indicated in theorem, we have to differentiate three cases, depending on how $\mathbf{g}$ and $\mathbf{C}_n$ is given:

1. We assume $\mathbf{g} \in \mathrm{R}(\mathbf{C}_n) \setminus \{\mathbf{0}\}$, i.e., $\mathbf{g}$ has no component(s) in $\mathrm{N}(\mathbf{C}_n)$. Then, we set

$$\mathbf{a}_1 = \mathbf{C}_n^\dagger \mathbf{g}. \tag{4.40}$$

First, we note that the application of $\mathbf{a}_1^T$ reduces the vector $\mathbf{y}$ to the scalar random variable $\widetilde{y}$, given by

$$\widetilde{y} = \mathbf{a}_1^T \mathbf{y} = \mathbf{a}_1^T \mathbf{g}\theta + \mathbf{a}_1^T \mathbf{n} = \mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}\theta + \widetilde{n} = \widetilde{g}\theta + \widetilde{n}, \tag{4.41}$$

where $\widetilde{g} = \mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}$ and

$$\widetilde{n} \sim \mathcal{N}\left(0, \widetilde{\sigma}^2 = \mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}\right), \tag{4.42}$$

because $\widetilde{\sigma}^2 = \mathbf{a}_1^T \mathbf{C}_n \mathbf{a}_1 = \mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{C}_n \mathbf{C}_n^\dagger \mathbf{g} = \mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}$[1]. Thus

$$\mathbf{y}' = \mathbf{a}_2 \widetilde{y} + \mathbf{n}_l' \overset{(4.41)}{=} \mathbf{a}_2 \mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}\theta + \mathbf{a}_2 \widetilde{n} + \mathbf{n}_l', \tag{4.43}$$

[1]The pseudo-inverse $\mathbf{C}_n^\dagger$ of the covariance matrix $\mathbf{C}_n$ (symmetric, positive semi-definite) is symmetric and positive semi-definite - both have the same eigenspace (eigenvectors), in particular.

where $\mathbf{n}_l' \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_l')$. Choosing

$$\mathbf{a}_2 = \frac{1}{\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}} \mathbf{g}, \tag{4.44}$$

we obtain

$$\mathbf{y}' = \mathbf{g}\theta + \mathbf{n}', \tag{4.45}$$

where $\mathbf{n}' \sim \mathcal{N}(\mathbf{0}, \mathbf{C}')$, with

$$\begin{aligned} \mathbf{C}' &= \mathbf{a}_2 \widetilde{\sigma}^2 \mathbf{a}_2^T + \mathbf{C}_l' \\ &\overset{(4.44)}{=} \frac{1}{\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}} \mathbf{g}\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}\mathbf{g}^T \frac{1}{\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}} + \mathbf{C}_l' \\ &= \frac{1}{\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}} \mathbf{g}\mathbf{g}^T + \mathbf{C}_l'. \end{aligned} \tag{4.46}$$

Note that $\mathbf{a}_2$ from (4.44) is realizable, since $\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g} > 0$, due to the assumption that $\mathbf{g}$ is orthogonal to $\mathrm{N}(\mathbf{C}_n)$ and $\mathbf{g} \neq \mathbf{0}$. From (4.46) it is evident that the choice

$$\mathbf{C}_l' = \mathbf{C}_n - \frac{1}{\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}} \mathbf{g}\mathbf{g}^T, \tag{4.47}$$

yields

$$\mathbf{C}' \overset{(4.46)}{=} \frac{1}{\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}} \mathbf{g}\mathbf{g}^T + \mathbf{C}_l' = \mathbf{C}_n. \tag{4.48}$$

It can be argued that any noise $\mathbf{n}_l'$ with positive semi-definite covariance matrix $\mathbf{C}_l'$ can be simulated/generated. Consequently, we now examine whether $\mathbf{x}^T \mathbf{C}_l' \mathbf{x} \geq \mathbf{0}$ is valid for all $\mathbf{x} \in \mathbb{R}^m$. To that end, we decompose $\mathbf{x}$, analog to (4.37), into

$$\mathbf{x} = \mathbf{x}' + \mathbf{x}_\perp', \tag{4.49}$$

where $\mathbf{x}' \in \mathrm{R}(\mathbf{C}_n)$ and $\mathbf{x}_\perp' \in \mathrm{N}(\mathbf{C}_n) = \{\mathrm{R}(\mathbf{C}_n)\}^\perp$, i.e., $\mathbf{C}_n \mathbf{x}_\perp' = \mathbf{0}$. Because of our assumption that $\mathbf{g}$ is orthogonal to $\mathrm{N}(\mathbf{C}_n)$ and thus $\mathbf{g}^T \mathbf{x}_\perp' = \mathbf{0}$, it follows that

$$\mathbf{x}_\perp'^T \mathbf{C}_l' \mathbf{x}_\perp' \overset{(4.47)}{=} \mathbf{x}_\perp'^T \mathbf{C}_n \mathbf{x}_\perp' - \mathbf{x}_\perp'^T \frac{1}{\mathbf{g}^T \mathbf{C}_n^\dagger \mathbf{g}} \mathbf{g}\mathbf{g}^T \mathbf{x}_\perp' = 0 - 0 = 0. \tag{4.50}$$

Thus, it remains to verify that $\mathbf{x}'^T \mathbf{C}_l' \mathbf{x}' \geq 0$ for all $\mathbf{x}' \in \mathrm{R}(\mathbf{C}_n)$. In what follows, we use the fact that

$$\mathbf{x}' = \mathbf{C}_n^\dagger \mathbf{x}, \tag{4.51}$$

because $\mathrm{R}\left(\mathbf{C}_n^\dagger\right) = \mathrm{R}\left(\mathbf{C}_n\right)$. Thus,

$$\begin{aligned}
\mathbf{x}'^T\mathbf{C}_l'\mathbf{x}' &\overset{(4.51)}{=} \mathbf{x}^T\mathbf{C}_n^\dagger\mathbf{C}_l'\mathbf{C}_n^\dagger\mathbf{x} \\
&\overset{(4.47)}{=} \mathbf{x}^T\left(\mathbf{C}_n^\dagger\mathbf{C}_n\mathbf{C}_n^\dagger - \mathbf{C}_n^\dagger\frac{1}{\mathbf{g}^T\mathbf{C}_n^\dagger\mathbf{g}}\mathbf{g}\mathbf{g}^T\mathbf{C}_n^\dagger\right)\mathbf{x} \\
&= \mathbf{x}^T{\mathbf{C}_n^\dagger}^{1/2}{\mathbf{C}_n^\dagger}^{1/2}\mathbf{x} - \mathbf{x}^T{\mathbf{C}_n^\dagger}^{1/2}\frac{{\mathbf{C}_n^\dagger}^{1/2}\mathbf{g}\mathbf{g}^T{\mathbf{C}_n^\dagger}^{1/2}}{\mathbf{g}^T{\mathbf{C}_n^\dagger}^{1/2}{\mathbf{C}_n^\dagger}^{1/2}\mathbf{g}}{\mathbf{C}_n^\dagger}^{1/2}\mathbf{x} \\
&\overset{(a)}{=} \widetilde{\mathbf{x}}^T\widetilde{\mathbf{x}} - \widetilde{\mathbf{x}}^T\frac{\widetilde{\mathbf{g}}\widetilde{\mathbf{g}}^T}{\widetilde{\mathbf{g}}^T\widetilde{\mathbf{g}}}\widetilde{\mathbf{x}} \\
&= \|\widetilde{\mathbf{x}}\|^2 - \frac{\left(\widetilde{\mathbf{x}}^T\widetilde{\mathbf{g}}\right)^2}{\|\widetilde{\mathbf{g}}\|^2},
\end{aligned} \tag{4.52}$$

where in step (a) we introduced $\widetilde{\mathbf{x}} \triangleq {\mathbf{C}_n^\dagger}^{1/2}\mathbf{x}$ and $\widetilde{\mathbf{g}} \triangleq {\mathbf{C}_n^\dagger}^{1/2}\mathbf{g}$. Using the Cauchy-Schwarz inequality, i.e.,

$$\left(\widetilde{\mathbf{x}}^T\widetilde{\mathbf{g}}\right)^2 \leq \|\widetilde{\mathbf{x}}\|^2\|\widetilde{\mathbf{g}}\|^2, \tag{4.53}$$

we can show that

$$\begin{aligned}
\mathbf{x}'^T\mathbf{C}_l'\mathbf{x}' &\overset{(4.52)}{=} \|\widetilde{\mathbf{x}}\|^2 - \left(\widetilde{\mathbf{x}}^T\widetilde{\mathbf{g}}\right)^2/\|\widetilde{\mathbf{g}}\|^2 \\
&\overset{(4.53)}{\geq} \|\widetilde{\mathbf{x}}\|^2 - \|\widetilde{\mathbf{x}}\|^2\|\widetilde{\mathbf{g}}\|^2/\|\widetilde{\mathbf{g}}\|^2 = \|\widetilde{\mathbf{x}}\|^2 - \|\widetilde{\mathbf{x}}\|^2 = 0,
\end{aligned} \tag{4.54}$$

and thus $\mathbf{x}'^T\mathbf{C}_l'\mathbf{x}' \geq 0$ for all $\mathbf{x}' \in \mathrm{R}\left(\mathbf{C}_n\right)$. Hence, $\mathbf{C}_l'$ is positive semi-definite.

2. Now we consider the case, when $\mathbf{g}$ has at least one component in $\mathrm{N}\left(\mathbf{C}_n\right)$. In that case, we set

$$\mathbf{a}_1 = \mathbf{g}_\perp'. \tag{4.55}$$

The application of $\mathbf{a}_1^T$ reduces the vector $\mathbf{y}$ again to the scalar random variable $\widetilde{y}$, given by

$$\widetilde{y} = \mathbf{a}_1^T\mathbf{y} = \mathbf{a}_1^T\mathbf{g}\theta + \mathbf{a}_1^T\mathbf{n} \overset{(4.55)}{=} \mathbf{g}_\perp'^T\mathbf{g}\theta + \widetilde{n} = \widetilde{g}\theta + \widetilde{n}, \tag{4.56}$$

where $\widetilde{g} = \mathbf{g}_\perp'^T\mathbf{g}$ and $\widetilde{n} \sim \mathcal{N}\left(0, \widetilde{\sigma}^2 = \mathbf{g}_\perp'^T\mathbf{C}_n\mathbf{g}_\perp' = 0\right)$ and thus can be neglected, i.e., $\widetilde{n} = 0$. Thus

$$\mathbf{y}' = \mathbf{a}_2\widetilde{y} + \mathbf{n}_l' \overset{(4.56)}{=} \mathbf{a}_2\mathbf{g}^T\mathbf{g}\theta + \mathbf{n}_l' \tag{4.57}$$

and by choosing

$$\mathbf{a}_2 = \frac{\mathbf{g}}{\mathbf{g}_\perp'^T \mathbf{g}_\perp'}, \tag{4.58}$$

we obtain

$$\mathbf{y}' = \mathbf{g}\theta + \mathbf{n}'. \tag{4.59}$$

where $\mathbf{n}' \sim \mathcal{N}(\mathbf{0}, \mathbf{C}')$ with $\mathbf{C}' = \mathbf{C}_l'$. Thus, we can choose $\mathbf{C}_l' = \mathbf{C}_n$, such that $\mathbf{C}' = \mathbf{C}_n$, which is obviously positive semi-definite and thus can be simulated/generated.

3. Last but not least, we now consider the trivial case, when $\mathbf{g} = \mathbf{0}$. Setting $\mathbf{a}_1 = \mathbf{0}$, implies

$$\widetilde{y} = \mathbf{a}_1^T \mathbf{y} = \mathbf{a}_1^T \mathbf{g}\theta + \mathbf{a}_1^T \mathbf{n} = 0 = \widetilde{g}\theta + \widetilde{n}, \tag{4.60}$$

where $\widetilde{g} = 0$ and $\widetilde{n} = 0$. It is evident that $\mathbf{y}' = \mathbf{n}_l'$ and thus we obtain

$$\mathbf{y}' = \mathbf{g}\theta + \mathbf{n}', \tag{4.61}$$

where $\mathbf{g} = \mathbf{0}$ and $\mathbf{n}' \sim \mathcal{N}(\mathbf{0}, \mathbf{C}')$ with $\mathbf{C}' = \mathbf{C}_l'$. We can again choose $\mathbf{C}_l' = \mathbf{C}_n$, such that $\mathbf{C}' = \mathbf{C}_n$, which is obviously positive semi-definite and thus can be simulated/generated.

Hence, we have verified the equivalence of LO and LO′ - the equivalence of the sets (4.38) and (4.39) - according to Definition 4.4.15. Since for an arbitrary observation vector $\mathbf{g}$, we can determine deterministic vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and a zero-mean, Gaussian distributed systematic noise $\mathbf{n}_l'$ with covariance matrix $\mathbf{C}_l' \geq \mathbf{0}$, accordingly, for obtaining equivalence between $\mathbf{y}$ and $\mathbf{y}'$ and thus also between $\mathbf{s}$ and $\mathbf{s}'$ (cf. Fig. 4.4). □

In Theorem 4.4.16, we have showed that there always exist an equivalent LO′ for a given original LO, if we consider a scalar parameter θ. Let us again consider a specific observation model for sensor i in our original notation. As a corollary, we can design an equivalent local sensor rule, denoted by LO_i', also for an observation model according to (4.41) (or (4.56) or (4.60)). In what follows, we thus consider a scalar observation model y_i w.l.o.g., in order to simplify the optimization problem later, i.e.,

$$y_i = g_i\theta + n_i, \quad \text{with } n_i \sim \mathcal{N}\left(0, \sigma_{n_i}^2\right). \tag{4.62}$$

For the *first case*, when $\mathbf{g}_i \in \mathrm{R}(\mathbf{C}_{n_i}) \setminus \{\mathbf{0}\}$, the model parameters g_i and $\sigma_{n_i}^2$ for (4.62), can be obtained according to (4.41) and (4.42) as

$$g_i = \mathbf{g}_i^T \mathbf{C}_{n_i}^\dagger \mathbf{g}_i, \qquad \sigma_{n_i}^2 = \mathbf{g}_i^T \mathbf{C}_{n_i}^\dagger \mathbf{g}_i. \tag{4.63}$$

Note that for that case $g_i > 0$ and $\sigma_{n_i}^2 > 0$ is guaranteed. For the *second case*, when $\mathbf{g}_i$ has at least one component in $\mathrm{N}(\mathbf{C}_{n_i})$, i.e., $\mathbf{g}_{i\perp}' \neq \mathbf{0}$, the model parameters g_i and $\sigma^2_{n_i}$ for (4.62) can be obtained according to (4.56) for $\widetilde{n} = 0$ as

$$g_i = \mathbf{g}'^T_{i\perp}\mathbf{g}'_{i\perp}, \qquad \sigma^2_{n_i} = 0, \tag{4.64}$$

where again $g_i > 0$. Recently, the trivial *third case*, i.e., when $\mathbf{g}_i = \mathbf{0}$. Then the model parameters are both zero, i.e.,

$$g_i = 0, \qquad \sigma^2_{n_i} = 0. \tag{4.65}$$

Based on the assumption of non-correlation between the individual observation noise vectors $\mathbf{n}_i$ for all i in our original system model, i.e., $\mathrm{cov}\left\{\mathbf{n}_i, \mathbf{n}_j\right\} = \mathbf{0}$ for $1 \leq i, j \leq L$ with $i \neq j$, we conclude that also n_i from (4.62) for all i are uncorrelated across different sensors. Since n_i, follows by a linear mapping of $\mathbf{n}_i$ onto the real line, i.e., in the form $n_i = \mathbf{a}_1^T \mathbf{n}_i$, where $\mathbf{a}_1$ is a deterministic vector (cf. Theorem 4.4.16). However,

$$\mathrm{cov}\left\{\mathbf{a}_1^T\mathbf{n}_i, \mathbf{a}_2^T\mathbf{n}_j\right\} = \mathrm{E}\left\{\mathbf{a}_1^T\mathbf{n}_i\mathbf{n}_j^T\mathbf{a}_2\right\} = \mathbf{a}_1^T \mathrm{cov}\left\{\mathbf{n}_i, \mathbf{n}_j\right\}\mathbf{a}_2 = 0$$

for $1 \leq i, j \leq L$ with $i \neq j$ and for arbitrary deterministic vectors $\mathbf{a}_1$ and $\mathbf{a}_2$.

The equivalent local sensor rule LO'_i, based on the scalar observation model y_i from (4.62), can thus be described by linear mapping with a sensor vector $\mathbf{a}_i$ and additive systematic noise $\mathbf{n}'_{l_i}$ (cf. Fig. 4.5), i.e., $\mathrm{LO}'_i \triangleq \left(\mathbf{a}, \mathbf{C}'_{l_i}\right)$. The next corollary shows how we can determine the original $\mathrm{LO}_i = (\mathbf{A}_i, \mathbf{C}_{l_i})$

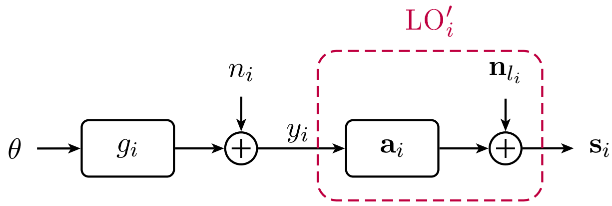


Figure 4.5: Equivalent (scalar) observation model for the ith local sensor for a scalar parameter.

from a given $\mathrm{LO}'_i = (\mathbf{a}_i, \mathbf{C}'_{l_i})$.

Corollary 4.4.17 *Two local sensor rules* $\mathrm{LO}_i = (\mathbf{A}_i, \mathbf{C}_{l_i})$ *and* $\mathrm{LO}'_i = \left(\mathbf{a}_i, \mathbf{C}'_{l_i}\right)$ *are equivalent iff it is of the form* $\mathbf{A}_i = \mathbf{a}_i\mathbf{a}_1^T$ *and* $\mathbf{C}_{l_i} = \mathbf{C}'_{l_i}$*, where* $\mathbf{a}_1$ *is defined in Theorem 4.4.16, i.e.,* $\mathbf{a}_1 = \mathbf{C}^{\dagger}_{n_i}\mathbf{g}_i$ *for the case, when* $\mathbf{g}_i \in \mathrm{R}\left(\mathbf{C}_{n_i}\right) \setminus \{\mathbf{0}\}$ *- and* $\mathbf{a}_1 = \mathbf{g}'_{i\perp}$*, when* $\mathbf{g}_i$ *has at least one component in* $\mathrm{N}\left(\mathbf{C}_{n_i}\right)$*, i.e., when* $\mathbf{g}'_{i\perp} \neq \mathbf{0}$ *- and finally* $\mathbf{a}_1 = \mathbf{0}$*, when* $\mathbf{g}_i = \mathbf{0}$*.*

Proof. The $\mathrm{LO}_i = (\mathbf{A}^*_i, \mathbf{C}_{l_i})$ performs a linear mapping of $\mathbf{y}_i$, given in (4.35), to the transmit data $\mathbf{s}_i$ as

$$\mathbf{s}_i = \mathbf{A}_i\mathbf{g}_i\theta + \mathbf{A}_i\mathbf{n}_i + \mathbf{n}_{l_i}, \quad \text{where } \mathbf{s}_i \sim \mathcal{N}\left(\mathbf{A}_i\mathbf{g}_i\theta,\ \mathbf{A}_i\mathbf{C}_{n_i}\mathbf{A}_i^T + \mathbf{C}_{l_i}\right). \tag{4.66}$$

Analog, the $\mathrm{LO}_i' = \left(\mathbf{a}_i, \mathbf{C}_{l_i}'\right)$ performs a linear mapping of y_i, given in (4.62), to the transmit data $\mathbf{s}_i'$ as

$$\mathbf{s}_i' = \mathbf{a}_i g_i \theta + \mathbf{a}_i n_i + \mathbf{n}_{l_i}'. \tag{4.67}$$

For the first case, i.e., when $\mathbf{g}_i \in \mathrm{R}\left(\mathbf{C}_{n_i}\right) \setminus \{\mathbf{0}\}$, (4.67) with (4.63) yields

$$\mathbf{s}_i' \sim \mathcal{N}\left(\mathbf{a}_i \mathbf{g}_i^T \mathbf{C}_{n_i}^\dagger \mathbf{g}_i \theta,\ \mathbf{a}_i \mathbf{g}_i^T \mathbf{C}_{n_i}^\dagger \mathbf{g}_i \mathbf{a}_i^T + \mathbf{C}_{l_i}'\right). \tag{4.68}$$

According to Definition 4.4.15, LO_i and LO_i' are equivalent iff $\mathbf{s}$ from (4.66) and $\mathbf{s}'$ from (4.68) have the same pdf for every θ - equivalence follows with $\mathbf{a}_1 = \mathbf{C}_{n_i}^\dagger \mathbf{g}_i$:

$$\begin{aligned}
\mathbf{A}_i &= \mathbf{a}_i \mathbf{g}_i^T \mathbf{C}_{n_i}^\dagger = \mathbf{a}_i \mathbf{a}_1^T \quad \text{and} \\
\mathbf{C}_{l_i} &= \mathbf{a}_i \mathbf{g}_i^T \mathbf{C}_{n_i}^\dagger \mathbf{g}_i \mathbf{a}_i^T + \mathbf{C}_{l_i}' - \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \\
&\overset{(a)}{=} \mathbf{a}_i \mathbf{g}_i^T \mathbf{C}_{n_i}^\dagger \mathbf{C}_{n_i} \mathbf{C}_{n_i}^\dagger \mathbf{g}_i \mathbf{a}_i^T + \mathbf{C}_{l_i}' - \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \\
&= \mathbf{a}_i \mathbf{a}_1^T \mathbf{C}_{n_i} \mathbf{a}_1 \mathbf{a}_i^T + \mathbf{C}_{l_i}' - \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T \\
&= \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T + \mathbf{C}_{l_i}' - \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T = \mathbf{C}_{l_i}',
\end{aligned} \tag{4.69}$$

where in step (a) we used $\mathbf{C}_{n_i}^\dagger = \mathbf{C}_{n_i}^\dagger \mathbf{C}_{n_i} \mathbf{C}_{n_i}^\dagger$. For the second case, i.e., when $\mathbf{g}_{i\perp}' \neq \mathbf{0}$, (4.67) with (4.64) yields

$$\mathbf{s}_i' \sim \mathcal{N}\left(\mathbf{a}_i \mathbf{g}_{i\perp}'^T \mathbf{g}_{i\perp}' \theta,\ \mathbf{C}_{l_i}'\right). \tag{4.70}$$

Again, LO_i and LO_i' are equivalent iff $\mathbf{s}$ from (4.66) and $\mathbf{s}'$ from (4.70) have the same pdf for every θ - equivalence thus follows with $\mathbf{a}_1 = \mathbf{g}_{i\perp}'$:

$$\begin{aligned}
\mathbf{A}_i &= \mathbf{a}_i \mathbf{g}_{i\perp}'^T = \mathbf{a}_i \mathbf{a}_1^T \quad \text{and} \\
\mathbf{C}_{l_i} &= \mathbf{C}_{l_i}' - \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T = \mathbf{C}_{l_i}' - \mathbf{a}_i \mathbf{g}_{i\perp}'^T \mathbf{C}_{n_i} \mathbf{g}_{i\perp}' \mathbf{a}_i^T \overset{(a)}{=} \mathbf{C}_{l_i}',
\end{aligned} \tag{4.71}$$

where in step (a) we used $\mathbf{C}_{n_i} \mathbf{g}_{i\perp}' = \mathbf{0}$, since $\mathbf{g}_{i\perp}' \in \mathrm{N}\left(\mathbf{C}_{n_i}\right)$. Finally for the case, when $\mathbf{g}_i = \mathbf{0}$, (4.67) with (4.65) yields

$$\mathbf{s}_i' \sim \mathcal{N}\left(\mathbf{0},\ \mathbf{C}_{l_i}'\right). \tag{4.72}$$

Again, LO_i and LO_i' are equivalent iff $\mathbf{s}$ from (4.66) and $\mathbf{s}'$ from (4.72) have the same pdf for every θ - equivalence thus follows with $\mathbf{a}_1 = \mathbf{0}$:

$$\begin{aligned}
\mathbf{A}_i &= \mathbf{0} = \mathbf{a}_i \mathbf{0} \quad \text{and} \\
\mathbf{C}_{l_i} &= \mathbf{C}_{l_i}' - \mathbf{A}_i \mathbf{C}_{n_i} \mathbf{A}_i^T = \mathbf{C}_{l_i}'.
\end{aligned} \tag{4.73}$$

Hence, we have verified the equivalence of LO_i and LO_i' for all (three) cases of $\mathbf{g}$. $\square$

4.4.2 System Model for the Equivalent Model Reformulated

Let us now consider the scalar observation model from (4.62) and the equivalent local sensor rule LO_i' for sensor i in what follows. Note assuming a scalar observation model, corresponds to the original system model with $m_i = 1$ for all i. We also recall that for the original LO_i: the optimal systematic noise covariance matrix $\mathbf{C}_{l_i}^* = \mathbf{0}$ for all i (cf. introduction of this section) and since $\mathbf{C}_{l_i}' = \mathbf{C}_{l_i}$ (cf. Corollary 4.4.17), we conclude that also $\mathbf{C}_{l_i}'^* = \mathbf{0}$, as expected. Thus, we restrict our equivalent local sensor rule LO_i' by the sensor vector $\mathbf{a}_i$ in what follows. Let us recall the transmit data vector $\mathbf{s}_i$ for the scalar observation model. The LO_i' performs a linear mapping of y_i, given in (4.62), to the transmit data $\mathbf{s}_i$ as

$$\mathbf{s}_i = \mathbf{a}_i y_i. \tag{4.74}$$

Depending on the different multiple access schemes, we will now customize the scalar-valued FIs $J_{\mathbf{z}_i}$ and $J_{\mathbf{z}}$, given in (4.32), (4.33) and (4.34) - especially to our equivalent model with scalar observation from (4.62) and to the equivalent local sensor rule LO_i'.

Orthogonal MAC: In that case, it is evident that the scalar-valued FI $J_{\mathbf{z}_i}$ from (4.32), then specializes to

$$J_{\mathbf{z}_i} = g_i^2 \mathbf{a}_i^T \mathbf{H}_i^T \left(\sigma_{h_i}^2 \mathbf{I} + \sigma_{n_i}^2 \mathbf{H}_i \mathbf{a}_i \mathbf{a}_i^T \mathbf{H}_i^T \right)^{-1} \mathbf{H}_i \mathbf{a}_i, \tag{4.75}$$

and thus, the FI $J_{\mathbf{z}}$ from (4.33) yields

$$J_{\mathbf{z}} = \sum_{i=1}^{L} g_i^2 \mathbf{a}_i^T \mathbf{H}_i^T \left(\sigma_{h_i}^2 \mathbf{I} + \sigma_{n_i}^2 \mathbf{H}_i \mathbf{a}_i \mathbf{a}_i^T \mathbf{H}_i^T \right)^{-1} \mathbf{H}_i \mathbf{a}_i, \tag{4.76}$$

where the parameters g_i and $\sigma_{n_i}^2$ are given in (4.63), (4.64) or (4.65), depending on the given original observation model parameters $\mathbf{g}_i$ and $\mathbf{C}_{n_i}$.

Coherent MAC: In that case, we customize, first, the notations to the system model with scalar observation. We recall the shorthand (3.15), which specializes to a vector as

$$\widetilde{\mathbf{a}}_i \triangleq \mathbf{H}_i \mathbf{a}_i. \tag{4.77}$$

Already made assumptions are, of course, also adopted for this special case, so $\mathbf{H}_i$ for all i are of full column-rank - thus $p \geq q_i$ for all i - we refer to $\widetilde{\mathbf{a}}_i$ as the sensor-channel vector, accordingly. Note that we can reclaim $\mathbf{a}_i$ from $\widetilde{\mathbf{a}}_i$ as $\mathbf{a}_i = \mathbf{H}_i^\dagger \widetilde{\mathbf{a}}_i$ in a unique manner, since $\mathbf{H}_i$ is of full column-rank. We also recall (3.16) and (3.17). Those variables, which are then used in the following, specialize to

$$\begin{aligned}
\widetilde{\mathbf{A}} &\triangleq \begin{pmatrix}\widetilde{\mathbf{a}}_1 & \widetilde{\mathbf{a}}_2 & \dots & \widetilde{\mathbf{a}}_L\end{pmatrix}, \quad \widetilde{\mathbf{A}} \in \mathbb{R}^{p\times L},\\
\mathbf{g} &\triangleq \begin{pmatrix}g_1 & g_2 & \dots & g_L\end{pmatrix}^T, \quad \mathbf{g} \in \mathbb{R}^{L},\\
\mathbf{n} &\triangleq \begin{pmatrix}n_1 & n_2 & \dots & n_L\end{pmatrix}^T, \quad \mathbf{n} \in \mathbb{R}^{L}.
\end{aligned} \tag{4.78}$$

Remember that $p = p_i$ for all i. Note that $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n\}$. It is evident that the covariance matrix $\mathbf{C}_n = \operatorname{diag}\left\{\sigma_{n_i}^2\right\}_{i=1}^L$ and is non-singular for the case, when $\sigma_{n_i}^2 \neq 0$ for all i, which follows when $\mathbf{g}_i$ from the original observation model for sensor i (4.35) holds: $\mathbf{g}_i \in \mathrm{R}(\mathbf{C}_{n_i}) \setminus \{\mathbf{0}\}$ for all i (for the coherent MAC, we have already restricted to this special case - cf. introduction of this section). To that end, we assume iid total observation noise for the scalar observation model $\mathbf{n}$ w.l.o.g., i.e., $\mathbf{C}_n = \sigma_n^2 \mathbf{I}$ with $\sigma_n^2 > 0$. Otherwise, we can define an equivalent model, in which $\mathbf{C}_{n'} = \sigma_{n'}^2 \mathbf{I}$ (cf. Theorem 4.3.12). It is evident that the scalar-valued FI $J_{\mathbf{z}}$ from (4.34), then specializes to

$$J_{\mathbf{z}} = \mathbf{g}^T \widetilde{\mathbf{A}}^T \left(\sigma_h^2 \mathbf{I} + \sigma_n^2 \widetilde{\mathbf{A}} \widetilde{\mathbf{A}}^T\right)^{-1} \widetilde{\mathbf{A}} \mathbf{g}, \tag{4.79}$$

where $\mathbf{g}$ and $\widetilde{\mathbf{A}}$ are given in (4.78).

Power Constraint: Let us recall both constraints (C1) and (C2) evaluated for $\mathbf{C}_{l_i} = \mathbf{0}$ for all i. For a scalar paramter and in particular, for the scalar observation model, (C1) specializes to

$$\mathrm{E}_\theta\left\{\|\mathbf{s}_i\|^2\right\} = \|\mathbf{a}_i\|^2 \left((g_i\theta)^2 + \sigma_{n_i}^2\right) \leq P_{0,i} \quad \text{for } 1 \leq i \leq L, \tag{C1-s}$$

and (C2) to

$$\mathrm{var}_\theta\left\{\mathbf{s}_i\right\} = \|\mathbf{a}_i\|^2 \sigma_{n_i}^2 \leq P'_{0,i} \quad \text{for } 1 \leq i \leq L. \tag{C2-s}$$

Note that constraint (C2-s) does not depend on the parameter θ, as expected.

Total Power Constraint for the Coherent MAC: Later, when we consider the case of a coherent MAC, we will simplify the optimization problem (P-II) by introducing a modified power constraint. Here, we will consider a total power constraint - by which the total transmit power for all sensors, i.e., the sum of all individual powers of $\mathbf{s}_i$ for all i is bounded above a given constant total power $P_0 = \sum_{i=1}^L P_{0,i}$ - i.e., especially adapted to the equivalent model with scalar observation:

$$\sum_{i=1}^{L} \left((g_i\theta)^2 + \sigma_n^2\right) \|\mathbf{a}_i\|^2 \leq P_0. \tag{C1-t}$$

Analog, we define the total variance constraint with $P'_0 = \sum_{i=1}^{L} P'_{0,i}$ as

$$\sum_{i=1}^{L} \|\mathbf{a}_i\|^2 \leq P'_0/\sigma_n^2. \tag{C2-t}$$

Note that for the coherent MAC, we have assumed $\mathbf{C}_n = \sigma_n^2 \mathbf{I}$ with $\sigma_n^2 > 0$ and thus $\sigma_{n_i}^2 = \sigma_n^2$ for all i.

4.4.3 Optimal Local Sensor Rules for a Scalar Parameter

So far, we have introduced an equivalent scalar observation model, in which we restricted an equivalent local sensor rule LO'_i by the sensor vector $\mathbf{a}_i$ - since we set $\mathbf{C}'_{l_i} = \mathbf{0}$ w.l.o.g. We have customized the FI $J_{\mathbf{z}}$ for both multiple access schemes to the equivalent system model with scalar observation (cf. (4.76) and (4.79))- especially, the FI $J_{\mathbf{z}_i}$ for the orthogonal MAC case (cf. (4.75)). Furthermore, we have also adapted both considered constraints (C1) and (C2) to the equivalent model - and deduced (C1-s) and (C2-s).

Let us recall that an optimal sensor matrix $\mathbf{A}_i$ solves (P-II), where (P-II) can be considered for both multiple access schemes. Adapted to our equivalent model with scalar obervation, it specializes to:

$$\begin{aligned}
\underset{\mathbf{a}_i,\ 1\leq i\leq L}{\text{maximize}} \quad & J_{\mathbf{z}} \\
\text{subject to} \quad & J_{\mathbf{z}} \text{ satisfies (4.76) or (4.79)}, \\
& \mathbf{a}_i \in \mathbb{R}^{q_i} \quad \text{for } 1 \leq i \leq L, \\
& \text{(C1-s) or (C2-s) is satisfied, i.e.,} \\
& \text{(C1-s)}: \ \|\mathbf{a}_i\|^2 \left((g_i\theta)^2 + \sigma_{n_i}^2\right) \leq P_{0,i} \ \text{ for } 1 \leq i \leq L, \text{ or} \\
& \text{(C2-s)}: \ \|\mathbf{a}_i\|^2 \sigma_{n_i}^2 \leq P'_{0,i} \ \text{ for } 1 \leq i \leq L,
\end{aligned} \tag{P-II-s}$$

where g_i and $\sigma_{n_i}^2$ are given in (4.63), (4.64) or (4.65), depending on the given original observation model parameters $\mathbf{g}_i$ and $\mathbf{C}_{n_i}$.

In particular, we consider problem (P-III) for determining the optimal LO_i, for the case of an orthogonal MAC (cf. Section 4.2). For a scalar parameter and especially for the equivalent model with scalar observation, (P-III) specializes with (4.75) to

$$\begin{aligned}
\underset{\mathbf{a}_i}{\text{maximize}} \quad & J_{\mathbf{z}_i} = g_i^2 \mathbf{a}_i^T \mathbf{H}_i^T \left(\sigma_{h_i}^2 \mathbf{I} + \sigma_{n_i}^2 \mathbf{H}_i \mathbf{a}_i \mathbf{a}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{H}_i \mathbf{a}_i \\
\text{subject to} \quad & \mathbf{a}_i \in \mathbb{R}^{q_i}, \\
& \text{(C1-s)}_i \text{ or } \text{(C2-s)}_i \text{ is satisfied, i.e.,} \\
& \text{(C1-s)}_i: \ \|\mathbf{a}_i\|^2 \left((g_i\theta)^2 + \sigma_{n_i}^2\right) \leq P_{0,i} \text{ or} \\
& \text{(C2-s)}_i: \ \|\mathbf{a}_i\|^2 \sigma_{n_i}^2 \leq P'_{0,i}.
\end{aligned} \tag{P-III-s}$$

The notation $(\text{C1-s})_i$ and $(\text{C2-s})_i$ mean the ith constraint of (C1-s) and (C2-s). In what follows, we first solve problem (P-III-s) - we consider the orthogonal MAC case.

4.4.3.1 Orthogonal MAC

Let us now consider the case of an orthogonal MAC. Remember, that we have already assumed for this case diagonal matrices $\mathbf{H}_i$ for all i w.l.o.g., i.e.,

$$(\mathbf{H}_i)_{k,l} = \begin{cases} h_{ik} & k = l \quad 1 \leq k, l \leq w \triangleq \min\{p_i, q_i\} \\ 0 & k \neq l, \end{cases} \tag{4.80}$$

where h_{ik} for $1 \leq k \leq w$ are the singular values of $\mathbf{H}_i$, in particular (cf. Theorem 4.3.13). In that case, we consider especially the optimization problem (P-III-s), for determining the optimal local sensor vector $\mathbf{a}_i$ for LO'_i. Thus, we consider the FI $J_{\mathbf{z}_i}$ from (4.75) and take note that $g_i \in \mathbb{R}$ and $\sigma_{n_i}^2 \geq 0$.

Let us first treat the trivial third case, when $g_i = 0$ and $\sigma_{n_i}^2 = 0$ (cf. (4.65)). Then, it is obvious that the FI $J_{\mathbf{z}_i}$ from (4.75) yields $J_{\mathbf{z}_i} = 0$ for all $\mathbf{a}_i \in \mathbb{R}^{q_i}$, since we assumed $\sigma_{h_i}^2 > 0$. Hence, for this special case, there exist no optimal LO'_i for (P-III-s) and (P-II-s) and thus no optimal LO_i for (P-I). For further proceed we distinguish the two remaining cases - $\sigma_{n_i}^2 > 0$ or $\sigma_{n_i}^2 = 0$. In what follows we assume $g_i > 0$, which is guaranteed for both cases (cf. (4.63) and (4.65)).

1. Case - $\sigma_{n_i}^2 \neq 0$: The FI $J_{\mathbf{z}_i}$ from (4.75) then yields

$$\begin{aligned} J_{\mathbf{z}_i} &= g_i^2 \mathbf{a}_i^T \mathbf{H}_i^T \left(\sigma_{h_i}^2 \mathbf{I} + \sigma_{n_i}^2 \mathbf{H}_i \mathbf{a}_i \mathbf{a}_i^T \mathbf{H}_i^T\right)^{-1} \mathbf{H}_i \mathbf{a}_i \\ &\overset{(a)}{=} g_i^2 \widetilde{\mathbf{a}}_i^T \left(\sigma_{h_i}^2 \mathbf{I} + \sigma_{n_i}^2 \widetilde{\mathbf{a}}_i \widetilde{\mathbf{a}}_i^T\right)^{-1} \widetilde{\mathbf{a}}_i \\ &= \frac{g_i^2}{\sigma_{n_i}^2} \widetilde{\mathbf{a}}_i^T \left(\frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} \mathbf{I} + \widetilde{\mathbf{a}}_i \widetilde{\mathbf{a}}_i^T\right)^{-1} \widetilde{\mathbf{a}}_i \quad \text{for } \sigma_{n_i}^2 > 0, \end{aligned} \tag{4.81}$$

where in step (a) we introduced the shorthand

$$\widetilde{\mathbf{a}}_i \triangleq \mathbf{H}_i \mathbf{a}_i. \tag{4.82}$$

We now introduce

$$\mathbf{P}_{\mathcal{A}_i} \triangleq \frac{1}{\|\widetilde{\mathbf{a}}_i\|^2} \widetilde{\mathbf{a}}_i \widetilde{\mathbf{a}}_i^T. \tag{4.83}$$

Note that $\mathbf{P}_{\mathcal{A}_i}$ is the projection matrix[1] associated to the linear subspace $\mathcal{A}_i = \{c\widetilde{\mathbf{a}}_i | c \in \mathbb{R}\}$. Furthermore, $\mathbf{P}_{\mathcal{A}_i}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathcal{A}_i}$ is the projection matrix associated to the orthogonal complement linear subspace $\mathcal{A}_i^{\perp}$ of $\mathcal{A}_i$, i.e., $\mathcal{A}_i^{\perp} = \left\{\mathbf{x} \in \mathbb{R}^{p_i} | \widetilde{\mathbf{a}}_i^T \mathbf{x} = 0\right\}$ [5]. With the identity

[1] A projection matrix $\mathbf{P}$ is symmetric ($\mathbf{P} = \mathbf{P}^T$) and indempotent ($\mathbf{P}^2 = \mathbf{P}$).

$\mathbf{I} = \mathbf{P}_{\mathcal{A}_i} + \mathbf{P}_{\mathcal{A}_i}^{\perp}$, (4.81) can also be written in terms of $\mathbf{P}_{\mathcal{A}_i}$ and $\mathbf{P}_{\mathcal{A}_i}^{\perp}$ as

$$J_{\mathbf{z}_i} = \frac{g_i^2}{\sigma_{n_i}^2}\widetilde{\mathbf{a}}_i^T \left(\left(\|\widetilde{\mathbf{a}}_i\|^2 + \frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} \right) \mathbf{P}_{\mathcal{A}_i} + \frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} \mathbf{P}_{\mathcal{A}_i}^{\perp} \right)^{-1} \widetilde{\mathbf{a}}_i \tag{4.84}$$

for $\sigma_{n_i}^2 > 0$. Invoking [11], we have the identity

$$\left(c_1 \mathbf{P}_{\mathcal{A}_i} + c_2 \mathbf{P}_{\mathcal{A}_i}^{\perp} \right)^{-1} = \frac{1}{c_1}\mathbf{P}_{\mathcal{A}_i} + \frac{1}{c_2}\mathbf{P}_{\mathcal{A}_i}^{\perp}, \tag{4.85}$$

for any $c_1 \in \mathbb{R}\backslash\{0\}$ and $c_2 \in \mathbb{R}\backslash\{0\}$, since

$$\begin{aligned}
&\left(c_1 \mathbf{P}_{\mathcal{A}_i} + c_2 \mathbf{P}_{\mathcal{A}_i}^{\perp} \right)\left(\frac{1}{c_1}\mathbf{P}_{\mathcal{A}_i} + \frac{1}{c_2}\mathbf{P}_{\mathcal{A}_i}^{\perp} \right) = \\
&= \mathbf{P}_{\mathcal{A}_i}\mathbf{P}_{\mathcal{A}_i} + \frac{c_2}{c_1}\mathbf{P}_{\mathcal{A}_i}^{\perp}\mathbf{P}_{\mathcal{A}_i} + \frac{c_1}{c_2}\mathbf{P}_{\mathcal{A}_i}\mathbf{P}_{\mathcal{A}_i}^{\perp} + \mathbf{P}_{\mathcal{A}_i}^{\perp}\mathbf{P}_{\mathcal{A}_i}^{\perp} \overset{(a)}{=} \mathbf{P}_{\mathcal{A}_i} + \mathbf{P}_{\mathcal{A}_i}^{\perp} = \mathbf{I},
\end{aligned}$$

where in step (a) we used the fact that $\mathbf{P}_{\mathcal{A}_i}\mathbf{P}_{\mathcal{A}_i}^{\perp} = \mathbf{P}_{\mathcal{A}_i}(\mathbf{I} - \mathbf{P}_{\mathcal{A}_i}) = \mathbf{P}_{\mathcal{A}_i} - \mathbf{P}_{\mathcal{A}_i} = \mathbf{0} = \mathbf{P}_{\mathcal{A}_i}^{\perp} - \mathbf{P}_{\mathcal{A}_i}^{\perp} = \mathbf{P}_{\mathcal{A}_i}^{\perp}\left(\mathbf{I} - \mathbf{P}_{\mathcal{A}_i}^{\perp}\right) = \mathbf{P}_{\mathcal{A}_i}^{\perp}\mathbf{P}_{\mathcal{A}_i}$. Since we assumed that $\sigma_{h_i}^2 > 0$ and $\sigma_{n_i}^2 > 0$, we can use (4.85) for $c_2 = \sigma_{h_i}^2/\sigma_{n_i}^2 > 0$ and $c_1 = \left(\|\widetilde{\mathbf{a}}_i\|^2 + \sigma_{h_i}^2/\sigma_{n_i}^2\right) > 0$ and thus, (4.84) yields

$$\begin{aligned}
J_{\mathbf{z}_i} &= \frac{g_i^2}{\sigma_{n_i}^2}\widetilde{\mathbf{a}}_i^T \left(\frac{1}{\frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} + \|\widetilde{\mathbf{a}}_i\|^2}\mathbf{P}_{\mathcal{A}_i} + \frac{\sigma_{n_i}^2}{\sigma_{h_i}^2}\mathbf{P}_{\mathcal{A}_i}^{\perp} \right)\widetilde{\mathbf{a}}_i \\
&= \frac{g_i^2}{\sigma_{n_i}^2}\frac{1}{\frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} + \|\widetilde{\mathbf{a}}_i\|^2}\widetilde{\mathbf{a}}_i^T\mathbf{P}_{\mathcal{A}_i}\widetilde{\mathbf{a}}_i + \frac{g_i^2}{\sigma_{h_i}^2}\widetilde{\mathbf{a}}_i^T\mathbf{P}_{\mathcal{A}_i}^{\perp}\widetilde{\mathbf{a}}_i \\
&\overset{(a)}{=} \frac{g_i^2}{\sigma_{n_i}^2}\frac{\widetilde{\mathbf{a}}_i^T\widetilde{\mathbf{a}}_i\widetilde{\mathbf{a}}_i^T\widetilde{\mathbf{a}}_i}{\|\widetilde{\mathbf{a}}_i\|^2\left(\frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} + \|\widetilde{\mathbf{a}}_i\|^2\right)} \\
&= \frac{g_i^2}{\sigma_{n_i}^2}\frac{\|\widetilde{\mathbf{a}}_i\|^2}{\frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} + \|\widetilde{\mathbf{a}}_i\|^2} \quad \text{for } \sigma_{n_i}^2 \neq 0,
\end{aligned} \tag{4.86}$$

where in step (a) we inserted (4.83) and used the fact that

$$\mathbf{P}_{\mathcal{A}_i}^{\perp}\widetilde{\mathbf{a}}_i = (\mathbf{I} - \mathbf{P}_{\mathcal{A}_i})\widetilde{\mathbf{a}}_i \overset{(4.83)}{=} \widetilde{\mathbf{a}}_i - \widetilde{\mathbf{a}}_i\widetilde{\mathbf{a}}_i^T\widetilde{\mathbf{a}}_i / \|\widetilde{\mathbf{a}}_i\|^2 = \widetilde{\mathbf{a}}_i - \widetilde{\mathbf{a}}_i = \mathbf{0}.$$

Take note that $\widetilde{\mathbf{a}}_i^T\widetilde{\mathbf{a}}_i = \|\widetilde{\mathbf{a}}_i\|^2$.

2. Case - $\sigma_{n_i}^2 = 0$: In that case, the FI $J_{\mathbf{z}_i}$ from (4.75) yields

$$J_{\mathbf{z}_i} = g_i^2\mathbf{a}_i^T\mathbf{H}_i^T\left(\sigma_{h_i}^2\mathbf{I}\right)^{-1}\mathbf{H}_i\mathbf{a}_i = \frac{g_i^2}{\sigma_{h_i}^2}\mathbf{a}_i^T\mathbf{H}_i^T\mathbf{H}_i\mathbf{a}_i \quad \text{for } \sigma_{n_i}^2 = 0. \tag{4.87}$$

Using the shorthand from (4.82), (4.87) can also be written in the form

$$J_{\mathbf{z}_i} = \frac{g_i^2}{\sigma_{h_i}^2} \|\widetilde{\mathbf{a}}_i\|^2 \quad \text{for } \sigma_{n_i}^2 = 0. \tag{4.88}$$

Let us summarize both derivatives (4.86) and (4.88), as follows:

$$J_{\mathbf{z}_i} = \begin{cases} \frac{g_i^2}{\sigma_{n_i}^2} \frac{\|\widetilde{\mathbf{a}}_i\|^2}{\frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} + \|\widetilde{\mathbf{a}}_i\|^2} & \sigma_{n_i}^2 \neq 0 \\ \frac{g_i^2}{\sigma_{h_i}^2} \|\widetilde{\mathbf{a}}_i\|^2 & \sigma_{n_i}^2 = 0, \end{cases} \tag{4.89}$$

which is in turn equivalent to (4.75). So we can replace the FI $J_{\mathbf{z}_i}$ of problem (P-III-s) by (4.89), without loss. It can be verified easily that the FI $J_{\mathbf{z},i}$, given in (4.89), is a monotonic increasing function in $\|\widetilde{\mathbf{a}}_i\|^2$, since with $x \triangleq \|\widetilde{\mathbf{a}}_i\|^2$ and $J_{\mathbf{z}_i} = J_{\mathbf{z}_i}(x)$, the first derivation

$$J'_{\mathbf{z}_i}(x) \triangleq \frac{\partial}{\partial x} J_{\mathbf{z}_i}(x) = \begin{cases} \frac{g_i^2}{\sigma_{n_i}^2} \frac{1}{\left(\frac{\sigma_{h_i}^2}{\sigma_{n_i}^2} + x\right)^2} & \sigma_{n_i}^2 \neq 0 \\ \frac{g_i^2}{\sigma_{h_i}^2} & \sigma_{n_i}^2 = 0, \end{cases}$$

is strictliy positve for all x, i.e., $J'_{\mathbf{z}_i}(x) > 0$ for all x. Therefore, we can equivalently maximize $\|\widetilde{\mathbf{a}}_i\|^2 = \mathbf{a}_i^T \mathbf{H}_i^T \mathbf{H}_i \mathbf{a}_i$ instead of $J_{\mathbf{z}_i}$ from (4.89), while respecting the constraint $(\text{C1-s})_i$ or $(\text{C2-s})_i$. The optimization problem (P-III-s) can thus be reformulated equivalently as

$$\begin{aligned} &\underset{\mathbf{a}_i \in \mathbb{R}^{q_i}}{\text{maximize}} \quad \mathbf{a}_i^T \mathbf{H}_i^T \mathbf{H}_i \mathbf{a}_i \\ &\text{subject to} \quad \begin{cases} \left((g_i\theta)^2 + \sigma_{n_i}^2\right) \|\mathbf{a}_i\|^2 \leq P_{0,i} & (\text{C1-s})_i \text{ or} \\ \sigma_{n_i}^2 \|\mathbf{a}_i\|^2 \leq P'_{0,i}. & (\text{C2-s})_i \end{cases} \end{aligned} \tag{4.90}$$

Theorem 4.4.18 *Consider a real symmetric maxtrix* $\mathbf{A} \in \mathrm{Sym}(s)$. *Let* $\lambda_{\max}$ *denotes the largest eigenvalue and* $\mathbf{v}_{\max}$ *the corresponding eigenvector of* $\mathbf{A}$, *i.e.,* $\mathbf{A}\lambda_{\max} = \mathbf{A}\mathbf{v}_{\max}$. *Then,*

$$\mathbf{a}^T \mathbf{A} \mathbf{a} \leq \lambda_{\max} \|\mathbf{a}\|^2 \quad \text{for all } \mathbf{a} \in \mathbb{R}^s.$$

Equality holds iff $\mathbf{a} = c\mathbf{v}_{\max}$ *for any* $c \in \mathbb{R}$.

Proof. Cf. [12, 6.2, p.110]. ☐

Let us recall that we assumed a diagonal channel matrix $\mathbf{H}_i$ with diagonal entries h_{ik} for $1 \leq k \leq w$ - the singular values of $\mathbf{H}_i$ (cf. (4.80)). It is evident that also $\mathbf{H}_i^T \mathbf{H}_i$ of size $q_i \times q_i$ is diagonal, where the w largest diagonal entries

- the w largest eigenvalues of $\mathbf{H}_i^T\mathbf{H}_i$ - are the squared singular values of $\mathbf{H}_i$, i.e., $h_{i\,k}^2$ for $1 \leq k \leq w$. Let $h_{i\,max}^2$ denotes the maximum value of the set $\left\{h_{i\,k}^2,\ 1 \leq k \leq w\right\}$. Let us write the sensor vector $\mathbf{a}_i$ in the form

$$\mathbf{a}_i = c_i\mathbf{v}_i, \tag{4.91}$$

where $c_i = \|\mathbf{a}\| \in \mathbb{R}$ is the length and $\mathbf{v}_i = \mathbf{a}/c_i \in \mathbb{R}^{q_i}$ denotes the normalized vector of $\mathbf{a}$ (direction), i.e., $\|\mathbf{v}_i\|^2 = 1$. We can reformulate (4.90) equivalently as

$$\begin{aligned} &\underset{\substack{\mathbf{v}_i\in\mathbb{R}^{q_i},\ c_i\in\mathbb{R} \\ \|\mathbf{v}_i\|^2=1}}{\text{maximize}} && c_i^2\mathbf{v}_i^T\mathbf{H}_i^T\mathbf{H}_i\mathbf{v}_i \\ &\text{subject to} && \begin{cases} \left((g_i\theta)^2 + \sigma_{n_i}^2\right)c_i^2 \leq P_{0,i} & (\text{C1-s})_i \ \text{ or} \\ \sigma_{n_i}^2 c_i^2 \leq P'_{0,i}. & (\text{C2-s})_i \end{cases} \end{aligned} \tag{4.92}$$

Take note that only the objective function of (4.92) depends on the normalized vector $\mathbf{v}_i$ - the constraints $(\text{C1-s})_i$ and $(\text{C2-s})_i$ are not affected by $\mathbf{v}_i$. For solving (4.92) w.r.t. $\mathbf{v}_i$, we can maximize the objective function of (4.92), without considering $(\text{C1-s})_i$ or $(\text{C2-s})_i$. According to Theorem 4.4.18, the objective function of (4.92), with $\|\mathbf{v}_i\| = 1$, is bounded above by

$$c_i^2\mathbf{v}_i^T\mathbf{H}_i^T\mathbf{H}_i\mathbf{v}_i \leq c_i^2 h_{i\,\max}^2 \quad \text{for all } \mathbf{v}_i \in \mathbb{R}^{q_i},\ \text{where } \|\mathbf{v}_i\| = 1, \tag{4.93}$$

where equality (maximum) holds when $\mathbf{v}_i = \mathbf{e}_{i\max}$, since $\mathbf{H}_i^T\mathbf{H}_i$ is diagonal. The vector $\mathbf{e}_{i\max}$ denotes those unit vector (eigenvector of $\mathbf{H}_i^T\mathbf{H}_i$), which corresponds to the largest eigenvalue $h_{i\,\max}^2$ of $\mathbf{H}_i^T\mathbf{H}_i$. Hence, the optimal $\mathbf{v}_i$ for (4.92) can be obtained with

$$\mathbf{v}_i^* = \mathbf{e}_{i\max}. \tag{4.94}$$

Inserting $\mathbf{v}_i = \mathbf{e}_{i\max}$ into (4.92) yields

$$\begin{aligned} &\underset{c_i\in\mathbb{R}}{\text{maximize}} && c_i^2 h_{i\,\max}^2 \\ &\text{subject to} && \begin{cases} \left((g_i\theta)^2 + \sigma_{n_i}^2\right)c_i^2 \leq P_{0,i} & (\text{C1-s})_i \ \text{ or} \\ \sigma_{n_i}^2 c_i^2 \leq P'_{0,i}. & (\text{C2-s})_i \end{cases} \end{aligned} \tag{4.95}$$

It remains to determine the optimal constant c_i for (4.92) or, equivalently, for (4.95). Let us consider constraint $(\text{C1-s})_i$. Then, for the case $(g_i\theta)^2 + \sigma_{n_i}^2 = 0$, the constraint is always fulfilled and c_i could be chosen arbitrarily high in order to maximize the objective function in (4.95). Similar holds for constraint $(\text{C2-s})_i$ when $\sigma_{n_i}^2 = 0$. However, both special cases can occur only if $\sigma_{n_i}^2 = 0$. To remain mathematically correct in what follows, we exclude the case $\sigma_{n_i}^2 = 0$. For the analysis we therefore use the limit value $\sigma_{n_i}^2 \to 0$.

So, we have to maximize c_i^2 in (4.95), while respecting the constraint $(\text{C1-s})_i$ or $(\text{C2-s})_i$, respectively. Thus, it is obvious that optimal c_i for (4.95) results when equality prevails in $(\text{C1-s})_i$ or $(\text{C2-s})_i$, in order to obtain maximal value of the objective $c_i^2 h_{i\,\max}^2$, i.e., for

$$c_i^* = \begin{cases} \sqrt{\frac{P_{0,i}}{(g_i\theta)^2+\sigma_{n_i}^2}} & (\text{C1-s})_i \\ \sqrt{\frac{P'_{0,i}}{\sigma_{n_i}^2}} & (\text{C2-s})_i. \end{cases} \tag{4.96}$$

In turn, optimal $\mathbf{a}_i$ for problem (4.90) can be obtained by inserting (4.96) and (4.94) into (4.91) as

$$\mathbf{a}_i^* = c_i^* \mathbf{e}_{i\max}, \quad \text{where } c_i^* = \begin{cases} \sqrt{\frac{P_{0,i}}{(g_i\theta)^2+\sigma_{n_i}^2}} & (\text{C1-s})_i \\ \sqrt{\frac{P'_{0,i}}{\sigma_{n_i}^2}} & (\text{C2-s})_i. \end{cases} \tag{4.97}$$

Hence, we have determined optimal LO'_i for an orthogonal MAC - for the equivalent model with scalar observation. Finally, we give the optimal FI $J_{\mathbf{z}_i}$, that follows for an optimal LO'_i. To that end, we insert $\mathbf{a}_i^*$ from (4.97) into (4.82), i.e., $\widetilde{\mathbf{a}}_i^* = \mathbf{H}_i \mathbf{a}_i^* = c_i^* \mathbf{H}_i \mathbf{e}_{i\max} = c_i^* h_{i\max} \mathbf{e}_{i\max}$, which in turn is used in (4.75), i.e.,

$$\begin{aligned} J_{\mathbf{z}_i}^* &= g_i^2 c_i^{*2} h_{i\,\max}^2 \mathbf{e}_{i\,\max}^T \left(\sigma_{h_i}^2 \mathbf{I} + \sigma_{n_i}^2 c_i^{*2} h_{i\,\max}^2 \mathbf{e}_{i\max} \mathbf{e}_{i\,\max}^T\right)^{-1} \mathbf{e}_{i\max} \\ &= g_i^2 c_i^{*2} h_{i\,\max}^2 \frac{1}{\sigma_{h_i}^2 + \sigma_{n_i}^2 c_i^{*2} h_{i\,\max}^2} \\ &\overset{(4.96)}{=} \begin{cases} \frac{g_i^2 P_{0,i} h_{i\,\max}^2}{(g_i\theta)^2+\sigma_{n_i}^2} \frac{1}{\sigma_{h_i}^2+\sigma_{n_i}^2 \frac{P_{0,i}}{(g_i\theta)^2+\sigma_{n_i}^2} h_{i\,\max}^2} & (\text{C1-s})_i \\ \frac{g_i^2 P'_{0,i} h_{i\,\max}^2}{\sigma_{n_i}^2} \frac{1}{\sigma_{h_i}^2+\sigma_{n_i}^2 \frac{P'_{0,i}}{\sigma_{n_i}^2} h_{i\,\max}^2} & (\text{C2-s})_i \end{cases} \\ &= \begin{cases} \frac{g_i^2 h_{i\,\max}^2 P_{0,i}}{\sigma_{h_i}^2\left((g_i\theta)^2+\sigma_{n_i}^2\right)+\sigma_{n_i}^2 h_{i\,\max}^2 P_{0,i}} & (\text{C1-s})_i \\ \frac{g_i^2 h_{i\,\max}^2}{\sigma_{n_i}^2} \frac{P'_{0,i}}{\sigma_{h_i}^2+h_{i\,\max}^2 P'_{0,i}} & (\text{C2-s})_i \end{cases} \quad \text{for } \sigma_{n_i}^2 \neq 0. \end{aligned} \tag{4.98}$$

Note that (4.98) only hold for the case when $g_i \neq 0$ and $\sigma_{n_i}^2 > 0$. For the trivial case, when $g_i = 0$ and $\sigma_{n_i}^2 = 0$ (cf. (4.65)), we have already mentioned that obviously $J_{\mathbf{z}_i}^* = 0$.

Let us now analyze (4.98) the one remaining case (cf. (4.64)), i.e., when $g_i \neq 0$ and $\sigma_{n_i}^2 = 0$, in more detail. Considering first the second constraint $(\text{C2-s})_i$, in particular. Here, we can use for the analysis, as already mentioned, the limit value $\sigma_{n_i}^2 \to 0$. Then, the FI $J_{\mathbf{z}_i}^*$ goes to inifity, since

$$\lim_{\sigma_{n_i}^2 \to 0} J_{\mathbf{z}_i}^* = \infty. \quad (\text{C2-s})_i$$

This is at first sight strange, but quite explainable. The second constraint $(\text{C2-s})_i$ bounds only the variance of the transmit data $\mathbf{s}_i$ - which goes to zero, when $\sigma_{n_i}^2 \to 0$. Therefore, we could theoretically provide an arbitrarily large power for sensor i without violating the constraint $(\text{C2-s})_i$, but this results in an arbirtraly large FI $J_{\mathbf{z}_i}^*$. This is the reason why $(\text{C1-s})_i$ has more practical relevance. Now, we consider the first constraint $(\text{C1-s})_i$. The optimal local sensor rule $\mathbf{a}_i$, given in (4.97), depend on the parameter θ. In practice, the computation of the optimal c_i from (4.96) for $(\text{C1-s})_i$, has to be solved with an estimate (locally) $\hat{\theta}\,(y_i)$, since the true parameter θ is unknown. In this special case, so even when $\sigma_{n_i}^2 = 0$, then we can estimate the unknown parameter without estimation error, since $\hat{\theta}\,(y_i) = y_i/g_i$ and the estimation error $e = \hat{\theta}\,(y_i) - \theta = 0$. Hence, at the local sensor we know the exact value of the parameter θ. However, we are interested at the resulting FI $J_{\mathbf{z}_i}$ at the FC for an optimum LO_i'.

$$J_{\mathbf{z}_i}^* = \frac{h_{i\,\max}^2 P_{0,i}}{\theta \sigma_{h_i}^2}.$$

Both just been treated cases coincides with (4.98) when we use the limiting case $\sigma_{n_i}^2 \to 0$. Last but not least, still indicate the total optimal FI $J_{\mathbf{z}}$ from (4.76) for the orthogonal MAC:

$$J_{\mathbf{z}}^* = \sum_{i=1}^{L} J_{\mathbf{z}_i}^*, \quad \text{where} \begin{cases} J_{\mathbf{z}_i}^* = 0 & \text{if } g_i = 0 \text{ and } \sigma_{n_i}^2 = 0 \\ J_{\mathbf{z}_i}^* \text{ is given in (4.98)} & \text{else,} \end{cases} \tag{4.99}$$

which is the optimum in (P-II-s) and also the global optimum in (P-I) for the special case of a scalar parameter and considering the orthogonal MAC with L local sensors, since $\mathbf{C}_{l_i}'^* = \mathbf{0}$ and in turn $\mathbf{C}_{l_i}^* = \mathbf{0}$.

4.4.3.2 Coherent MAC

Let us now consider the case of an coherent MAC. In that case, we consider especially the optimization problem (P-II-s), in order to determine the optimal local sensor vectors $\mathbf{a}_i$ (i.e., LO_i') for $1 \leq i \leq L$. Thus, we consider the

FI $J_{\mathbf{z}}$ from (4.79), which can also be written as

$$\begin{aligned}
J_{\mathbf{z}} &= \mathbf{g}^T \widetilde{\mathbf{A}}^T \left(\sigma_h^2 \mathbf{I} + \sigma_n^2 \widetilde{\mathbf{A}} \widetilde{\mathbf{A}}^T\right)^{-1} \widetilde{\mathbf{A}} \mathbf{g} \\
&\overset{(a)}{=} \mathbf{g}^T \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T \left(\sigma_h^2 \mathbf{I} + \sigma_n^2 \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T\right)^{-1} \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \mathbf{g} \\
&\overset{(b)}{=} \mathbf{g}^T \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T \left(\sigma_h^2 \mathbf{U} \mathbf{U}^T + \sigma_n^2 \mathbf{U} \mathbf{\Sigma} \mathbf{\Sigma}^T \mathbf{U}^T\right)^{-1} \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \mathbf{g} \\
&= \mathbf{g}^T \mathbf{V} \mathbf{\Sigma}^T \mathbf{U}^T \mathbf{U} \left(\sigma_h^2 \mathbf{I} + \sigma_n^2 \mathbf{\Sigma} \mathbf{\Sigma}^T\right)^{-1} \mathbf{U}^T \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \mathbf{g} \\
&= \frac{1}{\sigma_n^2} \mathbf{g}^T \mathbf{V} \mathbf{\Sigma}^T \left(\frac{\sigma_h^2}{\sigma_n^2} \mathbf{I} + \mathbf{\Sigma} \mathbf{\Sigma}^T\right)^{-1} \mathbf{\Sigma} \mathbf{V}^T \mathbf{g} \\
&\overset{(c)}{=} \frac{1}{\sigma_n^2} \mathbf{g}^T \mathbf{V} \mathbf{D} \mathbf{V}^T \mathbf{g},
\end{aligned} \tag{4.100}$$

where in step (a) we performed the SVD $\widetilde{\mathbf{A}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, with the unitary matrices $\mathbf{U} \in \mathbb{R}^{p \times p}$, $\mathbf{V} \in \mathbb{R}^{L \times L}$ and the (possibly rectangular) diagonal matrix $\mathbf{\Sigma}$ of size $p \times L$, which contains the singular values σ_j for $1 \leq j \leq w \triangleq \min\{p, L\}$ of $\widetilde{\mathbf{A}}$ on the main diagonal, i.e., $(\mathbf{\Sigma})_{j,j} = \sigma_i$ for $1 \leq j \leq w$. In step (b), we used the fact that $\mathbf{V}\mathbf{V}^T = \mathbf{I}$ and $\mathbf{U}\mathbf{U}^T = \mathbf{I}$, since $\mathbf{V}$ and $\mathbf{U}$ are unitary. Let us assume that the singular values $|\sigma_j|$ are ordered decreasingly (in turn of magnitude), i.e., $|\sigma_1| \geq |\sigma_2| \geq \cdots \geq |\sigma_w|$. Take note that, for the coherent MAC case, we assumed $\sigma_n^2 > 0$. In the last step (c), we introduced the diagonal matrix

$$\mathbf{D} \triangleq \mathbf{\Sigma}^T \left(\frac{\sigma_h^2}{\sigma_n^2} \mathbf{I} + \mathbf{\Sigma} \mathbf{\Sigma}^T\right)^{-1} \mathbf{\Sigma}, \tag{4.101}$$

which is squared and diagonal of size $L \times L$. The L elements on the main diagonal are thus given by

$$d_i \triangleq (\mathbf{D})_{j,j} = \begin{cases} \frac{\sigma_j^2}{\frac{\sigma_h^2}{\sigma_n^2} + \sigma_j^2} & \text{if } 1 \leq j \leq w \\ 0 & \text{else,} \end{cases} \tag{4.102}$$

for $1 \leq j \leq L$. Note that the diagonal elements d_j are also ordered decreasingly, i.e., $d_1 \geq d_2 \geq ...d_L \geq 0$, that follows by adopting the order of the set $\{|\sigma_j|\}$ for $1 \leq j \leq w$. As mentioned above, we consider for the coherent MAC case especially, a total power/variance constraint (C1-t) or (C2-t), respectively. Furthermore, we will treat only the case for constraint (C2-t). Let us recall the shorthand from (4.77) and the notations from (4.78). With the assumption made that $\mathbf{H}_i$ for all i have all full column-rank, we can uniquely reclaim $\mathbf{a}_i$ with $\mathbf{a}_i = \mathbf{H}_i^\dagger \widetilde{\mathbf{a}}_i$. Thus, we can reformulate (C2-t) with $\mathbf{a}_i = \mathbf{H}_i^\dagger \widetilde{\mathbf{a}}_i = \mathbf{H}_i^\dagger \widetilde{\mathbf{A}} \mathbf{e}_i$ ($\mathbf{e}_i$ denotes the ith unit vector), in terms of $\widetilde{\mathbf{A}}$ as

$$\sum_{i=1}^{L} \|\mathbf{a}_i\|^2 = \sum_{i=1}^{L} \left\|\mathbf{H}_i^\dagger \widetilde{\mathbf{A}} \mathbf{e}_i\right\|^2 = \sum_{i=1}^{L} \mathbf{e}_i^T \widetilde{\mathbf{A}}^T \left(\mathbf{H}_i \mathbf{H}_i^T\right)^\dagger \widetilde{\mathbf{A}} \mathbf{e}_i \leq P_0'/\sigma_n^2, \tag{4.103}$$

where we used the fact $\left(\mathbf{H}_i^T\right)^\dagger \mathbf{H}_i^\dagger = \left(\mathbf{H}_i^\dagger\right)^T \mathbf{H}_i^\dagger = \left(\mathbf{H}_i\mathbf{H}_i^T\right)^\dagger$ [13]. In terms of SVD $\widetilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, (4.103) yields

$$\begin{aligned} t\left(\mathbf{U},\boldsymbol{\Sigma},\mathbf{V}\right) &\triangleq \sum_{i=1}^{L} \mathbf{e}_i^T\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}_i\mathbf{H}_i^T\right)^\dagger \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{e}_i \\ &= \sum_{i=1}^{L}\left(\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}_i\mathbf{H}_i^T\right)^\dagger \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\right)_{i,i} \le P_0'/\sigma_n^2, \end{aligned} \tag{4.104}$$

where we introduced the constraint function $t(\cdot)$. With the derivations (4.100) (exluding the constant[1]) and (4.104), the optimization problem (P-II-s) can thus be reformulated equivalently as

$$\begin{aligned} \underset{\mathbf{V},\boldsymbol{\Sigma},\mathbf{U}}{\text{maximize}} \quad & \mathbf{g}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{g} \\ \text{subject to} \quad & \sum_{i=1}^{L}\left(\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}_i\mathbf{H}_i^T\right)^\dagger \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\right)_{i,i} \le P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ & \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T \ge \mathbf{0}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \\ & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}. \end{aligned} \tag{4.105}$$

Solving (4.105) with respect to $\mathbf{U}$, $\mathbf{V}$, and $\boldsymbol{\Sigma}$, further gives the optimal $\widetilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, and in turn, the optimal $\mathbf{a}_i = \mathbf{H}_i^\dagger\widetilde{\mathbf{A}}\mathbf{e}_i$ for (P-II-s) (since $\mathbf{H}_i$ is assumed to be of full column-rank). In what follows, we solve (4.105) for two special cases, with assumptions on the individual channel matrices $\mathbf{H}_i$ for all i.

Orthogonal Channels: In that case, we assume that all individual channel matrices $\mathbf{H}_i$ for all i are orthogonal (unitary), i.e., $\mathbf{H}_i\mathbf{H}_i^T = \mathbf{H}_i^T\mathbf{H}_i = \mathbf{I}$ for $1 \le i \le L$ - implying $p = q_i$ for $1 \le i \le L$. Then, the constraint function $t(\cdot)$ from (4.104) yields

$$\begin{aligned} t\left(\mathbf{U},\boldsymbol{\Sigma}\right) &= \sum_{i=1}^{L}\left(\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}_i\mathbf{H}_i^T\right)^\dagger \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\right)_{i,i} \\ &= \sum_{i=1}^{L}\left(\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\right)_{i,i} = \operatorname{tr}\left\{\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\right\} = \operatorname{tr}\left\{\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\right\}, \end{aligned}$$

where we inserted $\mathbf{H}_i\mathbf{H}_i^T = \mathbf{I}$ and used the cyclic property of the trace operator [8], further, the facts that $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. In turn,

[1] It is common to omit the constant in the objective function, since it does not affect the optimal solution.

(4.105) yields

$$\begin{aligned} \underset{\mathbf{V},\mathbf{\Sigma}}{\text{maximize}} \quad & \mathbf{g}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{g} \\ \text{subject to} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}^T\mathbf{\Sigma}\right\} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ & \mathbf{\Sigma}\mathbf{\Sigma}^T \geq \mathbf{0}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}. \end{aligned} \tag{4.106}$$

In what follows, we will solve (4.106) sequentially, by determining first the optimum $\mathbf{V}$ and then the optimum $\mathbf{\Sigma}$. A necessary condition for $\mathbf{V}$ to be optimum in (4.106), can be obtained by fixing $\mathbf{\Sigma}$. For a fixed $\mathbf{\Sigma}'$, the optimum $\mathbf{V}$ has to solve the problem

$$\begin{aligned} \underset{\mathbf{V}}{\text{maximize}} \quad & \mathbf{g}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{g} \\ \text{subject to} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}'^T\mathbf{\Sigma}'\right\} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ & \mathbf{\Sigma}'\mathbf{\Sigma}'^T \geq \mathbf{0}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \end{aligned} \tag{4.107}$$

where $\mathbf{D}'$ is obtained from (4.101) by inserting $\mathbf{\Sigma}'$ for $\mathbf{\Sigma}$. Accordingly, we denote d_j' for $1 \leq j \leq L$ as the elements of $\mathbf{D}'$. Now consider the problem (4.107) we recognize that it no longer depends on the unitary $\mathbf{U}$ - we thus can choose an arbitrarily unitary $\mathbf{U}$, e.g., $\mathbf{U} = \mathbf{I}$. A $\mathbf{V}$ is optimum for (4.107) if and only if it is optimum for

$$\begin{aligned} \underset{\mathbf{V}}{\text{maximize}} \quad & \mathbf{g}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{g} \\ \text{subject to} \quad & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \end{aligned} \tag{4.108}$$

as can be verified easily. Let us denote the orthonormal column vectors of the unitary matrix $\mathbf{V}$ by $\mathbf{v}_j$ for $1 \leq j \leq L$. According to Theorem 4.4.18, the objective function of (4.108) is bounded above ($\mathbf{g}$ is given) by

$$\mathbf{g}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{g} \leq d_1' \left\|\mathbf{g}\right\|^2, \tag{4.109}$$

since we assumed that the diagonal entries d_j' for $1 \leq j \leq L$ in $\mathbf{D}'$ are ordered decreasingly. Equality in (4.109) (maximum) holds when $\mathbf{v}_1 = c\mathbf{g}$ for any constant $c \in \mathbb{R}$. Since $\mathbf{v}_1$ must has unit norm, i.e., $\|\mathbf{v}_1\| = 1$, the optimal $\mathbf{v}_1$ is given by

$$\mathbf{v}_1^* = \frac{\mathbf{g}}{\|\mathbf{g}\|}. \tag{4.110}$$

The remaining $(L-1)$ column vectors of optimum $\mathbf{V}$, i.e., $\mathbf{v}_j$ for $2 \leq j \leq L$ can be choosen arbitrarily, such that the set $\{\mathbf{v}_j,\ 1 \leq j \leq L\}$ form an orthonormal basis, i.e., $\mathbf{V}$ is unitary.

After having determined the optimal $\mathbf{V}$ for (4.106), we will now characterize the optimum $\boldsymbol{\Sigma}$ for (4.106). To that end, we first insert the optimal $\mathbf{V}$ in (4.106), where the objective function then yields

$$\mathbf{g}^T\mathbf{V}^*\mathbf{D}\mathbf{V}^{*T}\mathbf{g} = \sum_{j=1}^{L}\mathbf{g}^T d_j \mathbf{v}_j^*\mathbf{v}_j^{*T}\mathbf{g} = \left(\mathbf{g}^T\mathbf{v}_1^*\right)^2 d_1 + \sum_{j=2}^{L}\left(\mathbf{g}^T\mathbf{v}_j^*\right)^2 d_j \overset{(a)}{=} \|\mathbf{g}\|^2 d_1, \tag{4.111}$$

where in step (a) we inserted (4.110) and used the fact that $\mathbf{g}^T\mathbf{v}_j^* = 0$ for $2 \leq j \leq L$, since $\mathbf{g} = \|\mathbf{g}\|\,\mathbf{v}_1^*$ (cf. (4.110)) and $\mathbf{v}_1$ is obviously orthogonal to each $\mathbf{v}_j$ for $2 \leq j \leq L$ by definition ($\mathbf{V}$ is unitary). Hence, with d_1 from (4.102), (4.106) by inserting optimum $\mathbf{V}$ yields

$$\begin{aligned} &\underset{\boldsymbol{\Sigma}}{\text{maximize}} && \|\mathbf{g}\|^2 \frac{\sigma_1^2}{\frac{\sigma_h^2}{\sigma_n^2} + \sigma_1^2} \\ &\text{subject to} && \operatorname{tr}\left\{\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\right\} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ &&& \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T \geq \mathbf{0}. \end{aligned} \tag{4.112}$$

Introducing the vector notation $\mathbf{s} = \left(s_1, s_2, \ldots, s_w\right)^T \triangleq \left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_w^2\right)^T \in \mathbb{R}^{+w}$, i.e., it has to be: $\mathbf{s} \succeq \mathbf{0}$[1], the constraint (C2-t) can thus be written by $\operatorname{tr}\left\{\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\right\} = \sum_{i=1}^{w} s_i = \mathbf{s}^T\mathbf{1} \leq P_0'/\sigma_n^2$, where $\mathbf{1}$ denotes a vector of ones. Thus, (4.112) can equivalently be written in terms of s_j for $1 \leq j \leq w$ as

$$\begin{aligned} &\underset{\mathbf{s}}{\text{maximize}} && f\left(s_1\right) \triangleq \frac{s_1}{\frac{\sigma_h^2}{\sigma_n^2} + s_1} \\ &\text{subject to} && \mathbf{s}^T\mathbf{1} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ &&& \mathbf{s} \succeq \mathbf{0}, \end{aligned} \tag{4.113}$$

where we introduced the function $f\left(s_1\right)$. It is obvious that the function $f\left(s_1\right)$ is a monotonic function in s_1, since the first derivation

$$\frac{\partial}{\partial s_1} f_{s_1}\left(s_1\right) = \frac{1}{\left(\frac{\sigma_h^2}{\sigma_n^2} + s_1\right)^2} > 0 \quad \text{for all } s_1 \in \mathbb{R}.$$

Thus, we can equivalentely maximize s_1, while respecting the constraint (C2-t). The optimization problem (4.113) can thus be reformulated equivalently as

$$\begin{aligned} &\underset{\mathbf{s}}{\text{maximize}} && s_1 \\ &\text{subject to} && \mathbf{s}^T\mathbf{1} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ &&& \mathbf{s} \succeq \mathbf{0}, \end{aligned} \tag{4.114}$$

[1]For two vectors $\mathbf{a}$ and $\mathbf{b}$, the relation $\mathbf{a} \succeq \mathbf{b}$ means elementwise inequality, i.e., $a_i \geq b_i$ for all i, where a_i and b_i denote the ith elements of $\mathbf{a}$ and $\mathbf{b}$, respectively.

which can be easily solved by

$$s_1^* = P_0'/\sigma_n^2 \quad \text{and } s_j^* = 0 \text{ for } 2 \leq j \leq w, \tag{4.115}$$

and therefore, with $\sigma_j = \sqrt{s_j}$, we finally obtain the optimal $\mathbf{\Sigma}$ for (4.106) as

$$(\mathbf{\Sigma}^*)_{j,j} = \sigma_j^* = \begin{cases} \sqrt{\frac{P_0'}{\sigma_n^2}} & \text{for } j = 1 \\ 0 & \text{else.} \end{cases} \tag{4.116}$$

After having determined the optimal $\mathbf{V}$ and $\mathbf{\Sigma}$ for (4.106), assuming $\mathbf{U} = \mathbf{I}$ (but could be an arbitrary unitary matrix), we are now able to compute optimal $\widetilde{\mathbf{A}}$ as

$$\widetilde{\mathbf{A}}^* = \mathbf{\Sigma}^* \mathbf{V}^{*T} = \sum_{j=1}^{w} \sigma_j^* \mathbf{e}_j \mathbf{v}_j^{*T} \overset{(4.116)}{=} \sqrt{\frac{P_0'}{\sigma_n^2}} \mathbf{e}_1 \mathbf{v}_1^{*T} \overset{(4.110)}{=} \sqrt{\frac{P_0'}{\sigma_n^2 \left\|\mathbf{g}\right\|^2}} \mathbf{e}_1 \mathbf{g}^T, \tag{4.117}$$

where again $\mathbf{e}_j$ denotes the jth unit vector. Note that the first unit vector $\mathbf{e}_1$ in (4.117) follows from the assumption that $\mathbf{U} = \mathbf{I}$. It remains to determine the optimal local sensor vectors $\mathbf{a}_i$ for $1 \leq i \leq L$, which can be obtained with $\mathbf{a}_i = \mathbf{H}_i^{-1} \widetilde{\mathbf{A}} \mathbf{e}_i$ as

$$\mathbf{a}_i^* = \mathbf{H}_i^{-1} \widetilde{\mathbf{A}}^* \mathbf{e}_i \overset{(4.117)}{=} \sqrt{\frac{P_0'}{\sigma_n^2 \left\|\mathbf{g}\right\|^2}} \mathbf{H}_i^T \mathbf{e}_1 \mathbf{g}^T \mathbf{e}_i = \sqrt{\frac{P_0'}{\sigma_n^2 \left\|\mathbf{g}\right\|^2}} g_i \mathbf{h}_{i\,1}^r \tag{4.118}$$

for $1 \leq i \leq L$, since $\mathbf{H}_i$ is assumed to be unitary, i.e., $\mathbf{H}_i^{-1} = \mathbf{H}_i^T$. The vector $\mathbf{h}_{i\,1}^r$ denotes the first row-vector of $\mathbf{H}_i$ and $g_i = \mathbf{g}^T \mathbf{e}_i$ denotes the ith element of $\mathbf{g}$ (cf. (4.78)). Note that $\mathbf{h}_{i\,1}^r$ has unit norm, i.e., $\left\|\mathbf{h}_{i\,1}^r\right\| = 1$, since $\mathbf{H}_i$ is unitary.

Hence, we have determined optimal LO_i' for a coherent MAC, where $\mathbf{H}_i$ for all i are assumed to be orthogonal - for the equivalent model with scalar observation. Finally, we give the optimal FI $J_{\mathbf{z}}$, that follows for an optimal LO_i'. To that end, we insert $\mathbf{V}^*$ and $\mathbf{D}^*$ - where $\mathbf{D}^*$, with diagonal entries d_i^*, is obtained from (4.101) by inserting $\mathbf{\Sigma}^*$ for $\mathbf{\Sigma}$ - into (4.100), i.e.,

$$J_{\mathbf{z}}^* = \frac{1}{\sigma_n^2} \mathbf{g}^T \mathbf{V}^* \mathbf{D}^* \mathbf{V}^{*T} \mathbf{g} \overset{(a)}{=} \frac{\left\|\mathbf{g}\right\|^2}{\sigma_n^2} d_1^* \overset{(b)}{=} \frac{\left\|\mathbf{g}\right\|^2}{\sigma_n^2} \frac{\sigma_1^{*2}}{\frac{\sigma_h^2}{\sigma_n^2} + \sigma_1^{*2}} \overset{(4.116)}{=} \frac{\left\|\mathbf{g}\right\|^2}{\sigma_n^2} \frac{P_0'}{\sigma_h^2 + P_0'}, \tag{4.119}$$

where in step (a) we used the derivation in (4.111) for $\mathbf{D} = \mathbf{D}^*$ (i.e., $d_1 = d_1^*$). In step (b) we insert (4.102) for $\sigma_1^2 = \sigma_1^{*2}$ into d_1^*.

Identical Channels: In that case, we assume that all individual channel matrices $\mathbf{H}_i$ for all i are identical and especially invertable, i.e., $\mathbf{H} \triangleq \mathbf{H}_j = \mathbf{H}_i$ for all $1 \leq i, j \leq L$ and it exists $\mathbf{H}^{-1}$. Then, the constraint function $t(\cdot)$ from (4.104) yields

$$\begin{aligned} t(\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}) &= \sum_{i=1}^{L} \left(\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T \left(\mathbf{H}_i\mathbf{H}_i^T \right)^{-1} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \right)_{i,i} \\ &= \sum_{i=1}^{L} \left(\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T \left(\mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \right)_{i,i} \\ &= \operatorname{tr}\left\{ \mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T \left(\mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \right\} \\ &= \operatorname{tr}\left\{ \left(\mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T \right\} = t(\mathbf{U}, \mathbf{\Sigma}), \end{aligned} \tag{4.120}$$

where in the last step, we used the cyclic property of the trace operator [8] and the fact that $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. Thus, (4.105) yields

$$\begin{aligned} \underset{\mathbf{U},\mathbf{\Sigma},\mathbf{V}}{\text{maximize}} \quad & \mathbf{g}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{g} \\ \text{subject to} \quad & \operatorname{tr}\left\{ \left(\mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T \right\} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ & \mathbf{\Sigma}\mathbf{\Sigma}^T \geq \mathbf{0}, \\ & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}. \end{aligned} \tag{4.121}$$

Note that the constraint (C2-t) in (4.121), now depends also on the unitary matrix $\mathbf{U}$, the left singular vectors of $\widetilde{\mathbf{A}}$.

In what follows, we will solve (4.121) sequentially, by determining first the optimum $\mathbf{V}$ and then the optimum $\mathbf{U}$. A necessary condition for $\mathbf{V}$ to be optimum in (4.121) can be obtained by fixing $\mathbf{U}$ and $\mathbf{\Sigma}$. For a fixed $\mathbf{U}'$ and $\mathbf{\Sigma}'$, the optimum $\mathbf{V}$ has to solve the problem

$$\begin{aligned} \underset{\mathbf{V}}{\text{maximize}} \quad & \mathbf{g}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{g} \\ \text{subject to} \quad & \operatorname{tr}\left\{ \left(\mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{U}'\mathbf{\Sigma}'\mathbf{\Sigma}'^T\mathbf{U}'^T \right\} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ & \mathbf{\Sigma}'\mathbf{\Sigma}'^T \geq \mathbf{0}, \\ & \mathbf{U}'\mathbf{U}'^T = \mathbf{U}'^T\mathbf{U}' = \mathbf{I}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}. \end{aligned} \tag{4.122}$$

where $\mathbf{D}'$ is obtained from (4.101) by inserting $\mathbf{\Sigma}'$ for $\mathbf{\Sigma}$. A $\mathbf{V}$ is optimum for (4.122) if and only if it is optimum for

$$\begin{aligned} \underset{\mathbf{V}}{\text{maximize}} \quad & \mathbf{g}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{g} \\ \text{subject to} \quad & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \end{aligned} \tag{4.123}$$

as can be verified easily. Problem (4.123) is exactly the same as (4.108), which we have already solved. Denoting again, the orthonormal column vectors of the unitary matrix $\mathbf{V}$ by $\mathbf{v}_j$ for $1 \leq j \leq L$. Then, since d'_j for all j (the diagonal elements in $\mathbf{D}'$) are assumed to be ordered decreasingly, we obtain the optimal $\mathbf{V}$ for (4.123) and thus also for (4.121) by choosing the first column vector $\mathbf{v}_1$ as in (4.110). The remaining $(L-1)$ column vectors of the optimal $\mathbf{V}$, i.e., $\mathbf{v}_j$ for $2 \leq j \leq L$ can again be choosen arbitrarily such that the set $\{\mathbf{v}_j,\ 1 \leq j \leq L\}$ forms an orthonormal basis, i.e., $\mathbf{V}$ is unitary.

After having determined the optimal $\mathbf{V}$ for (4.121), we will now characterize the optimum $\mathbf{U}$ for (4.121). To that end, we insert first the optimal $\mathbf{V}$ into (4.121), yielding:

$$\begin{aligned} \underset{\mathbf{U},\boldsymbol{\Sigma}}{\text{maximize}} \quad & \|\mathbf{g}\|^2 d_1 \\ \text{subject to} \quad & \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\mathbf{U}^T\right\} \leq P'_0/\sigma_n^2, \quad \text{(C2-t)} \\ & \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T \geq \mathbf{0}, \\ & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \end{aligned} \tag{4.124}$$

where we used exactly the derivation in (4.111). If $\mathbf{U}^*$ is optimal for (4.124), then it is also optimal for (4.121), as can be verified easily.

Let us consider an optimal pair $(\mathbf{U}', \boldsymbol{\Sigma}')$ solving (4.121) or, equivalently, (4.124). We will now show that necessarily $\mathbf{U}'$ has to be a minimizer of the constraint function $t(\mathbf{U}, \boldsymbol{\Sigma})$, given in (4.120), for the specific choice $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$, i.e.,

$$\begin{aligned} \mathbf{U}' = \arg \underset{\mathbf{U}}{\text{minimize}} \quad & t\left(\mathbf{U}, \boldsymbol{\Sigma}'\right) = \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{U}\boldsymbol{\Sigma}'\boldsymbol{\Sigma}'^T\mathbf{U}^T\right\} \\ \text{subject to} \quad & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \end{aligned} \tag{4.125}$$

Indeed assume that there is another unitary matrix $\mathbf{U}''$ such that $t(\mathbf{U}'', \boldsymbol{\Sigma}') < t(\mathbf{U}', \boldsymbol{\Sigma}')$. It follows that also $\mathbf{U}''$, $\boldsymbol{\Sigma}'$ is a feasible pair, since $\mathbf{U}''$ is unitary and

$$t\left(\mathbf{U}'', \boldsymbol{\Sigma}'\right) < t\left(\mathbf{U}', \boldsymbol{\Sigma}'\right) \leq P'_0/\sigma_n^2. \quad \text{(C2-t)}. \tag{4.126}$$

We can now construct another $\boldsymbol{\Sigma}$, i.e., $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}''$ by $\boldsymbol{\Sigma}'' \triangleq \sqrt{c}\boldsymbol{\Sigma}'$, where $c > 1$. Since, as can be verified easily, $t(\mathbf{U}, \sqrt{c}\boldsymbol{\Sigma}) = c \cdot t(\mathbf{U}, \boldsymbol{\Sigma})$, we can choose c small enough such that

$$t\left(\mathbf{U}'', \boldsymbol{\Sigma}''\right) = c \cdot t\left(\mathbf{U}'', \boldsymbol{\Sigma}'\right) \leq P'_0/\sigma_n^2 \quad \text{(C2-t)},$$

due to (4.126) implying that also $(\mathbf{U}'', \boldsymbol{\Sigma}'')$ is feasible. However, a simple computation shows that for the feasible pair $(\mathbf{U}'', \boldsymbol{\Sigma}'')$ the objective in (4.121) is strictly larger than for $(\mathbf{U}', \boldsymbol{\Sigma}')$. A contradiction to the assumption that $(\mathbf{U}', \boldsymbol{\Sigma}')$ is optimal.

Theorem 4.4.19 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *are two real symmetric* $(s \times s)$*-matrices and denoting* $\lambda_i(\mathbf{A})$ *and* $\lambda_i(\mathbf{B})$ *as the ith eigenvalue of* $\mathbf{A}$ *and* $\mathbf{B}$*, respectively. Assuming the eigenvalues* $\lambda_i(\mathbf{A})$ *and* $\lambda_i(\mathbf{B})$ *are arranged in decreasing order, i.e.,* $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq ... \geq \lambda_n(\mathbf{A}) \geq 0$ *and* $\lambda_1(\mathbf{B}) \geq \lambda_2(\mathbf{B}) \geq ... \geq \lambda_n(\mathbf{B}) \geq 0$*. Then*

$$\sum_{i=1}^{n} \lambda_i(\mathbf{A})\, \lambda_{s-i+1}(\mathbf{B}) \leq \operatorname{tr}\{\mathbf{AB}\} \leq \sum_{i=1}^{s} \lambda_i(\mathbf{A})\, \lambda_i(\mathbf{B}). \tag{4.127}$$

Proof. Cf. [14, Theorem II-1]. □

So, we can determine the optimal $\mathbf{U}$ for (4.121) by solving (4.125), where the optimal $\mathbf{U}$ is given by $\mathbf{U}^* = \mathbf{U}'$. An application of Theorem 4.4.19 reveals that the optimal $\mathbf{U}$ is given by the eigenvectors of $\left(\mathbf{HH}^T\right)^{-1}$ or $\mathbf{HH}^T$, in the order of increasing eigenvalues of $\left(\mathbf{HH}^T\right)^{-1}$ or decreasing eigenvalues of $\mathbf{HH}^T$, respectively. Thus, with eigenvalue decomposition (EVD) $\mathbf{HH}^T = \mathbf{U}_h \boldsymbol{\Lambda}_h \mathbf{U}_h^T$, where the unitary $\mathbf{U}_h$ contains the eigenvectors, and the diagonal $\boldsymbol{\Lambda}_h$ contains the positive eigenvalues of $\mathbf{HH}^T$, denoted by λ_{hj} for $1 \leq j \leq p$, in decreasing order, i.e., $\lambda_{h1} \geq \lambda_{h2} \geq, \ldots, \geq \lambda_{hp} > 0$, we obtain optimal $\mathbf{U}$ by

$$\mathbf{U}^* = \mathbf{U}_h, \quad \mathbf{u}_j^* = \mathbf{u}_{hj} \quad \text{for } 1 \leq j \leq p, \tag{4.128}$$

where the vectors $\mathbf{u}_j$ and $\mathbf{u}_{hj}$ denote the jth column vectors of $\mathbf{U}$ and $\mathbf{U}_h$, respectively.

So far, we have determined the optimal $\mathbf{U}$ and $\mathbf{V}$ for (4.121). It remains to determine the optimal $\boldsymbol{\Sigma}$ for (4.121). To that end, we insert $\mathbf{U}^*$ into (4.124) an together with d_1 from (4.102) in turn leads to

$$\begin{aligned} \underset{\boldsymbol{\Sigma}}{\text{maximize}} \quad & \|\mathbf{g}\|^2 \frac{\sigma_1^2}{\frac{\sigma_h^2}{\sigma_n^2} + \sigma_1^2} \\ \text{subject to} \quad & \operatorname{tr}\left\{\boldsymbol{\Lambda}_h^{-1} \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\right\} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ & \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T \geq \mathbf{0}. \end{aligned} \tag{4.129}$$

We now accept the notation $\mathbf{s} = \left(s_1, s_2, \ldots, s_L\right)^T \triangleq \left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2\right)^T \in \mathbb{R}^{+w}$ as for the last case, where we considered orthogonal channels, in turn: $\mathbf{s} \succeq \mathbf{0}$. Comparing (4.129) with (4.112), we recognize that only the constraint (C2-t) differs - the objective in both problems are equivalent. Writting (4.129) in terms of s_j for $1 \leq j \leq w$, we thus can use (4.114), whereby the constraint is still to be adapted accordingly. In the current case, the constraint (C2-t) in terms of $\mathbf{s}$ can be written as $\operatorname{tr}\left\{\boldsymbol{\Lambda}_h^{-1}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\right\} = \sum_{j=1}^{w} s_j/\lambda_{hj} \leq P_0'/\sigma_n^2$ and thus, (4.129) in terms of s_j for $1 \leq j \leq w$ finally yields

$$\begin{aligned} \underset{\mathbf{s}}{\text{maximize}} \quad & s_1 \\ \text{subject to} \quad & \sum_{i=1}^{w} \frac{s_j}{\lambda_{hj}} \leq P_0'/\sigma_n^2, \quad \text{(C2-t)} \\ & s_j \geq 0 \quad \text{for } 1 \leq j \leq w, \end{aligned} \tag{4.130}$$

which can be easily solved by

$$s_1^* = \frac{P_0' \lambda_{h1}}{\sigma_n^2} \quad \text{and } s_j^* = 0 \text{ for } 2 \leq j \leq w, \tag{4.131}$$

and therefore, with $\sigma_j = \sqrt{s_j}$, we finally obtain the optimal $\mathbf{\Sigma}$ for (4.121) as

$$(\mathbf{\Sigma}^*)_{j,j} = \sigma_j^* = \begin{cases} \sqrt{\frac{P_0'\lambda_{h1}}{\sigma_n^2}} & \text{for } j = 1 \\ 0 & \text{else.} \end{cases} \tag{4.132}$$

After having determined the optimal $\mathbf{V}$, $\mathbf{U}$ and $\mathbf{\Sigma}$ for (4.121), we are now able to compute optimal $\widetilde{\mathbf{A}}$ as

$$\widetilde{\mathbf{A}}^* = \mathbf{U}^* \mathbf{\Sigma}^* \mathbf{V}^{*T} = \sum_{j=1}^{w} \sigma_j^* \mathbf{u}_j \mathbf{v}_j^{*T} \overset{(4.132)}{=} \sqrt{\frac{P_0'\lambda_{h1}}{\sigma_n^2}} \mathbf{u}_1^* \mathbf{v}_1^{*T} \overset{(a)}{=} \sqrt{\frac{P_0'\lambda_{h1}}{\sigma_n^2 \|\mathbf{g}\|^2}} \mathbf{u}_{h1} \mathbf{g}^T, \tag{4.133}$$

where in step (a) we inserted (4.110) and (4.128). It remains to determine the optimal local sensor vectors $\mathbf{a}_i$ for $1 \leq i \leq L$, which can be obtained with $\mathbf{a}_i = \mathbf{H}_i^{-1} \widetilde{\mathbf{A}} \mathbf{e}_i = \mathbf{H}^{-1} \widetilde{\mathbf{A}} \mathbf{e}_i$ as

$$\mathbf{a}_i^* = \mathbf{H}^{-1} \widetilde{\mathbf{A}}^* \mathbf{e}_i \overset{(4.133)}{=} \sqrt{\frac{P_0'\lambda_{h1}}{\sigma_n^2 \|\mathbf{g}\|^2}} \mathbf{H}^{-1} \mathbf{u}_{h1} \mathbf{g}^T \mathbf{e}_i = \sqrt{\frac{P_0'\lambda_{h1}}{\sigma_n^2 \|\mathbf{g}\|^2}} g_i \frac{1}{\sigma_{h1}} \mathbf{v}_{h1} \tag{4.134}$$

for $1 \leq i \leq L$, where in step (a) we used, with SVD on $\mathbf{H} = \mathbf{V}_h \mathbf{\Sigma}_h \mathbf{U}_h^T$, the derivation $\mathbf{H}^{-1}\mathbf{u}_{h1} = \mathbf{V}_h \mathbf{\Sigma}_h^{-1} \mathbf{U}_h^T \mathbf{u}_{h1} = \mathbf{V}_h \mathbf{\Sigma}_h^{-1} \mathbf{e}_1 = \frac{1}{\sigma_{h1}} \mathbf{v}_{h1}$, where σ_{h1} denotes the first singular value of $\mathbf{H}$ - which in turn is the largest one in terms of magnitude - and $\mathbf{v}_{h1}$ denotes the first column vector of $\mathbf{V}_h$ - the corresponding first right singular vector of $\mathbf{H}$. The scalar $g_i = \mathbf{g}^T \mathbf{e}_i$ in (4.134) denotes again the ith element of $\mathbf{g}$. Note that $\sigma_{h1} \neq 0$, since we assumed that $\mathbf{H}$ is invertable. Writting[1] $\sigma_{h1} = \text{sign}(\sigma_{h1}) |\sigma_{h1}|$ and using the relation $\lambda_{h1} = |\sigma_{h1}|^2$, we can reformulate (4.134) as

$$\mathbf{a}_i^* = \text{sign}(\sigma_{h1}) \sqrt{\frac{P_0'}{\sigma_n^2 \|\mathbf{g}\|^2}} g_i \mathbf{v}_{h1}, \tag{4.135}$$

[1]The signum function $\text{sign}(a)$ on $a \in \mathbb{R}$ returns 1 for $a \geq 0$ and -1 for $a < 0$.

where we recognize that $\mathbf{a}_i^*$ does not depend on the magnitude of σ_{h1} or λ_{h1}.

Hence, we have determined optimal LO_i' for a coherent MAC, where all $\mathbf{H}_i$ for $1 \leq i \leq L$ are assumed to be identical - for the equivalent model with scalar observation. Finally, we give the optimal FI $J_{\mathbf{z}}$, that follows for an optimal LO_i'. To that end, we accept the steps in (4.119) - except the last step, where we used instead of (4.116), (4.132), i.e., we obtain

$$J_{\mathbf{z}}^* = \frac{\|\mathbf{g}\|^2}{\sigma_n^2} \frac{P_0'}{\sigma_h^2/\lambda_{h1} + P_0'}. \tag{4.136}$$

4.4.3.3 Optimal Power Scheduling for an Orthogonal MAC

Let us now study an optimal power scheduling strategie, especially for the orthogonal MAC case. To that end, we still consider the equivalent model with scalar observation and suppose optimal local sensors LO_i' for all i as already determined in closed-form (cf. (4.97)). Then, we have already derived the optimal, resulting FI $J_{\mathbf{z}}^*$ shown in (4.99) - for both constraints (C1-s) and (C2-s). Let us first consider (C1-s). It raises the question of how a given total power $P_0 = \sum_{i=1}^{L} P_{0,i}$ should be allocated optimally to the individual sensors. Similar holds, when we consider constraint (C2-s), i.e., how a given total variance $P_0' = \sum_{i=1}^{L} P_{0,i}'$ should be allocated optimally to the individual sensors.

We assume the case in which $\sigma_{n_i}^2 > 0$ and thus $g_i > 0$ for all i. Let us recall the FI $J_{\mathbf{z}}^*$ from (4.99), which can also be written as

$$\begin{aligned}
J_{\mathbf{z}}^* &= \sum_{i=1}^{L} b_i^{(1)} \frac{P_i}{b_i^{(2)} + b_i^{(3)} P_i}, \\
\text{where } b_i^{(1)} &\triangleq g_i^2 h_{i\,\max}^2, \\
b_i^{(3)} &\triangleq \sigma_{n_i}^2 h_{i\,\max}^2, \\
b_i^{(2)} &\triangleq \begin{cases} \sigma_{h_i}^2 \left((g_i\theta)^2 + \sigma_{n_i}^2 \right) & (\text{C1-s})_i \\ \sigma_{h_i}^2 \sigma_{n_i}^2, & (\text{C2-s})_i \end{cases} \\
P_i &\triangleq \begin{cases} P_{0,i} & (\text{C1-s})_i \\ P_{0,i}'. & (\text{C2-s})_i \end{cases}
\end{aligned} \tag{4.137}$$

Before we define the optimal power scheduling problem, we first treat the case of an uniform power scheduling strategie, in order to obtain a perfomance benchmark for the optimal power scheduling. To that end, we use the notation of (4.137) and still intruduce

$$P \triangleq \begin{cases} P_0 & (\text{C1-s})_i \\ P_0', & (\text{C2-s})_i \end{cases}$$

so that both constraints to be addressed simultaneously in what follows.

Uniform Power Scheduling Suppose all sensors use the same transmit power/variance, i.e., $P_i = P/L$ ($L \geq 1$). Then, (4.137) yields

$$J_{\mathbf{z},u}(P) \triangleq J_{\mathbf{z}}^* = \sum_{i=1}^{L} b_i^{(1)} \frac{P/L}{b_i^{(2)} + b_i^{(3)} P/L} = \sum_{i=1}^{L} b_i^{(1)} \frac{P}{b_i^{(2)} L + b_i^{(3)} P}, \quad (4.138)$$

where we introduced $J_{\mathbf{z},u}(P)$ - the FI $J_{\mathbf{z}}^*$ from (4.137) as a function on the total power/variance P for a uniform power scheduling strategie. We now analyse the asymptotic behaviour of $J_{\mathbf{z},u}(P)$ for $P \to \infty$. It is easy to verify that

$$J_{\mathbf{z},u}(P \to \infty) = \lim_{P\to\infty} \sum_{i=1}^{L} b_i^{(1)} \frac{P}{b_i^{(2)} L + b_i^{(3)} P} = \sum_{i=1}^{L} \frac{b_i^{(1)}}{b_i^{(3)}} = \sum_{i=1}^{L} \frac{g_i}{\sigma_{n_i}^2}, \quad (4.139)$$

the same result for both constraints. Since, $J_{\mathbf{z},u}(P)$ is a monotonic function in P, we have that

$$J_{\mathbf{z},u}(P \to \infty) > J_{\mathbf{z},u}(P),$$

for all $P \in \mathbb{R}^+$. Thus, (4.139) is an upper bound for $J_{\mathbf{z},u}(P)$.

Optimal Power Scheduling We assume at first that $h_{i\,\max}^2 > 0$ for all i, i.e., we exclude the case when $\mathbf{H}_i = \mathbf{0}$. Now we consider an optimal power allocation strategy, whereby transmit power is optimally scheduled among sensors to achieve the best estimation performance. We study the following problem under a total power/variance constraint:

$$\begin{aligned} \underset{P_0, P_1, \ldots, P_L}{\text{maximize}} \quad & \sum_{i=1}^{L} b_i^{(1)} \frac{P_i}{b_i^{(2)} + b_i^{(3)} P_i} \\ \text{subject to} \quad & \sum_{i=1}^{L} P_i \leq P, \\ & P_i \geq 0 \quad \text{for } 1 \leq i \leq L, \end{aligned} \quad (4.140)$$

i.e., maximizing the FI from (4.137) for a given total power/variance $P = \sum_{i=1}^{L} P_i$ (constraint), w.r.t. $P_i \geq 0$ for $1 \leq i \leq L$. We first, reformulate problem (4.140) equivalently into standard form [15] as

$$\begin{aligned} \underset{P_0, P_1, \ldots, P_L}{\text{minimize}} \quad & -\sum_{i=1}^{L} b_i^{(1)} \frac{P_i}{b_i^{(2)} + b_i^{(3)} P_i} \\ \text{subject to} \quad & \sum_{i=1}^{L} P_i - P \leq 0, \\ & P_i \geq 0 \quad \text{for } 1 \leq i \leq L, \end{aligned} \quad (4.141)$$

which is equivalent to problem (A.1) in Appendix A for $x_k = P_i$, $c_k^{(1)} = b_i^{(1)}$, $c_k^{(2)} = b_i^{(2)}$, $c_k^{(3)} = b_i^{(3)}$, $K = L$, $P = P$. The resulting optimum P_i^* for $1 \leq i \leq L$, can be obtained by a so called "water-filling" procedure and can be expressed as

$$P_i^* = \max\left\{0, \sqrt{\frac{b_i^{(2)} b_i^{(1)}}{(b_i^{(3)})^2} \frac{1}{\nu^*}} - \frac{b_i^{(2)}}{b_i^{(3)}}\right\}, \tag{4.142}$$

and

$$\sum_{i=1}^{L} \max\left\{0, \sqrt{\frac{b_i^{(2)} b_i^{(1)}}{(b_i^{(3)})^2} \frac{1}{\nu^*}} - \frac{b_i^{(2)}}{b_i^{(3)}}\right\} = P. \tag{4.143}$$

The optimal P_i for $1 \leq i \leq L$ for (4.141) and also for (4.140) can not be computed in closed-form. First, we have to determine the optimal variable ν from (4.143). Subsequently, the optimal P_i for $1 \leq i \leq L$ can then be computed using (4.142). This can be done by a so called "water-filling" algorithm (Cf. Algorithm A.1).

In Subsection 5.1.1, we will analyse the optimal power scheduling versus the uniform power scheduling performance in some numerical experiments.

4.4.3.4 Implementation of an Optimal Local Sensor

So far, we have solved (P-II-s) for an orthogonal MAC without any restriction. For the coherent MAC we modified (P-II-s) concerning the constraint, we considered a total power constraint (C1-t) and (C2-t) instead of (C1-s) and (C2-s), where we then have determined the optimal local sensor rule LO_i' for (C2-t) and for certain special cases on the channel matrix $\mathbf{H}_i$ for all i.

However, we have solved the local sensor rules LO_i' for the equivalent model with scalar observation, i.e., $\mathrm{LO}_i'^* = \left(\mathbf{a}^*, \mathbf{C}_{l_i}'^* = \mathbf{0}\right)$. According to Corollary 4.4.17, we finally obtain the optimal local sensor rule LO_i for our original model as

$$\begin{aligned}
&\mathrm{LO}_i^* \Leftrightarrow \mathrm{LO}_i'^* : \\
&\mathbf{A}_i^* = \mathbf{a}_i^* \widetilde{\mathbf{g}}_i^T, \quad \text{where } \widetilde{\mathbf{g}}_i = \begin{cases} \mathbf{C}_{n_i}^\dagger \mathbf{g}_i & \mathbf{g}_i \in \mathrm{R}\left(\mathbf{C}_{n_i}\right) \setminus \{\mathbf{0}\} \\ \mathbf{g}_{i\perp}' & \mathbf{g}_{i\perp}' \neq \mathbf{0} \\ \mathbf{0} & \mathbf{g}_i = \mathbf{0}, \end{cases} \\
&\mathbf{C}_{l_i}^* = \mathbf{C}_{l_i}'^* = \mathbf{0},
\end{aligned} \tag{4.144}$$

where $\mathbf{g}_{i\perp}'$ is defined in (4.37). Let us now, summarize all main results for a scalar parameter.

Orthogonal MAC: Let us recall (4.97) - the optimal local sensor vector $\mathbf{a}_i$ for sensor i, when we consider an orthogonal MAC. In what follows, we consider only the case, when $g_i > 0$ and $\sigma^2_{n_i} > 0$ - the trivial case, when $g_i = 0$ and $\sigma^2_{n_i} =$ has no optimal solution, since $J_{\mathbf{z}} = 0$ for choices of $\mathbf{a}$ - for the case, when $g_i > 0$ and $\sigma^2_{n_i} =$ cf. discussion in Subsubsection 4.4.3.1. With the already known solution $\mathbf{C}^*_{l_i} = \mathbf{0}$, we obtain the optimal LO_i, according to (4.144), as

$$\mathrm{LO}^*_i: \quad \left(\mathbf{A}^*_i = c^*_i \mathbf{e}_{i\max} \widetilde{\mathbf{g}}^T_i,\ \mathbf{C}^*_{l_i} = \mathbf{0}\right),$$

$$\text{where } c^*_i = \begin{cases} \sqrt{\frac{P_{0,i}}{(g_i\theta)^2 + \sigma^2_{n_i}}} & \text{(C1-s)} \\ \sqrt{\frac{P'_{0,i}}{\sigma^2_{n_i}}}, & \text{(C2-s)} \end{cases} \tag{4.145}$$

where $\widetilde{\mathbf{g}}_i$ is given in (4.144). The unit vector $\mathbf{e}_{i\max}$ corresponds to the largest diagonal entry of $\mathbf{H}^T_i\mathbf{H}_i$, so $h^2_{i\max}$. The model parameter g_i and σ_{n_i} are given in (4.63). An implementation is illustrated in Fig. 4.6, which can be regarded as a three stage implementation.

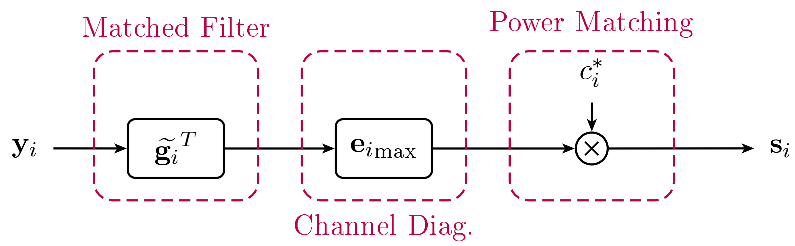


Figure 4.6: Optimal LO_i implementation for a scalar parameter and orthogonal MAC.

The first stage in Fig. 4.6, can be regarded as a *Matched Filter*, i.e., an optimal prefiltering (matching) in accordance with the local observation model $(\mathbf{g}_i, \mathbf{C}_{n_i})$. The *Channel Diagonalization* stage, forces the optimal direction for the transmit data $\mathbf{s}_i$ onto the strongest transmission path of the given channel $\mathbf{H}_i$. Finally, the task of the amplification stage is, to attain the maximum available power for the transmit data $\mathbf{s}_i$, i.e., a *Power Matching* for $\mathbf{s}_i$. Here, the gain is given by (4.96) and depends on the constraint (C1-s) or (C2-s), respectively.

Let us finally consider the implementation of an optimal LO_i for constraint (C1-s) in more detail. As already mentioned and as can be seen in (4.96), the optimal solution for an LO_i depends on the unknown paramter θ. At first glance, the sensor thus can not be implemented optimally. However, only the third stage in Fig. 4.6, i.e., only the optimal c_i denpends on θ. The optimal value of c_i occurs, when the power of $\mathbf{s}_i$, i.e., $\mathrm{E}\left\{\mathbf{s}^T_i\mathbf{s}_i\right\}$ reaches the given constant $P_{0,i}$. Thus, we can implement the optimal local sensor LO_i^* for constraint (C1-s) as follows: Choosing the first two stages of Fig. 4.6 as usual, i.e., a matching and the channel diagonalization, which do not depend on θ. The third stage can be implemented by means of a control loop as illustrated in Fig. 4.7. Starting with an arbitrary initial value $c_i = c_i^s$, a controller C increases the factor c_i until the deviation $d = P_{0,i} - \mathrm{E}\left\{\mathbf{s}_i^T\mathbf{s}_i\right\} = 0$, i.e., unitl the steady state is reached. After the steady state has been reached, the power on $\mathbf{s}_i$ yields $\mathrm{E}\left\{\mathbf{s}_i^T\mathbf{s}_i\right\} = P_{0,i}$ and we have determined the optimal $c_i = c_i^*$ and thus the optimal $\mathbf{A}_i$ for (C1-s).

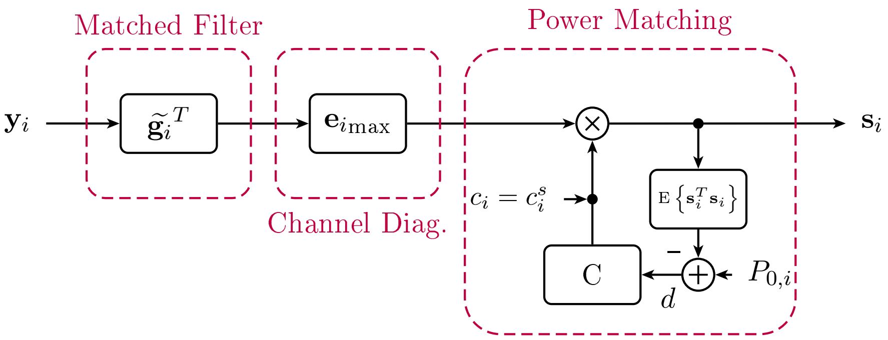


Figure 4.7: Optimal LO_i implementation for a scalar paramter - considering an orthogonal MAC and constraint (C1-s). A control loop with an controller C is implemented to obtain maximum transmit power for $\mathbf{s}_i$ in the steady state.

Coherent MAC: Let us finally, give the optimal LO_i for the coherent MAC case. To that end, we recall the optimal local sensor vector $\mathbf{a}_i$ for LO_i', given in (4.118) for the case of orthogonal individual channel matrices $\mathbf{H}_i$ for $1 \leq i \leq L$, and in (4.135) for the case of identical and invertable individual channel matrices $\mathbf{H} = \mathbf{H}_i$ for $1 \leq i \leq L$. With the already known solution $\mathbf{C}_{l_i}^* = \mathbf{0}$, we obtain the optimal LO_i for the orthogonal individual channel case, according to (4.144), as

$$\mathrm{LO}_i^*\colon \quad \left(\mathbf{A}_i^* = \sqrt{\frac{P_0'}{\sigma_n^2 \left\|\mathbf{g}\right\|^2}} g_i \mathbf{h}_{i\,1}^r \widetilde{\mathbf{g}}_i^T,\ \mathbf{C}_{l_i}^* = \mathbf{0}\right). \tag{4.146}$$

and for the identical individual channel case as

$$\mathrm{LO}_i^*\colon \quad \left(\mathbf{A}_i^* = \operatorname{sign}\left(\sigma_{h1}\right)\sqrt{\frac{P_0'}{\sigma_n^2 \left\|\mathbf{g}\right\|^2}} g_i \mathbf{v}_{h1} \widetilde{\mathbf{g}}_i^T,\ \mathbf{C}_{l_i}^* = \mathbf{0}\right). \tag{4.147}$$

The vector $\widetilde{\mathbf{g}}_i$ in (4.146) and (4.146), is again given in (4.144). Note that for the coherent MAC we only considered the first case for $\widetilde{\mathbf{g}}_i$ in (4.144), i.e., when $\mathbf{g}_i \in \mathrm{R}(\mathbf{C}_{n_i}) \setminus \{\mathbf{0}\}$. The vector $\mathbf{h}^r_{i\,1}$ in (4.146) denotes the first row-vector of the unitary $\mathbf{H}_i$. The vector $\mathbf{v}_{h1}$ in (4.147) denotes the largest right singular vector of $\mathbf{H} = \mathbf{H}_i$ for all i, which corresponds to the largest singular value σ_{h1} of $\mathbf{H}$ (in terms of magnitude). The model parameter g_i and σ_{n_i} are given in (4.63). Note that in the coherent MAC case, we assumed $g_i > 0$ and $\sigma_n^2 = \sigma_{n_i} > 0$. Finally P_0' denotes the total variance power including all sensors.

The implementation of the optimal LO_i for both cases is similar to Fig. 4.6. Only the second stage differs in choosing $\mathbf{h}^r_{i\,1}$ or $\mathbf{v}_{h1}$ instead of $\mathbf{e}_{i\max}$, respectively.

4.5 Vector Parameter Case

Let us now consider the general case of a vector parameter $\boldsymbol{\theta} \in \mathbb{R}^n$. Here, we exclusively use the standard model (cf. Definition 4.3.14), i.e., we consider only the case of an orthogonal MAC - the observation and channel noise are iid - the channel matrix $\mathbf{H}_i$ is diagonal for all i. Thus, we describe the ith observation noise covariance matrix by $\mathbf{C}_{n_i} = \sigma^2_{n_i}\mathbf{I}$ and the ith channel noise covariance matrix by $\mathbf{C}_{h_i} = \sigma^2_{h_i}\mathbf{I}$, where $\sigma^2_{n_i} > 0$ and $\sigma^2_{h_i} > 0$. The assumed diagonal channel matrix $\mathbf{H}_i$, can thus be written as

$$(\mathbf{H})_{k,l} = \begin{cases} h_l & l = k \\ 0 & k \neq l \end{cases} \quad \text{for } 1 \leq k \leq p_i \text{ and } 1 \leq l \leq q_i. \tag{4.148}$$

So far, we have already solved the basic optimization problem (P-I) w.r.t. $\mathbf{C}_{l_i}$, where the resulting optimum is given by $\mathbf{C}^*_{l_i} = \mathbf{0}$ (for all i) (cf. Section 4.1). We thus consider the optimization problem (P-II) to determine the still unknown sensor matrix $\mathbf{A}_i$ (for all i) ϕ-optimally. Let us first rewrite the FIMs $\mathbf{J}_{\mathbf{z}_i}$ from (3.5) and $\mathbf{J}_{\mathbf{z}}$ from (3.8) according to the standard model as

$$\mathbf{J}_{\mathbf{z}_i} = \mathbf{G}_i^T\mathbf{A}_i^T\mathbf{H}_i^T\left(\sigma^2_{h_i}\mathbf{I} + \sigma^2_{n_i}\mathbf{H}_i\mathbf{A}_i\mathbf{A}_i^T\mathbf{H}_i^T\right)^{-1}\mathbf{H}_i\mathbf{A}_i\mathbf{G}_i \tag{4.149}$$

and

$$\mathbf{J}_{\mathbf{z}} = \sum_{i=1}^{L}\mathbf{G}_i^T\mathbf{A}_i^T\mathbf{H}_i^T\left(\sigma^2_{h_i}\mathbf{I} + \sigma^2_{n_i}\mathbf{H}_i\mathbf{A}_i\mathbf{A}_i^T\mathbf{H}_i^T\right)^{-1}\mathbf{H}_i\mathbf{A}_i\mathbf{G}_i. \tag{4.150}$$

With these assumptions, we consider further problem (P-II), where $\mathbf{J}_{\mathbf{z}}$ is now given in (4.150). In Section 4.2, we have showed that for an orthogonal MAC and when the optimality criterion function ϕ is linear, we can solve an equivalent problem (P-III) in order to obtain the ϕ-optimal $\mathbf{A}_i$ for a specific sensor i. In what follows, we are interested on a T- and A-optimal design for a local sensor i.

4.5.1 T–Optimal Design

Let us first consider the T–optimality criterion ϕ_1, defined in (3.41). A T–optimal designed local sensor rule $\mathrm{LO}^*_{i\,\phi_1}$ maximizes the trace of the FIM $\mathbf{J_z}$, while respecting the constraint (C1) or (C2), respectively. Since, the trace and thus ϕ_1 is a linear function on NND(n) (cf. (3.42)), we can equivalently solve (P-III), where $\mathbf{J}_{\mathbf{z}_i}$ is now given in (4.149), in order to determine the T–optimal local sensor matrix $\mathbf{A}_i$ for sensor i. Thus, we consider a single-sensor model, since all L sensors can be determined independently of each other optimally. In particular, we treat only the case with constraint (C2). In the following, we let the index notation to address the ith sensor away, i.e., we set $\mathbf{G}_i = \mathbf{G}$, $\mathbf{A}_i = \mathbf{A}$, $\mathbf{H}_i = \mathbf{H}$, $\mathbf{C}_{n_i} = \mathbf{C}_n$, $\mathbf{C}_{h_i}$, $P_{0,i} = P_0$, $P'_{0,i} = P'_0$, $p_i = p$, $q_i = q$, $m_i = m$. Hence, we can state the following optimization problem:

$$\begin{aligned} &\underset{\mathbf{A}\in\mathbb{R}^{q\times m}}{\text{maximize}} && \phi_1\left\{\mathbf{A}\right\} = \frac{1}{n}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I} + \sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G}\right\} \\ &\text{subject to} && \mathrm{tr}\left\{\mathbf{A}\mathbf{A}^T\right\} \leq P'_0/\sigma_n^2, \qquad \text{(C2)} \end{aligned} \tag{4.151}$$

where we introduced the notation $\phi_1\left\{\mathbf{A}_i\right\} = \phi_1\left\{\mathbf{J}_{\mathbf{z}_i}\right\}$. In what follows, we solve (4.151) for certain special cases, where we make assumptions on the channel matrix $\mathbf{H}$. First, we assume an orthogonal channel matrix. Then we generalized it to a rectangular channel matrix, where we suppose full column-rank.

4.5.1.1 Orthogonal Channel

Here we assume that the channel matrix $\mathbf{H}$ is orthogonal (unitary), i.e., $\mathbf{H}^T\mathbf{H} = \mathbf{H}\mathbf{H}^T = \mathbf{I}$. Implying that $\mathbf{H}$ is a squared matrix, i.e., $p = q$.

Lemma 4.5.20 *Any unitary matrix* $\mathbf{U}$ *has singular values equal to one.*

Proof. Cf. [16, Theorem 6.2, p. 173] □

Since, we assumed that the channel matrix $\mathbf{H}$ is diagonal (standard model), we conclude according to Lemma 4.5.20 that the diagonal elements h_l (cf. (4.148)) are given by $h_l = 1$ for $1 \leq l \leq q$, which in turn yields that $\mathbf{H} = \mathbf{I}$. Thus, the objective function $\phi_1(\cdot)$ in (4.151), can be equivalently reformulated as

$$
\begin{aligned}
\phi_1\{\mathbf{A}\} &= \frac{1}{n}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I}+\sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G}\right\}\\
&\overset{\mathbf{H}=\mathbf{I}}{=} \frac{1}{n}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{A}^T\left(\sigma_h^2\mathbf{I}+\sigma_n^2\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{G}\right\}\\
&= \frac{1}{n}\frac{1}{\sigma_n^2}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{A}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{A}\mathbf{G}\right\}\\
&\overset{(a)}{=} \frac{1}{n}\frac{1}{\sigma_n^2}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T\right)^{-1}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{G}\right\}\\
&= \frac{1}{n}\frac{1}{\sigma_n^2}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T\mathbf{U}\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\mathbf{\Sigma}\mathbf{\Sigma}^T\right)^{-1}\mathbf{U}^T\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{G}\right\}\\
&= \frac{1}{n}\frac{1}{\sigma_n^2}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{\Sigma}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\mathbf{\Sigma}\mathbf{\Sigma}^T\right)^{-1}\mathbf{\Sigma}\mathbf{V}^T\mathbf{G}\right\}\\
&\overset{(b)}{=} \frac{1}{n}\frac{1}{\sigma_n^2}\mathrm{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{G}\right\},
\end{aligned}
\tag{4.152}
$$

In step (a) we performed the SVD $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, with unitary matrices $\mathbf{U} \in \mathbb{R}^{q\times q}$, $\mathbf{V} \in \mathbb{R}^{m\times m}$ and the rectangular diagonal matrix $\mathbf{\Sigma}$ of size $q \times m$, which contains the singular values σ_j for $1 \le j \le w \triangleq \min\{q, m\}$ of $\mathbf{A}$ on the main diagonal. We assume that the singular values σ_j are ordered decreasingly (in terms of magnitude), i.e., $|\sigma_1| \ge |\sigma_2| \ge \cdots \ge |\sigma_w| \ge 0$. In step (b) we introduced the diagonal matrix

$$
\mathbf{D} \triangleq \mathbf{\Sigma}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\mathbf{\Sigma}\mathbf{\Sigma}^T\right)^{-1}\mathbf{\Sigma}, \tag{4.153}
$$

which is indeed squared and diagonal of size $m \times m$. The m elements on the main diagonal are thus given by

$$
d_j \triangleq (\mathbf{D})_{j,j} = \begin{cases} \frac{\sigma_j^2}{\frac{\sigma_h^2}{\sigma_n^2}+\sigma_j^2} & 1 \le j \le w \\ 0 & w < j \le m. \end{cases} \tag{4.154}
$$

As can be verified easily, the diagonal elements d_j are ordered decreasingly, i.e., $d_1 \ge d_2 \ge ... \ge d_m \ge 0$, as a result of the adoption order on the set $\{|\sigma_j|\}_{j=1}^{w}$.

The constraint (C2) in (4.151), can also be written in terms of SVD $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ as

$$
\begin{aligned}
\mathrm{tr}\left\{\mathbf{A}\mathbf{A}^T\right\} &= \mathrm{tr}\left\{\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T\right\}\\
&= \mathrm{tr}\left\{\mathbf{U}^T\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\right\}\\
&= \mathrm{tr}\left\{\mathbf{\Sigma}\mathbf{\Sigma}^T\right\} \le P_0'/\sigma_n^2, \quad \text{(C2)}
\end{aligned}
\tag{4.155}
$$

where we used again the cyclic property of the trace operator [8] and the facts that $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ and $\mathbf{U}^T\mathbf{U} = \mathbf{I}$.

With the derivations in (4.152) (excluding constants, cf. p. 59 ftn. 1) and (4.155), the optimization problem (4.151) can thus be reformulated equivalently in terms of SVD $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ as

$$\begin{aligned} \underset{\mathbf{\Sigma},\mathbf{V}}{\text{maximize}} \quad & \operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{G}\right\} \\ \text{subject to} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}\mathbf{\Sigma}^T\right\} \le P_0'/\sigma_n^2, \\ & \mathbf{\Sigma}\mathbf{\Sigma}^T \ge \mathbf{0} \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}. \end{aligned} \tag{4.156}$$

Note that problem (4.156), does not depend on the channel matrix $\mathbf{H}$ as a consequence that $\mathbf{H}$ is assumed to be unitary. Furthermore, problem (4.156) does not dependend on the unitary $\mathbf{U}$ - the left singular vectors of the sensor matrix $\mathbf{A}$ - so that an arbitrary orthogonal (unitary) $\mathbf{U}$ can be chosen, e.g., $\mathbf{U} = \mathbf{I}$.

In what follows, we will solve (4.156) sequentially, by determining first the optimum $\mathbf{V}$ and then the optimum $\mathbf{\Sigma}$. A necessary condition for $\mathbf{V}$ to be optimum in (4.156), can be obtained by fixing $\mathbf{\Sigma}$. For a fixed $\mathbf{\Sigma}'$, the optimal $\mathbf{V}$ has to solve the problem

$$\begin{aligned} \underset{\mathbf{V}}{\text{maximize}} \quad & \operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{G}\right\} \\ \text{subject to} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}'\mathbf{\Sigma}'^T\right\} \le P_0'/\sigma_n^2 \\ & \mathbf{\Sigma}'\mathbf{\Sigma}'^T \ge \mathbf{0} \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \end{aligned} \tag{4.157}$$

where $\mathbf{D}'$ is obtained from (4.153) by inserting $\mathbf{\Sigma}'$ for $\mathbf{\Sigma}$. A $\mathbf{V}$ is optimum for (4.158) if and only if it is optimum for

$$\begin{aligned} \underset{\mathbf{V}}{\text{maximize}} \quad & \operatorname{tr}\left\{\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{G}\mathbf{G}^T\right\} \\ \text{subject to} \quad & \mathbf{V}^T\mathbf{V} = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \end{aligned} \tag{4.158}$$

as can be verified easily, where we used again

$$\operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{G}\right\} = \operatorname{tr}\left\{\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{G}\mathbf{G}^T\right\}.$$

Let us denote the EVD of $\mathbf{G}\mathbf{G}^T$ by $\mathbf{G}\mathbf{G}^T = \mathbf{U}_g\mathbf{\Lambda}_g\mathbf{U}_g^T$, with the unitary matrix $\mathbf{U}_g \in \mathbb{R}^{m\times m}$ (contains the eingenvectors of $\mathbf{G}\mathbf{G}^T$) and the diagonal matrix $\mathbf{\Lambda}_g$ of size $m \times m$, which contains the positive eigenvalues λ_{g_j} for $1 \le j \le m$ of $\mathbf{G}\mathbf{G}^T$ on the main diagonal. We assume that the eigenvalues λ_{g_j} are ordered decreasingly, i.e., $\lambda_{g_1} \ge \lambda_{g_2} \ge \cdots \ge \lambda_{g_m} \ge 0$, since the diagonal values of $\mathbf{D}$ are olso sorted decreasingly. Then, we have by Theorem 4.4.19 that the optimal $\mathbf{V}$, solving (4.158), is given by

$$\mathbf{V}^* = \mathbf{U}_g, \tag{4.159}$$

where the column vectors of $\mathbf{U}_g$ correspond to the eigenvectors of $\mathbf{G}\mathbf{G}^T$, sorted decreasingly[1].

After having determined the optimal $\mathbf{V}$ for (4.156), it remains to determine the optimal $\mathbf{\Sigma}$ for (4.156). To that end, we insert the optimal $\mathbf{V}$ from (4.159) into (4.156), yielding:

$$\begin{aligned} \underset{\mathbf{\Sigma}}{\text{maximize}} \quad & \operatorname{tr}\{\mathbf{\Lambda}_g \mathbf{D}\} \\ \text{subject to} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}\mathbf{\Sigma}^T\right\} \leq P_0'/\sigma_n^2, \\ & \mathbf{\Sigma}\mathbf{\Sigma}^T \geq \mathbf{0}, \end{aligned} \tag{4.160}$$

which has to be solved for the optimal $\mathbf{\Sigma}$, which then yields together with $\mathbf{V}^*$ the solution for (4.156). Let us introduce the vector notation $\mathbf{s} = (s_1, s_2, \ldots, s_w)^T \triangleq \left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_w^2\right)^T \in \mathbb{R}^{+w}$, i.e., $\mathbf{s} \succeq \mathbf{0}$. Then, we can reformulate the optimization problem (4.156) equivalently in standard form [15] as

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & -\sum_{j=1}^{w} \lambda'_{g_j} \frac{s_j}{\frac{\sigma_h^2}{\sigma_n^2} + s_j} \\ \text{subject to} \quad & \mathbf{1}^T\mathbf{s} - P_0' \leq 0, \\ & -\mathbf{s} \preceq \mathbf{0}, \end{aligned} \tag{4.161}$$

where

$$\lambda'_{g_j} \triangleq \begin{cases} \lambda_{g_j} & 1 \leq j \leq m \\ 0 & m < i \leq w. \end{cases} \tag{4.162}$$

The optimization problem (4.161) is equivalent to problem (A.1) in Appendix A for $x_k = s_j$, $c_k^{(1)} = \lambda'_{g_j}$, $c_k^{(2)} = \frac{\sigma_h^2}{\sigma_n^2}$, $c_k^{(3)} = 1$, $K = w$, $P = P_0'$. The resulting optimum s_j^* for $1 \leq j \leq w$, can be obtained by a so called "water-filling" procedure and can be expressed as

$$s_j^* = \max\left\{0, \sqrt{\frac{\sigma_h^2}{\sigma_n^2}\frac{\lambda'_{g_j}}{\nu^*}} - \frac{\sigma_h^2}{\sigma_n^2}\right\} \tag{4.163}$$

and

$$\sum_{j=1}^{w} \max\left\{0, \sqrt{\frac{\sigma_h^2}{\sigma_n^2}\frac{\lambda'_{g_j}}{\nu^*}} - \frac{\sigma_h^2}{\sigma_n^2}\right\} = P_0'. \tag{4.164}$$

[1]The eigenvectors of a symmetric matrix are sorted decreasingly/increasingly if the corresponding eingenvalues are sorted decreasingly/increasingly

The optimal s_j for $1 \leq j \leq w$ for (4.161) can not be computed in closed-form. However, we can determine the optimal ν, first numerically, according to (4.164). Subsequently, we can compute the optimal s_j for $1 \leq j \leq w$ using (4.163). This can be done by a so called "water-filling" algorithm (cf. Algorithm A.1). With the optimal s_j for (4.161), we can finally compute the optimal singular values σ_j of $\mathbf{A}$ with $\sigma_j^* = \sqrt{s_j^*}$ for $1 \leq j \leq w$, which are then arranged on the main diagonal of the optimal $\mathbf{\Sigma}$ in decreasing order.

After having determined the optimal $\mathbf{V}$ and $\mathbf{\Sigma}$ for (4.156), we are now able to compute the optimal local sensor matrix $\mathbf{A}$ with $\mathbf{U} = \mathbf{I}$ as

$$\mathbf{A}^* = \mathbf{\Sigma}^* \mathbf{V}^{*T} \overset{(4.159)}{=} \mathbf{\Sigma}^* \mathbf{U}_g. \tag{4.165}$$

Note again: the optimal $\mathbf{\Sigma}$ can not be expressed in closed-form - it has to be determined numerically ("water-filling" procedure); the unitary matrix $\mathbf{U}_g$ contains the eigenvectors of $\mathbf{G}\mathbf{G}^T$, sorted decreasingly.

Conclusions Let us return to our original index notation that indicates the ith sensor and recall that $\mathbf{C}_{l_i}^* = \mathbf{0}$ for sensor i. For an orthogonal MAC with L local sensors, the T-optimal ith LO - for the standard model - considering constraint (C2) - assuming an unitary channel matrix $\mathbf{H}_i$ - is given by

$$\mathrm{LO}_{i\,\phi_1}^*: \quad \left(\mathbf{A}_i^* = \mathbf{\Sigma}_i^* \mathbf{U}_{g_i},\ \mathbf{C}_{l_i}^* = \mathbf{0} \right), \tag{4.166}$$

where the unitary matrix $\mathbf{U}_{g_i}$ contains the eigenvectors of $\mathbf{G}_i\mathbf{G}_i^T$ sorted decreasingly, the diagonal $\mathbf{\Sigma}_i^*$ has to be determined in a "water-filling" principle in order to balance channel noise and sensor observation states of sensor i.

4.5.1.2 Invertible Channel

Still considering problem (4.151), we now allow for a general invertible channel matrix $\mathbf{H}$, i.e., it exist $\mathbf{H}^{-1}$. That implies $h_l \neq 0$ for $1 \leq l \leq q = p$ (cf. (4.148)).

Introducing

$$\widetilde{\mathbf{A}} \triangleq \mathbf{H}\mathbf{A}, \tag{4.167}$$

the objective function $\phi_1(\cdot)$ in (4.151), can then be equivalently reformulated as

$$
\begin{aligned}
\phi_1(\mathbf{A}) &= \operatorname{tr}\left\{\mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I}+\sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G}\right\} \\
&= \frac{1}{\sigma_n^2}\operatorname{tr}\left\{\mathbf{G}^T\widetilde{\mathbf{A}}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^T\right)^{-1}\widetilde{\mathbf{A}}\mathbf{G}\right\} \\
&\overset{(a)}{=} \frac{1}{\sigma_n^2}\operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\right)^{-1}\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{G}\right\} \\
&= \frac{1}{\sigma_n^2}\operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\mathbf{U}\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\right)^{-1}\mathbf{U}^T\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{G}\right\} \\
&= \frac{1}{\sigma_n^2}\operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\boldsymbol{\Sigma}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\right)^{-1}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{G}\right\} \\
&\overset{(b)}{=} \frac{1}{\sigma_n^2}\operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{G}\right\}.
\end{aligned}
\tag{4.168}
$$

In step (a) we performed the SVD $\widetilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, with unitary matrices $\mathbf{U} \in \mathbb{R}^{p\times p}$, $\mathbf{V} \in \mathbb{R}^{m\times m}$ and the rectangular diagonal matrix $\boldsymbol{\Sigma}$ of size $p \times m$, which contains the singular values σ_j for $1 \le j \le w \triangleq \min\{p, m\}$ of $\widetilde{\mathbf{A}}$ on the main diagonal. We assume again that the singular values σ_j are ordered decreasingly (in terms of magnitude), i.e., $|\sigma_1| \ge |\sigma_2| \ge \cdots \ge |\sigma_w| \ge 0$. In step (b) we introduced the diagonal matrix

$$
\mathbf{D} \triangleq \boldsymbol{\Sigma}^T\left(\frac{\sigma_h^2}{\sigma_n^2}\mathbf{I}+\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\right)^{-1}\boldsymbol{\Sigma}, \tag{4.169}
$$

which is indeed squared and diagonal of size $m \times m$. The m elements on the main diagonal are thus given by

$$
d_j \triangleq (\mathbf{D})_{j,j} = \begin{cases} \frac{\sigma_j^2}{\frac{\sigma_h^2}{\sigma_n^2}+\sigma_j^2} & 1 \le j \le w \\ 0 & w < j \le m, \end{cases} \tag{4.170}
$$

As can be verified easily, the diagonal elements d_j are ordered decreasingly, i.e., $d_1 \ge d_2 \ge ... \ge d_m \ge 0$, as a result of the adoption order on the set $\{|\sigma_j|\}_{j=1}^{w}$.

Consider the shorthand $\widetilde{\mathbf{A}}$ from (4.167). Since we assumed that $\mathbf{H}$ is invertible, we can uniquely reclaim $\mathbf{A}$ from $\widetilde{\mathbf{A}}$ with $\mathbf{A} = \mathbf{H}^{-1}\widetilde{\mathbf{A}}$. The constraint (C2) in (4.151), can thus be equivalently written in terms of SVD

$\widetilde{\mathbf{A}} = \mathbf{U\Sigma V}^T$ as

$$\begin{aligned}
\mathrm{tr}\left\{\mathbf{A}\mathbf{A}^T\right\} &= \mathrm{tr}\left\{\mathbf{H}^{-1}\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^T\left(\mathbf{H}^{-1}\right)^T\right\} \\
&= \mathrm{tr}\left\{\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^T\left(\mathbf{H}^{-1}\right)^T\mathbf{H}^{-1}\right\} \\
&= \mathrm{tr}\left\{\mathbf{U\Sigma V}^T\mathbf{V\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\right\} \\
&= \mathrm{tr}\left\{\mathbf{U\Sigma\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\right\} \le \frac{P_0'}{\sigma_n^2}, \quad \text{(C2)}
\end{aligned} \tag{4.171}$$

where we used again the cyclic property of the trace operator, the fact that $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ and $\left(\mathbf{H}^{-1}\right)^T\mathbf{H}^{-1} = \left(\mathbf{H}^T\right)^{-1}\mathbf{H}^{-1} = \left(\mathbf{H}\mathbf{H}^T\right)^{-1}$.

With the derivations in (4.168) (excluding constants, cf. p. 59 ftn. 1) and (4.171), the optimization problem (4.151) can thus be reformulated equivalently in terms of SVD $\widetilde{\mathbf{A}} = \mathbf{U\Sigma V}^T$ as

$$\begin{aligned}
\underset{\mathbf{U},\mathbf{\Sigma},\mathbf{V}}{\text{maximize}} \quad & \mathrm{tr}\left\{\mathbf{G}^T\mathbf{VDV}^T\mathbf{G}\right\} \\
\text{subject to} \quad & \mathrm{tr}\left\{\mathbf{U\Sigma\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\right\} \le \frac{P_0'}{\sigma_n^2}, \\
& \mathbf{\Sigma\Sigma}^T \succeq \mathbf{0}, \\
& \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \\
& \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}.
\end{aligned} \tag{4.172}$$

Note that the the unitary matrix $\mathbf{U}$, the left singular vectors of $\widetilde{\mathbf{A}}$, now enters the constraint in (4.172) and thus has to be chosen optimally for (4.172).

In what follows, we will solve (4.172) sequentially, by determining first the optimal $\mathbf{V}$ and then the optimal $\mathbf{U}$. A necessary condition for $\mathbf{V}$ to be optimum in (4.172), can be obtained by fixing $\mathbf{U}$ and $\mathbf{\Sigma}$. For a fixed $\mathbf{U}'$ and $\mathbf{\Sigma}'$, the optimal $\mathbf{V}$ has to solve the problem

$$\begin{aligned}
\underset{\mathbf{V}}{\text{maximize}} \quad & \mathrm{tr}\left\{\mathbf{G}^T\mathbf{VD}'\mathbf{V}^T\mathbf{G}\right\} \\
\text{subject to} \quad & \mathrm{tr}\left\{\mathbf{U}'\mathbf{\Sigma}'\mathbf{\Sigma}'^T\mathbf{U}'^T\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\right\} \le \frac{P_0'}{\sigma_n^2}, \\
& \mathbf{\Sigma}'\mathbf{\Sigma}'^T \succeq \mathbf{0}, \\
& \mathbf{U}'\mathbf{U}'^T = \mathbf{U}'^T\mathbf{U}' = \mathbf{I}, \\
& \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I},
\end{aligned} \tag{4.173}$$

where $\mathbf{D}'$ is obtained from (4.169), by inserting $\mathbf{\Sigma}'$ for $\mathbf{\Sigma}$. A $\mathbf{V}$ is optimum for (4.173), if and only if it is optimum for

$$\begin{aligned}
\underset{\mathbf{V}}{\text{maximize}} \quad & \mathrm{tr}\left\{\mathbf{VD}'\mathbf{V}^T\mathbf{G}\mathbf{G}^T\right\} \\
\text{subject to} \quad & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I},
\end{aligned} \tag{4.174}$$

as can be verified easily, where we used again

$$\operatorname{tr}\left\{\mathbf{G}^T\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{G}\right\} = \operatorname{tr}\left\{\mathbf{V}\mathbf{D}'\mathbf{V}^T\mathbf{G}\mathbf{G}^T\right\}.$$

Note that problem (4.174) is exactly the same as (4.158), which we have already solved. Thus, optimal $\mathbf{V}$ for problem (4.174) is given by

$$\mathbf{V}^* = \mathbf{U}_g, \tag{4.175}$$

where the column vectors of $\mathbf{U}_g$ correspond to the eigenvectors of $\mathbf{G}\mathbf{G}^T$, sorted decreasingly, i.e., the EVD $\mathbf{G}\mathbf{G}^T = \mathbf{U}_g\mathbf{\Lambda}_g\mathbf{U}_g^T$, where we assume that the eigenvalues λ_{g_j} for $1 \leq j \leq m$ of $\mathbf{G}\mathbf{G}^T$ are ordered decreasingly along the main diagonal in $\mathbf{\Lambda}_g$.

After having determined the optimal $\mathbf{V}$ for (4.172), we will now characterize the optimal $\mathbf{U}$ for (4.172). To that end, we insert the optimal $\mathbf{V}$ from (4.175) into (4.172), yielding:

$$\begin{aligned}
\underset{\mathbf{U},\mathbf{\Sigma}}{\text{maximize}} \quad & \operatorname{tr}\left\{\mathbf{\Lambda}_g\mathbf{D}\right\} \\
\text{subject to} \quad & \operatorname{tr}\left\{\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\right\} \leq \frac{P_0'}{\sigma_n^2} \\
& \mathbf{\Sigma}\mathbf{\Sigma}^T \succeq \mathbf{0} \\
& \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}.
\end{aligned} \tag{4.176}$$

If $\mathbf{U}^*$ is optimal for (4.176), then it is also optimal for (4.172), as can be verified easily. Invoking problem (4.124), we note that it has exactly the same constraint functions and a fairly similar objective. Where we have showed that the optimal $\mathbf{U}$ is determined by minimization of the constraint function

$$t\left(\mathbf{U},\mathbf{\Sigma}\right) \triangleq \operatorname{tr}\left\{\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\right\}.$$

We can closely follow the approach and recognize that this also applies to (4.176). Thus, we can determine the optimal $\mathbf{U}$ by solving

$$\begin{aligned}
\underset{\mathbf{U}}{\text{minimize}} \quad & \operatorname{tr}\left\{\mathbf{U}\mathbf{\Sigma}'\mathbf{\Sigma}'^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\right\} \\
\text{subject to} \quad & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I},
\end{aligned} \tag{4.177}$$

i.e., the optimal $\mathbf{U}$ will be a minimizer of the constraint function $t\left(\mathbf{U},\mathbf{\Sigma}\right)$, for the specific choice of $\mathbf{\Sigma} = \mathbf{\Sigma}'$. An application of Theorem 4.4.19 reveals, that the optimal $\mathbf{U}$ is given by the eigenvectors of $\left(\mathbf{H}\mathbf{H}^T\right)^{-1}$ or $\mathbf{H}\mathbf{H}^T$, respectively, in the order of increasing eigenvalues of $\left(\mathbf{H}\mathbf{H}^T\right)^{-1}$, or, decreasing eigenvalues of $\mathbf{H}\mathbf{H}^T$. However, since $\mathbf{H}\mathbf{H}^T$ and in turn $\left(\mathbf{H}\mathbf{H}^T\right)^{-1}$ is diagonal, the eigenvectors are given by the unit vectors $\{\mathbf{e}_k\}_{k=1}^p$. Moreover, since the eigenvalues of $\left(\mathbf{H}\mathbf{H}^T\right)^{-1}$ are the squared reciprocals of the diagonal values

$h_j \neq 0$ for $1 \leq j \leq p$, we have that the kth column of $\mathbf{U}^*$ is given by $\mathbf{e}_{j_k}$, where j_k is the index of the kth largest main diagonal entry h_{j_k} (in terms of magnitude). Howerver, we have assumed that the diagonal entries are in decreasing order in terms of magnitude w.l.o.g. Thus, the optimal $\mathbf{U}$ is given by

$$\mathbf{U}^* = \mathbf{I}, \tag{4.178}$$

So far, we have determined the optimal $\mathbf{U}$ and $\mathbf{V}$ for (4.172). It remains to determine the optimal singular values σ_j for $1 \leq j \leq w$. Inserting the optimal choices $\mathbf{U}^*$ and $\mathbf{V}^*$ in (4.172), yielding:

$$\begin{aligned} \underset{\mathbf{\Sigma}}{\text{maximize}} \quad & \operatorname{tr}\{\mathbf{\Lambda}_g \mathbf{D}\} \\ \text{subject to} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}\mathbf{\Sigma}^T \mathbf{\Lambda}_h^{-1}\right\} \leq P_0'/\sigma_n^2, \\ & \mathbf{\Sigma}\mathbf{\Sigma}^T \geq \mathbf{0}, \end{aligned} \tag{4.179}$$

Let us accept the notation $\mathbf{s} = \left(s_1, s_2, \ldots, s_w\right)^T \triangleq \left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_w^2\right)^T \in \mathbb{R}^{+w}$ as for the last case, where we considered an orthogonal channel matrix, in turn: $\mathbf{s} \succeq \mathbf{0}$. Further, we introduce the vector $\mathbf{b} = \left(h_1^{-2}, h_2^{-2}, \ldots, h_w^{-2}\right)^T$ and

$$\lambda'_{g_j} \triangleq \begin{cases} \lambda_{g_j} & 1 \leq j \leq m \\ 0 & m < i \leq w. \end{cases} \tag{4.180}$$

Then, we can reformulate (4.179) equivalently into standard form [15] as

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & -\sum_{j=1}^{w} \lambda'_{g_j} \frac{s_j}{\frac{\sigma_h^2}{\sigma_n^2} + s_j} \\ \text{subject to} \quad & \mathbf{b}^T\mathbf{s} - P_0' \leq 0, \\ & -\mathbf{s} \preceq \mathbf{0}, \end{aligned} \tag{4.181}$$

The only difference to problem (4.161) is that the one vector $\mathbf{1}$ is now replaced by the vector $\mathbf{b}$ in the constraint function. However, we can reformulate (4.181), by using scaled variables $s_i' \triangleq b_i s_i = s_i h_i^{-2}$ into an equivalent problem

$$\begin{aligned} \underset{\mathbf{s}'}{\text{minimize}} \quad & -\sum_{j=1}^{w} \lambda_{g_j}^2 \frac{s_j'}{\frac{\sigma_h^2}{h_j^2\sigma_n^2} + s_j'} \\ \text{subject to} \quad & \mathbf{1}^T\mathbf{s}' - P_0' \leq 0, \\ & -\mathbf{s}' \preceq \mathbf{0}, \end{aligned} \tag{4.182}$$

of which we already know the solution for the optimal $\mathbf{s}'$. The resulting optimum ${s_j'}^*$ for $1 \leq j \leq w$ for (4.182), can be obtained by a so called "water-filling" procedure according to (4.163) and (4.164). Thus, with $s_i^* = h_i^2 {s_i'}^*$, we finally obtain the resulting optimum ${s_j}^*$ for $1 \leq j \leq w$ for (4.181) as

$$s_j^* = \max\left\{0, h_j\sqrt{\frac{\sigma_h^2}{\sigma_n^2}\frac{\sigma_{g_j}^2}{\nu^*}} - \frac{\sigma_h^2}{\sigma_n^2}\right\} \tag{4.183}$$

and

$$\sum_{j=1}^{w} \max\left\{0, \sqrt{\frac{\sigma_h^2}{\sigma_n^2 h_j^2}\frac{\sigma_{g_j}^2}{\nu^*}} - \frac{\sigma_h^2}{\sigma_n^2 h_j^2}\right\} = P_0', \tag{4.184}$$

respectively. According to the "water-filling" algorithm (cf. Algorithm A.1), we can compute the optimal singular values of $\widetilde{\mathbf{A}}$ numerically. We can finally compute the optimal singular values σ_j of $\mathbf{A}$ with $\sigma_j^* = \sqrt{s_j^*}$ for $1 \leq j \leq w$, which are then arranged on the main diagonal of the optimal $\mathbf{\Sigma}$ in decreasing order.

So far, we have determined the optimum $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{\Sigma}$ for (4.172) and thus $\widetilde{\mathbf{A}}^* = \mathbf{U}^*\mathbf{\Sigma}^*\mathbf{V}^{*T}$. Now, we are able to compute the optimal local sensor matrix $\mathbf{A}$ for (4.151) as

$$\mathbf{A}^* = \mathbf{H}^{-1}\mathbf{U}^*\mathbf{\Sigma}^*\mathbf{V}^{*T} \overset{(a)}{=} \mathbf{H}^{-1}\mathbf{\Sigma}^*\mathbf{U}_g. \tag{4.185}$$

where in step (a) we inserted the optimal choices $\mathbf{U}^*$ from (4.178) and $\mathbf{V}^*$ from (4.175). The channel matrix $\mathbf{H}^{-1}$ is diagonal and contains their diagonal entries in decreasing order (in terms of magnitude). The unitary $\mathbf{U}_g$ contains the eigenvectors of $\mathbf{G}\mathbf{G}^T$ in decreasing order and the diagona $\mathbf{\Sigma}^*$ contains the optimal singular values of $\widetilde{\mathbf{A}}$, which has to be determined in a "water-filling" like manner (cf. (4.183) and (4.184)).

Conclusions Let us again return to our original index notation that indicates the ith sensor and recall that $\mathbf{C}_{l_i}^* = \mathbf{0}$ for sensor i. For an orthogonal MAC with L local sensors, the T-optimal ith LO - for the standard model - considering constraint (C2) - assuming an invertible channel matrix $\mathbf{H}_i$ - is given by

$$\mathrm{LO}_{i\,\phi_1}^*: \quad \left(\mathbf{A}_i^* = \mathbf{H}_i^{-1}\mathbf{\Sigma}_i^*\mathbf{U}_{g_i},\ \mathbf{C}_{l_i}^* = \mathbf{0}\right), \tag{4.186}$$

where the unitary matrix $\mathbf{U}_{g_i}$ contains the eigenvectors of $\mathbf{G}_i\mathbf{G}_i^T$ sorted decreasingly, the channel matrix $\mathbf{H}_i$ is assumed to be diagonal in decreasing order without loss, the diagonal $\mathbf{\Sigma}_i^*$ has to be determined in a "water-filling" principle in order to balance channel and sensor observation states and noise of sensor i.

4.5.2 A–Optimal Design

We now consider the A–optimality criterion ϕ_{-1}, defined in (3.39). Note that the A–optimality criterion ϕ_{-1} is not linear. Thus, we cannot solve the indidual sensor rules independendly, when we consider an orthogonal MAC. Throughout this thesis, we only consider a single sensor setup, i.e., when $L = 1$. For this special case, we no longer speak about multiple access schemes. Again, we let the subscript notation away and accept the same notation as for the T–optimal design. According to (3.40), an A–optimal designed local sensor rule $\mathrm{LO}^*{}_{\phi_{-1}}$ minimizes the trace of the inverse FIM $\mathbf{J_z}^{-1}$, which is indeed the CRLB, while respecting the constraint (C1) or (C2), respectively. Therefore, we introduce $\widetilde{\phi}_{-1}\{\mathbf{J_z}\} \triangleq 1/\left(\frac{1}{n}\phi_{-1}(\mathbf{J_z})\right)$ and consider the following optimization problem:

$$\begin{aligned} &\underset{\mathbf{A}\in\mathbb{R}^{q\times m}}{\text{minimize}} && \widetilde{\phi}_{-1}(\mathbf{A}) = \mathrm{tr}\left\{\left(\mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I} + \sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G}\right)^{-1}\right\} \\ &\text{subject to} && \begin{cases} \mathrm{tr}\left\{\mathbf{A}\mathbf{M}\mathbf{A}^T\right\} \le P_0/\sigma_n^2 & \text{(C1)} \quad \text{or} \\ \mathrm{tr}\left\{\mathbf{A}\mathbf{A}^T\right\} \le P_0'/\sigma_n^2, & \text{(C2)} \end{cases} \end{aligned} \tag{4.187}$$

where $\mathbf{M}$ is given in (3.28). Note that the A–optimality criterion only applies for a non-singular FIM as can be easily seen. Thus we study the A–optimal design for the case of a positiv definit FIM $\mathbf{J_z}$. The conditions on $\mathbf{H}$, $\mathbf{G}$ and $\mathbf{A}$ can be obtained from (3.23), i.e., $\mathbf{H}$, $\mathbf{G}$ and $\mathbf{A}$ must has at least rank n.

4.5.2.1 Invertible System Matrices

Let us first assume that the observation matrix $\mathbf{G}$, the local sensor matrix $\mathbf{A}$ and the channel matrix $\mathbf{H}$ are all invertable. Thus, we consider squared matrices, where $n = m = q = p$. In particular, that implies $h_i \neq 0$ for $1 \le i \le p$ (cf. (4.148)), since $\mathbf{H}$ is required to be invertible. In what follows, we assume that the diagonal entries h_l for all l are ordered decreasingly in terms of magnitude w.l.o.g. (cf. Section 4.3), i.e., $|h_1| \ge |h_2| \ge \cdots \ge |h_n| > 0$.

The objective $\widetilde{\phi}_{-1}(\mathbf{A})$ of (4.187) can equivalently reformulated as

$$\begin{aligned} \widetilde{\phi}_{-1}(\mathbf{A}) &= \mathrm{tr}\left\{\left(\mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I} + \sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G}\right)^{-1}\right\} \\ &= \mathrm{tr}\left\{\mathbf{G}^{-1}\mathbf{A}^{-1}\mathbf{H}^{-1}\left(\sigma_h^2\mathbf{I} + \sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)\mathbf{H}^{-T}\mathbf{A}^{-T}\mathbf{G}^{-T}\right\} \\ &= \sigma_h^2\mathrm{tr}\left\{\left(\mathbf{G}\mathbf{G}^T\right)^{-1}\mathbf{A}^{-1}\mathbf{H}^{-1}\mathbf{H}^{-T}\mathbf{A}^{-T}\right\} + \\ &\qquad \sigma_n^2\mathrm{tr}\left\{\left(\mathbf{G}\mathbf{G}^T\right)^{-1}\mathbf{A}^{-1}\mathbf{H}^{-1}\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\mathbf{H}^{-T}\mathbf{A}^{-T}\right\} \\ &= \sigma_h^2\mathrm{tr}\left\{\left(\mathbf{G}\mathbf{G}^T\right)^{-1}\mathbf{A}^{-1}\left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{A}^{-T}\right\} + \\ &\qquad \sigma_n^2\mathrm{tr}\left\{\left(\mathbf{G}\mathbf{G}^T\right)^{-1}\right\}, \end{aligned} \tag{4.188}$$

where we used again the cyclic property of the trace operator. The second term of the last equation in (4.188) is constant (i.e., it does not depend on $\mathbf{A}$) and can thus be neglected for solving (4.187). By neclecting also the remaining constant factors, we can reformulate problem (4.187) equivalently as

$$\begin{aligned} &\underset{\mathbf{A}}{\text{minimize}} \quad \operatorname{tr}\left\{\left(\mathbf{G}\mathbf{G}^T\right)^{-1}\mathbf{A}^{-1}\left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{A}^{-T}\right\} \\ &\text{subject to} \quad \begin{cases} \operatorname{tr}\left\{\mathbf{A}\mathbf{M}\mathbf{A}^T\right\} \le P_0 & \text{(C1)} \\ \operatorname{tr}\left\{\mathbf{A}\mathbf{A}^T\right\} \le P_0'/\sigma_n^2. & \text{(C2)} \end{cases} \end{aligned} \tag{4.189}$$

We first solve (4.189) for constraint (C1). The solution for constraint (C2) can then be obtained, by setting $\mathbf{M} = \sigma_n^2\mathbf{I}$ and $P_0 = P_0'$, respectively. First, we introduce the matrix

$$\widetilde{\mathbf{A}} \triangleq \mathbf{H}\mathbf{A}\mathbf{M}^{1/2}, \tag{4.190}$$

where $\mathbf{M}$, given in (3.28), is positive definite and thus also invertible for all $\boldsymbol{\theta}$, since we assumed the standard model, in which $\sigma_n^2\mathbf{I} > 0$. Note that since we assumed $\mathbf{H}$ to be invertible, we can uniquely reclaim the sensor matrix $\mathbf{A}$ from (4.190) by $\mathbf{A} = \mathbf{H}^{-1}\widetilde{\mathbf{A}}\mathbf{M}^{-1/2}$. Inserting $\mathbf{A} = \mathbf{H}^{-1}\widetilde{\mathbf{A}}\mathbf{M}^{-1/2}$ into the objective function of (4.189) yields

$$\begin{aligned} &\operatorname{tr}\left\{\mathbf{M}^{1/2}\left(\mathbf{G}\mathbf{G}^T\right)^{-1}\mathbf{M}^{1/2}\widetilde{\mathbf{A}}^{-1}\mathbf{H}\left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\widetilde{\mathbf{A}}^{-T}\right\} \\ &\qquad\qquad = \operatorname{tr}\left\{\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T\right)^{-1}\left(\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}}\right)^{-1}\right\}, \end{aligned}$$

where we introduced $\widetilde{\mathbf{G}} \triangleq \mathbf{M}^{-1/2}\mathbf{G}$, and into the constraint function (C1) in (4.189) yields

$$\begin{aligned} \operatorname{tr}\left\{\mathbf{A}\mathbf{M}\mathbf{A}^T\right\} &= \operatorname{tr}\left\{\mathbf{H}^{-1}\widetilde{\mathbf{A}}\mathbf{M}^{-1/2}\mathbf{M}\mathbf{M}^{-1/2}\widetilde{\mathbf{A}}^T\mathbf{H}^{-T}\right\} \\ &= \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^T\right\}, \end{aligned}$$

since $\mathbf{M}$ is symmetric and positive definite, i.e., it holds $\mathbf{M} = \mathbf{M}^{1/2}\mathbf{M}^{1/2}$. We can equivalently reformulate problem (4.189), for constraint (C1) as

$$\begin{aligned} &\underset{\widetilde{\mathbf{A}}}{\text{minimize}} \quad \operatorname{tr}\left\{\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T\right)^{-1}\left(\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}}\right)^{-1}\right\} \\ &\text{subject to} \quad \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^T\right\} \le P_0, \quad \text{(C1)} \end{aligned} \tag{4.191}$$

which now has to be solved with respect to $\widetilde{\mathbf{A}}$. Using the SVD $\widetilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, where we assume that the singular values σ_i for $1 \le i \le n$ of $\widetilde{\mathbf{A}}$ are ordered decreasingly (in terms of magnitude) on the main diagonal of $\boldsymbol{\Sigma}$, i.e.,

$|\sigma_1| \geq |\sigma_2| \geq \cdots \geq |\sigma_n| > 0$. Both unitary matrices $\mathbf{U}$ and $\mathbf{V}$, contain the corresponding left- and right singular vectors of $\widetilde{\mathbf{A}}$, respectively. Problem (4.191) can then be written in terms of SVD $\widetilde{\mathbf{A}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ as

$$\begin{aligned} \underset{\mathbf{U},\mathbf{\Sigma},\mathbf{V}}{\text{minimize}} \quad & \operatorname{tr}\left\{\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T\right)^{-1}\mathbf{V}\mathbf{\Sigma}^{-2}\mathbf{V}^T\right\} \\ \text{subject to} \quad & \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^T\right\} \leq P_0, \quad \text{(C1)} \\ & \mathbf{\Sigma}^2 \geq \mathbf{0}, \\ & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}. \end{aligned} \tag{4.192}$$

In what follows, we will solve (4.192) sequentially, by determining first the optimal $\mathbf{V}$ and then the optimal $\mathbf{U}$. A necessary condition for $\mathbf{V}$ to be optimum in (4.192) can be obtained by fixing $\mathbf{U}$ and $\mathbf{\Sigma}$. For a fixed $\mathbf{U}'$ and $\mathbf{\Sigma}'$, the optimum $\mathbf{V}$ has to solve the problem

$$\begin{aligned} \underset{\mathbf{V}}{\text{minimize}} \quad & \operatorname{tr}\left\{\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T\right)^{-1}\mathbf{V}\mathbf{\Sigma}'^{-2}\mathbf{V}^T\right\} \\ \text{subject to} \quad & \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{U}'\mathbf{\Sigma}'^2\mathbf{U}'^T\right\} \leq P_0, \quad \text{(C1)} \\ & \mathbf{\Sigma}'^2 \geq \mathbf{0}, \\ & \mathbf{U}'\mathbf{U}'^T = \mathbf{U}'^T\mathbf{U}' = \mathbf{I}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \end{aligned} \tag{4.193}$$

A $\mathbf{V}$ is optimum for (4.193) if and only if it is optimum for

$$\begin{aligned} \underset{\mathbf{V}}{\text{minimize}} \quad & \operatorname{tr}\left\{\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T\right)^{-1}\mathbf{V}\mathbf{\Sigma}'^{-2}\mathbf{V}^T\right\} \\ \text{subject to} \quad & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}. \end{aligned} \tag{4.194}$$

According to Theorem 4.4.19, the optimal $\mathbf{V}$, solving (4.194) is given by the matrix $\mathbf{U}_{\widetilde{g}}$, containing the eigenvectors of $\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T$, sorted increasingly. So, the EVD $\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T = \mathbf{U}_{\widetilde{g}}\mathbf{\Lambda}_{\widetilde{g}}\mathbf{U}_{\widetilde{g}}^T$, where the eigenvalues $\lambda_{\widetilde{g}_i}$ for $1 \leq i \leq n$ of $\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T$ are ordered increasingly along the main diagonal in $\mathbf{\Lambda}_{\widetilde{g}}$. Note that $\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T\right)^{-1} = \mathbf{U}_{\widetilde{g}}\mathbf{\Lambda}_{\widetilde{g}}^{-1}\mathbf{U}_{\widetilde{g}}^T$ and thus the eigenvalues of $\left(\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T\right)^{-1}$ are now ordered decreasingly along the main diagonal of $\mathbf{\Lambda}_{\widetilde{g}}^{-1}$, due to the inverse operation. Thus, the optimal $\mathbf{V}$ for problem (4.194) and thus also for (4.192), is given by

$$\mathbf{V}^* = \mathbf{U}_{\widetilde{g}}, \tag{4.195}$$

where the column vectors of $\mathbf{U}_{\widetilde{g}}$ correspond to the eigenvectors of $\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T$, sorted increasingly.

After having determined the optimal $\mathbf{V}$ for (4.192), we will now characterize the optimum $\mathbf{U}$ for (4.192). To this end, we insert the optimal $\mathbf{V}$ into (4.192), yielding:

$$\begin{aligned}
\underset{\mathbf{U},\boldsymbol{\Sigma}}{\text{minimize}} \quad & \operatorname{tr}\left\{\boldsymbol{\Lambda}_{\tilde{g}}^{-1}\boldsymbol{\Sigma}^{-2}\right\} \\
\text{subject to} \quad & \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^{T}\right)^{-1}\mathbf{U}\boldsymbol{\Sigma}^{2}\mathbf{U}^{T}\right\} \leq P_0, \quad \text{(C1)} \\
& \boldsymbol{\Sigma}^{2} \geq \mathbf{0}, \\
& \mathbf{U}\mathbf{U}^{T} = \mathbf{U}^{T}\mathbf{U} = \mathbf{I}.
\end{aligned} \tag{4.196}$$

If $\mathbf{U}^*$ is optimal for (4.196), then it is also optimal for (4.192), as can be verified easily.

Invoking problem (4.124), we note that it has the same constraint function (more specifically, the pseudo-inverse specializes now to the matrix inverse) and a fairly similar objective. Where we have showed that the optimal $\mathbf{U}$ is determined by minimization of the constraint function

$$t\left(\mathbf{U},\boldsymbol{\Sigma}\right) \triangleq \operatorname{tr}\left\{\mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{T}\mathbf{U}^{T}\left(\mathbf{H}\mathbf{H}^{T}\right)^{-1}\right\}.$$

We can closely follow the approach and recognize that this also applies to (4.196). Thus, we can determine the optimal $\mathbf{U}$ by solving

$$\begin{aligned}
\underset{\mathbf{U}}{\text{minimize}} \quad & \operatorname{tr}\left\{\left(\mathbf{H}\mathbf{H}^{T}\right)^{-1}\mathbf{U}\boldsymbol{\Sigma}'^{2}\mathbf{U}^{T}\right\} \\
\text{subject to} \quad & \mathbf{U}\mathbf{U}^{T} = \mathbf{U}^{T}\mathbf{U} = \mathbf{I},
\end{aligned} \tag{4.197}$$

i.e., the optimal $\mathbf{U}$ will be a minimizer of the constraint function $t\left(\mathbf{U},\boldsymbol{\Sigma}\right)$, for the specific choice of $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$. An application of Theorem 4.4.19 reveals, that the optimal $\mathbf{U}$ is given by the eigenvectors of $\left(\mathbf{H}\mathbf{H}^{T}\right)^{-1}$ or $\mathbf{H}\mathbf{H}^{T}$, respectively, in the order of increasing eigenvalues of $\left(\mathbf{H}\mathbf{H}^{T}\right)^{-1}$, or, decreasing eigenvalues of $\mathbf{H}\mathbf{H}^{T}$. However, since $\mathbf{H}\mathbf{H}^{T}$ and in turn $\left(\mathbf{H}\mathbf{H}^{T}\right)^{-1}$ is diagonal, the eigenvectors are given by the unit vectors $\{\mathbf{e}_k\}_{k=1}^{n}$. Moreover, since the eigenvalues of $\left(\mathbf{H}\mathbf{H}^{T}\right)^{-1}$ are the squared reciprocals of the diagonal values $h_j \neq 0$, we have that the kth column of $\mathbf{U}^*$ is given by $\mathbf{e}_{j_k}$, where j_k is the index of the kth largest main diagonal entry h_{j_k}. Since, we assumed that the diagonal entries of $\mathbf{H}$ are in decreasing order (in terms of magnitude), the optimal $\mathbf{U}$ is given by

$$\mathbf{U}^* = \mathbf{I}, \tag{4.198}$$

So far, we have determined the optimum $\mathbf{U}$ and $\mathbf{V}$ for (4.192). It remains to determine the optimal singular values σ_i for $i = 1, \ldots, n$. Inserting the optimal choices $\mathbf{U}^*$ and $\mathbf{V}^*$ into (4.192), yields to the optimization problem in standard form [15] as

$$\begin{aligned}
\underset{\mathbf{s}}{\text{minimize}} \quad & f(\mathbf{s}) \triangleq \sum_{i=1}^{n} \frac{1}{\lambda_{\widetilde{g}_i} s_i} \\
\text{subject to} \quad & g_1(\mathbf{s}) \triangleq \mathbf{b}^T \mathbf{s} - P_0 \leq 0 \\
& g_2(\mathbf{s}) \triangleq -\mathbf{s} \prec \mathbf{0},
\end{aligned} \tag{4.199}$$

with the new introduced objective function $f(\mathbf{s})$, where the vector $\mathbf{s}$, contains the squared singular values of $\widetilde{\mathbf{A}}$, i.e., $\mathbf{s} = (s_1, s_2, \ldots, s_n)^T \triangleq (\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)^T \in \mathbb{R}^{+n}$, i.e., $\mathbf{s} \succeq \mathbf{0}$. The new introduced vector $\mathbf{b}$ in (4.199) contains the reciprocals of the squared diagonal entries of $\mathbf{H}$, i.e., $\mathbf{b} \triangleq (h_1^{-2}, h_2^{-2}, \ldots, h_n^{-2})$, since $\mathbf{H}$ is assumed to be diagonal in decreasing order. Note that $\lambda_{\widetilde{g}_i} > 0$ and $h_i^2 > 0$ for all i, due to our assumption that $\mathbf{G}$, $\mathbf{H}$ and in turn $\widetilde{\mathbf{G}}$ are invertible.

We first verify, that (4.199) is a convex optimization problem. We can write the objective $f(\mathbf{s})$ of problem (4.199) as

$$f(\mathbf{s}) = \sum_{i=1}^{n} f_i(\mathbf{s}_i), \tag{4.200}$$

with $f_i(\mathbf{s}_i) \triangleq \frac{1}{\lambda_{\widetilde{g}_i} s_i}$. The first two derivatives of $f_i(\mathbf{s}_i)$ are given by

$$f_i'(s_i) \triangleq \frac{\partial}{\partial s_i} f(s_i) = -\frac{1}{\lambda_{\widetilde{g}_i} s_i^2} \tag{4.201}$$

and

$$f_i''(s_i) \triangleq \frac{\partial^2}{\partial s_i^2} f(s_i) = 2\frac{1}{\lambda_{\widetilde{g}_i} s_i^3} > 0, \tag{4.202}$$

and therefore $f_i(s_i)$ is convex. Since by (4.200), the objective is a sum of convex functions [15], we conclude $f(\mathbf{s})$ is convex. The convexity of the constraint functions g_1 and g_2 are obvious, since both are linear in $\mathbf{s}$. Hence, problem (4.199) is a convex optimization problem [15].

The Karush–Kuhn–Tucker (KKT) conditions (cf. [15]) for a solution $\mathbf{s}^*$ to the optimization problem (4.199) and corresponding Lagrange multipliers (cf. [15, p.244]), i.e., ν^* for the inequality constraint $g_1(\mathbf{s}) \leq 0$ and $\boldsymbol{\lambda}^* \in \mathbb{R}^n$

for the inequality constraint $g_2(\mathbf{s}) \prec \mathbf{0}$, are given as

$$\begin{aligned} \mathbf{b}^T\mathbf{s}^* - P_0 &\le 0 \\ \mathbf{s}^* &\succ 0 \\ \nu^* &\ge 0 \\ \boldsymbol{\lambda}^* &= 0 \\ \nu^*\left(\mathbf{b}^T\mathbf{s}^* - P_0\right) &= 0 \\ \lambda_i^* s_i^* &= 0,\ i = 1,2,\dots,n \\ -\frac{1}{\lambda_{\widetilde{g}_i} s_i^{*2}} + \nu^* - \lambda_i^* &= 0,\ i = 1,2,\dots,n. \end{aligned} \tag{4.203}$$

Combining the 4th and the last condition of (4.203), i.e., with $\lambda_i^* = 0$ for all $i = 1,\dots,n$, yields

$$\nu^* = \frac{1}{\lambda_{\widetilde{g}_i} s_i^{*2}} > 0 \quad \Rightarrow \quad s_i^* = \sqrt[+]{\frac{1}{\lambda_{\widetilde{g}_i}\nu^*}} > 0 \quad \text{for } i = 1,\dots,n$$

and by the 5th condition of (4.203), i.e., $\mathbf{b}^T\mathbf{s}^* = P_0$, since $\nu^* \neq 0$, we obtain

$$\nu^* = \frac{\left(\sum_{i=1}^n \frac{1}{h_i^2 \sqrt[+]{\lambda_{\widetilde{g}_i}}}\right)^2}{P_0^2}.$$

Therefore,

$$s_i^* = \frac{P_0}{\sum_{j=1}^n \frac{1}{h_j^2 \sqrt[+]{\lambda_{\widetilde{g}_j}}}} \sqrt[+]{\frac{1}{\lambda_{\widetilde{g}_i}}} \quad \text{for } i = 1,\dots,n \tag{4.204}$$

and finally, we obtain the optimum singular values of $\widetilde{\mathbf{A}}$ with $(\boldsymbol{\Sigma}^*)_{i,i} = \sigma_i^* = \sqrt{s_i^*}$, i.e., the optimal $\boldsymbol{\Sigma}$ can be computed in closed-form with

$$c^*_{(\mathrm{C1})} \triangleq \sqrt{\frac{P_0}{\sum_{j=1}^n \frac{1}{h_j^2 \sqrt[+]{\lambda_{\widetilde{g}_j}}}}}, \tag{4.205}$$

as

$$\boldsymbol{\Sigma}^* = c^*_{(\mathrm{C1})} \boldsymbol{\Lambda}_{\widetilde{g}}^{-1/4}. \tag{4.206}$$

So far, we have determined the optimum $\mathbf{U}$, $\mathbf{V}$ and $\boldsymbol{\Sigma}$ for (4.192) and thus optimum $\widetilde{\mathbf{A}}$ for (4.191), i.e., $\widetilde{\mathbf{A}}^* = \mathbf{U}^*\boldsymbol{\Sigma}^*\mathbf{V}^{*T}$. The unitary matrix $\mathbf{U}^* = \mathbf{I}$ (cf. (4.198)), since the diagonal $\mathbf{H}$ is assumed to be in decreasing order. The unitary matrix $\mathbf{V}^* = \mathbf{U}_{\widetilde{g}}$ (cf. (4.195)), where $\mathbf{U}_{\widetilde{g}}$ contains the eigenvectors of $\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T$, in increasing order and $\boldsymbol{\Sigma}^* = c^*_{(\mathrm{C1})}\boldsymbol{\Lambda}_{\widetilde{g}}^{-1/4}$ (cf. (4.206)), where the eigenvalues of $\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T$ in $\mathbf{\Lambda}_{\tilde{g}}$ are ordered increasingly. From (4.190), we obtain the optimal $\mathbf{A}$ by

$$\mathbf{A}^* = \mathbf{H}^{-1}\mathbf{U}^*\mathbf{\Sigma}^*\mathbf{V}^{*T}\mathbf{M}^{-1/2} = c^*_{(\mathrm{C1})}\mathbf{H}^{-1}\mathbf{\Lambda}_{\tilde{g}}^{-1/4}\mathbf{U}_{\tilde{g}}{}^T\mathbf{M}^{-1/2}$$

and in turn with the already known solution $\mathbf{C}_l^* = \mathbf{0}$, the optimal $\mathrm{LO}^*_{\phi_{-1}} = (\mathbf{A}^*,\ \mathbf{C}_l^*)$ is thus given by

$$\begin{aligned}
&\mathrm{LO}^*_{\phi_{-1}}\!: \quad \Big(\mathbf{A}^* = c^*_{(\mathrm{C1})}\mathbf{H}^{-1}\mathbf{\Lambda}_{\tilde{g}}^{-1/4}\mathbf{U}_{\tilde{g}}{}^T\mathbf{M}^{-1/2},\ \mathbf{C}_l^* = \mathbf{0}\Big)_{\phi_{-1}}\\
&\quad \text{for constraint (C1)}, \quad c^*_{(\mathrm{C1})} \text{ is given in (4.205)},\\
&\quad \text{EVD: } \widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T = \mathbf{U}_{\tilde{g}}\mathbf{\Lambda}_{\tilde{g}}\mathbf{U}_{\tilde{g}}^T,\\
&\qquad\quad \text{where } \mathbf{\Lambda}_{\tilde{g}},\ \mathbf{U}_{\tilde{g}} \text{ are sorted increasingly},\\
&\quad \mathbf{H} \text{ is sorted decreasingly.}
\end{aligned} \tag{4.207}$$

Let us recall $\mathbf{M}$, given in (3.28), i.e., with $\mathbf{C}_n = \sigma_n^2\mathbf{I}$ it follows $\mathbf{M} = \mathbf{G}\boldsymbol{\theta}\boldsymbol{\theta}^T\mathbf{G}^T + \sigma_n^2\mathbf{I}$. Thus, the optimal $\mathrm{LO}_{\phi_{-1}}$ for constraint (C1), depends on the parameter $\boldsymbol{\theta}$, which is indeed unknown.

Note that we have determined the optimal $\mathrm{LO}^*_{\phi_{-1}}$ for the constraint (C1). As already mentioned, we can determine the optimal $\mathrm{LO}^*_{\phi_{-1}}$ for constraint (C2), if we set $\mathbf{M} = \sigma_n^2\mathbf{I}$, $P_0 = P_0'$ and consequently $\widetilde{\mathbf{G}} = \mathbf{G}$ in (4.207), i.e., with

$$c^*_{(\mathrm{C2})} \triangleq \sqrt{\frac{P_0'/\sigma_n^2}{\sum_{j=1}^{n} \frac{1}{h_j^2 \sqrt[4]{\lambda_{g_j}}}}}, \tag{4.208}$$

as

$$\begin{aligned}
&\mathrm{LO}^*_{\phi_{-1}}\!: \quad \Big(\mathbf{A}^* = c^*_{(\mathrm{C2})}\mathbf{H}^{-1}\mathbf{\Lambda}_{g}^{-1/4}\mathbf{U}_{g}{}^T,\ \mathbf{C}_l^* = \mathbf{0}\Big)_{\phi_{-1}},\\
&\quad \text{for constraint (C2)}, \quad c^*_{(\mathrm{C2})} \text{ is given in (4.208)},\\
&\quad \text{EVD: } \mathbf{G}\mathbf{G}^T = \mathbf{U}_{g}\mathbf{\Lambda}_{g}\mathbf{U}_{g}^T,\\
&\qquad\quad \text{where } \mathbf{\Lambda}_{g},\ \mathbf{U}_{g} \text{ are sorted increasingly},\\
&\quad \mathbf{H} \text{ is sorted decreasingly.}
\end{aligned} \tag{4.209}$$

However, since we have found a closed-form solution for the A-optimal ith LO, we can still specify the resulting FIM $\mathbf{J}_{\mathbf{z}}^*$. To that end, we insert $\mathbf{A}^*$ from (4.207) into (4.149) (without the subscript notation) for $\mathbf{A} = \mathbf{A}^*$, yielding for constraint (C1):

$$
\begin{aligned}
\mathbf{J}_{\mathbf{z}}^{*} &= \mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I}+\sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G}\\
&\overset{(a)}{=}\Bigg(c^{*2}_{(\mathrm{C1})}\widetilde{\mathbf{G}}^T\mathbf{U}_{\widetilde{g}}\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\left(\sigma_h^2\mathbf{I}+\sigma_n^2c^{*2}_{(\mathrm{C1})}\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\mathbf{U}_{\widetilde{g}}{}^T\mathbf{M}^{-1}\mathbf{U}_{\widetilde{g}}\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\right)^{-1}\cdot\\
&\qquad\qquad\cdot\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\mathbf{U}_{\widetilde{g}}{}^T\widetilde{\mathbf{G}}\Bigg)\\
&\overset{(b)}{=}\mathbf{V}_{\widetilde{g}}\mathbf{\Sigma}_{\widetilde{g}}\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\left(\frac{\sigma_h^2}{c^{*2}_{(\mathrm{C1})}}\mathbf{I}+\sigma_n^2\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\mathbf{U}_{\widetilde{g}}{}^T\mathbf{M}^{-1}\mathbf{U}_{\widetilde{g}}\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\right)^{-1}\mathbf{\Lambda}_{\widetilde{g}}^{-1/4}\mathbf{\Sigma}_{\widetilde{g}}\mathbf{V}_{\widetilde{g}}^T,
\end{aligned}
\tag{4.210}
$$

where in step (a), we inserted $\mathbf{A}=\mathbf{A}^*$ from (4.207) and also $\widetilde{\mathbf{G}}=\mathbf{M}^{-1/2}\mathbf{G}$; in step (b), we performed the SVD $\widetilde{\mathbf{G}}=\mathbf{U}_{\widetilde{g}}\mathbf{\Sigma}_{\widetilde{g}}\mathbf{V}_{\widetilde{g}}^T$ - sorted increasingly in terms of magnitude, according to the EVD $\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^T=\mathbf{U}_{\widetilde{g}}\mathbf{\Lambda}_{\widetilde{g}}\mathbf{U}_{\widetilde{g}}^T$. Analog, we can still specify the resulting FIM $\mathbf{J}_{\mathbf{z}}^*$ for constraint (C2) with SVD $\mathbf{G}=\mathbf{U}_g\mathbf{\Sigma}_g\mathbf{V}_g^T$ (sorted increasingly):

$$
\begin{aligned}
\mathbf{J}_{\mathbf{z}}^{*} &= \mathbf{V}_g\mathbf{\Sigma}_g\mathbf{\Lambda}_g^{-1/4}\left(\frac{\sigma_h^2}{c^{*2}_{(\mathrm{C2})}}\mathbf{I}+\sigma_n^2\mathbf{\Lambda}_g^{-1/4}\mathbf{\Lambda}_g^{-1/4}\right)^{-1}\mathbf{\Lambda}_g^{-1/4}\mathbf{\Sigma}_g\mathbf{V}_g^T.\\
&= \mathbf{V}_g\mathbf{\Sigma}_g\mathbf{\Lambda}_g^{-1/2}\mathbf{\Sigma}_g\left(\frac{\sigma_h^2}{c^{*2}_{(\mathrm{C2})}}\mathbf{I}+\sigma_n^2\mathbf{\Lambda}_g^{-1/2}\right)^{-1}\mathbf{V}_g^T\\
&\overset{(a)}{=}\mathbf{V}_g\mathbf{\Lambda}_g^{1/2}\left(\frac{\sigma_h^2}{c^{*2}_{(\mathrm{C2})}}\mathbf{I}+\sigma_n^2\mathbf{\Lambda}_g^{-1/2}\right)^{-1}\mathbf{V}_g^T,
\end{aligned}
\tag{4.211}
$$

where in step (a), we used the fact that $\mathbf{\Lambda}_g=\mathbf{\Sigma}_g^2$.

4.5.2.2 Full Column–Rank Channel Matrix

We now consider only constraint (C2). Note that we already assumed that observation-, channel- and sensor matrix $\mathbf{G}$, $\mathbf{H}$ and $\mathbf{A}$ has at least rank n (cf. (3.23) in order to obtain a non-singular FIM $\mathbf{J}_{\mathbf{z}}$). In particular, we assume that $\mathbf{H}$ is of full column-rank, implying $p \geq q$ and also $h_i \neq 0$ for $1 \leq i \leq q$ (cf. (4.148)). In what follows, we assume that the diagonal entries h_l for $1 \leq l \leq q$ are ordered increasingly in terms of magnitude w.l.o.g. (cf. Section 4.3), i.e., $0 < |h_1| \leq |h_2| \leq \cdots \leq |h_q|$. Finally, we assume that the channel input dimension is equal to the the parameter dimension, i.e., $n = q$.

Again, we first substitude the local sensor matrix $\mathbf{A}$ into the channel matrix $\mathbf{H}$, i.e., we introduce

$$
\widetilde{\mathbf{A}} \triangleq \mathbf{H}\mathbf{A}. \tag{4.212}
$$

Performing the SVD $\widetilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, where we again assume that the singular values σ_i for $1 \leq i \leq w \triangleq \min\{p, m\}$ of $\widetilde{\mathbf{A}}$ are ordered decreasingly on the main diagonal of $\boldsymbol{\Sigma}$ (in terms of magnitude), i.e., $|\sigma_1| \geq |\sigma_2| \geq \cdots \geq |\sigma_w| \geq 0$. Both unitary matrices $\mathbf{U}$ and $\mathbf{V}$, contain the corresponding left- and right singular vectors of $\widetilde{\mathbf{A}}$, respectively. Invoking the derivation from (4.168). We can fully accept the derivation of the FIM

$$\mathbf{J_z} = \mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I} + \sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G}$$

in terms of SVD $\widetilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, i.e.,

$$\mathbf{J_z} = \mathbf{G}^T\mathbf{A}^T\mathbf{H}^T\left(\sigma_h^2\mathbf{I} + \sigma_n^2\mathbf{H}\mathbf{A}\mathbf{A}^T\mathbf{H}^T\right)^{-1}\mathbf{H}\mathbf{A}\mathbf{G} = \frac{1}{\sigma_n^2}\mathbf{G}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{G},$$

where $\mathbf{D}$ is given by

$$\mathbf{D} \triangleq \boldsymbol{\Sigma}^T\left(\sigma_h^2\mathbf{I} + \sigma_n^2\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\right)^{-1}\boldsymbol{\Sigma}, \tag{4.213}$$

which is squared and diagonal of size $m \times m$. The m elements on the main diagonal are given by

$$d_i \triangleq (\mathbf{D})_{i,i} = \begin{cases} \frac{\sigma_i^2}{\frac{\sigma_h^2}{\sigma_n^2} + \sigma_i^2} & 1 \leq i \leq w \\ 0 & \text{else.} \end{cases} \tag{4.214}$$

Hence, the objective function $\widetilde{\phi}_{-1}(\mathbf{A})$ of problem (4.187), can then be equivalently reformulated in terms of SVD $\widetilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ as

$$\begin{aligned} \widetilde{\phi}_{-1}(\boldsymbol{\Sigma}, \mathbf{V}) &= \sigma_n^2 \mathrm{tr}\left\{\left(\mathbf{G}^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{G}\right)^{-1}\right\} \\ &\overset{(a)}{=} \sigma_n^2 \mathrm{tr}\left\{\mathbf{V}_g\left(\boldsymbol{\Sigma}_g^T\mathbf{U}_g^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{U}_g\boldsymbol{\Sigma}_g\right)^{-1}\mathbf{V}_g^T\right\} \\ &= \sigma_n^2 \mathrm{tr}\left\{\left(\boldsymbol{\Sigma}_g^T\mathbf{U}_g^T\mathbf{V}\mathbf{D}\mathbf{V}^T\mathbf{U}_g\boldsymbol{\Sigma}_g\right)^{-1}\right\}, \end{aligned} \tag{4.215}$$

where in step (a) we inserted the SVD $\mathbf{G} = \mathbf{U}_g\boldsymbol{\Sigma}_g\mathbf{V}_g^T$. As can be seen, the unitary matrix $\mathbf{V}_g$ vanishes in (4.215), due to the cyclic properity of the trace operator. Note that $\mathrm{rank}\left(\widetilde{\mathbf{A}}\right) = n$, which follows from our assumption that $q = n$ and the conditions for a non-singular FIM $\mathbf{J_z}$ from (3.23). Thus we can write (4.215) also in a partitioned form, with

$$\mathbf{U}_g\boldsymbol{\Sigma}_g = \begin{bmatrix}\mathbf{U}_{g,1} & \mathbf{U}_{g,2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_{g,1} \\ \mathbf{0}\end{bmatrix} = \mathbf{U}_{g,1}\boldsymbol{\Sigma}_{g,1},$$

where $\boldsymbol{\Sigma}_{g,1}$ is squared of size $n \times n$, containing the n non-zero singular values of $\mathbf{G}$ and $\mathbf{U}_{g,1}$ contains the corresponding n left singular vectors (submatrix of the unitary matrix $\mathbf{U}_g$), and with

$$\mathbf{V}\mathbf{D}\mathbf{V}^T = \begin{bmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{bmatrix}\begin{bmatrix}\mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_1^T \\ \mathbf{V}_2^T\end{bmatrix} = \mathbf{V}_1\mathbf{D}_1\mathbf{V}_1^T$$

as

$$\begin{aligned}\widetilde{\phi}_{-1}(\mathbf{\Sigma},\mathbf{V}) &= \sigma_n^2 \mathrm{tr}\left\{\left(\mathbf{\Sigma}_{g,1}\mathbf{U}_{g,1}^T\mathbf{V}_1\mathbf{D}_1\mathbf{V}_1^T\mathbf{U}_{g,1}\mathbf{\Sigma}_{g,1}\right)^{-1}\right\} \\ &= \sigma_n^2 \mathrm{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\left(\mathbf{U}_{g,1}^T\mathbf{V}_1\right)^{-1}\mathbf{D}_1^{-1}\left(\mathbf{V}_1^T\mathbf{U}_{g,1}\right)^{-1}\right\}.\end{aligned} \tag{4.216}$$

Note that the n column vectors of $\mathbf{U}_{g,1}$ are also the n eigenvectors of $\mathbf{G}\mathbf{G}^T$, which corresponds to the n non-zero eigenvalues of $\mathbf{G}\mathbf{G}^T$, in particular. The submatrix $\mathbf{V}_1$ of $\mathbf{V}$, contains the first n right singular vectors, which corresponds to the n non-zero singular values of $\mathbf{A}$, already assumed to be ordered decreasingly. The diagonal $\mathbf{D}_1$ is the submatrix of $\mathbf{D}$ from (4.213), which contains the n non-zero diagonal entries d_i for $1 \le i \le n$ of $\mathbf{D}$, in decreasing order, as a consequence of the order in $\mathbf{\Sigma}$.

We can rewrite the constraint of (4.187) in terms of SVD $\widetilde{\mathbf{A}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, analog to (4.171), as

$$\mathrm{tr}\left\{\mathbf{A}\mathbf{A}^T\right\} = \mathrm{tr}\left\{\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}\right\} \le \frac{P_0'}{\sigma_n^2}, \tag{4.217}$$

where $\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}$ denotes the pseudo inverse of $\mathbf{H}\mathbf{H}^T$. In turn, the optimization problem (4.187) in terms of SVD $\widetilde{\mathbf{A}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ with (4.216) and (4.217) then yields

$$\begin{aligned} \underset{\mathbf{U},\mathbf{\Sigma},\mathbf{V}}{\text{minimize}} \quad & \mathrm{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\left(\mathbf{U}_{g,1}^T\mathbf{V}_1\right)^{-1}\mathbf{D}_1^{-1}\left(\mathbf{V}_1^T\mathbf{U}_{g,1}\right)^{-1}\right\} \\ \text{subject to} \quad & \mathrm{tr}\left\{\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}\right\} \le \frac{P_0'}{\sigma_n^2}, \\ & \mathbf{\Sigma}\mathbf{\Sigma}^T \ge \mathbf{0}, \\ & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \\ & \mathbf{V} = \begin{bmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{bmatrix}, \end{aligned} \tag{4.218}$$

where we neglect the constant factor σ_n^2 of the objective function in (4.215). In what follows, we will solve (4.218) sequentially, by determining first the optimum $\mathbf{V}$ and then the optimum $\mathbf{U}$.

Theorem 4.5.21 *Let $\mathbf{A}$ be a $n \times n$ real matrix with singular values $\sigma_1(\mathbf{A}) \ge \sigma_2(\mathbf{A}) \cdots \ge \sigma_n(\mathbf{A})$ and $\mathbf{B}$ be a $(n-k)\times(n-k)$ submatrix of $\mathbf{A}$ obtained by deleting a total of k rows and columns from $\mathbf{A}$, with singular values $\sigma_1(\mathbf{B}) \ge \sigma_2(\mathbf{B}) \cdots \ge \sigma_{n-k}(\mathbf{B})$, then*

$$\sigma_j(\mathbf{A}) \ge \sigma_j(\mathbf{B}) \ge \sigma'_{j+k}(\mathbf{A}) \quad \textit{for } j = 1, \ldots, n,$$

where

$$\sigma'_j(\mathbf{A}) = \begin{cases}\sigma_j(\mathbf{A}) & j \le n \\ 0 & \textit{else.}\end{cases}$$

.

Proof. Cf. [17, Corollary 3.1.3, p. 149]. □

A necessary condition for $\mathbf{V}$ to be optimum in (4.218) can be obtained by fixing $\mathbf{U}$ and $\mathbf{\Sigma}$. For a fixed $\mathbf{U}'$ and $\mathbf{\Sigma}'$, the optimum $\mathbf{V}$ has to solve the problem

$$\begin{aligned} \underset{\mathbf{V}}{\text{minimize}} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\left(\mathbf{U}_{g,1}^T\mathbf{V}_1\right)^{-1}\mathbf{D}_1'^{-1}\left(\mathbf{V}_1^T\mathbf{U}_{g,1}^T\right)^{-1}\right\} \\ \text{subject to} \quad & \operatorname{tr}\left\{\mathbf{U}'\mathbf{\Sigma}'\mathbf{\Sigma}'^T\mathbf{U}'^T\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}\right\} \le \frac{P_0'}{\sigma_n^2}, \\ & \mathbf{\Sigma}'\mathbf{\Sigma}'^T \ge \mathbf{0}, \\ & \mathbf{U}'\mathbf{U}'^T = \mathbf{U}'^T\mathbf{U}' = \mathbf{I}, \\ & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \\ & \mathbf{V} = \begin{bmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{bmatrix}, \end{aligned} \tag{4.219}$$

where $\mathbf{D}'$ and thus $\mathbf{D}_1'$, are obtained from (4.213) by inserting $\mathbf{\Sigma}'$ for $\mathbf{\Sigma}$. A $\mathbf{V}$ is optimum for (4.219), if and only if it is optimum for

$$\begin{aligned} \underset{\mathbf{V}}{\text{minimize}} \quad & \operatorname{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\left(\mathbf{U}_{g,1}^T\mathbf{V}_1\right)^{-1}\mathbf{D}_1'^{-1}\left(\mathbf{V}_1^T\mathbf{U}_{g,1}^T\right)^{-1}\right\} \\ \text{subject to} \quad & \mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}, \\ & \mathbf{V} = \begin{bmatrix}\mathbf{V}_1 & \mathbf{V}_2\end{bmatrix}, \end{aligned} \tag{4.220}$$

Note that $\mathbf{V}_1^T\mathbf{V}_1 = \mathbf{I}$, but $\mathbf{V}_1\mathbf{V}_1^T \neq \mathbf{I}$. Let us introduce $\mathbf{Z}_1 \triangleq \mathbf{U}_{g,1}^T\mathbf{V}_1$ of size $n \times n$. It is easy to verfiy that $\mathbf{Z}_1$ is a submatrix of the unitary $\mathbf{Z} \triangleq \mathbf{U}_g^T\mathbf{V}$ of size $m \times m$ (we assume $n \le m$). Performing the SVD $\mathbf{Z}_1 = \mathbf{U}_{z_1}\mathbf{\Sigma}_{z_1}\mathbf{V}_{z_1}^T$, where the unitaries $\mathbf{U}_{z_1}$ and $\mathbf{V}_{z_1}$ contain the left- and right singular vectors and the diagonal $\mathbf{\Sigma}_{z_1}$ contains the singular values of $\mathbf{Z}_1$, on the main diagonal. All singular values of $\mathbf{Z}$ are equal to one according to Lemma 4.5.20. Invoking Theorem 4.5.21, we thus conclude that

$$0 < \sigma_j\left(\mathbf{Z}_1\right) \le 1 \quad \text{for } j = 1, \dots, n, \tag{4.221}$$

where $\sigma_j\left(\mathbf{Z}_1\right)$ denotes the singular values of the submatrix $\mathbf{Z}_1$. Hence, the objective function of (4.220), can be bounded below by

$$\begin{aligned} \operatorname{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\mathbf{Z}_1^{-1}\mathbf{D}_1'^{-1}\mathbf{Z}_1^{-T}\right\} &= \operatorname{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\mathbf{V}_{z_1}\mathbf{\Sigma}_{z_1}^{-1}\mathbf{U}_{z_1}^T\mathbf{D}_1'^{-1}\mathbf{U}_{z_1}\mathbf{\Sigma}_{z_1}^{-1}\mathbf{V}_{z_1}^T\right\} \\ &\overset{(a)}{\ge} \operatorname{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\mathbf{V}_{z_1}\mathbf{U}_{z_1}^T\mathbf{D}_1'^{-1}\mathbf{U}_{z_1}\mathbf{\Sigma}_{z_1}^{-1}\mathbf{V}_{z_1}^T\right\} \\ &\overset{(b)}{\ge} \operatorname{tr}\left\{\mathbf{\Sigma}_{g,1}^{-2}\mathbf{V}_{z_1}\mathbf{U}_{z_1}^T\mathbf{D}_1'^{-1}\mathbf{U}_{z_1}\mathbf{V}_{z_1}^T\right\} \\ &= \operatorname{tr}\left\{\left(\mathbf{V}_{z_1}^T\mathbf{\Sigma}_{g,1}^{-2}\mathbf{V}_{z_1}\right)\left(\mathbf{U}_{z_1}^T\mathbf{D}_1'^{-1}\mathbf{U}_{z_1}\right)\right\} \\ &\overset{(c)}{\ge} \operatorname{tr}\left\{\mathbf{\Sigma}_{g,1}'^{-2}\mathbf{D}_1'^{-1}\right\} = \operatorname{tr}\left\{\mathbf{\Lambda}_{g,1}'^{-1}\mathbf{D}_1'^{-1}\right\}, \end{aligned} \tag{4.222}$$

where in step (a) and (b), we used first the cyclic property of the trace-operator and second the fact that

$$\mathrm{tr}\left\{\boldsymbol{\Sigma}_{z_1}^{-1}\mathbf{X}\right\} = \sum_{j=1}^{n} \sigma_j^{-1}\left(\mathbf{Z}_1\right)\left(\mathbf{X}\right)_{j,j} \overset{(4.221)}{\geq} \sum_{j=1}^{n} \left(\mathbf{X}\right)_{j,j} = \mathrm{tr}\left\{\mathbf{X}\right\},$$

for all σ_j, which satisfies (4.221) and for all $\mathbf{X} \in \mathrm{NND}\left(n\right)$ and $\left(\mathbf{X}\right)_{j,j} \geq 0$ denotes the jth diagonal entry of $\mathbf{X}$. Note that all diagonal entries of a positive semi-definite matrix are positive [9]. The inequalities in (a) and (b) are satisfied, since it is evident that $\mathbf{U}_{z_1}^T\mathbf{D}_1'^{-1}\mathbf{U}_{z_1}\boldsymbol{\Sigma}_{z_1}^{-1}\mathbf{V}_{z_1}^T\boldsymbol{\Sigma}_{g,1}^{-2}\mathbf{V}_{z_1} > 0$ and $\mathbf{V}_{z_1}^T\boldsymbol{\Sigma}_{g,1}^{-2}\mathbf{V}_{z_1}\mathbf{U}_{z_1}^T\mathbf{D}_1'^{-1}\mathbf{U}_{z_1} > 0$. In step (c), we applied Theorem 4.4.19, where $\boldsymbol{\Sigma}_{g,1}'$ contains the n non-zero singular values σ_{g_i} in increasing order (in terms of magnitude). Note that we already assumed that the singular values σ_i' of $\widetilde{\mathbf{A}}$ and in turn d_i' (elements in $\mathbf{D}'$), are ordered decreasingly. In the last step, we introduced the diagonal matrix $\boldsymbol{\Lambda}_{g,1}'$, which in particular, contains the n non-zero eigenvalues of $\mathbf{G}\mathbf{G}^T$, i.e, $\boldsymbol{\Lambda}_{g,1}'$ is the $n \times n$ upper left diagonal submatrix of $\boldsymbol{\Lambda}_g'$, which follows from the EVD $\mathbf{G}\mathbf{G}^T = \mathbf{U}_g'\boldsymbol{\Lambda}_g'\mathbf{U}_g'^T$, where the eigenvalues are sorted increasingly. In fact, the eigenvalues in $\boldsymbol{\Lambda}_{g,1}'$ are thus ordered increasingly.

Hence, the optimal $\mathbf{V}$, solving (4.220), is given by $\mathbf{V}_1^* = \mathbf{U}_{g,1}$, which contains the first n eigenvectors of $\mathbf{G}\mathbf{G}^T$, i.e., the first n orthonromal column vectors of $\mathbf{U}_g$, sorted increasingly. The remaining $m-n$ singular vectors in $\mathbf{V}_2$ for completition the optimal $\mathbf{V}$, can be chosen arbitrary in order to obtain a unitary $\mathbf{V}$ (orthonormal basis). Moreover, the optimal $\mathbf{V}$ for problem (4.218) is also given by

$$\mathbf{V}^* = \mathbf{U}_g, \tag{4.223}$$

where the columns of $\mathbf{U}_g$ correspond to the eigenvectors of $\mathbf{G}\mathbf{G}^T$, sorted increasingly.

After having determined the optimal $\mathbf{V}$ for (4.218), we will now characterize the optimal $\mathbf{U}$ for (4.218). To this end, we insert the optimum $\mathbf{V}$ in (4.218), yielding:

$$\begin{aligned} \underset{\mathbf{U},\boldsymbol{\Sigma}}{\text{maximize}} \quad & \mathrm{tr}\left\{\boldsymbol{\Lambda}_{g,1}^{-1}\mathbf{D}_1^{-1}\right\} \\ \text{subject to} \quad & \mathrm{tr}\left\{\mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}\right\} \leq \frac{P_0'}{\sigma_n^2}, \\ & \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T \succeq \mathbf{0}, \\ & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \end{aligned} \tag{4.224}$$

If $\mathbf{U}^*$ is optimal for (4.224), then it is also optimal for (4.218), as can be verified easily.

Definition 4.5.22 *A row- and column-swapping permutation matrix or a reflection matrix is given by*

$$\mathbf{\Xi} \triangleq \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}. \tag{4.225}$$

Invoking again problem (4.124), we note that it has exactly the same constraint function and a fairly similar objective. Where we have showed that the optimal $\mathbf{U}$ is determined by minimization of the constraint function

$$t(\mathbf{U}, \mathbf{\Sigma}) \triangleq \operatorname{tr}\left\{\mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}\right\}.$$

We can closely follow the approach and recognize that this also applies to (4.224). Thus, we can determine the optimal $\mathbf{U}$ by solving

$$\begin{aligned} \underset{\mathbf{U}}{\text{minimize}} \quad & \operatorname{tr}\left\{\mathbf{U}\mathbf{\Sigma}'\mathbf{\Sigma}'^T\mathbf{U}^T\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}\right\} \\ \text{subject to} \quad & \mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}, \end{aligned} \tag{4.226}$$

i.e., the optimal $\mathbf{U}$ will be a minimizer of the constraint function $t(\mathbf{U}, \mathbf{\Sigma})$, for the specific choice of $\mathbf{\Sigma} = \mathbf{\Sigma}'$. An application of Theorem 4.4.19 reveals, that the optimal $\mathbf{U}$ is given by the eigenvectors of $\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}$ or $\mathbf{H}\mathbf{H}^T$, respectively, in the order of increasing eigenvalues of $\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}$, or, decreasing eigenvalues of $\mathbf{H}\mathbf{H}^T$. However, since $\mathbf{H}\mathbf{H}^T$ and in turn $\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}$ is diagonal, the eigenvectors are given by the unit vectors $\{\mathbf{e}_k\}_{k=1}^{p}$. Moreover, since the eigenvalues of $\left(\mathbf{H}\mathbf{H}^T\right)^{\dagger}$ are the squared reciprocals of the diagonal values $h_j \neq 0$, we have that the kth column of $\mathbf{U}^*$ is given by $\mathbf{e}_{p-k+1}$ for $1 \leq k \leq p$, since we assumed that h_j are ordered increasingly in terms of magnitude. Thus, the optimal $\mathbf{U}$ is given by

$$\mathbf{U}^* = \mathbf{\Xi}, \tag{4.227}$$

where $\mathbf{\Xi}$ is defined in (4.225).

So far, we have determined the optimal $\mathbf{U}$ and $\mathbf{V}$ for (4.218). It remains to determine the optimal $\mathbf{\Sigma}$ for (4.218). To that end, we insert the optimum choices $\mathbf{U}^*$ from (4.227) and $\mathbf{V}^*$ from (4.223) into 4.218, yields to the optimization problem in standard form [15] as

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \sum_{i=1}^{n} \frac{1}{\lambda_{g_i} s_i}, \\ \text{subject to} \quad & \mathbf{b}'^T\mathbf{s} - P_0' \leq 0, \\ & -\mathbf{s} \preceq \mathbf{0}, \end{aligned} \tag{4.228}$$

again with $\mathbf{s} = \left(s_1, s_2, \ldots, s_n\right)^T \triangleq \left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\right)^T \in \mathbb{R}^{+n}$, i.e., $\mathbf{s} \succeq \mathbf{0}$. The vector $\mathbf{b}$ in the constraint function of (4.228) is defined as $\mathbf{b}' \triangleq \left({h'_1}^{-2}, {h'_2}^{-2}, \ldots, {h'_n}^{-2}\right)^T$, where

$$h'_j \triangleq \begin{cases} \infty & 1 \leq j \leq p-q \\ h_{p-j+1} & p-q < j \leq p. \end{cases} \tag{4.229}$$

Note that $h_i \neq 0$ for $1 \leq i \leq n \leq q$ and $\sigma_j = 0$ for $n < j \leq w$, since $\widetilde{\mathbf{A}}$ has rank n and thus $\mathbf{s}$ and $\mathbf{b}'$ has dimension n (cf. (4.228)). We recognize that problem (4.228) is equivalent to (4.199) for $\lambda_{\widetilde{g}_i} = \lambda_{g_i}$, $P_0 = P'_0$ and $\mathbf{b} = \mathbf{b}'$. Hence, the optimal s_i of problem (4.228), is given by (4.204) for $\lambda_{\widetilde{g}_i} = \lambda_{g_i}$ and $P_0 = P'_0$. Therefore, we finally obtain with

$$\sigma_i^* = \begin{cases} \sqrt{s_i^*} & \text{for } 1 \leq i \leq n \\ 0 & n < i \leq w, \end{cases}$$

the optimal $\mathbf{\Sigma}$ for (4.218) as

$$\mathbf{\Sigma}^* = c^* \begin{bmatrix} \mathbf{\Lambda}_{g,1}^{-1/4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{4.230}$$

where

$$c^* = \sqrt{\frac{P'_0}{\sum_{j=1}^{n} \frac{1}{{h'_j}^2 \sqrt[4]{\lambda_{g_j}}}}}. \tag{4.231}$$

So far, we have determined the optimum $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{\Sigma}$ for (4.218) and thus the optimal $\widetilde{\mathbf{A}}$ with $\widetilde{\mathbf{A}}^* = \mathbf{U}^*\mathbf{\Sigma}^*\mathbf{V}^{*T}$. The unitary matrix $\mathbf{U}^*$ contains the unit vectors $\{\mathbf{e}_i\}_{i=1}^p$, since $\mathbf{H}$ is assumed to be diagonal. If we further assume, that the the diagonal $\mathbf{H}$ is in decreasing order (in terms of magnitude), then $\mathbf{U}^* = \mathbf{I}$. The unitary matrix $\mathbf{V}^* = \mathbf{U}_g$ (cf. (4.223)), where $\mathbf{U}_g$ contains the eigenvectors of $\mathbf{G}\mathbf{G}^T$, in increasing order and $\mathbf{\Sigma}^*$ is given in (4.230), where the eigenvalues of $\mathbf{G}\mathbf{G}^T$ in $\mathbf{\Lambda}_g$ are ordered increasingly.

Finally, with (4.212) and the fact that $\mathbf{H}$ is of full column-rank, we obtain the optimal local sensor matrix $\mathbf{A}$ as

$$\mathbf{A}^* = \mathbf{H}^\dagger\mathbf{U}^*\mathbf{\Sigma}^*\mathbf{V}^{*T} = c^*\mathbf{H}^\dagger \begin{bmatrix} \mathbf{\Lambda}_{g,1}^{-1/4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}_g^T,$$

and in turn with the already known solution $\mathbf{C}_l^* = \mathbf{0}$, the optimal $\mathrm{LO}_{\phi_{-1}}^* =$

$(\mathbf{A}^*,\ \mathbf{C}_l^*)$ is thus given by

$$\begin{aligned}
&\mathrm{LO}^*_{\phi_{-1}}: \quad \left(\mathbf{A}^* = c^*\mathbf{H}^\dagger \begin{bmatrix} \mathbf{\Lambda}_{g,1}^{-1/4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}_g^T,\ \mathbf{C}_l^* = \mathbf{0}\right)_{\phi_{-1}} \\
&\quad \text{for constraint (C2)}, \quad c^* \text{ is given in (4.231)}, \\
&\quad \text{EVD: } \mathbf{G}\mathbf{G}^T = \mathbf{U}_g\mathbf{\Lambda}_g\mathbf{U}_g^T, \\
&\qquad \text{where } \mathbf{\Lambda}_g,\ \mathbf{U}_g \text{ are sorted increasingly}, \\
&\qquad \mathbf{\Lambda}_{g,1} \text{ is the } n \times n \text{ left upper submatrix of } \mathbf{\Lambda}_g, \\
&\quad \mathbf{H} \text{ is sorted decreasingly.}
\end{aligned} \tag{4.232}$$

However, since we have found a closed-form solution for the A-optimal ith LO, we can still specify the resulting FIM $\mathbf{J}_{\mathbf{z}}^*$. To that end, we insert $\mathbf{A}^*$ from (4.232) into (4.149) (without the subscript notation) for $\mathbf{A} = \mathbf{A}^*$, yielding:

$$\begin{aligned}
\mathbf{J}_{\mathbf{z}}^* &= \left(\mathbf{V}_g\mathbf{\Sigma}_g^T \begin{bmatrix} \mathbf{\Lambda}_{g,1}^{-1/4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \left(\frac{\sigma_h^2}{c^{*2}}\mathbf{I} + \sigma_n^2 \begin{bmatrix} \mathbf{\Lambda}_{g,1}^{-1/4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}^2\right)^{-1} \cdot \right. \\
&\qquad\qquad \left. \cdot \begin{bmatrix} \mathbf{\Lambda}_{g,1}^{-1/4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{\Sigma}_g\mathbf{V}_g^T\right) \\
&= \mathbf{V}_g\mathbf{\Sigma}_{g,1}\mathbf{\Lambda}_{g,1}^{-1/4}\left(\frac{\sigma_h^2}{c^{*2}}\mathbf{I} + \sigma_n^2\mathbf{\Lambda}_{g,1}^{-1/2}\right)^{-1}\mathbf{\Lambda}_{g,1}^{-1/4}\mathbf{\Sigma}_{g,1}\mathbf{V}_g^T \\
&\overset{(a)}{=} \mathbf{V}_g\mathbf{\Lambda}_{g,1}^{1/2}\left(\frac{\sigma_h^2}{c^{*2}}\mathbf{I} + \sigma_n^2\mathbf{\Lambda}_{g,1}^{-1/2}\right)^{-1}\mathbf{V}_g^T,
\end{aligned} \tag{4.233}$$

where in step (a), we used the fact that $\mathbf{\Lambda}_{g,1} = \mathbf{\Sigma}_{g,1}^2$.

Chapter 5

Numerical Experiments

In this chapter, we do some numerical experiments to study the performance behaviour for optimal designed local sensors. Supposing, we use an MVU estimator for our performance analysis. First, we use result for the special case of a scalar parameter, where we compare the optimal power scheduling (cf. Subsubsection 4.4.3.3) versus the uniform power scheduling performance for an orhtogonal MAC. Then, we restrict our channel model to the cases, where we also derived optimal solutions for the coherent MAC case, where we then analyse the performance of orthogonal versus coherent MAC for an optimal power scheduling strategie (total power constraint). Finally, we will consider the general case of a vector-valued parameter, where we will analyse the MSE performance for a single sensor setup (i.e., $L = 1$) for an T-optimal versus an A-optimal design.

Assuming ideal channel models, i.e., when the local sensor observations $\mathbf{y}_i$ for $1 \leq i \leq L$ are directly available to the FC, the FIM $\mathbf{J}_\mathbf{z}$ is then given by

$$\mathbf{J}_{\mathbf{z},0} \triangleq \mathbf{J}_\mathbf{z} = \sum_{i=1}^{L} \mathbf{G}_i^T \mathbf{C}_{n_i}^{-1} \mathbf{G}_i. \tag{5.1}$$

Note that here we assume that $\mathbf{C}_{n_i}$ is non-singular for $1 \leq i \leq L$. The FIM $\mathbf{J}_{\mathbf{z},0}$ is our central performance benchmark for non-ideal channel models and it is obvious, that it holds for both multiple access schemes, i.e, for (3.8) and (3.20). Note, that the FIM $\mathbf{J}_\mathbf{z}$ for both MAC schemes could exist, even if $\mathbf{J}_{\mathbf{z},0}$ do not exist in that form. This fact results from the existence of the channel noise, and the assumption that the covariance matrix is non-singular.

Let us first recall the MSE definition, given in (2.9). Since the CRLB $\mathbf{J}_\mathbf{z}^{-1}$ is the covariance matrix of an efficient MVU, which exists for a LGM in particular (provided that the FIM $\mathbf{J}_\mathbf{z}$ is not singular, cf. Subsection 2.2.5), the MSE can be computed by invoking (2.9) as

$$\text{MSE} = \frac{1}{n} \text{tr} \left\{ \mathbf{J}_\mathbf{z}^{-1} \right\}, \tag{5.2}$$

which is the arithmetic average of the scalar variances $\operatorname{var}\left\{\hat{\theta}_k\right\}$ for $1 \leq k \leq n$.

Let us now introduce some additional definitions, which are used in the following numerical experiments. We denote the total channel noise power by

$$P_h \triangleq \sum_{i=1}^{L} \sigma_{h_i}^2, \tag{5.3}$$

and the signal to noise ratio (SNR) for constraint (C1) by

$$\mathrm{SNR} \triangleq P_0/P_h, \quad \text{with } P_0 \triangleq \sum_{i=1}^{L} P_{0,i}, \tag{5.4}$$

and the SNR for constraint (C2) by

$$\mathrm{SNR}' \triangleq P_0'/P_h, \quad \text{with } P_0' \triangleq \sum_{i=1}^{L} P_{0,i}'. \tag{5.5}$$

5.1 Scalar Parameter

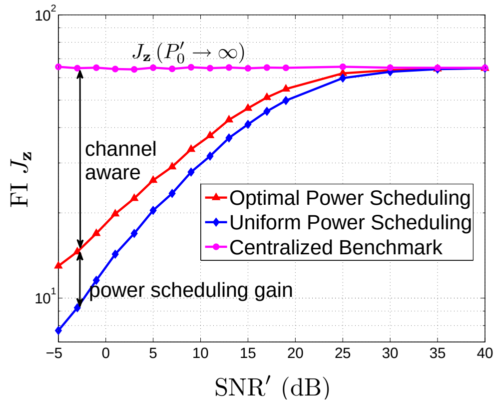
We first consider the scalar parameter case. In what follows, we will analyse the performance for the optimal power scheduling compared to the uniform power scheduling strategie and their asymptotic behaviour for an orthogonal MAC, i.e., on the one hand, when the total power/variance (P_0 or P_0') increases and on the other hand, when the number of sensors L increases. Note in simulations we consider the equivalent model with scalar observation at each local sensor. Therefore, we consider the observation model paramters g_i and σ_i^2, respectively. For performance analysis of the channel aware, we recall the centralized performance benchmark from (5.1), which specializes for a scalar parameter and using the equivalent model with scalar observation to

$$J_{\mathbf{z},0} = \sum_{i=1}^{L} \frac{g_i^2}{\sigma_{n_i}^2}, \quad \sigma_{n_i}^2 \neq 0. \tag{5.6}$$

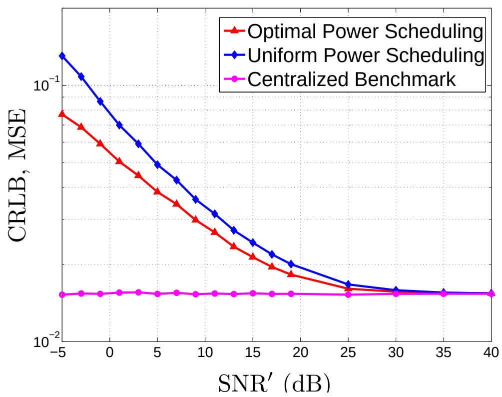
For the following simulations and performance analysis, we will consider only constraint (C2). Also, we adhere strongly to the simulations made by the authors of [1] to finally carry out a comparison.

5.1.1 Optimal Power Scheduling for an Orthogonal MAC

As discussed in Subsubsection 4.4.3.3, we have found a "water-filling" solution for the optimal power scheduling (cf. (4.142) and (4.143)), i.e., for a given total transmit variance power P_0' - and assuming optimal LO_i for $1 \leq i \leq L$, we derived optimal power scheduling among all L sensors, in order to achieve the best performance, i.e., the maximum FI $J_{\mathbf{z}}$, which is given in (4.137) for optimal designed sensors. In contrast, we also discussed the uniform power scheduling, where the total variance power P_0', are uniformly distributed among all L sensors - the resulting FI $J_{\mathbf{z},u}$ is given in (4.138).



(a)



(b)

Figure 5.1: Uniform power scheduling vs. optimal power scheduling, when P_0' or SNR$'$ increases, for a fixed total number of sensors $L = 15$. As can be seen, both scheduling strategies converge to the centralized benchmark, when SNR$'$ increases. The lower the SNR$'$, the more significant the performance gain, due to optimal power scheduling.

We will now compare both power scheduling strategies in a simple numerical experiment, where we compare both the FI $J_{\mathbf{z}}$ and the CRLB (MSE of the MVU) by varying the total transmit variance power P_0', while the channel noise power is constant (in simulation we used unit variance $P_h = 1$).

In Fig. 5.1, we plot the curves for the FI $J_{\mathbf{z}}$ and the CRLB $J_{\mathbf{z}}^{-1}$ versus the $\mathrm{SNR}' = P_0'/P_h$ (cf. (5.5)), under both uniform and optimal power schedules, where the total number of sensors is constant with $L = 15$. Further, observation noise variances $\sigma_{n_i}^2$ are uniformly taken from the real interval $[1, 1.5]$ - the observation gains g_i are uniformly taken from the real interval $(0, 4]$. The squared channel gain $h_{i\,\max}^2$, i.e., the largest eigenvalue of $\mathbf{H}_i^T\mathbf{H}_i$ are taken as $h_{i\,\max}^2 = c_g \cdot d^{-3.5}$, where d is uniformly taken from the real interval $[1, 10]$ and c_g is a normalization constant such that $\mathrm{E}\left\{h_{i\,\max}^2\right\} = 1$. In the simulation, the simulated FI $J_{\mathbf{z}}$ or, equivalently, the CRLB, is averaged over 1000 realizations of the set $\left\{\sigma_{n_i}^2, g_i, h_{i\,\max}^2 : 1 \le i \le L\right\}$ and is actually the expected $J_{\mathbf{z}}$ and the expected CRLB, respectively.

As can be seen in Fig. 5.1, when the SNR' increases, both uniform an optimal power scheduling converges to the centralized benchmark, given in (5.6), i.e.,

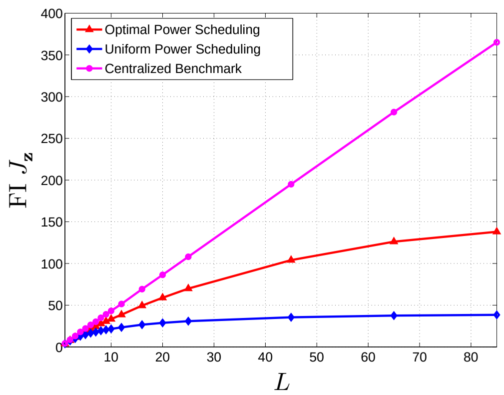
$$J_{\mathbf{z},o}\left(P_0' \to \infty\right) = J_{\mathbf{z},u}\left(P_0' \to \infty\right) = \sum_{i=1}^{L} \frac{g_i^2}{\sigma_{ni}^2}, \tag{5.7}$$

when we denote $J_{\mathbf{z},o}$ as the achieved FI for optimal power scheduling (cf. red curve in Fig 5.1) and $J_{\mathbf{z},u}$ as the achieved FI for uniform power scheduling (cf. blue curve in Fig 5.1 and (4.138)). Note that the asymptotic behaviour for the uniform case when P_0' increases, yields to (4.139), which coincides of course with the central performance benchmark in (5.6). On the other hand, the optimal power scheduling gain, i.e., the difference between uniform an optimum power scheduling in a logarithmic plot, becomes more significant as the SNR' decreases.

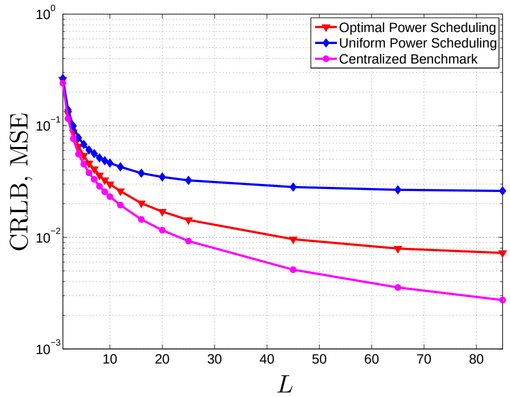
We now fix the total transmission variance power P_0' such that we obtain an $\mathrm{SNR}' = 15\mathrm{dB}$ and varying the total number of sensors L. In Fig 5.2, we plot the curves for the FI $J_{\mathbf{z}}$ and the CRLB versus the total number of sensors L under both, uniform and optimal power schedules. Again, in the simulation, the FI $J_{\mathbf{z}}$ or equivalently the CRLB is averaged over 1000 realisations of $\left\{\sigma_{n_i}^2, g_i, h_i : 1 \le i \le L\right\}$ for $1 \le L \le L_{\max} = 45$, and is actually the expected $J_{\mathbf{z}}$ and the expected CRLB or MSE, respectively. As can be seen, the optimum power scheduling gain increases, as the total number of sensors L increases.

5.1.2 Optimal Power Scheduling for an Orthogonal MAC and a Coherent MAC

Let us now compare the performance of orthogonal MAC and coherent MAC under an optimal power scheduling startegie. We restrict our numerical experiment by assuming orthogonal channel matrices $\mathbf{H}_i$ for all i, since for this special case we derived an optimal local sensor rule for the coherent MAC case (cf. Subsubsection 4.4.3.2).



(a)



(b)

Figure 5.2: Uniform Power Scheduling vs. Optimal Power Scheduling, when L increases for $\mathrm{SNR}' = 15\mathrm{dB}$.

For a unitary channel matrix $\mathbf{H}_i$, all eigenvalues of $\mathbf{H}_i^T\mathbf{H}_i$ are equal to one, thus $h_{i\,\max}^2 = 1$. In the following simulation we take the same setting for observation parameters g_i and $\sigma_{n_i}^2$ as before - the channel noise power is again assumed to have unit variance, i.e., $\sigma_h^2 = 1$. In Fig. 5.3, we plot the curves of the FI $J_{\mathbf{z}}$ and the CRLB versus the total number of sensors L for the orthogonal MAC case, as before, and in addition the coherent MAC case. Again, we fixed the $\mathrm{SNR}' = 10\mathrm{dB}$. Note that since we solved the otpimum LO_i for the coherent MAC case with respect to a total power constraint, the observed FI $J_{\mathbf{z}}$ and the corresponding CRLB or MSE can be perceived as an optimal power scheduling solution. In the simulation, the FI $J_{\mathbf{z}}$ and the CRLB is again averaged over 1000 realisations of $\left\{\sigma_{n_i}^2, g_i : 1 \leq i \leq L\right\}$ for $1 \leq L \leq L_{\max}$, and is actually the expected $J_{\mathbf{z}}$ and the expected CRLB or MSE, respectively.

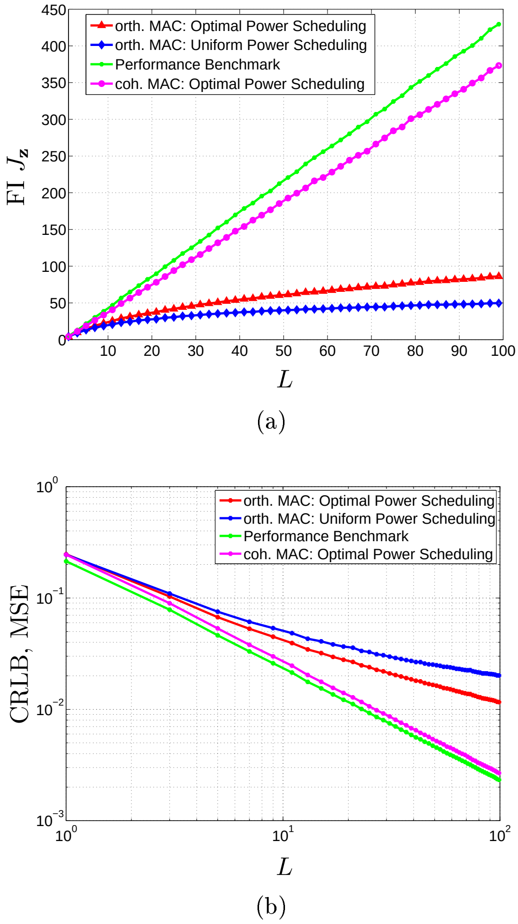


Figure 5.3: Uniform power scheduling vs. optimal power scheduling, when L increases for $\mathrm{SNR}' = 10\mathrm{dB}$.

We perceive that for an orthogonal MAC with finite amount on P_0', the overall CRLB or MSE does not decreases to zero, even if L, the total number of sensors, aproaches infinity. This fact results from the orthogonality of each link from sensor to the FC, which leads to L different and independent channel noise vectors $\mathbf{n}_{h_i}$ for $1 \leq i \leq L$. Therefore, the corruption of channel noise cannot be eliminated even when L goes to infinity. In the coherent MAC case, only one channel noise $\mathbf{n}_h$ is generated per transmission unit. As a result of the coherent combination, the SNR for the received data scales with L, since all transmitted data vectors are correlated to each other, even though when P_0' is finite.

5.1.3 Comparison to Existing Results

In conclusion, we show a comparison between our simulation results for the optimal powers scheduling performance and the simulations results by the authors of [1]. In Fig. 5.4 are illustrated the optimal power scheduling gains compared to the uniform power scheduling for both multiple access schemes. As already mentioned they considered a Baysian setting, where they minimized the MSE of the MMSE estimator. However, the performance results, with regard to asymptotic behaviours are basically the same insights.

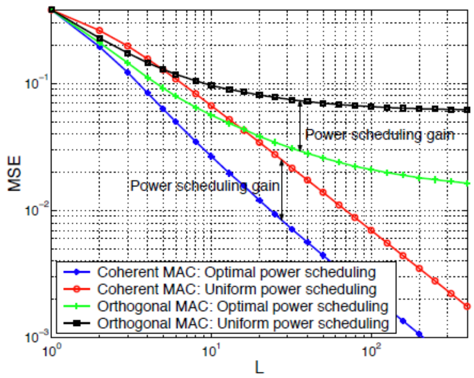


Figure 5.4: MSE performance comparision between orthogonal and coherent MACs [1].

5.2 Vector Parameter

In what follows, we analyse the performance for the vector paramter case, where T- and A-optimal designed local sensors (designed for constraint (C2)), will be compared with regard to the MSE performance - the MSE of an efficient MVU is given in (5.2). We consider the standard model in our simulation setup (cf. Definition 4.3.14).

5.2.1 T–Optimal and A–Optimal MSE Performance

In the following simulation we used a system setup, where observation- and channel matrix are both invertable. An A-optimal LO is given in closed-form - the T-optimal LO has to be computed in a water-filling like manner. Since, the A-optimal design minimizes the MSE of an efficient unbiased estimator MVU, we expect a significant performance gain against the T-optimal design. We suppose a single sensor setup, i.e., $L = 1$. In Fig. 5.5, we plot the curves for the MSE of the MVU versus the SNR$'$ from (5.5), for a T- and a A-optimal design.

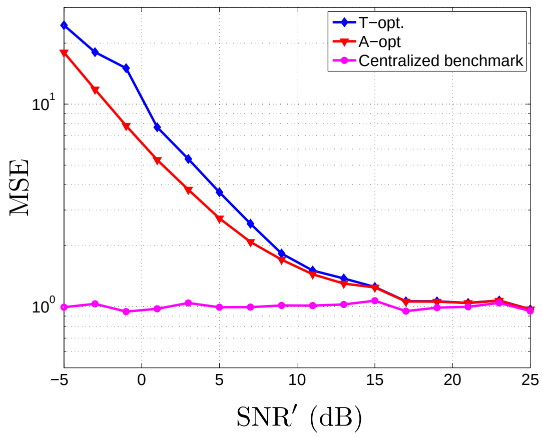


Figure 5.5: MSE-performance for a T- and A-optimal design over SNR′ in a single sensor setup, $L = 1$.

In the simulation, we used constant eigenvalues of observation matrix $\mathbf{G}^T\mathbf{G}$, i.e., $\lambda_{g_i} = 1$ for all i - a Chi-squared (degree 1) distributed variance σ_n^2 and the eigenvalues for $\mathbf{H}\mathbf{H}^T$ are uniformly taken from the real interval $[0.5, 1]$. For the channel noise covariance, we set $\sigma_h^2 = 1$. Again, we averaged over 1000 realisations on the set $\left\{\sigma_n^2, h_i : 1 \leq i \leq L\right\}$, and is actually the expected MSE.

As can be seen, a design under the A-criteria, performs better than that under the T-criteria, in terms of the MSE. Consider the asymptotic behaviour of both designs for increasing SNR′. We conclude that both converges to the cetralized benchmark, i.e., to the minimum achievable MSE, wich results for an ideal channel. On the other extrema, when SNR′ goes very small, then the difference between A- and T-optimal performance gets larger. Hence, a T-optimal design is quite sufficient, just when the SNR′ is large enough.

Chapter 6

Conclusions

We considered a WSN, where sensors and a FC collaborate to estimate an unknown deterministic vector parameter. Due to bandwidth and/or power limitations, each local sensor has to encode and/or compress (local sensor rule) their measurement data of the unknown parameter first, before transmitting it over an imperfect channel to the FC. This encoding should be, such that the FC can estimate the parameter of interest most accurately. We used the FI as our performance metric, due to their relation to the CRLB. We considered a linear Gaussian setup, where each local sensor rule and the fusion rule (estimator function at the FC) are described by linear mappings. We investigated two types of channel usage, an orthogonal and a coherent MAC. The main goal of this thesis was to determine optimal local sensor rules, in the sense of maximizing the FI, subject to bandwidth and/or power constraints of the transmit signals.

First we have described our local sensor rule more generally by a linear transformation and additive systematic Gaussian noise, whereby we have showed that the systematic noise can be neglected.

For the scalar case, we have shown that we can reduce our system model to an equivalent model in which all local observations are scalar-valued. Based on this equivalent model, we derived optimal local sensor rules for an orthogonal MAC in closed form. We also studied the coherent MAC case and derived optimal local sensor rules under a total power constraint for certain special cases of the channel states. Based on these optimal local sensor rules, we have considered the optimal power allocation among sensors. We derived a water-filling based solution for the optimal power scheduling under a given total power constraint for the orthogonal MAC case. Simulations showed that the proposed power scheduling strategy significantly improves the performance when compared to the uniform power scheduling. We have also shown that the performance has significantly different asymptotic behaviors when the number of sensors L is large for orthogonal and coherent MACs.

For a vector parameter, we first discussed some fundamental notions of optimal experiment designs. In particular, we introduced the T- and A-optimality criteria, which we then used for a vector-valued parameter. We derived T- and A-optimal local sensor rules for certain special cases of channel states. A final simulation showed a MSE-performance comparison of these two optimal designs.

Appendix A

A Convex Optimization Problem

A.1 Water–filling Solution

In solving the optimal power scheduling for a scalar paramter in the orthogonal MAC case (cf. Subsubsection 4.4.3.3), and for solving the T-optimal Design of a local sensor LO_i in the vector parameter case (cf. Subsection 4.5.1), we have to solve an equivalent optimization problem in the form:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) \triangleq = -\sum_{k=1}^{K} c_k^{(1)} \frac{x_k}{c_k^{(2)} + c_k^{(3)} x_k} \\ \text{subject to} \quad & g_1(\mathbf{x}) \triangleq \mathbf{1}^T \mathbf{x} - P \leq 0 \\ & g_2(\mathbf{x}) \triangleq -\mathbf{x} \preceq 0, \end{aligned} \tag{A.1}$$

where the vector $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_K \end{bmatrix}^T$; We assume that the constants $P \geq 0$, $c_k^{(1)} \geq 0$, $c_k^{(2)} \geq 0$ and $c_k^{3)} > 0$. First, we verify that (A.1) is a convex optimization problem. We can write the objective $f(\mathbf{x})$ of problem (A.1) as

$$f(\mathbf{x}) = \sum_{k=1}^{K} f_k(x_k), \tag{A.2}$$

with $f_k(x_k) \triangleq -c_k^{(1)} \frac{x_k}{c_k^{(2)} + c^{(3)} x_k}$. It can be verified easily that the first two derivatives of $f_k(x_k)$ are given by

$$f_k'(x_k) \triangleq \frac{\partial}{\partial x_k} f(x_k) = -\frac{c_k^{(1)} c_k^{(2)}}{\left(c_k^{(2)} + c_k^{(3)} x_k\right)^2} \tag{A.3}$$

and

$$f_k''(x_k) \triangleq \frac{\partial^2}{\partial x_k^2} f(x_k) = 2\frac{c_k^{(1)} c_k^{(2)} c_k^{(3)}}{\left(c_k^{(2)} + c_k^{(3)} x_k\right)^3} \geq 0, \tag{A.4}$$

and therefore $f_k(x_k)$ is convex. The convexity of the constraint functions $g_1(\cdot)$ and $g_2(\cdot)$ is obvious, since both are linear in $\mathbf{x}$. Hence, problem (A.1) is a convex optimization problem [15, Chapter 4.2.1].

The KKT conditions (cf. [15]) for a solution $\mathbf{x}^*$ to the optimization problem (A.1) and corresponding Lagrange multipliers (cf. [15, p.244]), i.e., ν^* for the inequality constraint $g_1(\cdot) \leq 0$ and $\boldsymbol{\lambda}^* \in \mathbb{R}^K$ for the inequality constraint $g_2(\cdot) \preceq \mathbf{0}$ are given as

$$\begin{aligned} \mathbf{1}^T \mathbf{x}^* - P &\leq 0 \\ \mathbf{x}^* &\succeq 0 \\ \nu^* &\geq 0 \\ \boldsymbol{\lambda}^* &\succeq 0 \\ \nu^* \left(\mathbf{1}^T \mathbf{x}^* - P\right) &= 0 \\ \lambda_i^* x_k^* &= 0, \ k = 1, 2, \ldots, K \\ -c1_k \frac{c_k^{(2)}}{\left(c_k^{(2)} + c_k^{(3)} x_k^*\right)^2} + \nu^* - \lambda_k^* &= 0, \ k = 1, 2, \ldots, K. \end{aligned} \tag{A.5}$$

The resulting optimum $\mathbf{x}^*$ can be obtained by a so called "water-filling" procedure. We note that the problem (A.1) is identical to the problem considered in [15, Ex.5.2], except for the objective functions $f_k(x_k)$. Therefore, we can closely follow the method in [15, Ex.5.2] to solve the KKT conditions (A.5). In particular, we obtain the following "watter-filling". The optimal values x_k can be expressed as

$$x_k^* = \max\left\{0, \sqrt{\frac{c_k^{(2)} c_k^{(1)}}{(c_k^{(3)})^2} \frac{1}{\nu^*}} - \frac{c_k^{(2)}}{c_k^{(3)}}\right\}, \tag{A.6}$$

and

$$\sum_{i=1}^{K} \max\left\{0, \sqrt{\frac{c_k^{(2)} c_k^{(1)}}{(c_k^{(3)})^2} \frac{1}{\nu^*}} - \frac{c_k^{(2)}}{c_k^{(3)}}\right\} = P. \tag{A.7}$$

Hence, the optimal x_k for $1 \leq k \leq K$ for (A.1) can not be computed in closed-form. First, we have to determine the optimal variable ν from (A.7). Subsequently, the optimal x_k for $1 \leq k \leq K$ can then be computed according to (A.6). This can be done by a so called "water-filling" algorithm (Cf. Algorithm A.1).

Algorithm 1 Water-filling Algorithm

$P,\ c_k^{(1)},\ c_k^{(2)},\ c_k^{(3)}$ given for all k
$tol = 1e - 5$
$wline = 0 \quad \%wline = 1/\nu$
$Ptot = \sum_k max(0, \sqrt{\frac{c_k^{(2)} c_k^{(1)}}{(c_k^{(3)})^2} wline} - \frac{c_k^{(2)}}{c_k^{(3)}})$
$z = 0$
while $abs(x_k - Ptot) > tol$ && $z < 10000$ **do**
  $wline = wline + (x_k - Ptot)/300$
  $Ptot = \sum_k max(0, \sqrt{\frac{c_k^{(2)} c_k^{(1)}}{(c_k^{(3)})^2} wline} - \frac{c_k^{(2)}}{c_k^{(3)}})$
  $z = z + 1$
end while
$x_k^* = max(0, \sqrt{\frac{c_k^{(2)} c_k^{(1)}}{(c_k^{(3)})^2} \frac{1}{\nu^*}} - \frac{c_k^{(2)}}{c_k^{(3)}})$

Bibliography

[1] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Processing*, vol. 56, pp. 757 –770, Feb. 2008.

[2] I. D. Schizas, G. B. Giannakis, and Z.-Q. Luo, "Distributed estimation using reduced-dimensionality sensor observations," *IEEE Trans. Signal Processing*, vol. 55, pp. 4284 –4299, Aug. 2007.

[3] P. Z. K. Zhang, X. R. Li and H. Li, "Optimal linear estimation fusion part vi: sensor data compressionn," *Proc. 2003 International Conf. Information Fusion*, pp. 221–228, July 2003.

[4] F. Pukelsheim, *Optimal Design of Experiments.* University of Augsburg: Society for Industrial and Applied Mathematics, classic ed. ed., 2006.

[5] F. Hlawatsch, *Statistical Signal Processing: Parameter Estimation Methods.* Lecture notes; Institut für Nachrichtentechnik und Hochfrequenztechnik, Technische Universität Wien, March 2001.

[6] Z. Ben-Haim and Y. Eldar, "On the constrained Cramér–Rao bound with a singular Fisher information matrix," *IEEE Signal Processing Letters*, vol. 16, pp. 453–456, June 2009.

[7] V. V. Michel Verhaegen, *Filtering and System Identification: A Least Squares Approach.* USA by Cambridge University Press, New York, 2007.

[8] K. Schmidt and G. Trenkler, *Moderne Matrix-Algebra.* Berlin-Heidelberg: Springer, 1998.

[9] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge (UK): Cambridge Univ. Press, 1999.

[10] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis.* Cambridge, UK: Cambridge Univ. Press, 1991.

[11] G. H. Golub and C. F. Van Loan, *Matrix Computations.* Baltimore, MD: Johns Hopkins University Press, 3rd ed., 1996.

[12] D. Serre, *Matrices: Theory and Applications.* France by Dunod, Paris, 2001.

[13] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications.* Wiley, 1974.

[14] J. Lasserre, "A trace inequality for matrix product," *Automatic Control, IEEE Transactions*, vol. 40, Issue: 8, pp. 1500 – 1501, Aug. 1995.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, UK: Cambridge Univ. Press, 2004.

[16] F. Zhang, *Matrix Theory: Basic Results and Techniques.* Springer-Verlag, 2011.

[17] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis.* Cambridge (UK): Cambridge Univ. Press, 1994.