Using Kullback-Leibler Divergence to Analyze the Performance of Collaborative Positioning

Jeannette D. Nounagnon

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

Timothy Pratt, Chair
R. Michael Buehrer, Co-Chair
Charles W. Bostian
Steven W. Ellingson
Yvan J. Beliveau

April 22, 2016
Blacksburg, Virginia

Keywords: Collaborative positioning, Kullback-Leibler Divergence, Position Mean Squared Error, Performance Metric, Mutual Information

Using Kullback-Leibler Divergence to Analyze the Performance of Collaborative Positioning

Jeannette D. Nounagnon

(ABSTRACT)

Geolocation accuracy is a very crucial and a life-or-death factor for rescue teams. Natural disasters or man-made disasters are just a few convincing reasons why fast and accurate position location is necessary. One way to unleash the potential of positioning systems is through the use of collaborative positioning. It consists of simultaneously solving for the position of two nodes that need to locate themselves. Although the literature has addressed the benefits of collaborative positioning in terms of accuracy, a theoretical foundation on the performance of collaborative positioning has been disproportionally lacking.

This dissertation uses information theory to perform a theoretical analysis of the value of collaborative positioning.The main research problem addressed states: 'Is collaboration always beneficial? If not, can we determine *theoretically* when it is and when it is not?'
We show that the immediate advantage of collaborative estimation is in the acquisition of another set of information between the collaborating nodes. This acquisition of new information reduces the uncertainty on the localization of both nodes. Under certain conditions, this reduction in uncertainty occurs for both nodes by the same amount. Hence collaboration is beneficial in terms of uncertainty.

However, reduced uncertainty does not necessarily imply improved accuracy. So, we define a novel theoretical model to analyze the improvement in accuracy due to collaboration. Using this model, we introduce a variational analysis of collaborative positioning to determine factors that affect the improvement in accuracy due to collaboration. We derive range conditions when collaborative positioning starts to degrade the performance of standalone positioning. We derive and test criteria to determine on-the-fly (ahead of time) whether it is worth collaborating or not in order to improve accuracy.

The potential applications of this research include, but are not limited to: intelligent positioning systems, collaborating manned and unmanned vehicles, and improvement of GPS applications.

Dedication

To my special research lab team: my husband, children and family
Words are not enough...

Acknowledgments

I would like to thank all my committee members for their support and their time in assisting me all these years, both in person and remotely. Thanks to all of you for the time reviewing this work and for your appreciated feedback. Special thanks to Dr. Pratt and Dr. Buehrer. Each of you has contributed in a special way to the completion of this research topic and its review. I have been working as a full-time engineer while working on this research. I did not have a lab and other graduate students to exchange ideas with. So the result of this research is from direct interactions and technical exchanges with you. So thank you.

Thank you Dr. Pratt for your unwavering support, guidance and encouragement through the years. I learned from you that life is also about the relationships we build along the way. These are the important things that we carry with us after the work is completed. I am grateful that our paths crossed via my interest in the satellite industry and specifically for unmanned vehicles. You helped me see how I could use my digital design background in an area I dreamed to work in since I was a child. That led me to take a very drastic turn in my academic course by going from a masters in the semiconductor field, to a doctorate in positioning applications. Your encouragement and mentoring through the years have allowed me to successfully transition to a technical area that my heart always desired to work on. Thank you again.

A special thanks to you Dr. Buehrer for your particular investment in this research, for your technical guidance and for challenging me with very tough questions. This has justified my numerous clarification emails to you in regard to questions you had asked *a while back*. Thank you for going through all revisions of all my long chapters and drafts. Please know that this is not the end of my research career, but the beginning of another season that will produce even more technical results in many more technical arenas. Interestingly, I recently remembered that you were the first professor that I talked to when I came to visit the VT campus as a prospective student. I am happy and grateful that our paths have crossed again in such a significant manner. I am honored to have worked with you. As one of the lead researchers on this research topic, your feedback and technical questions have been critical to the completion of this work, and have meant a lot to me. Thank you.

Dr. Bostian, I joined your lab for a short season as a graduate student and learned the most on topics that have fascinated me since: software defined radios, and public safety. Even in my busy schedule, as I reflect about tomorrow, I have since taken extra time to ensure that whatever I do technologically-speaking goes hand-in-hand with guaranteeing the safety and the betterment of the community I live in.

Thank you Dr. Ellingson and Dr. Beliveau for your time and involvement in the success of this defense. Thank you Dr. Beliveau for making yourself always available. Dr. Ellingson, thank you very much for your advice on how to strategically approach the doctorate like a business.

The completion of this work would not have happened without the unwavering crowd of witnesses that have encouraged me, prayed with and for me all these years. To my sisters, brothers, parents, in-laws, extended family, my managers and mentors at work, our local friends and colleagues, my undergraduate professors and advisers from the University of Kentucky who kept on passing words of encouragement. Special thanks to Cindy H. and John H. in the VT ECE department, for helping me through my rather rough transition to life on the VT campus; I will not forget your friendship at such a critical time of my life and even afterward. To the VT International/Graduate chapter of Intervarsity Christian Fellowship led by Karen and Robert, thanks for introducing me to another life outside of school and especially for taking me along to Urbana! You are well aware of how that year has forever changed my life. I really wish one dear friend was still here to see the beginning of this new season and sing and dance with us. I am just grateful to have known her and can't wait to sing and dance with her again.

To my Kentucky family and friends, you saw me from my very first months as an undergraduate student, then as a graduate student, engineer, mom and still a student. You were there through the rough years of war in my home country when I still had to find a way to concentrate on my research while working full-time. Thank you for being our family away from home. I will forever be grateful for your support, opened doors and your opened arms through the tough times and through the happy seasons of our lives when our parents could not be here to celebrate with us.

I do not have enough words to thank you Abba, my God, the Creator of the Universe, your Son, and Counselor, for giving me understanding, inspiration when I hit dead-ends, and above all the strength to keep on going. A lot has happened through these years that words cannot describe. But through it all, you have been my rock and my shield in whom I find shelter, the power that saves me, and my firm foundation. The completion of this work is a miracle in itself. I will forever be grateful to you Abba.

Contents

List of Figures

Table 1: Table of Acronyms

Acronyms	Definitions
CDF	Cumulative density function
CGF	Cumulant generating function
COL	Collaborative positioning
CRB	Cramer-Rao Bound
CRB_{COL}	The Cramer-Rao Lower Bound of the collaborative positioning estimator
CRB_{STD}	The Cramer-Rao Lower Bound of the standalone positioning estimator
CRLB	Cramer-Rao Lower Bound
ΔKLD	Novel performance metric introduced in this work
GDOP	Geometric dilution of precision
GGDOP	Generalized Geometric dilution of precision
FCC	Federal Communications Commission
FIM	Fisher Information Matrix
KLD	Kullback-Leibler divergence
LOS	Line of sight
MGF	Moment generating function
MI	Mutual information
MSE	Mean squared error
NLOS	Non-line of sight
PDF	Probability density function
PMF	Probability mass function
PMSE	Position mean squared error
RMSE	Root mean squared error
SNR	Signal-to-Noise Ratio
STD	Standalone positioning
w.r.t	with respect to

Notation and Nomenclature

In this dissertation, the following notation and nomenclature is adopted. Normal font lower case letters (e.g: ϵ, x) are scalars (deterministic but with unknown values). Normal font upper case letters are random variables. Bold face lower case letters are vectors, and bold face upper case letters are matrices. The table below illustrates the use of this nomenclature and includes other definitions for mathematical and statistical notations.

μ	The mean of a scalar distribution
σ^2	The variance of a scalar distribution
$\boldsymbol{\mu}$	The mean vector of a multivariate distribution
$\mathbf{\Sigma}$	A covariance matrix
$\Sigma_{i=j}^{k}$	Summation sign for index i going from j to k with $j \leq k$
$\mathbf{I}$	The identity matrix
$\mathbf{x}_A$	The position vector of node A
$\mathbf{x}_B$	The position vector of node B

x	A scalar
$X\|Y$	The conditional random variable X given random variable $Y = y$
$\|x\|$	The absolute value of scalar x
$\|\|\mathbf{x}\|\|$	The Euclidean norm of vector $\mathbf{x}$
$\mathbf{\Sigma}^{-1}$	The inverse matrix of matrix $\mathbf{\Sigma}$
$\det(\mathbf{\Sigma})$	The determinant of matrix $\mathbf{\Sigma}$
$\mathbf{X}^T$	The transpose of matrix $\mathbf{X}$
$\boldsymbol{FIM}$	The Fisher Information Matrix
$\mathrm{Tr}(\mathbf{\Sigma})$	The trace of matrix $\mathbf{\Sigma}$
$\mathbf{0}_{m,n}$	An mxn null matrix
$\mathbf{I}_m$	An mxm identity matrix
$diag[\mathbf{M}]$	The diagonal matrix $\mathbf{M}$
$\log_z(x)$	The logarithm of x to base z
$\log(x)$	The natural logarithm of x (the logarithm to base e)
SNR_{dB}	The Signal-to-noise ratio in dB

$\exp(x)$	The exponential of x
min	The minimum of a set
sup	The supremum of a set
$f(x;\theta)$	The pdf $f(x)$ depending on parameter θ
$f_{X,Y}(x,y)$	The joint pdf of random variables X and Y
$E_[g(x)]$	The expected value of function $g(x)$ given that the pdf of X is $f_X(x)$
$n!$	The factorial of n
C_n^k	The number of k-combinations in a set of n elements

$N(\mu,\sigma^2)$	A scalar Gaussian distribution with mean μ and variance σ^2
$SN_k(\boldsymbol{\xi},\boldsymbol{\Omega},\boldsymbol{\eta})$	The multivariate skew-normal pdf of a k-dimensional random variable of location vector $\boldsymbol{\xi}$, dispersion matrix $\boldsymbol{\Omega}$, and shape/skewness parameter $\boldsymbol{\eta}$

$H(f_X)$	The entropy or uncertainty of a random variable X described by a distribution $f(x)$
$H(f_{XY})$	The entropy of the joint pdf $f_{XY}(x,y)$ of the random variables X and Y
$H(f_{X\|Y})$	The conditional entropy of $X\|Y=y$ described by a distribution $f_{X\|Y}$
$H(f_X\|\|g_X)$	The cross-entropy of two distributions $f(x)$ and $g(x)$ each describing a random variable X
$KLD(f_X\|\|g_X)$	The Kullback-Leibler Divergence between two distributions $f(x)$ and $g(x)$
$MI(X,Y)$	The mutual information between two random variables X and Y

Chapter 1

Introduction

The problem of geolocation has gained interest in recent years with the introduction of the FCCś E911 standard along with numerous efforts to improve position accuracy of GPS-based (Global Positioning System) systems and also to improve position and tracking accuracy in GPS-denied environments. Geolocation is useful not only for consumer applications, but also vital in many situations involving public safety. Timely responses to find individuals after bomb attacks or natural disasters are just a few reasons for such increased interest. Public safety and military applications also need improved and more accurate geolocation devices and techniques. Consequently, achieving accuracy in position location is a task of great importance.

Looking ahead, the overall technological trend in regard to positioning, navigation and tracking applications is that navigation and positioning applications are merging. In industry, the technological roadmap for commercial technologies in the field of inertial navigation systems (INS) includes performance objectives such as: providing less than three to five meters in both horizontal and vertical accuracy, and locating indoors to a floor and a room. Mission duration requirements are moving towards more than 30 minutes of sustainable performance in a fully GPS-denied environment, and 24 hours performance with intermittent GPS. The main technical challenge is to be able to navigate through significant GPS signal degradation or denial for longer periods of time [97].

Although they provide tracking information indoors, Inertial Navigation Systems are known to date for degrading accuracy over time. As a result, they are only reliable for a short period. That was until the INS+GPS combination made its way to the technological market. The combination of INS and GPS has proven itself very reliable especially in areas where GPS is accessible because accurate position estimates from GPS minimize the tracking errors in INS. However, in GPS-denied environments, relying solely on INS with no way to improve position estimates has and will continue to be a problem unless appropriate solutions are considered.

As collaborative positioning is better understood, we believe that in the near future, the combination of INS+Collaborative Positioning will be one answer that will improve the performance of indoor navigation and tracking applications. Nonetheless, it is evident that as sensors increase in numbers around us, understanding the theoretical fundamentals about how to make efficient use of such available resources is a ground-breaking effort that will open the way to improve the tracking of our first responders in emergency situations.

According to the United States Department of Homeland Security (DHS) Science and Technology Group on March 3, 2015, 'Developing sensors and communications devices that can track first responders in a variety of indoor environments remains a critical need that previous research and development efforts have not been able to address'. Consequently, adequate indoor tracking in a GPS-denied environment is still a problem that we have yet to resolve.

One solution that has been considered in some applications is to exchange data (collaborate) between devices that need to be located or tracked. By exchanging data, devices that could not locate themselves can acquire some information about their position from neighboring devices with the hope to improve overall performance. Exchanging data however brings additional complexity to the communication layer between devices. Provided the limited bandwidth that all devices have to share to communicate, the old adage: 'Think before you speak' is worth considering in this context; that is, assessing ahead of time whether it is worth sharing data, is worth understanding to efficiently use the limited communication bandwidth and power resources available, especially in emergency scenarios.

The problem of collaboration is not trivial. In an emergency scenario with first responders on a scene indoors, exchanging data with another first responder carries the potential to worsen one's own position accuracy. How then can one tell when it is worth exchanging data or not to improve accuracy? This is the problem we address in this research.

One way to approach the analysis of collaboration is the use of simulations. However, simulation results are limited because they cannot be easily adapted to different environments, especially unpredictable environments like emergency situations. Hence, the proper approach that we selected is to address the problem theoretically.

The problem of knowing ahead of time whether collaborating will be beneficial or not is the problem that we address in this research, specifically for the positioning problem. The problem of collaborative navigation is an extension of this work. In fact, by combining collaborative positioning with INS systems, we believe that we can address the positioning errors within tracking systems in GPS-denied environment. Consequently, in this research, we focus on the following question: 'Is collaboration always beneficial? If not, can we determine *theoretically* when it is and when it is not?'

In essence, collaboration is about acquiring new information. Thus, if we can mathematically describe the new information acquired, we would be able to mathematically describe the impact of the new information on position estimation accuracy. This is the main foundation of this research: we address the problem of collaborative positioning (an estimation problem) from an information theory perspective. In essence, in this research, we analyze the information acquired during collaboration and model its impact on the position accuracy. The goals of this research are mainly two-fold: a) derive a theoretical framework that provides a theoretical understanding of how collaborative positioning works and b) provide practical criteria that can tell ahead of time whether it is worth collaborating or not, with the intent to improve positioning accuracy.

1.1 Literature Review

Numerous algorithms have been implemented for geolocation. The generic approach to estimate position consists of placing sensors in fixed known locations, and using techniques such as *trilateration* to locate a transmitter (Tx) or receiver (Rx). On one hand, one transmitter to be located (or node) may transmit the appropriate location parameter to receivers of known positions (or anchors) which process all range estimates to the node to determine its position. On the other hand, a transmitter can transmit a signal with timestamps, for example, to a set of receivers of known positions. As a convention, we shall refer to nodes as sensors that need to be located and refer to anchors as sensors of known positions used to help locate a node. All nodes determine their range with respect to anchors; a central unit combines all range information to estimate the position of the node. This is essentially how range-based geolocation systems operate.

One common technique, namely trilateration, converts estimated distances between anchors and the node into radii of circles centered at sensor nodes. Ideally, when there are no errors, all circles intersect at a point: the location of the node. However, this is never the case due to errors that reduce the accuracy of the location information. The accuracy of the position estimate depends greatly on the type of information transmitted between a radio and sensors, and on the perturbations the signal suffers along the flight.

If time of arrival is used to estimate distances, distance can be obtained by time of flight. This requires that all sensors' clocks in the system be synchronized. Since electromagnetic waves travel in free space at the constant speed of light, the relationship between time of flight and distance is the product of time of flight and the speed of light. In such cases, the receiver must be capable of identifying the first arriving signal. Due to the presence of obstacles along the path traveled by a transmitted signal, the signal is reflected and diffracted. Hence, time of flight is increased. This phenomenon is called multipath. As a result of multipath, multiple copies of the same signal arrive at different times at the receiver. These copies can combine constructively, or destructively. Therefore, in time of arrival systems which estimate time of flight, the accuracy of estimation relies on the ability of the receiver to estimate the first arriving signal, that is, the signal that traveled the direct line of sight (LOS) path between the transmitter and receiver. In some cases, this direct path is blocked such that the signal has to travel a longer distance to reach the receiver. The additional path length introduced into the blocked LOS path is defined as non line of sight (NLOS) error. Therefore, NLOS errors are a major issue in position location systems as they cause positive unknown range errors. Another type of error is measurement noise. In order to detect time of arrival (TOA) of the first arriving signal, the receiver must have the capabilities to do so. Receiver bandwidth is one parameter that determines how accurately a receiver can detect the first arriving signal. The larger the receiver's bandwidth, the higher the resolution in time of arrival estimates.

Once the processing station acquires all sensors' data, a position of the node can be estimated using different types of position estimators. Position estimators can be linear (*e.g*: Linear Least Squares Estimator) or non-linear (*e.g:* Maximum Likelihood Estimator). In the linear case, the distribution of position estimates is known to be the same as the distribution of observed range measurements. In the non-linear case, one cannot predict the distribution of position estimates. Nonetheless, maintaining or improving the accuracy of positioning systems is very critical for many applications today.

While requirements for improved accuracy are becoming more stringent, the number of nodes to locate in positioning systems keeps increasing. In the well-established standalone positioning framework, each node location is evaluated individually, with the use of reference anchors of known locations. This is the position estimation system aforementioned.

As moving nodes become less accessible to their anchors (*e.g:* indoors), location availability of tracked devices can quickly become a serious concern in standalone positioning. To circumvent this challenge, we have observed the emergence of *collaborative positioning* [11]. It consists of solving *simultaneously* for the location of more than a single node by making use of additional resources available (*i.e* an additional node that also needs to be located). By gathering more data about a node's location through its neighboring nodes, collaborative positioning has two great potential advantages over standalone positioning: (1) increased availability: it increases the capability to locate nodes that could not be located otherwise (inability to 'hear' a sufficient number of anchors); (2) increased accuracy: it provides the opportunity to improve the position accuracy of nodes that could locate themselves without collaboration.

To illustrate this point, we consider the case of GPS applications. GPS makes use of pseudo-ranges and requires at least four satellites to resolve a position. The fourth satellite is used for clock synchronization. Wide Area Augmentation System (WAAS) is an augmented version of the GPS that is only available in North America. It uses an additional ground station to improve by at least a factor of 5 the position accuracy of standard GPS systems. This indicates that being able to augment positioning systems through the use of other nearby nodes has the potential to improve GPS accuracy. This advantage is even more obvious indoors in cases where there are too few visible satellites to locate a node. In such cases, collaborating with other nodes to acquire the minimum number of communication links has an immediate advantage of increased availability. Similarly to WAAS in relation to GPS, for nodes that could locate themselves, collaborating with other nodes that could also locate themselves has a promising positive impact for improved accuracy.

Although the advantage of collaborative positioning in the sense of increased availability is easy to understand and forecast, its advantage in terms of improved accuracy is not always guaranteed. Indeed, the communication link to the collaborating node could be so noisy to the point of making the collaborative estimation accuracy worse than what could be achieved with standalone positioning. Hence, it is important to understand the conditions under which collaborative positioning outperforms standalone estimation, when it does not, and why. This is the main overall problem we intend to address in this research.

Numerous aspects of collaborative positioning have been covered in the literature; however, they mostly target how to solve (from an estimation theoretic perspective) the collaborative problem and describe the performance of the resulting algorithm. Reference [78] introduces two approaches for solving the collaborative problem and concludes that the maximum likelihood estimator (MLE) is more appropriate. References [22–26] propose different approaches for the MLE problem, but tend to identify solutions that are not proven optimal in the MLE sense. Jia and Buehrer [14] then introduce a branch-and-bound solution to identify the global optimum, the MLE solution.

In regard to assessing the performance of position estimators, it is common practice to compute the mean squared error (MSE) of position (PMSE). However, this information alone only indicates how inaccurate an estimator is; it does not provide information on how well the estimator performed given the data at hand. The Cramer-Rao Lower Bound (CRLB, or CRB) has been defined as the lower bound to the variance of any estimator that yields on average the true position (also called an unbiased estimator). Since PMSE is the sum of the squared bias and the variance of an estimator, if an estimator is unbiased, PMSE reduces to just the variance that will be greater than or equal to the CRB. When the variance of an unbiased estimator reaches the CRB, the estimator is described as *efficient*. The Cramer Rao Bound is equal to the trace of the inverse of the Fisher Information Matrix (FIM) which is a measure of information on the location to estimate that can be extracted from the distributions (or pdfs) of observed range measurements.

There are two very common metrics used to measure information gained between distribution functions: a) The Fisher Information, which is a metric derived from the FIM, and b) the Kullback-Leibler Divergence, that we define later. We introduce two concepts that define the foundation of these information metrics: uncertainty and inaccuracy. Uncertainty has to do with the vagueness of an information while inaccuracy has to do with its incorrectness. To put it in more simple terms, we use Bell's simplified distinction between these terms [98]. If a confidence level is defined as 'the number expressing the degree of confidence in a result', then uncertainty has to do with the 'quantified doubt about the result of a measurement', and accuracy has to do with the 'closeness of the agreement between measurement result and true value'.

When it comes to comparing the performance of collaborative vs standalone positioning, it has been proven that the Cramer-Rao Bound (CRB) of position [1] using a set of collaborative range measurements, under some minor connectivity conditions, is smaller than the CRB using standalone range measurements [18], [12, 20]. Schloemann and Buehrer [13] introduce results on FIM-based summary statistics that can help identify if it is worth using collaborative positioning over standalone. Although these metrics perform well in line-of-sight (LOS) conditions, they do not in non-line-of-sight (NLOS) conditions, and duly so because the CRB, and other FIM-based statistics in general tend to apply to unbiased estimators [1] under stringent conditions. Another analysis on the value of collaborative estimation is presented by Schloemann and Buehrer [17] which addresses the fact that the CRB for collaborative positioning is also reduced in NLOS conditions when compared against the standalone positioning CRB. The main limitation of these theoretical analyses of the performance of collaborative positioning is that they have focused more on the Cramer Rao Bound, which is a measure of uncertainty rather than inaccuracy; although both uncertainty and inaccuracy can be equivalent when unbiased efficient estimators are used, (*i.e*, for the case of unbiased estimators that attain the CRB). Furthermore, analyses on factors that impact the accuracy of collaborative positioning in the literature have mainly been based on simulations rather than theory.

Having identified the limitations of using the CRB to address theoretically the performance of collaborative positioning, more recently, Schloemann, Dhillon and Buehrer [19] introduced an early theoretical model whose intent is to track the change in accuracy using stochastic geometry. This is a promising model that could provide more insight into identifying theoretically how geometric configurations affect changes in accuracy due to collaboration.

The state-of-the art contribution of this research is in the *theoretical* analysis of the improvement in accuracy due to collaboration from an information theoretic perspective. The essential result from this research is that we prove theoretically that even with a reduced CRB due to the acquisition of a collaborative link, it is still not guaranteed that collaboration will improve accuracy. We derive factors that affect the improvement in accuracy due to collaboration and also range conditions when collaboration starts to hurt standalone estimation. Then, we introduce criteria that assess ahead of time whether collaboration will be beneficial in terms of accuracy or not.

Using information theory, we explain the *difference* between uncertainty and inaccuracy and justify why analyzing both are needed to gain a better understanding of the performance of collaborative positioning. Although a reduced CRB was proven, an intuitive understanding for *why* the reduction in uncertainty occurs has yet to be provided. In this work, we rely on information theoretic concepts to address the *why*.

We establish that getting more information about a node location via collaboration only reduces its uncertainty. It does not always indicate that accuracy will be improved. Via theoretical analysis, we present 1) an uncertainty analysis of collaborative positioning and 2) an analysis of its accuracy. Rather than using the CRB or FIM, we use an information theoretic tool that can 1) measure the departure of inaccuracy from uncertainty, 2) measure the reduction in *uncertainty*, and 3) as we uncovered in this research, measure the improvement in *accuracy* due to collaboration. This tool, the Kullback-Leibler Divergence [10], is introduced in more detail, and a discussion on its suitability for our analysis is included. We show that in differential geometry, the Fisher Information Matrix defines a metric that represents a local version of the Kullback-Leibler Divergence (KLD) [10,27]. This shows that using KLD complements and extends the CRB analysis by removing all stringent conditions that traditionally made the FIM and CRB less practical.

Since the position estimator selected is a key contributor to the results from our theoretical analysis, we base our analysis using two position estimators, both linked to the Maximum Likelihood Estimator because it is the recommended estimator in collaborative positioning systems: the Newton-Raphson (NR) Maximum Likelihood Estimator(MLE), a recursive estimator, [67] and Torrieri's Linear estimator, which is a one-shot version of the MLE estimator [33].

1.2 Motivation for Selecting the Kullback-Leibler Divergence

The selection of the Kullback-Leibler Divergence as a tool for our theoretical analysis is justified by the fact that it has been used for more than 60 years for its sound foundation on decision-choice theory in numerous fields. One of the concerns that motivated the original paper by Kullback and Leibler [10] was to define a measure of the 'distance' between statistical populations in terms of information, in the Shannon's sense. The goal was to answer the question that any statistician, given two populations could ask: 'How difficult is it to discriminate between them with the best test?' The fact that this problem can be defined in almost all fields of study has justified the wide use of this tool over the years, and is also the reason why we consider it to compare standalone to collaborative positioning.

In macroeconomics, central banks and supervisory authorities derive models that they use to predict the future state of the economy. And when it comes to predicting future events (*e.g*: crash in the stock market) it is easy to be overwhelmed by the choice of models and their efficiency. The trivial issue in economics is that once an event occurs, it becomes the truth, and models that did not predict it accurately can then be considered less reliable. The not-so-trivial problem is that economists would really like to know ahead of time which model to trust for their decision making process, so that better policies can be put in place to avert upcoming problems. Therefore, the decision making process in model selection must account for a lot more parameters than just the average performance of the economy. The decision-making tool needs to simultaneously look at many events and account for the effect of surprises and not be biased by a surprising event. One tool that these institutions mostly rely on is the Kullback-Leibler Divergence for its decision-theoretic capabilities as mentioned in [34]. This capability to thoroughly make a decision by accounting for more than the average behavior of distributions, is the main reason that motivated our choice of KLD for this research and that warrants its popularity in decision-choice theory.

In addition to decision-choice theory, KLD has been used intensively in statistics, signal processing, pattern recognition, and coding theory, to name a few. For instance, reference [36] uses it for speech and image recognition applications. Burnham, et al., refer to it as an extension of likelihood theory that avoids the pitfalls of null hypothesis testing [42]. KLD has been used in numerous parametric estimation problems such as reference [43] where it helps identify if the relative maximum of a log-likelihood function is a global maximum. It has also been used in estimating Gaussian Mixture Model parameters [36]. KLD has been applied as a performance metric when designing detectors and more generally with hypothesis tests ([35], [37–39]). It has also been used in source coding, as a penalty term in Shannon coding [35, 40]. In Bayesian econometrics, it is considered in dynamic stochastic general equilibrium (DSGE) models as a sound tool for decision-choice theory [34]. Reference [41] uses KLD to measure the loss (or difference) between a true model and a derived model in Bayesian analysis. Hence, in our context, using the KLD to address the differences between standalone and collaborative positioning is an approach worth considering.

As a discriminating tool between two pdfs, the KLD has a great advantage when we want to compare the performance of collaborative and standalone estimators, because it compares the shapes for the pdfs of both standalone and collaborative estimates, hence encompassing all the statistical information that can be known about each distribution in the process, especially higher moments beyond the variance. In contrast with other performance metrics like MSE which tend to focus only on the first two moments (thus assuming Gaussian pdfs because the mean and variance describe fully a Gaussian distribution), KLD is, as we prove later in this paper, a more reliable tool than MSE, especially when the pdfs compared are asymmetric (that is, non-Gaussian). We also show that when one wishes to the performance of standalone and collaborative estimators, KLD proves itself more useful in practice, than the FIM and CRB.

When one compares the joint density to the product of marginal densities using KLD, one gets a particular KLD called *mutual information* [7, 9]. Mutual information provides the linear and non-linear dependence between two variables [49]. Cover, et al., define it as *a measure of the amount of information one random variable contains about another...the reduction in uncertainty of one random variable due to the knowledge of the other.* Mutual information tends to be even more popular than KLD in the literature. It has been used extensively in cases where one needs to know the dependencies between two pdfs. It has been thoroughly studied for multivariate normal distributions [51] and for particular cases like skew-elliptical distributions([45, 46]). Reference [48] uses mutual information to reconstruct the more-than-pair-wise interractions between molecules in biological networks. It has been used in natural language processing applications ([52, 53]), neuroscience ([54–56]), data fusion [57], and multiple sensor platforms [60]. In Simultaneous Localization and Mapping (SLAM) applications [59], which consist of a single node (a robot in most applications) attempting to locate itself while defining a map of its surrounding environment, mutual information is maximized as a cost function to minimize uncertainty [58]. It has been used as a bound to mean square error in Gaussian channels ([50, 61]). It has also been used in even more applications pertaining to MIMO and communication channels ([44, 61–65]).

1.3 Organization of Dissertation

The goal of this research is to use the Kullback-Leibler Divergence to assess the performance of collaborative positioning in comparison to the standalone estimator. The technical content of this research is organized as follows.

In Chapter 2, some background is provided on the different terminologies used in this work. We introduce information theoretic concepts and relate them to estimation theory. The distinction between uncertainty and inaccuracy is important to assess the performance of collaborative positioning adequately. We introduce a well-known information theoretic tool, the Kullback-Leibler Divergence. We briefly introduce how the Kullback-Leibler Divergence and its relationship with uncertainty and inaccuracy can be made applicable to the context of collaborative position estimation.

Chapter 3 introduces a novel performance metric that is used to detect the most accurate

estimator better than a comparison of position MSEs (PMSE) in the presence of heavy tails in the error distribution. A heavy tail in the distribution of position error occurs when an estimator tends to yield estimates that form a cluster of outliers such that a representation of the distribution (or pdf) of position error yields an asymmetric shape with a longer tail; hence the term *heavy tail*. Heavy tails are mainly caused by the presence of errors in the range measurements (*e.g*: NLOS errors). Heavy tails have been known as a case that PMSE (as a performance metric) tends not to assess well in general as mentioned by Zekavat and Buehrer [67]. This is because PMSE only considers the variance and bias of a distribution (which are defined as the first two moments), and is easily offset by the presence of just one outlier in the set of position estimates. In practical positioning systems, estimates tend to be averaged over time and outliers tend to be discarded to minimize the offset in the PMSE evaluation. A new problem arises: 'where is the cutoff point when discarding outliers?' The answer to this question has been dynamic over the years as the FCC continues to increase the limits of regions of interest from 67^{th} to 95^{th} percentile.
Knowing that regardless of FCC's cutoff limit, PMSE doesn't handle well data beyond the average behavior, we introduce in this chapter a novel performance metric that provides a solution to the problem of needing a cutoff point in the performance assessment when using PMSE. By accounting for all statistical information without pre-processing the data, the main contribution of this performance metric is that it allows the original data to be used without pre-processing. This is practical in the context of field-testing, but it is also practical for theoretical analysis because the original data is used as-is.

Chapter 4 defines our theoretical analysis framework and contains a theoretical analysis of the inaccuracy of collaborative positioning with respect to (w.r.t) standalone positioning. Using the performance metric introduced in the previous chapter, we derive a theoretical model of the improvement in accuracy due to collaboration. This model allows us to assess theoretically when collaborating starts to hurt the performance of standalone estimation. It allows us to prove theoretically, that although there is a benefit to collaboration, collaboration is not always accurate. This had not yet been addressed in the literature *theoretically*. It was only done via simulations.

Chapter 5 presents a theoretical analysis of the uncertainty of collaborative positioning. In this chapter, we explain why collaboration reduces the CRB and introduce an observable metric that can indicate the maximum gain in accuracy due to collaboration. The impact of this finding is that we use observable information to assess the maximum improvement in accuracy that we can get, if one chooses to collaborate. This has applications in numerous fields as it provides means to assess ahead of time whether one should collaborate or not.

Chapter 6 introduces criteria to identify on-the-fly whether it is worth collaborating or not with the intent to improve accuracy. The practical implication of this chapter are that it provides criteria that we have proven from our theoretical analysis to be valid observable indicators that collaborating with a given node will be beneficial in terms of accuracy. Knowing that poor geometric conditions and NLOS conditions can degrade the performance of acceptable estimators, the criteria are introduced and tested in both LOS and NLOS range conditions and in both good and poor geometric conditions as a means to validate and make efficient use of the theoretical analysis introduced in this research.

Chapter 7 includes a summary of contributions and addresses future research considerations.

Chapter 2

Background

When considering the performance of collaborative positioning, most theoretical analyses have focused on its *uncertainty* by analyzing the Cramer-Rao Lower Bound (CRB or CRLB). In regard to the *accuracy* of collaborative positioning with respect to (w.r.t) standalone positioning most analyses have been completed through simulations. In this chapter, we introduce concepts and tools that allow us to: 1) *theoretically* analyze the performance of collaborative positioning in terms of accuracy and 2) extend the CRB analysis.

The Cramer-Rao Bound (CRB) is a lower bound on the variance of an unbiased estimator. It is the trace of the inverse of the Fisher Information Matrix (FIM) which is a measure of how much information a data set provides about a parameter ([88], [89]). A CRB analysis of collaborative estimation by Schloemann and Buehrer has indicated that the collaborative CRB is less than the standalone CRB in the multivariate case [18]. For unbiased estimators, the CRB is a measure of uncertainty. It is a lower bound on the inaccuracy (mean squared error - MSE) of an unbiased estimator. Thus it is different from inaccuracy, although for efficient estimators, MSE reduces to the CRB. Therefore, a better understanding of uncertainty and inaccuracy is needed to get a deeper insight into what $CRB_{COL} \leq CRB_{STD}$ means, and why this inequality does not necessarily indicate that collaborative estimation will be more accurate than standalone positioning.

In this chapter, we define our analytical framework and explain the difference between two concepts: uncertainty and inaccuracy. Their mathematical difference yields the Kullback-Leibler Divergence (KLD), which is the main tool used in this research to analyze theoretically the performance of collaborative positioning w.r.t standalone positioning. The Kullback-Leibler Divergence and one of its special forms, mutual information, are introduced along with their applications. The Kullback-Leibler Divergence and the Fisher Information Matrix are related. The key relationship between them is also derived and interpreted.

2.1 Analytical framework

Our goal is to analyze the performance (in terms of accuracy) of the collaborative position estimator with respect to (w.r.t) the standalone position estimator when estimating the position vectors of two nodes $\mathbf{x}_A = [x_A, y_A]^T$ and $\mathbf{x}_B = [x_B, y_B]^T$ both fixed points in a two-dimensional Euclidean space using time of arrival (TOA). The results defined in this chapter are not restricted to neither TOA nor a two-dimensional space. We assume that all receivers and transmitters are synchronized in time, based on synchronization techniques [4, 5]. We also assume that at any time one needs to locate the two nodes, they are within communication range of their anchors. The position estimators considered in this chapter are: the Newton-Raphson (NR) Maximum Likelihood Estimator (MLE) [67] and Torrieri's Linear estimator, which is a one-shot version of the MLE estimator [33] described in Appendix 2.A.

In the standalone configuration, we assume each point can locate itself using three sensors of known fixed locations, specifically, we define anchors $\mathbf{x}_{Ai} = [x_{Ai}, y_{Ai}]^T$, $i = \{1, 2, 3\}$ for locating $\mathbf{x}_A$. Similarly, $\mathbf{x}_B$ can be located using anchors $\mathbf{x}_{Bj} = [x_{Bj}, y_{Bj}]^T$, $j = \{1, 2, 3\}$.

We consider both line-of-sight (LOS) and non-line-of-sight (NLOS) errors on all range measurements. In LOS, a zero-mean Gaussian distributed error $N(0, \sigma^2)$ is added to the true range measurements. In NLOS, an additional exponentially-distributed bias is added to all range measurements. For theoretical analyses, we assume that all range measurements are in LOS conditions. Given the true distance d between a point and its anchor, the variance σ^2 of observed range measurements is dependent on their signal-to-noise (SNR) ratio as follows:

$$SNR_{dB} = 10log_{10}\left(\frac{d^2}{\sigma^2}\right) \Longrightarrow \sigma^2 = \frac{d^2}{10^{\left(\frac{SNR_{dB}}{10}\right)}} \tag{2.1}$$

The set of range measurements observed between, for instance, $\mathbf{x}_A$ and its i^{th} anchor $\mathbf{x}_{Ai}$ is:

$$r_{Ai} = \sqrt{(x_{Ai} - x_A)^2 + (y_{Ai} - y_A)^2} + n_{Ai} + b_{Ai} \tag{2.2}$$

where $n_{Ai} \sim N(0, \sigma^2_{Ai})$ is Gaussian distributed with zero-mean and variance σ^2_{Ai}. The positive bias b_{Ai} is zero in LOS conditions. In our analytical framework, we assume LOS range conditions. During simulations, we also test for the presence of an NLOS bias in the range measurements. In NLOS conditions (for simulations only), we assume b_{Ai} has a pdf that is exponentially-distributed with pdf $f_B(b) = \lambda \exp(-\lambda b)$, and mean λ^{-1}, where λ is the parameter of the exponential distribution.

In collaborative positioning, an additional range measurement is acquired between $\mathbf{x}_A$ and $\mathbf{x}_B$ to simultaneously solve for the positions of both nodes. We call this additional range measurement *the collaborative link* r_{AB}. We assume that the observed range measurements on the collaborative link are symmetrical: $r_{AB} = r_{BA}$.

2.2 Uncertainty vs. Inaccuracy

In this section, we first define uncertainty and inaccuracy in general terms and illustrate their definitions in our research context. Kerridge [72] established that 'there are two ways to define the lack in precision in the occurrence of a statistical event: either a) there is not enough information - uncertainty of the information, or b) some of the information gathered is incorrect - inaccuracy of the information'. Thus, uncertainty has to do with the sufficiency of the data available to estimate a parameter. It indicates the vagueness of the estimation. Uncertainty can also be understood from the perspective of the law of large numbers: the more observables we acquire about a parameter, the more its uncertainty reduces. A large reduction in uncertainty eventually leads to more confidence in the accuracy of the data to be estimated, if the right estimator is used.

Inaccuracy has to do with the incorrectness of a set of estimates. It addresses how correct an estimator is. Given a set of range measurements that have the same uncertainty, two estimators can yield two very different position estimates out of the same set of observed range measurements: the 'inaccurate' estimator yields estimates that are farther from the true value than the 'accurate' estimator. The choice of the estimator is one example that can affect accuracy.

By definition, *Shannon's entropy*, or uncertainty, of any discrete random variable X described by its probability mass function (pmf) $p(x)$ is a measure of the randomness of a discrete random variable X with values $\{x_1, x_2, ..., x_n\}$:

$$H(p_X) = -\sum_{i=1}^{n} p_X(x_i) \log p(x_i) \tag{2.3}$$

It is important to note that using the terminology *entropy* in the case of continuous pdfs should be strictly referred to as *differential entropy*, because Shannon's entropy was originally applicable to discrete functions. The error between differential entropy and Shannon's entropy for discrete pmfs has been discussed in [7].
The cross-entropy of two pmfs $p(x)$ and $q(x)$ describing a discrete random variable X with values $\{x_1, x_2, ...x_n\}$ is:

$$H(p_X||q_X) = -\sum_{i=1}^{n} p(x_i) \log q(x_i) \tag{2.4}$$

Per Kerridge [72], cross-entropy can be defined as a measure of inaccuracy, or incorrectness of the information in pmf $q(x)$ if the data in pmf $p(x)$ is a reference for the best definition of accuracy available. For instance, if $q(x)$ describes a new model being tested that indicates the distribution of outcomes for a value on the stock market, $q(x)$ could be the distribution of outcomes from a better reference model that has been used in the past and that is known to be trustworthy. The cross-entropy of $q(x)$ in reference to $p(x)$ would measure the inaccuracy of $q(x)$ given that the reference for the truth is $p(x)$ based on the data at hand.

Nath then extended Kerridge's results to the continuous case [73, 74]. Cross-entropy has been used in the literature as an alternative to squared error, for example in neural networks applications [70,75]. Thus, going forward, we understand that cross-entropy is also a measure of inaccuracy if the reference pdf reflects accuracy based on the data at hand.
The Shannon-Gibbs inequality is defined as:

$$H(p_X) \leq H(p_X||q_X) \tag{2.5}$$

From the Shannon-Gibbs inequality, inaccuracy should reduce to uncertainty when $q(x) = p(x)$. This means that when estimating a parameter, one can only be as accurate as the vagueness of the information provided on average. Reference [71] provides a survey on how uncertainty and inaccuracy connect with each other.

2.2.1 In our research context

Looking at uncertainty and inaccuracy in the context of our research, uncertainty is based on the errors in the set of observed range measurements. Say we have a set of range measurements that we use to estimate a position. We pick two position estimators. Each of them uses the same set of observed range measurements. Hence, they use data that is based on the same uncertainty. The inaccurate estimator is the one that yields answers that are farthest from the true position.

Uncertainty does not depend on a position estimator, but inaccuracy does. Uncertainty is based on the noise in the range measurements that we acquire. If we get very noisy range measurements, we will end up with a larger uncertainty than if we had less noisy range measurements. Inaccuracy on the other hand depends on the position estimator selected. If you have two position estimators, the estimator that gives position estimates closest to the truth is the one that is most accurate. The estimator that gives answers farthest from the true position is the least accurate.

Uncertainty characterizes the noise in the range measurements regardless of the estimator that you pick. The noise that you have in your range measurements bounds the capabilities for your estimator to accurately make an assessment. This bounding effect describes in essence the Shannon-Gibb's inequality which states that inaccuracy is bounded by uncertainty. As mentioned in the next section, the mathematical positive difference between uncertainty and inaccuracy represents the Kullback-Leibler Divergence.

2.3 Kullback-Leibler Divergence

Two properties of cross-entropy, as a measure of inaccuracy, are that 1) it is minimum when both pdfs are equal, and 2), 'At any instant, the departure of inaccuracy from entropy represents the Kullback-Leibler Divergence' [73]. Therefore, both entropy (uncertainty) and cross-entropy (inaccuracy) are combined into the Kullback-Leibler Divergence as follows:

$$KLD(p_X||q_X) = H(p_X||q_X) - H(p_X) \tag{2.6}$$

The Kullback-Leibler Divergence (KLD) was introduced in 1951 by Kullback and Leibler as a means to measure, 'with the best test', how different a pdf is from a reference pdf [10]. As such, it is a discrimination tool. Comparing a probability density function $g_X(x)$ against a reference pdf $f_X(x)$, KLD is defined as:

$$KLD(f_X||g_X) = \int_x f_X(x) \log \frac{f_X(x)}{g_X(x)} dx \tag{2.7}$$

2.3.1 Properties

Sign of KLD

From Shannon-Gibb's inequality, we can conclude that the Kullback-Leibler Divergence is always positive or equal to zero. Equality only holds when the two pdfs being compared are equal.

$$0 \leq KLD(f_X||g_X) \tag{2.8}$$

Asymmetry of KLD

Another property of KLD is that it is not a symmetric metric in general, that is:

$$KLD(f_X||g_X) \neq KLD(g_X||f_X) \tag{2.9}$$

Triangle Inequality

The Kullback-Leibler divergence is not called a 'distance' mainly because it does not satisfy the triangle inequality. That is:

$$KLD(f_X||g_X) + KLD(g_X||h_X) \ngeq KLD(f_X||h_X) \tag{2.10}$$

2.3.2 Applications

The selection of the Kullback-Leibler Divergence as a tool to analyze theoretically factors that make the performance of collaborative positioning differ from the performance of standalone positioning is justified by the fact that this tool has been used for more than 65 years, for its sound foundation on decision-choice theory in numerous fields. In our application, we use the robust capability of KLD to make decisions, to generate a performance metric and a theoretical model for our analysis of the accuracy of collaborative positioning.

One of the concerns that motivated the original paper by Kullback and Leibler [10] was to define the 'distance' between statistical populations in terms of information, in the Shannon's sense. The goal was to answer the question that any statistician, given two populations could ask: 'How difficult is it to discriminate between them with the best test?' The fact that this problem can be defined in almost all fields of study has justified the wide use of this tool over the years, and is also the reason why we consider it to compare standalone to collaborative positioning.

In fact, KLD has been used in statistics, signal processing, pattern recognition, coding theory, econometrics, to name a few. By being a discrimination tool between two pdfs, KLD is a useful tool to evaluate the difference in performance between collaborative and standalone estimators. It compares the shapes of the pdfs of both standalone and collaborative estimates, encompassing all the statistical information that can be known about each distribution in the process, especially higher moments beyond the variance. In our research, we use this powerful discriminating feature to: 1) derive a robust performance metric, 2) identify factors that affect the improvement in the accuracy of collaborative positioning w.r.t standalone, 3) identify range conditions within which collaboration starts to hurt the accuracy of standalone estimation.

2.4 Mutual Information

When one compares the product of marginal densities $f_X(x)f_Y(y)$ to the joint density $f_{XY}(x,y)$ using KLD, one gets a particular KLD called *mutual information* (MI) [7, 9]:

$$MI(X,Y) = KLD(f_{XY}(x,y)||f_X(x)f_Y(y)) \tag{2.11}$$

Since mutual information is a Kullback-Leibler Divergence, it is positive. Unlike KLD, MI is symmetric (see Appendix 2.C), which means that:

$$MI(X,Y) = MI(Y,X) \tag{2.12}$$

Let $H(f_X)$ be the uncertainty (or Shannon's entropy) of the pdf $f_X(x)$ of a random variable X:

$$H(f_X) = \int_x f_X(x) \log \frac{1}{f_X(x)} dx \tag{2.13}$$

Given the definition of the Kullback Leibler Divergence, the mutual information between two variables X and Y is the reduction in uncertainty of X from knowing about Y and vice-versa [7]:

$$\begin{aligned} MI(X,Y) &= H(f_X) + H(f_Y) - H(f_{XY}) \\ &= H(f_X) - H(f_{X|Y}) \\ &= H(f_Y) - H(f_{Y|X}) \\ &\geq 0 \end{aligned} \tag{2.14}$$

where the conditional entropy is given by:

$$H(f_{X|Y}) = H(f_{XY}) - H(f_Y) \tag{2.15}$$

and the joint entropy of X and Y is the entropy of their joint distribution $f_{(XY)}(x,y)$:

$$H(f_{XY}) = \int_x \int_y f_{XY}(x,y) \log \frac{1}{f_{XY}(x,y)} dy dx \tag{2.16}$$

Another characteristic of mutual information is that it provides the linear and non-linear dependence between two variables [49]. In doing so, it has an advantage over linear correlation. In fact, we know that if two variables are independent, their linear correlation coefficient is zero. However, we also know that the converse does not hold: that is, if the correlation coefficient is zero, we cannot automatically conclude that both variables are independent, unless the two variables are Gaussian distributed [27]. Using mutual information addresses this limitation of linear correlations. That is, if two variables are independent, their mutual information is zero. Conversely, if mutual information is zero, then, both variables have to be independent. The converse is true because: 1) mutual information is defined using KLD, which only yields zero when both pdfs compared are equal, and 2) mutual information accounts for non-linear correlation.

Mutual information tends to be even more popular than KLD in the literature, as mentioned in the previous chapter. In our research context, using KLD to compare the pdf of collaborative estimates to the pdf of standalone estimates is very similar to deriving their mutual information. This special form of KLD is another tool that we later use to explain why the collaborative CRB is less than the standalone CRB. It is important to note that for multidimensional variables, mutual information can be negative because it accounts for all dependencies across all variables, which can yield unpredictable results that can be very complex to interpret. For 2 and 3-dimensional positioning systems with pdfs belonging to the multivariate elliptical family of distributions, which contains the multivariate Gaussian distribution, mutual information is derived by [46]. Additionally, mutual information is an average measure, which implies that a point-wise (non-average) computation of mutual information can also be negative.

2.5 Fisher Information Matrix vs. Kullback-Leibler Divergence

The Fisher Information Matrix (FIM) is used intensively in the literature because the trace of its inverse yields a lower bound on the variance of any unbiased estimator, namely, the Cramer-Rao Lower Bound (CRB). In their original paper, Kullback and Leibler [10] established that the Kullback-Leibler Divergence (KLD) and the Fisher Information Matrix (FIM) are related. However, as Zacks [89] indicated, they are designed to fulfill different roles. The FIM tends to be used in estimation problems, while KLD tends to be used in decision theory. In our application, KLD is considered a reliable statistical tool that bases its selection of the most accurate position estimator on all statistical information provided to make a decision. Additionally, many other references [27,91] consider the FIM a local version of the KLD. Numerous papers have used the relation between KLD and FIM as a bridge to connect estimation theory to information theory. For instance, reference [50] relates mutual information to the MMSE in Gaussian channels.

In this section, we show how the FIM and KLD are related and describe how they differ in the information they communicate. With this, we intend to show that KLD is a better tool than FIM for theoretically analyzing the performance of collaborative positioning w.r.t standalone positioning in general. If we consider the likelihood function for a parameter θ, $f(x;\theta)$, and another likelihood function $f(x;\theta+\epsilon)$, whose parameter differs from $f(x;\theta)$ by a small amount ϵ, the KLD between these two distributions is proportional to the FIM as derived in Appendix 2.B:

$$\begin{aligned} KLD(f(x;\theta)||f(x;\theta+\epsilon)) &= -\frac{\epsilon^2}{2}E_{f(x;\theta)}[\frac{\partial^2}{\partial\theta^2}\log f(x,\theta)] \\ &= \frac{\epsilon^2}{2}I_{Fisher} \end{aligned} \tag{2.17}$$

The FIM is a KLD between two likelihood functions of the same family, whose parameters are very close, up to a factor. This relationship provides deeper insight into what the FIM and CRB mean, and how they differ from the information provided by KLD:

- The FIM measures the curvature of the log-likelihood function. This makes the CRB (which is derived from the inverse of the FIM) the radius of curvature of the likelihood function [1]. Thus, when the likelihood function is sharp, its radius of curvature, the CRB, is reduced.

- KLD measures the relative entropy for using $f(x;\theta+\epsilon)$ rather than $f(x;\theta)$. In an estimation problem, the goal is to minimize ϵ. Minimizing ϵ yields efficient estimators. Thus the FIM is a good measure of how well we can estimate θ if we get $f(x;\theta+\epsilon)$ as long as ϵ is small. If ϵ increases, the FIM is no longer adequate.

- Putting this into context for our research problem on collaborative positioning: using the CRB to assess the performance of collaborative positioning does not tell the entire story. The FIM and CRB only assess performance adequately when the likelihood functions being considered are very close to the most accurate likelihood function (centered at the parameter we want to estimate). This explains why the CRB makes a good assessment on the lower bound of the inaccuracy (mean squared error) of unbiased estimators (that is, estimators that yield on average the true position), but does not assess well a lower bound on the inaccuracy of biased estimators (*e.g:* in biased estimators, the mean squared error can be lower than the CRB).

- The FIM is an approximation of KLD using a second order Taylor series. As the parameters of the likelihood functions go farther apart (as ϵ increases), an approximation to the second degree is no longer valid and the FIM no longer measures adequately the information gained between the distributions. KLD becomes the better-suited tool. It includes both linear and non-linear discrimination data.

2.6 Summary

In this chapter, we explained the difference between uncertainty and inaccuracy, both from a conceptual perspective and mathematically. Uncertainty (or entropy) has to do with the lack of confidence (or doubt) in the information at hand. Inaccuracy deals with the incorrectness of the information. Mathematically, the positive difference between inaccuracy and uncertainty is called the Kullback-Leibler Divergence (or relative entropy). It is a powerful tool that can be used to assess the discriminating power between likelihood functions.

KLD is a suitable, theoretically-tractable tool for three main reasons. First, it relates uncertainty to inaccuracy, giving us a tool to assess inaccuracy relative to a reference for the best accuracy achievable. Second, through mutual information, it is used to measure the reduction in uncertainty (information gain) due to collaboration. Third, as we show in this research, it is also a suitable tool to measure the reduction in inaccuracy due to collaboration.

In the next chapters, we use KLD to assess the performance of collaborative positioning w.r.t standalone positioning. First, KLD is used to derive a performance metric that we claim outperforms a comparison of position mean squared error (PMSE). Second, KLD is used to analyze the inaccuracy and uncertainty of collaborative positioning theoretically. Results from these analyses are then combined to define criteria that can tell ahead of time whether collaboration will be more accurate than standalone in both good and poor geometries, and in the lack or presence of non-line of sight (NLOS). In short using KLD to analyze the performance of collaborative positioning extends the theoretical CRB analyses available in the literature and provides a means to analyze *theoretically* the inaccuracy of collaborative positioning; analysis which was only done to-date through *simulations*.

Appendix

2.A Linear Position Estimator

2.A.1 Torrieri's Linear Position Estimator

In vector form, the set of observed range measurements from M anchors used to estimate a point (or node) can be written in the following vector format:

$$\mathbf{r} = h(\mathbf{x}) + \mathbf{n} \tag{2.18}$$

Where, $\mathbf{r}$ is a Mx1 vector and $h(\mathbf{x})$ represents the vector of the Euclidean distance between a node and its anchors.
Assuming these range measurements are independent, their joint density function is:

$$f_R(\mathbf{r}|\mathbf{x}) = [(2\pi)^{M/2} \det \mathbf{N}]^{-1} \exp\left[-(1/2)(\mathbf{r}\text{-}\mathbf{h}(\mathbf{x}))^T \mathbf{N}^{-1}(\mathbf{r}\text{-}\mathbf{h}(\mathbf{x}))\right] \tag{2.19}$$

Where $\mathbf{N}$ is the covariance matrix of the measurement errors. The solution that maximizes the pdf described above, minimizes $Q(\mathbf{x})$, the negative argument of the exponential function:

$$Q(\mathbf{x}) = (\mathbf{r}\text{-}\mathbf{h}(\mathbf{x}))^T \mathbf{N}^{-1}(\mathbf{r}\text{-}\mathbf{h}(\mathbf{x})) \tag{2.20}$$

Its estimate $\hat{\mathbf{x}}$, is obtained by setting the gradient of $Q(\mathbf{x})$ to 0. This is a non-linear Least Squares Estimator because $Q(\mathbf{x})$ is a non-linear function. To linearize the problem, we first derive the Taylor series expansion of $h(\mathbf{x})$ and disregard anything beyond the first two terms. We get:

$$h(\mathbf{x}) \cong h(\mathbf{x}_o) + \mathbf{G}(\mathbf{x} - \mathbf{x}_0) \tag{2.21}$$

Where $\mathbf{G}$ is the matrix of derivatives evaluated at a reference point $\mathbf{x}_0$. We assume that the one-shot reference point $\mathbf{x}_0$ has to be within the convex region of the global minimum.

$$\mathbf{G} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}\Big|_{x=x_0} & \cdots & \frac{\partial f_1}{\partial x_n}\Big|_{x=x_0} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_N}{\partial x_1}\Big|_{x=x_0} & \cdots & \frac{\partial f_N}{\partial x_n}\Big|_{x=x_0} \end{bmatrix} \tag{2.22}$$

Combining equation (2.21) and equation (2.20), we get:

$$Q(\mathbf{x}) = (\mathbf{r}_1 - \mathbf{G}\mathbf{x})^T \mathbf{N}^{-1} (\mathbf{r}_1 - \mathbf{G}\mathbf{x}) \tag{2.23}$$

where

$$\mathbf{r}_1 = \mathbf{r} - h(\mathbf{x}_0) + \mathbf{G}\mathbf{x}_0 \tag{2.24}$$

Taking the gradient of the linearized $Q(\mathbf{x})$ yields:

$$\nabla_x Q(\mathbf{x}) = 2\mathbf{G}^T \mathbf{N}^{-1} \mathbf{G}\mathbf{x} - 2\mathbf{G}^T \mathbf{N}^{-1} \mathbf{r}_1 \tag{2.25}$$

Setting the gradient of $Q(\mathbf{x})$ to zero yields:

$$2\mathbf{G}^T \mathbf{N}^{-1} \mathbf{G}\hat{\mathbf{x}} - 2\mathbf{G}^T \mathbf{N}^{-1} \mathbf{r}_1 = 0 \tag{2.26}$$

Solving for $\hat{\mathbf{x}}$, we get the following linear position estimator:

$$\hat{\mathbf{x}} = \mathbf{x}_0 + (\mathbf{G}^T \mathbf{N}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{N}^{-1} (\mathbf{r} - h(\mathbf{x}_0)) \tag{2.27}$$

2.A.2 Mean and bias of linear position estimator

The mean of the linear estimator derived in the previous equation is:

$$E[\hat{\mathbf{x}}] = E[\mathbf{x}_0] + E\left[(\mathbf{G}^T \mathbf{N}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{N}^{-1} (\mathbf{r} - h(\mathbf{x}_0))\right] \tag{2.28}$$

After rewriting the expression for $\hat{\mathbf{x}}$ as follow:

$$\hat{\mathbf{x}} = \mathbf{x} + (\mathbf{G}^T \mathbf{N}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{N}^{-1} [(h(\mathbf{x}) - h(\mathbf{x}_0) - \mathbf{G}(\mathbf{x} - \mathbf{x}_0) + \mathbf{n}] \tag{2.29}$$

The bias of the linear estimator derived is:

$$\mathbf{b} = E[\hat{\mathbf{x}}] - \mathbf{x} = (\mathbf{G}^T \mathbf{N}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{N}^{-1} (h(\mathbf{x}) - h(\mathbf{x}_0) - \mathbf{G}(\mathbf{x} - \mathbf{x}_0) + E[\mathbf{n}]) \tag{2.30}$$

2.A.3 Covariance matrix of linear position estimator

The covariance matrix for this estimator is $\mathbf{P}$, such that:

$$\mathbf{P} = E[(\hat{\mathbf{x}} - E[\hat{\mathbf{x}}])^T (\hat{\mathbf{x}} - E[\hat{\mathbf{x}}])] \tag{2.31}$$

Given that:

$$(\hat{\mathbf{x}} - E[\hat{\mathbf{x}}]) = \mathbf{A}(\mathbf{n} - E[\mathbf{n}]) \tag{2.32}$$

where

$$\mathbf{A} = (\mathbf{G}^T \mathbf{N}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{N}^{-1} \tag{2.33}$$

And, given that $\mathbf{N}$ is symmetric, $(\mathbf{N}^{-1})^T = (\mathbf{N}^T)^{-1} = \mathbf{N}^{-1}$, $\mathbf{P}$ is equal to:

$$\begin{aligned}
\mathbf{P} &= \mathbf{ANA}^T \\
&= \left((\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{N}^{-1}\right)(\mathbf{N})\left((\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\mathbf{N}^{-1}\right)^T \\
&= (\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\left[\mathbf{G}^T\mathbf{N}^{-1}\right]^T\left[(\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})^{-1}\right]^T \\
&= \left[\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G}\right]^{-1}\left[\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G}\right]\left[(\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})^{-1}\right]^T \\
&= \left[(\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})^{-1}\right]^T \\
&= \left[(\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})\right]^{-1}
\end{aligned} \tag{2.34}$$

Hence, the covariance matrix for this linear position estimator is:

$$\mathbf{P} = \left[(\mathbf{G}^T\mathbf{N}^{-1}\mathbf{G})\right]^{-1} \tag{2.35}$$

2.A.4 Fisher Information Matrices for standalone and collaborative estimation

The Cramer-Rao lower bound is the smallest variance achievable by any unbiased estimator given a set of range measurements. Since we are considering two groups of observed range measurements, namely the standalone set and the collaborative set, there exist two Cramer-Rao Bounds, one for each group of observed range measurements: CRB_{COL} for collaborative positioning and CRB_{STD} for standalone positioning. The CRB can be derived once we obtain the inverse of the Fisher information matrix for each set of range measurements.

Fisher Information Matrix (FIM) derivation for Collaborative

The conditional pdfs of the range observed between position vector $\mathbf{x}_A$ and its i^{th} anchor, and vector $\mathbf{x}_B$ and its j^{th} anchor are:

$$f_R(r_{Ai}|\mathbf{x}_A) = \frac{1}{\sqrt{2\pi\sigma^2_{AAi}}}\exp\left(-\frac{(r_{Ai} - h(\mathbf{x}_{Ai}))^2}{2\sigma^2_{AAi}}\right) \tag{2.36}$$

$$f_R(r_{Bj}|\mathbf{x}_B) = \frac{1}{\sqrt{2\pi\sigma^2_{BBj}}}\exp\left(-\frac{(r_{Bj} - h(\mathbf{x}_{Bj}))^2}{2\sigma^2_{BBj}}\right) \tag{2.37}$$

With observed range measurements corrupted by independent and Gaussian-distributed noise, the joint pdf of all observed range measurements is the product of pdfs of individual range measurements. The standalone pdf of observed range measurements is:

$$f_{RSTD}(r|\mathbf{x}_A, \mathbf{x}_B) = \prod_{i=1}^{3} f_R(r_{Ai}|\mathbf{x}_A)\prod_{j=1}^{3} f_R(r_{Bj}|\mathbf{x}_B) \tag{2.38}$$

If $f_R(r_{AB}|\mathbf{x}_A,\mathbf{x}_B)$ is the pdf of the range on the collaborative link, the joint pdf of observed range measurements in the collaborative positioning system is:

$$f_{RCOL}(r|\mathbf{x}_A,\mathbf{x}_B) = f_R(r_{AB}|\mathbf{x}_A,\mathbf{x}_B)\prod_{i=1}^{3} f_R(r_{Ai}|\mathbf{x}_A)\prod_{j=1}^{3} f_R(r_{Bj}|\mathbf{x}_B) \tag{2.39}$$

The Fisher Information Matrix (FIM) is defined as [1]:

$$FIM = -E\left[\frac{(\partial^2 \ln f_R(r|\mathbf{x}))}{\partial \mathbf{x}^2}\right] \tag{2.40}$$

The FIM can be derived to be:

$$FIM = \begin{bmatrix} \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial {x_A}^2} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_A \partial y_A} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_A \partial x_B} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_A \partial y_B} \\ \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_A \partial y_A} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial {y_A}^2} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_B \partial y_A} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial y_A \partial y_B} \\ \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_A \partial x_B} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial y_A \partial x_B} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial {x_B}^2} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_B \partial y_B} \\ \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_A \partial y_B} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial y_A \partial y_B} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial x_B \partial y_B} & \frac{\partial^2 \ln f_R(r|\mathbf{x}_A,\mathbf{x}_B)}{\partial {y_B}^2} \end{bmatrix} \tag{2.41}$$

Taking the natural log of $f_{RCOL}(r|\mathbf{x}_A,\mathbf{x}_B)$, we get:

$$\ln f_{RCOL}(r|\mathbf{x}_A,\mathbf{x}_B) = \ln f_R(r_{AB}|\mathbf{x}_A,\mathbf{x}_B) + \sum_{i=1}^{3}\ln f_R(r_{Ai}|\mathbf{x}_A) + \sum_{j=1}^{3}\ln f_R(r_{Bj}|\mathbf{x}_B) \tag{2.42}$$

Thus the FIM for the set of collaborative range measurements, after computations yields:

$$\mathbf{FIM}_{COL} = \begin{bmatrix} \mathbf{\Sigma}_{FACOL} & \mathbf{\Sigma}_{FAB} \\ \mathbf{\Sigma}_{FAB}^T & \mathbf{\Sigma}_{FBCOL} \end{bmatrix} \tag{2.43}$$

where, given θ_{AA_i}, the angle between $\mathbf{x}_A$ and its i^{th} anchor $\mathbf{x}_{Ai}$, and θ_{BB_i}, the angle between $\mathbf{x}_B$ and its i^{th} anchor $\mathbf{x}_{Bi}$:

$$\mathbf{\Sigma}_{FACOL} = \begin{bmatrix} \sum_{i=1}^{3}\frac{\cos^2(\theta_{AA_i})}{\sigma^2_{AAi}} + \frac{\cos^2(\theta_{AB})}{\sigma^2_{AB}} & \sum_{i=1}^{3}\frac{\cos(\theta_{AA_i})\sin(\theta_{AA_i})}{\sigma^2_{AAi}} + \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} \\ \sum_{i=1}^{3}\frac{\cos(\theta_{AA_i})\sin(\theta_{AA_i})}{\sigma^2_{AAi}} + \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} & \sum_{i=1}^{3}\frac{\sin^2(\theta_{AA_i})}{\sigma^2_{AAi}} + \frac{\sin^2\theta_{AB}}{\sigma^2_{AB}} \end{bmatrix} \tag{2.44}$$

$$\mathbf{\Sigma}_{FBCOL} = \begin{bmatrix} \sum_{i=1}^{3}\frac{\cos^2(\theta_{BBi})}{\sigma^2_{BBi}} + \frac{\cos^2(\theta_{AB})}{\sigma^2_{AB}} & \sum_{i=1}^{3}\frac{\cos(\theta_{BBi})\sin(\theta_{BBi})}{\sigma^2_{BBi}} + \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} \\ \sum_{i=1}^{3}\frac{\cos(\theta_{BBi})\sin(\theta_{BBi})}{\sigma^2_{BBi}} + \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} & \sum_{i=1}^{3}\frac{\sin^2(\theta_{BBi})}{\sigma^2_{BBi}} + \frac{\sin^2\theta_{AB}}{\sigma^2_{AB}} \end{bmatrix} \tag{2.45}$$

$$\mathbf{\Sigma}_{FAB} = \mathbf{\Sigma}_{AB}^T = \begin{bmatrix} -\frac{\cos^2(\theta_{AB})}{\sigma^2_{AB}} & -\frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} \\ -\frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} & -\frac{\sin^2\theta_{AB}}{\sigma^2_{AB}} \end{bmatrix} \tag{2.46}$$

Fisher Information Matrix (FIM) derivation for Standalone

In the standalone case, the collaborative link no longer exists. Hence, the FIM for the standalone problem yields:

$$\mathbf{FIM}_{STD} = \begin{bmatrix} \boldsymbol{\Sigma}_{FASTD} & 0 \\ 0 & \boldsymbol{\Sigma}_{FBSTD} \end{bmatrix} \tag{2.47}$$

where:

$$\boldsymbol{\Sigma}_{FASTD} = \begin{bmatrix} \sum_{i=1}^{3} \frac{\cos^2(\theta_{AA_i})}{\sigma_{AAi}^2} & \sum_{i=1}^{3} \frac{\cos(\theta_{AA_i})\sin(\theta_{AA_i})}{\sigma_{AAi}^2} \\ \sum_{i=1}^{3} \frac{\cos(\theta_{AA_i})\sin(\theta_{AA_i})}{\sigma_{AAi}^2} & \sum_{i=1}^{3} \frac{\sin^2(\theta_{AA_i})}{\sigma_{AAi}^2} \end{bmatrix} \tag{2.48}$$

$$\boldsymbol{\Sigma}_{FBSTD} = \begin{bmatrix} \sum_{i=1}^{3} \frac{\cos^2(\theta_{BBi})}{\sigma_{BBi}^2} & \sum_{i=1}^{3} \frac{\cos(\theta_{BBi})\sin(\theta_{BBi})}{\sigma_{BBi}^2} \\ \sum_{i=1}^{3} \frac{\cos(\theta_{BBi})\sin(\theta_{BBi})}{\sigma_{BBi}^2} & \sum_{i=1}^{3} \frac{\sin^2(\theta_{BBi})}{\sigma_{BBi}^2} \end{bmatrix} \tag{2.49}$$

It is important to note that the joint pdfs of observed range measurements for standalone and collaborative only differ by the presence of the collaborative link information. Thus, using the same position estimator, we expect both estimators to yield different pdfs. Therefore, as long as a collaborative link exists, one estimator will always be more accurate than the other.

2.A.5 Distributions of position estimates

Considering we assumed Gaussian range measurements and a linear position estimator, the pdf of position estimates is a Gaussian distribution. The vector of position coordinates to estimate is:

$$\mathbf{x} = [x_A, y_A, x_B, y_B]^T \tag{2.50}$$

Therefore, the pdf for the collaborative position estimator is $f_{XCOL}(\mathbf{x}) \sim N(\boldsymbol{\mu_{COL}}, \boldsymbol{\Sigma_{COL}})$ where the mean $\boldsymbol{\mu_{COL}}$ is a 4x1 vector, and $\boldsymbol{\Sigma}_{COL}$ is a 4-x-4 matrix. Similarly, the pdf of the standalone position estimator is $f_{XSTD}(\mathbf{x}) \sim N(\boldsymbol{\mu_{STD}}, \boldsymbol{\Sigma_{STD}})$ where the mean $\boldsymbol{\mu_{STD}}$ is a 4x1 vector, and $\boldsymbol{\Sigma}_{STD}$ is a 4-x-4 matrix. From the previous sections, the mean for each estimator is:

$$\boldsymbol{\mu_{STD}} = E[\mathbf{x}_0] + E\left[(\mathbf{G}_{STD}^T \mathbf{N}_{STD}^{-1} \mathbf{G}_{STD})^{-1} \mathbf{G}_{STD}^T \mathbf{N}_{STD}^{-1} (\mathbf{r}_{STD} - f(\mathbf{x}_0))\right] \tag{2.51}$$

$$\boldsymbol{\mu_{COL}} = E[\mathbf{x}_0] + E\left[(\mathbf{G}_{COL}^T \mathbf{N}_{COL}^{-1} \mathbf{G}_{COL})^{-1} \mathbf{G}_{COL}^T \mathbf{N}_{COL}^{-1} (\mathbf{r}_{COL} - f(\mathbf{x}_0))\right] \tag{2.52}$$

where:

$$\boldsymbol{N}_{STD} = diag[\sigma_{AA1}^2, \sigma_{AA2}^2, \sigma_{AA3}^2, \sigma_{BB1}^2, \sigma_{BB2}^2, \sigma_{BB3}^2] \tag{2.53}$$

$$\boldsymbol{N}_{COL} = diag[\sigma_{AA1}^2, \sigma_{AA2}^2, \sigma_{AA3}^2, \sigma_{BB1}^2, \sigma_{BB2}^2, \sigma_{BB3}^2, \sigma_{AB}^2] \tag{2.54}$$

$$\mathbf{G}_{STD} = \begin{bmatrix} \frac{\partial h_{A_0A1}}{\partial x_A} & \frac{\partial h_{A_0A1}}{\partial yA} & 0 & 0 \\ \frac{\partial h_{A_0A2}}{\partial x_A} & \frac{\partial h_{A_0A2}}{\partial yA} & 0 & 0 \\ \frac{\partial h_{A_0A3}}{\partial x_A} & \frac{\partial h_{A_0A3}}{\partial yA} & 0 & 0 \\ 0 & 0 & \frac{\partial h_{B_0B1}}{\partial x_B} & \frac{\partial h_{B_0B1}}{\partial yB} \\ 0 & 0 & \frac{\partial h_{B_0B2}}{\partial x_B} & \frac{\partial h_{B_0B2}}{\partial yB} \\ 0 & 0 & \frac{\partial h_{B_0B3}}{\partial x_B} & \frac{\partial h_{B_0B3}}{\partial yB} \end{bmatrix} \tag{2.55}$$

where indices A_0 and B_0 represent the initial guesses at reference vector $\mathbf{x}_0$. h_{A_0A1} is the Euclidean distance between points $\mathbf{x}_{A_0}$ and anchor $\mathbf{x}_{A_1}$. Given that:

$$\frac{\partial h_{A_0A1}}{\partial x_A} = \frac{x_{A0} - x_{A1}}{\sqrt{(x_{A0} - x_{A1})^2 + (y_{A0} - y_{A1})^2}} = \cos(\theta A_0 A_1) \tag{2.56}$$

$\mathbf{G}_{STD}$ yields:

$$\mathbf{G}_{STD} = \begin{bmatrix} \cos(\theta_{A_0A_1}) & \sin(\theta_{A_0A_1}) & 0 & 0 \\ \cos(\theta_{A_0A_2}) & \sin(\theta_{A_0A_2}) & 0 & 0 \\ \cos(\theta_{A_0A_3}) & \sin(\theta_{A_0A_3}) & 0 & 0 \\ 0 & 0 & \cos(\theta_{B_0B_1}) & \sin(\theta_{B_0B_1}) \\ 0 & 0 & \cos(\theta_{B_0B_2}) & \sin(\theta_{B_0B_2}) \\ 0 & 0 & \cos(\theta_{B_0B_3}) & \sin(\theta_{B_0B_3}) \end{bmatrix} \tag{2.57}$$

Similarly, in the collaborative framework, where there is a collaborative link, we get:

$$\mathbf{G}_{COL} = \begin{bmatrix} \cos(\theta_{A_0A_1}) & \sin(\theta_{A_0A_1}) & 0 & 0 \\ \cos(\theta_{A_0A_2}) & \sin(\theta_{A_0A_2}) & 0 & 0 \\ \cos(\theta_{A_0A_3}) & \sin(\theta_{A_0A_3}) & 0 & 0 \\ 0 & 0 & \cos(\theta_{B_0B_1}) & \sin(\theta_{B_0B_1}) \\ 0 & 0 & \cos(\theta_{B_0B_2}) & \sin(\theta_{B_0B_2}) \\ 0 & 0 & \cos(\theta_{B_0B_3}) & \sin(\theta_{B_0B_3}) \\ \cos(\theta_{A_0B_0}) & \sin(\theta_{A_0B_0}) & -\cos(\theta_{A_0B_0}) & -\sin(\theta_{A_0B_0}) \end{bmatrix} \tag{2.58}$$

Hence, the multivariate distributions for the collaborative and standalone estimators are:

$$f_{XCOL}(\mathbf{x}) = \left([2\pi]^4 \det(\boldsymbol{\Sigma}_{\boldsymbol{COL}})\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\boldsymbol{COL}})^T \boldsymbol{\Sigma}_{\boldsymbol{COL}}{}^{\mathbf{-1}}(\mathbf{x} - \boldsymbol{\mu}_{\boldsymbol{COL}})\right] \tag{2.59}$$

$$f_{XSTD}(\mathbf{x}) = \left([2\pi]^4 \det(\boldsymbol{\Sigma}_{\boldsymbol{STD}})\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\boldsymbol{STD}})^T \boldsymbol{\Sigma}_{\boldsymbol{STD}}{}^{\mathbf{-1}}(\mathbf{x} - \boldsymbol{\mu}_{\boldsymbol{STD}})\right] \tag{2.60}$$

In our theoretical framework, $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are Gaussian-distributed pdfs of position estimates with the same bias: the true position vector $\mathbf{x} = [x_A, y_A, x_B, y_B]^T$. Covariance matrices are $\mathbf{\Sigma}_{STD} = (\mathbf{G}_{STD}^T \mathbf{N}_{STD} \mathbf{G}_{STD})^{-1}$ and $\mathbf{\Sigma}_{COL} = (\mathbf{G}_{COL}^T \mathbf{N}_{COL} \mathbf{G}_{COL})^{-1}$.

$$\mathbf{\Sigma}_{STD}^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{FASTD} & 0 \\ 0 & \mathbf{\Sigma}_{FBSTD} \end{bmatrix} \tag{2.61}$$

where:

$$\mathbf{\Sigma}_{FASTD} = \begin{bmatrix} \sum_{i=1}^{3} \frac{\cos^2(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} & \sum_{i=1}^{3} \frac{\cos(\theta_{A_0A_i})\sin(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} \\ \sum_{i=1}^{3} \frac{\cos(\theta_{A_0A_i})\sin(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} & \sum_{i=1}^{3} \frac{\sin^2(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} \end{bmatrix} \tag{2.62}$$

$$\mathbf{\Sigma}_{FBSTD} = \begin{bmatrix} \sum_{i=1}^{3} \frac{\cos^2(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} & \sum_{i=1}^{3} \frac{\cos(\theta_{B_0B_i})\sin(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} \\ \sum_{i=1}^{3} \frac{\cos(\theta_{B_0B_i})\sin(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} & \sum_{i=1}^{3} \frac{\sin^2(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} \end{bmatrix} \tag{2.63}$$

$$\mathbf{\Sigma}_{COL}^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{FACOL} & \mathbf{\Sigma}_{FAB}^T \\ \mathbf{\Sigma}_{FAB} & \mathbf{\Sigma}_{FBCOL} \end{bmatrix} \tag{2.64}$$

$$\mathbf{\Sigma}_{FACOL} = \mathbf{\Sigma}_{FASTD} + \begin{bmatrix} \frac{\cos^2(\theta_{A_0B_0})}{\sigma^2_{AB}} & \frac{\cos(\theta_{A_0B_0})\sin(\theta_{A_0B_0})}{\sigma^2_{AB}} \\ \frac{\cos(\theta_{A_0B_0})\sin(\theta_{A_0B_0})}{\sigma^2_{AB}} & \frac{\sin^2\theta_{A_0B_0}}{\sigma^2_{AB}} \end{bmatrix} \tag{2.65}$$

$$\mathbf{\Sigma}_{FBCOL} = \mathbf{\Sigma}_{FBSTD} + \begin{bmatrix} \frac{\cos^2(\theta_{AB})}{\sigma^2_{AB}} & \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} \\ \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} & \frac{\sin^2\theta_{AB}}{\sigma^2_{AB}} \end{bmatrix} \tag{2.66}$$

$$\mathbf{\Sigma}_{FAB} = \mathbf{\Sigma}_{AB}^T = \begin{bmatrix} -\frac{\cos^2(\theta_{A_0B_0})}{\sigma^2_{AB}} & -\frac{\cos(\theta_{A_0B_0})\sin(\theta_{A_0B_0})}{\sigma^2_{AB}} \\ -\frac{\cos(\theta_{A_0B_0})\sin(\theta_{A_0B_0})}{\sigma^2_{AB}} & -\frac{\sin^2\theta_{A_0B_0}}{\sigma^2_{AB}} \end{bmatrix} \tag{2.67}$$

From these equations, we can establish that the equation below. For simplicity, we refer to indices A_0 and B_0 as A and B hereafter.

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \mathbf{R} \tag{2.68}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{P} & -\mathbf{P} \\ -\mathbf{P} & \mathbf{P} \end{bmatrix} \tag{2.69}$$

where $\mathbf{P}$ is:

$$\mathbf{P} = \begin{bmatrix} \frac{\cos^2(\theta_{AB})}{\sigma^2_{AB}} & \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} \\ \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} & \frac{\sin^2\theta_{AB}}{\sigma^2_{AB}} \end{bmatrix} \tag{2.70}$$

These pdfs define the distributions of position estimates for the collaborative and standalone etimators using Torrieri's estimator. It is worth noting that in our scalar analysis and uncertainty analysis (Chapters 4 and 5), we consider a special case for our model: one where the difference between the diagonal of the covariance matrices for both standalone and collaborative positioning is infinitesimally small. This was done to 1) simplify the scalar analysis and 2) allow us to see how mutual information between the distributions of collaboration and standalone estimates affects the performance of collaborative position estimation.

The best estimator achievable, in the CRB sense, for $\mathbf{x}_A$ and $\mathbf{x}_B$ has a Gaussian distribution $f_{Xref}(\mathbf{x}) = f_{CRB}(\mathbf{x}) \sim N(\boldsymbol{\mu_{CRB}}, \boldsymbol{\Sigma_{CRB}})$ where the mean $\boldsymbol{\mu_{CRB}}$ is a 4x1 vector, and $\boldsymbol{\Sigma}_{CRB}$ is a 4-x-4 matrix equal to the inverse of the Fisher Information Matrix.

$$f_{XCRB}(\mathbf{x}) = \Big([2\pi]^4 \det(\boldsymbol{\Sigma_{CRB}})\Big)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu_{CRB}})^T \boldsymbol{\Sigma_{CRB}}^{\mathbf{-1}}(\mathbf{x} - \boldsymbol{\mu_{CRB}})\right] \tag{2.71}$$

$$\boldsymbol{\mu_{CRB}} = [x_A, y_A, x_B, y_B]^T \tag{2.72}$$

The covariance matrix of the best unbiased estimator has the smallest CRB between the collaborative or standalone estimators. From [17] the covariance matrix of this best estimator is a matrix that yields the collaborative Cramer-Rao Bound (CRB_{COL}). This best covariance matrix is therefore the inverse of the Fisher Information Matrix using the collaborative set of range measurements.

$$\boldsymbol{\Sigma_{CRB}} = \mathbf{FIM}_{COL}^{-1} \tag{2.73}$$

2.B Relating Fisher Information Matrix to Kullback-Leibler Divergence

If we consider a distribution for a parameter θ, $f(x;\theta)$, and its neighboring distribution $f(x;\theta+\epsilon)$, whose parameter differs from $f(x;\theta)$ by a small amount ϵ, the KLD between these two nearly similar distributions is proportional to the FIM as shown below.

$$\begin{aligned} KLD(f(x;\theta)||f(x;\theta+\epsilon)) &= \int_x f(x;\theta)\log f(x;\theta) - f(x;\theta)\log f(x;\theta+\epsilon)dx \\ &= E_{f(x;\theta)}[\log f(x;\theta)] - E_{f(x;\theta)}[\log f(x;\theta+\epsilon)] \end{aligned} \tag{2.74}$$

Knowing that by definition, the Taylor series for any given function at a value x_0 is:

$$f(x_0) + (x-x_0)f'(x_0) + \frac{1}{2}(x-x_0)^2 f''(x_0) + ... = \sum_{n=0}^{\infty}(x-x_0)^n\frac{f^{(n)}(x_0)}{n!} \tag{2.75}$$

Taking the Taylor series of $\log f(x;\theta+\epsilon)$ up to the second order, we get:

$$\log f(x;\theta+\epsilon) \cong \log f(x;\theta) + \epsilon\frac{\partial}{\partial\theta}\log f(x;\theta) + \epsilon^2\frac{\partial^2}{2\partial\theta^2}\log f(x;\theta) \tag{2.76}$$

Since $f(x;\theta)$ is a pdf, expectation and differentiation can be exchanged. Thus, taking the expected value of the second term in the equation above yields:

$$E_{f(x;\theta)}[\epsilon\frac{\partial}{\partial\theta}\log f(x;\theta)] = \epsilon\frac{\partial}{\partial\theta}\int f_\theta(x,\theta)\log f(x;\theta)dx = 0 \tag{2.77}$$

In the expression above, $\int f_\theta(x,\theta)\log f(x;\theta)dx$ is equal to unity, given that the log function is the natural log. Thus, its derivative is zero. Hence,

$$E_{f(x;\theta)}[\log f(x;\theta+\epsilon)] \cong E_{f(x;\theta)}[\log f(x;\theta) + \epsilon^2\frac{\partial^2}{2\partial\theta^2}\log f(x;\theta)] \tag{2.78}$$

The KLD yields:

$$\begin{aligned} KLD(f(x;\theta)||f(x;\theta+\epsilon)) &= E_{f(x;\theta)}[\log f(x;\theta)] - E_{f_\theta}[\log f(x;\theta+\epsilon)] \\ &= -\frac{\epsilon^2}{2}E_{f(x;\theta)}[\frac{\partial^2}{\partial\theta^2}\log f(x;\theta)] \\ &= \frac{\epsilon^2}{2}I_{Fisher} \end{aligned} \tag{2.79}$$

where matrix I_{Fisher} is referred in the literature [27] as expected Fisher information, and identified as 'providing a local metric in the space of distributions indexed by the parameter vector θ...it is formally a *metric tensor* describing the manifold of models.' [27]. In fact, this matrix represents a metric in an n-dimensional Riemannian manifold.Reference [96] explains it a bit more clearly: 'in the limit of infinitesimal deformations, the Kullback-Leibler distance between the two densities is a quadratic form with the Fisher information matrix playing the role of the metric tensor.'

2.C Mutual Information Symmetry

In this appendix, we prove that mutual information is symmetric.
Let P and Q be two random variables with joint distribution function. $f_{P,Q}(p,q)$ and standalone pdfs $f_P(p)$ and $f_Q(q)$. By definition, the mutual information between P and Q is $MI(P,Q)$ such that:

$$\begin{aligned}
MI(P,Q) &= \int_p \int_q f_{PQ}(p,q) \log \frac{f_{PQ}(p,q)}{f_P(p) f_Q(q)} dqdp \\
&= \int_p \int_q f_{PQ}(p,q) \log \frac{f_{P|Q}(p|Q=q) f_Q(q)}{f_P(p) f_Q(q)} dqdp = \int_p \int_q f_{PQ}(p,q) \log \frac{f_{Q|P}(q|P=p) f_P(p)}{f_P(p) f_Q(q)} dqdp \\
&= \int_p \int_q f_{PQ}(p,q) \log \frac{f_{P|Q}(p|Q=q)}{f_P(p)} dqdp = \int_p \int_q f_{PQ}(p,q) \log \frac{f_{Q|P}(q|P=p)}{f_Q(q)} dqdp \\
&= \int_p \int_q f_{PQ}(p,q) \log f_{P|Q}(p|Q=q) dqdp - \int_p f_{PQ}(p,q) \log f_P(p) dp \\
&= \int_p \int_q f_{PQ}(p,q) \log f_{Q|P}(q|P=p) dqdp - \int_q f_{PQ}(p,q) \log f_Q(q) dq \\
&= -H(f_{P|Q}) + H(f_P) = -H(f_{Q|P}) + H(f_Q)
\end{aligned} \tag{2.80}$$

Hence,

$$MI(P,Q) = H(f_P) - H(f_{P|Q}) = H(f_Q) - H(f_{Q|P}) = MI(Q,P) \tag{2.81}$$

Consequently, mutual information is symmetric.

Chapter 3

Kullback-Leibler Divergence as a Performance Metric

When it comes to assessing the performance of position estimators (*e.g:* during field testing) in terms of accuracy, the correct performance metric to choose is the one that meets application requirements. Hence performance metrics are application-specific. A common technique to assess the performance of position estimators compares their position mean squared errors (PMSE) or its square root: Root Mean Squared Error (RMSE). However, this performance metric is limited because it focuses on the central information (mean and variance) of the distribution of position error. In this chapter we show that the Kullback-Leibler Divergence can be used to address this limitation for applications where end-tail behaviors in the error distributions matter.

Kullback-Leibler Divergence (KLD) measures the distance between two distributions. It achieves this by comparing the shapes of two pdfs, one of which being the reference for accuracy based on the data at hand. There are two important features about KLD that we utilize in this chapter. Firstly, KLD encompasses all the statistical information that can be known about each distribution in its comparison. This means that its comparison is not restricted to the average behavior of the distributions (the first two moments). Secondly, it is a relative measure of accuracy. It achieves this by comparing distributions to a reference for accuracy.

In this chapter, we use KLD to introduce a performance metric which outperforms a comparison of RMSE, especially when distributions of position error have a heavy-tail. A heavy-tail skews the first two moments, making them poor descriptors of the overall distribution.

To summarize the results discussed in this chapter, using the Kullback-Leibler Divergence, we introduce a novel performance metric that accounts for all statistical information of the error distributions being compared. This novel performance metric is a robust option for applications like E911, where a comparison of the central information (mean and variance) of error distributions is not sufficient.

3.1 Motivation and Literature Review

Selecting a performance metric to assess the accuracy of a position estimator is mainly application-specific. Li, et al. in [79]- [83] discuss different performance metrics that can be selected based on the driving requirement for measuring performance. In some applications, one may need to pay more attention to estimates that have a good behavior. In such cases, metrics like Harmonic Average Error are proposed by the authors. In other applications, it is more important to focus on estimates with a bad behavior (outliers) [80]; root mean squared of position error (RMSE) is then recommended. Other metrics like Average Euclidean Error (AEE) and Geometric Average Error (GAE) are also presented for a more intuitive look at the errors (AEE) and to provide a more balanced measure of the performance (GAE). Of all the metrics indicated above, RMSE is the one that is theoretically tractable. This feature has warranted its very common adoption as a performance benchmark. Given that our goal is to theoretically analyze the performance of collaborative positioning, it was the only metric worth considering for further study.

Given a set of M position estimate vectors $\mathbf{x'}_k$ with values $\{\mathbf{x'}_1, \mathbf{x'}_2, ...\mathbf{x'}_M\}$, and a position vector $\mathbf{x}$ of a true position, RMSE is defined as:

$$RMSE = \sqrt{E[\delta_k^2]} = \sqrt{E[||\mathbf{x'_k} - \mathbf{x})||^2]} \tag{3.1}$$

where δ_k represents the k^{th} position error for position estimate $\mathbf{x'_k}$. In practice, the expected value term in the equation above is estimated using the sample mean.

RMSE is an absolute error. This means that it measures error in comparison to the truth (the true position). Hence, it is only useful when compared against another absolute error, or against a reference for accuracy (like the CRB). The latter comparison is called relative error assessment. In reference [81], Li et al address the advantages of relative measures of performance. The authors indicate that 'relative error reveals better the inherent error characteristics of an estimator rather than the absolute error'. It is common practice to either compare RMSE to another RMSE or its squared to the Cramer-Rao Bound. In an RMSE-to-CRB comparison, the efficiency of the estimator is evaluated; that is, its performance with respect to (w.r.t) the best performance achievable if the estimator is unbiased ([1], [21], [84], [85], [31], [86], [87], [12]). In an RMSE-to-RMSE comparison, the estimator with the smallest RMSE is deemed most accurate.

RMSE is the square root of the mean of all squared position errors. Thus, a comparison of RMSEs only assesses performance based on the central information (mean and variance) of the error distributions. As a result, a comparison of RMSEs does not account well for end-tail and heavy-tailed distributions [67]. In fact, the presence of a single outlier can offset an RMSE estimate and lead to an erroneous assessment of performance.

In our research context, the inputs to the standalone and collaborative estimators only differ by the addition of the collaborative link (set of range measurements between the collaborating nodes). It is possible that differences between the pdfs of their position errors occur beyond the first two moments (mean and variance). In such cases, a comparison of RMSEs would not be adequate because it would only compare the first two moments of the error distributions. For this reason, we need to identify a metric that will account for all moments of the pdf of position error.

In this chapter, we introduce a novel performance metric that accounts for all statistical information in the error distributions. We achieve this using KLD. The usefulness of this approach is that it not only compares the entire statistical content of the error distributions, but it also compares them to the best performance achievable based on the position estimates at hand.

3.2 Novel Performance Metric: ΔKLD

$KLD(f_X||g_X)$ can be used to compare two pdfs, with $f_X(x)$ being the reference for estimation accuracy based on the information at hand. In this context, it can be used to select the most accurate pdf by accounting for all statistical information (all moments) in the pdfs compared. Let δ represent the set of position errors from all position estimators. Let δ_{ref} represent the smallest position error at any point in time based on the performance of the standalone and collaborative estimators. Given $\mathbf{x}$ the true position vector of a node, $\hat{\mathbf{x}}_{kSTD}$ the k^{th} standalone position estimate and $\hat{\mathbf{x}}_{kCOL}$ the k^{th} collaborative position estimate, the k^{th} smallest position error between standalone and collaborative estimators is:

$$\delta_{kref} = \min\{||\hat{\mathbf{x}}_{kSTD} - \mathbf{x}||, ||\hat{\mathbf{x}}_{kCOL} - \mathbf{x}||\} = \min\{\delta_{kSTD}, \delta_{kCOL}\} \tag{3.2}$$

Let COL represent the collaborative estimator and STD the standalone estimator. If $f_{\Delta ref}(\delta)$ is the pdf that describes the set of smallest position errors between the standalone and collaborative position estimators, and if $f_{\Delta STD}(\delta)$ and $f_{\Delta COL}(\delta)$ represent the pdfs of position error from the standalone and collaborative estimators, respectively, then, ΔKLD identifies the most accurate estimator as follows:

$$\begin{array}{ll} KLD(f_{\Delta ref}||f_{\Delta STD}) > KLD(f_{\Delta ref}||f_{\Delta COL}) & \text{, COL is more accurate than STD} \\ KLD(f_{\Delta ref}||f_{\Delta STD}) < KLD(f_{\Delta ref}||f_{\Delta COL}) & \text{, STD is more accurate than COL} \end{array} \tag{3.3}$$

Therefore, a valid candidate for a performance metric is ΔKLD:

$$\begin{aligned}\Delta KLD &= KLD(f_{\Delta ref}||f_{\Delta STD}) - KLD(f_{\Delta ref}||f_{\Delta COL}) \\ &= H(f_{\Delta ref}||f_{\Delta STD}) - H(f_{\Delta ref}||f_{\Delta COL})\end{aligned} \tag{3.4}$$

$$\begin{aligned}\Delta KLD > 0 &\quad \text{, COL is more accurate than STD} \\ \Delta KLD < 0 &\quad \text{, STD is more accurate than COL}\end{aligned} \tag{3.5}$$

Since cross-entropies $H(f_{\Delta ref}||f_{\Delta STD})$ and $H(f_{\Delta ref}||f_{\Delta COL})$ are measures of inaccuracy, equation (3.4) indicates that ΔKLD measures the reduction in inaccuracy due to collaboration; this is also the improvement in accuracy due to collaboration. Thus, ΔKLD identifies the most accurate estimator based on a comparison of all statistical information from the pdfs of position error for both standalone and collaborative estimators to a reference distribution. We show in the simulation section that this approach is more robust than a comparison of position RMSEs.

It is important to note that a definition for $f_{\Delta ref}(\delta)$ must exist when using ΔKLD. $f_{\Delta ref}(\delta)$ is a reference for accuracy based on the data available. Thus omitting $f_{\Delta ref}(\delta)$ will be equivalent to computing either $KLD(f_{\Delta COL}||f_{\Delta STD})$ or $KLD(f_{\Delta STD}||f_{\Delta COL})$ which only provides discriminating information regarding how different the distributions $f_{\Delta COL}(\delta)$ and $f_{\Delta STD}(\delta)$ are from each other, with no additional information on *which* distribution is most accurate. Therefore, ΔKLD requires a definition for $f_{\Delta ref}(\delta)$, the reference pdf that is used to select which pdfs of position error (between $f_{\Delta STD}(\delta)$ and $f_{\Delta COL}(\delta)$) is most accurate.

$f_{\Delta ref}(\delta)$ can be defined either theoretically or artificially. Deriving $f_{\Delta ref}(\delta)$ theoretically is equivalent to identifying the distribution that represents the most accurate distribution for position error given an estimator, and a set of range measurements. For unbiased linear estimators, we know that the variance of the most accurate estimator will not be smaller than the CRB. Thus, if the range measurements are for instance Gaussian-distributed, the pdf of position estimates will have a known mean (the true position), and a variance equal to the CRB. Consequently, $f_{\Delta ref}(\delta)$ will have a Rayleigh-like distribution whose parameters can be defined. However, for non-linear estimators, the theoretical distribution of position error cannot be determined because the distribution of position estimates is unknown. Since the MLE is the recommended position estimator for collaborative positioning [78], a theoretical derivation for $f_{\Delta ref}(\delta)$ is not possible.

Alternatively, deriving $f_{\Delta ref}(\delta)$ in practical cases, for empirical data, has been accomplished in the literature by making an artificial definition for $f_{\Delta ref}(\delta)$ based on the data observed in reality. Burnham and Anderson indicated that 'there are no true models' and establish that 'in the analysis of empirical data, one must face the question: 'What model should be used to best approximate reality given the data at hand?' (the best model depends on sample size).' [42]. We consider this recommendation to circumvent the limitation on the knowledge of the 'true' $f_{\Delta ref}(\delta)$, and refer to $f_{\Delta ref}(\delta)$ as a 'reference pdf for the most accurate position estimator'. We consider a novel approach to define the reference distribution $f_{\Delta ref}(\delta)$ that is based on the requirement that $f_{\Delta ref}(\delta)$ is the distribution of the smallest position error between the standalone and collaborative estimators for any estimate. In this regard, it is an artificially derived distribution of the best definition of accuracy based on the data at hand at any point in time. The need to artificially derive a reference distribution when comparing two models has also been discussed by Degroot and Feinberg in the context of comparing two model predictions of rain to identify the most reliable one [92]. The implications of our novel approach for defining $f_{\Delta ref}(\delta)$ are that:

- $f_{\Delta ref}(\delta)$ needs not be equal to the theoretical reference pdf achievable (in the CRB-sense). It only needs to be the pdf that will identify which pdf ($f_{\Delta STD}(\delta)$ or $f_{\Delta COL}(\delta)$) is most accurate, at the occurrence of every statistical event, based on the data at hand.
- a definition for $f_{\Delta ref}(\delta)$ requires prior knowledge of the true position to assess which estimator yields the smallest position error: this requirement is acceptable because we intend through this metric to provide a performance metric more reliable than RMSE which also requires prior knowledge of the true position.
- when one estimator yields the smallest position error all the time, its pdf will be equal to $f_{\Delta ref}(\delta)$.
- the smallest position error from both collaborative and standalone estimators is the observable for $f_{\Delta ref}(\delta)$.

As an example, $f_{\Delta refA}(\delta)$ is the pdf describing the smallest position error for node A obtained from both standalone and collaborative estimators. The k^{th} observable of $f_{\Delta refA}(\delta)$ is δ_{refAk}:

$$\delta_{refAk} = \min\{||\hat{\mathbf{x}}_{kSTD} - \mathbf{x}_A||, ||\hat{\mathbf{x}}_{kCOL} - \mathbf{x}_A||\} \tag{3.6}$$

$f_{\Delta ref}(\delta)$ was selected such that, if one estimator, say the collaborative estimator, is more accurate all the time, then $f_{\Delta ref}(\delta) = f_{\Delta COL}(\delta)$ and ΔKLD will be positive. If ΔKLD is positive, the collaborative estimator will be deemed more accurate than the standalone estimator. To the best of our knowledge, there is no precedent for this approach in the literature. This proposed metric defines the most accurate position estimator as the one whose pdf of position error is the closest, in the Kullback-Leibler sense, to $f_{\Delta ref}(\delta)$. The performance of this metric is further addressed in the simulation section where we confirm that our proposed metric is able to identify the most accurate estimator better than a comparison of position RMSEs, in the presence of strong heavy tails in the error distribution.

3.3 Simulation Results

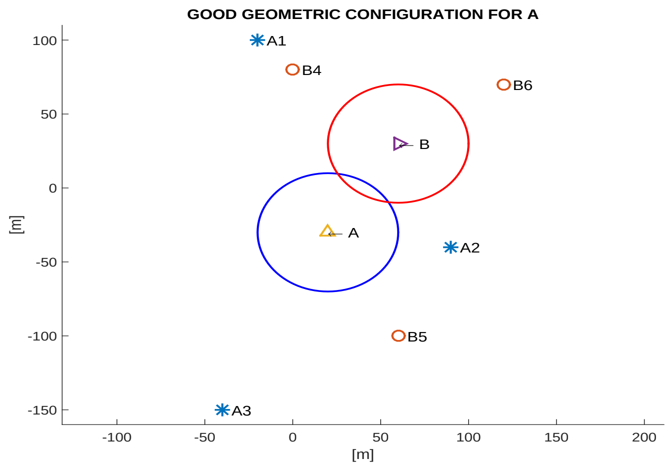
We consider a 2-D synchronized Time of Arrival (TOA) system in a 300mx300m Matlab simulation environment. We test the theoretical results presented in the previous sections in both LOS and NLOS cases on all range measurements. The SNR on all non-collaborative links is fixed to 20 dB, unless otherwise specified. The SNR on the collaborative link varies from 0 to 40 dB, unless otherwise specified. We use the Newton-Raphson Maximum Likelihood estimator (MLE) [67] in the simulation cases presented. We also consider both good and poor geometric configurations for node A as shown on Fig. 3.1(a) and Fig. 3.1(b), respectively. In the good geometric configuration, node A is within the convex hull of its standalone anchors. It is outside the convex hull in the poor geometric configuration. Node B is always within the convex hull of its anchors. To test the proposed performance metric in very good and poor conditions, we consider two cases: 1) we assume LOS on all range measurements in a good geometric configuration as a best case scenario and 2) we assume non-line of sight (NLOS) errors on all range measurements in a poor geometric configuration as a worst case scenario. In NLOS range conditions, NLOS errors are added to all range measurements (both collaborative and standalone range links) for both nodes A and B. All NLOS errors are exponentially-distributed with a mean of 25 meters. The exponential distribution was selected because it is the most common distribution used to describe NLOS errors in the literature. The poor geometric configuration was selected to analyze a case where one node (A) can gain more from collaboration than node (B) due to its standalone geometric configuration. Fig. 3.1 shows the geometric configurations described above. The circles around nodes A and B delimit the area within which 400 different simulation cases are executed. For each simulation case considered, 1000 position estimates are generated for each of nodes A and B. For each of the 400 simulation cases, we vary the SNR relative to the range measurements on the collaborative link from 0 to 40 dB while keeping the SNR on all standalone links to 20 dB. The SNR and the variance σ^2_{AB} of observed range measurements on the collaborative link are related as follows:

$$SNR_{dB} = 10log_{10}\left(\frac{d^2}{\sigma^2_{AB}}\right) \Longrightarrow \sigma^2_{AB} = \frac{d^2}{10^{\left(\frac{SNR_{dB}}{10}\right)}} \tag{3.7}$$

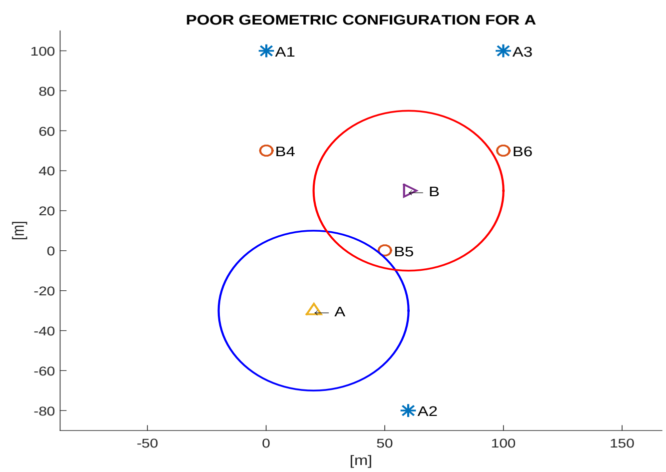
In the next sub-sections, we present simulation results from individual simulation cases and later present simulation results based on the average of the 400 simulation cases.

3.3.1 Simulation Results for Proposed Performance Metric

The proposed performance metric ΔKLD is examined in this section. We compare the performance of ΔKLD and Root Mean Squared Error (RMSE) in four specific examples. For reference, the plots for the pdfs of position error are provided alongside the 67% and 90% values for the CDF of position error. As discussed in the theoretical section, ΔKLD as a



(a) Good geometric condition for nodes A and B



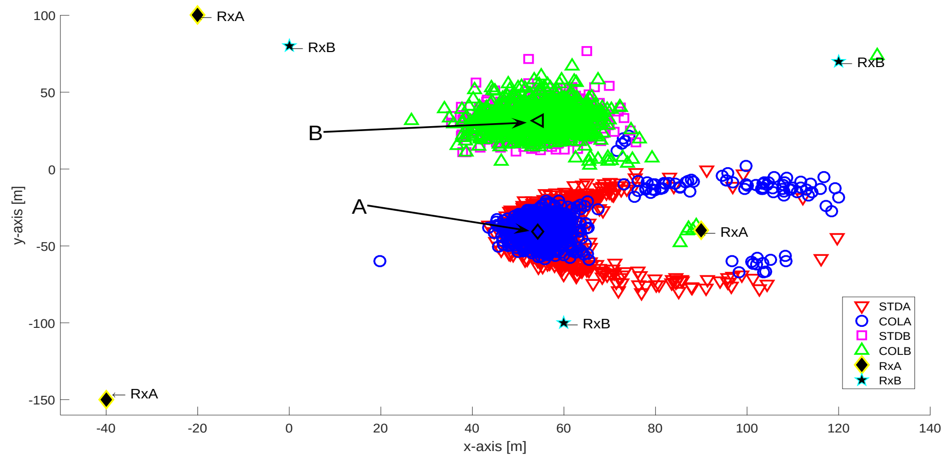
(b) Poor geometric configuration for node A alone

Figure 3.1: Simulation geometric configurations: Overview of test configuration. The circles centered at A and B are delimiting regions containing 400 different positions of nodes A and B, respectively, for 400 different simulation cases. A1, A2, A3 are anchors for A. B4, B5, B6 are anchors for B. (a) Good geometric configuration for A and B with respect to their anchors. LOS with 20 dB-SNR is assumed on all range measurements in this configuration. (b) Poor geometric configuration for node A with respect to its anchors. Exponentially-distributed NLOS errors of mean 25 m are added to all range measurements for both nodes A and B.

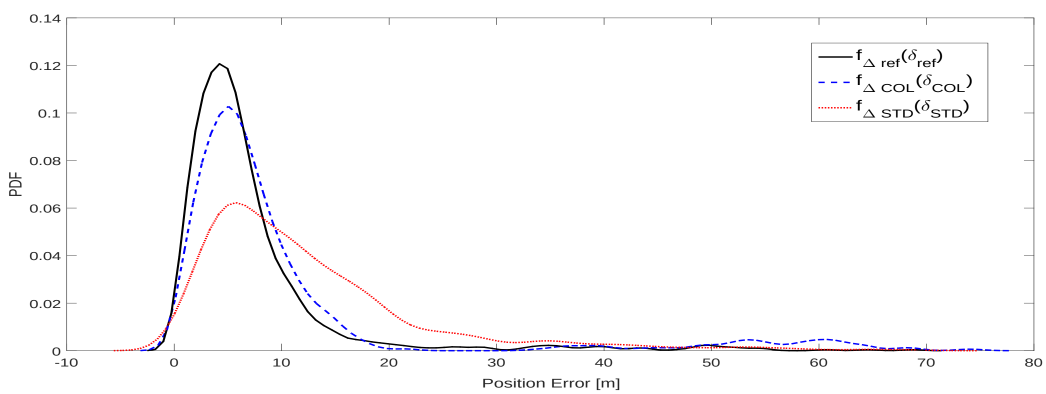
performance metric identifies an accurate estimator, as the one whose pdf of position error is closest to the reference pdf of position error. A negative ΔKLD means that the standalone estimator is more accurate, while a positive ΔKLD indicates that the collaborative estimator is more accurate than the standalone estimator.

Test Case #1: ΔKLD and RMSE disagree

The first example on Fig. 3.2 shows one simulation case where the MLE estimator is used in a good geometric configuration (Fig. 3.1(a)) with LOS on all range measurements for both nodes A and B, 20 dB-SNR on all standalone range measurements and 30 dB-SNR on the collaborative link. From visual observation of the pdfs of position errors on Fig.3.2(b), we note that the pdf of position error for the collaborative estimator ($f_{\Delta COL}(\delta)$) is closer to the reference pdf ($f_{\Delta ref}(\delta)$) than the pdf of position error for the standalone estimator ($f_{\Delta STD}(\delta)$). This indicates that if we look at all statistical information provided by the pdfs of position error, the collaborative estimator is more accurate than the standalone estimator. This explains why ΔKLD identifies collaborative estimation as most accurate, while RMSE disagrees as shown on Fig.3.2(c). The presence of a heavy tail on the error distribution for estimates of point $\mathbf{x}_A$ makes RMSE's assessment incorrect. Per Fig.3.2(c), the CDF 67% and 90% values agree with ΔKLD that collaboration is more accurate.



(a) Good geometric configuration (Fig. 3.1(a)) in one simulation case: R_{xA}, R_{xB} are anchors used to locate nodes A and B, respectively.



(b) Pdf of position error for node A. $f_{\Delta ref}(\delta)$, $f_{\Delta COL}(\delta)$ and $f_{\Delta STD}(\delta)$ represent pdfs of position error for the reference pdf, the collaborative and standalone estimators, respectively.

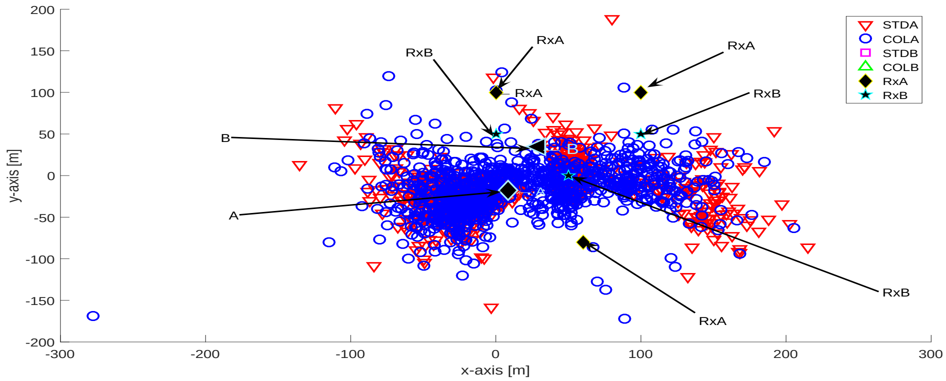
Metrics	**STD**	**COL**	**More Accurate**
RMSE	16.4108	17.5802	STD
CDF67	13.44	8.2154	COL
CDF90	25.1023	16.2203	COL
DKLD		0.30892	COL

(c) Performance metrics summary for node A

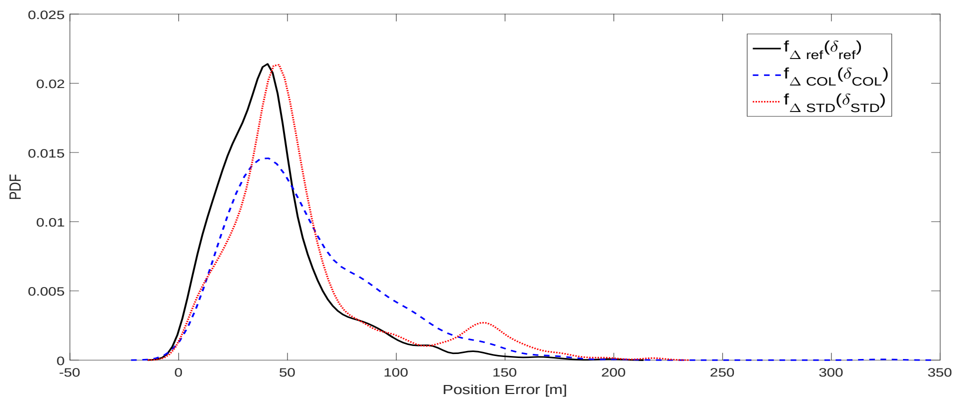
Figure 3.2: Specific Example #1: (a) Good geometric configuration (Fig. 3.1(a)) in one test case with 1000 position estimates for each node, (b) pdf of position error for node A, (c) performance metrics summary for node A. Using MLE in LOS range conditions with good geometric conditions (Fig. 3.1(a)), 20 dB on all standalone links and 30 dB on the collaborative link. ΔKLD and RMSE disagree on identifying accurate estimator. The presence of a heavy tail on the pdf of position error makes RMSE's assessment incorrect.

Test Case #2: ΔKLD and RMSE disagree

The second example in Fig.3.3 shows one of 400 simulation cases where the MLE estimator is used in a poor geometric configuration (Fig. 3.1(b)) with an exponentially-distributed NLOS error of mean 25 meters on all range measurements for both nodes A and B, 20 dB-SNR on all standalone range measurements and 20 dB-SNR on the collaborative link. Based on a comparison of pdfs of position errors to the reference pdf (Fig.3.3(b)), the standalone estimator is most accurate. RMSE and ΔKLD disagree on which estimator is most accurate. Per Fig.3.3(c), RMSE and ΔKLD disagree in identifying the most accurate estimator because near the 100m mark on the pdf plot, the collaborative estimator has a set of outliers. These outliers also offset the 90% CDF value. ΔKLD accounts for these outliers differently than RMSE does. By definition, since KLD accounts for all statistical moments of the distributions being compared, we can conclude that ΔKLD makes the most accurate assessment of all the metrics provided in Fig.3.3(c). Furthermore, in this test case, the CDF67 value indicates that if we only look at the 67% error bound to make an assessment of which estimator is most accurate, we end up selecting the standalone estimator. However, when we look at the 90% error bound to make an assessment, we end up selecting the collaborative estimator. This confirms that discarding some points to make an assessment on which estimator is most accurate does not reflect the 'actual' performance of the estimator. This further confirms that ΔKLD's assessment is more robust than the other metrics proposed because it accounts well for the outliers and does not require discarding any data obtained to make an assessment.



(a) Poor geometric configuration for node A in one test case (Fig. 3.1(b)): R_{xA}, R_{xB} are anchors used to locate nodes A and B, respectively.



(b) Pdf of position error for node A. $f_{\Delta ref}(\delta)$, $f_{\Delta COL}(\delta)$ and $f_{\Delta STD}(\delta)$ represent pdfs of position error for the reference pdf, the collaborative and standalone estimators, respectively.

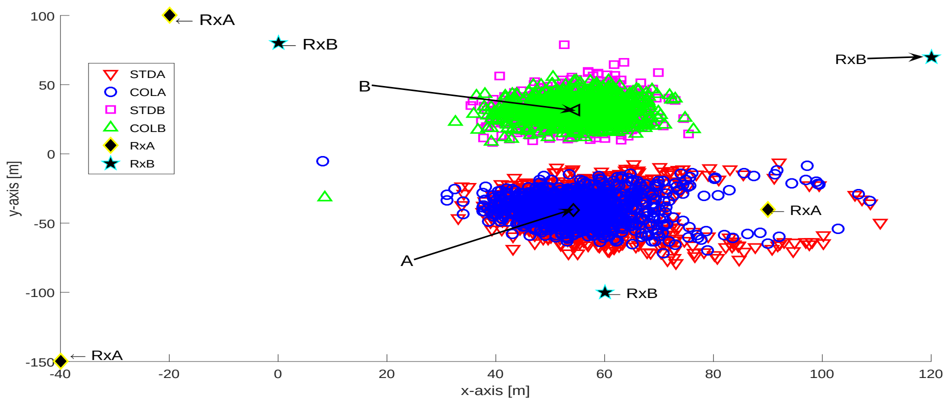
Metrics	**STD**	**COL**	**More Accurate**
RMSE	64.5782	63.5365	COL
CDF67	57.8436	62.0478	STD
CDF90	110.1587	95.7022	COL
DKLD	-0.06202		STD

(c) Performance metrics summary for node A

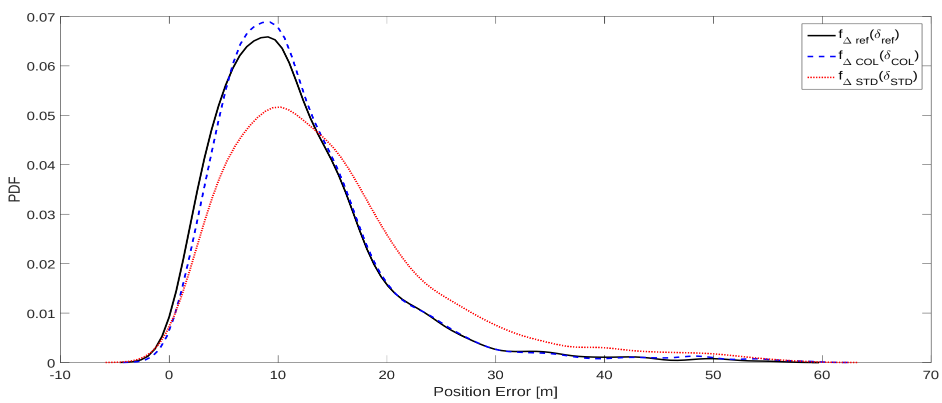
Figure 3.3: Specific Example #2: (a) Poor geometric configuration for node A in one test case with 1000 position estimates for each node, (b) pdf of position error for node A, (c) Performance metrics summary for node A. Using MLE with poor geometry (Fig. 3.1(b)) with an exponentially-distributed NLOS error of mean 25 meters on all range measurements for both nodes A and B, 20 dB-SNR on all standalone range measurements and 20 dB-SNR on the collaborative link. ΔKLD and RMSE disagree on identifying the most accurate estimator.

Test Case #3: ΔKLD and RMSE agree

The third example in Fig.3.4 shows the NLS estimator in a good geometric configuration (Fig. 3.1(a)) with LOS on all range measurements for both nodes A and B, 20 dB-SNR on all standalone range measurements and 30 dB on the collaborative link. In this case scenario, both RMSE and ΔKLD agree that collaborative positioning works as shown on Fig.3.4(c). The pdfs on Fig.3.4(b) confirm through visual inspection that the collaborative estimator is more accurate than the standalone estimator because the collaborative estimator pdf of position error is closer to the reference pdf than the standalone pdf.



(a) Good geometric configuration (Fig. 3.1(a)) for a single test case: R_{xA}, R_{xB} are anchors used to locate nodes A and B, respectively.



(b) Pdf of position error for node A. $f_{\Delta ref}(\delta)$, $f_{\Delta COL}(\delta)$ and $f_{\Delta STD}(\delta)$ represent pdfs of position error for the reference pdf, the collaborative and standalone estimators, respectively.

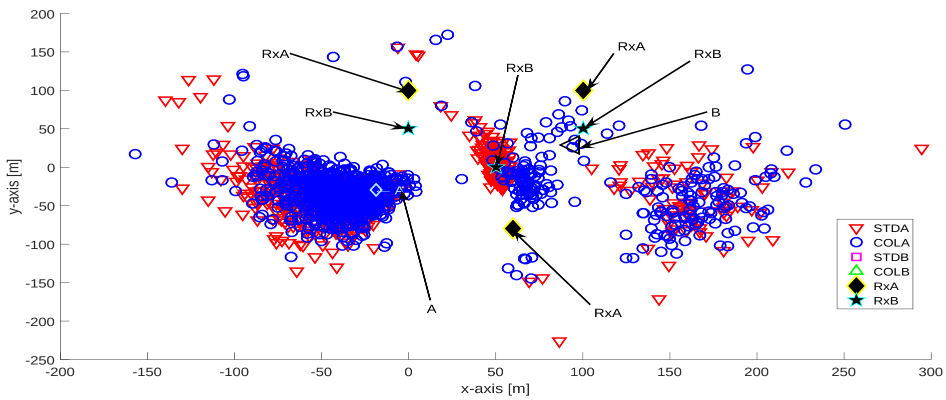
Metrics	**STD**	**COL**	**More Accurate**
RMSE	15.6492	14.2602	COL
CDF67	15.3155	14.3981	COL
CDF90	23.5202	21.9643	COL
DKLD		0.09113	COL

(c) Performance metrics summary for node A

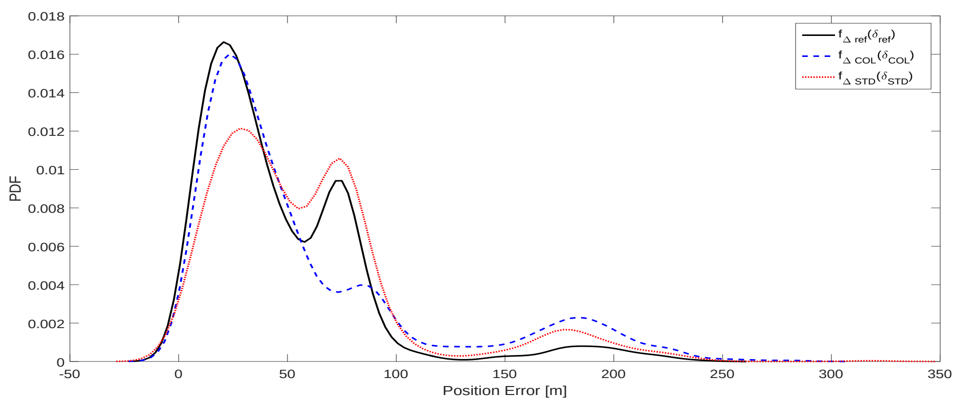
Figure 3.4: Specific Example #3: (a) Good geometric configuration (Fig. 3.1(a)) in one test case with 1000 position estimates for each node, (b) pdf of position error for node A, (c) performance metrics summary for node A. Using the NLS estimator in good geometric configuration (Fig. 3.1(a)) with LOS on all range measurements for both nodes A and B, 20 dB-SNR on all standalone range measurements and 30 dB on the collaborative link. Both ΔKLD and RMSE agree that collaborative estimation works.

Test Case #4: ΔKLD and RMSE agree

In the fourth example, we show a simulation case where it is not obvious to identify the most accurate estimator by visual inspection. Fig.3.5 shows the MLE estimator in a poor geometric configuration (Fig. 3.1(b)) with exponentially-distributed NLOS errors with a mean of 25 meters on all range measurements for both nodes A and B, 20 dB-SNR on all standalone range measurements and 20 dB-SNR on the collaborative link. In this case, both RMSE and ΔKLD agree that standalone is the most accurate estimator.



(a) Poor geometric configuration for node A (Fig. 3.1(b)) in one simulation case: R_{xA}, R_{xB} are anchors used to locate nodes A and B, respectively.



(b) Pdf of position error for node A. $f_{\Delta ref}(\delta)$, $f_{\Delta COL}(\delta)$ and $f_{\Delta STD}(\delta)$ represent pdfs of position error for the reference pdf, the collaborative and standalone estimators, respectively.

Metrics	STD	COL	More Accurate
RMSE	77.7491	85.2644	STD
CDF67	71.6957	57.8219	COL
CDF90	115.4308	177.1944	STD
DKLD	-0.010381		STD

(c) Performance metrics summary for node A

Figure 3.5: Specific Example #4: (a) Poor geometric configuration (Fig. 3.1(b)) for node A in one simulation case with 1000 position estimates for each node, (b) pdf of position error for node A, (c) performance metrics summary for node A. Using the MLE estimator in poor geometric configuration (Fig. 3.1(b)) with exponentially-distributed NLOS errors of mean 25 meters on range measurements for both nodes A and B, 20 dB-SNR on all standalone range measurements and 20 dB-SNR on the collaborative link. Both ΔKLD and RMSE agree that collaborative estimation is not beneficial.

3.3.2 Summary of individual test cases results

From the four individual simulation cases discussed above, we have shown cases where RMSE and ΔKLD both agree and disagree on their assessment of which estimator is most accurate. We confirmed through simulations that RMSE is not reliable because it accounts for only the first two moments in error distributions. Furthermore, we also provided the CDF 67% and 90% values for the distribution of position error. These two values confirmed that traditional techniques that discard outliers by pre-processing data prior to making an assessment on the most accurate estimator, yield local assessments on the performance of estimators and do not account for end-tail behaviors well. Since the definition of an outlier is very subjective and usually depends on the user or the application, such techniques are limited in their usefulness. ΔKLD on the other hand, does not require pre-processing of data and accounts well for outliers. In this regard, ΔKLD is a robust performance metric in applications where we need to account for end-tail behaviors in the error distribution.

3.3.3 Simulation Results on Accuracy of Collaborative Positioning

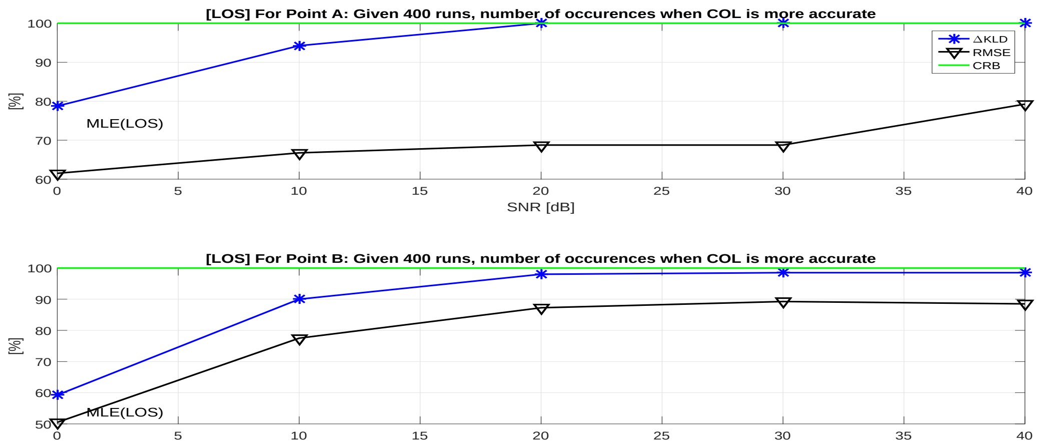
Having validated through specific simulation testcases that ΔKLD accounts for end-tail behaviors from the error distributions, we now compare the performance of standalone vs. collaborative positioning using ΔKLD, a comparison of position RMSEs and a comparison of CRBs, in both poor and good geometric configurations, under both LOS and NLOS conditions, over a set of 400 simulation cases.

We show in the following plots (Fig. (3.6)), the number of times collaborative positioning is more accurate than standalone, based on 400 different test cases where the positions of A and B are randomly selected within the circles shown on Fig. 3.1.

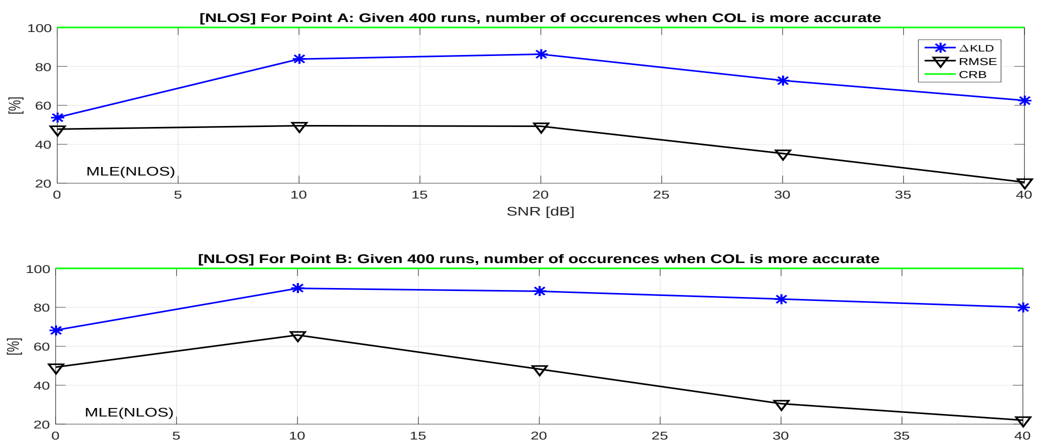
For each of the 400 test cases, we estimate 1000 points to measure RMSE of position and ΔKLD, as well as CRB. Thus we generate 400,000 points each time we vary the SNR on the collaborative link. The SNR on the collaborative link varies from 0 to 40 dB, while the standalone SNR is maintained at 20 dB. On each plot, we show how comparisons of a) RMSEs, b) KLDs and c) CRBs indicate that collaborative positioning is more accurate than standalone. The results confirm that using a comparison of CRBs to determine whether collaborative estimation is more accurate than standalone's does not always work, because a comparison of CRBs always indicates collaboration as more accurate, while accuracy metrics like ΔKLD and RMSE indicate that collaborative estimation is not always accurate.

A comparison of position RMSEs or ΔKLD, on the other hand can identify when collaborative estimation is more accurate than standalone or not. However, a comparison of position RMSEs may incorrectly identify the most accurate estimator, because RMSE is sensitive to outliers, and does not account well in general for end-tail behaviors. In this regard, RMSE turns out to be more pessimistic in believing collaborative positioning to be more accurate, than our proposed metric ΔKLD. This can be seen in the line of sight plots at SNR=40dB (Fig. 3.6(a)). In fact, although we expect, that at such a high SNR with LOS on all range measurements and good geometric configuration, collaborative positioning will be more accurate, RMSE is still more pessimistic than our proposed metric. Therefore, in our context of comparing the accuracy between standalone and collaborative estimation, ΔKLD is a more robust performance metric than a comparison of CRBs or RMSEs.

As expected in Fig. 3.6(a), in LOS range conditions, collaboration improves position accuracy as the SNR on the collaborative link increases. NLOS errors in range measurements tend to cause outliers in the error distribution. In Fig. 3.6(b), in the presence of NLOS errors, the combination of a large range variance (low SNR) and NLOS errors can actually improve accuracy because the Gaussian range variance and always positive NLOS bias compensate each other. As the SNR increases, the noise variance reduces considerably and the effect of NLOS errors can no longer be compensated by the noise variance. As a result, the presence of NLOS errors yields more outliers, which results (in general) in degraded accuracy as the SNR increases. The considerably low value for the RMSE plots at high SNRs in NLOS range conditions (in Fig. 3.6(b)) is due to the fact that RMSE does not handle well outliers.



(a) Case of LOS range conditions and good geometric configuration



(b) Case of NLOS range conditions and poor geometric configuration

Figure 3.6: Percentage of times collaboration is more accurate vs. SNR on collaborative link. 400 simulation cases are considered. This plot indicates how often different performance metrics (a comparison of CRBs, a comparison of RMSEs and the proposed performance metric using ΔKLD) indicate that collaborative estimation is more accurate than standalone estimation. A comparison of CRBs is the most optimistic on the performance of collaborative estimation, it always shows that collaboration is more accurate, even when collaboration is not expected to be accurate at very low SNRs on the collaborative link. A comparison of standalone and collaborative RMSEs of position is the most pessimistic metric on the performance of collaboration, it indicates that collaborative estimation does not perform well even when it is expected to perform well at high SNRs on the collaborative link. Of all the metrics, the proposed ΔKLD selects more reliably the collaborative estimator as intuitively expected based on the SNR on the collaborative link. It does so by accounting for all statistical moments in the error distributions.

3.4 Summary

Performance metrics are used to target different purposes, and the selection of performance metrics must always be application-specific. For applications that care about the first two moments of the distribution of error and nothing else, a comparison of position root mean squared errors (RMSEs) is an adequate option. However, some applications like E911 have different requirements; position estimators have to be assessed based on their 67% and 95% bounds, and these requirements are getting more stringent over the years. For such applications that need to compare more than the central information in the error distribution, RMSE is no longer the best metric because it tends to not handle end-tail behaviors well.

In this chapter, we introduced a novel performance metric that uses the Kullback-Leibler Divergence to identify the most accurate position estimator. Two features characterize our metric: 1) it is based on an assessment of all statistical moments of the distributions, and 2) it is relative to the best performance that can be achieved based on the data at hand.

The performance metric introduced in this chapter is a more robust option for our application because, unlike RMSE, it accounts for all statistical information in its assessment of performance. Furthermore, if a theoretical distribution for its reference distribution is available, it is also theoretically tractable like RMSE. The new performance metric was successfully tested and compared against a comparison of RMSEs in both good and poor geometries, with both LOS and NLOS conditions.

After validating the performance metric through simulations, we tested the accuracy of collaborative positioning using ΔKLD, a comparison of position RMSEs and a comparison of CRBs. These simulation results confirm that the collaborative CRB is less than the standalone CRB. However, this reduction in CRBs does not necessarily mean that collaborative positioning is more accurate. In the next chapters, we analyze the implications of this reduction in CRB w.r.t the accuracy of collaborative positioning.

Chapter 4

Inaccuracy Analysis of Collaborative Positioning

In the previous chapter, we established that collaborative positioning is not always beneficial in terms of accuracy. In the literature, analyses of factors that affect the accuracy of collaborative positioning have mainly been completed via simulations. We need to define a theoretical framework within which we can analyze the accuracy of collaborative positioning. This is the focus of this chapter.

This chapter provides a theoretical analysis (using Gaussian-distributions of position estimates) of the inaccuracy of collaborative positioning. We provide a variational analysis of the improvement in accuracy due to collaboration, along with range conditions when collaborative positioning starts hurting the performance of standalone positioning.

To complete these theoretical analyses, we rely on the Kullback-Leibler Divergence (KLD). In the previous chapter, we introduced a novel model using ΔKLD and established that using ΔKLD to assess the performance of collaborative positioning is more robust than a comparison of position mean squared errors (PMSEs) because ΔKLD is a relative metric that accounts for all statistical information in the error distribution. In this chapter, we take advantage of the theoretical tractability of ΔKLD to analyze (1) factors that contribute to an increase in accuracy due to collaboration and (2) range conditions under which collaborative positioning outperforms standalone positioning.

In this chapter, we introduce a theoretical model to analyze the improvement in accuracy due to collaboration. We analyze the proposed theoretical model and identify factors that affect the accuracy of collaborative positioning. These factors are: the geometry of each node with respect to (w.r.t) its anchors, the quality of the collaborative link and the geometry of the collaborating nodes w.r.t each other. Under some conditions, when the collaborating nodes share a coordinate, the improvement in accuracy due to collaboration reaches local extrema. Range conditions on when collaboration starts to hurt the performance of standalone positioning are derived.

4.1 Using ΔKLD for theoretical analysis

Recall the definition of ΔKLD from Chapter 3.

$$\Delta KLD = KLD(f_{\Delta ref}||f_{\Delta STD}) - KLD(f_{\Delta ref}||f_{\Delta COL}) \tag{4.1}$$

ΔKLD indicates the reduction in inaccuracy, or improvement in accuracy due to collaboration. In the previous chapter, we used ΔKLD as a performance metric to identify the estimator that is most accurate in terms of position error. In that case, all pdfs used represented pdfs of position error. $f_{\Delta ref}(\delta)$ was derived empirically.

In this chapter, our goal is to identify via theoretical analysis, factors and conditions that affect the improvement in accuracy due to collaborative estimation. In this context, we derive ΔKLD using pdfs of position estimates:

$$\Delta KLD = KLD(f_{Xref}||f_{XSTD}) - KLD(f_{Xref}||f_{XCOL}) \tag{4.2}$$

f_{Xref} is theoretically derived as the pdf of position estimates from the most accurate estimator attainable. It is defined as $f_{Xref}(\mathbf{x})$ in this analytical framework. We first simplify our analysis model and later derive theoretical conclusions that we generalize.

Considering that accuracy is tightly connected to the selected position estimator, identifying the position estimator to use in our analysis is critical. We assume that the position estimator used is Torrieri's linear estimator because (1) it is linear and (2) it is in essence a one-shot version of the Maximum Likelihood Estimator [33]. This selection is justified by reference [78] which indicates that the MLE is the recommended estimator for collaborative position estimation. Hence, by selecting both a one-shot version of the MLE, and a linear estimator, we are able to predict the distribution of position estimates and theoretically analyze the performance of collaborative positioning w.r.t standalone positioning.

To simplify our framework, we also assume that all range measurements used are Gaussian-distributed in line-of-sight (LOS) range conditions, which implies that the distribution of position estimates is also Gaussian if the position estimator used is linear. Also, for simplicity, the bias difference in the multivariate definition of ΔKLD (equation (4.84)) is assumed negligible. This means that we assume both collaborative and standalone estimators to have the same bias; more specifically, we assume that both estimators are unbiased.

One of the challenges in using ΔKLD is that KLD requires prior knowledge of the pdfs of position estimates and of a reference pdf which represents the best estimator attainable. The best position estimator achievable, however, may be non-linear. So defining an accurate pdf for its position estimate is non-trivial.

Nonetheless, we assume that we can define $f_{Xref}(\mathbf{x})$. Theoretically, our assumed definition of $f_{Xref}(\mathbf{x})$ is that it is the pdf of an unbiased estimator whose covariance matrix is the inverse Fisher Information Matrix that we identify later. The choice of $f_{Xref}(\mathbf{x})$ as a Gaussian distribution can also be assumed appropriate based on the maximum entropy principle. This principle indicates that when one has no prior knowledge of a distribution, it is acceptable to select the distribution whose sufficient moments are known (*e.g:* mean, variance), and the distribution that yields the largest entropy. We know that the Gaussian distribution is fully represented by the mean and variance. Furthermore, in the family of continuous exponential pdfs, the Gaussian distribution has the largest entropy. Hence, assuming $f_{Xref}(\mathbf{x})$ to be Gaussian is an acceptable assumption.

In this work, we first focus on the Gaussian distribution and later address how other exponential distributions (specifically, the skew-normal distribution) are affected by the results presented. In regard to the distributions of position estimates, the selection of the Gaussian distribution for the pdfs of position estimates for the collaborative and standalone estimators (*i.e*: $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$, respectively) is fundamentally because it is the most common distribution used to define models and lends itself to analysis.

Finally, KLD was estimated using the Information Theoretical Estimators (ITE) toolbox [28]. This toolbox estimates all information metrics used in our simulation results. Since KLD will be estimated, we expect that some of the errors in its estimation during simulation will be due in part to the fact that the pdfs in our theoretical framework are assumed continuous while in actuality, we used discrete observables of a continuous function to generate estimates of their density functions.

4.1.1 Deriving ΔKLD

From the discussion above, the pdfs for position estimates from the collaborative ($f_{XCOL}(\mathbf{x})$) and standalone ($f_{XSTD}(\mathbf{x})$) estimators are derived in Appendix 2.A and yield:

$$f_{XCOL}(\mathbf{x}) = \left([2\pi]^4 \det(\boldsymbol{\Sigma_{COL}})\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu_{COL}})^T \boldsymbol{\Sigma_{COL}}^{\mathbf{-1}} (\mathbf{x} - \boldsymbol{\mu_{COL}})\right] \tag{4.3}$$

$$f_{XSTD}(\mathbf{x}) = \left([2\pi]^4 \det(\boldsymbol{\Sigma_{STD}})\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu_{STD}})^T \boldsymbol{\Sigma_{STD}}^{\mathbf{-1}}(\mathbf{x}-\boldsymbol{\mu_{STD}})\right] \tag{4.4}$$

The best estimator achievable, in the CRB sense, for $\mathbf{x}_A$ and $\mathbf{x}_B$ has a Gaussian distribution $f_{Xref}(\mathbf{x}) = f_{XCRB}(\mathbf{x}) \sim N(\boldsymbol{\mu_{CRB}, \Sigma_{CRB}})$ where the mean $\boldsymbol{\mu_{CRB}}$ is a 4x1 vector, and $\boldsymbol{\Sigma_{CRB}}$ is a 4-x-4 matrix equal to the inverse of the Fisher Information Matrix.

$$f_{XCRB}(\mathbf{x}) = \left([2\pi]^4 \det(\boldsymbol{\Sigma_{CRB}})\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu_{CRB}})^T \boldsymbol{\Sigma_{CRB}}^{\mathbf{-1}}(\mathbf{x}-\boldsymbol{\mu_{CRB}})\right] \tag{4.5}$$

$$\boldsymbol{\mu_{CRB}} = [x_A, y_A, x_B, y_B]^T \tag{4.6}$$

The covariance matrix of the best unbiased estimator yields the smallest CRB between the collaborative or standalone estimators. From [17] the covariance matrix of this best estimator is therefore $\boldsymbol{FIM_{COL}^{-1}}$, the inverse of the Fisher information matrix derived from the set of range measurements for the collaborative position estimator.

$$\boldsymbol{\Sigma_{CRB}} = \boldsymbol{FIM}_{COL}^{-1} \tag{4.7}$$

The expression for ΔKLD then becomes:

$$\Delta KLD = KLD(f_{XCRB}||f_{XSTD}) - KLD(f_{XCRB}||f_{XCOL}) \tag{4.8}$$

In Appendix 4.B, we derive the multivariate equation for ΔKLD when the pdfs of position estimates have multivariate Gaussian distributions. It is equal to:

$$\begin{aligned} \Delta KLD = &-\frac{1}{2}\log [2\pi]^n \det(\boldsymbol{\Sigma}_{COL}) - \frac{1}{2}\left[(\mu_{CRB}-\mu_{COL})^T \boldsymbol{\Sigma}_{COL}^{-1}(\mu_{CRB}-\mu_{COL})\right] \\ &+\frac{1}{2}\log\left([2\pi]^n \det(\boldsymbol{\Sigma}_{STD})\right) + \frac{1}{2}\left[(\mu_{CRB}-\mu_{STD})^T \boldsymbol{\Sigma}_{STD}^{-1}(\mu_{CRB}-\mu_{STD})\right] \\ &+\frac{1}{2}\left(\mathrm{tr}\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{CRB}\right\} - \mathrm{tr}\left\{\boldsymbol{\Sigma}_{COL}^{-1}\boldsymbol{\Sigma}_{CRB}\right\}\right) \end{aligned} \tag{4.9}$$

Assuming both standalone and collaborative positioning estimators are unbiased (that is, their mean-vectors are both equal to the true position vector), the bias terms cancel and ΔKLD for this analysis of Gaussian-distributed pdfs of position estimates is therefore equal to:

$$\Delta KLD = \frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} + \frac{1}{2}\mathrm{tr}\left\{\left[\boldsymbol{\Sigma}_{STD}^{-1} - \boldsymbol{\Sigma}_{COL}^{-1}\right]\boldsymbol{\Sigma}_{CRB}\right\} \tag{4.10}$$

where the reference covariance matrix is such that: $(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB})$ and $(\boldsymbol{\Sigma}_{STD} - \boldsymbol{\Sigma}_{CRB})$ are positive semi-definite.

4.2 Overview of the sign of ΔKLD

The sign of ΔKLD indicates when collaborative positioning degrades the accuracy of standalone estimation ($\Delta KLD \leq 0$), and when it improves standalone accuracy ($\Delta KLD \geq 0$). In this section, we provide an initial top-level summary of conditions that affect the sign of ΔKLD. Our theoretic analytical framework consists of analyzing the reduction in inaccuracy due to collaboration by considering ΔKLD as derived in equation (4.10).

This analysis proves that the sign of ΔKLD is determined by how the following factors interact with each other: (1) the geometric configuration of the nodes w.r.t their anchors, and (2) the efficiency of both standalone and collaborative estimators relative to the collaborative CRB.

From equation (4.10), there are two main terms in the expression of ΔKLD: the LogDet term and the Trace term. Brogan [77] refers to the covariance matrices $\mathbf{\Sigma}_{COL}$ and $\mathbf{\Sigma}_{STD}$ as GDOP matrices because they provide information on the effect of geometry in the distribution of position estimates. Since ΔKLD is heavily-dependent on the GDOP matrices, we expect geometry to play a key-role in the sign of ΔKLD.

The first term in the expression of ΔKLD (the LogDet term) represents the difference between the log determinants of the GDOP-matrices [77]. The determinant of the GDOP-matrix is related to the volume of the hyperellipsoid whose principal axes are in the direction of the eigenvectors of the GDOP-matrices, and whose semi-major axes have length equal to the eigenvalues of the matrices [77]. Since the determinant of a matrix describes its geometric volume, the difference of LogDets measures a change in the volume described by the covariance matrices. As such, it describes how the geometric configuration of the anchors w.r.t. their targets contributes to an improvement or reduction in accuracy in collaborative estimation. Specifically, the smaller the determinant of the collaborative GDOP-matrix w.r.t the determinant of the standalone GDOP-matrix, the more positive ΔKLD becomes.

We now consider the second term, namely the Trace term. By definition, an unbiased estimator is efficient if its variance reaches its CRB. Thus, the Trace term is a measure of the efficiency of both estimators relative to the collaborative CRB (CRB_{COL}). Minimizing this term is highly dependent on the efficiency of the standalone estimator w.r.t CRB_{COL}, and is bound by the difference between the standalone CRB (CRB_{STD}) and CRB_{COL}.

It is worth noting in equation (4.10) that the LogDet term turns out to be positive while the Trace terms turns out to be negative. Hence, the LogDet term determines the maximum value achievable by ΔKLD. As a result, the Trace term represents the loss in improvement in accuracy due to collaboration because it reduces ΔKLD from a positive value, to a negative value under some range conditions that are discussed later in this chapter.

This result is important because it generates the following conclusions based on our starting assumptions:

- For unbiased estimators, the maximum improvement achievable by collaborative positioning is completely defined by the geometric configuration of the nodes w.r.t their anchors, and w.r.t each other.
- For unbiased estimators, any loss in improvement in accuracy is bound by the relative efficiency of the standalone estimator w.r.t to the collaborative Cramer-Rao Bound. For unbiased efficient estimators, this relative efficiency is a function of the difference between the standalone and collaborative CRBs.
- For unbiased estimators, when the reduction in log volume described by the covariance matrices becomes less than the loss in estimator efficiency of both estimators, collaborative estimation starts to hurt the accuracy of standalone estimation.

4.3 Variational analysis

Recall the expression for ΔKLD:

$$\Delta KLD = \frac{1}{2}\log\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})} + \frac{1}{2}\mathrm{tr}\left\{\left[\mathbf{\Sigma}_{STD}^{-1} - \mathbf{\Sigma}_{COL}^{-1}\right]\mathbf{\Sigma}_{CRB}\right\} \tag{4.11}$$

From Appendix 2.A, equation (2.68), we have:

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \boldsymbol{R} \tag{4.12}$$

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{P} & -\boldsymbol{P} \\ -\boldsymbol{P} & \boldsymbol{P} \end{bmatrix} \tag{4.13}$$

where matrix $\boldsymbol{P}$ is:

$$\boldsymbol{P} = \begin{bmatrix} \frac{\cos^2(\theta_{AB})}{\sigma_{AB}^2} & \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma_{AB}^2} \\ \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma_{AB}^2} & \frac{\sin^2\theta_{AB}}{\sigma_{AB}^2} \end{bmatrix} \tag{4.14}$$

$\mathbf{R}$ depends on two terms: σ_{AB}^2 and angle θ_{AB}. Angle θ_{AB} is defined as: $\tan^{-1}\frac{(y_B-y_A)}{(x_B-x_A)}$. σ_{AB}^2 depends on the SNR on the collaborative link (SNR_{dBCOL}) as follows:

$$SNR_{dBCOL} = 10log_{10}\left(\frac{d_{AB}^2}{\sigma_{AB}^2}\right) \Longrightarrow \sigma_{AB}^2 = \frac{(x_A - x_B)^2 + (y_A - y_B)^2}{10^{\left(\frac{SNR_{dBCOL}}{10}\right)}} \tag{4.15}$$

In addition to the geometric configurations of the targets w.r.t their anchors, the expression of $\boldsymbol{R}$ implies that, in terms of accuracy, collaborative estimation differs from standalone estimation based on two more factors: the quality of the collaborative link, and the angle between the two points to locate collaboratively: $\mathbf{x}_A$, and $\mathbf{x}_B$. To identify how each of these factors affect the accuracy of collaboration, we analyze the sign of the first derivative of ΔKLD w.r.t. each of these parameters. We start by deriving ΔKLD w.r.t $\mathbf{R}$ and later use the chain rule to obtain the derivative w.r.t σ^2_{AB} and angle θ_{AB}. The sign of the first derivative of ΔKLD is used to analyze factors that increase or decrease ΔKLD.

By definition, for a matrix $\mathbf{U}(r)$, the derivative of the LogDet of $\mathbf{U}(r)$ w.r.t r is:

$$\frac{\partial \log \det \mathbf{U}(r)}{\partial r} = \operatorname{Tr}\left(\mathbf{U}(r)^{-1}\frac{\partial \mathbf{U}(r)}{\partial r}\right) \tag{4.16}$$

Thus, the derivative of ΔKLD w.r.t matrix $\mathbf{R}$ is a matrix whose elements are $\frac{\partial \Delta KLD}{\partial_{\{i,j\}}}$ where $R_{\{i,j\}}$ is the element of matrix $\mathbf{R}$ at row i and column j:

$$\frac{\partial \Delta KLD}{\partial R_{\{i,j\}}} = \frac{1}{2}\left[\operatorname{Tr}(\boldsymbol{\Sigma}_{COL}.\frac{\partial(\boldsymbol{\Sigma}^{-1}_{STD} + R_{\{i,j\}})}{\partial R_{\{i,j\}}}) - \operatorname{Tr}(\boldsymbol{\Sigma}_{CRB})\right] = \frac{1}{2}\left[\operatorname{Tr}(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB})\right] \tag{4.17}$$

Given that we established earlier that $\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB}$ is positive semi-definite, $[\operatorname{Tr}(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB})]$ is always positive. Since we are only concerned with the sign of the derivative of ΔKLD, there is no need to expand this expression any further.

4.3.1 Change in accuracy as σ_{AB} varies

Using the chain rule to compute the first derivative of ΔKLD w.r.t σ_{AB}, we get:

$$\frac{\partial \Delta KLD}{\partial \sigma_{AB}} = \frac{\partial \Delta KLD}{\partial \mathbf{R}}.\frac{\partial \mathbf{R}}{\partial \sigma_{AB}} \tag{4.18}$$

$$\frac{\partial \mathbf{R}}{\partial \sigma_{AB}} = \begin{bmatrix} \frac{\partial \boldsymbol{P}}{\partial \sigma_{AB}} & -\frac{\partial \boldsymbol{P}}{\partial \sigma_{AB}} \\ -\frac{\partial \boldsymbol{P}}{\partial \sigma_{AB}} & \frac{\partial \boldsymbol{P}}{\partial \sigma_{AB}} \end{bmatrix} \tag{4.19}$$

$$\frac{\partial \boldsymbol{P}}{\partial \sigma_{AB}} = \frac{-2}{\sigma^3_{AB}}\begin{bmatrix} \cos^2\theta_{AB} & \cos\theta_{AB}\sin\theta_{AB} \\ \cos\theta_{AB}\sin\theta_{AB} & \sin^2\theta_{AB} \end{bmatrix} \tag{4.20}$$

Thus:

$$\frac{\partial \Delta KLD}{\partial \sigma_{AB}} = -\frac{\mathbf{R}'}{\sigma^3_{AB}}\left[\operatorname{Tr}(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{ref})\right] \tag{4.21}$$

where

$$\mathbf{R}' = \begin{bmatrix} \boldsymbol{P}' & -\boldsymbol{P}' \\ -\boldsymbol{P}' & \boldsymbol{P}' \end{bmatrix} \tag{4.22}$$

$$\boldsymbol{P'} = \begin{bmatrix} \cos^2(\theta_{AB}) & \cos(\theta_{AB})\sin(\theta_{AB}) \\ \cos(\theta_{AB})\sin(\theta_{AB}) & \sin^2\theta_{AB} \end{bmatrix} \tag{4.23}$$

Given that the trace of matrix $\mathbf{R}$' includes elements on x_A, y_A, x_B, y_B, and that $\boldsymbol{P}$' along the trace is always positive and equal to two ($\mathrm{Tr}(\mathbf{R}') = 2$), we can conclude that the sign of $\frac{\partial \Delta KLD}{\partial \sigma_{AB}}$ is negative along the trace.

These results show that keeping all other unknowns constant, and increasing σ_{AB} reduces the improvement of collaborative estimation. Thus, the improvement of collaborative positioning w.r.t standalone positioning decreases as the noise on the collaborative link increases. Intuitively, this result is expected because one would expect the performance of collaborative estimation to degrade as the noise on the collaborative link increases.

To illustrate this result, we consider the following test case (Fig.4.1): the SNR on the collaborative link varies, while the SNR on all standalone links is maintained at 20 dB. Although our theoretical results were based on Torrieri's linear algorithm, which is in essence a one-shot MLE estimator, we use the standard MLE and assume LOS on all range measurements. The selection of the MLE estimator is to show that our theoretical results can be generalized to non-linear estimators and non-Gaussian distributions of position estimates. This was done to provide a practical understanding of these theoretical results, especially since the MLE is the recommended estimator in collaborative positioning [78]. Fig.4.1 indicates that our theoretical results also apply to the MLE estimator.

ΔKLD which reflects the improvement in accuracy of collaborative positioning, increases as the SNR increases; its gradient w.r.t σ_{AB} is always positive. These results validate the intuitive understanding that as the accuracy of the collaborative link increases, the accuracy of collaborative estimation improves.

4.3.2 Change in accuracy as angle θ_{AB} varies

Recall that the derivative of ΔKLD w.r.t matrix $\mathbf{R}$ is a matrix whose elements are $\frac{\partial \Delta KLD}{\partial \mathbf{R}_{\{i,j\}}}$ where $R_{\{i,j\}}$ is the element of matrix $\mathbf{R}$ at row i and column j:

$$\frac{\partial \Delta KLD}{\partial R_{\{i,j\}}} = \frac{1}{2}\left[\mathrm{Tr}(\boldsymbol{\Sigma}_{COL}.\frac{\partial(\boldsymbol{\Sigma}_{STD}^{-1} + R_{\{i,j\}})}{\partial R_{\{i,j\}}}) - \mathrm{Tr}(\boldsymbol{\Sigma}_{CRB})\right] = \frac{1}{2}\left[\mathrm{Tr}(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB})\right] \tag{4.24}$$

Using the chain rule to compute the first derivative of ΔKLD w.r.t angle θ_{AB}, we evaluate how θ_{AB} affects the accuracy of collaborative estimation:

$$\frac{\partial \Delta KLD}{\partial \theta_{AB}} = \frac{\partial \Delta KLD}{\partial \mathbf{R}}.\frac{\partial \mathbf{R}}{\partial \theta_{AB}} \tag{4.25}$$

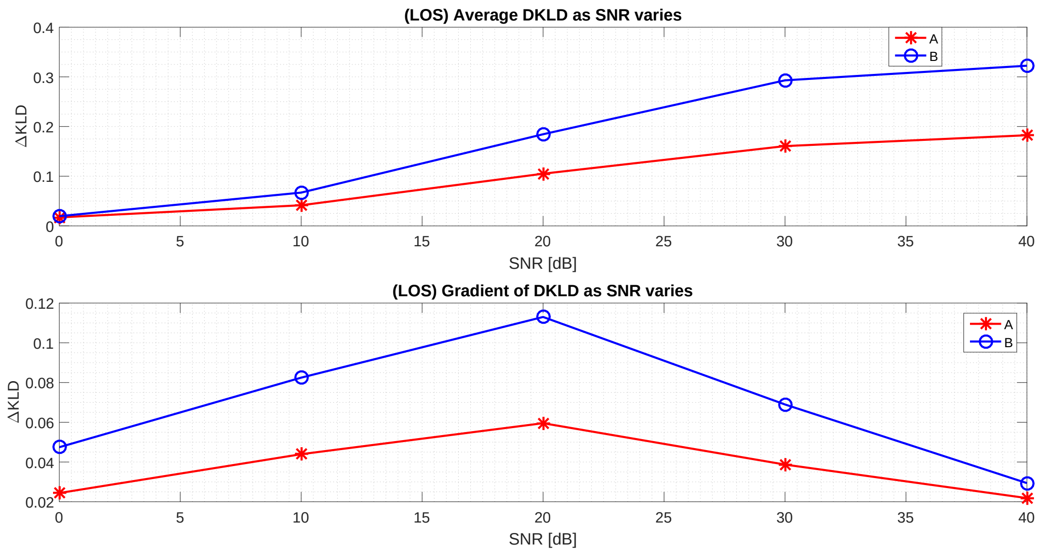


Figure 4.1: The upper curve represents the average change in accuracy (average ΔKLD) from using collaboration between nodes A and B as the SNR on the collaborative link varies. The lower plot is the derivative of the average ΔKLD as the SNR on the collaborative link varies. It is not monotonic and maximal at 20 dB, which means that ΔKLD has an inflection point at 20 dB. This shows that as the SNR on the collaborative link increases, the average ΔKLD becomes convex down, thus indicating that ΔKLD is heading towards a local maximum. Given that all standalone links have an SNR of 20 dB on their LOS range measurements, the inflection point indicates that the improvement in accuracy due to collaboration switches from a local minimum to a local maximum when the SNR on the collaborative link is the same as the SNR on all the other links. As the SNR increases, the collaborative link becomes more accurate than all other links, thus improving position accuracy due to collaboration. Improvements in accuracy due to collaboration are therefore a function of the noise on the collaborative link relative to the noise on all standalone links.

$$\frac{\partial \mathbf{R}}{\partial \theta_{AB}} = \begin{bmatrix} \frac{\partial \boldsymbol{P}}{\partial \theta_{AB}} & -\frac{\partial \boldsymbol{P}}{\partial \theta_{AB}} \\ -\frac{\partial \boldsymbol{P}}{\partial \theta_{AB}} & \frac{\partial \boldsymbol{P}}{\partial \theta_{AB}} \end{bmatrix} \tag{4.26}$$

$$\frac{\partial \boldsymbol{P}}{\partial \theta_{AB}} = \frac{1}{\sigma_{AB}^2} \begin{bmatrix} -\sin(2\theta_{AB}) & \cos(2\theta_{AB}) \\ \cos(2\theta_{AB}) & \sin(2\theta_{AB}) \end{bmatrix} \tag{4.27}$$

Thus:

$$\frac{\partial \Delta KLD}{\partial \theta_{AB}} = \frac{\mathbf{R}'}{\sigma_{AB}^2} \left[\mathrm{Tr}(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB}) \right] \tag{4.28}$$

where

$$\mathbf{R}' = \begin{bmatrix} \boldsymbol{P}' & -\boldsymbol{P}' \\ -\boldsymbol{P}' & \boldsymbol{P}' \end{bmatrix} \tag{4.29}$$

$$\boldsymbol{P}' = \begin{bmatrix} -\sin(2\theta_{AB}) & \cos(2\theta_{AB}) \\ \cos(2\theta_{AB}) & \sin(2\theta_{AB}) \end{bmatrix} \tag{4.30}$$

We consider the trace of matrix $\mathbf{R}'$ which describes changes on x_A, y_A, x_B and y_B. From the expression of $\frac{\partial \Delta KLD}{\partial \theta_{AB}}$, its magnitude varies as a function of: a) θ_{AB} and b) the magnitude of the scalar $\mathrm{Tr}(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB})$. The implications of this result are:

- Recall that $\frac{\partial \Delta KLD}{\partial \theta_{AB}}$ is a function of $\mathrm{Tr}(\boldsymbol{\Sigma}_{COL} - \boldsymbol{\Sigma}_{CRB})$, which is a function of the geometry of the nodes w.r.t their anchors. This means that as angle θ_{AB} varies, the geometry of the nodes w.r.t their anchors determines the magnitude of changes in ΔKLD.

- If variations in ΔKLD can be observed, the trace of $\mathbf{R}'$ indicates that an increase in ΔKLD for x_A results in a decrease in ΔKLD for y_A. This means that one expects the changes in accuracy for x_A and y_A to be symmetric across the x-axis. This interesting result is observed and further discussed in the simulation section of chapter 5 where we show that changes in ΔKLD for the x and y coordinates of each node vary asymmetrically along the x-axis w.r.t θ_{AB}. This also confirms simulation results on collaborative positioning that were derived by Jia et al in Fig. 8 of reference [14]: as the angle between collaborating nodes A and B varies between 0 and 360 degrees, the bias plots for node A's x_A and y_A coordinates were shown to be symmetric across the x-axis.

- If variations in ΔKLD can be observed, the trace of $\mathbf{R}'$ indicates that for both nodes, the increase and decrease in ΔKLD due to collaboration w.r.t θ_{AB} is a function of $\sin(2\theta_{AB})$. When $\sin(2\theta_{AB}) = 0$, the increase or decrease in ΔKLD is at a local extremum (minimum or maximum) because the first derivative of ΔKLD w.r.t θ_{AB} is zero. Given that $\sin(2\theta_{AB}) = 0$ if and only if $\sin(\theta_{AB}) = 0$ or $\cos(\theta_{AB}) = 0$, we can conclude that the increase or decrease in ΔKLD w.r.t θ_{AB} reaches a local extremum when nodes A and B share a coordinate. This result is significant because it allows a user with angle-of-arrival information to identify ahead of time, on average, which points will yield local maximum or local minimum on the improvement in accuracy based on the geometric configuration of each node w.r.t its anchors.

We consider a special test case for illustration. In a 300 m-by-300 m Matlab simulation environment described in Fig. 4.2. We use the Newton-Raphson MLE to estimate positions for nodes A and B and assume LOS on all range measurements. The SNR on all standalone and collaborative links is fixed at 20 dB. We vary the value of angle θ_{AB} between nodes A and B, from 0 to 2π in 21 intervals, by fixing node A and varying the location of node B in a circle of fixed radius of 40 meters. We repeat the same experiment at 100 different locations for node A by maintaining good geometric configuration and the radius between A and B fixed. At each fixed position for node A and node B, we generate 1000 position estimates for

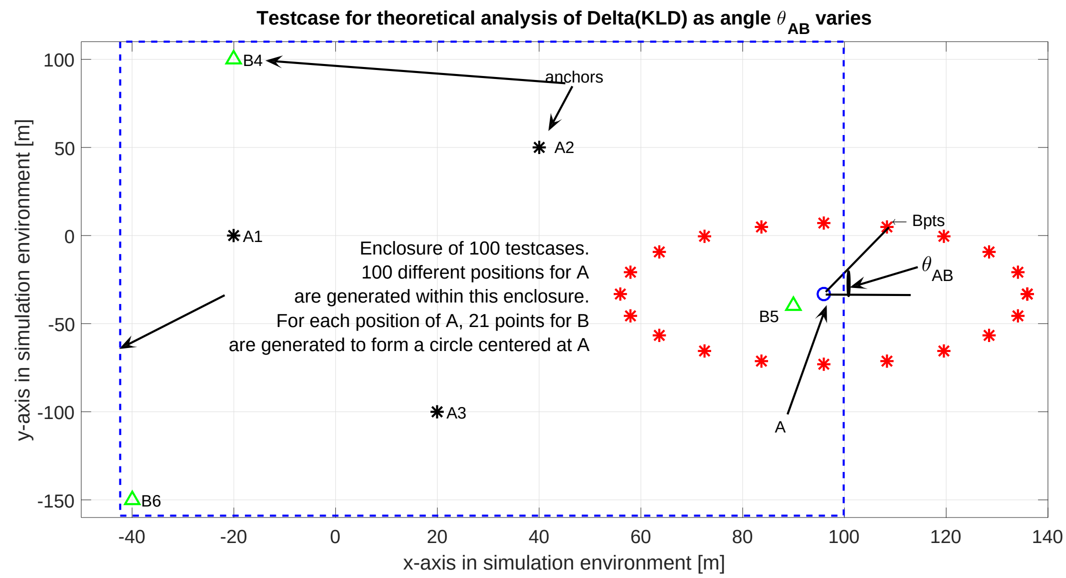


Figure 4.2: Analysis of geometry on the accuracy of collaborative estimation: Anchors vs. node positions. Point B rotates around point A in a circle such that angle θ_{AB} varies in different geometric configurations with respect to all anchors.

each node and compute ΔKLD. The results in Fig.4.3 show the average ΔKLD for the 100 different test cases. From the upper plot in Fig.4.3, we note that collaborative positioning is more accurate than standalone in the scenario considered, which means that collaboration improved the geometric configuration and contributed to an increase in accuracy. The lower plot in Fig.4.3 indicates that the sign of the derivative of ΔKLD is closely related to $-\sin(2\theta_{AB})$. Additionally, the linear correlation coefficient between the average ΔKLD of point A and $-\cos(2\theta_{AB})$ is 0.85, while the linear correlation coefficient for point B is 0.79.

In summary, factors that cause variations in ΔKLD are: a) the quality of the collaborative link, b) the geometry of the nodes with respect to their anchors, and c) the geometry of the nodes with respect to each other. Local extrema of the change in accuracy due to collaborative estimation occur whenever A and B share a coordinate, based on the geometric configuration of each node w.r.t its anchors. In other words, in collaboration, the lack of geometric diversity between the collaborating nodes can either improve standalone accuracy significantly, or degrade it significantly, based on the geometric configuration of each node w.r.t its anchors.

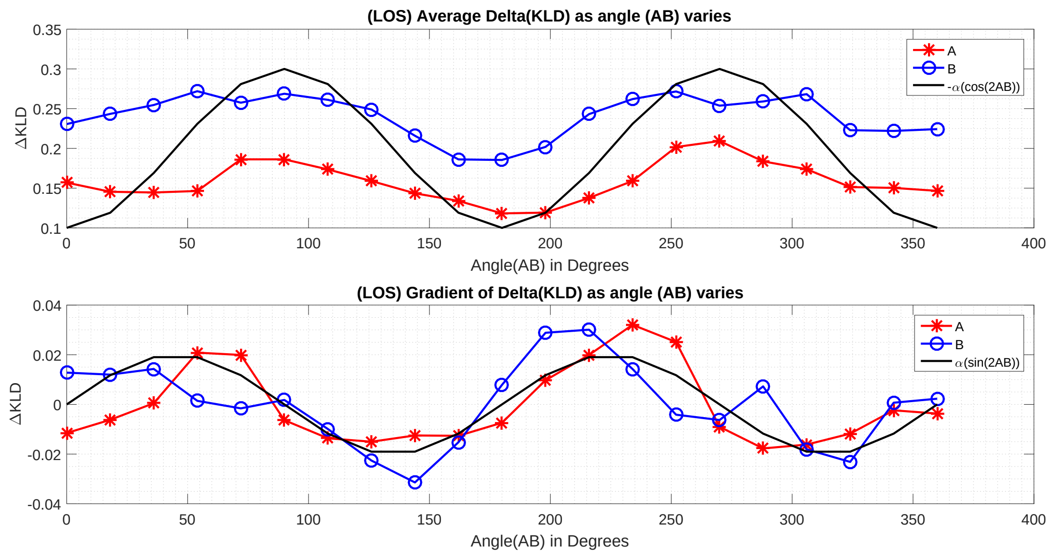


Figure 4.3: Analysis of geometry on the accuracy of collaborative estimation: Analyzing average change in accuracy as the angle between A and B varies from 0 to 2π. Average is computed over 100 different positions of A, with B rotating around A in a circle. The gradient plot confirms that when A and B share a coordinate, the increase or decrease in accuracy due to collaborative estimation is maximal.

4.4 Analyzing the sign of ΔKLD

4.4.1 Analyzing a simplified framework: the case of scalar unbiased efficient estimators

A negative ΔKLD indicates that collaborative estimation hurts standalone accuracy. A multivariate analysis of the range conditions when collaboration starts to hurt standalone positioning would consist of solving for the zeros of ΔKLD given factors that affect the covariance matrices in equation (4.10). This task is not trivial because it requires taking the derivative of the logDet of matrices. The multivariate version of ΔKLD being complex to analyze and interpret, we use its scalar version and analyze it to derive a more intuitive understanding of range conditions where collaborative positioning hurts the performance of standalone estimation. We assume that we want to estimate x_A and x_B, or y_A and y_B with standalone and collaborative positioning. In the remainder of the scalar analysis, we will only focus on the x-coordinates of nodes A and B: x_A and x_B.

In Appendix 4.C, we derive the scalar version for ΔKLD. For the sake of theoretical analysis, we consider that the estimates for scalars x_A and x_B both have a Gaussian distribution and are jointly Gaussian when estimated collaboratively. Their standalone estimator's variances are σ_A^2 and σ_B^2. In collaborative estimation, a correlation coefficient ρ is added to the off-diagonal of the covariance matrix of their 2x2 joint distribution. Hence, both standalone and collaborative estimators have the same diagonal as shown in the equations below. The reference variances σ_{AT}^2 and σ_{BT}^2 represent the collaborative CRBs for estimating scalars x_A and x_B collaboratively. Thus, having assumed unbiased estimators, we get: $\sigma_{AT}^2 \leq \sigma_A^2$ and $\sigma_{BT}^2 \leq \sigma_B^2$.

$$\boldsymbol{\Sigma_{STD}} = \begin{bmatrix} \sigma_A^2 & 0 \\ 0 & \sigma_B^2 \end{bmatrix} \tag{4.31}$$

$$\boldsymbol{\Sigma_{COL}} = \begin{bmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{bmatrix} \tag{4.32}$$

$$\boldsymbol{\Sigma_{T}} = \begin{bmatrix} \sigma_{AT}^2 & 0 \\ 0 & \sigma_{BT}^2 \end{bmatrix} \tag{4.33}$$

where ρ is a scalar that we interpret later in section 4.5.2. We assumed that the covariance matrices $\boldsymbol{\Sigma_{COL}}$ and $\boldsymbol{\Sigma_{STD}}$ from both standalone and collaborative estimators have nearly the same diagonal. This assumption in this part of the analysis is useful for two reasons: 1) it reduces the number of unknowns by making the diagonal terms of both standalone and collaborative the same, and 2) later, it allows us to address a special case in collaboration: how mutual information impacts the accuracy of collaborative positioning.

We now analyze the conditions when this assumption is acceptable. If ϕ_k is the additional term that should be added to the k^{th} element on the diagonal of $\boldsymbol{\Sigma_{COL}}$ above, then this assumption is feasible if $\phi_1 << \sigma_A^2$ such that $\sigma_A^2 + \phi_1 \approx \sigma_A^2$. Similarly, if we assume that $\phi_2 << \sigma_B^2$ such that $\sigma_B^2 + \phi_2 \approx \sigma_B^2$, the assumption above also holds. Hence, our assumption for this scalar is acceptable under such conditions.
The scalar ΔKLD is then equal to (see Appendix 4.C):

$$\Delta KLD = -\log\sqrt{1-\rho^2} - \frac{\rho^2\chi}{2(1-\rho^2)} \tag{4.34}$$

where the correlation coefficient ρ between the parameters is such that $-1 \leq \rho \leq 1$ and $0 \leq \chi \leq 2$, with:

$$\chi = \left[\frac{\sigma_{AT}^2}{\sigma_A^2} + \frac{\sigma_{BT}^2}{\sigma_B^2}\right] \tag{4.35}$$

with σ_{AT}^2 and σ_{BT}^2 being the variances of the CRBs for estimating scalars x_A and x_B collaboratively.

It is important to note here that the mutual information between x_A and x_B is always positive and equal to $-\log\sqrt{1-\rho^2}$, which is the first term in equation (4.34). This was proven in Appendix 4.D.

Equation (4.34) is displayed on Fig.4.4 as a function of ρ for different values of χ. It represents the overall change in accuracy for both nodes A and B due to collaboration.

The plot indicates that the change (improvement or reduction) in accuracy due to collaboration relies mainly on two factors: 1) the value of χ, indicative of the relative efficiency of the estimators w.r.t the collaborative CRB, and 2) ρ, the adjustment to the standalone covariance matrix due to the addition of one or more collaborative links as explained in Section 4.5. The three extrema for this plot occur at: $\rho = \{0; -\sqrt{1-\chi}; \sqrt{1-\chi}\}$.

Values $\rho = -\sqrt{1-\chi}$ and $\rho = \sqrt{1-\chi}$ are the two maxima for ΔKLD as long as $\chi < 1$. This means that as long as $\chi < 1$, collaborative estimation will be more accurate than standalone, within an interval slightly greater than or equal to $\rho = [-\sqrt{1-\chi}, \sqrt{1-\chi}]$, with ρ not exceeding interval $[-1, 1]$. This implies that when $\chi < 1$, there is always a potential for collaborative positioning to be more accurate than standalone, or as accurate. Furthermore, when χ is very small, function ΔKLD is positive over a greater interval for ρ, hence, collaborative estimation accuracy depends less on ρ. χ represents the relative efficiencies of the estimators w.r.t the collaborative CRB. This infers that when the standalone estimator efficiency is very poor relative to the collaborative CRB, the collaborative estimator tends to be more accurate, and this accuracy depends less on the value of ρ, as long as ρ is non-zero. If the standalone estimator is very accurate, then, χ increases, and collaborative estimation accuracy is better over a smaller interval for ρ. This infers that as χ increases, the accuracy of collaborative estimation depends more on the value of ρ. If χ is greater than unity, collaboration will mostly be less accurate than standalone overall, or at best perform as accurately as standalone estimation.

A $\rho = 0$ condition is essentially indicative that we gained nothing from collaboration, because when $\rho = 0$, both standalone and collaborative covariance matrices are equal.
In practice however, $\rho = 0$ cannot exist as long as there is a collaborative link between nodes A and B. To be more specific, as long as there exists a minimum level of communication connectivity between the collaborating nodes that can cause the existence of a collaborative link, ρ cannot equal zero. However, a small ρ can exist.

For unbiased estimators, a small ρ can yield a small or near zero-ΔKLD. A very small or near-zero ΔKLD indicates that collaboration did not change the accuracy of the standalone estimator. As shown in Section 4.6, a zero ΔKLD exists when the bias of the standalone estimator is maintained. In other words, if the collaborative and standalone estimators have the same bias, then, it is possible that we gain nothing from collaborating ($\Delta KLD = 0$).

In Section 4.5, we show that beyond the scalar problem, ρ describes in general the adjustment to the standalone covariance matrix due to the addition of one or more collaborative links. Hence a small ρ indicates that the collaborative matrix did not change much, while a large ρ indicates that the collaborative matrix is very different from the standalone matrix. Hence, for a small ρ, we expect very little impact from collaboration. Conversely, as ρ increases, we expect the performance of the collaborative estimator to be very different from the standalone's. This large difference in performance can result in a significant improvement in accuracy or in a significant degradation. The determining factor is the value of χ which represents the relative efficiency of both estimators w.r.t the collaborative CRB. This result indicates that understanding how the collaborative covariance matrix varies as one or more collaborative links are added will provide useful insight into assessing the benefits of collaborative positioning. In Section 4.5, we derive closed-form expressions for the change in the standalone matrix due to the addition of one or more collaborative links.

Since the collaborative CRB CRB_{COL} is smaller than the standalone CRB CRB_{STD}, the reference variances σ^2_{AT} and σ^2_{BT} refer to CRB_{COL}. Then, when standalone is very efficient for both A and B, $\chi = 2$ for the scalar case, and is maximal. As a result, the plot shows that standalone estimation is accurate all the time (for $\chi = 2$). Interpreting this result, we get the following: if the standalone estimator is efficient, the trace of its covariance matrix $\mathrm{Tr}(\boldsymbol{\Sigma}_{STD})$ is closer to CRB_{STD}. For $\mathrm{Tr}(\boldsymbol{\Sigma}_{STD})$ to be near CRB_{COL} (in order for χ to equal 2) , the difference $(CRB_{STD} - CRB_{COL})$ must tend towards zero. In such a case, collaborative estimation did not bring any benefit to standalone estimation. If both ρ and χ are large, collaboration will always worsen the standalone accuracy. In the uncertainty analysis, we show that the difference between the CRBs is a measure of the reduction in uncertainty due to collaboration. The smaller this reduction, the larger χ and the less collaboration is beneficial.

Furthermore, for unbiased estimators, the mutual information between x_A and x_B is the maximum increase in accuracy that collaboration can yield. In this scalar analysis, we note that the term $-\log\sqrt{1-\rho^2}$ is the mutual information for the distributions of x_A and x_B. Interestingly, in the 2-dimensional Euclidean space, this is equivalent to the LogDet term in equation (4.10) which reflects the geometric configuration of the nodes and their respective anchors. Thus, we can conclude the following: if the collaborative estimator is more accurate than the standalone estimator, the geometric configuration of the nodes w.r.t their anchors plays a major role in establishing the bound on the improvement in accuracy due to collaboration. In the multi-dimensional case, the geometric configuration of the nodes w.r.t their anchors contributes to the maximum improvement in accuracy due to collaboration. In this scalar analysis, under the theoretical model defined in this section, the LogDet term *is* the maximum improvement in accuracy due to collaboration.

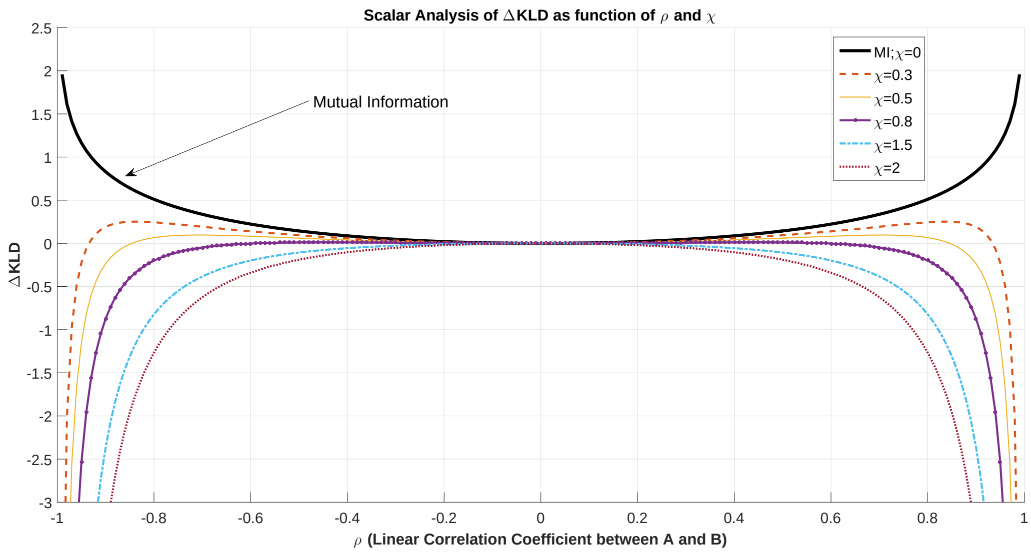


Figure 4.4: (Zoomed-in view) Sign of scalar ΔKLD for unbiased estimators. ΔKLD indicates improvement in accuracy due to collaboration. When ΔKLD is negative, standalone is more accurate than collaborative estimation. χ reflects the efficiency of the standalone estimator w.r.t the collaborative CRB: a small χ indicates a very inefficient standalone estimator w.r.t the collaborative CRB. ρ represents the adjustment to the standalone covariance matrix due to the addition of a collaborative link. A small ρ means that the collaborative covariance matrix is almost equal to the standalone covariance matrix. Hence, for small ρ, we don't expect a large change in accuracy due to collaboration. A large ρ indicates that both standalone and collaborative covariance matrices are very different. Thus we expect a significant change in accuracy (either good or bad) due to collaboration. The significant change in accuracy is beneficial (it yields an improvement in accuracy) if χ is small (*i.e*, the standalone estimator is not efficient). The significant change in accuracy is not beneficial (it yields a degradation in accuracy) if χ is large (a large χ indicates an efficient standalone estimator w.r.t to the collaborative CRB). This graph also indicates that the mutual information between estimates of A and B is the maximum improvement in accuracy achievable by the collaborative estimator for unbiased estimators when the diagonal of both collaborative and standalone covariance matrices are the same. It shows that mutual information alone cannot tell that collaborative estimation will be more accurate than standalone. It simply yields the potential in accuracy improvement that collaboration entails.

From our original assumption that the covariance matrices from the pdfs of position estimates have the same diagonal, we can derive that $KLD(f_{XCOL}||f_{XSTD})$ is the localization mutual information. This mutual information $KLD(f_{XCOL}||f_{XSTD})$ is the maximum improvement in accuracy due to collaboration if both estimators are efficient. This complements an observation from equation (4.8) of ΔKLD, that if collaborative positioning is more accurate than standalone such that $f_{Xref}(\mathbf{x}) = f_{XCOL}(\mathbf{x})$, and the standalone estimator is efficient, then, the maximum ΔKLD achievable will be $KLD(f_{XCOL}||f_{XSTD})$. This result is further addressed in Chapter 5.

Hence the following relationship holds if a) the standalone estimator is efficient (w.r.t to the standalone CRB), or has a lower variance than the standalone CRB, and b) the collaborative positioning estimator is more accurate than the standalone estimator:

$$\Delta KLD \leq KLD(f_{XCOL}||f_{XSTD}) \tag{4.36}$$

Although the conditions aforementioned are very stringent, it is worth noting that from simulation results, we observed that this relationship holds on average if a good estimator (like the MLE) is used. To put it simply, if a very good estimator (like the MLE) is used, and collaboration tends to be more accurate than standalone estimation, then, *on average* we will observe the relationship above. Furthermore, the closer ΔKLD is from $KLD(f_{XCOL}||f_{XSTD})$, the more confident we can be that the estimator used maximized the potential that collaborative positioning has to offer, given the data at hand.

One of the goals of this paper is to identify metrics that can help assess whether collaboration will improve accuracy ahead of time. The best increase in accuracy achievable defined in this section is based on prior knowledge of the distributions of position estimates. Hence, it requires prior knowledge of the distributions of the collaborative estimates. In this regard, $KLD(f_{XCOL}||f_{XSTD})$ is not a practical metric to tell ahead of time, prior to collaborating, whether collaboration will improve accuracy significantly or not. Thus, we need to assess other ways to measure the impact of collaboration based on observables. We address this in the next chapter in section 5.2.

So far, we have addressed the change in accuracy due to collaboration under a constrained theoretical framework. We now consider the impact of other conditions on our theoretical results.

4.4.2 Analyzing the case of biased estimators

If we consider the case where all estimators are biased, regardless of the conditions that generated the estimator bias (*e.g:* NLOS on range measurements, choice of position estimator, etc...), ΔKLD yields:

$$\begin{aligned}\Delta KLD_{Biased} = \frac{1}{2}\log\frac{\det(\Sigma_{STD})}{\det(\Sigma_{COL})} - \frac{1}{2}\left[(\mu_{CRB}-\mu_{COL})^T\Sigma_{COL}^{-1}(\mu_{CRB}-\mu_{COL})\right] \\ +\frac{1}{2}\left[(\mu_{CRB}-\mu_{STD})^T\Sigma_{STD}^{-1}(\mu_{CRB}-\mu_{STD})\right] \\ +\frac{1}{2}\left(\mathrm{tr}\left\{\left(\Sigma_{STD}^{-1}-\Sigma_{COL}^{-1}\right)\Sigma_{ref}\right\}\right)\end{aligned} \tag{4.37}$$

For Gaussian distributions, comparing the equation above to the unbiased form of ΔKLD from equation (4.10) indicates that when the standalone and collaborative estimators have the same bias, the improvement in accuracy due to collaboration is the same as in the unbiased case. When they each have a different bias, the difference in biases affect the value of ΔKLD by the following term, which can be either positive or negative:

$$\begin{aligned}\Delta KLD_{biased} - \Delta KLD_{unbiased} = -\frac{1}{2}\left[(\mu_{CRB}-\mu_{COL})^T\Sigma_{COL}^{-1}(\mu_{CRB}-\mu_{COL})\right] \\ +\frac{1}{2}\left[(\mu_{CRB}-\mu_{STD})^T\Sigma_{STD}^{-1}(\mu_{CRB}-\mu_{STD})\right]\end{aligned} \tag{4.38}$$

Consequently, for Gaussian distributions, if the standalone and collaborative estimators have the same bias, the improvement in accuracy due to collaboration is no longer dependent on the bias of both estimators. It is solely dependent on the covariance matrices of the estimators and the reference distribution. Hence, if both estimators share the same bias, an analysis of the improvement in accuracy due to collaboration can be restrained to analyzing the change in covariance matrices due to the acquisition of the collaborative link. This analysis of the collaborative covariance matrix is further addressed in Section 4.5.

4.4.3 Relating ΔKLD to Generalized Geometric Dilution of Precision (GGDOP)

Geometry contributes significantly to the improvement in accuracy due to collaboration. In Appendix 4.F, we derived the closed-form expression relating the multivariate ΔKLD to the Generalized GDOP (GGDOP).

By definition, if nodes A and B collaborate, the collaborative GGDOP for node A is defined as $GGDOP_{ACOL}$ such that:

$$GGDOP_{ACOL}^{-1} = \frac{\Omega}{\left(\frac{1}{\sigma_{AB}^2}+\sum_{j=1}^{w}\frac{1}{\sigma_{AAj}^2}\right)^2} \tag{4.39}$$

where w is the number of anchors of known location used to locate node A. In our defined system, $w = 3$. σ^2_{AB} is the noise variance on the collaborative link. Given angle θ_{AA_i} defined between node A and its i^{th} anchor, and angle θ_{BB_j} between node B and its j^{th} anchor, we have:

$$\begin{aligned}\Omega = &\frac{\sin^2(\theta_{AA_1} - \theta_{AA_2})}{\sigma^2_{AA_1}\sigma^2_{AA_2}} + \frac{\sin^2(\theta_{AA_1} - \theta_{AA_3})}{\sigma^2_{AA_1}\sigma^2_{AA_3}} + \frac{\sin^2(\theta_{AA_2} - \theta_{AA_3})}{\sigma^2_{AA_2}\sigma^2_{AA_3}} \\ &+ \frac{\sin^2(\theta_{AA_1} - \theta_{AB})}{\sigma^2_{AA_1}\sigma^2_{AB}} + \frac{\sin^2(\theta_{AA_2} - \theta_{AB})}{\sigma^2_{AA_2}\sigma^2_{AB}} + \frac{\sin^2(\theta_{AA_3} - \theta_{AB})}{\sigma^2_{AA_3}\sigma^2_{AB}}\end{aligned} \tag{4.40}$$

Since the logDet term in the expression of ΔKLD is a positive value, the sign of ΔKLD is mainly dependent on the magnitude of the trace term, which is negative and derived in Appendix 4.F. We note that the trace term is a function of the collaborative Generalized GDOPs of both nodes A and B.

This result confirms that the accuracy of collaborative estimation is dependent on the collaborative Generalized GDOP. We note that when both $GGDOP_{ACOL}$ and $GGDOP_{BCOL}$ are small, the trace term in the expression of ΔKLD is minimized. This implies that a small collaborative Generalized GDOP will tend to minimize the trace term and increase ΔKLD. Hence, a valid criteria to consider for maximizing ΔKLD is to minimize the collaborative Generalized GDOP. In Chapter 6, we derive and test a criteria based on this conclusion.

4.5 Adjustment to covariance matrix due to collaboration

As discussed in section 4.4.2, for Gaussian distributions of position estimates, when a linear estimator is used and the collaborative and standalone estimators share the same bias, an analysis of ΔKLD can be constrained to the analysis of covariance matrices only. The problem with a covariance matrix analysis is that although deriving $\mathbf{\Sigma}^{-1}_{COL}$ is known (equation (2.64) in Appendix 2.A), deriving its inverse ($\mathbf{\Sigma}_{COL}$) can be very computational intensive, especially when we have more than 2 collaborating nodes. Hence it is worth identifying a more efficient way to predict $\mathbf{\Sigma}_{COL}$ from: a) the standalone covariance matrix and b) the set of collaborative links range measurements.
In this section, we address this concern by deriving the closed-form expression of the collaborative covariance matrix $\mathbf{\Sigma}_{COL}$ based on the addition of a collaborative link to the standalone problem.

The results from this section are useful because they provide the tools to analyze the change between standalone and collaborative matrices for the case of linear estimators. In cases where the collaborative problem involves more than two nodes, this approach yields a significant processing cost reduction when it comes to analyzing how different the standalone and collaborative covariance matrices are. The results in this section can also be used to derive the collaborative CRB more efficiently than a traditional computation for the inverse of the FIM.

Furthermore, in section 4.4, we claimed that the more different the covariance matrices are, the larger ρ is, and the greater of an impact collaboration has on the standalone positioning accuracy, whether good or bad. By deriving closed-form expressions of $\mathbf{\Sigma}_{COL}$ in terms of $\mathbf{\Sigma}_{STD}$, we can assess this impact in practical applications, assuming a one-shot MLE (like Torrieri's) is used. Hence the contributions of this section are two-fold: 1) derive a closed-form expression for the covariance matrix $\mathbf{\Sigma}_{COL}$, given the change in the standalone positioning problem due to the acquisition of one or more collaborative links, and 2) derive an interpretation for ρ beyond the scalar case.

4.5.1 Deriving a closed-form expression for $\mathbf{\Sigma}_{COL}$

In Appendix 4.E, we derive a closed form expression for $\mathbf{\Sigma}_{COL}$, given the change in the standalone covariance matrix. We use the Sherman-Morrison equation to achieve this goal. In this section, we summarize the results detailed in the appendix.

We observe that:

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \mathbf{R} \tag{4.41}$$

where matrix $\mathbf{R} = \frac{\mathbf{b}\mathbf{b}^T}{\sigma_{AB}^2}$ with vector $\mathbf{b}$ equal to:

$$\mathbf{b} = [\cos(\theta_{A_0B_0}), \sin(\theta_{A_0B_0}), -\cos(\theta_{A_0B_0}), -\sin(\theta_{A_0B_0})]^T \tag{4.42}$$

$\theta_{A_0B_0}$ represents the angle between the original guesses for nodes A ($\mathbf{x}_{A_0}$) and B ($\mathbf{x}_{B_0}$) as defined in Torrieri's estimator in Appendix 2.A. σ_{AB}^2 represents the variance of the noise on the observed range measurements across the collaborative link.

Using the Sherman-Morrison's formula, we get:

$$\begin{aligned}\mathbf{\Sigma}_{COL} &= (\mathbf{\Sigma}_{STD}^{-1} + \ \mathbf{R})^{-1} \\ &= \mathbf{\Sigma}_{STD} - \frac{r}{\sigma_{AB}^2}(\mathbf{\Sigma}_{STD}\mathbf{b}\mathbf{b}^T\mathbf{\Sigma}_{STD}) \\ &= \mathbf{\Sigma}_{STD} - \frac{r}{\sigma_{AB}^2}(\mathbf{\Sigma}_{STD}\mathbf{b}\mathbf{b}^T\mathbf{\Sigma}_{STD})\end{aligned} \tag{4.43}$$

with scalar r equal to:

$$r = \frac{1}{1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b}} \tag{4.44}$$

To summarize, using the Torrieri's linear estimator, when we collaborate between two nodes, $\mathbf{\Sigma}_{STD}^{-1}$ is updated by adding matrix $\mathbf{R}$ to yield a new matrix $\mathbf{\Sigma}_{COL}^{-1}$. The Sherman-Morrison's formula allows us to predict $\mathbf{\Sigma}_{COL}$ from $\mathbf{\Sigma}_{STD}$ and $\mathbf{R}$ for the case of collaboration between two nodes at a lower computation cost than generating the inverse of $\mathbf{\Sigma}_{COL}^{-1}$. For two or more collaborating nodes, the Morrison formula is the applicable generalized formula to derive $\mathbf{\Sigma}_{COL}$. This generalization-case is more detailed in Appendix 4.E, and summarized below.

Consider that we want to assess the impact of collaborative positioning on standalone positioning when we collaborate between three nodes A, B and D. We assume that both standalone and collaborative estimators are unbiased. Thus, they have the same bias. From section 4.4.2 this implies that an analysis of the impact of collaborative positioning on the accuracy of standalone positioning can be constrained to an analysis of covariance matrices. We define C_n^k as the number of k-combinations in a set of n elements:

$$C_n^k = \frac{n!}{k!(n-k)!} \tag{4.45}$$

where $n!$ stands for factorial n.

If C_3^2 is the number of 2-combinations in a set of 3 elements, in the collaborative problem, there can be up to $C_3^2 = 3$ collaborative links between any two nodes. Hence, $\mathbf{\Sigma}_{COL}^{-1}$ then takes the form:

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \mathbf{R} \tag{4.46}$$

where the new generalized matrix $\mathbf{R}$ is of the form

$$\mathbf{R} = \mathbf{UVU}^T \tag{4.47}$$

In a 2-dimensional Euclidean space, if there are 3 collaborative links between nodes A, B and D, matrix $\mathbf{U}$ is a 6x3 matrix and V is a 3x3. More generally, in an h-dimensional Euclidean space, if there are t collaborating nodes, there can be up to C_t^2 collaborative links between any two nodes.
This results in matrix $\mathbf{R} = \mathbf{UVU}^T$ where matrix $\mathbf{U}$ is an ht-by-C_t^2 matrix and $\mathbf{V}$ is a C_t^2-by-C_t^2 matrix. Matrix $\mathbf{U}$ contains information on the geometric configuration between the collaborating nodes and matrix $\mathbf{V}$ contains information about the variance of the range measurements for all collaborative links.

When the collaborative positioning problem concerns three or more nodes, the generalized form of the Sherman-Morrison equation is used. It is called the Woodbury identity. Hence, using the Woodbury identity, we get $\mathbf{\Sigma}_{COL}$ as follows:

$$\begin{aligned}\mathbf{\Sigma}_{COL} &= (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{R})^{-1} \\ &= (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{U}\mathbf{V}\mathbf{U}^T)^{-1} \\ &= \mathbf{\Sigma}_{STD} - \mathbf{\Sigma}_{STD}\mathbf{U}(\mathbf{V}^{-1} + \mathbf{U}^T\mathbf{\Sigma}_{STD}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Sigma}_{STD}\end{aligned} \tag{4.48}$$

The importance of this result is that, when using a linear estimator, the Sherman-Morrison formula, and in general the Woodbury identiy allow us to predict the covariance matrix of the collaborative estimator given: a) $\mathbf{\Sigma}_{STD}$, and b) the set of acquired collaborative links. In cases where the standalone and collaborative estimators have the same bias, deriving this new covariance matrix ahead of time, can provide useful information on the impact that collaboration has on standalone estimation, at a significantly reduced computation cost.

4.5.2 Interpreting ρ beyond the scalar analysis case

We integrate the expression of $\mathbf{\Sigma}_{COL}$ from equation (4.43) into the definition of ΔKLD when nodes A and B collaborate. We get a simplified form of ΔKLD that can be proven to be:

$$\Delta KLD = -\frac{1}{2}\left[\log\det\left(\mathbf{I}_4 - r\mathbf{R}\mathbf{\Sigma}_{STD}\right) - \mathrm{Tr}(\mathbf{R}\mathbf{\Sigma}_{CRB})\right] \tag{4.49}$$

where $\mathbf{I}_4$ is the 4x4 identity matrix, and matrix $\mathbf{R} = \frac{\mathbf{b}\mathbf{b}^T}{\sigma_{AB}^2}$ with:

$$\mathbf{b} = [\cos(A_0B_0), \sin(A_0B_0), -\cos(A_0B_0), -\sin(A_0B_0)]^T \tag{4.50}$$

Recall the scalar form of ΔKLD from Section 4.4, equation (4.34).

$$\Delta KLD = -\log\sqrt{1-\rho^2} - \frac{\rho^2\chi}{2(1-\rho^2)} \tag{4.51}$$

By observation of the first two terms in equations (4.49) and (4.51), we get that $(1-\rho^2)$ in the scalar case, is equivalent to the determinant of matrix $(\mathbf{I}_4 - r\mathbf{R}\mathbf{\Sigma}_{STD})$ in the generalized multivariate case. From this equivalence, we can conclude that ρ describes the adjustment to the standalone matrix due to the acquisition of a collaborative link. Hence, the larger ρ, the greater the difference between the standalone and collaborative estimators, and the greater the impact that collaboration can have on standalone estimation. As discussed in Section 4.4, whether this impact is a significant improvement or a significant degradation in accuracy, is determined by the relative efficiency of both standalone and collaborative estimators. w.r.t the collaborative CRB.

4.6 A Lower Bound to $KLD(f_{XCOL}||f_{XSTD})$

Given the pdfs of position estimates for the standalone and collaborative estimator, we analyze the lower bound to $KLD(f_{XCOL}||f_{XSTD})$ because it provides more insight into the trades involved in the improvement in accuracy due to collaboration. Recall ΔKLD:

$$\Delta KLD = KLD(f_{Xref}||f_{XSTD}) - KLD(f_{Xref}||f_{XCOL}) \tag{4.52}$$

If the standalone estimator is efficient or better, and if the collaborative estimator is the best it can be (that is $f_{XCOL}(\mathbf{x}) = f_{Xref}(\mathbf{x})$), then, the following occurs:

$$\Delta KLD = KLD(f_{XCOL}||f_{XSTD}) - KLD(f_{XCOL}||f_{XCOL}) = KLD(f_{XCOL}||f_{XSTD}) \tag{4.53}$$

which implies that ΔKLD is positive, maximal and equal to $KLD(f_{XCOL}||f_{XSTD})$ under the conditions aforementioned.

This implies that if $f_{Xref}(\mathbf{x}) = f_{XCOL}(\mathbf{x})$, the best improvement in accuracy due to collaboration is $KLD(f_{XCOL}||f_{XSTD})$. Hence, the closer ΔKLD is to $KLD(f_{XCOL}||f_{XSTD})$, the more we can confirm that the estimator used has maximized the achievable gain due to collaboration. Similarly, the farther ΔKLD is from $KLD(f_{XCOL}||f_{XSTD})$, the more we know that collaboration is not yielding the best improvement achievable.

We now consider $KLD(f_{XCOL}||f_{XSTD})$ as the multivariate KLD between two Gaussian distributions with different biases and covariance matrices. In this section, we evaluate the lower bound to $KLD(f_{XCOL}||f_{XSTD})$ using the Kullback-Leibler inequality [29]. The Kullback-Leibler inequality for $KLD(f_{XCOL}||f_{XSTD})$ is defined as [29]:

$$\Psi^*_{STD}(\boldsymbol{\mu}_{COL}) \leq KLD(f_{XCOL}||f_{XSTD}) \tag{4.54}$$

where:

- $\boldsymbol{\mu}_{COL}$ is the first moment of $f_{XCOL}(\mathbf{x})$.
- Ψ^*_{STD} is the convex conjugate of the cumulant generating function of the standalone estimator defined in equation 4.183 in Appendix 4.I. The cumulant generating function is the log of the moment generating function.

The lower bound is derived in Appendix 4.I. It follows from that derivation that a lower bound to $KLD(f_{XCOL}||f_{XSTD})$ is:

$$0 \leq \frac{1}{2}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T \boldsymbol{\Sigma}^{-1}_{STD}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD}) \leq KLD(f_{XCOL}||f_{XSTD}) \tag{4.55}$$

This equation indicates that when the bias of the standalone and collaborative estimators are different, the lowest bound to the best performance due to collaboration is a function of 1) the geometric configuration of the nodes w.r.t their anchors and 2) the change in bias w.r.t the standalone covariance matrix.

Kullback Leibler's inequality can be applied in general to non-Gaussian pdfs of position estimates, to assess the lower bound to $KLD(f_{XCOL}||f_{XSTD})$. For Gaussian pdfs, some important conclusions follow from this inequality.

- The lower bound derived is in agreement with our theoretical analysis of the accuracy in Section 4.4. We had assumed that both estimators had the same bias (they were unbiased). As such, the lower bound to the best improvement in accuracy due to collaboration was zero as shown on Fig. 4.4.
- The lower bound to $KLD(f_{XCOL}||f_{XSTD})$ is zero when both standalone and collaborative estimators share the same bias.
- As long as the biases of the collaborative and standalone estimators are not equal, there will always be a non-zero positive lower bound to $KLD(f_{XCOL}||f_{XSTD})$. This implies that there will always be a non-zero potential to improve accuracy through collaboration as long as maintaining the standalone bias is not enforced. This is in accordance with the common trade-off between reducing an estimator variance while trying to maintain its bias.
- The minimum $KLD(f_{XCOL}||f_{XSTD})$ does not depend on the variance of the collaborative estimator, but on its mean and on the mean and variance of the standalone estimator. This implies that the minimal improvement in accuracy is dependent on a) the performance of the standalone estimator and b) the difference in bias of both standalone and collaborative estimators.

4.7 Generalization: Assessing the case of asymmetric distributions

In our system definition, and in our theoretical framework, we assumed Gaussian range measurements with LOS, which when used with a linear position estimator, yield Gaussian-distributed position estimates. Since our analysis so far was limited to an analysis of variance and mean, the advantage of using KLD instead of MSE is not evident because both use these same moments, and MSE yields an even lower computation cost. Nonetheless, it is worth mentioning that using KLD for theoretical analysis of Gaussian distributions was still more robust than a comparison of position MSEs because KLD compares the covariance matrices to the collaborative CRB in addition to comparing the covariance matrices to each other. Hence using KLD for Gaussian distributions still has an advantage because it additionally assesses the efficiency of the estimators. In this section, the advantages of using KLD over MSE is evident because KLD is shown to account for other moments beyond the mean and covariance.

Knowing that in practice, the distribution of position estimates does not generally yield a Gaussian distribution, we consider the case where the pdfs are asymmetric (non-Gaussian). It is of interest to characterize how using ΔKLD under normal (Gaussian) assumptions differs from using ΔKLD when pdfs being compared are non-Gaussian.

As Serfling [66] indicates, one way to approach the case of asymmetric pdfs (for the case of skewed pdfs) is to model the skewness using distributions that have been generated to model the skewness. Examples include: the skew-t, skew-elliptical and skew-normal distributions. Using this type of distributions allows us to parametrize features like the presence of 'heavy tails' in the pdf of position estimates, which have been known to degrade the location accuracy of position estimators as mentioned by Zekavat and Buehrer [67], and have also been hard to assess via the use of goodness measures. We show here that KLD lends itself even more useful with asymmetric pdfs as it accounts for all the statistical information of the pdfs, and assesses estimation error at all levels, including heavy-tails.
As an example of asymmetric pdfs, we select the skew-normal distribution introduced by [69], which has a non-zero skewness. For additional information and reference, reference [68] extends the definition of Azzalini's function to the case of non-zero kurtosis.

We consider that position estimates for both standalone and collaborative estimators have a multivariate skew-normal distribution. Such a distribution parametrizes the skewness, so that we can better visualize and understand how it is being considered in the comparison between pdfs of standalone and collaborative estimates. We use the results in [46] to derive the KLD for Multivariate Skew-Normal Distributions. Knowing that MSE would only compare the bias and covariance matrices, we show that using KLD compares the skewness in addition to the mean and covariance. We conclude by generalization that higher moments than the skewness would also be captured with a comparison of KLDs. Thus, we can say that if the pdfs of position estimates have non-zero higher moments like skewness and kurtosis, MSE will not account for them, while KLD will. Furthermore, we evaluate the changes in ΔKLD if we assume that the pdfs of positioning estimates for the collaborative and standalone estimators $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$, and the reference pdf $f_{Xref}(\mathbf{x})$ are no longer Gaussian.

4.7.1 The Skew-Normal Distribution

In 1996, Azzalini and Valle introduced the multivariate closed-form expression of the skew-normal (SN) distribution [69] that yields a normal multivariate pdf when the skewness is

zero, and also marginal pdfs that are skew-normal.
The multivariate skew-normal pdf of a k-dimensional random vector $\mathbf{Z}$ of location vector $\boldsymbol{\xi}$, dispersion matrix $\boldsymbol{\Omega}$, and shape/skewness parameter $\boldsymbol{\eta}$ is $\mathbf{Z} \sim SN_k(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\eta})$ and its pdf is as follow:

$$\boldsymbol{f_Z(z) = 2\phi_k(z; \xi, \Omega)\Phi[\eta^T \Omega^{-1/2}(z - \xi)]} \tag{4.56}$$

where Φ is the univariate cumulative distribution function of the normal distribution $N_1(0, 1)$:

$$\Phi(x) = \int_{\infty}^{x} \phi_1(t)dt = \frac{1}{2}\left(1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right) \tag{4.57}$$

$\phi_k(z; \boldsymbol{\xi}, \boldsymbol{\Omega}) = (\det(\boldsymbol{\Omega}))^{-1/2} \phi_k(z_0)$, with $z_0 = \boldsymbol{\Omega}^{-1/2}(z - \boldsymbol{\xi})$ is the pdf of the k-variate normal distribution $N_k(\boldsymbol{\xi}, \boldsymbol{\Omega})$ and with $(\det(\boldsymbol{\Omega}))^{-1/2}$ being the reciprocal of the square root for the determinant of the k-by-k covariance matrix.
It is worth noting that the mean, variance, and skewness of the skew-normal distribution are defined as follows:

$$\text{Mean} = \frac{\sqrt{\frac{2}{\pi}}\boldsymbol{\eta}\boldsymbol{\Omega}}{\sqrt{\boldsymbol{\eta}^2 + 1}} + \boldsymbol{\xi} \tag{4.58}$$

$$\text{Variance} = \left(1 - \frac{2\boldsymbol{\eta}^2}{\pi\left(\boldsymbol{\eta}^2 + 1\right)}\right)\boldsymbol{\Omega}^2 \tag{4.59}$$

$$\text{Skewness} = \frac{\sqrt{2}(4 - \pi)\boldsymbol{\eta}^3}{\left((\pi - 2)\boldsymbol{\eta}^2 + \pi\right)^{3/2}} \tag{4.60}$$

For Gaussian distributions, the skewness, or third moment is zero. In this section, we consider the case of non-zero skewness, specifically the case where pdfs of position estimates have a skew-normal distribution.

4.7.2 The Kullback-Leibler Divergence for multivariate Skew-Normal Distributions

Valle et al derived the multivariate Kullback-Leibler Divergence between the pdfs of two k-dimensional vectors such that: $f_{X1}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi_1, \Omega_1, \eta_1})$ and $f_{X2}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi_2, \Omega_2, \eta_2})$ [46] as follow:

$$KLD(f_{X1}||f_{X2}) = KLD(f_{X01}||f_{X02}) + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)^T\boldsymbol{\Omega}_2^{-1}\delta_1 + E[\log\{2\Phi(W_{1-1})\}] - E[\log\{2\Phi(W_{2-1})\}] \tag{4.61}$$

where $KLD(f_{X01}||f_{X02})$ is the KLD for normal multivariate distributions, with $f_{X01}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi}_1, \boldsymbol{\Omega}_1, 0)$ and with $f_{X02}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi}_2, \boldsymbol{\Omega}_2, 0)$. The thorough derivation of the KLD between normal multivariate distributions $KLD(f_{X01}||f_{X02})$ was provided in Appendix 4.A.

$$KLD(f_{X01}||f_{X02}) = \frac{1}{2}\left[\log\left(\frac{\det(\boldsymbol{\Omega}_2)}{\det(\boldsymbol{\Omega}_1)}\right) - k + \text{tr}(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Omega}_1) + (\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)^T\boldsymbol{\Omega}_2^{-1}(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2)\right] \tag{4.62}$$

where:

$$W_{i-j} \sim SN_1\left(\eta_i^T(\xi_j-\xi_i), \eta_i^T\Omega_j\eta_i, \frac{\eta_i^T\delta_j}{\sqrt{\eta_i^T\Omega_j\eta_i-(\eta_i^T\delta_j)^2}}\right) \tag{4.63}$$

where:

$$\delta = \frac{\Omega\eta}{\sqrt{1+\eta^T\Omega\eta}} \tag{4.64}$$

We have:

$$W_{1-1} \sim SN_1\left(0, \eta_1^T\Omega_1\eta_1, (\eta_1^T\Omega_1\eta_1)^{1/2}\right) \tag{4.65}$$

$$W_{1-2} \sim SN_1\left(\eta_1^T(\xi_2-\xi_1), \eta_1^T\Omega_2\eta_1, \frac{\eta_1^T\delta_2}{\sqrt{\eta_1^T\Omega_2\eta_1-(\eta_1^T\delta_2)^2}}\right) \tag{4.66}$$

$$W_{2-1} \sim SN_1\left(\eta_2^T(\xi_1-\xi_2), \eta_2^T\Omega_1\eta_2, \frac{\eta_2^T\delta_1}{\sqrt{\eta_2^T\Omega_1\eta_2-(\eta_2^T\delta_1)^2}}\right) \tag{4.67}$$

$$W_{2-2} \sim SN_1\left(0, \eta_2^T\Omega_2\eta_2, (\eta_2^T\Omega_2\eta_2)^{1/2}\right) \tag{4.68}$$

Given the position estimates for nodes A and B in a two-dimensional system, in both standalone and collaborative positioning, and given the standalone $f_{XSTD} \sim SN_k(\boldsymbol{\xi}_{STD}, \boldsymbol{\Omega}_{STD}, \boldsymbol{\eta}_{STD})$ and collaborative pdfs $f_{XCOL} \sim SN_k(\boldsymbol{\xi}_{COL}, \boldsymbol{\Omega}_{COL}, \boldsymbol{\eta}_{COL})$, as well as the pdf of the best estimator achievable (most accurate reference) $f_{XT} \sim SN_k(\boldsymbol{\xi}_T, \boldsymbol{\Omega}_T, \boldsymbol{\eta}_T)$, if $f_{XT0}(\mathbf{x})$, $f_{XSTD0}(\mathbf{x})$ and $f_{XCOL0}(\mathbf{x})$ represent multivariate normal distributions of position estimates, we get:

$$\begin{aligned} KLD(f_{Xref}||f_{XSTD}) = KLD(f_{XT0}||f_{XSTD0}) + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_T-\boldsymbol{\xi}_{STD})^T\boldsymbol{\Omega}_{STD}^{-1}\delta_T \\ +E[\log\{2\Phi(W_{T-T})\}] - E[\log\{2\Phi(W_{T-STD})\}] \end{aligned} \tag{4.69}$$

$$\begin{aligned} KLD(f_{Xref}||f_{XCOL}) = KLD(f_{XT0}||f_{XCOL0}) + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_T-\boldsymbol{\xi}_{COL})^T\boldsymbol{\Omega}_{COL}^{-1}\delta_T \\ +E[\log\{2\Phi(W_{T-T})\}] - E[\log\{2\Phi(W_{T-COL})\}] \end{aligned} \tag{4.70}$$

It is worth noting here that $KLD(f_{Xref}||f_{XCOL})$ and $KLD(f_{Xref}||f_{XSTD})$ listed below are both functions of: a) $KLD(f_{XT0}||f_{XCOL0})$ and $KLD(f_{XT0}||f_{XSTD0})$ which assumes all pdfs are Gaussian and b) additional terms that must address the asymmetry of the distributions being compared. This confirms that KLD and consequently ΔKLD, which is the difference between the two upper equations, will account for other moments beyond the mean and variance. Hence, ΔKLD will account for all moments available in the pdfs being compared.

4.7.3 Error in assuming $f_{Xref}(\mathbf{x})$ incorrectly

In our theoretical model, we assumed that all distributions were Gaussian and derived ΔKLD_0 that can also be defined as shown in equation (4.163).

In Appendix 4.G, we derive the errors due to incorrectly estimating ΔKLD for the case of skew-normal pdfs.

Let ΔKLD_{SN} represent ΔKLD assuming all pdfs used are skew-normal. Let ΔKLD_0 represent ΔKLD assuming all pdfs used are Gaussian-distributed. The errors in incorrectly assessing ΔKLD if we assume all pdfs are Gaussian, while they are all skew-normal is:

$$\begin{aligned}\Delta KLD_{SN} - \Delta KLD_0 = &\sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T \boldsymbol{\Omega}_{STD}^{-1}\delta_{ref} - E[\log\{2\Phi(W_{STD-ref})\}] \\ &- \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T \boldsymbol{\Omega}_{COL}^{-1}\delta_{ref} + E[\log\{2\Phi(W_{COL-ref})\}]\end{aligned} \tag{4.71}$$

If $f_{Xref}(\mathbf{x})$ is skew-normal but , while $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ remain Gaussian, the error in our theoretical model due to incorrectly assuming $f_{Xref}(\mathbf{x})$ is derived below.

Let $\Delta KLD_{SN,0}$ represent ΔKLD when $f_{Xref}(\mathbf{x})$ is skew-normal and $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are Gaussian distributed. When $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are Gaussian pdfs, from the previous equation, $\Phi(W_{COL-ref}) = 0$ and $\Phi(W_{COL-ref}) = 0$, which yields:

$$\begin{aligned}\Delta KLD_{SN,0} &= H(f_{ref}||f_{STD}) - H(f_{ref}||f_{COL}) \\ &= \Delta KLD_0 + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T \boldsymbol{\Omega}_{STD}^{-1}\delta_{ref} - \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T \boldsymbol{\Omega}_{COL}^{-1}\delta_{ref}\end{aligned} \tag{4.72}$$

The error in incorrectly assuming $f_{Xref}(\mathbf{x})$ while $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are assumed Gaussian is:

$$\Delta KLD_{SN,0} - \Delta KLD_0 = \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T \boldsymbol{\Omega}_{STD}^{-1}\delta_{ref} - \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T \boldsymbol{\Omega}_{COL}^{-1}\delta_{ref} \tag{4.73}$$

4.7.4 Summary

In this section, we note that our theoretical model ΔKLD_0 which assumes all pdfs are Gaussian is included in all combinations of ΔKLD considered (equations (4.71) and (4.72)). This implies that beyond the case where all distributions are Gaussian-distributed, ΔKLD

accounts for other moments of the distributions being compared. As such, it is more robust than a comparison of MSEs which only focuses on the first two moments of the distributions being considered.

4.8 Conclusion

In this chapter, we introduced a theoretical model that we use to analyze the improvement in accuracy due to collaboration. This model is based on the Kullback-Leibler Divergence. It uses ΔKLD with the pdfs of position estimates of both standalone and collaborative estimators.

Through a variational analysis, we derived factors that affect increase and decrease in ΔKLD assuming Gaussian distributions of position estimates. We also assessed range conditions within which collaborative positioning starts to hurt standalone accuracy.

From the inaccuracy analysis of ΔKLD, we found that factors that affect the improvement in accuracy due to collaborative estimation are 1) the geometric configuration of the nodes w.r.t their anchors, 2) the quality of the collaborative link and 3) the angle between the collaborating nodes.

We also addressed how the theoretical results from our analysis will be impacted in cases where some assumptions in our theoretical model do not hold. Our theoretical model assumed Gaussian distributions. Hence, we derived the incorrectness of our model in cases where all pdfs of position estimates are skew-normal (asymmetric). These derivations further confirmed that when the pdfs are non-Gaussian, using KLD is a more robust analytical approach because it provides the means to analyze how other moments (beyond mean and variance) can affect accuracy due to collaboration.

Appendix

4.A Kullback-Leibler Divergence between multivariate Gaussian distributions

In this appendix, we derive the Kullback-Leibler Divergence (KLD) between multivariate Gaussian distributions.

Given two multivariate Gaussian distributions $f_X(\mathbf{x})$ and $g_X(\mathbf{x})$ of nx1 mean vectors $\boldsymbol{\mu}_f$ and $\boldsymbol{\mu}_g$, and nxn covariance matrices $\boldsymbol{\Sigma}_f$ and $\boldsymbol{\Sigma}_g$, the Kullback-Leibler Divergence between both pdfs is:

$$\begin{aligned}
KLD(f_X||g_X) &= \int f_X(\mathbf{x}) \log \frac{f_X(\mathbf{x})}{g_X(\mathbf{x})} d\mathbf{x} \\
&= \int f_X(\mathbf{x}) \log \frac{([2\pi]^n \det(\boldsymbol{\Sigma}_f))^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_f})^T \boldsymbol{\Sigma}_f^{-1}(\boldsymbol{x}-\boldsymbol{\mu_f})\right]}{([2\pi]^n \det(\boldsymbol{\Sigma}_g))^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_g})^T \boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu_g})\right]} d\mathbf{x} \\
&= E_f\left[\log \frac{\det(\boldsymbol{\Sigma}_f)^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_f})^T \boldsymbol{\Sigma}_f^{-1}(\boldsymbol{x}-\boldsymbol{\mu_f})\right]}{\det(\boldsymbol{\Sigma}_g)^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_g})^T \boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu_g})\right]}\right] \\
&= -\frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_f)}{\det(\boldsymbol{\Sigma}_g)} - \frac{1}{2}E_f\left[(\boldsymbol{x}-\boldsymbol{\mu_f})^T \boldsymbol{\Sigma}_f^{-1}(\boldsymbol{x}-\boldsymbol{\mu_f}) - (\boldsymbol{x}-\boldsymbol{\mu_g})^T \boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu_g})\right] \\
&= -\frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_f)}{\det(\boldsymbol{\Sigma}_g)} - \frac{1}{2}E_f\left[(\boldsymbol{x}-\boldsymbol{\mu_f})^T \boldsymbol{\Sigma}_f^{-1}(\boldsymbol{x}-\boldsymbol{\mu_f})\right] + \frac{1}{2}E_f\left[(\boldsymbol{x}-\boldsymbol{\mu_g})^T \boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu_g})\right] \\
&= \frac{1}{2}\left[\log\frac{\det(\boldsymbol{\Sigma}_g)}{\det(\boldsymbol{\Sigma}_f)} - \operatorname{tr}\left\{\boldsymbol{\Sigma}_f^{-1}\boldsymbol{\Sigma}_f\right\} + E_f\left[(\boldsymbol{x}-\boldsymbol{\mu_g})^T \boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu_g})\right]\right]
\end{aligned} \tag{4.74}$$

By definition,

$$E_f\left[(\boldsymbol{x}-\boldsymbol{\mu_g})^T \boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu_g})\right] = (\boldsymbol{x}-\boldsymbol{\mu_g})^T \boldsymbol{\Sigma}_g^{-1}(\boldsymbol{x}-\boldsymbol{\mu_g}) + \operatorname{Tr}\left\{\boldsymbol{\Sigma}_g^{-1}\boldsymbol{\Sigma}_f\right\} \tag{4.75}$$

Given that $\mathbf{\Sigma}_f$ is a n-by-n matrix, we get:

$$\begin{aligned} \operatorname{tr}\left\{\mathbf{\Sigma}_f^{-1}\mathbf{\Sigma}_f\right\} &= \operatorname{Tr}\left\{\mathbf{I}_n\right\} \\ &= n \end{aligned} \tag{4.76}$$

Hence, the multivariate version of the Kullback-Leibler Divergence between two multivariate Gaussian distributions $g_X(\mathbf{x})$ and $f_X(\mathbf{x})$ is:

$$KLD(f||g) = \frac{1}{2}\left[\log\frac{\det(\mathbf{\Sigma}_g)}{\det(\mathbf{\Sigma}_f)} - n + (\boldsymbol{x} - \boldsymbol{\mu_g})^T\mathbf{\Sigma}_g^{-1}(\boldsymbol{x} - \boldsymbol{\mu_g}) + \operatorname{Tr}\left\{\mathbf{\Sigma}_g^{-1}\mathbf{\Sigma}_f\right\}\right] \tag{4.77}$$

4.B Deriving ΔKLD for multivariate Gaussian distributions

In this appendix, we derive ΔKLD for the case of multivariate Gaussian pdfs of position estimates. The multivariate Gaussian pdf with n-dimensions is defined as:

$$f_X(\mathbf{x}) = ([2\pi]^n \det(\boldsymbol{\Sigma}_f))^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_f)^T \boldsymbol{\Sigma}_f^{-1}(\mathbf{x}-\boldsymbol{\mu}_f)\right] \tag{4.78}$$

For our position estimation problem, each of the two nodes to locate is in 2-dimensional space. Hence, the multivariate pdfs considered are in 4-dimensions. The pdfs for $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are:

$$f_{XCOL}(\mathbf{x}) = \left([2\pi]^4 \det(\boldsymbol{\Sigma_{COL}})\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu_{COL}})^T \boldsymbol{\Sigma_{COL}}^{\mathbf{-1}}(\mathbf{x}-\boldsymbol{\mu_{COL}})\right] \tag{4.79}$$

$$f_{XSTD}(\mathbf{x}) = \left([2\pi]^4 \det(\boldsymbol{\Sigma_{STD}})\right)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu_{STD}})^T \boldsymbol{\Sigma_{STD}}^{\mathbf{-1}}(\mathbf{x}-\boldsymbol{\mu_{STD}})\right] \tag{4.80}$$

4.B.1 Kullback-Leibler Divergence between best and standalone estimators

Using the multivariate definition of the KLD from Appendix 4.A, the KLD between the pdf $f_{XSTD}(\mathbf{x})$ of position estimates from the standalone estimator and the pdf $f_{Xref}(\mathbf{x}) = f_{XCRB}(\mathbf{x})$ of position estimates from the best estimator available is:

$$\begin{aligned} KLD(f_{XCRB}||f_{XSTD}) &= \frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{CRB})} - \frac{n}{2} \\ &+ \frac{(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{STD})^T \boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{STD}) + \operatorname{tr}\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{CRB}\right\}}{2} \end{aligned} \tag{4.81}$$

4.B.2 Kullback-Leibler Divergence between best and collaborative estimators

Similarly, the KLD between the pdf $f_{XCOL}(\mathbf{x})$ of position estimates from the collaborative estimator and the pdf $f_{Xref}(\mathbf{x}) = f_{XCRB}(\mathbf{x})$ of position estimates from the best estimator available is:

$$\begin{aligned} KLD(f_{XCRB}||f_{XCOL}) &= \frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_{COL})}{\det(\boldsymbol{\Sigma}_{CRB})} - \frac{n}{2} \\ &+ \frac{(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{COL})^T \Sigma_{COL}^{-1}(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{COL}) + \operatorname{tr}\left\{\boldsymbol{\Sigma}_{COL}^{-1}\boldsymbol{\Sigma}_{CRB}\right\}}{2} \end{aligned} \tag{4.82}$$

4.B.3 Difference between KLDs: Estimation of ΔKLD

The reduction in inaccuracy (or improvement in accuracy) due to collaboration is ΔKLD. It is defined as:

$$\Delta KLD = KLD(f_{XCRB}||f_{XSTD}) - KLD(f_{XCRB}||f_{XCOL}) \tag{4.83}$$

$$\begin{aligned}\Delta KLD = \frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} &+ \frac{(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{STD}) + \mathrm{tr}\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{CRB}\right\}}{2} \\ &- \frac{(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{COL})^T\boldsymbol{\Sigma}_{COL}^{-1}(\boldsymbol{\mu}_{CRB}-\boldsymbol{\mu}_{COL}) + \mathrm{tr}\left\{\boldsymbol{\Sigma}_{COL}^{-1}\boldsymbol{\Sigma}_{CRB}\right\}}{2}\end{aligned} \tag{4.84}$$

4.C Scalar ΔKLD derivation for theoretical analysis

In this appendix, we derive the scalar ΔKLD for our theoretical analysis of collaborative positioning in section 4.4.

Recall ΔKLD:

$$\Delta KLD = KLD(f_{Xref}||f_{XSTD}) - KLD(f_{Xref}||f_{XCOL}) \tag{4.85}$$

$$\begin{aligned}\Delta KLD = \frac{1}{2}\log\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})} &+ \frac{(\boldsymbol{\mu}_{ref}-\boldsymbol{\mu}_{STD})^T\mathbf{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{ref}-\boldsymbol{\mu}_{STD}) + \mathrm{tr}\left\{\mathbf{\Sigma}_{STD}^{-1}\mathbf{\Sigma}_{ref}\right\}}{2} \\ &- \frac{(\boldsymbol{\mu}_{ref}-\boldsymbol{\mu}_{COL})^T\mathbf{\Sigma}_{COL}^{-1}(\boldsymbol{\mu}_{ref}-\boldsymbol{\mu}_{COL}) + \mathrm{tr}\left\{\mathbf{\Sigma}_{COL}^{-1}\mathbf{\Sigma}_{ref}\right\}}{2}\end{aligned} \tag{4.86}$$

The multivariate version of ΔKLD being very complex to analyze, we derive its scalar version. We assume that we want to estimate scalars x_A and x_B in standalone and collaborative context. Based on our assumptions defined in section 4.4, the covariance matrices yield:

$$\mathbf{\Sigma_{STD}} = \begin{bmatrix} \sigma_A^2 & 0 \\ 0 & \sigma_B^2 \end{bmatrix} \tag{4.87}$$

$$\mathbf{\Sigma_{COL}} = \begin{bmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{bmatrix} \tag{4.88}$$

$$\mathbf{\Sigma_{T}} = \begin{bmatrix} \sigma_{AT}^2 & 0 \\ 0 & \sigma_{BT}^2 \end{bmatrix} \tag{4.89}$$

where ρ is a scalar that we interpret later in section 4.5.2. For the sake of theoretical analysis, we consider that the reference variances σ_{AT}^2 and σ_{BT}^2 represent the collaborative CRB for estimating scalars x_A and x_B. Thus, having assumed unbiased estimators, we get: $\sigma_{AT}^2 \leq \sigma_A^2$ and $\sigma_{BT}^2 \leq \sigma_B^2$. It is also worth noting that in this theoretical framework, we assume that the diagonal terms of both covariance matrices are nearly equal.

Hence, ΔKLD_{scalar} is equivalent to:

$$\Delta KLD_{scalar} : \tau + \upsilon + \omega \geq 0 \tag{4.90}$$

where

$$\tau = \frac{1}{2}\log\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})} = -\log\sqrt{1-\rho^2} \tag{4.91}$$

$$\upsilon = \frac{\mathrm{tr}\left\{\mathbf{\Sigma}_{STD}^{-1}\mathbf{\Sigma}_{ref}\right\} - \mathrm{tr}\left\{\mathbf{\Sigma}_{COL}^{-1}\mathbf{\Sigma}_{ref}\right\}}{2} = -\frac{\rho^2}{2(1-\rho^2)}\left[\frac{\sigma_{AT}^2}{\sigma_A^2} + \frac{\sigma_{BT}^2}{\sigma_B^2}\right] \tag{4.92}$$

$$\begin{aligned}\omega = \frac{(\boldsymbol{\mu}_{ref} - \boldsymbol{\mu}_{STD})^T\mathbf{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{CRB} - \boldsymbol{\mu}_{STD}) - (\boldsymbol{\mu}_{CRB} - \boldsymbol{\mu}_{COL})^T\mathbf{\Sigma}_{COL}^{-1}(\boldsymbol{\mu}_{CRB} - \boldsymbol{\mu}_{COL})}{2} &= \\ \frac{1}{2}\left[\frac{(\mu_{AT} - \mu_{As})^2}{\sigma_A^2} + \frac{(\mu_{BT} - \mu_{Bs})^2}{\sigma_B^2}\right] & \\ -\frac{1}{2(1-\rho^2)}\left[\frac{(\mu_{AT} - \mu_{Ac})^2}{\sigma_A^2} + \frac{(\mu_{BT} - \mu_{Bc})^2}{\sigma_B^2}\right] + \frac{\rho}{1-\rho^2}\frac{(\mu_{BT} - \mu_{Bc})(\mu_{AT} - \mu_{Ac})}{\sigma_A\sigma_B} &\end{aligned} \tag{4.93}$$

where variables with endices ending in $*_{AT}$ refer to data from the best estimator. Variables with indices ending in $*_A$ and $*_{As}$ refer to data from the standalone estimator. Variables with indices ending in $*_{Ac}$ refers to data from the collaborative estimator. If both standalone and collaborative estimators are unbiased, the last equation above is equal to zero.

Hence, the scalar ΔKLD is:

$$\Delta KLD_{scalar} = -\log\sqrt{1-\rho^2} - \frac{\rho^2\chi}{2(1-\rho^2)} \tag{4.94}$$

where: $-1 \leq \rho \leq 1$ and $0 \leq \chi \leq 2$, with:

$$\chi = \left[\frac{\sigma_{AT}^2}{\sigma_A^2} + \frac{\sigma_{BT}^2}{\sigma_B^2}\right] \tag{4.95}$$

4.D Mutual Information between multivariate and between scalar Gaussian distributions

In this appendix, we derive the mutual information between two multivariate and two scalar variables with Gaussian distributions.

We consider the multivariate Gaussian pdfs $f_{XSTD}(\mathbf{x})$ and $f_{XCOL}(\mathbf{x})$ of respective covariance matrices:

$$\boldsymbol{\Sigma_{STD}} = \begin{bmatrix} \boldsymbol{\Sigma}_A & 0 \\ 0 & \boldsymbol{\Sigma}_B \end{bmatrix} \tag{4.96}$$

$$\boldsymbol{\Sigma_{COL}} = \begin{bmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{AB}^T & \boldsymbol{\Sigma}_B \end{bmatrix} \tag{4.97}$$

When the block-diagonal of $\boldsymbol{\Sigma}_{COL}$ and $\boldsymbol{\Sigma}_{STD}$ are nearly the same, a special case occurs when computing the reduction in uncertainty due to collaboration. The reduction in entropy when the block diagonal matrices are the same, is equal to the mutual information between the distributions of position estimates for the nodes.

The entropy $H(f_{XSTD})$ for pdf $f_{XSTD}(\mathbf{x})$ was proven to be (Appendix 5.A):

$$H(f_{XSTD}) = \frac{1}{2}\log\left((2\pi e)^n \det\left(\boldsymbol{\Sigma}_{STD}\right)\right) \tag{4.98}$$

Where $\log = \log_e$.

After collaboration, the entropy for pdf $f_{XCOL}(\mathbf{x})$ is defined as:

$$H(f_{XCOL}) = \frac{1}{2}\log\left((2\pi e)^n \det\left(\boldsymbol{\Sigma}_{COL}\right)\right) \tag{4.99}$$

4.D.1 Multivariate Mutual Information

The mutual information (MI) or reduction in uncertainty when the block diagonal matrices are equal is:

$$\begin{aligned} MI = H(f_{XSTD}) - H(f_{XCOL}) &= \frac{1}{2}\left(\log\det\boldsymbol{\Sigma}_{STD} - \log\det\boldsymbol{\Sigma}_{COL}\right) \\ &= \frac{1}{2}\log\frac{\det\boldsymbol{\Sigma}_{STD}}{\det\boldsymbol{\Sigma}_{COL}} \end{aligned} \tag{4.100}$$

Hence, the multivariate mutual information between multivariate Gaussian pdfs is:

$$MI = \frac{1}{2}\log\frac{\det \boldsymbol{\Sigma}_{STD}}{\det \boldsymbol{\Sigma}_{COL}} \tag{4.101}$$

4.D.2 Scalar Mutual Information

In the scalar case, we have:

$$\boldsymbol{\Sigma_{STD}} = \begin{bmatrix} \sigma_A^2 & 0 \\ 0 & \sigma_B^2 \end{bmatrix} \tag{4.102}$$

$$\boldsymbol{\Sigma_{COL}} = \begin{bmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{bmatrix} \tag{4.103}$$

where ρ is a scalar less than unity. Hence, the scalar mutual information yields:

$$\begin{aligned} MI_{scalar} &= \frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} \\ &= \frac{1}{2}\log\frac{\sigma_A^2\sigma_B^2}{\sigma_A^2\sigma_B^2 - (\rho^2\sigma_A^2\sigma_B^2)} \\ &= \frac{1}{2}\log\frac{\sigma_A^2\sigma_B^2}{\sigma_A^2\sigma_B^2(1-\rho^2))} \\ &= \frac{1}{2}\log\frac{1}{(1-\rho^2))} \\ &= -\log\sqrt{1-\rho^2} \end{aligned} \tag{4.104}$$

4.E Adjusting standalone matrix by rank update

Recall from Appendix 2.A that the definitions for $\mathbf{G}_{STD}$ and $\mathbf{G}_{COL}$ were derived in equations (2.58) and (2.57):

$$\mathbf{G}_{STD} = \begin{bmatrix} \cos(\theta_{A_0A_1}) & \sin(\theta_{A_0A_1}) & 0 & 0 \\ \cos(\theta_{A_0A_2}) & \sin(\theta_{A_0A_2}) & 0 & 0 \\ \cos(\theta_{A_0A_3}) & \sin(\theta_{A_0A_3}) & 0 & 0 \\ 0 & 0 & \cos(\theta_{B_0B_1}) & \sin(\theta_{B_0B_1}) \\ 0 & 0 & \cos(\theta_{B_0B_2}) & \sin(\theta_{B_0B_2}) \\ 0 & 0 & \cos(\theta_{B_0B_3}) & \sin(\theta_{B_0B_3}) \end{bmatrix} \tag{4.105}$$

Similarly, in the collaborative framework, where there is a collaborative link, we get:

$$\mathbf{G}_{COL} = \begin{bmatrix} \cos(\theta_{A_0A_1}) & \sin(\theta_{A_0A_1}) & 0 & 0 \\ \cos(\theta_{A_0A_2}) & \sin(\theta_{A_0A_2}) & 0 & 0 \\ \cos(\theta_{A_0A_3}) & \sin(\theta_{A_0A_3}) & 0 & 0 \\ 0 & 0 & \cos(\theta_{B_0B_1}) & \sin(\theta_{B_0B_1}) \\ 0 & 0 & \cos(\theta_{B_0B_2}) & \sin(\theta_{B_0B_2}) \\ 0 & 0 & \cos(\theta_{B_0B_3}) & \sin(\theta_{B_0B_3}) \\ \cos(\theta_{A_0B_0}) & \sin(\theta_{A_0B_0}) & -\cos(\theta_{A_0B_0}) & -\sin(\theta_{A_0B_0}) \end{bmatrix} \tag{4.106}$$

By observation, we note that $\mathbf{G}_{COL}$ and $\mathbf{G}_{STD}$ are related by a rank-1 update as follows:

$$\mathbf{G}_{COL} = \begin{bmatrix} \mathbf{G}_{STD} \\ \mathbf{b} \end{bmatrix} \tag{4.107}$$

where vector $\mathbf{b}$ is such that:

$$\mathbf{b} = [\cos(\theta_{A_0B_0}), \sin(\theta_{A_0B_0}), -\cos(\theta_{A_0B_0}), -\sin(\theta_{A_0B_0})]^T \tag{4.108}$$

This rank-1 update is generated by the acquisition of a collaborative link between nodes A and B. Hence, when nodes A and B collaborate, the acquisition of a set of new range measurements between nodes A and B modifies $\mathbf{G}_{STD}$ with a rank-1 update such that $\mathbf{G}_{STD}$ becomes $\mathbf{G}_{COL}$. Given that the covariance matrices are such that:

$$\mathbf{\Sigma}_{STD} = (\mathbf{G}_{STD}^T \mathbf{N}_{STD} \mathbf{G}_{STD})^{-1} \tag{4.109}$$

$$\mathbf{\Sigma}_{COL} = (\mathbf{G}_{COL}^T \mathbf{N}_{COL} \mathbf{G}_{COL})^{-1} \tag{4.110}$$

it can be proven (this was also shown via lengthy derivations in Appendix 2.A) that the update in $\mathbf{G}_{STD}$ results in a new $\mathbf{\Sigma}_{COL}^{-1}$ such that:

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \mathbf{R} \tag{4.111}$$

where we observe that 4x4 matrix $\mathbf{R}$ is also equal to:

$$\mathbf{R} = \frac{\mathbf{b}\mathbf{b}^T}{\sigma_{AB}^2} \tag{4.112}$$

We now introduce the Sherman-Morrison formula given an n-by-n matrix $\mathbf{K}$, and two n-by-1 vectors $\mathbf{b}$ and $\mathbf{c}$:

$$(\mathbf{K} + \mathbf{b}\mathbf{c}^T)^{-1} = \mathbf{K}^{-1} - r(\mathbf{K}^{-1}\mathbf{b}\mathbf{c}^T\mathbf{K}^{-1}) \tag{4.113}$$

where scalar r is equal to:

$$r = \frac{1}{1 + \mathbf{c}^T\mathbf{K}^{-1}\mathbf{b}} \tag{4.114}$$

This formula means that if we update a matrix $\mathbf{K}$, we can derive more easily the inverse of the updated matrix without having to go through the traditional computationally-expensive inverse derivation.

Since block matrix $\mathbf{\Sigma}_{STD}^{-1}$ is block-diagonal, deriving its inverse $\mathbf{\Sigma}_{STD}$ is trivial:

$$\mathbf{\Sigma}_{STD}^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{FASTD} & 0 \\ 0 & \mathbf{\Sigma}_{FBSTD} \end{bmatrix} \tag{4.115}$$

where:

$$\mathbf{\Sigma}_{FASTD} = \begin{bmatrix} \sum_{i=1}^{3} \frac{\cos^2(\theta_{AA_i})}{\sigma_{(AAi)}^2} & \sum_{i=1}^{3} \frac{\cos(\theta_{AA_i})\sin(\theta_{AA_i})}{\sigma_{(AAi)}^2} \\ \sum_{i=1}^{3} \frac{\cos(\theta_{AA_i})\sin(\theta_{AA_i})}{\sigma_{(AAi)}^2} & \sum_{i=1}^{3} \frac{\sin^2(\theta_{AA_i})}{\sigma_{AAi}^2} \end{bmatrix} \tag{4.116}$$

$$\mathbf{\Sigma}_{FBSTD} = \begin{bmatrix} \sum_{i=1}^{3} \frac{\cos^2(\theta_{BBi})}{\sigma_{(BBi)}^2} & \sum_{i=1}^{3} \frac{\cos(\theta_{BBi})\sin(\theta_{BBi})}{\sigma_{(BBi)}^2} \\ \sum_{i=1}^{3} \frac{\cos(\theta_{BBi})\sin(\theta_{BBi})}{\sigma_{(BBi)}^2} & \sum_{i=1}^{3} \frac{\sin^2(\theta_{BBi})}{\sigma_{BBi}^2} \end{bmatrix} \tag{4.117}$$

Hence, $\mathbf{\Sigma}_{STD}$ is simply:

$$\mathbf{\Sigma}_{STD} = \begin{bmatrix} \mathbf{\Sigma}_{FASTD}^{-1} & 0 \\ 0 & \mathbf{\Sigma}_{FBSTD}^{-1} \end{bmatrix} \tag{4.118}$$

However, deriving $\mathbf{\Sigma}_{COL}$ from $\mathbf{\Sigma}_{COL}^{-1}$ can be very computationally intensive. Hence, using the Sherman-Morrison formula can be beneficial in our research context.

Therefore, using the Sherman-Morrison formula with equation (4.111), we get:

$$\begin{aligned} \mathbf{\Sigma}_{COL} &= (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{R})^{-1} \\ &= \mathbf{\Sigma}_{STD} - \frac{r}{\sigma_{AB}^2}(\mathbf{\Sigma}_{STD}\mathbf{b}\mathbf{b}^T\mathbf{\Sigma}_{STD}) \\ &= \mathbf{\Sigma}_{STD} - \frac{r}{\sigma_{AB}^2}(\mathbf{\Sigma}_{STD}\mathbf{b}\mathbf{b}^T\mathbf{\Sigma}_{STD}) \end{aligned} \tag{4.119}$$

with scalar r equal to:

$$r = \frac{1}{1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b}} \tag{4.120}$$

To summarize, using the Torrieri's linear estimator, when we collaborate between two nodes, we update $\mathbf{G}_{STD}$ with a rank-1 update. $\mathbf{\Sigma}_{STD}^{-1}$ is then updated by adding matrix $\mathbf{R}$ to yield a new matrix $\mathbf{\Sigma}_{COL}^{-1}$. The Sherman-Morrison formula allows us to predict $\mathbf{\Sigma}_{COL}$ using $\mathbf{\Sigma}_{STD}$ and $\mathbf{R}$ for the case of collaboration between two nodes.

From equation (4.119), using the Sherman-Morrison formula allows us to generate a simpler closed-form expression relating $\mathbf{\Sigma}_{COL}$ to $\mathbf{\Sigma}_{STD}$ and $\mathbf{R}$. This result is useful because it allows us to derive an expression for $\mathbf{\Sigma}_{COL}$ without having to go through the computationally-expensive task to derive the inverse of $\mathbf{\Sigma}_{COL}^{-1}$. The cost saving of this formula is especially significant in real world applications if we consider collaboration between more than two collaborating nodes.

To illustrate this point, we now consider the case where three nodes A, B and D need to be located using standalone and collaborative positioning. If $\mathbf{0}_{m,n}$ is an m-by-n null matrix, the standalone matrix $\mathbf{G}_{STD}$ becomes a 9x6 matrix as follows:

$$\mathbf{G}_{STD} = \begin{bmatrix} \mathbf{G}_{ASTD} & \mathbf{0}_{3,2} & \mathbf{0}_{3,2} \\ \mathbf{0}_{3,2} & \mathbf{G}_{BSTD} & \mathbf{0}_{3,2} \\ \mathbf{0}_{3,2} & \mathbf{0}_{3,2} & \mathbf{G}_{DSTD} \end{bmatrix} \tag{4.121}$$

where $\mathbf{G}_{ASTD}$ is a 3x2 matrix equal to:

$$\mathbf{G}_{ASTD} = \begin{bmatrix} \cos(\theta_{A_0A_1}) & \sin(\theta_{A_0A_1}) \\ \cos(\theta_{A_0A_2}) & \sin(\theta_{A_0A_2}) \\ \cos(\theta_{A_0A_3}) & \sin(\theta_{A_0A_3}) \end{bmatrix} \tag{4.122}$$

We define C_n^k as the number of k-combinations in a set of n elements:

$$C_n^k = \frac{n!}{k!(n-k)!} \tag{4.123}$$

If C_3^2 is the number of 2-combinations in a set of 3 elements, in the collaborative problem, there can be up to $C_3^2 = 3$ collaborative links between nodes A, B and D. Hence, $\mathbf{G}_{STD}$ can be updated by a rank-n matrix where $1 \leq n \leq 3$. $\mathbf{\Sigma}_{COL}^{-1}$ then takes the form:

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \mathbf{R} \tag{4.124}$$

where the new generalized matrix $\mathbf{R}$ is of the form

$$\mathbf{R} = \mathbf{U}\mathbf{V}\mathbf{U}^T \tag{4.125}$$

In a 2-dimensional Euclidean space, if there are 3 collaborative links between nodes A, B and D, matrix $\mathbf{U}$ is a 6x3 matrix and V is a 3x3. More generally, in an h-dimensional Euclidean space, if there are t collaborating nodes, there can be up to C_t^2 collaborative links between any two nodes.
This results in matrix $\mathbf{R} = \mathbf{U}\mathbf{V}\mathbf{U}^T$ where matrix $\mathbf{U}$ is a ht-by-C_t^2 matrix and V is a C_t^2-by-C_t^2 matrix. Matrix U contains information on the geometric configuration between the collaborating nodes and V contains information about the variance of the range measurements for all collaborative links.

In the case where the collaborative positioning problem concerns three or more nodes, the generalized form of the Sherman-Morrison formula is used. It is called the Woodbury identity. Hence, using the Woodbury identity, we get:

$$\begin{aligned}\mathbf{\Sigma}_{COL} &= (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{R})^{-1} \\ &= (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{U}\mathbf{V}\mathbf{U}^T)^{-1} \\ &= \mathbf{\Sigma}_{STD} - \mathbf{\Sigma}_{STD} - \mathbf{\Sigma}_{STD}\mathbf{U}(\mathbf{V}^{-1} + \mathbf{U}^T\mathbf{\Sigma}_{STD}U)^{-1}U^T\mathbf{\Sigma}_{STD}\end{aligned} \tag{4.126}$$

To conclude, the importance of these results is that, when using a linear estimator, with the Woodbury identity, we are able to predict the covariance matrix of the collaborative estimator given the standalone covariance matrix, and its update due to the presence of collaborative links. In cases where the standalone and collaborative estimators have the same bias, deriving this new covariance matrix ahead of time, can provide useful information on the impact that collaboration has on standalone estimation, at a significantly reduced computation cost.

4.F ΔKLD vs. Generalized GDOP

Geometry contributes significantly to the improvement in accuracy due to collaboration. In this appendix, we derive the closed-form expression for ΔKLD when the standalone and collaborative estimators are unbiased and have a multivariate-Gaussian distribution. We also analyze how the Generalized GDOP (GGDOP) and ΔKLD are related.

Let $f_{XSTD}(\mathbf{x})$ and $f_{XCOL}(\mathbf{x})$ represent the pdf of position estimates from the standalone and collaborative estimators, respectively. Let $f_{Xref}(\mathbf{x})$ represent the pdf of the most accurate estimator. All pdfs are assumed Gaussian and multivariate.

For unbiased estimators, recall ΔKLD:

$$\Delta KLD = \frac{1}{2}\log\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})} + \frac{\mathrm{tr}\left\{\mathbf{\Sigma}_{STD}^{-1}\mathbf{\Sigma}_{ref}\right\}}{2} - \frac{\mathrm{tr}\left\{\mathbf{\Sigma}_{COL}^{-1}\mathbf{\Sigma}_{ref}\right\}}{2} \tag{4.127}$$

with:

$$\mathbf{\Sigma}_{STD}^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{FASTD} & 0 \\ 0 & \mathbf{\Sigma}_{FBSTD} \end{bmatrix} \tag{4.128}$$

where:

$$\mathbf{\Sigma}_{FASTD} = \begin{bmatrix} \sum_{i=1}^{3}\frac{\cos^2(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} & \sum_{i=1}^{3}\frac{\cos(\theta_{A_0A_i})\sin(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} \\ \sum_{i=1}^{3}\frac{\cos(\theta_{A_0A_i})\sin(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} & \sum_{i=1}^{3}\frac{\sin^2(\theta_{A_0A_i})}{\sigma^2_{A_0A_i}} \end{bmatrix} \tag{4.129}$$

$$\mathbf{\Sigma}_{FBSTD} = \begin{bmatrix} \sum_{i=1}^{3}\frac{\cos^2(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} & \sum_{i=1}^{3}\frac{\cos(\theta_{B_0B_i})\sin(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} \\ \sum_{i=1}^{3}\frac{\cos(\theta_{B_0B_i})\sin(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} & \sum_{i=1}^{3}\frac{\sin^2(\theta_{B_0B_i})}{\sigma^2_{B_0B_i}} \end{bmatrix} \tag{4.130}$$

where $\sigma^2_{B_0B_i}$ is the variance of the noise on the observed range measurements between node B and its i^{th} anchor. $\sigma^2_{A_0A_i}$ is the variance of the noise on the observed range measurements between node A and its i^{th} anchor. Angle $\theta_{B_0B_i}$ is the angle between node B and its i^{th} anchor. $\theta_{A_0A_i}$ is the angle between node A and its i^{th} anchor. $\sigma^2_{B_0B_i}$ is the variance of the noise on the observed range measurements between node B and its i^{th} anchor. $\sigma^2_{A_0A_i}$ is the variance of the noise on the observed range measurements between node A and its i^{th} anchor.

We showed in Appendix 2.A that:

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \mathbf{R} \tag{4.131}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{P} & -\mathbf{P} \\ -\mathbf{P} & \mathbf{P} \end{bmatrix} \tag{4.132}$$

where $\mathbf{P}$ is:

$$\mathbf{P} = \begin{bmatrix} \frac{\cos^2(\theta_{AB})}{\sigma^2_{AB}} & \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} \\ \frac{\cos(\theta_{AB})\sin(\theta_{AB})}{\sigma^2_{AB}} & \frac{\sin^2\theta_{AB}}{\sigma^2_{AB}} \end{bmatrix} \tag{4.133}$$

with σ^2_{AB} being the variance of the noise on the observed range measurements between nodes B and A, and θ_{AB} being the angle between nodes A and B.

Thus, ΔKLD yields:

$$\Delta KLD = \frac{1}{2}\log\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})} - \frac{\mathrm{Tr}\left(\mathbf{R}\mathbf{\Sigma}_{ref}\right)}{2} \tag{4.134}$$

The closed form expressions for $\left(\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})}\right)$ and $\mathrm{Tr}\left\{\mathbf{R}\mathbf{\Sigma}_{ref}\right\}$ in terms of nodes A, B are presented.

Deriving $\left(\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})}\right)$

$$\begin{aligned}\left(\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})}\right) &= \det(\mathbf{\Sigma}_{STD})\det(\mathbf{\Sigma}^{-1}_{COL})\\ &= \det(\mathbf{\Sigma}_{STD})\det(\mathbf{\Sigma}^{-1}_{STD}+\mathbf{R})\\ &= \det(\mathbf{I}_4+\mathbf{\Sigma}_{STD}\mathbf{R})\end{aligned} \tag{4.135}$$

where $\mathbf{\Sigma}_{STD}\mathbf{R}$ is a 4x4 matrix and $\mathbf{I}_4$ is the 4x4 identity matrix.

By definition, for a 4x4 matrix $\mathbf{M}$, we have:

$$\begin{gathered}\det(\mathbf{I}_4+\mathbf{M}) = 1+\det(\mathbf{M})+\mathrm{Tr}(\mathbf{M})+\frac{1}{2}[\mathrm{Tr}(\mathbf{M})]^2-\frac{1}{2}\mathrm{Tr}(\mathbf{M}^2)+\\ \frac{1}{6}[\mathrm{Tr}(\mathbf{M})]^3-\frac{1}{2}\mathrm{Tr}(\mathbf{M})\mathrm{Tr}(\mathbf{M}^2)+\frac{1}{3}\mathrm{Tr}(\mathbf{M}^3)\end{gathered} \tag{4.136}$$

For the case where $\mathbf{M} = \mathbf{\Sigma}_{STD}\mathbf{R}$, it can be proven that:

$$\begin{aligned}Tr(\mathbf{M}) &= m_{A1}+m_{A2}+m_{B1}+m_{B2}\\ [Tr(\mathbf{M})]^2 &= (m_{A1}+m_{A2}+m_{B1}+m_{B2})^2\\ [Tr(\mathbf{M})]^3 &= (m_{A1}+m_{A2}+m_{B1}+m_{B2})^3\\ Tr(\mathbf{M}^2) &= (m^2_{A1}+m^2_{A2}+m^2_{B1}+m^2_{B2})\\ Tr(\mathbf{M}^3) &= (m^3_{A1}+m^3_{A2}+m^3_{B1}+m^3_{B2})\end{aligned} \tag{4.137}$$

Where m_{A1}, m_{A2}, m_{B1}, m_{B2}, represent the main diagonal elements of matrix $\mathbf{M}$. Therefore, we get:

$$\left(\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})}\right) = \det(\mathbf{I}_4+\mathbf{\Sigma}_{STD}\mathbf{R}) = \gamma \tag{4.138}$$

where γ is:

$$\begin{aligned}\gamma = (1+m_{A1}+m_{A2})(m_{B1}.m_{B2})+&\\ (1+m_{B1}+m_{B2})(m_{A1}.m_{A2})+&\\ (1+m_{A1}+m_{A2})(1+m_{B1}+m_{B2})&\end{aligned} \tag{4.139}$$

γ is always greater than unity because the collaborative determinant was proven to be smaller than the standalone determinant under our assumptions.

Deriving tr $\{\mathbf{R\Sigma}_{ref}\}$

In this subsection, we derive the closed form expression for tr $\{\mathbf{R\Sigma}_{ref}\}$.

Given w anchors of known locations used to locate node A, and w anchors of known locations used to locate node B, the angle θ_{AAi} between points A and its i^{th} anchor is:

$$\theta_{AAi} = \tan^{-1}\frac{(y_i - y_A)}{(x_i - x_A)} \tag{4.140}$$

θ_{AAi} can be obtained in systems with angle-of-arrival information available, or using an initial estimate for A, namely $\mathbf{x}_{A_0}$ per Appendix 2.A. Thus, the Generalized Geometric Dilution of Precision (GGDOP) for point A is defined as:

$$GGDOP_A = \left[\frac{\sum_{i=1}^{w}\sum_{(j=1,j\geq i)}^{w}\frac{\sin^2(\theta_{AAi}-\theta_{AAj})}{\sigma_i^2\sigma_j^2}}{(\sum_{i=1}^{w}1/\sigma_i^2)^2}\right]^{-1} \tag{4.141}$$

After lengthy derivations, the closed form expression for tr $\{\mathbf{R\Sigma}_{ref}\}$ can be derived to be:

$$\mathrm{Tr}\{\mathbf{R\Sigma}_{ref}\} = TrC_1 + TrC_2 \tag{4.142}$$

$$\begin{aligned} TrC_2 = {} & \frac{1}{DetC_2}\left(\frac{sin^2(\theta_{A_0B_0})}{\sigma^2_{A_0B_0}}\right)\left(-bk_2 + \sum_{i=1}^{3}\frac{cos^2(\theta_{BBi})}{\sigma^2_{BBi}} + \frac{cos^2(\theta_{AB})}{\sigma^2_{AB}}\right) \\ & + \frac{1}{DetC_2}\left(\frac{cos^2(A_0B_0)}{\sigma^2_{A_0B_0}}\right)\left(-bk_2 + \sum_{i=1}^{3}\frac{sin^2(\theta_{BBi})}{\sigma^2_{BBi}} + \frac{sin^2(\theta_{AB})}{\sigma^2_{AB}}\right) \\ & + \frac{1}{DetC_2}\left(\frac{2cos(A_0B_0)sin(A_0B_0)}{\sigma^2_{A_0B_0}}\right)\left(\frac{bk_2}{2} - \sum_{i=1}^{3}\frac{cos(\theta_{BBi})sin(\theta_{BBi})}{\sigma^2_{BBi}} - \frac{sin(\theta_{AB})cos(\theta_{AB})}{\sigma^2_{AB}}\right) \end{aligned} \tag{4.143}$$

where:

$$bk_2 = \sum_{i=1}^{w}\frac{\frac{sin^2(\theta_{AAi}-\theta_{AB})}{\sigma^2_{AB}\sigma^2_{AAi}}}{\sigma^2_{AB}\det\mathbf{\Sigma}_{FACOL}} \tag{4.144}$$

If $GGDOP_{ACOL}$ is the collaborative GGDOP for node A, then:

$$\begin{aligned}\det \mathbf{\Sigma}_{FACOL} &= \frac{\sin^2(\theta_{AA_1}-\theta_{AA_2})}{\sigma^2_{AA_1}\sigma^2_{AA_2}} + \frac{\sin^2(\theta_{AA_1}-\theta_{AA_3})}{\sigma^2_{AA_1}\sigma^2_{AA_3}} + \frac{\sin^2(\theta_{AA_2}-\theta_{AA_3})}{\sigma^2_{AA_2}\sigma^2_{AA_3}} \\ &+ \frac{\sin^2(\theta_{AA_1}-\theta_{AB})}{\sigma^2_{AA_1}\sigma^2_{AB}} + \frac{\sin^2(\theta_{AA_2}-\theta_{AB})}{\sigma^2_{AA_2}\sigma^2_{AB}} + \frac{\sin^2(\theta_{AA_3}-\theta_{AB})}{\sigma^2_{AA_3}\sigma^2_{AB}} \\ &= GGDOP^{-1}_{ACOL}\cdot \left(\frac{1}{\sigma^2_{AB}} + \sum_{j=1}^{w}\frac{1}{\sigma^2_{AAj}}\right)^2\end{aligned} \tag{4.145}$$

If $GGDOP_{BCOL}$ is the collaborative GGDOP for node B, then:

$$\begin{aligned}DetC_2 = GGDOP^{-1}_{BCOL}\cdot &\left(\frac{1}{\sigma^2_{AB}} + \sum_{j=1}^{w}\frac{1}{\sigma^2_{BBj}}\right)^2 + \frac{3}{4}bk_2^2 - bk_2\left(\frac{1}{\sigma^2_{AB}}\sum_{i=1}^{i=3}\frac{1}{\sigma^2_{BBi}}\right) \\ &+bk_2\left(\frac{cos(\theta_{AB})sin(\theta_{AB})}{\sigma^2_{AB}}\sum_{i=1}^{3}\frac{cos(\theta_{BBi})sin(\theta_{BBi})}{\sigma^2_{BBi}}\right)\end{aligned} \tag{4.146}$$

Similarly for TrC_1, we get:

$$\begin{aligned}TrC_1 = &\frac{1}{DetC_1}\left(\frac{sin^2(A_0B_0)}{\sigma^2_{A_0B_0}}\right)\left(-bk_1 + \sum_{i=1}^{3}\frac{cos^2(\theta_{AAi})}{\sigma^2_{AAi}} + \frac{cos^2(\theta_{AB})}{\sigma^2_{AB}}\right) \\ &+\frac{1}{DetC_1}\left(\frac{cos^2(A_0B_0)}{\sigma^2_{A_0B_0}}\right)\left(-bk_1 + \sum_{i=1}^{3}\frac{sin^2(\theta_{AAi})}{\sigma^2_{AAi}} + \frac{sin^2(\theta_{AB})}{\sigma^2_{AB}}\right) \\ &+\frac{1}{DetC_1}\left(\frac{2cos(A_0B_0)sin(A_0B_0)}{\sigma^2_{A_0B_0}}\right)\left(\frac{bk_1}{2} - \sum_{i=1}^{3}\frac{cos(\theta_{AAi})sin(\theta_{AAi})}{\sigma^2_{AAi}} - \frac{sin(\theta_{AB})cos(\theta_{AB})}{\sigma^2_{AB}}\right)\end{aligned} \tag{4.147}$$

$$bk_1 = \sum_{i=1}^{w}\frac{\frac{sin^2(\theta_{BBi}-\theta_{AB})}{\sigma^2_{AB}\sigma^2_{BBi}}}{\sigma^2_{AB}\det \mathbf{\Sigma}_{FBCOL}} \tag{4.148}$$

$$\det \mathbf{\Sigma}_{FBCOL} = GGDOP^{-1}_{BCOL}\cdot\left(\frac{1}{\sigma^2_{AB}} + \sum_{j=1}^{w}\frac{1}{\sigma^2_{BBj}}\right)^2 \tag{4.149}$$

$$\begin{aligned}DetC_1 = GGDOP^{-1}_{ACOL}\cdot &\left(\frac{1}{\sigma^2_{AB}} + \sum_{j=1}^{w}\frac{1}{\sigma^2_{BBj}}\right)^2 + \frac{3}{4}bk_1^2 - bk_1\left(\frac{1}{\sigma^2_{AB}}\sum_{i=1}^{i=3}\frac{1}{\sigma^2_{AAi}}\right) \\ &+bk_1\left(\frac{cos(\theta_{AB})sin(\theta_{AB})}{\sigma^2_{AB}}\sum_{i=1}^{3}\frac{cos(\theta_{AAi})sin(AA)}{\sigma^2_{AAi}}\right)\end{aligned} \tag{4.150}$$

Since the logDet term in the expression of ΔKLD is a positive value, the sign of ΔKLD is mainly dependent on the magnitude of $\mathrm{tr}\{\mathbf{R\Sigma}_{ref}\}$, which is the sum of TrC_1 and TrC_2

derived above. In each expression for TrC_1 and TrC_2, indices A_0 and B_0 each refer to $\mathbf{x}_{A_0}$, and $\mathbf{x}_{B_0}$, respectively and describe the first guess in the one-shot estimator. Indices A and B refer to the true position vectors of nodes A and B: $\mathbf{x}_A$ and $\mathbf{x}_B$, respectively.

Each of the terms TrC_1 and TrC_2 are also functions of the collaborative Generalized GDOP (GGDOP) for both nodes A and B w.r.t their anchors. This confirms that the accuracy of collaborative estimation is dependent on the Generalized GDOP. We note that both $GGDOP_{ACOL}$ and $GGDOP_{BCOL}$ are proportional to the trace term. This implies that a small collaborative Generalized GDOP tends to minimize the loss in ΔKLD. Hence, a small collaborative GGDOP tends to maximize ΔKLD. Therefore, a valid criteria to consider for maximizing ΔKLD is to minimize the collaborative Generalized GDOP.

4.G ΔKLD Model inaccuracies in case of asymmetric distributions

In our theoretical analysis, we had assumed the case where all pdfs considered in ΔKLD were Gaussian. In this appendix, we consider the case of asymmetric distributions of position estimates. Specifically, we focus on skew-normal pdfs and derive errors in the estimation of ΔKLD if the pdfs used in its computation are non-Gaussian.

By definition, the cross-entropy between the pdfs of two k-dimensional vectors such that: $f_{X1}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi_1}, \boldsymbol{\Omega_1}, \boldsymbol{\eta_1})$ and $f_{X2}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi_2}, \boldsymbol{\Omega_2}, \boldsymbol{\eta_2})$ is [46]:

$$H(f_{X1}||f_{X2}) = H(f_{X01}||f_{X02}) + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi_1} - \boldsymbol{\xi_2})^T \boldsymbol{\Omega}_2^{-1}\boldsymbol{\delta}_1 - E[\log\{2\Phi(W_{2-1})\}] \tag{4.151}$$

where $H(f_{X01}||f_{X02})$ is the cross-entropy between the pdfs of two k-variate Gaussian distributions $f_{X01}(\mathbf{x})$ and $f_{X02}(\mathbf{x})$,

$$W_{i-j} \sim SN_1\left(\eta_i^T(\xi_j - \xi_i), \eta_i^T\Omega_j\eta_i, \frac{\eta_i^T\delta_j}{\sqrt{\eta_i^T\Omega_j\eta_i - (\eta_i^T\delta_j)^2}}\right) \tag{4.152}$$

with

$$\delta_j = \frac{\Omega_j\eta_j}{\sqrt{1+\eta_j^T\Omega_j\eta_j}} \tag{4.153}$$

$$W_{1-1} \sim SN_1\left(0, \eta_1^T\Omega_1\eta_1, (\eta_1^T\Omega_1\eta_1)^{1/2}\right) \tag{4.154}$$

$$W_{1-2} \sim SN_1\left(\eta_1^T(\xi_2 - \xi_1), \eta_1^T\Omega_2\eta_1, \frac{\eta_1^T\delta_2}{\sqrt{\eta_1^T\Omega_2\eta_1 - (\eta_1^T\delta_2)^2}}\right) \tag{4.155}$$

$$W_{2-1} \sim SN_1\left(\eta_2^T(\xi_1 - \xi_2), \eta_2^T\Omega_1\eta_2, \frac{\eta_2^T\delta_1}{\sqrt{\eta_2^T\Omega_1\eta_2 - (\eta_2^T\delta_1)^2}}\right) \tag{4.156}$$

$$W_{2-2} \sim SN_1\left(0, \eta_2^T\Omega_2\eta_2, (\eta_2^T\Omega_2\eta_2)^{1/2}\right) \tag{4.157}$$

Using the definition above, the cross-entropy between the pdfs of position estimates for the standalone estimator ($f_{XSTD}(\mathbf{x})$) and the reference estimator ($f_{Xref}(\mathbf{x})$) yields:

$$H(f_{Xref}||f_{XSTD}) = H(f_{X0,ref}||f_{X0,STD}) + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi_{ref}} - \boldsymbol{\xi_{STD}})^T \boldsymbol{\Omega}_{STD}^{-1}\boldsymbol{\delta}_{ref} - E[\log\{2\Phi(W_{STD-ref})\}] \tag{4.158}$$

Similarly, the cross-entropy between the pdfs of position estimates for the collaborative estimator ($f_{XCOL}(\mathbf{x})$) and the reference estimator ($f_{Xref}(\mathbf{x})$) yields:

$$H(f_{Xref}||f_{XCOL}) = H(f_{X0,ref}||f_{X0,COL})+\sqrt{\frac{2}{\pi}}(\boldsymbol{\xi_{ref}} - \boldsymbol{\xi_{COL}})^T\boldsymbol{\Omega}_{COL}^{-1}\boldsymbol{\delta}_{ref}-E[\log\{2\Phi(W_{COL-ref})\}] \tag{4.159}$$

where pdfs $f_{X0,ref}(\mathbf{x})$, $f_{X0,STD}(\mathbf{x})$, and $f_{X0,COL}(\mathbf{x})$ are Gaussian-distributions and:

$$\begin{aligned} H(f_{X0,ref}||f_{X0,STD}) = \frac{1}{2}\Big\{k\log(2\pi) + \log(\det(\boldsymbol{\Omega}_{STD})) + \mathrm{Tr}(\boldsymbol{\Omega}_{STD}^{-1}\boldsymbol{\Omega}_{ref}) + \\ (\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T\boldsymbol{\Omega}_{STD}^{-1}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD}) \end{aligned} \tag{4.160}$$

$$\begin{aligned} H(f_{X0,ref}||f_{X0,COL}) = \frac{1}{2}\Big\{k\log(2\pi) + \log(\det(\boldsymbol{\Omega}_{COL})) + \mathrm{Tr}(\boldsymbol{\Omega}_{COL}^{-1}\boldsymbol{\Omega}_{ref}) + \\ (\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T\boldsymbol{\Omega}_{COL}^{-1}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL}) \end{aligned} \tag{4.161}$$

From the definition of ΔKLD in equation (3.4), we have:

$$\Delta KLD = H(f_{Xref}||f_{XSTD}) - H(f_{Xref}||f_{XCOL}) \tag{4.162}$$

We define ΔKLD_0 as ΔKLD assuming all pdfs are Gaussian distributed. This is the expression of ΔKLD used in the multivariate case as derived in equation (4.77).

$$\begin{aligned} \Delta KLD_0 &= H(f_{X0,ref}||f_{X0,STD}) - H(f_{X0,ref}||f_{X0,COL}) \\ &= \frac{1}{2}\left\{\log(\frac{\det(\boldsymbol{\Omega}_{STD})}{\det(\boldsymbol{\Omega}_{COL})}) + \mathrm{Tr}\left((\boldsymbol{\Omega}_{STD}^{-1} - \boldsymbol{\Omega}_{COL}^{-1})\boldsymbol{\Omega}_{ref}\right)\right\} \\ &+ \frac{1}{2}\left\{(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T\boldsymbol{\Omega}_{STD}^{-1}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})\right\} - \frac{1}{2}\left\{(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T\boldsymbol{\Omega}_{COL}^{-1}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})\right\} \end{aligned} \tag{4.163}$$

In the next subsections, we compute different values for ΔKLD given different distributions families and compare them against ΔKLD_0.

4.G.1 ΔKLD with skew-normal distributions: $f_{XCOL}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi}_1, \boldsymbol{\Omega}_1, \eta_1)$, $f_{XSTD}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi}_2, \boldsymbol{\Omega}_2, \eta_2)$, $f_{Xref}(\mathbf{x}) \sim SN_k(\boldsymbol{\xi}_r, \boldsymbol{\Omega}_r, \eta_r)$

The expression for ΔKLD_{SN}, the value for ΔKLD assuming all pdfs are skew-normal is:

$$\begin{aligned} \Delta KLD_{SN} &= H(f_{Xref}||f_{XSTD}) - H(f_{Xref}||f_{XCOL}) \\ &= \Delta KLD_0 + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T\boldsymbol{\Omega}_{STD}^{-1}\delta_{ref} - E[\log\{2\Phi(W_{STD-ref})\}] \\ &- \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T\boldsymbol{\Omega}_{COL}^{-1}\delta_{ref} + E[\log\{2\Phi(W_{COL-ref})\}] \end{aligned} \tag{4.164}$$

Where ΔKLD_0 is ΔKLD assuming all pdfs are Gaussian distributed.

4.G.2 Errors in theoretical model of ΔKLD

In our theoretical model, we assumed that all distributions were Gaussian and derived ΔKLD_0 that can also be defined as shown on equation (4.163).

The errors in incorrectly assessing ΔKLD if we assume all pdfs are Gaussian, while they are all skew-normal is:

$$\begin{aligned}\Delta KLD_{SN} - \Delta KLD_0 = &\sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T \boldsymbol{\Omega}_{STD}^{-1}\delta_{ref} - E[\log\{2\Phi(W_{STD-ref})\}] \\ &- \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T \boldsymbol{\Omega}_{COL}^{-1}\delta_{ref} + E[\log\{2\Phi(W_{COL-ref})\}]\end{aligned} \tag{4.165}$$

If $f_{Xref}(\mathbf{x})$ is skew-normal, while $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ remain Gaussian, the error in our theoretical model due to incorrectly assuming $f_{Xref}(\mathbf{x})$ only is derived below.

Let $\Delta KLD_{SN,0}$ represent ΔKLD when the pdfs of $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are Gaussian, and $f_{Xref}(\mathbf{x})$ is skew-normal. When $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are Gaussian pdfs, from the previous equation, $\Phi(W_{COL-ref}) = 0$ and $\Phi(W_{COL-ref}) = 0$, which yields:

$$\begin{aligned}\Delta KLD_{SN,0} &= H(f_{ref}||f_{STD}) - H(f_{ref}||f_{COL}) \\ &= \Delta KLD_0 + \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T \boldsymbol{\Omega}_{STD}^{-1}\delta_{ref} - \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T \boldsymbol{\Omega}_{COL}^{-1}\delta_{ref}\end{aligned} \tag{4.166}$$

The error in incorrectly assuming $f_{Xref}(\mathbf{x})$ while $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ are assumed Gaussian is:

$$\Delta KLD_{SN,0} - \Delta KLD_0 = \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{STD})^T \boldsymbol{\Omega}_{STD}^{-1}\delta_{ref} - \sqrt{\frac{2}{\pi}}(\boldsymbol{\xi}_{ref} - \boldsymbol{\xi}_{COL})^T \boldsymbol{\Omega}_{COL}^{-1}\delta_{ref} \tag{4.167}$$

4.H Deriving $KLD(f_{XCOL}||f_{XSTD})$ for theoretical framework

In this appendix, we derive $KLD(f_{XCOL}||f_{XSTD})$ for the theoretical framework described in section 4.4.

Let $f_{XCOL}(\mathbf{x})$ represent the pdf of position estimates from the collaborative estimator, $f_{XSTD}(\mathbf{x})$ is the pdf of position estimates from the standalone estimator and $f_{Xref}(\mathbf{x})$ is the pdf of position estimates from the most accurate estimator as defined in Appendix 2.A. Recall ΔKLD:

$$\Delta KLD = KLD(f_{Xref}||f_{XSTD}) - KLD(f_{Xref}||f_{XCOL}) \tag{4.168}$$

If $f_{XCOL}(\mathbf{x}) = f_{Xref}(\mathbf{x})$,

$$\Delta KLD = KLD(f_{XCOL}||f_{XSTD}) - KLD(f_{XCOL}||f_{XCOL}) = KLD(f_{XCOL}||f_{XSTD}) \tag{4.169}$$

which implies that ΔKLD is positive, maximal and equal to $KLD(f_{XCOL}||f_{XSTD})$.

$$\text{If } f_{Xref} = f_{XCOL},\ \Delta KLD = KLD(f_{XCOL}||f_{XSTD}) \tag{4.170}$$

Using the general multivariate definition of KLD from equation (4.77), where $f_{XSTD}(\mathbf{x})$ and $f_{XCOL}(\mathbf{x})$ represent the pdfs of standalone and collaborative position estimates, $KLD(f_{XCOL}||f_{XSTD})$ is equal to:

$$\begin{gathered} KLD(f_{XCOL}||f_{XSTD})_{n=4} = \frac{1}{2}\left[\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} - n\right] \\ +\frac{1}{2}\left[(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD}) + \operatorname{tr}\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{COL}\right\}\right] \end{gathered} \tag{4.171}$$

Recalling our assumptions from our analytical framework (for the scalar analysis in section 4.4), let us consider that the standalone and collaborative position estimators have covariance matrices defined as shown below. We assume that the difference between their diagonals is infinitesimally small.

$$\boldsymbol{\Sigma_{STD}} = \begin{bmatrix} \boldsymbol{\Sigma}_A & 0 \\ 0 & \boldsymbol{\Sigma}_B \end{bmatrix} \tag{4.172}$$

$$\boldsymbol{\Sigma_{COL}} = \begin{bmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{AB}^T & \boldsymbol{\Sigma}_B \end{bmatrix} \tag{4.173}$$

4.H.1 Deriving tr $\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{COL}\right\}$

The product $\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{COL}\right\}$ yields:

$$\begin{aligned}\operatorname{tr}\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{COL}\right\} &= \operatorname{tr}\left\{\boldsymbol{\Sigma}_{A}^{-1}\boldsymbol{\Sigma}_{A}\right\} + \operatorname{tr}\left\{\boldsymbol{\Sigma}_{B}^{-1}\boldsymbol{\Sigma}_{B}\right\} \\ &= n\end{aligned} \tag{4.174}$$

Equation (4.171) can thus be rewritten as follows

$$KLD(f_{XCOL}||f_{XSTD})_{n=4} = \frac{1}{2}\left[\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} + (\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})\right] \tag{4.175}$$

This describes $KLD(f_{XCOL}||f_{XSTD})$ for our framework.

Based on the definitions of $\boldsymbol{\Sigma}_{COL}$ and $\boldsymbol{\Sigma}_{STD}$ such that they have nearly the same diagonal, the difference of LogDets is equal to the mutual information (MI) between the distributions of position estimates for nodes A and B as proven in equation (4.101).

$$\frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} = \text{Mutual Information} \tag{4.176}$$

Hence, for our theoretical framework in section 4.4, $KLD(f_{XCOL}||f_{XSTD})$ is:

$$KLD(f_{XCOL}||f_{XSTD}) = \text{MI} + \frac{1}{2}\left[(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})\right] \tag{4.177}$$

Since we assumed unbiased estimators, we get:

$$KLD(f_{XCOL}||f_{XSTD}) = \frac{1}{2}\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} = \text{Mutual Information} \tag{4.178}$$

However, in general, the expression for $KLD(f_{XCOL}||f_{XSTD})$ is:

$$\begin{aligned} KLD(f_{XCOL}||f_{XSTD})_{n=4} &= \frac{1}{2}\left[\log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} - n\right] \\ +\frac{1}{2}&\left[(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD}) + \operatorname{tr}\left\{\boldsymbol{\Sigma}_{STD}^{-1}\boldsymbol{\Sigma}_{COL}\right\}\right]\end{aligned} \tag{4.179}$$

When $KLD(f_{XCOL}||f_{XSTD})$ is a mutual information, we can conclude that the change in accuracy due to collaboration is the same for both nodes A and B. The results from this appendix prove that $KLD(f_{XCOL}||f_{XSTD})$ is equal to the mutual information for the distributions of position estimates of nodes A and B if: 1) the pdfs have the same bias, and

2) if the covariance matrices of both standalone and collaborative estimators have the same diagonal. When those conditions are not met, $KLD(f_{XCOL}||f_{XSTD})$ still represents the maximum improvement in accuracy due to collaboration (as long as $f_{Xref}(\mathbf{x}) = f_{XCOL}(\mathbf{x})$), but it is no longer a mutual information.

4.I Deriving a lower bound to $KLD(f_{XCOL}||f_{XSTD})$

Let $f_{XCOL}(\mathbf{x}) \sim N(\boldsymbol{\mu}_{COL}, \boldsymbol{\Sigma}_{COL})$ and $f_{XSTD}(\mathbf{x}) \sim N(\boldsymbol{\mu}_{STD}, \boldsymbol{\Sigma}_{STD})$ represent the pdfs of position estimates from the collaborative and standalone estimators, respectively. In this appendix, we evaluate the lower bound to $KLD(f_{XCOL}||f_{XSTD})$ using the Kullback-Leibler inequality [29] defined as:

$$\Psi^*_{STD}(\boldsymbol{\mu}_{COL}) \leq KLD(f_{XCOL}||f_{XSTD}) \tag{4.180}$$

where:

- $\boldsymbol{\mu}_{COL}$ is the first moment of $f_{XCOL}(\mathbf{x})$.
- Ψ^*_{STD} is the convex conjugate of the cumulant generating function of the standalone estimator.

4.I.1 Deriving the multivariate convex conjugate of the cumulant generating function

We consider the case of a Gaussian pdf $f_X(\mathbf{x})$ with mean vector $\boldsymbol{\mu}$, and covariance matrix $\boldsymbol{\Sigma}$. The joint moment-generating function of this Gaussian distributed function $f_X(\mathbf{x})$ is defined as [2,3]:

$$M_{\mathbf{X}}(\mathbf{t}) = E[\exp(\mathbf{t}^T\mathbf{X})] = \exp(\mathbf{t}^T\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}) \tag{4.181}$$

By definition, the cumulant generating function (CGF) $\Gamma(\mathbf{t})$ is equal to the log of the moment generating function (MGF). For a Gaussian distribution, we obtain:

$$\Gamma(\mathbf{t}) = \mathbf{t}^T\boldsymbol{\mu} + \frac{1}{2}t^T\boldsymbol{\Sigma}\mathbf{t} \tag{4.182}$$

By definition, $\Psi^*_f(\mathbf{y})$, the convex conjugate of the CGF is derived by taking the supremum of: $\mathbf{y}^T\mathbf{t} - \mathbf{t}^T\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}$.

$$\Psi^*_f(\mathbf{y}) = \sup(\mathbf{y}^T\mathbf{t} - \mathbf{t}^T\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}) \tag{4.183}$$

We now derive $\Psi^*_f(\mathbf{y})$, the convex conjugate of the CGF $\Gamma(\mathbf{t})$. Since the CGF is derived by taking the log of the MGF, it is differential and convex. Hence, the supremum can be derived as follows.

Taking the derivative of the argument of the supremum, we get:

$$\begin{aligned}\frac{\partial}{\partial\mathbf{t}}(\mathbf{y}^T\mathbf{t} - \mathbf{t}^T\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}) &= \frac{\partial}{\partial\mathbf{t}}\left((\mathbf{y}-\boldsymbol{\mu})^T\mathbf{t} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right)\\ &= (\mathbf{y}-\boldsymbol{\mu}) - \boldsymbol{\Sigma}\mathbf{t}\end{aligned} \tag{4.184}$$

Setting the derivative above to zero, and solving for $\mathbf{t}$, we get $\mathbf{t}_0$:

$$\mathbf{t}_0 = \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu}) \tag{4.185}$$

Plugging this value of $\mathbf{t}_0$ for $\mathbf{t}$ back into $\left((\mathbf{y} - \boldsymbol{\mu})^T\mathbf{t} - \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right)$, we get $\Psi_f^*(\mathbf{y})$ as shown in the following equations. We consider that since $\boldsymbol{\Sigma}$ is symmetric, $(\boldsymbol{\Sigma}^{-1})^T = \boldsymbol{\Sigma}^{-1}$.

$$\begin{aligned}\Psi_f^*(\mathbf{y}) &= \left((\mathbf{y} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu}) - \frac{1}{2}(\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu}))^T\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right)\\ &= \frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})\end{aligned} \tag{4.186}$$

4.I.2 Deriving the multivariate lower bound to $KLD(f_{XCOL}||f_{XSTD})$

Using equations (4.180) and (4.186), the convex conjugate of the cumulant generating function for the standalone pdf, evaluated at $\boldsymbol{\mu}_{COL}$ is:

$$\Psi_{STD}^*(\boldsymbol{\mu}_{COL}) = \frac{1}{2}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD}) \tag{4.187}$$

Hence, a lower bound to $KLD(f_{XCOL}||f_{XSTD})$ is:

$$\frac{1}{2}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD}) \leq KLD(f_{XCOL}||f_{XSTD}) \tag{4.188}$$

We now derive an alternate version for this lower bound. We note that the natural log of a Gaussian distribution is:

$$\log\left(([2\pi]^n \det(\boldsymbol{\Sigma}))^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right]\right) \tag{4.189}$$

which is equivalent to:

$$-\frac{1}{2}\log\left([2\pi]^n \det(\boldsymbol{\Sigma})\right) - \left[\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right] \tag{4.190}$$

Hence, the lower bound can be rewritten as follow:

$$-\log f_{XSTD}(\boldsymbol{\mu}_{COL})\sqrt{([2\pi]^n \det(\boldsymbol{\Sigma}_{STD}))} \leq KLD(f_{XCOL}||f_{XSTD}) \tag{4.191}$$

If the inverse of the covariance matrix is readily available, the following alternative for evaluating a lower bound to $KLD(f_{XCOL}||f_{XSTD})$ is:

$$\log\frac{\sqrt{\det(\boldsymbol{\Sigma}_{STD}^{-1})}}{f_{XSTD}(\boldsymbol{\mu}_{COL})\sqrt{([2\pi]^n)}} \leq KLD(f_{XCOL}||f_{XSTD}) \tag{4.192}$$

Chapter 5

Uncertainty analysis of collaboration

In Chapter 2, we showed that when it comes to assessing the performance of estimators, there is a clear distinction between *uncertainty* and *inaccuracy* and that both terms can be related. In Chapter 4, we presented an analysis of the *inaccuracy* of collaborative positioning from an information-theoretic perspective. In this chapter, we introduce an information-theoretic perspective on the *uncertainty* of collaborative positioning. More specifically, the focus of this chapter is to present an intuitive information-theoretic explanation for *why* the CRB is reduced with collaboration and address how a) the change in uncertainty due to collaboration and b) the change in inaccuracy due to collaboration relate.

In Chapter 3, we showed via simulations that collaboration reduces the positioning Cramer-Rao Bound (CRB) [1] relative to standalone. This result has been proven theoretically in the literature. Specifically, it has been proven that the Cramer-Rao Bound (CRB) of position using a set of collaborative range measurements, under certain mild connectivity conditions, is smaller than the CRB using standalone range measurements ([18], [20], [12]). Another analysis on the value of collaborative estimation was presented by [17] which addresses the fact that the CRB for collaborative positioning is also reduced in NLOS conditions when compared against the standalone CRB.

Reference [13] introduced results on Fisher Information Matrix summary statistics that could help identify if it is worth collaborating. When it comes to predicting whether it is worth collaborating or not, these metrics perform well in line-of-sight (LOS) conditions. However, they do not in non-line-of-sight (NLOS) conditions, and duly so because the CRB applies to unbiased estimators [1].

If CRB-based criteria do not directly affect accuracy due to collaboration, and if a reduced CRB due to collaboration does not always make collaborative positioning more accurate than standalone positioning, we are faced with one fundamental question: 'What does a reduced CRB due to collaboration really mean?' To address this question we need to first understand *why collaboration reduces the CRB*. This is the main problem we address in this chapter.

In this chapter, we use the Kullback-Leibler Divergence to explain *why* collaboration reduces the CRB from an information-theoretic perspective. We show how a reduced CRB due to collaboration relates to our accuracy analysis presented in chapter 4. We also derive an upper bound based on observables for the reduction in uncertainty due to collaboration.

5.1 Why is $CRB_{COL} \leq CRB_{STD}$? An information theoretic perspective

In information theory, *information gain* is also called *reduction in uncertainty*. When we collaborate, we acquire new information, or gain information. This information gain is acquired through the addition of a set of range measurements from the collaborative link. The acquisition of new range measurements is the reason why a reduction in uncertainty occurs when we collaborate. Since the Cramer-Rao Bound is a measure of localization uncertainty, the acquisition of a new set of range measurements through the collaborative link is the reason why the collaborative Cramer-Rao Bound is smaller than the standalone CRB.

The Cramer Rao Bound is usually referred to as a measure of localization uncertainty. There are two main reasons for this valid interpretation. Firstly, the CRB is derived from the Fisher Information Matrix, hence it is obtained using the set of observed range measurements and is directly affected when we acquire a new set of range measurements through collaboration. Like uncertainty as defined in Chapter 2, it is not dependent on the position estimator used, but on the set of observed range measurements. Secondly, the CRB is a measure of uncertainty because it meets Gibb's inequality; that is, it represents a lower bound to the inaccuracy (*i.e*: the variance) of any unbiased estimator, if the pdfs of position estimates have a Gaussian distribution. It is more specifically known as a measure of *localization* uncertainty because it describes the radius of curvature of the likelihood function derived from the sets of observed range measurements used to locate a node. To summarize, the fact that we acquire a new set of range measurements when we collaborate is the reason why the collaborative CRB is smaller than the standalone CRB.

We now look at the same problem from a different perspective. We prove theoretically that the reduction in the CRB is due to the acquisition of a new set of observed range measurements from the collaborative link.

In general, the pdf of an n-dimensional Gaussian distribution $f_X(\mathbf{x})$ of mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$ is defined as:

$$f_X(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\mathbf{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) \tag{5.1}$$

The entropy for a multivariate Gaussian distribution $f_X(\mathbf{x})$ was proven to be (Appendix 5.A):

$$H(f_X) = \frac{1}{2}\log\left((2\pi e)^n \det(\mathbf{\Sigma})\right) \tag{5.2}$$

Where $\log = \log_e$ ($i.e$: the natural log function)

This equation indicates that if observables are Gaussian-distributed, uncertainty or entropy is a function of the covariance matrix; more precisely, a function of the logDet of the covariance matrix. It is worth noting that entropy is not based on the mean of a distribution. Essentially, uncertainty measures the lack in confidence in a set of range measurements. One particular advantage of looking at uncertainty rather than just a variance is that for non-Gaussian distributions, uncertainty reflects better the lack in confidence than the variance. This is because the maximum entropy principle indicates that in the family of exponential functions, the Gaussian distribution has the largest entropy. Hence, looking only at the variance is a reflection of uncertainty strictly for Gaussian distributions (up to a factor). However, for non-Gaussian distributions, entropy or uncertainty is a better way to assess the lack in confidence in a set of data.

We assume that the covariance matrices of both standalone and collaborative position estimators are equal to the inverse of their respective Fisher Information Matrices. That is, we assume $\mathbf{\Sigma}_{STD} = \mathbf{FIM}_{STD}^{-1}$ and $\mathbf{\Sigma}_{COL} = \mathbf{FIM}_{COL}^{-1}$. Let $H(f_{XCOL})$ be the entropy of $f_{XCOL}(\mathbf{x})$, the pdf of position estimates from the collaborative estimator. Let $H(f_{XSTD})$ be the entropy of $f_{XSTD}(\mathbf{x})$, the pdf of position estimates from the standalone estimator. We also refer to $H(f_{XCOL})$ and $H(f_{XSTD})$ as localization uncertainties, because they describe the entropies for pdfs of position estimates.

The change in localization uncertainty due to collaboration is defined as:

$$\begin{aligned} H(f_{XSTD}) - H(f_{XCOL}) &= \log\det\mathbf{\Sigma}_{STD} - \log\det\mathbf{\Sigma}_{COL} \\ &= \log\frac{\det\mathbf{\Sigma}_{STD}}{\det\mathbf{\Sigma}_{COL}} \end{aligned} \tag{5.3}$$

We prove in Appendix 5.B that if we acquire new information through collaboration, and the covariance matrices $\mathbf{\Sigma}_{STD}$ and $\mathbf{\Sigma}_{COL}$ represent inverses of $\mathbf{FIM}_{STD}$ and $\mathbf{FIM}_{COL}$, then the difference between localization uncertainties $H(f_{XSTD})$ and $H(f_{XCOL})$ is positive.

$$H(f_{XSTD}) - H(f_{XCOL}) \geq 0 \tag{5.4}$$

Hence, the collaborative CRB is reduced because of the acquisition of a new set of range measurements through the collaborative link.

Collaborative positioning has a distinguishing characteristic from a simple acquisition of new information. In addition to being characterized by an acquisition of new information, collaboration creates a dependence between the variables being estimated. In creating this dependence, collaboration conditions the position vectors being estimated. If the dependence between the nodes can be expressed such that the standalone and collaborative covariance matrices share the same diagonal, then the reduction in localization uncertainty would be the same for both nodes A and B. In this specific case, this reduction in uncertainty is called mutual information.

5.1.1 Reduction in uncertainty as mutual information

In this sub-section, we discuss conditions when the reduction in uncertainty is equal to mutual information. Recall from Chapter 2 that the mutual information MI between two random variables P and Q with marginal pdfs $f_P(p)$, $f_Q(q)$ and joint pdf $f_{P,Q}(p,q)$ is a special form of Kullback-Leibler Divergence [7]. It is defined as:

$$MI(P,Q) = KLD(f_{P,Q}(p,q)||f_P(p)f_Q(q)) = H(f_P) - H(f_{P|Q}) = H(f_Q) - H(f_{Q|P}) \quad (5.5)$$

It is positive and symmetric as proven in Appendix 2.C. It also confirms the results from Dembo, et al which indicate in [94] that conditioning reduces entropy.

For Gaussian distributions $f_{XSTD}(\mathbf{x})$ and $f_{XCOL}(\mathbf{x})$ describing the position estimates for nodes A and B, mutual information can be derived using KLD when the covariance matrices of the collaborative and standalone estimators have the same diagonal components, and the estimators have the same bias.

$$KLD(f_{XCOL}||f_{XSTD}) = \frac{1}{2}\left[\log\frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})}\right] = MI(\mathbf{x}_A, \mathbf{x}_B) \quad (5.6)$$

We prove this result in Appendix 4.H where we show that the KLD between the collaborative pdf of position estimates $f_{XCOL}(\mathbf{x})$ and the standalone pdf of position estimates $f_{XSTD}(\mathbf{x})$ is:

$$KLD(f_{XCOL}||f_{XSTD}) = \frac{1}{2}\left[MI(\mathbf{x}_A, \mathbf{x}_B) + (\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})^T \mathbf{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL} - \boldsymbol{\mu}_{STD})\right] \quad (5.7)$$

with

$$\mathbf{\Sigma}_{\boldsymbol{STD}} = \begin{bmatrix} \mathbf{\Sigma}_A & 0 \\ 0 & \mathbf{\Sigma}_B \end{bmatrix} \quad (5.8)$$

$$\boldsymbol{\Sigma_{COL}} = \begin{bmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{AB}^T & \boldsymbol{\Sigma}_B \end{bmatrix} \tag{5.9}$$

We showed in section 4.4 that if both standalone and collaborative position estimators are unbiased and efficient, $KLD(f_{XCOL}||f_{XSTD})$ reflects the maximum improvement in accuracy due to collaboration. This is intuitively true because the covariance matrices represent the inverse of FIMs. Hence, $KLD(f_{XCOL}||f_{XSTD})$ is a measure of the difference between the best performance of standalone and collaborative estimators, provided the collaborative estimator is more accurate than the standalone.

If the covariance matrices have the same diagonal, this maximum improvement in accuracy due to collaboration is the same for both nodes A and B. If the diagonals are not equal, $KLD(f_{XCOL}||f_{XSTD})$ reflects only the maximum improvement in accuracy if the estimator used is a robust estimator (like the MLE); it is no longer symmetric like the mutual information. This result is significant because it indicates that if we use a very good estimator (like the MLE), we can assess the maximum improvement in accuracy due to collaboration from the pdfs of $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$. The inconvenience of this metric is that it requires prior knowledge of $f_{XCOL}(\mathbf{x})$, hence it is not observable prior to collaboration and is not a valid metric to assess ahead of time whether one should collaborate or not. However, in practical systems, after collaboration, it is a valid metric to assess whether the estimator used maximized the potential of collaborative positioning, if we expect collaborative positioning to be more accurate than standalone positioning. The closer ΔKLD is from $KLD(f_{XCOL}||f_{XSTD})$, the more we can be assured that the collaborative estimator has reached the maximum potential it can achieve due to collaboration.

5.1.2 Summary

To summarize, when we use collaborative estimation, we acquire more information about the parameters we need to estimate. This acquisition of more information reduces the uncertainty of collaborative estimation. The *reduction in uncertainty* generated by collaboration is also called *information gain.* This reduction in uncertainty is the reason *why* the collaborative CRB is less than the standalone CRB. Under certain conditions (described earlier), this information gain is also called mutual information.

Intuitively, it is expected that for unbiased, efficient estimators with Gaussian distributions of position estimates, the maximum improvement in accuracy would be a function of the difference of the inverse FIMs. This intuitive result is confirmed through $KLD(f_{XCOL}||f_{XSTD})$ because it represents the difference of logDet between the standalone and covariance matrices. Therefore, if a very good estimator is used (one that is efficient or better), $KLD(f_{XCOL}||f_{XSTD})$ is a good indicator of the maximum improvement in accuracy due to collaboration, provided the collaborative estimator is more accurate than the standalone estimator. $KLD(f_{XCOL}||f_{XSTD})$ is more useful than a difference of logDet of covariance matrices because: 1) it is not limited to Gaussian distributions, thus it provides more useful information than a difference of covariance matrices, and 2) it incorporates the difference in estimator biases in its assessment of the maximum improvement in accuracy due to collaboration.

For Gaussian distributions, by being a function of covariance matrices for unbiased estimators, $KLD(f_{XCOL}||f_{XSTD})$ is mainly based on the geometry of the nodes with respect to (w.r.t) their anchors. This means that the maximum improvement that can be achieved due to collaboration is mainly dependent on the geometric configuration of each node w.r.t its anchors.

Although $KLD(f_{XCOL}||f_{XSTD})$ can be used to assess the maximum improvement in accuracy due to collaboration if the collaborative estimator is more accurate than the standalone estimator, it relies on prior knowledge of $f_{XCOL}(\mathbf{x})$. As a result, it is not observable prior to collaboration and is not a valid metric to assess ahead of time whether one should collaborate or not. In the next sections, we derive an upper bound to $KLD(f_{XCOL}||f_{XSTD})$. This upper bound is based on observables. Hence, its major advantage is that it can be used to assess the best improvement in accuracy due to collaboration *prior* to collaborating.

5.2 Evaluating the observable reduction in uncertainty due to collaboration

The logDet difference between the inverse FIMs from the standalone and collaborative sets of range measurements yields the reduction in localization uncertainty due to collaboration for unbiased efficient estimators with Gaussian-distributed position estimates. The reduction of the standalone CRB (CRB_{STD}) to the collaborative CRB (CRB_{COL}) is a corollary of this reduction in localization uncertainty if both standalone and collaborative estimators are efficient. In section 4.4, we established that the difference $CRB_{STD} - CRB_{COL}$ bounds the loss in accuracy due to collaboration. The larger this difference, the smaller the loss in accuracy due to collaboration and the more collaboration improves in accuracy. Hence, the difference $CRB_{STD} - CRB_{COL}$ is important in assessing the potential impact that collaboration has on the standalone accuracy. However, this difference requires prior knowledge of the true position. Hence it is not observable and cannot help in telling ahead of time whether it is worth collaborating or not.

In this section, coming from an information theoretical perspective, we introduce a simpler approach to estimate another form of reduction in uncertainty due to collaboration: the reduction in uncertainty from *the set of observed range measurements*. This approach is more suitable for practical systems because it is based on observable range measurements used to estimate the locations of nodes A and B, and can be computed prior to collaborating. We show that this reduction in range uncertainty is an upper bound to $KLD(f_{XCOL}||f_{XSTD})$ which tends to represent the maximum improvement in accuracy due to collaboration if the collaborative estimator is more accurate than the standalone estimator.

In our framework defined in Appendix 2.A, we defined the pdf of range measurements used to locate $\mathbf{x}_A$ and $\mathbf{x}_B$ in the standalone format as:

$$f_{RSTD}(r|\mathbf{x}_A, \mathbf{x}_B) = \prod_{i=1}^{3} f_{RAi}(r_{Ai}|\mathbf{x}_A) \prod_{j=1}^{3} f_{RBj}(r_{Bj}|\mathbf{x}_B) \tag{5.10}$$

where r_{Ai} is the set of observed range measurements between node A's position vector $\mathbf{x}_A$ and its i^{th} anchor $\mathbf{x}_{Ai}$, and r_{Bj} is the set of observed range measurements between node B's position vector $\mathbf{x}_B$ and its j^{th} anchor $\mathbf{x}_{Bj}$. $f_{RAi}(r_{Ai}|\mathbf{x}_A)$ and $f_{RBj}(r_{Bj}|\mathbf{x}_B)$ represent the pdfs for the sets of observed range measurements between node A and its anchors, and node B and its anchors, respectively.

The pdf of range measurements used to locate $\mathbf{x}_A$ and $\mathbf{x}_B$ in a collaborative framework is:

$$f_{RCOL}(r|\mathbf{x}_A, \mathbf{x}_B) = f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B) \prod_{i=1}^{3} f_{RAi}(r_{Ai}|\mathbf{x}_A) \prod_{j=1}^{3} f_{RBj}(r_{Bj}|\mathbf{x}_B) \tag{5.11}$$

It can be noted that $f_{RSTD}(r|\mathbf{x}_A, \mathbf{x}_B)$ is the product of marginal pdfs between pdfs of observed range measurements for nodes A and B. $f_{RCOL}(r|\mathbf{x}_A, \mathbf{x}_B)$ is the joint pdf of observed range measurements for nodes A and B. Therefore, the KLD between $f_{RCOL}(r|\mathbf{x}_A, \mathbf{x}_B)$ and $f_{RSTD}(r|\mathbf{x}_A, \mathbf{x}_B)$ is by definition a mutual information.

The mutual information between the sets of observed range measurements is therefore:

$$\begin{aligned} KLD(f_{RCOL}||f_{RSTD}) &= \int f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B) \log f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B) dr_{AB} \\ &= -H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B)) \end{aligned} \tag{5.12}$$

It is worth pointing that this reduction in uncertainty,
$\sum f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B) \log f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B)$, is a negative value when plotted originally, because $\log f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B)$ is negative as $f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B)$ is a pdf (with values less than one). The absolute value of the inverse of this negative metric is what actually represents the information gain. By dropping the negative sign of $\sum f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B) \log f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B)$, and by generating an inverse-proportional relationship, we get a larger positive information gain when the SNR is high and a small information gain when the SNR on the collaborative link is low. Intuitively, this adaptation of the information gain makes sense because

it indicates that when the collaborative link is corrupted by a noise of small variance, we acquire more information from collaboration. Hence, this is how we approach describing the observable information gain in our context of position estimation. This approach is further justified and elaborated by Lindley [95] in his adaptation of Shannon's communication theory to statistical theory. He explains: 'The reason for this is as follows: the maximum information, in a statistician's sense, will be obtained when the probability distribution is concentrated on a single value of θ, and the information will be reduced as the distribution of θ 'spreads'; this is exactly the reverse of the situation faced by a communications engineer, where the concentration on a single value would allow no choice in his messages. The two scales are therefore *reversed*.' [95]

Therefore, the **reduction in range uncertainty due to collaboration** is shown below. It is inversely proportional to the uncertainty of the pdf of the collaborative link's range measurements.

$$\begin{aligned}\text{Information Gain} &= |\sum f_{RAB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B)\log f_{RAB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B)| \\ &= |H(f_{RAB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B))|^{-1}\end{aligned} \tag{5.13}$$

The information gain provided in equation (5.13) is based on observable range measurements rather than a distribution of position estimates. It is the same for both nodes A and B because it is a mutual information. We showed in the previous section that the reduction in localization uncertainty (the LogDet difference), which indicates the maximum increase in ΔKLD when collaboration is more accurate, is based on prior knowledge of the distributions of position estimates. Thus, a question arises: 'How does the information gain (from observed range measurements) relate to the reduction in localization uncertainty from the distribution of position estimates?' From the data processing inequality, we expect the mutual information from equation (5.13) to be an upper bound to the information gain from the distribution of position estimates on average. Hence, the following inequality holds **if collaborative positioning is more accurate than standalone**:

$$KLD(f_{XCOL}||f_{XSTD}) \leq |H(f_{RAB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B))|^{-1} \tag{5.14}$$

where $KLD(f_{XCOL}||f_{XSTD})$ relies heavily on the geometry of the nodes w.r.t their anchors and requires prior knowledge of the distributions of both standalone and collaborative estimators. $|H(f_{RAB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B))|^{-1}$ relies heavily on the noise on the collaborative link and only requires prior knowledge of the distribution of range measurements on the collaborative link.

5.3 Reconciling inaccuracy and uncertainty analyses

How do uncertainty and inaccuracy affect the improvement in accuracy due to collaboration? In this section, we reconcile both theoretical uncertainty and inaccuracy analyses. We simulate the theoretical inequalities derived so far using the MLE estimator which is non-linear, but is the recommended estimator in collaborative position estimation problems.

5.3.1 Summary of Inequalities in Collaborative Positioning

Recall that $f_{XCOL}(\mathbf{x})$ and $f_{XSTD}(\mathbf{x})$ represent the pdfs of position estimates for both collaborative and standalone estimators, respectively. $f_{Xref}(\mathbf{x})$ represents the pdf of estimates from the most accurate position estimator.
Combining the lower bound from Section 4.6 equation (4.55) and the upper bound from equation (5.14) above, $KLD(f_{XCOL}||f_{XSTD})$ is bounded as follows:

$$\frac{1}{2}(\boldsymbol{\mu}_{COL}-\boldsymbol{\mu}_{STD})^T\boldsymbol{\Sigma}_{STD}^{-1}(\boldsymbol{\mu}_{COL}-\boldsymbol{\mu}_{STD}) \leq KLD(f_{XCOL}||f_{XSTD}) \leq |H(f_{AB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B))|^{-1} \tag{5.15}$$

From the inaccuracy analysis in Chapter 4, we showed that if $f_{Xref}(\mathbf{x}) = f_{XCOL}(\mathbf{x})$, the best performance of the collaborative position estimator occurs when
$\Delta KLD = KLD(f_{XCOL}||f_{XSTD})$. Hence, a ΔKLD value very close to $KLD(f_{XCOL}||f_{XSTD})$ shows that the collaborative estimator has reached its best performance. To put it simply, if a good estimator (like the MLE) is used, and collaborative positioning tends to be more accurate than standalone, then, we expect that on average:

$$\Delta KLD \leq KLD(f_{XCOL}||f_{XSTD}) \tag{5.16}$$

We validate this through our simulation results.

As a result:

$$\Delta KLD \leq KLD(f_{XCOL}||f_{XSTD}) \leq |H(f_{AB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B))|^{-1} \tag{5.17}$$

Where $KLD(f_{XCOL}||f_{XSTD})$ is a tight bound on average, and only applies when collaborative positioning is more accurate than standalone. $KLD(f_{XCOL}||f_{XSTD})$ can only be observed post-collaboration, while $|H(f_{AB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B))|^{-1}$ is an observable upper bound on average, that we expect to be looser than $KLD(f_{XCOL}||f_{XSTD})$. The simulation results validate these expectations. The bounds derived are probabilistic upper bounds that hold on average. This implies that for a single individual test case, we may not be able to observe this inequality. However, these inequalities hold over an average of different tests.

A practical bound to take away from these theoretical analyses is:

$$\Delta KLD \leq |H(f_{AB}(r_{AB}|\mathbf{x}_A,\mathbf{x}_B))|^{-1} \tag{5.18}$$

This bound indicates that when a good estimator (like the MLE) is used, then, we expect the improvement in accuracy due to collaboration to be less than the information gain acquired from the set of range measurements on the collaborative link.

5.3.2 Simulation results and discussion

We test the bounds from the previous section using the simulation conditions described in Section 4.3.2. In a 300 m-by-300 m Matlab simulation environment described on Fig. 5.1, the SNR on all standalone and collaborative links is fixed at 20 dB. A 20 dB SNR is a fairly realistic assumption; it indicates that the standard deviation for the noise on a set of range measurements is a tenth of the true distance. We use the Newton-Raphson MLE to estimate positions for nodes A and B and assume LOS on all range measurements. We vary the value of angle θ_{AB}, the angle between nodes A and B, from 0 to 2π in 21 intervals, by fixing node A and varying the location of node B in a circle of fixed radius of 40 meters. We repeat the same experiment at 100 different locations for point A by maintaining good geometric configuration and the radius between A and B fixed. At each fixed position for node A and node B, we generate 1000 position estimates and compute the following metrics: ΔKLD, $KLD(f_{XCOL}||f_{XSTD})$ that we name *'UpperBound'*, and $|H(f_{AB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$ that we name *'InfoGain'* in Fig.5.2-Fig.5.6. We average the values of these metrics over the 100 different test cases. The goal of these simulations is to validate the bounds derived in equation (5.17).

Our simulation results validate our theoretical results on average (by averaging the results from 100 simulation testcases) as shown in Fig.5.2, Fig.5.3 and Fig.5.5. Therefore, although the theoretical results derived were obtained based on unbiased, efficient estimators with Gaussian distributions, we note that on average these bounds hold when using the non-linear MLE estimator in LOS conditions as shown in the simulation results included.

Fig.5.2, Fig.5.3 and Fig.5.5 display the average of ΔKLD, $KLD(f_{XCOL}||f_{XSTD})$ and the information gain $|H(f_{AB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$ over 100 test cases. These values were computed for just the scalar x and y coordinates of nodes A (Fig.5.2) and B (Fig.5.3), and also for the combined coordinates of both nodes (Fig.5.5). The plots confirm that the bounds from equation (5.17) hold on average.

Fig.5.4 and Fig.5.6 show the percentage of times $KLD(f_{XCOL}||f_{XSTD})$ and $|H(f_{AB}(r|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$ are upper bounds to ΔKLD for 100 individual cases. These plots reflect that for individual test cases, the bounds do not always hold. However, on average, they do.

As discussed earlier, the fact that ΔKLD is very close to $KLD(f_{XCOL}||f_{XSTD})$ indicates that the MLE estimator tends to maximize the attainable gain due to collaboration. The simulation plots confirm the results from the literature that the MLE is the recommended estimator for collaborative positioning. Fig.5.2-Fig.5.3 show that ΔKLD attains $KLD(f_{XCOL}||f_{XSTD})$ on average. As such, when collaborative positioning is more accurate than standalone positioning, and the MLE is used, collaboration tends to provide on average the maximum

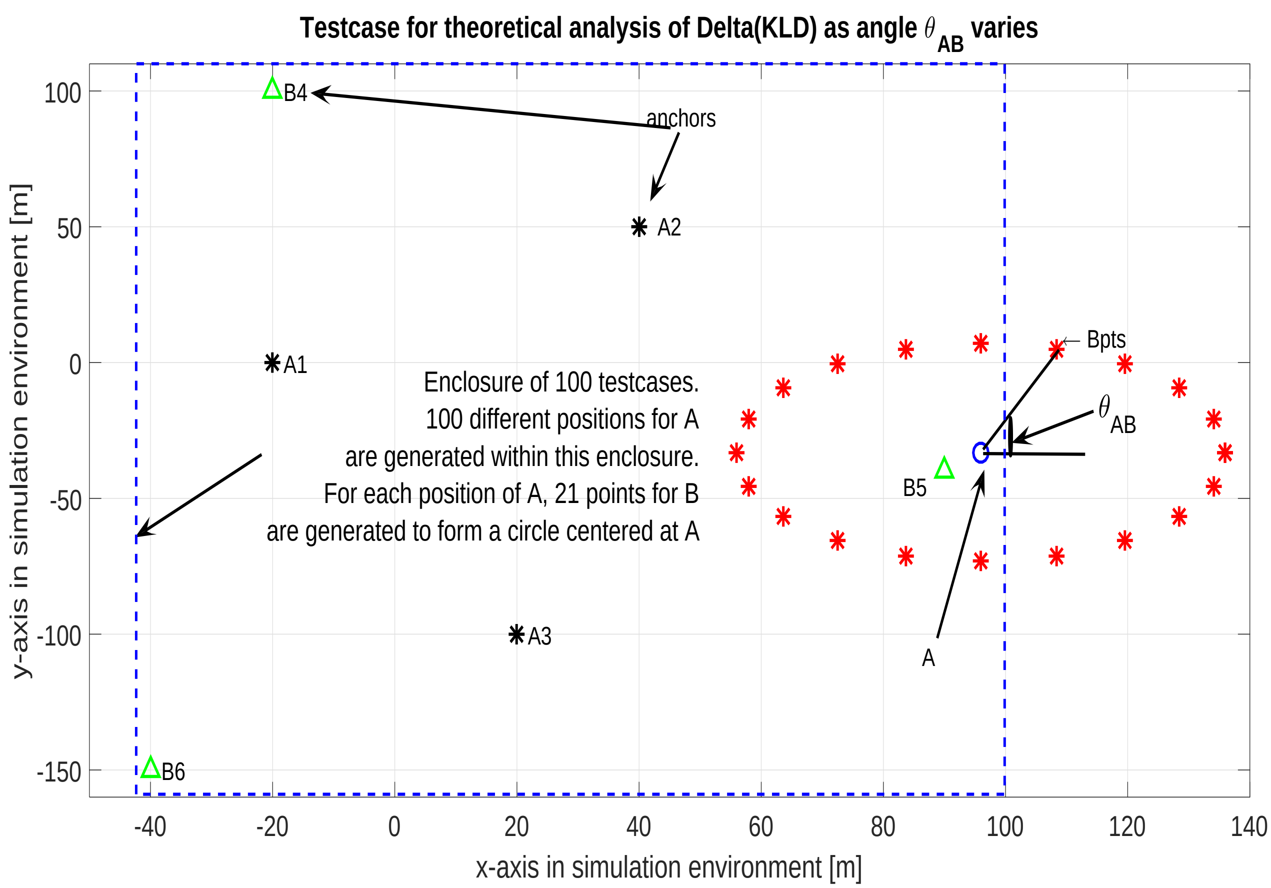


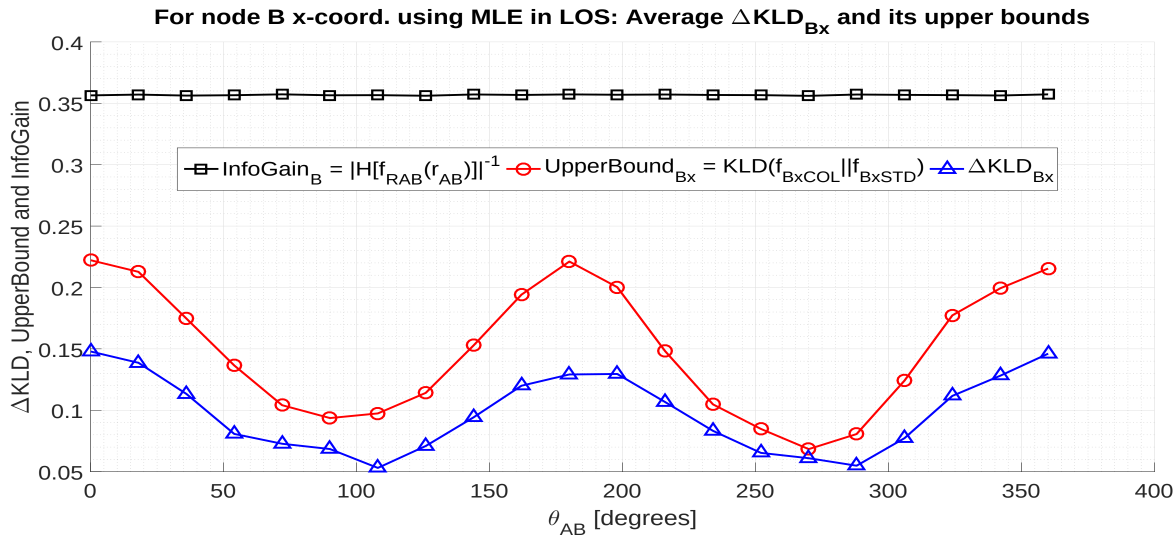
Figure 5.1: Simulation setup to validate the performance of the bounds relating ΔKLD, $KLD(f_{XCOL}||f_{XSTD})$ and $|H(f_{AB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$. Node B rotates around node A in a circle such that angle θ_{AB} varies in different geometric configurations with respect to all anchors. 100 testcases are considered. For each test case, node B rotates around node A in 21 intervals of angles ranging from 0 to 2π, at a radius of 40 meters. At each fixed position of nodes A and B, 1000 position estimates for both nodes are generated in both standalone and collaborative frameworks.

improvement in accuracy for x and y coordinates in LOS conditions.

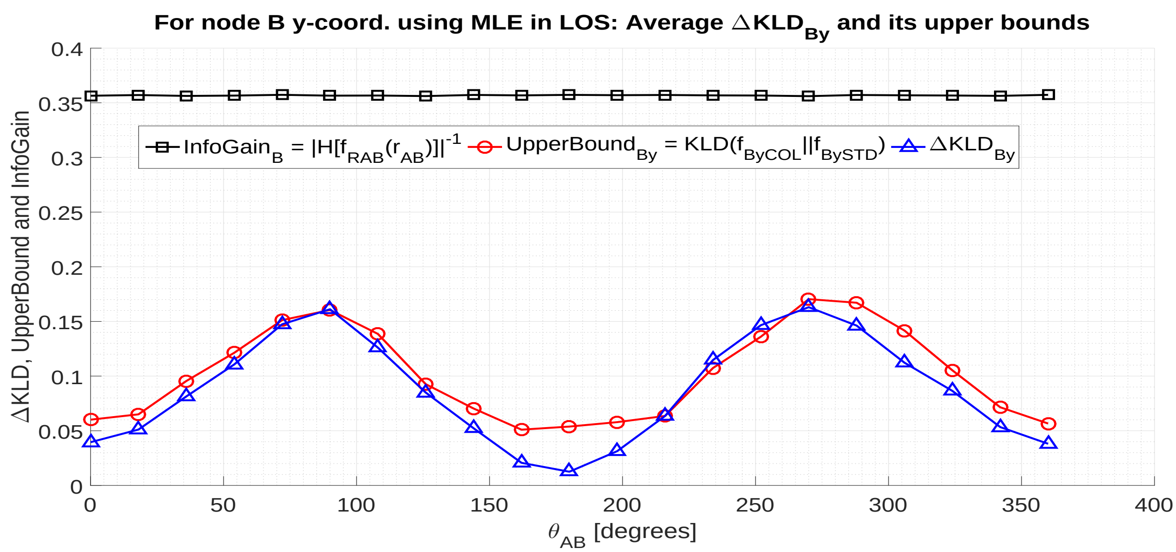
Subplots (a) and (b) in Fig.5.2-Fig.5.3 and Fig.5.5 also confirm that on average, as mentioned in the inaccuracy analysis, when two nodes share a coordinate, the lack of geometric diversity yields extrema in ΔKLD. We also observed from the simulation results that changes in accuracy for node A's x and y coordinates are opposite. The same occurs for node B's x and y coordinates. These results confirm our variational analysis results from Chapter 4.

It is also worth noting that the scalar bounds (Fig.5.2(b)-Fig.5.3(b)) perform better than the multivariate bounds (Fig.5.6); that is, the scalar bounds have a higher accuracy than the multivariate bounds. This difference in performance between scalar and multivariate KLD-based metrics has been noticed in the literature and addressed in Chapter 2 in particular for the case of mutual information where its scalar form tends to be more intuitive to interpret than its multivariate form.

The bound labelled in all plots as $InfoGain$ is as expected a looser upper bound to ΔKLD. In practical systems however, it is observable. Therefore, it is a valid criteria to assess the maximum improvement in accuracy due to collaboration, **prior to collaborating**. In the next chapter, we use it as a criteria to tell ahead-of-time whether collaboration will be beneficial in terms of accuracy or not.

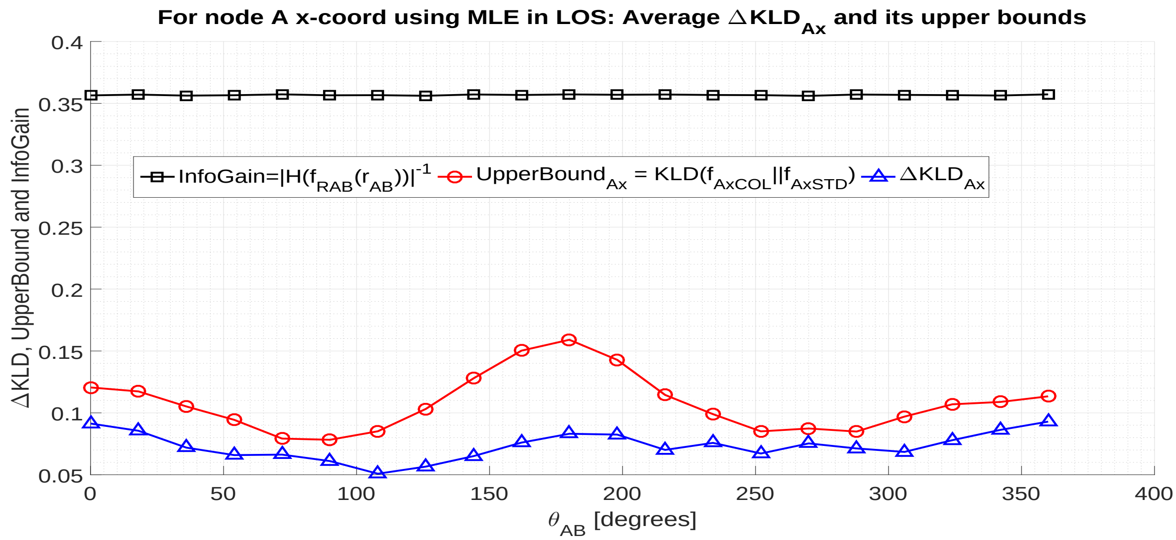


(a) Node B: x-coordinate - Average results for different bounds

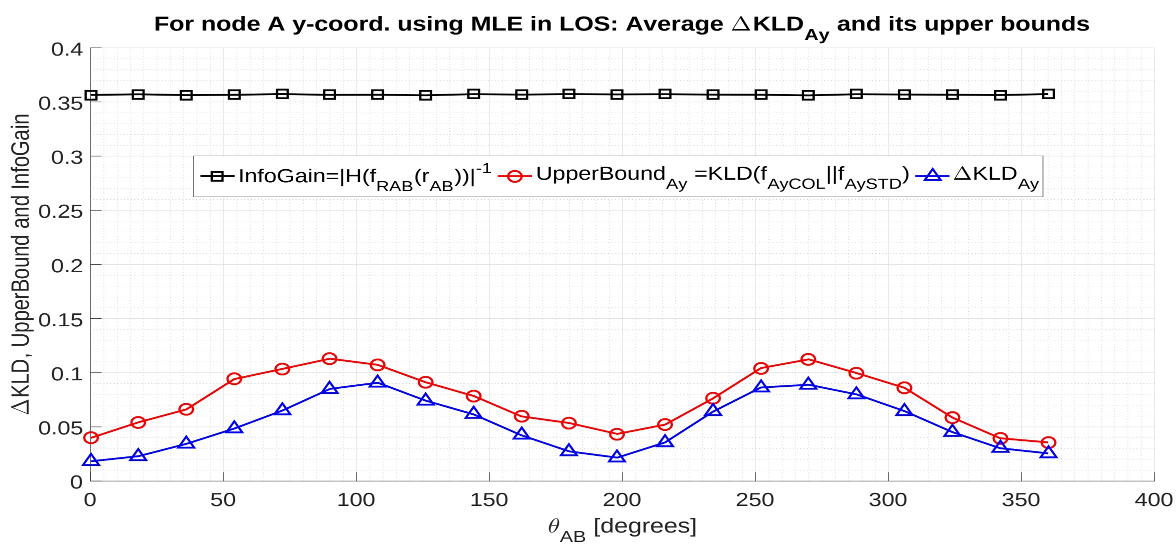


(b) Node B: y-coordinate - Average results for different bounds

Figure 5.2: Upper bounds to ΔKLD for node B's x and y coordinates. These results show node B's x (a) and y (b) coordinates derived using the MLE in LOS conditions based on Fig. 5.1 setup. The metrics shown are averaged over 100 test cases. This figure describes how the following inequality holds: $\Delta KLD \leq KLD(f_{XCOL}||f_{XSTD}) \leq |H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$. ΔKLD is the change in accuracy due to collaboration. $InfoGain = |H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$ is the observable reduction in range uncertainty from the set of observed range measurements. It is a looser bound than $UpperBound = KLD(f_{XCOL}||f_{XSTD})$ which represents the best improvement in position accuracy if the collaborative estimator is more accurate than the standalone estimator.

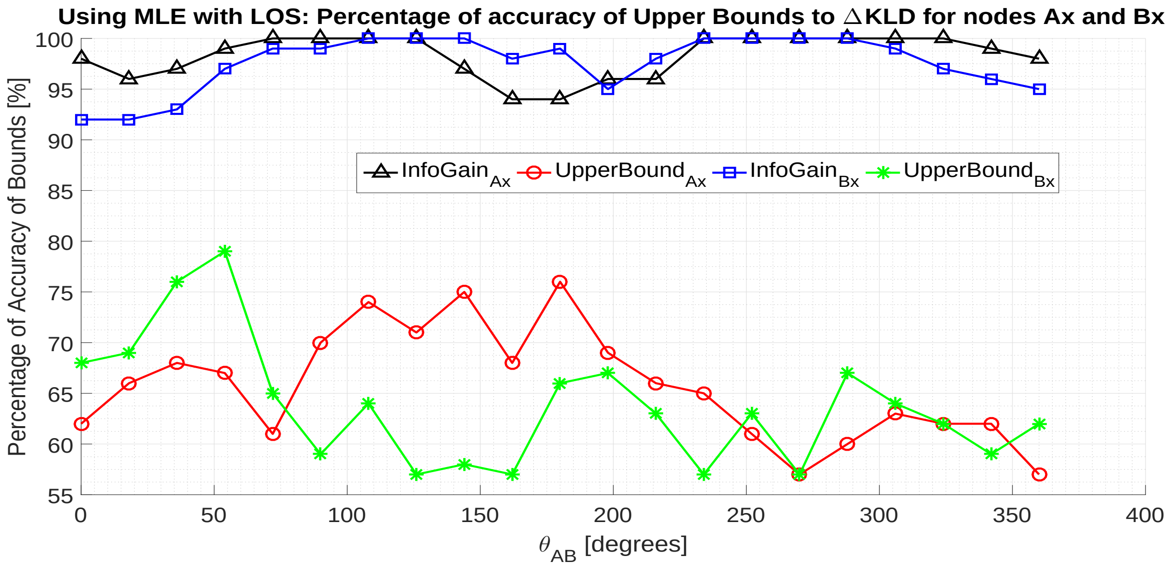


(a) Node A: x-coordinate - Average results for different bounds

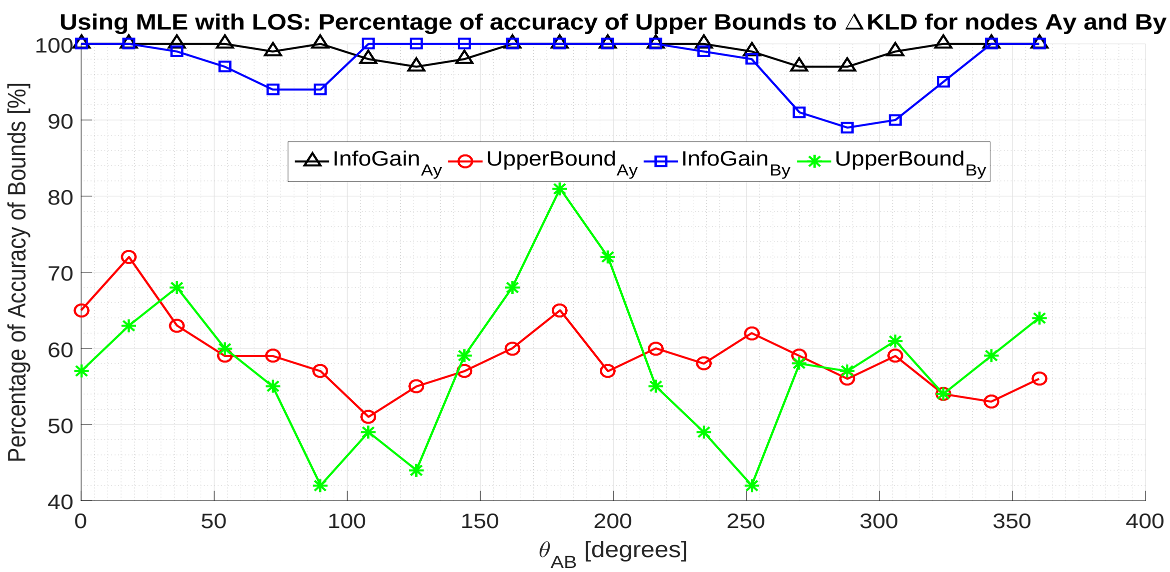


(b) Node A: y-coordinate - Average results for different bounds

Figure 5.3: Upper bounds to ΔKLD for node A's x and y coordinates. These results show node A's x (a) and y (b) coordinates derived using the MLE in LOS conditions based on Fig. 5.1 setup. The metrics shown are averaged over 100 test cases. This figure describes how the following inequality holds: $\Delta KLD \leq KLD(f_{XCOL}||f_{XSTD}) \leq |H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$. ΔKLD is the change in accuracy due to collaboration. $InfoGain = |H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$ is the observable reduction in range uncertainty from the set of observed range measurements. It is a looser bound than $UpperBound = KLD(f_{XCOL}||f_{XSTD})$ which represents the best improvement in position accuracy if the collaborative estimator is more accurate than the standalone estimator.

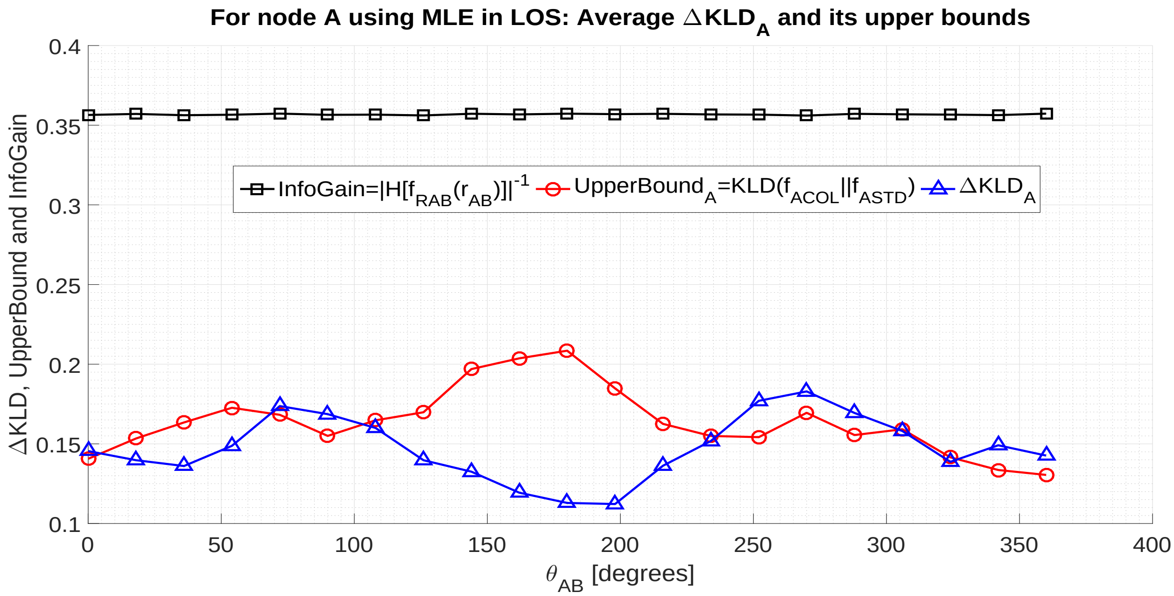


(a) Node A & B: x-coordinates - Accuracy of Upper Bounds

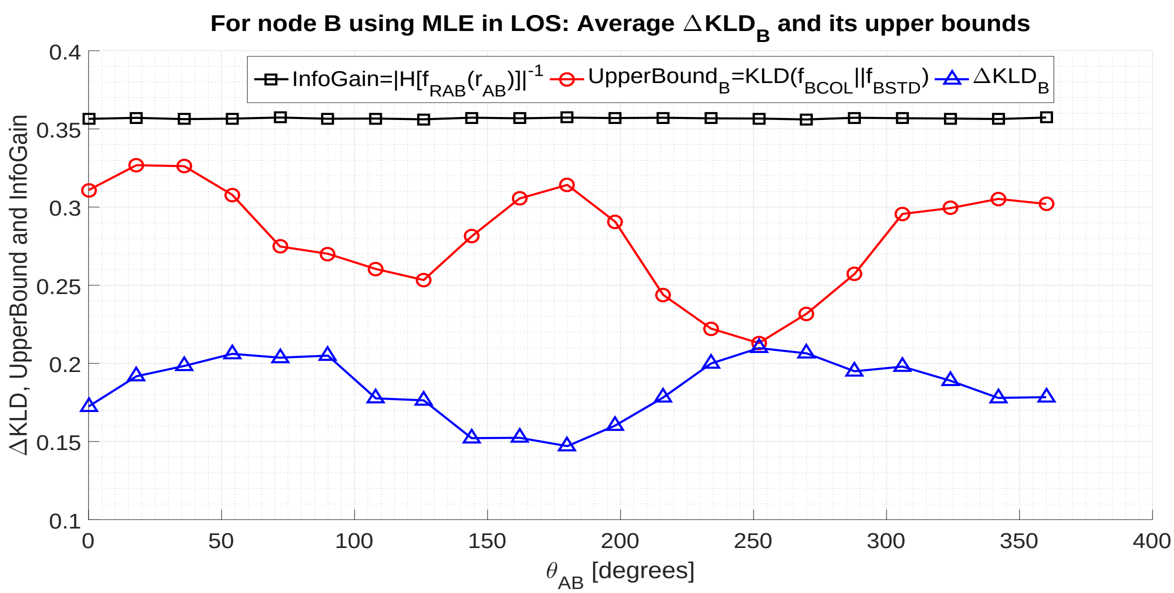


(b) Node A & B: y-coordinates - Accuracy of Upper Bounds

Figure 5.4: Accuracy of Upper Bounds to ΔKLD for nodes A and B, x-y coordinates. These plots describe the percentage of times the proposed upper bounds are accurate for 100 test cases. Results are based on position estimates for node A's x and y coordinates using the MLE in LOS range conditions based on Fig. 5.1 setup. $InfoGain = |H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$ is the observable reduction in range uncertainty from the set of observed range measurements. $UpperBound$ represents $KLD(f_{XCOL}||f_{XSTD})$. These plots confirm that the bounds defined Fig.5.2 and Fig.5.3 are valid on average or better.



(a) Node A - Average results for different bounds



(b) Node B - Average results for different bounds

Figure 5.5: Upper Bounds to ΔKLD for nodes A and B - multivariate case. In this figure we use the MLE in LOS conditions (based on Fig. 5.1 setup). The metrics shown are averaged over 100 test cases. This figure describes how the following inequality holds: $\Delta KLD \leq KLD(f_{XCOL}||f_{XSTD}) \leq |H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1}$. The fact that ΔKLD tends towards $KLD(f_{XCOL}||f_{XSTD})$ indicates that on average, the estimator used maximizes the potential that collaborative positioning can provide.

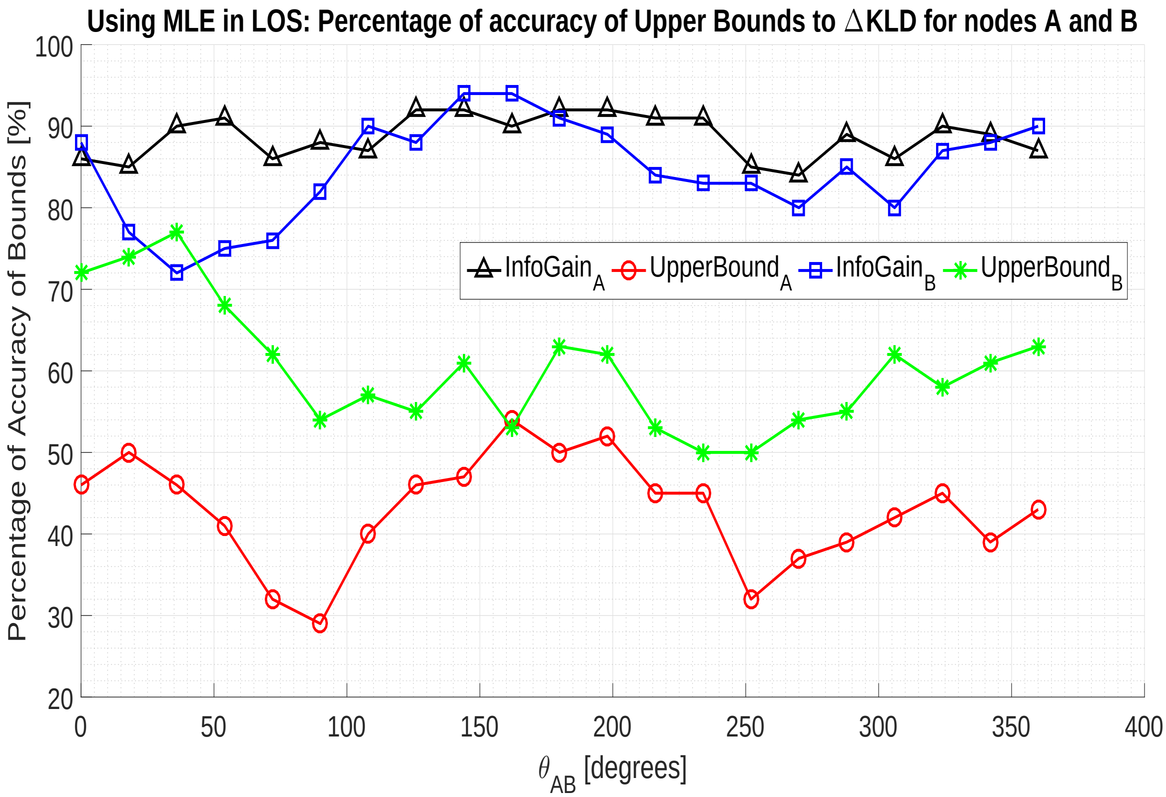


Figure 5.6: Accuracy of Upper Bounds to ΔKLD for nodes A and B - multivariate case. In this figure we use the MLE in LOS range conditions (based on Fig. 5.1 setup). The metrics shown are averaged over 100 test cases for the 2-dimensional position estimates for both nodes A and B. This plot indicates that the 2-dimensional metrics do not perform as well as the x-coordinates and y-coordinates metrics derived in Fig. 5.2 and Fig. 5.3

5.4 Conclusion

In this chapter, we explained *why* the collaborative Cramer-Rao Bound (CRB) is less than the standalone CRB. When we estimate collaboratively, we acquire new information. As a result, the *information gain* from collaboration yields a reduction in localization uncertainty that we derived for the case of unbiased efficient estimators with Gaussian distributions. The fact that localization uncertainty reduces with collaboration explains why the collaborative CRB is smaller than the standalone CRB. For unbiased efficient estimators with Gaussian distributions, the reduction in localization uncertainty is strictly equal to the positive difference of LogDets of inverse FIMs.

From our theoretical framework, we derived bounds for the reduction in localization uncertainty due to collaboration. We reconciled the uncertainty and inaccuracy results by deriving inequalities between ΔKLD (which represents the improvement in accuracy) and uncertainty bounds.

By simulating these theoretical inequalities using the Maximum Likelihood estimator, we were able to generalize our theoretical results. Via simulations, we showed that the information gain acquired by the addition of the collaborative link is an acceptable probabilistic upper bound to ΔKLD for the MLE in LOS range conditions. As such, it is a valid criteria to assess ahead of time whether it is worth collaborating to improve accuracy. We address the use of this criteria in the next chapter and validate it in both LOS and NLOS conditions, and in both good and poor geometric configurations.

Appendix

5.A Entropy for multivariate Gaussian Distributions

In this appendix, we derive the entropy or uncertainty of a multivariate Gaussian distribution. We consider that the log function has base e: $\log = \log_e$ ($i.e$: the natural log).

The pdf $f_X(\mathbf{x})$ of an n-dimensional Gaussian distribution of mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is defined as:

$$f_X(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) \tag{5.19}$$

We prove that the entropy or uncertainty for $f_X(\mathbf{x})$ is:

$$H(f_X) = \frac{1}{2}\log\left((2\pi e)^n \det(\boldsymbol{\Sigma})\right) \tag{5.20}$$

It is worth noting that entropy does not depend on the bias of the distribution.

Proof:

$$\begin{aligned}
H(f_X) &= -\int f_X(\mathbf{x}) \log(f_X(\mathbf{x})) d\mathbf{x} \\
&= -E_{f_X}\left[\log(f_X(\mathbf{x}))\right] \\
&= E_{f_X}\left[\frac{1}{2}\log\left((2\pi)^n \det(\boldsymbol{\Sigma})\right) + \left(\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)\right] \\
&= \frac{1}{2}\log\left((2\pi)^n \det(\boldsymbol{\Sigma})\right) + \frac{1}{2}E_{f_X}\left[\left((\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)\right] \\
&= \frac{1}{2}\log\left((2\pi)^n \det(\boldsymbol{\Sigma}) + n\right) \\
&= \frac{1}{2}\log\left((2\pi e)^n \det(\boldsymbol{\Sigma})\right)
\end{aligned} \tag{5.21}$$

5.B Reduction in localization uncertainty due to collaboration

In this appendix, we derive the reduction in localization uncertainty due to collaboration. Our goal is to show that the acquisition of new information through collaboration reduces uncertainty and is the main reason why the CRB reduces with collaboration. We also show that under certain conditions, the reduction in localization uncertainty reduces by the same amount for each of the collaborating nodes.

5.B.1 Proof: Collaboration reduces the CRB

Let $f_{XCOL}(\mathbf{x})$ represent the multivariate Gaussian distribution of position estimates from the collaborative estimator. Let $f_{XSTD}(\mathbf{x})$ be the multivariate Gaussian pdf of position estimates from the standalone estimator. We assume that the covariance matrices of both standalone and collaborative estimators are equal to the inverse of their respective Fisher Information Matrices. That is, we assume $\mathbf{\Sigma}_{STD} = \mathbf{FIM}_{STD}^{-1}$ and $\mathbf{\Sigma}_{COL} = \mathbf{FIM}_{COL}^{-1}$.

From (Appendix 5.A), the entropy $H(f_{XSTD})$ of $f_{XSTD}(\mathbf{x})$ is:

$$H(f_{XSTD}) = \frac{1}{2} \log\left((2\pi e)^n \det\left(\mathbf{\Sigma}_{STD}\right)\right) \tag{5.22}$$

Where $\log = \log_e$ ($i.e$: the natural log function)

The entropy $H(f_{XCOL})$ of $f_{XCOL}(\mathbf{x})$ is:

$$H(f_{XCOL}) = \frac{1}{2} \log\left((2\pi e)^n \det\left(\mathbf{\Sigma}_{COL}\right)\right) \tag{5.23}$$

Hence, the change in entropy due to collaboration is:

$$\begin{aligned} H(f_{XSTD}) - H(f_{XCOL}) &= \log\det\mathbf{\Sigma}_{STD} - \log\det\mathbf{\Sigma}_{COL} \\ &= \log\frac{\det\mathbf{\Sigma}_{STD}}{\det\mathbf{\Sigma}_{COL}} \end{aligned} \tag{5.24}$$

Recall from Appendix 2.A equation 2.68 that:

$$\mathbf{\Sigma}_{COL}^{-1} = \mathbf{\Sigma}_{STD}^{-1} + \mathbf{R} \tag{5.25}$$

Therefore, the change in entropy due to collaboration is:

$$\begin{aligned} H(f_{XSTD}) - H(f_{XCOL}) &= \log \frac{\det \mathbf{\Sigma}_{COL}^{-1}}{\det \mathbf{\Sigma}_{STD}^{-1}} \\ &= \log \frac{\det (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{R})}{\det \mathbf{\Sigma}_{STD}^{-1}} \end{aligned} \tag{5.26}$$

where matrix $\mathbf{R} = \frac{\mathbf{b}\mathbf{b}^T}{\sigma_{AB}^2}$, with σ_{AB}^2 being the variance of the set of observed range measurements on the collaborative link, and vector $\mathbf{b}$ is such that:

$$\mathbf{b} = [\cos(\theta_{AB}), \sin(\theta_{AB}), -\cos(\theta_{AB}), -\sin(\theta_{AB})]^T \tag{5.27}$$

where θ_{AB} is the angle between nodes A and B. Hence,

$$\begin{aligned} H(f_{XSTD}) - H(f_{XCOL}) &= \log \frac{\det (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{R})}{\det \mathbf{\Sigma}_{STD}^{-1}} \\ &= \log \frac{\det (\mathbf{\Sigma}_{STD}^{-1} + \sigma_{AB}^{-2}\mathbf{b}\mathbf{b}^T)}{\det \mathbf{\Sigma}_{STD}^{-1}} \end{aligned} \tag{5.28}$$

It can be proven that:

$$\det (\mathbf{\Sigma}_{STD}^{-1} + \sigma_{AB}^{-2}\mathbf{b}\mathbf{b}^T) = (1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b}) \det (\mathbf{\Sigma}_{STD}^{-1}) \tag{5.29}$$

So,

$$\begin{aligned} H(f_{XSTD}) - H(f_{XCOL}) &= \log \frac{(1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b}) \det (\mathbf{\Sigma}_{STD}^{-1})}{\det \mathbf{\Sigma}_{STD}^{-1}} \\ &= \log (1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b}) \end{aligned} \tag{5.30}$$

By definition, an n-by-n matrix $\mathbf{M}$ is positive semi-definite **if and only if** for all real n-dimensional vector $\mathbf{p}$, the following equation holds:

$$\mathbf{p}^T\mathbf{M}\mathbf{p} \geq 0 \tag{5.31}$$

Since the condition **'if and only if'** holds in both directions, we can conclude that because matrix $\mathbf{\Sigma}_{STD}$ is a covariance matrix, it is positive semi-definite. Hence, for any real vector $\mathbf{b}$, the following occurs:

$$\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b} \geq 0 \tag{5.32}$$

Hence, $\sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b} \geq 0$ which infers that the argument of the log function in $[\log (1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b})]$ is greater than unity. As a result, the change in entropy is positive.

$$\begin{aligned} H(f_{XSTD}) - H(f_{XCOL}) &= \log (1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b}) \\ &\geq 0 \end{aligned} \tag{5.33}$$

We note that vector $\mathbf{b}$ represents the rank-1 update to matrix $\mathbf{G}_{STD}$ as discussed in Appendix 4.E. Hence, $\mathbf{b}$ reflects the acquisition of new information. From the equation above, this new information acquired through collaboration reduces the localization uncertainty of the standalone estimator by an amount equal to: $\log(1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b})$.

Furthermore, in Appendix 4.E, we derived a closed-form expression relating the collaborative covariance to the standalone covariance matrix.

$$\begin{aligned}\mathbf{\Sigma}_{COL} &= (\mathbf{\Sigma}_{STD}^{-1} + \mathbf{R})^{-1} \\ &= \mathbf{\Sigma}_{STD} - \frac{r}{\sigma_{AB}^2}(\mathbf{\Sigma}_{STD}\mathbf{b}\mathbf{b}^T\mathbf{\Sigma}_{STD})\end{aligned} \tag{5.34}$$

with scalar r positive and equal to:

$$r = \frac{1}{1 + \sigma_{AB}^{-2}\mathbf{b}^T\mathbf{\Sigma}_{STD}\mathbf{b}} \tag{5.35}$$

Taking the trace of both sides of equation (5.34), we get:

$$\mathrm{Tr}(\mathbf{\Sigma}_{COL}) = \mathrm{Tr}(\mathbf{\Sigma}_{STD}) - \mathrm{Tr}(\frac{r}{\sigma_{AB}^2}(\mathbf{\Sigma}_{STD}\mathbf{b}\mathbf{b}^T\mathbf{\Sigma}_{STD})) \tag{5.36}$$

If $\mathbf{\Sigma}_{COL} = \mathbf{FIM}_{COL}^{-1}$, and $\mathbf{\Sigma}_{STD} = \mathbf{FIM}_{STD}^{-1}$, then their traces yield the collaborative and standalone Cramer-Rao Bounds, respectively. As a result, the reduction in CRBs due to the acquisition of the collaborative link yields:

$$\mathrm{Tr}(\mathbf{\Sigma}_{STDCRB}) - \mathrm{Tr}(\mathbf{\Sigma}_{COLCRB}) = \mathrm{Tr}(\frac{r}{\sigma_{AB}^2}(\mathbf{\Sigma}_{STD}\mathbf{b}\mathbf{b}^T\mathbf{\Sigma}_{STD})) \tag{5.37}$$

which is a quadratic form that is always positive. Therefore, the acquisition of new information through collaboration yields a reduction in localization uncertainty. The reduction in CRBs due to collaboration is a corollary of this result.

5.B.2 Uncertainty reduction by same amount under certain conditions

In the previous section, we showed that for Gaussian distributions of position estimates, the LogDet Ratio reflects the reduction in localization uncertainty due to collaboration. For unbiased efficient estimators, this reduction in uncertainty is also the maximum improvement in accuracy due to collaboration.
In this appendix, we compute the LogDet Ratio of covariance matrices $\boldsymbol{\Sigma_{STD}}$ and $\boldsymbol{\Sigma_{COL}}$ assuming that the covariance matrices have the same diagonal. This result is significant

because it implies that when collaborating, the potential to improve position accuracy increases by the same amount for the collaborating nodes if the standalone and collaborative covariance matrices have the same diagonal.

Let us consider that the standalone and collaborative position estimators have covariance matrices defined as shown below. We assume that their diagonals are equal.

$$\boldsymbol{\Sigma_{STD}} = \begin{bmatrix} \boldsymbol{\Sigma}_A & 0 \\ 0 & \boldsymbol{\Sigma}_B \end{bmatrix} \tag{5.38}$$

$$\boldsymbol{\Sigma_{COL}} = \begin{bmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{AB}^T & \boldsymbol{\Sigma}_B \end{bmatrix} \tag{5.39}$$

Using block matrices, and Schur complements, the determinant for the standalone covariance matrix is:

$$\det \boldsymbol{\Sigma_{STD}} = \det(\boldsymbol{\Sigma}_A)\det(\boldsymbol{\Sigma}_B) \tag{5.40}$$

The determinant for the collaborative covariance matrix is:

$$\det \boldsymbol{\Sigma_{COL}} = \det(\boldsymbol{\Sigma}_A)\det\left(\boldsymbol{\Sigma}_B - \boldsymbol{\Sigma}_{AB}^T\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{AB}\right) \tag{5.41}$$

where $\left(\boldsymbol{\Sigma}_B - \boldsymbol{\Sigma}_{AB}^T\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{AB}\right)$ is the Schur complement of $\boldsymbol{\Sigma}_A$. By definition, since $\boldsymbol{\Sigma}_{COL}$ is a covariance matrix, the Schur complement is also equal to the conditional covariance, which means:

$$\boldsymbol{\Sigma}_{B|A} = \left(\boldsymbol{\Sigma}_B - \boldsymbol{\Sigma}_{AB}^T\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{AB}\right) \tag{5.42}$$

Similarly, the determinant for the collaborative covariance matrix is also equal to:

$$\det \boldsymbol{\Sigma_{COL}} = \det(\boldsymbol{\Sigma}_B)\det\left(\boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{AB}^T\right) \tag{5.43}$$

where $\left(\boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{AB}^T\right)$ is the Schur complement of $\boldsymbol{\Sigma}_B$. By definition, since $\boldsymbol{\Sigma}_{COL}$ is a covariance matrix, the Schur complement is also equal to the conditional covariance, which means:

$$\boldsymbol{\Sigma}_{A|B} = \left(\boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{AB}^T\right) \tag{5.44}$$

A sanity check on the condition that A and B estimations are independent, would make the previous equation yield: $\boldsymbol{\Sigma}_{B|A} = \boldsymbol{\Sigma}_B$ (and $\boldsymbol{\Sigma}_{A|B} = \boldsymbol{\Sigma}_A$) because $\boldsymbol{\Sigma}_{AB}$ is zero in standalone positioning.

Hence, the difference of log determinants yields the same reduction in uncertainty for nodes A and B:

$$\begin{aligned} \log\frac{\det(\boldsymbol{\Sigma}_{STD})}{\det(\boldsymbol{\Sigma}_{COL})} &= \log\frac{\det(\boldsymbol{\Sigma}_A)\det(\boldsymbol{\Sigma}_B)}{\det(\boldsymbol{\Sigma}_A)\det(\boldsymbol{\Sigma}_{B|A})} \\ &= \log\frac{\det(\boldsymbol{\Sigma}_B)}{\det(\boldsymbol{\Sigma}_{B|A})} = \log\frac{\det(\boldsymbol{\Sigma}_A)}{\det(\boldsymbol{\Sigma}_{A|B})} \end{aligned} \tag{5.45}$$

As a covariance matrix, $\det(\mathbf{\Sigma}_{COL})$ is positive semidefinite. Given that $\mathbf{\Sigma}_A$ is also a covariance matrix, it is positive semidefinite. Hence, its Schur complement $\mathbf{\Sigma}_{B|A}$ is also positive semidefinite, and the following inequalities occurs, from equation (5.44):

$$\det\left(\mathbf{\Sigma}_{B|A}\right) \leq \det\left(\mathbf{\Sigma}_B\right) \tag{5.46}$$

$$\det\left(\mathbf{\Sigma}_{A|B}\right) \leq \det\left(\mathbf{\Sigma}_A\right) \tag{5.47}$$

Consequently, the ratio of determinants is always greater than or equal to unity. So:

$$\log \frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})} \geq 0 \tag{5.48}$$

Most importantly, the logDet difference, which represents the reduction in uncertainty due to collaboration is the same for node A and B.

$$\begin{aligned} \log \frac{\det(\mathbf{\Sigma}_{STD})}{\det(\mathbf{\Sigma}_{COL})} &= \log \frac{\det(\mathbf{\Sigma}_B)}{\det(\mathbf{\Sigma}_{B|A})} \\ &= \log \frac{\det(\mathbf{\Sigma}_A)}{\det(\mathbf{\Sigma}_{A|B})} \end{aligned} \tag{5.49}$$

Chapter 6

Criteria to recommend collaboration on-the-fly

Having identified theoretically factors that affect the accuracy of collaborative positioning, can we tell ahead of time whether collaboration will improve accuracy or not? This is the main research problem addressed in this chapter.

We address the following topic: 'How can we assess on-the-fly, that is, without prior knowledge of the true position, and without estimating collaboratively, whether collaborating will be beneficial in terms of accuracy?' From the results of the previous chapters: a) geometry and b) the quality of the collaborative link affect the accuracy of collaborative positioning. We use these two factors to identify criteria that can help assess whether it is worth collaborating or not on-the-fly.

To summarize the results discussed in this chapter, we derive criteria to tell ahead of time when to collaborate for improved accuracy. The criteria selected are based on the uncertainty of the collaborative link and on the generalized geometric dilution of precision (GGDOP). They are tested for both good and poor geometric configurations in both LOS and NLOS conditions.

6.1 Criteria Definitions

Two major factors affect the accuracy of collaborative positioning: the geometric configuration of the nodes with respect to (w.r.t) their anchors, and the quality of the collaborative link.

6.1.1 Geometric Criterion

In regard to geometry, we consider the Generalized GDOP (GGDOP) defined below ([99], [16], [76]). It was selected because it partially dissociates the impact of range variances and geometry on the localization error as elaborated by Buehrer et al [15]. In addition, we showed in Chapter 4 that GGDOP affects the sign of ΔKLD. Given m range measurements obtained from m nodes of known location, the angle α_{Ai} between points A and the i^{th} node:

$$\alpha_i = \tan^{-1} \frac{(y_i - y_A)}{(x_i - x_A)} \tag{6.1}$$

α_i can be obtained in systems with angle-of-arrival information available, or by using an initial estimate for A. Thus, the standalone GGDOP for point A is defined as [99]:

$$GGDOP_{ASTD} = \left[\frac{\sum_{i=1}^{m} \sum_{(j=1,j\geq i)}^{m} \frac{\sin^2(\alpha_{Ai} - \alpha_{Aj})}{\sigma_i^2 \sigma_j^2}}{(\sum_{i=1}^{m} 1/\sigma_i^2)^2} \right]^{-1} \tag{6.2}$$

GGDOP is referred to as *generalized* because it includes the standard GDOP when the range variances are equal [99]. It follows from this definition that if nodes A and B collaborate, the collaborative GGDOP for node A is defined as $GGDOP_{ACOL}$ such that:

$$GGDOP_{ACOL}^{-1} = \frac{\Omega}{\left(\frac{1}{\sigma_{AB}^2} + \sum_{j=1}^{w} \frac{1}{\sigma_{AAj}^2} \right)^2} \tag{6.3}$$

where w is the number of anchors of known location used to locate node A. In our defined system, $w = 3$.

$$\begin{aligned} \Omega = & \frac{\sin^2(\theta_{AA_1} - \theta_{AA_2})}{\sigma_{AA_1}^2 \sigma_{AA_2}^2} + \frac{\sin^2(\theta_{AA_1} - \theta_{AA_3})}{\sigma_{AA_1}^2 \sigma_{AA_3}^2} + \frac{\sin^2(\theta_{AA_2} - \theta_{AA_3})}{\sigma_{AA_2}^2 \sigma_{AA_3}^2} \\ & + \frac{\sin^2(\theta_{AA_1} - \theta_{AB})}{\sigma_{AA_1}^2 \sigma_{AB}^2} + \frac{\sin^2(\theta_{AA_2} - \theta_{AB})}{\sigma_{AA_2}^2 \sigma_{AB}^2} + \frac{\sin^2(\theta_{AA_3} - \theta_{AB})}{\sigma_{AA_3}^2 \sigma_{AB}^2} \end{aligned} \tag{6.4}$$

The on-the-fly criterion derived to account for geometry is as follows: *if the standalone GGDOP is greater than the collaborative GGDOP, collaborate*. Obviously, this criteria can only be used if angle-of-arrival information is available, or if an initial estimate for A and B are available. A valid initial estimate can be the standalone estimate for A and B. If a non-linear estimator is used, the initial guess of the non-linear estimator for A and B is also a valid candidate to provide an estimate of α_i as discussed in [93].

$$\begin{bmatrix} GGDOP_{CRITA}: & GGDOP_{STDA} \geq GGDOP_{COLA} & \text{Collaborate for A} \\ GGDOP_{CRITB}: & GGDOP_{STDB} \geq GGDOP_{COLB} & \text{Collaborate for B} \end{bmatrix} \tag{6.5}$$

6.1.2 Uncertainty Criterion

In regard to the quality of the collaborative link, there are two reasons why we consider the magnitude of the information gain (reduction in range uncertainty) as a valid component that can tell ahead of time whether collaboration will improve accuracy: 1) it is based on observables, and 2) most importantly, the information gain is a function of σ_{AB} (for Gaussian distributions), which affects the accuracy of the collaborative estimation, as shown in our variational analysis in Chapter 4. Furthermore, by being a measure of uncertainty rather than just a variance, the information gain will be more reliable than the variance when the set of range measurements on the collaborative link has a non-Gaussian distribution.
Recall from Chapter 5 equation (5.13) that the information gain from the acquisition of the collaborative link is a function of the uncertainty $H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))$ of the range measurements on the collaborative link. This information gain was defined as:

$$\begin{aligned} \text{Information Gain} &= |\sum f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B) \log f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B)| \\ &= |H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))|^{-1} \end{aligned} \tag{6.6}$$

We established in equation (5.18) that this information gain is an upper bound to the maximum improvement in position accuracy due to collaborative positioning, if the collaborative positioning estimator reaches its optimal performance. Hence, a 'large information gain' indicates that collaborative positioning has the potential to be more accurate than standalone, if a very good estimator is used. However, just assessing the magnitude of the information gain does not easily provide a way to quantify the qualitative term: 'large information gain'. To assess whether the information gain from collaboration is large enough, we propose a comparative approach. We compare the uncertainty on the collaborative link $H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))$ to the uncertainty of the set of standalone observed range measurements. Since we proved that a small $H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))$ yields a large information gain, the comparative approach should target minimizing $H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B))$. We thus propose the following criteria for on-the-fly assessment of whether it is worth collaborating or not for improved accuracy.

In the criteria defined below, $H(f_{RAB}(r_{AB}|\mathbf{x}_A, \mathbf{x}_B)) = H_{RAB}$, and the different uncertainties of range measurements from the standalone links are represented as follows: H_{RAi} is the uncertainty of the set of observed range measurements between node A and its i^{th} anchor; H_{RBj} is the uncertainty of the set of observed range measurements between node B and its j^{th} anchor. In the criteria represented below, we assume that each node has three anchors in the standalone configuration.

Mean Uncertainty Criteria

We introduce the following criterion (Mean Uncertainty Criterion, or MEANUC): If the uncertainty of the collaborative link is smaller than the average uncertainty from the standalone range measurements, collaborate.

$$\begin{bmatrix} MEANUC_A: & \text{mean}\,[H_{RA1}, H_{RA2}, H_{RA2}] \geq H_{RAB} & \text{Collaborate for A} \\ MEANUC_B: & \text{mean}[H_{RB1}, H_{RB2}, H_{RB3}] \geq H_{RAB} & \text{Collaborate for B} \\ MEANUC_{All}: & \text{mean}[H_{RA1}, H_{RA2}, H_{RA3}, H_{RB1}, H_{RB2}, H_{RB3}] \geq H_{RAB} & \text{Coll. for A and B} \end{bmatrix} \tag{6.7}$$

Max Uncertainty Criteria

Another criterion is to compare the uncertainty of the collaborative link to the maximum uncertainty on all standalone links. If the uncertainty on the collaborative link is smaller than the maximum uncertainty available in the set of standalone range measurements, then collaborate:

$$\begin{bmatrix} MAXUC_A: & \max[H_{RA1}, H_{RA2}, H_{RA3}] \geq H_{RAB} & \text{Collaborate for A} \\ MAXUC_B: & \max[H_{RB1}, H_{RB2}, H_{RB3}] \geq H_{RAB} & \text{Collaborate for B} \\ MAXUC_{All}: & \max[H_{RA1}, H_{RA2}, H_{RA3}, H_{RB1}, H_{RB2}, H_{RB3}] \geq H_{RAB} & \text{Coll. for A and B} \end{bmatrix} \tag{6.8}$$

It is worth noting that the derivation of these criteria does not account for the estimator used, which also contributes to the accuracy of collaborative estimation. In the simulation section, we test the proposed criteria with the recommended estimator for collaborative positioning: the Maximum Likelihood Estimator (MLE), as indicated by [78].

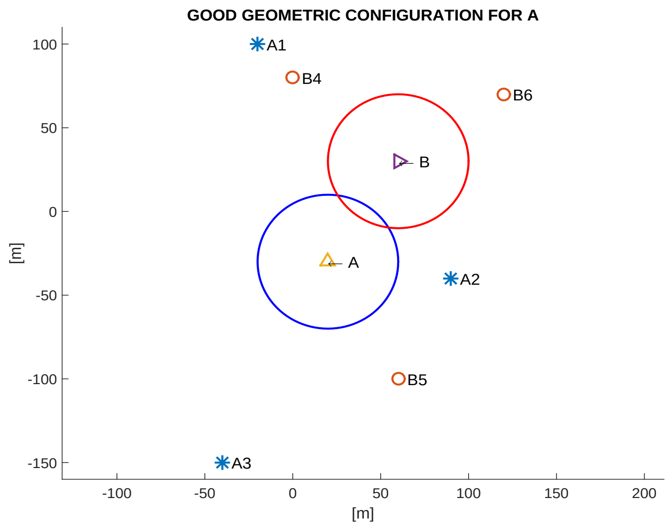
6.2 Simulation Results

6.2.1 Simulation environment

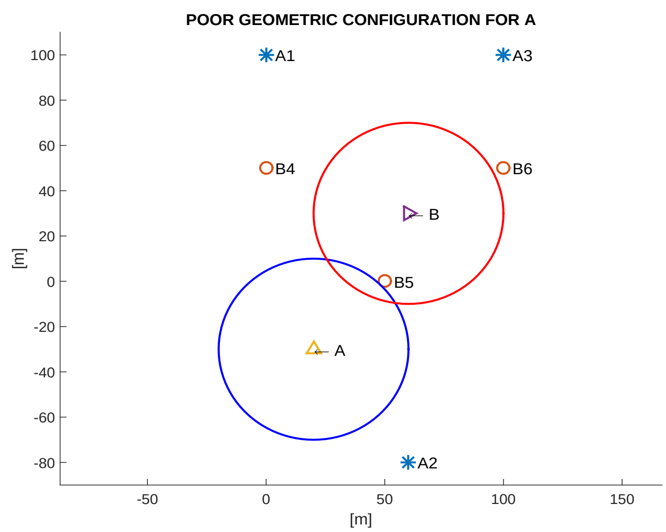
We consider a 2-D synchronized Time of Arrival (TOA) system in a 300mx300m Matlab simulation environment. We test the criteria presented in the previous sections in two configurations that represent a best case scenario and a worst case scenario. The best case scenario includes LOS range conditions on both standalone and collaborative links and a good geometric configuration between the nodes w.r.t their anchors. It is shown on Fig. 6.1(a). The worst case scenario includes NLOS range conditions on both standalone and collaborative links with poor geometric configuration for node A. It is shown on Fig. 6.1(b). In NLOS range conditions, an exponentially-distributed NLOS bias of mean 25 meters is added to both standalone and collaborative range measurements. The use of the exponential distribution to represent NLOS errors was selected because this distribution is commonly used to describe NLOS errors in the literature.

The SNRs on all standalone links are fixed to 20 dB. The SNR on the collaborative link varies from 10 to 40 dB. For each SNR on the collaborative link, we generate 400 test cases by selecting 400 different positions for nodes A and B within the delimiting circles shown on Fig. 6.1. The delimiting circles were used to maintain the desired geometric configurations. For each pair of fixed nodes A and B, 1000 position estimates are derived to assess whether collaborative positioning is more accurate than standalone positioning. Hence, for each SNR on the collaborative link, 400,000 position estimates are analyzed for each node. We use the Newton-Raphson Maximum Likelihood estimator (MLE) to estimate positions for both nodes A and B.

We indicated that the Generalized GDOP requires a true position, or an initial estimate. In these simulations, the Weighted Linear Least Squares (WLLS) estimator is used to generate initial estimates for nodes A and B. The initial estimates used to generate GGDOP are the average of 1000 position estimates from the WLLS estimator in the standalone configuration. In the good geometric configuration, node A is within the convex hull defined by its anchors. In the poor geometric configuration, node A is outside the convex hull. Node B is always within the convex hull defined by its anchors in either configurations. The poor geometric configuration was selected to analyze a case where one node (*i.e*: node A) can gain more from collaboration than node (B) due to its standalone geometric configuration. Fig. 6.1 shows the geometric configuration described above.



(a) Good geometric configuration for both nodes



(b) Poor geometric configuration for node A only

Figure 6.1: Simulation geometric configurations: Overview of test configuration. The circles centered at A and B are delimiting regions containing 400 different positions of A and B, respectively, for 400 different simulation cases. A1, A2, A3 are anchors for A. B4, B5, B6 are anchors for B. (a) Good geometric configuration for A and B with respect to their sensors. LOS is assumed on all range measurements in this configuration. (b) Poor geometric configuration for A with respect to its sensors. NLOS is assumed on all range measurements in this configuration. Exponentially- distributed NLOS errors of mean 25 m are added to all range measurements for both nodes A and B

6.2.2 Performance metric of criteria

We assess the performance of the proposed criteria using the accuracy rate of the criteria. The accuracy rate (labeled 'Accu' in Fig. 6.2-6.6) determines the number of times a criterion was accurate, out of the number of times the criterion recommended to collaborate.

$$\text{Accu} = \frac{\text{Number of times a criteria's recommendation was accurate}}{\text{Number of recommendations from criteria}} \tag{6.9}$$

In Fig. 6.2-6.6, label 'Accu' indicates the accuracy rate (as a percentage) of different criteria. The higher the accuracy rate, the more accurate the criterion is when it recommends to collaborate.

6.2.3 Performance of geometric criteria

In regard to geometry, we assess the performance of the GGDOP criterion from equation (6.5) in the next plot. As discussed earlier, to compute GGDOP on-the-fly requires an initial estimate of each node's location. We use the Weighted Linear Least Squares estimator (WLLS) in the standalone configuration to generate initial estimates for each node's location. In LOS range conditions, in Fig. 6.2 the geometric criterion accuracy ranges between 90% and 100% for both nodes; in NLOS range conditions, the criterion ranges from 65% to 90%. These results indicate that in LOS range conditions, and good geometric configurations, GGDOP can be considered as a reasonably good metric to identify whether collaborative positioning will improve accuracy or not. In NLOS range conditions, the GGDOP criterion may yield unpredictable results.

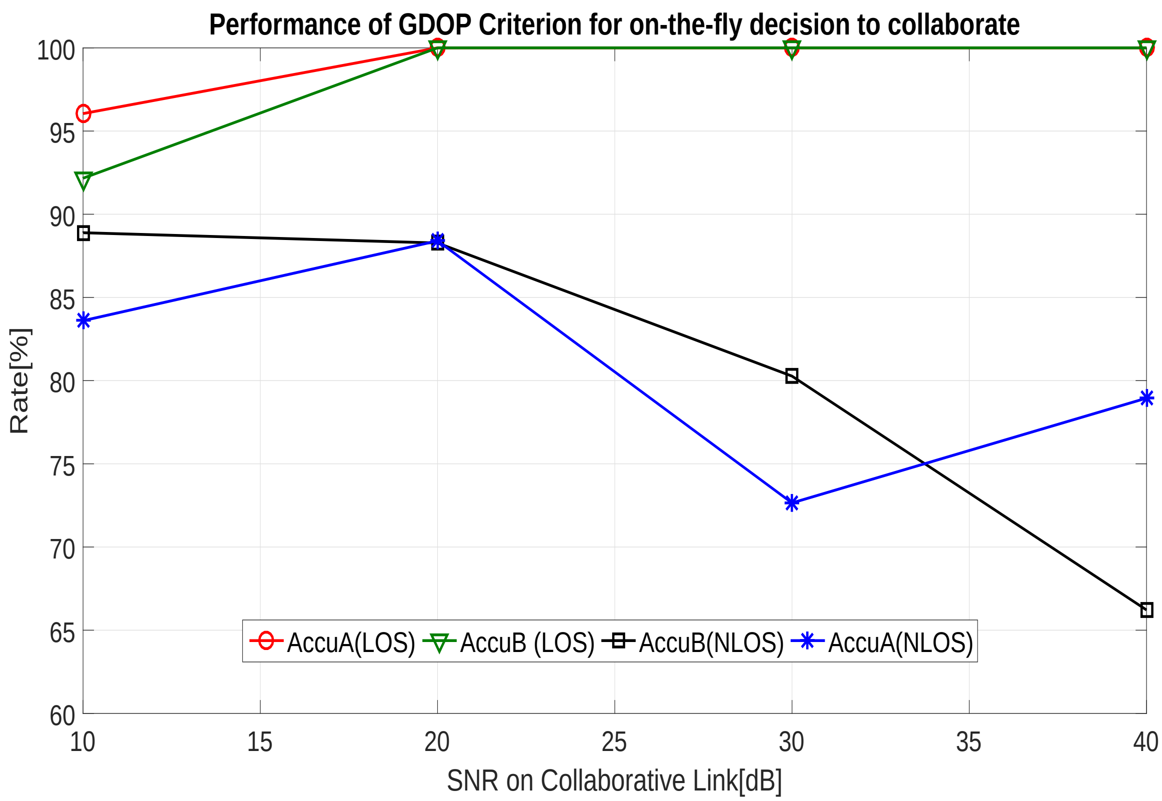


Figure 6.2: Performance of GGDOP criteria in two case scenarios: 1) a best case scenario with LOS range conditions and good geometric conditions per Fig. 6.1(a) and 2) a worst-case scenario with NLOS range conditions and poor geometric conditions for node A per Fig. 6.1(b). *AccuA* and *AccuB* represent the accuracy of criterion $GGDOP_{CRIT}$ for nodes A, and B, respectively.

6.2.4 Performance of uncertainty criteria

In regard to uncertainty, we proposed a set of uncertainty criteria based on a comparison between the uncertainty of the collaborative link range measurements relative to the uncertainty of all other range measurements. The plots from Fig. 6.3 and Fig. 6.5 show the results for the MEAN Uncertainty criteria (MEANUC) in LOS and NLOS range conditions, respectively. The plots from Fig. 6.4 and Fig. 6.6 show the results for the MAX Uncertainty (MAXUC) in LOS and NLOS range conditions, respectively. Differences in performance between the MAX and MEAN criteria are negligible. Hence, either criterion can be used.

In LOS range conditions and good geometric configurations, when the SNR on the collaborative link is equal to or greater than the average SNR of all standalone links (20 dB in this testcase), the uncertainty criteria (both MAX and MEAN) have accuracy rates of at least 90% up to 100% whenever they recommend to collaborate. In NLOS range conditions, the accuracy of both uncertainty criteria depends more on the geometry of the nodes with respect to (w.r.t) their anchors. For node B, which in NLOS range conditions has a good geometric configuration w.r.t its standalone anchors (per Fig. 6.1(b), the uncertainty criteria have accuracy rates between 80% and 90%. For node A which has a poor geometric configuration w.r.t its standalone anchor, the uncertainty criteria accuracy rates vary between 60% and 90%. This implies that the uncertainty criterion yields reasonable results even in NLOS range conditions, if the geometric configuration of the node w.r.t its standalone anchors is good. It is worth noting that it is very likely that the uncertainty criteria does not recommend to collaborate often. We expect that the uncertainty criteria will be making more recommendations at high SNRs, and less recommendations at low SNRs.

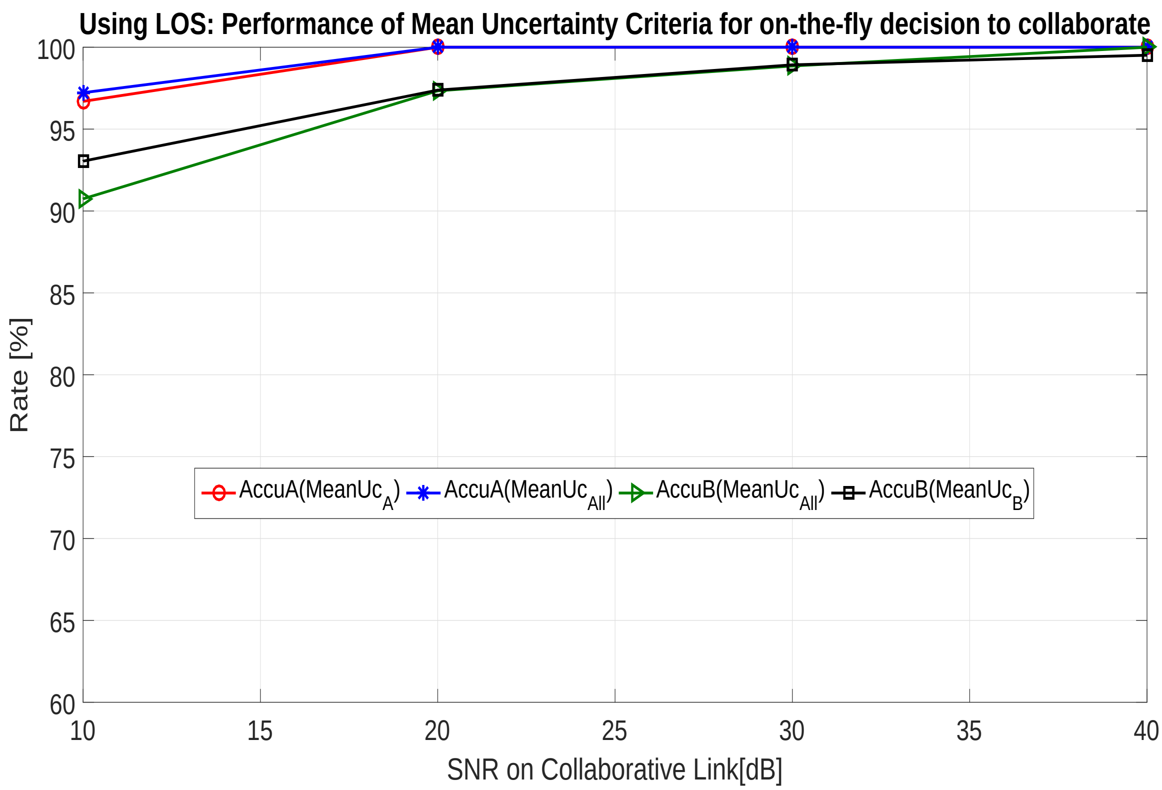


Figure 6.3: Performance of uncertainty criteria in (line-of-sight) LOS range conditions and good geometric configuration per Fig 6.1(a). Performance of MEANUC: $AccuA(MEANUC_A)$ represents the percentage of accuracy of criterion $MEANUC_A$. Similarly, $AccuA(MEANUC_{All})$ represents the accuracy rate for the $MEANUC_{All}$ criterion.

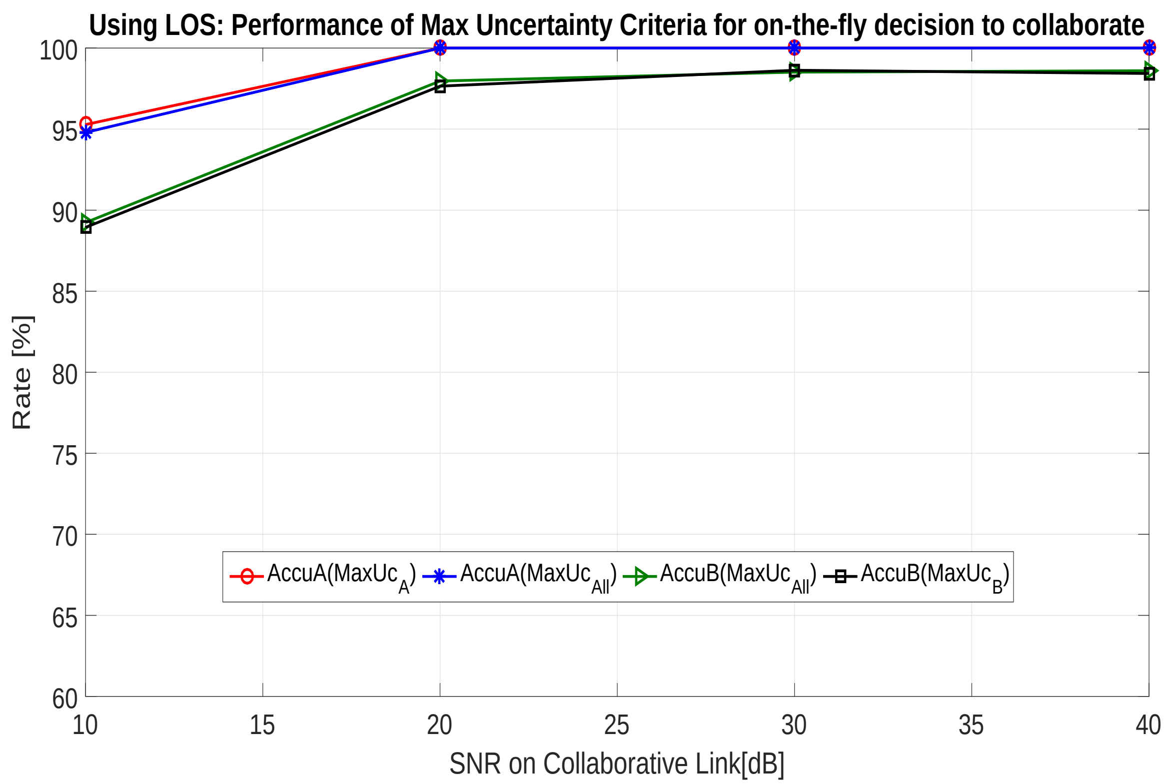


Figure 6.4: Performance of uncertainty criteria in (line-of-sight) LOS range conditions and good geometric configuration per Fig 6.1(a). Performance of MAXUC: $AccuA(MAXUC_A)$ represents the percentage of accuracy of criterion $MAXUC_A$. Similarly, $AccuA(MAXUC_{All})$ represents the accuracy rate for the $MAXUC_{All}$ criterion.

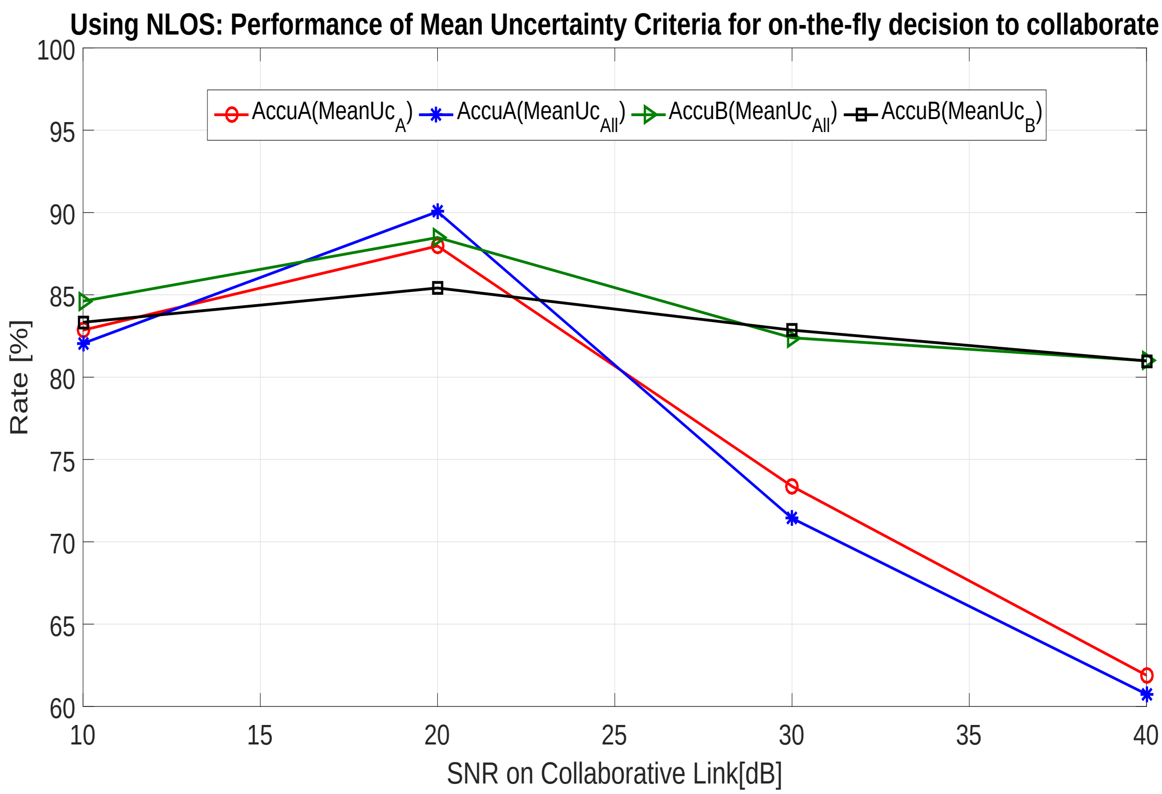


Figure 6.5: Performance of uncertainty criteria in non-line-of-sight (NLOS) range conditions and poor geometric conditions per Fig 6.1(b). Performance of MEANUC: $AccuA(MEANUC_A)$ represents the percentage of accuracy of criterion $MEANUC_A$. Similarly, $AccuA(MEANUC_{All})$ represents the accuracy rate for the $MEANUC_{All}$ criterion. The discontinuities on the plot at zero-dB (for node B) indicate that both criteria did not recommend to collaborate at such a low SNR for any of the simulation cases considered. This means that the uncertainty criteria does not recommend often at low SNRs.

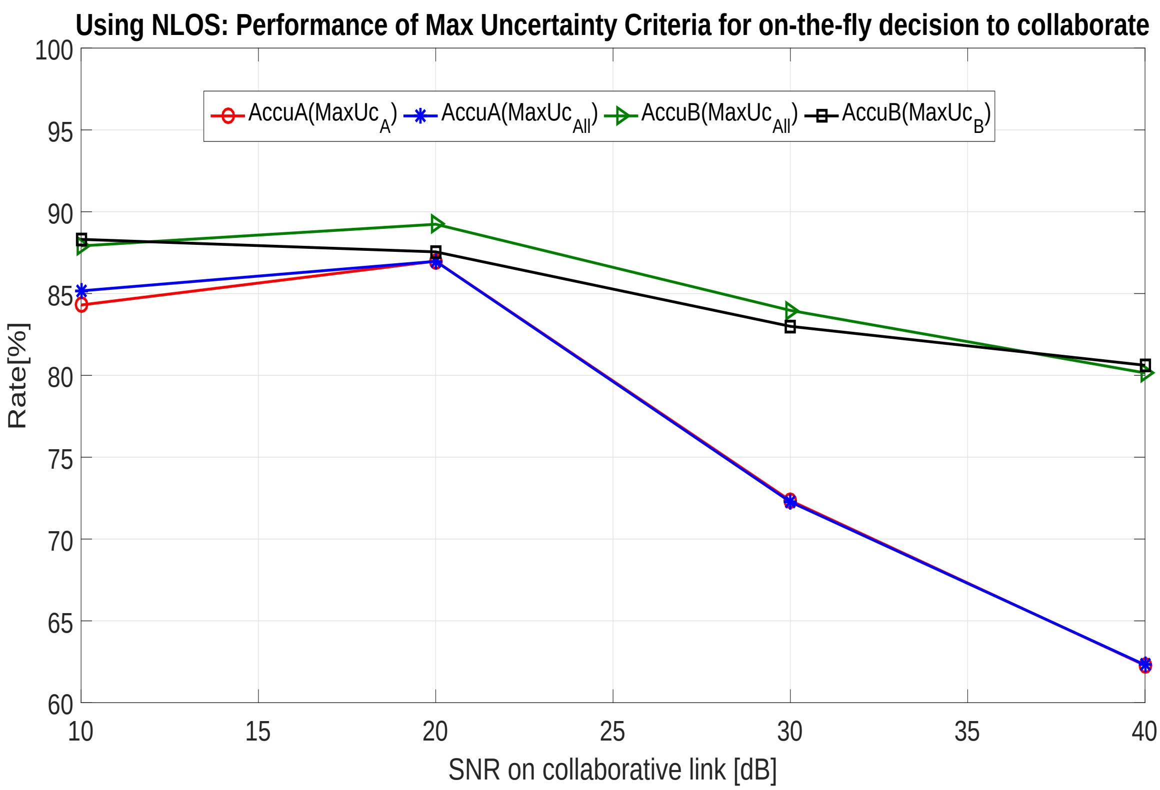


Figure 6.6: Performance of uncertainty criteria in non-line-of-sight (NLOS) range conditions and poor geometric conditions per Fig 6.1(b). Performance of MAXUC: $AccuA(MAXUC_A)$ represents the percentage of accuracy of criterion $MAXUC_A$. Similarly, $AccuA(MAXUC_{All})$ represents the accuracy rate for the $MAXUC_{All}$ criterion.

6.3 Simulation summary and conclusion

In previous chapters we evaluated theoretically factors that affect the improvement in accuracy due to collaboration. In this chapter, we focused on the practical aspect of our work by introducing criteria that tell ahead of time whether collaboration will improve accuracy. The two types of criteria are uncertainty-based and geometrically-based.

From the simulation results in this section, both uncertainty and GGDOP criteria are reasonable metrics to tell ahead of time whether to collaborate or not. In LOS range conditions, when they recommend to collaborate, the two types of criteria are nearly 100 % accurate if the SNR on the collaborative link is above the average SNR from all standalone links. In NLOS range conditions, when the SNR on the collaborative link is above the average SNR from all standalone links, their accuracy is reduced, but still ranging from 60% to nearly 90%. When the geometric configuration of the node to locate w.r.t its standalone anchors is good, the uncertainty criteria yield an accuracy rate between 80% and 90% in NLOS range conditions. Furthermore, regardless of NLOS or LOS range conditions, when the geometric configuration of the nodes w.r.t their anchors is good, the uncertainty criteria yield an accuracy of at least 80% up to 100%. In LOS range conditions, the accuracy rate of the uncertainty criteria is at least 90%, and in NLOS range conditions, it is at least 80%. This implies that the proposed uncertainty criteria are reasonable metrics to consider in both LOS and NLOS range conditions with good geometric configuration.

Chapter 7

Summary and Conclusions

7.1 Summary

The problem of collaborative positioning is a current challenge that needs more attention in both academia and industry. The research community still has more ground-breaking fundamentals to uncover about this positioning concept. Technical challenges that collaborative positioning entails include: a) managing the communication layer between collaborating devices, b) identifying how much data to share, and c) identifying whom to collaborate with. These are very essential questions that need to be answered prior to making this positioning application very useful in numerous critical domains pending its deployment (*e.g*: first responders tracking and unmanned vehicles collaborative actions). It is therefore a topic worthy of continued investigation.

The essential *philosophical* question asked in this research is: 'Is collaborative positioning *always* beneficial? If not, can we theoretically analyze when it is or not beneficial?' One of the major *technical* questions that collaborative positioning entails and that we addressed in this research is 'How can we tell ahead of time whether collaboration will improve accuracy or not?' Analyzing these questions via simulations is not adequate and would not yield conclusions that can be easily adapted to different environments. Hence, in this research, we considered a theoretical approach to address these questions. Our research motivation was to provide the theoretical tools to understand how collaborative positioning works; and from this understanding, assess its performance before and after collaborating. Hence, this research is an analysis of: 1) why collaborative positioning works, 2) factors that impact its performance and make it worse than the traditional standalone positioning, and 3) criteria that can tell ahead of time whether collaboration will improve position accuracy.

The main research challenge consisted of identifying a way to mathematically describe the new information acquired through collaboration and assess its impact on standalone accuracy. Previous research efforts on this topic have mainly focused on the Cramer-Rao Bound. We explained in this research that this approach was limited because it only addressed the *uncertainty* of collaborative positioning and not its *accuracy* directly. This is because the Cramer-Rao Bound is a measure of localization uncertainty and as such only describes the lack in confidence in the set of position estimates acquired. It does not provide direct information on the inaccuracy of collaborative positioning, although, if both standalone and collaborative estimators attain their Cramer-Rao Bounds, the maximum improvement in accuracy can be derived from the reduction in uncertainty due to collaboration. Our research results distinguish themselves from the previous literature by 1) dissociating **uncertainty** from **inaccuracy**, two terms that have been used alternatively without distinction in most applications, and 2) focusing on the accuracy analysis of collaborative positioning.

In this research, we addressed this theoretical problem as follows. First, we defined a theoretical framework that can be used to model the improvement in accuracy due to collaboration. Second, we used this model to analyze factors that vary and conditions that worsen position accuracy in collaboration. Third, we used the results from our theoretical analysis to identify criteria that can tell ahead of time whether collaboration will improve accuracy. Since collaborative positioning is a costly application both in terms of computation power and communication bandwidth, these criteria that tell ahead of time whether to collaborate or not will provide tools to efficiently collaborate.

Although most of the results uncovered during this research were expected (*i.e*: geometry between nodes and between nodes w.r.t their anchors, and the quality on the collaborating nodes are the main drivers to the change in accuracy in collaborative positioning), there were some not-so-obvious but important results as well. First, the improvement in accuracy reaches extrema when the collaborating nodes share a coordinate, under certain conditions. Second, one cannot expect to guarantee a minimum improvement in collaborative positioning accuracy while maintaining the standalone bias. Third, the geometric configuration and the change in bias between standalone and collaborative estimators determine the maximum improvement in accuracy. If both estimators share the same bias, then the geometric configuration of the nodes to locate w.r.t their anchors pre-determines the maximum improvement in accuracy due to collaboration. Finally, the loss term in the improvement in position accuracy is bound by the relative efficiencies of the standalone and collaborative estimators w.r.t the collaborative Cramer-Rao Bound, for Gaussian distributions of position estimates.

We also derived closed-form expressions that can be used to predict the covariance matrix of the collaborative estimator, given the standalone covariance matrix and the collaborative range measurements of one or more collaborating nodes, if a linear estimator is used. Furthermore, we provided a means to determine whether the improvement in accuracy due to collaboration has reached its maximum potential. We also introduced a novel performance metric that is a practical and more robust performance benchmark when in need to select an accurate position estimator. With FCC's E-911 requirements increasing in complexity, our proposed performance metric is worth considering because it accounts well for end-tail behaviors in the distributions of position errors. Contrarily to position root mean squared error (RMSE), which requires pre-processing of the data to remove outliers, our proposed novel metric makes its assessment without the need to pre-process data. It is not only theoretically tractable (like RMSE), but also more robust and more accurate than RMSE in the presence of outliers.

Since collaboration is about acquiring new information, we used an information-theoretic tool, the Kullback-Leibler Divergence, to assess the impact of collaborative positioning on position accuracy. We showed that the acquisition of new information through collaboration is beneficial in terms of uncertainty, but not necessarily in terms on inaccuracy. The analysis results from this research can be summarized in the following five points.

First, we introduced a performance metric based on the Kullback-Leibler Divergence (KLD), that compares the entire statistical distributions of position error between collaborative and standalone estimators, to identify the most accurate estimator. Via simulations, we confirmed that the proposed metric outperforms a comparison of position root mean squared error (RMSE) in the presence of strong heavy tails in the error distributions.

Second, using KLD, we introduced an uncertainty analysis of collaborative positioning by explaining *why* collaboration reduces localization uncertainty. We derived an observable measure of the reduction in uncertainty due to collaboration and indicated that for consistently good estimators (like the Maximum Likelihood Estimator), this observable reduction in uncertainty (or information gain) is an upper bound to the improvement in accuracy due to collaboration.

Third, we validated theoretically and through simulations that the fact that collaboration reduces uncertainty does not guarantee that accuracy is improved. In regard to accuracy, we showed that aside from selecting the right position estimator, an improvement in accuracy due to collaboration is dependent on three main factors: 1) the quality of the collaborative link range measurements relative to the other range measurements, 2) the geometric configuration of the nodes with respect to their anchors, and 3) the angle between the collaborating nodes.

Fourth, we derived theoretically the threshold intervals within which collaboration starts to hurt the accuracy of standalone estimation for scalar unbiased estimators. This analysis indicates that an improvement in accuracy due to collaboration is dependent on 1) the efficiency of the standalone estimator, 2) the geometric configuration of the nodes w.r.t their anchors, and 3) the difference in biases between the standalone and collaborative estimators.

And fifth, we derived two types of criteria to detect on-the-fly and ahead of time whether collaborative estimation will be more accurate than standalone based on: 1) a reduction in Generalized GDOP due to collaboration, and 2) the uncertainty of the collaborative link range measurements relative to all other range measurements. We tested our results in both non-line of sight and line of sight range conditions, with both poor and good geometric configurations.

7.2 Conclusions

The originality of this research lies in the fact that we have been able to use information theory to address an estimation theoretic problem, in order to derive a theory that explains the changes in the accuracy of collaborative positioning w.r.t standalone positioning.

We found that collaboration is certainly beneficial because it creates new information that can increase our confidence in the data at hand. However, collaboration does not guarantee that accuracy in position estimation will improve. We derived factors that make the change in accuracy due to collaboration vary, and also range conditions when collaboration starts to hurt the performance of standalone estimation. We explained that the information gain from the observable range measurements due to collaboration contributes to the maximum increase in localization accuracy due to collaboration. We also explained *why* collaboration reduces localization uncertainty. We determined a formula for the maximum improvement in accuracy due to collaboration under some conditions, defined its bounds and interpreted them. Finally, we selected criteria to tell ahead of time whether collaboration will improve accuracy or not. Hence, we have achieved the goal of this research which was to assess whether collaboration is always beneficial or not. If not, could we determine theoretically when it is, and when it is not?

Potential users that can benefit from these research results are mainly: 1) designers of indoor navigation and GPS-denied positioning systems (*e.g*: for tracking first responders, civilians in a mall, etc...), 2) developers of navigation systems for unmanned vehicles, and 3) designers of spacecraft navigation and control systems for constellations of satellites.

We believe this research contributed to the state-of-the-art by providing 1) the tools and models to better understand how collaborative positioning impacts position accuracy, 2) criteria which identify whether to collaborate or not, thus making the communication between collaborating nodes more efficient, and 3) tools to assess the efficiency of collaborative positioning.

Looking ahead, future research considerations could focus on: 1) deriving a range condition model that assesses when collaboration starts to hurt the accuracy of standalone as more nodes are added, 2) extending the proposed model in this research with skew-elliptical distributions of position estimates as a more generic theoretical model to analyze the change in accuracy due to collaboration, and 3) combining the stochastic-geometric model from Schloemann and Buehrer to the information-theoretic model from this research in order to improve our understanding of the impact of geometry on the accuracy of collaboration when multiple nodes from different geometric configurations are combined.

References

[1] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[2] P. Peyton, *Probability, Statistics, and Random Signal Principles*, McGraw-Hill, NY, USA: 2001.

[3] A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering (3rd Edition)*, Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2008.

[4] Y.-C. Wu, Q. Chaudhari, and E. Serpedin. *Clock synchronization of wireless sensor networks.* Signal Processing Magazine, IEEE 28.1 (2011): 124-138.

[5] B. M. Sadler and A. Swami, *Synchronization in sensor networks: an overview.* IEEE Military Communications Conference, MILCOM 2006.

[6] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* Wiley, 1991.

[7] T.M. Cover and J.A. Thomas, *Entropy, Relative Entropy and Mutual Information*, Elements of Information Theory, pp. 12-49. John Wiley and Sons, 1991.

[8] T.M. Cover and J.A. Thomas, *Inequalities in Information Theory*, Elements of Information Theory, pp. 482-509. John Wiley and Sons, 1991.

[9] R. Fano, *Transmission of information.* MIT Press, Cambridge, MA. 1961.

[10] S. Kullback and R. A. Leibler, *On Information and Sufficiency.* The Annals of Mathematical Statistics 22 (1951), no. 1, 79-86.

[11] Patwari, N.; J. N. Ash, S. Kyperountas, A. O. Hero, R. L Moses, and N. S. Correal, *Locating the nodes: cooperative localization in wireless sensor networks,* Signal Processing Magazine, IEEE , vol.22, no.4, pp.54,69, July 2005.

[12] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correal, and R. J. O'Dea, *Relative location estimation in wireless sensor networks*, IEEE Transactions on signal processing, vol. 51, no. 8, pp. 2137-2148, 2003.

[13] J. Schloemann and R. M. Buehrer, *Using Fisher Information Matrix Summary Statistics to Assess the Value of Collaborative Positioning Opportunities.* IEEE Military Communications Conference, MILCOM 2013.

[14] T. Jia and R. M. Buehrer, *On the optimal performance of collaborative position location,* Wireless Communications, IEEE Transactions on , vol.9, no.1, pp.374,383, January 2010.

[15] S. Venkatesh and R. M. Buehrer, *Power control in UWB position-location networks,* Communications, 2006. ICC'06. IEEE International Conference on. Vol. 9. IEEE, 2006.

[16] S. Venkatesh and R. M. Buehrer, *Multiple-access design for ad-hoc UWB position-location networks,* in Proc. IEEE Wireless Commun. and Networking Conf., pp. 18661873, Apr. 2006.

[17] J. Schloemann and R. M. Buehrer, *On the Value of Collaboration in Non-Line-of-Sight Location Estimation,* Military Communications Conference (MILCOM), 2014 IEEE , vol., no., pp.1060,1065, 6-8 Oct. 2014

[18] J. Schloemann and R. M. Buehrer, *On the Value of Collaboration In Location Estimation,* Vehicular Technology, IEEE Transactions on , vol.PP, no.99

[19] J. Schloemann, H. Dhillon, and R. Buehrer, *A Tractable Analysis of the Improvement in Unique Localizability Through Collaboration,* in IEEE Transactions on Wireless Communications , vol.PP, no.99

[20] J. Schloemann and R. M. Buehrer, *On the value of collaboration in multidimensional location estimation,* Global Communications Conference (GLOBECOM), 2014 IEEE , vol., no., pp.498,504, 8-12 Dec. 2014

[21] T. Jia and R. M. Buehrer, *A new Cramer-Rao lower bound for TOA-based localization,*IEEE Military Communications Conference (MILCOM), vol., no., pp.1,5, 16-19 Nov. 2008

[22] P. Biswas and Y. Ye, *Semidefinite programming for ad hoc wireless sensor network localization*, International Symp. Inf. Process. Sensor Netw., pp.46 -54 2004.

[23] P. Tseng, *Second-order cone programming relaxation of sensor network localization*, SIAM J. Optimization, vol. 18, no. 1, pp.156 -185 2006.

[24] D. Dardari and A. Conti, *A sub-optimal hierarchical maximum likelihood algorithm for collaborative localization in ad-hoc network*, Proc. IEEE SECON 2004, pp.425 -429 2004.

[25] A. A. Kannan , G. Mao, and B. Vucetic, *Simulated annealing based wireless sensor network localization*, J. Computers, vol. 1, no. 2, pp.15 -22 2006.

[26] P.-J. Chuang and C.-P. Wu, *An effective PSO-based node localization scheme for wireless sensor networks*, Proc. Int. Conf. Parallel Distributed Computing, Appl. Tech., pp.187 -194 2008.

[27] K. Strimmer, *Statistical Thinking*, Chapter 4 - 'What is information?", Draft June 13, 2013, Web. < *http : //strimmerlab.org/korbinian.html* >

[28] Z. Szabo, *Information Theoretical Estimators (ITE) Toolbox,* Computer software. Vers. 0.60. N.p., n.d. Web. < *https : //bitbucket.org/szzoli/ite/* >.

[29] A. Fuchs and G. Letta, *L'inegalite de Kullback. Application a la theorie de l'estimation.* Sminaire de probabilites de Strasbourg 4 (1970): 108-131.

[30] Y. Qi and H. Kobayashi. *Cramer-Rao lower bound for geolocation in non-line-of-sight environment.* Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on. Vol. 3. IEEE, 2002.

[31] Y. Qi, H. Kobayashi, and H. Suda, *On Time-of-Arrival Positioning in a Multipath Environment*, IEEE Trans. Vehicular Technology, vol. 55, pp. 1516-1526, Sept. 2006.

[32] S. Gezici, I. Guvenc, and Z. Sahinoglu, *On the performance of linear least-squares estimation in wireless positioning systems.* Communications, 2008. ICC'08. IEEE International Conference on. IEEE, 2008.

[33] D.J. Torrieri, *Statistical Theory of Passive Location Systems.* Aerospace and Electronic Systems, IEEE Transactions on , vol.AES-20, no.2, pp.183,198, March 1984.

[34] J. Fernandez-Villaverde and J. F. Rubio-Ramirez. *Comparing dynamic equilibrium models to data: a Bayesian approach,* Journal of Econometrics 123.1 (2004): 153-187.

[35] M. Lexa, *Useful Facts about the Kullback-Leibler discrimination distance.* None 2004.

[36] J.R. Hershey and P.A. Olsen, *Approximating the Kullback Leibler Divergence Between Gaussian Mixture Models,* Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on , vol.4, no., pp.IV-317,IV-320, 15-20 April 2007.

[37] J.-F. Chamberland and V. V. Veeravalli. *Decentralized detection in sensor networks.* Signal Processing, IEEE Transactions on 51.2 (2003): 407-416.

[38] D. Kazakos and P. Papantoni-Kazakos. *Detection and estimation*, Computer Science Press, 1990.

[39] J. N. Tsitsiklis, *Decentralized detection by a large number of sensors,* Mathematics of Control, Signals and Systems 1.2 (1988): 167-182.

[40] S. Sinanovic and D. Johnson, *Toward a theory of information processing,* Journal of the Franklin Institute, 2004.

[41] B. S. Clarke and A. R. Barron. *Information-theoretic asymptotics of Bayes methods,* Information Theory, IEEE Transactions on 36.3 (1990): 453-471.

[42] K. P. Burnham and D. R. Anderson. *Kullback-Leibler information as a basis for strong inference in ecological studies,* Wildlife research 28.2 (2001): 111-119.

[43] B. Chen, et al. *Parameter identifiability with KullbackLeibler information divergence criterion,* International Journal of Adaptive Control and Signal Processing 23.10 (2009): 940-960.

[44] M. Zakai, *On mutual information, likelihood ratios, and estimation error for the additive Gaussian channel,* Information Theory, IEEE Transactions on 51.9 (2005): 3017-3024.

[45] K. Fang, S. Kotz, and K. W. Ng. *Symmetric multivariate and related distributions,* Vol. 36. Chapman & Hall/CRC, 1990.

[46] R. B. Arellano-Valle, J. E. Contreras-Reyes, and M. G. Genton. *Shannon Entropy and Mutual Information for Multivariate Skew-Elliptical Distributions,* Scandinavian Journal of Statistics 40.1 (2013): 42-62.

[47] J. E. Contreras-Reyes, and R. B. Arellano-Valle. *KullbackLeibler divergence measure for multivariate skew-normal distributions.* Entropy 14.9 (2012): 1606-1626.

[48] T. H. Pham, et al. *Multivariate Mutual Information Measures for Discovering Biological Networks,* RIVF. 2012.

[49] S. Lu, *Measuring dependence via mutual information*, Master's thesis, Queen's University. 2011.

[50] D. Guo, S. Shamai, and S. Verdu, *Mutual information and minimum mean-square error in Gaussian channels,* Information Theory, IEEE Transactions on 51.4 (2005): 1261-1282.

[51] W. McGill, *Multivariate information transmission,* Information Theory, Transactions of the IRE Professional Group on , vol.4, no.4, pp.93,111, September 1954.

[52] T. Van de Cruys, *Two multivariate generalizations of pointwise mutual information,* Proceedings of the Workshop on Distributional Semantics and Compositionality, Association for Computational Linguistics, p.16-20, June 24-24, 2011. Portland, Oregon.

[53] K. W. Church and P. Hanks. *Word association norms, mutual information, and lexicography,* Computational linguistics 16.1 (1990): 22-29.

[54] N. Brunel and J. Nadal, *Mutual information, Fisher information, and population coding,*Neural Computation,10, pp. 17311757, 1998.

[55] S. Laughlin, *A Simple Coding Procedure Enhances a Neuron's Information Capacity,* Zeitschrift fr Naturforschung.Section C: Biosciences 36.9-10 (1981): 910.

[56] A.J. Bell, *The co-information lattice,* Proc. Fourth Intl Symp, Independent Component Analysis and Blind Source Separation, ICA 2003.

[57] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, *A non-reference image fusion metric based on mutual information of image features*, Computers & Electrical Engineering, Volume 37, Issue 5, September 2011, Pages 744-756.

[58] F. Bourgault, A.A. Makarenko, S.B Williams, B. Grocholsky, and H.F. Durrant-Whyte, *Information based adaptive robotic exploration,* Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on , vol.1, no., pp.540,545 vol.1, 2002.

[59] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F.Durrant-Whyte, and M. Csobra. *A solution to the simultaneous localization and map building (slam) problem.* Robotics and Automation, 17(3):229241, 2001.

[60] B. Grocholsky, *Information-Theoretic Control of Mutiple Sensor Platforms*, Dissertation, The University of Sydney, March 2002.

[61] J. Seidler, *Bounds on the mean-square error and the quality of domain decisions based on mutual information,* Information Theory, IEEE Transactions on , vol.17, no.6, pp.655,665, Nov 1971.

[62] N. Prasad and M.K. Varanasi, *Outage Theorems for MIMO Block-Fading Channels,* Information Theory, IEEE Transactions on , vol.52, no.12, pp.5284,5296, Dec. 2006.

[63] M. Ding and S.D. Blostein, *Maximum Mutual Information Design for MIMO Systems With Imperfect Channel Knowledge,* Information Theory, IEEE Transactions on , vol.56, no.10, pp.4793,4801, Oct. 2010.

[64] Y. Zhu, P. Kam, and Y. Xin, *On the Mutual Information Distribution of MIMO Rician Fading Channels,* Communications, IEEE Transactions on , vol.57, no.5, pp.1453,1462, May 2009.

[65] W. Hachem, P. Loubaton, and J. Najim, *The mutual information of a MIMO channel: A survey,* ISCCSP 2006.

[66] R. J. Serfling, *Multivariate symmetry and asymmetry*, Encyclopedia of statistical sciences. 2006.

[67] R. Zekavat and R. M. Buehrer, *Handbook of Position Location: Theory, Practice, and Advances,* Wiley, 2012.

[68] J. E. Contreras-Reyes, *Asymptotic Form Of The KullbackLeibler Divergence For Multivariate Asymmetric Heavy-Tailed Distributions,* Physica A 395.(2014): 200-208.

[69] A. Azzalini and A. Dalla-Valle, *The multivariate skew-normal distribution,* Biometrika, 1996 83: 715-726.

[70] D. M. Kline, and V. L. Berardi. *Revisiting squared-error and cross-entropy functions for training neural network classifiers* Neural Computing & Applications 14.4 (2005): 310-318.

[71] I. J. Taneja, L. Pardo, D. M. Gonzlez, and M. L. Menndez, *On generalized information and divergence measures and their applications: a brief review*, Questiio. 1989, vol. 13, nm. 1-3 (1989).

[72] D. F. Kerridge, *Inaccuracy and inference*, Journal of the Royal Statistical Society. Series B (Methodological) (1961): 184-194.

[73] P. Nath, *Inaccuracy and coding theory* Metrika 13.1 (1968): 123-135.

[74] S. Smitha, *A study on the Kerridge's inaccuracy measure and related concepts,* Dissertation. Cochin University of Science and Technology, 2010.

[75] H. A. K. Rady, *Shannon Entropy and Mean Square Errors for speeding the convergence of Multilayer Neural Networks: A comparative approach*, Egyptian Informatics Journal 12.3 (2011): 197-209.

[76] C. Chang and A. Sahai, *Estimation bounds for localization,* First Annual IEEE Communications Society Conference on Sensor and Ad-Hoc Communications and Networks, (IEEE SECON 2004), pp. 415 424, 4-7 Oct. 2004.

[77] W. L. Brogan, *Improvements and Extensions of the Geometrical Dilution of Precision (GDOP) Concept for Selecting Navigation Measurements,* Nebraska Univ. Lincoln Dept of Electrical Engineering, 1981.

[78] X.R. Li, *Collaborative Localization With Received-Signal Strength in Wireless Sensor Networks,* Vehicular Technology, IEEE Transactions on , vol.56, no.6, pp.3807,3817, Nov. 2007.

[79] X.R. Li and Z. Zhao, *Testing Estimator's Credibility - Part I: Tests for MSE*, Information Fusion, 2006 9th International Conference on , vol., no., pp.1,8, 10-13 July 2006.

[80] X.R. Li and Z. Zhao, *Evaluation of estimation algorithms part I: incomprehensive measures of performance*, Aerospace and Electronic Systems, IEEE Transactions on , vol.42, no.4, pp.1340,1358, October 2006.

[81] X.R. Li and Z. Zhao, *Relative error measures for evaluation of estimation algorithms*, Information Fusion, 2005 8th International Conference on , vol.1, no., pp.8 pp.,, 25-28 July 2005.

[82] X.R. Li, Z. Zhao, and X.B. Li, *Evaluation of Estimation Algorithms: Credibility Tests*, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on , vol.42, no.1, pp.147,163, Jan. 2012.

[83] X. R. Li and Z. Zhao, *Measure of performance for evaluation of estimators and filters*, Proceedings of the 2001 SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, USA, July-August 2001, pp. 530-541.

[84] A. Urruela, J. Sala, and J. Riba, *Average performance analysis of circular and hyperbolic geolocation*, Vehicular Technology, IEEE Transactions on, On page(s): 52 - 66 Volume: 55, Issue: 1, Jan. 2006.

[85] F. Yin, C. Fritsche, F. Gustafsson, and A. M. Zoubir, *TOA-Based Robust Wireless Geolocation and Cramr-Rao Lower Bound Analysis in Harsh LOS/NLOS Environments*, Signal Processing, IEEE Transactions on , vol.61, no.9, pp.2243,2255, May1, 2013.

[86] K. Yu, *3-D Localization Error Analysis in Wireless Networks*, IEEE Trans. Wireless Comm., vol. 6, no. 10, pp. 3473-3481, Oct. 2007.

[87] H. Miao, K. Yu, and M. Juntti *Positioning for NLOS propagation algorithm derivations and CramerRao bounds*, IEEE Trans. Veh. Technol., vol. 56, no. 5, pp.2568-2580 2007.

[88] M. J. Schervish, *Theory of statistics,* Springer Science Business Media, 1995.

[89] J. K. Lindsey, *Parametric statistical inference,* Oxford University Press, 1981.

[90] C. Wang, *Kullback-Leibler Information and its applications in multi-dimensional adaptive testing,*Psychometrika 76.1 (2011): 13-39.

[91] Y. Chang, *A global information approach to computerized Adaptive Testing,* Applied Psychological Measurement 20.3 (1996): 213-229.

[92] M. H. Degroot, and S. E. Fienberg, *The comparison and evaluation of forecasters,* The Statistician, 32, 14-22, 1983.

[93] V.S. Chernyak, *Fundamentals of Multisite Radar Systems: Multistatic Radars and Multistatic Radar Systems,* Gordon and Breach Science Publishers, OPA, 1998.

[94] A. Dembo, T. M. Cover, and J. A. Thomas, *Information theoretic inequalities,* in IEEE Transactions on Information Theory, vol. 37, no. 6, pp. 1501-1518, Nov 1991.

[95] D.V. Lindley, *The Annals of Mathematical Statistics,* Vol. 27, No. 4 (Dec., 1956), pp. 986-1005.

[96] A. Peter and A. Rangarajan, *Shape analysis using the Fisher-Rao Riemannian metric: unifying shape representation and deformation,* Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on, Arlington, VA, 2006, pp. 1164-1167.

[97] R. Hopkins, *Contemporary and Emerging Inertial Sensor Technologies,* ION PLANS 2016 - Tutorial, Draper Labs, April 2016.

[98] S. Bell, *A beginner's guide to uncertainty of measurement,* 2001.

[99] S. Venkatesh, *The Design and Modeling of Ultra-Wideband Position-Location Networks*, Dissertation, Virginia Tech, 2007.