

**SIGNAL STRENGTH BASED LOCALIZATION AND
PATH-LOSS EXPONENT SELF-ESTIMATION IN
WIRELESS NETWORKS**

SIGNAL STRENGTH BASED LOCALIZATION AND PATH-LOSS EXPONENT SELF-ESTIMATION IN WIRELESS NETWORKS

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 24 oktober 2017 om 10:00 uur

door

Yongchang HU

Master of Science in Electrical Engineering,
Northwestern Polytechnical University, Xian, China,
geboren te Xian, China.

Dit proefschrift is goedgekeurd door de

promotor: Prof. dr. ir. G. J. T. Leus

Samenstelling promotiecommissie:

Rector Magnificus,
Prof. dr. ir. G. J. T. Leus,

voorzitter
Delft University of Technology, promotor,

Onafhankelijke leden:

Prof. dr. ir. A. -J. van der Veen,
Prof. dr. H. Wymeersch
Dr. S. Chatzinotas
Dr. ir. J. Romme
Dr. ir. C. Tiberius
Prof. K. Bertels

Delft University of Technology,
Chalmers University of Technology
Université du Luxembourg
IMEC-NL
Delft University of Technology,
Delft University of Technology, reservelid



The work presented in this thesis was financially supported by China Scholarship Council (CSC).

Chapter 1,3,6 Copyright © 2017 by Yongchang Hu
Chapter 2,4,5 Copyright © 2014-2017 by IEEE

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission of the copyright holder.

Printed by: Proefschriftmaken

ISBN: 978-94-6295-758-9

An electronic version of this dissertation is available at
<http://repository.tudelft.nl/>.

CONTENTS

Summary	ix
Acknowledgements	xi
I PROLOGUE	1
1 Introduction	3
1.1 Background and Motivations	3
1.1.1 Localization	4
1.1.2 Path-loss Exponent Estimation	18
1.1.3 Channel Effects	27
1.2 Outline and Contributions	29
References	34
II SIGNAL STRENGTH BASED SOURCE LOCALIZATION	43
2 Robust Differential Received Signal Strength-Based Localization	45
2.1 Introduction	47
2.2 Whitened Model for DRSS-Based Localization	50
2.3 Estimators For Known Path-Loss Model.	52
2.3.1 Advanced Best Linear Unbiased Estimator.	53
2.3.2 Lagrangian Estimator	54
2.3.3 Robust Semidefinite Programming Based Estimator	55
2.3.4 Complexity Analysis	58
2.3.5 Numerical Results	58
2.3.6 Discussions	66
2.4 Estimator For Unknown Path-Loss Model.	67
2.4.1 Handling Unknown Path-Loss Exponent.	67
2.4.2 Prototype of the Proposed Iterative Estimator	68
2.4.3 Robust Semidefinite Programming Based Block Coordinate Descent Estimator	69
2.4.4 Numerical results	71
2.5 Conclusions.	72
2.6 Appendices	73
2.6.1 RSS Collection	73
2.6.2 Derivation from RSS-Based Model	74
2.6.3 Cramér-Rao Lower Bounds	75
References	76

3	On A Unified Framework for Linear Nuisance Parameters	81
3.1	Introduction	82
3.2	Handling Linear Nuisance Parameters	83
3.2.1	Joint Estimation	83
3.2.2	OSP-Based Estimation	84
3.2.3	Differential Signal Processing	85
3.2.4	Discussion	89
3.3	Localization Examples	92
3.3.1	Time-Based Localization	92
3.3.2	Received Signal Strength Based Localization	99
3.3.3	Other Examples	101
3.4	Conclusions.	103
	References	103
III	PATH-LOSS EXPONENET SELF-ESTIMATION IN WIRELESS NETWORKS	109
4	Self-Estimation of Path-Loss Exponent in Wireless Networks and Applications	111
4.1	Introduction	113
4.2	System Model.	114
4.2.1	Node Distribution	114
4.2.2	Channel Model.	114
4.2.3	Problem Statement	116
4.3	Linear Regression Model for the path-loss exponent	117
4.3.1	Ranking Received Signal Strengths	117
4.3.2	Linear Regression Model for the Path-Loss Exponent	117
4.3.3	Estimation of $L_{i,j}$	118
4.4	Path-loss Exponent Estimation	119
4.4.1	Total Least Squares Solution	119
4.4.2	Closed-Form Total Least Squares Estimation.	120
4.4.3	Closed-Form Weighted Total Least Squares Estimation	122
4.4.4	Discussions and Future Works	123
4.5	Simulations	125
4.5.1	The Impact of the Shadowing	126
4.5.2	The Impact of the Actual Density	129
4.6	Applications	129
4.6.1	Secure RSS-Based Localization.	129
4.6.2	Energy-Efficient Routing	132
4.6.3	Other Applications.	134
4.7	Conclusions.	136
	References	136
5	Directional Maximum Likelihood Self-Estimation of the Path-Loss Exponent	139
5.1	Introduction	140
5.2	RSS distribution in wireless random networks	141

5.3	Directional Maximum Likelihood Self-Estimation of the PLE	143
5.3.1	CRLB.	143
5.3.2	Two ML Self-Estimators for the PLE	144
5.4	Numerical Results.	145
5.4.1	First Simulation	147
5.4.2	Second Simulation	147
5.5	Applications and Future Works	147
5.6	Conclusion	147
5.7	Appendix	148
	References	148
IV	EPILOGUE	151
6	Conclusions and Future Work	153
6.1	Conclusions.	153
6.1.1	SS Based Localization	153
6.1.2	PLE Self-Estimation in Wireless Networks	155
6.2	Suggestions for Future Work	155
6.2.1	Localization	156
6.2.2	PLE Self-Estimation	157
	References	158
	Curriculum Vitae	161
	List of Publications	163

SUMMARY

WIRELESS communications and networking are gradually permeating our life and substantially influencing every corner of this world. Wireless devices, particularly those of small size, will take part in this trend more widely, efficiently, seamlessly and smartly. Techniques requiring only limited resources, especially in terms of hardware, are becoming more important and urgently needed. That is why we focus this thesis around analyzing wireless communications and networking based on signal strength (SS) measurements, since these are easy and convenient to gather. SS-based techniques can be incorporated into any device that is equipped with a wireless chip.

More specifically, this thesis studies **SS-based localization** and **path-loss exponent (PLE) self-estimation**. Although these two research lines might seem unrelated, they are actually marching towards the same goal. The former can easily enable a very simple wireless chip to infer its location. But to solve that localization problem, the PLE is required, which is one of the key parameters in wireless propagation channels that decides the SS level. This makes the PLE very crucial to SS-based localization, although it is often unknown. Therefore, we need to develop accurate and robust PLE self-estimation approaches, which will eventually contribute to the improvement of the localization performance. Additionally, our work also provides very useful links to possible applications in other related fields.

In this thesis, we start with the first research line, where we try to cope with all possible issues that we encounter in solving the localization problem. To eliminate the unknown transmit power issue, we adopt differential received signal strength (DRSS) measurements. Colored noise, non-linearity and non-convexity are the next three major issues. To deal with the first two, we introduce a whitened linear data model for DRSS-based localization. Based on that and assuming the PLE is known, three different approaches are respectively proposed to tackle the non-convexity issue: an advanced best linear unbiased estimator (A-BLUE), a Lagrangian estimator (LE) and a robust semi-definite programming (SDP)-based estimator (RSDPE). Note that the RSDPE is particularly designed to be robust against the model uncertainties (imperfect PLE and inaccurate anchor location information) while the A-BLUE and the LE are based on an exact data model. We thoroughly compare them from different perspectives and conclude they have their own advantages: the A-BLUE has the lowest computational complexity; the LE holds the best accuracy for a small measurement noise; and the RSDPE yields the best performance under a large measurement noise and possesses a very good robustness against model uncertainties. Moreover, to cope with an unknown PLE, we propose a robust SDP-based block coordinate descent estimator (RSDP-BCDE) that jointly estimates the PLE and the target location. Its performance iteratively converges to that of the RSDPE with a known PLE.

As mentioned earlier, while generating DRSS measurements, we eliminate the unknown transmit power. This is very similar to the way time-difference-of-arrival (TDOA)

methods cope with an unknown transmit time. Both of them use a differencing process to cope with an unknown linear nuisance parameter. Our DRSS study shows the differencing process does not cause any information loss and hence the selection of the reference is not important. However, this apparently contradicts what is commonly known in TDOA-based localization, where selecting a good reference is very crucial. To resolve this conflict, we introduce a unified framework for linear nuisance parameters such that all our conclusions apply to any kind of problem that can be written into this form. Three methods that can cope with linear nuisance parameters are considered by investigating their best linear unbiased estimators (BLUEs): joint estimation, orthogonal subspace projection (OSP) method and differential method. The results coincide with those obtained in our DRSS study. For TDOA-based localization, it is actually the modelling process that causes a reference dependent information loss, not the differencing process. Many other interesting conclusions are also drawn here.

Next, we turn our attention to the second research line. Undoubtedly, knowledge of the PLE is decisive to SS-based localization and hence accurately estimating the PLE will lead to a better localization performance. However, estimating the PLE also has benefits for other applications. If each node can self-estimate the PLE in a distributed fashion without any external assistance or information, it might be very helpful for efficiently designing some wireless communication and networking systems, since the PLE yields a multi-faceted influence therein. Driven by this idea, we propose two closed-form (weighted) total least squares (TLS) methods for self-estimating the PLE, which are merely based on the locally collected SS measurements. To solve the unknown nodal distance issue, we particularly extract information from the random placement of neighbours in order to facilitate the derivations. We also elaborate on many possible applications thereafter, since this kind of PLE self-estimation has never been introduced before.

Although the previous two methods estimate the PLE by minimizing some residue, we also want to introduce Bayesian methods, such as maximizing the likelihood. Some obstacles related to such approaches are the totally unknown distribution for the SS measurements and the mathematical difficulties of computing it, since the SS is subject to not only the wireless channel effects but also the geometric dynamics (the random node placement). To deal with that, we start with a simple case that only considers the geometric path-loss for wireless channels. We are the first to discover that in this case the SS measurements in random networks are *Pareto* distributed. Based on that, we derive the CRLB and introduce two maximum likelihood (ML) estimators for PLE self-estimation. Although we considered a simplified setting, finding the general SS distribution would still be very useful for studying wireless communications and networking.

Finally, we wrap up this thesis by summarizing our research results and providing suggestions for future work.

ACKNOWLEDGEMENTS

MY Ph.D journey in Delft University of Technology (TU Delft) that spans more than four years finally comes to the end. This thesis embodies not only the essence of our work but also all the efforts that have been made during the past years. Hereby, I would like to express my most sincere gratitude to all those people who have encouraged, supported and accompanied me throughout this journey.

First of all, I would like to wholeheartedly thank my dear advisor, Prof. dr. ir. Geert Leus. He is the saviour of my Ph.D. Less than one year since I came to this university in 2011, I had to face the decision of department restructuring. After that, I had stayed in two groups until I finally settled down in the current one, the Circuits and Systems (CAS) group. Geert is the one who brought me here, set my Ph.D journey on the right path and saved me from the toughest moments in the recent years. He showed up just like a shining beacon to a fisherman who is sailing alone on a ferocious ocean in the dark night, lost, broken and desperate. Just for that, I will always be indebted to him. Moreover, he is very rigorous and exceptionally smart; he is also a generous gentleman with a nice sense of humour; more importantly, he is a patient listener and an incredibly open-minded learner who can quickly understand the speaker, accept his new idea and accurately grasp its essence. In spite of my often absence in the office, I really enjoy talking, making jokes and discussing with him. I deeply thank you Geert for the wise guidance and the precious supports these years. They will be my most cherished treasure and memories in the future.

I would also like to thank Prof. Alle-Jan van der Veen, the head of CAS group, and all the other members for creating such a friendly and nice research environment and providing the facilities that I could access during my Ph.D period. I treasure every second of the wonderful discussions in the office, the interesting chats during coffee breaks and the unforgettable activities in the group outings. Those are all the invaluable memories that will accompany me for the rest of my life and the endless resources, from which I can always gain strength. I would like to thank Yan Xie, Shahrzad Naghibzadeh, Jiani Liu, Venkat Roy, Sundeep Prabhakar Chepuri, Jie Zhang, Millad Sardarabadi, Seyran Khademi, Jorge Martinez-Castaneda, Raj Thilak Rajan, Dyonisius Dony Ariananda and etc.. I do extend my appreciation to Prof. dr. Tadashi Ebihara, Dr. Jing Han and etc., those who have shortly visited our group. Particularly, I want to thank our secretariat Minaksie Ramsoekh for helping everyone not only in the office manners but also in many of the life aspects and Antoon Frehe for his amazing technical assistance. Their presence makes this group more reliable and united.

I much appreciate all the other members in my defence committee, Dr. ir. J. P. A. Romme, Dr. S. Chatzinotas, Prof. dr. H. Wymeersch, Dr. ir. C. C. J. M. Tiberius and Prof. dr. ir. K. L. M. Bertels, for carefully reviewing my thesis and providing the insightful and constructive comments, which are very helpful for improving the quality of my thesis. It is also my greatest honour that they can attend my Ph.D defence and be part of what is

definitively one of the most important moments in my life.

Next, my gratitude goes to those people who had helped me before I joined the CAS group. Hereby, I would like to thank Prof. dr. ir. Ignas Niemegeers and Dr. RangaRao Venkatesha Prasad (VP) for bringing me to this prestigious university in 2011. This was a critical point that indicates the beginning of my Ph.D journey. They kindly introduced me to the Telecommunications group, where I have met many talented colleagues. I would like to thank Dr. Ertan Onur, who is my first advisor in TU Delft and now an associate Professor at Middle East Technical University. He not only instructed me on the academic fields but also helped and protected me in all aspects of life. I would also like to thank Qing Wang, Huazhou Shi, Kishor Chandra, Vijay Rao, Diptanil Debbarma and etc., for their company and all the happy and memorable time together. Then, I want to thank Prof. dr. Koen Langendoen, the chairman of the Embedded Software group, for accepting Dr. Ertan Onur and me in your group after the department restructuring and allowing me to transfer to the CAS group later. Finally, I appreciate the help from the Graduate school of Faculty of Electrical Engineering, Mathematics, and Computer Science.

I also take this opportunity to thank quite a few friends that I am very fortunate to have in the Netherlands. I want to thank Yao Liu and Hao Huang who are my first roommates since I came to this country. We together overcame many difficulties while adapting to a totally new society with a different culture. A special thank goes to Nahaheya who is now Hao Huang's wife for her helps in the process. I would like to thank Qing Wang for allowing me to kidnap his X-Box 360 playstation for more than two years. I want to thank Shizhe Zhang for his tremendous help in my daily life. He is a very loyal and wonderful roommate and, although it may sound weird, the time that we lived under the same roof is even longer than I did with my wife so far. I would like to thank Yuanjie Yu for his always positive life attitude that keeps influencing the others around. I am specially thankful to Yan Xie who sat next to me in the office. We both have witnessed the efforts that each other has been made on the work in order to pursue the Ph.D. My gratitude list contains a lot more than the aforementioned people, though I will not refer to all of them to save space.

The most fortunate thing happened in my Ph.D journey, probably even for my whole life, was meeting a beautiful angel who sat next to me on the flight KL987 from Amsterdam to Beijing on January 20th, 2014. The whole things on the plane and thereafter were like intentionally, carefully and romantically orchestrated by God. Both studying for Ph.D in the Netherlands, being a Chinese yet interested in the western history, craving for travelling around the world and enjoying the life full of music and freedom, she suddenly came to my prosaic life and made it full of colours and meanings. In the last year at Place de la Concorde in Paris, when the sky wheel rose us to the top, she accepted my proposal and became my beloved wife. The angel has a graceful and elegant name, Sarah (Xiao) Xun, and my gratitude to her would never be enough from the day we first met to the end of time.

Back to China, I would like to thank my family for their many-year selfless supports. Your constant encouragements always keep me stepping forward bravely and determinedly. Particularly, I am genuinely grateful to my parents, Junliang Hu and Yuee Long, for their unconditional love and dedication. Moreover, I want to express my deepest

apology to my grandfather for not being there when he passed away in 2014. I was not informed of this devastating news until the next year when I came back to China. My family kept the unimaginable sadness to themselves in order to protect me, a grandson in a foreign country who could do nothing at the moment. So far, all I know is that he left so suddenly, but very peaceful. Maybe he is blessing everyone from the heaven now. By the end of this year, I would have embraced two more lovely nephews or nieces. I wish these two little babies always safe and healthy. Finally, I also want to extend my appreciation to my relatives and their families.

Delft, September 2017
Yongchang Hu

I

PROLOGUE

1

INTRODUCTION

None is of freedom or of life deserving unless he daily conquers it anew.

Erasmus

SINCE Guglielmo Marconi successfully telegraphed the first message in Morse code, a new era of wireless communication has been opened. Driven by the simple goal of improving the quality of our life, generations of scientists and engineers have been following his legacy and never ceased to contribute in this field. With great respect to their contributions, we would also like to follow this path and offer our strength for a better future.

Therefore, our work mainly contributes to the fields of localization and wireless channel sensing. Accordingly, this thesis is respectively comprised of two research lines: **signal strength (SS) based localization** and **path-loss exponent (PLE) self-estimation in wireless networks**. For the convenience of reading, this chapter introduces the background and the motivations of our research, where we will also highlight our research challenges in red blocks. Furthermore, while reading this thesis, the readers are also recommended to refer to our research diagram in Fig. 1.8 at page 30 for a general perspective. Finally, based on that, we outline this thesis and our contributions.

1.1. BACKGROUND AND MOTIVATIONS

LOCATION-awareness has become an indispensable feature for many aspects of commercial, public service and military sectors [1, 2]. Hence, it is already a wide and essential concern for enabling applications such as tourist guiding [3], health-care monitoring [4], animal and asset tracking [5, 6], emergency service [7], etc.. Driven by this urgent requirement, much effort has been put on the topic of localization or positioning. As one of the solutions, the global positioning system (GPS), which is assisted by satellites, provides a considerable estimation accuracy in most outdoor scenarios [8]. However, in some unattended, hostile or very severe environments, the performance of

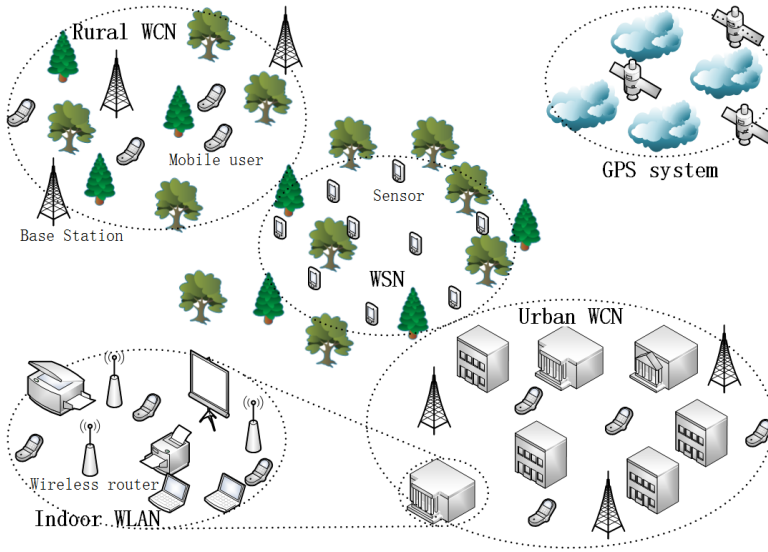


Figure 1.1: Demonstration of different localization scenarios: The localization is assisted by nearby anchors while the GPS system heavily relies on faraway satellites in space. Any kind of node with known location can act as an anchor, e.g., the base station in WCNs, the wireless router in WLANs and even the sensor equipped with a GPS receiver.

GPS will deteriorate significantly [9, 10]. Moreover, if indoor scenarios are considered, the GPS signal will become even worse and hence unreliable for localization.

As an alternative solution, localization that tackles GPS-denied scenarios is emerging rapidly and showing prominent features with many notable advantages in recent years [1, 11–14]. To be specific, unlike GPS, this kind of localization does not rely on orbiting satellites in space, whose locations are changing and should be periodically calculated in time by the ephemeris. Instead, some surrounding nodes with *a priori* known locations, i.e., the anchors, take the role of the satellites. Since the anchors are relatively near, as shown in Fig. 1.1, the localization signal is less distorted and weakened, which might lead to a better localization performance. Besides, constructing a localization network is very easy, cheap and rather scalable in size, such that any kind of network can feature and accordingly benefit from the location-awareness.

1.1.1. LOCALIZATION

LET us focus on the first line of our research: **signal strength (SS) based localization**. For a better understanding of localization, we will elaborate on scenarios, basic concepts, measurements and some mathematical issues while solving the localization problem. Finally, we make a further step and extend our research to differential signal processing, which is closely related to localization and provides some more insights.

DIFFERENT SCENARIOS

WE would like to start with some discussions about practical scenarios. As depicted in Fig. 1.1, localization is mostly implemented in wireless cellular networks (WCNs),

wireless local area networks (WLANs) and wireless sensor networks (WSNs).

- **WCNs** primarily depend upon wide area localization, e.g. for emergency services, since the infrastructure of the underlying networks covers a large geographic area [15]. The communication environment varies from the urban area to the rural area, where the former yields relatively more reflections and attenuation of the localization signal. In both cases, the base station (BS) acts as the anchor. To facilitate the localization, the network is often well-equipped and cooperative.
- Source localization for **WLANs** is actively considered for indoor scenarios such as tourist guiding in museums [16, 17]. Basically, any wireless device with known location can be chosen as the anchor. Though, the wireless router is the most favorite, since it is often carefully attended and networked for a large amount of data exchange. Compared with WCNs, WLANs encounter more complicated and severe communication environments.
- If communication environments are hostile or unattended, **WSNs** are often constructed temporarily [18, 19]. For instance, a large number of sensors are randomly scattered for environmental monitoring, where the geographical location of the sensors needs to be known. In this case, localization has to be carried out with very limited resources and cope with rather severe signal distortions. To be specific, unlike WCNs and WLANs, many more issues have to be taken into consideration such as battery constraints, limited data exchange, etc..

In a nutshell, different underlying features and configurations of wireless localization networks yield different practical concerns, which keep motivating us to extend our research to more realistic cases. For instance, the reliability of the anchor location information is just one of our concerns.

Research Challenge No. 1: Inaccurate Anchor Location Information

The anchor location information can be provided by the GPS system, which might not be accurate enough especially in indoor scenarios. In military scenarios, such important information might also be faked, spoofed and tampered with to sabotage the localization, as shown in Fig. 1.2. Therefore, inaccurate anchor location information is a rather critical issue, which needs to be taken into account.

Question: How does the inaccurate anchor location affect the localization and how do we deal with that?

Answer: When solving a localization problem, inaccurate anchor location information actually results into uncertainties in our data model. Therefore, in Section 2.3.3 of Chapter 2, we introduce a robust localization method that copes with the model uncertainties. Note that model uncertainties can also be caused by other inaccurate information such as an inaccurate PLE estimate, which will be discussed later.

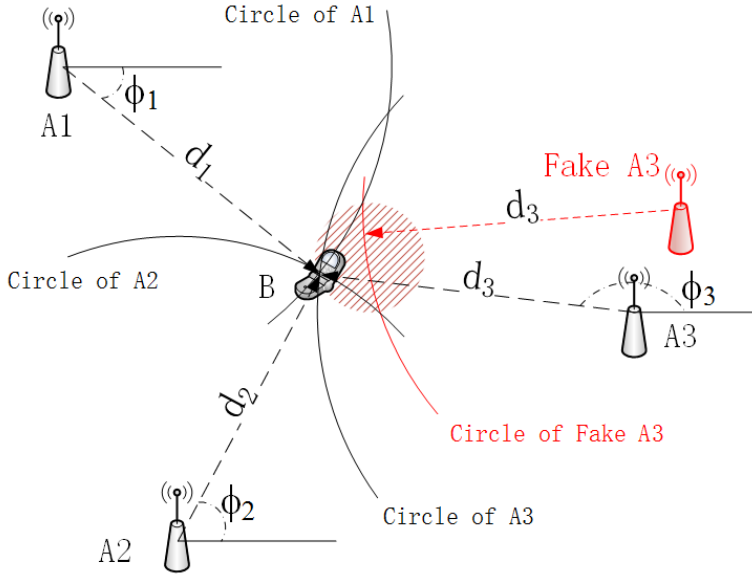


Figure 1.2: Demonstration in \mathbb{R}^2 of two localization principles and the impact of inaccurate anchor location information: the nodes A_1 , A_2 and A_3 are three anchors with known locations; the corresponding transmission distances to the target node B are d_1 , d_2 and d_3 ; the corresponding transmission angles are ϕ_1 , ϕ_2 and ϕ_3 . For *trilateration*, the location B is determined at the intersection of the solid circles centered at A_1 , A_2 and A_3 with radii d_1 , d_2 and d_3 , respectively. For *triangulation*, the location B is determined at the intersection of the dashed radial lines originated at A_1 , A_2 and A_3 with angles ϕ_1 , ϕ_2 and ϕ_3 , respectively. An adversary might attack the localization system by reporting the wrong localization information. The target node B will then be estimated much deviated from the true location, e.g., around the red shaded area.

BASIC CONCEPTS AND MEASUREMENTS

NEXT, we discuss some basic concepts of localization. In most scenarios, two localization principles are commonly considered: *trilateration* [20] and *triangulation* [21]. To be specific, the former relies on the distance information while the latter depends on the angular information, as depicted in Fig. 1.2. *Triangulation* is based on angular measurements such as angle-of-arrival (AOA) information, which normally requires multiple antennas. Although it is possible for WCNs and WLANs, it is still very expensive, especially for WSNs, to use the AOA for localization. In contrast, as long as there are measurements that contain distance information, *trilateration* is applicable and hence preferably chosen in most cases. Therefore, from now on, the term “localization” will only refer to *trilateration* for convenience, since this is the main focus of our research.

Before elaborating on the measurements, we first notice that localization can be carried out in the two following fashions: *centralized* and *distributed* [13], as depicted in Fig. 1.3.

- In a *centralized* fashion, each anchor collects measurements from the target node and then a central processor gathers all the measurements to calculate the location of the target node. Usually, this kind of localization network is not very scalable,

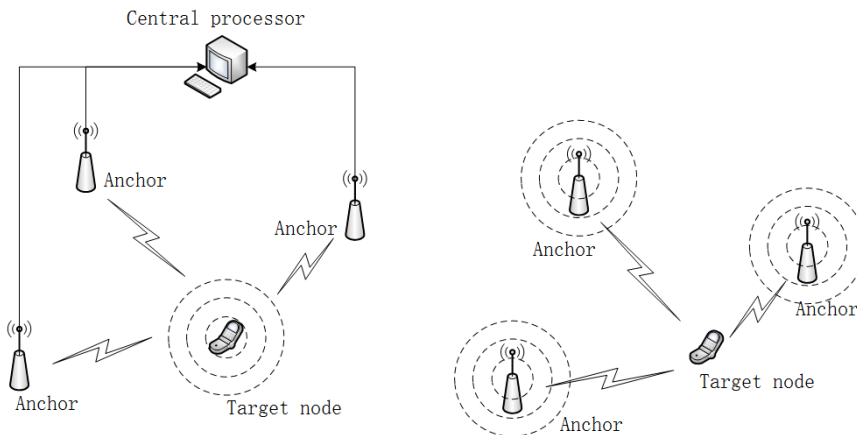


Figure 1.3: *Centralized* localization is demonstrated on the left side, where a single localization signal is broadcast by the target node and received by all the anchors. In this case, all measurements should be aggregated to the central processor. *Distributed* localization is shown on the right side, where each anchor transmits its own localization signal and hence the target node receives the measurements from different transmitters. In this case, the target node can self-estimate its location.

since aggregating the required information such as anchor locations and measurements might need a great deal of node collaborations, thus causing unnecessary overheads and even congestions especially in large networks.

- *Distributed* localization is often more attractive, since there is no central controller and the target node can infer its own position only based on locally collected information, also called *self-localization*. *Self-localization* works individually and independently with less external assistance. The target can even just listen and not participate into the localization network. As a result, it can not only reduce the networking load but also be invisible to any other node if required.

These two kinds of methods can be simply distinguished by observing who transmits the localization signal, the anchor or the target. This difference yields a notable influence, which will be frequently observed later. Note that the signal receiver can also send back a response signal to the transmitter, which might be helpful for improving the localization performance. However, this kind of two-way (TW) based localization requires intensive node cooperation and communications [22]. More importantly, the unknown response time might also become a serious issue. Therefore, in this thesis, we only consider the previous one-way fashion for localization, where no response signal is sent, and leave the study of TW based localization for future work.

Now, we discuss the measurements for localization, where the popular types include time-of-arrival (TOA), time-difference-of-arrival (TDOA) and signal strength (SS).

TOA The TOA is the measured time, at which a signal first arrives at a receiver. For synchronized networks, the transmit time, say t_0 , needs to be known to the receiver in

order to calculate the time of flight (TOF) or equivalently the transmission distance. To be specific, let t_i be the TOA measurement associated with the i -th anchor located at $\mathbf{s}_i \in \mathbb{R}^2$. Then, the estimate of the corresponding transmission distance $d_i = \|\mathbf{x} - \mathbf{s}_i\|_2$ can be expressed as

$$\hat{d}_i \triangleq c \left(\underbrace{t_i}_{\text{referenced to the receiver clock}} - \underbrace{t_0}_{\text{referenced to the transmitter clock}} \right) = \|\mathbf{x} - \mathbf{s}_i\|_2 + n_i, \quad (1.1)$$

where the target location $\mathbf{x} \in \mathbb{R}^2$ is the parameter to estimate, c is the speed of light, $\|\cdot\|_2$ indicates the Euclidean norm and n_i is the measurement noise. Obviously, clock synchronization plays a very significant role, since t_i and t_0 are related to different clocks.

- In a *centralized* localization network, where the localization signal is transmitted only by the target node, there only exists a single transmit time t_0 for all measurements. However, clock synchronization should be carried out, not only among the anchors, but also between the target and the anchors.
- In a *distributed* system, multiple localization signals are transmitted by the anchors. Obviously, the synchronization requirement does not become any less strict. Moreover, the anchors also have to guarantee the same t_0 , which might be practically infeasible, otherwise different transmit times need to be considered.

Additionally, in whatever case, the information of t_0 should also be conveyed to the signal receiver for calculating the location, which causes extra network load.

TDOA In order to relax the aforementioned constraints, TDOA is widely considered, allowing an unknown transmit time t_0 . By selecting a reference, say the j -th TOA, the TDOA is obtained as

$$t_{i,j} \triangleq \underbrace{t_i - t_j}_{\text{referenced to the receiver clock}}, \quad i \neq j, \quad (1.2)$$

where the unknown t_0 is cancelled out. Obviously, this makes TDOA based localization more independent of the localization signal transmitter, since clock synchronization between the target node and the anchors is not required any more. Note that TDOA as in (1.2) can be directly computed using signal correlations without computing any TOA [23, 24]. Then, the estimate of the distance can be expressed as

$$\hat{d}_{i,j} = c t_{i,j} = \|\mathbf{x} - \mathbf{s}_i\|_2 - \|\mathbf{x} - \mathbf{s}_j\|_2 + n_{i,j}, \quad (1.3)$$

where $\hat{d}_{i,j} \triangleq \hat{d}_i - \hat{d}_j$ and $n_{i,j} \triangleq n_i - n_j$.

- In a *centralized* network, clock synchronization is required only between the anchors.
- In a *distributed* network, since both t_i and t_j are measured at the target node, it appears that no clock synchronization is required this time. But, recall that this is under the condition of the same transmit time t_0 . As already mentioned, in order to meet this condition, the anchors need to be synchronized again.

After all, it is readily observed that clock synchronization is the Achilles' heel of TOA or TDOA based localization. Once the network is not synchronized, this will incur an unimaginable impact on the localization performance. Furthermore, a reliable clock synchronization requires frequent information exchanges and intensive cooperations between nodes, which is also a rather tough requirement.

On account of practical simplicity and convenience, our work is mainly focused on localization using the SS, the last type of measurement. Compared with the former two measurements, employing SS measurements has the following advantages:

- The SS is rather convenient to collect and we will discuss this in Section 1.1.3. Moreover, most wireless devices are also equipped with a received signal strength indicator (RSSI).
- It does not rely on any external assistance like clock synchronization or peak detection (for determining the signal arrival if ultra-wideband (UWB) is used).
- Unlike the transmit time, the transmit power is often standardized in wireless devices, constant over time and hence relatively easy to obtain.

Therefore, this topic has gradually become an emerging and popular topic in recent years. In this thesis, we further categorize the SS as received signal strength (RSS) and differential received signal strength (DRSS).

RSS The RSS is the measured power of the transmitted signal at the receiver, which is mainly subject to the geometric path-loss and the *shadowing* effect. Commonly considered to be *log-normally* distributed [25], the RSS associated with the i -th anchor can be expressed in *decibels* as

$$P_i \triangleq \underbrace{P_0}_{\text{transmitter configuration}} - \underbrace{10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{d_0} \right)}_{\text{geometric path-loss}} + \underbrace{\chi_i}_{\text{shadowing effect}}, \quad (1.4)$$

where P_0 is the received power at the reference distance d_0 , γ is the path-loss exponent (PLE) and $\chi_i \sim \mathcal{N}(0, \sigma^2)$ is the shadowing effect. Without any loss of generality, we assume that $d_0 = 1$ m for convenience and hence P_0 can equivalently be regarded as the transmit power. Although the value of P_0 still has to be conveyed to the signal receiver for calculating the target location, it is already more convenient than coping with the transmit time t_0 , since P_0 might be standardized for all transmitters and hence is easily accessible as already mentioned.

In either a *centralized* or *distributed* localization network, the value of P_0 can be piggybacked on the localization signal and conveyed to the receivers. However, this will still cause some burden for networking.

DRSS In our work, we actually consider the DRSS measurement rather than the RSS measurement for localization. Because, while preserving all the advantages of RSS-based localization, DRSS-based localization is more independent of the localization signal transmitter, similar to the TDOA case. To be specific, we can expect the following advantages.

- The implicit dependence of localization on the signal transmitter, which could be defective, malicious or uncooperative, is significantly alleviated.
- It can minimize or even does not require an overhead control message between anchors and the target node. This will save energy, bandwidth and throughput, which are very scarce resources for wireless networking.
- The localization process can be well concealed from the signal transmitter, which is very beneficial to surveillance or military applications.

In order to gain the aforementioned benefits, we must allow for an unknown P_0 . Then, selecting the j -th RSS as reference, we obtain the DRSS as

$$P_{i,j} \triangleq P_i - P_j = \underbrace{-10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{\|\mathbf{x} - \mathbf{s}_j\|_2} \right)}_{\text{geometric path-loss}} + \underbrace{\chi_{i,j}}_{\text{shadowing effect}}, \quad (1.5)$$

where $\chi_{i,j} \triangleq \chi_i - \chi_j$. Note that the unknown P_0 is eliminated.

It is worth noting that the DRSS corresponds to the distance ratio whereas the TDOA corresponds to the distance difference. Unlike TDOA-based source localization, which has already been intensively studied, research on DRSS-based localization is still in its infancy, thus requiring more attention. This is another reason why we consider DRSS measurements for localization in Chapter 2.

Research Challenge No. 2: Different Unknown Transmit Powers

Although the transmit power might be standardized for wireless communications, some techniques like transmit power control could still be carried out in order to save energy or guarantee signal coverage [26–28]. This results into different transmit powers. Moreover, some unpredictable power surge or system instabilities might also lead to the same issue.

Question: How do we tackle different unknown transmit powers and what is the impact on DRSS-based localization?

Answer: In order to be more realistic, different unknown transmit powers are taken into account for DRSS-based localization in Chapter 2. We reasonably assume that the different unknown transmit powers are *normal* distributed around an unknown nominal averaged power in *decibels*. Then, we show that those deviations of the unknown transmit powers can be incorporated into the *shadowing* effect and thus can be regraded as measurement noise.

OPTIMIZATION PROBLEMS AND DISCUSSIONS

As far as the DRSS localization problem is concerned, there exist two common facts:

- The target location is hidden inside a distance norm, i.e., $d_i = \|\mathbf{x} - \mathbf{s}_i\|_2, \forall i$.
- The target node cannot physically overlap with the anchors, i.e., $\mathbf{x} \neq \mathbf{s}_i, \forall i$.

These two facts respectively incur two mathematical issues when solving the DRSS localization problem: non-linearity and non-convexity.

Non-linearity The non-linearity issue is readily observed from the DRSS measurement in (1.5).

Research Challenge No. 3: Non-Linearity

Apparently, (1.5) is not a linear function w.r.t. \mathbf{x} , otherwise the target location \mathbf{x} can be directly estimated by applying a (weighted) least squares (LS) estimator.

Question: In order to estimate the target location \mathbf{x} , how can we deal with the non-linearity issue?

Answers: A linearization procedure can be considered, though it may proceed in different fashions, which will be discussed below. We will not point out and explain which way we consider for DRSS-based localization at this point, since the impact of different linearizations will also be passed on to the non-convexity issue. Therefore, our solutions to the non-linearity and the non-convexity will be discussed together thereafter.

Here, we elaborate on two common ways for linearization as follows.

- The first way applies a first-order Taylor series expansion around a given expansion point. In order to keep a small approximation error, this expansion point should be chosen as close as possible to the true target location \mathbf{x} and thus can also be regarded as an estimate. Obviously, some information is lost when dumping the high-order terms, which however can be reduced with a better expansion point. Therefore, the estimate is often iteratively updated to reduce the information loss and hence to obtain a better estimation accuracy. Famous examples are maximum likelihood (ML) estimators that use the Newton method, and those using the extended Kalman filter (EKF) for mobile scenarios [29–39].
- The other way unfolds the distance norm as $d_i^2 = \|\mathbf{x}\|_2^2 + 2\mathbf{s}_i^T \mathbf{x} + \|\mathbf{s}_i\|_2^2$ and treats $R \triangleq \|\mathbf{x}\|_2^2$ as a new unknown parameter [40–50]. As a result, a new linear localization problem can be formulated after some manipulations. In most cases, in order to construct a zero-mean model noise, some approximations have to be used, which cause some information loss.

These two kinds of linearization for localization are studied in Section 3.3 of Chapter. 3.

Non-convexity For a better demonstration, we directly apply the least squares (LS) criterion to (1.5) and formulate the optimization problem as

$$\min_{\mathbf{x}} \sum_{i=1}^{N-1} \left[P_{i,j} + 10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{\|\mathbf{x} - \mathbf{s}_j\|_2} \right) \right]^2, i \neq j, \quad (1.6)$$

where N anchors are considered. Here comes the other mathematical issue, the non-convexity.

Research Challenge No. 4: Non-Convexity

There exists a singularity issue at every anchor location. More specifically, if $\mathbf{x} = \mathbf{s}_i$, for some i , the logarithmic term in (1.6) will be either positive or negative infinity. As a result, the cost function will have multiple minima, resulting into non-convexity [51]. Also see the examples demonstrated in Fig. 2.1 of Chapter 2.

Question: How does the non-convexity affect the localization problem and what are the solutions?

Answers: The impact of the non-convexity on the localization problem varies with the linearization procedure. We will elaborate on that below and also discuss our solutions to the non-linearity and the non-convexity together.

Taking the earlier linearization into account, the non-convexity might manifest itself differently:

- If the linearization is carried out by a Taylor series expansion, the non-convexity lies in the choice of the initial expansion point. Since there exist multiple minima for the cost function in (1.6), the target location estimate might converge to a local solution if an inappropriate initial expansion point is selected, which also leads to a large information loss caused by the linearization.
- When a new linear localization problem is constructed by unfolding the distance norm, the new parameter vector that includes R is bound to a non-convex set. This means that we also have to consider the relation between R and \mathbf{x} when solving the localization problem.

Obviously, if we choose the first way to linearize (1.5), the only solution to the non-convexity is guaranteeing an appropriate initialization to avoid the local solutions. However, this is rather difficult in practice. Moreover, in order to reach a certain estimation accuracy, a large number of iterations is required resulting into a high computational complexity.

In Chapter 2, we consider the second kind of linearization for DRSS-based localization, since a new linear data model is readily and immediately to use although the new parameter vector is still bound to a non-convex set. Ignoring this constraint, we can directly obtain an unconstrained (weighted) LS estimator, which is quite convenient for a real-life implementation.

Of course, we will not be satisfied without further resolving the non-convexity issue. Recall that this is equivalent to considering the relation between R and \mathbf{x} . To do so, we present three kinds of methods:

- **Two-step Estimation:** The first step of the estimation uses the above unconstrained (weighted) LS estimator such that the second step can consider the relation between R and \mathbf{x} by fine-tuning the first-step estimate of \mathbf{x} . For more details, please refer to Section 2.3.1.
- **Lagrangian Estimation:** First, a Lagrangian multiplier is introduced to incorporate a transformed constraint that considers the relation between R and \mathbf{x} into the

cost function for the localization problem [51]. Then, in order to minimize the cost function and estimate \mathbf{x} , a trust region is also provided for efficiently finding the unknown Lagrangian multiplier [52–54]. Please refer to Section 2.3.2 for more details.

- **Semi-Definite Relaxation (SDR):** First, the relation between R and \mathbf{x} is equivalently reformulated into a linear matrix inequality (LMI) and a rank constraint [55]. Then, dropping this rank constraint can relax the new parameter vector onto a convex set, such that semi-definite programming (SDP) can be applied [56]. More details are presented in Section 2.3.3.

We also compare and study these three methods in a more general context in Chapter 2.

Correlated Noise Besides the two aforementioned issues, we need to notice another issue here, i.e., the measurement noise $\chi_{i,j}$ in (1.5) is correlated.

Research Challenge No. 5: Correlated Noise

Compared with the shadowing effect χ_i in (1.4), the differencing process that generates the DRSS measurement not only removes the unknown transmit power P_0 but also colours the measurement noise $\chi_{i,j}$. The coloured measurement noise significantly degrades the localization performance, which however is often ignored in literature [57–60].

Question: How do we eliminate the impact of the coloured measurement noise?

Answers: The coloured measurement noise results into a coloured model noise for the localization problem. Therefore, whitening the model noise is very important for a better performance. For that, we particularly introduce a whitened model for DRSS-based localization in Section 2.2 of Chapter 2 and further discuss the significance of the whitening procedure in Chapter 3, especially for the DRSS and TDOA measurements which are generated by a differencing process.

Finally, we would like to discuss some practical concerns that also influence the optimization problem for DRSS-based localization. They are the unknown path-loss exponent (PLE) γ and the unknown transmit power P_0 .

Unknown Path-Loss Exponent According to the geometric path-loss, the signal power (in *watts*) exponentially decays over the distance. Obviously, the PLE is a key parameter in the radio propagation channel, which is mostly unknown in real-life unless in the ideal free space ($\gamma = 2$). Moreover, since communication environments are complicated and time-varying, the PLE often varies over time, location and scenario [25]. Therefore, getting grip on the PLE information is very essential not only for localization using SS measurements, but also for many other wireless communications and networking designs.

Here, we only focus on the localization and note that the unknown PLE is usually estimated beforehand. In fact, this kind of PLE estimation belongs to our second research line, i.e., **path-loss exponent (PLE) self-estimation in wireless networks**. Therefore, we will not further discuss it here, but only emphasize the fact that, if this kind of PLE estimation is not reliable or maliciously sabotaged, an inaccurate PLE estimate might be used in the localization phase, incurring a considerable impact.

Research Challenge No. 6: Inaccurate PLE Estimate

Currently, most PLE estimation approaches heavily rely on the assistance of anchors. Once the anchor location information is inaccurate or tampered with by adversaries, this might lead to a terrible estimation result. On the one hand, we should improve those PLE estimation methods, which is one of the motivations of our second research line. But on the other hand, we should also consider the impact of an inaccurate PLE estimate when designing localization techniques.

Question: How does an inaccurate PLE estimate affect the localization problem and what is our solution?

Answers: While solving the localization problem, we notice that an inaccurate PLE estimate causes model uncertainties, which is similar to the mentioned case of inaccurate anchor location information. Therefore, we tackle an inaccurate PLE estimate and anchor location information together by proposing a robust approach against general model uncertainties for DRSS-based localization in Section 2.3.3 of Chapter 2

Next, we make a further step to consider the case when the PLE estimation is not available before the localization phase, i.e., the PLE is totally unknown to the localization problem.

Research Challenge No. 7: Unknown PLE

First note that estimating the unknown PLE is actually the fundamental challenge for our second research line and we will come back to that in Section 1.1.2. However, in some cases, the PLE estimation before the localization phase might still be very costly, difficult or unreliable, especially in unattended and hostile communication environments. Although our second research line is exactly aimed at preventing this kind of situation, it would be still better for us to exploit a DRSS-based localization method that copes with a totally unknown PLE.

Question: How do we locate the target node using DRSS measurements without knowing the PLE?

Answers: The solutions that estimate the unknown PLE before the localization phase will be discussed later in our second research line. During the localization process, we need to jointly estimate the unknown target location and PLE. To achieve that, a new approach is proposed in Section 2.4 of Chapter 2.

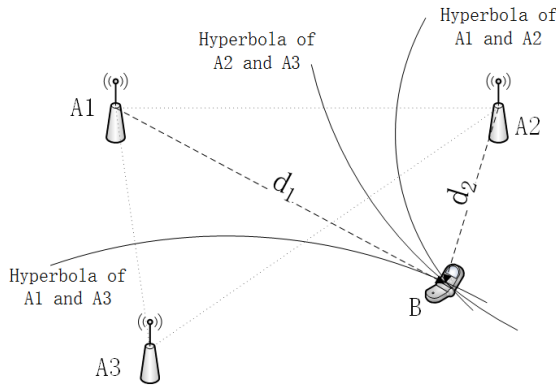


Figure 1.4: Demonstration in \mathbb{R}^2 of hyperbolic localization: the nodes A1, A2 and A3 are three anchors with known locations; d_1 , d_2 are the transmission distances from A1, A2 to the target node B, respectively. The location B is determined at the intersection of three hyperbolas with foci at A1 and A2, A1 and A3, and A2 and A3, respectively.

Unknown Transmit Power Coping with an unknown transmit power is relatively easy, since this nuisance parameter is linear and additive to the measurement P_i . Some RSS-based localization approaches jointly estimate the unknown P_0 and target location \mathbf{x} [40, 61]. As for the localization based on DRSS measurements $P_{i,j}$, there is no such issue, since the unknown P_0 is already eliminated by the differencing process in order to gain more independence of the signal transmitter for localization. In fact, what interests us here is which kind of measurement is more effective in case of an unknown transmit power.

Research Challenge No. 8: Information Loss?

In terms of information theory, the RSS sample set holds the full data information. The DRSS sample set is constructed by taking differences between the RSS measurements. Therefore, the mentioned problem boils down to one intuitive question below.

Question: Is there any information loss when using a DRSS sample set?

Answers: For the introduced DRSS-based model in Chapter 2, we show that there is no information loss compared with the RSS measurement set by bridging them with an orthogonal operator. Moreover, this result intrigues us to study a more general problem with multiple linear nuisance parameters, thus leading to our work in Chapter 3. We will brief on it next.

DIFFERENTIAL SIGNAL PROCESSING

OUR research did not cease to move forward when the localization problem had been solved. While studying the unknown transmit power issue, we started to wonder whether or not taking differences between observations, referred to as the differencing

process hereafter, will cause any information loss. Or does the differencing process influence the localization performance? At the moment, what we had concluded from DRSS-based localization says no. However, what is commonly known in TDOA based localization implies otherwise, since a similar differencing process is used there to generate a TDOA measurement without the unknown transmit time t_0 . The following facts unveil the above contradiction.

1. First, TOA and TDOA based localization are apparently studied separately following different paths for decades [1, 62]. They even diverge in the basic principle of localization, since, unlike *trilateration*, TDOA localization finds the target location at the intersections of different hyperbolas [58, 59, 63] as depicted in Fig 1.4, thus also called the hyperbolic localization. So far, there is no obvious indication that these two different localization principles can be linked together.
2. Second, tremendous literature heavily exists about TDOA based localization on topics like reference selection [64–66] or constructing an optimal differential observation set [67–69]. Those imply that the choice of the reference node used for taking TDOA measurements is very important. Since TOA based localization does not suffer from any reference-related issue, some reference-related information must be lost when using TDOA measurements.

Obviously, the differencing process is a widely used manipulative trick, and undoubtedly there exist many applications other than localization. Therefore, resolving the above conflictive issue is very important, which might provide some useful insights.

For that, we conduct some more research. Now, our focus is not just limited to a single unknown parameter, but a general model with multiple linear nuisance parameters.

Research Challenge No. 9: Multiple Linear Nuisance Parameters

In the aforementioned localization cases, only a single nuisance parameter is considered. However, estimation problems with multiple nuisance parameters widely exist [42, 45–50, 70, 71]. Since we would like to study the differencing process with multiple nuisance parameters, some questions arise immediately as below.

Question: How can we use the differencing process to eliminate multiple linear nuisance parameters? And, in this general case, will this procedure cause any information loss?

Answers: The general idea will be briefly presented as below and elaborated on in Chapter 3. We further prove that the differencing process will not cause any information loss as in DRSS-based localization. Based on this fact, some more interesting and insightful conclusions have also been drawn therein.

We consider a unified framework for linear nuisance parameters and hence let us denote a linear (or linearized) model with the measurement vector $\mathbf{y} \in \mathbb{R}^N$ and the parameter vector \mathbf{x} transformed by \mathbf{H} as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{G}\mathbf{u} + \mathbf{n}, \quad (1.7)$$

where \mathbf{n} is the zero-mean (whitened) white noise vector (not necessarily *Gaussian*), and the multiple linear nuisance parameters are stacked in $\mathbf{u} \in \mathbb{R}^M$, ($N > M$) that enter the observations through \mathbf{G} . The concatenation of \mathbf{H} and \mathbf{G} is assumed to have full column rank. Also note that, when $\mathbf{G} = \mathbf{1}_N$, (1.7) is reduced to the case of a single nuisance parameter. Obviously, all the obtained conclusions apply to any kind of problem that can be written in the general form of (1.7).

In Chapter 3, we introduce and study a general differential method that tackles M linear nuisance parameters. We will not present all the details here, but only brief the general idea.

First, we write $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_M]$ with \mathbf{g}_k the k -th column vector of \mathbf{G} related to the k -th nuisance parameter u_k ($1 \leq k \leq M$) such that (1.7) can be rewritten as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \underbrace{\mathbf{g}_1 u_1 + \dots + \mathbf{g}_M u_M}_{M \text{ nuisance parameters}} + \mathbf{n}. \quad (1.8)$$

Then, we eliminate the nuisance parameters recursively in the order of u_1, \dots, u_M by the differencing process, although the explicit ordering is not important.

To be more specific, let us focus on the differencing process. For instance, if the j -th observation y_j is selected as the reference, the differencing process is presented as

$$\mathbf{d}_j \triangleq \begin{bmatrix} \vdots \\ y_i - y_j \\ \vdots \end{bmatrix}_{(N-1) \times 1} = \mathbf{\Gamma}_j \mathbf{y}, \quad i \neq j, \quad (1.9)$$

where

$$\mathbf{\Gamma}_j \triangleq \begin{bmatrix} \mathbf{I}_{j-1} & -\mathbf{1}_{(j-1) \times 1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_{(N-j) \times 1} & \mathbf{I}_{N-j} \end{bmatrix}_{(N-1) \times N} \quad (1.10)$$

can be seen as the differential operator with $\mathbf{1}$ the all-one matrix (sizes are mentioned in subscript if needed), and \mathbf{d}_j is the new differential observation set with size reduced to $N - 1$ since j is fixed for every element in \mathbf{d}_j . However, if we want to cancel u_1 , the above differencing process cannot be applied directly unless \mathbf{g}_1 becomes a (scaled) all-one vector $\mathbf{1}_N$, since $\mathbf{\Gamma}_j \mathbf{1}_N = \mathbf{0}$. Therefore, we need to find an operator \mathbf{O}_1 that satisfies $\mathbf{O}_1 \mathbf{g}_1 = \mathbf{1}_N$, and pre-multiply it with (1.8) before the differencing process.

Obviously, the rest of the nuisance parameters should also be eliminated one by one in a similar fashion. Thus, the following important facts should be noted.

- When eliminating u_1 , both the operator \mathbf{O}_1 and the differencing process affect the other parts in (1.8). This will be the same when eliminating the rest of the nuisance parameters.
- The model noise \mathbf{n} becomes coloured as soon as the first differencing process is carried out. Moreover, the covariance matrix of the coloured model noise will only become more complicated with multiple differencing processes.

Therefore, even though we can cancel all the nuisance parameters, how do we express and study this complicated procedure? In Chapter 3, we elaborate on that in details and,

in order to study this differential method, we also investigate two other methods, i.e., the joint estimation method and the orthogonal subspace projection (OSP) based method, which can also cope with multiple nuisance parameters.

- The joint estimation method reformulates (1.7) as $\mathbf{y} = [\mathbf{H} \quad \mathbf{G}] \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} + \mathbf{n}$ and treats $[\mathbf{x} \quad \mathbf{u}]^T$ as a new parameter vector of interest. Obviously, this method estimates the unknown nuisance parameters \mathbf{u} jointly with \mathbf{x} . More importantly, no pre-processing of the original observation is considered, thus preserving the full data information.
- The OSP based method eliminates the impact of \mathbf{u} by pre-multiplying (1.7) with an orthogonal projector, say $\mathbf{P}_{\mathbf{G}}^{\perp}$, since $\mathbf{P}_{\mathbf{G}}^{\perp} \mathbf{G} = \mathbf{0}$. See Chapter 3 for the construction of $\mathbf{P}_{\mathbf{G}}^{\perp}$.

The corresponding best linear unbiased estimators (BLUEs) are exploited to bridge these three methods. Although the relation between the joint estimation method and the OSP based method has been reported before [72, 73], the differential method has never been linked to them in literature.

In short, we can now briefly summarize our research about differential signal processing by answering the following questions:

- How does the differential signal process cope with multiple linear nuisance parameters?
- Compared with the joint estimation and the OSP-based estimation, which preserve the full data information, is there any information loss for the differential method?
- In the differencing process, a reference is chosen for subtracting the observations. Does the choice of the reference affect the estimation performance? If not, how do we explain the TDOA localization case.
- What is the optimal differential observation set for differential signal processing? The differential subset of size $N - 1$ associated with a single reference or the full differential set of size $N(N - 1)/2$ that considers every possible reference?

1.1.2. PATH-LOSS EXPONENT ESTIMATION

OUR two research lines are faced with the same issue as described in *Research Challenge No. 7*, but cope with an unknown PLE γ in (1.4) in different manners. The previously discussed method forces the localization method to jointly estimate the PLE and target location. Now, we pay our attention to the other one that estimates the PLE before the localization phase.

As mentioned before, if this kind of PLE estimation is not reliable, an inaccurate PLE estimate might be passed on to the localization phase, severely deteriorating the localization performance. To deal with that, on the one hand, we should develop a robust localization approach that tackles an inaccurate PLE estimate, which is already included

Table 1.1: Comparison of different kinds of PLE estimation

Methods \ Drawbacks	Anchor Dependence	Intensive node cooperation	Not Pervasive
Anchor-Based [74–76]	✓	✓	✓
Anchor-Free [77, 78]	✗	✓	✓
Self-Estimation [79]	✗	✗	✓
Collective Self-Estimation [80, 81]	✗	✗	✗

in our first line of research. On the other hand, this kind of PLE estimation could be made more reliable. To achieve that, our second line of research, i.e., **PLE self-estimation in wireless networks**, is mainly devoted to seek out a better solution, which is not only more accurate but also less susceptible to external factors such as inaccurate location information or attacks from adversaries.

RELATED WORKS

THE existing methods for estimating the PLE can be categorized into anchor-based, anchor-free, self-estimation and collective self-estimation. We compare them in Table 1.1 and hereby elaborate on them as follows.

Anchor-Based In order to estimate the PLE, the intuitive way is observing the SS measurements between anchors. With known locations, the transmission distance can easily be calculated such that estimating the PLE becomes possible [74–76, 82]. However, this kind of method mainly suffers from the following critical drawbacks.

1. **Anchor Dependence:** This kind of PLE estimation heavily relies on the anchors' known locations. Once the location information is not reliable or even impossible to obtain, it would be very difficult to calculate the true transmission distance, let alone an accurate PLE. For instance, the anchor location information may be given by GPS, which can be very inaccurate in indoor scenarios; further, this information could also be tampered with by adversaries in military cases; finally, in some WSNs, employing anchors might be very expensive or rather difficult.
2. **Intensive node cooperation:** In order to estimate the PLE of the targeted area, this kind of PLE estimation often requires intensive node cooperation for sharing critical information such as anchor locations and collecting the samples from multiple transmission links. Also, the PLE estimation process often works in a centralized rather than distributed fashion. Obviously, this will consume a lot of network resources like throughput, bandwidth and battery energy, which are often very scarce in practice. Furthermore, certain node topologies are sometimes particularly required for performance enhancement, which is practically even more difficult.
3. **Not Pervasive:** The above mentioned drawbacks boil down to the most significant disadvantage that this kind of PLE estimation is not pervasive. To be more clear, it can only be used in a network that is well-infrastructure, well-equipped and cooperative. Once the network is temporarily formed, with very limited network

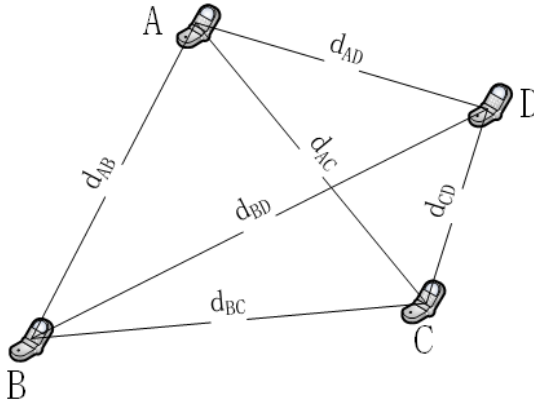


Figure 1.5: Demonstration in \mathbb{R}^2 : all possible nodal distances are constrained by the *Cayley-Menger* determinant being zero as mentioned in (1.11.)

resources or uncooperative, implementing this kind of PLE estimation would be very difficult.

Anchor-Free Some other scientists have also noticed those drawbacks and put considerable effort on solving them.

Research Challenge No. 10: Anchor-free

To cope with the anchor dependence issue, an intuitive solution is estimating the PLE in an anchor-free manner, i.e., without relying on any known location or transmission distance.

Question: How can the PLE be estimated in an anchor-free fashion?

Answers: We will present the anchor-free methods below. Although the methods following thereafter, the self-estimation and the collective self-estimation, also work in the same fashion, they are named due to some other notable features, which will be discussed later.

Following this idea, G. Mao et al introduced some anchor-free approaches, which estimate the PLE merely based on some geometric constraints [77, 78]. To be more specific, they are based on the fact that all the possible nodal distances between the considered nodes are subject to the *Cayley-Menger* determinant being zero [83].

For instance, assume there are 4 nodes in a plane, yielding $\binom{4}{2} = 6$ different nodal distances, as shown in Fig. 1.5. Then setting the *Cayley-Menger* determinant to zero in

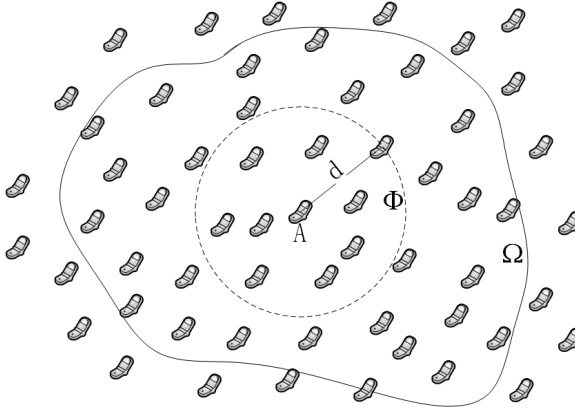


Figure 1.6: Without any external assistance, the considered node A should be able to self-estimate the PLE individually and solely. Due to the unknown node topology, the neighbours of the node A are ideally considered to be randomly deployed. In \mathbb{R}^2 , the dotted circle Φ indicates a bounded set with radius d and Ω is a finite space with n nodes randomly deployed inside.

this case can be expressed as

$$\begin{vmatrix} 0 & d_{AB}^2 & d_{AC}^2 & d_{AD}^2 & 1 \\ d_{AB}^2 & 0 & d_{BC}^2 & d_{BD}^2 & 1 \\ d_{AC}^2 & d_{BC}^2 & 0 & d_{CD}^2 & 1 \\ d_{AD}^2 & d_{BD}^2 & d_{CD}^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix} = |\mathbf{M}(A, B, C, D)| = 0, \quad (1.11)$$

where $\mathbf{M}(A, B, C, D)$ is the *Cayley-Menger* matrix. Combining (1.4) and (1.11) results into a complicated and highly non-linear function w.r.t the unknown PLE γ in [78, eq. 14], which considers the SS measurements between all pairs of given nodes. In the absence of the shadowing effect, the PLE can be obtained via a bisection method, while some smart pattern matching techniques are required to deal with a noisy case. To avoid tautology, we will not present more details here.

Undoubtedly, the anchor-free property is a very notable merit for PLE estimation, making this approach less constrained. Without known locations or transmission distances, a group of nodes can cooperate to estimate the PLE, which is rather convenient. However, this approach still requires intensive node cooperation for aggregating the samples. Therefore, it cannot be universally adopted, especially not in *ad hoc* and uncooperative scenarios.

Self-Estimation Next, while keeping the anchor-free property, there exist some other approaches that can estimate the PLE without intensive node cooperation.

Research Challenge No. 11: Self-Estimation

Nowadays, the tendency of wireless networks is towards high-efficiency and low-cost. To accommodate that, the PLE estimation should not be based on intensive node cooperation. The best solution may be self-estimating the unknown PLE such that the PLE estimation process can be set free and carried out in a distributed fashion. To be specific, without any external assistance, we want any kind of node to be able to estimate the PLE individually and solely, as demonstrated in Fig. 1.6.

Question: How can we self-estimate the PLE as described above? Is there any other benefit of doing so?

Answers: Self-estimating the PLE implies that the considered node should be blind to the surrounding node topology, not require any external assistance and even better conceal the estimation process itself if required. In this case, the considered node ideally assumes the surrounding neighbours to be randomly placed. Therefore, the problem of PLE self-estimation can basically be interpreted as "how do we use the information of the node random placement for estimating the PLE?" We will next first present how the random node placement is modelled, and then the existing methods for PLE self-estimation and finally ours will be discussed.

Besides the benefit of working in a distributed fashion, we should also notice that getting grip on the PLE has actually been an urgent demand, since the PLE yields a multifaceted influence on wireless communications and networking. To be more clear, it might enable smart designs of protocols and techniques in an adaptive fashion for coping with dynamic communication environments, e.g., a changing PLE, which directly gives rise to our second research line, i.e., **PLE self-estimation in wireless networks**.

In most cases, networks are constructed in an *ad hoc* manner. Due to the unknown node topology, the nodes in wireless networks are ideally considered to be randomly placed. According to stochastic geometry [84, 85], which treats a random node placement as a stochastic process, two spatial point processes are commonly considered: a binomial point process (BPP) and a Poisson point process (PPP) [86]. The BPP describes a finite random node placement. In other words, if n nodes are constrained in $\Omega \subset \mathbb{R}^2$ as depicted in Fig. 1.6, the probability of finding k nodes in a bounded Borel set $\Phi \subset \mathbb{R}^2$ follows a binomial distribution $\mathcal{B}(n, \frac{\mathcal{V}(\Phi \cap \Omega)}{\mathcal{V}(\Omega)})$ as

$$\Pr(k \text{ nodes in } \Phi) = \binom{n}{k} \left[\frac{\mathcal{V}(\Phi \cap \Omega)}{\mathcal{V}(\Omega)} \right]^k \left[\frac{\mathcal{V}(\Omega \setminus \Phi)}{\mathcal{V}(\Omega)} \right]^{n-k}, \quad (1.12)$$

where $\mathcal{V}(\cdot)$ indicates the standard Lebesgue measure. The PPP, which is parameterized by the node density λ , is a limit case of the BPP, where Ω is infinite. In other words, according to the Poisson limit theorem [87], if n approaches ∞ while $n/\mathcal{V}(\Omega)$ remains fixed at $\lambda > 0$, the Binomial distribution in (1.12) becomes the Poisson distribution $\mathcal{P}(\lambda)$,

which is given by

$$\Pr(k \text{ nodes in } \Phi) = e^{-\lambda \mathcal{V}(\Phi)} \frac{(\lambda \mathcal{V}(\Phi))^k}{k!}. \quad (1.13)$$

The BPP and the PPP are mostly used for studying the wireless network performance such as the effects of routing and interference, but they rarely appear in the field of signal processing. In our second research line, they serve as the basic concept to characterize a random node placement.

Based on this idea, [79] introduced three PLE self-estimation approaches. The first one observes the mean interference at the receiver to calculate the PLE, but the network node density λ should be known *a priori* or estimated beforehand [88, 89]. The other two estimate the PLE respectively from the virtual outage probabilities and the cardinalities of the transmitting set at two different receiver sensitivities. However, these methods obviously require frequently changing the receiver configuration.

Admittedly, they do save a lot of network resources by self-estimating the PLE. However, the aforementioned constraints still make those approaches rather difficult for a practical implementation. For instance, the network node density λ changes over time, especially when new participants join the network, and hence frequently updating the estimate of λ would be very troublesome. Furthermore, the receiver sensitivity is often standardized and cannot be changed in most cases.

Although the three aforementioned methods are called "self-estimation", they are still not up to what is expected. At the very least, their application is somewhat limited due to the constraints and hence they are not very pervasive. In the next section, we will present our expected criteria for "PLE self-estimation" and brief our proposed collective PLE self-estimation methods.

COLLECTIVE PATH-LOSS EXPONENT SELF-ESTIMATION

WE are the first to study the collective PLE self-estimation that is free of all the mentioned drawbacks, as shown in Table 1.1.

In summary, the collective PLE self-estimation must hold the following features.

- **Simple**: The method is very easy to carry out, i.e., the requirements for a practical implementation are not very difficult.
- **Pervasive**: The method is applicable to any kind of wireless network regardless of its design.
- **Local**: The method works locally such that each node in the network can independently estimate the PLE in a distributed fashion.
- **Sole**: The method requires neither any external assistance nor any other information such as exact locations or the node density.
- **Collective**: The only freedom left for us is utilizing the SS measurements. Therefore, this method is designed to estimate the PLE merely by collecting the locally received SS measurements from neighboring nodes.
- **Directional**: The method self-estimates the PLE from a given angular window.

- **Secure** : Owing to the above merits, the method is very secure and hence sabotaging it would be very difficult. Furthermore, we can even conceal the estimation process from the neighbours if required, which is particularly beneficial in military scenarios.

Compared with the previous PLE self-estimation, although also assuming a random node placement, the desired method can only use the locally collected SS measurements in (1.4) for self-estimating the PLE.

Research Challenge No. 12: Ranking the Measurements

Instead of coping with the unknown transmission distance, we circumvent this issue by introducing another observation that also contains the distance information. To be specific, we rank the locally collected SS measurements and use the rank indices for estimating the PLE. As a result, the *shadowing* effect could disturb the ranking procedure, i.e., the SS measurement rank might not exactly map the distance rank, thus causing some mismatch in the rank indices.

Question: How can we use the rank indices for PLE estimation and how can we cope with the mismatch of the rank indices?

Answers: We define a new distance-related intermediate parameter, which can be estimated using the rank index based on the distribution in (1.12) that models the random node placement. This results into a total least squares optimization problem for the collective PLE self-estimation. Besides the traditional low-rank approximation solution that requires the singular value decomposition (SVD) [90], we also propose a closed-form solution that greatly reduces the computational complexity. Furthermore, we introduce a weighted closed-form solution, which particularly copes with the mismatch of the rank indices caused by the *shadowing* effect and which show a good performance. Please see Chapter 4 for more details.

The proposed methods in Chapter 4 possess all the desired features. Accordingly, wireless engineers can reasonably assume that the PLE can be accurately self-estimated while designing some protocols or algorithms in any kind of wireless network.

Research Challenge No. 13: Applications

Owing to all the earlier mentioned features, our proposed methods for collective PLE self-estimation can easily be incorporated into any kind of wireless network. The benefits of doing so have already been answered before. Nevertheless, one might still ask the the following question.

Question: How exactly can we use the PLE information to improve the performance of wireless communications and networking?

Answers: Obviously, there are many potential uses of the PLE information in smart designs for wireless communication and networking. In Chapter 4, we elaborate on some practical applications and also shed light on the possible future paths.

Although two collective PLE self-estimation approaches will be proposed in Chapter 4, there still exist some remaining problems such as the Cramér-Rao lower bound (CRLB) and the true maximum likelihood (ML) solution [91].

Research Challenge No. 14: CRLB and ML Solution

The CRLB provides a fundamental limit of the estimation accuracy and the ML solution yields the performance that can approach that limit. However, obtaining them becomes rather difficult in Chapter 4.

Question: What are the obstacles to obtaining the CRLB and the ML solution?

Answers: If we intend to obtain the CRLB and the ML solution, we have to first know the observation distribution. Obviously, the only observations are the SS measurements, although in Chapter 4 the rank indices are also used to self-estimate the PLE. However, the distribution of the SS measurements in a random node placement is totally unknown, which will be explained in details next. In Chapter 5, we try to derive the CRLB and the ML solution for some simplified wireless channels and we will also discuss how we do that later on.

The difficulties in obtaining the CRLB and the ML solution are already explained in Chapter 4. Therefore, in order to solve this issue, we can do nothing but investigating the distribution of the SS measurements in a random node placement. Note that this might also lead to a PLE self-estimation method that does not require ranking the SS measurements and thus is free of any unnecessary factors such as the rank mismatch.

Research Challenge No. 15: Observation Distribution

Given a random node placement, the SS measurement is not only impacted by wireless channel effects, but also subject to the spatial dynamics (random node placement). In fact, the unknown nodal distance should be regarded as a stochastic random variable, and some distributions were derived for the ranked nodal distances [92, 93]. This implies that the geometric path-loss of the SS measurement in (1.4) also becomes stochastic. More importantly, no distribution has been reported before in order to characterize that. Accordingly, obtaining the observation distribution for the collective PLE self-estimation is rather difficult.

Question: How do we cope with this issue?

Answers: In Chapter 5, we try to deal with the above issue and start with a simple case that considers only a geometric path-loss for wireless channels. We are the first to discover that in this case the SS measurement in random networks is *Pareto* distributed. Accordingly, the CRLB and two ML solutions can be derived. This finding would also be very useful for studying wireless communications and networking.

Although there exist some methods that strip out the complicated channel effects for SS measurements and leave only the geometric path-loss [94] in the model, we would still like to further consider more realistic wireless channels.

Research Challenge No. 16: Shadowing Effect

In order to be more realistic, the *shadowing* effect should also be considered for the SS measurements.

Question: In this case, how can we obtain the observation distribution including the *shadowing* effect?

Answers: To answer this question, we first need to realize that the SS measurement in this case is actually *log-normal* distributed with the mean subject to a *Pareto* distribution. Therefore, obtaining the desired distribution requires blending these two distributions.

For instance, let us denote \bar{P} as the SS measurement that only considers the geometric path-loss. Then, the distribution that also considers the *shadowing* effect can be obtained [87] as

$$\mathbb{P}(P) = \int \mathbb{P}_{\lognormal}(P|\bar{P}) \mathbb{P}_{pareto}(\bar{P}) d\bar{P}, \quad (1.14)$$

where \bar{P} is the mean of the *log-normal* distribution $\mathbb{P}_{\lognormal}(P|\bar{P})$ and $\mathbb{P}_{pareto}(\bar{P})$ is the *Pareto* distribution. Since this distribution is mathematically very difficult to calculate, we leave it for our future research. Fortunately, although the ML solution in Chapter 5 is derived only considering a geometric path-loss, it is surprisingly very resilient to the shadowing effect if considered.

1.1.3. CHANNEL EFFECTS

OUR work is mainly focused on utilizing SS measurements for localization and PLE self-estimation. We hereby discuss in details how the radio propagation channel influences the SS measurement.

Let us first denote the receiver signal at time t as

$$y(t) = g_{tx}g_{rx}g_{pl}x(t) * h(t) + n(t), \quad (1.15)$$

where g_{tx} and g_{rx} are respectively the transmitter and receiver antenna voltage gains, g_{pl} indicates the voltage attenuation caused by the geometric path-loss, $x(t)$ is the transmitted signal, $h(t)$ models random constructive and destructive self-interference and shadowing in a multi-path channel, $n(t)$ is additive white Gaussian noise (AWGN) (also called the background noise (BGN)) and $*$ denotes the convolution operator. Also, note that the geometric path-loss in power can be expressed as

$$g_{pl}^2 \triangleq \frac{\text{power at } d}{\text{power at } d_0} = \left(\frac{d}{d_0}\right)^{-\gamma},$$

where γ is the path-loss exponent, d is the transmission distance and d_0 is the reference distance.

For convenience, we do not present the frequency impact such as the *Doppler* effect here and only observe that, if the signal $y(t)$ is successfully demodulated, the BGN $n(t)$ can be well segregated or suppressed, hence obtaining the denoised signal envelope $r(t) \triangleq |g_{tx}g_{rx}g_{pl}x(t) * h(t)|$.

As shown in Fig. 1.7, the radio propagation channel affects $r(t)$ in two ways: the *small-scale* fading and the *large-scale* fading [25].

- The *small-scale* fading is mainly caused by signal reflections, e.g., on buildings or moving cars. Besides the line-of-sight (LOS) signal, multiple signal replica also arrive at the receiver and are superimposed, which results in a rapid fluctuation of $r(t)$ on a very small scale. It changes very fast in time, thus also called the *fast* fading. To model that, $r(t)$ can be considered to be *Nakagami- m* distributed [95].
- The *large-scale* fading includes the geometric path-loss and the *shadowing* effect, which dominantly decides the general level of instantaneous signal power $E(r(t)^2)$, equivalently the SS in *watts*. The geometric path-loss indicates that the signal power decays exponentially over distance as d^γ . Due to some large obstacles like buildings and hills, which obscure the main signal path, the *shadowing* effect causes some *Gaussian* variations on the attenuated signal power in *dB*. The *large-scale* fading affects $r(t)$ on a very large scale and changes very slowly, thus also referred to as *slow* fading. To model that, $E(r(t)^2)$ is commonly considered to be *log-normal* distributed. In other words, expressing $E(r(t)^2)$ into *decibels* results into the SS in (1.4) and we will discuss the collection of the SS measurements.

COLLECTING SS MEASUREMENTS

IF we want to collect SS measurements and assume the *small-scale* is modelled by a *Nakagami- m* distribution, we sample the instantaneous received power $p(t) \triangleq r(t)^2$,

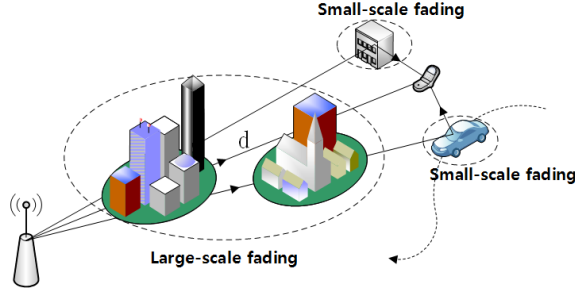


Figure 1.7: Demonstration of the *small-scale* fading and the *large-scale* fading in \mathbb{R}^2

which is accordingly *Gamma* distributed as

$$\mathbb{P}(p|\Omega) = \frac{1}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m p^{m-1} e^{-\frac{mp}{\Omega}}. \quad (1.16)$$

where $\Omega \triangleq E(p)$ is the SS to collect in *watts*, m is the fading parameter and a small value of m indicates a stronger fading. Therefore, collecting SS measurements is actually the procedure that estimates Ω and removes the impact of the *small-scale* fading.

In order to do that, an unbiased ML estimator for Ω can be readily given by

$$\hat{\Omega} = \frac{1}{K} \sum_{i=1}^K p_k, \quad (1.17)$$

where $p_k, \forall k = 1, \dots, K$ are K consecutive samples of $p(t)$. Obviously, denoting $\hat{\Omega} = \Omega + \Delta\Omega$, the SS collection error $\Delta\Omega$ is asymptotically *Normal* distributed as $\Delta\Omega \sim \mathcal{N}(0, \sigma_{\Delta\Omega}^2)$ and, since $Var(p) = \frac{\Omega^2}{m}$, $\sigma_{\Delta\Omega}^2$ can be obtained as

$$\sigma_{\Delta\Omega}^2 \triangleq Var(\hat{\Omega}) = \frac{\Omega^2}{Km}, \quad (1.18)$$

which can be reduced with more $p(t)$ samples, i.e., a large k .

Expressing $\hat{\Omega}$ in *decibels* and applying the first-order Taylor series expansion w.r.t. $\Delta\Omega$ result in

$$\hat{P} = 10 \log_{10}(\hat{\Omega}) = 10 \log_{10}(\Omega + \Delta\Omega) \approx P + \Delta P, \quad (1.19)$$

where $\Delta P \triangleq \frac{10}{\ln(10)} \frac{\Delta\Omega}{\Omega}$ is the SS collection error in *decibels* and $\Delta P \sim \mathcal{N}(0, \sigma_{\Delta P}^2)$ with

$$\sigma_{\Delta P}^2 = \frac{100}{\ln(10)^2 Km}.$$

We notice that, compared with $\sigma_{\Delta\Omega}^2$, $\sigma_{\Delta P}^2$ does not depend on Ω any more, i.e., the SS collection error in *decibels* is independent of the anchors. This means that, even if we do not collect enough $p(t)$ samples, the impact of ΔP is still similar to that of the shadowing effect in (1.4). Therefore, we can assume the SS is perfectly collected in this paper without loss of generality.

BACKGROUND NOISE

ALTHOUGH the background noise (BGN) is segregated in the demodulation part and hence does not appear in the SS measurement in (1.4), it still has some influence here.

Sample Size In fact, the signal can only be demodulated if the signal-to-noise ratio (SNR) exceeds a system-dependent threshold. This means, if the BGN becomes severe, collecting the SS measurements might become rather difficult due to a possible failure of the demodulation. Therefore, the BGN leads to transmission range limitations and hence decides how many SS measurements can be collected in real-life.

Power Floor The BGN also suggests a power floor for the SS measurement and hence we should re-formulate (1.4) as

$$P_i = \begin{cases} P_0 - 10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{d_0} \right) + \chi_i & P_r \geq \tau_{thres}, \\ \text{Nan} & \text{otherwise,} \end{cases} \quad (1.20)$$

where τ_{thres} is the lowest power level in dB , only above which the packet can be demodulated, and Nan indicates “not a number”.

For the presented work in this thesis, we simply assume that N distinct SS measurements are collected and all of them exceed the threshold. More complicated scenarios are left for future works, which will be discussed in Chapter 6.

1.2. OUTLINE AND CONTRIBUTIONS

AFTER the background and the motivations of our research, we would like to outline this thesis and elaborate on our contributions. Recalling that our research challenges have been highlighted in red blocks, we summarize them in Fig. 1.8 under different research topics of the chapters such that readers can better understand this thesis.

As mentioned, our work is mainly focused on two lines of research: **signal strength (SS) based localization** and **path-loss exponent (PLE) self-estimation in wireless networks**. Therefore, this thesis is comprised of four main parts. Part I contains this chapter with motivations and preliminaries. Our work of the first research line, i.e., **signal strength (SS) based localization**, is presented in Part II, which contains Chapter 2 and Chapter 3. Part III is focused on our second research line, i.e., **path-loss exponent (PLE) self-estimation in wireless networks**, and this part includes Chapter 4 and Chapter 5. Finally, we summarize our work and provide some suggestions for future challenges in Part IV, which contains Chapter 6.

Chapter 2 This chapter mainly exploits DRSS-based localization methods, of which the advantages have already been mentioned earlier. To be specific again, similar to TDOA-based localization, DRSS-based localization significantly alleviates the implicit dependence on the localization signal transmitter without dropping any advantage of RSS-based localization. As a result, the localization process becomes more robust against

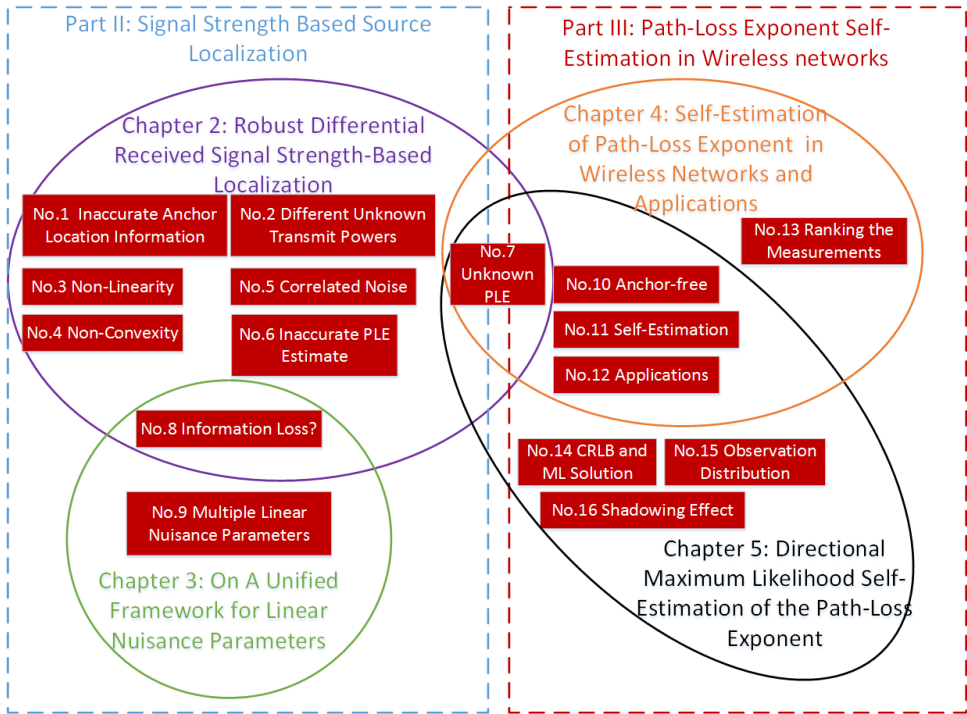


Figure 1.8: Diagram of our two research lines: The dashed rectangles separate our two research lines. Each solid ellipse stands for our research topic in the corresponding chapter, where the highlighted research challenges in the red blocks are also included. These research challenges have already been explained previously in this chapter. For clarity, a) Research Challenge No. 7 exists in both research lines, but is dealt with in different manners; b) Research Challenge No. 8 originates from Chapter 2, but is also investigated in a general manner in Chapter 3. Both of them link some of the chapters in this thesis.

defective, malicious or uncooperative signal transmitters. The network load can also be reduced with less or even no unnecessary overhead control messages, thus saving precious networking resources. Moreover, the localization process can easily be invisible to the neighbours if required.

Nonetheless, unlike TDOA-based localization that corrects the distance difference errors, the optimization problem for DRSS-based localization actually corrects the errors of the distance ratio. Accordingly, the research on DRSS-based localization is still in its fancy and hence requires more attention.

Our contributions in this chapter are listed as follows.

- We propose a new whitened data model for DRSS-based localization, since, in most cases, the coloured model noise is overlooked, especially the one brought by the differencing process that generates the DRSS measurements. In order to be more realistic, different unknown transmit powers are also taken into account. Moreover, this model is proven to suffer no information loss, compared with another properly whitened RSS-based model.

- Based on the introduced whitened model, three different kinds of methods are first introduced assuming that the PLE has already been estimated before the localization phase, i.e., a PLE estimation approach is assumed to be carried out beforehand. They are the advanced best linear unbiased estimator (A-BLUE), the Lagrangian estimator (LE) and the robust semi-definite programming estimator (RSDPE). In particular, the RSDPE is robust against the model uncertainties (caused by the imperfect PLE estimate or the inaccurate location information).
- The computational complexity of these three methods is studied. Numerical simulations have been conducted to evaluate their performance. We also discuss and compare the different localization methods in a more general context in Table 2.1, which is expected to be helpful for practical implementation purposes.
- Finally, we deal with the unknown PLE issue during the localization phase. In other words, a robust SDP-based block coordinate descent estimator (RSDP-BCDE) is proposed for DRSS-based localization to cope with a completely unknown PLE.

Chapter 2 has also been submitted as:

- **Y. Hu** and G. Leus, "Robust Differential Received Signal Strength-Based Localization", Signal Processing, IEEE Transactions on, 2016, Accepted

Chapter 3 In the previous chapter, the introduced whitened DRSS-based model for localization has been proven to bridge another properly whitened RSS-based model adopting an orthogonal operator, which implies that the differencing process that generates the DRSS measurement causes no information loss. However, this is in sharp contrast with what is commonly known in TDOA-based localization. This motivates us to study general differential signal processing with linear nuisance parameters.

We below list the most important results from this chapter.

- The differencing process that generates DRSS or TDOA measurements only cancels a single nuisance parameter. Nonetheless, in this chapter, we consider a general framework with multiple linear nuisance parameters such that all the conclusions apply to any kind of problem that can be written into this form. For the first time, we introduce a new differential approach to cope with multiple linear nuisance parameters while most existing works are merely dealing with a single nuisance parameter.
- We also investigate two other methods that can cope with nuisance parameters: the joint estimation and the orthogonal subspace projection (OSP) based estimation. Surprisingly, the BLUEs of the three considered methods are proven rigorously to be identical to each other if an appropriate preprocessing step is used.
- Compared with the joint estimation, which directly utilizes the original observations, none of the other two methods suffers any information loss.
- Although the differential method seems to rely on the selected reference, selecting the right reference is not important since there is no actual trace of the selected reference in the corresponding BLUE.

- As far as the differencing process is concerned, the differential observation set associated with a single reference already preserves the full data information.
- Two kinds of localization examples are demonstrated to verify our conclusions. We also explain the conflicts between our conclusions and what is commonly considered in localization literature.

Chapter 3 has also been submitted as:

- **Y. Hu** and G. Leus, "On A Unified Framework for Linear Nuisance Parameters," EURASIP Journal on Advances in Signal Processing, vol. 2017, no. 1, p. 4, 2017.

Chapter 4 This chapter is mainly aimed at proposing a directional self-estimator that possesses all the required features mentioned before. We list the important results from this chapter below.

- A new linear model for the PLE is first introduced, for which we particularly design an intermediate auxiliary parameter that is distance related. Without any exact information of nodal distance, we rank the locally collected SS measurements and regard the rank index as a new observation. Based on stochastic geometry, we obtain the ML estimate for the intermediate auxiliary parameter based on the rank indices in order to construct this model.
- Based on this model, a traditional total least squares (TLS) method for PLE self-estimation is first presented using the singular value decomposition (SVD), which however holds a high computational complexity.
- To guarantee a low computational complexity, we propose a closed-form TLS solution, which yields the same result but saves a great deal of computational time.
- In order to further improve the performance, we carefully analyze the model errors and propose a closed-form weighted TLS (WTLS) solution.
- Numerical simulations have been conducted to evaluate and study the proposed estimators.
- Some potential applications of this PLE self-estimation are carefully elaborated on, e.g., secure RSS-based localization, energy-efficient routing.

Chapter 4 has also been published as:

- **Y. Hu** and G. Leus, "Self-Estimation of Path-Loss Exponent in Wireless Networks and Applications," in IEEE Transactions on Vehicular Technology, vol. 64, no. 11, pp. 5091-5102, Nov. 2015.

Chapter 5 The previous chapter leaves some unfinished works such as the CRLB and the ML solution, which are very important and provide an in-depth understanding of the PLE self-estimation. However, it is hard to obtain them. Therefore, this chapter tries to solve this issue.

The important results from this chapter are listed below.

- As mentioned before, the distribution of the observation is the key to further obtain the CRLB and ML solution. However, the distribution of the SS measurement in random networks is hard to obtain. In order to solve this issue, we first only assume the geometric path-loss for the radio propagation channel and discover that the SS measurement is subject to a *Pareto* distribution in this case. Although the radio propagation channel is simplified, it is still very helpful for studying wireless communications and networking problems. Particularly, this distribution also applies to the case of node clusters.
- Based on the obtained distribution, we derive and study the CRLB for PLE self-estimation.
- Accordingly, we also introduce two ML solutions, which are evaluated later on by numerical simulations. Simulation results show that their performances are both very close to the CRLB.
- Although the two ML solutions are derived only assuming the geometric path-loss for the radio propagation channel, we surprisingly find that they are also very resilient to the shadowing effect if considered.

Chapter 5 has also been published as:

- **Y. Hu** and G. Leus, "Directional maximum likelihood self-estimation of the path-loss exponent," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 2016, pp. 3806-3810.

Chapter 6 This chapter summarizes this thesis and discusses some future research challenges as mentioned before.

Next to the content presented in this thesis, I also made additional contributions during my time as a Ph.D researcher at the Delft University of Technology.

For the first time, we extended SS-based localization to an underwater acoustic scenario, where the SS measurement is subject to a different underwater propagation channel. This work has been published as

- T. Xu, **Y. Hu**, B. Zhang and G. Leus, "RSS-based sensor localization in underwater acoustic sensor networks," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 2016, pp. 3906-3910.

I co-worked with my friend Yao Liu on the analysis of a phase-domain ADC and an amplitude-domain IQ ADC. This work has been published as

- Y. Liu, R. Lotfi, **Y. Hu** and W. A. Serdijn, "A Comparative Analysis of Phase-Domain ADC and Amplitude-Domain IQ ADC," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 3, pp. 671-679, March 2015.

In my early years as a Ph.D researcher at the Delft University of Technology, I worked with Assoc. Prof. Ertan Onur on the topic of "Density Estimation in Wireless Random Networks" and implemented a wireless networking platform with the OPnet simulator. I also corrected a theoretical error in a related previous publication [88], which has been published as

- **Y. Hu**, "Comments on "Cooperative Density Estimation in Random Wireless Ad Hoc Networks"," in *IEEE Communications Letters*, vol. 20, no. 4, pp. 832-835, April 2016.

In the meantime, I also supervised two MSc students, Fengju An and Tarikul Islam, respectively leading to the theses entitled

- An, Fengju. "Density Adaptive Sleep Scheduling in Wireless Sensor Networks". Diss. TU Delft, Delft University of Technology, 2013.

and

- Islam, T. "Statistical Modeling of Intelligent Transportation Systems Communication Channels". Diss. TU Delft, Delft University of Technology, 2013.

The latter also resulted in a conference publication:

- T. Islam, **Y. Hu**, E. Onur, B. Boltjes and J. F. C. M. de Jongh, "Realistic simulation of IEEE 802.11p channel in mobile Vehicle to Vehicle communication," *Microwave Techniques (COMITE)*, 2013 Conference on, Pardubice, 2013, pp. 156-161.

REFERENCES

- [1] S. Gezici, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, *Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks*, *IEEE Signal Processing Magazine* **22**, 70 (2005).
- [2] F. Gustafsson and F. Gunnarsson, *Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements*, *IEEE Signal Processing Magazine* **22**, 41 (2005).
- [3] R. Thrapp, C. Westbrook, and D. Subramanian, *Robust localization algorithms for an autonomous campus tour guide*, in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, Vol. 2 (2001) pp. 2065–2071 vol.2.
- [4] T. E. Budinger, *Biomonitoring with wireless communications*, *Annual Review of Biomedical Engineering* **5**, 383 (2003), PMID: 14527317, <http://dx.doi.org/10.1146/annurev.bioeng.5.040202.121653>.

- [5] G. Schofield, C. M. Bishop, G. MacLean, P. Brown, M. Baker, K. A. Katselidis, P. Dimopoulos, J. D. Pantis, and G. C. Hays, *Novel {GPS} tracking of sea turtles as a tool for conservation management*, *Journal of Experimental Marine Biology and Ecology* **347**, 58 (2007).
- [6] R. J. Fontana, E. Richley, and J. Barney, *Commercialization of an ultra wideband precision asset location system*, in *Ultra Wideband Systems and Technologies, 2003 IEEE Conference on* (2003) pp. 369–373.
- [7] J. Warrior, E. McHenry, and K. McGee, *They know where you are [location detection]*, *IEEE Spectrum* **40**, 20 (2003).
- [8] E. Kaplan and C. Hegarty, *Understanding GPS: Principles and Applications* (Artech House, 2005).
- [9] B. Parkinson and J. Spilker, *Global Positioning System: Theory and Applications*, Progress in astronautics and aeronautics No. v. 1 (American Institute of Aeronautics & Astronautics, 1996).
- [10] P. Misra and P. Enge, *Global positioning system: signals, measurements, and performance* (Ganga-Jamuna Press, 2006).
- [11] D. Liu, B. Sheng, F. Hou, W. Rao, and H. Liu, *From Wireless Positioning to Mobile Positioning: An Overview of Recent Advances*, *IEEE Systems Journal* **8**, 1249 (2014).
- [12] N. Bulusu, J. Heidemann, and D. Estrin, *GPS-less low-cost outdoor localization for very small devices*, *IEEE Personal Communications* **7**, 28 (2000), [arXiv:arXiv:1011.1669v3](https://arxiv.org/abs/1011.1669v3).
- [13] H. Wymeersch, J. Lien, and M. Z. Win, *Cooperative Localization in Wireless Networks*, *Proceedings of the IEEE* **97**, 427 (2009).
- [14] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, *Locating the nodes: cooperative localization in wireless sensor networks*, *Signal Processing Magazine, IEEE* **22**, 54 (2005).
- [15] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, *A survey on 3gpp heterogeneous networks*, *IEEE Wireless Communications* **18**, 10 (2011).
- [16] K. Pahlavan, X. Li, and J. P. Makela, *Indoor geolocation science and technology*, *IEEE Communications Magazine* **40**, 112 (2002).
- [17] B. P. Crow, I. Widjaja, L. G. Kim, and P. T. Sakai, *Ieee 802.11 wireless local area networks*, *IEEE Communications Magazine* **35**, 116 (1997).
- [18] C.-Y. Chong and S. P. Kumar, *Sensor networks: evolution, opportunities, and challenges*, *Proceedings of the IEEE* **91**, 1247 (2003).
- [19] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, *A survey on sensor networks*, *IEEE Communications Magazine* **40**, 102 (2002).

- [20] F. Thomas and L. Ros, *Revisiting trilateration for robot localization*, *IEEE Transactions on Robotics* **21**, 93 (2005).
- [21] B. F. Chen, J. P. Yue¹, and K. Shi, *Research on least squares adjustment of high precision network of triangulation*, *Survey Review* **45**, 111 (2013), <http://www.tandfonline.com/doi/pdf/10.1179/1752270611Y.0000000032>.
- [22] R. M. Vaghefi and R. M. Buehrer, *Cooperative joint synchronization and localization in wireless sensor networks*, *IEEE Transactions on Signal Processing* **63**, 3615 (2015).
- [23] W. A. Gardner and C. M. Spooner, *Comparison of autocorrelation and cross-correlation methods for signal-selective tdoa estimation*, *IEEE Transactions on Signal Processing* **40**, 2606 (1992).
- [24] H. Shi, H. Zhang, and X. Wang, *A tdoa technique with super-resolution based on the volume cross-correlation function*, *IEEE Transactions on Signal Processing* **64**, 5682 (2016).
- [25] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001).
- [26] Y. Kwon, T. Hwang, and X. Wang, *Energy-efficient transmit power control for multi-tier mimo hetnets*, *IEEE Journal on Selected Areas in Communications* **33**, 2070 (2015).
- [27] B. Kloiber, J. Harri, T. Strang, S. Sand, and C. R. Garcia, *Random transmit power control for dsrc and its application to cooperative safety*, *IEEE Transactions on Dependable and Secure Computing* **13**, 18 (2016).
- [28] C. Sun, Y. D. Alemseged, H. N. Tran, and H. Harada, *Transmit power control for cognitive radio over a rayleigh fading channel*, *IEEE Transactions on Vehicular Technology* **59**, 1847 (2010).
- [29] X. Li, *Collaborative Localization With Received-Signal Strength in Wireless Sensor Networks*, *Vehicular Technology, IEEE Transactions on* **56**, 3807 (2007).
- [30] X. Li, *RSS-Based Location Estimation with Unknown Pathloss Model*, *Wireless Communications, IEEE Transactions on* **5**, 3626 (2006).
- [31] N. Patwari, A. Hero, M. Perkins, N. Correal, and R. O'Dea, *Relative location estimation in wireless sensor networks*, *Signal Processing, IEEE Transactions on* **51**, 2137 (2003).
- [32] T. Qiao, Y. Zhang, and H. Liu, *Nonlinear Expectation Maximization Estimator for TDOA Localization*, *IEEE Wireless Communications Letters* **3**, 637 (2014).
- [33] W. Foy, *Position-Location Solutions by Taylor-Series Estimation*, *Aerospace and Electronic Systems, IEEE Transactions on* **AES-12**, 187 (1976).

- [34] C. Mensing and S. Plass, *Positioning Algorithms for Cellular Networks Using TDOA*, in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, Vol. 4 (2006) pp. IV–IV.
- [35] K. Yu, Y. Guo, and M. Hedley, *TOA-based distributed localisation with unknown internal delays and clock frequency offsets in wireless sensor networks*, *Signal Processing, IET* **3**, 106 (2009).
- [36] X. Lu and K. Ho, *Taylor-series technique for moving source localization in the presence of sensor location errors*, in *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on* (2006) pp. 4 pp.–.
- [37] S. Yousefi, X.-W. Chang, and B. Champagne, *A joint localization and synchronization technique using time of arrival at multiple antenna receivers*, in *Signals, Systems and Computers, 2013 Asilomar Conference on* (2013) pp. 2017–2021.
- [38] M. Sun and L. Yang, *On the joint time synchronization and source localization using TOA measurements*, *International Journal of Distributed Sensor Networks* **2013** (2013), 10.1155/2013/794805.
- [39] A. E. Assaf, S. Zaidi, S. Affes, and N. Kandil, *Low-Cost Localization for Multihop Heterogeneous Wireless Sensor Networks*, *IEEE Transactions on Wireless Communications* **15**, 472 (2016).
- [40] R. Vaghefi, M. Gholami, and E. Strom, *RSS-based sensor localization with unknown transmit power*, in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on* (2011) pp. 2480–2483.
- [41] D. Li and Y. H. Hu, *Least square solutions of energy based acoustic source localization problems*, in *Parallel Processing Workshops, 2004. ICPP 2004 Workshops. Proceedings. 2004 International Conference on* (2004) pp. 443–446.
- [42] D. B. Haddad, W. A. Martins, M. d. V. M. da Costa, L. W. P. Biscainho, L. O. Nunes, and B. Lee, *Robust Acoustic Self-Localization of Mobile Devices*, *IEEE Transactions on Mobile Computing* **15**, 982 (2016).
- [43] G. Wang and K. Yang, *A New Approach to Sensor Node Localization Using RSS Measurements in Wireless Sensor Networks*, *Wireless Communications, IEEE Transactions on* **10**, 1389 (2011).
- [44] G. Wang, H. Chen, Y. Li, and M. Jin, *On Received-Signal-Strength Based Localization with Unknown Transmit Power and Path Loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 536 (2012).
- [45] J. C. Chen, R. E. Hudson, and K. Yao, *Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field*, *IEEE Transactions on Signal Processing* **50**, 1843 (2002).

- [46] C.-H. Park, S. Lee, and J.-H. Chang, *Robust closed-form time-of-arrival source localization based on alpha-trimmed mean and Hodges Lehmann estimator under NLOS environments*, *Signal Processing* **111**, 113 (2015).
- [47] M. Sun and K. C. Ho, *Successive and Asymptotically Efficient Localization of Sensor Nodes in Closed-Form*, *IEEE Transactions on Signal Processing* **57**, 4522 (2009).
- [48] N. D. Gaubitch, W. B. Kleijn, and R. Heusdens, *Auto-localization in ad-hoc microphone arrays*, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (2013) pp. 106–110.
- [49] L. Wang, T. K. Hon, J. D. Reiss, and A. Cavallaro, *Self-Localization of Ad-Hoc Arrays Using Time Difference of Arrivals*, *IEEE Transactions on Signal Processing* **64**, 1018 (2016).
- [50] K. Liu, X. Liu, and X. Li, *Guoguo: Enabling Fine-Grained Smartphone Localization via Acoustic Anchors*, *IEEE Transactions on Mobile Computing* **15**, 1144 (2016).
- [51] S. Boyd and L. Vandenberghe, *Convex Optimization*, Berichte über verteilte messsysteme (Cambridge University Press, 2004).
- [52] J. J. More, *Generalizations Of The Trust Region Problem*, *OPTIMIZATION METHODS AND SOFTWARE* **2**, 189 (1993).
- [53] T. Pong and H. Wolkowicz, *The generalized trust region subproblem*, *Computational Optimization and Applications* **58**, 273 (2014).
- [54] J. More, J. J. and D. Sorensen, *Computing a Trust Region Step*, *SIAM Journal on Scientific and Statistical Computing* **4**, 553 (1983), <http://dx.doi.org/10.1137/0904038>.
- [55] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, Vol. 15 (SIAM, 1994).
- [56] L. Vandenberghe and S. Boyd, *Semidefinite Programming*, *SIAM Review* **38**, 49 (1994).
- [57] N. Salman, A. Kemp, and M. Ghogho, *Low Complexity Joint Estimation of Location and Path-Loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 364 (2012).
- [58] B.-C. Liu and K.-H. Lin, *Distance Difference Error Correction by Least Square for Stationary Signal-Strength-Difference-Based Hyperbolic Location in Cellular Communications*, *Vehicular Technology, IEEE Transactions on* **57**, 227 (2008).
- [59] B.-C. Liu, K.-H. Lin, and J.-C. Wu, *Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless communications*, *Vehicular Technology, IEEE Transactions on* **55**, 499 (2006).
- [60] A. Sayed, A. Tarighat, and N. Khajehnouri, *Network-based wireless location: challenges faced in developing techniques for accurate wireless location information*, *Signal Processing Magazine, IEEE* **22**, 24 (2005).

- [61] M. Gholami, R. Vaghefi, and E. Strom, *RSS-Based Sensor Localization in the Presence of Unknown Channel Parameters*, *Signal Processing, IEEE Transactions on* **61**, 3752 (2013).
- [62] M. Vossiek, L. Wiebking, P. Gulden, J. Wieghardt, C. Hoffmann, and P. Heide, *Wireless local positioning*, *IEEE Microwave Magazine* **4**, 77 (2003).
- [63] Y. Chan and K. Ho, *A simple and efficient estimator for hyperbolic location*, *Signal Processing, IEEE Transactions on* **42**, 1905 (1994).
- [64] Q. Xu, Y. Lei, J. Cao, and H. Wei, *An improved algorithm based on reference selection for time difference of arrival location*, in *Image and Signal Processing (CISP), 2014 7th International Congress on* (2014) pp. 953–957.
- [65] Y. Wang, F. Zheng, M. Wiemeler, W. Xiong, and T. Kaiser, *Reference selection for hybrid toa/rss linear least squares localization*, in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th* (2013) pp. 1–5.
- [66] I. Guvenc, S. Gezici, F. Watanabe, and H. Inamura, *Enhancements to linear least squares localization through reference selection and ml estimation*, in *2008 IEEE Wireless Communications and Networking Conference* (2008) pp. 284–289.
- [67] H. C. So, Y. T. Chan, and F. Chan, *Closed-form formulae for time-difference-of-arrival estimation*, *Signal Processing, IEEE Transactions on* **56**, 2614 (2008).
- [68] S. C. K. Herath and P. N. Pathirana, *Robust localization with minimum number of tdoa measurements*, *IEEE Signal Processing Letters* **20**, 949 (2013).
- [69] Y. Huang, J. Benesty, G. W. Elko, and R. M. Mersereati, *Real-time passive source localization: a practical linear-correction least-squares approach*, *IEEE Transactions on Speech and Audio Processing* **9**, 943 (2001).
- [70] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, *Least squares algorithms for time-of-arrival-based mobile location*, *IEEE Transactions on Signal Processing* **52**, 1121 (2004).
- [71] J. J. Caffery, *A new approach to the geometry of TOA location*, in *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, Vol. 4 (2000) pp. 1943–1949 vol.4.
- [72] C.-I. Chang, *Orthogonal subspace projection (osp) revisited: a comprehensive study and analysis*, *Geoscience and Remote Sensing, IEEE Transactions on* **43**, 502 (2005).
- [73] M. Song and C. I. Chang, *A theory of recursive orthogonal subspace projection for hyperspectral imaging*, *IEEE Transactions on Geoscience and Remote Sensing* **53**, 3055 (2015).
- [74] F. Santucci and N. Benvenuto, *A least squares path loss estimation approach to hand-over algorithms*, in *Proceedings of ICC/SUPERCOMM '96 - International Conference on Communications*, Vol. 2 (IEEE, 1996) pp. 802–806.

- [75] C. C. Pu, S. Y. Lim, and P. C. Ooi, *Measurement arrangement for the estimation of path loss exponent in wireless sensor network*, in *Computing and Convergence Technology (ICCCCT), 2012 7th International Conference on* (2012) pp. 807–812.
- [76] L. Razoumov and L. Greenstein, *Path loss estimation algorithms and results for RF sensor networks*, in *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004*, Vol. 7 (IEEE, 2004) pp. 4593–4596.
- [77] G. Mao, B. D. O. Anderson, and B. Fidan, *WSN06-4: Online Calibration of Path Loss Exponent in Wireless Sensor Networks*, in *IEEE Globecom 2006* (IEEE, 2006) pp. 1–6.
- [78] G. Mao, B. D. O. Anderson, and B. Fidan, *Path Loss Exponent Estimation for Wireless Sensor Network Localization*, *Comput. Netw.* **51**, 2467 (2007).
- [79] S. Srinivasa and M. Haenggi, *Path Loss Exponent Estimation in Large Wireless Networks*, in *Information Theory and Applications Workshop* (2009) pp. 124–129.
- [80] Y. Hu and G. Leus, *Self-estimation of path-loss exponent in wireless networks and applications*, *Vehicular Technology, IEEE Transactions on* **PP**, 1 (2014).
- [81] Y. Hu and G. Leus, *Directional maximum likelihood self-estimation of the path-loss exponent*, in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2016) pp. 3806–3810.
- [82] N. Benvenuto and F. Santucci, *A least squares path-loss estimation approach to handover algorithms*, *IEEE Transactions on Vehicular Technology* **48**, 437 (1999).
- [83] L. Blumenthal, *Theory and applications of distance geometry* (Chelsea Pub. Co., 1970).
- [84] M. Haenggi, *Stochastic Geometry for Wireless Networks* (Cambridge University Press, 2012).
- [85] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, *Stochastic geometry and random graphs for the analysis and design of wireless networks*, *IEEE Journal on Selected Areas in Communications* **27**, 1029 (2009).
- [86] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*, *An Introduction to the Theory of Point Processes* (Springer, 2007).
- [87] A. Papoulis and S. Pillai, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill series in electrical engineering: Communications and signal processing (Tata McGraw-Hill, 2002).
- [88] E. Onur, Y. Durmus, and I. Niemegeers, *Cooperative density estimation in random wireless ad hoc networks*, *IEEE Communications Letters* **16**, 331 (2012).
- [89] Y. Hu, *Comments on cooperative density estimation in random wireless ad hoc networks*, *IEEE Communications Letters* **20**, 832 (2016).

- [90] C. Eckart and G. Young, *The Approximation of One Matrix by Another of Lower Rank*, *Psychometrika* **1**, 211 (1936).
- [91] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993).
- [92] M. Haenggi, *On distances in uniformly random networks*, *Information Theory, IEEE Transactions on* **51**, 3584 (2005).
- [93] S. Srinivasa and M. Haenggi, *Distance distributions in finite uniformly random networks: Theory and applications*, *Vehicular Technology, IEEE Transactions on* **59**, 940 (2010).
- [94] J. Reig and L. Rubio, *Estimation of the composite fast fading and shadowing distribution using the log-moments in wireless communications*, *IEEE Transactions on Wireless Communications* **12**, 3672 (2013).
- [95] N. Nakagami, *The m-distribution, a general formula for intensity distribution of rapid fading*, in *Statistical Methods in Radio Wave Propagation*, edited by W. G. Hoffman (Oxford, England: Pergamon, 1960).

II

SIGNAL STRENGTH BASED SOURCE LOCALIZATION

2

ROBUST DIFFERENTIAL RECEIVED SIGNAL STRENGTH-BASED LOCALIZATION

Yongchang HU and Geert LEUS

*Fear not that the life shall come to an end,
but rather fear that it shall never have a beginning.*

J. H. Newman

Source localization is recently drawing a lot of attention and signal strength based methods have gradually thrived as a very popular topic due to their practical simplicity. However, the severe non-linearity and non-convexity make the related optimization problem mathematically difficult to solve, especially when the transmit power or the path-loss exponent (PLE) is unknown. Moreover, even if the PLE is known but not perfectly estimated or the anchor location information is not accurate, the constructed data model will become uncertain, making the problem again hard to solve.

This paper particularly focuses on differential received signal strength (DRSS)-based localization with model uncertainties in case of unknown transmit power and PLE. A new whitened model for DRSS-based localization with unknown transmit powers is first presented and investigated, on which all our proposed estimators are based. When assuming the PLE is known, we introduce two estimators based on an exact data model, an advanced best linear unbiased estimator (A-BLUE) and a Lagrangian estimator (LE), and then we present a robust semi-definite programming (SDP)-based estimator (RSDPE), which

can cope with model uncertainties (imperfect PLE and inaccurate anchor location information). The three proposed estimators have their own advantages from different perspectives: the A-BLUE has the lowest computational complexity; the LE holds the best accuracy for a small measurement noise; and the RSDPE yields the best performance under a large measurement noise and possesses a very good robustness against model uncertainties. Finally, we propose a robust SDP-based block coordinate descent estimator (RSDP-BCDE) to deal with a completely unknown PLE. The RSDP-BCDE jointly estimates the unknown PLE and the target location iteratively and its performance converges to that of the RSDPE using a perfectly known PLE.

2.1. INTRODUCTION

RECENTLY source localization is a rather prevalent technique aimed at locating a target based on measurements related to pre-deployed distributed sensors with *prior* known locations [1], i.e., anchor nodes. Briefly speaking, the commonly used measurements include, for example, time-of-arrival (TOA), time-difference-of-arrival (TDOA), angle-of-arrival (AOA) and signal strength. Among those, signal strength, such as received signal strength (RSS) [2] and differential RSS (DRSS)[3], gradually becomes the primary concern of numerous engineers owing to its implementation simplicity. Compared with other kinds of measurements, employing the signal strength as a measurement requires neither clock synchronization as for TOA-based or TDOA-based localization nor an antenna array which is indispensable for AOA-based localization. Therefore, this kind of source localization is more cost-effective in terms of both hardware and software. Besides, sensors usually have very scarce resources like limited computational abilities, constrained communication capabilities and depletable batteries, which further emphasizes its significance.

The signal strength measurement is determined by the signal power after successful demodulation[4, 5], which is still subject to a complicated radio propagation channel [6]. Without elaborating on the details, observe that the term “RSS”, in most literature, actually refers to the *large-scale* fading, the average of the instantaneous received signal power over several consecutive time slots, such that the *small-scale* fading, which is usually considered to be *Rayleigh* [7] or *Nakagami* [8] distributed, can be neglected. Please also refer to Appendix 2.6.1 for details on the RSS collection. Based on such an underlying assumption, the *log-normal* shadowing model can be used to characterize the RSS. Therefore, in \mathbb{R}^d , the RSS between the i -th anchor node, located at \mathbf{s}_i , and the target node, located at \mathbf{x} , can be presented in *dB* by

$$P_i = P_{0,i} - 10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{d_0} \right) + \chi_i, \quad i = 1, 2, \dots, N, \quad (2.1)$$

where $P_{0,i}$ is the received power related to the i -th anchor node at the reference distance d_0 , γ is the path-loss exponent (PLE), $\chi_i \sim \mathcal{N}(0, \sigma_\chi^2)$ represents the shadowing effect and N is the number of anchor nodes. Without loss of generality, we assume $d_0 = 1$ *m*. Note that $P_{0,i}$ can also be considered to be equivalent to the transmit power of the RSS related to the i -th anchor node. Then, the ultimate goal is to estimate the target location \mathbf{x} from the RSS samples P_i and known anchor locations \mathbf{s}_i .

To achieve this goal, source localization techniques using RSS measurements can be divided into three categories: maximum likelihood (ML), least squares (LS) based and semidefinite programming (SDP) based. The ML method is asymptotically optimal, but the related ML optimization problem is highly non-linear and non-convex [9]. Admittedly speaking, it can be iteratively solved [10–14]. However, this actually comes at the price of a high computational complexity. Moreover, the non-convexity also implies multiple local minima and hence an appropriate initialization is very important. The LS-based method relies on tackling the non-linearity by converting the non-linear ML optimization problem into a linear form such that some (weighted) LS-based solutions can be easily obtained [15–18]. However, these estimators are very susceptible to a large shadowing effect. The SDP-based method deals with the non-convexity by relaxing the

non-convex optimization problem to a convex one such that a global minimum can be effectively found [15, 19–23]. However, this method still requires a high complexity as well as a tight relaxation to guarantee an accurate estimate.

Besides those aforementioned issues, it is worth noting that the RSS measurements can be collected either by anchors in a distributed fashion or locally by the target node. To be specific, the former indicates that the localization signal is broadcast only by the target node and hence the transmit power $P_{0,i}, \forall i$ is obviously the same for all the RSS measurements. However, in the latter case, when several localization signals are emitted by the anchors, the related transmit power $P_{0,i}, \forall i$ should be considered different. This is because, even if anchors are equipped with stable and sustainable power supplies to guarantee a consistent transmit power, a deviation $\Delta P_{0,i}$ can still occur due to some unexpected power surges or system instabilities and hence we have $P_{0,i} \triangleq \bar{P}_0 + \Delta P_{0,i}$ with \bar{P}_0 the nominal transmit power. Besides, some transmit power control techniques are often carried out for energy saving purpose, which could also result in a $\Delta P_{0,i}$. Compared with the former case, which might require particular networking protocols to aggregate the collected RSS measurements to a computation center (CC) for localization, the latter is more convenient and widely assumed, since the target node can just listen and then self-estimate its location based on the locally collected RSS measurements without increasing any workload related to the wireless networking. However, to the best of our knowledge, current RSS-based localization techniques rarely consider the case of different $P_{0,i}$.

In either one of the aforementioned cases, if the network is not very cooperative or the signal transmitter intentionally withholds information (e.g., for military scenarios), $P_{0,i}$ will be unknown. Similarly, the PLE γ is often unknown as well, since it might be very difficult or expensive to acquire, especially in dynamic communication environments. Yet, many works simply assume that they are perfectly known [12, 19, 20, 23, 24]. To tackle the problem of an unknown PLE γ , a pre-calibration procedure of the PLE can be carried out among the anchor nodes before the actual localization phase [25–28]. However, this will consume extra resources and will make the implementation more cumbersome. Consequently, some joint estimators of \mathbf{x} and γ appear in [10, 17, 18, 22, 29–32]. To handle the issue of an unknown $P_{0,i}$, there are also some joint estimators for \mathbf{x} and $P_{0,i}$ [15, 16, 21, 22, 30, 31].

In this paper, instead of utilizing RSS measurements, we consider DRSS measurements for localization. The practical advantages for using the DRSS measurements are similar to those of TDOA-based localization. While preserving all the advantages of RSS-based localization, it can significantly alleviate the passive dependence of localization on the signal transmitter, which could be defective, malicious or uncooperative. Moreover, control overhead message between anchors and target node is minimized or even no longer required, which saves energy, bandwidth and throughput. This also conceals the localization process from the signal transmitter, which is very beneficial to surveillance or military applications. Therefore, DRSS-based localization is very promising. Considering different unknown transmit powers $P_{0,i}$, the DRSS measurements can be obtained from (2.1) as

$$P_{i,1} = -10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{\|\mathbf{x} - \mathbf{s}_1\|_2} \right) + \Delta P_{0,i,1} + \chi_{i,1}, \quad i \neq 1, \quad (2.2)$$

where $P_{i,1} = P_i - P_1$, $\Delta P_{0,i,1} \triangleq \Delta P_{0,i} - \Delta P_{0,1}$ and $\chi_{i,1} = \chi_i - \chi_1$. To construct a DRSS sample set, a reference node (RN) is chosen and the measurements are taken w.r.t. that RN. For convenience, the RN is appointed as the first anchor. Note that, in such a case, the size of the DRSS sample set becomes $N - 1$. In spite of the fact that (2.2) still remains non-linear and non-convex, the benefit of using a DRSS sample set is that the unknown nominal transmit power \bar{P}_0 vanishes. However, compared with (2.1), the inevitable price is that the shadowing effect and the transmit power deviation are exacerbated since $\chi_{i,1}$ and $\Delta P_{0,i,1}$ become correlated and (2.2) gets even more complicated to solve. This is also the reason why very few papers study this type of localization. To the best of our knowledge, some early results occurred in [33, 34]. In [15], some initial DRSS-based localization techniques were presented, yet having a worse accuracy than the corresponding RSS-based localization techniques, except for a simple least squares (LS) estimator which merely is slightly better. Recently, [3] presented a two-step weighted LS estimator, yet it requires perfect knowledge of the variance of the shadowing effect. Moreover, in practice, if the PLE and anchor location information is inaccurate (e.g., especially in military scenarios, some critical information might be unreliable), uncertainties have to be considered into the constructed data model for DRSS-based localization. However, very few results exist in this area, even for RSS-based localization. In a nutshell, the research on DRSS-based localization is still in its infancy and requires more attention.

To enrich the research on DRSS-based localization and to tackle the earlier mentioned problems, the first contribution of this paper is to introduce a new whitened model for DRSS-based localization with different unknown transmit powers. Based on this model, an advanced best linear unbiased estimator (A-BLUE), a Lagrangian estimator (LE) and a robust SDP-based estimator (RSDPE) are respectively proposed, assuming the PLE is known, in which the RSDPE is particularly designed to cope with model uncertainties. Their computational complexities are discussed and verified by experiments. We also conduct simulations to study their performances under different noise conditions, different PLEs, imperfect PLE knowledge and inaccurate anchor location information. Finally, after accumulating enough insights by studying the three proposed estimators, we take a step further and develop an RSDP-based block coordinate descent estimator (RSDP-BCDE) to cope with the case when the PLE is totally unknown. Some issues related to a real-life implementation are also considered and discussed.

After this brief introduction, Section 2.2 elaborates on our new whitened model for DRSS-based localization with different unknown transmit powers, which is used throughout this paper. Then, three different kinds of estimators based on a known PLE (i.e., the A-BLUE, the LE and the RSDPE) are proposed in Section 2.3. Their complexities and performances in different situations are also analyzed and studied by numerical simulations. Based on those studies, Section 2.4 presents a solution (i.e., the RSDP-BCDE) to the DRSS-based localization problem when the PLE is completely unknown. We also simulate and discuss this solution at the end of Section 2.4. Finally, Section 2.5 summarizes the results of this paper.

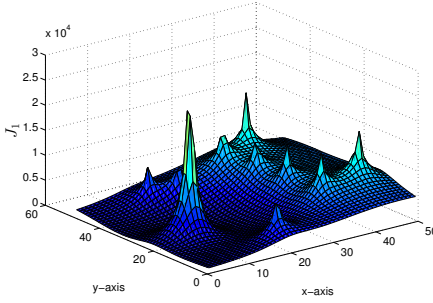
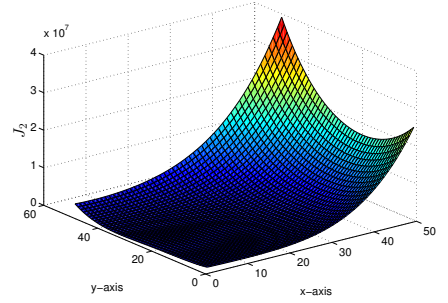
(a) The cost function J_1 .(b) The cost function J_2 .

Figure 2.1: Comparison between the least squares cost functions related to the models in (2.2) and (2.4): In \mathbb{R}^2 , the target node is at (28.7, 16.3); the anchor nodes are located at (22.5, 10.2), (44.9, 38.1), (44.1, 14.2), (33.6, 33.2), (6.1, 20.3), (13.7, 35.8), (14.1, 44.8), (41.3, 19.5), (24.9, 34.7) and (41.7, 30.5), of which the RN is selected as the one at (22.5, 10.2). Obviously, the target node cannot overlap with the anchor nodes and hence every location of the anchor node becomes a singular point in J_1 yielding multiple minima while J_2 has only a single optimal point.

2.2. WHITENED MODEL FOR DRSS-BASED LOCALIZATION

WE would firstly like to cope with the non-linearity issue of (2.2). To do this, we transform (2.2) into

$$\|\mathbf{x} - \mathbf{s}_i\|_2^2 P'_{i,1} = \Delta P'_{0,i,1} \chi'_{i,1} \|\mathbf{x} - \mathbf{s}_1\|_2^2, \quad i \neq 1, \quad (2.3)$$

where $P'_{i,1} \triangleq 10^{\frac{P_{i,1}}{5\gamma}}$, $\Delta P'_{0,i,1} \triangleq 10^{\frac{\Delta P_{0,i,1}}{5\gamma}}$, and $\chi'_{i,1} \triangleq 10^{\frac{\chi_{i,1}}{5\gamma}}$. Then, unfolding the Euclidean norm in (2.3), introducing $d_1^2 \triangleq \|\mathbf{x} - \mathbf{s}_1\|_2^2$ and stacking equations into matrices, our linear model for DRSS-based localization can be written as

$$\mathbf{p} = \Psi \boldsymbol{\theta} + \boldsymbol{\epsilon}, \quad (2.4)$$

where $\Psi \triangleq \begin{bmatrix} \vdots & \vdots \\ 2\mathbf{s}_1^T - 2P'_{i,1}\mathbf{s}_i^T & P'_{i,1} - 1 \\ \vdots & \vdots \end{bmatrix}$, $\boldsymbol{\theta} \triangleq [\mathbf{x}^T, \|\mathbf{x}\|_2^2]^T$, $\mathbf{p} \triangleq \begin{bmatrix} \vdots \\ \|\mathbf{s}_1\|_2^2 - \|\mathbf{s}_i\|_2^2 P'_{i,1} \\ \vdots \end{bmatrix}$, and $\boldsymbol{\epsilon} \triangleq \begin{bmatrix} \vdots \\ d_1^2(1 - \Delta P'_{0,i,1}\chi'_{i,1}) \\ \vdots \end{bmatrix}$. By respectively denoting $[\cdot]_i$ as the i -th element of a vector and

$[\cdot]_{1:i}$ as the subvector containing the first i elements of a vector, we observe that $[\boldsymbol{\theta}]_{1:d}$ corresponds to the target location \mathbf{x} and, more importantly, a new parameter is introduced at $[\boldsymbol{\theta}]_{d+1}$ which corresponds to $\|\mathbf{x}\|_2^2$. Our optimization problem w.r.t. $\boldsymbol{\theta}$ obviously becomes easier and any estimate of $\boldsymbol{\theta}$ leads to an estimate of \mathbf{x} , i.e., $\hat{\mathbf{x}} = [\hat{\boldsymbol{\theta}}]_{1:d}$.

To be more explicit, the model (2.4) is smoother than (2.2). To illustrate that, let us apply the least squares criterion to (2.2) and (2.4), leading to the respective cost functions

in \mathbf{x} :

$$J_1 = \sum_{i=2}^N \left[P_{i,1} + 10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{\|\mathbf{x} - \mathbf{s}_1\|_2} \right) \right]^2$$

and

$$J_2 = \|\Psi \begin{bmatrix} \mathbf{x} \\ \|\mathbf{x}\|_2^2 \end{bmatrix} - \mathbf{p}\|_2^2.$$

As depicted in Fig. 2.1, J_1 has multiple minima while J_2 becomes convex w.r.t. \mathbf{x} yielding only a single optimal point. Note that we explicitly take the dependence in $\boldsymbol{\theta}$ into account when formulating J_2 . In other words, we assume

$$[\boldsymbol{\theta}]_{1:d}^T [\boldsymbol{\theta}]_{1:d} = [\boldsymbol{\theta}]_{d+1},$$

which also implies that $\boldsymbol{\theta}$ is bound to a non-convex set since the dependence in $\boldsymbol{\theta}$ is considered.

To obtain our whitened model for DRSS-based localization, let us denote an element of $\boldsymbol{\epsilon}$ as $\epsilon_i = d_1^2(1 - \Delta P'_{0,i,1} \chi'_{i,1})$, $i \neq 1$. For a sufficiently small shadowing effect and transmit power deviation, ϵ_i can be approximated by its first-order Taylor series expansion¹

$$\epsilon_i = d_1^2 \left(1 - 10^{\frac{\Delta P_{0,i,1} + \chi_{i,1}}{5\gamma}} \right) = C(\Delta P_{0,i,1} + \chi_{i,1}), \quad (2.5)$$

which is apparently zero-mean yet mutually correlated, where $C \triangleq -\frac{\ln(10)d_1^2}{5\gamma}$ is a scaling factor. For the record, when the shadowing effect or the transmit power deviation grows very large, the approximation in (2.5) might become inaccurate. Notice that ϵ_i is subject to the PLE γ as well as the distance d_1 from the target to the RN. We cannot do much about the PLE, but it is clear that choosing a close RN will suppress the model error $\boldsymbol{\epsilon}$. Therefore, in this paper, the RN is chosen as the anchor node that has the highest RSS, since that anchor node is most likely the one that is closest to the target node. Note that in a mobile scenario, the RN should be updated in time, but we will not consider that in this paper.

For convenience, we respectively define the correlated DRSS measurement noise in (2.2) as $n_{i,1} \triangleq \Delta P_{0,i,1} + \chi_{i,1}$ and the independent measurement noise as $n_i \triangleq \Delta P_{0,i} + \chi_i$. Recalling that $\Delta P_{0,i,1} = \Delta P_{0,i} - \Delta P_{0,1}$ and $\chi_{i,1} = \chi_i - \chi_1$, we have $n_{i,1} = n_i - n_1$. Hence, from (2.5), the unwhitened model error $\boldsymbol{\epsilon}$ can be approximated as

$$\boldsymbol{\epsilon} = C\boldsymbol{\Gamma}\mathbf{n},$$

where \mathbf{n} stacks all independent DRSS measurement noise terms n_i and

$$\boldsymbol{\Gamma} \triangleq \begin{bmatrix} -\mathbf{I}_{(N-1) \times 1} & \mathbf{I}_{N-1} \end{bmatrix}_{(N-1) \times N}, \quad (2.6)$$

with \mathbf{I} the identity matrix, $\mathbf{0}$ the zero matrix and $\mathbf{1}$ the all-one matrix (sizes are mentioned in the subscript if needed). In this paper, we assume that $\Delta P_{0,i}$ is a zero-mean Gaussian

¹ $a^x = 1 + x \ln(a) + \dots + \frac{(x \ln(a))^n}{n!} + \dots$, $-\infty < x < \infty$.

variable with variance $\sigma_{P_0}^2$. Therefore, we can obtain $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})$ and the covariance matrix of $\boldsymbol{\epsilon}$ can be computed as

$$\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = C^2(\sigma_{P_0}^2 + \sigma_{\chi}^2)\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T = C^2\sigma_n^2\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T,$$

where $\sigma_n^2 \triangleq \sigma_{P_0}^2 + \sigma_{\chi}^2$ is the variance of the independent measurement noise n_i (or simply called measurement noise from now on), i.e., $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2\mathbf{I}_N)$.

Finally, from (2.4), we can obtain the whitened model as

$$\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1/2}\boldsymbol{p} = \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1/2}\boldsymbol{\Psi}\boldsymbol{\theta} + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1/2}\boldsymbol{\epsilon} \quad (2.7a)$$

$$\Rightarrow (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2}\boldsymbol{p} = (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2}\boldsymbol{\Psi}\boldsymbol{\theta} + (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2}\boldsymbol{\epsilon} \quad (2.7b)$$

$$\Rightarrow \boldsymbol{\rho} = \boldsymbol{\Phi}\boldsymbol{\theta} + \boldsymbol{v} \quad (2.7c)$$

where $\boldsymbol{\rho} \triangleq (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2}\boldsymbol{p}$, $\boldsymbol{\Phi} \triangleq (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2}\boldsymbol{\Psi}$ and $\boldsymbol{v} \triangleq (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2}\boldsymbol{\epsilon}$. Obviously, the model error \boldsymbol{v} in (2.7c) is whitened, since its covariance matrix is a scaled identity, i.e., $\boldsymbol{\Sigma}_{\boldsymbol{v}} = C^2\sigma_n^2\mathbf{I}_{N-1}$.

An important observation that we would like to make about our whitened DRSS-based data model is that no information is lost by taking differences of RSSs, since our model can be alternatively derived from a properly whitened RSS-based model after orthogonally projecting out the unknown average power \bar{P}_0 (see Appendix 2.6.2 for details). As a result, the choice of the RN has no effect on the performance of the localization accuracy.

2.3. ESTIMATORS FOR KNOWN PATH-LOSS MODEL

IN this section, we assume that the PLE γ is known and our derivations start from an exactly known data model. Considering our whitened model (2.7c) and ignoring the dependence in $\boldsymbol{\theta}$, it is possible to formulate the following unconstrained optimization problem

$$\min_{\boldsymbol{\theta}} \|\boldsymbol{\Phi}\boldsymbol{\theta} - \boldsymbol{\rho}\|_2^2. \quad (2.8)$$

This leads to the unconstrained best linear unbiased estimator (U-BLUE) for \mathbf{x} , which can be presented as

$$\begin{aligned} \hat{\mathbf{x}}_{u-blue} &= [\hat{\boldsymbol{\theta}}_{u-blue}]_{1:d} \\ &= [(\boldsymbol{\Phi}^T\boldsymbol{\Sigma}_{\boldsymbol{v}}^{-1}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\boldsymbol{\Sigma}_{\boldsymbol{v}}^{-1}\boldsymbol{\rho}]_{1:d} \\ &= [(\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\boldsymbol{\rho}]_{1:d}. \end{aligned} \quad (2.9)$$

Note that the unknowns C and σ_n^2 are eliminated in this solution. Although there are other similar least squares (LS) solutions [15, 17, 18], none of them is the BLUE since their data models are still coloured. Here, the U-BLUE will not perform very well as we will illustrate later on. Hence, in this section, we introduce two alternative methods based on an exactly known data model and then take some model uncertainties into account, which finally leads to a robust estimator for DRSS-based localization. We conduct simulations to study their performances under different noise conditions, different PLEs, imperfect PLE knowledge and inaccurate anchor location information. Their complexities are also studied and numerical results are presented. We end this section by discussing some practical issues.

2.3.1. ADVANCED BEST LINEAR UNBIASED ESTIMATOR

To boost the performance of the U-BLUE, we will take the dependence in $\boldsymbol{\theta}$ into account and hence our optimization problem has to be reformulated as

$$\min_{\boldsymbol{\theta}} \|\boldsymbol{\Phi}\boldsymbol{\theta} - \boldsymbol{\rho}\|_2^2 \quad (2.10a)$$

$$\text{subject to } [\boldsymbol{\theta}]_{1:d}^T [\boldsymbol{\theta}]_{1:d} = [\boldsymbol{\theta}]_{d+1}. \quad (2.10b)$$

The commonly known method to solve this problem indirectly is by constructing a new data model [3, 16, 33–36]. For instance, the new model can be given by

$$\mathbf{g} = \mathbf{Q}\mathbf{z} + \mathbf{m}, \quad (2.11)$$

where $\mathbf{g} \triangleq [[\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_1^2, \dots, [\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_d^2, [\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_{d+1}]^T$, $\mathbf{Q} \triangleq [\mathbf{I}_d, \mathbf{1}_{d \times 1}]^T$, $\mathbf{z} \triangleq [[\mathbf{x}]_1^2, \dots, [\mathbf{x}]_d^2]^T$ and

$$\mathbf{m} \triangleq \begin{bmatrix} [\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_1^2 - [\mathbf{x}]_1^2 \\ \vdots \\ [\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_d^2 - [\mathbf{x}]_d^2 \\ [\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_{d+1} - \|\mathbf{x}\|_2^2 \end{bmatrix} \approx \begin{bmatrix} 2[\mathbf{x}]_1([\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_1 - [\mathbf{x}]_1) \\ \vdots \\ 2[\mathbf{x}]_d([\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_d - [\mathbf{x}]_d) \\ [\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_{d+1} - \|\mathbf{x}\|_2^2 \end{bmatrix}. \quad (2.12)$$

Based on this model, the location estimate considering the constraint (2.10b) can be obtained as

$$\hat{\mathbf{x}} = [\text{sign}([\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_1) \sqrt{[\hat{\mathbf{z}}]_1}, \dots, \text{sign}([\hat{\boldsymbol{\theta}}_{u\text{-blue}}]_d) \sqrt{[\hat{\mathbf{z}}]_d}]^T,$$

where $\text{sign}(\cdot)$ indicates the signum function and $\hat{\mathbf{z}}$ is an estimate of \mathbf{z} . However, note that this method actually estimates the squared element of the target location \mathbf{x} and the squaring procedure on $\hat{\boldsymbol{\theta}}_{u\text{-blue}}$, which leads to the new observation vector \mathbf{g} , might exacerbates the estimation error in $\hat{\boldsymbol{\theta}}_{u\text{-blue}}$.

Here, we propose an advanced best linear unbiased estimator (A-BLUE) to solve (2.10) directly, which fine-tunes $\hat{\boldsymbol{\theta}}_{u\text{-blue}}$ without any squaring procedure. Recalling from (2.9) that $\hat{\boldsymbol{\theta}}_{u\text{-blue}} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \boldsymbol{\rho}$, the cost function in (2.10a) can be rewritten as

$$\begin{aligned} J &= (\boldsymbol{\Phi}\boldsymbol{\theta} - \boldsymbol{\rho})^T (\boldsymbol{\Phi}\boldsymbol{\theta} - \boldsymbol{\rho}) \\ &= (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{u\text{-blue}})^T \boldsymbol{\Phi}^T \boldsymbol{\Phi} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{u\text{-blue}}). \end{aligned} \quad (2.13)$$

In order to take the constraint (2.10b) into account, $\boldsymbol{\theta}$ has to be reformulated as a function of $[\boldsymbol{\theta}]_{1:d}$, i.e.,

$$\boldsymbol{\theta} = \begin{bmatrix} [\boldsymbol{\theta}]_{1:d} \\ [\boldsymbol{\theta}]_{1:d}^T [\boldsymbol{\theta}]_{1:d} \end{bmatrix} \quad (2.14)$$

. By now using the first-order Taylor series expansion of $\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{u\text{-blue}}$ for $[\boldsymbol{\theta}]_{1:d}$ in the vicinity of $\hat{\mathbf{x}}_{u\text{-blue}}$, we obtain

$$\begin{aligned} \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{u\text{-blue}} &= \boldsymbol{\theta}|_{[\boldsymbol{\theta}]_{1:d} = \hat{\mathbf{x}}_{u\text{-blue}}} - \hat{\boldsymbol{\theta}}_{u\text{-blue}} \\ &\quad + \left. \frac{\partial \boldsymbol{\theta}}{\partial [\boldsymbol{\theta}]_{1:d}^T} \right|_{[\boldsymbol{\theta}]_{1:d} = \hat{\mathbf{x}}_{u\text{-blue}}} ([\boldsymbol{\theta}]_{1:d} - \hat{\mathbf{x}}_{u\text{-blue}}) \\ &= \boldsymbol{\tau} + \mathbf{G}([\boldsymbol{\theta}]_{1:d} - \hat{\mathbf{x}}_{u\text{-blue}}), \end{aligned} \quad (2.15)$$

where

$$\boldsymbol{\tau} \triangleq \boldsymbol{\theta}|_{\boldsymbol{\theta}|_{1:d}=\hat{\mathbf{x}}_{u-blue}} - \hat{\boldsymbol{\theta}}_{u-blue} = \begin{bmatrix} \mathbf{0}_{d \times 1} \\ \|\hat{\mathbf{x}}_{u-blue}\|_2^2 - [\hat{\boldsymbol{\theta}}_{u-blue}]_{d+1} \end{bmatrix}$$

and $\mathbf{G} \triangleq \frac{\partial \boldsymbol{\theta}}{\partial [\boldsymbol{\theta}]_{1:d}} \Big|_{\boldsymbol{\theta}|_{1:d}=\hat{\mathbf{x}}_{u-blue}} = \begin{bmatrix} \mathbf{I}_d \\ 2\hat{\mathbf{x}}_{u-blue}^T \end{bmatrix}$.

Substituting (2.15) into (2.13), we obtain

$$J = (\boldsymbol{\tau} + \mathbf{G}([\boldsymbol{\theta}]_{1:d} - \hat{\mathbf{x}}_{u-blue}))^T \boldsymbol{\Phi}^T \boldsymbol{\Phi} (\boldsymbol{\tau} + \mathbf{G}([\boldsymbol{\theta}]_{1:d} - \hat{\mathbf{x}}_{u-blue})). \quad (2.16)$$

Taking the derivative of (2.16) w.r.t. $[\boldsymbol{\theta}]_{1:d}$, we have

$$\frac{\partial J}{\partial [\boldsymbol{\theta}]_{1:d}} = 2\mathbf{G}^T \boldsymbol{\Phi}^T \boldsymbol{\Phi} \mathbf{G}([\boldsymbol{\theta}]_{1:d} - \hat{\mathbf{x}}_{u-blue}) + 2\mathbf{G}^T \boldsymbol{\Phi}^T \boldsymbol{\Phi} \boldsymbol{\tau}. \quad (2.17)$$

Finally, by forcing (2.17) to $\mathbf{0}$, the A-BLUE can be expressed as

$$\hat{\mathbf{x}}_{a-blue} = \hat{\mathbf{x}}_{u-blue} - (\mathbf{G}^T \boldsymbol{\Phi}^T \boldsymbol{\Phi} \mathbf{G})^{-1} \mathbf{G}^T \boldsymbol{\Phi}^T \boldsymbol{\Phi} \boldsymbol{\tau}. \quad (2.18)$$

2.3.2. LAGRANGIAN ESTIMATOR

THE A-BLUE approximates (2.15) by linearizing it around $[\boldsymbol{\theta}]_{1:d} = \hat{\mathbf{x}}_{u-blue}$, which implies that its accuracy will certainly be degraded if there is a large estimation error in the U-BLUE. In this subsection, we would like to go one step further to find an estimator without any approximation.

In order to do so, we need to rewrite the constraint in (2.10b) and reformulate our optimization problem (2.10) as

$$\min_{\boldsymbol{\theta}} \|\boldsymbol{\Phi}\boldsymbol{\theta} - \boldsymbol{\rho}\|_2^2 \quad (2.19a)$$

$$\text{subject to } \boldsymbol{\theta}^T \mathbf{A}\boldsymbol{\theta} + 2\mathbf{b}^T \boldsymbol{\theta} = 0, \quad (2.19b)$$

where $\mathbf{A} \triangleq \begin{bmatrix} \mathbf{I}_d & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}$ and $\mathbf{b} \triangleq \begin{bmatrix} \mathbf{0}_{d \times 1} \\ -\frac{1}{2} \end{bmatrix}$. The Lagrangian of (2.19) is

$$L(\boldsymbol{\theta}; \lambda) = (\boldsymbol{\Phi}\boldsymbol{\theta} - \boldsymbol{\rho})^T (\boldsymbol{\Phi}\boldsymbol{\theta} - \boldsymbol{\rho}) + \lambda(\boldsymbol{\theta}^T \mathbf{A}\boldsymbol{\theta} + 2\mathbf{b}^T \boldsymbol{\theta}), \quad (2.20)$$

where λ is the Lagrangian multiplier. Taking the derivative of (2.20) w.r.t. $\boldsymbol{\theta}$, we have

$$\frac{\partial L(\boldsymbol{\theta}; \lambda)}{\partial \boldsymbol{\theta}} = 2\boldsymbol{\Phi}^T \boldsymbol{\Phi}\boldsymbol{\theta} - 2\boldsymbol{\Phi}^T \boldsymbol{\rho} + 2\lambda\mathbf{A}\boldsymbol{\theta} + 2\lambda\mathbf{b} \quad (2.21)$$

and forcing (2.21) to 0 leads to our Lagrangian estimator (LE) which is given by

$$\hat{\boldsymbol{\theta}}_{le}(\lambda) = (\boldsymbol{\Phi}^T \boldsymbol{\Phi} + \lambda\mathbf{A})^{-1} (\boldsymbol{\Phi}^T \boldsymbol{\rho} + \lambda\mathbf{b}). \quad (2.22)$$

Since λ is unknown, it is required to find an appropriate value for λ . A similar problem also appears in [37, 38], where all possible values of λ should be calculated to determine the desired one. Note that some of those values might lead to a maximum of the

Lagrangian in (2.20), since the second-order optimality conditions are not examined[9]. Besides, the above method is very cumbersome and, recalling the fact that $\boldsymbol{\theta}$ is bound to a non-convex set, a suboptimal value of λ might be selected, yielding a local solution. Without going into many details, we will not further discuss it. Here, the idea is to firstly pinpoint an interval for λ , in which only one single global solution is guaranteed, and then to search for that solution..

To find such an interval, note that the solution in (2.22) is a minimum of the Lagrangian in (2.20) if the Hessian of (2.20) is positive semidefinite, i.e.,

$$\begin{aligned} & \boldsymbol{\Phi}^T \boldsymbol{\Phi} + \lambda \mathbf{A} \geq 0 \\ \Rightarrow & (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{\frac{1}{2}} (\mathbf{I}_{N-1} + \lambda (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}} \mathbf{A} (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}}) (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{\frac{1}{2}} \geq 0 \end{aligned} \quad (2.23)$$

In order to guarantee (2.23), the eigenvalues of $\mathbf{I}_{N-1} + \lambda (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}} \mathbf{A} (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}}$ should be all non-negative. Obviously, all the eigenvalues of $(\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}} \mathbf{A} (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}}$ are non-negative. Then denoting the largest eigenvalue of $(\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}} \mathbf{A} (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-\frac{1}{2}}$ as λ_{max} , we need $1 + \lambda \lambda_{max} \geq 0$, which provides a useful interval for λ as

$$\mathcal{I} = (-1/\lambda_{max}, \infty)$$

. On such an interval, we can find the desired value of λ , say $\hat{\lambda}_{le}$, such that

$$\hat{\boldsymbol{\theta}}_{le}(\hat{\lambda}_{le})^T \mathbf{A} \hat{\boldsymbol{\theta}}_{le}(\hat{\lambda}_{le}) + 2\mathbf{b}^T \hat{\boldsymbol{\theta}}_{le}(\hat{\lambda}_{le}) = 0. \quad (2.24)$$

Then, the Lagrangian estimator (LE) for \mathbf{x} can be obtained as $\hat{\mathbf{x}}_{le} = [\hat{\boldsymbol{\theta}}_{le}(\hat{\lambda}_{le})]_{1:d}$.

Now the problems left are how to search for $\hat{\lambda}_{le}$ and whether or not the LE yields the global solution. Before going into the details, it is important to firstly realize that the problem (2.19) is a quadratically constrained quadratic program (QCQP) which can be cast as a generalized trust region subproblem (GTRS) [39], for which an optimal solution can be found within a bounded interval, i.e., the interval \mathcal{I} . In this paper, we actually consider a simpler case with an equality constraint (2.19b) rather than an inequality constraint, yet some results can still be used to support the following discussions.

To search for $\hat{\lambda}_{le}$, let us define a function $f(\lambda)$ as $f(\lambda) \triangleq \hat{\boldsymbol{\theta}}_{le}(\lambda)^T \mathbf{A} \hat{\boldsymbol{\theta}}_{le}(\lambda) + 2\mathbf{b}^T \hat{\boldsymbol{\theta}}_{le}(\lambda)$, which is already known to be strictly decreasing on the interval \mathcal{I} [40, Theorem 5.2], such that $\hat{\lambda}_{le}$, which satisfies the constraint (2.24), can be effectively found by the bisection method. Next, the LE is guaranteed as a global solution [40, Theorem 3.2], since it follows the *Karush-Kuhn-Tucker* (KKT) conditions. This also indicates that there only exists one solution, i.e., the global solution, in the interval \mathcal{I} , which is the reason why it is called the trust region. Besides, note that the case $\hat{\lambda}_{le} = -1/\lambda_{max}$ is called the *hard case* [41] (since it is relatively difficult to solve), which is very rare and has never been seen in our numerous simulations. The *hard case* is also found to be very rare in other papers, e.g., in [31, 42].

2.3.3. ROBUST SEMIDEFINITE PROGRAMMING BASED ESTIMATOR

THE previously proposed estimators are both based on an exactly known data model. However, when the data model is not perfectly known due to an imperfect PLE estimate or inaccurate anchor location information, a huge bias will obviously occur in these

estimates. Therefore, in this subsection, we present a robust semidefinite programming based estimator (RSDPE) that can cope with such model uncertainties.

First, after using the *Schur complement* [43] and forming some linear matrix inequalities (LMIs), we equivalently rewrite the constraint in (2.10b) as

$$\begin{bmatrix} \mathbf{I}_d & [\boldsymbol{\theta}]_{1:d} \\ [\boldsymbol{\theta}]_{1:d}^T & [\boldsymbol{\theta}]_{d+1} \end{bmatrix} \succeq \mathbf{0}, \quad (2.25a)$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{I}_d & [\boldsymbol{\theta}]_{1:d} \\ [\boldsymbol{\theta}]_{1:d}^T & [\boldsymbol{\theta}]_{d+1} \end{bmatrix} \right) = d. \quad (2.25b)$$

The semidefinite relaxation (SDR) approach then relaxes the set of $\boldsymbol{\theta}$ by dropping the rank constraint in (2.25b). This procedure is also used in [15, 19–23], but they all assume an exactly known data model.

We want to go one step further and consider an uncertain Φ as $\Phi^\circ \triangleq \Phi + \Delta_\Phi$, where the perturbation matrix Δ_Φ collects the uncertainties caused by an imperfect PLE estimate or inaccurate anchor location information. Although the data model is not exactly known, a known upper bound ζ for $\|\Delta_\Phi\|_2$ could be very helpful, i.e., $\|\Delta_\Phi\|_2 \leq \zeta$, where $\|\cdot\|_2$ denotes the spectral norm, i.e., the largest singular value of the corresponding matrix.

The idea of the RSDPE is to cope with the worst-case model uncertainties using the SDP procedure. Therefore, we reformulate (2.10) as a minmax SDP optimization problem

$$\min_{\boldsymbol{\theta}, t} \max_{\|\Delta_\Phi\|_2 \leq \zeta} t \quad (2.26a)$$

$$\text{subject to} \begin{bmatrix} \mathbf{I}_{N-1} & (\Phi^\circ - \Delta_\Phi)\boldsymbol{\theta} - \boldsymbol{\rho} \\ ((\Phi^\circ - \Delta_\Phi)\boldsymbol{\theta} - \boldsymbol{\rho})^T & t \end{bmatrix} \succeq \mathbf{0}, \quad (2.26b)$$

$$\begin{bmatrix} \mathbf{I}_d & [\boldsymbol{\theta}]_{1:d} \\ [\boldsymbol{\theta}]_{1:d}^T & [\boldsymbol{\theta}]_{d+1} \end{bmatrix} \succeq \mathbf{0}. \quad (2.26c)$$

where t is an auxiliary slack variable.

Note that Δ_Φ only affects the constraint (2.26b) and hence we can isolate Δ_Φ in (2.26b) as

$$\begin{aligned} \mathbf{B}(\boldsymbol{\theta}, t) &\geq \begin{bmatrix} \mathbf{0} & \Delta_\Phi \boldsymbol{\theta} \\ \mathbf{0} & 0 \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ (\Delta_\Phi \boldsymbol{\theta})^T & 0 \end{bmatrix} \\ \Rightarrow \mathbf{B}(\boldsymbol{\theta}, t) &\geq \mathbf{T}^T \Delta_\Phi \mathbf{L}(\boldsymbol{\theta}) + \mathbf{L}(\boldsymbol{\theta})^T \Delta_\Phi^T \mathbf{T}, \end{aligned} \quad (2.27)$$

where

$$\mathbf{B}(\boldsymbol{\theta}, t) \triangleq \begin{bmatrix} \mathbf{I}_{N-1} & (\Phi^\circ \boldsymbol{\theta} - \boldsymbol{\rho}) \\ (\Phi^\circ \boldsymbol{\theta} - \boldsymbol{\rho})^T & t \end{bmatrix},$$

$\mathbf{T} \triangleq [\mathbf{I}_{N-1} \quad \mathbf{0}]$ and $\mathbf{L}(\boldsymbol{\theta}) \triangleq [\mathbf{0} \quad \boldsymbol{\theta}]$. Obviously, for the maximization in (2.26), the constraint (2.27) has to be reformulated considering the worst-case Δ_Φ .

To do so, we can easily state that

$$\mathbf{B}(\boldsymbol{\theta}, t) \geq \mathbf{T}^T \Delta_\Phi \mathbf{L}(\boldsymbol{\theta}) + \mathbf{L}(\boldsymbol{\theta})^T \Delta_\Phi^T \mathbf{T}, \quad \forall \Delta_\Phi : \|\Delta_\Phi\|_2 \leq \zeta \quad (2.28)$$

if and only if

$$\begin{aligned}\check{\mathbf{x}}^T \mathbf{B}(\boldsymbol{\theta}, t) \check{\mathbf{x}} &\geq \max_{\|\Delta_{\Phi}\|_2 \leq \zeta} \{\check{\mathbf{x}}^T \mathbf{T}^T \Delta_{\Phi} \mathbf{L}(\boldsymbol{\theta}) \check{\mathbf{x}} + \check{\mathbf{x}}^T \mathbf{L}(\boldsymbol{\theta})^T \Delta_{\Phi}^T \mathbf{T} \check{\mathbf{x}}\} \\ &= \max_{\|\Delta_{\Phi}\|_2 \leq \zeta} \{2\|\Delta_{\Phi} \mathbf{L}(\boldsymbol{\theta}) \check{\mathbf{x}}\|_2 \|\mathbf{T} \check{\mathbf{x}}\|_2\} \\ &= 2\zeta \|\mathbf{L}(\boldsymbol{\theta}) \check{\mathbf{x}}\|_2 \|\mathbf{T} \check{\mathbf{x}}\|_2, \forall \check{\mathbf{x}} \in \mathbb{R}^N.\end{aligned}\quad (2.29)$$

After introducing the bound ζ into (2.29), a new problem arises since the vector $\mathbf{T} \check{\mathbf{x}} \in \mathbb{R}^{N-1}$ does not have the same size as the vector $\mathbf{L}(\boldsymbol{\theta}) \check{\mathbf{x}} \in \mathbb{R}^{d+1}$. To bypass this issue, we introduce a new auxiliary vector $\check{\mathbf{y}} \in \mathbb{R}^{d+1}$, which is bounded using $\check{\mathbf{x}}$, such that we can use the Cauchy-Schwarz inequality on (2.29) to unfold the norm. To be specific, only after the worst-case constraint (2.29) is reformulated as

$$\check{\mathbf{x}}^T \mathbf{B}(\boldsymbol{\theta}, t) \check{\mathbf{x}} \geq 2\zeta \|\check{\mathbf{y}}\|_2 \|\mathbf{L}(\boldsymbol{\theta}) \check{\mathbf{x}}\|_2, \forall \check{\mathbf{x}}, \check{\mathbf{y}}: \|\mathbf{T} \check{\mathbf{x}}\|_2 \geq \|\check{\mathbf{y}}\|_2, \quad (2.30)$$

we can obtain a new constraint without the norm from (2.30) as

$$\check{\mathbf{x}}^T \mathbf{B}(\boldsymbol{\theta}, t) \check{\mathbf{x}} \geq \zeta (\check{\mathbf{y}}^T \mathbf{L}(\boldsymbol{\theta}) \check{\mathbf{x}} + \check{\mathbf{x}}^T \mathbf{L}(\boldsymbol{\theta})^T \check{\mathbf{y}}), \forall \check{\mathbf{x}}, \check{\mathbf{y}}: \|\mathbf{T} \check{\mathbf{x}}\|_2 \geq \|\check{\mathbf{y}}\|_2. \quad (2.31)$$

Although both (2.30) and (2.31) consider the worst-case Δ_{Φ} , we have to use the latter one to facilitate the derivations, which is actually a weaker condition due to the Cauchy-Schwarz inequality. Then, for convenience, we respectively rewrite $\|\mathbf{T} \check{\mathbf{x}}\|_2 \geq \|\check{\mathbf{y}}\|_2$ as

$$\begin{bmatrix} \check{\mathbf{x}} \\ \check{\mathbf{y}} \end{bmatrix}^T \begin{bmatrix} \mathbf{T}^T \mathbf{T} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_{d+1} \end{bmatrix} \begin{bmatrix} \check{\mathbf{x}} \\ \check{\mathbf{y}} \end{bmatrix} \geq 0 \quad (2.32)$$

and (2.31) as

$$\begin{bmatrix} \check{\mathbf{x}} \\ \check{\mathbf{y}} \end{bmatrix}^T \begin{bmatrix} \mathbf{B}(\boldsymbol{\theta}, t) & -\zeta \mathbf{L}(\boldsymbol{\theta})^T \\ -\zeta \mathbf{L}(\boldsymbol{\theta}) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \check{\mathbf{x}} \\ \check{\mathbf{y}} \end{bmatrix} \geq 0, \quad (2.33)$$

where note that (2.33) is a necessary condition to (2.32).

Finally, according to the S-procedure [44, p. 23], the implication that (2.32) leads to (2.33) holds true if and only if there exists an α such that

$$\begin{bmatrix} \mathbf{B}(\boldsymbol{\theta}, t) & -\zeta \mathbf{L}(\boldsymbol{\theta})^T \\ -\zeta \mathbf{L}(\boldsymbol{\theta}) & \mathbf{0} \end{bmatrix} - \alpha \begin{bmatrix} \mathbf{T}^T \mathbf{T} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_{d+1} \end{bmatrix} \geq \mathbf{0} \quad (2.34a)$$

$$\Leftrightarrow \begin{bmatrix} (1-\alpha)\mathbf{I}_{N-1} & \Phi^\circ \boldsymbol{\theta} - \boldsymbol{\rho} & \mathbf{0} \\ (\Phi^\circ \boldsymbol{\theta} - \boldsymbol{\rho})^T & t & -\zeta \boldsymbol{\theta}^T \\ \mathbf{0} & -\zeta \boldsymbol{\theta} & \alpha \mathbf{I}_{d+1} \end{bmatrix} \geq \mathbf{0}. \quad (2.34b)$$

Replacing (2.26b) by the new constraint (2.34b) leads to the following SDP optimization problem

$$\min_{\boldsymbol{\theta}, t, \alpha} t \quad (2.35a)$$

$$\text{subject to } \begin{bmatrix} (1-\alpha)\mathbf{I}_{N-1} & \Phi^\circ \boldsymbol{\theta} - \boldsymbol{\rho} & \mathbf{0} \\ (\Phi^\circ \boldsymbol{\theta} - \boldsymbol{\rho})^T & t & -\zeta \boldsymbol{\theta}^T \\ \mathbf{0} & -\zeta \boldsymbol{\theta} & \alpha \mathbf{I}_{d+1} \end{bmatrix} \geq \mathbf{0}, \quad (2.35b)$$

$$\begin{bmatrix} \mathbf{I}_d & [\boldsymbol{\theta}]_{1:d} \\ [\boldsymbol{\theta}]_{1:d}^T & [\boldsymbol{\theta}]_{d+1} \end{bmatrix} \geq \mathbf{0}, \quad (2.35c)$$

which can be solved by CVX[45, 46]. The solution is our new RSDPE.

To end this subsection, we discuss how to determine the value of ζ . One possibility is that ζ can be computed from the total least squares (TLS) method[47]. More specifically, we can compute the singular value decomposition (SVD) of the augmented matrix $[\Phi^\circ \ \rho] = \mathbf{U}\Sigma\mathbf{V}^T$ and the corrected $[\Phi^\circ \ \rho]$ is given by $[\hat{\Phi} \ \hat{\rho}] = \mathbf{U}\hat{\Sigma}\mathbf{V}^T$, where $\hat{\Sigma}$ is obtained by forcing the $(d+2)$ -th diagonal of Σ to 0, which is the typical low-rank approximation[48]. In fact, $\hat{\Phi}$ can be viewed as an estimate of the exact Φ and also observe that $\|\Delta_\Phi\|_2 = \|\Phi^\circ - \Phi\|_2 \leq \|\Phi^\circ - \hat{\Phi}\|_F$, where $\|\cdot\|_F$ indicates the Frobenius norm or the Hilbert-Schmidt norm. Therefore, in this paper, ζ is computed as $\zeta = \|\Phi^\circ - \hat{\Phi}\|_F$, which will also be used in our simulations later.

2.3.4. COMPLEXITY ANALYSIS

WE now calculate the computational complexity of the different methods without considering the whitening procedure [49]. It is easy to derive that the complexity of the U-BLUE is $O[d^2N]$. As for the A-BLUE, its complexity is $O[d^2N^2]$ considering that the extra cost is mainly comes from the second step in (2.18).

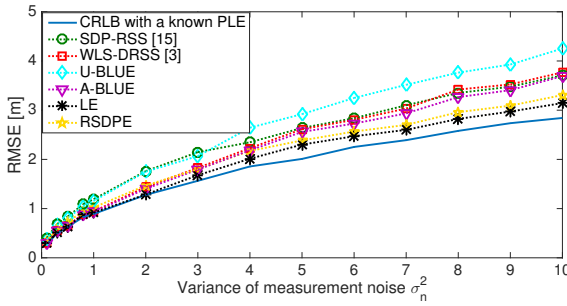
For the LE, the complexity is mostly due to the bisection method. Suppose that the bisection method takes K steps to find an appropriate λ , which has already been observed to be around 20. In each iteration, first the $\hat{\theta}_{le}(\lambda)$ in (2.22) is computed and then $f(\lambda)$ is calculated to check if the outcome is smaller than the tolerance. As a result, the cost of each iteration is $O[d^2N^2]$ and hence the complexity of the LE is $O[Kd^2N^2]$.

Finally, let us focus on the RSDPE. We consider the *worst-case* complexity for solving (2.35), which can be derived from employing the interior-point algorithm [9]. This implies that the complexity for each iteration is $O[d^2N^2]$ and the iteration number is bounded by $O[\sqrt{N}\ln(1/\xi)]$ [43], where ξ is the iteration tolerance. Therefore, the complexity of the RSDPE in this paper is $O[d^2N^{2.5}\ln(1/\xi)]$.

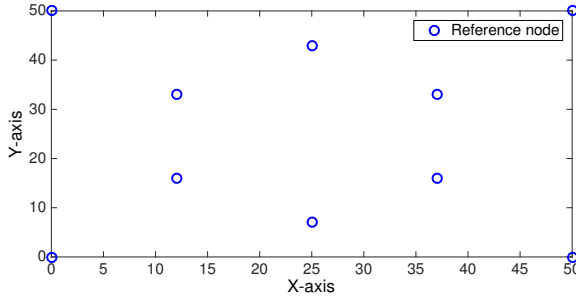
Obviously, the RSDPE has the highest complexity among all the proposed estimators. To verify the complexities, we conduct an experiment in a 2-D space with 10 anchor nodes and use the average computational time as a complexity measure. The experiment is implemented in Matlab R2013b on a Lenovo IdeaPad Y570 (Processor 2.0 GHz Intel Core i7, Memory 8GB). We observe that the U-BLUE and the A-BLUE respectively have the least and the second least average computational time of 0.026 *ms* and 0.049 *ms* while the RSDPE holds the highest one with 314.0 *ms*. Compared with the others, the complexity of the LE is reasonable with a computational time of 4.8 *ms*.

2.3.5. NUMERICAL RESULTS

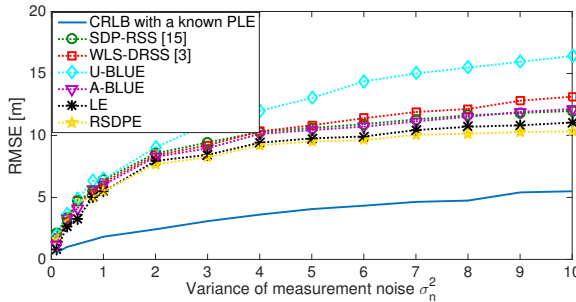
WE have conducted a Monte Carlo (MC) simulation using 1000 trials on a $50\ m \times 50\ m$ field, where one target node is randomly deployed for each trial. Our proposed estimators are compared against two existing methods: the RSS-based joint estimator (SDP-RSS) which is the best estimator from [15] and applies the SDP procedure on a ℓ_1 -norm approximation to jointly estimate the transmit power and the target location; and the recent DRSS-based two-step weighted least squares estimator (WLS-DRSS) of [3] which requires perfect knowledge of the variance of the measurement noise σ_n^2 . Recall that, in this paper, the measurement noise includes the shadowing effect and transmit power derivations. For computing the *Cramér-Rao lower bound* (CRLB), see Ap-



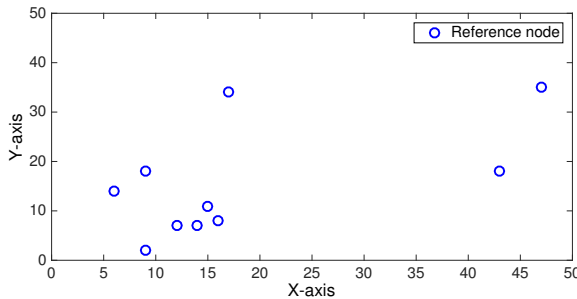
(a) Performance of different estimators with a good anchor node placement.



(b) A good anchor node placement.



(c) Performance of different estimators with a bad anchor node placement.



(d) A bad anchor node placement.

Figure 2.2: Impact of the anchor node placement: In \mathbb{R}^2 , 10 anchor nodes are considered with different placements and the target node is randomly deployed within a $50\text{ m} \times 50\text{ m}$ field where the path-loss exponent is considered $\gamma = 4$. The anchor location information is accurate.

pendix 2.6.3. The root mean square error (RMSE) is used to evaluate the performance of all estimators.

IMPACT OF THE ANCHOR NODE PLACEMENT

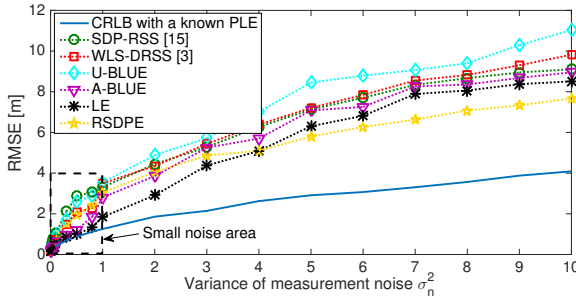
WE first discuss the impact of the anchor node placement, where the simulation is conducted with a perfectly known data model. Obviously, a good anchor node placement is very significant for any kind of localization. To be specific, if the anchor nodes get clustered, the measurements and the locations of those anchor nodes are both very close to each other, which easily leads to an ill-posed optimization problem for localization. For example, the cluster of anchor nodes in Fig. 2.2(d) causes the matrix Φ to be ill-conditioned, thus making our optimization problem very susceptible to the measurement noise. To verify that, two simulations are conducted one with a good and one with a bad anchor node placement. These two anchor node placements and the numerical results are shown in Fig. 2.2. Clearly, when a good anchor node placement is considered, our proposed estimators, especially the LE, can yield a performance very close to the CRLB with a known PLE, i.e., $CRLB_3$ in Appendix 2.6.3. However, a bad anchor node placement causes a considerable gap between our proposed estimators and $CRLB_3$.

To include the effect of different anchor node placements, in the following simulations, 10 anchor nodes will be randomly deployed within the $50\ m \times 50\ m$ field in each simulation trial and, hence, an average CRLB will be considered since the CRLB varies over the anchor node placement.

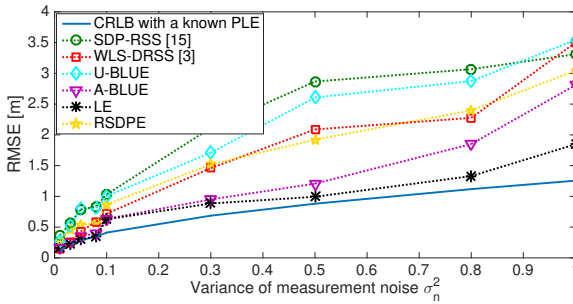
IMPACT OF THE MEASUREMENT NOISE

IN Fig. 2.3, we study all estimators with a perfectly known data model under large and small measurement noise when the PLE is known and fixed at $\gamma = 4$. The following observations can be made:

1. *U-BLUE*: Even based on a whitened data model and being a BLUE, the U-BLUE still yields a very bad performance especially under a large measurement noise since it does not consider the dependence in the parameter vector θ .
2. *WLS-DRSS*: Even requiring perfect knowledge of the variance of the measurement noise σ_n^2 to construct its system model and its weighting matrices, the WLS-DRSS is still no better than any of our proposed DRSS-based estimators except for the U-BLUE. This is because many approximations are used in its derivation and the DRSS measurements are used themselves to construct the weighting matrices. Therefore, when the measurement noise grows more severe, those approximations become more inaccurate and the DRSS measurements are more corrupted, making the weighting matrices less effective as they are in a small noise situation.
3. *A-BLUE*: Even without any knowledge of the variance of the measurement noise, the A-BLUE still has a better performance than the WLS-DRSS under a large measurement noise, as shown in Fig. 2.3(a). Under a small measurement noise, the A-BLUE becomes very accurate and only worse than the LE, as shown in Fig. 2.3(b). To explain this, the approximation in the second step of the A-BLUE is taken in the vicinity of the estimate $\hat{\mathbf{x}}_{u-blue}$ from the U-BLUE, which remains accurate under



(a) Impact of a large measurement noise.



(b) Impact of a small measurement noise.

Figure 2.3: Performance comparison of different estimators under different noise conditions when the actual PLE is known and fixed at $\gamma = 4$. The anchor location information is accurate.

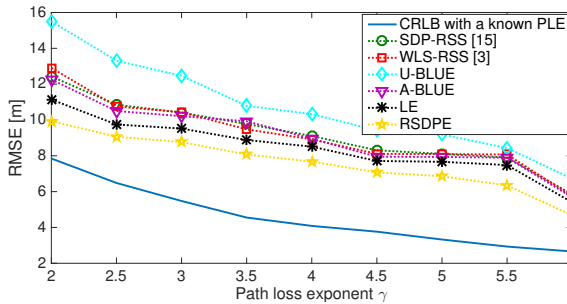


Figure 2.4: Performance of different estimators under different PLEs when the variance of the measurement noise is $\sigma_n^2 = 10$. The anchor location information is accurate.

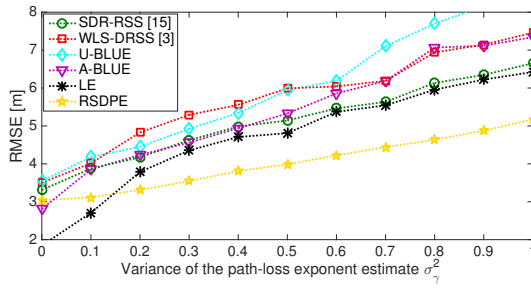
a small measurement noise. However, under a severe measurement noise, the U-BLUE yields a very bad performance and hence it becomes more difficult for the A-BLUE to fine-tune the U-BLUE estimate.

2

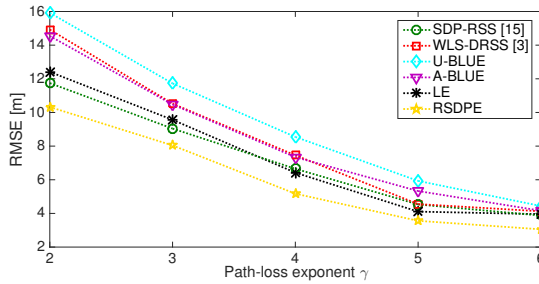
4. *LE*: The LE outperforms all the other estimators under a small measurement noise due to the fact that it requires neither any approximation nor dropping a constraint. In fact, the LE is the exact solution to our optimization problem in (2.19) if our model error \mathbf{v} is perfectly whitened. Therefore, we can observe from Fig. 2.3(b) that its performance is very close to the CRLB. However, the LE becomes only the second best estimator under a large measurement noise. To explain that, we need to recall that the approximation in (2.5) might become inaccurate under a large measurement noise, thus making our proposed whitened model not as effective as it is under a small measurement noise. On the other hand, a bad anchor node placement can also exacerbate the impact of a large measurement noise, which heavily deteriorates the performance of our proposed estimators.
5. *RSDPE*: Unlike the A-BLUE, the RSDPE does not use any approximation to deal with the non-linearity issue. Instead, the RSDPE uses the SDR procedure in (2.26) to guarantee a global yet suboptimal solution at a price of dropping the rank constraint in (2.25b). This explains why the RSDPE under a small measurement noise can not have a very accurate performance, which is merely close to the WLS-DRSS as shown in Fig. 2.3(b). However, the RSDPE surprisingly becomes the best estimator under a large measurement noise, as shown in Fig. 2.3(a). It seems that the RSDPE possesses a very good robustness to the deviation of the whitening procedure caused by the approximation inaccuracy in (2.5) under a large measurement noise. An interpretation for this is that this deviation yields the same impact as that of Δ_{Φ} . And the robustness to a bad anchor node placement is also shown in Fig. 2.2(c).
6. *SDP-RSS*: The SDP-RSS yields the worst performance under a small measurement noise. Besides the fact that the SDP procedure yields a suboptimal solution, this is also because using the ℓ_1 -norm might not be the best choice for the SDP-RSS due to a lack of ML optimality. Under a large measurement noise, the SDP-RSS becomes better, almost the same as the A-BLUE. However, the high computational complexity brought by the SDP procedure and a lack of robustness make it lose its advantage over our proposed estimators.

IMPACT OF THE PATH-LOSS EXPONENT

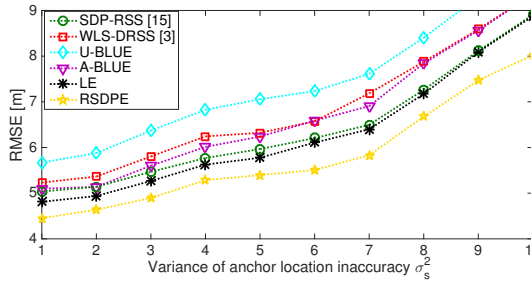
WE are also interested in how the PLE impacts our proposed estimators and hence we study our proposed estimators with a perfectly known data model under different PLEs when considering a large measurement noise. In fact, the PLE increases when the surrounding environment becomes more severe. Interestingly though, all the estimators grow more accurate in a more severe surrounding environment, as clearly depicted in Fig. 2.4. The performance of our proposed estimators can also be interpreted from our model error \mathbf{v} in (2.7c), where the covariance of \mathbf{v} obviously drops with an increasing PLE.



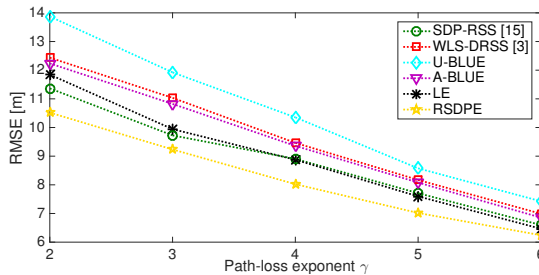
(a) Performance of different estimators under imperfect PLE knowledge when the actual PLE is $\gamma = 4$. The anchor location information is accurate.



(b) Performance of different estimators with an imperfect PLE estimate under different PLEs when the variance of the PLE estimate is $\sigma_\gamma^2 = 1$. The anchor location information is accurate.



(c) Performance of different estimators with inaccurate anchor location information when the actual PLE is known and $\gamma = 4$.



(d) Performance of different estimators under different PLEs when the actual PLE is known and the variance of the anchor location inaccuracy is $\sigma_s^2 = 10$.

Figure 2.5: Performance comparison of different estimators with model uncertainties when the variance of the measurement noise is $\sigma_n^2 = 1$.

IMPACT OF IMPERFECT PATH-LOSS EXPONENT ESTIMATE

WE previously assumed that the PLE γ is perfectly known. However, in practice, the PLE is calibrated or estimated before the localization phase [27, 28]. Hence, we have to consider the case where the PLE is not perfectly known, i.e., the model uncertainty is considered. Therefore, to study the performance of our proposed estimators in such a case, we have also conducted two MC simulations, where for each trial an imperfect PLE $\tilde{\gamma}$ is used to carry out the localization. The deviation $\Delta\gamma$, i.e., $\tilde{\gamma} \triangleq \gamma + \Delta\gamma$, of the imperfect PLE from the actual PLE is considered to be zero-mean Gaussian distributed with variance σ_γ^2 .

As shown in Fig. 2.5(a), all the estimators become worse with an increasing variance of the PLE estimate. The U-BLUE, the A-BLUE and the LE are all heavily impacted, while the RSDPE behaves relatively better, especially under a worse PLE estimate, due to its robust design.

To explain this in more detail, by using the imperfect PLE $\tilde{\gamma}$, the imperfect $P'_{i,1}$ used to construct our data model in (2.4) is given by $\tilde{P}'_{i,1} = 10^{\frac{P_{i,1}}{5(\gamma+\Delta\gamma)}}$. Using the first order Taylor series expansion of $\tilde{P}'_{i,1}$ w.r.t. $\Delta\gamma$, we obtain

$$\tilde{P}'_{i,1} = P'_{i,1} \left[1 - \frac{\ln(10)P_{i,1}}{5\gamma^2} \Delta\gamma \right]. \quad (2.36)$$

Then, for a sufficiently small noise, (2.2) can be presented as $P_{i,1} \approx -5\gamma \log_{10} \left(\frac{d_i^2}{d_1^2} \right)$ and hence (2.36) can be rewritten as

$$\tilde{P}'_{i,1} \approx P'_{i,1} \left[1 + \ln \left(\frac{d_i^2}{d_1^2} \right) \frac{\Delta\gamma}{\gamma} \right]. \quad (2.37)$$

Since $\Delta\gamma \sim \mathcal{N}(0, \sigma_\gamma^2)$, from (2.37), we can clearly see that an increasing variance σ_γ^2 of the imperfect PLE $\tilde{\gamma}$ incurs a more severe impact on our proposed estimators. Fortunately, under a large PLE γ , the impact of σ_γ^2 becomes less severe than under a small PLE, which can also be seen from Fig. 2.5(b). Finally, to better serve the following discussions, we should emphasize again that, among all the estimators, the RSDPE yields the best performance in case of an imperfect PLE.

IMPACT OF INACCURATE ANCHOR LOCATION INFORMATION

IN real life, the anchor location information might be inaccurate, if obtained by the global positioning system (GPS). Especially in military scenarios, this kind of information might be even more difficult to obtain, unreliable or tampered with by attackers. Therefore, we have to consider the model uncertainty in case of inaccurate anchor location information. Two MC simulations have been conducted, where for each trial every anchor location is given with a deviation, i.e., $\tilde{\mathbf{s}}_i \triangleq \mathbf{s}_i + \boldsymbol{\delta}_s, \forall i$, where $\boldsymbol{\delta}_s \sim \mathcal{N}(\mathbf{0}, \sigma_s^2 \mathbf{I}_d)$.

As shown in Fig. 2.5(c), all the estimators behave worse with an increasing variance of the anchor location inaccuracy, but the RSDPE again yields the best performance, due to its design for coping with model uncertainties. Finally, we notice that, if an inaccurate anchor location $\tilde{\mathbf{s}}_i$ is used for constructing our data model in (2.4), considering the fact that $\boldsymbol{\delta}_s$ is scaled by $P'_{i,1}$, a large PLE will lead to a small value of $P'_{i,1}$ and hence can suppress the impact of $\boldsymbol{\delta}_s$, which can be observed in Fig. 2.5(d).

Table 2.1: Comparison of different kinds of estimation for localization problem

Examples ^a	ULLSE U-BLUE, [15], [17] ^b , [18]	TLLSE A-BLUE, [3], [16], [33], [34]	SDRE RSDPE, [15], [19], [20-23]	EE LE, [31]
Considering $R = \mathbf{x}^T \mathbf{x}$	✗	✓	✓	✓
No approximation error ^c	-	✗	✓	✓
Considering all constraints	-	✓	✗	✓
Accuracy under a small measurement noise	Bad	Good	Medium	Very good
Accuracy under a large measurement noise	Very bad	Medium	Good	Good
Robustness to model uncertainties	Very bad	Bad	Good	Medium
Complexity	Very low	Low	High	Medium

^a For convenience, the examples are provided only in the field of RSS/DRSS-based localization, but the conclusions are not limited to this field.

^b [17] eliminates R by taking differences between a selected reference and the other nodes, which is equivalent to ignoring $R = \mathbf{x}^T \mathbf{x}$.

^c The approximation here only refers to the one related to the relation $R = \mathbf{x}^T \mathbf{x}$.

2.3.6. DISCUSSIONS

IN this subsection, we present the proposed estimators in a more general context. This discussion is also suitable for other localization problems, since there exist some common issues between the proposed localization problems and other ones.

For optimal localization problems, non-linearity and non-convexity issues are inevitable, both of which are due to the distance norm $d_i = \|\mathbf{x} - \mathbf{s}_i\|_2$. To be more specific, the distance norm is obviously non-linear w.r.t. \mathbf{x} , and the target cannot physically overlap with the anchors, i.e., $\mathbf{x} \neq \mathbf{s}_i, \forall i$, which explains the non-convexity. Most localization techniques first cope with non-linearity, either by directly applying a Taylor series expansion (TSE) around an appropriate value of \mathbf{x} , or by squaring and unfolding the distance norm. The former leads to some iterative ML methods, where a good initiation is critical for coping with the non-convexity as shown in Fig. 2.1(a). The latter one, which is our main focus, requires squaring the distance norm as $d_i^2 = \|\mathbf{x}\|_2^2 - 2\mathbf{s}_i^T \mathbf{x} + \|\mathbf{s}_i\|_2^2$, where $R \triangleq \|\mathbf{x}\|_2^2$ has to be considered as a new unknown parameter to avoid non-linearity. As a result, a linear unconstrained localization problem can be formulated, which has $[\mathbf{x}^T, R]^T$ as a new unknown parameter vector (other unknown parameters could be jointly estimated as well), directly leading to a closed-form (weighted) LS solution. We categorize this kind of estimator as the unconstrained linear least squares estimator (ULLSE), and obviously the ULLSE ignores the fact that the new parameter vector $[\mathbf{x}^T, R]^T$ (or the one that contains it) is still bound to a non-convex set. To cope with that, the relation $R = \mathbf{x}^T \mathbf{x}$ should be considered, and accordingly other localization techniques can be considered:

1. The two-step linear least squares estimation (TLLSE) first obtains an initial estimate from the ULLSE and then fine-tunes it in the second step based on $R = \mathbf{x}^T \mathbf{x}$, equivalently the constraints in (2.10b) and (2.14). The Achilles' heel of the TLLSE are approximations like (2.12) or (2.15), which are often carried out to facilitate the update of the estimate. The goodness of such approximations often relies on the ULLSE. Under a small measurement noise and an exactly known data model, the ULLSE and hence the approximations are reliable, leading to a very good performance of the TLLSE. However, when the measurement noise becomes severe or there exist considerable model uncertainties, the approximations deteriorate, thus significantly undermining the performance of the TLLSE. In this paper, the A-BLUE tries to minimize the impact of the approximations as much as possible, e.g., by sticking to the original data model instead of constructing a new one. Please refer to Section 2.3.1 for details and references therein.
2. The semidefinite relaxation (SDR)-based estimator (SDRE) reformulates $R = \mathbf{x}^T \mathbf{x}$ as an LMI as in (2.26c) such that an SDP problem can be constructed. More importantly though, this also requires introducing slack variables so as to change the optimization problem from minimizing the cost function to its upper bound. All those procedures lead to a relatively worse performance of the SDRE under a small measurement noise, but guarantee a very good estimation accuracy under a large noise. In this paper, the RSDPE is particularly improved with a robust design against model uncertainties. Please refer to Section 2.3.3 for details and references therein.
3. The exact estimator (EE) is the theoretically optimal solution when considering the

relation $R = \mathbf{x}^T \mathbf{x}$. The EE translates it into a new constraint as in (2.19b) without any approximation or dropping a constraint. Therefore, given an exactly known data model, if the solution that meets the KKT conditions can be precisely found, the EE should perform the best under both small and large measurement noises. It is worth noting that (2.19b) is still a non-convex constraint, which makes the search for the global solution very important. In this paper, the LE provides a useful interval, in which only the global solution resides. However, when the data model is uncertain, the global solution will be more difficult to find and hence the EE will not perform as good as expected. Please refer to Section 2.3.2 for details and references therein.

Localization techniques from the same category have a similar level of computational complexity and hence we can refer to Subsection 2.3.4. We give a general comparison of the ULLSE, the TLLSE, the SDRE and the EE in Table 2.1, where also some other examples beyond the proposed estimators are listed.

Also, it is very important to notice that most localization techniques (not limited to the RSS/DRSS-based) use a colored data model, which will generally degrade the localization performance and also explains why our proposed estimators are relatively better. Furthermore, some data models are very difficult or even impossible to whiten, since the true nodal distances might be required for whitening like the one in [3] and the famous Chan algorithm [17, 33–35]. Additionally, taking differences between the observations, e.g., generating TDOA or DRSS measurements, also leads to a colored model noise, which is often ignored in literature [17, 33, 34, 50].

After all, it is hard to say which kind of estimator is overall the best. Based on Table 2.1, we can choose the most suitable estimator or adaptively switch from one to another according to the demands. For example, if a low complexity is the most important consideration, the TLLSE could be the best choice. Under a severe measurement noise or given an uncertain data model, the SDRE is recommended. If there is no particular requirement, the EE is a good choice, since it has a good performance and yields the best accuracy under a small measurement noise.

2.4. ESTIMATOR FOR UNKNOWN PATH-LOSS MODEL

IN the previous sections, we have introduced robust DRSS-based localization for a known PLE. Based on these studies, we want to take one step further and explore a new iterative estimator which can jointly estimate the unknown PLE γ and the unknown location \mathbf{x} .

2.4.1. HANDLING UNKNOWN PATH-LOSS EXPONENT

Before introducing our new method, we would like to first discuss the current techniques to cope with an unknown PLE. Presently, many RSS/DRSS based localization methods assume a perfect pre-calibration stage without any PLE estimation error. Ironically though, PLE calibration techniques are still rarely studied. Here, we try to collect and summarize them in Table 2.2. The anchor-based methods [11, 51–53] have to be carried out between the anchors and hence are very susceptible to inaccurate anchor location information. Based on some geometric constraints, the anchor-free meth-

Table 2.2: Comparison of different methods for PLE calibration

Methods \ Drawbacks	Anchor Dependence	Intensive Node Cooperation	Not Pervasive ^a
Anchor-Based [11, 51–53]	✓	✓	✓
Anchor-Free [26, 54]	✗	✓	✓
Self-Estimation [25]	✗	✗	✓ ^b
Collective Self-Estimation [27, 28]	✗	✗	✗

^a A pervasive method is a method that can be implemented in any kind of wireless network, i.e., without any external assistance or information.

^b They still require some external information (e.g., network density) or a frequently changing receiver configuration and hence are not pervasive.

ods [26, 54] can estimate the unknown PLE for temporarily grouped nodes without any location information. But, they still require intensive node cooperation and might cause a heavy network load. Therefore, if each node can self-estimate the PLE in a distributed fashion, this could solve the aforementioned issues [25]. Pervasiveness is another shortcoming which we have to conquer, since the PLE is a very crucial wireless channel parameter and we want to enable a collective PLE self-estimation [27, 28] that can be used in any kind of wireless device for facilitating efficient communication and networking designs. In a nutshell, a more robust and cost effective PLE pre-calibration stage can undoubtedly benefit the localization procedure.

Alternatively, we can conveniently skip the PLE pre-calibration when it is not available or reliable. Then, we have to jointly estimate the unknown PLE and the target location, which could also save a lot of resources. In this section, we are particularly interested in this kind of solution. Commonly, an initial guess for the unknown PLE γ has to be adopted to obtain a quasi-estimate of the target location, which can then be used to update the PLE estimate [10, 17, 18, 22, 29–32]. Obviously, this will cause model uncertainties for the localization problem, which are often ignored however. Therefore, based on the previous studies, we want to seek a new robust DRSS localization approach in case of an unknown PLE.

2.4.2. PROTOTYPE OF THE PROPOSED ITERATIVE ESTIMATOR

IN addition to the model in (2.4), if given a known target location, we can obtain another linear model from (2.2) as

$$\boldsymbol{\pi} = \boldsymbol{\lambda}\gamma + \mathbf{v}, \quad (2.38)$$

where $\boldsymbol{\pi} \triangleq [\dots, P_{i,1}, \dots]^T$, $\mathbf{v} \triangleq [\dots, n_{i,1}, \dots]^T$ and $\boldsymbol{\lambda} \triangleq [\dots, -10\log_{10}\left(\frac{\|\mathbf{x}-\mathbf{s}_i\|_2}{\|\mathbf{x}-\mathbf{s}_1\|_2}\right), \dots]^T$. Again, we stack the equations for a fixed RN and all anchor nodes $i \neq 1$. However, it is very difficult to obtain a single linear model for both an unknown target location and an unknown PLE. This enlightens us that a block coordinate descent (BCD) method might be applicable to this problem [55]. In order to do so, we need to redefine the parameter vector $\boldsymbol{\theta}$ to be estimated as $\boldsymbol{\theta} \triangleq [\mathbf{x}^T, \|\mathbf{x}\|_2^2, \gamma]^T$. The BCD is implemented by partitioning $\boldsymbol{\theta}$ into two blocks, $[\mathbf{x}^T \|\mathbf{x}\|_2^2]^T$ and γ , and then at each iteration a cost function is minimized with respect to one of the blocks while the other is held fixed. We denote the $\boldsymbol{\theta}$ estimate at the k -th iteration as $\hat{\boldsymbol{\theta}}^{(k)}$, the iteration tolerance as ξ and the cost functions

for estimating the target location and the PLE respectively as $J'(\cdot)$ and $J''(\cdot)$. The prototype of our proposed estimator is presented in Algorithm 1.

Algorithm 1: PROTOTYPE of proposed iterative estimator.

- 1 *Initialization:* Choose the initial value $\hat{\boldsymbol{\theta}}^{(0)}$;
- 2 *Loop:* Given $\hat{\boldsymbol{\theta}}^{(k)} = [[\hat{\boldsymbol{\theta}}^{(k)}]_{1:d+1}^T, [\hat{\boldsymbol{\theta}}^{(k)}]_{d+2}^T]^T$, solve

$$[\hat{\boldsymbol{\theta}}^{(k+1)}]_{1:d+1} = \arg \min_{[\boldsymbol{\theta}]_{1:d+1}} J'([\boldsymbol{\theta}]_{1:d+1}, [\hat{\boldsymbol{\theta}}^{(k)}]_{d+2}); \quad (2.39)$$

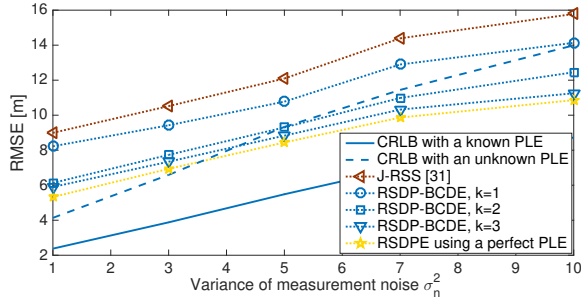
- 3 Given $[[\hat{\boldsymbol{\theta}}^{(k+1)}]_{1:d+1}^T, [\hat{\boldsymbol{\theta}}^{(k)}]_{d+2}^T]^T$, solve

$$[\hat{\boldsymbol{\theta}}^{(k+1)}]_{d+2} = \arg \min_{[\boldsymbol{\theta}]_{d+2}} J''([\hat{\boldsymbol{\theta}}^{(k+1)}]_{1:d+1}, [\boldsymbol{\theta}]_{d+2}); \quad (2.40)$$

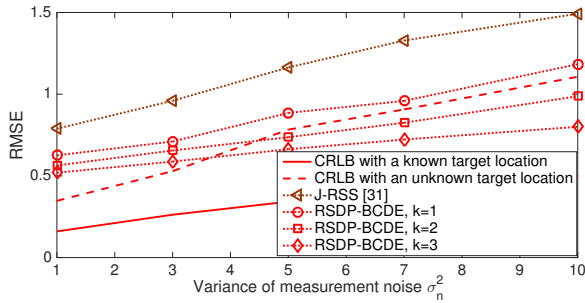
- 4 Let $\hat{\boldsymbol{\theta}}^{(k+1)} = [[\hat{\boldsymbol{\theta}}^{(k+1)}]_{1:d+1}^T, [\hat{\boldsymbol{\theta}}^{(k+1)}]_{d+2}^T]^T$;
 - 5 If $\|[\hat{\boldsymbol{\theta}}^{(k+1)}]_{1:d} - [\hat{\boldsymbol{\theta}}^{(k)}]_{1:d}\|_2 \leq \xi$, continue. Otherwise go back to *Loop*;
 - 6 **return** $\hat{\boldsymbol{\theta}}^{(k+1)}$
-

2.4.3. ROBUST SEMIDEFINITE PROGRAMMING BASED BLOCK COORDINATE DESCENT ESTIMATOR

To fully describe our method, we need to elaborate on the minimizations in (2.39) and (2.40). Since the RSDPE has a very good robustness to imperfect PLE knowledge, applying a similar method to (2.38) might also result in a good robustness to imperfect target location knowledge. Therefore, the idea behind our robust SDP-based block coordinate descent estimator (RSDP-BCDE) is to utilize this method to update both the location and the PLE. Considering that our method introduces two new auxiliary variables, next to the parameter vector $\boldsymbol{\theta}$, we introduce the slack variables t_1 , α_1 and t_2 , α_2 to update the target location estimate and the PLE estimate, respectively. Additionally, two bounds ζ_1 and ζ_2 are also needed, which are both computed in the same way as the RSDPE does for ζ .



(a) Accuracy of the location estimate.



(b) Accuracy of the PLE estimate.

Figure 2.6: Performance of our proposed RSDP-BCDE under different noise conditions: the PLE is $\gamma = 2$; the initial value of the PLE estimate is $\hat{\gamma}^{(0)} = 4$; k is the iteration number when the RSDP-BCDE stops the iterative procedure.

For updating the block $[\mathbf{x}^T, \|\mathbf{x}\|_2^2]^T$, we use

$$\begin{aligned} & \left([\hat{\boldsymbol{\theta}}^{(k+1)}]_{1:d+1}, \hat{t}_1^{(k+1)}, \hat{\alpha}_1^{(k+1)} \right) = \arg \min_{[\boldsymbol{\theta}]_{1:d+1}, t_1, \alpha_1} t_1 \\ & \text{subject to} \\ & \begin{bmatrix} (1 - \alpha_1) \mathbf{I}_{N-1} & \tilde{\boldsymbol{\Phi}}^{(k)} [\boldsymbol{\theta}]_{1:d+1} - \tilde{\boldsymbol{\rho}}^{(k)} & \mathbf{0} \\ (\tilde{\boldsymbol{\Phi}}^{(k)} [\boldsymbol{\theta}]_{1:d+1} - \tilde{\boldsymbol{\rho}}^{(k)})^T & t_1 & -\zeta_1 [\boldsymbol{\theta}]_{1:d+1}^T \\ \mathbf{0} & -\zeta_1 [\boldsymbol{\theta}]_{1:d+1} & \alpha_1 \mathbf{I}_{d+1} \end{bmatrix} \\ & \succeq \mathbf{0}, \\ & \begin{bmatrix} \mathbf{I}_{d \times d} & [\boldsymbol{\theta}]_{1:d} \\ [\boldsymbol{\theta}]_{1:d}^T & [\boldsymbol{\theta}]_{d+1} \end{bmatrix} \succeq \mathbf{0}, \end{aligned} \quad (2.41)$$

where $\hat{t}_1^{(k+1)}$ and $\hat{\alpha}_1^{(k+1)}$ are respectively the estimates of t_1 and α_1 at the $(k+1)$ -th iteration, $\tilde{\boldsymbol{\Phi}}^{(k)}$ and $\tilde{\boldsymbol{\rho}}^{(k)}$ are respectively the $\boldsymbol{\Phi}$ and the $\boldsymbol{\rho}$ constructed by the imperfect PLE estimate at the k -th iteration, i.e., $[\hat{\boldsymbol{\theta}}^{(k)}]_{d+2}$.

For updating γ , we notice from (2.38) that $\mathbf{v} = \boldsymbol{\Gamma} \mathbf{n}$ and hence the covariance matrix of \mathbf{v} is $\boldsymbol{\Sigma}_{\mathbf{v}} = \sigma_n^2 \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T$. Hence, the whitened model of (2.38) can be expressed as

$$\boldsymbol{\Sigma}_{\mathbf{v}}^{-1/2} \boldsymbol{\pi} = \boldsymbol{\Sigma}_{\mathbf{v}}^{-1/2} \boldsymbol{\lambda} \gamma + \boldsymbol{\Sigma}_{\mathbf{v}}^{-1/2} \mathbf{v} \quad (2.42a)$$

$$\Rightarrow (\boldsymbol{\Gamma} \boldsymbol{\Gamma}^T)^{-1/2} \boldsymbol{\pi} = (\boldsymbol{\Gamma} \boldsymbol{\Gamma}^T)^{-1/2} \boldsymbol{\lambda} \gamma + (\boldsymbol{\Gamma} \boldsymbol{\Gamma}^T)^{-1/2} \mathbf{v} \quad (2.42b)$$

$$\Rightarrow \mathbf{c} = \mathbf{d} \gamma + \mathbf{e}, \quad (2.42c)$$

where $\mathbf{c} \triangleq (\boldsymbol{\Gamma} \boldsymbol{\Gamma}^T)^{-1/2} \boldsymbol{\pi}$, $\mathbf{d} \triangleq (\boldsymbol{\Gamma} \boldsymbol{\Gamma}^T)^{-1/2} \boldsymbol{\lambda}$ and $\mathbf{e} \triangleq (\boldsymbol{\Gamma} \boldsymbol{\Gamma}^T)^{-1/2} \mathbf{v}$. Note that now the covariance matrix of \mathbf{e} is $\boldsymbol{\Sigma}_{\mathbf{e}} = \sigma_n^2 \mathbf{I}_N$. Based on the whitened data model (2.42c), we update γ as

$$\begin{aligned} & \left([\hat{\boldsymbol{\theta}}^{(k+1)}]_{d+2}, \hat{t}_2^{(k+1)}, \hat{\alpha}_2^{(k+1)} \right) = \arg \min_{[\boldsymbol{\theta}]_{d+2}, t_2, \alpha_2} t_2 \\ & \text{subject to} \end{aligned} \quad (2.43)$$

$$\begin{bmatrix} (1 - \alpha_2) \mathbf{I}_{N-1} & \mathbf{c} - \tilde{\mathbf{d}}^{(k+1)} [\boldsymbol{\theta}]_{d+2} & \mathbf{0} \\ (\mathbf{c} - \tilde{\mathbf{d}}^{(k+1)} [\boldsymbol{\theta}]_{d+2})^T & t_2 & -\zeta_2 [\boldsymbol{\theta}]_{d+2}^T \\ \mathbf{0} & -\zeta_2 [\boldsymbol{\theta}]_{d+2} & \alpha_2 \mathbf{I}_{d+1} \end{bmatrix} \succeq \mathbf{0},$$

where $\hat{t}_2^{(k+1)}$ and $\hat{\alpha}_2^{(k+1)}$ are respectively the estimates of t_2 and α_2 at the $(k+1)$ -th iteration, $\tilde{\mathbf{d}}^{(k+1)}$ is the \mathbf{d} constructed by the imperfect target location estimate at the $(k+1)$ -th iteration, i.e., $[\hat{\boldsymbol{\theta}}^{(k+1)}]_{1:d}$.

Finally, the optimization problems (2.41) and (2.43) are solved by CVX and the complexity of the RSDP-BCDE after k iterations is $O[kd^2 N^{2.5} \ln(1/\xi)]$.

2.4.4. NUMERICAL RESULTS

TO study the performance of the RSDP-BCDE, we have conducted an MC simulation. We select the initial value of the PLE estimate as 4 considering that the PLE normally

ranges from 2 to 6 [6]. In the simulation, the PLE is set to 2 and the rest of the MC simulation settings are the same as before. The numerical results are shown in Fig. 2.6.

The RSDP-BCDE is studied for different iteration numbers k and compared against one of the RSS-based estimators (J-RSS) from [31], which jointly estimates the unknown transmit power and PLE. According to the simulation results, our proposed method outperforms the J-RSS. As shown in Fig. 2.6(a), with more iterations, the performance of the RSDP-BCDE becomes better and gradually approaches that of the RSDPE using a perfect γ . The PLE estimate also becomes more accurate with an increasing iteration number k , as shown in Fig. 2.6(b). We also notice that, after the first iteration, the performance of the RSDP-BCDE is already very close to the CRLB with an unknown PLE or target location due to a good initial value of the PLE. Then, with more iterations, the knowledge of the target location and the PLE becomes better, thus improving the performance of the RSDP-BCDE over the CRLB with unknown PLE or target location. Additionally, the RSDP-BCDE converges quickly under a small measurement noise.

To end this section, we can conclude from the numerical results that even if the path-loss model is unknown, the RSDP-BCDE is still able to obtain an accurate location estimate. However, note that the SDP procedure has a very large complexity in each iteration. Hence, if the PLE is already accurate enough, we can similarly replace the SDP procedure with the A-BLUE or the LE to estimate the location such that the total computational complexity can be greatly reduced.

2.5. CONCLUSIONS

A whitened model for DRSS-based localization has been introduced and studied. Based on such a model, we have proposed and analyzed three different estimators for a known path-loss model (i.e., the A-BLUE, the LE and the RSDPE), where the latter is robust against an imperfect PLE estimate or inaccurate anchor location information. We have also proposed one robust iterative estimator for an unknown path-loss model (i.e., the RSDP-BCDE).

Simulation results have shown that, when the PLE is known, our three proposed estimators outperform an RSS-based joint estimator (SDP-RSS), which applies the SDP procedure on an ℓ_1 -norm approximation, as well as a recent weighted least squares estimator (WLS-DRSS), which requires perfect knowledge of the variance of the measurement noise. The performance of our three proposed estimators for a known PLE is studied under different noise conditions, different PLEs, imperfect PLE knowledge and inaccurate anchor location information. Their computational complexities are also investigated. Each estimator has its own advantages: the A-BLUE has the lowest computational complexity; the LE yields the best performance for a small measurement noise; and the RSDPE holds the best accuracy under a large measurement noise, an imperfect PLE and inaccurate anchor location information. Besides, in case of an unknown PLE, it is finally shown that, with more iterations, the performance of the RSDP-BCDE can approach that of the RSDPE with a known path-loss model. In real-life, to meet different practical demands when encountering different situations, different proposed estimators are provided as options.

2.6. APPENDICES

2.6.1. RSS COLLECTION

ASSUME the received signal $y(t)$ with the time index t can be expressed as

$$y(t) = x(t) \star h(t) + n(t), \quad (2.44)$$

where \star denotes the convolution operator, $x(t)$ is the transmitted signal, $h(t)$ indicates the channel response and $n(t)$ is the additive zero-mean white *Gaussian* noise. In most literature, the RSS refers to the signal power after a successful demodulation. To be specific, if the signal $y(t)$ can successfully be demodulated, $n(t)$ is cancelled and hence we can easily observe the signal envelope $r(t) \triangleq |x(t) \star h(t)|$, from which the RSS can be computed. In [4, 5], real-life experiments have been conducted to collect RSS measurements from demodulated signals, thereby demonstrating that the noise can be ignored. Moreover, the case of no signal demodulation is also investigated therein, but we feel this is beyond the scope of this paper.

To further explain the RSS collection procedure, we notice that $r(t)$ is affected by *small-scale* fading. For simplicity, we will remove the time index t from now on to represent an instantaneous value. If the *Nakagami- m* distribution is considered, which actually characterizes the instantaneous signal envelope r , the instantaneous received power $p \triangleq r^2$ is *Gamma* distributed as

$$\mathbb{P}(p|\Omega) = \frac{1}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m p^{m-1} e^{-\frac{mp}{\Omega}}. \quad (2.45)$$

where m is the fading parameter and a small value of m indicates a severe fading. The other parameter Ω , defined as $\Omega \triangleq E(p)$, is the RSS to collect, but expressed in *watts*. Therefore, collecting RSS measurements corresponds to estimating Ω .

Denoting Ω_i as the Ω associated with the i -th anchor, the maximum likelihood (ML) estimate of Ω_i is readily given by

$$\hat{\Omega}_i = \frac{1}{K} \sum_{k=1}^K p_i^{(k)}, \quad (2.46)$$

where $p_i^{(k)}, \forall k = 1, \dots, K$, represent K consecutive samples of p related to the i -th anchor. Due to the fact that $\text{Var}(p) = \frac{\Omega^2}{m}$, we can easily obtain

$$\text{Var}(\hat{\Omega}_i) = \frac{\Omega_i^2}{Km}, \quad (2.47)$$

which indicates that the estimation error can be reduced by taking more samples. Obviously, the ML estimate $\hat{\Omega}_i$ is unbiased and, denoting $\hat{\Omega}_i = \Omega_i + \Delta\Omega_i$, $\Delta\Omega_i$ is asymptotically *Normal* distributed as $\Delta\Omega_i \sim \mathcal{N}(0, \sigma_{\Delta\Omega_i}^2)$ with $\sigma_{\Delta\Omega_i}^2 = \frac{\Omega_i^2}{Km}$.

Expressing the RSS estimate in *dB* and applying the first-order Taylor series expansion w.r.t. $\Delta\Omega_i$ results in

$$\hat{P}_i = 10\log_{10}(\hat{\Omega}_i) = 10\log_{10}(\Omega_i + \Delta\Omega_i) \approx P_i + \Delta P_i, \quad (2.48)$$

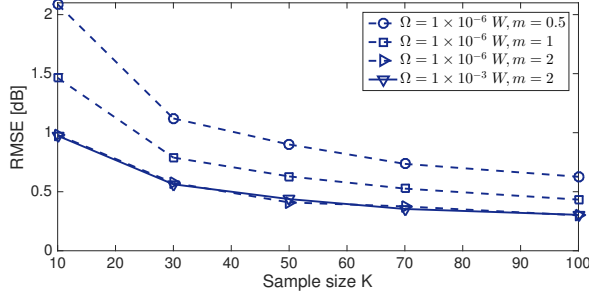


Figure 2.7: Performance of RSS collection under different sample sizes and different values of m and Ω .

where the estimation error of \hat{P}_i is denoted as $\Delta P_i \triangleq \frac{10}{\ln(10)} \frac{\Delta \Omega_i}{\Omega_i}$ and hence $\Delta P_i \sim \mathcal{N}(0, \sigma_{\Delta P_i}^2)$ with $\sigma_{\Delta P_i}^2 = \frac{100}{\ln(10)^2 K m}$. We notice that, compared with $\sigma_{\Delta \Omega_i}^2$, $\sigma_{\Delta P_i}^2$ does not depend on Ω_i any more, i.e., the RSS estimation error in dB is independent of the anchors. This means that, even if not enough samples are collected, the impact of ΔP_i is still similar to that of the shadowing effect χ_i . We have also conducted a simple simulation for RSS collection. As shown in Fig. 2.7, the collection error decreases with a large m and more samples. But more importantly, different values of Ω yield no significant impact on the collection error if considered in dB. In a nutshell, we can assume the RSS is perfectly collected in this paper without loss of generality.

2.6.2. DERIVATION FROM RSS-BASED MODEL

IN this appendix, we show that our whitened DRSS-based model can also be derived from a properly whitened RSS-based model after orthogonally projecting out the unknown \tilde{P}_0 .

To show that, let us first rewrite (2.1) as

$$\|\mathbf{x} - \mathbf{s}_i\|_2^2 = \frac{\tilde{P}'_0 \Delta P'_{0,i} \chi'_i}{P'_i}, \quad (2.49)$$

where $P'_i \triangleq 10^{\frac{P_i}{5\gamma}}$, $\tilde{P}'_0 \triangleq 10^{\frac{\tilde{P}_0}{5\gamma}}$, $\Delta P'_{0,i} \triangleq 10^{\frac{\Delta P_{0,i}}{5\gamma}}$, $\chi'_i \triangleq 10^{\frac{\chi_i}{5\gamma}}$ and the reference distance is again $d_0 = 1$ m without loss of generality. For a sufficiently small noise and using the first-order Taylor series expansion on (2.49), we obtain

$$\|\mathbf{x}\|_2^2 - 2\mathbf{s}_i^T \mathbf{x} + \|\mathbf{s}_i\|_2^2 = \frac{\tilde{P}'_0}{P'_i} \left[1 + \frac{\ln(10)}{5\gamma} n_i \right], \quad (2.50)$$

where $n_i = \Delta P_{0,i} + \chi_i$. Then, we can formulate a linear model as

$$\mathbf{B}\boldsymbol{\phi} = \mathbf{h} + \boldsymbol{\zeta} \quad (2.51)$$

where

$$\mathbf{B} \triangleq \begin{bmatrix} \vdots & \vdots & \vdots \\ 2\mathbf{s}_i^T & -1 & 1/P'_i \\ \vdots & \vdots & \vdots \end{bmatrix},$$

$\boldsymbol{\phi} \triangleq [\mathbf{x}, \|\mathbf{x}\|_2^2, \bar{P}'_0]^T$, $\mathbf{h} \triangleq [\dots, \|\mathbf{s}_i\|_2^2, \dots]^T$ and $\boldsymbol{\zeta} \triangleq [\dots, -\frac{\ln(10)\bar{P}'_0}{5\gamma P'_i} n_i, \dots]^T$. Every element of $\boldsymbol{\zeta}$, say ζ_i , is a zero-mean Gaussian variable with variance $\frac{[\ln(10)]^2 \bar{P}'_0{}^2 \sigma_n^2}{25\gamma^2 P_i'^2}$ and hence the covariance matrix of $\boldsymbol{\zeta}$ can be expressed as $\boldsymbol{\Sigma}_\zeta = \frac{\ln(10)^2 \bar{P}'_0{}^2 \sigma_n^2}{25\gamma^2} \mathbf{D}'^{-2}$, where $\mathbf{D}' = \text{diag}([P'_1, \dots, P'_N]^T)$ with $\text{diag}(\cdot)$ a diagonal matrix with its argument on the diagonal.

Let us first describe the relation between the model in (2.4) and the RSS-based model in (2.51). By recalling the definitions of \mathbf{p} and \mathbf{h} , we can easily observe that $\mathbf{p} = \mathbf{P}\mathbf{h}$, where $\mathbf{P} \triangleq \begin{bmatrix} \mathbf{1}_{(N-1) \times 1} & \text{diag}([-P'_{2,1}, \dots, -P'_{N,1}]^T) \end{bmatrix} = -\frac{1}{P'_1} \boldsymbol{\Gamma} \mathbf{D}'$, and similarly, $\mathbf{P}\mathbf{B}\boldsymbol{\phi} = \boldsymbol{\Psi}\boldsymbol{\theta}$ and $\mathbf{P}\boldsymbol{\zeta} = \boldsymbol{\epsilon}$.

Hence, before the whitening procedure, the DRSS-based model (2.4) can be viewed as the RSS-based model (2.51), where we remove the influence of \bar{P}'_0 by applying the transformation matrix \mathbf{P} . However, it is hard to judge at this point whether this operation will cause a loss of information or not.

In order to do that, let us first whiten the RSS-based model (2.51), which leads to

$$\boldsymbol{\Sigma}_\zeta^{-1/2} \mathbf{B}\boldsymbol{\phi} = \boldsymbol{\Sigma}_\zeta^{-1/2} \mathbf{h} + \boldsymbol{\Sigma}_\zeta^{-1/2} \boldsymbol{\zeta} \quad (2.52a)$$

$$\Rightarrow \mathbf{D}' \mathbf{B}\boldsymbol{\phi} = \mathbf{D}' \mathbf{h} + \mathbf{D}' \boldsymbol{\zeta} \quad (2.52b)$$

$$\Rightarrow \mathbf{B}' \boldsymbol{\phi} = \mathbf{h}' + \boldsymbol{\zeta}', \quad (2.52c)$$

where $\mathbf{B}' \triangleq \mathbf{D}' \mathbf{B}$, $\mathbf{h}' \triangleq \mathbf{D}' \mathbf{h}$ and $\boldsymbol{\zeta}' \triangleq \mathbf{D}' \boldsymbol{\zeta}$ with the covariance matrix $\boldsymbol{\Sigma}_{\zeta'} = \frac{\ln(10)^2 \bar{P}'_0{}^2 \sigma_n^2}{25\gamma^2} \mathbf{I}_N$.

Now, to see the relation between our whitened DRSS-based model (2.7) and this whitened RSS-based model (2.52), we can show that $\boldsymbol{\Phi}\boldsymbol{\theta} = (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2} \mathbf{P}\mathbf{B}\boldsymbol{\phi} = \mathbf{P}' \mathbf{B}' \boldsymbol{\phi}$, where $\mathbf{P}' \triangleq -\frac{1}{P'_1} (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T)^{-1/2} \boldsymbol{\Gamma}$. So, after the whitening procedure, the whitened DRSS-based model (2.7) can be viewed as the whitened RSS-based model (2.52), where we remove the influence of \bar{P}'_0 by applying the transformation matrix \mathbf{P}' . The crucial observation now is that this transformation matrix \mathbf{P}' is a (scaled) unitary operator, i.e., $\mathbf{P}' \mathbf{P}'^T = \frac{1}{P_1'^2} \mathbf{I}_{N \times N}$, and hence by taking differences of RSSs to eliminate the unknown transmit power, our whitened DRSS-based model does not entail any loss of information compared to the whitened RSS-based model.

2.6.3. CRAMÉR-RAO LOWER BOUNDS

To derive the Cramér-Rao lower bounds (CRLBs) used in this paper, we recall from (2.2) that the vector of DRSS samples, say $\boldsymbol{\pi}$, is Gaussian distributed as $\boldsymbol{\pi} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_\pi)$, where for $i \neq 1$ we have $\boldsymbol{\pi} \triangleq [\dots, P_{i,1}, \dots]^T$, $\boldsymbol{\mu} = [\dots, \mu_i, \dots]^T$ with $\mu_i = -10\gamma \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{\|\mathbf{x} - \mathbf{s}_1\|_2} \right)$ and according to (2.6), $\boldsymbol{\Sigma}_\pi = \sigma_n^2 \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T$.

To obtain the CRLB, the Fisher information matrix (FIM) can be computed as [56]

$$\mathbb{J}]_{n,m} = \left[\frac{\partial \boldsymbol{\mu}}{\partial \theta_n} \right]^T \boldsymbol{\Sigma}_\pi^{-1} \left[\frac{\partial \boldsymbol{\mu}}{\partial \theta_m} \right] + \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}_\pi^{-1} \frac{\partial \boldsymbol{\Sigma}_\pi}{\partial \theta_n} \boldsymbol{\Sigma}_\pi^{-1} \frac{\partial \boldsymbol{\Sigma}_\pi}{\partial \theta_m} \right], \quad (2.53)$$

where depending on the scenarios $\boldsymbol{\theta} = \mathbf{x}$, $\boldsymbol{\theta} = [\mathbf{x}^T, \gamma]^T$, or θ is the scalar $\theta = \gamma$, and $\frac{\partial \boldsymbol{\mu}}{\partial \theta_n} \triangleq [\dots, \frac{\partial \mu_i}{\partial \theta_n}, \dots]^T$. Since $\boldsymbol{\Sigma}_\pi$ does not depend on $\boldsymbol{\theta}$, we can simplify (2.53) as $\mathbb{J}]_{n,m} =$

$\left[\frac{\partial \boldsymbol{\mu}}{\partial \theta_n} \right]^T \boldsymbol{\Sigma}^{-1} \left[\frac{\partial \boldsymbol{\mu}}{\partial \theta_m} \right]$. Letting $\mathbf{x} = [x_1, \dots, x_d]^T$ and $\mathbf{s}_i = [s_{i,1}, \dots, s_{i,d}]^T$, we obtain

$$\frac{\partial [\boldsymbol{\mu}]_i}{\partial x_k} = -\frac{10\gamma}{\ln(10)} \frac{(x_k - s_{i,k}) \|\mathbf{x} - \mathbf{s}_1\|_2^2 - (x_k - s_{1,k}) \|\mathbf{x} - \mathbf{s}_i\|_2^2}{\|\mathbf{x} - \mathbf{s}_i\|_2^2 \|\mathbf{x} - \mathbf{s}_1\|_2^2}, \quad (2.54)$$

$$k = 1, \dots, d$$

and $\frac{\partial [\boldsymbol{\mu}]_i}{\partial \gamma} = -10 \log_{10} \left(\frac{\|\mathbf{x} - \mathbf{s}_i\|_2}{\|\mathbf{x} - \mathbf{s}_1\|_2} \right)$.

CRLBs on Joint Location Estimate and PLE Estimate In this case, $\boldsymbol{\theta} = [\mathbf{x}^T, \gamma]^T$ in \mathbb{R}^{d+1} and the CRLB for the location estimate is obtained as

$$CRLB_1 = \sqrt{\sum_{k=1}^d [\mathbf{J}^{-1}]_{k,k}}$$

while the CRLB for the PLE estimate is obtained as

$$CRLB_2 = \sqrt{[\mathbf{J}^{-1}]_{d+1,d+1}}$$

CRLB on Location Estimate with a Known PLE In this case, $\boldsymbol{\theta} = \mathbf{x}$ in \mathbb{R}^d and the CRLB for the location estimate with a known PLE is obtained as

$$CRLB_3 = \sqrt{\sum_{k=1}^d [\mathbf{J}^{-1}]_{k,k}}$$

CRLB on PLE estimate with a known location In this case, $\theta = \gamma$ and the CRLB for the PLE estimate with a known location is simply given by

$$CRLB_4 = \sqrt{1/[\mathbf{J}]_{1,1}},$$

where we note that \mathbf{J} is just a scalar here.

REFERENCES

- [1] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, *Locating the nodes: cooperative localization in wireless sensor networks*, [Signal Processing Magazine, IEEE 22, 54 \(2005\)](#).
- [2] K. Whitehouse, C. Karlof, and D. Culler, *A Practical Evaluation of Radio Signal Strength for Ranging-based Localization*, [SIGMOBILE Mob. Comput. Commun. Rev. 11, 41 \(2007\)](#).
- [3] L. Lin, H. So, and Y. Chan, *Accurate and simple source localization using differential received signal strength*, [Digital Signal Processing 23, 736 \(2013\)](#).

- [4] R. Martin, A. King, J. Pennington, R. Thomas, R. Lenahan, and C. Lawyer, *Modeling and mitigating noise and nuisance parameters in received signal strength positioning*, *Signal Processing, IEEE Transactions on* **60**, 5451 (2012).
- [5] R. K. Martin, R. W. Thomas, and Z. Wu, *Using spectral correlation for non-cooperative rssi-based positioning*, in *2011 IEEE Statistical Signal Processing Workshop (SSP)* (2011) pp. 241–244.
- [6] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001).
- [7] B. Sklar, *Rayleigh fading channels in mobile digital communication systems .I. Characterization*, *Communications Magazine, IEEE* **35**, 90 (1997).
- [8] N. Nakagami, *The m-distribution, a general formula for intensity distribution of rapid fading*, in *Statistical Methods in Radio Wave Propagation*, edited by W. G. Hoffman (Oxford, England: Pergamon, 1960).
- [9] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).
- [10] X. Li, *RSS-Based Location Estimation with Unknown Pathloss Model*, *Wireless Communications, IEEE Transactions on* **5**, 3626 (2006).
- [11] A. Coluccia and F. Ricciato, *On ML estimation for automatic RSS-based indoor localization*, in *Wireless Pervasive Computing (ISWPC), 2010 5th IEEE International Symposium on* (2010) pp. 495–502.
- [12] N. Patwari, A. Hero, M. Perkins, N. Correal, and R. O’Dea, *Relative location estimation in wireless sensor networks*, *Signal Processing, IEEE Transactions on* **51**, 2137 (2003).
- [13] X. Li, *Collaborative Localization With Received-Signal Strength in Wireless Sensor Networks*, *Vehicular Technology, IEEE Transactions on* **56**, 3807 (2007).
- [14] J. H. Lee and R. M. Buehrer, *Location estimation using differential rssi with spatially correlated shadowing*, in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE* (2009) pp. 1–6.
- [15] R. Vaghefi, M. Gholami, and E. Strom, *RSS-based sensor localization with unknown transmit power*, in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on* (2011) pp. 2480–2483.
- [16] H. C. So and L. Lin, *Linear Least Squares Approach for Accurate Received Signal Strength Based Source Localization*, *Signal Processing, IEEE Transactions on* **59**, 4035 (2011).
- [17] N. Salman, A. Kemp, and M. Ghogho, *Low Complexity Joint Estimation of Location and Path-Loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 364 (2012).

- [18] Y. Xu, J. Zhou, and P. Zhang, *Rss-based source localization when path-loss model parameters are unknown*, *IEEE Communications Letters* **18**, 1055 (2014).
- [19] C. Meng, Z. Ding, and S. Dasgupta, *A Semidefinite Programming Approach to Source Localization in Wireless Sensor Networks*, *Signal Processing Letters, IEEE* **15**, 253 (2008).
- [20] R. Ouyang, A.-S. Wong, and C.-T. Lea, *Received Signal Strength-Based Wireless Localization via Semidefinite Programming: Noncooperative and Cooperative Schemes*, *Vehicular Technology, IEEE Transactions on* **59**, 1307 (2010).
- [21] R. Vaghefi, M. Gholami, R. Buehrer, and E. Strom, *Cooperative Received Signal Strength-Based Sensor Localization With Unknown Transmit Powers*, *Signal Processing, IEEE Transactions on* **61**, 1389 (2013).
- [22] G. Wang, H. Chen, Y. Li, and M. Jin, *On Received-Signal-Strength Based Localization with Unknown Transmit Power and Path Loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 536 (2012).
- [23] G. Wang and K. Yang, *A New Approach to Sensor Node Localization Using RSS Measurements in Wireless Sensor Networks*, *Wireless Communications, IEEE Transactions on* **10**, 1389 (2011).
- [24] Y.-Y. Cheng and Y.-Y. Lin, *A new received signal strength based location estimation scheme for wireless sensor network*, *Consumer Electronics, IEEE Transactions on* **55**, 1295 (2009).
- [25] S. Srinivasa and M. Haenggi, *Path Loss Exponent Estimation in Large Wireless Networks*, in *Information Theory and Applications Workshop* (2009) pp. 124–129.
- [26] G. Mao, B. D. O. Anderson, and B. Fidan, *Path Loss Exponent Estimation for Wireless Sensor Network Localization*, *Comput. Netw.* **51**, 2467 (2007).
- [27] Y. Hu and G. Leus, *Self-estimation of path-loss exponent in wireless networks and applications*, *Vehicular Technology, IEEE Transactions on* **PP**, 1 (2014).
- [28] Y. Hu and G. Leus, *Directional maximum likelihood self-estimation of the path-loss exponent*, in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2016) pp. 3806–3810.
- [29] N. Salman, M. Ghogho, and A. Kemp, *On the Joint Estimation of the RSS-Based Location and Path-loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 34 (2012).
- [30] S. Tomic, M. Beko, and R. Dinis, *Rss-based localization in wireless sensor networks using convex relaxation: Noncooperative and cooperative schemes*, *IEEE Transactions on Vehicular Technology* **64**, 2037 (2015).
- [31] M. Gholami, R. Vaghefi, and E. Strom, *RSS-Based Sensor Localization in the Presence of Unknown Channel Parameters*, *Signal Processing, IEEE Transactions on* **61**, 3752 (2013).

- [32] C. Liang and F. Wen, *Received signal strength-based robust cooperative localization with dynamic path loss model*, *IEEE Sensors Journal* **16**, 1265 (2016).
- [33] B.-C. Liu and K.-H. Lin, *Distance Difference Error Correction by Least Square for Stationary Signal-Strength-Difference-Based Hyperbolic Location in Cellular Communications*, *Vehicular Technology, IEEE Transactions on* **57**, 227 (2008).
- [34] B.-C. Liu, K.-H. Lin, and J.-C. Wu, *Analysis of hyperbolic and circular positioning algorithms using stationary signal-strength-difference measurements in wireless communications*, *Vehicular Technology, IEEE Transactions on* **55**, 499 (2006).
- [35] Y. Chan and K. Ho, *A simple and efficient estimator for hyperbolic location*, *Signal Processing, IEEE Transactions on* **42**, 1905 (1994).
- [36] K. Ho, *Bias reduction for an explicit solution of source localization using tdoa*, *Signal Processing, IEEE Transactions on* **60**, 2101 (2012).
- [37] K. Cheung, H. So, W.-K. Ma, and Y. Chan, *Least squares algorithms for time-of-arrival-based mobile location*, *Signal Processing, IEEE Transactions on* **52**, 1121 (2004).
- [38] Y. Wang and G. Leus, *Reference-free time-based localization for an asynchronous target*, *EURASIP Journal on Advances in Signal Processing* **2012**, 19 (2012), 10.1186/1687-6180-2012-19.
- [39] T. Pong and H. Wolkowicz, *The generalized trust region subproblem*, *Computational Optimization and Applications* **58**, 273 (2014).
- [40] J. J. More, *Generalizations Of The Trust Region Problem*, *OPTIMIZATION METHODS AND SOFTWARE* **2**, 189 (1993).
- [41] J. More, J. J. and D. Sorensen, *Computing a Trust Region Step*, *SIAM Journal on Scientific and Statistical Computing* **4**, 553 (1983), <http://dx.doi.org/10.1137/0904038>.
- [42] A. Beck, P. Stoica, and J. Li, *Exact and Approximate Solutions of Source Localization Problems*, *Signal Processing, IEEE Transactions on* **56**, 1770 (2008).
- [43] L. Vandenberghe and S. Boyd, *Semidefinite Programming*, *SIAM Review* **38**, 49 (1994).
- [44] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, Vol. 15 (SIAM, 1994).
- [45] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 2.1*, <http://cvxr.com/cvx> (2014).
- [46] M. Grant and S. Boyd, *Graph implementations for nonsmooth convex programs*, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag Limited, 2008) pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.

- [47] I. Markovsky and S. Van Huffel, *Overview of total least-squares methods*, [Signal Process.](#) **87**, 2283 (2007).
- [48] C. Eckart and G. Young, *The Approximation of One Matrix by Another of Lower Rank*, [Psychometrika](#) **1**, 211 (1936).
- [49] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)* (Johns Hopkins University Press, Baltimore, MD, USA, 1996).
- [50] A. H. Sayed, A. Tarighat, and N. Khajehnouri, *Network-based wireless location: challenges faced in developing techniques for accurate wireless location information*, [IEEE Signal Processing Magazine](#) **22**, 24 (2005).
- [51] F. Santucci and N. Benvenuto, *A least squares path loss estimation approach to handover algorithms*, in [Proceedings of ICC/SUPERCOMM '96 - International Conference on Communications](#), Vol. 2 (IEEE, 1996) pp. 802–806.
- [52] C. C. Pu, S. Y. Lim, and P. C. Ooi, *Measurement arrangement for the estimation of path loss exponent in wireless sensor network*, in [Computing and Convergence Technology \(ICCT\), 2012 7th International Conference on](#) (2012) pp. 807–812.
- [53] L. Razoumov and L. Greenstein, *Path loss estimation algorithms and results for RF sensor networks*, in [IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004](#), Vol. 7 (IEEE, 2004) pp. 4593–4596.
- [54] G. Mao, B. D. O. Anderson, and B. Fidan, *WSN06-4: Online Calibration of Path Loss Exponent in Wireless Sensor Networks*, in [IEEE Globecom 2006](#) (IEEE, 2006) pp. 1–6.
- [55] P. Tseng, *Convergence of a Block Coordinate Descent Method for Nondifferentiable Minimization*, [Journal of Optimization Theory and Applications](#) **109**, 475 (2001).
- [56] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993).

3

ON A UNIFIED FRAMEWORK FOR LINEAR NUISANCE PARAMETERS

Yongchang HU and Geert LEUS

Cease to struggle and you cease to live.

Thomas Carlyle

Estimation problems in the presence of deterministic linear nuisance parameters arise in a variety of fields. To cope with those, three common methods are widely considered: 1) jointly estimating the parameters of interest and the nuisance parameters; 2) projecting out the nuisance parameters; 3) selecting a reference and then taking differences between the reference and the observations, which we will refer to as "differential signal processing". A lot of literature has been devoted to these methods, yet all following separate paths.

Based on a unified framework, we analytically explore the relations between these three methods, where we particularly focus on the third one and introduce a general differential approach to cope with multiple distinct nuisance parameters. After a proper whitening procedure, the corresponding best linear unbiased estimators (BLUEs) are shown to be all equivalent to each other. Accordingly, we unveil some surprising facts, which are in contrast to what is commonly considered in literature, e.g., the reference choice is actually not important for the differencing process. Since this paper formulates the problem in a general manner, one may specialize our conclusions to any particular application. Some localization examples are also presented in this paper to verify our conclusions.

3.1. INTRODUCTION

THE problem of estimating unknown parameters of interest $\mathbf{x} \in \mathbb{R}^{L \times 1}$ observed through a linear transformation $\mathbf{H} \in \mathbb{R}^{N \times L}$ ($N > L$), and corrupted by additive noise $\mathbf{n} \in \mathbb{R}^{N \times 1}$, has been well studied and considered in a wide variety of fields [1]. However, the observations $\mathbf{y} \in \mathbb{R}^{N \times 1}$ are sometimes also influenced by unknown linear nuisance parameters, denoted by $\mathbf{u} \in \mathbb{R}^{M \times 1}$ which enter \mathbf{y} through the linear transformation $\mathbf{G} \in \mathbb{R}^{N \times M}$ ($N > M$). For instance, these nuisance parameters could be some common offsets such as the transmit time, the clock bias and the transmit power in time-of-arrival (TOA) or received signal strength (RSS) based localization [2], or they could represent some redundant signals like the undesired signatures in hyperspectral imaging [3]. In fact, estimation problem with linear nuisance parameters widely exist in many other fields such as communications [4–6], source separation [7] and machine learning [8, 9]. Though, only Bayesian approaches are generally studied in case of nuisance parameters [1, 10, 11]. In this paper, we mainly adopt deterministic approaches, for which we first formulate our general model with linear nuisance parameters as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{G}\mathbf{u} + \mathbf{n}, \quad (3.1)$$

where we assume that

1. the concatenation of \mathbf{H} and \mathbf{G} has full column rank, i.e., $\text{Rank}([\mathbf{H} \ \mathbf{G}]) = L + M$;
2. the noise \mathbf{n} is zero-mean, i.e., the expected value of \mathbf{n} is $E(\mathbf{n}) = \mathbf{0}$;
3. the noise \mathbf{n} is white (e.g. after whitening), i.e., the covariance matrix $\Sigma_{\mathbf{n}}$ is (scaled) identity $\Sigma_{\mathbf{n}} = \sigma^2 \mathbf{I}_N$, where \mathbf{I}_N is the $N \times N$ identity matrix.

Note the noise \mathbf{n} does not have to be Gaussian distributed¹, although it is true for many cases.

To cope with this kind of problem in case \mathbf{u} is deterministic, three methods are often considered: 1) the joint estimation approach estimates the unknown \mathbf{x} together with the unknown nuisance term \mathbf{u} (e.g., the location and the unknown clock bias in [13]); 2) the orthogonal subspace projection (OSP) approach projects out the nuisance term \mathbf{u} such that the resulting observation vector is only subject to \mathbf{x} (e.g., the extraction of the desired signature in [14]); 3) the differential signal processing approach firstly chooses a reference and then estimates \mathbf{x} from the differences between the reference and the observations [15–19]. Note that these methods obviously result in three distinct observation sets with different signal-to-noise ratios (SNRs), which will greatly influence the estimation performance. Therefore, a vast amount of research has been conducted on these methods, though all following separate paths. Admittedly, some early results have been reported bridging the first two methods. For instance, the famous OSP-based solution using a matched filter to maximize the output SNR proposed in [20] was later on proven to be equivalent to the least squares (LS) approach based on the joint estimation [21, 22]. However, the proposed differential approaches are still widely regarded as a common but distinct way to cope with linear nuisance parameters. One of the most famous applications is time-based localization (TOA or time-difference-of-arrival (TDOA)), where

¹For example, the noise \mathbf{n} could also be uniform, Laplace or student's t-distributed [12]

many papers exist on selecting an optimal reference [23–25], constructing an optimal observation subset [26–28], or just using the full observation set adopting each sample as a reference [29–31]. All these issues never occur in the first two methods due to the fact that they are free of a reference. In a nutshell, there still seems to be a huge and inevitable gap between the differential approaches and the other two.

This paper analytically investigates the relations between all three methods, where the corresponding best linear unbiased estimators (BLUEs) are presented and discussed. Since the general framework in (3.1) is used throughout this paper, all the conclusions apply to any kind of problem that can be written in this form, which is exactly the strength of this paper. We also present some localization examples to verify our conclusions. To summarize, the main contributions of this paper are listed below.

1. For the first time, we extend the differential signal processing approach to a more general framework, which can cope with multiple nuisance parameters, whereas most existing methods consider a single nuisance parameter.
2. Surprisingly, the BLUEs of the three considered methods are proven rigorously to be identical to each other if an appropriate preprocessing step is used. This might be expected or known w.r.t. the first two methods, but the equivalence with differential methods has never been reported before.
3. Compared with the joint estimation method, which directly utilizes all the original observations, none of the other two methods suffers any information loss.
4. Although differential methods seem to rely on the selected reference, selecting the right reference is not important since there is no actual trace of the selected reference in the corresponding BLUE. This is in sharp contrast to what is commonly considered in literature.
5. As far as the differencing process is concerned, the differential observation set associated with a single reference already preserves the full data information.

The rest of this paper is organized as follows. Section 3.2 presents the relations between the three considered methods. Some examples of source localization are shown and numerically studied to support our conclusions in Section 3.3. Finally, Section 3.4 summarizes this paper.

3.2. HANDLING LINEAR NUISANCE PARAMETERS

IN this section, we study the relations between the joint estimation, the OSP-based estimation and the differential estimation by investigating their corresponding BLUEs, where for the first time a general differential approach is introduced coping with multiple nuisance parameters.

3.2.1. JOINT ESTIMATION

THE joint least squares (JLS) estimate of \mathbf{x} and \mathbf{u} , based on the model (3.1), is given by

$$\begin{bmatrix} \hat{\mathbf{x}}_{\text{jls}} \\ \hat{\mathbf{u}}_{\text{jls}} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{H}^T \\ \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{H} & \mathbf{G} \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{H}^T \\ \mathbf{G}^T \end{bmatrix} \mathbf{y}, \quad (3.2)$$

where we have used the fact that the augmented matrix $[\mathbf{H} \ \mathbf{G}]$ has full column rank. Obviously, $\hat{\mathbf{x}}_{\text{js}}$ is the BLUE, since \mathbf{n} is zero-mean white noise, according to the Gauss-Markov Theorem [1].

3.2.2. OSP-BASED ESTIMATION

If we prefer to project out the nuisance term \mathbf{u} , an orthogonal subspace projector can be formulated [20] as

$$\mathbf{P}_{\mathbf{G}}^{\perp} \triangleq \mathbf{I}_N - \mathbf{G}\mathbf{G}^{\dagger}, \quad (3.3)$$

where $[\cdot]^{\dagger}$ indicates the pseudo-inverse which is given by $\mathbf{G}^{\dagger} \triangleq (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T$, since \mathbf{G} is assumed to have full column rank. Applying $\mathbf{P}_{\mathbf{G}}^{\perp}$ to our original model in (3.1) results in a new model

$$\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{y} = \mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{H}\mathbf{x} + \mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{n}, \quad (3.4)$$

where the impact of the nuisance term \mathbf{u} is eliminated. Due to the symmetry and the idempotence of an orthogonal subspace projector, i.e., $\mathbf{P}_{\mathbf{G}}^{\perp} = \mathbf{P}_{\mathbf{G}}^{\perp T}$ and $\mathbf{P}_{\mathbf{G}}^{\perp} = \mathbf{P}_{\mathbf{G}}^{\perp 2}$, we obtain the covariance matrix of the model noise in (3.4) as $\Sigma_{\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{n}} = \sigma^2\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{P}_{\mathbf{G}}^{\perp T} = \sigma^2\mathbf{P}_{\mathbf{G}}^{\perp}$. Then, following the OSP-based model (3.4), the corresponding LS optimization problem can be formulated as

$$\min_{\mathbf{x}} \|\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{y} - \mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{H}\mathbf{x}\|_2^2, \quad (3.5)$$

which leads to the following OSP-based LS estimate Type I of \mathbf{x}

$$\begin{aligned} \hat{\mathbf{x}}_{\text{osp-1}} &= (\mathbf{H}^T\mathbf{P}_{\mathbf{G}}^{\perp T}\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{P}_{\mathbf{G}}^{\perp T}\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{y} \\ &= (\mathbf{H}^T\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{y}. \end{aligned} \quad (3.6)$$

However, the model noise $\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{n}$ in (3.4) is not white, i.e., $\Sigma_{\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{n}}$ is not a (scaled) identity. Moreover, the orthogonal subspace projector $\mathbf{P}_{\mathbf{G}}^{\perp}$ is obviously singular, which implies that the covariance matrix $\Sigma_{\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{n}}$ is not invertible and hence can not be used to whiten the model (3.4). Therefore, it is very difficult to decide at this point whether $\hat{\mathbf{x}}_{\text{osp-1}}$ is the BLUE or not.

To cope with that, we need to introduce another type of OSP-based LS estimator for \mathbf{x} . If this estimator can be shown to be the BLUE and can also be proven equivalent to $\hat{\mathbf{x}}_{\text{osp-1}}$, then we can conclude that both of them are the BLUE.

Assume that $\mathbf{U}_n \in \mathbb{R}^{N \times (N-M)}$ contains orthonormal basis vectors spanning the null space of \mathbf{G} . Then, the idea of this second OSP-based estimator is to adopt the null space of \mathbf{G} to remove the impact of \mathbf{u} . More specifically, pre-multiplying \mathbf{U}_n^T on both sides of our original model leads to

$$\mathbf{U}_n^T\mathbf{y} = \mathbf{U}_n^T\mathbf{H}\mathbf{x} + \mathbf{U}_n^T\mathbf{n}. \quad (3.7)$$

Note that (3.4) can be obtained from (3.7) by multiplying it on both sides with \mathbf{U}_n since $\mathbf{U}_n\mathbf{U}_n^T = \mathbf{P}_{\mathbf{G}}^{\perp}$ [32], and hence these two models are basically equivalent. We can also see that, since \mathbf{U}_n is an isometry, the model noise $\mathbf{U}_n^T\mathbf{n}$ remains white, i.e., the covariance matrix of $\mathbf{U}_n^T\mathbf{n}$ is $\Sigma_{\mathbf{U}_n^T\mathbf{n}} = \sigma^2\mathbf{U}_n^T\mathbf{U}_n = \sigma^2\mathbf{I}_{N-M}$, which means that the LS estimate of this model is the BLUE.

Applying the LS criterion to the model (3.7) results in the optimization problem

$$\min_{\mathbf{x}} \|\mathbf{U}_n^T \mathbf{y} - \mathbf{U}_n^T \mathbf{H} \mathbf{x}\|_2^2, \quad (3.8)$$

from which we can obtain the OSP-based LS estimate type II of \mathbf{x} as

$$\hat{\mathbf{x}}_{\text{osp-2}} = (\mathbf{H}^T \mathbf{U}_n \mathbf{U}_n^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{U}_n \mathbf{U}_n^T \mathbf{y}. \quad (3.9)$$

Due to the fact that $\mathbf{U}_n \mathbf{U}_n^T = \mathbf{P}_G^\perp$, we obtain the equivalence $\hat{\mathbf{x}}_{\text{osp-1}} \equiv \hat{\mathbf{x}}_{\text{osp-2}}$ and hence both estimators represent the BLUE. In the later simulations, these two OSP-based BLUEs will be considered together for convenience.

Finally, to end this subsection, we would like to focus on the equivalence between the joint estimation and the OSP-based estimation approaches. In fact, the equivalence between $\hat{\mathbf{x}}_{\text{jls}}$ and $\hat{\mathbf{x}}_{\text{osp-1}}$ is already known [21, 22, 33], but we found it useful to revisit this result from a different viewpoint. To be explicit, applying the block-wise inversion to (3.2), we can easily rewrite the joint LS estimate of \mathbf{x} and \mathbf{u} as

$$\begin{aligned} \begin{bmatrix} \hat{\mathbf{x}}_{\text{jls}} \\ \hat{\mathbf{u}}_{\text{jls}} \end{bmatrix} &= \begin{bmatrix} \mathbf{M}_G & -\mathbf{M}_G \mathbf{H}^T (\mathbf{G}^\dagger)^T \\ -\mathbf{M}_H \mathbf{G}^T (\mathbf{H}^\dagger)^T & \mathbf{M}_H \end{bmatrix} \begin{bmatrix} \mathbf{H}^T \\ \mathbf{G}^T \end{bmatrix} \mathbf{y}, \\ &= \begin{bmatrix} \mathbf{M}_G \mathbf{H}^T - \mathbf{M}_G \mathbf{H}^T (\mathbf{G}^\dagger)^T \mathbf{G}^T \\ \mathbf{M}_H \mathbf{G}^T - \mathbf{M}_H \mathbf{G}^T (\mathbf{H}^\dagger)^T \mathbf{H}^T \end{bmatrix} \mathbf{y}, \\ &= \begin{bmatrix} \mathbf{M}_G \mathbf{H}^T \mathbf{P}_G^\perp \\ \mathbf{M}_H \mathbf{G}^T \mathbf{P}_H^\perp \end{bmatrix} \mathbf{y}, \end{aligned} \quad (3.10)$$

where $\mathbf{M}_G \triangleq (\mathbf{H}^T \mathbf{P}_G^\perp \mathbf{H})^{-1}$ and $\mathbf{M}_H \triangleq (\mathbf{G}^T \mathbf{P}_H^\perp \mathbf{G})^{-1}$ with $\mathbf{P}_H^\perp \triangleq \mathbf{I} - \mathbf{H} \mathbf{H}^\dagger$. From (3.10), we can directly observe that $\hat{\mathbf{x}}_{\text{jls}} = \mathbf{M}_G \mathbf{H}^T \mathbf{P}_G^\perp \mathbf{y}$ and hence

$$\hat{\mathbf{x}}_{\text{jls}} \equiv \hat{\mathbf{x}}_{\text{osp-1}} \equiv \hat{\mathbf{x}}_{\text{osp-2}},$$

where the equivalence between $\hat{\mathbf{x}}_{\text{jls}}$ and $\hat{\mathbf{x}}_{\text{osp-2}}$ is an interesting observation that has never been directly reported before, to the best of our knowledge.

3.2.3. DIFFERENTIAL SIGNAL PROCESSING

IN this subsection, we would like to examine differential approaches. This method firstly selects a reference and then removes the impact of \mathbf{u} by taking differences between the observations and the reference. To be specific, if the j -th observation y_j is selected as the reference, a new differential observation set can be constructed as

$$\mathbf{d}_j \triangleq \begin{bmatrix} \vdots \\ y_i - y_j \\ \vdots \end{bmatrix}_{(N-1) \times 1} = \mathbf{\Gamma}_j \mathbf{y}, \quad i \neq j, \quad (3.11)$$

where

$$\mathbf{\Gamma}_j \triangleq \begin{bmatrix} \mathbf{I}_{j-1} & -\mathbf{1}_{(j-1) \times 1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_{(N-j) \times 1} & \mathbf{I}_{N-j} \end{bmatrix}_{(N-1) \times N} \quad (3.12)$$

with $\mathbf{1}$ the all-one matrix (sizes are mentioned in subscript if needed) and the size of the observation set is reduced to $N - 1$ since j is fixed for every element in \mathbf{d}_j . This type of observation set is very popular and has wide applications in source localization and many other areas. Clearly, it can only be used to remove a single nuisance parameter in case $\mathbf{G} = \mathbf{1}_{N \times 1}$.

One may also suggest to select the average of the observations as the reference [17, eq. (28)], thus leading to another kind of differential observation set, given by

$$\mathbf{d}_{\text{avg}} \triangleq \begin{bmatrix} \vdots \\ y_i - \bar{y} \\ \vdots \end{bmatrix}_{N \times 1} = \mathbf{P}_{\mathbf{1}_{N \times 1}}^\perp \mathbf{y} \quad (3.13)$$

where $\mathbf{P}_{\mathbf{1}_{N \times 1}}^\perp \triangleq \mathbf{I} - \mathbf{1}_{N \times 1} \mathbf{1}_{N \times 1}^\dagger = \mathbf{I}_N - \frac{1}{N} \mathbf{1}_{N \times N}$. Sometimes, the use of this type of observation set to eliminate the nuisance parameters can be implicit[4], i.e., taking the average of the observations is not clearly pointed out. However, this case can obviously be linked to the OSP-based estimation with a single nuisance parameter in case $\mathbf{G} = \mathbf{1}_{N \times 1}$. Therefore, we are more interested in the simple differencing process of (3.11), where the reference index j seems to play a significant role.

As already pointed out, (3.11) only eliminates one nuisance parameter. Nevertheless, we would like to extend this to tackle multiple nuisance parameters, i.e., we would like to relax the constraint $\mathbf{G} = \mathbf{1}_{N \times 1}$ to $\text{rank}(\mathbf{G}) = M \geq 1$. The idea we will adopt here is based on eliminating the impact of the nuisance parameters one by one, which requires M differencing steps.

To achieve that, we write $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_M]$ with \mathbf{g}_k the k -th column vector of \mathbf{G} related to the k -th nuisance parameter u_k ($1 \leq k \leq M$). Thus, our original model in (3.1) can be rewritten as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \underbrace{\mathbf{g}_1 u_1 + \dots + \mathbf{g}_M u_M}_{M \text{ nuisance parameters}} + \mathbf{n}. \quad (3.14)$$

We then eliminate the nuisance parameters recursively in the order of u_1, \dots, u_M , although the explicit ordering is not important. At the k -th iteration, when $k - 1$ nuisance parameters have already been canceled, the observation vector containing the remaining nuisance parameters can be written as

$$\mathbf{d}^{(k-1)} = \mathbf{H}^{(k-1)} \mathbf{x} + \underbrace{\mathbf{g}_k^{(k-1)} u_k + \dots + \mathbf{g}_M^{(k-1)} u_M}_{M-k+1 \text{ nuisance parameters}} + \mathbf{n}^{(k-1)}, \quad (3.15)$$

where the superscript $(\cdot)^{(k-1)}$ indicates the variables after $k - 1$ differencing steps, $\mathbf{y}^{(k-1)}$, $\mathbf{g}_k^{(k-1)}, \dots, \mathbf{g}_M^{(k-1)}, \mathbf{n}^{(k-1)} \in \mathbb{R}^{(N-k+1) \times 1}$ and $\mathbf{H}^{(k-1)} \in \mathbb{R}^{(N-k+1) \times L}$. We also assume that, for $k = 1$, $\mathbf{d}^{(0)} = \mathbf{y}$ and similarly $\mathbf{H}^{(0)} = \mathbf{H}, \mathbf{g}_k^{(0)} = \mathbf{g}_k$ and $\mathbf{n}^{(0)} = \mathbf{n}$.

To cancel u_k , we first notice that some elements of $\mathbf{g}_k^{(k-1)}$ might be zero, i.e., u_k yields no impact on the corresponding observations in $\mathbf{d}^{(k-1)}$ and hence these observations should not be involved in the differencing process at this iteration. Without loss of generality, we assume that the first K elements of $\mathbf{g}_k^{(k-1)}$ are zero, where $1 \leq K \leq N - k - 1$

(there should be at least 2 non-zero elements for executing the differencing process). Then, among the remaining observations impacted by u_k , we select the j -th element as the reference, $K+1 \leq j \leq N-k+1$, and perform the following differencing step

$$\mathbf{d}^{(k)} = \begin{bmatrix} [\mathbf{d}^{(k-1)}]_1 \\ \vdots \\ [\mathbf{d}^{(k-1)}]_K \\ \frac{[\mathbf{d}^{(k-1)}]_{K+1}}{[\mathbf{g}_k^{(k-1)}]_{K+1}} - \frac{[\mathbf{d}^{(k-1)}]_j}{[\mathbf{g}_k^{(k-1)}]_j} \\ \vdots \\ \frac{[\mathbf{d}^{(k-1)}]_{j-1}}{[\mathbf{g}_k^{(k-1)}]_{j-1}} - \frac{[\mathbf{d}^{(k-1)}]_j}{[\mathbf{g}_k^{(k-1)}]_j} \\ \frac{[\mathbf{d}^{(k-1)}]_{j+1}}{[\mathbf{g}_k^{(k-1)}]_{j+1}} - \frac{[\mathbf{d}^{(k-1)}]_j}{[\mathbf{g}_k^{(k-1)}]_j} \\ \vdots \\ \frac{[\mathbf{d}^{(k-1)}]_{N-k+1}}{[\mathbf{g}_k^{(k-1)}]_{N-k+1}} - \frac{[\mathbf{d}^{(k-1)}]_j}{[\mathbf{g}_k^{(k-1)}]_j} \end{bmatrix}_{(N-k) \times 1} = \mathbf{\Gamma}^{(k)} \mathbf{d}^{(k-1)}, \quad (3.16)$$

where $\mathbf{\Gamma}^{(k)} \triangleq \begin{bmatrix} \mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_\perp^{(k)} \text{diag}\left(\frac{1}{[\mathbf{g}_k^{(k-1)}]_{K+1}}, \dots, \frac{1}{[\mathbf{g}_k^{(k-1)}]_{N-k+1}}\right)^T \end{bmatrix}$ is the $(N-k) \times (N-k+1)$ differencing operator for $\mathbf{d}^{(k-1)}$ with

$$\mathbf{\Gamma}_\perp^{(k)} \triangleq \begin{bmatrix} \mathbf{I}_{j-K-1} & -\mathbf{1}_{(j-K-1) \times 1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_{(N-k-j+1) \times 1} & \mathbf{I}_{N-k-j+1} \end{bmatrix}, \quad (3.17)$$

and obviously $\mathbf{\Gamma}^{(k)} \mathbf{g}_k^{(k-1)} = \mathbf{0}$. Accordingly, the new differential observation vector $\mathbf{d}^{(k)}$ can be formulated as

$$\mathbf{d}^{(k)} = \underbrace{\mathbf{\Gamma}^{(k)} \mathbf{H}^{(k-1)}}_{\mathbf{H}^{(k)}} \mathbf{x} + \underbrace{\mathbf{\Gamma}^{(k)} \mathbf{g}_{k+1}^{(k-1)} u_{k+1} + \dots + \mathbf{\Gamma}^{(k)} \mathbf{g}_M^{(k-1)} u_M}_{\substack{M-k \text{ nuisance parameters} \\ \mathbf{g}_{k+1}^{(k)} \quad \mathbf{g}_M^{(k)}}} + \underbrace{\mathbf{\Gamma}^{(k)} \mathbf{n}^{(k-1)}}_{\mathbf{n}^{(k)}}, \quad (3.18)$$

where u_k has been canceled.

We can see that (3.18) is similar to (3.15) with $k-1$ replaced by k . So it is clear that this recursive process can remove all nuisance parameters. Note that the number of zero values K as well as the reference index j could be different in every step, but for simplicity we use the same notation in every step.

To understand the interaction of the successive differencing steps, let us introduce the total differencing operator $\mathbf{\Gamma} = \mathbf{\Gamma}^{(M)} \dots \mathbf{\Gamma}^{(1)}$, where obviously

$$\text{rank}(\mathbf{\Gamma}^{(k)} \mathbf{\Gamma}^{(k-1)}) = \text{rank}(\mathbf{\Gamma}^{(k)}) = N - k$$

and hence $\mathbf{\Gamma}$ has full row rank. Since it is clear that $\mathbf{\Gamma} \mathbf{G} = \mathbf{0}$, the final differential observation vector $\mathbf{d}^{(M)}$ can be expressed as

$$\mathbf{d}^{(M)} = \mathbf{\Gamma} \mathbf{y} = \mathbf{\Gamma} \mathbf{H} \mathbf{x} + \mathbf{\Gamma} \mathbf{n}, \quad (3.19)$$

where the covariance matrix of $\Gamma \mathbf{n}$ is $\Sigma_{\Gamma \mathbf{n}} = \sigma^2 \Gamma \Gamma^T$.

Observe that the model noise has become correlated ever since the first step of the differencing process. Therefore, we need to whiten the model in (3.19) as

$$\begin{aligned} \Sigma_{\Gamma \mathbf{n}}^{-1/2} \mathbf{d}^{(M)} &= \Sigma_{\Gamma \mathbf{n}}^{-1/2} \Gamma \mathbf{H} \mathbf{x} + \Sigma_{\Gamma \mathbf{n}}^{-1/2} \Gamma \mathbf{n}, \\ \Rightarrow (\Gamma \Gamma^T)^{-1/2} \mathbf{d}^{(M)} &= (\Gamma \Gamma^T)^{-1/2} \Gamma \mathbf{H} \mathbf{x} + (\Gamma \Gamma^T)^{-1/2} \Gamma \mathbf{n}, \\ &\Rightarrow \mathbf{P} \mathbf{y} = \mathbf{P} \mathbf{H} \mathbf{x} + \mathbf{P} \mathbf{n}, \end{aligned} \quad (3.20)$$

where the unknown σ^2 is cancelled out at both sides of the equation and $\mathbf{P} \triangleq (\Gamma \Gamma^T)^{-1/2} \Gamma$ which exists since Γ has full row rank. Note that \mathbf{P} , as well as Γ and $\mathbf{d}^{(k)}$, depend on the reference indices j that have been chosen in the successive differencing steps, although this has not been explicitly stated.

Applying the LS criterion, the corresponding optimization problem is now obtained as

$$\min_{\mathbf{x}} \|\mathbf{P} \mathbf{y} - \mathbf{P} \mathbf{H} \mathbf{x}\|_2^2, \quad (3.21)$$

which leads to the following BLUE for model (3.19)

$$\hat{\mathbf{x}}_{\mathbf{d}} = (\mathbf{H}^T \mathbf{P}^T \mathbf{P} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{P}^T \mathbf{P} \mathbf{y}. \quad (3.22)$$

Finally, to prove the equivalence of the estimate $\hat{\mathbf{x}}_{\mathbf{d}}$ to the previous estimates, i.e., to prove that

$$\hat{\mathbf{x}}_{\text{jls}} \equiv \hat{\mathbf{x}}_{\text{osp-1}} \equiv \hat{\mathbf{x}}_{\text{osp-2}} \equiv \hat{\mathbf{x}}_{\mathbf{d}},$$

we need to establish the relation $\mathbf{P}^T \mathbf{P} = \mathbf{U}_n \mathbf{U}_n^T = \mathbf{P}_{\mathbf{G}}^{\perp}$. To do that, we first recall that $\Gamma \mathbf{G} = \mathbf{0}$ and that Γ has full row rank. Hence, Γ can always be written as $\Gamma = \mathbf{Q} \mathbf{U}_n^T$, where \mathbf{Q} is an $(N - M) \times (N - M)$ invertible matrix and \mathbf{U}_n has already been defined before as a basis that spans the null space of \mathbf{G} . The proof is completed by computing

$$\begin{aligned} \mathbf{P}^T \mathbf{P} &= \Gamma^T (\Gamma \Gamma^T)^{-1} \Gamma \\ &= \mathbf{U}_n \mathbf{Q}^T (\mathbf{Q} \mathbf{U}_n^T \mathbf{U}_n \mathbf{Q}^T)^{-1} \mathbf{Q} \mathbf{U}_n^T \\ &= \mathbf{U}_n \mathbf{Q}^T (\mathbf{Q}^T)^{-1} (\mathbf{U}_n^T \mathbf{U}_n)^{-1} \mathbf{Q}^{-1} \mathbf{Q} \mathbf{U}_n^T \\ &= \mathbf{U}_n \mathbf{U}_n^T = \mathbf{P}_{\mathbf{G}}^{\perp}, \end{aligned} \quad (3.23)$$

where we surprisingly notice that, even though \mathbf{P} and Γ are subject to possibly different reference indices j , there is no trace of any j in $\mathbf{P}^T \mathbf{P}$ and hence in $\hat{\mathbf{x}}_{\mathbf{d}}$.

A Simple Illustrative Case: We would like to demonstrate these three different methods, particularly the differential signal processing, with a simple example. Given $N = 3$ samples, we only assume a single parameter of interest ($L = 1$), but with two linear nuisance parameters ($M = 2$). We also know that $\mathbf{H} = [3 \ 6 \ 7]^T$ and $\mathbf{G} = \begin{bmatrix} 3 & 5 & 2 \\ 2 & 4 & 8 \end{bmatrix}^T$

and hence the joint estimator in (3.2) results into $\begin{bmatrix} \hat{\mathbf{x}}_{\text{jls}} \\ \hat{\mathbf{u}}_{\text{jls}} \end{bmatrix} = \begin{bmatrix} -3.2 & 2 & -0.2 \\ 2 & -1 & 0 \\ 2.3 & -1.5 & 0.3 \end{bmatrix} \mathbf{y}$, where

Table 3.1: Relations between the BLUEs related to the joint estimation, the OSP-based estimation and the differential estimation.

Models	BLUEs	Equality Conditions
Original Model in (3.1)	Joint Estimator in (3.2)	$[\mathbf{I}_L \mathbf{0}_{L \times M}]^a, \mathbf{P}_G^\perp \triangleq \mathbf{I}_N - \mathbf{G}\mathbf{G}^\dagger$
OSP Model Type I in (3.4)	OSP Estimator Type I in (3.6)	$\mathbf{P}_G^{\perp T} \mathbf{P}_G^\perp = \mathbf{P}_G^\perp$
OSP Model Type II in (3.7)	OSP Estimator Type II in (3.9)	$\mathbf{U}_n \mathbf{U}_n^T = \mathbf{P}_G^\perp$
Differential Model in (3.19) or the Whitenen One in (3.20)	Differential Estimator in (3.22)	$\mathbf{P}^T \mathbf{P} = \mathbf{P}_G^\perp$

^a $[\mathbf{I}_L \mathbf{0}_{L \times M}]$ is used for extracting $\hat{\mathbf{x}}_{jls}$ in (3.22).

the parameter estimate of interest is given by $\hat{\mathbf{x}}_{jls} = [-3.2 \quad 2 \quad -0.2] \mathbf{y}$. Then, we calculate $\mathbf{P}_G^\perp = \begin{bmatrix} 0.7171 & -0.4482 & 0.0448 \\ -0.4482 & 0.2801 & -0.0280 \\ 0.0448 & -0.0280 & 0.0028 \end{bmatrix}$ and $\mathbf{U}_n = [-0.8468 \quad 0.5293 \quad -0.0529]^T$ such that two OSP-based estimators in (3.6) and (3.9) can easily be carried out and proved to be equal to $\hat{\mathbf{x}}_{jls}$. We will not present more details for simplicity, but particularly focus on the differential method. Since there exist two linear nuisance parameters, it would take two steps for eliminating all of them:

1. In the first step ($k = 1$), we arbitrarily select the third element of \mathbf{y} as the reference ($j = 3$). Splitting \mathbf{G} by columns, we have $\mathbf{g}_1^{(0)} = [3 \quad 5 \quad 2]^T$, $\mathbf{g}_2^{(0)} = [2 \quad 4 \quad 8]^T$. According to (3.16), the new differential observation vector can be obtained as $\mathbf{d}^{(1)} = [y_1/3 - y_3/2 \quad y_2/5 - y_3/2]^T = \mathbf{\Gamma}^{(1)} \mathbf{y}$, where $\mathbf{\Gamma}^{(1)} = \begin{bmatrix} 1/3 & 0 & -1/2 \\ 0 & 1/5 & -1/2 \end{bmatrix}$. We can observe from $\mathbf{\Gamma}^{(1)} \mathbf{G} = \begin{bmatrix} 0 & -10/3 \\ 0 & -16/5 \end{bmatrix}$ that the impact of the first nuisance parameter u_1 is already eliminated. Also $\mathbf{g}_2^{(1)} = \mathbf{\Gamma}^{(1)} \mathbf{g}_2^{(0)}$ corresponds to the last column and the next nuisance parameter u_2 .
2. In the second step ($k = 2$), the first element of $\mathbf{d}^{(1)}$ is selected as the reference ($j = 1$). The differential observation becomes a scalar as $\mathbf{d}^{(2)} = -\frac{5}{16}(y_2/5 - y_3/2) + \frac{3}{10}(y_1/3 - y_3/2) = \mathbf{\Gamma}^{(2)} \mathbf{d}^{(1)} = \mathbf{\Gamma} \mathbf{y}$, where $\mathbf{\Gamma}^{(2)} = [3/10 \quad -5/16]$ and $\mathbf{\Gamma} = \mathbf{\Gamma}^{(2)} \mathbf{\Gamma}^{(1)} = [1/10 \quad -1/16 \quad 1/160]$. Now, we can readily observe that all the nuisance parameters are eliminated, since $\mathbf{\Gamma} \mathbf{G} = \mathbf{0}$.

With a known $\mathbf{\Gamma}$, we can easily whiten the model in (3.19) and obtain the differential estimator in (3.22). Moreover, the equivalence of the differential estimation can also be proved by observing $\mathbf{P}^T \mathbf{P} = \mathbf{\Gamma}^T (\mathbf{\Gamma} \mathbf{\Gamma}^T)^{-1} \mathbf{\Gamma} = \mathbf{P}_G^\perp$.

3.2.4. DISCUSSION

WE have studied estimation problems in the presence of deterministic linear nuisance parameters based on a general model. Therefore, all the conclusions drawn in this paper are applicable to any optimization problem with a data model that matches our general model (3.1). The equivalences between the BLUEs of the joint estimation, the OSP-based estimation and the differential estimation are summarized in Table. 3.1

and also in Fig. 3.1. Some interesting observations from these equivalences are listed below:

1. The joint estimation has to estimate both \mathbf{x} and the nuisance term \mathbf{u} while the other two estimation approaches remove the impact of \mathbf{u} before estimating \mathbf{x} .
2. For the OSP-based estimation, in order to remove the impact of \mathbf{u} , using \mathbf{P}_G^\perp actually colors the noise, but using \mathbf{U}_n^T keeps the model noise white. Interestingly though, the corresponding LS estimates for those two models are theoretically equivalent and hence they are both the BLUE.
3. In many applications, the differential processing is commonly considered as a separate and independent approach. But, in this paper, we have generally proven its equivalence to the joint estimation and the OSP-based estimation. The differential approach removes the impact of the nuisance parameters by taking differences between the reference and the observations. If one of the observations is selected as a reference, the obtained differential observation set has to be properly whitened in order to obtain the BLUE for this model.
4. From an information theoretic perspective, the joint estimation, which directly utilizes the observations \mathbf{y} , preserves the full data information, and any preprocessing on the observations might cause an information loss. However, in this paper, all the other BLUEs have been proven to be equivalent to the BLUE of the joint estimation, which implies that neither the OSP-based estimation nor the differential estimation suffers any information loss by removing the impact of the nuisance parameters.
5. It is also worth noting that, for the differential approach, selecting which observation will function as a reference is not important, since the reference index j yields no impact on the BLUE. This is in sharp contrast to what is commonly considered in literature.
6. One might notice that, in the differencing process, N observations can generate a maximum of $N(N-1)/2$ distinct observation differences. In contrast, we only study the estimation problem based on a subset, which is associated with a single reference and corresponds to $N-1$ observation differences. However, from the above conclusions, it is clear that the considered subset already preserves all the information (independent of the reference), which implies that the full set obtains no more information than any subset does. Also this is a novel observation.

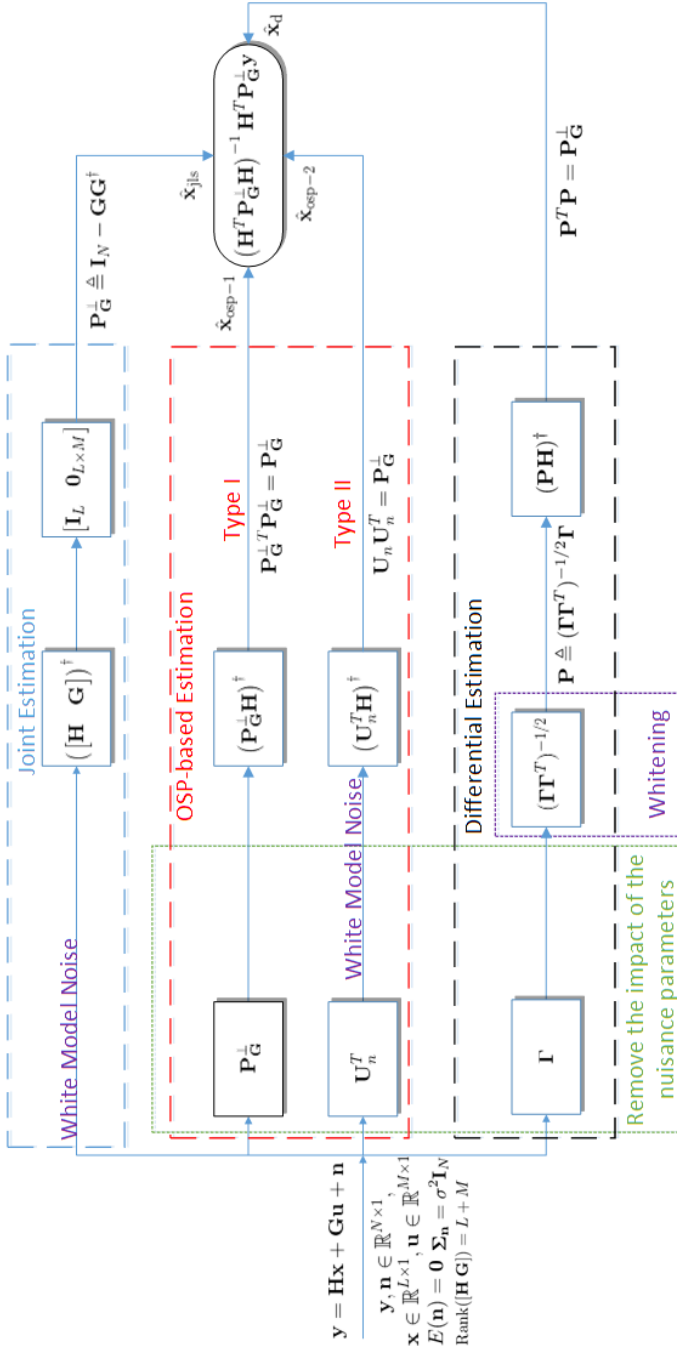


Figure 3.1: Diagram to illustrate the relations between the BLUEs related to the joint estimation, the OSP-based estimation and the differential estimation. Note that the noise \mathbf{n} is not necessarily Gaussian distributed and the operator $[\mathbf{I}_L \ \mathbf{0}_{L \times M}]$ is used to extract the first L elements of a vector, i.e., $\hat{\mathbf{x}}_{\text{js}}$.

3.3. LOCALIZATION EXAMPLES

By studying the relations between the BLUE of the joint estimation, the OSP-based estimation and the differential estimation, the essence of this paper is to provide some in-depth understanding of coping with unknown nuisance parameters. Some important underlying equivalences have been unveiled, especially the one related to the differential method, since, in many applications, this approach is still considered as a separate optimization problem. Owing to the generality of this paper, one may easily apply our analyses and conclusions to some particular applications, if the data model can be (re)formulated to match our general model (3.1). Some specific localization examples are detailed next.

3.3.1. TIME-BASED LOCALIZATION

BOTH TOA and TDOA based localization are called time-based localization[2], since they both rely on time measurements (either the global time or the local time). The essence of this kind of localization problem is how to accurately extract distance-related information (e.g., the time of flight (TOF)). Directly using TOA measurements requires not only perfect clock synchronization between the emitters and the receivers but also the knowledge of the transmitting time[34]. In cooperative networks, where clock synchronization is frequently carried out (because the inner clock might drift over time) and the transmitting times are also piggybacked with the transmitted signals, one can precisely calculate the TOFs from the TOA measurements and then localize the target node. However, it is often very expensive to meet those requirements and most networks are constrained by limited resources and capabilities. Therefore, in most cases, sensors suffer from two linear nuisance parameters, i.e., the unknown clock biases to the global time and the unknown transmitting times.

In this example, we assume N anchor nodes that are perfectly synchronized with the global time and there exists only a clock bias in the target node, which broadcasts beacon signals at unknown local transmit times. We denote $\mathbf{x}_t \in \mathbb{R}^d$ as the target location and $\mathbf{s}_i \in \mathbb{R}^d$ as the i -th anchor location. For convenience, a single unknown global transmit time t_0 is considered for the target node, instead of the local transmit time plus the clock bias. Taking the speed of light c into account, we obtain the TOA measurements as

$$\mathbf{d} = \mathbf{r}(\mathbf{x}_t) + \mathbf{1}_{N \times 1} r_o + \mathbf{n}, \quad (3.24)$$

where the element d_i of \mathbf{d} indicates the TOA measurement from the i -th anchor, $\mathbf{r}(\mathbf{x}_t)$ stacks $r_i \triangleq \|\mathbf{x}_t - \mathbf{s}_i\|_2$, $r_o \triangleq c t_0$ and \mathbf{n} is the vector of the measurement noise n_i with $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$. Note that, compared with more realistic scenarios, the model (3.24) is simplified for convenience, but still adequate to make our point.

TAYLOR SERIES EXPANSION

OBVIOUSLY, the non-linearity of (3.24) is a very serious issue for localization, other than the nuisance parameter. Many methods, especially those considering mobile scenarios, directly linearize (3.24) by a Taylor series expansion (TSE) [35]. Note that this kind of method is very similar to the Gauss-Newton (GN) method[36] and holds the maximum likelihood (ML) property. Since we can obtain the estimate of \mathbf{x}_t by iteratively updating the previous iteration, we first have to apply the TSE to (3.24) around the target

location estimate $\hat{\mathbf{x}}_t^{(k-1)}$ at the $(k-1)$ -th iteration, thus resulting into

$$\mathbf{d} = \mathbf{r}(\hat{\mathbf{x}}_t^{(k-1)}) + \left. \frac{\partial \mathbf{r}}{\partial \mathbf{x}_t^T} \right|_{\mathbf{x}_t = \hat{\mathbf{x}}_t^{(k-1)}} (\mathbf{x}_t - \hat{\mathbf{x}}_t^{(k-1)}) + \mathbf{1}_{N \times 1} r_o + \mathbf{n}.$$

Then, we rearrange the above equation and present the TSE model for iteration step k as

$$\begin{aligned} \mathbf{d} - \mathbf{r}(\hat{\mathbf{x}}_t^{(k-1)}) + \left. \frac{\partial \mathbf{r}}{\partial \mathbf{x}_t^T} \right|_{\mathbf{x}_t = \hat{\mathbf{x}}_t^{(k-1)}} \hat{\mathbf{x}}_t^{(k-1)} &= \left. \frac{\partial \mathbf{r}}{\partial \mathbf{x}_t^T} \right|_{\mathbf{x}_t = \hat{\mathbf{x}}_t^{(k-1)}} \mathbf{x}_t + \mathbf{1}_{N \times 1} r_o + \mathbf{n} \\ \Rightarrow \boldsymbol{\delta}^{(k-1)} = \boldsymbol{\Delta}^{(k-1)} \mathbf{x}_t + \mathbf{1}_{N \times 1} r_o + \mathbf{n}, \end{aligned} \quad (3.25)$$

where $\boldsymbol{\delta}^{(k-1)} \triangleq \mathbf{d} - \mathbf{r}(\hat{\mathbf{x}}_t^{(k-1)}) + \boldsymbol{\Delta}^{(k-1)} \hat{\mathbf{x}}_t^{(k-1)}$, and

$$\boldsymbol{\Delta}^{(k-1)} \triangleq \left. \frac{\partial \mathbf{r}}{\partial \mathbf{x}_t^T} \right|_{\mathbf{x}_t = \hat{\mathbf{x}}_t^{(k-1)}} = \left[\dots, \frac{(\hat{\mathbf{x}}_t^{(k-1)} - \mathbf{s}_i)^T}{\|\hat{\mathbf{x}}_t^{(k-1)} - \mathbf{s}_i\|_2}, \dots \right]^T.$$

The localization problem at the k -th iteration boils down to estimating \mathbf{x}_t from (3.25) to update the location estimate from the $(k-1)$ -th iteration. The relation between the TSE model and the general model (3.1) is presented in Table 3.2 on page 98. Note that since the discussed approaches can directly be applied to the TOA measurements with a single nuisance parameter ($M = 1$), the differential approach applied to the TOA measurements actually corresponds to working with the TDOA measurements, i.e.,

$$\begin{bmatrix} \vdots \\ d_{i,j} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ d_i - d_j \\ \vdots \end{bmatrix}_{(N-1) \times 1}, \quad i \neq j. \quad (3.26)$$

However, to avoid any confusion with the TDOA methods we will discuss later on, we will refer to this method as the differential approach applied to the TSE model of the TOA measurements.

SQUARED DISTANCE

THE TSE method highly relies on an appropriate initialization that is near the global solution, otherwise it might converge to a local minimum. Thus, some closed-form solutions were proposed to solve this non-convex problem, which requires squaring the distance norm (SD) for linearization[37]. Unlike the TSE method, the SD method depends on the type of measurements, since different modeling steps are carried out for TOA and TDOA measurements.

TOA: Let us first focus on the SD method based on the TOA measurements which can be expressed as

$$d_i = \|\mathbf{x}_t - \mathbf{s}_i\|_2 + r_o + n_i. \quad (3.27)$$

Moving r_0 to the other side and squaring both sides of the equation, we obtain

$$\begin{aligned} (d_i - r_0)^2 &= (\|\mathbf{x}_t - \mathbf{s}_i\|_2 + n_i)^2 \\ \Rightarrow -2\mathbf{s}_i^T \mathbf{x}_t + \|\mathbf{x}_t\|_2^2 - r_0^2 - 2d_i r_0 &= d_i^2 - \|\mathbf{s}_i\|_2^2 - 2r_i n_i - n_i^2, \end{aligned} \quad (3.28)$$

where r_0^2 is viewed as a new nuisance parameter. As a result, a linear model with two nuisance parameters ($M = 2$) can be formulated as

$$\mathbf{z}_1 = \mathbf{A}_1 \boldsymbol{\theta}_1 + \boldsymbol{\epsilon}_1, \quad (3.29)$$

where $\mathbf{A}_1 \triangleq \begin{bmatrix} \vdots & \vdots & \vdots \\ -2\mathbf{s}_i^T & 1 & -2d_i \\ \vdots & \vdots & \vdots \end{bmatrix}$, $\boldsymbol{\theta}_1 \triangleq \begin{bmatrix} \mathbf{x}_t \\ \|\mathbf{x}_t\|_2^2 - r_0^2 \\ r_0 \end{bmatrix}$, $\mathbf{z}_1 \triangleq \begin{bmatrix} \vdots \\ d_i^2 - \|\mathbf{s}_i\|_2^2 \\ \vdots \end{bmatrix}$ and

$$\boldsymbol{\epsilon}_1 \triangleq \begin{bmatrix} \vdots \\ 2r_i n_i + n_i^2 \\ \vdots \end{bmatrix} \approx \begin{bmatrix} \vdots \\ 2r_i n_i \\ \vdots \end{bmatrix} = 2\mathbf{D}_1 \mathbf{n}. \quad (3.30a)$$

Here, we denote $\mathbf{D}_1 = \text{diag}([r_1, \dots, r_N]^T)$ with $\text{diag}(\cdot)$ as a diagonal matrix with its argument on the diagonal, and hence $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_1} = 4\sigma^2 \mathbf{D}_1^2$. This SD-TOA model is widely considered [38–42]. Some researchers apply the differencing process to remove the nuisance parameters [25, 34, 43–46] while some others use the OSP method [17, 47]. Note that the model noise in (3.30a) is still not white and hence an appropriate whitening procedure is required. Assuming \mathbf{D}_1 is perfectly known, we can whiten the model (3.29) as

$$\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_1}^{-1/2} \mathbf{z}_1 = \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_1}^{-1/2} \mathbf{A}_1 \boldsymbol{\theta}_1 + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_1}^{-1/2} \boldsymbol{\epsilon}_1 \quad (3.31a)$$

$$\Rightarrow \mathbf{D}'_1 \mathbf{z}_1 = \mathbf{D}'_1 \mathbf{A}_1 \boldsymbol{\theta}_1 + \mathbf{D}'_1 \boldsymbol{\epsilon}_1 \quad (3.31b)$$

where $\mathbf{D}'_1 \triangleq \mathbf{D}_1^{-1}$ and the covariance matrix of $\mathbf{D}'_1 \boldsymbol{\epsilon}_1$ is now a scaled identity, i.e., $\boldsymbol{\Sigma}_{\mathbf{D}'_1 \boldsymbol{\epsilon}_1} = 4\sigma^2 \mathbf{I}_N$. In practice, a LS estimate based on the model (3.29) can first be used to construct an estimate of \mathbf{D}_1 for carrying out the whitening. Then, the estimate of \mathbf{D}_1 can be repeatedly updated to approach the true \mathbf{D}_1 with a more accurate location estimate. In this paper though, we only want to evaluate its best performance and hence directly use the true \mathbf{D}_1 . Finally, expressing $\mathbf{A}_1 = [\mathbf{A}'_1, \mathbf{A}''_1]$ with \mathbf{A}'_1 and \mathbf{A}''_1 respectively containing the first d and the remaining columns, the relation between the whitened SD-TOA model and the general model (3.1) is presented in Table 3.2 on page 98.

TDOA: Directly applying the differencing process on the TOA observations \mathbf{d} removes the unknown nuisance parameter r_0 , resulting in the TDOA measurements

$$d_{i,j} = \|\mathbf{x}_t - \mathbf{s}_i\|_2 - \|\mathbf{x}_t - \mathbf{s}_j\|_2 + n_{i,j}, \quad i \neq j, \quad (3.32)$$

where $n_{i,j} = n_i - n_j$. Introducing $r_j = \|\mathbf{x}_t - \mathbf{s}_j\|_2$ as a new unknown parameter, we can linearize (1.2) using the following squaring operation

$$\begin{aligned} (d_{i,j} + r_j)^2 &= (\|\mathbf{x}_t - \mathbf{s}_i - (\mathbf{s}_i - \mathbf{s}_j)\|_2 + n_{i,j})^2 \\ \Rightarrow -2(\mathbf{s}_i - \mathbf{s}_j)^T \mathbf{x}_t - 2d_{i,j} r_j &= d_{i,j}^2 + \|\mathbf{s}_j\|_2^2 - \|\mathbf{s}_i\|_2^2 - 2r_i n_{i,j} - n_{i,j}^2. \end{aligned} \quad (3.33)$$

As a result, a linear model with a single unknown nuisance parameter r_j ($M = 1$) can be formulated as

$$\mathbf{z}_2 = \mathbf{A}_2 \boldsymbol{\theta}_2 + \boldsymbol{\epsilon}_2, \quad (3.34)$$

$$\text{where } \mathbf{A}_2 \triangleq -2 \begin{bmatrix} \vdots & \vdots \\ (\mathbf{s}_i - \mathbf{s}_j)^T & d_{i,j} \\ \vdots & \vdots \end{bmatrix}, \boldsymbol{\theta}_2 \triangleq \begin{bmatrix} \mathbf{x}_t \\ r_j \end{bmatrix}, \mathbf{z}_2 \triangleq \begin{bmatrix} \vdots \\ d_{i,j}^2 + \|\mathbf{s}_j\|_2^2 - \|\mathbf{s}_i\|_2^2 \\ \vdots \end{bmatrix} \text{ and}$$

$$\boldsymbol{\epsilon}_2 \triangleq \begin{bmatrix} \vdots \\ 2r_i n_{i,j} + n_{i,j}^2 \\ \vdots \end{bmatrix} \approx \begin{bmatrix} \vdots \\ 2r_i n_{i,j} \\ \vdots \end{bmatrix} = 2\mathbf{D}_2 \boldsymbol{\Gamma}_j \mathbf{n}. \quad (3.35a)$$

Here, we denote $\mathbf{D}_2 = \text{diag}([\dots, r_i, \dots]^T)$, $i \neq j$ and hence $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_2} = 4\sigma^2 \mathbf{D}_2 \boldsymbol{\Gamma}_j \boldsymbol{\Gamma}_j^T \mathbf{D}_2^T$. Also this SD-TDOA model has been commonly adopted in literature [15, 34, 48–52]. Among the TDOA localization techniques based on this model, the famous Chan algorithm [15], from which many others stem, is actually equivalent to some earlier works [53–55], where the unknown r_j is simply removed by the OSP method. Again, note that the model noise (3.35a) is not white. Assuming \mathbf{D}_2 is perfectly known (as already explained for \mathbf{D}_1 , in practice, \mathbf{D}_2 should be iteratively estimated), we can whiten the model (3.34) as

$$\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_2}^{-1/2} \mathbf{z}_2 = \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_2}^{-1/2} \mathbf{A}_2 \boldsymbol{\theta}_2 + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_2}^{-1/2} \boldsymbol{\epsilon}_2 \quad (3.36a)$$

$$\Rightarrow \mathbf{D}'_2 \mathbf{z}_2 = \mathbf{D}'_2 \mathbf{A}_2 \boldsymbol{\theta}_2 + \mathbf{D}'_2 \boldsymbol{\epsilon}_2, \quad (3.36b)$$

where $\mathbf{D}'_2 \triangleq (\mathbf{D}_2 \boldsymbol{\Gamma}_j \boldsymbol{\Gamma}_j^T \mathbf{D}_2^T)^{-1/2}$ and the covariance matrix of $\mathbf{D}'_2 \boldsymbol{\epsilon}_2$ is now a scaled identity, i.e., $\boldsymbol{\Sigma}_{\mathbf{D}'_2 \boldsymbol{\epsilon}_2} = 4\sigma^2 \mathbf{I}_{N-1}$. Finally, we split \mathbf{A}_2 into $\mathbf{A}_2 = [\mathbf{A}'_2, \mathbf{A}''_2]$ with \mathbf{A}'_2 and \mathbf{A}''_2 respectively containing the first d and the remaining columns. The relation between the whitened SD-TDOA model and the general model (3.1) is finally presented in Table 3.2 on page 98.

Numerical results: We have conducted a Monte Carlo simulation with 1000 trials to verify our conclusions, where the BLUEs of the joint estimation, the OSP-based estimation and the differential estimation are carried out for each one of the discussed time-based models. Some LS estimators without a proper whitening process are also presented for comparison. The acronyms of all estimators used in the simulations are summarized in Table 3.3 on page 98. We also calculate the Cramér-Rao lower bound (CRLB) with an unknown r_0 based on the original model (3.24) [1, Chapter 3], since the TSE, SD-TOA and SD-TDOA models all lose some information by ignoring some high-order terms. The root mean square error (RMSE) of the location estimate, which is defined as $\sqrt{E[(\hat{\mathbf{x}} - \mathbf{x})^2]}$ in general, is used as a performance measure in this paper. From the numerical results in Fig. 3.2, we can draw the following conclusions.

1. For each model, the corresponding BLUEs yield the same performance as expected.
2. Without a proper whitening, it can be observed that the performance of the LS estimators deteriorates. The D-LS-TSE-TOA, J-LS-SD-TOA and J-LS-SD-TDOA clearly perform worse than their corresponding BLUEs.

3. The TSE model ignores $\mathcal{O}((\mathbf{x}_t - \hat{\mathbf{x}}_t^{(k-1)})^2)$ and accordingly suffers some information loss in modeling. However, the information loss can be reduced with a more accurate $\hat{\mathbf{x}}_t^{(k-1)}$. Therefore, with more iterations, the BLUEs for the TSE model approach the CRLB, which is in fact the essence of the ML property.
4. The SD-TOA model ignores $n_i^2, \forall i$ while the SD-TDOA model ignores $n_{i,j}^2, \forall i, i \neq j$. Ignoring these terms will cause an increasing information loss as the measurement noise gets larger.
5. Even though the BLUEs of the SD-TOA model outperform those of the SD-TDOA model in our simulation, we still cannot decide at this point which model is the best. This is because an optimal localization problem for the SD models should also include any dependence between the (nuisance) parameters, e.g., between \mathbf{x}_t and $\|\mathbf{x}_t\|_2^2$, and between r_0 and r_0^2 in $\boldsymbol{\theta}_1$, or between \mathbf{x}_t and r_j in $\boldsymbol{\theta}_2$, which explains the huge gap between the CRLB and the BLUEs for the SD models. By contrast, the TSE model obviously does not have this kind of issue. Nevertheless, including these dependencies is beyond the scope of this paper and we will not further consider this.
6. In practice, both the TSE and SD methods require iterations to obtain an accurate location estimate. However, note that, even after several iterations, the estimators based on the SD models still need to cope with the above mentioned dependency issue. Therefore, in real-life, one often combines those two models, i.e., one uses the TSE model with the J-LS-SD-TDOA or the J-LS-SD-TOA as an initialization.
7. For the SD-TDOA model, ignoring the terms $n_{i,j}^2, \forall i, i \neq j$ implies that the information loss depends on the reference choice of the differencing process in (1.2). However, this is only because of the SD modeling thereafter, not because of the differencing process itself. Note that, for any other differencing process in this paper, the reference index is not important as long as the model is properly whitened.

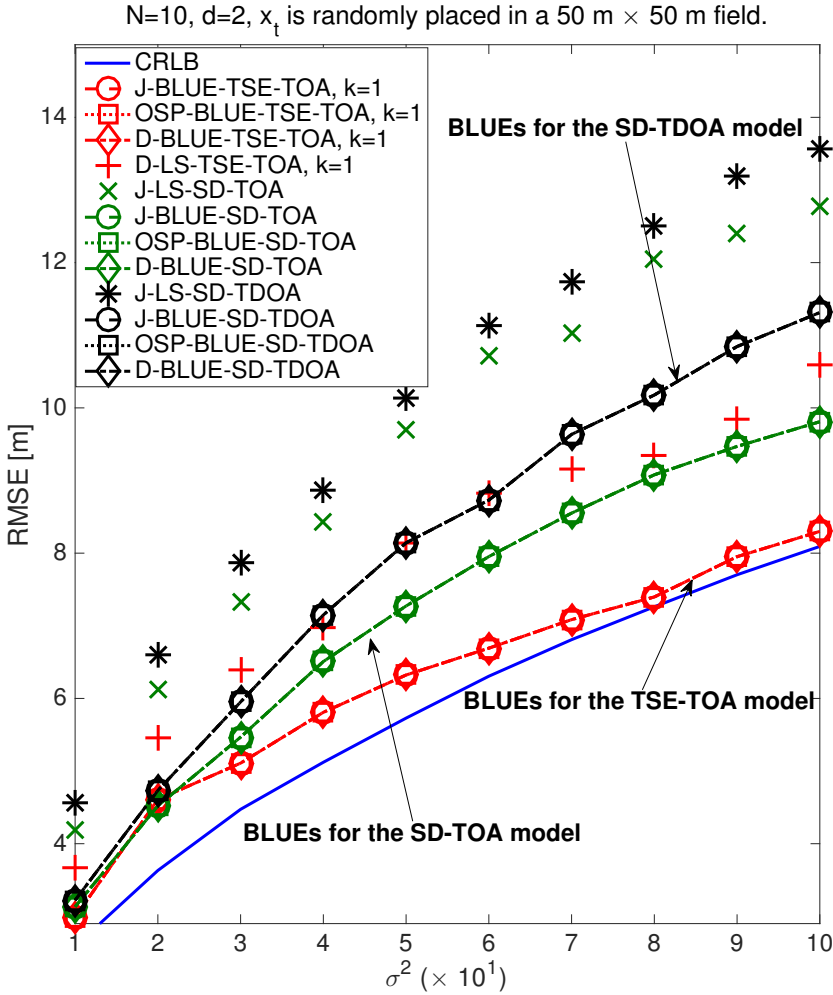


Figure 3.2: Performance of different time-based estimators: the target node is randomly placed in a 50 \times 50 field and 10 anchors are deployed with coordinates (50, 50), (50, 0), (0, 50), (0, 0), (25, 7), (25, 43), (12, 33), (12, 16), (37, 33) and (37, 16).

Table 3.2: Relations between the general model (3.1) and the considered time-based and RSS-based localization models^a.

General model (3.1)	\mathbf{y}	\mathbf{H}	\mathbf{x}	\mathbf{G}	\mathbf{u}
TSE model (3.25)	$\delta^{(k-1)}$	$\Delta^{(k-1)}$	\mathbf{x}_t	$\mathbf{I}_{N \times 1}$	r_0
SD-TOA model (3.31b)	$\mathbf{D}'_1 \mathbf{z}_1$	$\mathbf{D}'_1 \mathbf{A}'_1$	\mathbf{x}_t	$\mathbf{D}'_1 \mathbf{A}'_1$	$\ \mathbf{x}_t\ _2^2 - r_0^2, r_0\}^T$
SD-TDOA model (3.36b)	$\mathbf{D}'_2 \mathbf{z}_2$	$\mathbf{D}'_2 \mathbf{A}'_2$	\mathbf{x}_t	$\mathbf{D}'_2 \mathbf{A}'_2$	r_j
SD-RSS model (3.43b)	$\mathbf{D}\mathbf{h}$	\mathbf{F}	$\ \mathbf{x}_t\ _2^2, \mathbf{x}_t^T$	$\mathbf{I}_{N \times 1}$	p_0

^aAll the considered models must be white or whitened, i.e., the covariance of the model noise should be a (scaled) identity.

Table 3.3: Acronyms of the estimators used in the localization simulations.

Notations	Data Models	Estimation Methods
J-BLUE-TSE-TOA, $k = 1$	White TSE model (3.25) ^a , $M = 1$	Joint Estimation (3.2)
OSP-BLUE-TSE-TOA, $k = 1$	"	OSP-based Estimation (3.6) or (3.9)
D-BLUE-TSE-TOA ^b , $k = 1$	"	Differential Estimation (3.22)
D-LS-TSE-TOA ^b , $k = 1$	"	Differential Estimation (3.22)
J-LS-SD-TOA	Unwhitened SD-TOA model (3.29), $M = 2$	LS estimator based on the unwhitened differential observations in (3.19)
J-BLUE-SD-TOA	Whitened SD-TOA model (3.31b), $M = 2$	LS estimator with correlated model noise
OSP-BLUE-SD-TOA	"	OSP-based Estimation (3.6) or (3.9)
D-BLUE-SD-TOA	"	Differential Estimation (3.22)
J-LS-SD-TDOA	Unwhitened SD-TDOA model (3.34), $M = 1$	LS estimator with correlated model noise
J-BLUE-SD-TDOA	Whitened SD-TDOA model (3.36b), $M = 1$	Joint Estimation (3.2)
OSP-BLUE-SD-TDOA	"	OSP-based Estimation (3.6) or (3.9)
D-BLUE-SD-TDOA	"	Differential Estimation (3.22)
J-LS-SD-RSS	Unwhitened SD-RSS model (3.40), $M = 1$	LS estimator with correlated model noise
J-BLUE-SD-RSS	Whitened SD-RSS model (3.43b), $M = 1$	Joint Estimation (3.2)
OSP-BLUE-SD-RSS	"	OSP-based Estimation (3.6) or (3.9)
D-LS-SD-RSS ^c	"	Differential Estimation (3.22)
D-BLUE-SD-RSS ^c	"	Differential Estimation (3.22)

^aThe J-LS-SD-TDOA is used as an initial value (i.e., $k = 0$), which is guaranteed to be near the global solution.

^bD-BLUE-TSE-TOA and D-LS-TSE-TOA can equivalently be considered to work with the TDOA measurements.

^cD-LS-SD-RSS and D-BLUE-SD-RSS can equivalently be considered to work with the DRSS measurements.

3.3.2. RECEIVED SIGNAL STRENGTH BASED LOCALIZATION

DU^E to the simplicity of utilizing received signal strength (RSS) measurements, wireless networks with very constrained resources preferably rely on RSS-based localization [2]. Therefore, it gradually became very popular in recent years and many efforts have already been put on this topic[56–59].

RSS-based localization mainly suffers from the complicated radio propagation channel. As before, assume that the target node is located at \mathbf{x}_t and the i -th anchor at \mathbf{s}_i . Based on a large-scale log-normal fading model [60], the RSS measurement can then be modeled as

$$P_i = P_0 - 10\gamma \log_{10} \left(\frac{\|\mathbf{x}_t - \mathbf{s}_i\|_2}{d_0} \right) + n_i, \quad i = 1, 2, \dots, N, \quad (3.37)$$

where P_0 is the received power at the reference distance d_0 , γ is the path-loss exponent (PLE), $n_i \sim \mathcal{N}(0, \sigma^2)$ is the shadowing effect and N is the number of anchor nodes. RSS-based localization is aimed at estimating the target location \mathbf{x}_t from the RSS measurements. However, in some military or hostile scenarios, the transmit power might be unknown. Therefore, without loss of generality, we assume the reference distance d_0 to be 1 m and then the problem of the unknown transmit power can be equivalently converted into that of an unknown P_0 . Note that (3.37) also has the non-linearity issue and, obviously, the iterative TSE model for RSS-based localization will be very similar to that developed for time-based localization. Therefore, to save space, we do not consider directly applying the TSE model, but only focus on the SD method here.

To construct a linear data model, we rewrite (3.37) as

$$\|\mathbf{x}_t - \mathbf{s}_i\|_2^2 = \frac{P'_0 n'_i}{P'_i}, \quad (3.38)$$

where $P'_i \triangleq 10^{\frac{P_i}{5\gamma}}$, $P'_0 \triangleq 10^{\frac{P_0}{5\gamma}}$ and $n'_i \triangleq 10^{\frac{n_i}{5\gamma}}$. Interestingly though, we still need to apply the TSE to n'_i here², such that (3.38) can further be approximated as

$$\|\mathbf{x}_t\|_2^2 - 2\mathbf{s}_i^T \mathbf{x}_t + \|\mathbf{s}_i\|_2^2 = \frac{P'_0}{P'_i} \left[1 + \frac{\ln(10)}{5\gamma} n_i \right]. \quad (3.39)$$

Then, a linear SD-RSS model for localization can be formulated from (3.39) as

$$\mathbf{h} = \mathbf{F}\boldsymbol{\phi} + \boldsymbol{\zeta} \quad (3.40)$$

²We use $a^x = 1 + x \ln(a) + \dots + \frac{(x \ln(a))^n}{n!} + \dots$, $-\infty < x < \infty$ [61]. Note that the right hand side of (3.39) is an approximation, but it is regarded to be exact in this paper.

where

$$\mathbf{F} \triangleq \begin{bmatrix} \vdots & \vdots & \vdots \\ 2\mathbf{s}_i^T & -1 & 1/P'_i \\ \vdots & \vdots & \vdots \end{bmatrix}_{N \times (d+2)}, \quad (3.41a)$$

$$\boldsymbol{\phi} \triangleq \begin{bmatrix} \mathbf{x}_t \\ \|\mathbf{x}_t\|_2^2 \\ P'_0 \end{bmatrix}_{(d+2) \times 1}, \quad (3.41b)$$

$$\mathbf{h} \triangleq \begin{bmatrix} \vdots \\ \|\mathbf{s}_i\|_2^2 \\ \vdots \end{bmatrix}_{N \times 1}, \quad (3.41c)$$

$$\boldsymbol{\zeta} \triangleq \begin{bmatrix} \vdots \\ \frac{\ln(10)P'_0}{5\gamma P'_i} n_i \\ \vdots \end{bmatrix}_{N \times 1}. \quad (3.41d)$$

This model was firstly presented in [58, eq. (18)], but in the absence of the shadowing effect. If we whiten the model (3.40) utilizing the covariance matrix of $\boldsymbol{\zeta}$, i.e.,

$$\boldsymbol{\Sigma}_{\boldsymbol{\zeta}} = \frac{[\ln(10)]^2 P_0'^2 \sigma^2}{25\gamma^2} \mathbf{D}^{-2}, \quad (3.42)$$

where $\mathbf{D} = \text{diag}([P'_1, \dots, P'_N]^T)$, we can obtain

$$\boldsymbol{\Sigma}_{\boldsymbol{\zeta}}^{-1/2} \mathbf{h} = \boldsymbol{\Sigma}_{\boldsymbol{\zeta}}^{-1/2} \mathbf{F} \boldsymbol{\phi} + \boldsymbol{\Sigma}_{\boldsymbol{\zeta}}^{-1/2} \boldsymbol{\zeta} \quad (3.43a)$$

$$\Rightarrow \mathbf{D} \mathbf{h} = \mathbf{D} \mathbf{F} \boldsymbol{\phi} + \mathbf{D} \boldsymbol{\zeta} \quad (3.43b)$$

where the covariance matrix of $\mathbf{D} \boldsymbol{\zeta}$ becomes a scaled identity matrix, i.e.,

$$\boldsymbol{\Sigma}_{\mathbf{D} \boldsymbol{\zeta}} = \frac{\ln(10)^2 P_0'^2 \sigma^2}{25\gamma^2} \mathbf{I}_N.$$

Note that this whitening step simply corresponds to an appropriate scaling of every entry of (3.40).

The whitened model (3.43b) is found to match our general model (3.1), since we notice that $\mathbf{D} \mathbf{F}$ can be split into

$$\mathbf{D} \mathbf{F} = \begin{bmatrix} \vdots & \vdots & \vdots \\ 2\mathbf{s}_i^T P'_i & -P'_i & 1 \\ \vdots & \vdots & \vdots \end{bmatrix} = [\mathbf{F}' \quad \mathbf{1}_{N \times 1}], \quad (3.44)$$

where \mathbf{F}' contains the first $d+1$ columns of $\mathbf{D} \mathbf{F}$. The relation between this model and the general model (3.1) is presented in Table 3.2 on page 98. Note that we only consider a

single nuisance parameter P'_0 in this model ($M = 1$). Although we could consider both $\|\mathbf{x}_t\|_2^2$ and P'_0 as nuisance parameters ($M = 2$), which would lead to the same performance after using the correct preprocessing steps, the reason why we take $M = 1$ here is to connect this model to the existing literature. For instance, after removing P'_0 using a single differencing step, the model for $\mathbf{\Gamma}_j \mathbf{D} \mathbf{h}$ is equal to the SD-DRSS model used in [58, eq. (22)]. However, without an appropriate whitening procedure, the LS estimators of the SD-RSS and SD-DRSS models yield a different performance, which is why they were treated and studied separately. Now, we realize that they actually are identical to each other as long as the model noise is properly whitened.

Numerical results: A simulation has also been conducted to verify our conclusions for this example. As before, the BLUEs of the joint estimation, OSP-based estimation, and the differential estimation for the SD-RSS model are evaluated and compared with some LS estimators without a proper whitening. Based on the original model in (3.37), the CRLB with an unknown P_0 is easy to calculate [1, Chapter 3]. From the numerical results in Fig. 3.3, the critical observation is that all the BLUEs here yield exactly the same performance as expected. Due to the colored model noise, the J-LS-SD-RSS and the D-LS-SD-RSS are relatively worse. Finally, denoting $R \triangleq \|\mathbf{x}_t\|_2^2$, we again point out that neglecting the dependence between R and \mathbf{x}_t results in the gap between the CRLB and the estimators presented here.

3.3.3. OTHER EXAMPLES

WE believe that there are many other examples with linear nuisance parameters for our results. However, due to the limited space, we will only point out some of them. Besides the aforementioned localization examples, if anchors are separated into groups with different central clocks, multiple relative clock biases might exist in the TDOA measurements for localization, which can be removed by the OSP method [62, eq. (3)]. In cooperative localization, the multidimensional scaling (MDS) also uses the OSP-based method to eliminate the unknown terms [63, eq. (3)]. An acoustic source localization model, which also matches our general model (3.1), was presented in [64, eq. (6)]. In [4, eq. (2)], the transmission times and clock offsets are the unknown nuisance parameters for the considered clock synchronization problem. The authors claim that those unknown parameters are systematically ML estimated before the synchronization. However, in fact, those nuisance parameters are equivalently removed by using respectively the observations \mathbf{d}_{avg} in (3.13) or the OSP procedure. In hyperspectral imaging, OSP is also a very common procedure to extract the desired signals [20]. And when tracking mobile targets, frequency-difference-of-arrival measurements are often measured to cope with the Doppler effect [18, 19, 65, 66]. Furthermore, multiple-input-multiple-output (MIMO) receiver design might be affected by some nuisance parameters like I-Q imbalance and DC offset [5, eq. (7)]. In machine learning, a well-designed OSP is desired for dimensionality reduction [8, 9]. Extracting and working on the signal space is a strong need for signal separation [7] and underwater communication [6], which can be facilitated by OSP. At last, the famous differential global positioning system (DGPS) introduces a reference station on the ground and constructs a new differential observation set for positioning [67], where even the double differencing process is considered [68–70].

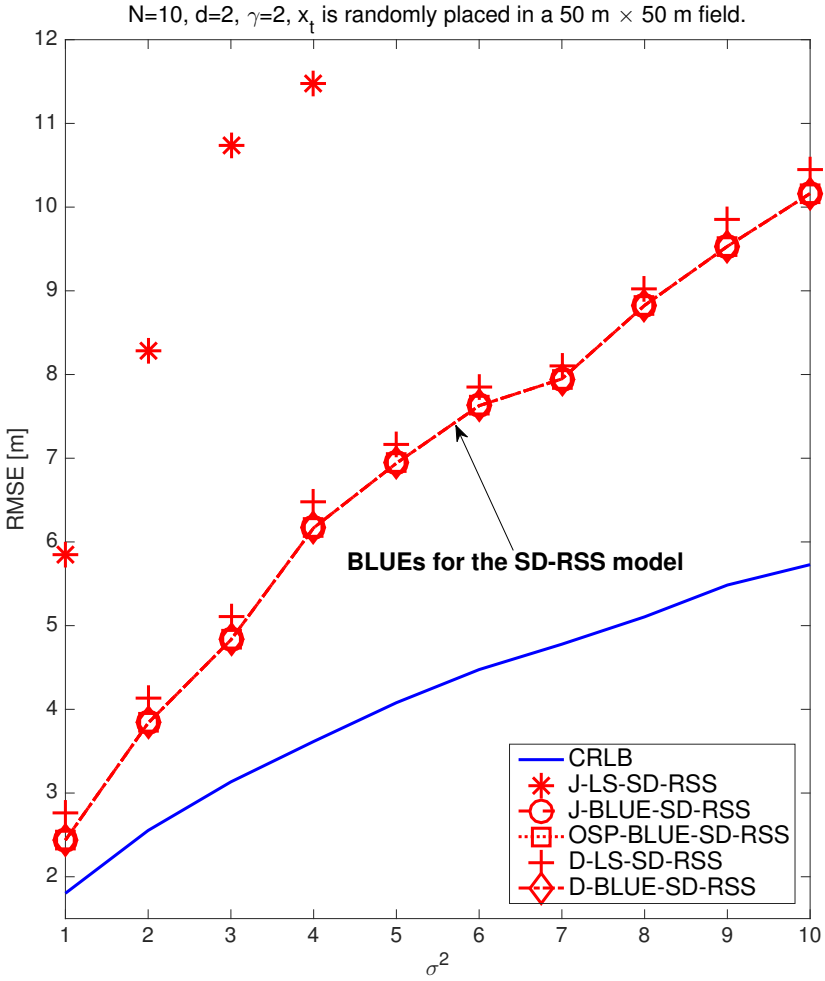


Figure 3.3: Performance of different RSS-based estimators: the target node is randomly placed in a 50 \times 50 field and 10 anchors are deployed with coordinates of (50,50), (50,0), (0,50), (0,0), (25,7), (25,43), (12,33), (12,16), (37,33) and (37,16). The transmit power is set to 10 dBm and the PLE is set to 2.

3.4. CONCLUSIONS

IN this paper, we have introduced a general framework for estimation in the presence of unknown linear nuisance parameters. Three different kinds of methods to cope with the unknown nuisance parameters have been studied, i.e., the joint estimation, the OSP-based estimation and the differential estimation. These approaches have been analyzed by investigating their corresponding BLUEs, where a new differential method has been introduced to cope with multiple nuisance parameters. We have discovered that, after a proper whitening procedure, all the BLUEs are equivalent to each other. From this interesting fact, one can draw some useful conclusions:

1. there only exists one unique BLUE for all these methods proposed to cope with unknown nuisance parameters.
2. compared with the joint estimation, which directly utilizes all the original observations, none of the other two methods suffers any information loss.
3. for the differential approach, which requires selecting some references, the choice of the references is not important since there is no actual trace of the selected references in the corresponding BLUE.
4. In the differencing process, compared with the full differential observation set, any subset related to a single reference already preserves the full data information.

The presented analyses of the general model can be projected onto many practical applications, e.g., hyperspectral imaging, source localization and synchronization. Some localization examples have also been demonstrated, simulated and discussed to verify our conclusions.

REFERENCES

- [1] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993).
- [2] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, *Locating the nodes: cooperative localization in wireless sensor networks*, *Signal Processing Magazine, IEEE* **22**, 54 (2005).
- [3] S. Lopez, T. Vladimirova, C. Gonzalez, J. Resano, D. Mozos, and A. Plaza, *The Promise of Reconfigurable Computing for Hyperspectral Imaging Onboard Systems: A Review and Trends*, *Proc. IEEE* **101**, 698 (2013).
- [4] O. Jean and A. Weiss, *Passive Localization and Synchronization Using Arbitrary Signals*, *Signal Processing, IEEE Transactions on* **62**, 2143 (2014).
- [5] C. J. Hsu, R. Cheng, and W. H. Sheen, *Joint least squares estimation of frequency, dc offset, i-q imbalance, and channel in mimo receivers*, *IEEE Trans. Veh. Technol.* **58**, 2201 (2009).

- [6] J. He, M. N. S. Swamy, and M. O. Ahmad, *Joint space-time parameter estimation for underwater communication channels with velocity vector sensor arrays*, *IEEE Trans. Wireless Commun.* **11**, 3869 (2012).
- [7] M. A. Uusitalo and R. J. Ilmoniemi, *Signal-space projection method for separating meg or eeg into components*, *Med. Biol. Eng. Comput.* **35**, 135 (1997).
- [8] E. Kokiopoulou and Y. Saad, *Orthogonal neighborhood preserving projections: A projection-based dimensionality reduction technique*, *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 2143 (2007).
- [9] C. X. Ren and D. Q. Dai, *2d-onpp: Two dimensional extension of orthogonal neighborhood preserving projections for face recognition*, in *Pattern Recognition, 2008. CCPR '08. Chinese Conference on* (2008) pp. 1–6.
- [10] S. Bar and J. Tabrikian, *Bayesian estimation in the presence of deterministic nuisance parameters-part i: Performance bounds*, *IEEE Transactions on Signal Processing* **63**, 6632 (2015).
- [11] S. Bar and J. Tabrikian, *Bayesian estimation in the presence of deterministic nuisance parameters-part ii: Estimation methods*, *IEEE Transactions on Signal Processing* **63**, 6647 (2015).
- [12] M. Abramowitz, I. A. Stegun, *et al.*, *Handbook of mathematical functions*, Applied mathematics series **55**, 62 (1966).
- [13] S. Zhu and Z. Ding, *Joint synchronization and localization using TOAs: A linearization based WLS solution*, *Selected Areas in Communications, IEEE Journal on* **28**, 1017 (2010).
- [14] T.-M. Tu, C.-H. Chen, and C.-I. Chang, *A posteriori least squares orthogonal subspace projection approach to desired signature extraction and detection*, *Geoscience and Remote Sensing, IEEE Transactions on* **35**, 127 (1997).
- [15] Y. Chan and K. Ho, *A simple and efficient estimator for hyperbolic location*, *Signal Processing, IEEE Transactions on* **42**, 1905 (1994).
- [16] K. Ho, *Bias Reduction for an Explicit Solution of Source Localization Using TDOA*, *Signal Processing, IEEE Transactions on* **60**, 2101 (2012).
- [17] Y. Wang and G. Leus, *Reference-free time-based localization for an asynchronous target*, *EURASIP Journal on Advances in Signal Processing* **2012**, 19 (2012).
- [18] K. Ho, X. Lu, and L. Kovavisaruch, *Source Localization Using TDOA and FDOA Measurements in the Presence of Receiver Location Errors: Analysis and Solution*, *Signal Processing, IEEE Transactions on* **55**, 684 (2007).
- [19] D. Musicki and W. Koch, *Geolocation using TDOA and FDOA measurements*, in *Information Fusion, 2008 11th International Conference on* (2008) pp. 1–8.

- [20] J. Harsanyi and C.-I. Chang, *Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach*, *Geoscience and Remote Sensing, IEEE Transactions on* **32**, 779 (1994).
- [21] C.-I. Chang, *Orthogonal subspace projection (OSP) revisited: a comprehensive study and analysis*, *Geoscience and Remote Sensing, IEEE Transactions on* **43**, 502 (2005).
- [22] M. Song and C. I. Chang, *A Theory of Recursive Orthogonal Subspace Projection for Hyperspectral Imaging*, *IEEE Trans. Geosci. Remote Sens.* **53**, 3055 (2015).
- [23] Q. Xu, Y. Lei, J. Cao, and H. Wei, *An improved algorithm based on reference selection for time difference of arrival location*, in *Image and Signal Processing (CISP), 2014 7th International Congress on* (2014) pp. 953–957.
- [24] Y. Wang, F. Zheng, M. Wiemeler, W. Xiong, and T. Kaiser, *Reference Selection for Hybrid TOA/RSS Linear Least Squares Localization*, in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th* (2013) pp. 1–5.
- [25] I. Guvenc, S. Gezici, F. Watanabe, and H. Inamura, *Enhancements to Linear Least Squares Localization Through Reference Selection and ML Estimation*, in *2008 IEEE Wireless Communications and Networking Conference* (2008) pp. 284–289.
- [26] H. C. So, Y. T. Chan, and F. Chan, *Closed-Form Formulae for Time-Difference-of-Arrival Estimation*, *Signal Processing, IEEE Transactions on* **56**, 2614 (2008).
- [27] S. C. K. Herath and P. N. Pathirana, *Robust Localization With Minimum Number of TDoA Measurements*, *IEEE Signal. Proc. Let.* **20**, 949 (2013).
- [28] Y. Huang, J. Benesty, G. W. Elko, and R. M. Mersereati, *Real-time passive source localization: a practical linear-correction least-squares approach*, *IEEE T. Speech. Audi. P.* **9**, 943 (2001).
- [29] R. O. Schmidt, *A New Approach to Geometry of Range Difference Location*, *IEEE Trans. Aerosp. Electron. Syst.* **AES-8**, 821 (1972).
- [30] R. Schmidt, *Least squares range difference location*, *IEEE Trans. Aerosp. Electron. Syst.* **32**, 234 (1996).
- [31] S. Venkatesh and R. M. Buehrer, *A Linear Programming Approach to NLOS Error Mitigation in Sensor Networks*, in *Proceedings of the 5th International Conference on Information Processing in Sensor Networks*, IPSN '06 (ACM, New York, NY, USA, 2006) pp. 301–308.
- [32] A.-J. van der Veen, E. Deprettere, and A. Swindlehurst, *Subspace-based signal analysis using singular value decomposition*, *Proc. IEEE* **81**, 1277 (1993).
- [33] L. Scharf and M. McCloud, *Blind adaptation of zero forcing projections and oblique pseudo-inverses for subspace detection and estimation when interference dominates noise*, *Signal Processing, IEEE Transactions on* **50**, 2938 (2002).

- [34] A. Sayed, A. Tarighat, and N. Khajehnouri, *Network-based wireless location: challenges faced in developing techniques for accurate wireless location information*, *Signal Processing Magazine, IEEE* **22**, 24 (2005).
- [35] W. Foy, *Position-Location Solutions by Taylor-Series Estimation*, *Aerospace and Electronic Systems, IEEE Transactions on AES-12*, 187 (1976).
- [36] C. Kelley, *Iterative Methods for Optimization*, Frontiers in Applied Mathematics (Society for Industrial and Applied Mathematics, 1999).
- [37] A. Beck, P. Stoica, and J. Li, *Exact and Approximate Solutions of Source Localization Problems*, *Signal Processing, IEEE Transactions on* **56**, 1770 (2008).
- [38] D. B. Haddad, W. A. Martins, M. d. V. M. da Costa, L. W. P. Biscainho, L. O. Nunes, and B. Lee, *Robust Acoustic Self-Localization of Mobile Devices*, *IEEE Trans. Mob. Comput.* **15**, 982 (2016).
- [39] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, *Least squares algorithms for time-of-arrival-based mobile location*, *IEEE Trans. Signal Process.* **52**, 1121 (2004).
- [40] J. C. Chen, R. E. Hudson, and K. Yao, *Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field*, *IEEE Trans. Signal Process.* **50**, 1843 (2002).
- [41] C.-H. Park, S. Lee, and J.-H. Chang, *Robust closed-form time-of-arrival source localization based on alpha-trimmed mean and Hodges-Lehmann estimator under NLOS environments*, *Signal Process.* **111**, 113 (2015).
- [42] M. Sun and K. C. Ho, *Successive and Asymptotically Efficient Localization of Sensor Nodes in Closed-Form*, *IEEE Trans. Signal Process.* **57**, 4522 (2009).
- [43] N. D. Gaubitch, W. B. Kleijn, and R. Heusdens, *Auto-localization in ad-hoc microphone arrays*, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (2013) pp. 106–110.
- [44] L. Wang, T. K. Hon, J. D. Reiss, and A. Cavallaro, *Self-Localization of Ad-Hoc Arrays Using Time Difference of Arrivals*, *IEEE Trans. Signal Process.* **64**, 1018 (2016).
- [45] K. Liu, X. Liu, and X. Li, *Guoguo: Enabling Fine-Grained Smartphone Localization via Acoustic Anchors*, *IEEE Trans. Mob. Comput.* **15**, 1144 (2016).
- [46] J. J. Caffery, *A new approach to the geometry of TOA location*, in *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, Vol. 4 (2000) pp. 1943–1949 vol.4.
- [47] Y. Wang, G. Leus, and X. Ma, *Time-based localization for asynchronous wireless sensor networks*, in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2011) pp. 3284–3287.
- [48] P. Stoica and J. Li, *Lecture Notes - Source Localization from Range-Difference Measurements*, *Signal Processing Magazine, IEEE* **23**, 63 (2006).

- [49] Y. Liu, F. Guo, L. Yang, and W. Jiang, *An Improved Algebraic Solution for TDOA Localization With Sensor Position Errors*, *IEEE Commun. Lett.* **19**, 2218 (2015).
- [50] J. Liu, Z. Wang, J. H. Cui, S. Zhou, and B. Yang, *A Joint Time Synchronization and Localization Design for Mobile Underwater Sensor Networks*, *IEEE Trans. Mob. Comput.* **15**, 530 (2016).
- [51] B. Huang, L. Xie, and Z. Yang, *TDOA-Based Source Localization With Distance-Dependent Noises*, *IEEE Trans. Wireless Commun.* **14**, 468 (2015).
- [52] H. Yang, J. Chun, and D. Chae, *Hyperbolic Localization in MIMO Radar Systems*, *IEEE Antennas Wirel. Propag. Lett.* **14**, 618 (2015).
- [53] J. Smith and J. Abel, *The spherical interpolation method of source localization*, *IEEE J. Oceanic. Eng.* **12**, 246 (1987).
- [54] B. Friedlander, *A passive localization algorithm and its accuracy analysis*, *IEEE J. Oceanic. Eng.* **12**, 234 (1987).
- [55] J. Smith and J. Abel, *Closed-form least-squares source location estimation from range-difference measurements*, *IEEE Trans. Acoust. Speech Signal Process.* **35**, 1661 (1987).
- [56] X. Li, *RSS-Based Location Estimation with Unknown Pathloss Model*, *Wireless Communications, IEEE Transactions on* **5**, 3626 (2006).
- [57] H. C. So and L. Lin, *Linear Least Squares Approach for Accurate Received Signal Strength Based Source Localization*, *Signal Processing, IEEE Transactions on* **59**, 4035 (2011).
- [58] R. Vaghefi, M. Gholami, and E. Strom, *RSS-based sensor localization with unknown transmit power*, in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on* (2011) pp. 2480–2483.
- [59] M. Gholami, R. Vaghefi, and E. Strom, *RSS-Based Sensor Localization in the Presence of Unknown Channel Parameters*, *Signal Processing, IEEE Transactions on* **61**, 3752 (2013).
- [60] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001).
- [61] M. Abramowitz, *[Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables]* (Dover Publications, Incorporated, 1974).
- [62] Y. Wang and K. Ho, *TDOA Source Localization in the Presence of Synchronization Clock Bias and Sensor Position Errors*, *Signal Processing, IEEE Transactions on* **61**, 4532 (2013).
- [63] S. Kumar, R. Kumar, and K. Rajawat, *Cooperative localization of mobile networks via velocity-assisted multidimensional scaling*, *IEEE Trans. Signal Process.* **64**, 1744 (2016).

- [64] D. Li and Y. H. Hu, *Least square solutions of energy based acoustic source localization problems*, in *Parallel Processing Workshops, 2004. ICPP 2004 Workshops. Proceedings. 2004 International Conference on* (2004) pp. 443–446.
- [65] H. W. Wei, R. Peng, Q. Wan, Z. X. Chen, and S. F. Ye, *Multidimensional Scaling Analysis for Passive Moving Target Localization With TDOA and FDOA Measurements*, *IEEE Trans. Signal Process.* **58**, 1677 (2010).
- [66] K. C. Ho and W. Xu, *An accurate algebraic solution for moving source location using TDOA and FDOA measurements*, *IEEE Trans. Signal Process.* **52**, 2453 (2004).
- [67] B. Parkinson and J. Spilker, *Global Positioning System: Theory and Applications*, Progress in astronautics and aeronautics No. v. 1 (American Institute of Aeronautics & Astronautics, 1996).
- [68] R. O. Nielsen, *Relationship between dilution of precision for point positioning and for relative positioning with gps*, *IEEE Trans. Aerosp. Electron. Syst.* **33**, 333 (1997).
- [69] P. J. G. Teunissen, *A proof of nielsen's conjecture on the gps dilution of precision*, *IEEE Trans. Aerosp. Electron. Syst.* **34**, 693 (1998).
- [70] C. Park and I. Kim, *Comments on "relationships between dilution of precision for point positioning and for relative positioning with gps"*, *IEEE Trans. Aerosp. Electron. Syst.* **36**, 315 (2000).

III

PATH-LOSS EXPONENT SELF-ESTIMATION IN WIRELESS NETWORKS

4

SELF-ESTIMATION OF PATH-LOSS EXPONENT IN WIRELESS NETWORKS AND APPLICATIONS

Yongchang HU and Geert LEUS

Man errs as long as he strives.

Goethe

The path-loss exponent (PLE) is one of the most crucial parameters in wireless communications to characterize the propagation of fading channels. It is currently adopted for many different kinds of wireless network problems such as power consumption issues, modeling the communication environment, and received signal strength (RSS)-based localization. PLE estimation is thus of great use to assist wireless networking. However, a majority of methods to estimate the PLE requires either some particular information of the wireless network, which might be unknown, or some external auxiliary devices, such as anchor nodes or the global positioning system (GPS). Moreover, this external information might sometimes be unreliable, spoofed or difficult to obtain. Therefore, a self-estimator for the PLE, which is able to work independently, becomes an urgent demand to robustly and securely get a grip on the PLE for various wireless network applications.

This paper is the first to introduce two methods which can solely and locally estimate the PLE. To start, a new linear regression model for the PLE is presented. Based on this model, a closed-form total least squares method to estimate the PLE is firstly proposed,

Parts of this chapter have been published in *Annalen der Physik*.

in which, without any other assistance or external information, each node can estimate the path-loss exponent merely by collecting RSSs. Secondly, in order to suppress the estimation errors, a closed-form weighted total least squares method is further developed having a better performance. Due to their simplicity and independence of any auxiliary system, our two proposed methods can be easily incorporated into any kind of wireless communication stack. Simulation results show that our estimators are reliable even in harsh environments, where the PLE is high. Many potential applications are also explicitly illustrated in this paper, such as secure RSS-based localization, k -th nearest neighbor routing, etc. Those applications detail the significance of self-estimation of the PLE.

4.1. INTRODUCTION

IN wireless communications, the received instantaneous signal powers at receivers are commonly modeled as the product of the large-scale path-loss and the small-scale fading. The large-scale path-loss assumes that the attenuation of the average received power is subject to the transmitter-receiver distance r as r^γ , where γ is the path-loss exponent (PLE). Due to the dynamics of the communication channel, the PLE varies over different scenarios and different locations. At the same time, the small-scale fading constitutes a rapid fluctuation around the average of the received powers and follows a stochastic process. It is mainly due to the multi-path effect and changes over very small distances and very small time intervals. However, it can generally be well-suppressed by means of some special receiver designs and digital signal processing (DSP). Therefore, the PLE becomes a key parameter to characterize the propagation channel, which significantly determines power consumption, quality of a transmission link, efficiency of packet delivery, etc.

It is of importance to accurately estimate the PLE so that the wireless communication stack can be dynamically adapted to the PLE changes in order to yield a better performance. For instance, a path with a relatively low PLE can be chosen to route messages in order to save power. The PLE is also significant for some other applications. For instance, to calculate the location of a target node in received signal strength (RSS)-based localization, accurate PLE estimation is required, which is mostly provided by reference nodes with known positions. However, in some cases, the reference nodes might be broken and cannot talk to the target node or the location information of the reference nodes might be unreliable, or spoofed by an adversary. Then, accurately estimating the PLE will become a difficult task.

Current methods to estimate the PLE either require some information of the wireless network, which is unknown in most cases, or the assistance from auxiliary systems. Three algorithms are presented in [1]: firstly, when the network density is known, the PLE can be estimated by observing RSSs during several time slots and by calculating the mean interference; as regards to the other two algorithms, by changing the receiver's sensitivity, the PLE can be estimated either from the corresponding virtual outage probabilities or from the corresponding neighborhood sizes. All three algorithms require the knowledge of the network density or the receiver settings, and even require changing them. Other methods to estimate the PLE mostly lie in the area of RSS-based localization. As already mentioned, using the RSSs for localization requires an accurate estimate of the PLE, which is tightly related with the transmitter-receiver distance. Therefore, special reference nodes with known positions, namely anchor nodes, are strategically pre-deployed with the purpose of calibrating the PLE [2]. Considering that the transmitter-receiver distances between anchor nodes can be difficult or expensive to accurately measure in some environments, the PLE can also be estimated by using received power measurements and geometric constraints of anchor nodes to avoid the distance calculation [3]. In the mean time, many efforts have been put to jointly estimate the location and the PLE [4–6]. Some other methods start with an initial guess of the PLE to approximate the location which is then used to update the PLE estimate [7, 8]. However, all those methods basically rely on the information from anchor nodes or other auxiliary systems. Once such systems are attacked, unavailable or generate large errors, the

impact on the whole system will be unimaginable. Furthermore, the above methods are also not feasible for many kinds of wireless networks, in which communications and information exchanges might be highly restricted. Therefore, a new self-estimator of the PLE is urgently required, which can solely and locally estimate the PLE without relying on any external assistance. Such an estimator should not only be able to serve localization techniques, but can also act as a general method which can be easily incorporated into any kind of wireless network and any layer of the communication stack.

The rest of the paper is structured as follows. In Section 4.2, we present the system model considered in this paper and discuss the problem statement. Some new parameters are introduced in Section 4.3 to build a linear regression model for the PLE. Section 4.4 presents and discusses the derivation of our two proposed path-loss exponent estimators. Simulation results are given and analyzed in Section 4.5. Many potential applications are discussed in Section 4.6. Section 4.7 finally summarizes the paper.

4

4.2. SYSTEM MODEL

IN this section, we introduce some important system model concepts and additionally describe the problem statement.

4.2.1. NODE DISTRIBUTION

DUE to the unknown topology of wireless networks, especially in wireless ad hoc networks, neighbors of a node are ideally considered randomly deployed within the transmission range, indicated by W . In other words, a local random region around the considered node is assumed. Therefore, the probability of finding k nodes in a subset $\Omega \subset W$ is given by

$$\mathbb{P}[k \text{ nodes in } \Omega] = \frac{n!}{k!(n-k)!} \left(\frac{\mu(\Omega)}{\mu(W)} \right)^k \left(1 - \frac{\mu(\Omega)}{\mu(W)} \right)^{n-k}, \quad (4.1)$$

where \mathbb{P} denotes probability, n is the neighborhood size in W and $\mu(\cdot)$ is the standard Lebesgue measure. If we let Ω be a d -dimensional ball of radius r originating at the considered node, $\mu(\Omega)$ is the volume of Ω and is given by $\mu(\Omega) = c_d r^d$, where

$$c_d = \frac{\pi^{d/2}}{\Gamma(1+d/2)}, \quad (4.2)$$

with $\Gamma(\cdot)$ the gamma function. When $d = 1, 2$ or 3 , $c_d = 2, \pi$ and $\frac{4}{3}\pi$, respectively. For example, wireless vehicular networks can be modeled in a 1-dimensional space, a flat-earth model requires $d = 2$, and wireless unmanned aerial vehicle communications requires $d = 3$. In this paper, all formulae are generalized in a d -dimensional manner for the sake of theoretical consistency.

4.2.2. CHANNEL MODEL

THE attenuation of the channel can be modeled as comprised of the large-scale fading, the shadowing effect and the small-scale fading. The large-scale fading indicates that the empirical deterministic reduction in power density of an electromagnetic wave

is exponentially associated with the distance when it propagates through space. We assume that the transmitted power P_t is reduced through the propagation channel over a distance r , such that the received signal strength P_r is given by

$$P_r = C_1 P_t \left(\frac{r_0}{r} \right)^\gamma, \quad (4.3)$$

where $r_0 \ll r$ is the reference distance related to far-field and C_1 is a non-distance-related constant that depends on the carrier frequency, the antenna gain and the speed of light. P_r and P_t are both expressed in *Watts*.

Depending on the environment, the path-loss exponent (PLE) γ ranges from 2 to 6 [9]. Obstacles, such as trees, buildings and so forth, cause the actual attenuation of the received power to follow a log-normal distribution, also called the shadowing effect. As such, (4.3) has to be changed into

$$\Delta P = 10\gamma \log_{10}(r) - 10\log_{10}(C_1) - 10\gamma \log_{10}(r_0) + \chi, \quad (4.4)$$

where $\Delta P = 10\log_{10}\left(\frac{P_t}{P_r}\right)$ in *dB* indicates the attenuation of the signal strength and χ follows a zero-mean *Gaussian distribution* with standard deviation $2 < \sigma < 12$. To serve the following derivations, two severe consequences of the shadowing effect should be mentioned:

1. The theoretical neighborhood size n is different from the actual neighborhood size $\hat{n} = n + \Delta n$. As shown in Fig. 4.1 for $d = 2$, the dashed regular circle is the theoretical transmission range of node A . In fact, packets can be successfully received under the condition that $P_r > P_{thres}$, where P_{thres} is the receiver's sensitivity. Due to the shadowing effect, the actual transmission range is irregular, as indicated by the solid line.
2. Another consequence caused by the shadowing effect is that after ranking all the received powers at node A , the node with the \hat{i} -th strongest received power $P_{r,\hat{i}}$ corresponds to the i -th nearest neighbor at distance r_i , where $\hat{i} = i + \Delta i$.

When signals are being transmitted, scatterers and reflectors create several reflected paths that reach the receiver, besides the line-of-sight (LOS). This is called the small-scale fading, which is non-distance-related. The instantaneous received signal envelope follows the Nakagami- m distribution [10] and the distribution of the instantaneous received power p is hence given by

$$\mathbb{P}(p) = \frac{\left(\frac{m}{E(p)}\right)^m p^{m-1} e^{-\frac{mp}{E(p)}}}{\Gamma(m)}, \quad (4.5)$$

where m is the fading parameter and a small value of m indicates more fading. The measured received power P_r can be obtained by taking the average over K consecutive time slots of instantaneous received powers p_k , i.e. $P_r = \frac{1}{K} \sum_{k=1}^K p_k$ and thus, $Var(P_r) = \frac{[E(p_k)]^2}{Km}$. When K is large enough, the impact of the small-scale fading can be greatly eliminated. Additionally, a well-designed receiver is able to suppress the multi-path effect to a great degree by using special antenna designs such as a choke ring antenna,

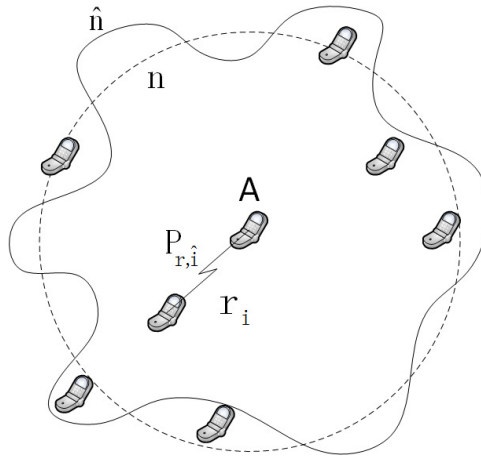


Figure 4.1: The impact of the shadowing effect on node A : \hat{n} is the estimate of the theoretical neighborhood size n by counting the reachable neighbors, $\hat{n} = n + \Delta n$. By ranking the received powers at A , the corresponding ranking numbers \hat{i} are the estimate of the ranking numbers i of the ranges, where $\hat{i} = i + \Delta i$.

a right-hand-circular polarized (RHCP) antenna, etc. Therefore, the power attenuation model in this paper is mostly subject to the large-scale fading and the shadowing, and hence we will rely on (4.4) in the rest of this paper.

4.2.3. PROBLEM STATEMENT

WE are now aiming at developing a new self-estimator of the PLE. The desired properties of the proposed estimator can be summarized as: *simple*, *pervasive*, *local*, *sole*, *collective* and *secure*. *Simple* indicates that the proposed estimator should be easy to implement and carry out. *Pervasive* signals that it can be incorporated into any kind of network regardless of its design. Therefore, the only freedom left for us is to utilize the received signal strength. Some kind of networks might not have any external auxiliary system or access to external information and their mutual nodal cooperations might be severely constrained. And even if there are no such constraints, adversaries can easily tamper with or forge the exchanged critical information. This requires that the estimator has to run *solely* on a single node by *collecting* the *locally* received signal strengths. By this means, a path-loss exponent can be *securely* and *locally* estimated.

As is shown in (4.3), the path-loss exponent γ is strictly subject to the power attenuation and the transmitter-receiver distance. Therefore, conventional estimators in wireless localization try to obtain the path-loss exponent by introducing anchor nodes to fix the transmitter-receiver distance and by observing power attenuations. However, the desired properties of the proposed estimator determine that it is not possible to fix or to know exact transmitter-receiver distances of some of the collected received signal strengths. As such, we can define the problem as “How can we estimate the path-loss exponent γ without knowing transmitter-receiver distances, i.e., merely from the local received signal strengths?”

4.3. LINEAR REGRESSION MODEL FOR THE PATH-LOSS EXPONENT

IN order to solve the earlier mentioned problem, we introduce some new parameters. After estimating those parameters, a new linear regression model for the PLE is presented.

4.3.1. RANKING RECEIVED SIGNAL STRENGTHS

LET us focus on a single node and denote $P_{r,\hat{i}}$ as the \hat{i} -th strongest power received at the considered node where $\hat{i} = 1, 2, \dots, \hat{n}$, i.e., $P_{r,1} \geq P_{r,2} \geq \dots \geq P_{r,\hat{n}}$ and r_i as the i -th closest range to the considered node, where $i = 1, 2, \dots, n$, i.e., $r_1 \leq r_2 \leq \dots \leq r_n$. As we mentioned earlier, $\hat{i} = i + \Delta i$ is considered as an estimate of i , where Δi is called the mismatch.

From (4.4), we can then write

$$\Delta P_{\hat{i}} = 10\gamma \log_{10}(r_i) - C_2 + \chi_i, \quad (4.6)$$

where $\chi_i \sim \mathcal{N}(0, \sigma^2)$, $\Delta P_{\hat{i}} = 10 \log_{10}(P_t / P_{r,\hat{i}})$ and $C_2 = 10 \log_{10}(C_1) + 10\gamma \log_{10}(r_0)$ is a constant. We assume that all neighboring nodes transmit signals with the same power P_t such that the ordered values of $P_{r,\hat{i}}$ lead to the ordered values of $\Delta P_{\hat{i}}$, i.e., we can assume that $\Delta P_1 \leq \Delta P_2 \leq \dots \leq \Delta P_{\hat{n}}$. Admittedly, in a more realistic situation, the transmit power P_t at each neighboring node might be different. But our proposed estimators can still remain feasible in such a case and we will come back to this issue in Section 4.4.4.

4.3.2. LINEAR REGRESSION MODEL FOR THE PATH-LOSS EXPONENT

FROM (4.6), we notice that $\Delta P_{\hat{i}}$ is a function of P_t and C_2 , which are both unknown. But these can be canceled by subtracting $\Delta P_{\hat{j}}$ from $\Delta P_{\hat{i}}$ leading to $\Delta P_{\hat{i},\hat{j}} = \Delta P_{\hat{i}} - \Delta P_{\hat{j}} = 10 \log_{10}(P_{r,\hat{j}} / P_{r,\hat{i}})$ which can further be written as

$$\begin{aligned} \Delta P_{\hat{i},\hat{j}} &= 10\gamma \log_{10}(r_i) - 10\gamma \log_{10}(r_j) + \chi_{i,j} \\ &= 10\gamma \log_{10}\left(\frac{r_i}{r_j}\right) + \chi_{i,j} \end{aligned} \quad (4.7)$$

where $\chi_{i,j} \sim \mathcal{N}(0, 2\sigma^2)$.

Now, we define $L_i = 10 \log_{10}(r_i)$ as a logarithmic function of r_i , and hence $L_{i,j} = L_i - L_j = 10 \log_{10}(\frac{r_i}{r_j})$. Thus (4.7) becomes

$$\Delta P_{\hat{i},\hat{j}} = \gamma L_{i,j} + \chi_{i,j}. \quad (4.8)$$

It is already apparent that if $L_{i,j}$ can be estimated, a linear regression model for the path-loss exponent can be constructed from (4.8). Let us denote $\hat{L}_{\hat{i},\hat{j}}$ as the estimate of $L_{i,j}$ and $\varepsilon_{\hat{i},\hat{j}}$ as the corresponding estimation error. The linear regression model is then given by

$$\Delta P_{\hat{i},\hat{j}} = \gamma(\hat{L}_{\hat{i},\hat{j}} - \varepsilon_{\hat{i},\hat{j}}) + \chi_{i,j}. \quad (4.9)$$

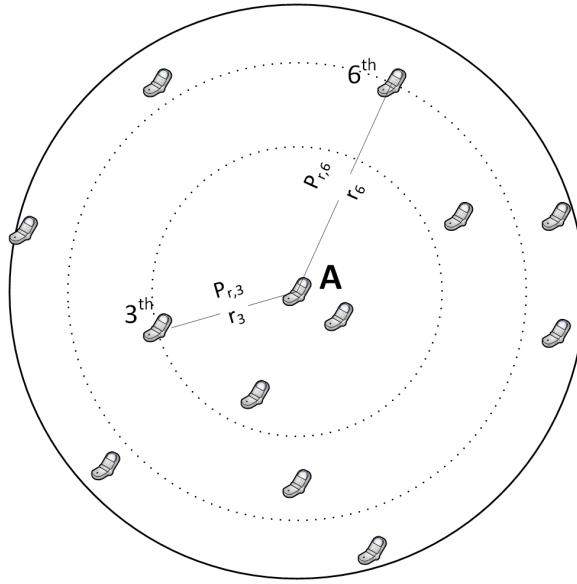


Figure 4.2: In 2-dimensional space when the shadowing effect does not impact the ranking, i.e. $\hat{i} = i$, the solid circle shows the transmit range of random node A , where A receives 12 signal strengths from its neighbors. Its 3^{th} and 6^{th} closest neighbors lie on the dotted circles which have r_3 and r_6 as radii, respectively. Therefore, r_3 has 3 nodes inside, while r_6 has 6 nodes inside. $P_{r,3}$ and $P_{r,6}$ are respectively the 3^{th} and the 6^{th} strongest received powers. $\Delta P_{3,6} = 10 \log_{10}(P_{r,6}/P_{r,3})$ and $\hat{L}_{3,6} = \frac{10}{2} \log_{10}(\frac{3}{6}) \approx -1.505$. Likewise, other pairs of $\Delta P_{\hat{i}, \hat{j}}$ and $\hat{L}_{\hat{i}, \hat{j}}$ can be obtained.

4.3.3. ESTIMATION OF $L_{i,j}$

As discussed in the problem statement, it is not possible to directly obtain the transmitter-receiver distances if the estimating node solely and locally collects the received signal strengths. Therefore, the idea of ranking the received signal strengths is crucial to our method.

By ranking the values of $P_{r,\hat{i}}$, we obtain the ranking number \hat{i} which will be further used to estimate the ranking numbers i of the ranges, where we recall that $\hat{i} = i + \Delta i$. Additionally, it is obvious that i indicates the number of nodes within the ball of radius r_i , which can be further exemplified in Fig. 4.2. Therefore, the essence of the proposed method is to use the rank numbers of \hat{i} as new measurements to estimate the values of $L_{i,j}$.

Note that $L_{i,j}$ is a linear combination of L_i and L_j . We focus on estimating L_i and the estimate of L_j can be obtained likewise.

Considering (4.1) and (4.2), the probability mass function of finding i nodes within the d -ball of radius r_i , which is parameterized by $L_i = 10 \log_{10}(r_i)$, can be written as

$$\mathbb{P}[i | L_i] = \frac{n!}{i!(n-i)!} \left(\frac{c_d 10^{\frac{dL_i}{10}}}{\mu(W)} \right)^i \left(1 - \frac{c_d 10^{\frac{dL_i}{10}}}{\mu(W)} \right)^{n-i}. \quad (4.10)$$

Based on (4.10), to find the maximum likelihood estimator \hat{L}_i , we need to force the derivative of our likelihood function to zero by

$$\frac{\partial \ln(\mathbb{P}[i | L_i])}{\partial L_i} = 0. \quad (4.11)$$

Therefore, by solving (4.11), the maximum likelihood estimator \hat{L}_i can be easily obtained as

$$\hat{L}_i = \frac{10}{d} \log_{10} \left(\frac{i\mu(W)}{nc_d} \right). \quad (4.12)$$

Likewise, \hat{L}_j can be obtained and the estimate of $L_{i,j}$ is hence given by

$$\hat{L}_{i,j} = \frac{10}{d} \log_{10} \left(\frac{i}{j} \right) = L_{i,j} + \varepsilon_{i,j}, \quad (4.13)$$

where $\varepsilon_{i,j}$ is the estimation error of $\hat{L}_{i,j}$. Plugging $\hat{i} = i + \Delta i$ and $\hat{j} = j + \Delta j$ into (4.13), we have

$$\hat{L}_{\hat{i},\hat{j}} = \frac{10}{d} \log_{10} \left(\frac{\hat{i}}{\hat{j}} \right) = L_{i,j} + \varepsilon_{\hat{i},\hat{j}} \quad (4.14)$$

and

$$\varepsilon_{\hat{i},\hat{j}} = \varepsilon_{i,j} + \Delta \varepsilon_{i,j}, \quad (4.15)$$

where $\Delta \varepsilon_{i,j} = \hat{L}_{\hat{i},\hat{j}} - \hat{L}_{i,j} = \frac{10}{d} \log_{10} \left(\frac{i+\Delta i}{j+\Delta j} \right)$.

From (4.13) and (4.14), we even notice that $\mu(W)$, n and c_d disappear after subtraction. This makes the proposed estimators only subject to the received signal strengths and the rank numbers in a d -dimensional space.

4.4. PATH-LOSS EXPONENT ESTIMATION

To solve the linear regression model, the total least squares (TLS) method helps us to obtain the estimate of the path-loss exponent γ . However, the general solution to the total least squares method turns out to be time-consuming. Therefore, a closed-form solution is provided saving computational time tremendously. Moreover, a closed-form weighted total least squares method is further proposed to suppress the estimation errors yielding a better performance.

4.4.1. TOTAL LEAST SQUARES SOLUTION

As for the example in Fig. 4.2, node A computes $\Delta P_{\hat{i},\hat{j}}$ and estimates $\hat{L}_{\hat{i},\hat{j}}$ for all pairs of nodes within its range, i.e., $\hat{i}, \hat{j} = 1, 2, 3, \dots, \hat{n}$. However, from (4.9), we notice that the received signal strengths are corrupted by the shadowing and the values of $\hat{L}_{\hat{i},\hat{j}}$ are measured with errors. Therefore, the total least squares method is utilized to obtain our estimate, $\hat{\gamma}_{tls}$ [11].

We assume that the considered node has \hat{n} neighbors and all RSSs from its neighbors are ranked. Thus, we have a sample set of $\Delta P_{\hat{i},\hat{j}}$ values whose size is $N = \binom{\hat{n}}{2}$ in total. We

vectorize the collected samples of $\Delta P_{i,j}$ and the corresponding values of $\widehat{L}_{i,j}$, which are respectively represented by the $N \times 1$ vectors $\Delta \mathbf{P}$ and $\widehat{\mathbf{L}}$. Then, (4.9) can be rewritten as

$$\Delta \mathbf{P} = \gamma(\widehat{\mathbf{L}} - \mathbf{E}) + \mathbf{X}, \quad (4.16)$$

where \mathbf{E} and \mathbf{X} are respectively the $N \times 1$ vectors obtained by stacking the estimation errors $\varepsilon_{i,j}$ on $\widehat{L}_{i,j}$ and the shadowing parameters $\chi_{i,j}$. The basic idea of the total least squares method is to find an optimally corrected system of equations $\Delta \mathbf{P}_{tls} = \gamma \widehat{\mathbf{L}}_{tls}$ with $\Delta \mathbf{P}_{tls} := \Delta \mathbf{P} - \mathbf{X}_{tls}$, $\widehat{\mathbf{L}}_{tls} := \widehat{\mathbf{L}} - \mathbf{E}_{tls}$, where \mathbf{X}_{tls} and \mathbf{E}_{tls} are respectively optimal perturbation vectors. Therefore, the path-loss exponent estimate $\widehat{\gamma}_{tls}$ for γ is the solution to the optimization problem

$$\{\widehat{\gamma}_{tls}, \mathbf{X}_{tls}, \mathbf{E}_{tls}\} := \arg \min_{\gamma, \mathbf{X}, \mathbf{E}} \|\mathbf{X}\mathbf{E}\|_F^2 \quad (4.17)$$

subject to (4.16), where $\|\cdot\|_F$ is the Frobenius norm.

By changing (4.16) into

$$[(\widehat{\mathbf{L}} - \mathbf{E}) \quad (\Delta \mathbf{P} - \mathbf{X})] \begin{bmatrix} \gamma \\ -1 \end{bmatrix} = 0, \quad (4.18)$$

we see that this is a typical low-rank approximation problem where the rank of the augmented matrix $[\widehat{\mathbf{L}} \Delta \mathbf{P}]$ should be optimally reduced to 1.

Therefore, we compute the singular value decomposition (SVD) of $[\widehat{\mathbf{L}} \Delta \mathbf{P}]$ resulting in

$$[\widehat{\mathbf{L}} \Delta \mathbf{P}] = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

where \mathbf{V} can be explicitly expressed as

$$\mathbf{V} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}.$$

Based on the Eckart-Young theorem [12], the estimated path-loss exponent is then given by

$$\widehat{\gamma}_{tls} = -\frac{1}{V_{22}} V_{21}. \quad (4.19)$$

4.4.2. CLOSED-FORM TOTAL LEAST SQUARES ESTIMATION

THE SVD-based method discussed in the previous section provides a general solution to the total least squares problem. However, considering the complexity brought by the SVD when processing a tremendous number of samples, a simplified method is required.

Noting the linearity of (4.16) and the fact that the total least squares method minimizes the orthogonal residuals, we can reformulate the TLS cost function as

$$J_{tls} = \frac{\|\Delta \mathbf{P} - \gamma \widehat{\mathbf{L}}\|^2}{1 + \gamma^2}. \quad (4.20)$$

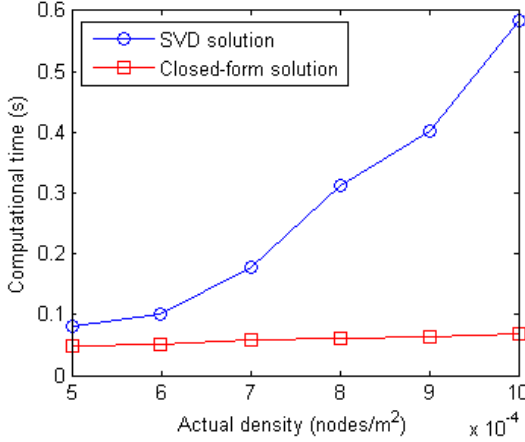


Figure 4.3: The computational time of the traditional solution and the closed-form solution.

By solving

$$\frac{\partial J_{tls}}{\partial \gamma} = \frac{\gamma^2 \hat{\mathbf{L}}^T \Delta \mathbf{P} + \gamma (\hat{\mathbf{L}}^T \hat{\mathbf{L}} - \Delta \mathbf{P}^T \Delta \mathbf{P}) - \hat{\mathbf{L}}^T \Delta \mathbf{P}}{(1 + \gamma^2)^2} = 0, \quad (4.21)$$

we obtain two solutions, which are respectively given by

$$\hat{\gamma}_1 = \eta + \sqrt{1 + \eta^2} > 0 \quad (4.22)$$

and

$$\hat{\gamma}_2 = \eta - \sqrt{1 + \eta^2} < 0, \quad (4.23)$$

where

$$\eta = \frac{\Delta \mathbf{P}^T \Delta \mathbf{P} - \hat{\mathbf{L}}^T \hat{\mathbf{L}}}{2 \hat{\mathbf{L}}^T \Delta \mathbf{P}}.$$

Actually, optimizing (4.20) can also be viewed as finding a linear curve with slope γ through the origin, in which the values of $P_{\hat{i}, \hat{j}}$ and the values of $\hat{L}_{\hat{i}, \hat{j}}$ are respectively on the y-axis and x-axis. Please also refer to [13] for some other total least squares solutions to different modified linear regression models. Therefore, it is evident that two perpendicular curves are obtained, i.e.,

$$\hat{\gamma}_1 \hat{\gamma}_2 = -1.$$

One of the solutions minimizes J_{tls} while the other maximizes it. Considering that $\hat{\gamma}_{tls} > 0$, the total least squares path-loss exponent (TLS-PLE) estimate is obviously given by

$$\hat{\gamma}_{tls} = \hat{\gamma}_1.$$

As far as the computational complexity is concerned, the SVD procedure on $[\hat{\mathbf{L}} \Delta \mathbf{P}]$ requires a complexity of approximately $8N^2$ to obtain \mathbf{U} , Σ and \mathbf{V} [14]. If only \mathbf{V} is required to estimate the PLE, the SVD-based method still has a complexity of approximately $16N$. However, our closed-form solution has only a complexity of approximately $3N$.

Compared with the SVD-based solution, we also measure the average computational time when the transmission range is 200 *m*. The methods are implemented in Matlab 2012b on a Lenovo IdeaPad Y570 Laptop (Processor 2.0GHz Intel Core i7, Memory 8 GB). As shown in Fig. 4.3, the computational time of the closed-form solution is greatly reduced especially when the sample size is increased.

4.4.3. CLOSED-FORM WEIGHTED TOTAL LEAST SQUARES ESTIMATION

FROM the aforementioned analyses, we can conclude that there are three kinds of errors impacting the path-loss exponent estimate:

1. The estimation error $\varepsilon_{i,j}$ on $\hat{L}_{i,j}$ is subject to the spatial dynamics of the node deployment. Therefore, when increasing the actual density, such errors will be decreased.
2. The shadowing effect introduces a Gaussian error $\chi_{i,j}$ which will decrease when the sample size is increased.
3. The last kind of error is the $\Delta\varepsilon_{i,j}$ which represents the mismatch between the ranking numbers of the received powers and the ranges. This kind of error is subject not only to the shadowing but also to the spatial dynamics of the nodes. When the actual density is increased and the nodes get closer to each other, the differences of the received powers become relatively small which leads to a large impact of the shadowing on the ranking.

We propose a weighted total least squares method targeting the suppression of the $\Delta\varepsilon_{i,j}$. Plugging $\hat{i} = i + \Delta i$ and $\hat{j} = j + \Delta j$ into $\Delta\varepsilon_{i,j}$, we have

$$\Delta\varepsilon_{i,j} = \frac{10}{d} \log_{10} \left(\frac{\hat{i}}{\hat{i} - \Delta i} \right) - \frac{10}{d} \log_{10} \left(\frac{\hat{j}}{\hat{j} - \Delta j} \right). \quad (4.24)$$

By using some bounds of the natural logarithm

$$1 - \frac{\hat{i} - \Delta i}{\hat{i}} \leq \ln \left(\frac{\hat{i}}{\hat{i} - \Delta i} \right) \leq \frac{\hat{i}}{\hat{i} - \Delta i} - 1, \quad (4.25)$$

where equality is obtained when $\Delta i = 0$, bounds for $\Delta\varepsilon_{i,j}$ can be computed as

$$\frac{10 \ln(10)}{d} \left(2 - \frac{\hat{i} - \Delta i}{\hat{i}} - \frac{\hat{j}}{\hat{j} - \Delta j} \right) \leq \Delta\varepsilon_{i,j} \leq \frac{10 \ln(10)}{d} \left(\frac{\hat{i}}{\hat{i} - \Delta i} + \frac{\hat{j} - \Delta j}{\hat{j}} - 2 \right). \quad (4.26)$$

Considering that $1 \leq \hat{i} - \Delta i \leq \hat{n}$ and $1 \leq \hat{j} - \Delta j \leq \hat{n}$, we can further bound $\Delta\varepsilon_{i,j}$ as

$$\frac{10 \ln(10)}{d} \left(2 - \hat{i} - \frac{\hat{n}}{\hat{j}} \right) \leq \Delta\varepsilon_{i,j} \leq \frac{10 \ln(10)}{d} \left(\frac{\hat{n}}{\hat{i}} + \hat{j} - 2 \right). \quad (4.27)$$

From (4.27), we can finally find an upper bound of $\Delta\varepsilon_{i,j}^2$ as

$$\Delta\varepsilon_{i,j}^2 \leq \frac{100 \ln(10)^2}{d^2} \max \left\{ \left(\frac{\hat{n}}{\hat{i}} + \hat{j} - 2 \right)^2, \left(\frac{\hat{n}}{\hat{j}} + \hat{i} - 2 \right)^2 \right\}. \quad (4.28)$$

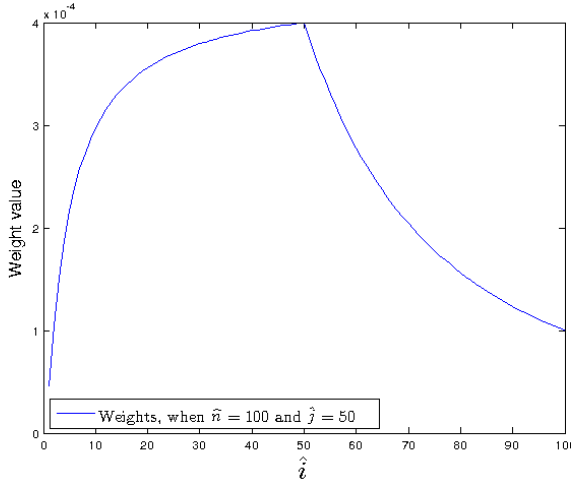


Figure 4.4: The total least squares weights as a function of \hat{i} for $\hat{n} = 100$ and $\hat{j} = 50$.

The idea is now to assign a large weight to a sample with a small upper bound of the mismatch $\Delta \varepsilon_{i,j}^2$. Therefore, based on (4.28), the weights can be constructed by

$$\omega_{i,j} = \frac{1}{\max\{(\frac{\hat{n}}{i} + \hat{j} - 2)^2, (\frac{\hat{n}}{j} + \hat{i} - 2)^2\}}. \quad (4.29)$$

We plot the weights when $\hat{j} = 50$ and $\hat{n} = 100$ in Fig. 4.4.

By stacking the values of $\omega_{i,j}$ on the diagonal of a diagonal matrix in the same way we stack the values of $\Delta P_{i,\hat{j}}$ and the values of $\hat{L}_{i,\hat{j}}$, we construct the $N \times N$ weight matrix \mathbf{W} and then the weighted TLS cost function can be constructed by

$$J_{wtls} = \frac{(\Delta \mathbf{P} - \gamma \hat{\mathbf{L}})^T \mathbf{W} (\Delta \mathbf{P} - \gamma \hat{\mathbf{L}})}{1 + \gamma^2}. \quad (4.30)$$

As before, the closed-form weighted total least squares path-loss exponent (WTLS-PLE) estimate is then easily given by

$$\hat{\gamma}_{wtls} = \eta' + \sqrt{1 + \eta'^2}, \quad (4.31)$$

where $\eta' = \frac{\Delta \mathbf{P}^T \mathbf{W} \Delta \mathbf{P} - \hat{\mathbf{L}}^T \mathbf{W} \hat{\mathbf{L}}}{2 \hat{\mathbf{L}}^T \mathbf{W} \Delta \mathbf{P}}$.

4.4.4. DISCUSSIONS AND FUTURE WORKS

IN this section, we discuss some remaining theoretical problems and some possible issues related to real-life environments. Meanwhile, we cast light on our future works.

CRAMÉR-RAO LOWER BOUND

THE Cramér-Rao lower bound (CRLB) is very difficult to obtain for this problem. This is due to the fact that the estimation accuracy of the PLE is subject to the spatial dynamics, the shadowing and the rank number estimate. They are all mutually related, especially for the ranking number estimate which does not follow any known probability density function (PDF). That is also why we selected a bound on the error to construct the weights in order to suppress the mismatch of the ranking numbers.

In our future work, we are looking for one-step estimation methods which can directly utilize the RSSs without the ranking procedure. To achieve that, a PDF of the RSS in an ad-hoc environment is required, which considers the spatial dynamics and the shadowing. Based on such a PDF, a better estimator, such as the maximum likelihood (ML) estimator, and the CRLB can be introduced.

4

DIFFERENT TRANSMIT POWERS

PREVIOUSLY, we assume the same transmit power P_t for all the neighboring nodes, which might not be so realistic. But assume now that the transmit powers are different. We then have to particularly estimate the transmit power $P_{t,\hat{i}}$ from the \hat{i} -th node to calculate the path-loss $\Delta P_{\hat{i}} := 10 \log_{10}(P_{t,\hat{i}}/P_{r,\hat{i}})$ and further compute the $\Delta P_{\hat{i},\hat{j}} := \Delta P_{\hat{i}} - \Delta P_{\hat{j}}$. Otherwise, if we still compute $\Delta P_{\hat{i},\hat{j}} := 10 \log_{10}(P_{r,\hat{j}}/P_{r,\hat{i}})$, our estimators will become worse yet still feasible. To see that, we firstly need to assume an unknown average transmit power \bar{P}_t and hence use $10 \log_{10}(P_{t,\hat{i}}) = 10 \log_{10}(\bar{P}_t) + \Delta P_{t,\hat{i}}$, where $\Delta P_{t,\hat{i}}$ is the deviation in dB of the transmit power from the \hat{i} -th node. Then (4.9) has to be changed into

$$\Delta P_{\hat{i},\hat{j}} = \gamma(\widehat{L}_{\hat{i},\hat{j}} - \varepsilon_{\hat{i},\hat{j}}) + \chi_{i,j} + \Delta P_{t,\hat{i},\hat{j}}, \quad (4.32)$$

where \bar{P}_t can still be cancelled and $\Delta P_{t,\hat{i},\hat{j}} := \Delta P_{t,\hat{i}} - \Delta P_{t,\hat{j}}$. Obviously, although \mathbf{X} in (4.16) has to become the vector of $\chi_{i,j} + \Delta P_{t,\hat{i},\hat{j}}$ values, our proposed estimators can still estimate the PLE since the general form of (4.16) remains the same.

So if we assume that $\Delta P_{t,\hat{i},\hat{j}}$ is Gaussian distributed, $\chi_{i,j} + \Delta P_{t,\hat{i},\hat{j}}$ is still a zero-mean Gaussian variable, which means that the different transmit powers can equivalently be considered as a more severe shadowing impact. Therefore, for convenience, we still assume the same transmit power in this paper.

DIRECTIONAL PLE ESTIMATION

ANOTHER practical problem is that the PLE sometimes varies over different directions while we previously assume that the PLE is omni-directionally the same. To cope with this problem, we discuss and can extend our proposed estimators with a directional PLE estimation.

As shown in Fig. 4.5, we assume that only the RSSs from the nodes within the angular window ϕ are subject to the same PLE. Hereby in (4.1), W has to become the actual transmission range bounded by the angle ϕ , say W_ϕ , while Ω becomes the corresponding sector Ω_ϕ with radius r . The volume of Ω_ϕ then becomes $\mu(\Omega_\phi) := c_{d,\phi} r^d$, where for $d = 1, 2, 3$ we have $c_{1,\phi} := 1$, $c_{2,\phi} := \phi/2$ and $c_{3,\phi} := \frac{2\pi}{3}(1 - \cos\phi)$. Since the nodes are still

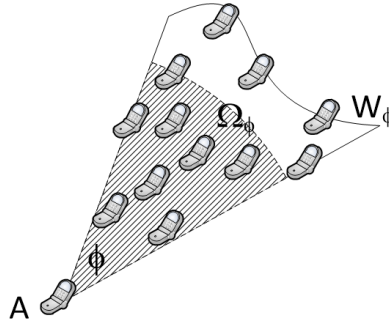


Figure 4.5: The demonstration of the directional PLE estimation in \mathbb{R}^2 : A is the considered node collecting the RSSs from within the angle ϕ . W_ϕ is the actual transmission range bounded by ϕ and the shaded area Ω_ϕ is the corresponding sector with radius r .

randomly deployed within W_ϕ , compared with (4.10), we can hence similarly write

$$\mathbb{P}[i | L_i] = \frac{n!}{i!(n-i)!} \left(\frac{c_{d,\phi} 10^{\frac{dL_i}{10}}}{\mu(W_\phi)} \right)^i \left(1 - \frac{c_{d,\phi} 10^{\frac{dL_i}{10}}}{\mu(W_\phi)} \right)^{n-i}. \quad (4.33)$$

Even though the estimate of L_i has to be changed into

$$\hat{L}_i = \frac{10}{d} \log_{10} \left(\frac{i\mu(W_\phi)}{nc_{d,\phi}} \right), \quad (4.34)$$

the estimate of $L_{i,j}$ however remains the same, i.e., $\hat{L}_{i,j} := \hat{L}_i - \hat{L}_j$, since $\mu(W_\phi)$, n and $c_{d,\phi}$ will be canceled. Therefore, the rest of the theoretical derivation remain the same and our estimators are still feasible.

To achieve a directional PLE estimate, we only have to constrain the RSS sample set within a certain angular window ϕ and our proposed estimators can estimate the PLE for the given direction. Of course to achieve the same accuracy, the directional PLE estimator has to collect more samples than the omni-directional PLE estimator. Again in this paper, for convenience, we still assume the same PLE for all directions.

4.5. SIMULATIONS

IN this section, we simulate our two proposed path-loss exponent estimators in a 2-dimensional space and we leave real-life experiments as future work. Two simulations are conducted to study their performance, with different shadowing impacts and with different actual densities.

We also compare them with the path-loss exponent estimator based on the cardinality of the transmitting set (C-PLE), proposed in [1]. The C-PLE requires changing the receiver's sensitivity from P_{thres1} to P_{thres2} and evaluating the corresponding cardinalities n_1 , n_2 of the transmitting set, namely the different theoretical neighborhood sizes.

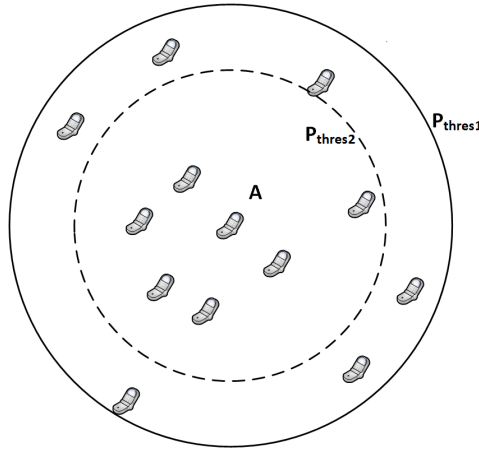


Figure 4.6: Demonstration of the C-PLE estimator: node A changes its receiver's sensitivity from P_{thres1} to P_{thres2} . The solid circle and the dashed circle are respectively the transmission ranges related to P_{thres1} and P_{thres2} . The corresponding neighborhood sizes are $\hat{n}_1 = 12$ and $\hat{n}_2 = 6$ in this figure. The estimated path-loss exponent can be obtained from (4.35).

Thus, considering the shadowing, C-PLE is given in a 2-dimensional space by

$$\hat{\gamma}_c = 2 \ln \left(\frac{P_{thres2}}{P_{thres1}} \right) / \ln \left(\frac{\hat{n}_1}{\hat{n}_2} \right), \quad (4.35)$$

where \hat{n}_1 and \hat{n}_2 are the corresponding actual neighborhood sizes. Fig. 4.6 gives an example of the C-PLE estimator. In our simulations, we set $P_{thres2} = 2P_{thres1}$.

To avoid any border effect, our simulations take place in a very large area, where nodes are randomly deployed. The estimated PLE is only considered for a single node somewhere in the center of the network, rather than for every node in the wireless network. The Monte Carlo method is used to generate the results by repeatedly deploying nodes. The general settings are shown in Table 4.1.

The normalized root mean square error (RMSE) is adopted to present the accuracy of the estimator. In this paper, the normalized RMSE is defined by $\sqrt{\frac{1}{N_{trials}} \sum_{i=1}^{N_{trials}} \left[\frac{\hat{\gamma}(i) - \gamma}{\gamma} \right]^2}$, where N_{trials} is the number of simulation trials, $\hat{\gamma}(i)$ is the estimate of the PLE in the i -th trial, and γ is the actual PLE.

4.5.1. THE IMPACT OF THE SHADOWING

THIS simulation is conducted when the actual density is set as $0.005 \text{ node}/m^2$. Three estimators are studied with an increasing standard deviation of the shadowing and an increasing actual path-loss exponent. Observing Fig. 4.7, we can conclude the following:

1. Our two proposed methods outperform the C-PLE estimator. This can be easily understood from the fact that our methods consider received powers from all

Table 4.1: Values of the parameters used in the simulations.

Parameter	Value
Dimension	$d = 2$
Carrier frequency	2401 MHz
Receiver sensitivity	For TLS-PLE and WTLS-PLE, P_{thres} is adjusted to have a theoretical transmission range of 200 m. For C-PLE, $P_{thres1} = P_{thres}$ and $P_{thres2} = 2P_{thres}$.
Number of trials	100

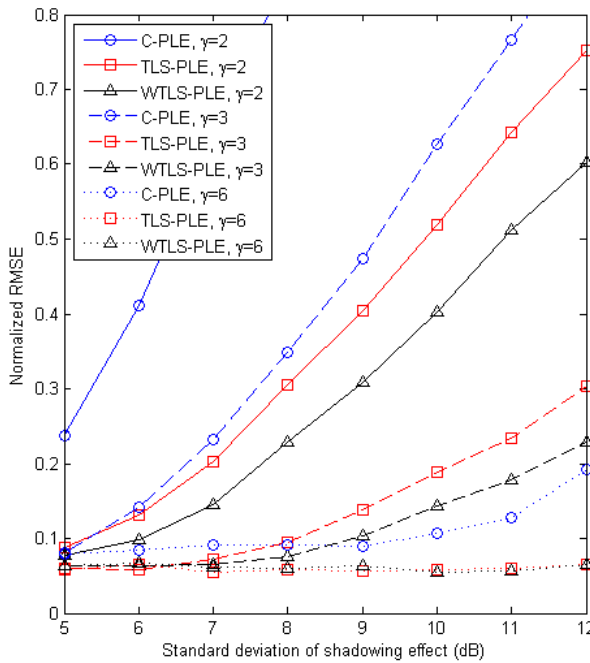


Figure 4.7: The performance of different PLE estimators with an increasing standard deviation of the shadowing.

neighbors rather than only using two neighborhood sizes. Besides, the total least squares procedure helps to minimize the three kinds of errors mentioned earlier.

2. When the shadowing effect becomes more severe, the accuracy of the three esti-

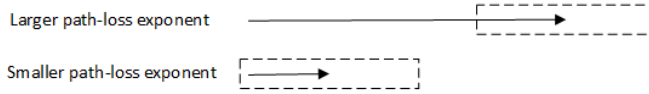


Figure 4.8: The length of the arrow indicates the received signal strength reduction ΔP and the dashed rectangles show the shadowing effect χ . Considering the shadowing means that the arrows can end up anywhere within the rectangles. The width of the rectangle indicates the severity of the shadowing. Under the same transmitter-receiver distance, the arrow with a smaller path-loss exponent is shorter and thus easier to be impacted by the shadowing effect. Therefore, under a high PLE, the matching between the ranking numbers of the received powers and the ranges is not so easily disrupted in the TLS-PLE and the WTLS-PLE. Likewise, the shadowing also becomes more tolerable when estimating the theoretical neighborhood size in the C-PLE.

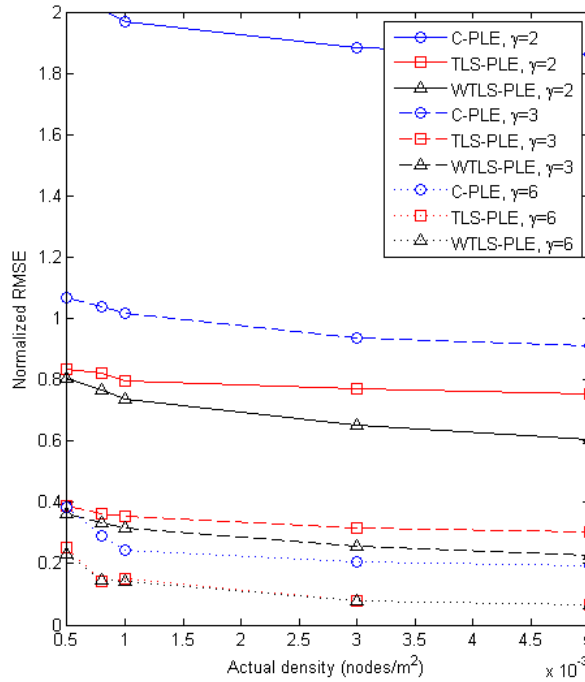


Figure 4.9: The performance of three considered estimators with an increasing actual density.

meters decreases. For the C-PLE, the accuracy mainly depends on the absolute deviation of the actual neighborhood size $|\Delta n| = |\hat{n} - n|$. The shadowing increases such an absolute deviation thus leading to a worse accuracy. For our methods, the shadowing impacts the accuracy by increasing $|\chi_{i,j}|$ and by disrupting the matches between the rank numbers of the received powers and the ranges, i.e, by increasing $|\Delta \varepsilon_{i,j}|$.

3. Surprisingly, the performance of the estimators becomes better in a hasher environment, i.e, when the actual path-loss exponent is high. This is due to the fact

that a high PLE causes relatively large differences between the received powers which makes the shadowing effect more tolerable. It is better explained in Fig. 4.8. To be specific for our methods, when the PLE is small, the accuracy is subject to the three kinds of errors $\chi_{i,j}$, $\varepsilon_{i,j}$ and $\Delta\varepsilon_{i,j}$. However, when the actual PLE is increased, the matches of the rank numbers are more accurate, i.e. $|\Delta\varepsilon_{i,j}|$ decreases.

4. The WTLS-PLE has a better performance than the TLS-PLE, especially under a small PLE. Meanwhile, the improvement of the WTLS-PLE is not so obvious compared with the TLS-PLE when the PLE is high. This is understandable from the fact that the WTLS-PLE is especially targeted at suppressing $\Delta\varepsilon_{i,j}$, the improvement is hence insignificant when $\Delta\varepsilon_{i,j}$ is decreased which has already been pointed out in the previous conclusion.

4.5.2. THE IMPACT OF THE ACTUAL DENSITY

SINCE the estimation error $\varepsilon_{i,j}$ of $\hat{L}_{i,j}$ is related to the actual density, we are interested in how the actual density impacts the accuracy in this section. The transmission range is fixed at 200 m and a 12 dB standard deviation of the shadowing is considered. From Fig. 4.9, we can see that, compared with the impact of the shadowing, the impact of the actual density is relatively small. Additionally, when more samples are collected, the WTLS-PLE has a larger improvement on the accuracy by suppressing $\Delta\varepsilon_{i,j}$.

4.6. APPLICATIONS

THE path-loss exponent plays a very significant role in many kinds of wireless networks. Due to the difficulties to locally and solely estimate the path-loss exponent though, only a few techniques are able to utilize path-loss exponent measurements in their designs. However, the proposed path-loss exponent estimation approaches tackle such issues. In this section, we detail some applications and discuss the significance of our path-loss exponent self-estimation schemes.

4.6.1. SECURE RSS-BASED LOCALIZATION

DUE to our PLE self-estimation schemes, either the reference node or the target node can solely and independently estimate the PLE. Therefore, an adversary cannot launch an attack on the PLE estimation by spoofing. For instance, as shown in Fig. 4.10, even if there is a cheating reference node maliciously reporting its fake location, e.g., *attacker C* registering itself at *fake C*, the PLE can still be estimated accurately. Besides making the RSS-based localization more robust to the spoofing attack, this also enables every node to detect and locate the cheating reference node.

STRATEGY FOR DETECTING CHEATING REFERENCE NODES

TO explicitly illustrate the strategy, we firstly explain each one's role and the detection algorithm will be described afterward:

- Each *reference node* knows its own location and is skeptical to any reported location from the other reference nodes.

- It periodically broadcasts its own location and self-estimates the PLE simultaneously.
 - It keeps listening to the messages broadcasted by the other reference nodes, reading the RSSs and their corresponding reported locations.
 - It detects the attackers according to the self-estimated PLE, the RSSs, the reported locations and its own location. The detection algorithm will be discussed later. As soon as an attacker is detected, it will announce the detection as well as the corresponding RSS from the attacker by broadcasting.
 - In case some cheating reference nodes spoof the attacker announcement, an announced attacker needs to be further confirmed as a true attacker. To be confirmed as a true attacker, the announced attacker has to be announced more than T times, where T depends on the total number of reference nodes and the detection sensitivity. When the announced attacker is confirmed as a true attacker, the corresponding announced RSSs from the attacker at least $d + 1$ different reference nodes can further be used to locate the attacker.
- Each *target node* only listens and is invisible to the other nodes.
 - It keeps listening to all information broadcasted by the reference nodes. In the meantime, the PLE is self-estimated.
 - It discovers the true attackers from the message broadcasted by the reference nodes and discards the RSSs from the true attackers.
 - Then, it can accurately and safely locate itself with the rest of the RSSs.

ALGORITHM FOR DETECTING CHEATING REFERENCE NODES

To complete the strategy, the algorithm for detecting the cheating reference nodes is essential. For an explicit demonstration, an example is shown in Fig. 4.10. Let us denote the locations of *reference A*, *reference B*, *attacker C*, *fake C* and *target D* respectively as \mathbf{s}_A , \mathbf{s}_B , \mathbf{s}_C , $\mathbf{s}_{C'}$ and \mathbf{s}_D . To detect *attacker C*, we need to test two hypotheses, which are respectively defined as

$$\mathcal{H}_0 : \mathbf{s}_C \text{ and } \mathbf{s}_{C'} \text{ are the same location.} \quad (4.36)$$

and

$$\mathcal{H}_1 : \mathbf{s}_C \text{ and } \mathbf{s}_{C'} \text{ are different locations.} \quad (4.37)$$

The detection algorithm can be carried out with the following procedure:

1. Firstly, a reference RSS from the suspected reference node needs to be calculated based on the self-estimated PLE, the reported location and the own location of the detecting node. For example, recalling the definition of RSS, the reference RSS at *reference B* from *attacker C* can be calculated in dB as

$$P'_{r,C'B} = C_3 - 10\hat{\gamma}_B \log_{10}(\|\mathbf{s}_{C'} - \mathbf{s}_B\|), \quad (4.38)$$

where

$$C_3 = 10 \log_{10}(P_t) + 10 \log_{10}(C_1) + 10\hat{\gamma}_B \log_{10}(r_0)$$

and $\hat{\gamma}_B$ is the self-estimated PLE at \mathbf{s}_B .

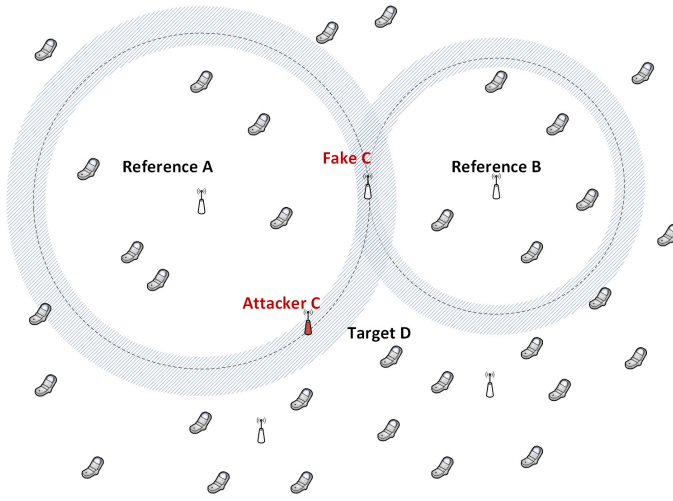


Figure 4.10: *Attacker C* reports its fake location at *fake C*. Both reference nodes like *reference A* and *reference B* and the target node *target D* can self-estimate the PLE. Based on the self-estimated PLE and the location information, the shaded area can be constructed as the trust region for detecting an attacker, outside which *attacker C* will be detected.

- Secondly, the actual RSSs from the suspected reference node are recorded over time to construct our observation set by subtracting the reference RSS. For example, *reference B* records the observation at time i , which is given by

$$\Delta P_{r,CB}^i = P_{r,CB}^i - P_{r,C'B}^i, \quad (4.39)$$

where $P_{r,CB}^i$ is the actual RSS in dB at time i from *attacker C* and $\Delta P_{r,CB}^i \sim \mathcal{N}(\mu_B, \sigma^2)$. If *attacker C* and *fake C* have the same range, then $\mu_B = 0$, otherwise $\mu_B \neq 0$.

Since only the range can be tested, we need two different hypotheses for range testing, which are given by

$$\mathcal{H}'_0 : \mu_B = 0. \quad (4.40)$$

and

$$\mathcal{H}'_1 : \mu_B \neq 0. \quad (4.41)$$

Considering the fact that *attacker C* and *fake C* might also have the same range to a reference node, e.g., to *reference A* in Fig. 4.10, we hence have the relations $\mathcal{H}_0 \subset \mathcal{H}'_0$ and $\mathcal{H}'_1 \subset \mathcal{H}_1$. This means if \mathcal{H}'_1 is tested, *attacker C* is certainly detected while if \mathcal{H}'_0 is tested, we might fail to detect the attacker. But, we now focus on testing \mathcal{H}'_1 and the detection failure in \mathcal{H}'_0 will be discussed later.

- Finally, by using the Neyman-Pearson lemma [15], \mathcal{H}'_1 can be tested from the average observation over I time slots. For example, the observation at *reference B* is given by $\rho = (\sum_{i=1}^I \Delta P_{r,CB}^i) / I$. If we wish to test at 95% accuracy, the critical region

for the observation is given by

$$C = \{(P_{r,CB}^1, P_{r,CB}^2, \dots, P_{r,CB}^I) : \rho \leq -1.96\sigma/\sqrt{I}, \rho \geq 1.96\sigma/\sqrt{I}\} \quad (4.42)$$

Equivalently, we can also use the critical region

$$C = \{(P_{r,CB}^1, P_{r,CB}^2, \dots, P_{r,CB}^I) : \rho^2 \geq 3.84\sigma^2/I\}, \quad (4.43)$$

which considers the Chi-squared distribution with ρ^2 as observation.

DISCUSSIONS

1. The shadowing deviation σ is required for the Neyman-Pearson test, which can be obtained by empirical training.
2. The detection failure in \mathcal{H}'_0 can easily be noticed when reference nodes work in a cooperative fashion according to the detection strategy. Since every reference node detects and announces the attackers, such a detection failure can be somehow corrected by listening to the announced information flooding in the network. Therefore, the detection algorithm can be improved, by introducing a new cooperative algorithm. For example, according to the observations from multiple nodes, an attacker can still be detected even if such a detection failure in \mathcal{H}'_0 occurs.
3. Considering the shadowing, the complement of the critical region corresponds to a trust region of the detecting node in space, in which the detected node will be trusted. As shown in Fig. 4.10, two shaded areas respectively indicate the trust regions of *reference A* and *reference B*. *Attacker C* resides outside the trust region of *reference B* but inside that of *reference A*. Therefore, *attacker C* will be detected by *reference B* but not by *reference A*. The size of the trust region depends on the severity of the shadowing.
4. The cheating node can also jeopardize this system by maliciously announcing a credible reference node as an attacker. In most cases, the credible reference nodes outnumber the attackers. Hence, the attackers can still be smartly distinguished. However, if the attackers have the majority, a more robust strategy might be required.

4.6.2. ENERGY-EFFICIENT ROUTING

SINCE the path loss over a channel exponentially increases with the distance, multi-hop communications becomes a better option than single-hop to prolong the network lifetime. Routing is hence aimed at finding an efficient path to the destination in order to minimize the power consumption. It is well-known that a routing path is better to be chosen through an area where the PLE is small. But alternatively, in this section, we consider the k -th nearest neighbor routing protocol to illustrate the significance of the PLE.

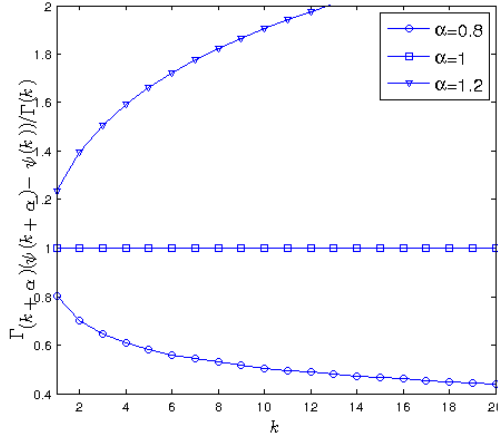


Figure 4.11: The efficiency of single-hop communications: a small value indicates a smaller power cost by increasing k , i.e., a high power efficiency.

From (4.1), if considering the local random region W around the considered node A as a d -dimensional ball of radius R , i.e. $\mu(W) = c_d R^d$, the distribution of the distance r_k to the k -th nearest neighbor is given by [16]

$$\mathbb{P}(r_k|k) = \frac{d}{r_k B(n-k+1, k)} \left(\frac{r_k^d}{R^d} \right)^k \left(1 - \frac{r_k^d}{R^d} \right)^{n-k}, \quad (4.44)$$

where $B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the beta function. To avoid the singularity issue of (4.3), the received power at the k -th nearest neighbor can also be given by

$$P_{r,k} = P_{r,0} \left(\frac{r_0}{r_k} \right)^{\gamma_A} \quad (4.45)$$

where $P_{r,0}$ is the received power at the reference distance $r_0 < r_k, \forall k$ and γ_A is the PLE at the location of node A . Let us denote the path-loss to the k -th nearest neighbor as $\mathfrak{L}_k := \frac{P_{r,0}}{P_{r,k}} = \frac{r_k^{\gamma_A}}{r_0^{\gamma_A}}$. We commonly assume $r_0 = 1$ m and thus $\mathfrak{L}_k := r_k^{\gamma_A}$. From (4.44), we can obtain the expectation of \mathfrak{L}_k for a single hop to the k -th nearest neighbor, which can be given by

$$\begin{aligned} E(\mathfrak{L}_k) &= \frac{R^{\gamma_A} B(k + \gamma_A/d, n - k + 1)}{B(n - k + 1, k)} \\ &= \frac{R^{\gamma_A} \Gamma(n + 1)}{\Gamma(n + \gamma_A/d + 1)} \frac{\Gamma(k + \gamma_A/d)}{\Gamma(k)}. \end{aligned} \quad (4.46)$$

From (4.46), we especially focus on $\frac{\partial E(\mathfrak{L}_k)}{\partial k}$ to study the efficiency of increasing k , which

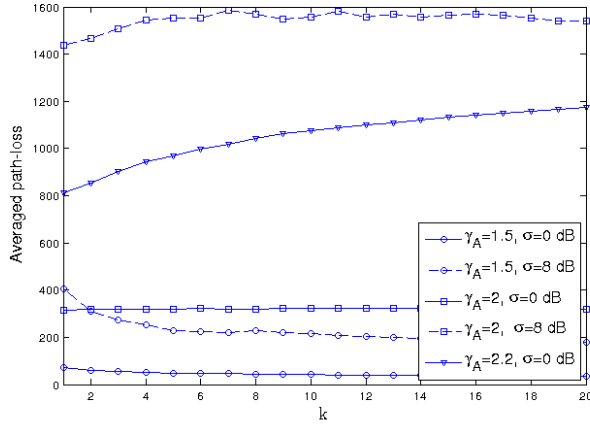


Figure 4.12: The numerical results of the k -th nearest neighbor routing in a 2-dimensional space.

is given by

$$\frac{\partial E(\mathcal{L}_k)}{\partial k} = \frac{R^{\gamma_A} \Gamma(n+1)}{\Gamma(n + \frac{\gamma_A}{d} + 1)} \frac{\Gamma(k + \frac{\gamma_A}{d})(\psi(k + \frac{\gamma_A}{d}) - \psi(k))}{\Gamma(k)}, \quad (4.47)$$

where $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is the ploggamma function. We denote $\alpha = \gamma_A/d$ and plot the k -related part of (4.47), i.e. $f(k) = \frac{\Gamma(k+\alpha)(\psi(k+\alpha)-\psi(k))}{\Gamma(k)}$ in Fig. 4.11. When $\alpha < 1$, $\frac{\partial E(\mathcal{L}_k)}{\partial k}$ decreases with k , which means that it takes less extra power every time k is increased. As a conclusion, a single long-hop communication link is more energy-efficient as long as $\gamma_A < d$, which is also briefly pointed out in [16].

To be more realistic, we also conduct a numerical simulation for the k -th nearest neighbor routing, in which the shadowing effect is also considered. We introduce the average path-loss for a single link, denoted by $\bar{\mathcal{L}}_k = \mathcal{L}_k/k$. A 2-dimensional space is considered with a density of $0.001 \text{ nodes}/m^2$. As is shown in Fig. 4.12, as long as the PLE is smaller than 2, the average path-loss decreases with k and a single long-hop link becomes energy efficient. Additionally, in the presence of log-normal shadowing, $\bar{\mathcal{L}}_k$ becomes larger than when there is no shadowing. Such an increase also becomes larger with a large PLE.

Many other interesting results have been obtained. However, due to the limited space, no more tautology will be presented. It is already evident that the efficiency of the k -th nearest neighbor routing protocol highly relies on the actual PLE. Therefore, the principles for designing such a routing protocol should involve the PLE estimation. In a nutshell, an accurate estimate of γ_A is hence necessary for designing an efficient routing protocol.

4.6.3. OTHER APPLICATIONS

To further illustrate some applications of the proposed PLE estimators, we need to explicitly explain how the PLE affects the network operation. The PLE has a multidimensional

dimensional effect on the performance of the whole system for wireless networking:

1. It determines the quality of the signals at the receivers and thus impacts the physical (PHY) layer. This is because the PLE controls not only the received signal strength but also the interference the nodes create for the other receivers. Since the signal-to-interference-plus-noise ratio (SINR) is decisive for the channel capacity and the performance of decoding, the PLE is essential for designing the PHY layer.
2. It determines the transmission range and thus impacts the network (NET) layer and the media access control (MAC) layer. The transmission range, together with the neighborhood size, which is also determined by the PLE, affects the performance of routing and the connectivity in the NET layer. When the number of nodes within the transmission range of a node increases, the contention in the MAC layer consequently becomes more severe and thus congestion of the network will occur. As a consequence, the ability of delivering the packet will be affected.
3. It determines the energy consumption for transmission links and thus impacts the lifetime of networking. In order to guarantee the efficiency of wireless networking, the transmit power should be smartly controlled to compensate for the energy loss on the transmission links. Considering the battery is strictly limited in e.g. wireless sensor networks, the PLE is rather significant to those protocols aiming at prolonging the network lifetime.

Based on the above mentioned reasons, some other applications can be listed:

1. *Relay nodes* are recently drawing much attention [17] and the mobile ones are even more flexible and more convenient [18]. Since the path-loss exponent is one of the key criteria for energy-efficient routing, relay nodes should be deployed or move to the place where the path-loss exponent is relatively small. The relay nodes can also benefit from the low PLE location to save the battery. Therefore, relay nodes have to be able to estimate the PLE.
2. *Energy harvesting* relies on ambient sources such as solar, wind and kinetic activities, aiming at prolonging the network lifetime. Particularly, among those sources, radio-frequency (RF) signals, can also be used to charge the battery of wireless sensors [19]. Its application is also extended to cognitive radio in [20]. The PLE directly determines the efficiency of harvesting and the size of the harvesting zone. The time slot for harvesting could be adaptive according to the PLE changes. Therefore, the PLE estimation is very significant when the surrounding communication environment is changing or the harvesting node is mobile.
3. *Power control* requires distributedly choosing an appropriate transmit power for each packet at each node. This is because of the fact that the transmit power affects the wireless networking in the same way as the PLE does [21]. Since the PLE is different at different locations, an efficient power control scheme also needs to distributedly and locally consider the PLE. Therefore, our proposed estimators can be integrated into power control to yield a better performance.

4.7. CONCLUSIONS

TWO self-estimators for the path-loss exponent are proposed in this paper, in which each node can solely and locally estimate the path-loss exponent merely by collecting the received signal strengths. They rely neither on external auxiliary systems nor on any information of the wireless network. Their simplicity makes them feasible for any kind of wireless network.

In order to better describe our estimators, a new linear regression model for the path-loss exponent has been introduced. Our closed-form total least squares method can solve this linear regression model. Compared with the SVD-based solution, our estimator tremendously saves computational time. Moreover, a weighted total least squares method is also designed to better suppress the estimation errors.

Simulations present the accuracy of our estimators and demonstrate that the shadowing effect dominantly influences the estimation error. By analyzing the performance of the estimators, it is interesting to observe that the estimators work better in harsh communication environments, where the path-loss exponent is high.

We have also discussed the significance of our PLE self-estimators by illustrating some potential applications and have brought the dawn to some relevant future researches.

REFERENCES

- [1] S. Srinivasa and M. Haenggi, *Path Loss Exponent Estimation in Large Wireless Networks*, in *Information Theory and Applications Workshop* (2009) pp. 124–129.
- [2] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, *Locating the Nodes: Cooperative Localization in Wireless Sensor Networks*, *Signal Processing Magazine, IEEE* **22**, 54 (2005).
- [3] G. Mao, B. D. O. Anderson, and B. Fidan, *Path Loss Exponent Estimation for Wireless Sensor Network Localization*, *Comput. Netw.* **51**, 2467 (2007).
- [4] N. Salman, M. Ghogho, and A. Kemp, *On the Joint Estimation of the RSS-Based Location and Path-loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 34 (2012).
- [5] N. Salman, A. Kemp, and M. Ghogho, *Low Complexity Joint Estimation of Location and Path-Loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 364 (2012).
- [6] X. Li, *RSS-Based Location Estimation with Unknown Pathloss Model*, *Wireless Communications, IEEE Transactions on* **5**, 3626 (2006).
- [7] M. Gholami, R. Vaghefi, and E. Strom, *Rss-based sensor localization in the presence of unknown channel parameters*, *Signal Processing, IEEE Transactions on* **61**, 3752 (2013).
- [8] G. Wang, H. Chen, Y. Li, and M. Jin, *On Received-Signal-Strength Based Localization with Unknown Transmit Power and Path Loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 536 (2012).

- [9] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001).
- [10] N. Nakagami, *The m-distribution, a general formula for intensity distribution of rapid fading*, in *Statistical Methods in Radio Wave Propagation*, edited by W. G. Hoffman (Oxford, England: Pergamon, 1960).
- [11] I. Markovsky and S. Van Huffel, *Overview of total least-squares methods*, [Signal Process.](#) **87**, 2283 (2007).
- [12] C. Eckart and G. Young, *The Approximation of One Matrix by Another of Lower Rank*, [Psychometrika](#) **1**, 211 (1936).
- [13] I. Petras and D. Bednarova, *Total Least Squares Approach to Modeling: A Matlab Toolbox*, *Acta Montanistica Slovaca* **15**, 158 (2010).
- [14] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)* (Johns Hopkins University Press, Baltimore, MD, USA, 1996).
- [15] P. Hoel, S. Port, and C. Stone, *Introduction to statistical theory*, Houghton Mifflin research series (Houghton-Mifflin, 1971).
- [16] S. Srinivasa and M. Haenggi, *Distance Distributions in Finite Uniformly Random Networks: Theory and Applications*, [Vehicular Technology, IEEE Transactions on](#) **59**, 940 (2010).
- [17] S. Misra, S. D. Hong, G. Xue, and J. Tang, *Constrained Relay Node Placement in Wireless Sensor Networks: Formulation and Approximations*, [Networking, IEEE/ACM Transactions on](#) **18**, 434 (2010).
- [18] A. Venkateswaran, V. Sarangan, T. La Porta, and R. Acharya, *A Mobility-Prediction-Based Relay Deployment Framework for Conserving Power in MANETs*, [Mobile Computing, IEEE Transactions on](#) **8**, 750 (2009).
- [19] T. Le, K. Mayaram, and T. Fiez, *Efficient Far-Field Radio Frequency Energy Harvesting for Passively Powered Sensor Networks*, [Solid-State Circuits, IEEE Journal of](#) **43**, 1287 (2008).
- [20] S. Lee, R. Zhang, and K. Huang, *Opportunistic wireless energy harvesting in cognitive radio networks*, [Wireless Communications, IEEE Transactions on](#) **12**, 4788 (2013).
- [21] V. Kawadia and P. Kumar, *Principles and protocols for power control in wireless ad hoc networks*, [Selected Areas in Communications, IEEE Journal on](#) **23**, 76 (2005).

5

DIRECTIONAL MAXIMUM LIKELIHOOD SELF-ESTIMATION OF THE PATH-LOSS EXPONENT

Yongchang HU and Geert LEUS

We must accept finite disappointment, but we must never lose infinite hope.

Martin Luther King

The path-loss exponent (PLE) is a key parameter in wireless propagation channels. Therefore, obtaining the knowledge of the PLE is rather significant for assisting wireless communications and networking to achieve a better performance. Most existing methods for estimating the PLE not only require nodes with known locations but also assume an omnidirectional PLE. However, the location information might be unavailable or unreliable and, in practice, the PLE might change with the direction.

In this paper, we are the first to introduce two directional maximum likelihood (ML) self-estimators for the PLE in wireless networks. They can individually estimate the PLE in any direction merely by locally collecting the related received signal strength (RSS) measurements. The corresponding Cramér-Rao lower bound (CRLB) is also obtained. Simulation results show that the performance of the proposed estimators is very close to the CRLB. Additionally, also for the first time, the RSSs based on only a geometric path loss are found to follow a truncated Pareto distribution in wireless random networks. This might be of great help in the analysis of wireless communications and networking.

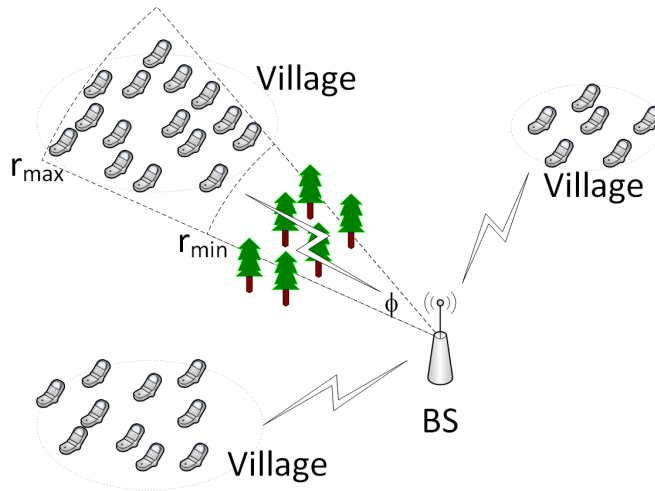


Figure 5.1: Example in \mathbb{R}^2 based on a spherical coordinate system: mobile users cluster in different separate villages, but remotely connect to the BS; the PLEs to different villages are often different (e.g., some forests might result in a large PLE).

5

5.1. INTRODUCTION

THE path-loss exponent (PLE) is very crucial for efficiently designing wireless communications and networking systems. For instance, the information-theoretic capacity of large ad hoc networks highly depends on the PLE, which might lead to different routing strategies [1]. Source localization based on the received signal strength (RSS) measurements requires the knowledge of the PLE to estimate the target location [2]. The interference in wireless ad hoc networks is greatly affected by the PLE [3], which has a strong impact on the quality of the transmission link. Therefore, the PLE needs to be accurately estimated.

Current methods for estimating the PLE can mostly be found in the field of RSS-based localization [4–6], where some nodes with known locations, i.e., anchors, are required. However, the anchors are sometimes not available and the location information might also be unreliable, especially in military scenarios where adversaries can maliciously sabotage wireless networks by spoofing some specific information. Some other estimators of the PLE are presented in [7], which however require the network density or the receiver sensitivity, and even require changing them. Besides, in practice, the PLE might change with direction, as depicted in Fig. 5.1. Yet, all existing estimators assume the PLE to be omni-directionally the same.

Therefore, a directional self-estimator, which can solely and individually estimate the PLE merely by locally collecting the RSSs, is urgently required. Driven by this motivation, two (weighted) total least squares self-estimators of this kind have already been proposed in our previous work [8]. However, obtaining the maximum likelihood (ML) solution and the Cramér-Rao lower bound (CRLB) still remains a problem. Additionally, our previous work assumes a homogeneously random deployment of the surrounding

nodes, i.e., a homogeneous random network (HRN), and such an assumption can easily be violated, when nodes are clustered. For example, considering cellular networks, the mobile users tend to cluster in different villages with different PLEs, as shown in Fig. 5.1. In this case, all the aforementioned estimators become unfeasible and certainly cannot estimate the PLE to every village. A possible solution is to consider the mobile users in each village to be locally randomly deployed, i.e., those villages are viewed as several locally random networks (LRNs), which remotely connect to the base station (BS). Then, this issue can be well-resolved if we propose a self-estimator for the PLE based on the RSS measurements from those LRNs. In fact, if considering a spherical coordinate system, the LRN is more general (and includes the HRN), thus leading to more general solutions.

The contributions of this paper can be listed as follows:

1. The RSSs based only a geometric path-loss are first found to follow a truncated *Pareto* distribution. This is derived using properties of LRNs. This finding might be of great help in the analysis of wireless communications and networking.
2. Based on the RSS distribution, two ML self-estimators of the PLE are derived, which meet the mentioned requirements. Further, the CRLB for this kind of estimator is computed.
3. The two proposed ML self-estimators are both close to the CRLB. For comparison, we especially consider the case of a HRN and both our estimators outperform two existing ones: one weighted total least squares estimator from our previous work (WTLS-PLE) and another estimator based on the cardinality of the transmitting set (C-PLE).

The rest of the paper is organized as follows. Section 5.2 introduces the truncated *Pareto* distribution for the RSS in the wireless random networks. Based on this RSS distribution, the CRLB and the proposed directional ML self-estimators for the PLE are presented in Section 5.3. Numerical results are shown in Section 5.4 and finally we conclude this paper in Section 5.6.

5.2. RSS DISTRIBUTION IN WIRELESS RANDOM NETWORKS

SINCE the self-estimation of the PLE only relies on the RSS measurements, it is obviously significant to obtain the RSS distribution in wireless random networks. However, this has never been studied before, to the best of our knowledge. If this distribution can be found, it might not only help to obtain the CRLB as well as the ML solution for the self-estimation of the PLE, but also lead to other insightful properties of wireless networks.

To begin, we first have to study the distribution of the nodal distance r for a random node deployment. Two distributions for ordered nodal distances were already given in [9, 10]. However, they were limited to (infinite) HRNs. Therefore, for the remote LRNs depicted in Fig. 5.1, we actually need a more general distribution.

A random deployment of nodes implies that every node holds an equal chance ρ to reside in a considered area. Therefore, if all the nodes are bounded by an LRN in \mathbb{R}^m , we

can obtain

$$\rho = 1/(c_{m,\phi} r_{max}^m - c_{m,\phi} r_{min}^m),$$

where ϕ is the angular window, r_{min} and r_{max} are considered the smallest and the largest nodal distances, and for $m = 1, 2, 3$ we have $c_{1,\phi} = 1$, $c_{2,\phi} = \phi/2$ and $c_{3,\phi} = \frac{2\pi}{3}(1 - \cos\phi)$. Therefore, the cumulative density function (CDF) of r is given by

$$\mathbb{F}(r) = \rho c_{m,\phi} (r^m - r_{min}^m) = \frac{r^m - r_{min}^m}{r_{max}^m - r_{min}^m}, \quad (5.1)$$

for $r \in [r_{min}, r_{max})$.

and hence the probability density function (PDF) of r can be obtained as

$$\mathbb{P}(r) = \frac{\partial \mathbb{F}(r)}{\partial r} = \frac{m r^{m-1}}{r_{max}^m - r_{min}^m}, \quad \text{for } r \in [r_{min}, r_{max}). \quad (5.2)$$

Then, for the wireless propagation channel, we currently only consider the geometric path loss[11], i.e., the RSS can be presented (in *Watt*) by

$$P_r = Cr^{-\gamma}, \quad (5.3)$$

where γ is the PLE and $C \triangleq G_t G_r P_t$ with G_t the transmitter antenna gain, G_r the receiver antenna gain and P_t the transmit power. Admittedly, the *shadowing* effect is very important, yet considering it will complicate the following derivations. Besides, the proposed ML solutions are also very resilient to the *shadowing* effect if considered, which will be discussed later on.

One may also consider the *small-scale* fading, which mainly decides the instantaneous received power. In fact, the instantaneous received signal envelope follows a *Nakagami* distribution [12] and accordingly the distribution of the instantaneous received power p follows a *Gamma* distribution, which is given by

$$\mathbb{P}(p) = \frac{(\frac{d}{E(p)})^d p^{d-1} e^{-\frac{dp}{E(p)}}}{\Gamma(d)}, \quad (5.4)$$

where d is the fading parameter and a small value of d indicates a stronger fading. Precisely speaking, the *small-scale fading* just causes the instantaneous power p to rapidly fluctuate within a very small scale around the expectation that is determined by the RSS, i.e., $E(p) = P_r$. Therefore, compared with the geometric path-loss, the impact of *small-scale fading* is relatively small. In practice, the RSS P_r is obtained by taking the average over K consecutive time slots of instantaneous received powers p_k , i.e. $P_r = \frac{1}{K} \sum_{i=1}^K p_k$. From (5.4), we have $Var(P_r) = \frac{E(p_k)^2}{Kd}$, which implies that, when K is large enough, the impact of the small-scale fading almost vanishes. Therefore, the term “received signal strength (RSS)” does not consider the *small-scale fading*, i.e., the RSS in this paper refers to P_r .

Obviously, the geometric path-loss in (5.3) follows the Zipf’s law, which enlightens us that, in this case, the RSS in wireless random networks might be subject to one of

$$\begin{aligned} \mathcal{J}(\gamma) = & -\frac{n}{\gamma^2} - \frac{2mnl\ln(P_{r,\min})}{\gamma^3} + \frac{2n[(\gamma + m\ln(P_{r,\max}))(\frac{P_{r,\min}}{P_{r,\max}})^{\frac{m}{\gamma}} - (\gamma + m\ln(P_{r,\min}))]}{\gamma^3 \left(\left(\frac{P_{r,\min}}{P_{r,\max}}\right)^{\frac{m}{\gamma}} - 1\right)} \\ & + \frac{nm\left(\frac{P_{r,\min}}{P_{r,\max}}\right)^{\frac{m}{\gamma}} \ln\left(\frac{P_{r,\min}}{P_{r,\max}}\right) [2\gamma\left(\frac{P_{r,\min}}{P_{r,\max}}\right)^{\frac{m}{\gamma}} - 2\gamma - m\ln\left(\frac{P_{r,\min}}{P_{r,\max}}\right)]}{\left(1 - \left(\frac{P_{r,\min}}{P_{r,\max}}\right)^{\frac{m}{\gamma}}\right)^2 \gamma^4} \end{aligned} \quad (5.7)$$

the power-law distributions [13], e.g., the *Pareto* distribution, but this has never been observed before. Note that this kind of distribution has rather wide applications in research on the city population [14], the sizes of earthquakes [15], etc., yet so far not in the field of wireless networks.

Based on (5.1) and (5.3), the CDF of the RSS can be obtained after a simple transformation of variables as

$$\begin{aligned} \mathbb{F}(P_r|m, \gamma, P_{r,\min}, P_{r,\max}) \\ = \begin{cases} \frac{1 - (P_{r,\min}/P_r)^{m/\gamma}}{1 - (P_{r,\min}/P_{r,\max})^{m/\gamma}}, & \text{for } P_{r,\min} \leq P_r \leq P_{r,\max}, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \quad (5.5)$$

where $P_{r,\min} \triangleq Cr_{\max}^{-\gamma}$, and $P_{r,\max} \triangleq Cr_{\min}^{-\gamma}$ in the LRN ($r_{\min} \gg 0$), or $P_{r,\max} \triangleq P_t$ in the HRN to avoid the singularity issue in (5.3). And, the PDF can finally be obtained as

$$\begin{aligned} \mathbb{P}(P_r|m, \gamma, P_{r,\min}, P_{r,\max}) &= \frac{\partial \mathbb{F}(P_r|m, \gamma, P_{r,\min}, P_{r,\max})}{\partial P_r} \\ &= \begin{cases} \frac{m}{\gamma} \frac{P_{r,\min}^{m/\gamma} P_r^{-m/\gamma-1}}{1 - (P_{r,\min}/P_{r,\max})^{m/\gamma}}, & \text{for } P_{r,\min} \leq P_r \leq P_{r,\max}, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \quad (5.6)$$

which apparently follows a truncated *Pareto* distribution Type I [16].

5.3. DIRECTIONAL MAXIMUM LIKELIHOOD SELF-ESTIMATION OF THE PLE

AFTER obtaining the distribution for the RSS measurements, we can introduce the CRLB for the self-estimation of the PLE and our proposed ML solutions.

5.3.1. CRLB

IF n RSS samples are locally collected from an LRN, where the i -th sample is denoted by P_i , the truncated *Pareto* distribution (5.6) directly leads to the CRLB for the self-estimation of the PLE, which can be given by

$$\text{CRLB}(\gamma) = \frac{1}{\mathcal{J}(\gamma)},$$

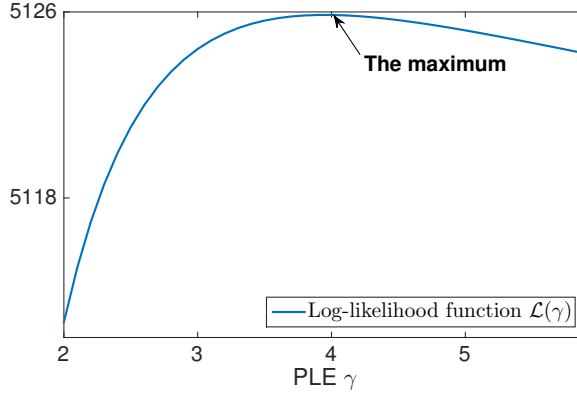


Figure 5.2: Demonstration of the convexity of the log-likelihood function $\mathcal{L}(\gamma)$, when the PLE is set to 4.

5

where

$$\mathcal{I}(\gamma) = -E \left[\sum_{i=1}^n \frac{\partial^2 \ln(\mathbb{P}(P_i | m, \gamma, P_{r,min}, P_{r,max}))}{\partial \gamma^2} \right]$$

is the Fisher information shown in (5.7) on the top of page 143. As shown in Fig. 5.3(a), the CRLB decreases with a large sample size or a small PLE. We also notice that, the farther the LRN is located from the considered node, the larger the CRLB becomes.

5.3.2. TWO ML SELF-ESTIMATORS FOR THE PLE

NOW, let us focus on the ML solution to the self-estimation of the PLE. Based on the truncated *Pareto* distribution in (5.6), the log-likelihood function can be expressed as

$$\begin{aligned} \mathcal{L}(\gamma) &= \sum_{i=1}^n \ln(\mathbb{P}(P_i | m, \gamma, P_{r,min}, P_{r,max})) \\ &= n \ln\left(\frac{m}{\gamma}\right) + \frac{nm}{\gamma} \ln(P_{r,min}) - \left(\frac{m}{\gamma} + 1\right) \sum_{i=1}^n \ln(P_i) \\ &\quad - n \ln(1 - (P_{r,min}/P_{r,max})^{m/\gamma}), \end{aligned} \tag{5.8}$$

which is required to be convex to facilitate the proposed ML estimators. To prove that, the derivative of $\mathcal{L}(\gamma)$ should be a strictly decreasing function. For convenience, it would also be sufficient to prove the monotonicity of

$$f(\gamma) \triangleq \gamma^2 \frac{\partial \mathcal{L}(\gamma)}{\partial \gamma}, \forall \gamma: \gamma > 0,$$

which is what we will do next. Considering any two values of γ that satisfy $\forall \gamma_1, \gamma_2: \gamma_1 > \gamma_2 > 0$, we have

$$f(\gamma_1) - f(\gamma_2) = -n \left(\gamma_1 - \gamma_2 + \frac{\gamma_1 \ln(t^{\frac{m}{\gamma_1}}) t^{\frac{m}{\gamma_1}}}{1 - t^{\frac{m}{\gamma_1}}} - \frac{\gamma_2 \ln(t^{\frac{m}{\gamma_2}}) t^{\frac{m}{\gamma_2}}}{1 - t^{\frac{m}{\gamma_2}}} \right), \tag{5.9}$$

where $t \triangleq \frac{P_{r,min}}{P_{r,max}}$. Finally, noticing that $t^{\frac{m}{\gamma}} \in (0, 1)$, we complete the proof of convexity by using some bounds on the natural logarithm, i.e.,

$$1 - 1/t^{\frac{m}{\gamma}} \leq \ln(t^{\frac{m}{\gamma}}) \leq t^{\frac{m}{\gamma}} - 1,$$

and observing that (5.9) is always negative as

$$f(\gamma_1) - f(\gamma_2) \leq -n \left(2\gamma_1 - \gamma_2 + \gamma_2 t^{\frac{m}{\gamma}} \right) < 0.$$

The convexity of $\mathcal{L}(\gamma)$ is also demonstrated in Fig. 5.2.

When $P_{r,min}$ and $P_{r,max}$ can be calculated based on some prior knowledge, the ML self-estimate of the PLE can be obtained by forcing the derivative of $\mathcal{L}(\gamma)$ to 0, i.e., the ML solution solves

$$\frac{n\gamma}{m} - \sum_{i=1}^n \left(\ln \frac{P_i}{P_{r,min}} \right) + \frac{n \left(\frac{P_{r,min}}{P_{r,max}} \right)^{m/\gamma} \ln \left(\frac{P_{r,min}}{P_{r,max}} \right)}{1 - \left(\frac{P_{r,min}}{P_{r,max}} \right)^{m/\gamma}} = 0. \quad (5.10)$$

When $P_{r,min}$ and $P_{r,max}$ are unknown, we can rank the RSSs, leading to the following set of ordered RSSs: $P_{(1)} < \dots < P_{(n)}$. We further notice that the log-likelihood function in (5.8) is an increasing function of $P_{r,min}$ for $P_{r,min} \leq P_{(1)}$ and a decreasing function of $P_{r,max}$ for $P_{r,max} \geq P_{(n)}$. Therefore, for a fixed γ , this log-likelihood function is maximized when $P_{r,min} = P_{(1)}$ and $P_{r,max} = P_{(n)}$.

By respectively using the weakest RSS $P_{(1)}$ and the strongest RSS $P_{(n)}$ to replace the unknown $P_{r,min}$ and $P_{r,max}$, this ML self-estimate of the PLE can be obtained by solving

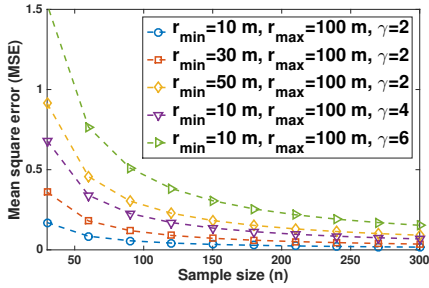
$$\frac{n\gamma}{m} - \sum_{i=1}^n \left(\ln \frac{P_i}{P_{(1)}} \right) + \frac{n \left(\frac{P_{(1)}}{P_{(n)}} \right)^{m/\gamma} \ln \left(\frac{P_{(1)}}{P_{(n)}} \right)}{1 - \left(\frac{P_{(1)}}{P_{(n)}} \right)^{m/\gamma}} = 0. \quad (5.11)$$

Both (5.10) and (5.11) can be efficiently solved by a simple bisection method. In our Matlab simulations, the function *fzero* helps us to calculate the solution.

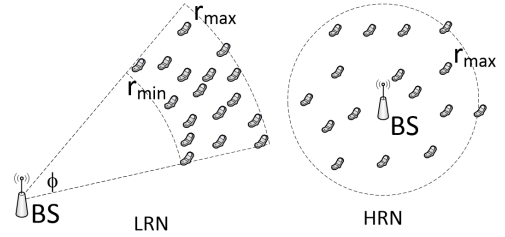
Finally, it is worth noting that, even if the *shadowing* effect is considered, the term $\sum_{i=1}^n \ln(P_i)$, which is the only sample-related part in our proposed ML solutions, becomes $\sum_{i=1}^n \ln(P_i) + \sum_{i=1}^n \xi_i$, where ξ_i is a zero-mean *Gaussian* variable, i.e., the *shadowing* effect by definition. Obviously, compared to $\sum_{i=1}^n \ln(P_i)$, the impact of $\sum_{i=1}^n \xi_i$ is relatively small, when the sample size n increases. Therefore, due to the limited space, we will not consider the case of the *shadowing* effect in the following simulations.

5.4. NUMERICAL RESULTS

WE have conducted two simulations to evaluate the performance of our two proposed ML estimators. The first simulation assumes an LRN and our two ML estimators are compared with the CRLB. Since no existing method is capable to estimate the PLE in an LRN, we decide to conduct the second simulation for an HRN, where our two ML estimators can be compared with two existing methods: our previously proposed weighted total least squares estimator (WTLS-PLE) of [8] and the estimator based on the cardinality of the transmitting set (C-PLE) of [7] (see also the Appendix). The two node deployments are shown in Fig. 5.3(b). The mean square error (MSE) is adopted to determine the accuracy of the estimators.

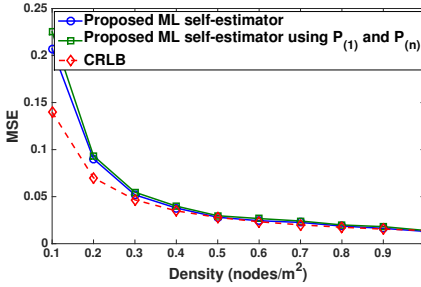


(a) CRLB for the self-estimation of the PLE in \mathbb{R}^2 .

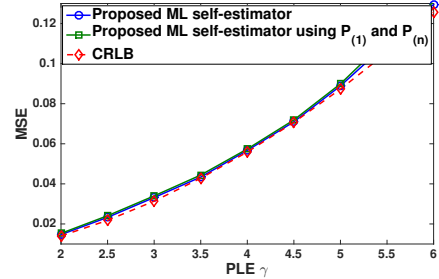


(b) The first simulation considers in an LRN, which is shown on the left side. The second simulation considers in an HRN, which is shown on the right side. Note that the HRN is a special case of the LRN.

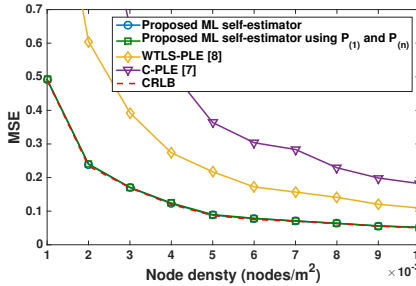
5



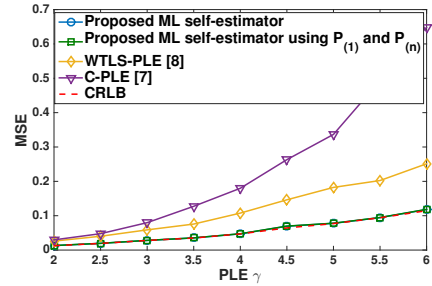
(c) First simulation: impact of the node density when the PLE is set to 2.



(d) First simulation: impact of the PLE when the node density is set to 1 nodes/m².



(e) Second simulation: impact of the node density when the PLE is set to 4.



(f) Second simulation: impact of the PLE when the node density is set to 0.01 nodes/m².

Figure 5.3: The simulations assume a 2-dimensional space, where nodes are randomly deployed. The carrier frequency is 2.4 GHz. The transmit power is 1 Watt. The antenna gains G_t and G_r are both 1. For the first simulation, $r_{min} = 50$ m, $r_{max} = 100$ m and $\phi = \pi/6$. For the second simulation, $r_{max} = 100$ m.

5.4.1. FIRST SIMULATION

THE numerical results are shown in Fig. 5.3(c) and Fig. 5.3(d), from which we can observe that both proposed ML self-estimators yield a very good performance that is very close to the CRLB. Even without the exact knowledge of $P_{r,min}$ and $P_{r,max}$ and using $P_{(1)}$ and $P_{(n)}$ instead, our ML estimator barely suffers any notable decrease in accuracy. Additionally, the performance of our two proposed estimators becomes better with a high node density and a small PLE.

5.4.2. SECOND SIMULATION

FOR comparison, the HRN, a special case of the LRN, is considered in this simulation to allow the use of existing estimators. In this case, $P_{r,max}$ is set to the transmit power P_t for our proposed ML self-estimators. As shown in Fig. 5.3(e) and Fig. 5.3(f), our ML self-estimators remarkably outperform the WTLS-PLE and the C-PLE. This can be explained by the fact that the WTLS-PLE requires ranking the RSSs, which adopts the rank numbers as a new set of observations. This incurs an extra impact on the estimation quality. The C-PLE, on the other hand, requires changing the receiver sensitivity. However, it simply depends on only two observations, i.e., the neighborhood size before and after the receiver sensitivity change, which makes this estimator very inaccurate and vulnerable.

5.5. APPLICATIONS AND FUTURE WORKS

DUe to their simplicity, the proposed ML self-estimators can be incorporated into any kind of wireless network. Hence, adapting existing wireless networking and communication designs to this change in PLE might lead to a better performance. We have already elaborated on many applications in [8]. Also note that the proposed self-estimators in this paper can also deal with the case when there exist node clusters, which might lead to broader applications. For example, as shown in Fig. 5.1, the BS can directionally adjust the transmit power to different remote villages according to the estimated PLEs such that the coverage of the signal or the energy efficiency can be guaranteed.

In this paper, the *shadowing* effect is ignored for convenience. To be more realistic, if it is considered, then the RSSs in wireless random networks are *log-normally* distributed with the expectation subject to the truncated *Pareto* distribution of (5.6). Therefore, if we intend to propose ML self-estimators for the PLE over *log-normal shadowing* fading channels, first a new distribution of the RSSs has to be obtained by blending the truncated *Pareto* distribution of (5.6) with the *log-normal* distribution, which might be mathematically very difficult and complicated.

5.6. CONCLUSION

TWO directional ML self-estimators for the PLE are proposed: one with known $P_{r,min}$ and $P_{r,max}$ and another one using $P_{(1)}$ and $P_{(n)}$ instead. The CRLB is also obtained. Only by locally collecting the RSSs, this kind of estimator can solely and individually estimate the PLE without any external information. Superior to all existing estimators, our two proposed ML self-estimators not only have a very good performance but are also feasible when nodes appear in clusters (all the existing methods assume a homogeneously

random node deployment). Two simulations have been conducted: the first one shows that the performance of our two proposed ML self-estimators is very close to the CRLB; the second one shows that they outperform two existing methods, i.e., our previously proposed WTLS-PLE and the C-PLE.

Most importantly, it is the first time that the RSSs based only a geometric path-loss in wireless random networks are found to follow a truncated *Pareto* distribution, which might be of great help for the analysis of future wireless networking and communication systems.

5.7. APPENDIX

THE PLE estimator based on the cardinality of the transmitting set (C-PLE) is proposed in [7], and requires changing the receiver sensitivity for a successful communication. More specifically, when the SINR of a nodal link at the considered receiver exceeds a certain threshold Θ , i.e., $\Theta \leq \frac{P_r}{I+N_0}$ where I is the interference and N_0 is the background noise, this communication link can be determined successful. The cardinality of the transmitting set is simply the number of successful communication links, which is also called the neighborhood size.

By changing the receiver sensitivity from Θ_1 to Θ_2 , the transmission range of the considered receiver changes and hence the cardinality of the transmitting set also varies from $N_{T,1}$ to $N_{T,2}$. Then, in \mathbb{R}^2 , the PLE can be estimated by

$$\hat{\gamma}_{\text{C-PLE}} = \frac{2 \ln(\Theta_2/\Theta_1)}{\ln(N_{T,1}/N_{T,2})}. \quad (5.12)$$

The C-PLE is only feasible for the HRN, where Θ_1 and Θ_2 are respectively calculated when the transmission ranges are 50 m and 100 m.

REFERENCES

- [1] A. Ozgur, O. Leveque, and D. Tse, *Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks*, [Information Theory, IEEE Transactions on](#) **53**, 3549 (2007).
- [2] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, *Locating the nodes: cooperative localization in wireless sensor networks*, [Signal Processing Magazine, IEEE](#) **22**, 54 (2005).
- [3] P. Cardieri, *Modeling interference in wireless ad hoc networks*, [Communications Surveys Tutorials, IEEE](#) **12**, 551 (2010).
- [4] G. Mao, B. D. O. Anderson, and B. Fidan, *Path Loss Exponent Estimation for Wireless Sensor Network Localization*, [Comput. Netw.](#) **51**, 2467 (2007).
- [5] G. Wang, H. Chen, Y. Li, and M. Jin, *On Received-Signal-Strength Based Localization with Unknown Transmit Power and Path Loss Exponent*, [Wireless Communications Letters, IEEE](#) **1**, 536 (2012).

- [6] N. Salman, M. Ghogho, and A. Kemp, *On the Joint Estimation of the RSS-Based Location and Path-loss Exponent*, [Wireless Communications Letters, IEEE 1, 34 \(2012\)](#).
- [7] S. Srinivasa and M. Haenggi, *Path loss exponent estimation in large wireless networks*, in *Information Theory and Applications Workshop, 2009* (IEEE, 2009) pp. 124–129.
- [8] Y. Hu and G. Leus, *Self-estimation of path-loss exponent in wireless networks and applications*, [Vehicular Technology, IEEE Transactions on PP, 1 \(2014\)](#).
- [9] M. Haenggi, *On distances in uniformly random networks*, [Information Theory, IEEE Transactions on 51, 3584 \(2005\)](#).
- [10] S. Srinivasa and M. Haenggi, *Distance distributions in finite uniformly random networks: Theory and applications*, [Vehicular Technology, IEEE Transactions on 59, 940 \(2010\)](#).
- [11] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001).
- [12] N. Nakagami, *The m -distribution, a general formula for intensity distribution of rapid fading*, in *Statistical Methods in Radio Wave Propagation*, edited by W. G. Hoffman (Oxford, England: Pergamon, 1960).
- [13] M. E. Newman, *Power laws, pareto distributions and zipf's law*, *Contemporary physics* **46**, 323 (2005).
- [14] Y. Ioannides and S. Skouras, *{US} city size distribution: Robustly pareto, but only in the tail*, [Journal of Urban Economics 73, 18 \(2013\)](#).
- [15] V. F. Pisarenko and D. Sornette, *Characterization of the frequency of extreme earthquake events by the generalized pareto distribution*, [pure and applied geophysics 160, 2343 \(2003\)](#).
- [16] B. Arnold, *Pareto distributions*, Statistical distributions in scientific work series (International Co-operative Publishing House, 1983).

IV

EPILOGUE

6

CONCLUSIONS AND FUTURE WORK

In this chapter, we conclude this thesis and provide some suggestions for future work.

6.1. CONCLUSIONS

THIS thesis is comprised of two lines of research: **signal strength (SS) based localization** and **path-loss exponent (PLE) self-estimation in wireless networks**. At first glance, these two research lines do not seem to relate to each other. Nevertheless, the PLE is a critical parameter of radio propagation channels, which is generally unknown. Therefore, estimating the unknown PLE is rather significant for SS-based localization. In practice, either the PLE is accurately estimated before the localization phase, or SS-based localization techniques have to jointly estimate the unknown PLE and the target location. However, the ultimate goal is nothing but obtaining an accurate location estimate. This is actually the main reason why we set these two lines for our research since they all stem from that same single purpose.

Furthermore, since scientific research should never be limited and innovations always originate from an open mind, we have also extended the PLE estimation towards a more general use in wireless networks, rather than only in the field of localization. Obviously, getting a grip on the PLE, a key parameter of radio propagation channels, is beneficial for designing more efficient wireless communication techniques and networking protocols. Therefore, this kind of method should be designed as an independent and self-driven entity, which can easily be incorporated into any kind of wireless network. This is also the reason why we use the term “**self-estimation**” for our second research line and note that we are the first to initiate this topic.

6.1.1. SS BASED LOCALIZATION

THE first research line has been discussed in Part II. In Chapter 1, we have explained the practical advantages of using the SS measurements (including the received signal strength (RSS) and the differential received signal strength (DRSS)) for localization. In short, employing SS measurements is very convenient for any wireless device in real-life.

Unlike other measurements such as time-of-arrival (TOA) and angle-of-arrival (AOA), SS based localization does not rely on any external system like clock synchronization or an antenna array. Furthermore, we select the DRSS measurements for localization, which can be viewed as analogous to time-difference-of-arrival (TDOA) based localization. In doing so, we not only eliminate the unknown transmit power by taking differences between SS measurements, but also gain the independence of the signal transmitter for localization. Other merits have also discussed in Chapter 1.

In Chapter 2, in order to construct the optimization problem for DRSS-based localization, we have first introduced a whitened linear model, which can be linked to a whitened RSS-based model through an orthogonal operator, thus suffering no information loss comparatively. Next, we have coped with the non-convexity issue and presented three different kinds of methods. They are the advanced best linear unbiased estimator (A-BLUE), the Lagrangian estimator (LE) and the robust semi-definite programming estimator (RSDPE). For the latter, we also considered the impact of the inaccurate anchor location information and the inaccurate PLE estimate, which actually results into model uncertainties for our localization problem. Hence, the RSDPE is particularly designed to be robust against those model uncertainties. Then, to study these three methods, we have not only conducted numerical simulations under different kinds of impacts but also fully compared them in a general context. Finally, when the PLE is totally unknown, we have also proposed a robust SDP-based block coordinate descent estimator (RSDP-BCDE) for DRSS-based localization.

When solving the localization problem, two conflictive facts kept haunting us: a) in SS-based localization, taking differences between measurements seems to cause no information loss as already indicated above; b) there exists a lot of TDOA localization literature studying the impact of the differencing processing, which implies otherwise. Therefore, in order to unveil the underlying truth, we make a further step to particularly study the differencing process. In Chapter 3, we have extended our study to a general case of multiple linear nuisance parameters, for which we present a unified framework and introduced a general differential method that can cope with that. We have compare the general differential method with the other two methods that can also tackle multiple linear nuisance parameters: joint estimation and the orthogonal subspace projection (OSP) estimation. Their corresponding best linear unbiased estimators (BLUEs) were used to bridge them. We noticed that, regardless of the complicated procedure of the differential method, the three BLUEs are actually equivalent to each other after a proper preprocessing. Considering the fact that the joint estimation directly uses the measurements and preserves the full data information, the differencing process hence causes no information loss, which coincides with the first fact mentioned earlier. To explain the second fact, we dug into some particular localization cases and pointed out that the modelling process actually causes the information loss. In some TDOA localization cases, this information loss may be subject to the reference choice in the differencing process, which is why the differencing process seems to be responsible for that. More importantly, this finding leads to many other interesting conclusions. For instance, the differential observation set associated with a single reference preserves the full data information, while many literature still tries to collect more differential observations with multiple references.

6.1.2. PLE SELF-ESTIMATION IN WIRELESS NETWORKS

THE second research line is discussed in Part III. Although this research line originated from the SS based localization, the significance of estimating the PLE is not merely limited to this particular case, but impacts a broader field. In fact, the PLE is so crucial and decisive for radio propagation channels that any kind of wireless network must get a grip of it for smart self-adaptation according to dynamically changing environments. Therefore, in order to do that, the PLE estimation cannot come with many constraints, but with a great freedom to be pervasively adopted. However, as a trade-off, that will lead to many difficulties when designing the desired PLE estimation method. In Chapter 1, we have elaborated on the motivations, the related work and our expected features for the PLE self-estimation.

In Chapter 4, we have initiated a new research topic: PLE self-estimation, which is merely based on the locally collected SS measurements. In order to cope with the unknown distance issue, we rank the SS measurements and introduce the rank indices as new observations that also contain the desired distance information. Next, a distance-related intermediate parameter is introduced and ML estimated using those rank indices such that we can formulate a total least squares (TLS) optimization problem for PLE self-estimation. This requires a low-rank approximation and hence the traditional solution can be obtained using the singular value decomposition (SVD). However, the SVD procedure results into a high computational complexity, which could be practically inconvenient. To cope with that, we have also proposed a closed-form solution that greatly reduces the computational time. We have further noticed the fact that the ranking procedure mentioned earlier could be disturbed due to some wireless channel effects. In other words, the SS rank might not exactly map the distance rank because of the *shadowing* effects, thus resulting into some mismatches in the rank indices. To eliminate the impact of those mismatches, we have further proposed a weighted closed-form solution. After that, we have elaborated on the potential applications of using this kind of collective PLE self-estimation.

Although the PLE self-estimation methods mentioned above yield a good performance, there still exist many unresolved issues such as the CRLB and the full ML solution. Obviously, all those issues can be attributed to one single obstacle, the unknown observation distribution, where the observation refers to the SS measurement. The problem is that the SS measurement is not only impacted by the wireless channel effects but also subject to the spatial dynamics (a random node placement). Unfortunately, there is no distribution that suffices to fully characterize the SS measurement. In Chapter 5, we start with a simple case that only considers the geometric path-loss of wireless channels. Surprisingly, we discover that the SS measurement now is *Pareto* distributed, which has never been reported before. Based on that, we obtain two ML solutions and the CRLB. More importantly, this might open a new perspective of studying wireless communication and networking, since the SS yields a multi-faceted influence such as on energy consumption, signal coverage and the signal-to-interference-plus-noise ratio (SINR).

6.2. SUGGESTIONS FOR FUTURE WORK

Here, we discuss some directions for the future work and hope that they could be of any help to those who are interested.

6.2.1. LOCALIZATION

- **Mobile Scenarios:** So far, our work in Chapter 2 only considers a stationary localization network. However, if we want to track a moving target (even with mobile anchors) using SS measurements, the whole optimization problem would become rather different. In mobile scenarios, the tracking problems are often formulated using the Taylor expansion, the first solution to the non-linearity issue discussed in Chapter 1. Due to the advantages of using SS measurements, the SS-based mobile localization problem is worthy of a deep investigation. In particular, not only the target location but also the target speed can be estimated. And, some unexpected external impacts require attention, e.g., sudden link disconnections and weak localization signals.
- **A Better Linearization?** Undoubtedly, the non-linearity issue widely exists in many fields. In Chapter 1, we have discussed two solutions for the localization case: Taylor expansion and unfolding the distance norm. Here, we would like to present another method that can deal with the non-linearity, the unscented transformation (UT) [1]. The UT is widely used and the most famous case is the unscented Kalman filter (UKF) [2], which yields a very good performance especially in a highly non-linear situation. Considering the severe non-linearity issue in SS based localization, the UT might be a good choice and there are already some applications [3, 4]. However, the number of studies on this topic are still far from enough. For instance, the impact of an inaccurate PLE estimate and inaccurate anchor location information has not been studied yet for this kind of method. Additionally, the UT is also applicable in the mobile scenarios discussed above.
- **Cooperative Localization with Multiple Targets:** In this thesis, we have only considered non-cooperative localization, where the SS measurements are only collected between the anchors and a single target. Nonetheless, if there exist multiple targets, SS measurements between these targets can also be used for localization, which is called cooperative (or collaborative) localization. The multidimensional scaling (MDS) method is the conventional solution for this case [5], but it removes all terms of the form $R_i \triangleq \|\mathbf{x}_i\|^2$ using the orthogonal subspace projection (OSP), where \mathbf{x}_i is the i -th target node location. As already pointed out in both Chapter 2 and Chapter 3, this is equivalently to assuming that R_i and \mathbf{x}_i are independent unknowns. However, we know that all the relations between R_i and \mathbf{x}_i should be taken into account for a better localization performance. Anyway, there still exists a large room for improvement in this area, where the breakthroughs are certainly not limited to the one mentioned above.
- **Optimal TDOA based Localization:** The idea of this direction mainly stems from the localization example in Chapter 3. As already concluded, the modelling process drops some high-order reference-dependent noise terms, thus causing an information loss as well as making the performance of most TDOA based localization methods rely on the reference selection. In other words, by studying the reference choice, engineers actually try to minimize the information loss. However, the conclusions in Chapter 3 have clearly pointed out that the differencing process that generates the TDOA measurements does not cause any information loss.

This implies that if we can somehow well compensate the information loss from the modelling process or just not drop those terms, the TDOA based localization should be free of the impact of the reference choice and might reach its optimal performance.

- **New SS Measurements with Background Noise:** As indicated in Section 1.1.3 of Chapter 1, the SS measurements used in this thesis are assumed to be obtained after successful demodulation and segregation of the background noise (BGN). However, if we consider hostile military scenarios, the signal demodulation might be very difficult or expensive. In this case, an alternative solution is obtaining new SS measurements by integrating the received power, i.e., $\frac{1}{T} \int |y(t)|^2 dt$, which obviously includes the impact of the BGN. A reliable model for this kind of SS measurement has been introduced in [6, eq. (17)]. However, the study of localization based on this kind of SS measurement has still been left almost blank. Moreover, the same holds for the PLE self-estimation, though we will not mention this direction later on.

6.2.2. PLE SELF-ESTIMATION

- **Applications of the PLE Self-Estimation:** Some potential applications of the PLE self-estimation have been carefully elaborated on in Section 4.6 of Chapter 4 and Section 5.5 of Chapter 5. Still, due to the significance of the PLE in wireless communications and networking, the uses of the PLE self-estimation are obviously far more than what we have listed. For instance, the PLE estimate can be used to improve the communication quality for the patient monitoring process in clinical environments [7]. The environmental sensing system can also detect the changing PLE to discover a good communication area [8], which is particularly helpful in military scenarios.
- **Impact of the Shadowing Effect:** Although the ML solutions derived in Chapter 5 are obviously robust against the *shadowing* effect, this has not been studied numerically. The impact might still be considerable under a large *shadowing* effect. For instance, if the two parameters $P_{r,min}$ and $P_{r,max}$ are not known *a priori*, $P_{(1)}$ and $P_{(n)}$ might not be good estimates respectively, since the *shadowing* effect would disturb the ranking procedure. Furthermore, in this case, it is questionable that the proposed estimators are still ML solutions.
- **Applications of the Pareto Distribution:** In our second research line, the finding of the SS being *Pareto* distributed in simplified wireless channels (with only a geometric path-loss) is viewed to be as significant as the new PLE self-estimation topic itself. This is the reason why we kept emphasizing that this might be of great help for studying wireless communications and networking. Due to the fact that this may go beyond the scope of this thesis, we will only list some potential fields that can benefit from this finding:
 1. **Interference:** Most literature about the interference in wireless networks assumes a homogeneous random node placement. Yet, the *Pareto* distribution we offered in Chapter 5 can also cope with node clusters. Coupled with other

wireless channel effects such as the *small-scale* fading, the *Pareto* distribution can be used to study the interference in the presence of node clusters, which has never been done before.

2. **Energy Consumption:** The *Pareto* distribution actually models the main power loss on the transmission link in random wireless networks. Accordingly, we have a new perspective for studying the energy consumption for propagation and the topology could be optimized or adjusted to minimize the total power loss for networking.
 3. **Routing:** Optimal routing must be energy-efficient, though most power is consumed for propagating the message. Therefore, every hop should be carefully designed. For instance, assume we can particularly point out a certain area, where the next hop would be. The nodes in this area can be considered as a node cluster and hence we can easily use the *Pareto* distribution to calculate the expected power loss for jumping to this area.
- **Realistic distribution of the SS measurements including the shadowing effect:** This direction has been frequently mentioned in Chapter 1, Chapter 4 and Chapter 5. Obtaining a more realistic distribution would be rather mathematically difficult, since it involves blending the *Pareto* distribution with the *log-normal* distribution. Even if it can be done, using this distribution would become the next problem, since the form might be very complicated. On the other hand, we should also notice the potential benefits of doing so. More realistic ML solutions and CRLBs for the PLE self-estimation could be obtained. The impact caused by the ranking procedure might be removed once and for all. Finally and more importantly, all the directions discussed above for the *Pareto* Distribution can be studied or solved more realistically and accurately.

REFERENCES

- [1] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, *A new method for the nonlinear transformation of means and covariances in filters and estimators*, *IEEE Transactions on Automatic Control* **45**, 477 (2000).
- [2] S. J. Julier and J. K. Uhlmann, *New extension of the kalman filter to nonlinear systems*, in *AeroSense'97* (International Society for Optics and Photonics, 1997) pp. 182–193.
- [3] G. Wang and K. Yang, *A New Approach to Sensor Node Localization Using RSS Measurements in Wireless Sensor Networks*, *Wireless Communications, IEEE Transactions on* **10**, 1389 (2011).
- [4] G. Wang, H. Chen, Y. Li, and M. Jin, *On Received-Signal-Strength Based Localization with Unknown Transmit Power and Path Loss Exponent*, *Wireless Communications Letters, IEEE* **1**, 536 (2012).
- [5] J. A. Costa, N. Patwari, and A. O. Hero, III, *Distributed weighted-multidimensional scaling for node localization in sensor networks*, *ACM Trans. Sen. Netw.* **2**, 39 (2006).

- [6] R. Martin, A. King, J. Pennington, R. Thomas, R. Lenahan, and C. Lawyer, *Modeling and mitigating noise and nuisance parameters in received signal strength positioning*, *Signal Processing, IEEE Transactions on* **60**, 5451 (2012).
- [7] J. Ko, T. Gao, R. Rothman, and A. Terzis, *Wireless sensing systems in clinical environments: Improving the efficiency of the patient monitoring process*, *IEEE Engineering in Medicine and Biology Magazine* **29**, 103 (2010).
- [8] M. F. Othman and K. Shazali, *Wireless sensor network applications: A study in environment monitoring system*, *Procedia Engineering* **41**, 1204 (2012).

CURRICULUM VITÆ

Yongchang HU

Yongchang Hu was born in Xi'an, China, in 1988. He received the B.Sc. and M.Sc. degrees from Northwestern Polytechnical University, Xi'an, China, in 2010 and 2013, respectively. From October 2011, he worked as a Ph.D researcher in the Telecommunications group and then, in April 2013, he joined the Circuits and Systems (CAS) Group, Department of Microelectronics, Delft University of Technology, Delft, The Netherlands. His Ph.D study was funded by China Scholarship Council (CSC). During this period, he was under the supervision of Prof. dr. ir. Geert Leus with the topic of "Signal Strength Based Source Localization and Path-Loss Exponent Self-Estimation in Wireless Networks". He has been actively serving as the technical reviewer for several IEEE journals and major conferences. His research interests lie in the board area of signal processing and machine learning for communication and networking, particularly in localization/tracking, synchronization, OFDM, MIMO, mmWave Communication, compressive sensing, modelling radio propagation channel, random network analysis and etc..

EDUCATION

- | | |
|-----------|--|
| 2006–2010 | Bachelor of Science in Engineering
School of Electronics and Information
Northwestern Polytechnical University, Xian, China
<i>Thesis:</i> The Design, Analysis and Simulation of GPS Receiver
Based on Matlab/Simulink
<i>Promotor:</i> Prof. Baowang Lian |
| 2010–2013 | Master of Science in Engineering
Texas Instruments-Northwestern Polytechnical University DSPs Laboratory,
School of Electronics and Information,
Northwestern Polytechnical University (NWPU), Xian, China
<i>Thesis:</i> The study of Multi-path Estimation on GPS Receiver
using Particle Filter
<i>Promotor:</i> Prof. Baowang Lian |
| 2011-2017 | Ph.D of Science in Engineering
Circuit and System (CAS) group,
Faculty of Mechanical, Maritime and Materials Engineering,
Delft University of Technology, Delft, the Netherlands
<i>Thesis:</i> Signal Strength Based Source Localization and Path-
Loss Exponent Self-Estimation in Wireless Networks
<i>Promotor:</i> Prof.dr.ir. Geert Leus |



These two pictures (from top to bottom) were respectively taken in 2011 and 2017 while I was hiking and climbing the Qin Mountains in China whose peak, Mount Taibai, is at 3767 meters. These two trips were specially arranged to mark the beginning and the end of my Ph.D journey.

*If you surrender halfway, you would never reach the peak;
you actually do not even have this option if you want to live in the wild.*

Yongchang Hu

LIST OF PUBLICATIONS

8. **Y. Hu** and G. Leus, "Robust Differential Received Signal Strength-Based Localization," in IEEE Transactions on Signal Processing, vol. 65, no. 12, pp. 3261-3276, June 15, 2017.
7. **Y. Hu** and G. Leus, "On A Unified Framework for Linear Nuisance Parameters," EURASIP Journal on Advances in Signal Processing, vol. 2017, no. 1, p. 4, 2017.
6. **Y. Hu**, "Comments on "Cooperative Density Estimation in Random Wireless Ad Hoc Networks", " in IEEE Communications Letters, vol. 20, no. 4, pp. 832-835, April 2016.
5. **Y. Hu** and G. Leus, "Directional Maximum Likelihood Self-Estimation of the Path-Loss Exponent", Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on, 2016
4. T. Xu, **Y. Hu**, B. Zhang, G. Leus, "RSS-Based Sensor Localization in Underwater Acoustic Sensor Networks", Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on, 2016
3. **Y. Hu** and G. Leus, "Self-Estimation of Path-Loss Exponent in Wireless Networks and Applications," in IEEE Transactions on Vehicular Technology, vol. 64, no. 11, pp. 5091-5102, Nov. 2015.
2. Y. Liu, R. Lotfi, **Y. Hu** and W. A. Serdijn, "A Comparative Analysis of Phase-Domain ADC and Amplitude-Domain IQ ADC," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 62, no. 3, pp. 671-679, March 2015.
1. T. Islam, **Y. Hu**, E. Onur, B. Boltjes and J. F. C. M. de Jongh, "Realistic simulation of IEEE 802.11p channel in mobile Vehicle to Vehicle communication," Microwave Techniques (COMITE), 2013 Conference on, Pardubice, 2013, pp. 156-161.