Geometric Autoconfiguration for Precision Personnel Location

by

Benjamin W. Woodacre

A Dissertation
Submitted to the Faculty
of the
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy
in
Electrical and Computer Engineering
by

_____

May 7, 2010

APPROVED:

_____

Professor David Cyganski, Major Advisor

_____

Professor R. James Duckworth

_____

Professor Arthur C. Heinricher Jr.

Abstract

The goal in a location system is knowledge, to within a desired accuracy, of the position of a mobile user based on signals propagated between that user and fixed stations. For example, in emergency response situations such information would assist search and rescue operations and provide improved situational awareness. In general, location estimation is possible given the signal measured at, and the position of, each receiver. In the case of a location system where such receivers are installed on vehicles, such as for fire trucks, no external infrastructure or prior characterization of the area of operations can be assumed and the estimation of the (relative) positions of the receiving stations must be repeated each time the system is deployed at a new site presulting in the geometry of the receiving antennas being changed.

This dissertation presents work towards an accurate and automatic method for determination of the geometric configuration of such receiving stations based on sampled frequency data using both a "classical" ranging method and a novel technique based on a singular value decomposition method for multilateralization. We compare the performance of our approaches to the Cramer-Rao bound for antenna geometry error for distance and frequency-data based geometry estimators, and provide experimental performance results for these methods tested in real multipath envirionments.

Acknowledgements

First and foremost, I thank my parents and my two sisters for their love and support. Without them, none of this would have been possible.

I would like to thank my advisor Professor David Cyganski for his patience and willingness to put up with me throughout the process of completing this dissertation, and for recognizing the thesis potential in GAC and assigning it to me.

Thanks to Professors R. James Duckworth and Arthur Heinricher for sitting on my committee and enduring the deluge of printed thesis copy, a number of schedule changes, and for sitting through my presentations.

I also would like to thank Prof. Fred Looft for assigning me to a teaching assistant position at the outset of my graduate studies, which has led to me completing (pending committee approval, of course!) this dissertation.

I acknowledge all the other grad students I have come to know in the Machine Vision Laboratory, listed roughly in order of appearance: David Holl (who taught me to read the man pages), Nick Hatch, Pavan Reddy, Dan Breen, Nick Sherwood (who used an extremely angry photo of me for an image processing project), Darius Kazemi (for his self-defense skills), Ryan Angilly, Jason Huang, Jay Farmer, Vincent Amendolare (for introducing me to LOST), Hemish Parikh (for being extremely relaxed), Jack Coyne, Hauke Dampfling, Bob Boisse (the super tech), Vivek Varshney (for off-color comments), Matt Campbell (for song improvisation), Jorge Alejandro, Matt Lowe (for causing most of us to be afflicted with *cube vision*), Andrew Cavanaugh, and Jamie Mitchell. If it wasn't for all these folks, or Jorge Cham (of `phdcomics.com`), grad school would have been much less enjoyable.

While I still have space, I acknowledge the software that made this thesis possible: MATLAB, Maple, Debian GNU/Linux, vim, TeX Live, GNU screen, Firefox, Amarok, and The Who.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this research, we set out to investigate signal processing methods and performance bounds for estimating the geometry of the positioning antennas used in a precision location system. (Here, precision refers to meter-level position accuracy.) Knowledge of the location system's antenna geometry, and measurements of the signals from a mobile transmitter as received at these antennas are the the two pieces of information needed to estimate the transmitter's position.

While some systems exist which can locate individuals within buildings and their position within interior rooms, many are fixed infrastructure systems which rely on incorporating sensors into the building itself. This is clearly unsuitable for firefighters, whose itinerant duties will not always bring them to places where such a system is installed, in working order, and not under threat of fire. Thus a positioning system for firefighters must be portable and deployed with minimal user effort.

The WPI Precision Personnel Location (PPL) system is one such (prototype) precision location system. It is a software radio-based architecture for the location and tracking of first responders in indoor environments, which is designed to require no pre-installed infrastructure and minimal setup by users. Tracking of users requires personnel each carry a transmitter emitting a multi-carrier wideband (MC-WB) signal which is sensed at receiving antennas which are fixed upon emergency response vehicles. Upon arrival at an incident, the receiving stations form an ad-hoc network and estimate their sensor (antenna) position

geometry (the subject of this thesis), thereby establishing a local coordinate system. Given this geometry, and measurements of transmitter signal data as captured at each receiver, time-difference-of-arrival techniques may be employed to estimate position.

The central challenge for RF-based indoor positioning is mitigation of the effects of building structure and contents upon location performance. Such interfering structures affect the signal measured at each receiver by adding significant amounts of reflected (multipath) signals and direct path signal attenuation. Multipath signals, which are not useful for positioning, interfere with the direct path signal component that is used for positioning; structures also cause signal attenuation which reduces signal-to-noise ratio and thus location performance. Further, the character of these effects varies with the point of signal reception as well as the position and orientation of a transmitter within the building.

Existing RF technologies for positioning include those based on GPS, which use signals from orbiting satellites or ground-based pseudolites, and those which use ultrawide bandwidth signals. Unfortunately such technologies are simply not robust in the indoor environment. Often, GPS satellite signals are simply too weak to be received indoors; even with repeating local pseudolites [1] which provide sufficient signal strength, multipath signals induce unacceptable levels of error. Ultrawideband approaches, which enjoy precise timing due to large signal bandwidths, are likely more robust in the presence of multipath, but are handicapped to detect attenuated direct path signals due to regulatory limitations on transmitted signal power and resulting long signal integration times.

The deficiencies of these technologies for indoor location necessitated the development, for the PPL system, of the multi-carrier wideband (MC-WB) signal structure. By the use of multiple discrete carriers over a large bandwidth, signal power may be concentrated in those carriers to thwart attenuation, and multipath effects may be mitigated by frequency diversity and real-time algorithms in software.

To date, field testing of the PPL system has relied on manual surveying to establish the local coordinate system and measurement of the receiving antennas' positions. This thesis investigates automatic, non-surveyed methods to determine antenna geometry and performance bounds for this estimation process, termed *geometric auto-configuration* (GAC), which is a prerequisite to any location solution in a practical deployment situation.

The dissertation is organized as follows. Chapter 2 presents background on the PPL system architecture, the multipath signal model, and signal processing methods for precision location. Chapter 3 presents the problem of Geometric Auto Configuration within the context of the PPL system, defines necessary nomenclature, and discusses two solution approaches.

Chapter 4 develops the Cramer-Rao bound (CRB) for GAC, which provides a lower limit on the achievable performance of any estimator for antenna location. The CRB techniques and theory from related literature are presented, and a significant deficiency in the method of constraints, as a novel contribution, is identified, characterized, and repaired. This repair is validated by a differently-constructed form of the bound which generates identical results.

Finally, in Chapter 5 the results of two methods devised for the solution of GAC are presented. The methods are applied to experimental data collected during system field tests, showing the accuracy of geometry estimation and the resulting transmitter location errors when using GAC geometry solutions; the methods are also applied using simulated data, and performance comparison with the CRB is made.

Additionally, this dissertation uses many abbreviations and notations, and defines many of them only once. For the reader's convenience, Tables 1.1 and 1.2 list the frequently-abused abbreviations and notations, respectively.

CRB	Cramer-Rao bound
DFT	Discrete Fourier transform
DSS	Direct state-space
EVD	Eigenvalue decomposition
FFT	Fast Fourier transform
FIM	Fisher information matrix
FPGA	Field-programmable gate array
GAC	Geometric auto configuration
KB	Kilobyte, i.e. $2^{10} = 1024$ bytes
MB	Megabyte, i.e. $2^{20} = 1024^2$ bytes
MDS	Multidimensional scaling
MC-WB	Multi-carrier wideband
PPL	Precision personnel locator
RF	Radio frequency
RMS	Root mean square
SART	Singular-value array reconciliation tomography
SNR	Signal-to-noise ratio
SVD	Singular value decomposition
TOA	Time of arrival
TDOA	Time-difference of arrival
UHF	Ultra-high frequency (300–1000 MHz)
UWB	Ultra-wide bandwidth
WPI	Worcester Polytechnic Institute

Table 1.1: Abbreviations used in this dissertation.

j	Square root of -1, $j^2 = -1$
a	Scalar or vector quantity
$\hat{a}$	a is an estimate
$\tilde{a}$	a is a random variable
$a_\star$	True value of variable a
$\vec{a}, \mathbf{a}$	Vector (usually a *column* vector)
$\vec{1}, \mathbf{1}$	Column vector with all entries equal to 1
A	Matrix
$A \in \mathbf{C}^{m \times n}$	A is a complex-valued matrix with m rows and n columns
$A \in \mathbf{R}^{m \times n}$	A is a real-valued matrix with m rows and n columns
a_{ij}	A single entry of matrix A at the ith row and jth column
$\mathrm{Re}(A)$	Real-valued component of A, e.g., $\mathrm{Re}(a + jb) = a$
$\mathrm{Im}(A)$	Imaginary-valued component of A, e.g., $\mathrm{Im}(a + jb) = b$
A^*	Complex conjugate of A, ($A^* = \mathrm{Re}(A) - j\mathrm{Im}(A)$)
A^T	Transpose of A, (if $B = A^T$, $b_{ij} = a_{ji}$)
A^H	Hermitian of A, (if $B = A^H$, $b_{ij} = a_{ji}^*$)
$A^\dagger$	Moore-Penrose pseudo-inverse of A [2, p. 257]
I	Identity matrix
I_m	Identity matrix having m rows and m columns
$A \circ B$	Entry-wise matrix product, $(A \circ B)_{ij} = a_{ij}b_{ij}$
$A \stackrel{\text{svd}}{\Rightarrow} U\Sigma V^H$	Singular value decomposition of A into the product $U\Sigma V^H$, $U^H U = I$, $V^H V = I$, $\Sigma = \mathrm{diag}(\sigma_i)$, $\sigma_i \geq 0$, $\sigma_i \geq \sigma_{i+1} \geq \cdots \geq 0$
$A \stackrel{\text{eig}}{\Rightarrow} V\Lambda V^{-1}$	Eigenvalue decomposition of A into the product $V\Lambda V^{-1}$, $V^{-1}V = VV^{-1} = I$, $\Lambda = \mathrm{diag}(\lambda_i)$, $\lambda_i \in \mathbf{C}$
$\mathrm{expm}(A)$	Matrix exponential of A, $\mathrm{expm}(A) = e^A = Ve^\Lambda V^{-1}$
$\mathrm{exp}(A)$	Elementwise exponential of A, $\exp(A) = \exp(a_{ij})$
$\ln(a)$	Natural (base e) logarithm of a
$\mathrm{diag}(A)$	Retrieve the diagonal elements of A as a column vector
$\mathrm{diag}(\mathbf{a})$	Form a diagonal matrix from the elements of vector $\mathbf{a}$
$\mathrm{vec}(A)$	Stack the columns of A into a single column vector
$|a|$	Absolute value of a; $|a| = \sqrt{a \cdot a^*}$
$\|\mathbf{a}\|_2$	Euclidean length, or 2-norm of $\mathbf{a}$; $\|\mathbf{a}\|_2 = \sqrt{\mathbf{a}^H \mathbf{a}} = \sqrt{\sum_i |a_i|^2}$
$\|A\|_F$	Frobenius norm of A, $\|A\|_F = \|\mathrm{vec}(A)\|_2$, for vectors: $\|a\|_2 = \|a\|_F$
$\mathcal{F}$	Fourier transform operator

Table 1.2: Mathematical notation.

Chapter 2

Signal Processing for Precision Location

This chapter discusses background material relevant to the understanding of the signal processing methods applied in later chapters. In the following sections,

- the architecture of the PPL system is described to highlight the existing hardware capabilities and field testing results,
- the Direct State Space approach to ranging is presented and difficulties in practice are addressed together with the abandoned DSS+TDOA location method, and
- the SART algorithm is presented in the context of the precision location problem and its invariance to certain system asynchronies which are imposed by the reality of a realizable system is described.

2.1   PPL System Architecture

On December 4, 1999, six firefighters lost their lives in a fire at the abandoned Worcester Cold Storage warehouse. Two firefighters who had gained roof access descended into the interior of the building to search for occupants and became lost in the windowless structure. As their situation deteriorated, two teams of two firefighters also became lost searching for the first team. Despite being in radio contact, all six firefighters perished. The difficulty of

navigating the interior of the large windowless warehouse, during accelerating fire conditions, played a significant role in their deaths.

The U.S. Fire Administration's report [3] on the Cold Storage fire states that "...without good radio reports, the exact movements and locations of crews are uncertain at best." Yet, in the presence of functional radio communications, self-reported positions may be inaccurate due to confusion or disorientation caused by lack of visibility, the stress of fire operations, and irregular, "organic" building characteristics. It is clear that objective position knowledge for firefighters could reduce the number of fatalities by accelerating search-and-rescue (including self-rescue) operations[1].

The WPI PPL system is designed with this firefighter rescue scenario in mind; the goals are to track

- multiple responders,
- in three dimensions,
- to within one foot accuracy,
- requiring no pre-installed infrastructure,
- and minimal setup.

Figure 2.1 illustrates the target scenario. The system consists of a transmitter worn by individuals to be tracked, and multiple receivers consisting of antennas and receiver electronics mounted on fire trucks. Upon arrival at the scene, the receivers form an ad-hoc network and establish their positions within a local coordinate system (which is the subject of this dissertation). Each receiver senses the signal from each transmitter and relays those measurements to a central location, where a joint solution for each transmitter's position is generated. As the firefighters move throughout the building, their accumulated position estimates viewed on a commander's display provide valuable information on building layout and location history for the purposes of egress and search-and-rescue.

The transmitters worn by individuals to be tracked continuously broadcast a multi-carrier wideband signal consisting of a span of unmodulated *sub-carriers* at known frequencies. In this context, wideband refers to the total span of frequencies occupied rather than fractional

---

[1]The leading cause of on-duty firefighter death is cardiac-related distress [4]. Physiological monitoring capabilities should also be a component of any such location technology.
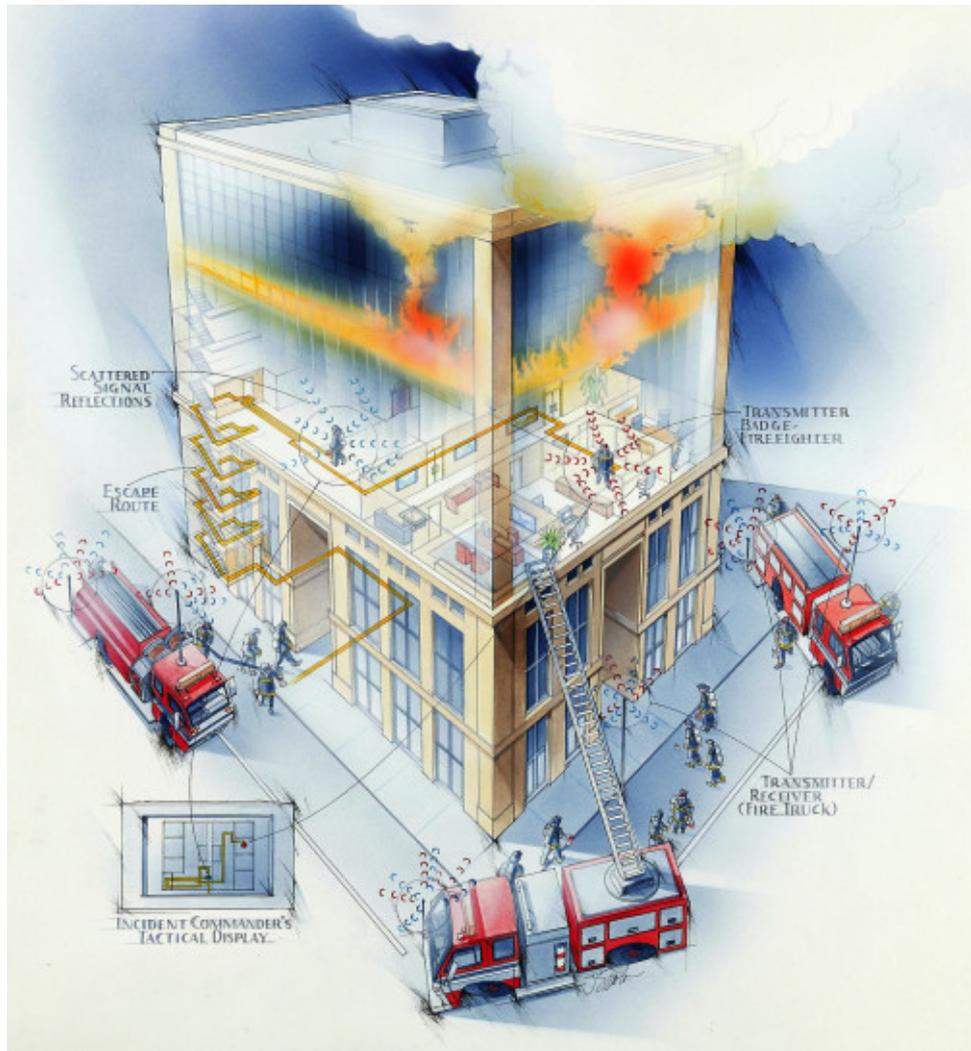
Figure 2.1: PPL concept illustration
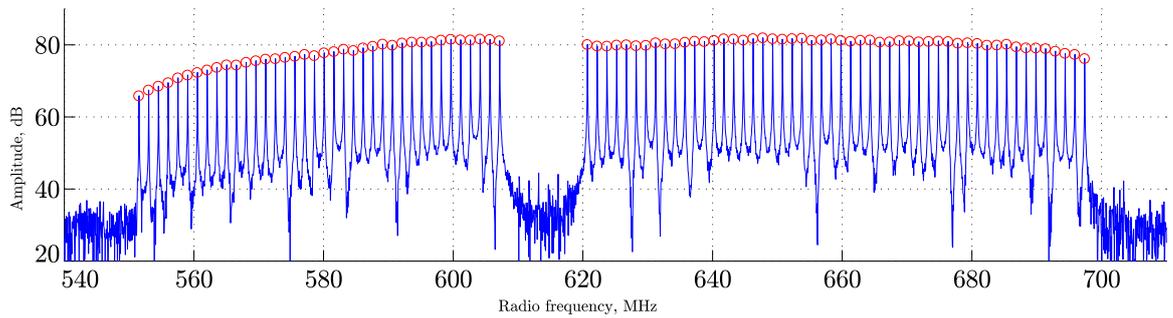
(a) Pristine channel response
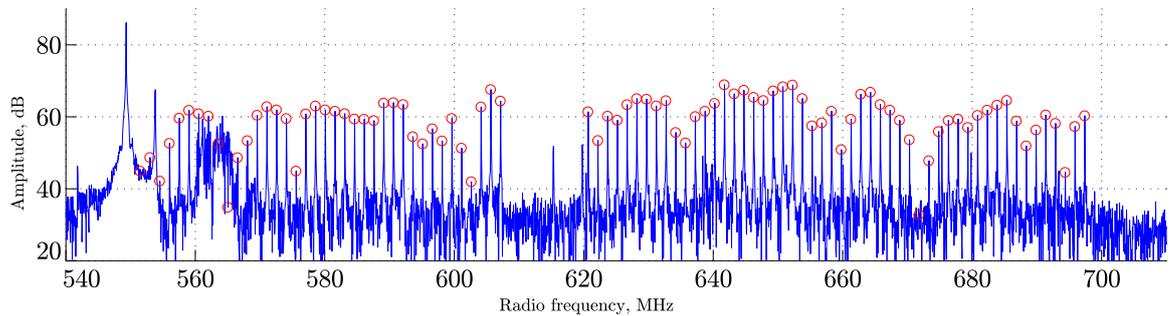
(b) Realistic channel response

Figure 2.2: Frequency spectra of two MC-WB signals. (a) is the spectrum of a signal propagated over a coaxial cable, (b) is the spectrum resulting from a signal propagated between two antennas in an indoor environment. The signals were captured on January 18, 2008 at WPI.

occupancy of the band. Figure 2.2 depicts two example MC-WB signals' spectral content as a function of radio frequency (with a gap from 608-620 MHz; a prohibited band reserved for emergency communications); the red circles indicate known sub-carrier locations. It can be seen that while a signal sent over a coaxial cable has a nearly flat magnitude response over 150 MHz, a signal propagated between two antennas over an indoor channel is subject to frequency-selective fading due to multipath and the reception of interfering signals; in this case, two observed interferers are located at 550 MHz (an analog television signal) and 565 MHz (a digital television signal).

As the transmitter will be carried on firefighters' persons, it is a simple beacon that contains enough digital and RF electronics to repeatedly broadcast the waveform encoded into its memory. Specifically, the hardware consists of a field-programmable gate array

(FPGA) which plays out the waveform through a digital-to-analog converter driving an RF amplification stage [5, 6] which feeds antennas integrated into the firefighter's turnout gear. Multiple transmitters may be supported via time- or frequency-division multiplexing. The receiver hardware stack is the reverse of the transmitter, consisting of an antenna, single sideband RF downmixer and filters, and an analog-to-digital converter controlled by an FPGA. The received signal is stored in onboard memory. The design of the RF portion of the transmitter and receiver to minimize impact on location performance is covered in detail in [6]. The FPGAs at each end of the receive/transmit chains introduce a significant software-radio component to the architecture of the system. At the receiver, the FPGAs accumulate fixed-length contiguous blocks of ADC samples, or *symbols* into memory, which are then relayed to a PC running Matlab, where signal processing and location estimation is done.

2.2   Time- and Time-difference of arrival

In this section, we present the basic hardware configurations necessary for obtaining signals which can be used for time-of-arrival and time-difference-of-arrival estimation. The terms *time of arrival*, *range*, and *delay* are used interchangeably to refer to the separation between a transmitter and receiver.

Figure 2.3 shows a hardware block diagram for simple delay estimation between a synchronized receiver and transmitter. The transmitter, consisting of a simple digital-to-analog and RF amplification chain, transmits $x(t)$ which propagates d_0 meters at the speed of light c, and arrives at the receiver $t_0 = d_0/c$ seconds later. Processing of the received signal may then make an estimate of the time delay between transmission and reception of the signal.

While the scheme in Figure 2.3 to estimate propagation delay, or for *ranging*, works for a system which transmits and receives its own signal such as radar, it does not work in the location scenario as the transmitter must be untethered, and thus cannot be synchronized to the receiver. When the synchronies are dropped as shown in Figure 2.4, the delay estimates proceed as before but contain an unknown time offset $\tilde{t}$. Ranging between an unsynchronized
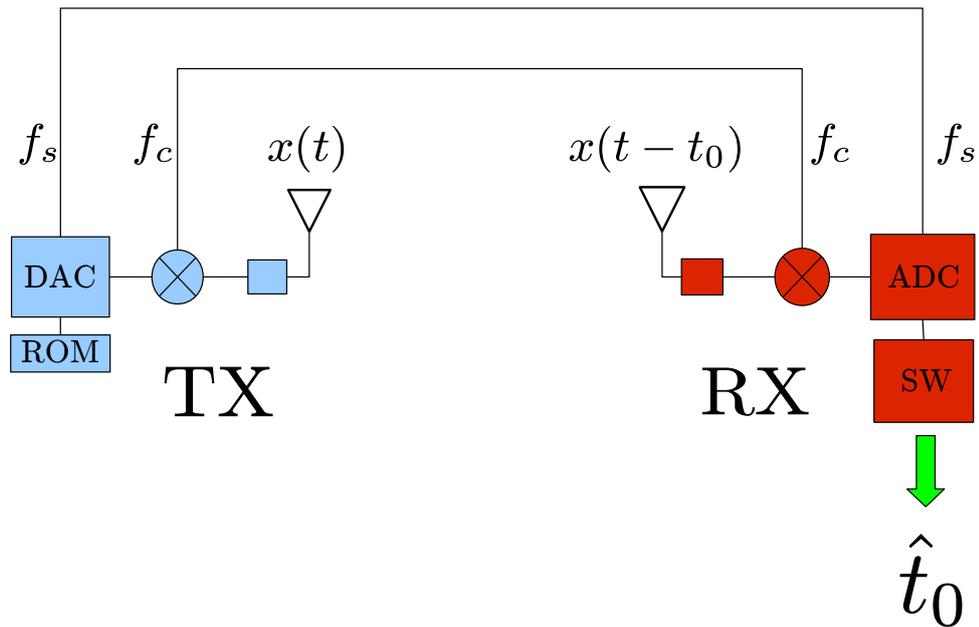
Figure 2.3: Time of arrival estimation, synchronized case. The signal $x(t)$ is generated internally by a DAC reading new samples from a ROM at clock rate f_s, and then upconverted to carrier frequency f_c and transmitted via its antenna. The signal arrives at the receiver t_0 seconds later, having propagated through the channel; there it is downconverted and sampled using an f_s and f_c perfectly synchronized with the transmitter. The samples processed by software to produce delay estimate $\hat{t}_0$.

transmit/receive pair thus requires two problems to be solved to the accuracy desired of the location system: estimation of the time offset $\tilde{t}$, and an estimate of the delay-plus-offset $t_0 + \tilde{t}$. Subtraction of the two results in a delay estimate:

$$\hat{t}_0 = (\hat{t_0 + \tilde{t}}) - \hat{\tilde{t}} \tag{2.1}$$

As we are interested in achieving location accuracy on the order of one to a few feet, nanosecond-level synchronization and delay estimation accuracies are necessary. Achieving synchronization implies that both clocks "tick" at the same rate and read the same value of time at a particular instant. In the case of repeatedly transmitting a symbol N samples long at f_s Hertz, the requirement may be relaxed such that the value of time must be correct modulo $N \cdot f_s$ seconds.

To avoid the transmitter-receiver synchronization problem, another receiver is introduced
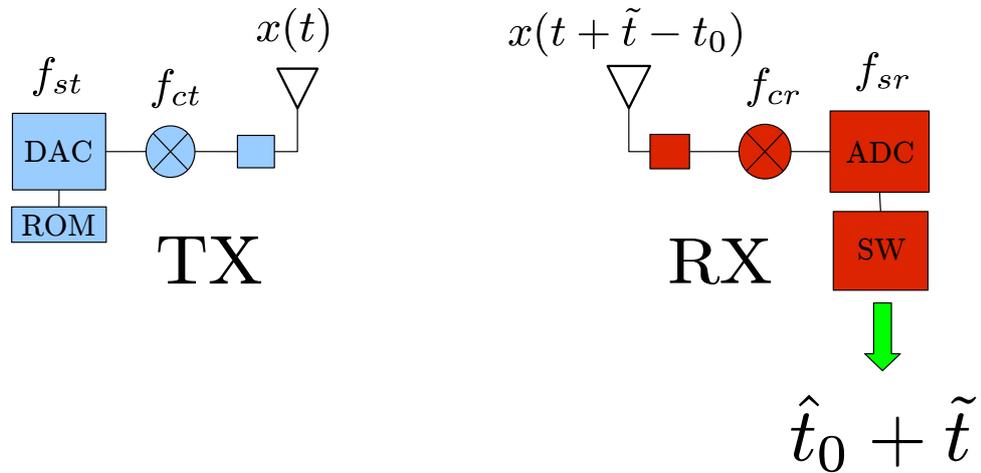
Figure 2.4: Time of arrival estimation, unsynchronized case. Here, the scheme Fig. 2.3, has been modified such that transmitter and receiver no longer share the same signals driving f_s and f_c, and so are given different names. Because the clocks f_{st} and f_{sr} will now vary independently in frequency, the resulting delay estimate at the receiver will have a random time offset component.
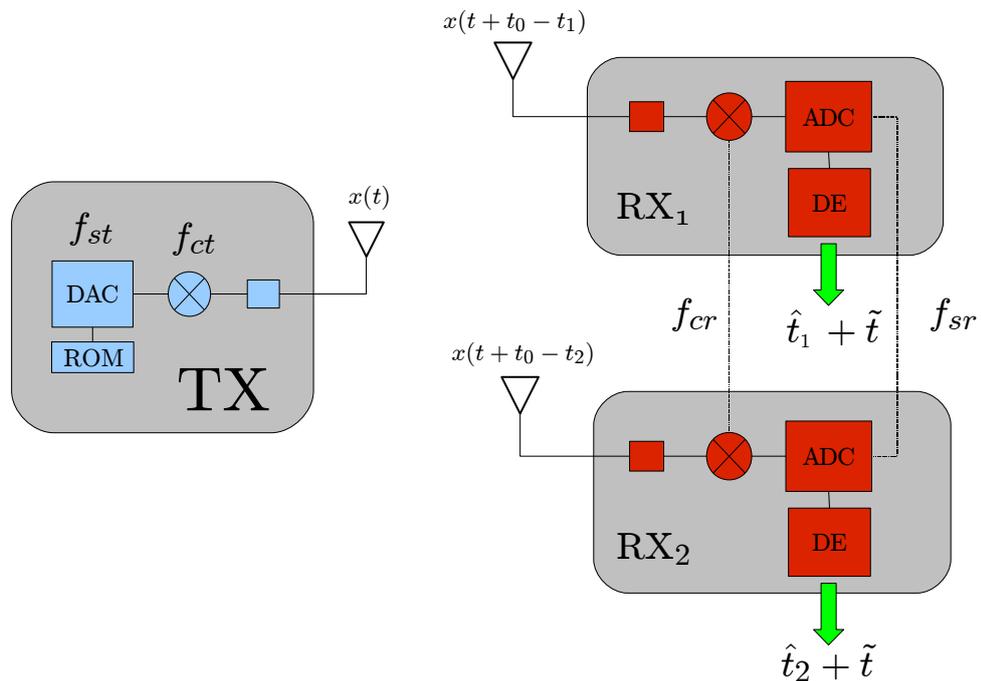
Figure 2.5: Time-difference of arrival estimation with two synchronized receivers. If two receivers share a sample clock f_{sr} and oscillator frequency f_{cr}, then due to the shared clock their delay estimates will also share the same random time offset. By subtracting the two delay estimates, we may obtain an uncorrupted estimate for time-difference of arrival.

as in Figure 2.5 which shares a digital clock and local oscillator (LO). Sharing a clock signal and commencing sampling at the same instant allows two delay estimates to be made; as the estimates have identical, unknown time offsets, differencing them eliminates the shared unknown resulting in the time-difference of arrival between the two receivers, that is

$$\Delta t_{12} = t_1 - t_2 = t_1 + \tilde{t} - (t_2 + \tilde{t}) \tag{2.2}$$

This approach of synchronizing the receivers to eliminate the transmitter offset allows location of the transmitter using TDOA methods.

2.3   Multipath Signal Model

This section introduces the notation used to describe our received signals and the physical and mathematical relationships created by the multipath channel that are exploited by signal processing presented later. Also the basics of time-of-arrival and time-difference-of-arrival and the hardware realities such methods reflect are presented.
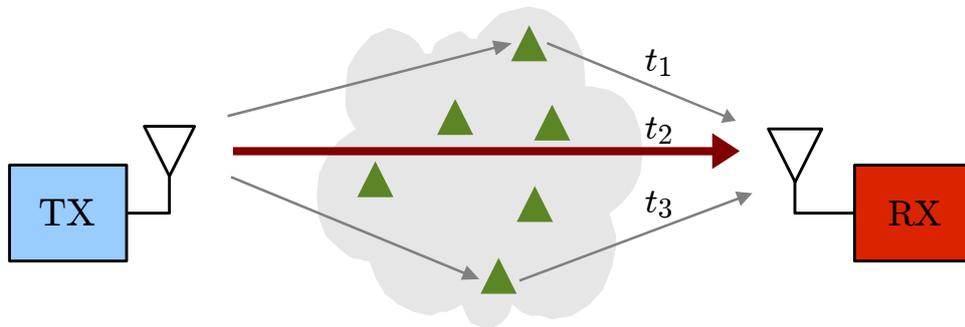
Figure 2.6: Discrete multipath signal model

Figure 2.6 shows a graphical depiction of the signal model we adopt; the space between a transmitter and receiver in our problem is occupied by material which transmitted signals reflect from and propagate through. For location estimation, our interest is only in the *direct path* signal, which contains, in its phase, information about the distance between transmitter and receiver. However, the effects of multipath signals cannot be ignored, and their possible existence must be accounted for in our model so as to not adversely affect estimates of the

direct path signal. We adopt a discrete multipath model in which the signal measured at a receiver consists of delayed and attenuated copies of the transmitted signal.

Denote a transmitter output signal as $x(t)$ and its Fourier transform, $X(\omega)$. The effect of a propagation delay of t_0 seconds applied to $x(t)$ has the effect of multiplying the transmitted spectrum by a complex exponential: (where $\overset{\mathcal{F}}{\Leftrightarrow}$ indicates a Fourier transform pair)

$$x(t - t_0) \overset{\mathcal{F}}{\Leftrightarrow} X(\omega)e^{-j\omega t_0}. \tag{2.3}$$

Each signal path incorporates another copy of the original transmitted signal with an unknown attenuation γ_i and propagation delay t_i:

$$\sum_i^{N_{\text{sig}}} \gamma_i\, x(t - t_i) \overset{\mathcal{F}}{\Leftrightarrow} X(\omega) \sum_i^{N_{\text{sig}}} \gamma_i\, e^{-j\omega t_i} = X(\omega) \cdot V(\omega) \tag{2.4}$$

where N_{sig} is the number of signals, and $V(\omega)$ represents a sum of sinusoids which are functions of frequency. Knowledge of the transmitted signal $X(\omega)$ then permits estimation of each $\{\gamma_i,\ t_i\}$ parameter pair from the channel response $V(\omega)$.

In practice, the channel response of Eq. 2.4 is measured at a finite set of carrier frequencies f_k through receiver hardware; thus the signals available to be processed are of the form

$$X(\omega_k) \cdot V(\omega_k) \cdot H(\omega_k), \tag{2.5}$$

where $X(\omega_k)$ is the spectrum of the transmitted signal consisting of the subcarriers' amplitudes and phases and $H(\omega_k)$ is the receiver frequency response; both of these quantities are assumed known, and thus they may be calibrated out, leaving the measured channel response V to be processed. As the multipath parameters of the channel are described by the sinusoids that comprise its impulse response, we wish to estimate the spectrum of V from its samples $V(\omega_k)$. The sum-of-sinusoids model is exploited by all subsequent methods presented in this dissertation. Specifically, the separation of signal paths can be achieved by taking advantage of the orthogonality of these sinusoids in the frequency domain.

This problem of identifying the parameters of multiple sinusoids in the presence of noise is known as a *spectral estimation* problem, which appears in many forms throughout the field of signal processing. Reference [7] presents a thorough survey on historical and contemporary techniques for the solution of this problem, presenting *state-space* methods to be the most

attractive; these methods take advantage of (assumed) orthogonality of the signal and noise subspaces through a robust direct matrix factorization-based solution for the parameters of interest. While some inferior methods have nearly equivalent estimator performance in some cases, many suffer from difficulties relating to high-degree polynomial approximations and computation time proportional to the amount of desired solution precision.

2.4   Direct State Space

Originally developed by Kung [8], *direct state space* (DSS) methods have been since rediscovered under the names ESPRIT [9] and Matrix Pencil [10]; a comparison of these methods is presented in [11] where no significant difference between the methods is observed. DSS can achieve the Cramer-Rao bound [12] (a theoretical lower bound for estimator error) at suitable SNRs, and thus it is *efficient*. We follow with a presentation of the DSS method and conclude with a discussion of practical issues which limit the utility of DSS for precision indoor location and motivate the development of SART.

2.4.1   Algorithm

The measured channel response V may be expressed as a matrix-vector product:

$$V(\omega) = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_i \end{pmatrix}^T \begin{pmatrix} e^{-j\omega t_1} & & \\ & \ddots & \\ & & e^{-j\omega t_i} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = cA(\omega)b \qquad (2.6)$$

Assuming that the subcarriers ω_k are evenly spaced $\Delta\omega$ apart, the kth sample of V can be expressed in terms of powers of A

$$e^{-j(k\cdot\Delta\omega)t} = (e^{-j\,\Delta\omega\,t})^k \quad \Rightarrow \quad V(k\cdot\Delta\omega) = cA^k b \qquad (2.7)$$

and V may be expressed as the output of an underlying linear time-invariant system:

$$\vec{x}_{k+1} = A\vec{x}_k + bu_k \qquad (2.8)$$

$$y_k = c\vec{x}_k \quad +du_k = V(\omega_k) \qquad (2.9)$$

Where the first equation describes how the state vector $\vec{x}_k$ changes, as a function of state transition matrix A and input term bu_k. The second equation describes how the output y_k is derived from the current state and input. As V is assumed to be an impulse response, the input u_k is an impulse (i.e., $u_k = 1, 0, 0, 0, \ldots$ and $b = \mathbf{1}$) such that at the first k, $\vec{x} = \mathbf{1}$, giving all sinusoids an initial phase of zero.

In terms of the multicarrier signal, $\vec{x}_k$ is the vector containing the phase of each multipath signal component at frequency ω_k. The output y_k is the resulting amplitude and phase measurement at each subcarrier, and is obtained via a sum of the complex-valued phases $\vec{x}_k$ weighted by the multipath signal strengths c. Thus, by indexing $k = 0, 1, 2, \ldots, N - 1$, the state-space formulation of the multipath signal model is completely described by

$$V(k \, \Delta\omega) = y_k = cA^k b. \tag{2.10}$$

The matrix A is the *state transition matrix* whose eigenvalues each correspond to the rate of phase change of each multipath signal with respect to index k. In this way the problem of estimating the multipath delays and amplitudes of a measured channel response is formulated as a *system identification* problem whose goal is to identify c and A based on measurements of $V(k \, \Delta\omega) = y_k$.

To estimate the state transition matrix A, the data are put into a Hankel matrix, which admits to a factorization into observability and controllability matrices:

$$H \;=\; \begin{pmatrix} cA^0 b & cA^1 b & \cdots \\ cA^1 b & cA^2 b & \\ \vdots & & \ddots \end{pmatrix} = \begin{pmatrix} cA^0 \\ cA^1 \\ \vdots \end{pmatrix} \begin{pmatrix} A^0 b \\ A^1 b \\ \vdots \end{pmatrix}^T = \mathcal{O}_1 \mathcal{C}_1$$

$$H_+ \;=\; \begin{pmatrix} cA^1 b & cA^2 b & \cdots \\ cA^2 b & cA^3 b & \\ \vdots & & \ddots \end{pmatrix} = \mathcal{O}_2 A \, \mathcal{C}_2 \tag{2.11}$$

Forming the "advanced" Hankel matrix starting with the second data sample reveals A within the product of another pair of observability and controllability matrices, which are close, but not identically equal to the first pair. To obtain A, H may be factored via a

singular value decomposition [2, 13]:

$$H = U\Sigma V^H = \left(U\Sigma^{1/2}\right)\left(\Sigma^{1/2}V^H\right) = \mathcal{O}_1 \cdot \mathcal{C}_1 \tag{2.12}$$

Given that H is size $m \times n$, U and V are unitary matrices, each forming an orthonormal basis for the columns of A and A^T respectively:

$$U^H U = I_m, \quad V^H V = I_n, \tag{2.13}$$

and Σ is a nonnegative diagonal matrix containing the *singular values* of H, that is:

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n), \quad \sigma_1 \geq \sigma_2 \geq \cdots \sigma_n \geq 0, \tag{2.14}$$

and thus $\Sigma^{1/2} = \mathrm{diag}\left(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \ldots\right)$. In the noiseless case, the rank of H indicates the number of multipath signals that are present, including the direct path. The columns of U and V corresponding to nonzero σ_i span the signal subspace. In the presence of noise, H becomes full rank, and the previously zero-valued σ_i terms become nonzero. Under the assumption that our signals are stronger than the noise, the SVD of H may then be partitioned into signal and noise subspaces,

$$H = \begin{pmatrix} U_s & U_n \end{pmatrix} \begin{pmatrix} \Sigma_s & \\ & \Sigma_n \end{pmatrix} \begin{pmatrix} V_s^H \\ V_n^H \end{pmatrix} \tag{2.15}$$

and estimates of the factors $\mathcal{O}_1$ and $\mathcal{C}_1$ may be formed using only terms from the signal subspace:

$$\left(U_s\Sigma_s^{1/2}\right) \cdot \left(\Sigma_s^{1/2}V_s^H\right) = \widehat{\mathcal{O}}_1\widehat{\mathcal{C}}_1 \tag{2.16}$$

Finally, to obtain an estimate of A, the pseudo inverses of the noise-discarded factors of H are applied to H_+:

$$\widehat{A} = \widehat{\mathcal{O}}_1^{\dagger}H_+\widehat{\mathcal{C}}_1^{\dagger} = \widehat{\mathcal{O}}_1^{\dagger}\mathcal{O}_2 A \mathcal{C}_2\widehat{\mathcal{C}}_1^{\dagger} = \left(\Sigma_s^{-1/2}U_s^H\right)\mathcal{O}_2 A \mathcal{C}_2\left(V_s\Sigma_s^{-1/2}\right) \tag{2.17}$$

As $\widehat{\mathcal{O}}_1^{\dagger}$ and $\mathcal{O}_2$ do not come from the same Hankel matrix, $\widehat{A}$ is not diagonal, but is similar to the A we are searching for and thus has the same eigenvalues. The time delays of interest may then be computed from these eigenvalues:

$$\lambda\left(\widehat{A}\right) = \{\lambda_i, \ i = 1, 2, \ldots, N_{\mathrm{sig}}\}, \quad \lambda_i = e^{-j\Delta\omega t_i}, \quad t_i = \ln(\lambda_i)/(-j\Delta\omega) \tag{2.18}$$

The signal amplitudes are best estimated using a least-squares solution of the system:

$$\left(e^{-j\omega_k t_n}\right)\vec{c} = V(\omega_k) \tag{2.19}$$

fitting the measured data to a basis consisting of sinusoids of the frequencies just estimated.

Other implementation details relating to DSS are covered in [14], where methods for weighting samples and optimum Hankel size[2] are investigated. Reference [15] also has a detailed treatment of DSS, discussing different approaches to the estimation of A, the factorization of H, and estimation of higher-order derivatives of eigenvalues of A.

An extension to DSS exists [16, 17] that utilizes multiple received signals from multiple locations, enabling estimation of the range and angle of each arriving signal. Unfortunately such an approach adds a new challenge for each benefit it offers. Multiple-receiver processing for angle estimation works best with narrowband signals and requires a data interpolation step if a fixed antenna array is to be used. In addition, as range and angle parameters are estimated separately, the task of properly pairing solution components suffers when the signals are closely spaced in both range and angle; methods based on a simultaneous Schur decomposition guarantee pairing but produce worse parameter estimates [18].

2.4.2 Difficulties in Practice

While DSS can achieve the Cramer-Rao bound, some steps of the algorithm are complicated by the realities of the indoor radio channel and realizable hardware considerations for single-receiver time-of-arrival or dual-receiver time-difference-of-arrival schemes.

One issue is the choice of signal rank, or *model order* of the Hankel matrix H. In order to discard the noise subspace, the implementation must *guess* how many multipath signals are present in the data. In less multipath-rich settings, such as anechoic chambers or outdoor locations with clear line-of-sight between antennas, heuristics for choosing the signal rank (the partitioning of Eq. 2.15) such as finding an amplitude "knee" in the singular value progression or just using a fixed value for the signal rank have been shown to yield reasonable model order estimates. However in more realistic settings, model order estimation is still

---

[2]the ratio of number of rows to number of columns, as a Hankel matrix formed from N samples must have size $(m \times N - m + 1)$ where $1 < m < N$.

considered to be an open problem [19], especially in the case of real-time processing where computation resources and number of available data samples are restricted. Choosing the model order is not straightforward as the vectors in the signal basis U_s merely span the signal subspace rather than exactly, individually, representing the sinusoidal signals of interest, preventing a simple "counting of sinusoids" from producing a model order estimate. Incorrect choice of model order results in a signal subspace which either includes vectors from the noise subspace or omits vectors from the signal subspace, thus adding error to estimates of the multipath signal parameters.

Another difficulty in practice using DSS arises when choosing the direct path signal from the set of solutions. Given a model order choice of N_{sig} signals, DSS returns N_{sig} solution pairs $\{t_i, \gamma_i\}$, one of which corresponds to the direct path signal. With perfect synchronization derived as depicted in Figure 2.3, the direct path estimate is simply the signal with the smallest delay. However, realistic channel and multipath assumptions and transmitter-receiver asynchrony (Figures 2.3 and 2.4) in the indoor location problem raise the question of how to choose the direct path signal when the parameter estimates for a multipath channel have a random time offset and random attenuation, i.e., parameter estimates of the form

$$\{t_i + \widetilde{t}, \ \widetilde{\alpha}_i \cdot \gamma_i\} \tag{2.20}$$

which prevent reliable identification of the direct path signal by simple examination of time delay or signal strength. In more forgiving conditions, such as with unobstructed line-of-sight, the signal with the strongest amplitude is chosen for lack of any alternative. In addition, the time offset $\widetilde{t}$ is time-varying as it is a function of the difference between transmitter and receiver sampling clocks, which drift as a function of time.

A third difficulty in practice relates to ill-behavior of DSS in certain multipath situations. In the presence of short-delay multipath which arrives very closely spaced in time and similar in amplitude to the direct path signal, DSS has been shown [20] to exhibit ill-behavior. While a graceful degradation to a parameter estimate that is the mean of two nearly identical signals' underlying parameters may be acceptable, it cannot be guaranteed that ill-behavior will not occur in realistic channels that feature interfering signals and non-flat frequency responses.

2.4.3   The DSS+TDOA approach to location

Given time-difference (TDOA) estimates from DSS and knowledge of receiver coordinates, we may then estimate the position of the transmitter using so-called *multilateration* [21] methods that pursue location solutions using TDOA data. Graphically, multilateration involves solving for the intersection of hyperbolic curves, each satisfying the measured TDOA between the reference antenna and another antenna. Unfortunately, the general solution for the intersection involves a coupled set of nonlinear equations, and an exact solution is generally intractable, especially for arbitrary numbers of receivers. Approximate solution methods amenable to real-time implementation include spherical intersection [22, 23], spherical interpolation [24], and linear-correction least-squares [25] approaches.

Any position estimator using TDOA estimates will suffer if, for any receivers, the DSS process returns an erroneous time-difference value beyond that expected by statistical estimate variation[3]. A realistic indoor environment could cause DSS to result in erroneous values for a number of reasons: from using a sub-optimal model order, choosing a multipath signal as the direct path, or due to ill-behavior in the presence of short-delay multipath. In some cases, the direct path signal may be so highly attenuated that is undetected, in which case a multipath signal delay will be chosen as the DSS result. The DSS+TDOA approach is thus unsuitable for indoor location due the unreliable performance of DSS in high-multipath environments.

Next, we present the SART algorithm and motivate its utility with presentation of location estimation results in real multipath settings.

2.5   SART

The *singular-value array reconciliation tomography* (SART) method was developed in response to the deficiencies of the DSS+TDOA approach in indoor settings, as described in Section 2.4.2, which contains a number of intermediate estimation steps which cause information loss and incur error in the final location solution. SART avoids such intermediate estimation steps and processes data from all antennas simultaneously to estimate location.

---

[3]as in choosing wrong model order or wrong signals will greatly increase your error variance

2.5.1 Algorithm

SART begins by forming a matrix M of frequency data whose columns are the measured channel responses $V_i(\omega_k)$ from each receiving antenna:

$$M = \begin{pmatrix} V_1(\omega_1) & \dots & V_n(\omega_1) \\ \vdots & \ddots & \vdots \\ V_1(\omega_k) & \dots & V_n(\omega_k) \end{pmatrix} = v_{nk} \qquad (2.21)$$

Considering only direct path signal energy from a transmitter, the data matrix is composed of n column vectors of frequency-domain sinusoids sampled at k frequencies:

$$M = \gamma_n e^{-j\omega_k t_n} = \begin{pmatrix} \gamma_1 e^{-j\omega_1 t_1} & \dots & \gamma_n e^{-j\omega_1 t_n} \\ \vdots & \ddots & \vdots \\ \gamma_1 e^{-j\omega_k t_1} & \dots & \gamma_n e^{-j\omega_k t_n} \end{pmatrix} \qquad (2.22)$$

Given knowledge of the receiver positions $\vec{p}_n$ (determined by a manual setup procedure or by GAC) and the frequencies ω_k, the received signals may be rephased for a test location $\vec{x}$ by removing the proper amount of signal delay (where c is the speed of light):

$$M(\vec{x}) = v_{nk} \cdot e^{j\omega_k \|\vec{p}_n - \vec{x}\|_2 / c} = v_{nk} \cdot e^{j\omega_k \Delta \hat{t}_n}, \qquad (2.23)$$

as illustrated in Figure 2.7. Thus a matrix of received signals with propagation delays t_n from a transmitter, rephased to $\vec{x}$ can be written as

$$M(\vec{x}) = \gamma_n \, e^{-j\omega_k t_n} \cdot e^{j\omega_k \Delta \hat{t}_n} = \gamma_n \, e^{-j\omega_k (t_n - \Delta \hat{t}_n)}, \qquad (2.24)$$

which may be interpreted as adjusting the phase of the carriers such that the spectral content of the channel response is frequency-shifted by an amount proportional to the change in time delay $\Delta \hat{t}_n$.

At the transmitter location $\vec{x}_\star$, the delays removed by rephasing are the same as the delays imparted by line-of-sight propagation, and the column vectors in M become a repeated constant, that is, each column is proportional to (the vector of ones) $\mathbf{1}$:

$$M(\vec{x}_\star) = \gamma_n e^{-j\theta_n} \qquad (2.25)$$
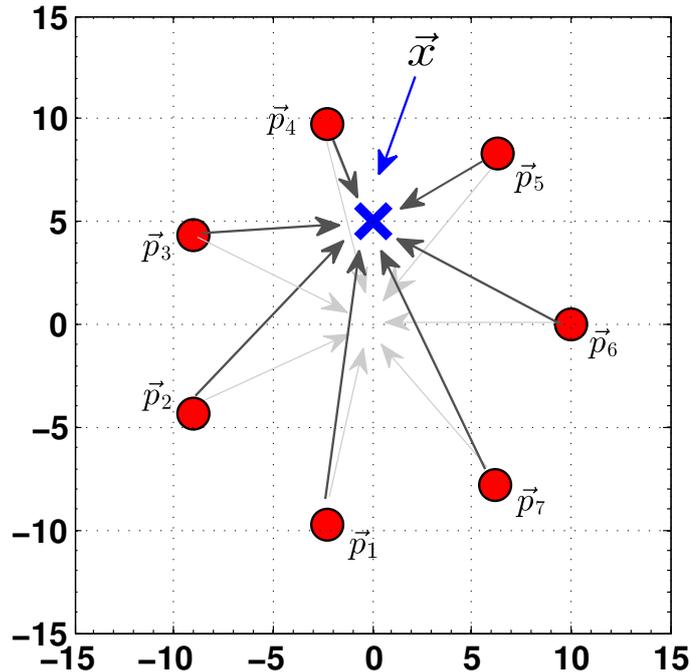
Figure 2.7: Illustration of rephasing in SART. Given receivers (in red) with known positions $\vec{p}_n, n = 1 \cdots 7$, a metric may be computed for any hypothetical transmitter position $\vec{x}$. When the location at $(0, 5)$ is tested, the signals captured at each receiver are rephased so as to remove or "dial back" signal delay commensurate with length of each vector $\vec{x} - \vec{p}_n$ (in gray) that points to the test location.

and thus $M(\vec{x}_\star)$ has a rank of one, as the matrix columns have become linearly dependant. When $\vec{x} \neq \vec{x}_\star$, the receivers' signals will no longer contribute to a single component of rank of $M(\vec{x})$, and further, multipath signals and noise will make $M(\vec{x})$ full rank for any such choice of $\vec{x}$. With the data matrix rephased for a test location, the numerical rank of the matrix may be examined via the singular value decomposition:

$$M = U \Sigma V^H, \qquad \Sigma = \mathrm{diag}(\sigma_1, \cdots \sigma_n), \ \sigma_n \geq \sigma_{n+1}, \qquad U^H U, V^H V = I. \qquad (2.26)$$

where the singular values σ_i are the magnitude of the ith strongest rank-one component of the rephased M at that point:

$$M(\vec{x}) = \sum_i \sigma_i u_i v_i^H \qquad (2.27)$$

If our direct path signals are not overwhelmed by other reflected, interfering signals, then the first singular value of $M(\vec{x})$ is a measure of the strength of the linear dependence of the

received signals' direct-path components. If the rephased data matrix is rank one, this linear dependence is maximized, and all signal energy is represented by the first singular value, and is thus equal to the Frobenius norm of the matrix:

$$\|M(\vec{x}_\star)\|_2 = \|M\|_F \tag{2.28}$$

as the rephasing operation preserves signal energy, and thus the Frobenius norm. At other re-phasing test locations and in the presence of multipath, the received signal energy will be distributed among the rest of the singular values. Thus, the first singular value reaches a maximum at the correct transmitter location where the received signals exhibit high "agreement," and decreases away from that location as the direct path signals components exhibit less pronounced linear dependence.

As the SVD is a highly nonlinear operation, it is impossible to analytically estimate what test location would maximize the first singular value. However, the ability to evaluate the SART metric at any point in space admits to a brute-force imaging approach: by exhaustively computing and recording the first singular value over a rectangular grid of points, a map of first singular value amplitude as intensity may be obtained over that region.

Figure 2.8 shows the result of a 2-D SART *scan* using a simulated dataset in the presence of multipath due to attenuative isotropic reflectors; the transmitter location estimate is chosen as the scan location where the metric is maximized.

2.5.2 Hardware Realization

SART has been demonstrated in field testing using the hardware architecture of Figure 2.9, which consists of a single transmitter with its own local oscillator and sampling clock, and multiple receivers. Each receiver may support up to four co-located receive antennas and consists of an RF downmixer fed by a local oscillator and other filters. The downconverted signals are fed over coaxial cable to the base station where the signals are sampled and sent to a PC running MATLAB.

While the presentation of the SART algorithm in the previous section presumed ideal conditions, the realities of hardware implementation impose asynchronies to which SART must be immune. In terms of the hardware as depicted in Figure 2.9, the measured
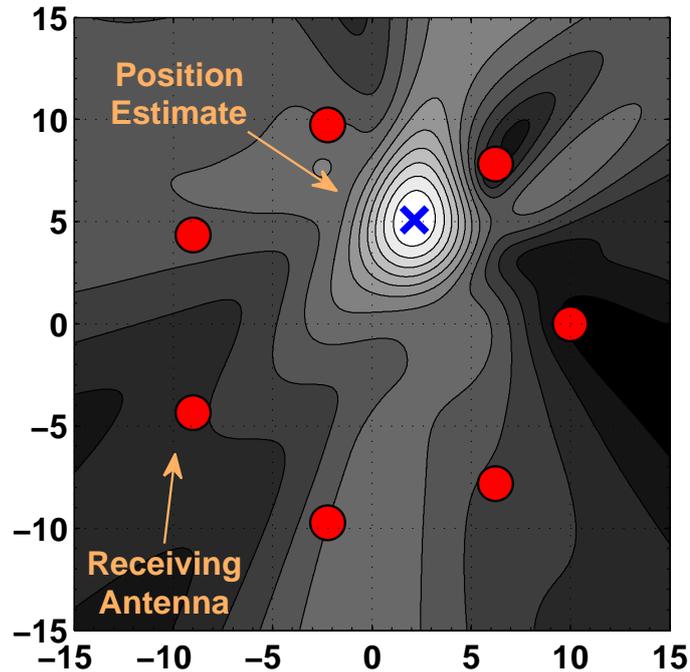
Figure 2.8: Two-dimensional SART image resulting from a simulation of positioning involving seven receiving antennas (red circles) and a single transmitter (blue $\times$). Metric amplitude is indicated by the underlaid contour plot and decreases away from the transmitter as the shading becomes darker.

channel response, including multipath, contains additional parameters to account for these asynchronies:

$$V_{kne} = \sum_{i}^{N_{\text{sig}}} \gamma_{i,n}\, e^{-j\left(\left[(\omega_{k,bb}+\Delta\omega_{bb})+(\omega_c-\omega_{c,n})\right]\left(t_{i,n}+\Delta\tilde{t}_e\right)+\phi_k+\phi_n\right)} \tag{2.29}$$

where k is the frequency sample index, n is the antenna index, and e denotes dependence on receiver antenna port (element) number[4]. The terms are:

- $\{\gamma_{i,n},\ t_{i,n}\}$, the amplitude-delay pair of the ith signal at antenna n;
- $\Delta\tilde{t}_e$, the random time offset at eth antenna port;
- $\omega_{k,bb}$, the frequency of the kth carrier at baseband;
- $\Delta\omega_{bb}$, the deviation between transmit and receive sample clocks;
- ω_c, the ideal RF carrier frequency;

---

[4]The index e can be replaced with n, but is used here to distinguish between spatial effects dependent on antenna position indexed by n, and time-varying effects due to multiplexing.
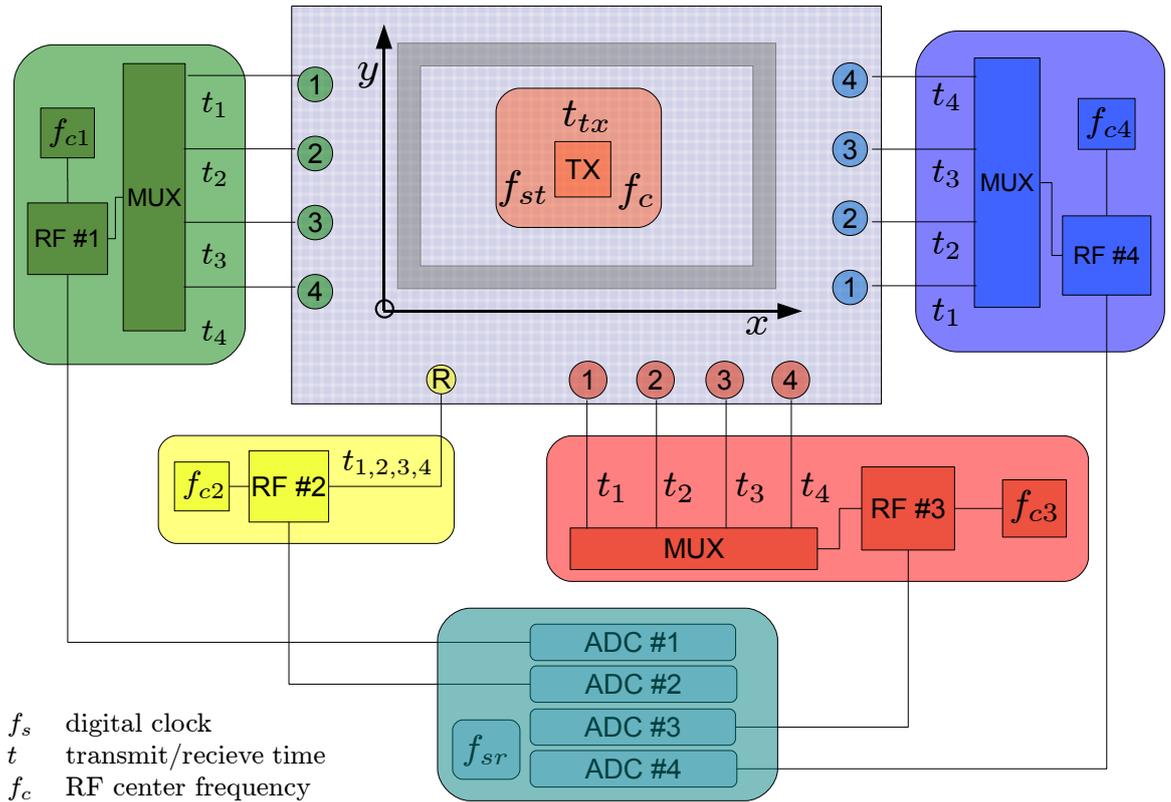
Figure 2.9: PPL system hardware block diagram. The four RF receivers surround the area of operations and downconvert the signals captured at the connected antennas to baseband before relaying them to the base station, at bottom, where they are sampled and delivered to a PC for processing. While receivers 1,3, and 4 each have four ports which are time-multiplexed during sampling, receiver #2 utilizes a single non-multiplexed antenna and may process the transmitter's signal so as to keep its random time offset constant for the duration of the signal capture. The signals are sampled using four co-located ADCs which share the digital clock f_s.

- $\omega_{c,n}$, the actual RF carrier frequency at the nth antenna;
- ϕ_k, the phase of the transmitter's kth carrier; and
- ϕ_n, the local oscillator phase at the nth receiver.

The following analysis will show how these hardware effects do not effect the SART location solution. First, the untethered transmitter will have a random time offset relative to the receivers; as the antenna ports are sampled at successive but different times, each signal has a time offset corresponding to the receiver port it was connected to. By devoting a single receiver (the second, yellow receiver in Figure 2.9) to sample during all time slots using the same antenna, it is possible to adjust the received signals such that they share a single unknown time offset. This "drift tracking" process [26] removes the effect of element number, driving $\Delta \tilde{t}_e \to \Delta \tilde{t}$. Thus, when the signal delays are properly adjusted for the transmitter location, their column vector content is not driven to $\mathbf{1}$ (which may be interpreted as a sampled DC signal), but to a sampled complex sinusoid whose frequency is $\Delta \tilde{t}$, making the rephasing in effect a TDOA process as the signals become *relatively* aligned by removal of their differences.

With the dependence on element number removed, the equivalence of the hardware-complicated SART data formulation to the ideal case is now straightforward to show. First, terms which do not depend on i may be factored outside the summation, where $\Delta\omega_n$ has been substituted for $(\omega_c - \omega_{c,n})$ and the tilde dropped from $\Delta\tilde{t}$:

$$V_{kn} = e^{-j(\phi_k + \phi_n)} \left(\sum_i^{N_{\mathrm{sig}}} \gamma_{i,n}\, e^{-j\omega_k(t_{i,n} + \Delta t)} \cdot e^{-j\Delta\omega_n(t_{i,n} + \Delta t)} \right) \tag{2.30}$$

Transmitted carrier phases ϕ_k and individual receivers' local oscillator phases ϕ_n may be expressed as (diagonal) unitary pre- and post- multipliers to the above summation (denoted M_{kn}). Expressed in this way, the multipliers are absorbed into the left and right singular vectors of the underlying signal matrix $M_{kn} = U\Sigma V^H$:

$$V_{kn} = \mathrm{diag}(e^{-j\phi_k}) \cdot M_{kn} \cdot \mathrm{diag}(e^{-j\phi_n}) = \Phi_k \left(U\Sigma V^H \right) \Phi_n = \tilde{U}\Sigma\tilde{V}^H \tag{2.31}$$

and thus do not affect the singular values of M_{kn}, nor the location solution. This may also be shown to hold for local oscillator frequency offset when direct path only energy is
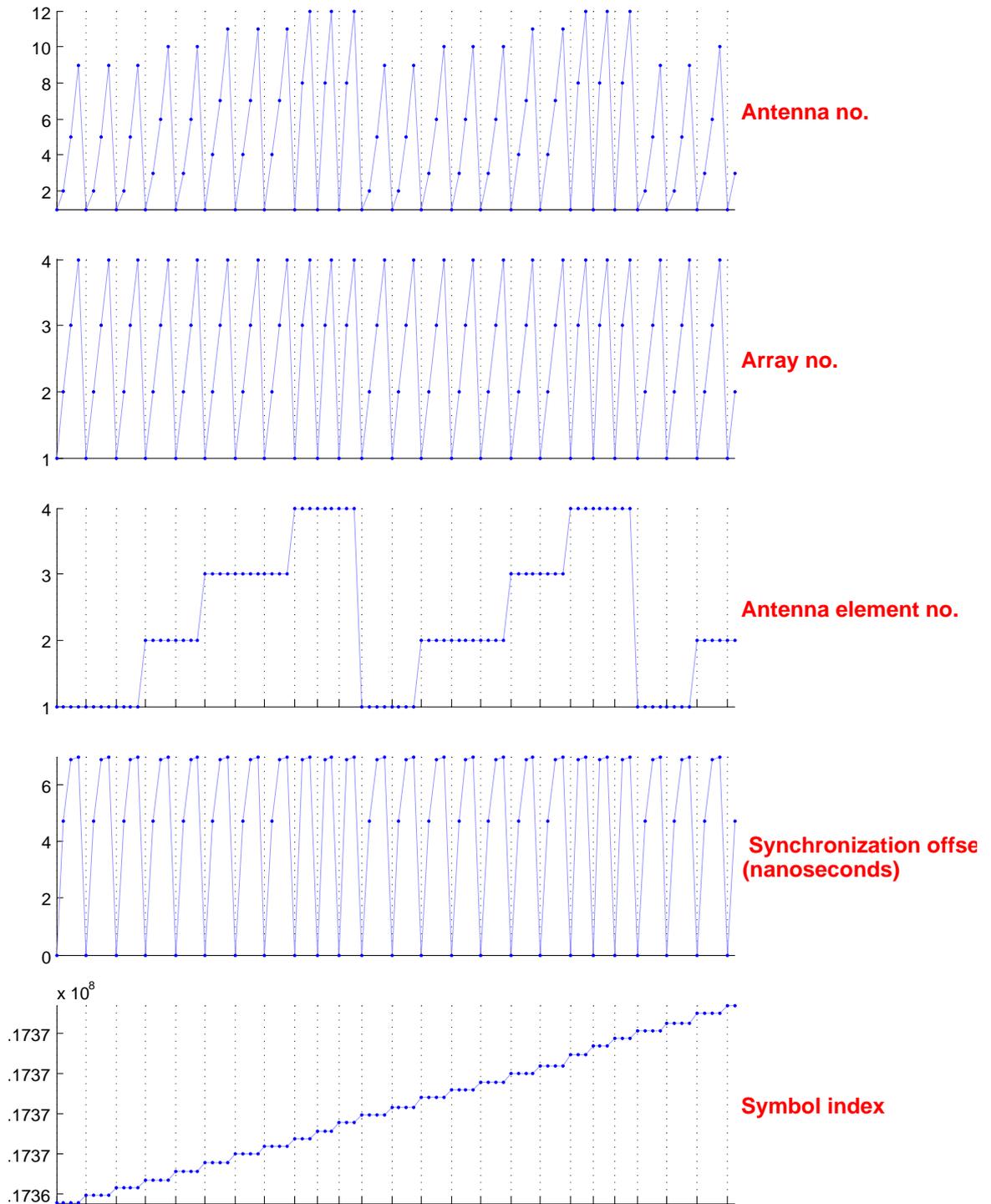
Figure 2.10: Properties of each captured symbol during system multiplexing. During signal capture, the multi-antenna receivers of Fig. 2.9 multiplex their four channels, each connected to a different antenna. For each symbol index during which signals from a particular multiplexer port (1–4) are captured, a signal from the reference antenna is always captured. This allows the synchronization offsets of each receiver to be tracked for the duration of signal capture. Here, only a short duration of the entire signal capture is illustrated, so the synchronization offsets vary little.

considered, allowing $\Delta\omega_n$ to be factored out of the signal matrix:

$$V_{kn} = \Phi_k \left(\gamma_n e^{-j\omega_k(t_n+\Delta t)} \right) \left(e^{-j\Delta\omega_n(t_n+\Delta t)} \right) \Phi_n = V_{kn} = \Phi_k M_{kn} \Omega_n \Phi_n \qquad (2.32)$$

as the effect of the offset may be expressed as a unitary matrix multiplication which does not affect the singular values. If multipath signals are included in the analysis, it can be seen that LO offset simply applies a random phase to each multipath component of a column of the signal matrix:

$$V_{k1} = e^{-j(\phi_k+\phi_1)} \left(\gamma_1 e^{-j\omega_{bb}t_1} \cdot e^{-j\Delta\omega t_1} \;+\; \gamma_2 e^{-j\omega_{bb}t_2} \cdot e^{-j\Delta\omega t_2} \;+\; \cdots \right) \qquad (2.33)$$

which cannot affect any linear dependence relationships measured by the SVD.

Finally, two factors contribute to the random time offset Δt. The first factor is lack of knowledge of the start of signal transmission, while the second factor is a time-varying frequency offset (drift) between the transmitter and receiver sample clocks. To show that the clock offset results in a random time offset, the carrier frequencies ω_k may be written in terms of carrier index $k = \{0, 1, 2, \ldots\}$, DFT bin spacing Δb, starting bin index b_0, sample rate $\omega_s = 2\pi f_s$, and the DFT block (symbol) size of N samples:

$$\omega_k = (k\,\Delta b + b_0)\frac{\omega_s}{N} \qquad (2.34)$$

A clock frequency error $\Delta\omega_s$ would shift the sample rate, and skew the frequencies of the carriers being transmitted. The kth carrier may be written as a function of time t:

$$\exp\left(-jk\Delta b \left(\frac{\omega_s + \Delta\omega_s}{N} \right) t + \phi_k \right) \qquad (2.35)$$

Upon transformation to the frequency domain, the signal consists of the measured phase of every carrier as a function of carrier index k. This may be factored into two components of phase corresponding to the desired clock frequency and the undesired clock offset:

$$\exp\left(-j\left(k\Delta b\frac{\omega_s}{N} + \phi_k \right) \right) \cdot \exp\left(-j\left(k\Delta b\frac{\Delta\omega_s}{N} \right) \right) \qquad (2.36)$$

Rewriting $\Delta\omega_s = \omega_s/M$, with $M \gg 1$, the additional multiplier due to the clock frequency offset is a complex sinusoid over frequency:

$$\exp\left(-jk\Delta b\frac{\omega_s}{N}\frac{1}{M} \right) = \exp\left(-j\omega_k\frac{1}{M} \right) \qquad (2.37)$$

When applied to a received signal which is a sum of sinusoids, this factor increases the multipath delays experienced at all receivers by $1/M$. Thus, a clock offset between receiver and transmitter, as may be expected from any practical oscillator-based system has a benign impact on a TDOA system by inducing a uniform time offset upon all received signals.

Frequency Domain Formulation We may also show the immunity of SART to digital clock offset with an argument in the frequency domain. In the presence of such a frequency offset, ω_s becomes $(\omega_s + \Delta\omega_s)$, and the received signal may be written as (ignoring LO offset, phase effects, and dropping b_0):

$$V_{kn} = \sum_i^{N_{\text{sig}}} \gamma_{i,n} \exp\left(-jk\Delta b\left(\frac{\omega_s + \Delta\omega_s}{N}\right)t_{i,n}\right) \tag{2.38}$$

Rewriting $\Delta\omega_s$ as a fraction $(M \gg 1)$ of ω_s shows the frequency drift may be expressed as an offset to the delay $t_{i,n}$:

$$V_{kn} = \sum_i^{N_{\text{sig}}} \gamma_{i,n} \exp\left(-jk\Delta b\frac{\omega_s}{N}\left(1+\frac{1}{M}\right)t_{i,n}\right) = \sum_i^{N_{\text{sig}}} \gamma_{i,n} \exp\left(-jk\Delta b\frac{\omega_s}{N}\left(t_{i,n}+\frac{t_{i,n}}{M}\right)\right) \tag{2.39}$$

such that the underlying delays are offset proportional to the relative frequency deviation and the underlying delay:

$$t_{i,n,\text{offset}} = t_{i,n}\left(1+\frac{\Delta\omega_s}{\omega_s}\right) \tag{2.40}$$

A reasonable sample clock frequency offset will skew the measured delays, but by a very small amount. Assuming a conservative estimate of clock stability of 100 parts per million, the underlying signal delays will change only a few hundredths of a percent ($\pm 0.01\%$). This delay skew effect is overwhelmed by the time offset due to the unknown start time of symbol transmission. Thus any TDOA method such as SART is unaffected by clock offset.

2.5.3 Location Performance Results

With SART shown to be tolerant to the asynchronies present in the PPL system architecture, in this section we present location results from actual field tests of the PPL system in realistic multipath environments. In such environments, SART has been demonstrated to achieve sub-meter location performance in a variety of indoor multipath settings, as

summarized in Table 2.1, and presented in further detail in [27]. All results involve location of a transmitter inside the building by signals received at antennas placed outside of it, except in the Atwater Kent indoor case where the receiving antennas were also placed inside the building with internal walls separating the transmitter and receiving antennas. The Kaven Hall test location comprises brick and steel-beam construction and houses a geotechnical lab on the WPI campus; the Atwater Kent location is an indoor-to-indoor test centered around an undergraduate laboratory, passing through steel-studded walls and under metal-corrugated ceilings; and the Campus Ministry location is a typical wood-construction three-story residential structure complete with furniture and metal appliances in the kitchen.

Table 2.1: Location Testing Mean Absolute Horizontal Error

Test Location	Error	Bandwidth
Kaven Hall	0.37 m	60 MHz
Atwater Kent, indoor	0.71m	60 MHz
Atwater Kent	1.08 m	60 MHz
Campus Ministry 1st fl.	0.59 m	60 MHz
Campus Ministry 2nd fl.	0.72 m	60 MHz
Campus Ministry 1st fl.	0.72 m	150 MHz
Campus Ministry 2nd fl.	0.30 m	150 MHz

Thus, in the context of this dissertation, the performance Figures of Table 2.1 set a goal for the performance level of GAC: to not increase transmitter location error significantly beyond the levels achieved with manually measured receiver coordinates. The receiver positions are measured at the location of signal reception, nearest to the geometric center of the antenna to an accuracy less than 10 cm.

2.5.4 Contrast with DSS+TDOA

Given SART's ability to function in realistic multipath environments in which DSS+TDOA was ill-suited, the attractive features of the SART method deserve mention.

Single algorithm SART is a single algorithm which directly estimates transmitter location, and does not contain intermediate decision or estimation steps which cause information loss or introduce error.

Uses all data jointly   Because SART does not make intermediate estimates, it operates on all frequency data from all antennas jointly. This benefits the location estimate when some receiving antennas contain no direct path signal energy, in which case they simply do not contribute to the location solution; as the amount of direct path energy increases at an antenna, so does its contribution to the location solution. In multi-step methods such as DSS+TDOA, any line-of-sight TOA estimate made for a direct-path-occluded antenna would be extremely error-prone and impact a TOA- or TDOA-based location solver greatly as it is impossible to distinguish direct path from multipath signals using one antenna.

Arbitrary multi-carrier signals   Further, while SART processing resembles the step of DSS in which frequency data occupies matrix columns, it does not use the Hankel structure which requires frequency samples to be evenly spaced. In SART, the frequency samples may be arbitrarily located, limited only by spectral compliance (reception or generation of RF interference) concerns and DFT bin locations.

Computation time   The imaging approach used by SART is computationally intensive as it must rephase the signal matrix and compute its SVD at each test location; a SART scan on a 10×10 meter grid at 0.25m resolution requires roughly 1,700 evaluations. DSS+TDOA, in comparison, requires only one SVD per antenna and a single SVD for the final location estimate. SART is an algorithm which has become feasible for real-time implementation on general-purpose microprocessors only within the past few years, and achieving location update intervals of a few seconds is the central practical challenge.

In the next chapter, we introduce Geometric Auto Configuration (GAC), whose goal is to automatically estimate receiver coordinates without user intervention, and discuss adaptation of the signal processing approaches presented in the current chapter to its solution. As we have shown SART to be immune to common hardware effects, and achieve meter-level error in real multipath environments, our efforts will focus on adapting it in particular.

Chapter 3

Geometric Auto Configuration

In this chapter, we introduce geometric auto configuration (GAC) as applied to the PPL system, discussing

- the general scheme of GAC, distinguishing it from the transmitter location problem,
- a model for the signals to be processed,
- the interchangeable expression of coordinate and distance information,
- the form of error metrics and theoretical performance bounds, and
- possible solution approaches.

GAC is the process by which a location system, in an unattended fashion, automatically determines the locations of the receiving antennas, before location estimation may proceed. The ultimate goal of GAC is to determine receiver locations to a degree of accuracy such that transmitter location error is not increased significantly beyond the levels achievable in realistic multipath settings with well-known (manually surveyed) receiver coordinates.

The terms *antenna*, *receiver*, and *station* (and combinations thereof) are used interchangeably to refer to the physical location of signal reception as well as the capabilities of the RF electronics at each antenna. Also, the terms *mobile* and *fixed* distinguish between the transmitter worn by a user to be tracked, and the MC-WB transmit capability used at each station to pursue a GAC solution.
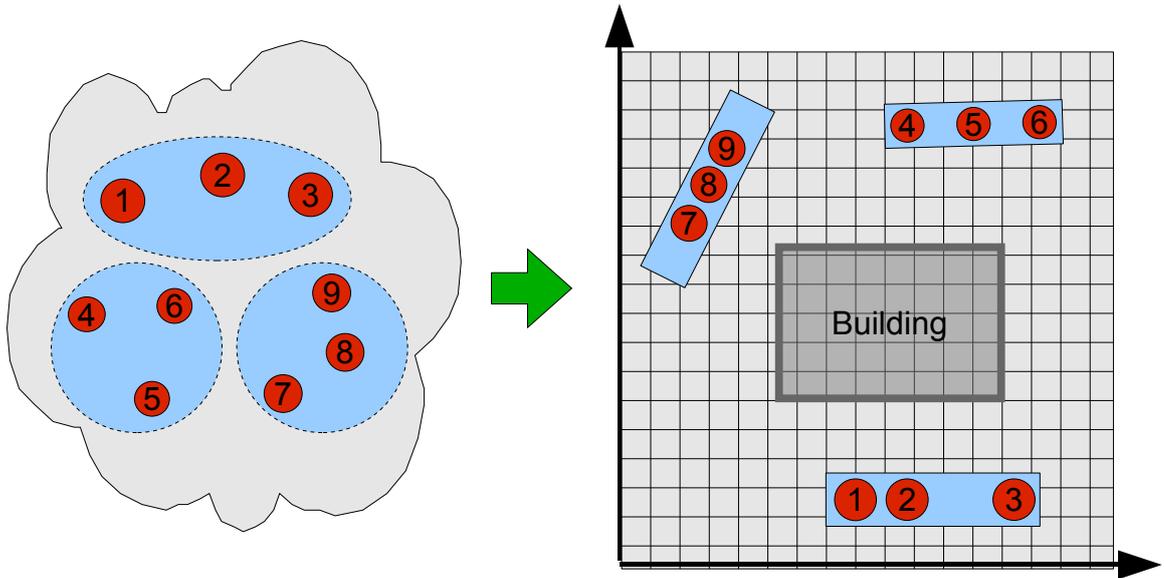
Figure 3.1: Illustration of GAC scenario. Antennas (in red) are part of arrays (in blue) with known geometries, whose overall positions and orientations must be discovered within a local coordinate system that may be aligned to the area of operations.

3.1 GAC Scenario

While fixed infrastructure location systems enjoy well-known, fixed, sensor locations, the PPL system must be deployed at every scene, and thus reliable methods for GAC are necessary. Figure 3.1 graphically illustrates the GAC process; groups of antennas which are co-located on a vehicle or other equipment must have their positions resolved within a local coordinate system defined on scene relative to landmarks of interest. In an ideal deployment, the antennas are physically distributed along the perimeter of the area of operations with as much horizontal and vertical diversity as possible.

From an end-user's perspective, GAC may be divided into two sequential parts. The first part is the accurate determination of the relative locations of the receiving antennas, which is the main research challenge considered in this dissertation. The second part of GAC concerns the mapping of the antenna configuration and transmitter positions within the physical context of the area of operations and conforming with the common conventions of relative and absolute position used by response personnel. For example, firefighters typically label the four sides of the building A, B, C, and D, proceeding clockwise relative to the

incident command post. Thus any positioning information is of diminished utility unless aligned to a street map, floor plan, or A/B/C/D designations.

Upon arrival at the scene of operations and the deployment of antennas used for positioning, GAC may commence. First, the receiver hardware uses an RF signalling capability to establish an ad-hoc network to coordinate the collection of signal data (i.e., MC-WB signal data as illustrated in Fig. 2.2). This is sufficient for normal transmitter location, but for GAC, the receivers themselves are also equipped with MC-WB transmit capability. One by one, each receiver's antenna takes part in transmitting, from its location, a MC-WB signal which all other antennas receive and record. Given N antennas, $N(N-1)/2$ unique signals will be collected, one for each pair of antennas.

Unlike methods in the previous chapter for transmitter location which must use TDOA methods due to an unknown clock offset between the mobile transmitter and receivers, in GAC we may use methods which assume absolute time information since the fixed transmitter hardware is co-located with receiver hardware. This co-location allows direct measurement of the fixed transmitter's clock offset at the moment of signal transmission, and the subsequent elimination of that offset in the signals captured at the receivers. With the offset removed, the ensemble of interantenna signals collected for GAC represent absolute time information, and a solution for the antennas' geometry may be pursued.

While knowledge of the subcarrier frequencies is the only other piece of information required in order to estimate geometry, other supplemental information may be reasonably assumed available for GAC, to both reduce the amount of work required for a solution and improve the theoretical performance bound for error variance.

The first type of supplemental information available arises from grouping the antennas into arrays. For the PPL system, since we assume the receivers are permanently mounted upon emergency vehicles, we may also assume knowledge of the geometry (but not position and orientation) of each subarray. Having knowledge of subarray geometry is beneficial in general because closely situated antennas which are part of the same array are likely to have edge-on look angles, and, as a result, suffer worsened antenna response effects in terms of high phase response distortion and reduced gain.

A second type of supplemental information related to orientation can further reduce the

complexity of finding a geometry solution, and simplify the second manual phase of GAC in which the geometry solution must be aligned to a map. If the subarrays are equipped with magnetometers and accelerometers, then the orientation angles of each array with respect to gravity and magnetic north can be measured. This orientation information would shrink the GAC solution space such that only the translational position of each array was unknown, and allow the resulting geometry solution to be correctly oriented with respect to the cardinal directions.

A final supplemental piece of information could be a rough position fix (possible via GPS) which places the correctly oriented antenna geometry in the vicinity of the area of operations, with the only action required of the user (likely an incident commander) would be a translational alignment of the geometry to the area of operations.

3.2   Existing Literature

Recently, there has been much attention in the literature on the topic of ad-hoc sensor network localization. Such sensor networks consist of a physically dispersed collection of low-power embedded systems equipped with short-range radio links which are used for communication as well as ranging; sensors are capable of range estimation to neighbors via received signal strength or time-of-arrival.

Proposed methods for sensor network localization resemble GAC in their objective, but are too specialized to the challenges of fault tolerance, distributed operation, or limited energy and computation resources to be directly applicable. In addition, proposed methods often assume a dense planar network of sensors which helps to amortize ranging and subsequent position error, and many methods assume the existence of sensors whose position is already known, which may not be assumed by GAC.

Further, such proposed methods and existing implementations make optimistic assumptions about sensors' ranging ability as some demonstrations in "simple" multipath environments exhibit surprising sensitivity to environmental factors [28, 29] and models of ranging error in simulation [30]. The need for robust range estimates despite the use of statistical ranging error models in real multipath situations is frequently acknowledged

[28, 31]. Despite this criticism, it would be foolish to ignore ranging altogether for a solution to GAC, and an approach to estimate range which utilizes *multidimensional scaling* (reviewed favorably in [28]) is presented in Section 3.4.

In contrast to sensor network localization, GAC is not limited to "black-box" inter-antenna range estimates, as the signal from which such estimates are generated (sometimes by hardware) is available for processing by software. Indeed, the MC-WB signal in use for GAC is wideband[1] and consists of multiple carriers; much more signal data is available for processing. As the PPL system has the central goal of *precision* location and is not subject to the hardware, spectral, or computational limitations of such sensor networks, other approaches beyond processing range estimates should be pursued. Exploitation of such an approach based on the signals and processing notions introduced by the WPI PPL system (roughly summarized in Chapter2) will be a focus of the methods to be presented.

3.3   GAC Signal Model

For a PPL-inspired solution to GAC, we retain the sum-of-sinusoids channel response model as presented in Section 2.3. Unlike the transmitter location problem, which has a single-transmitter-multiple-receiver signal model, GAC treats each receiver as a transmitter, or station, to be located and thus takes up a multiple-transmitter-multiple-receiver signal model in which the received signal, as relayed between every pair of antennas, is recorded. To accomplish this, each of the N antennas which normally just receive the mobile transmitter's signal are equipped with a MC-WB transmit capability. Upon arrival at an incident, the stations each in turn switch to transmit mode and emit the multi-carrier signal which is received at all other stations from their deployed locations. The receiving stations simultaneously record the received signal associated with each transmitter, resulting in $N^2 - N$ recorded signals in total. Some of these signals may be ignored during processing as certain pairs of antennas are members of the same fixed-configuration antenna array and thus have known separations.

---

[1]Signals which are ultra wideband are generally defined as having a *fractional bandwidth* (ratio of bandwidth to center frequency) of 20% or more. The current PPL RF signal has carriers spanning 550–700MHz and a fractional bandwidth of 24%.

As the multicarrier signal to be used for this implementation of GAC does not originate from an untethered transmitter but instead is co-located with receiver hardware, clock offset and drift effects may be completely removed from signals acquired in this fashion. Cancellation of these drift effects discards any random time offset, resulting in signals with absolute time information. Thus the direct-path only signal model for the inter-antenna signals may be written as

$$v_{ijk} = \gamma_{ij} \exp\left(-j\omega_k \|\vec{x}_i - \vec{x}_j\|_2/c\right) = \gamma_{ij} \exp\left(-j\omega_k d_{ij}/c\right) \tag{3.1}$$

where $\vec{x}_i$ is the position vector of the ith antenna, d_{ij} is the resulting distance (in units of meters) between the ith and jth antennas, and c is the speed of light. As Eq. 3.1 expresses the received signals in terms of either antenna coordinates or the distances between them, the entire antenna configuration itself can likewise be expressed in two ways: as a set of coordinates or a set of distances.

In coordinate form, the matrix describing a particular antenna configuration may be written as a list of column vectors (in three dimensions, for N antennas):

$$X = \begin{pmatrix} \vec{x}_1 & \vec{x}_2 & \vec{x}_3 & \cdots & \vec{x}_N \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_N \\ y_1 & y_2 & y_3 & \cdots & y_N \\ z_1 & z_2 & z_3 & \cdots & z_N \end{pmatrix} \in \mathbf{R}^{3 \times N} \tag{3.2}$$

In distance form, a configuration may be written as a *distance matrix*:

$$D = \begin{pmatrix} 0 & d_{12} & d_{13} & \cdots & d_{1N} \\ d_{21} & 0 & d_{23} & \cdots & d_{2N} \\ d_{31} & d_{32} & 0 & \cdots & d_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{N1} & d_{N2} & d_{N3} & \cdots & 0 \end{pmatrix} \in \mathbf{R}^{N \times N} \tag{3.3}$$

which has the properties of being symmetric, "hollow", and nonnegative:

$$d_{ij} = d_{ji}, \quad d_{ii} = 0, \quad d_{ij} \geq 0. \tag{3.4}$$

The entries of D are computed from appropriate columns of X,

$$d_{ij} = \|x_i - x_j\|_2 = \sqrt{(x_i - x_j)^T (x_i - x_j)} \tag{3.5}$$

and are invariant to any rotation or translation of X as a whole:

$$d_{ij} = \|R(x_i + t) - R(x_j + t)\|_2 = \|x_i - x_j\|_2 \qquad (3.6)$$

where R is a 3×3 orthogonal rotation matrix, and t is a column vector representing a translation. Thus in forming D, the orientation of the configuration and its absolute position (perhaps described by its center of gravity) are lost, but the relative positions of the antennas, henceforth known as *configuration* or *geometry*, are not. In mathematical terms, all coordinate matrices which generate an identical distance matrix are members of the same *equivalence class*. That is, if Y is a particular coordinate matrix, then its equivalence class is defined as the set of all matrices

$$\left\{ X \in \mathbf{R}^{3 \times N} \mid \|x_i - x_j\|_2 = \|y_i - y_j\|_2, \ i, j = 1, \cdots, N \right\}. \qquad (3.7)$$

The distance matrix formulation of the antenna configuration is useful for GAC, because, via a direct "change of basis" method known as *multidimensional scaling*, antenna coordinates may be directly computed from a distance matrix. Thus, as a distance matrix may be derived from relative coordinates and vice versa, the two representations are interchangeable and equally useful for any solution method for GAC.

3.4 Multidimensional Scaling

The technique of multidimensional scaling (MDS) is, in general, used to form low-dimensional representations of high-dimensional data [32]. MDS is a widely employed technique in psychology, statistics, and other fields, for the analysis of "objective" datasets (such as economic or demographic data) as well as for softer "subjective" data generated by surveys of human perception of similarity or ranking, in order to discover and visualize the strongest components of data variation. MDS may be thought of as an SVD for relational data, and its direct solution form (when working with Euclidean distances) is identical to principal components analysis apart from a change of sign.

While MDS may be written directly in the form of a matrix decomposition [33], modern MDS methods use iterative optimization techniques to accommodate non-linear notions

of distance, different measures of error, and weighted or missing data [34]. For our imple-
mentation of GAC, we are restricted to the physically intuitive, "textbook" application of
MDS in which the input data, Euclidean interantenna distances, are used to estimate a
three-dimensional antenna configuration, which best fits the original data in the least-squares
sense.

Both the direct and iterative approaches are useful for GAC; the direct method is suitable
for estimating X from an exact D matrix, while the iterative method may be preferred when
the elements of D are expected to have errors, and must be used when data are missing.
Direct MDS relates the symmetric product of the desired coordinate matrix to the matrix of
squared and "centered" distances:

$$X^T X = -\frac{1}{2} C (D \circ D) C \qquad (3.8)$$

where the "$\circ$" symbol denotes the Schur (elementwise) product, and C is the centering
matrix, $C = I_n - (1/n)\mathbf{1}\mathbf{1}^T$, which upon pre- or post-multiplication, subtracts average column
or row values, respectively. The matrix of squared distances, $E = D \circ D = d_{ij}^2$, is known as
a *Euclidean distance matrix* [35]. Using E, the equivalence of the two sides of Eq. 3.14 may
be shown. The elements of $X^T X$ are

$$X^T X = \begin{pmatrix} x_1^T x_1 & x_1^T x_2 & x_1^T x_3 & \cdots & x_1^T x_N \\ x_2^T x_1 & x_2^T x_2 & x_2^T x_3 & \cdots & x_2^T x_N \\ x_3^T x_1 & x_3^T x_2 & x_3^T x_3 & \cdots & x_3^T x_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_N^T x_1 & x_N^T x_2 & x_N^T x_3 & \cdots & x_N^T x_N \end{pmatrix} \qquad (3.9)$$

and the elements of E may be written in terms of the columns of X,

$$e_{ij} = (x_i - x_j)^T (x_i - x_j) = x_i^T x_i - 2x_i^T x_j + x_j^T x_j \qquad (3.10)$$

which may be expanded to full matrix form:

$$E = \operatorname{diag}(X^T X)\mathbf{1}^T - 2X^T X + \mathbf{1}\operatorname{diag}(X^T X)^T. \qquad (3.11)$$

Substituting this expression for E into Eq. 3.14, we obtain

$$X^T X = -\frac{1}{2} C \left(\operatorname{diag}(X^T X)\mathbf{1}^T - 2X^T X + \mathbf{1}\operatorname{diag}(X^T X)^T \right) C \qquad (3.12)$$

Centering retains only the second term of E (since $C\mathbf{1} = 0$), showing the result,

$$X^TX = C(X^TX)C = (XC)^T(XC) \tag{3.13}$$

assuming that X is geometrically centered at the origin, that is, $XC = X$ and $X\mathbf{1} = 0$. Thus, from any D, we may compute a symmetric positive semidefinite matrix equal to X^TX, and subsequently estimate X via SVD:

$$-\frac{1}{2}C(D \circ D)C = U\Sigma U^T = \left(\Sigma^{1/2}U^T\right)^T \left(\Sigma^{1/2}U^T\right) = X^TX \tag{3.14}$$

In the case of three-dimensional reconstruction, X^TX must be rank three, and thus the reconstruction product $\Sigma^{1/2}U^T$ will use the first three singular values and corresponding singular vectors to estimate X.

Unlike direct MDS, iterative MDS allows weighting of the entries in D, and by assigning a low or zero weight, allows missing data. These routines search for a configuration $\hat{X}$ which minimizes the least-squared error between the imperfect input data $D = \hat{d}_{ij}$, and the distances generated from the current candidate $\hat{X}$:

$$\mathrm{STRESS}(\hat{X}, D) = \sum_{1 < i < j < N} w_{ij}(d_{ij}(\hat{X}) - \hat{d}_{ij})^2 \tag{3.15}$$

where $w_{ij} = W$ is the weighting matrix. To minimize the error, iterative MDS solvers use a form of gradient search which has global linear convergence [32]. The significant advantage of iterative over direct MDS is the ability to converge to a solution despite missing a number of the entries in D. This advantage is due to overdetermination of the elements of D when there are greater than $N = 7$ antennas, as the number of unique elements in D grows quadratically as $N(N-1)/2$, while the number of parameters ultimately sought for X grows linearly as $3N$. Thus, for reconstruction in three dimensions, D must have at least $3N$ entries specified to match the number of unknowns in X. In D, these $3N$ values correspond to distances from each antenna to three others, which are the minimum number of distances necessary to make the configuration physically rigid and thus a unique minimizer of stress. Interestingly, as the number of antennas increases, so does the advantage of overdetermination, allowing reconstruction of very noisy datasets [35, §5.13].

(a) Antenna layout

(b) $N = 1$, 4 sig.

(c) $N \leq 2$, 12 sig.

(d) $N \leq 3$, 24 sig.
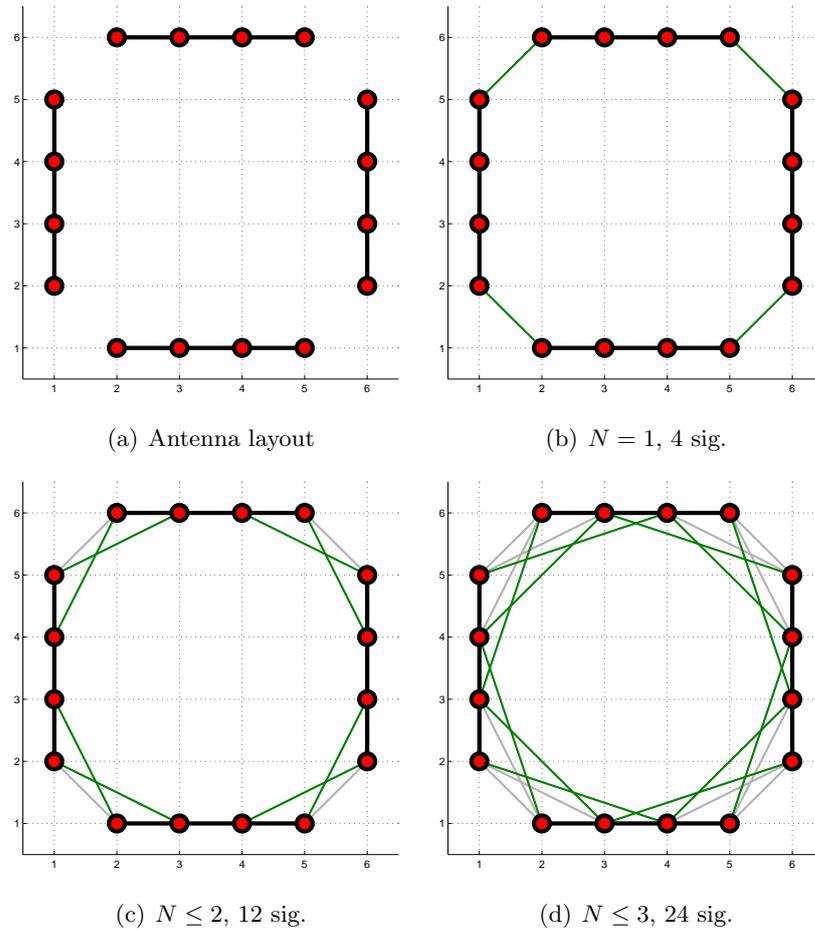
Figure 3.2: Illustration of neighbor distance metric. Signals included at each level are in green, signals with lower neighbor distance in gray.

3.5 Error metrics for GAC

For GAC, as antennas are to be distributed along a perimeter, this minimum threshold on the number of distances necessary for solution convergence suggests a natural ordering of the distance information in D by an integer-valued *neighbor distance*, which is illustrated in detail in Fig. 3.2. This ordering, relative to each antenna, assigns immediate neighbors a distance of 1, subsequent neighbors-of-neighbors a distance of 2, etc. Thus we may evaluate the performance of iterative MDS for GAC as a function of neighbor distance, successively including data from antenna pairs with higher neighbor distances, starting from 3, until there are no missing entries. This notion of neighbor distance is also useful because it gives

an ordering to the interantenna signals roughly based on signal propagation distance; by doing this the signals which are more likely to be strongly attenuated and accompanied by large amounts of multipath are generally the signals which have more distance to travel and thus a higher neighbor distance. By ordering the signals in this way we may attempt to compute geometry solutions using the "best" signals first.

Upon obtaining a relative coordinate solution $\hat{X}$ for each value of neighbor distance from a measurement matrix $\hat{D}$, the RMS ranging error r_{err} and antenna location error $x_{\mathrm{err}}(\hat{X})$ may also be computed as a function of *neighbor distance*:

$$r_{\mathrm{err}}(n) = \sqrt{\frac{1}{N} \sum_{i-j<n} \left(\hat{d}_{ij} - d_{ij}\right)^2} \qquad x_{\mathrm{err}}(\hat{X}) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left\|\hat{\mathbf{x}}_i - \mathbf{x}_i\right\|_2^2} \qquad (3.16)$$

In order to compute x_{err}, $\hat{X}$ and $X_\star$ must first be aligned via an (orthogonal) Procrustes problem solution technique, which solves for the unknown rotation matrix between two sets of points. A general solution for the Procrustes problem is presented in [2, §12.4.1], and a more robust solution which disallows rotations that include reflection (for which $\det(R) = -1$) is proposed in [36].

Considering performance by using these error metrics allows MDS to take on a diagnostic role in system deployment by generating curves showing the tradeoff between the amount of distance information used and antenna location error. If, for example, it were discovered that only distance estimates or signals with a neighbor distance of 4 were necessary for solution, then the number of signals to be captured and the size of the input data to a solution method, would be linear with respect the number of antennas, rather than quadratic when all measurements or signals are required.

While MDS may be used to solve for an antenna configuration given distance information, it is also useful in enabling transformation between (relative) coordinate and distance matrix formulations for antenna configuration, such that any GAC solution method may use either, or both, formulations.

3.6 Cramer-Rao bound for antenna location error

For evaluation of the performance of methods presented later in this dissertation for GAC, two Cramer-Rao bounds are derived in Chapter 4 for antenna location error in two dimensions.

The first CRB result bounds the performance of methods which estimate the antenna coordinates contained in the parameter vector θ based on measurements $\tilde{z}_{ij}$ of the interantenna distances d_{ij} corrupted by zero-mean Gaussian errors:

$$
\begin{aligned}
\tilde{z}_{ij} &= d_{ij} + \tilde{n}_{ij} \\
\tilde{z}_{ij} &= \mathcal{N}(d_{ij}, \sigma_{ij}^2) \\
d_{ij} &= \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \\
\theta &= (x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N)^T
\end{aligned}
\tag{3.17}
$$

This result will be used in evaluating the performance of methods which primarily use MDS to determine the antenna configuration.

The second Cramer-Rao result seeks the same bound on antenna geometry estimates, but assumes the input data are frequency-domain samples of a multipath-free channel response y_{ijk} between antennas i and j, and at carrier index k. (Using the notation of the previous sections, $y_{ijk} = V_{ij}(\omega_k)$). The measurements $\tilde{z}_{ijk}$ are assumed corrupted by zero-mean complex Gaussian errors:

$$
\begin{aligned}
\tilde{z}_{ijk} &= y_{ijk} + \tilde{n}_{ijk} \\
\tilde{z}_{ijk} &= \mathcal{N}(y_{ijk}, \sigma_{ijk}^2) \\
y_{ijk} &= a_{ij} \exp\left(-j\,2\pi f_k\, d_{ij}/c + j\phi_{ij}\right) \\
\{a_{ij}\} &= (a_{12}, a_{13}, \ldots, a_{1N},\ a_{23}, a_{24}, \ldots, a_{2N},\ \ldots,\ a_{N-1,N}) \\
\{\phi_{ij}\} &= (\phi_{12}, \phi_{13}, \ldots, \phi_{1N},\ \phi_{23}, \phi_{24}, \ldots, \phi_{2N},\ \ldots,\ \phi_{N-1,N}) \\
\theta &= \left(\{a_{ij}\}, \{\phi_{ij}\}, x_1 \ldots x_N, y_1 \ldots y_N\right)^T
\end{aligned}
\tag{3.18}
$$

and the parameter vector θ also includes the amplitude a_{ij} and overall phase ϕ_{ij} of each direct path signal. This approach will be used to evaluate the performance of SART-based techniques for GAC which process frequency-domain samples directly.

3.7   Solution Approaches for GAC

In the transmitter location problem, we seek to identify the transmitter position based on knowledge of the signals received at each antenna and the antenna configuration X. Knowledge of the antenna configuration happily allows reduction in the dimensionality of a SART scan from the space of all delays in $\mathbf{R}^n$ (given n received signals) to a three-dimensional manifold (embedded in $\mathbf{R}^n$) of physically meaningful delays corresponding to hypothetical transmitter locations.

In the case of GAC, $N(N-1)$ inter-antenna signals are available, and the dimensionality of the space of possible delays is $N(N-1)/2$, as there are two signals for every antenna pair[2]. While some of the interantenna delays, or other supplemental information regarding the geometry (as outlined in Section 3.1) may be considered known, the space of unknown delays remains very large. For example, given a system of 16 antennas, there are 120 possible delays, and even when those 16 antennas are grouped into four arrays of four antennas each of whose geometries are assumed known, the number of unknown delays remains high, at 96. Even in the most extreme case in which the orientation and geometry of each array is assumed known, the solution space has 8 degrees of freedom in the two dimensional problem and 12 degrees of freedom in the three dimensional problem.

In situations where much supplemental information is available, the solution space for GAC (assuming four or more antennas, and more than one array) remains larger than for transmitter location. While an exhaustive, brute force metric imaging scan is feasible (such as for the SART location method) when imaging a 2-D plane or 3-D volume of space on contemporary computer hardware a single brute force scan for GAC would be impossibly large and simply infeasible in terms of computation time and available memory,; storage requirements for the multidimensional imaging result would exceed the memory capacities of contemporary computers. Thus GAC cannot be solved with exhaustive imaging as with transmitter location, and other methods must be found.

One natural alternative to brute-force, scan-based metric imaging is iterative search. Such a technique, might, for example, follow the gradient of the SART singular value metric

---

[2]Corresponding to when antenna A transmitted to B, and the reverse, when B transmitted to A.

to a local maximum. This could be useful for getting near the solution, but the SART image in the neighborhood of the solution can have many local maxima in addition to the "true" maximum nearest to the transmitter's location. In the presence of noise and multipath signals, such maxima are difficult to distinguish between when choosing a position solution, because the first singular value can no longer be expected to reach its maximum value $\|M\|_F$ as was the case for the noise and multipath-free example in Section 2.5.

While the first singular value $\sigma_1 = \|M(d)\|_2$ of an $m \times n$ signal matrix rephased by a vector of delays d has its value bounded above by the Frobenius norm, and below in the extreme case in which all singular values are equal,

$$\frac{\|M(d)\|_F}{\min(m,n)} \leq \|M(d)\|_2 \leq \|M(d)\|_F \tag{3.19}$$

it cannot be known what maximum value within this range the first singular value will achieve at the transmitter location, and any blind iterative search method is likely to get stuck on local maxima because the global maximum while bounded, is yet unknown. Thus, due to the the intractability of brute-force global optimization, the solution to GAC must be sought among other possibilities.

For the two methods developed for this thesis, and presented in Chapter 5, we may take advantage of two aspects of GAC not shared with the transmitter location problem, which provides a path to a tractable solution approach. The first aspect relates to the concept of neighbor distance: signals which have low neighbor distance are generally assumed to have less severe multipath conditions than those with a higher neighbor distance. Higher neighbor distance signals are more likely to pass through buildings and other intervening structures which induce attenuation and add undesirable multipath signals. This potential for favorable propagation conditions between "nearby" antennas is the primary motivation for the ordering of signal/distance information by neighbor distance.

The second aspect of GAC which may be exploited concerns knowledge of the configuration of antenna arrays. If array geometry and orientation (relative to gravity and magnetic north) are known, then (as previously stated) the solution space reduces significantly in size, but may still be too large for a single metric imaging exercise.

Since absolute time information is available for signals captured during GAC we may

pursue a solution based on traditional ranging techniques which estimate the time-delay of each signal individually. This is the first method presented for GAC. Given estimates for the interantenna distances, we may attempt to find a coordinate solution using both direct MDS which utilizes all distances, and itertive MDS, which may generate a solution with incomplete distance information. Since the concept of neighbor distance gives a rough ordering to our distance information, we may attempt a number of iterative MDS solutions parameterized by neighbor distance.

The second method for the solution of GAC which we will present is a SART-based technique which consists of a series of imaging scans, utilizing a matrix extension which favors relative as well as absolute alignment of signals delays, thus respecting absolute time information. Since array geometries and orientations are assumed known, this method involves a number of SART scans, locating each array one at a time relative to the others, until all arrays' positions have been estimated.

In this chapter, we have described the general scheme of the GAC problem and how its signals are acquired, distinguishing it from the transmitter location problem. While the solution space in transmitter location is 2-D or 3-D and easily imaged, the GAC solution space is much larger, and a single exhaustive imaging optimization over it is infeasible. We may reduce the dimension of the solution space with assumption of array rigity and orientation information, but even with this information the space remains large and a multi-step optimization approach must be considered.

The following chapter details the construction of the Cramer-Rao bound for antenna location error for both distance-based and sampled frequency data-based methods, against which the performance of these methods will be compared.

Chapter 4

Cramer-Rao Bound for Antenna Location

In this chapter, we describe two Cramer-Rao bound (CRB) formulations which will be used to evaluate the performance of methods for GAC using simulated data. The first CRB is derived for methods which process interantenna distance estimates (perhaps derived from samples of frequency-domain data), and the second CRB is derived for methods which estimate antenna geometry from sampled frequency-domain data directly.

4.1   General CRB theory

The CRB is used in estimation theory to derive a lower limit on the achievable variance of an unbiased estimator. Given measurements of a signal with unknown parameter values (such as frequency or phase) and an assumed probability density for measurement errors, a CRB analysis gives lower bounds on the variances of unbiased estimates for any or all parameters, but, unfortunately, does not address *how* to obtain such estimates. CRBs are useful for problem analysis: not only for obtaining a numerical variance bound for a particular problem instance, but also for symbolic purposes such as understanding the influence of system and signal parameters on estimator variance.

Useful examples of generic Cramer-Rao analysis are presented by Van Trees for the

single-parameter case [12, §2.4] and in the multiple-parameter case [37, §8.3.2] assuming Gaussian measurement errors. Section A.7 of the latter reference and Graham [38] present useful reviews of the mechanics and notation necessary for the matrix-vector calculus in the multi-parameter case. The remainder of this introduction outlines the general form of a CRB, and the two sections that follow describe the two particular CRB formulations developed for this thesis.

Mathematically, a CRB analysis proceeds from the definition of the probability density $p_z(z|\theta)$, which is a function of the (random) vector of measurements z and vector of parameters θ. The vector θ contains all unknown parameters of the problem, even if they are not to be estimated. From this probability density function we may form the Fisher information matrix with respect to θ:

$$F_\theta = \mathrm{E}_z \left[\left(\nabla_\theta \ell_z(z|\theta) \right) \left(\nabla_\theta \ell_z(z|\theta) \right)^T \right] = -\mathrm{E}_z \left[\nabla_\theta \left(\nabla_\theta \ell_z(z|\theta) \right)^T \right]. \tag{4.1}$$

where $\ell_z(z|\theta) = \ln p_z(z|\theta)$, and the gradient operator ∇_θ is the column vector of partial derivatives $\nabla_\theta = \left(\frac{\partial}{\partial \theta_1} \cdots \frac{\partial}{\partial \theta_n} \right)^T$. Assuming unbiased estimation, the two forms in Eq. 4.1 are equivalent [12, p. 67]; the first, which expresses the Fisher matrix in terms of an outer product of first-order derivatives (gradients), is useful in a later analysis, while the second form, which is more straightforward to compute, expresses F_θ using the Hessian matrix of second-order derivatives. From F_θ, the lower limit on the estimator covariance Σ_θ is obtained via matrix inversion:

$$\Sigma_\theta \geq F_\theta^{-1}. \tag{4.2}$$

In the multi-parameter case, the inequality in Eq. 4.2 is a *matrix inequality* [39], meaning that $\left(\Sigma_\theta - F_\theta^{-1} \right)$ is positive semidefinite, but the scalar interpretation remains valid along the diagonal, that is [37, p. 926],

$$\sigma_{\theta_i}^2 \geq (F_\theta^{-1})_{ii} \tag{4.3}$$

for the local, perturbative, application of the bound.

A noteworthy complication to the inversion in Eq. 4.2 involves the case of a singular information matrix. For a well-posed estimation problem,[1] this indicates that some functions of the parameters cannot be estimated from the available measurements, and the matrix

---

[1] An example of an ill-posed problem is the task of estimating the positions of the vertices of a triangle given

inverse based bound of Eq. 4.2 must be replaced by a more general notion of the CRB which uses the pseudoinverse. This "modified" CRB [40] is valid provided that the gradient vectors of the functions to be estimated are not part of the nullspace of F_θ. That is, if H is the matrix whose row vectors are the gradients of each function of the parameters to be estimated, then the condition of Stoica [40]

$$H = HF_\theta F_\theta^\dagger \tag{4.4}$$

must be satisfied for the modified CRB to be valid. When this criterion is met, the inestimable parameters are either nuisance parameters[2] or are allowed (and presumably expected) degrees of freedom among the parameters.

In the next two sections, we present the particular bound derivations for distance and frequency data-based estimators, and develop the theory of the modified CRB and make several new and useful observations.

4.2 Distance data CRB formulation

Based on a similar analysis presented by Patwari [31], we derive the bound on location error given distance (inter-antenna) data, extending the derivation to allow rigid collinear sub-arrays, for an arbitrary number of antennas, and set of available measurements.

First, define the parameter vector θ, which vectorizes the (transposed) matrix of position vectors X (defined in Eq. 3.2), in two dimensions:

$$\theta = \text{vec}\left(X^T\right) = \begin{pmatrix} x_1 & x_2 & \dots & x_N & y_1 & y_2 & \dots & y_N \end{pmatrix}^T \in \mathbf{R}^{2N \times 1}. \tag{4.5}$$

The measurements z_{ij} are the interantenna distances d_{ij} between the ith and the jth antennas corrupted by independent, identically distributed, zero-mean Gaussian noise with standard deviation σ_d meters,

$$\tilde{z}_{ij} = d_{ij}(\theta) + \tilde{n}_{ij}, \qquad \tilde{n}_{ij} = \mathcal{N}(0, \sigma_d^2), \qquad \text{E}[\tilde{n}\tilde{n}^T] = \sigma_d^2 I. \tag{4.6}$$

---

only two vertex distance measurements: the triangle's shape is not uniquely determined and an estimator may choose arbitrary interior angles. In this case, the CRB is returning the variance bound for some *other* estimation problem in which the estimator makes the optimal choice for interior angle, and so the original problem is not well-posed.

[2] They are a nuisance in the sense that their values are unknown and thus must be modeled as free parameters in the CRB, but neither their estimates, if available, nor estimate variances, are of interest.

Thus we may write the probability density of z as a product of Gaussian density functions, and account for different problem instances in which some distance measurements are unavailable,

$$p_z(z|\theta) = \prod_{i=1}^{N} \prod_{\substack{j \\ K_{ij}=1}} \frac{1}{\sigma_d\sqrt{2\pi}} \exp\left(\frac{-(\tilde{z}_{ij} - d_{ij}(\theta))^2}{2\sigma_d^2}\right), \tag{4.7}$$

where K is the $N{\times}N$ symmetric matrix (with a zero diagonal) whose ij^{th} entry is 1 when d_{ij} has a measurement available, and is zero when unavailable, discarded, or whose value is known *a priori*. The first product iterates over each antenna, and the second product selects only the antennas to which antenna i makes measurements, as indicated by the entries of K.

4.2.1   Fisher matrix blocks

Taking the logarithm of Eq. 4.7 and writing $d_{ij}(\theta)$ in terms of the coordinates (the parameters), we obtain the log-likelihood:

$$\ell_z(z|\theta) = -\ln\left(\sigma_d\sqrt{2\pi}\right) + \sum_{i=1}^{N} \sum_{\substack{j \\ K_{ij}=1}} \left(\frac{-\left(\tilde{z}_{ij} - \sqrt{(x_i-x_j)^2 + (y_i-y_j)^2}\right)^2}{2\sigma_d^2}\right). \tag{4.8}$$

from which the individual elements of the Fisher matrix may be computed. First, the Hessian is obtained

$$H_{ij} = \frac{\partial^2 \ell_z(z|\theta)}{\partial\theta_i\partial\theta_j}, \tag{4.9}$$

followed by formation of the negative expectation, in which instances of the expectations of the random variables $\tilde{z}_{ij}$ are substituted with their respective statistics (e.g. $\mathrm{E}_z[\tilde{z}_{ij}] = d_{ij}$, $\mathrm{E}_z[\tilde{z}_{ij}^2] = \sigma_{ij}^2 + d_{ij}^2$):

$$(F_\theta)_{ij} = -\mathrm{E}_z\left[H_{ij}\right] = -H_{ij}\Big|_{\tilde{z}_{ij}=d_{ij}} \tag{4.10}$$

In F_θ, two types of entries arise: those in which the two differentiation parameters θ_i and θ_j belong to the same antenna (e.g., x_5, y_5), and those belonging to different antennas (e.g., x_1, y_2). In the former case, the argument to the expectation in Eq. 4.8 retains the summation over j, while the latter case reduces to a single term involving the different-antenna parameters. Based on the ordering of the parameters, the $2N{\times}2N$ Fisher matrix

may be partitioned as

$$F_\theta = \begin{pmatrix} F_{xx} & F_{xy} \\ F_{xy}^T & F_{yy} \end{pmatrix}, \tag{4.11}$$

where each $N \times N$ block contains "same-antenna summation" terms on its diagonal

$$(F_{xy})_{kl} = \sum_{\substack{j \\ K_{kj} \neq 0}} \frac{(x_k - x_j)(y_k - y_j)}{\sigma_{kj}^2 \, d_{kj}^2}, \quad k = l \tag{4.12}$$

and "singleton" terms on off-diagonal elements,

$$(F_{xy})_{kl} = -K_{kl} \, \frac{(x_k - x_l)(y_k - y_l)}{\sigma_{kl}^2 \, d_{kl}^2}, \quad k \neq l. \tag{4.13}$$

The remaining blocks are obtained by substitution of x and y by the coordinate subscripts identifying each block:

$$(F_{xx})_{kl} = \sum_{\substack{j \\ K_{kj} \neq 0}} \frac{(x_k - x_j)^2}{\sigma_{kj}^2 \, d_{kj}^2}, \quad k = l, \tag{4.14}$$

$$(F_{xx})_{kl} = -K_{kl} \, \frac{(x_k - x_l)^2}{\sigma_{kl}^2 \, d_{kl}^2}, \quad k \neq l, \tag{4.15}$$

and

$$(F_{yy})_{kl} = \sum_{\substack{j \\ K_{kj} \neq 0}} \frac{(y_k - y_j)^2}{\sigma_{kj}^2 \, d_{kj}^2}, \quad k = l \tag{4.16}$$

$$(F_{yy})_{kl} = -K_{kl} \, \frac{(y_k - y_l)^2}{\sigma_{kl}^2 \, d_{kl}^2}, \quad k \neq l. \tag{4.17}$$

This structure may be extended to three dimensions by adding F_{xz}, F_{yz}, and F_{zz} blocks, again replacing x and y in Eqs. 4.12 and 4.13 with the two subscripts identifying the desired blocks.

4.2.2   Singularity of the Fisher matrix

While distance measurements are enough to determine the geometry of the antennas (via some estimation procedure), they give no information on absolute position and orientation, and thus individual parameter values are not uniquely determined. Two estimators, given distance measurements only, may return the same geometry solution, but in different translations and orientations. That is, in general, solutions are only unique in the sense that
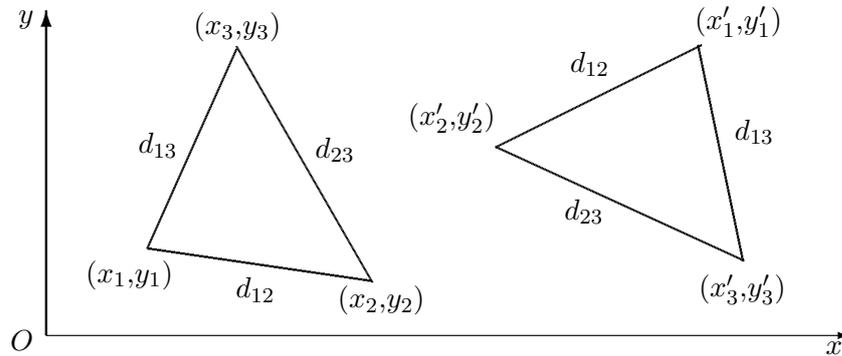
Figure 4.1: Two identical triangles with different position and orientation

they belong to the same equivalence class of relative coordinates. A concrete example of this is the case of a triangle in the plane, shown in two different orientations and locations in Fig. 4.1. It is clear that a different position and rotation angle about its center point cannot alter the distances d_{12}, d_{23}, and d_{13}. In other words, knowledge of the sides' lengths determines its shape completely, but provides no information as to where each of the vertices are located.

This underdeterminism of coordinate values causes the Fisher information matrix F_θ for a many-antenna system to be singular, and the resulting nullspace corresponds in entirety to the parameter subspace which encodes information about the position and orientation of the otherwise completely determined antenna geometry. Such parameter nullspaces for ranging and other types of interantenna measurement in two dimensions are characterized in Ash [41]. For distance measurements, the dimension of the nullspace is a basic geometric fact: in D dimensions there will be D translational degrees of freedom and $\binom{D}{2} = D(D-1)/2$ rotational degrees of freedom for a rigid body.[3]

We may demonstrate the singularity of the Fisher matrix with the simple example of a triangle in the plane. Given the coordinate matrix for such a (noncollinear) triangle,

$$X_\triangle = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \tag{4.18}$$

---

[3]For a single point there are only $D - 1$ rotational degrees of freedom, the extra rotations orient the body properly.

and assuming a measurement standard deviation of $\sigma = 0.10$, we obtain the following information matrix with respect to the six coordinate parameters, following the construction of Eqs. 4.12–4.17:

$$
F_\triangle = \begin{pmatrix}
100 & -80 & -20 & 80 & -40 & -40 \\
-80 & 130 & -50 & -40 & -10 & 50 \\
-20 & -50 & 70 & -40 & 50 & -10 \\
80 & -40 & -40 & 100 & -20 & -80 \\
-40 & -10 & 50 & -20 & 70 & -50 \\
-40 & 50 & -10 & -80 & -50 & 130
\end{pmatrix}
\tag{4.19}
$$

Computing the SVD $V\Sigma V^T$ of this matrix, its singular values are

$$
\mathrm{diag}\,(\Sigma) = \begin{pmatrix} 300 & 180 & 120 & 0 & 0 & 0 \end{pmatrix}^T,
\tag{4.20}
$$

and $F_\triangle$ is indeed singular. The three zero singular values correspond to the two translational and one rotational degrees of freedom that are present when only distance measurements are considered.

4.2.3   Validity of the modified CRB for a singular F_θ

While generally, a singular Fisher matrix invalidates evaluation of the CRB, as some parameters or combinations of them are underdetermined, a valid bound may still be obtained for the remaining estimable quantities, provided certain conditions are upheld. Stoica [40] justifies a "modified" CRB that simply utilizes the pseudoinverse $F_\theta^\dagger$ in place of the normal matrix inverse. This new CRB is accompanied by a condition to determine whether bound results are valid, given any particular function (or functions) of the parameters to be estimated. (Herein the CRB is *defined* as $F_\theta^\dagger$, since in the full-rank case, $F_\theta^{-1} = F_\theta^\dagger$.)

To determine the validity of the bound for a set of scalar functions $\{f_i(\theta)\}$ to be estimated, Stoica's condition involves the gradient of those functions and the information matrix F_θ. If

H is the matrix whose rows are the gradient of each function,

$$H = \begin{pmatrix} \nabla_\theta f_1(\theta) \\ \nabla_\theta f_2(\theta) \\ \nabla_\theta f_3(\theta) \\ \nabla_\theta f_4(\theta) \\ \vdots \end{pmatrix} = \begin{pmatrix} \partial/\partial\theta_1 f_1(\theta) & \cdots & \partial/\partial\theta_n f_1(\theta) \\ \partial/\partial\theta_1 f_2(\theta) & \cdots & \partial/\partial\theta_n f_2(\theta) \\ \partial/\partial\theta_1 f_3(\theta) & \cdots & \partial/\partial\theta_n f_3(\theta) \\ \partial/\partial\theta_1 f_4(\theta) & \cdots & \partial/\partial\theta_n f_4(\theta) \\ \vdots & \vdots & \vdots \end{pmatrix} \tag{4.21}$$

then the validity of the CRB for those functions' estimates requires that

$$H F_\theta F_\theta^\dagger = H. \tag{4.22}$$

In other words, the gradient (row) vectors of the functions to be estimated must "survive the trip" when mapped onto and back from the range of F_θ; that is, the vector space spanned by the rows of H must be a subset of the range of F_θ.

It is possible to deconstruct Eq. 4.22 to generate a simpler, equivalent test. If we express F_θ in terms of its symmetric SVD[4] as $V\Sigma V^T$, then the condition becomes

$$H = H F_\theta F_\theta^\dagger = H \left(V\Sigma V^T \right) \left(V\Sigma^{-1} V^T \right)^T = H \left(V\Sigma V^T \right) \left(V\Sigma^{-1} V^T \right) \tag{4.23}$$

where V is a full-rank, square, orthogonal matrix, and the singular nature of F_θ is encoded in the matrix of singular values Σ:

$$\Sigma = \text{diag}\left(\sigma_1, \sigma_2, \cdots, \sigma_r, 0, \cdots, 0 \right) \tag{4.24}$$

That is, F_θ is rank r, and Σ^{-1} is the diagonal matrix whose entries are the reciprocal singular values:

$$\Sigma^{-1} = \text{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \cdots, \frac{1}{\sigma_r}, 0, \cdots, 0 \right) \tag{4.25}$$

Eq. 4.23 then becomes

$$H = H \left(V\Sigma^0 V^T \right) \tag{4.26}$$

since $V^T V = I$, and the matrix Σ^0 signifies normalization of the nonzero singular values:

$$\Sigma^0 = \Sigma\Sigma^{-1} = \text{diag}\left(1, \cdots, 1, 0, \cdots, 0 \right) \tag{4.27}$$

---

[4]For a symmetric matrix (which must have real eigenvalues), the SVD can be written as $U\Sigma V^T$ with $U = V$, if F_θ is known to have no negative eigenvalues. If some eigenvalues are negative, however, U and V are no longer equal, but differ only by sign (e.g., $U = V\Pi$, where $\Pi = \text{diag}(\pm1, \pm1, \cdots, \pm1)$) and the results of this analysis are unchanged.

Since some singular values are zero, we may partition the n columns of V between those spanning the range of F_θ, and those spanning the nullspace[5] as

$$V = \begin{pmatrix} v_1 & v_2 & \cdots & v_r & \bigg| & v_{r+1} & \cdots & v_n \end{pmatrix} = \begin{pmatrix} V_r & \big| & V_n \end{pmatrix} \tag{4.28}$$

Using this partitioning, the condition becomes

$$H = H\left(V_r V_r^T\right). \tag{4.29}$$

The product $V_r V_r^T$ may be interpreted in two ways: first, as a projection matrix which removes components of each row vector of H not in the range of F_θ, and second, as a "normalized" Fisher information matrix U_θ whose nonzero singular values have been set to one. Replacing $V_r V_r^T$ with U_θ and transposing, the condition takes on a more familiar "$Ax = b$" matrix-vector product form:

$$U_\theta H^T = H^T \tag{4.30}$$

which is the simpler statement of Eq. 4.23 that we seek. Thus, if H is unchanged upon projection through the range of U_θ (and thus F_θ), then the CRB for the functions $\{f_i(\theta)\}$ whose gradients are the columns of H^T, is valid.

Considering a single function $f_1(\theta)$ to be estimated, we may construct a simple test to determine the validity of its CRB. The matrix H^T then becomes a single gradient vector, and we may bound the norm[6] of the product $U_\theta\big(\nabla_\theta f_1(\theta)\big)$ using the coefficient c:

$$\left\| U_\theta\big(\nabla_\theta f_1(\theta)\big) \right\|_2 = c \left\| \nabla_\theta f_1(\theta) \right\|_2, \qquad 0 \leq c \leq 1. \tag{4.31}$$

From this test, three cases may result: $c = 1$, $c = 0$, or $0 < c < 1$. First, when $c = 1$, $f_1(\theta)$ is uniquely estimable from the available measurements and the CRB for it is valid.

Second, when $c = 0$, $f_1(\theta)$ is completely inestimable from the available measurements, and the CRB for its estimate, given the stated information matrix, is undefined. In a sense (emphasized by Stoica) the CRB in this case is "infinite"; the complete lack of information lets an estimator choose arbitrary values for $f_1(\theta)$, and thus have arbitrary, "infinite" variance.

---

[5]For a general matrix, column vectors which are not in the range (column space) come from the *left* nullspace; since F_θ is symmetric, the nullspace and left nullspace are equal and we use the shorter form.

[6]Since U_θ is square but not full rank, we must use the vector 2-norm rather than a general norm.

In the third case, $0 < c < 1$, and $f_1(\theta)$ is "marginally" estimable, in the sense that $f_1(\theta)$ may be estimated from the available data, but not uniquely, as there exist some degree(s) of freedom among the parameters it depends on. Essentially, this means that $f_1(\theta)$ may be written as the sum of two other functions, i.e., $f_1(\theta) = f_{1a}(\theta) + f_{1b}(\theta)$, for which $c = 1$ and $c = 0$, respectively. Thus the CRB gives a variance bound for $f_{1a}(\theta)$ and is undefined for $f_{1b}(\theta)$.

To test CRB validity for multiple functions to be estimated, Eq. 4.31 needs only slight modification, by replacing $\nabla_\theta f_1(\theta)$ with H^T, and using the Frobenius norm:

$$\left\| U_\theta H^T \right\|_F = c \left\| H^T \right\|_F, \qquad 0 \le c \le 1. \tag{4.32}$$

Thus the test measures how much of the Frobenius norm of H is preserved when it is projected onto the range of U_θ. The value of c carries the same interpretation as in Eq. 4.31, but applies to the set of functions as a whole.

For antenna location in GAC, the functions of the parameters (defined in Eq. 4.5) we wish to estimate are simply the parameters themselves, the coordinates. Thus the individual functions $f_1 \cdots f_{2N}$ for obtaining each parameter are

$$f_1(\theta) = x_1, \quad \cdots \quad f_N(\theta) = x_N, \quad f_{N+1}(\theta) = y_1, \quad \cdots \quad f_{2N}(\theta) = y_N \tag{4.33}$$

and may be written together as the vector-valued function

$$f_t(\theta) = I\theta \tag{4.34}$$

where the subscript t is meant to indicate that it extracts "total" and complete information about the parameters. Applying the CRB validity test, we see that we cannot expect unique estimates for $f_t(\theta)$:

$$\left\| U_\theta \left(\nabla_\theta I\theta \right) \right\|_F = \left\| U_\theta I \right\|_F = c \| I \|_F, \qquad 0 < c < 1. \tag{4.35}$$

Since U_θ is singular, I will be diminished in Frobenius norm upon projection, and thus c cannot be zero or one. To find the functions which are completely estimable, we may decompose $f_t(\theta) = I\theta$ using the partitioning of V in Eq. 4.28 between the range and nullspace of U_θ:

$$f_t(\theta) = I\theta = VV^T\theta = \begin{pmatrix} V_r & V_n \end{pmatrix} \begin{pmatrix} V_r^T \\ V_n^T \end{pmatrix} \theta = \left(V_r V_r^T + V_n V_n^T \right) \theta \tag{4.36}$$

From this partitioning, we can see that the identity matrix I has been decomposed into the sum of two matrices $V_r V_r^T$ and $V_n V_n^T$, whose basis vectors (by definition) span the spaces of gradient vectors of estimable and non-estimable functions, respectively. Thus $f_t(\theta)$ is the sum of two functions f_r and f_n:

$$f_t(\theta) = f_r(\theta) + f_n(\theta) = \left(V_r V_r^T \right) \theta + \left(V_n V_n^T \right) \theta \tag{4.37}$$

Given this statement and decomposition of $f_t(\theta)$, we may derive a function for which $c = 1$, by simply subtracting the subspace of gradients of inestimable functions $V_n V_n^T$ from I:

$$f_r(\theta) = (I - V_n V_n^T)\theta. \tag{4.38}$$

Since the basis for V_n which we may derive may be rotated in any fashion, we may write the completely estimable function f_r in terms of an equally sized matrix Q whose columns span the same vector subspace as V_n (but is arbitrarily rotated):

$$f_r(\theta) = (I - QQ^T)\theta. \tag{4.39}$$

To find a suitable Q and verify CRB validity for $f_r(\theta)$ we must find vectors which span the nullspace V_n. Given distance measurements, the parameter nullspace corresponding to V_n has been characterized in Ash [41]. In two dimensions, V_n will have three orthogonal columns; the two vectors $(n_{\bar{x}}, n_{\bar{y}})$ which correspond to translation have fixed entries, and the single rotational degree of freedom vector n_ϕ has variable entries which depend on the particular coordinates (Eq. 4.18) of the geometry being considered:

$$n_{\bar{x}} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad n_{\bar{y}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad n_\phi = \begin{pmatrix} -(y_1 - \bar{y}) \\ -(y_2 - \bar{y}) \\ -(y_3 - \bar{y}) \\ x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}. \tag{4.40}$$

These columns, once normalized, may be assembled into the three-column matrix Q

$$Q = \begin{pmatrix} \dfrac{n_{\bar{x}}}{\|n_{\bar{x}}\|_2^2} & \dfrac{n_{\bar{y}}}{\|n_{\bar{y}}\|_2^2} & \dfrac{n_\phi}{\|n_\phi\|_2^2} \end{pmatrix}. \tag{4.41}$$

which, while not equal to V_n, spans the same vector subspace, and thus must be related by some rotation matrix R:

$$Q = RV_n. \tag{4.42}$$

To show the validity of the CRB of Eq. 4.39, we may now apply the gradient of Eq. 4.32:

$$
\begin{aligned}
U_\theta \left(\nabla_\theta (I - QQ^T)\theta \right) &= U_\theta (I - QQ^T) \\
&= U_\theta (V_r V_r^T + V_n V_n^T - QQ^T) \\
&= U_\theta (U_\theta + V_n V_n^T - (RQ)(RQ)^T) \\
&= U_\theta (U_\theta + V_n V_n^T - V_n V_n^T) \\
&= U_\theta U_\theta
\end{aligned}
\tag{4.43}
$$

and the so the norm-based test my be written as

$$\|U_\theta(U_\theta)\|_F = c\|U_\theta\| \tag{4.44}$$

and so c must be one. In forming the linear functions $f(\theta) = (I - QQ^T)\theta$ which ignore inestimable translation and rotation information, we have shown that the remaining information, corresponding to relative geometry, is uniquely estimable and has a valid CRB.

In summary, applying the modified CRB which involves taking the pseudoinverse of F_θ to obtain a covariance matrix Σ_θ yields a result with a valid interpretation, but when F_θ is singular, not every function of the parameters has a valid CRB for its estimates, and the range and nullspace of F_θ must be scrutinized so that the functions of the parameters for which the CRB is valid, and for which it is invalid, are understood.

4.2.4 Parameter constraints

The previous development was couched in terms that are especially well motivated by problems which have nuisance parameters such as the overall rotation and translation of an otherwise fixed geometry. In this section we will expand this notion to explicitly include unknowns that affect the geometry itself which may be fixed by outside constraints rather than by measurements.

In an alternative to the Stoica modified CRB approach, Ash [41] investigating absolute positioning methods, demonstrates incorporation of absolute position information into

the bound via parameter constraints. In their approach, applying a sufficient number of parameter constraints encoding this information enables formation of a full-rank information matrix, and thus a valid CRB for both relative and absolute geometry information. In characterizing the roles of absolute and relative position information in the CRB, the work of Ash makes an important observation: that the total location error, as measured by the trace of the covariance matrix, is separable into relative and absolute components:

$$\text{tr}(\Sigma_\theta) = \text{tr}(\Sigma_r) + \text{tr}(\Sigma_a) \tag{4.45}$$

since the parameter spaces shown to represent the two types of information are orthogonal and together span the entire parameter space. Until the amount of absolute position information supplied to an estimator surpasses the rank deficiency of F_θ, the two matrix traces in Eq. 4.45 bound two independent estimation problems; $\text{tr}(\Sigma_r)$ supplies the bound for geometry error assuming an optimal choice of position and orientation, and $\text{tr}(\Sigma_a)$ bounds the minimum error based on the absolute position information chosen to be used. When more absolute information than the minimum required for alignment is used, the subspaces begin to overlap and $\text{tr}(\Sigma_r)$ may decrease.

Thus there is a threshold with absolute position information that must be met before *any* reduction in the relative position bound is realized. For example, in our problem's context, to immediately realize such gains we may apply constraints to relative position information by assuming some interantenna distances to be known *a priori*.

The constrained CRB approach, originally due to Gorman and Hero [42], in general, allows introduction of functional equality constraints between parameters of the form $f(\theta) = 0$ given that f is a "smooth" function. As our antennas are naturally grouped together into rigid arrays with known geometries, we may formulate a bound which incorporates such restrictions by constraining all same-array distances to be fixed. Since we wish to constrain the measurements themselves (and not some combination), an alternative approach to applying such constraints would involve taking the limit of the Fisher matrix as the variance of the distance measurements we wish to constrain approaches zero; however the constraint projection approach to be described achieves such constraints *exactly* and is a direct and more flexible approach.

For GAC we consider two types of constraints: coordinate constraints in the case in which we wish to fix the coordinates of an antenna, such as when some antennas' coordinates happen to be known, and second, distance constraints to fix the distance between antennas, such as when a fixed array of antennas is permanently mounted upon a vehicle. To apply the constraints, we build a gradient matrix, one row per constraint. For coordinate constraints such as $y_1 = 5$, the gradient has a single nonzero entry:

$$G_{i\bullet} = \nabla(y_1 - 5) = \begin{pmatrix} \mathbf{0}_{1 \times N} & 1 & \mathbf{0}_{1 \times N-1} \end{pmatrix}, \tag{4.46}$$

and distance constraints such as

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} = 3.048, \tag{4.47}$$

have a gradient with four nonzero entries:

$$G_{i\bullet} = \begin{pmatrix} 0 & \cdots & 0 & \frac{x_i-x_j}{d_{ij}} & 0 & \cdots & 0 & \frac{x_j-x_i}{d_{ij}} & 0 & \cdots & 0 & \frac{y_i-y_j}{d_{ij}} & 0 & \cdots & 0 & \frac{y_j-y_i}{d_{ij}} \end{pmatrix}. \tag{4.48}$$

With the constraint gradients comprising the rows of the matrix G,

$$G = \frac{\partial h_i}{\partial \theta_j} = \begin{pmatrix} \frac{\partial h_1}{\partial \theta_1} & \cdots & \frac{\partial h_1}{\partial \theta_{2N}} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_k}{\partial \theta_1} & \cdots & \frac{\partial h_k}{\partial \theta_{2N}} \end{pmatrix} \in \mathbf{R}^{m \times 2N} \tag{4.49}$$

we may obtain the constrained bound, by projecting the unconstrained Fisher matrix onto the nullspace of G (the subspace where constraints have no effect), take the pseudoinverse, and then "reverse" the projection back into the original parameter space:

$$\tilde{V} \left(\tilde{V}^T F_\theta \tilde{V} \right)^\dagger \tilde{V}^T \leq \Sigma_{\theta,c} \tag{4.50}$$

where $\tilde{V}$ is a matrix whose columns form an orthonormal basis for the nullspace of G [43].

The utility of this constrained CRB formulation (similarly emphasized by Gorman [42]) is that from the unconstrained Fisher matrix one may also obtain a constrained bound with little additional effort, compared to a reparameterization of the problem such as in the "custom CRB" of [12, p. 83] which incorporates the constraints by design.

Unfortunately, a deficiency exists in the constrained CRB for some problem instances, as we have discovered in efforts related to this thesis; the problem arises when subarrays of antennas are constrained to be collinear.
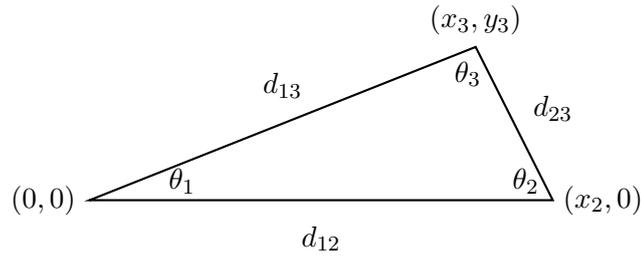
Figure 4.2: Triangle in two dimensions.

4.2.5   Deficiency of the constrained CRB

While it is a reasonable, practical, and common experimental convention to place arrays of antennas in a collinear arrangement, doing so when one wishes to *constrain* the geometry of those same arrays to be fixed (via the method of the previous section) causes the resulting constrained CRB to be deficient. For fixed collinear arrays, the CRB is deficient in the sense that it does not bound the same estimation problem as when the arrays are constrained to be noncollinear.

To demonstrate this deficiency, consider an arrangement of three antennas located at the vertices of a triangle, as shown in Fig. 4.2, representing an array whose geometry we wish to fix within a larger estimation problem. To constrain the geometry we form the three distance equations which effect the fixing of the geometry for this particular instance,

$$d_{12} = x_2 \qquad d_{23} = \sqrt{(x_2 - x_3)^2 + y_3^2} \qquad d_{13} = \sqrt{x_3^2 + y_3^2}, \tag{4.51}$$

and the gradient G_t of this "triangle problem" with respect to the coordinate parameters $\theta = \left(x_1,\, x_2,\, x_3,\, y_1,\, y_2,\, y_3\right)^T$, ignoring row scaling factors, is

$$G_t = \nabla_\theta \begin{pmatrix} d_{12} \\ d_{23} \\ d_{13} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & x_2 - x_3 & x_3 - x_2 & 0 & -y_3 & y_3 \\ -x_3 & 0 & x_3 & -y_3 & 0 & y_3 \end{pmatrix}. \tag{4.52}$$

The row vectors of this matrix are used to form $\tilde{V}$ in Eq. 4.50, allowing the constrained bound to be computed. Thus, the gradient G_t, a first-order quantity, is all that the CRB construction "knows" about the constraint equations, and we must verify that G_t is a faithful

representation of the constraints that are intended, so as not to obtain a bound for some other (differently-constrained) estimation problem.

To see which parameter restrictions G_t are put into effect for our three-antenna array, we may form the total differential of the distance measurements by multiplying G_t by the differential of our parameters $d\theta = \begin{pmatrix} dx_1, & dx_2, & dx_3, & dy_1, & dy_2, & dy_3 \end{pmatrix}^T$. If the distances are to remain constant, then each of the three rows must sum to zero:

$$G_t \ d\theta = \begin{pmatrix} -dx_1 + dx_2 \\ (x_2 - x_3)dx_2 + (x_3 - x_2)dx_3 - y_3 dy_2 + y_3 dy_3 \\ -x_3 dx_1 + x_3 dx_3 - y_3 dy_1 + y_3 dx_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \qquad (4.53)$$

In the noncollinear case $(x_2, x_3, y_3 \neq 0)$, manipulation of these equations provides a geometric interpretation to the allowed perturbations:

$$dx_1 = dx_2 \qquad \frac{y_3}{x_2 - x_3} = \frac{dx_2 - dx_3}{dy_2 - dy_3} \qquad \frac{y_3}{x_3} = \frac{dx_3 - dx_1}{dy_1 - dy_3}. \qquad (4.54)$$

Since x_2, x_3, and y_3 are assumed known, the only allowed perturbations are those that do not change d_{12}, $\tan\theta_2$, or $\tan\theta_1$, respectively, thus constraining the triangle to a single geometry as intended. In the collinear case, y_3 becomes zero and the three points lie on the x axis; the total differential expression of Eq. 4.53 becomes

$$G_t \ d\theta = \begin{pmatrix} -dx_1 + dx_2 \\ (x_2 - x_3)(dx_2 - dx_3) \\ x_3(dx_3 - dx_1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \qquad (4.55)$$

resulting in the simpler set of perturbation equations

$$dx_1 = dx_2 \qquad dx_2 = dx_3 \qquad dx_3 = dx_1 \qquad (4.56)$$

which only restrict the perturbations of the x coordinates. Due to the absence of dy terms, any perturbations in y are allowed by Eq. 4.56, including those which change the array's shape. This clearly violates the intent of the constraints, as only perturbations in y that preserve the geometry should be allowed. Here, instead of fully constraining the geometry, the distance constraints in the collinear case (as represented by their gradient) only respect as fixed the distances between the points as projected onto the axis of collinearity.

This deficiency is not limited to collinearity, however, as we may show by using the above approach that when any antenna arrays have a dimension smaller than than the ambient space in which the GAC problem is embedded, the constrained bound which was intended to fix such "low rank" arrays so as to be rigid using the gradient of the pairwise distance equations, is deficient and underconstrained. Thus, in two dimensions collinear arrays may cause the deficiency, and in three dimensions, collinear and coplanar arrays cause the deficiency, etc.

While it was convenient to enumerate the perturbation variables for the analysis of Eqs. 4.53–4.56, a more straightforward way to characterize constraint deficiency and to count degrees of freedom is to examine the rank of the constraint gradient G_t. Given C constraints and N antennas in D dimensions, a gradient matrix G will have C rows and ND columns, and thus have a maximum rank of $\min\{C, ND\}$. For the relative geometry constraints that we are interested in, among N antennas there will be $C = \binom{N}{2}$ rows corresponding to the pairwise distance constraints and ND columns for the coordinates. However, since the constraints contain only relative geometry information, the maximum rank of the gradient of such constraints becomes

$$\min \left\{ \binom{N}{2}, \ ND - D - \binom{D}{2} \right\} \tag{4.57}$$

since the column rank of G must then exclude the degrees of freedom corresponding to translation and rotation, which are part of the nullspace. Measuring a lower rank than this generally indicates that either the constraint equations are redundant in some fashion (and the bound is valid), or that the constraints are deficient.

When the constraints are deficient, each missing component of gradient rank (missing vectors from the row space) accounts for a constraint that is not being applied. For the gradient matrix of Eq. 4.52, when y_3 is set to zero, G_t will be

$$G_t = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & x_2 - x_3 & x_3 - x_2 & 0 & 0 & 0 \\ -x_3 & 0 & x_3 & 0 & 0 & 0 \end{pmatrix} \tag{4.58}$$

which is obviously rank two, falling one short of the expected rank of three (achieved when $y_3 \neq 0$). This missing component of rank in G_t represents the degree of freedom identified
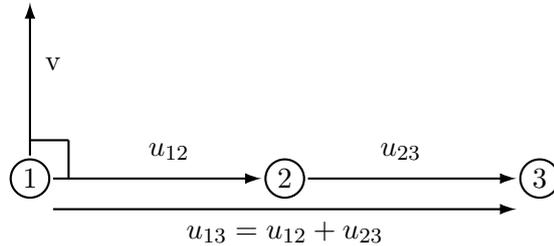
Figure 4.3: Vectors used to form constraints for three collinear antennas.

in Eq. 4.56, which has moved from the row space of G_t to its nullspace.

To eliminate the rank deficiency, a new constraint which preserves the array geometry in directions off the axis of collinearity must be devised, and its gradient appended to G_t.

4.2.6 Repairing the deficiency

If the existing distance constraints are considered as constraints on the lengths of vectors u_{ij} in the xy-plane which point between the elements of the array, e.g.,

$$u_{ij} = \left(x_j - x_i, \ y_j - y_i \right)^T \in \mathbf{R}^2, \tag{4.59}$$

then Fig. 4.3 illustrates the familiar distance constraints between three collinear antennas, along with a new vector v introduced to form the new constraint. The vector v is chosen so as to be perpendicular to u_{12}, and if the array is to remain collinear under perturbation, then u_{13} and u_{23} should also be perpendicular to it as well. This requirement of shared orthogonality with v is the constraint that we seek; while the existing distance equations constrain the length of the u vectors,[7]

$$u_{12}^T u_{12} = d_{12}^2, \qquad u_{23}^T u_{23} = d_{23}^2, \qquad u_{13}^T u_{13} = d_{13}^2 \tag{4.60}$$

the additional "inner product" constraint, to enforce rigidity, should require v to be orthogonal to one of either u_{23} or u_{13}, since only one new constraint is needed. Choosing u_{23} results in the orthogonality constraint

$$v^T u_{23} = 0. \tag{4.61}$$

---

[7]Taking the square root is unnecessary, as the distance-squared constraint has the same gradient, with just a different magnitude, as the normal constraint. The gradients' magnitudes are not important since the constraint approach only seeks to form the orthonormal basis ($\tilde{V}$ in Eq. 4.50) for the constrained subspace, and so only the directions of vectors spanning the space matter.

Dropping in the coordinates from Fig. 4.2, $v = (y_1 - y_2 \quad x_2 - x_1)^T$, the new constraint has gradient form (in which the substitution for the collinear case $y_3 = y_2 = y_1 = x_1 = 0$ has been made):

$$\nabla \left(v^T u_{23} \right) = \begin{pmatrix} 0 & 0 & 0 & x_3 - x_2 & x_1 - x_3 & x_2 - x_1 \end{pmatrix}. \tag{4.62}$$

Equating the total differential to zero as in Eq. 4.53 results in the perturbation restriction

$$\frac{dy_1 - dy_2}{dy_1 - dy_3} = \frac{x_3}{x_2} = \frac{d_{13}}{d_{12}} \tag{4.63}$$

which fixes the ratio x_3/x_2; that is, perturbations perpendicular to the axis of collinearity in the y direction are allowed only such that the ratio of distances x_3/x_2 along the axis of collinearity are kept constant. For the collinear case, this fixes the ratio d_{12}/d_{23}, and thus fixes the geometry. The new constraint couples the existing constraints with the space in which the geometry is embedded, and restores the gradient to the proper rank.

For the general case of N collinear antennas in two dimensions, a nondeficient set of constraints need only include the $N - 1$ nearest-neighbor distance constraints

$$u_{i,i+1}^T u_{i,i+1} = d_{i,i+1}^2 \quad i = 1 \ldots N \tag{4.64}$$

in addition to $N - 2$ inner product constraints

$$v^T u_{i,i+1} = 0, \quad i = 2 \ldots N \tag{4.65}$$

For each additional spatial dimension (beyond the plane), more inner product constraints are required to couple the geometry to the embedding space, and we simply add another set of inner product constraints of the form of Eq. 4.65, for each vector v_i that we find which is orthogonal to the array and the v vectors which came before it. This approach also works for coplanar geometries (with more than three antennas) in three dimensions, in which case one v vector (perpendicular to the plane) is required. In Section 4.2.9, we will demonstrate that the repair prescribed here is necessary and correct when we wish to constrain antenna arrays to be collinear.

In the next section, we show that the unconstrained Fisher matrix has the same rank deficiency in its range as that shown for the constraint gradient in the previous section. However, the rank deficiency for the unconstrained bound does not need repair.

4.2.7 Deficiency of the unconstrained bound

With the deficiency in the constraints for the collinear case demonstrated, it may also be shown that the unconstrained Fisher matrix constructed for a group of antennas (isolated from the larger whole), before any application of constraints, must have the same rank deficiency as the gradient matrix which encodes rigid geometry constraints for those antennas. In contrast to the constrained bound however, the unconstrained CRB remains valid when antennas become collinear, as the rank deficiency simply indicates that some coordinate information has become inestimable.

Proof that the unconstrained Fisher matrix shares the same rank deficiency is not obvious, but follows from the statement of the Fisher information matrix (in Eq. 4.1), using the outer product form of the bound:

$$F_\theta = \mathrm{E}_z \left[\left(\nabla_\theta \ell_z(z|\theta) \right) \left(\nabla_\theta \ell_z(z|\theta) \right)^T \right] = \mathrm{E}_z \left[\left(\nabla \ell \right) \left(\nabla \ell \right)^T \right] \tag{4.66}$$

Because our measurements in this form of the bound are the interantenna distances z and we are taking the gradient with respect to the coordinate parameters θ, the CRB construction, in effect, is composing F_θ from the same vectors $\nabla \ell$ as are produced by the gradient of the distance constraints in the constrained bound. In the following analysis, we will show that $\nabla \ell$ is indeed composed of these distance gradients, and that by the properties of the outer product and expectation, F_θ can have no components of rank beyond what is provided by those vectors. First, we express the log-likelihood function ℓ in terms of the measurements z, parameters θ, and covariances σ_{ij}^2. In the case of independent measurements, ℓ is simply a summation and may be written (ignoring vanishing terms):

$$\ell(z|\theta) = \sum_{i,j} \frac{-(\tilde{z}_{ij} - \mu_{ij}(\theta))^2}{2\sigma_{ij}^2}, \qquad \text{where } \mu_{ij}(\theta) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \tag{4.67}$$

Setting $\sigma_{ij}^2 = 1$ (to be restored later), and distributing the gradient inside the summation results in

$$\nabla_\theta \ell(z|\theta) = \sum_{ij} \tilde{z}_{ij} \nabla \mu_{ij}(\theta) - \frac{1}{2} \nabla \mu_{ij}^2(\theta). \tag{4.68}$$

Applying the gradient to the distance (μ_{ij}) and distance-square (μ_{ij}^2) terms shows that $\nabla \ell$

is a sum of vectors:

$$
\nabla\ell = \sum_{i,j}\left[\tilde{z}_{ij}\begin{pmatrix}\vdots\\(x_i-x_j)/D_{ij}\\\vdots\\(x_j-x_i)/D_{ij}\\\vdots\\(y_j-y_j)/D_{ij}\\\vdots\\(y_i-y_j)/D_{ij}\\\vdots\end{pmatrix}-\frac{1}{2}\begin{pmatrix}\vdots\\(2x_i-2x_j)\\\vdots\\(2x_j-2i_i)\\\vdots\\(2y_j-2y_j)\\\vdots\\(2y_i-2y_j)\\\vdots\end{pmatrix}\right] = \sum_{i,j}\left[\left(\frac{\tilde{z}_{ij}}{D_{ij}}-1\right)\begin{pmatrix}\vdots\\x_i-x_j\\\vdots\\x_j-x_i\\\vdots\\y_j-y_j\\\vdots\\y_i-y_j\\\vdots\end{pmatrix}\right],
$$
(4.69)

where the column vectors have zero entries where not specified, and D_{ij} is the non-random form of the distance measurement arising from the $\nabla\mu_{ij}(\theta)$ term. These are the same vectors that comprise the rows of the gradient matrix for distance constraints and from which we wish to show the Fisher matrix is constructed.

To compute the outer product and expectation, consider the individual likelihood terms of $\nabla\ell$ for a three antenna system, in two dimensions. There will be three unique nonzero terms corresponding to the pairwise distances:

$$
\nabla\ell = \left(\frac{\tilde{z}_{12}}{D_{12}}-1\right)\begin{pmatrix}x_1-x_2\\x_2-x_1\\0\\y_1-y_2\\y_2-y_1\\0\end{pmatrix}+\left(\frac{\tilde{z}_{13}}{D_{13}}-1\right)\begin{pmatrix}x_1-x_3\\0\\x_3-x_1\\y_1-y_3\\0\\y_3-y_1\end{pmatrix}+\left(\frac{\tilde{z}_{23}}{D_{23}}-1\right)\begin{pmatrix}0\\x_2-x_3\\x_3-x_2\\0\\y_2-y_3\\y_3-y_2\end{pmatrix}
$$
(4.70)

Using an abbreviation g_{ij} for each term, $\nabla\ell$ may be shortened to

$$
\nabla\ell = g_{12} + g_{13} + g_{23},
$$
(4.71)

and the argument to the expectation may then be stated as

$$
\begin{aligned}
(\nabla\ell)(\nabla\ell)^T &= (g_{12} + g_{13} + g_{23})(g_{12} + g_{13} + g_{23})^T \\
&= (g_{12} + g_{13} + g_{23})(g_{12}^T + g_{13}^T + g_{23}^T) \\
&= g_{12}g_{12}^T + g_{13}g_{13}^T + g_{23}g_{23}^T \\
&\quad + g_{12}g_{13}^T + g_{12}g_{23}^T \\
&\quad + g_{13}g_{12}^T + g_{13}g_{23}^T \\
&\quad + g_{23}g_{12}^T + g_{23}g_{13}^T.
\end{aligned}
\tag{4.72}
$$

Each of the $g_{ij}g_{kl}^T$ terms results in the product of their leading scale factors multiplied by the outer product of the respective column vectors (subsequently abbreviated to q, and restoring the $1/\sigma_{ij}^2$ factor dropped in Eq. 4.68):

$$
\begin{aligned}
g_{ij}g_{kl}^T &= \frac{1}{\sigma_{ij}^2}\left(\frac{\tilde{z}_{ij}}{D_{ij}} - 1\right)\frac{1}{\sigma_{kl}^2}\left(\frac{\tilde{z}_{kl}}{D_{kl}} - 1\right)q_{ij}q_{kl}^T \\
&= \frac{1}{\sigma_{ij}^2\,\sigma_{kl}^2}\left(\frac{\tilde{z}_{ij}\tilde{z}_{kl}}{D_{ij}D_{kl}} - \frac{\tilde{z}_{ij}}{D_{ij}} - \frac{\tilde{z}_{kl}}{D_{kl}} + 1\right)q_{ij}q_{kl}^T
\end{aligned}
\tag{4.73}
$$

And the expectation of any one term is in general (recall $\mathrm{var}(\tilde{z}_{ij}) = \sigma_{ij}^2$ and $\mathrm{E}[\tilde{z}_{ij}] = D_{ij}$),

$$
\begin{aligned}
E_z\left[g_{ij}g_{kl}^T\right] &= \frac{1}{\sigma_{ij}^2\,\sigma_{kl}^2}\left(\frac{\rho_{ij,kl}\sigma_{ij}\sigma_{kl} + \mu_{ij}\mu_{kl}}{D_{ij}\,D_{kl}} - \frac{\mu_{ij}}{D_{ij}} - \frac{\mu_{kl}}{D_{kl}} + 1\right)q_{ij}q_{kl}^T \\
&= \frac{1}{\sigma_{ij}^2\,\sigma_{kl}^2}\left(\frac{\rho_{ij,kl}\sigma_{ij}\sigma_{kl}}{D_{ij}\,D_{kl}}\right)q_{ij}q_{kl}^T.
\end{aligned}
\tag{4.74}
$$

For the entire expectation, we note that the nine terms of Eq. 4.72 fall into two groups based on whether their ij and kl indices match, that is

$$
E_z\left[(\nabla\ell)(\nabla\ell)^T\right] = E_z\left[\sum_{ij}g_{ij}g_{ij}^T + \sum_{\substack{ijkl \\ i\neq k, j\neq l}}g_{ij}g_{kl}^T\right].
\tag{4.75}
$$

For the matching case (three terms) the expected value is,

$$
E_z\left[g_{ij}g_{ij}^T\right] = \left(\frac{1}{\sigma_{ij}^2\,D_{ij}^2}\right)q_{ij}q_{ij}^T,
\tag{4.76}
$$

and in the nonmatching case we have, in general,

$$
E_z\left[g_{ij}g_{kl}^T\right] = \left(\frac{\rho_{ij,kl}}{\sigma_{ij}\,\sigma_{kl}\,D_{ij}\,D_{kl}}\right)q_{ij}q_{kl}^T,
\tag{4.77}
$$

which will be zero in GAC since we have assumed uncorrelated distance measurements in the definition of $\ell(z|\theta)$ in Eq. 4.67. If measurement correlation is assumed, however, the expansion of $\ell(z|\theta)$ will have more terms, and the coefficients of Eqs. 4.76–4.77 will change, but the vectors $q_{ij} = \nabla\mu_{ij}(\theta)$ will not.

Thus, in general, the Fisher matrix is simply a sum of outer products of vectors,

$$\mathrm{E}_z\left[(\nabla\ell)(\nabla\ell)^T\right] = F_\theta = \sum_{ijkl} C_{ijkl}\ q_{ij}q_{kl}^T \tag{4.78}$$

where C_{ijkl} is the appropriate coefficient for each term. A summation of outer products of vectors cannot result in a matrix whose range does not contain the vectors composing those outer products (the vectors in Eq. 4.69) [13, p. 6].

The unconstrained Fisher matrix is thus composed from the same vectors as those which are the gradient of the distance constraints, as in Eq. 4.48, and therefore must be missing the same vectors from its construction.

While deficiency in the constrained bound necessitated repair, for the unconstrained bound the Fisher matrix is rank deficient for good reason and does not need repair, since the estimation problem being modeled may simply be underdetermined by nature, for geometric reasons, when collinearity occurs. For example, the unconstrained Fisher matrix for three collinear antennas in the plane, arranged along the x axis, is singular because the distance measurements simply offer no information about the antennas' off-axis y coordinates, when the antennas are collinear. In cases where collinear antennas are part of a larger system whose geometry is to be estimated, the estimation problem's deficiency depends on the set of measurements which are available. In Section 4.2.9, the cases in which the unconstrained bound is and is not deficient are demonstrated.

4.2.8 Reparameterization of the constrained bound

To verify that the augmented constraints for the collinear case indeed are correct, an alternative method which computes the CRB assuming fixed collinear subarrays was pursued, which incorporates the constraints by design via a reparameterization. This reparameterization, from an approach in Van Trees [12, p. 83] (whose notation is used here), uses the same distance measurement model but chooses a different set of "basic" parameters

to model the linear arrays. The approach essentially computes the Fisher information matrix with respect to these basic parameters, and maps the result back to the desired parameter space, the coordinates, avoiding the collinear deficiency entirely.

The basic variable approach models a collection of antennas as set of linear arrays with spacing of Δ between array members, which is assumed known. Each array's unknown parameters are its coordinates in the plane as marked by the leading element, and the array's angle with respect to the positive X axis. Thus the vector of basic parameters θ_b is:

$$\theta_b = \begin{pmatrix} ex_1 & ey_1 & \phi_1 & \cdots & ex_M & ey_M & \phi_M \end{pmatrix}^T \tag{4.79}$$

where M is the number of arrays. Next, we define g as the vector of transformation functions to the "target" estimation parameters:

$$\theta_t = g(\theta_b) = \begin{pmatrix} x_1 & = & ex_1 \\ y_1 & = & ey_1 \\ x_2 & = & ex_1 + \Delta \sin(\phi_1) \\ y_2 & = & ey_1 + \Delta \cos(\phi_1) \\ x_3 & = & ex_1 + 2\Delta \sin(\phi_1) \\ y_3 & = & ey_1 + 2\Delta \cos(\phi_1) \\ x_4 & = & ex_1 + 3\Delta \sin(\phi_1) \\ y_4 & = & ey_1 + 3\Delta \cos(\phi_1) \\ & \vdots & \\ x_{N-1} & = & ex_M \\ y_{N-1} & = & ey_M \\ x_N & = & ex_M + \Delta \sin(\phi_M) \\ y_N & = & ey_M + \Delta \cos(\phi_M) \end{pmatrix} \tag{4.80}$$

with M arrays and N antennas. Given this expression of the antenna coordinates in terms of the basic variables, the log-likelihood function of the measurements may be formed and the Fisher information matrix with respect to θ_b will be

$$F_{\theta_b} = \begin{pmatrix} F_{ex} & F_{ex,ey} & F_{bx,\phi} \\ F_{ex,ey}^T & F_{ey} & F_{by,\phi} \\ F_{ex,\phi}^T & F_{by,\phi}^T & F_{\phi} \end{pmatrix} \tag{4.81}$$

The coordinate-coordinate blocks are composed of terms similar to the vanilla distance data formulation, but with a less intuitive indexing. Each of the diagonal terms has the form

$$(F_{ex,ey})_{kk} = \sum_{i \in A(k)} \sum_{j} K_{ij} \frac{(x_i - x_j)(y_i - y_j)}{\sigma_{ij}^2 \, d_{ij}^2} \tag{4.82}$$

where K_{ij} is the correspondence matrix indicating which interantenna signals are available, and the set $A(k)$ contains the indices of antennas comprising array k; in short, the summation is taken over all those signals which involve array k. For the off-diagonal coordinate-coordinate terms the summation is negated and proceeds over a smaller set,

$$(F_{ex,ey})_{kl} = \sum_{i \in A(k)} \sum_{j \in A(l)} -K_{ij} \frac{(x_i - x_j)(y_i - y_j)}{\sigma_{ij}^2 \, d_{ij}^2}, \quad k \neq l \tag{4.83}$$

such that only the signals between array k and array l take part. For the (ex, ex) or (ey, ey) blocks, the numerators become $(x_i - x_j)^2$ and $(y_i - y_j)^2$, respectively. For the coordinate-angle block, the diagonal terms are

$$(F_{ex,\phi})_{kk} = \sum_{i \in A(k)} \sum_{j} K_{ij} \frac{(x_i - x_j)\Delta\Big((y_i - y_j)\cos(\phi_k) - (x_i - x_j)\sin(\phi_k)\Big)}{\sigma_{ij}^2 \, d_{ij}^2} \tag{4.84}$$

and for the off-diagonal the summation again is negated and proceeds over a smaller set,

$$(F_{ex,\phi})_{kl} = \sum_{i \in A(k)} \sum_{j \in A(l)} -K_{ij} \frac{(x_i - x_j)\Delta\Big((y_i - y_j)\cos(\phi_l) - (x_i - x_j)\sin(\phi_l)\Big)}{\sigma_{ij}^2 \, d_{ij}^2}. \tag{4.85}$$

For the angle-angle terms, both multiplicands become $(\cos - \sin)$:

$$(F_{\phi})_{kk} = \sum_{i \in A(k)} \sum_{j} K_{ij} \frac{\Big(\Delta\big((y_i - y_j)\cos(\phi_k) - (x_i - x_j)\sin(\phi_k)\big)\Big)^2}{\sigma_{ij}^2 \, d_{ij}^2} \tag{4.86}$$

and for the off-diagonal the summation again is negated and proceeds over a smaller set,

$$(F_{\phi})_{kl} = \sum_{i \in A(k)} \sum_{j \in A(l)} -K_{ij} \frac{\Delta\Big(y_{ij}\cos(\phi_k) - x_{ij}\sin(\phi_k)\Big)\Delta\Big(y_{ij}\cos(\phi_l) - x_{ij}\sin(\phi_l)\Big)}{\sigma_{ij}^2 \, d_{ij}^2} \tag{4.87}$$

where the shorthand terms $y_{ij} = y_i - y_j$ and $x_{ij} = x_i - x_j$ have been used so that the equation fits on the page. With all the blocks of F_{θ_b} now defined, the entire Fisher information matrix of the basic variables may be constructed. To obtain the desired covariance bound in terms

of the target variables, we make a slight modification of [12, Eq. 286] by moving the Jacobian of the transformation inside the pseudoinverse,

$$\Sigma_{\theta_t} \geq \left(H F_{\theta_b} H^T\right)^{\dagger}, \qquad \text{where } H = \left(\nabla_{\theta_b} g^T(\theta_b)\right)^{\dagger} \tag{4.88}$$

which is necessary due to the nonlinear relationship between the angle and coordinate parameters [44, p. 121]. This final step first maps F_{θ_b} into the target variable space to obtain F_{θ_t} before computing the desired covariance bound with a pseudoinverse.

To ensure that the CRB in this construction is valid for relative geometry, we may again apply the test of Stoica [40] as presented in Section 4.2.3: that the terms we wish to estimate are part of the range of F_{θ_b}. Consider a six-antenna system comprising two three-antenna arrays, with coordinate matrix X

$$X = \begin{pmatrix} 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \tag{4.89}$$

Thus in terms of the basic variables, the parameter vector θ is

$$\theta_b^T = \begin{pmatrix} ex_1 & ex_2 & ey_1 & ey_2 & \phi_1 & \phi_2 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 & 1 & 0 & -180° \end{pmatrix}, \tag{4.90}$$

with separation $\Delta = 1$. Assuming a measurement standard deviation $\sigma = 1/10$, we may compute the Fisher matrix with respect to θ_b:

$$F_{\theta_b} = \begin{pmatrix} 360 & -360 & 0 & 0 & -180 & -180 \\ -360 & 360 & 0 & 0 & 180 & 180 \\ 0 & 0 & 540 & -540 & 540 & 540 \\ 0 & 0 & -540 & 540 & -540 & -540 \\ -180 & 180 & 540 & -540 & 880 & 380 \\ -180 & 180 & 540 & -540 & 380 & 880 \end{pmatrix} \tag{4.91}$$

This matrix is rank three, and just as with the relative geometry CRB, has a nullspace spanned by three vectors: the translation and rotation degrees of freedom which we know

cannot be estimated using only distance measurements:

$$
n_{\bar{x}} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad n_{\bar{y}} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad n_{\phi} = \begin{pmatrix} -(ey_1 - \overline{ey}) \\ -(ey_2 - \overline{ey}) \\ 1 \\ ex_1 - \overline{ex} \\ ex_2 - \overline{ex} \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/2 \\ -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \tag{4.92}
$$

With a set of vectors now defined which span the entire nullspace of F_{θ_b}, we have shown that the CRB is valid for this reparameterization.

4.2.9 Verification of the repaired construction

To provide a demonstration in support of our claims and constructions made regarding the different CRB approaches, Fig. 4.4 illustrates the relative geometry error predicted by four different distance-data CRB methods in the case of an antenna geometry of four 4-element collinear arrays, as shown in Fig. 4.5(a). The four CRB curves are plotted as a function of the set of distance measurements (and thus information) used to construct the Fisher matrix in each case grows. The set of distance measurements incorporated at each stage are parameterized by the maximum allowed neighbor distance N, such that only those signals with a neighbor distance less than or equal to N are assumed available; this is depicted in Figs. 4.5(b)–(i). For the unconstrained bound in which we can only approximately constrain our arrays' geometries, distance measurements within each array are always assumed available and are assigned a "very low" variance, regardless of the value of neighbor distance being tested.

The four approaches to the CRB shown in Fig. 4.4 result in two pairs of overlapping curves. The first pair of curves, U1 and C1, represent the CRBs obtained with collinear arrays for the unconstrained and deficient constrained bounds. Computed in exact arithmetic, the two curves' error bounds are quite close in value, to roughly five significant digits, as listed in Table 4.1. For $N > 2$, U1 predicts the highest error of all CRBs; closest to it is C1, which achieves a small reduction in error via application of (deficient) constraints. While C1
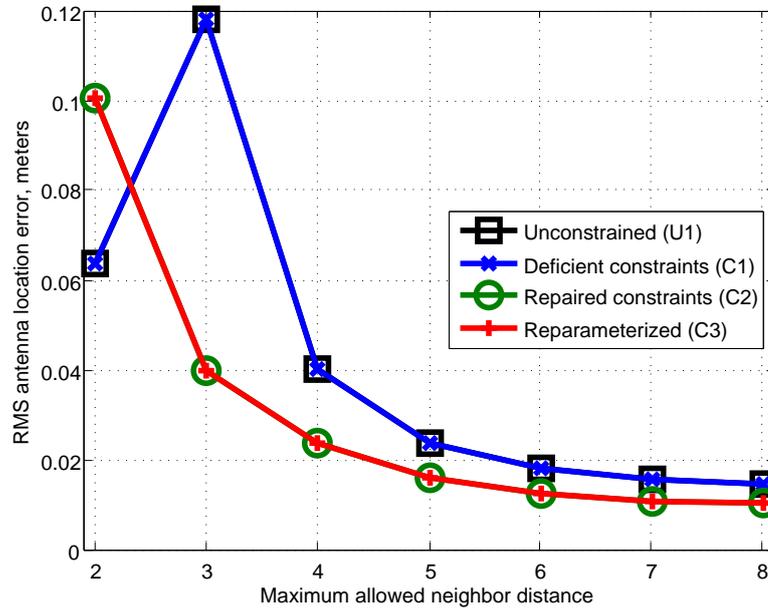
Figure 4.4: Bound results for a 16-antenna system, composed of four 4-element linear arrays, no two arrays mutually collinear. The deficient constrained bound offers negligible bound improvement; when the constraints are repaired bound results improve, and are identical to the results from a reparameterization of the problem. Measurement standard deviation was $\sigma = 1/100$ (1 cm) and same-array measurements required only for U1 were $\sigma = 1/100000$.

N	U1	C1		C2	C3
2	0.063568154̲04̲	0.063568144̲527̲		0.10045400870	0.10045400870
3	0.117922808̲40	0.117922786̲90		0.03973290455	0.03973290455
4	0.040210077̲48	0.040210067̲84		0.02381469115	0.02381469115
5	0.023551722̲697	0.023551711̲892		0.01595714872	0.01595714872
6	0.018123700̲39	0.018123691̲85		0.01251614847	0.01251614847
7	0.015455790̲64	0.015455780̲84		0.01082623961	0.01082623961
8	0.014536842̲47	0.014536832̲02		0.01029776704	0.01029776704

Table 4.1: Bound comparison numerical results. Each column shows the numbers used to generate Fig. 4.4 for increasing value of neighbor distance N. The figures shown are floating-point representations of the different CRB constructions computed in exact arithmetic. Differences between C1 & U1 are underlined, with U1 always having lower error; in exact arithmetic, U2 & U3 do not differ.

(a) Antenna layout

(b) $N = 1$, 4 sig.

(c) $N \leq 2$, 12 sig.

(d) $N \leq 3$, 24 sig.

(e) $N \leq 4$, 40 sig.

(f) $N \leq 5$, 56 sig.

(g) $N \leq 6$, 72 sig.

(h) $N \leq 7$, 88 sig.

(i) $N \leq 8$, 96 sig.
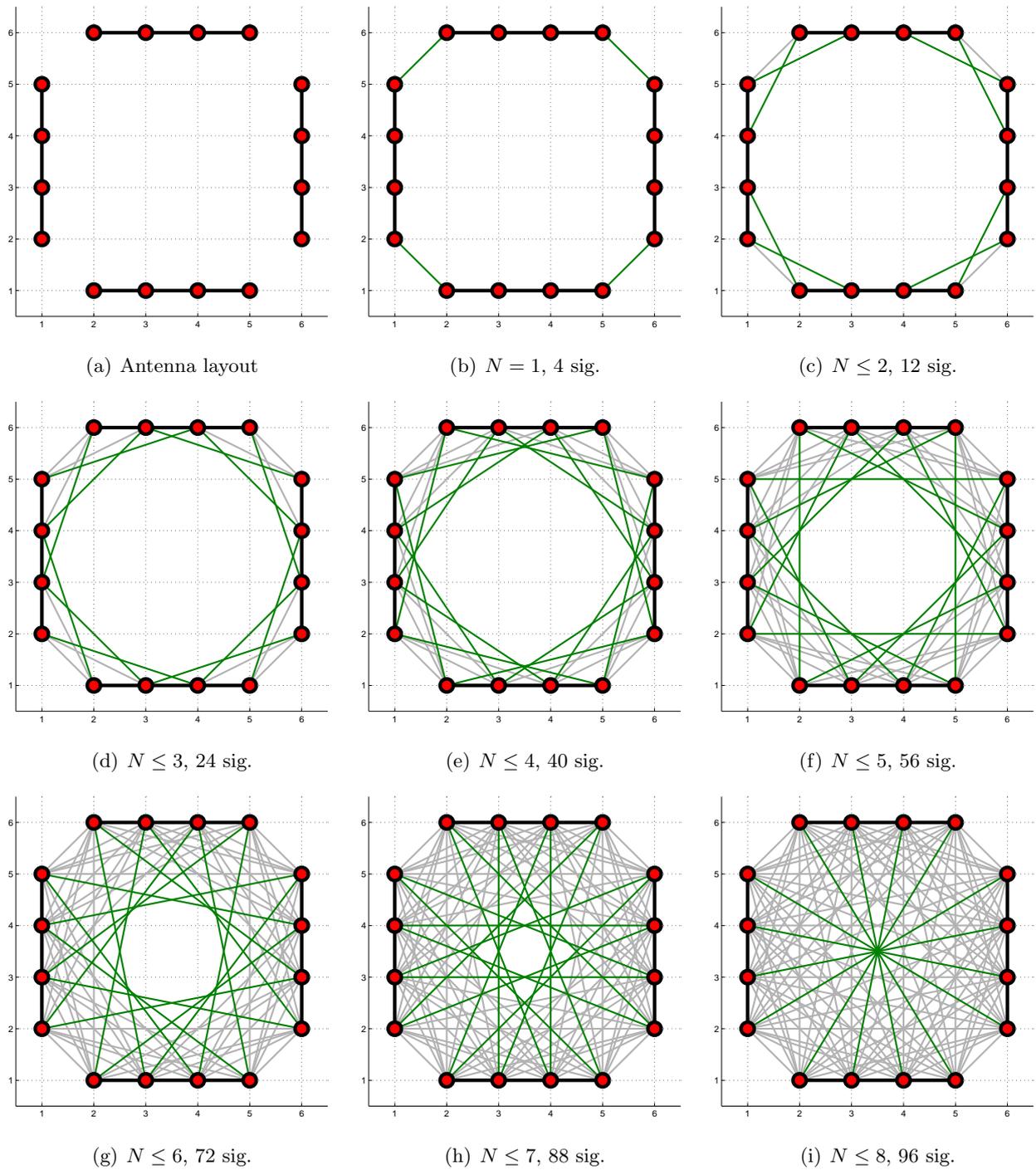
Figure 4.5: Interantenna signal groupings sorted by neighbor distance: (a) illustrates the 16-antenna layout used for simulation of the CRB; antenna positions are indicated by red circles, and black lines represent rigid arrays; (b)–(i) illustrate the signals available at each stage of increasing N, with newly added signals in green (meeting N with equality) and previously incorporated ones (less than N) in gray. Starting at $N \leq 2$, a unique solution for geometry is possible in the plane.

is invalid due to utilization of the deficient constraints, the unconstrained bound is valid for $N \geq 3$, where the 32×32 Fisher matrix has its proper rank of 29, accounting for one rotation and two translation degrees of freedom in 2-D. For $N = 2$, the rank of the Fisher matrix is 24, indicating the estimation problem is underdetermined beyond the expected degrees of freedom, and the unconstrained CRB for that case is invalid.

Further, for C1, and U1 in the $N = 2$ case (herein "U1-2"), the U1-2 & C1 bounds should be immediately suspect as CRBs based on their behavior at the lowest value of neighbor distance. The increase in the bound value between $N = 2$ to $N = 3$ is impossible statistical behavior, given the fact that the amount of information available for solution only increases with N; a valid CRB cannot increase as information is added.

The second pair of curves, C2 and C3, also overlap. This overlap demonstrates that the repaired constraints yield a result identical (in exact arithmetic) to a complete reparameterization of the problem which has the collinear array geometry constraints built in. As would be expected, as a function of neighbor distance, C2 and C3 are properly decreasing, which is proper statistical behavior: as the set of measurements available for geometry solution grows with N, the bound value must decrease, regardless of the quality (variance) of the measurements incorporated at each value of neighbor distance.

When many measurements are assumed available and N increases, the gap between the two pairs of curves shrinks, since the collinearities in the antenna arrangement of Fig. 4.5(a) that cause the deficiency only occur between antennas whose signals have neighbor distances of 2–4. Above $N = 4$, none of the measurements that are incorporated in the bound are between antennas which are part of a larger collinear arrangement, and so the discrepancy is less pronounced there. In this $N > 4$ region, we observe the difference between approximate modeling of the constraints by U1 and exact modeling of them by curves C2 and C3.

In summary, we have now defined variations on the Cramer-Rao bound (in two dimensions) based on distance measurements for four distinct cases:

1. The unconstrained bound. When absolute position information is not assumed, the Fisher matrix is necessarily singular, and a valid covariance bound for *relative* geometry estimates may still be obtained from the Fisher information matrix constructed via Eqs. 4.12–4.17. When nontrivial collinearity is achieved among any set of antennas,

the unconstrained bound is missing the same vector summands from its construction as were identified to be missing from the gradient of the distance constraints for those collinear antennas. However, these missing vectors do not indicate bound deficiency, but only the potential for deficiency if a high proportion of the measurements used for a geometry solution are between antennas which are part of the same collinear arrangements. In the cases studied in this thesis, this occurs only at the lowest value of neighbor distance, where the $N = 2$ case was shown to be deficient in this manner.

2. The deficient constrained bound. To handle the case in which our antenna arrays are assumed to have known geometries, we may use the constrained CRB construction of Eqs. 4.48–4.50. In this construction, the constraints are applied via the gradient of distance equations for the antennas within each array. Unfortunately, when the arrays, whose geometries are to be fixed, become collinear, this construction fails to apply the constraints in full, and is deficient.

3. The repaired constrained bound. To repair the demonstrated deficieny in the constrained bound for collinear arrays, a new constraint, Eq. 4.61, was introduced by this thesis. This new constraint augments the gradient matrix, and restores it to proper (column) rank. The repair applies in general to "low rank" geometries whose dimension is smaller than the ambient space in which they are embedded, such as arrays constrained to be planar in three dimensions.

4. A reparameterized constrained bound. By reparameterizing the estimation problem posed by GAC which assumes rigid arrays, we may incorporate the collinear array constraints by design and avoid the deficiency, while also verifying that the repairs made to the constraints in the collinear case give identical results.

These conclusions also apply to the parameter subspace corresponding to coordinate information in the sampled frequency-domain data CRB developed in the next section.

4.3 Sampled frequency-domain data CRB formulation

In this section, we derive the bound for antenna location using GAC assuming the measurements are the frequency domain samples of the complex amplitudes of multicarrier signals (defined in section 2.3) exchanged between every pair of antennas.

First, we extend the parameter vector θ to contain the real-valued amplitude and phase of each interantenna signal, and the antenna coordinates:

$$\theta = \begin{pmatrix} a_{ij} & \phi_{ij} & x_i & y_i \end{pmatrix}, \quad i, j = 1 \cdots N \tag{4.93}$$

for a total of $2N$ coordinate parameters and $N(N-1)$ unique signal parameters given N antennas in two dimensions. The direct-path signal between antennas i and j is written as

$$y_{ijk} = a_{ij} \exp\left(-j\, 2\pi f_k\, d_{ij}/c - j\phi_{ij}\right), \tag{4.94}$$

where i and j index antennas and the new subscript k corresponds to frequency (carrier) index. The complex-valued measurement is assumed corrupted by complex-valued zero-mean white Gaussian noise with variance σ_F^2:

$$\tilde{z}_{ijk} = y_{ijk}(\theta) + \tilde{n}_{ijk}, \qquad \tilde{n} = \mathcal{N}(0, \sigma_F^2), \qquad \mathrm{E}[\tilde{n}\tilde{n}^H] = \sigma_F^2 I. \tag{4.95}$$

resulting in probability density function

$$p_z(z|\theta) = \prod_{i,j,k} \frac{1}{\pi\sigma_F^2} \exp\left(\frac{-(z_{ijk} - y_{ijk})^*(z_{ijk} - y_{ijk})}{\sigma_F^2} \right). \tag{4.96}$$

Upon substitution of y_{ijk}, elimination of terms which will vanish in the Hessian, and replacing $2\pi f_k$ with ω_k, the log-likelihood becomes

$$\ell_z(z|\theta) = \frac{-1}{\sigma_F^2} \sum_{i,j,k} a_{ij}^2 - a_{ij} \exp\left(j\omega_k d_{ij}/c + j\phi_{ij}\right) z_{ijk} - a_{ij} \exp\left(-j\omega_k d_{ij}/c - j\phi_{ij}\right) z_{ijk}^*. \tag{4.97}$$

The Fisher information matrix takes on the following (symmetric) block structure with some of the partial derivatives having zero value by inspection:

$$F_\theta = \begin{pmatrix} F_a & 0 & 0 & 0 \\ 0 & F_\phi & F_{\phi x} & F_{\phi y} \\ 0 & F_{x\phi} & F_x & F_{xy} \\ 0 & F_{y\phi} & F_{yx} & F_y \end{pmatrix} = \begin{pmatrix} F_a & 0 & 0 & 0 \\ 0 & F_\phi & F_{\phi x} & F_{\phi y} \\ 0 & F_{\phi x}^T & F_x & F_{xy} \\ 0 & F_{\phi y}^T & F_{xy}^T & F_y \end{pmatrix} \tag{4.98}$$

Starting with a and ϕ, the first two blocks along the diagonal of F_θ are themselves diagonal:

$$(F_a)_{ll} = K(l)\frac{2M}{\sigma_F^2}, \qquad (F_\phi)_{ll} = K(l)\frac{2\,a_l^2 M}{\sigma_F^2}, \qquad (4.99)$$

where $K(l)$ is one when the lth term, as counted along the lower triangle of K_{ij}, means the ij^{th} signal is used for processing, and M is the *number* of carriers. In other words, each value of l indicates the only values for i and j for which the lj^{th} term is nonzero. Here, l and m are used to index over the $N(N-1)/2$ interantenna signal parameters $\{a, \phi\}$, and index variables i and j are used for indexing the N antenna parameters $\{x, y, z\}$.

The off-diagonal blocks involving ϕ are expressed in general as,

$$(F_{\phi x})_{lj} = K(l)\,\frac{2a_l^2}{\sigma_F^2}\,\frac{(x_i - x_j)}{c\,d_{ij}}\left(\sum_k \omega_k\right), \qquad (4.100)$$

The $F_{\phi y}$ term follows from $F_{\phi x}$ by replacing x with y:

$$(F_{\phi y})_{lj} = K(l)\,\frac{2a_l^2}{\sigma_F^2}\,\frac{(y_i - y_j)}{c\,d_{ij}}\left(\sum_k \omega_k\right). \qquad (4.101)$$

Next, the Fisher blocks with respect to coordinate parameters are,

$$(F_{xy})_{ij} = \sum_{\substack{n \\ K_{in}\neq 0}}\left(\frac{2a_{in}^2}{\sigma_F^2}\,\frac{(x_i - x_n)(y_i - y_n)}{c^2\,d_{in}^2}\sum_k \omega_k^2\right), \quad i = j, \qquad (4.102)$$

and for $i \neq j$,

$$(F_{xy})_{ij} = K_{ij}\left(\frac{-2a_{ij}^2}{\sigma_F^2}\,\frac{(x_i - x_j)(y_i - y_j)}{c^2\,d_{ij}^2}\sum_k \omega_k^2\right), \quad i \neq j, \qquad (4.103)$$

where the $i = j$ case corresponds to "summation" likelihoods involving differentiation with respect to parameters corresponding to the same antenna, and the $i \neq j$ case corresponding to "singleton" likelihoods involving parameters from different antennas. For the diagonal blocks, we simply replace $(x_i - x_n)(y_i - y_n)$ with $(x_i - x_n)^2$ or $(y_i - y_n)^2$ to obtain F_x or F_y, respectively. For F_x we have

$$(F_x)_{ij} = \sum_{\substack{n \\ K_{in}\neq 0}}\left(\frac{2a_{in}^2}{\sigma_F^2}\,\frac{(x_i - x_n)^2}{c^2\,d_{in}^2}\sum_k \omega_k^2\right), \quad i = j, \qquad (4.104)$$

and for $i \neq j$,

$$(F_x)_{ij} = K_{ij}\left(\frac{-2a_{ij}^2}{\sigma_F^2}\,\frac{(x_i - x_j)^2}{c^2\,d_{ij}^2}\sum_k \omega_k^2\right), \quad i \neq j; \qquad (4.105)$$

and for F_y we have

$$(F_y)_{ij} = \sum_{\substack{n \\ K_{in} \neq 0}} \left(\frac{2a_{in}^2}{\sigma_F^2} \frac{(y_i - y_n)^2}{c^2 \, d_{in}^2} \sum_k \omega_k^2 \right), \quad i = j, \tag{4.106}$$

and for $i \neq j$,

$$(F_y)_{ij} = K_{ij} \left(\frac{-2a_{ij}^2}{\sigma_F^2} \frac{(y_i - y_j)^2}{c^2 \, d_{ij}^2} \sum_k \omega_k^2 \right), \quad i \neq j. \tag{4.107}$$

With the elements of the Fisher matrix now defined, it can be seen that the form of these elements contains the same geometry dependence terms as the distance-data analysis. Among the two sets of amplitude and phase nuisance[8] parameters, only the interantenna signal amplitudes' parameter values impact matrix entries, while the value of the signal phases does not appear, since it has no relationship to the signal's period. Not surprisingly, the amount of Fisher information in each entry's summands is proportional to signal amplitude-squared, and thus signal-to-noise ratio a_{ij}^2/σ_F^2. Second, the entries are all proportional to the summation of subcarrier frequencies $\sum_k \omega_k$, which is a proxy for system bandwidth. Thus, any increase in signal amplitude, bandwidth, or number of carriers, will result in lower parameter estimate variances.

While this frequency-domain analysis involves the most natural input data (for SART) and desired solution parameters of the GAC problem, it only addresses the simple case of uncorrelated Gaussian errors imposed upon the samples of the channel response. This allows a useful assessment of expected positioning error as a function of signal to noise ratio within a particular antenna geometry, but does not address the principal source of error in the PPL system which Gaussian errors are ill-suited to impersonate: multipath signals.

While Gaussian errors in the above analysis are modeled by a single variance parameter, the complex character of multipath signals and rapid spatial variation make such effects difficult to realistically model without significant augmentation of the measurement model and parameter set in this analysis. As this thesis presents solution methods for antenna location which make no attempt to model multipath signal characteristics nor their impact upon position solutions, such an analysis is beyond the scope of this work.

---

[8]They are a nuisance in the sense that we must differentiate with respect to them in the CRB analysis, but bounds on variance of their estimates are not of interest.

Chapter 5

GAC Solution Methods

In this chapter we present results from novel methods that have been developed for the solution of GAC in pursuit of this thesis. The first results presented focus on algorithm performance measured in simulation, as compared to the Cramer-Rao bound developed in Chapter 4. The simulations independently model two different sources of error: frequency-domain sample measurement error, as a proxy for received signal-to-noise ratio, and time-delay error, which models variation in signal delay due to antenna pattern effects.

Following the simulation results, we present experimental results for the methods in so-called "indoor," "outdoor," and "building" GAC scenarios, using data collected with the PPL system. The results give proof that such methods, although not error-free, do not fail in the face of multipath in these realistic scenarios, and can provide accurate enough geometry solutions so as not to add significant amounts of error to transmitter location estimates in those same environments.

5.1 Algorithms

As described in Section 3.7, methods for the solution of GAC seek to produce an antenna geometry estimate, derived from the available signals, that allows transmitter location to proceed. While it is desirable for the geometry estimate to be as close as possible (in the least-squares sense) to the true geometry, ultimately the geometry error of a GAC solution need only be small enough so as not to add significant error to transmitter location

estimates. The two methods developed for this thesis are analogous to the two approaches for transmitter location developed in Chapter 2.

The first method, *interantenna ranging*, is similar to DSS+TDOA in that it takes a multi-step estimation approach. It first estimates the distance between pairs of antennas, and then passes these distance estimates to an MDS solution algorithm, which then estimates the participating antennas' geometry.

The second method, *spatial scanning*, is an extension of the SART (see Section 2.5) method. While a brute-force scan of the entire solution space is impossibly large and effectively impossible, we may employ an iterative gradient-descent method to find a geometry estimate which maximizes the SART metric. This approach utilizes a modified SART matrix which is sensitive to absolute time (distance) information.

The procedures for these two methods are described in the following two subsections.

5.1.1 Interantenna Ranging

Adopting the notation of Section 4.3, we may write the (complex-valued) measured samples of frequency-domain data y_{ijk} as a signal matrix M whose rows correspond to frequency-sample index k and columns correspond to signal "ij" between antennas i and j:

$$M = \begin{pmatrix} y_{121} & y_{131} & \cdots & y_{1N1} & y_{231} & y_{241} & \cdots & y_{2N1} & \cdots & y_{N-1,N,1} \\ y_{122} & y_{132} & \cdots & y_{1N2} & y_{232} & y_{242} & \cdots & y_{2N2} & \cdots & y_{N-1,N,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{12k} & y_{13k} & \cdots & y_{1Nk} & y_{23k} & y_{24k} & \cdots & y_{2Nk} & \cdots & y_{N-1,N,k} \end{pmatrix}. \tag{5.1}$$

Given K carriers and N antennas, M will have K rows and $N(N-1)/2$ columns (one for each antenna pair). Using a vector s to denote each column, M may be written as

$$M = \begin{pmatrix} s_{12} & s_{13} & \cdots & s_{1N} & s_{23} & s_{24} & \cdots & s_{2N} & \cdots & s_{N-1,N} \end{pmatrix}. \tag{5.2}$$

The approach starts by conducting a series of one-dimensional SART scans to determine the time delay of each signal s_{ij}. In this scan, the SART matrix M consists of two columns,

$$M = \begin{pmatrix} s_{ij} & s_{ij}^* \end{pmatrix} \tag{5.3}$$

where $*$ indicates complex conjugate. Taking the conjugate has the effect of negating the time delay information that s_{ij} represents, and so if we consider each column as a single complex sinusoid, M would be

$$M = \Big(\exp(-j\omega_k t_{ij} - j\phi_{ij}) \quad \exp(+j\omega_k t_{ij} + j\phi_{ij})\Big). \tag{5.4}$$

that is, the delays represented by s_{ij} are positive, and the delays represented by s_{ij}^* are the negative of those represented by s_{ij}. We may exploit this knowledge in a one-dimensional SART scan, which we use to determine the time delay of s_{ij}. (This analysis parallels the SART scan described in Section 2.5.) At each rephasing to determine the metric value for a particular hypothetical delay value $\hat{t}_{ij}$, M will become

$$M(\hat{t}_{ij}) = \Big(\exp(-j\omega_k(t_{ij} - \hat{t}_{ij}) - j\phi_{ij}) \quad \exp(+j\omega_k(t_{ij} - \hat{t}_{ij}) + j\phi_{ij})\Big). \tag{5.5}$$

As $\hat{t}_{ij}$ approaches the correct delay value, the two columns' direct path signal components become closer, but not equal, in periodicity (frequency), and thus cannot be linearly dependant. Linear dependence between the two columns is only achieved when $\hat{t}_{ij} = t_{ij}$ and the signals' direct-path components' periodicities are equal to one another, and the periodicities are zero (they are constant-valued DC signals).

Due to the sampled nature of the signals being processed (and the finite number of samples) the values of propagation delay that can be represented by these signals are limited to a particular interval. Since absolute time information is available, the minimum value for delay is zero, and the maximum value time delay that may be (unambiguously) measured is determined in a fashion to that relating to the equivalent of the time-domain Nyquist sampling criterion. For sampled frequency-domain data with carrier spacing Δf Hz, the range of admissible delays scanned by this approach are thus between 0 and $1/\Delta f$ seconds. For the current PPL signal spanning 550–700 MHz, $\Delta f = 1.5$ MHz, and so the ranging "window" allows measuring signal delays between zero and 665 nanoseconds (0-200 meters).

When the time delay t_{ij} of each signal s_{ij} is estimated by individual SART scans, the resulting delay estimates may be assembled into a distance measurement matrix D for which

$\hat{d}_{ij} = \hat{t}_{ij} \cdot c$ (where c is the speed of light) :

$$\hat{D} = \begin{pmatrix} 0 & \hat{d}_{12} & \hat{d}_{13} & \cdots & \hat{d}_{1N} \\ \hat{d}_{12} & 0 & \hat{d}_{23} & \cdots & \hat{d}_{2N} \\ \hat{d}_{13} & \hat{d}_{23} & 0 & \cdots & \hat{d}_{4N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{d}_{1N} & \hat{d}_{2N} & \hat{d}_{3N} & \cdots & 0 \end{pmatrix} \tag{5.6}$$

These distance estimates may then be processed via a MDS solution methods to obtain estimates for the antennas' geometry. Herein, we will consider the outcomes on application of two different MDS methods.

First, we may compute a direct MDS solution using the method summarized by Eq. 3.14. This procedure computes a coordinate estimate $\hat{X}$ derived from the (symmetric) SVD of the squared, centered, (and scaled) distance matrix:

$$-\frac{1}{2} C(\hat{D} \circ \hat{D})C \overset{\text{SVD}}{\Rightarrow} U\Sigma U^T. \tag{5.7}$$

Writing the singular values Σ and column vectors of U as

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_N) \qquad U = \begin{pmatrix} u_1 & u_2 & \cdots & u_N \end{pmatrix} \tag{5.8}$$

then the direct MDS coordinate estimate $\hat{X}$ is

$$\hat{X} = \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}) \begin{pmatrix} u_1 & u_2 \end{pmatrix}^T = \begin{pmatrix} \sqrt{\sigma_1}\, u_1^T \\ \sqrt{\sigma_2}\, u_2^T \end{pmatrix}. \tag{5.9}$$

for a two-dimensional geometry. For three dimensions, the vector $\sqrt{\sigma_3}\, u_3^T$ becomes the third row of $\hat{X}$.

In the case of the second coordinate estimation approach, the distance estimates of Eq. 5.6 are passed to an iterative MDS solution method. This iterative algorithm, provided by MATLAB's statistics toolbox as `mdscale.m`, is a gradient descent solver, which allows weighting of the entries in $\hat{D}$, missing data (signified by zero weight), and an initial solution estimate used as the starting point of iteration.

In a typical GAC problem, as we will generally have more antennas than the dimensionality of their geometry, there is a redundancy in the information contained within a distance

matrix. Assuming $N = 16$ antennas in $D = 2$ dimensions, we have $ND = 32$ total degrees of freedom; 29 of these degrees of freedom correspond to relative geometry information and completely determine the distance matrix, which in this case, will have $\binom{16}{2} = 120$ pieces of information. Thus, when the number of antennas exceeds the dimension of their geometry, there is redundancy in the distance information that may be taken advantage of.

Thus in using iterative MDS solution approach, we may generate multiple solution results as a function of available distance measurements. Many permutations of available measurements are possible; in the results presented in this thesis, we have generated one iterative MDS solution result for every valid value of neighbor distance (illustrated in Fig. 4.5). By doing so, we may investigate the character of coordinate error as a function of (maximum) neighbor distance. For example, given a system of 8 total antennas, grouped into four pairwise arrays, each entry of the distance matrix could be given a neighbor distance value (main diagonal zero "self" distances are always ignored) in a matrix B:

$$B = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 3 & 4 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 & 4 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 4 & 3 & 2 & 1 & 0 & 1 & 2 \\ 2 & 3 & 4 & 3 & 2 & 1 & 0 & 1 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 & 0 \end{pmatrix} \tag{5.10}$$

and so for a maximum neighbor distance value of two, for example, the distance measurements

considered missing may be illustrated by omitting them from the distance matrix:

$$
B_{ij} \leq 2 \quad \Rightarrow \quad
\begin{pmatrix}
\cdot & \hat{d}_{12} & \hat{d}_{13} & \cdot & \cdot & \cdot & \hat{d}_{17} & \hat{d}_{18} \\
\hat{d}_{12} & \cdot & \hat{d}_{23} & \hat{d}_{24} & \cdot & \cdot & \cdot & \hat{d}_{28} \\
\hat{d}_{13} & \hat{d}_{23} & \cdot & \hat{d}_{34} & \hat{d}_{35} & \cdot & \cdot & \cdot \\
\cdot & \hat{d}_{24} & \hat{d}_{34} & \cdot & \hat{d}_{45} & \hat{d}_{46} & \cdot & \cdot \\
\cdot & \cdot & \hat{d}_{35} & \hat{d}_{45} & \cdot & \hat{d}_{56} & \hat{d}_{57} & \cdot \\
\cdot & \cdot & \cdot & \hat{d}_{46} & \hat{d}_{56} & \cdot & \hat{d}_{67} & \hat{d}_{68} \\
\hat{d}_{17} & \cdot & \cdot & \cdot & \hat{d}_{57} & \hat{d}_{67} & \cdot & \hat{d}_{78} \\
\hat{d}_{18} & \hat{d}_{28} & \cdot & \cdot & \cdot & \hat{d}_{68} & \hat{d}_{78} & \cdot
\end{pmatrix} .
\tag{5.11}
$$

These missing measurements (indicated by a single dot) are given zero weight in the iterative MDS procedure; that is, the measurements may be passed to the algorithm, but since they have a zero weighting factor they have no impact on the outcome.

The remaining measurements passed to the routine are then either same-array ranges, or different-array ranges. Since we assume knowledge of the arrays' individual geometries, all ranges between antennas within the same array are known. While the solution method being utilized for iterative MDS does not allow such ranges to be declared fixed, they instead are assigned a higher weighting relative to the estimated ranges.

In the weighting matrix passed to `mdscale.m` we may indicate which distance measurements are missing with zero weight, those which are measured, having a weight of one, and those which are assumed known, which are assigned a weight of 5. Thus for a maximum neighbor distance of 2 in our eight-antenna example, with four two-antenna arrays (antennas

1&2, 3&4, 5&6, 7&8), the weighting matrix W would be

$$B_{ij} \leq 2 \quad \Rightarrow \quad W = \begin{pmatrix} 0 & 5 & 1 & 0 & 0 & 0 & 1 & 1 \\ 5 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 5 & 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 5 \\ 1 & 1 & 0 & 0 & 0 & 1 & 5 & 0 \end{pmatrix}. \tag{5.12}$$

This weighting between measured and known ranges was chosen empirically.

The final input to the iterative MDS procedure after the distance measurements and weighting matrix is the initial starting configuration for iteration. For the simulated and experimental results presented in this chapter, the initial solution configuration is a circular arrangement of the antennas in the plane $z = 0$ as seen in Fig. 5.1. In this geometry, the antennas are also arranged along the circumference of the circle in the same order as they were when the data was generated. In a global sense, this starting configuration is "close" to the true solution.

When coordinate estimates $\hat{X}$ become available, and if the true geometry $X_\star$ is known, then we may compute the relative geometry error between the two configurations by geometrically centering them (i.e., $X \leftarrow XC$ such that $X\vec{1} = 0$) and finding the rotation R that best fits $\hat{X}$ to $X_\star$. To do this, we solve the Procrustes problem as described in Golub & Van Loan [2, §12.4.1]. The first step is to compute the SVD of the product $\hat{X}X_\star^T$:

$$\hat{X}X_\star^T \quad \overset{\text{SVD}}{\Rightarrow} \quad U\Sigma V^T. \tag{5.13}$$

from which the optimal rotation may be derived. To produce a *pure* rotation which does not include reflection, Umeyama [36] presents an important modification to the method. Instead of the usual solution $R = VU^T$, a diagonal matrix is introduced between V and U^T:

$$R = VDU^T, \qquad D = \text{diag}(1, \cdots, 1, \det(V)\det(U)) \tag{5.14}$$
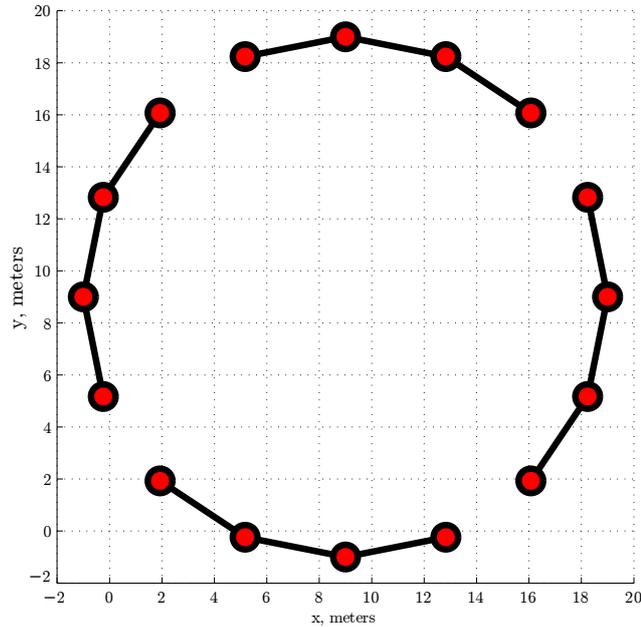
Figure 5.1: Circular configuration used as the starting point for iterative MDS. Red circles indicate antenna positions and solid lines indicate array membership.

where the sign of the last element of the main diagonal of D depends on the determinants of U and V. With the best fit rotation now determined, the (relative) RMS antenna location error is

$$e_{\mathrm{rms}} = \sqrt{\frac{1}{N} \vec{1} \left(X_\star C - \hat{X}CR \right)^2 \vec{1}}. \tag{5.15}$$

If x_i are the column vectors of $X_\star C$ and $\hat{x}_i$ are the column vectors of $\hat{X}CR$, then the RMS antenna location may also be expressed as

$$e_{\mathrm{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2}. \tag{5.16}$$

5.1.2 Spatial Scanning

The solution space for the GAC estimation problem, in general, is impossibly large for exhaustive metric imaging, as is done for transmitter location in the PPL system. Even if the availability of array geometry and orientation (as described in Section 3.1) information is assumed, the solution space is reduced in size to the point that a series of 2-D or 3-D brute force metric imagings are possible, but since such imagings must be performed sequentially,

such an approach is highly sensitive to the outcomes of the initial estimation steps and so is disadvantaged.

Thus a method which uses a SART metric for GAC, and simultaneously optimizes across all free parameters must be some other technique, such as an iterative gradient-descent method, which is the method described in this section.

The spatial scanning procedure for GAC is an extension of the SART method which directly estimates antenna geometry from sampled frequency-domain data, via an iterative optimization. (The word *spatial* in this context is used so as to distinguish it from imaging done with respect to time delay or other parameters.)

In the spatial scanning method, we search for an antenna geometry which maximizes the SART metric, given the set of all interantenna sampled frequency-data signals. Since exhaustive imaging is no longer being performed, we can no longer be sure that any local maximum reached via iteration is the global maximum value of the SART metric in the parameter space.

Because of such local maxima, it is desirable to make use of supplemental information available about the true antenna geometry to reduce the error of solutions found via such an iterative search. Utilizing information on array geometry and orientation can reduce the size of the solution space, thus eliminating many local maxima as possible solutions. Also, such information can be used to commence iteration at a geometry as close as possible to the true geometry being sought such that the iterations are likely to converge to the SART metric peak nearest to the correct solution.

In our implementation of this technique, the `fmincon` function from Matlab's optimization toolbox, which implements a constrained nonlinear minimization routine, is used to find a geometry estimate which results in a peak in the SART metric. Because our arrays are assumed fixed, we need only optimize across each array's translation and rotation angle, and so the optimization problem which `fmincon` solves is

$$\max_{\{x_1,y_1,\theta_1,\cdots\}} S(X_{\{x_1,y_1,\theta_1,\cdots\}}) \qquad (5.17)$$

where $\{x_1, y_1, \theta_1, \cdots\}$ are the translation and angle of each array in the geometry, $X_{\{x_1,y_1,\theta_1,\cdots\}}$ is the geometry whose arrays are translated and oriented according to $\{x_1, y_1, \theta_1, \cdots\}$, and

$S(\cdot)$ denotes the function which computes the SART metric on the signal data which is rephased according to the current geometry iterate. While the normal SART metric for transmitter (TDOA) location is the norm of the rephased data matrix $S(\vec{x}) = \|M(\vec{x})\|$, in GAC we may take advantage of absolute time information by computing the metric from an "extended" data matrix:

$$S(\hat{X}) = \|M(\hat{X}) \; M^*(\hat{X})\| \tag{5.18}$$

which appends a conjugated copy of all signal columns to the original rephased data, so as to take advantage of absolute time information. This stacking was described in the previous section for the case of a single signal; the multi-signal case is similar, but deserves discussion.

Were the SART metric computed only using the rephased signals $M(\hat{X})$, then any relative alignment of the signals' time delays would cause increased linear dependence within each set of signal columns $M(\hat{X})$ and $M^*(\hat{X})$. Because these two blocks of signals have time delay components which are negatives of one another (i.e., as with $e^{j\omega t}$ and $e^{-j\omega t}$), the two blocks will not be mutually linearly dependent and will contribute, as revealed by the SVD, to two separate singular values.

However, when the two sets of signals become relatively *and* absolutely aligned (as occurs when rephased in accordance with the correct array position), the rephased signals' direct path time delay components will have periodicities of zero. Despite being conjugates, the direct path signal components are now be linearly dependent, due to both sharing a periodicity of zero (the direct path signal is constant, or DC, signal), and both signal blocks will contribute, in the SVD, to the same singular value.

By extending the signal matrix in this way before the computation of the SART metric, we have provided for formation of an even larger singular value as when the signals' direct path delays are aligned to their true, absolute value at the true array translation. When the translation is nearly correct, the metric still responds with an increase in singular value when the signals become relatively aligned. Advantages to exploiting absolute time information have also been explored and demonstrated in Amendolare [45].

Thus by exploiting absolute time information, by extending the SART matrix, the array geometry solution benefits from a metric which increases when the signal delays are relatively aligned, and increases even further when such relative alignment becomes absolute

at the correct position. This effect has important consequences, as it likely allows geometry estimates to better ignore spurious linear dependencies between signal columns. But more importantly, the effect has useful geometric consequences, as the loci of signal alignments which cause metric peaks become circular when absolute time information is available, rather than hyperbolic as with relative (TDOA) time information.

This matrix extension is important because the location problem posed by GAC is more geometrically challenged (in a geometric dilution of precision sense) as compared to the case of transmitter location. In transmitter location, the receivers very nearly surround the transmitter providing low dilution of precision, while during GAC, the objects being located are always on the perimeter of the geometry, and can especially benefit from the availability of absolute time information.

5.2 Simulations

In the following two sections, we present the outcomes of simulations, which subject the two methods presented in Section 5.1 to two types of measurement error. The first simulation subjects the methods to Gaussian frequency-domain sample error as a proxy for receiver signal-to-noise ratio. The second simulation subjects the interantenna signals to random time-delay errors, which approximate, to a first order, the effects of antenna pattern-induced signal distortion.

These types of error are well-suited to simulation in the sense that they are common sources of system error, and they have simple models which allow them to be simulated in a straightforward fashion. Furthermore, the results of such simulation can be interpreted unambiguously and compared to the previously derived Cramer-Rao bounds.

5.2.1 Frequency-domain sampling error

This simulation examines GAC antenna location error as a function of frequency-domain sample measurement error variance. The antenna geometry used is a 16-antenna system, consisting of four arrays of four antennas each, as illustrated in Fig. 5.2.

In the simulation, synthetic data samples z_{ijk} are generated which model the direct path
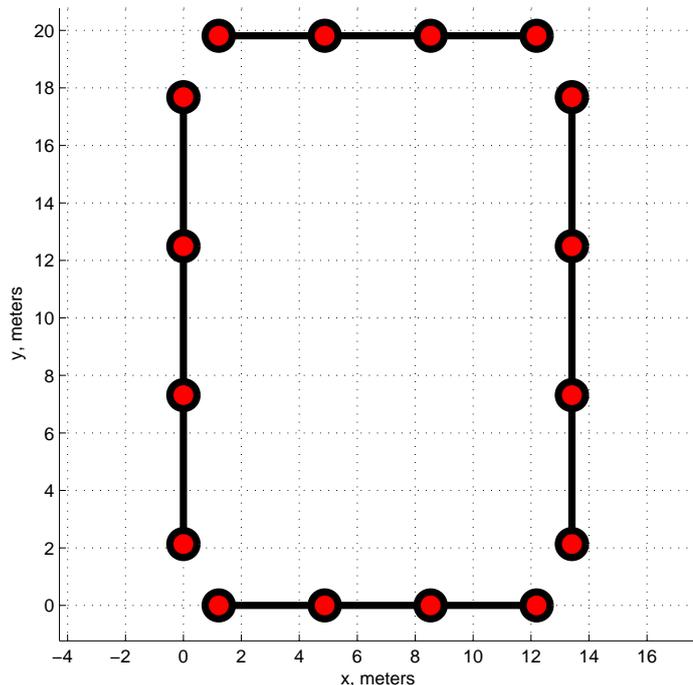
Figure 5.2: Geometry used for GAC simulations

channel frequency response corrupted by noise $\tilde{n}_{ijk}$:

$$z_{ijk} = \left[\frac{1}{d_{ij}} \exp\left(-j\,\omega_k\,d_{ij}/c - j\phi_{ij}\right) \right] + \tilde{n}_{ijk}. \tag{5.19}$$

As occurs in free-space propagation, the samples are attenuated by the inverse of the propagation distance. The noise samples $\tilde{n}_{ijk}$ are each drawn from independent, identically distributed, zero-mean, complex Gaussian random variables with variance σ^2. Each sample is constructed as the sum of two real random variables with variance $\sigma^2/2$:

$$\tilde{n}_{ijk} = \mathcal{N}(0, \sigma^2/2) + j\mathcal{N}(0, \sigma^2/2) \tag{5.20}$$

The variance σ^2 is a function of the desired signal-to-noise ratio, SNR:

$$\text{SNR} = 10\log_{10}(P_s/P_n) \qquad \Rightarrow \qquad P_n/P_s = \sigma^2 = 10^{-\text{SNR}/10} \tag{5.21}$$

which is measured at a single carrier, in units of decibels (dB). In the results to follow, the antenna location error is plotted as a function of this SNR. Due to the attenuation of $1/d_{ij}$ in Eq. 5.19, the interantenna signals will have varying SNRs as measured at each receiver,
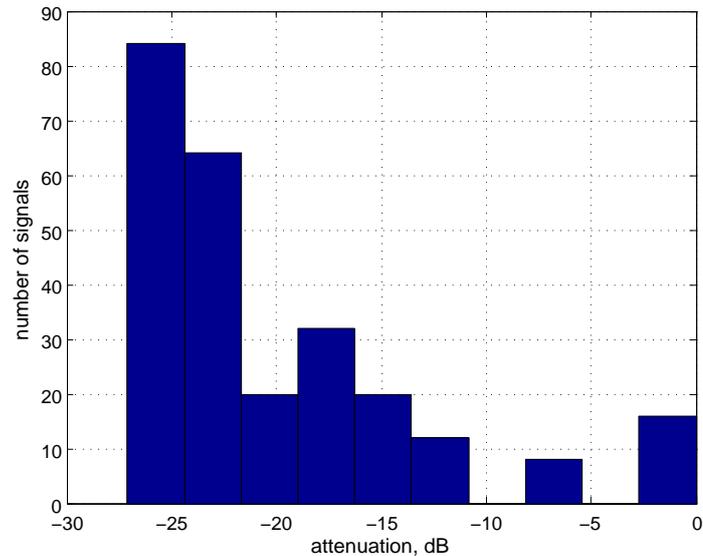
Figure 5.3: Distribution of signal attenuations applied in simulation, in the geometry of Fig. 5.2. Mean attenuation is -20 dB.

and the SNR values against which the following results are presented are given with respect to a "reference" value, measured at 1 meter distance from the transmitter, that is, with no attenuation; the SNR measured at every antenna will be less. To illustrate the distribution of the different attenuations applied to the 120 interantenna signals, Fig. 5.3 illustrates a histogram of their values, specific to the geometry of Fig. 5.2. Across all signals, the mean attenuation applied is roughly -20 dB, and so the average SNR experienced between all antennas is 20 dB lower than the 1 meter reference SNR.

The remaining parameters of the noise-free signal of Eq. 5.19 are the subcarrier frequencies ω_k and overall signal phases ϕ_{ij}. The frequencies ω_k used mimic the current PPL multicarrier signal, using the same number of carriers, with the same spacing (roughly 100 carriers spaced by 1.5 MHz), over the 550–700 MHz UHF band. Finally, the overall phase ϕ_{ij} of each signal is set to zero in simulation. While experimental data will have arbitrary overall phase, the value of ϕ_{ij} has no influence on the results of solution methods presented here since overall phase cannot alter signals' periodicities (time delay), and most importantly, the value of ϕ_{ij} does not appear in any of the terms of the frequency-data Cramer-Rao bound of section 4.3.
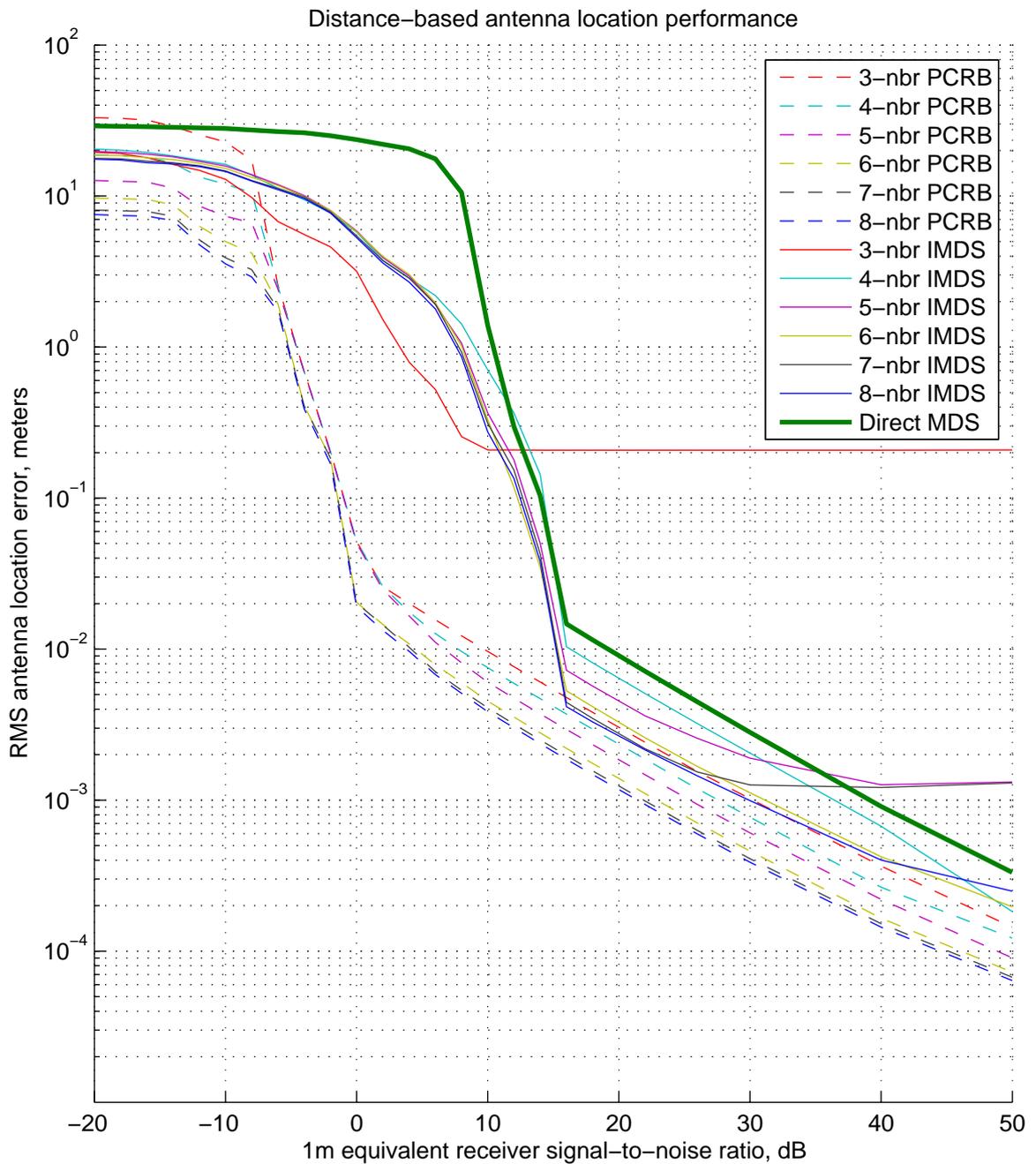
Figure 5.4: Simulation of Interantenna ranging with frequency sample error. Curves shown are performance figures for methods that process distance estimates derived from sampled frequency-domain data. Dashed lines indicate the respective constrained Cramer-Rao bound (CRB) curves, and solid lines indicate iterative MDS (IMDS) result curves, one curve for each value of maximum neighbor distance, "n-nbr", tested. The solid green line shows the result of direct MDS, equivalent to the 8-nbr case, which uses all available data.
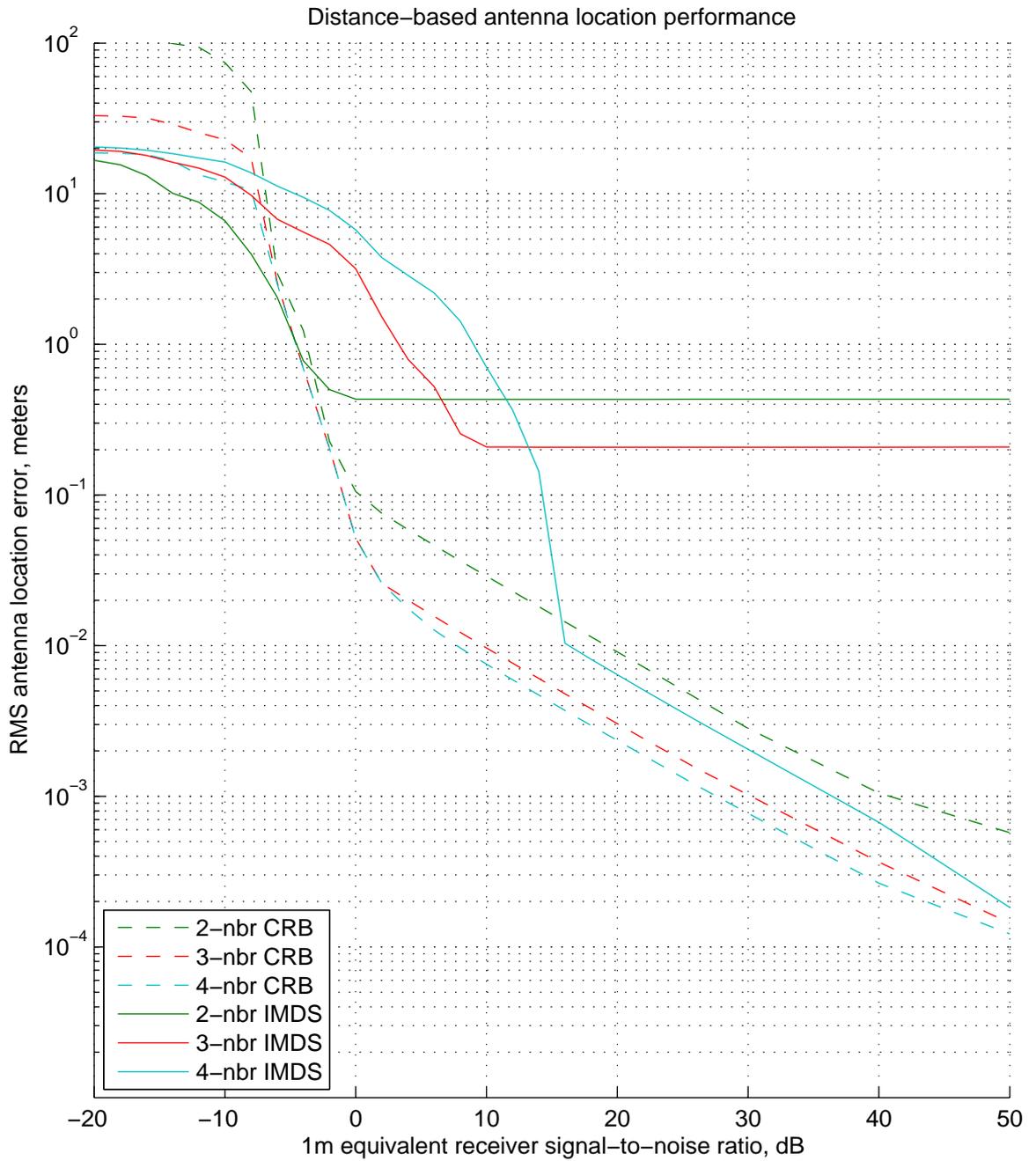
Figure 5.5: Performance in simulation of Interantenna ranging with frequency sample error for neighbor distance values 2–4. Dotted lines indicate Cramer-Rao bound (CRB) curves, and solid lines iterative MDS (IMDS) results, one curve per value of neighbor distance tested. For a maximum neighbor distances of 2, the estimation problem is underdetermined.

Simulation with frequency-sample error: Interantenna ranging Figure 5.4 shows the root-mean-square (RMS) antenna geometry errors for our interantenna ranging method as a function of signal-to-noise ratio. At each value of SNR, 100 Monte Carlo trials were used to measure average performance. Because the iterative MDS (IMDS) method may converge to a solution with missing data, multiple IMDS curves are visible for neighbor distance (abbreviated "n-nbr") values of interest.

In the figure, there are two curves of the same *color* for each value of maximum neighbor distance: a "pseudo" Cramer-Rao bound (CRB) curve, represented as a dashed line, illustrating a theoretical lower bound to performance, and an iterative MDS (IMDS) result as an unbroken line. The remaining curve (in thick green) illustrates the results of direct MDS, and uses the same data as the 8-neighbor (no missing data) case.

The CRB curves in the figure are each labelled as a pseudo CRB (PCRB) because the distance-data CRB construction used is not perfectly matched to the actual estimation process under simulation. That is, the distance-data CRB in a sense only captures the second part of the interantenna ranging method which processes distance estimates, and has no knowledge of the prior range estimation step which processes frequency samples and is the point where system error is actually injected. Also, while the errors injected at the frequency sample level themselves have a Gaussian distribution, the distribution of distance estimate error from this data is not necessarily Gaussian as the CRB assumes, since it is the output of an intermediate estimator, and not Gaussian by design. (In a forthcoming analysis, these curves will be compared with a more proper frequency-data CRB.) A final discrepancy in the comparison of IMDS results to the CRB is that IMDS only approximately constrains the arrays to be fixed; same-array distances are given higher relative weighting and their distance measurements are input without error, while the CRB results presented assume (perfectly) fixed arrays. Despite these deficits, the distance-data CRB, in nearly all cases, bounds the performance of these methods from below, and predicts better performance when higher values of neighbor distance are used. Right at 0 dB, a grouping of the bound curves can be seen, where 3-,4-, and 5-nbr PCRBs share roughly the same value, and the higher values of neighbor distance have a lower variance bound. This is due to our particular choice of geometry and array grouping in Fig. 5.2, in which no measurements between arrays

opposite one another are available for processing until a neighbor distances greater than 5 is allowed.

The behavior of the IMDS result curves may be broken down into two regions. The first region is at the values of SNR below 5 dB in which the RMS antenna geometry error exceeds 1 meter; as SNR worsens below 5 dB, the distance estimates being supplied to MDS are dominated by noise and the solutions plateau to an error commensurate with random antenna placement. In the second region, above 5 dB SNR, antenna geometry errors rapidly decrease to levels around 1 cm RMS error. Above 5 dB SNR, geometry errors parallel the CRBs, with a performance gap of no greater than 10 dB from the CRB.

It can be seen that the iterative MDS approach, in nearly all cases, surpasses the performance of direct MDS. While direct MDS is computationally simpler, iterative MDS enjoys two advantages which contribute to its improvement. The first advantage is the added weight given to the same array-distances in the distance matrix. While both iterative and direct approaches receive the same distance measurements as input, with truth values used for same-array distances, the iterative approach allows weighting of measurements, and since we have placed a higher relative weight upon such measurements which are known a priori to be error-free, this conveys an advantage. The second advantage of iterative MDS over direct MDS is also a feature of an iterative process: the use of an initial starting point for the iteration is another form of a priori information which conveys a statistical advantage.

For nbr≥ 4, it can be seen that antenna location error breaks away from the CRB curves when the SNR drops below 15 dB. This is likely due to a breakdown in the range estimation procedure, as many signals from which distances estimates are derived have SNRs below 0 dB and have begun to produce increasingly erroneous distance estimates, which in turn affects the resulting geometry estimates. Above 15 dB SNR, the IMDS and direct MDS curves parallel the CRBs.

The performance of the 3-nbr curve plateaus out beginning at 10 dB, at 2 cm error due to the behavior of the iterative MDS method. Because of the relatively small number of measurements being used for solution, the optimizer makes slow progress with each step, and terminates due to reaching the maximum number of iterations. Were IMDS executed in a 3-nbr situation with a higher limit on allowable number of iterations, the RMS antenna

location error at which performance no longer increases with SNR may be lowered. It can be seen that between 0 dB and 10 dB SNR, performance levels go below 1 m in the 3-nbr case before any other method. This behavior is not predicted by the CRB, and is due to the usage of an initialization point close to the desired optimum for the iterative method, from which the performance of the 3-nbr method benefits heavily upon.

In these curves we have omitted results from the 2-nbr case; the results for neighbor values of 2–4 compared to their respective CRBs are shown in Fig. 5.5. While the 2-nbr case is not an undesirable situation in which to solve GAC, since the signals used in solution will likely have the highest SNR and most favorable propagation conditions as compared to other signals, it is omitted from Fig. 5.4 because the unconstrained estimation problem which IMDS must solve in the 2-nbr case is underdetermined (solely due to array collinearity), and thus comparison to the CRB (for the 2-nbr case only) is invalid.

Simulation with frequency-sample error: Spatial scanning Figure 5.6 shows the results (on a logarithmic scale) of the frequency-sample error simulation using the spatial scanning method. As in the previous figure the curves are paired by color: the CRB for a particular neighbor distance is a dashed line, and the spatial scanning RMS antenna location error is a solid line. As with the previous method tested, spatial scanning was run several times, one for each value of neighbor distance.

Here, it can be seen that the CRB curves are consistent: their error variance bound decreases with increasing SNR, and also decreases with increasing neighbor distance (when more signals are used). In comparison to their CRB, the spatial scan result curves have a number of noteworthy features. First, while the curves generally have the same slope as the CRBs above 20 dB SNR, there is an obvious offset between the bound and the method results. One possible explanation for the offset is simply that the spatial scanning procedure is not taking advantage of all the information contained within the signals themselves. That is, a superior method that effectively uses all information may be identified with further investigation.

A second noteworthy feature is the range of signal-to-noise ratio over which the spatial scanning method exhibits sub-meter error, in comparison to the results from interantenna
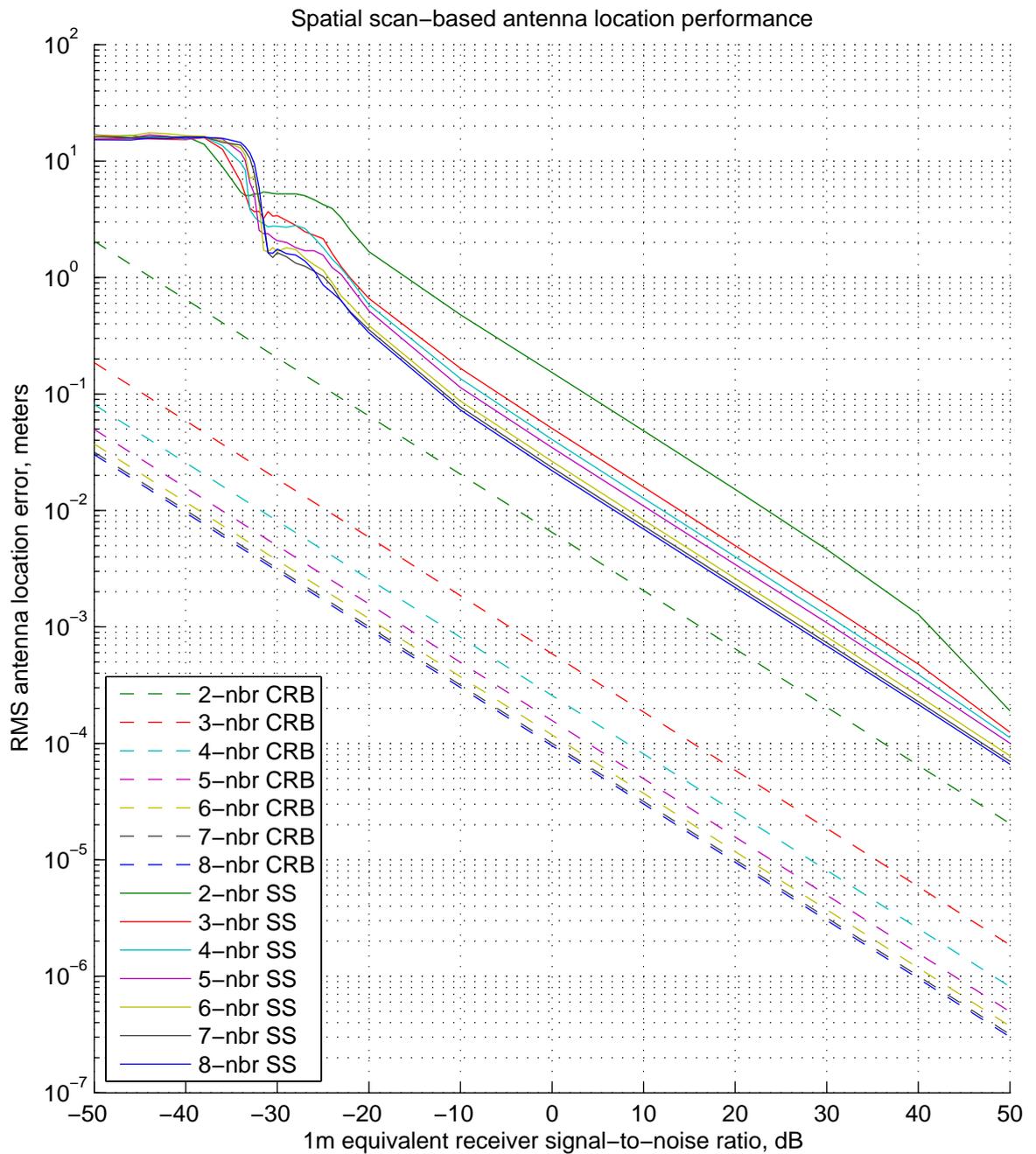
Figure 5.6: Simulation of Spatial scanning with frequency sample error. Curves shown are performance figures for methods that process sampled frequency-domain data. Dotted lines indicate the respective constrained Cramer-Rao bound (CRB) curves, and solid lines indicate Spatial scanning result curves, one curve for each value of maximum neighbor distance tested.

ranging, whose results are overlaid in Fig. 5.7. The points at which both methods generally achieve sub-meter error are separated by 30 dB. Further, spatial scanning achieves sub-meter error starting at -20 dB SNR; since the average SNR at each receiver is 20 dB lower than the 1 m equivalent SNR, the method enjoys a significant amount of processing gain.

Below -20 dB SNR, method performance exceeds 1 m geometry error, and the method begins to break down; at -30 dB the results are essentially noise, and errors are commensurate with a random geometry solution.

Regarding the gap between spatial scan performance and its respective CRB, such gaps, between Cramer-Rao bounds and the performance of particular estimators, for otherwise well formed estimators (that is, estimators that properly yield the minimum of a distance metric between the space of observables and the manifold of the observable space formed by error free observations over the range of all model parameter values, the solution manifold) can be explained as follows. If the dimensionality of the observation space is reduced by some mapping of observed data into a lower dimensional representation then locally this mapping is effectively a projection. And the minimizing parametric solution can be greatly affected by the projection away of some component of the distance from the point in observation space to the solution manifold. That is, if a point x^* on the solution manifold minimizes $d^* = \min\limits_{x} \sum_{n=1}^{N} m_n(z)$ given measurements z, and x^{**} minimizes $d^{**} = \min\limits_{x} \sum_{n=1}^{M} m_n(z)$, where $M < N$, then in general $x^* \neq x^{**}$ and $d^* \leq d^{**}$. The mapping that gives rise to this decrease in performance is induced by any data reduction that reduces the effective dimensionality of the observation space.

Yet, despite this draconian departure from full optimization complexity, for SNRs below 10 dB, it can be seen in Fig. 5.7 that this handicap is less detrimental to performance than using the iterative MDS approach, where the parameter space is not reduced, and the 270 dimensional measurement space is replaced by a 120 (maximum) dimensional distance data set. A comparison of the results of interantenna ranging versus spatial scanning seen in Fig. 5.7 shows that the spatial scanning approach yields significantly improved results over interantenna ranging in this region below 10 dB SNR.

Finally, Fig. 5.7 shows both methods' results, compared to the frequency-data CRB. While the distance-data CRB was valid to compare against the interantenna ranging method, that
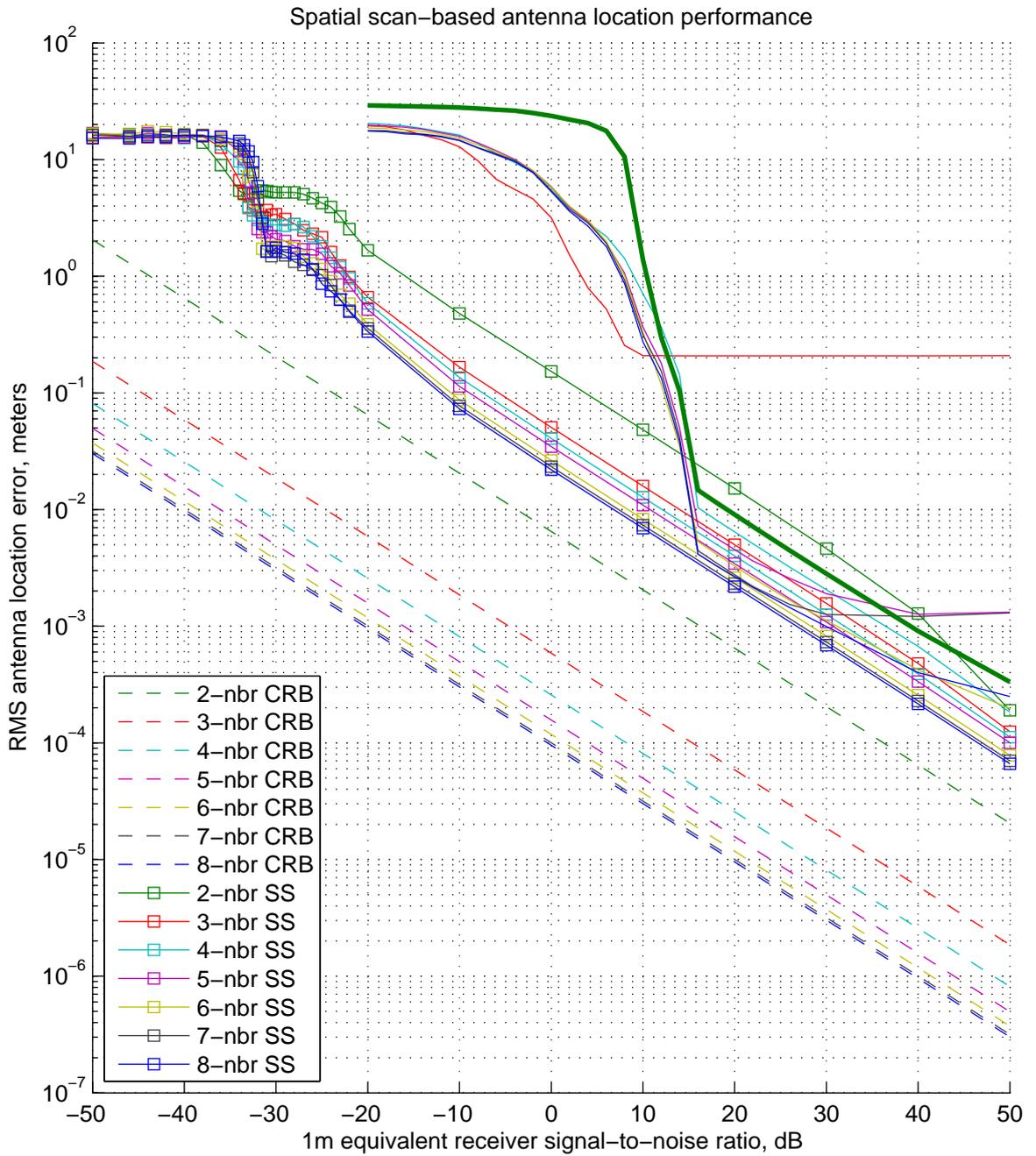
Figure 5.7: Performance of interantenna ranging and spatial scanning, compared to the CRB. Dotted line indicates the sampled frequency-data CRBs, solid lines represent interantenna ranging results (for 4-nbr and higher neighbor distances), and the square markers indicate spatial scan curves.

CRB is concerned with distance measurements and is ignorant of the underlying frequency data from which the distance estimates are made. The distance-data CRB effectively only measures the performance of iterative MDS, as parameterized by the frequency-sample signal-to-noise ratio. Because the interantenna ranging methods' results are ultimately derived from these frequency samples, it is also valid, and ultimately more correct, to compare it against the frequency-data CRB.

As was evident in the earlier graphs and discussions, both interantenna ranging and spatial scanning perform significantly more poorly than allowed by the CRB, for all test SNRs, in the sense that no method achieves the CRB. This indicates that significantly better performance, on the order of an improvement equivalent to a 30-40 dB SNR improvement is possible in theory via some technique that fully exploited the information in the frequency-sample data. Furthermore, the spatial-scanning technique, because of the parameter space reduced nature of its sequential scanning, enjoys lower RMS error than interantenna ranging at SNRs greater than 10 dB. However, interantenna ranging suffers an early break from asymptotic behavior below 10 dB and is significantly bested by spatial scanning. Spatial scanning performance suffers a similar break from its asymptotic behavior, but at progressively higher SNRs as the number of neighbors is increased; the break from the asymptotically linear (on a log-log scale) performance curve occurring at -20 dB for 3-nbr and at still lower SNRs for larger numbers of neighbors.

Note that the break from asymptotic behavior for interantenna ranging occurs at the 1 cm level of positioning accuracy and degrades to 1 m levels of error at an SNR of 3 dB for the smaller 4-nbr case, while the spatial scanning adherence to asymptotic behavior extends to the 1 m level for 3-nbr or larger, at SNRs of -15 dB or larger. Thus, we can say that in the important context of maintaining a sub-meter error, that the spatial scanning technique contributes performance that is equivalent to an improvement of 18 dB in input SNR when compared to classic interantenna ranging approaches.

5.2.2 Simulation of time-delay error

In this section, we present the results of our methods in the presence of simulated time-delay error.

While the PPL system admits many hardware effects that introduce error to the received signals, many of these effects may be calibrated away so as to not impact the location solution. However, antennas used for signal reception can induce signal distortion that is more difficult to completely eliminate, as their frequency response varies as a function of the angle of signal reception. To a first order, this signal distortion may be modeled as a superfluous signal delay which varies with angle of signal arrival.

The model for the synthetic signals to be generated in this simulation applies only a random additive delay to the true propagation delay:

$$z_{ijk} = \frac{1}{d_{ij}} \exp\left(-j\,\omega_k\,(d_{ij} + \tilde{d}_{ij})/c - j\phi_{ij}\right). \tag{5.22}$$

The random time delays are applied to each interantenna signal and are independent, identically distributed Gaussian random variables:

$$\tilde{d}_{ij} = \mathcal{N}(0, \sigma_d^2) \tag{5.23}$$

with a mean of zero and variance σ_d^2. In the simulation, we tested time delay errors of up to 1 meter standard deviation, performing 100 Monte Carlo tests at each value.

Timing error simulation results   Figure 5.8 shows the simulation results for the interantenna ranging and spatial scanning methods, plotted in linear units; spatial scanning results use square symbol markers, and the dotted line indicates the error variance predicted by the constrained distance-data CRB. Unlike the previous section, the distance-data CRB is a more proper comparison to the interantenna ranging method, since the distance (timing) errors being injected are indeed Gaussian. Since frequency-sample errors were not part of this simulation, no comparison to the associated sampled frequency-data result can be made. (No CRB was derived for frequency data solely for timing errors.)

In these results we can see that not only do the distance-data result curves closely match their respective CRBs, but that their performance in this simulation clearly and significantly surpasses the performance of the spatial scanning method. One factor in the error in the spatial scanning can be attributed to the reduced solution space which benefited it in the previous simulation. In the case of the current simulation which induces timing (or distance)
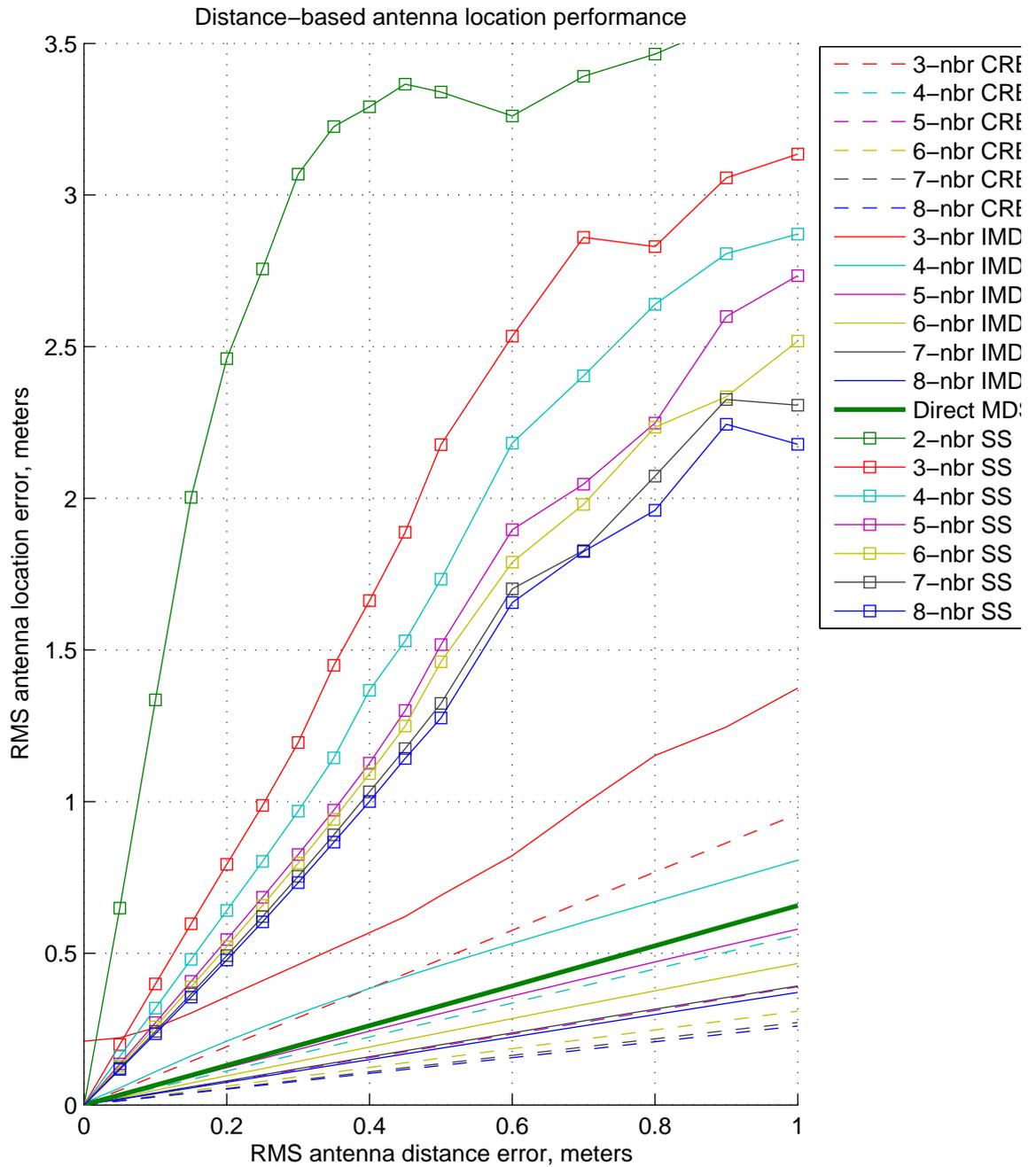
Figure 5.8: Simulation results of GAC methods with time delay error. There are three curves for each value of neighbor distance; the solid line indicates the interantenna ranging result, solid lines marked by squares indicate spatial scanning results, and the dotted lines are the error variances predicted by the constrained distance-data Cramer-Rao bound.

errors, application of these errors to the input data is very likely to have the effect of moving the position of the solution point which spatial scanning is searching for, via a SART metric, *outside of the solution space.*

For these two sets of curves, we may infer an approximate linear model for antenna geometry error as a function of RMS timing error for both methods. For neighbor distances of 4 and above, interantenna ranging incurs, on average, 0.50 m per meter of RMS timing error, while spatial scanning incurs 2.50 m of error for each meter of RMS timing error.

Clearly, if the major source of error in practical applications is angle-dependent phase response of antennas or other antenna-specific delay mechanisms, then spatial scanning is by far the less desirable choice for GAC implementation. In the following section, we look to experimental data in the indoor location context to determine which is the dominant factor to be considered, independent frequency-data noise or time delay error.

While we have simulated levels of RMS antenna timing error out to 1 meter in value, revealing two roughly linear regions, with different rates of response to increasing levels of timing error, in the current PPL system it is reasonable to assume that our RMS timing error due to antenna effects is at or below 0.3 m. We may assume this limit to RMS timing errors simply because all antennas in use have either low signal distortion as a function of angle, or have been explicitly designed with this maximum timing error of 1 foot as an antenna design constraint.

We may conclude from this result that spatial scanning for GAC is highly sensitive to timing errors, more so than interantenna ranging. It is possible that the intermediate steps in the interantenna ranging method help it mitigate timing error and produce a solution with less error.

Table 5.1: GAC Experimental Data Sets

Test Location	Setting	#	Date	RX Antenna Qty., Type	
Kaven Lawn	Outdoor	O1	6/21/2007	16	Bowtie with ground plane
Practice Field	Outdoor	O2	11/28/2007	16	Bowtie with ground plane
	Outdoor	O3	11/28/2007	16	Bowtie
	Indoor	I1	12/18/2007	16	Patch
	Indoor	I2	12/18/2007	16	Patch, alt. cabling
	Indoor	I3	12/18/2007	16	Bowtie
Odeum	Indoor	I4	12/18/2007	16	Bowtie with ground plane
	Indoor	I5	1/17/2008	16	Patch
	Indoor	I6	1/17/2008	16	Patch, intervening wall
Kaven Hall	Building	B1	6/22/2007	13	Bowtie with ground plane
Campus Ministry	Building	B2	7/12/2007	16	Bowtie with ground plane
Campus Ministry	Building	B3	3/13/2008	16	Patch
Campus Ministry	Building	B4	3/28/2008	16	Bowtie

5.3   Experimental results

In this section we present the results of testing our methods using experimental data. Using the PPL system, we have been able to capture data during location testing experiments performed at multiple locations over the course of nine months between 2007 and 2008. Table 5.1 lists the experimental data sets available to us for testing our methods for GAC.

In all of the tests, transmitter location data was collected, so we may see the accuracy of transmitter location estimates when driven by the GAC geometry solution as compared to the (very-low error, < 5cm RMS) manually measured antenna geometry.

The Outdoor and Football field test locations were both outdoor tests on level ground with unobstructed line-of-sight between antennas; Kaven Hall is a brick, concrete, and steel-beam building; the Campus Ministry is a typical three-level wooden-frame residential structure including metal appliances in the kitchen; finally, the Odeum was an indoor-to-indoor test conducted in an empty function hall.

Figures 5.9 and 5.10 show performance figures for GAC and GAC-driven transmitter location, respectively, using spatial scanning. It can be seen that GAC and transmitter location errors are the lowest in the outdoor scenario, due to the lack of significant multipath signal components and low direct path signal attenuation. GAC errors in this case are

always below 0.80 m; when using signals with a neighbor distance of three or higher, the O1 test achieves 0.30 m of error.

GAC errors for the indoor scenario become slightly worse. While, as in the outdoor tests, antennas in the indoor scenario have clear line-of-sight to one another (except for test I5), there are more multipath signals to deal with in this case due to the surrounding building structures, and so GAC is more challenged in this environment. This is likely the reason that more of the curves in this scenario exhibit higher error at a neighbor distance of 2. At higher values of neighbor distance however, their errors flatten out, except for curve I5, in which an antenna array was behind an intervening wall. Due to this intervening wall, and the more challenging propagation conditions for the signals that pass through it, more signals are required to achieve the same error levels as the other cases, and 0.80 m error is reached by $N = 5$.

In the third scenario, where our antenna arrays are placed around a building structure, and only have clear line-of-sight to some of their neighbors, GAC errors are higher than the two previous scenarios. This is likely due to less favorable propagation conditions between all pairs of antennas, in which more multipath signals are present, and attenuation of the direct path is guaranteed. Due to these worsened conditions, GAC errors are higher, and more signals are required to achieve the lowest errors of roughly between 1.00 m and 0.80 m.

Generally, when comparing the GAC and transmitter location errors of Figs. 5.9 and 5.10, lower GAC errors result in lower transmitter location errors, as can be seen in the similar shape shared between the GAC and transmitter location curves from the same test. However, due to the nonlinear nature of (TDOA) multilateration in general, this is not always guaranteed, as can be beween the two curves in Figs. 5.9(c) and 5.10(c) for tests B4 and B4-2, in which the $N = 2$ case achieves the highest GAC error but the lowest transmitter location error.

Figures 5.11 and 5.12 show GAC and transmitter location errors using the interantenna ranging method. In the outdoor scenario, GAC and transmitter location errors meet or exceed those resulting from spatial scanning. In the indoor scenario, GAC errors are nearly as good as for spatial scanning, again reaching 0.60 m of error, but requiring $N > 3$, rather than $N > 2$ as with spatial scanning. While GAC error for test I5 exceeds that of spatial

scanning at low values of neighbor distance, error instead worsens slightly at higher values of neighbor distance in which more through-wall distance estimates are input to MDS; spatial scanning, in contrast, achieves monotone improvement in GAC error as neighbor distance increases. However, the more unruly behavior of interantenna ranging GAC error curve I5 results in high transmitter location error at $N = 4$, but performance is equal to that of spatial scanning for $N \geq 6$.

While in outdoor and indoor scenarios, interantenna ranging is competitive with spatial scanning in terms of GAC and transmitter location error, interantenna ranging suffers generally worse GAC performance in the building scenario, which is most relevant scenario to the PPL problem. However, despite these higher GAC errors, transmitter location performance, such as in the B3 test, is nearly comparable to GAC.

We may better understand some aspects of the performance of these two methods by examining the character of the GAC and the transmitter location errors in three tests, one from each scenario, in which acceptable performance is achieved. Figures 5.13–5.15 show the results of tests O1, I6, & B3, respectively when using spatial scanning, and Figs. 5.16–5.18 show results for the same tests using interantenna ranging. (Throughout, circular symbols indicate GAC results; triangles indicate transmitter location.)

Figures 5.13 & 5.16 depict both methods' performance for the O1 test. Spatial scanning achieves a transmitter location error of 0.70 m; without GAC the RMS transmitter location error is 0.41 m. This superfluous error is due to the slight translational and angle error in the geometry solution produced by spatial scanning consistently for $N \geq 3$. Interantenna ranging does not suffer from this error, and for $N \geq 5$ has an RMS transmitter location error only a few centimeters worse than when using the hand-measured antenna geometry.

Figures 5.14 & 5.17 show both methods' performance in the I6 indoor test, where one array was behind an intervening wall. While the performance of spatial scanning increases with increasing neighbor distance, interantenna ranging does well at $N = 2$, but does not return to the same 0.50 m level of transmitter location performance until neighbor distances of 6 or higher are used. The character of the geometry solutions produced by interantenna ranging in this scenario are examples of how certain types of geometry errors incur less transmitter location error than others. Here, for example, in the interantenna ranging case

in Fig. 5.17 many antennas' errors are translations directed outward from the center of the geometry; this is an error which has less impact upon the resulting SART (TDOA) location solution than other types of geometry errors. Were a positioning method employed which used absolute time information in addition to TDOA information, the positioning errors in this case would be higher when using this particular geometry solution.

This type of geometry error which is more forgiving to transmitter location performance is also seen in the interantenna ranging results in Fig. 5.18, in which the positioning errors are quite comparable to those without GAC. (the RMS error does not reach $0.50\,\mathrm{m}$ due to the singular large error pointing to the upper left.) For spatial scanning results as shown in Fig. 5.15, geometry error is mainly composed of angle errors on the top, left, and right sides, which seem to be responsible for the majority of positioning errors for transmitter locations along the $x = 2$ line.
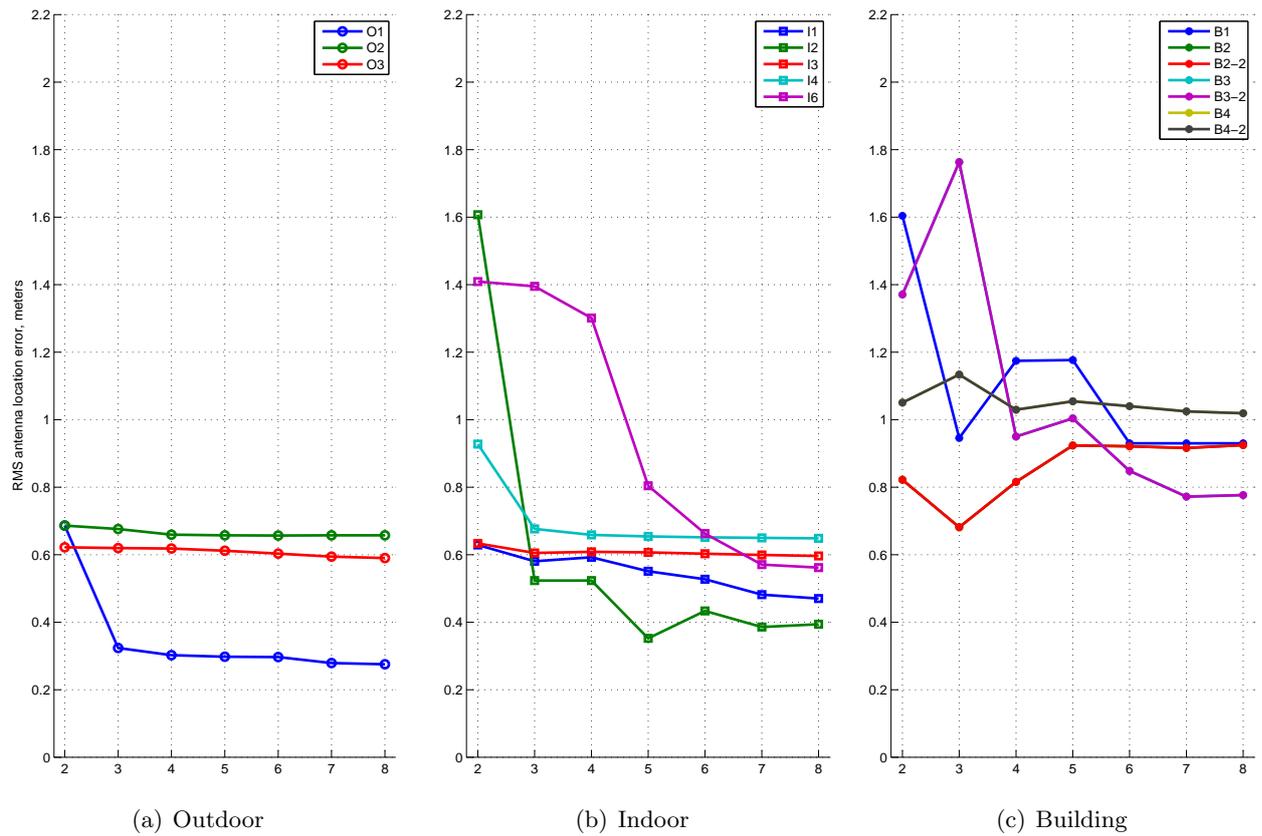
(a) Outdoor (b) Indoor (c) Building

Figure 5.9: Experimental GAC error, spatial scanning method. Curves show RMS antenna geometry error as a function of neighbor distance for (a) outdoor, (b) indoor-to-indoor, and (c) around-building testing.
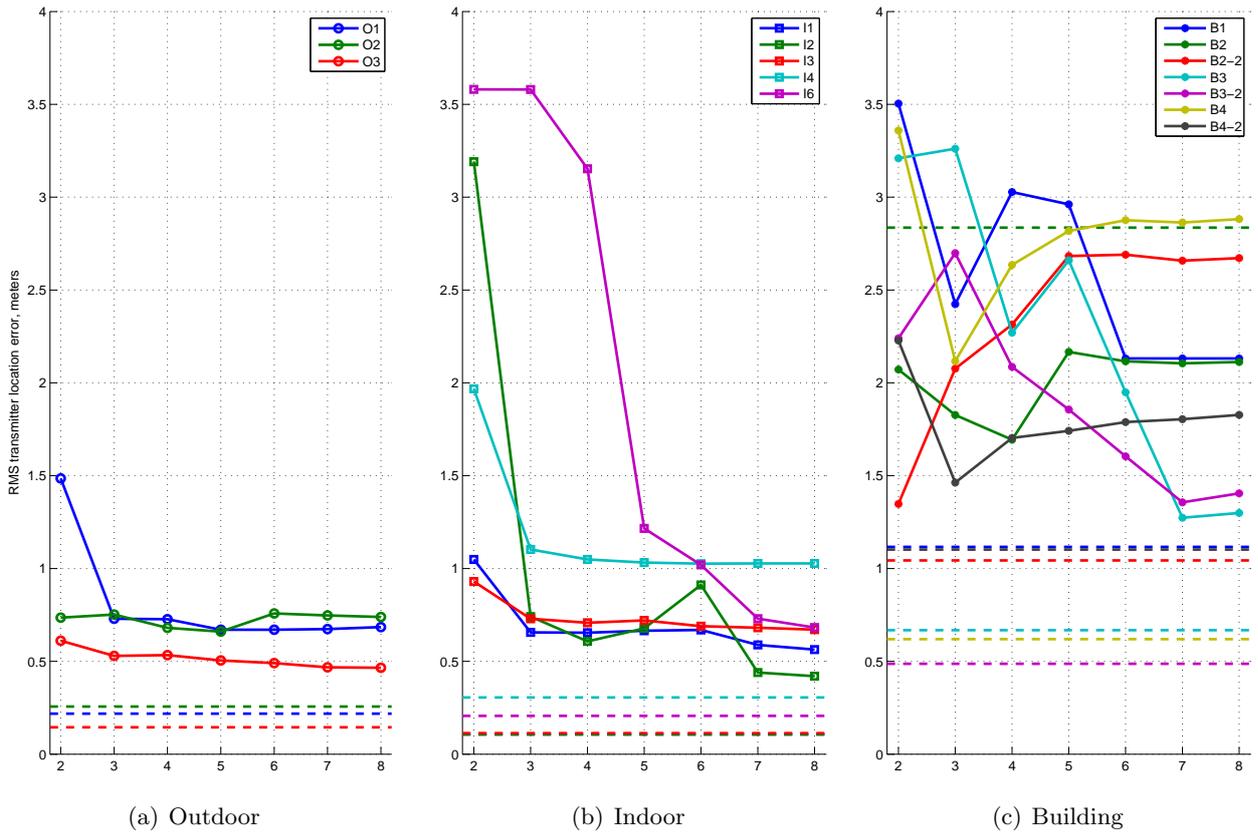
Figure 5.10: Transmitter location error using spatial scanning geometry solutions on experimental data. For (a) outdoor, (b) indoor, and (c) around-building testing, curves illustrate the RMS transmitter location error as a function of neighbor distance when SART multilateration is driven by GAC; dotted lines indicate error levels achieved when the true geometry is used.
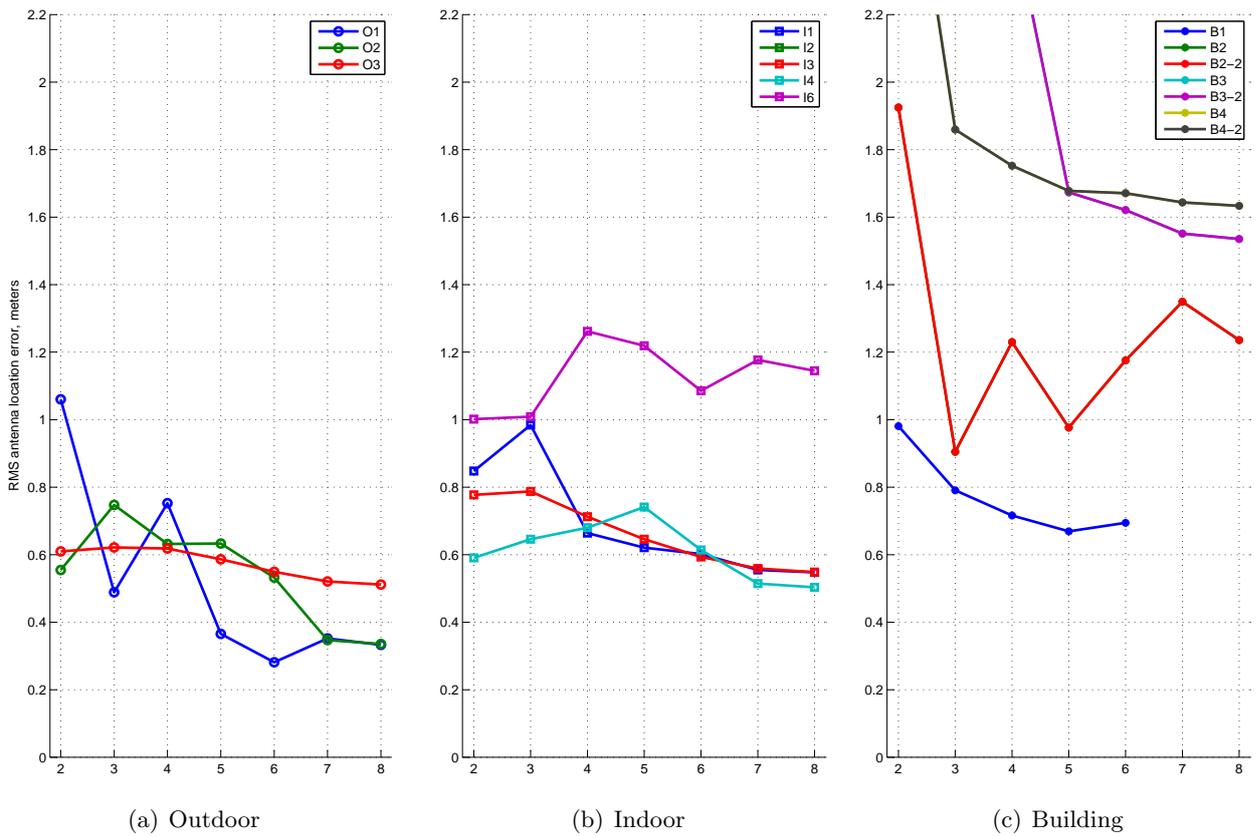
Figure 5.11: Experimental GAC error, interantenna ranging method. Curves show RMS antenna geometry error as a function of neighbor distance for (a) outdoor, (b) indoor-to-indoor, and (c) around-building testing.
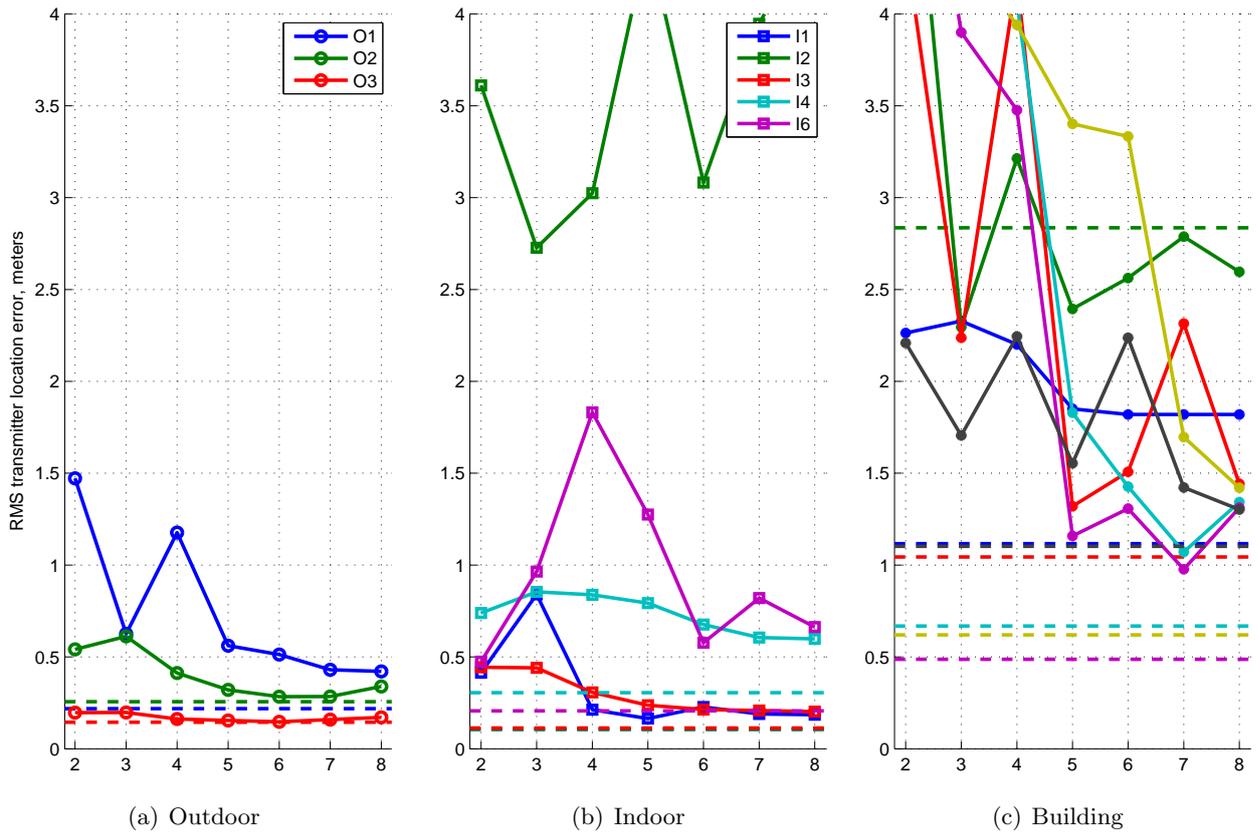
Figure 5.12: Transmitter location error using interantenna ranging-based geometry solutions on experimental data. For (a) outdoor, (b) indoor, and (c) around-building testing, curves illustrate the RMS transmitter location error, as a function of neighbor distance when SART multilateration is driven by GAC; dotted lines indicate error levels when the true geometry is used.
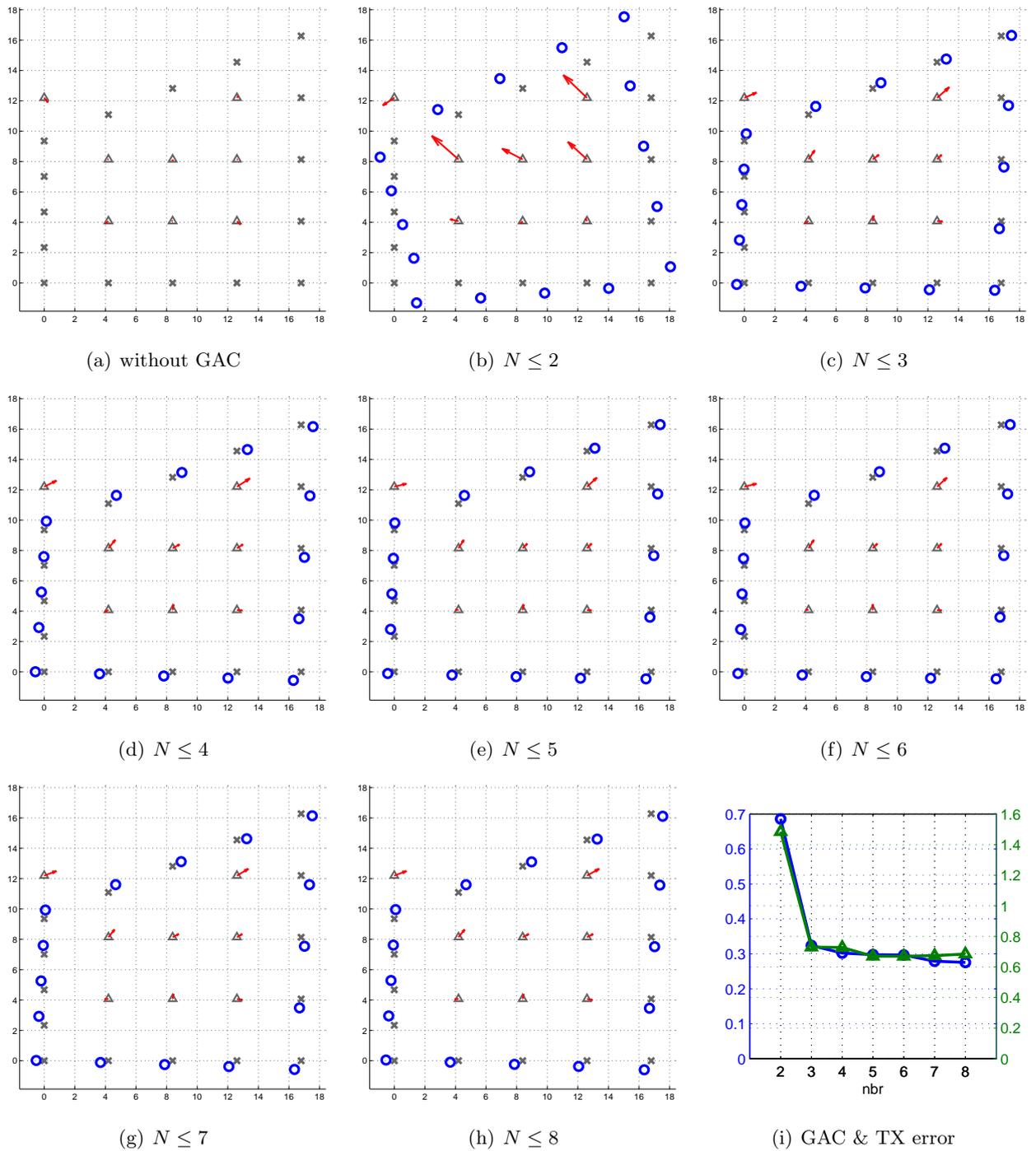
(a) without GAC

(b) $N \leq 2$

(c) $N \leq 3$

(d) $N \leq 4$

(e) $N \leq 5$

(f) $N \leq 6$

(g) $N \leq 7$

(h) $N \leq 8$

(i) GAC & TX error

Figure 5.13: Detailed spatial scanning results for 2007/06/21 "O1" test. (a) shows the true geometry and transmitter location error vectors. (b)–(h) show the geometry solutions for each value of neighbor distance. In this case, little marginal improvement in GAC and transmitter location error is observed for $N > 3$. (i) plots the errors numerically for each value of N; the left y-axis indicates GAC error levels (curve with circular symbols), and the right-side y-axis indicates TX location error for the curve using triangle markers.

Figure 5.14: Detailed spatial scanning results for 2008/01/17 "O1" test. (a) shows the true geometry and transmitter location error vectors. (b)–(h) show the geometry solutions for each value of neighbor distance. In this case, little marginal improvement in GAC and transmitter location error is observed for $N > 3$. (i) plots the errors numerically for each value of N; the left y-axis indicates GAC error levels (curve with circular symbols), and the right-side y-axis indicates TX location error for the curve using triangle markers.
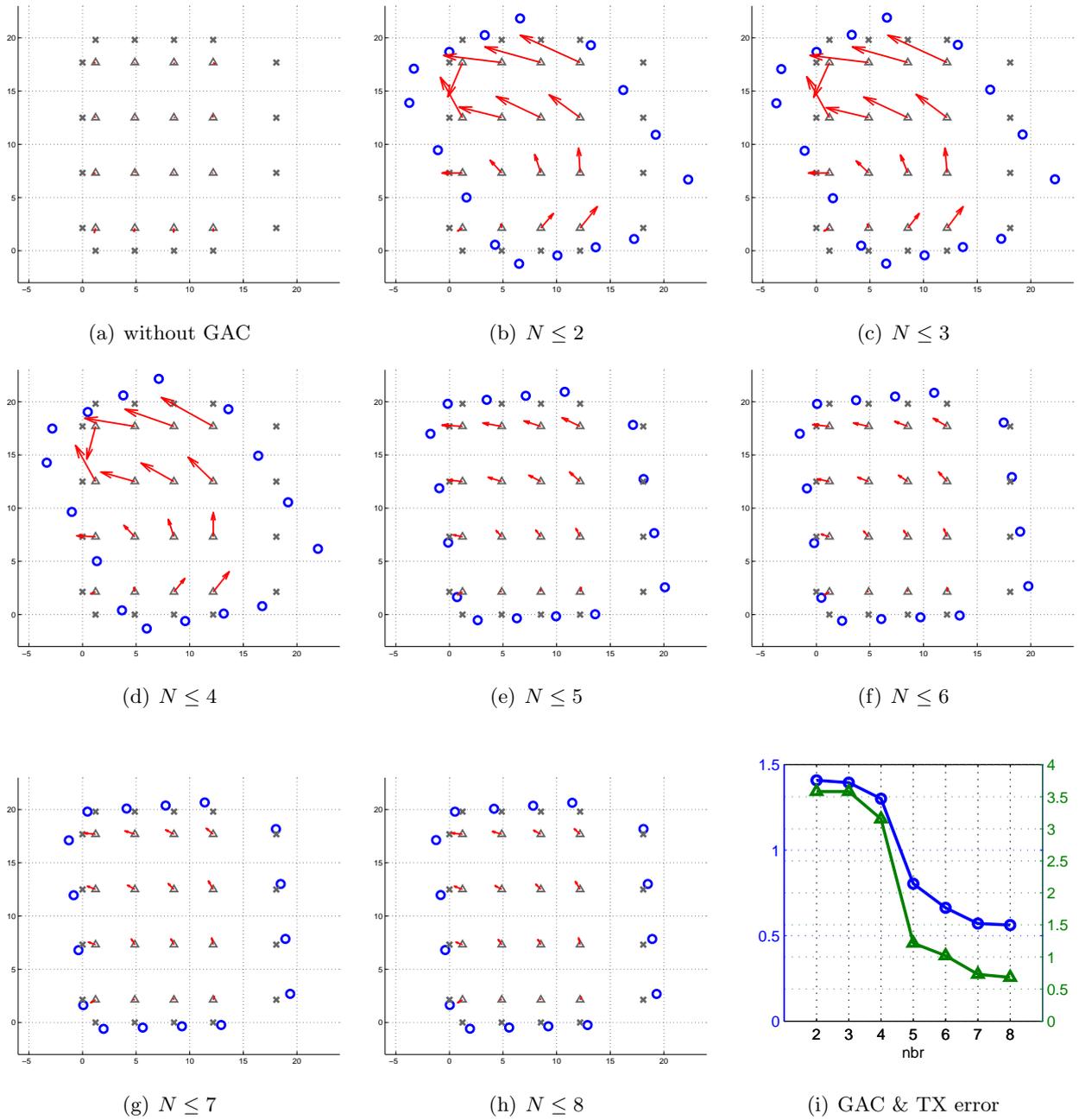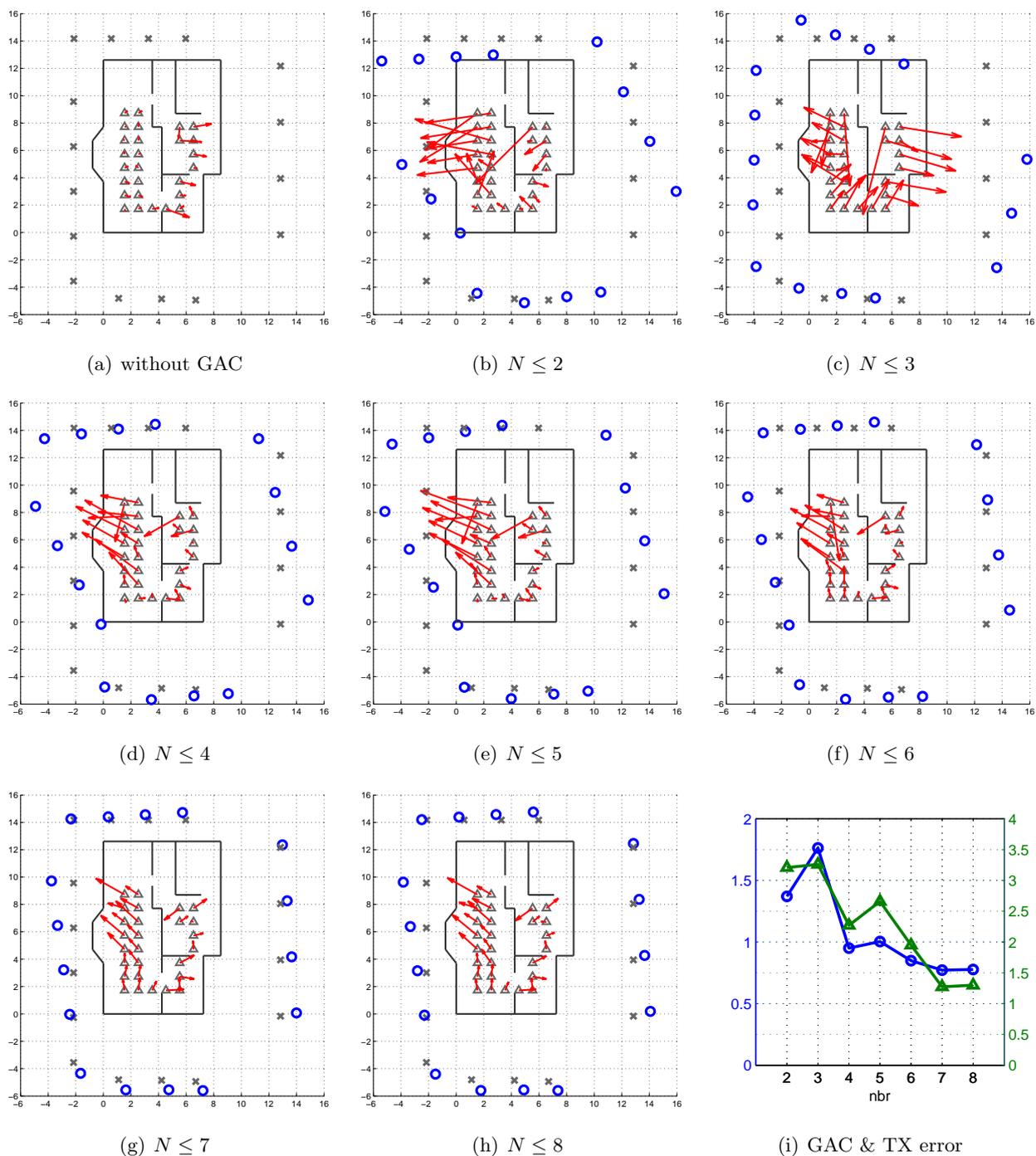
Figure 5.15: Detailed spatial scanning results for 2008/03/13 "B3" test. (a) shows the true geometry and transmitter location error vectors. (b)–(h) show the geometry solutions for each value of neighbor distance. Here it can be seen that GAC and transmitter location error experience improvement until $N = 7$; the left y-axis indicates GAC error levels (curve with circular symbols), and the right-side y-axis indicates TX location error for the curve using triangle markers.

(a) without GAC

(b) $N \leq 2$

(c) $N \leq 3$

(d) $N \leq 4$

(e) $N \leq 5$

(f) $N \leq 6$

(g) $N \leq 7$

(h) $N \leq 8$

(i) GAC & TX error

Figure 5.16: Detailed interantenna ranging results for 2007/06/21 "O1" test. (a) shows the true geometry and transmitter location error vectors. (b)–(h) show the geometry solutions for each value of neighbor distance. In this case, little marginal improvement in GAC and transmitter location error is observed for $N > 3$. (i) plots the errors numerically for each value of N; the left y-axis indicates GAC error levels (curve with circular symbols), and the right-side y-axis indicates TX location error for the curve using triangle markers.
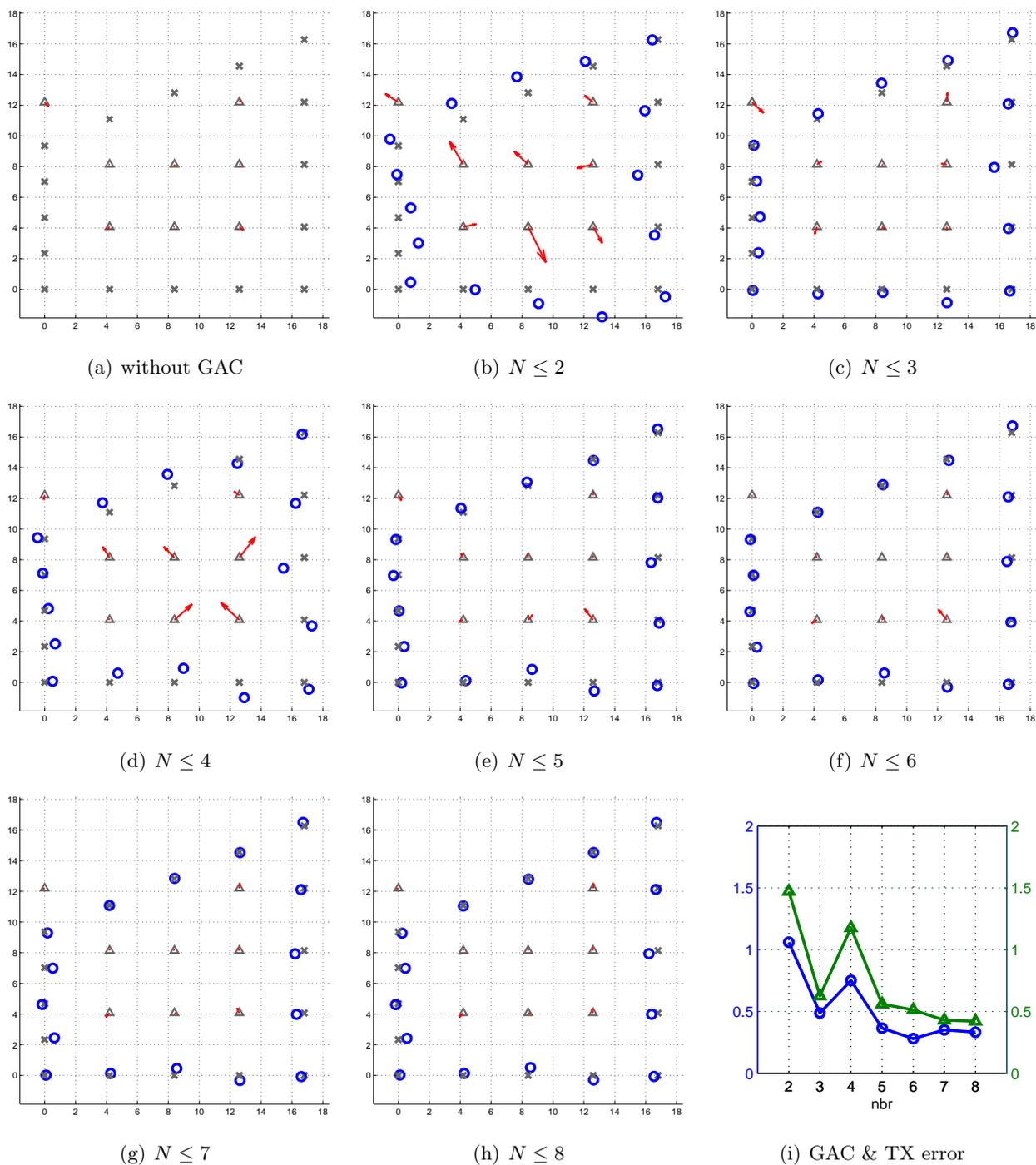
(a) without GAC (b) $N \le 2$ (c) $N \le 3$

(d) $N \le 4$ (e) $N \le 5$ (f) $N \le 6$

(g) $N \le 7$ (h) $N \le 8$ (i) GAC & TX error

Figure 5.17: Detailed interantenna ranging results for 2008/01/17 "O1" test. (a) shows the true geometry and transmitter location error vectors. (b)–(h) show the geometry solutions for each value of neighbor distance. In this case, little marginal improvement in GAC and transmitter location error is observed for $N > 3$. (i) plots the errors numerically for each value of N; the left y-axis indicates GAC error levels (curve with circular symbols), and the right-side y-axis indicates TX location error for the curve using triangle markers.
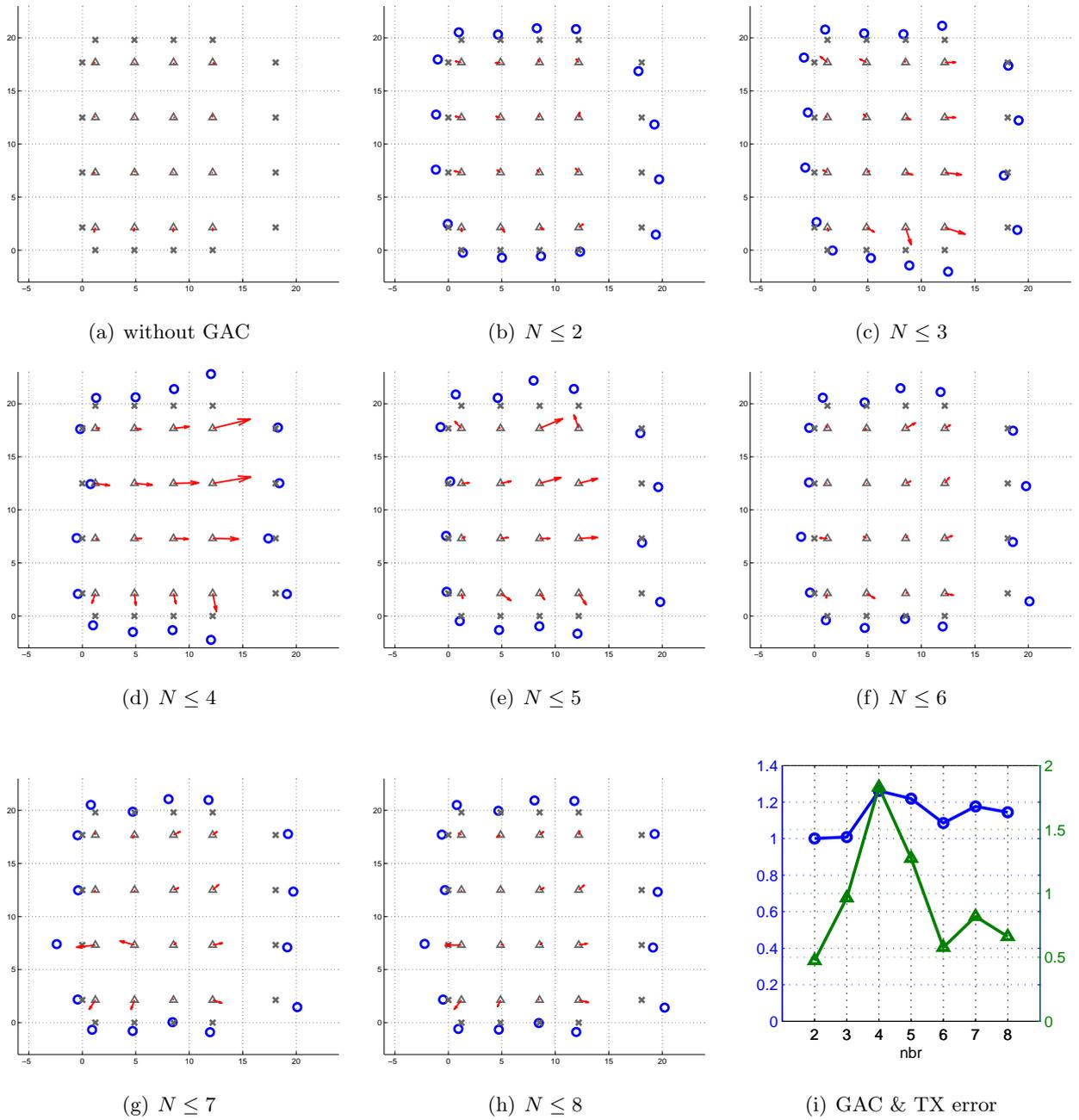
Figure 5.18: Detailed interantenna ranging results for 2008/03/13 "B3" test. (a) shows the true geometry and transmitter location error vectors. (b)–(h) show the geometry solutions for each value of neighbor distance. Here it can be seen that GAC and transmitter location error experience improvement until $N = 7$; the left y-axis indicates GAC error levels (curve with circular symbols), and the right-side y-axis indicates TX location error for the curve using triangle markers.
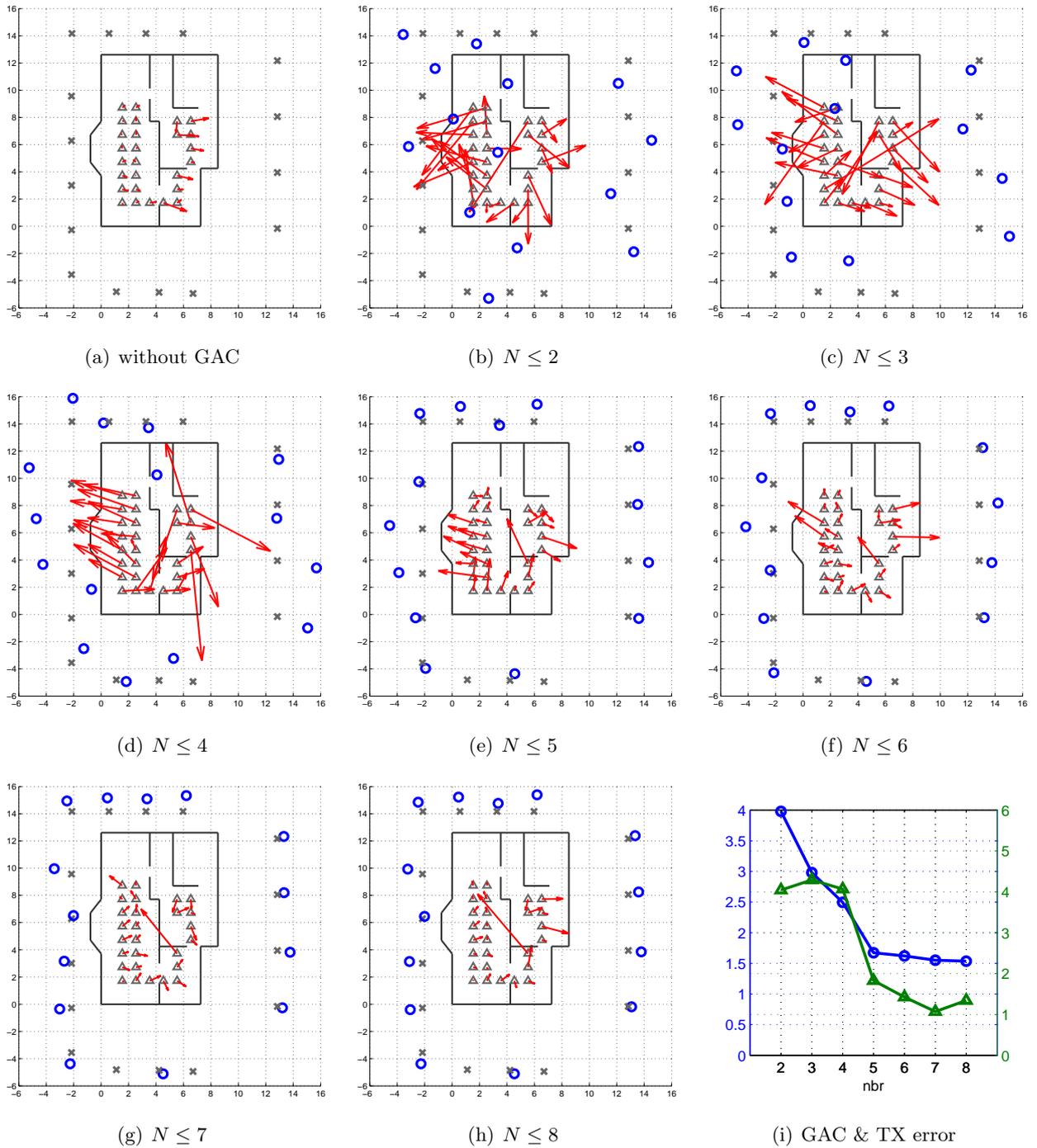
Chapter 6

Conclusions

In this thesis, we have described Geometric Autoconfiguration (GAC), a process by which we may discover the geometry of the receiving antennas used in a positioning system. This is a vital function of any "portable" location system which may be deployed to provide positioning services at a moment's notice for first responders, especially firefighters.

Our approach to GAC is tied to the hardware architecture and multicarrier signal structure of the PPL system. Since the challenge of RF-based indoor location sufficiently accurate to locate an individual, in the face of severe multipath and signal attenuation is still considered an open problem by many, our methods for GAC are extensions of techniques previously considered or currently employed in the PPL system (e.g., SART), as they have been shown robust in the face multipath signal interference.

However, while our techniques for GAC have been drawn from those used for transmitter location, there are also aspects of the GAC problem that set it apart from transmitter location which may be exploited to reduce solution error. One aspect that may be exploited is the availability of supplemental information, which may be used to reduce the size of the solution space and thus the number of suboptimal solution points to which iterative methods may converge. Such supplemental information includes array geometry constraints as demonstrated in this thesis, used to reduce the dimension of the GAC solution space. Such an improvement is predicted by the CRB, as shown in Fig. 4.4 in which the reduction in predicted error variance is observed between the unconstrained (U1) and properly constrained

(C2,C3) CRB curves. Measuring other pieces of supplemental information about the antenna geometry such as array orientation (azimuth and elevation angle) could also benefit solution error when used to reduce the GAC solution space.

A second aspect of GAC which we have explored in this thesis, which is not as pronounced in the transmitter location problem is the variation in channel propagation conditions between nearby (low neighbor distance) and far-away (high neighbor distance) antennas. In the GAC scenario, the antennas surround a building, with nearby antennas likely having clear line-of-sight and little multipath, while far-away antennas' signals must pass through at least two sets of walls when entering *and* exiting the building, experiencing propagation conditions similar to, or possibly worse than transmitter location. Since a geometry solution is possible using a sparse set of signals, we have investigated the effects upon GAC error of using a progressively larger number of signals, parameterized by neighbor distance, a rough metric of assumed signal favorability. While it was seen that higher neighbor distances always improved GAC error in theory and simulation, many of our experimental results, for the SART-based spatial scanning, showed no improvement beyond a neighbor distance of 5. In the least, it is clear (when there are many antennas) that not all interantenna signals are required for geometry to be estimated, and signals which may be detrimental to solution error may simply be omitted. We believe the failure to see improvement

We have outlined in detail the construction of two Cramer-Rao bounds against which the optimality of our methods may be gauged in two different simulation scenarios which model two common types of system error. In another novel development, we have demonstrated, repaired, and verified the repair of a significant deficiency in the Gorman constrained CRB [42] when using distance constraints to fix the geometry of a collinear antenna array, which has not yet been reported previously in the literature. We have also shown that the same vector components of Fisher information which the deficient constraints fail to observe are also missing from the unconstrained Fisher matrix, and this causes the unconstrained CRB to be properly underdetermined in cases when few measurements are assumed available and antennas are collinear.

We then conducted simulations of two methods, spatial scanning and interantenna ranging, with respect to two well-understood and common sources of error and compared the

outcomes to the Cramer-Rao bound. A significant conclusion with the first simulation with AWGN frequency-sample error is that the performance of neither of the proposed methods approaches the CRB at any point, and instead both exhibit a constant offset (on a log-log scale) in performance from the CRB. This indicates that there is much potential room for improvement in spatial scan-like methods which process frequency samples directly, at least with respect to signal-to-noise ratio. In the second simulation, which models time delay error as might be introduced by antenna pattern based signal distortion, it was observed that while interantenna ranging techniques accumulated roughly $0.50\,\mathrm{m}$ of RMS position error for every meter of RMS timing error, the spatial scanning method had approximately four times worse sensitivity to these timing errors, incurring roughly $2.0\,\mathrm{m}$ of position error per meter of RMS timing error. As one significant source of timing error is the angular dependence of antenna phase centers, clearly, antenna pattern design plays an important role in location system error, especially in GAC, in which case the nearby signals which are most likely to have favorable propagation conditions and play a majority role in a sparse solution using low neighbor distance (2,3,4) signals will all have most extreme antenna arrival and departure angles and thus suffer the highest antenna pattern distortion.

We have also presented the results from experimental tests to gauge the performance of GAC in real environments, and demonstrated that in the cases presented, transmitter location when using a GAC-driven geometry solution can result in a sub-meter RMS positioning error. GAC was shown to work in an indoor environment in which as many as half the signals processed passed through an intervening wall, as well as in a scenario in which the antennas surrounded a wood-frame building. Beyond each signal's value of neighbor distance, no other parameter was used when deciding what signals to include in processing. While generally, the level of GAC error and transmitter location error follow one another, there is clearly a nonlinear relationship between the two which is likely exacerbated by the use of TDOA processing for location estimation for the available data sets. Because of this nonlinear relation between geometry and transmitter location error, certain GAC errors whose character resembles a scaling of the entire geometry via an outward translation of all arrays can have less impact upon positioning error than array angle errors for a given overall RMS error.

While historically, antennas for the PPL system have been designed with only transmitter location in mind, it is clear that future designs must be conceived to minimize those antenna effects which also disadvantage a GAC geometry solution. Signals utilized for GAC processing must not suffer from significant additive time delay error nor superfluous attenuation due to such extreme angles. Future system versions must be designed with GAC in mind.

Future work which pursues methods for GAC could improve performance of the methods presented in this thesis by exploiting as much supplemental information regarding antenna geometry as possible. Such additional solution space reductions could be investigated in context of the constrained CRB, as well as pursued via addition of an orientation sensing capability to receiver hardware using devices such as inclinometers, accelerometers, and magnetometers. As in this work, the goal would be to also evaulate the efficacy of such information in improving performance of the entire system in real multipath conditions.

A third direction for future work would be investigation of the impact of antenna pattern effects on GAC performance and ultimately the mitigation of these effects for a given antenna by signal processing. While we have partly done so in this thesis by introducing an analysis of the impact of pure timing errors; it is unclear from experimental results which of multipath signals or antenna pattern effects are the dominant source of geometry error, however, given the high arrival and departure angles, and thus likely high distortion, of the GAC exchanged signals we believe this problem was responsible for the departure of our experimental outcomes from the expectations set by simulation. Were a detailed model of antenna pattern distortion as a function of angle and frequency available, such a model could be integrated into iterative methods to attempt to cancel antenna pattern effects at each step of iteration and possibly eliminate it as a source of error.

Finally, the gap between theoretical and specific GAC algorithm performance indicates that more complex signalling and processing may exist that could further improve geometry estimation. One could liken this to the performance gap between the CRB for data communications over a channel with a given bandwidth and signal-to-noise ratio versus the lesser performance of a particular scheme such as QPSK. The theoretical work in this thesis, we believe, lays the foundation for a methodical decomposition of error contributions and signal representation as needed to begin such an effort.

Bibliography

[1] A. Brown and J. Nordlie, "Integrated GPS/TOA navigation using a positioning and communication software defined radio," in *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, pp. 147–152, 25–27 2006.

[2] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Johns Hopkins Univ. Press, 3rd ed., 1996.

[3] J. R. Anderson, "Abandoned cold storage warehouse multi-firefighter fatality fire," technical report series, United States Fire Administration, December 1999.

[4] R. F. Fahy, P. R. LeBlanc, and J. L. Molis, "Firefighter fatalities in the United States," tech. rep., United States National Fire Protection Association, June 2006.

[5] H. C. Dampfling, "Design and implementation of the precision personnel locator digital transmitter system," Master's thesis, Worcester Polytechnic Institute, December 2006.

[6] H. K. Parikh, *An RF System Design for an Ultra Wideband Indoor Positioning System*. PhD thesis, Worcester Polytechnic Institute, December 2007.

[7] B. D. Rao and K. S. Arun, "Model based processing of signals: A state space approach," *Proceedings of the IEEE*, vol. 80, pp. 283–309, Feb. 1992. doi:10.1109/5.123298.

[8] S. Y. Kung, "A new identification and model reduction algorithm via singular value decomposition," in *Proc. Twelfth Ann. Asilomar Conference on Circuits, Systems, and Computers*, (Pacific Grove, CA), pp. 705–714, IEEE, 1978.

[9] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 984–995, Jul. 1989. `doi:10.1109/29.32276`.

[10] Y. Hua and T. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 814–824, May 1990. `doi:10.1109/29.56027`.

[11] Y. Hua and T. Sarkar, "On SVD for estimating generalized eigenvalues of singular matrix pencil in noise," *IEEE Transactions on Signal Processing*, vol. 39, pp. 892–900, Apr. 1991. `doi:10.1109/78.80911`.

[12] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, pp. 63–85. New York: John Wiley & Sons, Inc., 1968.

[13] L. Trefethen and D. Bau, *Numerical Linear Algebra*. Society for Industrial Mathematics, 1997.

[14] C. R. Wright, *Multidimensional Direction of Arrival Estimator Performance Bounds and Optimization for Non-Stationary Noise*. PhD thesis, Worcester Polytechnic Institute, 1994.

[15] D. J. Holl Jr., *State-Space Approaches to Ultra-Wideband Doppler Processing*. PhD thesis, Worcester Polytechnic Institute, April 2007.

[16] A.-J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Transactions on Signal Processing*, vol. 46, pp. 405–418, Feb. 1998.

[17] M. L. Burrows, "Two-dimensional ESPRIT with tracking for radar imaging and feature extraction," *IEEE Transactions on Antennas and Propagation*, vol. 52, pp. 524–532, Feb. 2004. `doi:10.1109/TAP.2003.822411`.

[18] M. Haardt and J. A. Nossek, "Simultaneous schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems,"

IEEE Transactions on Signal Processing, vol. 46, pp. 161–169, Jan. 1998. `doi:10.1109/78.651206`.

[19] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 608–619, Mar. 2004. `doi:10.1109/TGRS.2003.819189`.

[20] D. Cyganski, J. Orr, R. Angilly, and B. Woodacre, "Performance limits and field tests of a precision indoor positioning system using a multi-carrier approach," in *Proceceedings of the Institute of Navigation National Technical Meeting*, (San Diego, California), January 2005.

[21] Wikipedia, "Multilateration — wikipedia, the free encyclopedia," 2010. [Online; accessed 2-March-2010]. [link].

[22] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 1223–1225, Aug. 1987.

[23] J. Bard, F. Ham, and W. Jones, "An algebraic solution to the time difference of arrival equations," in *Southeastcon '96: Bringing Together Education, Science and Technology', Proceedings of the IEEE*, pp. 313–319, April 1996. `doi:10.1109/SECON.1996.510082`.

[24] J. Smith and J. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 1661–1669, Dec. 1987.

[25] Y. Huang, J. Benesty, G. Elko, and R. Mersereau, "Real-time passive source localization: a practical linear-correction least-squares approach," *IEEE Transactions on Speech and Audio Processing*, vol. 9, pp. 943–956, Nov. 2001. `doi:10.1109/89.966097`.

[26] V. T. Amendolare, "Synchronization in an indoor precision location system," Master's thesis, Worcester Polytechnic Institute, May 2007.

[27] D. Cyganski, R. J. Duckworth, S. Makarov, W. Michalson, J. Orr, V. Amendolare, J. Coyne, H. Daempfling, S. Kulkarni, H. Parikh, and B. Woodacre, "WPI Precision

Personnel Locator System: Demonstrations and RF Design Improvements," in *Proceedings of the Institute of Navigation Annual Meeting*, (Cambridge, Massachusetts), April 2007.

[28] K. Whitehouse and D. Culler, "A robustness analysis of multi-hop ranging-based localization approximations," in *IPSN '06: Proceedings of the 5th International Conference on Information Processing in Sensor Networks*, (New York, NY, USA), pp. 317–325, ACM, April 2006. `doi:10.1145/1127777.1127825`.

[29] K. Whitehouse, F. Jiang, C. Karlof, A. Woo, and D. Culler, "Sensor field localization: A deployment and empirical analysis," Tech. Rep. UCB/CSD-04-1349, EECS Department, University of California, Berkeley, April 2004.

[30] K. Whitehouse, C. Karlof, and D. Culler, "A practical evaluation of radio signal strength for ranging-based localization," *SIGMOBILE Mobile Computing and Communications Review*, vol. 11, no. 1, pp. 41–52, 2007. `doi:10.1145/1234822.1234829`.

[31] N. Patwari, A. Hero, M. Perkins, N. Correal, and R. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2137–2148, Aug. 2003. `doi:10.1109/TSP.2003.814469`.

[32] J. De Leeuw, "Multidimensional scaling," 2001. [link].

[33] W. Torgerson, "Multidimensional scaling: I. theory and method," *Psychometrika*, vol. 17, pp. 401–419, 1952.

[34] I. Borg and P. J. Groenen, *Modern Multidimensional Scaling*. Springer, 2nd ed., 2005.

[35] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*. Meboo, June 2008.

[36] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 376–380, April 1991. `doi:10.1109/34.88573`.

[37] H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*, pp. 925–928. New York: John Wiley & Sons, Inc., 2002.

[38] A. Graham, *Kronecker Products and Matrix Calculus With Applications.* Chichester, England: Ellis Horwood Ltd., 1981.

[39] S. Boyd and L. Vandenberghe, *Convex Optimization*, p. 43. New York: Cambridge University Press, 2004.

[40] P. Stoica and T. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Transactions on Signal Processing*, vol. 49, pp. 87–90, Jan. 2001. `doi:10.1109/78.890346`.

[41] J. Ash and R. Moses, "On the relative and absolute positioning errors in self-localization systems," *IEEE Transactions on Signal Processing*, vol. 56, pp. 5668–5679, Nov. 2008. `doi:10.1109/TSP.2008.927072`.

[42] J. Gorman and A. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Transactions on Information Theory*, vol. 36, pp. 1285–1301, Nov. 1990. `doi:10.1109/18.59929`.

[43] P. Stoica and B. C. Ng, "On the Cramer-Rao bound under parametric constraints," *IEEE Signal Processing Letters*, vol. 5, pp. 177–179, Jul. 1998. `doi:10.1109/97.700921`.

[44] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications.* New York: Wiley-Interscience [John Wiley & Sons], 1st ed., 1974.

[45] V. T. Amendolare, *Transactional Array Reconciliation Tomography for Precision Indoor Location.* PhD thesis, Worcester Polytechnic Institute, April 2010.