



**A GENERALIZED PHASE GRADIENT
AUTOFOCUS ALGORITHM**

DISSERTATION

Aaron Evers

AFIT-ENG-DS-19-M-025

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this document are those of the author and do not reflect the official policy or position of the United States Air Force, the United States Department of Defense or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENG-DS-19-M-025

A GENERALIZED PHASE GRADIENT AUTOFOCUS ALGORITHM

DISSERTATION

Presented to the Faculty
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

Aaron Evers, B.S.E.E., M.S.E.E.

March 2019

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENG-DS-19-M-025

A GENERALIZED PHASE GRADIENT AUTOFOCUS ALGORITHM
DISSERTATION

Aaron Evers, B.S.E.E., M.S.E.E.

Committee Membership:

Dr. Julie Jackson
Chairman

Dr. Peter Collins
Member

Dr. Matthew Fickus
Member

Lt Col Kenneth Burgi, PhD
Dean's Representative

Abstract

The phase gradient autofocus (PGA) algorithm has seen widespread use and success within the synthetic aperture radar (SAR) imaging community. However, its use and success has largely been limited to collection geometries where either the polar format algorithm (PFA) or range migration algorithm is suitable for SAR image formation. In this work, a generalized phase gradient autofocus (GPGA) algorithm is developed which is applicable with both the PFA and backprojection algorithm (BPA), thereby directly supporting a wide range of collection geometries and SAR imaging modalities. The GPGA algorithm preserves the four crucial signal processing steps comprising the PGA algorithm, while alleviating the constraint of using a single scatterer per range cut for phase error estimation which exists with the PGA algorithm. Moreover, the GPGA algorithm, whether using the PFA or BPA, yields an approximate maximum marginal likelihood estimate (MMLE) of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers. Also, in this work a new approximate MMLE, termed the max-semidefinite relaxation (Max-SDR) phase estimator, is proposed for use with the GPGA algorithm. The Max-SDR phase estimator provides a phase error estimate with a worst-case approximation bound compared to the solution set of MMLEs (i.e., solution set to the non-deterministic polynomial-time hard (NP-hard) GPGA phase estimation problem). Moreover, in this work a specialized interior-point method is presented for more efficiently performing Max-SDR phase estimation by exploiting low-rank structure typically associated with the GPGA phase estimation problem. Lastly, simulation and experimental results produced by applying the GPGA algorithm with the PFA and BPA are presented.

Acknowledgements

First and foremost, I would like to thank my wife for her love and support. Without it, none of this would have been possible.

I would like to thank my committee members, Dr. Julie Jackson, Dr. Peter Collins, and Dr. Matthew Fickus, for their valuable insights and recommendations. I would also like to extend a special thanks to Dr. Julie Jackson for the experience and opportunities she has afforded me, dating back to my undergraduate studies. I would like to thank both Matrix Research and Radial Research and Development for both the opportunity and financial support needed to finish my degree. Lastly, I would like to thank individuals who have provided me the continued opportunity to finish my degree: Dr. Lee Patton, Dr. Matthew Ferrara, and Dr. Michael Bryant.

Aaron Evers

Table of Contents

	Page
Abstract	iv
Acknowledgements	v
List of Figures	x
List of Tables	xiii
List of Acronyms	xv
I. Introduction	1
1.1 Scope	2
1.2 Contributions	5
1.2.1 Spatially Invariant Defocus Model	6
1.2.2 Generalized Phase Gradient Autofocus Algorithm	6
1.2.3 Max-Semidefinite Relaxation Phase Estimation	7
1.3 Outline	7
1.4 Notation	8
II. Background	10
2.1 Phase-Only Autofocus Signal Model	10
2.2 Phase Estimation for Phase-Only Autofocus	19
2.2.1 Generalized Phase Gradient Autofocus	21
2.2.2 Multichannel Autofocus	22
2.2.3 Maximum Sharpness Autofocus	24

	Page
2.2.4 Minimum Entropy Autofocus	25
2.2.5 Summary.....	26
2.3 Conclusion	27
III. Matched Filter and Image Formation Analysis	30
3.1 Matched Filter Analysis	30
3.1.1 Example Results	33
3.2 Image Formation Analysis	35
3.2.1 Differential Range Error	36
3.2.2 Approximate Differential Range Error Bound	38
3.2.3 Example Results	42
3.3 Conclusion	51
IV. Generalized Phase Gradient Autofocus Algorithm.....	52
4.1 Using the Polar Format Algorithm	52
4.2 Using the Backprojection Algorithm	55
4.3 Scatterer Selection	57
4.4 Low-Pass Filter Design	57
4.5 Iterations	59
4.6 Summary	60
4.7 Maximum Marginal Likelihood Estimate	61
4.8 Conclusion	66

	Page
V. Generalized Phase Gradient Autofocus Phase Estimation	67
5.1 Phase Differencing	67
5.2 Eigenvector Relaxation	68
5.3 Max-Semidefinite Relaxation	70
5.4 Efficient Max-Semidefinite Relaxation	73
5.4.1 Low-Rank Data	74
5.4.2 Specialized Interior Point Method	76
5.4.3 Newton's Method	78
5.4.4 Max-Semidefinite Relaxation Feasible Point	81
5.4.5 Computational Complexity	82
5.5 Conclusion	83
VI. Results	85
6.1 Phase Error Estimate Mean Squared Error	86
6.2 Average Run-Time	88
6.3 Near-Field Imaging	90
6.4 Far-Field Imaging	92
6.4.1 Using the Polar Format Algorithm	93
6.4.2 Using the Backprojection Algorithm	94
6.5 Conclusion	97
VII. Conclusion	100
7.1 Spatially Invariant Defocus Model	100
7.2 Generalized Phase Gradient Autofocus Algorithm	101

	Page
7.3 Max-Semidefinite Relaxation Phase Estimation	101
Appendix A. Marginal Likelihood	103
Appendix B. Fisher Information Matrix	106
2.1 Argument of the Phase Error Vector	107
2.2 Cross Terms	109
2.3 Hyperparameters	109
2.4 Conclusion	110
Appendix C. Lagrange Dual Problem	112
Appendix D. Logarithmic Barrier Gradient and Hessian	114
Bibliography	116

List of Figures

Figure		Page
1	The scope of this research.	3
2	Spotlight-mode bistatic SAR scenario.	10
3	High-Level SAR imaging processing chain.	13
4	Example ideal and defocused SAR images.	15
5	PFA and BPA SAR image formation processes.	17
6	Bistatic scenario for a single slow-time pulse of a SAR collect.	33
7	The surface and contours of a scatterer's ambiguity function, and the desired and observed samples in (15), given the position and velocity uncertainties in (45) to (48).....	35
8	Bistatic geometry.	36
9	Near-field imaging scenario.	43
10	Convergence of (56) and (57).	43
11	Spatially invariant regions identified using (63) and (64) with $\beta = 1$, $k_{\max} = 4$, and $\ \delta\mathbf{p}_t\ , \ \delta\mathbf{p}_r\ = \frac{3^{3/2}}{2}\sigma$	44
12	Minimum, maximum, and mean differential phase error for each point located within the identified spatially invariant region in Figure 11.	45
13	Far-field imaging scenario.	46

Figure		Page
14	Convergence of (56) and (57).	46
15	Spatially invariant regions identified using (63) and (64) with $\beta = 1$, $k_{\max} = 4$, and $\ \delta \mathbf{p}_t\ , \ \delta \mathbf{p}_r\ = \frac{3^{3/2}}{2}\sigma$	47
16	Minimum, maximum, and mean differential phase error for each point located within the identified spatially invariant region in Figure 15.	48
17	Near-field imaging scenario.	49
18	Comparison of phase-only autofocus SAR image quality as the spatially invariant phase error model breaks down.	50
19	PGA algorithm [19].	53
20	GPGA algorithm.	54
21	Number of scatterers, P , needed during the scatterer selection process for an efficient estimator to produce a phase error estimate with an MSE of $(\alpha\pi/4)^2$, with $\alpha = 1/4$	76
22	MSE of the argument of the phase error estimate for the PD, EVR, and Max-SDR phase estimators versus SINR and the size of the GPGA data (i.e., number of slow-time pulses N and number of selected scatterers P).	87
23	Average run-time for different GPGA phase estimators versus the size of the GPGA data (i.e., number of slow-time pulses N and number of selected scatterers P).	89
24	Experimental system and setup from [60, 94–97].	91

Figure		Page
25	Measured and autofocused BPA images produced using the GPGA algorithm with different phase estimators, for the experimental setup in Figure 24.	92
26	Autofocus results produced using the GPGA algorithm with different phase estimators and the PFA for SAR image formation.	98
27	Autofocus results produced using the GPGA algorithm with different phase estimators and the BPA for SAR image formation.	99

List of Tables

Table		Page
1	Summary of trade-offs between different autofocus algorithms.	29
2	Bistatic scenario.	33
3	Computational complexity comparison between the specialized IPM in Algorithm 1 and a generic IPM for solving (99).	83
4	Summary of trade-offs between different GPGA phase estimators.	84
5	MSE of PD phase error estimates in radians ² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.	94
6	MSE of EVR phase error estimates in radians ² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.	94
7	MSE of Max-SDR phase error estimates in radians ² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.	95
8	MSE of PD phase error estimates in radians ² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.	96
9	MSE of EVR phase error estimates in radians ² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.	96

Table		Page
10	MSE of Max-SDR phase error estimates in radians ² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.	97

List of Acronyms

AFIT	Air Force Institute of Technology.
AWG	arbitrary waveform generator.
AWGN	additive white Gaussian noise.
BPA	backprojection algorithm.
CFAR	constant false alarm rate.
CPI	coherent processing interval.
CRLB	Cramér-Rao lower bound.
DFT	discrete Fourier transform.
DSDR	dual semidefinite relaxation.
EVR	eigenvector relaxation.
FFT	fast Fourier transform.
FIM	Fisher information matrix.
GPGA	generalized phase gradient autofocus.
GSF	generalized sharpness function.
GUI	graphical user interface.
INS	inertial navigation system.
IPM	interior-point method.

LMI	linear matrix inequality.
LPF	low-pass filter.
Max-SDR	max-semidefinite relaxation.
MCA	multichannel autofocus.
MMLE	maximum marginal likelihood estimate.
MSE	mean squared error.
NP-hard	non-deterministic polynomial-time hard.
PD	phase difference.
PFA	polar format algorithm.
PGA	phase gradient autofocus.
RAIL	Radar Instrumentation Laboratory.
SAR	synthetic aperture radar.
SDR	semidefinite relaxation.
SINR	signal-to-interference-plus-noise ratio.
UAV	unmanned aerial vehicle.

I. Introduction

Synthetic aperture radar (SAR) imaging is a special form of radar where a scene or area of interest is interrogated with radio waves over a set of diverse viewing angles, and the scattered waves are observed and used to estimate the scene's electromagnetic reflectivity. Being a special form of radar, SAR imaging shares operational benefits associated with general radar systems including: 1) day/night operation, 2) all-weather operation, and 3) short/long range operation [1–9]. The earliest development of SAR imaging is commonly attributed to Carl Wiley of Goodyear Aircraft Corporation, in June 1951 [10, 11]. Since its early development, SAR imaging has been employed for a wide variety of military and civilian applications including: topographic mapping, land-use monitoring, agricultural monitoring, environmental monitoring, planetary or celestial investigations, and reconnaissance and acquisition of targeting information [1–3, 6–9, 12, 13].

Although there are a number of applications and operational advantages for SAR imaging systems, SAR imaging is not without its own challenges. In order to produce high-quality, focused SAR images accurate propagation models, high-grade transmit and receive hardware, and accurate motion measurements of the collection geometry are needed [2, 3, 14]. The level of accuracy needed is on the order of a fraction of the transmitted signal's carrier wavelength [2, 3]. Meeting this level of accuracy tends toward costly studies and/or hardware. Typically, a more cost-effective algorithmic approach is considered consisting of a data-driven automatic refocusing technique,

termed “autofocus” [2, 3, 14].

Numerous autofocus algorithms exist [1–3, 13, 15–63]. Of these many existing autofocus algorithms one is arguably considered the “gold-standard” — the phase gradient autofocus (PGA) algorithm¹ [14, 33]. The PGA algorithm was first introduced in the late 1980s and was formulated for use with the polar format algorithm (PFA) [16]. Since its early development, the PGA algorithm has seen widespread use and success within the SAR imaging community [2, 3, 14, 16, 19–24, 29, 31, 33, 35–37, 42, 44–46, 48, 51, 53, 55, 58, 62]. This success is largely driven by the balance the PGA algorithm provides in terms of being relatively efficient, yet proving to perform well by providing quality autofocus solutions for a number of scene types, scene content, and scenarios. As a result of such widespread use and success, a natural question has been raised in the literature: can the PGA algorithm be applied directly with the backprojection algorithm (BPA)², which directly supports a wide range of collection geometries and SAR imaging modalities [44]? It is this question, along with other closely related questions, which are answered in this work. Specifically, *in this research a generalized phase gradient autofocus (GPGA) algorithm is developed which preserves the four crucial signal processing steps comprising the PGA algorithm, while being applicable with both the PFA and BPA.*

1.1 Scope

Figure 1 depicts the scope of this research. As indicated in Figure 1, focus is given solely to spotlight-mode SAR imaging. Focus is limited to spotlight-mode for developing and studying the GPGA algorithm to enable side-by-side comparisons to be

¹Herein, PGA algorithm refers to the class of PGA algorithms, as variations of PGA exist using different phase estimators [3, 16, 17, 19, 20].

²Herein, BPA refers to the backprojection and filtered/convolution backprojection algorithms.

made with the PGA algorithm, which was initially developed for spotlight-mode SAR imaging. Although, as mentioned in [44], the BPA is invariant to the SAR operating mode (i.e., is performed the same regardless of the SAR imaging modality). Thus, the GPGA algorithm is expected to be applicable with other SAR imaging modalities, given it is shown to be applicable with the BPA.

As noted previously, SAR image defocus is generally attributed to inaccurate propagation models, transmit and receive hardware limitations, and/or inaccurate motion measurements of the collection geometry. Regardless of the source of defocus, SAR data is defocused in two ways: 1) range profile misalignment (i.e., range-cell migration) and 2) pulse-to-pulse phase corruption (i.e., phase-only defocus). These two types of SAR data defocus are listed in Figure 1 and constitute two different classes of SAR autofocus algorithms based on the assumed defocus model.

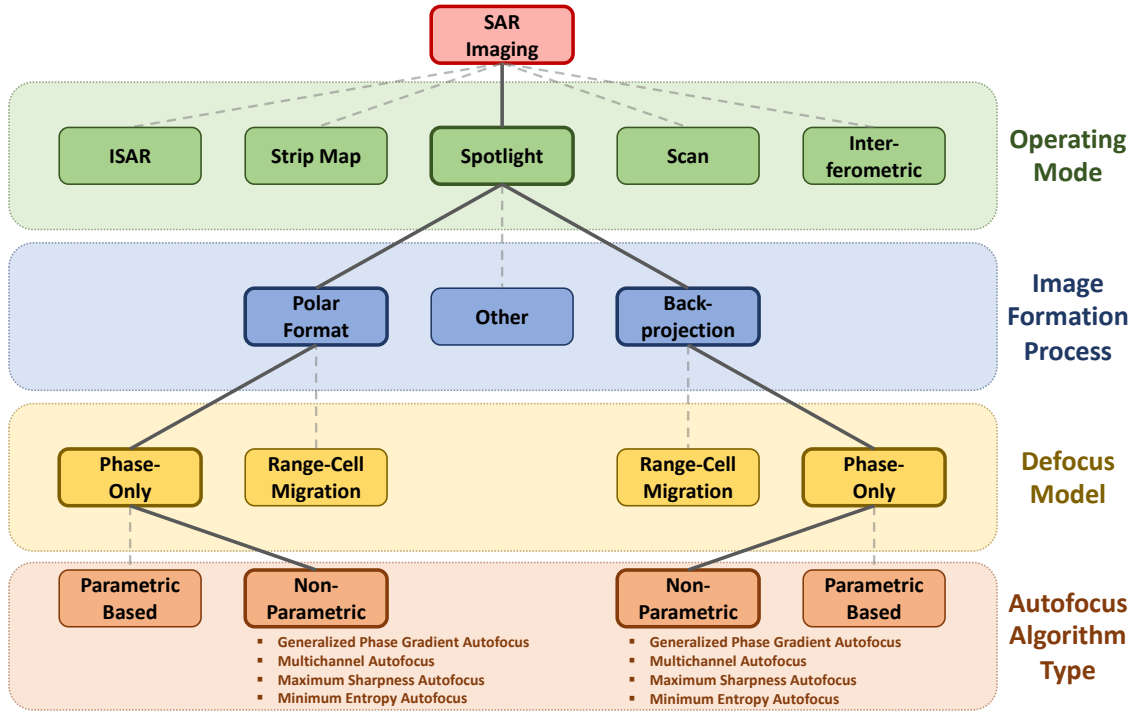


Figure 1. The scope of this research.

The scope of this research is limited to the phase-only defocus model, which is applicable when motion measurement errors are less than the range and velocity resolution of the SAR system [3, 64, 65]. Thanks in part to advanced propagation modeling, improved transmit and receive hardware, and inertial navigation systems (INSs) with improved accuracy, the phase-only defocus model is widely applicable. In fact, phase-only autofocus algorithms are so prevalent they are also referred to as traditional autofocus algorithms [65]. Although the phase-only defocus model is applicable in most cases, the range-cell migration defocus model is necessary for: 1) higher-resolution SAR imaging; 2) in cases where the SAR system is constrained by strict size-weight-and-power requirements, which indirectly lead to inaccurate INS measurements; and 3) the SAR system is deployed on a less stable platform, such as an unmanned aerial vehicle (UAV), where large motion measurement errors are common [64, 65]. Since the scope of this research is limited to the phase-only defocus model, for the remainder of this work the term “autofocus” is to be interpreted as meaning “phase-only autofocus.”

As shown in Figure 1, numerous non-parametric based autofocus algorithms exist for spotlight-mode SAR imaging using either the PFA or BPA [1–3, 13, 15–63]³. At first glance it may seem odd so many autofocus algorithms exist for a single defocus model. However, at the core of each of the listed autofocus algorithms is a non-deterministic polynomial-time hard (NP-hard) phase estimation problem, whose solution set consists of estimates of the phase errors causing SAR image defocus. For SAR imaging applications, where the size of the data involved with the NP-hard phase estimation problem is typically large, it is not reasonable to expect to always solve the phase estimation problem (i.e., find a global optimal point). Instead, heuristics are needed

³Figure 1 does not include popular autofocus algorithms such as map-drift and prominent point processing, as these algorithms are either parametric based or not automated [2].

to obtain an approximate solution to the NP-hard phase estimation problem in a reasonably efficient manner. As a result of having to resort to heuristics, an excess of autofocus algorithms is not so surprising. For instance, many of the existing autofocus algorithms provide a trade-off between performance/robustness, computational complexity, and suboptimality of the phase error estimate (i.e., suboptimality of an approximate solution to the NP-hard phase estimation problem).

As mentioned previously, of the many existing autofocus algorithms available, the PGA algorithm is arguably considered the “gold-standard” due in large part to the balance it provides between performance and computational complexity [14, 33]. Although considered the “gold-standard,” a number of areas exist where the PGA algorithm may be improved or generalized. In the next section, contributions made with this research, consisting of advancements made to the PGA algorithm through the development of the GPGA algorithm, are discussed.

1.2 Contributions

In this research, advancements to the current state-of-the-art in SAR autofocus are made by extending or generalizing the well-known PGA algorithm, resulting in what is termed the GPGA algorithm. The material from this research consists of three primary contributions and a number of secondary contributions, which have culminated in four publications [60, 66–68]. Together, the primary and secondary contributions of this research provide greater insight into topics concerning: 1) how the GPGA algorithm relates to other existing SAR autofocus algorithms, 2) when and where the GPGA algorithm is applicable, 3) strategies for configuring/applying the GPGA algorithm, 4) statistical motivation for the GPGA algorithm, and 5) trade-space study for three different GPGA phase estimators. In the sections to follow each of the three

primary contributions are summarized, with each section primarily concerned with material presented in each of the respective publications [66–68].

1.2.1 Spatially Invariant Defocus Model

Many existing autofocus algorithms, including the GPGA algorithm, rely on the assumption that phase errors, which cause SAR image defocus, are spatially invariant [1–3, 13, 15–17, 19, 20, 23, 26, 28, 30–33, 35, 38, 39, 42, 45, 46, 48, 50, 55, 59, 60]. In many cases this assumption holds, but for larger scene sizes or scenarios where the far-field approximation does not hold, the spatially invariant phase error approximation breaks down. In Section 3.2, an approximate upper bound is derived on the spatially variant phase error component of phase errors. The derived bound applies with near- and far-field bistatic geometries, with monostatic as a special case, and is expressed in terms of either known quantities or quantities which typically have a known upper bound. The derived bound enables SAR scene size limits to be computed, such that phase errors are approximately spatially invariant according to a specified phase threshold.

1.2.2 Generalized Phase Gradient Autofocus Algorithm

In Chapter IV, the GPGA algorithm is developed which is applicable with both the PFA and BPA, thereby directly supporting a wide range of collection geometries and SAR imaging modalities. The GPGA algorithm preserves the four crucial signal processing steps comprising the PGA algorithm, while alleviating the constraint of using a single scatterer per range cut for phase error estimation which exists with the PGA algorithm. Moreover, it is shown in Chapter IV that the GPGA algorithm, whether using the PFA or BPA, yields an approximate maximum marginal likelihood

estimate (MMLE) of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers.

1.2.3 Max-Semidefinite Relaxation Phase Estimation

In Section 5.3, a new approximate MMLE, termed the max-semidefinite relaxation (Max-SDR) phase estimator, is proposed for use with the GPGA algorithm. Leveraging recent work on SDR, the Max-SDR phase estimator provides a phase error estimate with a worst-case approximation bound compared to the solution set of MMLEs (i.e., worst-case suboptimality for a feasible point to the NP-hard GPGA phase estimation problem). Additionally, in Section 5.3 a specialized interior-point method (IPM) is presented for more efficiently performing Max-SDR phase estimation by exploiting low-rank structure typically associated with the GPGA phase estimation problem. When the size of the data involved with the phase estimation problem is large, which is typical for SAR imaging applications, the computational complexity and run-time are reduced by roughly N^2 and two orders of magnitude, respectively, as compared to a generic IPM. Hence, the presented specialized IPM makes Max-SDR phase estimation more practical for SAR imaging applications leveraging the GPGA algorithm.

1.3 Outline

The remainder of this work is outlined as follows. In Chapter II the phase-only autofocus signal model is formulated beginning with the baseband, received signal model for a bistatic collection geometry. Additionally, in Chapter II a high-level description of the objective of phase-only autofocus is provided, and a brief summary of many

existing phase-only autofocus algorithms is given. In Chapter III matched filter and image formation analysis is performed to identify bounds on motion measurement errors and SAR scene size limits such that the phase-only defocus model and spatially invariant phase error assumption are applicable. In Chapter IV the GPGA algorithm is formulated by first providing a side-by-side description of the PGA and GPGA algorithms, making evident the GPGA algorithm includes the PGA algorithm as a special case. Then, it is shown that whether using the PFA or BPA, the GPGA algorithm provides an approximate MMLE of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers. In Chapter V three phase estimators or approximate MMLEs are described for use with the GPGA algorithm, and their respective trade-offs discussed. In Chapter VI results produced from synthetic and measured data sets using the GPGA algorithm are presented and discussed. In Chapter VII concluding remarks are made.

1.4 Notation

The real and complex fields are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. The imaginary unit is denoted by $j = \sqrt{-1}$, the complex exponential function by $e : \mathbb{C} \rightarrow \mathbb{C}$, the argument/angle of a complex value by $\text{Arg} : \mathbb{C} \rightarrow \mathbb{R}$, the real operator by $\text{Re} : \mathbb{C} \rightarrow \mathbb{R}$, and the complex conjugate by $(\cdot)^*$. The set of $N \times N$ real-valued symmetric and complex-valued Hermitian matrices are denoted by $\mathbb{S}^N$ and $\mathbb{H}^N$, respectively. Column vectors are denoted by boldface lowercase letters and matrices by boldface uppercase letters. The n th entry of a vector $\mathbf{x}$ is denoted by $[\mathbf{x}]_n$ using 0-based indexing. The superscripts $(\cdot)^T$ and $(\cdot)^H$ represent the transpose and conjugate transpose, respectively. A square diagonal matrix with main diagonal elements $\mathbf{x}$ is denoted by $\text{Diag}(\mathbf{x})$, and $\text{diag}(\mathbf{X})$ is a column vector with elements given by the main diagonal elements of

$\mathbf{X}$. The Hadamard or element-wise product is denoted by $\odot$. The identity matrix is denoted by $\mathbf{I}$, and vectors or matrices of all zeros or ones are denoted by $\mathbf{0}$ and $\mathbf{1}$, respectively. The Euclidean norm is denoted by $\|\cdot\|$, the Frobenius norm by $\|\cdot\|_F$, the ceiling function by $\lceil \cdot \rceil$, and the absolute value of a scalar and determinant of a matrix by $|\cdot|$. The trace of a square matrix is denoted by $\text{Tr}(\cdot)$ and the rank of a matrix is denoted by $\text{Rank}(\cdot)$. The generalized inequalities $\mathbf{X} \succ \mathbf{0}$ and $\mathbf{X} \succeq \mathbf{0}$ denote the square matrix $\mathbf{X}$ is positive definite and positive semidefinite, respectively. The expectation of a random variable is denoted by $\text{E}[\cdot]$. Lastly, $\mathcal{U}$ and $\mathcal{CN}$ denote uniform and circularly symmetric complex normal distributions, respectively.

II. Background

In this section, the phase-only autofocus signal model is formulated beginning with the baseband, received signal model for a bistatic collection geometry, which includes monostatic as a special case. In addition, in this section a high-level description of the objective of phase-only autofocus is provided, and a brief summary of many existing phase-only autofocus algorithms is given.

2.1 Phase-Only Autofocus Signal Model

Consider the spotlight-mode bistatic SAR imaging scenario depicted in Figure 2 consisting of a transmitter-receiver pair and stationary point scatterers. In Figure 2, $\mathbf{p}_i, \mathbf{p}_t, \mathbf{p}_r : \mathbb{R} \rightarrow \mathbb{R}^3$ are vector-valued functions of time providing the true position state of the i th scatterer, transmitter, and receiver, respectively, and $t_n \in \mathbb{R}$ indicates the time of the n th slow-time pulse. Additionally, let $\tilde{\mathbf{p}}_t, \tilde{\mathbf{p}}_r : \mathbb{R} \rightarrow \mathbb{R}^3$ denote

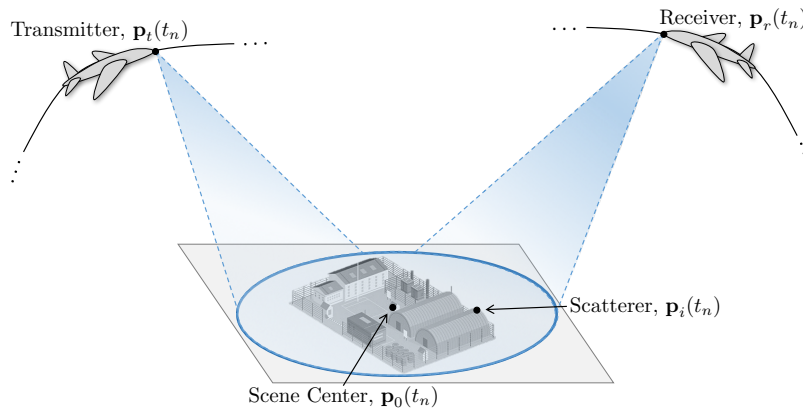


Figure 2. Spotlight-mode bistatic SAR scenario.

vector-valued functions of time providing the measured/estimated position state of

the transmitter and receiver

$$\tilde{\mathbf{p}}_t(t) = \mathbf{p}_t(t) + \boldsymbol{\delta}\mathbf{p}_t(t), \quad (1)$$

$$\tilde{\mathbf{p}}_r(t) = \mathbf{p}_r(t) + \boldsymbol{\delta}\mathbf{p}_r(t), \quad (2)$$

where $\boldsymbol{\delta}\mathbf{p}_t, \boldsymbol{\delta}\mathbf{p}_r : \mathbb{R} \rightarrow \mathbb{R}^3$ are the transmitter and receiver position errors, respectively. Then, the true and measured vectors from the transmitter-to- i th scatterer and i th scatterer-to-receiver, $\mathbf{r}_{t,i}, \mathbf{r}_{i,r}, \tilde{\mathbf{r}}_{t,i}, \tilde{\mathbf{r}}_{i,r} : \mathbb{R} \rightarrow \mathbb{R}^3$, are

$$\mathbf{r}_{t,i}(t) = \mathbf{p}_i(t) - \mathbf{p}_t(t) = r_{t,i}(t) \hat{\mathbf{r}}_{t,i}(t), \quad (3)$$

$$\mathbf{r}_{i,r}(t) = \mathbf{p}_r(t) - \mathbf{p}_i(t) = r_{i,r}(t) \hat{\mathbf{r}}_{i,r}(t), \quad (4)$$

$$\tilde{\mathbf{r}}_{t,i}(t) = \mathbf{p}_i(t) - \tilde{\mathbf{p}}_t(t) = \tilde{r}_{t,i}(t) \hat{\tilde{\mathbf{r}}}_{t,i}(t), \quad (5)$$

$$\tilde{\mathbf{r}}_{i,r}(t) = \tilde{\mathbf{p}}_r(t) - \mathbf{p}_i(t) = \tilde{r}_{i,r}(t) \hat{\tilde{\mathbf{r}}}_{i,r}(t), \quad (6)$$

where $r_{t,i}, r_{i,r}, \tilde{r}_{t,i}, \tilde{r}_{i,r} : \mathbb{R} \rightarrow \mathbb{R}$ are the respective vector magnitudes and $\hat{\mathbf{r}}_{t,i}, \hat{\mathbf{r}}_{i,r}, \hat{\tilde{\mathbf{r}}}_{t,i}, \hat{\tilde{\mathbf{r}}}_{i,r} : \mathbb{R} \rightarrow \mathbb{R}^3$ are the respective unit vector-valued functions. Using (3) to (6), the true and measured bistatic delay and Doppler from the transmitter-to- i th scatterer-to-receiver, $\tau_i, \nu_i, \tilde{\tau}_i, \tilde{\nu}_i : \mathbb{R} \rightarrow \mathbb{R}$, at slow-time t_n are

$$\tau_{i,n} = \tau_i(t_n) = \frac{1}{c} \left(\|\mathbf{r}_{t,i}(t_n)\| + \|\mathbf{r}_{i,r}(t_n)\| \right), \quad (7)$$

$$\nu_{i,n} = \nu_i(t_n) = -\frac{1}{\lambda} \left(\hat{\mathbf{r}}_{t,i}^H(t_n) \left(\dot{\mathbf{p}}_i(t_n) - \dot{\mathbf{p}}_t(t_n) \right) + \hat{\mathbf{r}}_{i,r}^H(t_n) \left(\dot{\mathbf{p}}_r(t_n) - \dot{\mathbf{p}}_i(t_n) \right) \right), \quad (8)$$

$$\tilde{\tau}_{i,n} = \tilde{\tau}_i(t_n) = \frac{1}{c} \left(\|\tilde{\mathbf{r}}_{t,i}(t_n)\| + \|\tilde{\mathbf{r}}_{i,r}(t_n)\| \right), \quad (9)$$

$$\tilde{\nu}_{i,n} = \tilde{\nu}_i(t_n) = -\frac{1}{\lambda} \left(\hat{\tilde{\mathbf{r}}}_{t,i}^H(t_n) \left(\dot{\mathbf{p}}_i(t_n) - \dot{\tilde{\mathbf{p}}}_t(t_n) \right) + \hat{\tilde{\mathbf{r}}}_{i,r}^H(t_n) \left(\dot{\tilde{\mathbf{p}}}_r(t_n) - \dot{\mathbf{p}}_i(t_n) \right) \right), \quad (10)$$

where c is the speed of light, $\lambda \in \mathbb{R}$ is the transmitted signal's carrier wavelength, and $\dot{\mathbf{p}}_i, \dot{\mathbf{p}}_t, \dot{\mathbf{p}}_r, \dot{\tilde{\mathbf{p}}}_t, \dot{\tilde{\mathbf{p}}}_r : \mathbb{R} \rightarrow \mathbb{R}^3$ are vector-valued functions of time providing the true

and measured velocity of the i th scatterer, transmitter, and receiver, respectively.

Next, let the n th baseband signal emitted by the emitter be an L -length complex-valued signal denoted by $\mathbf{u}_n \in \mathbb{C}^L$. Then, making stop-and-hop and narrowband signal approximations, the discretized baseband received signal, $\mathbf{s}_n \in \mathbb{C}^L$, for the n th slow-time pulse may be modeled as

$$\mathbf{s}_n = \sum_{m=0}^{M-1} \gamma_{m,n} e^{-j\omega_o \tau_{m,n}} \mathcal{D}(\tau_{m,n}, \nu_{m,n}) \mathbf{u}_n + \mathbf{n}_n, \quad (11)$$

where M is the number of stationary point scatterers, $\gamma_{m,n} \in \mathbb{C}$ is the complex-valued reflectivity for the m th scatterer during the n th slow-time pulse, $\omega_o = 2\pi f_o \in \mathbb{R}$ is the angular carrier frequency of the transmitted signal, $\mathcal{D} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}^{L \times L}$ is the delay-Doppler operator, $\mathbf{n}_n \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I})$ is circularly symmetric additive white Gaussian noise (AWGN), and $\sigma_n^2 \in \mathbb{R}$ is the received signal noise power (assumed constant across all slow-time pulses). The unitary delay-Doppler operator for an L -length signal sampled with a sampling rate of f_s is

$$\mathcal{D}(\tau, \nu) = e^{j2\pi\nu\tau} \mathbf{F}^H \text{Diag}(\boldsymbol{\tau}) \mathbf{F} \text{Diag}(\boldsymbol{\nu}), \quad (12)$$

where $\mathbf{F} \in \mathbb{C}^{L \times L}$ is a unitary discrete Fourier transform (DFT) matrix, the ℓ th element of the delay linear phase ramp, $\boldsymbol{\tau} \in \mathbb{C}^L$, is

$$[\boldsymbol{\tau}]_\ell = e^{-j2\pi \frac{f_s}{L} [\ell - L/2] \tau}, \quad (13)$$

and the ℓ th element of the Doppler shift linear phase ramp, $\boldsymbol{\nu} \in \mathbb{C}^L$, is

$$[\boldsymbol{\nu}]_\ell = e^{j2\pi\nu \frac{\ell}{f_s}}. \quad (14)$$

Figure 3 depicts a high-level processing chain for forming a SAR image given the baseband, received signal model in (11). In the discussion to follow each of the shown processing steps are discussed.

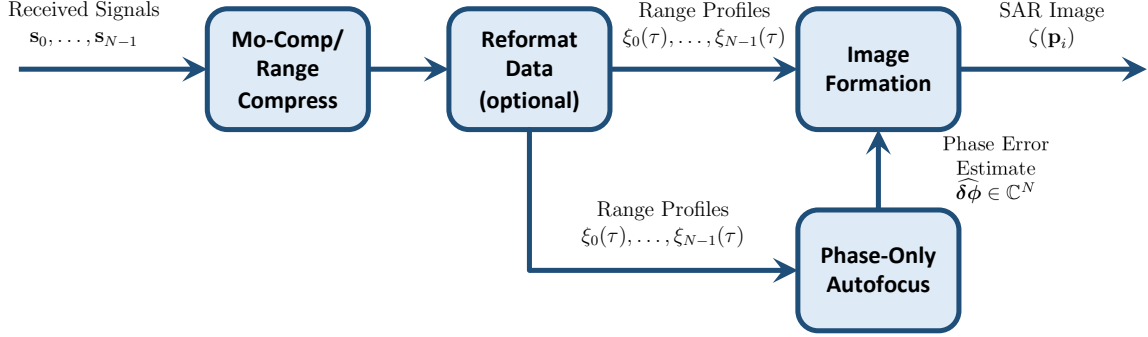


Figure 3. High-Level SAR imaging processing chain.

To form a SAR image, the received signal is first motion compensated (i.e., mo-comp) to the measured bistatic delay to the scene center, $\tilde{\tau}_{0,n}$, and range compressed via matched filtering. Assuming measurement errors are present, the n th range profile or range compressed signal, $\xi_n : \mathbb{R} \rightarrow \mathbb{C}$, at the measured bistatic delay to the i th scatterer, $\tilde{\tau}_{i,n}$, is

$$\begin{aligned} \xi_n(\tilde{\tau}_{i,n}) &= \left(e^{-j\omega_o \tilde{\tau}_{0,n}} \mathcal{D}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}) \mathbf{u}_n \right)^H \mathbf{s}_n \\ &= \gamma_{i,n} e^{-j\omega_o(\tau_{i,n} - \tilde{\tau}_{0,n})} \chi_{i,n}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}) + v_{i,n} + w_{i,n}, \end{aligned} \quad (15)$$

where the matched filtered signal is Doppler compensated to the measured bistatic Doppler to the scene center $\tilde{\nu}_{0,n}$, $\chi_{i,n} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$ is the n th transmitted signal's ambiguity function whose peak response is located at $(\tau_{i,n}, \nu_{i,n})$, and $v_{i,n}, w_{i,n} \in \mathbb{C}$ are interference and noise terms for the i th scatterer during the n th slow-time pulse, respectively. The n th transmitted signal's ambiguity function, whose peak response

is located at $(\tau_{i,n}, \nu_{i,n})$, is

$$\chi_{i,n}(\tau, \nu) = (\mathcal{D}(\tau, \nu) \mathbf{u}_n)^H \mathcal{D}(\tau_{i,n}, \nu_{i,n}) \mathbf{u}_n. \quad (16)$$

The interference and noise terms are

$$v_{i,n} = \sum_{\substack{m=0 \\ m \neq i}}^M \gamma_{m,n} e^{-j\omega_o(\tau_{m,n} - \tilde{\tau}_{0,n})} \chi_{m,n}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}), \quad (17)$$

and

$$w_{i,n} = (e^{-j\omega_o \tilde{\tau}_{0,n}} \mathcal{D}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}) \mathbf{u}_n)^H \mathbf{n}_n, \quad (18)$$

where $w_{i,n} \sim \mathcal{CN}(0, \sigma_w^2)$ and $\sigma_w^2 = \sigma_{\mathbf{n}}^2 \|\mathbf{u}_n\|^2$ is assumed constant across all slow-time pulses (i.e., assume $\|\mathbf{u}_n\|^2 = \|\mathbf{u}_{n'}\|^2$ for $n, n' \in \{0, \dots, N-1\}$).

Upon performing motion compensation and range compression, the next step is phase-only autofocus. The key step of phase-only autofocus is the estimation of phase errors which manifest in the range compressed data during the motion compensation step due to motion measurement errors. To convey the detrimental effect phase errors may potentially have on a SAR image, example ideal and defocused SAR images are shown in Figure 4. The defocused image in Figure 4 is produced by applying synthetic phase errors consisting of white phase errors for each slow-time pulse chosen from a uniform distribution over the interval $[-\pi, \pi)$. From the example images depicted in Figure 4, the potential need for performing autofocus is made apparent.

From (15), the ideal phase error estimate, $\widehat{\delta\phi} \in \mathbb{C}^N$, is (this result is made more

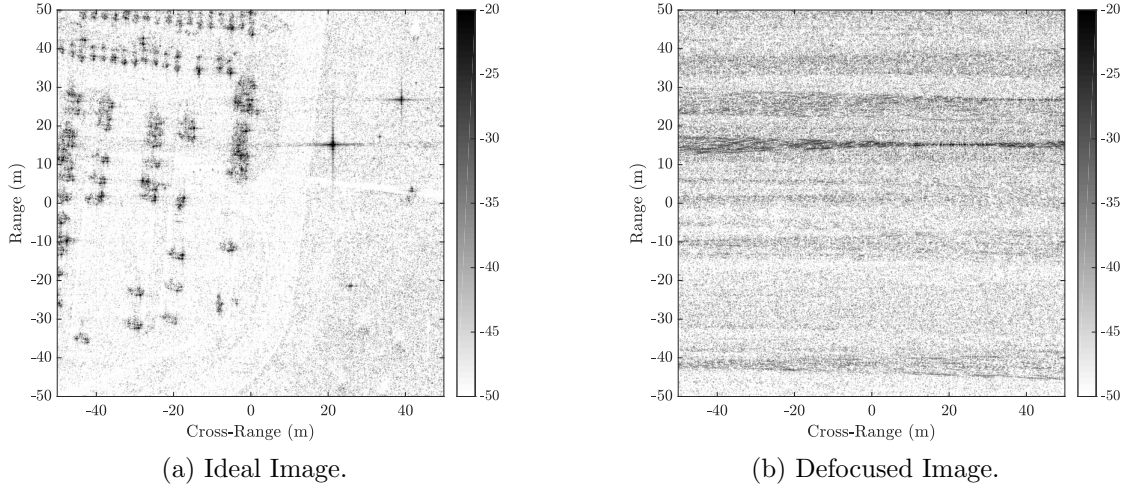


Figure 4. Example ideal and defocused SAR images.

evident in Section 3.2)

$$\left[\widehat{\delta\phi} \right]_n = e^{j\omega_o(\tilde{\tau}_{0,n} - \tau_{0,n})}, \quad (19)$$

and applied to the range compressed data as

$$\xi_n(\tilde{\tau}_{i,n}) := \left[\widehat{\delta\phi} \right]_n^* \xi_n(\tilde{\tau}_{i,n}). \quad (20)$$

From inspection of (15), (19) and (20), it is seen that only the leading phase error in (15) is corrected or compensated for, while the motion measurement errors present in the delay-Doppler operator of the matched filtered signal are ignored. Ignoring the motion measurement errors present in the matched filtered signal is what draws the distinction between phase-only and range-cell migration autofocus algorithms. More specifically, range-cell migration autofocus attempts to correct for both the leading phase error and the motion measurement errors in the matched filtered signal, while phase-only autofocus settles for correcting only the leading phase error. In Section 3.2 analysis is provided to identify when the motion measurement errors present in the

matched filtered signal may be ignored (i.e., when phase-only autofocus is applicable).

To better see how phase errors manifest in a formed SAR image, for the remainder of this section the phase error estimate is assumed to be $\widehat{\delta\phi} = \mathbf{1}$ (i.e., the discussion proceeds as though no autofocus is performed). In Section 2.2 additional discussion is provided on the phase estimation step of phase-only autofocus.

Upon performing autofocus, the next step is to form a SAR image. The SAR image, $\zeta : \mathbb{R}^3 \rightarrow \mathbb{C}$, at location $\mathbf{p}_i$ is

$$\zeta(\mathbf{p}_i) = \tilde{\phi}_i^H \boldsymbol{\xi}_i = \mathbf{1}^H \tilde{\boldsymbol{\xi}}_i, \quad (21)$$

where $\tilde{\phi}_i \in \mathbb{C}^N$ is the measured complex exponential phase vector applied to produce the SAR image at location $\mathbf{p}_i$, $\boldsymbol{\xi}_i \in \mathbb{C}^N$ consists of the N range profile samples contributing to the SAR image at location $\mathbf{p}_i$, and $\tilde{\boldsymbol{\xi}}_i = \boldsymbol{\xi}_i \odot \tilde{\phi}_i^*$ contains the phase compensated range profile samples contributing to the SAR image at location $\mathbf{p}_i$. In (21) the dependence of the SAR image on the range profile samples, phase error estimate, and estimated complex exponential phase vector is made implicit.

The expression in (21) encompasses multiple image formation processes, as each point of the SAR image is expressed as a coherent sum of phase compensated range profile samples. Figure 5 depicts how $\tilde{\phi}_i$, $\boldsymbol{\xi}_i$, and $\tilde{\boldsymbol{\xi}}_i$ are defined for two of the most popular image formation processes: 1) PFA and 2) BPA. For the PFA, the phase vector $\tilde{\phi}_i$ is a Fourier basis vector, and $\boldsymbol{\xi}_i$ are the range profile samples obtained after resampling the phase history data from a polar to a uniformly sampled Cartesian grid (i.e., the data is reformatted in Figure 3 and Figure 5a). For the BPA, the n th element of the phase vector is $[\tilde{\phi}_i]_n = e^{-j\omega_o(\tilde{\tau}_{i,n} - \tilde{\tau}_{0,n})}$, and the n th element of the range profile samples is $[\boldsymbol{\xi}_i]_n = \xi_n(\tilde{\tau}_{i,n})$ (i.e., the data is not reformatted in Figure 3 and Figure 5b).

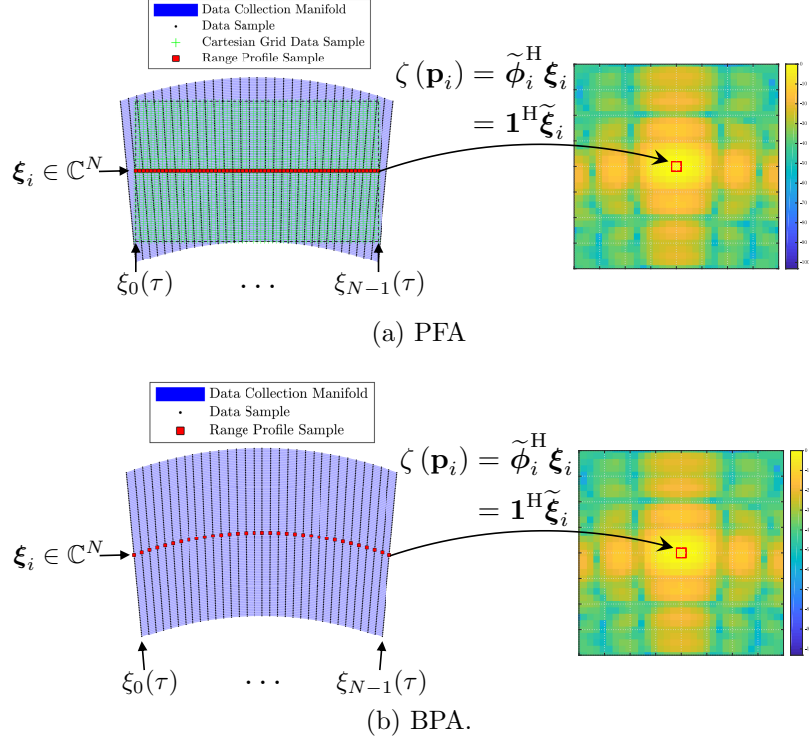


Figure 5. PFA and BPA SAR image formation processes.

In the remainder of this section, the expressions introduced for the BPA are used to build the phase-only autofocus signal model. It is important to note that using the BPA for analysis does not preclude the PFA, since loosely speaking, the PFA is a special case of the BPA. In this special case, the stand-off ranges of the transmitter and receiver to the scene center approaches infinity such that the data collection manifold maps out a uniformly sampled Cartesian grid. Hence, the need for performing polar-to-Cartesian resampling of the phase history data is alleviated (i.e., there is zero wavefront curvature such that the received data already lies on a uniformly sampled Cartesian grid).

Using the expressions from (15) to (21) for the BPA, the n th phase compensated

range profile sample contributing to the SAR image at location $\mathbf{p}_i$ is

$$\left[\tilde{\boldsymbol{\xi}}_i\right]_n = e^{j\omega_o\delta\tau_{i,n}} \left(\gamma_{i,n}\chi_{i,n}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}) + v_{i,n}\right) + w_{i,n}, \quad (22)$$

where $\delta\tau_{i,n} = \tilde{\tau}_{i,n} - \tau_{i,n}$ is the measurement error for the bistatic delay to the i th scatterer during the n th slow-time pulse, and the interference and noise terms after phase compensation (i.e., $\tilde{\boldsymbol{\xi}}_i = \boldsymbol{\xi}_i \odot \tilde{\boldsymbol{\phi}}_i^*$) are redefined as

$$v_{i,n} := e^{-j\omega_o\delta\tau_{i,n}} \left[\tilde{\boldsymbol{\phi}}_i\right]_n^* v_{i,n} = \sum_{\substack{m=0 \\ m \neq i}}^M \gamma_{m,n} e^{j\omega_o(\tau_{i,n} - \tau_{m,n})} \chi_{m,n}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}), \quad (23)$$

and

$$w_{i,n} := \left[\tilde{\boldsymbol{\phi}}_i\right]_n^* w_{i,n} = \left(e^{-j\omega_o\tilde{\tau}_{i,n}} \boldsymbol{\mathcal{D}}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}) \mathbf{u}_n\right)^H \mathbf{n}_n. \quad (24)$$

Given the expression in (22), the vector of phase compensated range profile samples contributing to the SAR image at location $\mathbf{p}_i$ is

$$\tilde{\boldsymbol{\xi}}_i = \text{Diag}(\boldsymbol{\delta\phi}) \boldsymbol{\gamma}_i + \mathbf{w}_i, \quad (25)$$

where each of the terms is defined as follows:

$$[\boldsymbol{\delta\phi}]_n = e^{j\omega_o\delta\tau_{i,n}}, \quad (26)$$

$$[\boldsymbol{\gamma}_i]_n = \gamma_{i,n}\chi_{i,n}(\tilde{\tau}_{i,n}, \tilde{\nu}_{0,n}) + v_{i,n}, \quad (27)$$

$$[\mathbf{w}_i]_n = w_{i,n}. \quad (28)$$

In (26), the phase error vector is spatially variant as the quantity depends upon i (i.e., the phase error vector changes for each location $\mathbf{p}_i$ in the SAR image). However, in

Section 3.2 it is shown that for appropriate SAR scene sizes the spatially variant phase error vector is approximately spatially invariant (i.e., the phase error vector is approximately independent of the location $\mathbf{p}_i$ in the SAR image for reasonable SAR scene sizes). Herein, it is assumed an appropriate SAR scene size is chosen so the spatially invariant phase error model is applicable. Hence, a subscript i is omitted from the phase error vector $\boldsymbol{\delta\phi}$ in (25) and (26).

The model in (25) represents the phase-only autofocus signal model assumed for phase-only autofocus algorithms, which includes the GPGA algorithm. In the next section, the phase estimation step for phase-only autofocus is described in greater detail and the relationship between the GPGA and other existing phase-only autofocus algorithms is discussed.

2.2 Phase Estimation for Phase-Only Autofocus

In [30] it is shown that generally speaking, phase error estimation for phase-only autofocus requires solving or finding an approximate solution (typically a local optimal point) to an optimization problem, whose objective function consists of an image mask or weighting function and a point transformation (possibly nonlinear) of the SAR image intensities. Herein, this optimization problem is referred to as the phase estimation problem.

Let the image intensity, $I : \mathbb{R}^3 \times \mathbb{C}^N \rightarrow \mathbb{R}$, be

$$I(\mathbf{p}_i, \widehat{\boldsymbol{\delta\phi}}) = |\zeta(\mathbf{p}_i)|^2 = \left| \widehat{\boldsymbol{\delta\phi}}^H \tilde{\boldsymbol{\xi}}_i \right|^2, \quad (29)$$

recalling the dependence of the SAR image on the phase error estimate is made

implicit in (21) and $\tilde{\boldsymbol{\xi}}_i$ is the vector of phase compensated range profile samples defined in (25) (i.e., is the range profile samples omitting the phase error correction for autofocus in (20)). Then, in general the phase estimation problem for performing phase-only autofocus is

$$\begin{aligned} & \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} && \sum_{i=0}^{M-1} m(\mathbf{p}_i, \widehat{\boldsymbol{\delta\phi}}) \Gamma(I(\mathbf{p}_i, \widehat{\boldsymbol{\delta\phi}})), \\ & \text{subject to} && \left| \left[\widehat{\boldsymbol{\delta\phi}} \right]_n \right|^2 = 1, \quad \text{for } n \in \{0, \dots, N-1\}, \end{aligned} \quad (30)$$

where $m : \mathbb{R}^3 \times \mathbb{C}^N \rightarrow \mathbb{R}$ is the image mask and $\Gamma : \mathbb{R} \rightarrow \mathbb{R}$ is the point transformation of the SAR image intensities referred to as the generalized sharpness function (GSF) [30]. From the phase estimation problem in (30) it is not clear how the GSF is chosen. In a few cases the GSF can be derived as a particular estimator for a given statistical model. In most other cases, the GSF is chosen heuristically [39]. The heuristics involved with choosing a GSF largely follow from intuition about designer metrics which are most suitable for a given scene type, scene content, and/or scenario (see [30] and references therein for discussion on common designer metrics). In the four sections to follow, the GSF for the GPGA and other existing autofocus algorithms is described.

In general, the phase estimation problem in (30) is NP-hard [69, 70]. Thus, for SAR imaging applications, where the number of slow-time pulses is typically on the order of $N \sim 10^2$ or more, it is not reasonable to expect to always solve (30) (i.e., find a global optimal point). Instead, heuristics are needed to obtain an approximate solution to (30) in a reasonably efficient manner. Given heuristics are necessary to perform autofocus, it is sensible to evaluate autofocus algorithms using three measures: 1) *performance/robustness*, in terms of accurately estimating the phase errors for a number of scene types, scene content, and scenarios; 2) *efficiency*, in terms of

the run-time and computational complexity of the autofocus algorithm; and 3) *sub-optimality*, in terms of the suboptimality of the phase error estimate with respect to a global optimal point of (30). In the four sections to follow, the GPGA and other existing autofocus algorithms are described and discussed in terms of these three measures.

2.2.1 Generalized Phase Gradient Autofocus

The GPGA algorithm (and thereby the PGA algorithm) is one of the few autofocus algorithms where the GSF can be derived as a particular estimator for a given statistical model. In Section 4.7 this derivation is presented in detail, showing the GSF is the identity function, $\Gamma(x) = x$. The image mask or weighting function for the GPGA algorithm is designed through a scatterer selection process applied to a formed defocused SAR image yielding

$$m(\mathbf{p}_i, \widehat{\boldsymbol{\delta\phi}}) = \begin{cases} 1, & i \in \mathcal{T}, \\ 0, & i \notin \mathcal{T}, \end{cases} \quad (31)$$

where $\mathcal{T}$ is an index set for moderate or nearly constant signal-to-interference-plus-noise ratio (SINR) scatterers. In Chapter IV strategies for performing scatterer selection are described in detail. Additionally, in Chapter IV it is shown that with the PGA algorithm, an additional constraint is placed on the indexing set used in (31) such that only a single scatterer per range cut may be selected during the scatterer selection process. Using the defined GSF and image mask, the GPGA phase estimation

problem is

$$\begin{aligned}
& \underset{\widehat{\delta\phi} \in \mathbb{C}^N}{\text{maximize}} && \sum_{i \in \mathcal{T}} \left| \widehat{\delta\phi}^H \tilde{\xi}_i \right|^2, \\
& \text{subject to} && \left| \left[\widehat{\delta\phi} \right]_n \right|^2 = 1, \quad \text{for } n \in \{0, \dots, N-1\}.
\end{aligned} \tag{32}$$

In Section 4.7 it is shown that a global optimal point to the GPGA phase estimation problem is an MMLE of the phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers. Thus, an approximate solution to the GPGA phase estimation problem is as an approximate MMLE of the phase errors. In Chapter V, three GPGA phase estimators or approximate MMLEs, two of which are derived from existing PGA phase estimators, are described in detail. These three phase estimators are referred to as: 1) phase difference (PD), 2) eigenvector relaxation (EVR), and 3) Max-SDR phase estimators, with the first two being derived from existing PGA phase estimators. In Chapter V the Max-SDR phase estimator is shown to yield an approximate MMLE of the phase errors with a worst-case suboptimality compared to the solution set of MMLEs (i.e., provides a strong statement regarding the suboptimality of the phase estimator). In Chapter VI results produced from synthetic and measured data sets using each of the three mentioned phase estimators are presented and discussed.

2.2.2 Multichannel Autofocus

With the obvious exception of the PGA algorithm, the multichannel autofocus (MCA) algorithm¹ is arguably the most closely related existing autofocus algorithm to the

¹Herein, MCA algorithm refers to the class of MCA algorithms developed in [42, 45, 46, 48, 49, 51, 55].

GPGA algorithm². For instance, the GPGA and MCA algorithms share the same generalized sharpness function, $\Gamma(x) = x$ [42, 45, 46, 48, 49, 51, 55]. Moreover, for both the GPGA and MCA algorithms the GSF can be derived as a particular estimator for a given statistical model. For details regarding the MCA statistical model the reader is referred to [51]. In large part, where the GPGA and MCA algorithms differ is in how the image mask is designed. With the GPGA algorithm the image mask is designed to maximize the GSF over moderate or nearly constant SINR scatterers, whereas with the MCA algorithm the image mask is designed to minimize the GSF over low-return regions [42, 45, 46, 48, 49, 51, 55]. In particular, the image mask for the MCA algorithm is

$$m(\mathbf{p}_i, \widehat{\boldsymbol{\delta\phi}}) = \begin{cases} -1, & i \in \mathcal{T}, \\ 0, & i \notin \mathcal{T}, \end{cases} \quad (33)$$

where $\mathcal{T}$ is an index set for low-return regions in the SAR image. In [42, 45, 46, 48, 49, 51, 55] it is assumed low-return regions are known *a priori* based on the antenna patterns of the transmit and receive antennas. Using the defined GSF and image mask, the MCA phase estimation problem is

$$\begin{aligned} & \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} && - \sum_{i \in \mathcal{T}} \left| \widehat{\boldsymbol{\delta\phi}}^H \tilde{\boldsymbol{\xi}}_i \right|^2, \\ & \text{subject to} && \left| \left[\widehat{\boldsymbol{\delta\phi}} \right]_n \right|^2 = 1, \quad \text{for } n \in \{0, \dots, N-1\}. \end{aligned} \quad (34)$$

In addition to sharing the same GSF, the MCA and GPGA algorithms also share two similar phase estimators, which provide an approximate solution to their respective phase estimation problems. These shared phase estimators are the EVR and SDR phase estimators. For specifics regarding how these phase estimators are applied with

²In fact, the work in [55] initially sparked both the author's and advisor's interest into studying autofocus algorithms.

the MCA algorithm the reader is referred to [55]. Similar to the Max-SDR phase estimator mentioned with the GPGA algorithm, the SDR phase estimator described in [55] also provides a phase error estimate with a worst-case approximation bound compared to the solution set of the MCA phase estimation problem in (34) (i.e., provides a strong statement regarding the suboptimality of the phase estimator).

2.2.3 Maximum Sharpness Autofocus

The maximum sharpness autofocus algorithm is another autofocus algorithm whose GSF can be derived as a particular estimator for a given statistical model³. For details regarding the given statistical model the reader is referred to [39]. The GSF derived for maximum sharpness autofocus is $\Gamma(x) = x^2$. In the literature, the image mask or weighting function for maximum sharpness autofocus is generally treated as being optional (i.e., $m(\mathbf{p}_i, \widehat{\boldsymbol{\delta\phi}}) = 1$) [26, 28, 30, 38, 50]. Although, as suggested in [26] employing a strategy for designing and applying a non-constant image mask could potentially lead to an improved maximum sharpness autofocus algorithm. Using the defined GSF and image mask, the maximum sharpness phase estimation problem is

$$\begin{aligned} & \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} && \sum_{i=0}^{M-1} \left| \widehat{\boldsymbol{\delta\phi}}^H \tilde{\boldsymbol{\xi}}_i \right|^4, \\ & \text{subject to} && \left| \left[\widehat{\boldsymbol{\delta\phi}} \right]_n \right|^2 = 1, \quad \text{for } n \in \{0, \dots, N-1\}. \end{aligned} \tag{35}$$

For maximum sharpness phase estimation, [26, 28, 30, 38, 50] propose using either gradient or coordinate descent to obtain a local optimal point to the maximum sharpness phase estimation problem. Regardless of the chosen descent technique, a local optimal point is obtained using only a first-order optimality condition. Thus, little can be

³This is true at least in the case where the image is formed using the PFA. It is not clear or evident from the work in [39] that the same derived result holds for an image formed using the BPA.

said about the suboptimality of the phase error estimate with respect to the solution set of the maximum sharpness phase estimation problem. Moreover, it is difficult to make statements regarding the efficiency or computational complexity of either descent technique, as their efficiency and computational complexity largely depends upon the quality of an initialization point.

2.2.4 Minimum Entropy Autofocus

For minimum entropy autofocus the GSF is $\Gamma(x) = -x \ln(x)$, and the image mask is

$$m(\mathbf{p}_i, \widehat{\boldsymbol{\delta\phi}}) = \frac{-1}{\sum_{i=0}^{M-1} \left| \widehat{\boldsymbol{\delta\phi}}^{\text{H}\sim} \boldsymbol{\xi}_i \right|^2}. \quad (36)$$

Substituting these functions in (30) forms what is termed the minimum entropy phase estimation problem

$$\begin{aligned} & \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} \quad \frac{1}{\sum_{i=0}^{M-1} \left| \widehat{\boldsymbol{\delta\phi}}^{\text{H}\sim} \boldsymbol{\xi}_i \right|^2} \sum_{i=0}^{M-1} \left| \widehat{\boldsymbol{\delta\phi}}^{\text{H}\sim} \boldsymbol{\xi}_i \right|^2 \ln \left(\left| \widehat{\boldsymbol{\delta\phi}}^{\text{H}\sim} \boldsymbol{\xi}_i \right|^2 \right), \\ & \text{subject to} \quad \left| \left[\widehat{\boldsymbol{\delta\phi}} \right]_n \right|^2 = 1, \quad \text{for } n \in \{0, \dots, N-1\}. \end{aligned} \quad (37)$$

In [33] it is noted that there is no information-theoretic connection for the minimum entropy GSF. Rather, the minimum entropy GSF is merely chosen as a mathematically convenient measure of the focus of a SAR image (i.e., is chosen as a heuristic).

To attempt to solve the minimum entropy phase estimation problem in (37), [33] opts to use optimization transfer. That is, [33] constructs a surrogate GSF $\Gamma(x; x^{(\ell)}) = -x \ln(x^{(\ell)})$, where x is the original value and $x^{(\ell)}$ is the value obtained after the ℓ th iteration of an optimization algorithm. Replacing the GSF in (37) with the surrogate

GSF produces the surrogate minimum entropy phase estimation problem

$$\begin{aligned}
& \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} && \frac{1}{\sum_{i=0}^{M-1} \left| \widehat{\boldsymbol{\delta\phi}}^H \widetilde{\boldsymbol{\xi}}_i \right|^2} \sum_{i=0}^{M-1} \left| \widehat{\boldsymbol{\delta\phi}}^H \widetilde{\boldsymbol{\xi}}_i \right|^2 \ln \left(\left| \widetilde{\boldsymbol{\xi}}_i^H \widehat{\boldsymbol{\delta\phi}}^{(\ell)} \right|^2 \right), \\
& \text{subject to} && \left| \left[\widehat{\boldsymbol{\delta\phi}} \right]_n \right|^2 = 1, \quad \text{for } n \in \{0, \dots, N-1\},
\end{aligned} \tag{38}$$

where $\widehat{\boldsymbol{\delta\phi}}^{(\ell)}$ is the phase error estimate after the ℓ th iteration of the minimum entropy autofocus algorithm.

In [33] coordinate descent is first proposed for obtaining a local optimal point to the surrogate minimum entropy phase estimation problem. However, since SAR imaging applications typically involve a large number of slow-time pulses, coordinate descent tends to be computationally inefficient. To improve computational efficiency, [33] proposes a heuristic where all the coordinates of the phase error estimate are updated simultaneously. In [38], this simultaneous update approach is shown to yield the same local optimal point obtained with coordinate descent, assuming the initialization point is within some local region around the local optimal point.

2.2.5 Summary

Table 1 provides a summary of the trade-offs for each of the autofocus algorithms described in this section. Due to the nature of the phase estimation problem (i.e., NP-hard), it is difficult (if not impossible) to make sweeping claims such as “*autofocus algorithm X outperforms one or all other autofocus algorithms.*” Moreover, the performance of each of the autofocus algorithms depends on design parameters including but not limited to designing an image mask (i.e., scatterer selection) and/or stopping criteria for an optimization routine [14]. Thus, barring tests are run with each autofocus algorithm covering all combinations of design parameters and differ-

ent data sets representing different scene types, scene content, and scenarios, general claims comparing the performance of each of the autofocus algorithms are inappropriate. Considering testing over such a large space is impractical, a more reasonable approach to selecting an autofocus algorithm is: 1) identify a set of autofocus algorithms which are expected or assumed to provide adequate performance for expected scenarios/testing environments then, 2) identify an autofocus algorithm from the set of considered autofocus algorithms by weighing their respective trade-offs.

From the trade-offs listed in Table 1, one may see the GPGA algorithm has utility in a number of cases based on the following properties: 1) the GPGA algorithm has clear statistical motivation as it is an approximate MMLE of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers; 2) results in Chapter VI suggest that for scatterers selected with moderate SINRs the GPGA algorithm is an efficient estimator of phase errors, meaning the CRLB derived in Chapter V may be used to predict expected performance; 3) the GPGA algorithm using Max-SDR phase estimation provides a phase error estimate with a worst-case suboptimality compared to the solution set of the NP-hard phase estimation problem; 4) the computational complexity and run-time performance for the GPGA algorithm is independent of an initialization point (i.e., is deterministic); 5) the GPGA algorithm is one of the most computationally efficient algorithms (by way of comparisons made in the literature between other autofocus algorithms and the PGA algorithm). Moreover, even in the event maximum sharpness or minimum entropy autofocus is the autofocus algorithm of choice, due to the relative efficiency of the GPGA algorithm, the GPGA algorithm may be used for warm-starting the gradient descent methods (i.e., to provide an initialization point) used with the maximum sharpness or minimum entropy autofocus algorithms (similar to the suggestion made in [33] regarding

the PGA algorithm).

2.3 Conclusion

In this section, the phase-only autofocus signal model was formulated beginning with the baseband, received signal model for a bistatic collection geometry. In addition, in this section a high-level description of the objective of phase-only autofocus was provided, and a brief summary of many existing phase-only autofocus algorithms was given. In the next section, analysis is performed to identify when the phase-only autofocus signal model is applicable and the spatially invariant phase error modeling assumption holds.

Table 1. Summary of trade-offs between different autofocus algorithms.

Autofocus Algorithm	Summary
GPGA	- GSF derived from a given statistical model (see Chapter IV) - Derived CRLB provides insight into expected performance (see Chapter V) - Worst-case approximation bound to NP-hard phase estimation problem (see Chapter V) - Deterministic computational complexity and run-time performance (see Chapter V) - One of the most efficient (based on comparability to PGA algorithm)
MCA	- GSF derived from a given statistical model (see [51]) - Little or no insight into expected performance - Worst-case approximation bound to NP-hard phase estimation problem (see [55]) - Deterministic computational complexity and run-time performance (see [55]) - One of the most efficient (comparable with PGA algorithm and hence, GPGA algorithm) [55]
Maximum Sharpness	- GSF derived from a given statistical model (see [39]) - No worst-case approximation bound to NP-hard phase estimation problem - Little or no insight into expected performance - Non-deterministic computational complexity and run-time performance - Tends to be less efficient than MCA [51]
Minimum Entropy	- GSF chosen as a heuristic (see [33]) - No worst-case approximation bound to NP-hard phase estimation problem - Little or no insight into expected performance - Non-deterministic computational complexity and run-time performance - Tends to be less efficient than PGA algorithm (and hence, GPGA algorithm) [33]

III. Matched Filter and Image Formation Analysis

In this section matched filter and image formation analysis is performed to identify bounds on motion measurement errors and SAR scene size limits such that the phase-only defocus model and spatially invariant phase error assumption are applicable.

3.1 Matched Filter Analysis

As mentioned in Section 2.1 regarding the matched filtered response in (15), phase-only autofocus is applicable when motion measurement errors in the delay-Doppler operator of the matched filtered signal have negligible effect on the focus of a SAR image (i.e., may be ignored). In order to ignore these motion measurement errors, the ambiguity function or range compressed scattered return of a scatterer in (15) must coherently combine across slow-time pulses. Thus, at a minimum the ambiguity function samples must correspond to the mainlobe of a scatterer's response at each slow-time pulse. In other words, for phase-only autofocus to be applicable, position and velocity measurement errors must be limited such that the mainlobe response of each scatterer's ambiguity function is sampled in the matched filtered return.

For a band-limited signal, the first null of a scatterer's mainlobe response in range occurs at $\frac{c}{B}$, where B is the two-sided bandwidth of the transmitted signal. Thus, the bound on position uncertainty ensuring the mainlobe of a scatterer's response is sampled in range is

$$\begin{aligned} \beta \frac{c}{B} &\geq \max_{n \in \{0, \dots, N-1\}} \|\delta \mathbf{p}_t(t_n)\| + \|\delta \mathbf{p}_r(t_n)\| \\ &\geq \max_{\substack{i \in \{0, \dots, M-1\} \\ n \in \{0, \dots, N-1\}}} |c \delta \tau_{i,n}|, \end{aligned} \tag{39}$$

where $\beta \in [0, 1]$ is a proportionality constant set by the system designer for an allowable range aberration from the peak of a scatterer's mainlobe response. Using the expressions in (3) to (6), the inequality in (39) is readily obtained by adding the following inequalities

$$|\tilde{r}_{t,i}(t) - r_{t,i}(t)| \leq \|\delta \mathbf{p}_t(t)\|, \quad (40)$$

$$|\tilde{r}_{i,r}(t) - r_{i,r}(t)| \leq \|\delta \mathbf{p}_r(t)\|. \quad (41)$$

The intention of the proportionality constant is to enable the system designer to bound the amount of matched filtered loss due to position uncertainty. For example, letting $\beta = 0.45$ results in no more than roughly 3 dB of matched filtered loss due to position uncertainty. Example results illustrating this case are provided in Section 3.1.1.

To gain more intuition about (39), consider a monostatic case where $\delta \mathbf{p}_t(t) = \delta \mathbf{p}_r(t)$ and $\beta = 1$. Then, the result from (39) simplifies to

$$\frac{c}{2B} \geq \max_{n \in \{0, \dots, N-1\}} \|\delta \mathbf{p}_t(t_n)\|, \quad (42)$$

which implies for the monostatic case that the Euclidean norm of the position uncertainty should be limited to the monostatic range resolution. The derived bound in (42) is equivalent to the position uncertainty bound provided in [3].

For a time-limited signal, the first null of a scatterer's mainlobe response in Doppler occurs at $\frac{1}{T}$, where T is the coherent processing interval (CPI). Assuming no position uncertainty to isolate the effects of velocity uncertainty, the bound on velocity

uncertainty ensuring the mainlobe of a scatterer's response is sampled in Doppler is

$$\begin{aligned} \beta \frac{\lambda}{T} &\geq \max_{\substack{i \in \{0, \dots, M-1\} \\ n \in \{0, \dots, N-1\}}} \left| \hat{\mathbf{r}}_{t,i}^H(t_n) \boldsymbol{\delta} \dot{\mathbf{p}}_t(t_n) - \hat{\mathbf{r}}_{i,r}^H(t_n) \boldsymbol{\delta} \dot{\mathbf{p}}_r(t_n) \right| \\ &= \max_{\substack{i \in \{0, \dots, M-1\} \\ n \in \{0, \dots, N-1\}}} \left| \lambda (\tilde{\nu}_{i,n} - \nu_{i,n}) \right|, \end{aligned} \quad (43)$$

where $\beta \in [0, 1]$ is a proportionality constant set by the system designer for an allowable velocity aberration from the peak of a scatterer's mainlobe response and $\boldsymbol{\delta} \dot{\mathbf{p}}_t, \boldsymbol{\delta} \dot{\mathbf{p}}_r : \mathbb{R} \rightarrow \mathbb{R}^3$ are the transmitter and receiver velocity errors, respectively. Again, the intention of the proportionality constant is to enable the system designer to bound the amount of matched filtered loss due to velocity uncertainty. For example, letting $\beta = 0.45$ results in no more than roughly 3 dB of matched filtered loss due to velocity uncertainty. Example results illustrating this case are also provided in Section 3.1.1.

To gain more intuition about (43), consider a monostatic case where $\hat{\mathbf{r}}_{i,r}(t) = -\hat{\mathbf{r}}_{t,i}(t)$, $\boldsymbol{\delta} \dot{\mathbf{p}}_t(t) = \boldsymbol{\delta} \dot{\mathbf{p}}_r(t)$, and $\beta = 1$. Then, the result from (43) simplifies to

$$\frac{\lambda}{2T} \geq \max_{\substack{i \in \{0, \dots, M-1\} \\ n \in \{0, \dots, N-1\}}} \left| \hat{\mathbf{r}}_{t,i}^H(t_n) \boldsymbol{\delta} \dot{\mathbf{p}}_t(t_n) \right|, \quad (44)$$

which implies for the monostatic case that the magnitude of the velocity uncertainty along the radial component should be limited to the monostatic velocity resolution (see [5] for discussion on the monostatic velocity resolution).

To conclude this section, example results illustrating derived bounds on position and velocity uncertainty for the phase-only defocus model to be applicable are presented and discussed.

3.1.1 Example Results

Consider the bistatic scenario shown in Figure 6 for a single slow-time pulse of a SAR collect. Table 2 is a detailed listing of the parameters for the bistatic scenario shown in Figure 6. With the depicted scenario, a single stationary scatterer located at the scene center is assumed. Such a simple collection geometry is assumed to enable the effects of position and velocity uncertainty to easily be examined for verifying bounds derived in Section 3.1.

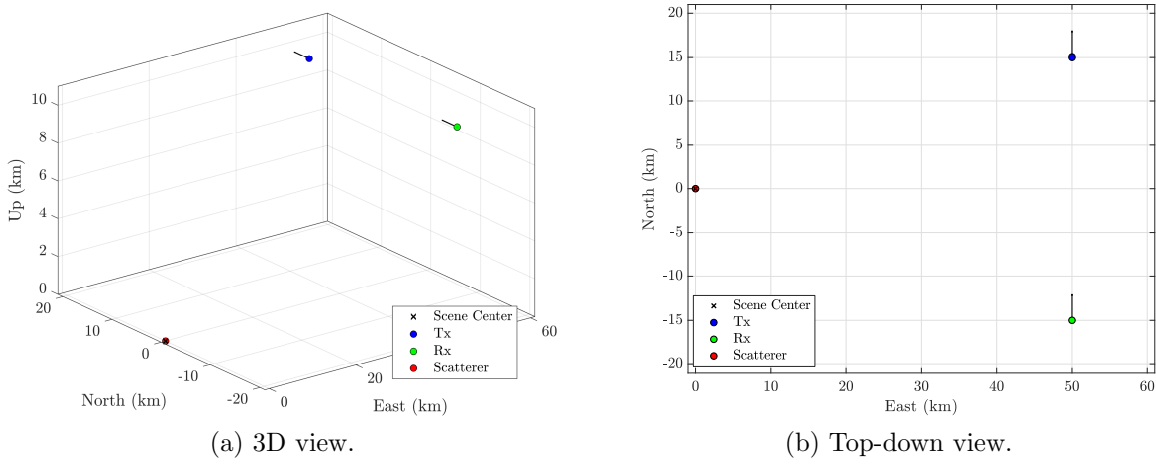


Figure 6. Bistatic scenario for a single slow-time pulse of a SAR collect.

Table 2. Bistatic scenario.

Parameter	Symbol	Value	Units
Transmitter Position	$\mathbf{p}_t$	$[50.00, 15.00, 10.00]^T$	km
Transmitter Position Error	$\delta\mathbf{p}_t$	$[12.69, 3.81, 2.54]^T$	m
Transmitter Velocity	$\dot{\mathbf{p}}_t$	$[0.00, 90.00, 0.00]^T$	m/s
Transmitter Velocity Error	$\delta\dot{\mathbf{p}}_t$	$[2.70, 0.81, 0.54]^T$	m/s
Receiver Position	$\mathbf{p}_r$	$[50.00, -15.00, 10.00]^T$	km
Receiver Position Error	$\delta\mathbf{p}_r$	$[12.69, -3.81, 2.54]^T$	m
Receiver Velocity	$\dot{\mathbf{p}}_r$	$[0.00, 90.00, 0.00]^T$	m/s
Receiver Velocity Error	$\delta\dot{\mathbf{p}}_r$	$[2.70, -0.81, 0.54]^T$	m/s
Scatterer '0' Position	$\mathbf{p}_0$	$[0.00, 0.00, 0.00]^T$	km
Scatterer '0' Velocity	$\dot{\mathbf{p}}_0$	$[0.00, 0.00, 0.00]^T$	m/s

The position and velocity uncertainties for the transmitter and receiver are derived from (39) and (43), where $\beta = 0.45$ to limit the individual matched filtered losses from position and velocity uncertainty to no more than roughly 3 dB. Specifically, the position and velocity uncertainties are

$$\delta \mathbf{p}_t = \beta \frac{c}{2B} \hat{\mathbf{r}}_{t,0}(t_n), \quad (45)$$

$$\delta \dot{\mathbf{p}}_t = \beta \frac{\lambda}{2T} \hat{\mathbf{r}}_{t,0}(t_n), \quad (46)$$

$$\delta \mathbf{p}_r = \beta \frac{c}{2B} \hat{\mathbf{r}}_{0,r}(t_n), \quad (47)$$

$$\delta \dot{\mathbf{p}}_r = \beta \frac{\lambda}{2T} \hat{\mathbf{r}}_{0,r}(t_n). \quad (48)$$

Figure 7 depicts the surface and contours of the scatterer's ambiguity function, $\chi_0(\tau, \nu)$, for the bistatic scenario depicted in Figure 6, where the transmitted signal is a 50 millisecond duration, 5 megahertz bandwidth random phase modulated signal (i.e., $T = 50$ ms and $B = 5$ MHz). Figure 7 also depicts the sample of the scatterer's ambiguity function observed in the matched filtered response in (15), for the given position and velocity uncertainties. As expected, the chosen position and velocity uncertainties in (45) to (48) result in roughly 3 dB of matched filtered loss along both the range and Doppler dimensions.

In this section, bounds on motion measurement errors were derived such that the phase-only defocus model is applicable. In the next section, SAR scene size limits are derived such that the spatially invariant defocus assumption is valid.

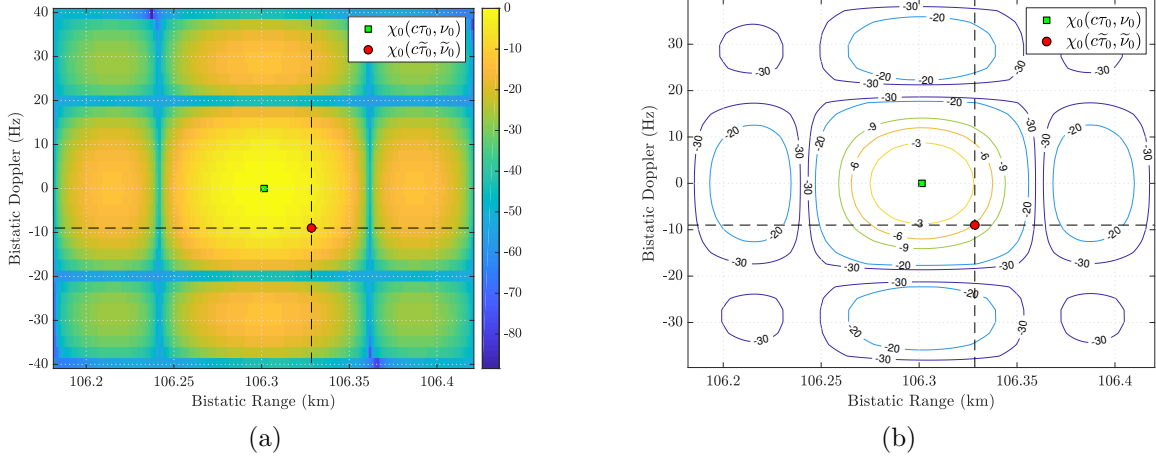


Figure 7. The surface and contours of a scatterer's ambiguity function, and the desired and observed samples in (15), given the position and velocity uncertainties in (45) to (48).

3.2 Image Formation Analysis

A common modeling assumption made with existing phase-only autofocus algorithms, including the GPGA algorithm, is that phase errors, which cause SAR image defocus, are spatially invariant [1–3, 13, 15–17, 19, 20, 23, 26, 28, 30–33, 35, 38, 39, 42, 45, 46, 48, 50, 55, 59, 60]. For far-field imaging with reasonably sized scenes the phase errors are approximately spatially invariant. However, as larger scene sizes or scenarios where the far-field approximation does not hold are considered, the spatially invariant phase error approximation breaks down.

A common approach for handling spatially variant defocus is to divide a larger scene into smaller sub-images over which the phase errors are approximately spatially invariant [2]. Then, autofocus is performed on each sub-image and each corresponding SAR sub-image is formed. To conclude the process, the full SAR image is formed by mosaicing together the individual SAR sub-images. Although this divide and conquer method has proven useful, a question still remains — how large should the sub-images be such that the phase errors are spatially invariant? In the proceeding discussion,

this question is answered by deriving an approximate upper bound on differential bistatic range errors for near- and far-field bistatic geometries, with monostatic as a special case. The derived bound is expressed in terms of known quantities or quantities which typically have a known upper bound.

3.2.1 Differential Range Error

To simplify notation in this section, focus is given to the bistatic geometry for an arbitrarily chosen pulse or slow-time recording of a SAR scenario, so that any time dependence of the transmitter or receiver position functions may be omitted (see the notation in Figure 8). By letting the slow-time pulse be arbitrarily chosen, a similar analysis follows for each pulse of a SAR collect.

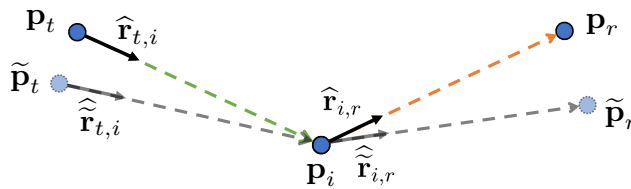


Figure 8. Bistatic geometry.

As shown in Section 2.1, the spatial variance of the phase errors is a direct result of the measured bistatic delay error $\delta\tau_i$. The measured bistatic delay error may be expressed in terms of a spatially variant and invariant component as

$$\begin{aligned} \delta\tau_i &= \widetilde{\tau}_i - \tau_i \\ &= (\widetilde{\tau}_{d,i} + \widetilde{\tau}_0) - (\tau_{d,i} + \tau_0) \\ &= \underbrace{\delta\tau_{d,i}}_{\text{spatially variant}} + \underbrace{\delta\tau_0}_{\text{spatially invariant}}, \end{aligned} \quad (49)$$

where $\tilde{\tau}_{d,i}, \tau_{d,i} \in \mathbb{R}$ are the measured and true differential bistatic delays to $\mathbf{p}_i$, respectively, $\tilde{\tau}_0, \tau_0 \in \mathbb{R}$ are the measured and true bistatic delays to the scene center, respectively, $\delta\tau_{d,i} = \tilde{\tau}_{d,i} - \tau_{d,i}$ is the differential bistatic delay error to $\mathbf{p}_i$, and $\delta\tau_0 = \tilde{\tau}_0 - \tau_0$ is the bistatic delay error to the scene center. The measured and true differential bistatic delays to $\mathbf{p}_i$ are

$$\tilde{\tau}_{d,i} = \tilde{\tau}_i - \tilde{\tau}_0, \quad (50)$$

$$\tau_{d,i} = \tau_i - \tau_0. \quad (51)$$

Using the vectors from (3) to (6), the differential bistatic range error is

$$\begin{aligned} c\delta\tau_{d,i} &= c(\tilde{\tau}_{d,i} - \tau_{d,i}) \\ &= (\|\tilde{\mathbf{r}}_{t,i}\| + \|\tilde{\mathbf{r}}_{i,r}\| - \|\tilde{\mathbf{r}}_{t,0}\| - \|\tilde{\mathbf{r}}_{0,r}\|) - (\|\mathbf{r}_{t,i}\| + \|\mathbf{r}_{i,r}\| - \|\mathbf{r}_{t,0}\| - \|\mathbf{r}_{0,r}\|). \end{aligned} \quad (52)$$

The measured differential range from the transmitter-to- $\mathbf{p}_i$ is

$$\begin{aligned} \|\tilde{\mathbf{r}}_{t,d,i}\| &= \|\tilde{\mathbf{r}}_{t,i}\| - \|\tilde{\mathbf{r}}_{t,0}\| \\ &= \|\mathbf{p}_i - \tilde{\mathbf{p}}_t\| - \|\tilde{\mathbf{p}}_t\|, \end{aligned} \quad (53)$$

where the coordinate system is assumed to be fixed such that $\mathbf{p}_0 = \mathbf{0}$ (i.e., the scene center is the origin of the coordinate system). Similar definitions follow for the measured differential range from $\mathbf{p}_i$ -to-receiver, as well as when no measurement errors are present on the transmitter and receiver positions.

In the next section, a Maclaurin series expansion is used for $\|\tilde{\mathbf{r}}_{t,i}\|$ to express (53) as a convergent series. Then, using a similar expression for when no measurement errors are present, an approximate upper bound on the differential range error from (52) is derived.

3.2.2 Approximate Differential Range Error Bound

Using the Binomial theorem and letting $z = x + y$, the Maclaurin series expansion for $\sqrt{1+z}$ is

$$\begin{aligned}\sqrt{1+z} &= 1 + \frac{z}{2} + \sum_{k=2}^{+\infty} \binom{1/2}{k} z^k \\ &= 1 + \frac{x+y}{2} + \sum_{k=2}^{+\infty} \sum_{\ell=0}^k \binom{1/2}{k} \binom{k}{\ell} x^{k-\ell} y^\ell,\end{aligned}\quad (54)$$

where $\binom{1/2}{k}$ is the generalized binomial coefficient [71]. Recognizing $\sqrt{1+z}$ is real analytic on $(-1, +\infty)$ and using the ratio test, one may see the series in (54) converges to $\sqrt{1+z}$ for $z \in (-1, 1)$.

To leverage the Maclaurin series in (54), the Euclidean distance between $\mathbf{p}_i$ and $\tilde{\mathbf{p}}_t$ is expressed as

$$\begin{aligned}\|\tilde{\mathbf{r}}_{t,i}\| &= \sqrt{\|\tilde{\mathbf{p}}_t\|^2 + \|\mathbf{p}_i\|^2 - 2\mathbf{p}_i^T \tilde{\mathbf{p}}_t} \\ &= \|\tilde{\mathbf{p}}_t\| \sqrt{1 + \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2} - 2\frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|^2}}.\end{aligned}\quad (55)$$

Applying (54) to (55) with $x = \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2}$ and $y = -2\frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|^2}$, and substituting back in (53) produces¹

$$\|\tilde{\mathbf{r}}_{t,d,i}\| = -\frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|} + \frac{1}{2} \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|} + \sum_{k=2}^{+\infty} \sum_{\ell=0}^k a_{k,\ell} \frac{\|\mathbf{p}_i\|^{2k-2\ell}}{\|\tilde{\mathbf{p}}_t\|^{2k-1}} (\mathbf{p}_i^T \tilde{\mathbf{p}}_t)^\ell, \quad (56)$$

where $a_{k,\ell} = (-2)^\ell \binom{1/2}{k} \binom{k}{\ell}$. The series in (56) converges as long as $\|\mathbf{p}_i\| < (\sqrt{2} - 1) \|\tilde{\mathbf{p}}_t\|$,

¹The function $\sqrt{1+z}$ is real analytic on $(-1, +\infty)$ and $\frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2} - 2\frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|^2} > -1$, since it is assumed the transmitter and receiver do not coincide with a SAR image scene position. This result is more evident from the inequality $\frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2} - 2\frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|^2} \geq \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2} - 2\frac{\|\mathbf{p}_i\|}{\|\tilde{\mathbf{p}}_t\|} = x^2 - 2x \geq -1$, with the quadratic reaching a minimum of -1 when $x = 1$ (i.e., $\|\mathbf{p}_i\| = \|\tilde{\mathbf{p}}_t\|$).

so that $|z| = |x + y| = \left| \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2} - 2 \frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|^2} \right| < 1$ is satisfied².

Recognizing $\mathbf{p}_i^T \tilde{\mathbf{p}}_t = \mathbf{p}_i^T \mathbf{p}_t + \mathbf{p}_i^T \boldsymbol{\delta} \mathbf{p}_t$ and applying the Binomial theorem to $(\mathbf{p}_i^T \tilde{\mathbf{p}}_t)^\ell$, (56) may be expressed as

$$\begin{aligned} \|\tilde{\mathbf{r}}_{t,d,i}\| = & -\frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|} + \frac{1}{2} \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|} + \sum_{k=2}^{+\infty} \sum_{\ell=0}^k a_{k,\ell,0} \frac{\|\mathbf{p}_i\|^{2k-2\ell}}{\|\tilde{\mathbf{p}}_t\|^{2k-1}} (\mathbf{p}_i^T \mathbf{p}_t)^\ell \\ & + \sum_{k=2}^{+\infty} \sum_{\ell=1}^k \sum_{m=1}^{\ell} a_{k,\ell,m} \frac{\|\mathbf{p}_i\|^{2k-2\ell}}{\|\tilde{\mathbf{p}}_t\|^{2k-1}} \frac{(\mathbf{p}_i^T \boldsymbol{\delta} \mathbf{p}_t)^m}{(\mathbf{p}_i^T \mathbf{p}_t)^{m-\ell}}, \end{aligned} \quad (57)$$

where $a_{k,\ell,m} = \binom{\ell}{m} a_{k,\ell}$.

In the case no measurement errors are present $\boldsymbol{\delta} \mathbf{p}_t = \mathbf{0}$ and $\tilde{\mathbf{p}}_t$ is replaced with $\mathbf{p}_t$ in (57), producing

$$\|\mathbf{r}_{t,d,i}\| = -\frac{\mathbf{p}_i^T \mathbf{p}_t}{\|\mathbf{p}_t\|} + \frac{1}{2} \frac{\|\mathbf{p}_i\|^2}{\|\mathbf{p}_t\|} + \sum_{k=2}^{+\infty} \sum_{\ell=0}^k a_{k,\ell,0} \frac{\|\mathbf{p}_i\|^{2k-2\ell}}{\|\mathbf{p}_t\|^{2k-1}} (\mathbf{p}_i^T \mathbf{p}_t)^\ell. \quad (58)$$

Subtracting (58) from (57) and assuming $\|\boldsymbol{\delta} \mathbf{p}_t\| \ll \|\mathbf{p}_t\|$ (i.e., $\|\tilde{\mathbf{p}}_t\| \approx \|\mathbf{p}_t\|$ and $\mathbf{p}_i^T \mathbf{p}_t \approx \mathbf{p}_i^T \tilde{\mathbf{p}}_t$), the differential bistatic range error from the transmitter-to- $\mathbf{p}_i$ may be approximated as

$$\|\tilde{\mathbf{r}}_{t,d,i}\| - \|\mathbf{r}_{t,d,i}\| \approx -\frac{\mathbf{p}_i^T \boldsymbol{\delta} \mathbf{p}_t}{\|\tilde{\mathbf{p}}_t\|} + \sum_{k=2}^{+\infty} \sum_{\ell=1}^k \sum_{m=1}^{\ell} a_{k,\ell,m} \frac{\|\mathbf{p}_i\|^{2k-2\ell}}{\|\tilde{\mathbf{p}}_t\|^{2k-1}} \frac{(\mathbf{p}_i^T \boldsymbol{\delta} \mathbf{p}_t)^m}{(\mathbf{p}_i^T \tilde{\mathbf{p}}_t)^{m-\ell}}. \quad (59)$$

Applying the Cauchy-Schwarz inequality to $\mathbf{p}_i^T \boldsymbol{\delta} \mathbf{p}_t$, the approximate differential bistatic

²The presented bound is more evident from the inequality $\left| \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2} - 2 \frac{\mathbf{p}_i^T \tilde{\mathbf{p}}_t}{\|\tilde{\mathbf{p}}_t\|^2} \right| \leq \frac{\|\mathbf{p}_i\|^2}{\|\tilde{\mathbf{p}}_t\|^2} + 2 \frac{\|\mathbf{p}_i\|}{\|\tilde{\mathbf{p}}_t\|} = x^2 + 2x$, with the quadratic on the right hand side being less than one on the interval $x = (-1 - \sqrt{2}, -1 + \sqrt{2})$.

range error from (59) is upper bounded by

$$\left| \|\tilde{\mathbf{r}}_{t,d,i}\| - \|\mathbf{r}_{t,d,i}\| \right| \lesssim \frac{\|\mathbf{p}_i\| \|\delta \mathbf{p}_t\|}{\|\tilde{\mathbf{p}}_t\|} + \sum_{k=2}^{+\infty} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\mathbf{p}_i\|^{2k-2\ell+m} \|\delta \mathbf{p}_t\|^m}{\|\tilde{\mathbf{p}}_t\|^{2k-1} |\mathbf{p}_i^T \tilde{\mathbf{p}}_t|^{m-\ell}}. \quad (60)$$

A similar bound follows for the differential range error from $\mathbf{p}_i$ -to-receiver

$$\left| \|\tilde{\mathbf{r}}_{r,d,i}\| - \|\mathbf{r}_{r,d,i}\| \right| \lesssim \frac{\|\mathbf{p}_i\| \|\delta \mathbf{p}_r\|}{\|\tilde{\mathbf{p}}_r\|} + \sum_{k=2}^{+\infty} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\mathbf{p}_i\|^{2k-2\ell+m} \|\delta \mathbf{p}_r\|^m}{\|\tilde{\mathbf{p}}_r\|^{2k-1} |\mathbf{p}_i^T \tilde{\mathbf{p}}_r|^{m-\ell}}. \quad (61)$$

Combining the results from (60) and (61), an inequality for the spatially variant phase error component from (49) is obtained

$$\begin{aligned} \beta \frac{\pi}{2} &\geq \frac{2\pi f_o}{c} \left(\frac{\|\mathbf{p}_i\| \|\delta \mathbf{p}_t\|}{\|\tilde{\mathbf{p}}_t\|} + \frac{\|\mathbf{p}_i\| \|\delta \mathbf{p}_r\|}{\|\tilde{\mathbf{p}}_r\|} \right. \\ &\quad + \sum_{k=2}^{+\infty} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\mathbf{p}_i\|^{2k-2\ell+m} \|\delta \mathbf{p}_t\|^m}{\|\tilde{\mathbf{p}}_t\|^{2k-1} |\mathbf{p}_i^T \tilde{\mathbf{p}}_t|^{m-\ell}} \\ &\quad \left. + \sum_{k=2}^{+\infty} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\mathbf{p}_i\|^{2k-2\ell+m} \|\delta \mathbf{p}_r\|^m}{\|\tilde{\mathbf{p}}_r\|^{2k-1} |\mathbf{p}_i^T \tilde{\mathbf{p}}_r|^{m-\ell}} \right) \\ &\gtrsim 2\pi f_o |\delta \tau_{d,i}|, \end{aligned} \quad (62)$$

where $\beta \in [0, 1]$ is a proportionality constant set by the system designer for specifying the allowable amount of change in phase for the spatially variant phase error term to be considered spatially invariant. The conditions for when the inequality in (62) holds are as follows:

1. $\frac{1}{\sqrt{2}-1} \|\mathbf{p}_i\| < \|\tilde{\mathbf{p}}_t\|, \|\tilde{\mathbf{p}}_r\|,$
2. $\|\delta \mathbf{p}_t\| \ll \|\mathbf{p}_t\|,$
3. $\|\delta \mathbf{p}_r\| \ll \|\mathbf{p}_r\|.$

Assuming known upper bounds exist on the position uncertainty, which is typical, all

quantities in (62) are known, with $\|\mathbf{p}_i\|$ representing a free parameter which specifies the SAR scene size limit. Thus, the SAR scene size limit, $\|\mathbf{p}_i\|$, for spatially invariant defocus may be computed numerically using a root finding algorithm. To make this point more evident, let $\mathbf{p}_i = r_i \hat{\mathbf{p}}_i$, where r_i is the magnitude of $\mathbf{p}_i$ and $\hat{\mathbf{p}}_i$ is a unit vector to $\mathbf{p}_i$. Then, using equality in (62), one obtains

$$\begin{aligned}
0 = & -\beta \frac{c}{4f_o} + \left(\frac{\|\delta \mathbf{p}_t\|}{\|\tilde{\mathbf{p}}_t\|} + \frac{\|\delta \mathbf{p}_r\|}{\|\tilde{\mathbf{p}}_r\|} \right) r_i \\
& + \sum_{k=2}^{k_{\max}} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\delta \mathbf{p}_t\|^m \left| \hat{\mathbf{p}}_i^T \hat{\mathbf{r}}_{t,0} \right|^{\ell-m}}{\|\tilde{\mathbf{p}}_t\|^{2k-\ell-1+m}} r_i^{2k-\ell} \\
& + \sum_{k=2}^{k_{\max}} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\delta \mathbf{p}_r\|^m \left| \hat{\mathbf{p}}_i^T \hat{\mathbf{r}}_{0,r} \right|^{\ell-m}}{\|\tilde{\mathbf{p}}_r\|^{2k-\ell-1+m}} r_i^{2k-\ell}. \tag{63}
\end{aligned}$$

For chosen/measured values of β , $k_{\max}$, $\|\delta \mathbf{p}_t\|$, $\|\delta \mathbf{p}_r\|$, $\|\tilde{\mathbf{p}}_t\|$ and $\|\tilde{\mathbf{p}}_r\|$, $2k_{\max} - 1$ roots may be found corresponding to the solutions of (63). Of these roots, at least one is guaranteed to be purely real since the degree of the real polynomial in (63) is always odd [72]. Thus, finding the largest purely real root for (63) yields the maximum radial extent in the direction $\hat{\mathbf{p}}_i$ for the SAR scene such that the phase errors are spatially invariant. Repeating this process with a set of $\hat{\mathbf{p}}_i$ sampled across the unit circle provides a sampling of a bounding region or polygon over which the phase errors are approximately spatially invariant.

In the case a single radius is desired for specifying the SAR scene size limit (instead of a polygon), the following may be solved to obtain the radius to a ball which is

contained within the bounding region determined from (63)

$$\begin{aligned}
0 = & -\beta \frac{c}{4f_o} + \left(\frac{\|\delta \mathbf{p}_t\|}{\|\tilde{\mathbf{p}}_t\|} + \frac{\|\delta \mathbf{p}_r\|}{\|\tilde{\mathbf{p}}_r\|} \right) r_i \\
& + \sum_{k=2}^{k_{\max}} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\delta \mathbf{p}_t\|^m}{\|\tilde{\mathbf{p}}_t\|^{2k-\ell-1+m}} r_i^{2k-\ell} \\
& + \sum_{k=2}^{k_{\max}} \sum_{\ell=1}^k \sum_{m=1}^{\ell} |a_{k,\ell,m}| \frac{\|\delta \mathbf{p}_r\|^m}{\|\tilde{\mathbf{p}}_r\|^{2k-\ell-1+m}} r_i^{2k-\ell}.
\end{aligned} \tag{64}$$

In the next section example results are shown for near- and far-field imaging scenarios, where (63) and (64) are used to identify SAR scene sizes for which the phase errors are approximately spatially invariant.

3.2.3 Example Results

Near-Field Geometry Figure 9 depicts a near-field bistatic SAR imaging scenario with a center frequency of $f_o = 2.5$ GHz. For the given scenario, the transmitter and receiver position errors are chosen as realizations of Gaussian random variables

$$\delta \mathbf{p}_t, \delta \mathbf{p}_r \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}), \tag{65}$$

where $\sigma^2 = 2.5^2 \text{ cm}^2$ is the variance of the position errors.

Figure 10 shows the residual between the true differential range error and that computed using (56) and (57). As expected, as the number of terms considered in the series expansion is increased, the residual decreases (i.e., the series is converging to the true value).

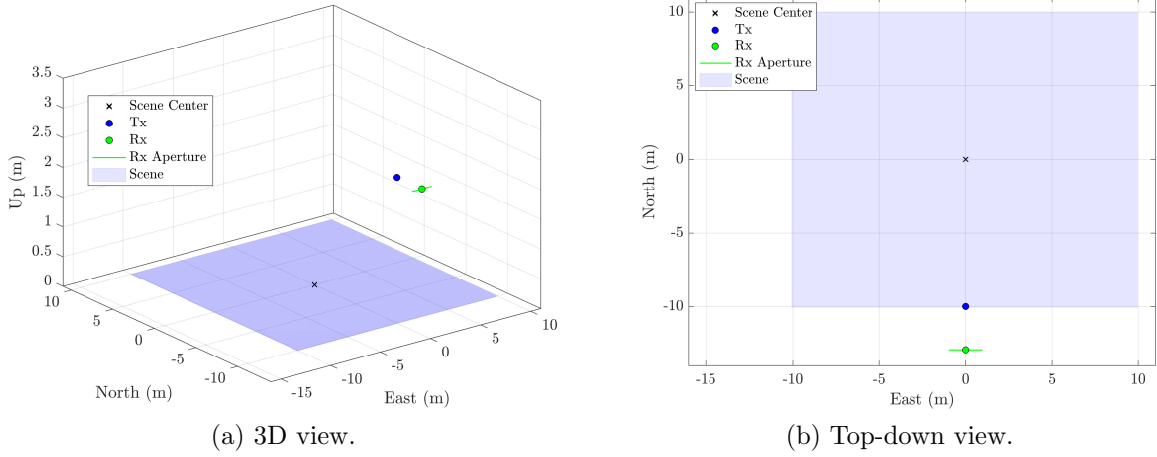


Figure 9. Near-field imaging scenario.

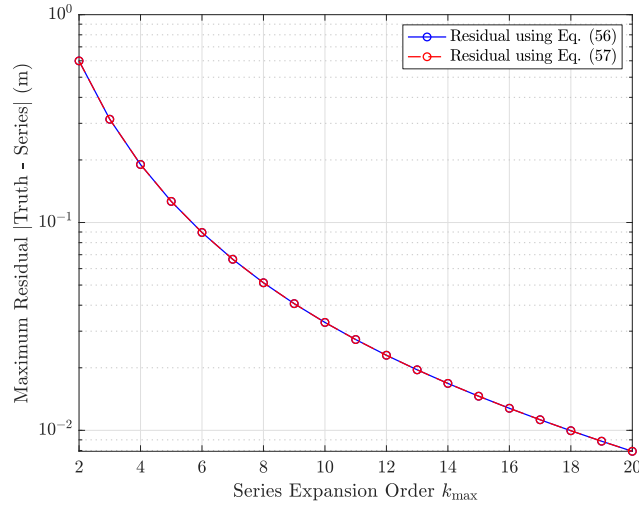


Figure 10. Convergence of (56) and (57).

Figure 11 depicts spatially invariant regions identified using (63) and (64), with $\beta = 1$, $k_{\max} = 4$, and $\|\delta \mathbf{p}_t\|, \|\delta \mathbf{p}_r\| = \frac{3^{3/2}}{2} \sigma$. The value used for $\|\delta \mathbf{p}_t\|$ and $\|\delta \mathbf{p}_r\|$ roughly corresponds to the 90% confidence interval of the Euclidean norm of the Gaussian distributed position errors³. In Figure 11, the radius of the spatially invariant ball (the red contour) is ≈ 2 m. The blue contours in Figure 11 show the maximum

³The upper bound assumed for the Euclidean norm of the position uncertainty may be relaxed or tightened based on the willingness of the system designer to accept SAR scene sizes, which may include scene positions that do not have spatially invariant phase errors according to the phase threshold specified in (62).

differential phase error at each point across the synthetic aperture. Mathematically, this is

$$\max_{n \in \{0, \dots, N-1\}} 2\pi f_o |\delta\tau_{d,i,n}| \quad (66)$$

at each point $\mathbf{p}_i$. As expected, the predicted spatially invariant regions fall within the $\pi/2$ differential phase error contour. The predicted spatially invariant regions are not tight with the $\pi/2$ contour primarily due to the application of the Cauchy-Schwarz inequality from (59) to (60), which will rarely hold with equality given random position errors.

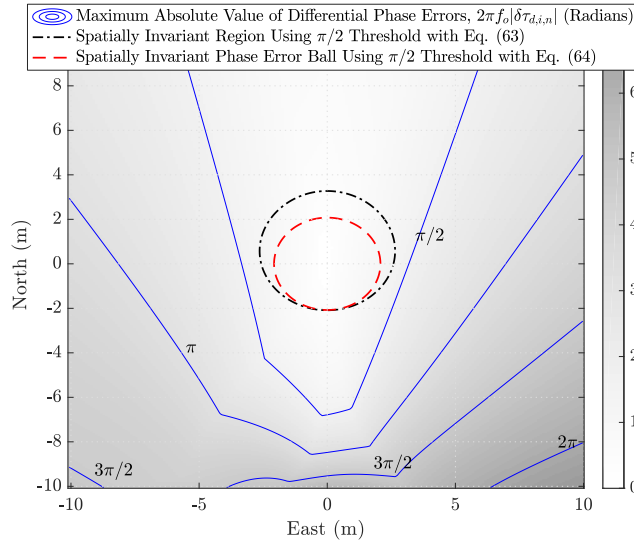


Figure 11. Spatially invariant regions identified using (63) and (64) with $\beta = 1$, $k_{\max} = 4$, and $\|\delta\mathbf{p}_t\|, \|\delta\mathbf{p}_r\| = \frac{3^{3/2}}{2}\sigma$.

Figure 12 depicts the minimum, maximum, and mean differential phase error for each point located within the identified spatially invariant region in Figure 11, for each pulse of the synthetic aperture. Mathematically, this is

$$\min_{i \in \mathcal{S}} / \text{mean} / \max 2\pi f_o \delta\tau_{d,i,n} \quad (67)$$

for each slow-time pulse across the synthetic aperture, where $\mathcal{S}$ is the set of points within the identified spatially invariant region in Figure 11. In addition, Figure 12 depicts the $\pi/2$ differential phase error bound, making it evident no points within the identified spatially invariant region exceed the specified threshold.

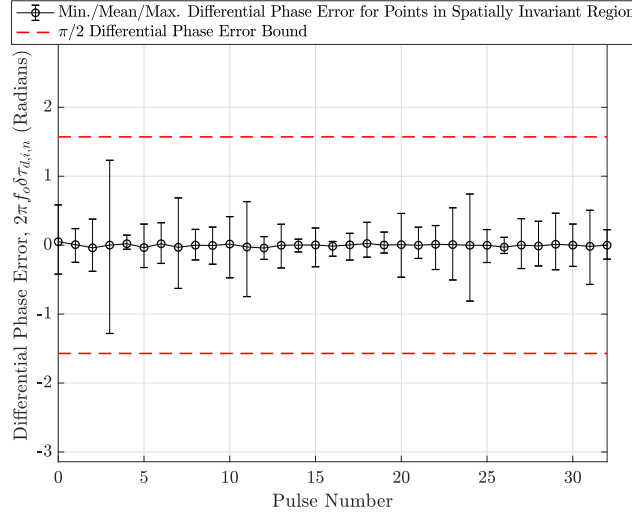


Figure 12. Minimum, maximum, and mean differential phase error for each point located within the identified spatially invariant region in Figure 11.

Far-Field Geometry Figure 13 depicts a far-field bistatic SAR imaging scenario with a center frequency of $f_o = 10$ GHz. Again, the transmitter and receiver position errors are chosen as realizations of Gaussian random variables

$$\delta \mathbf{p}_t, \delta \mathbf{p}_r \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}), \quad (68)$$

where $\sigma^2 = 0.15^2$ m² is the variance of the position errors.

Figure 14 shows the residual between the true differential range error and that computed using (56) and (57). Again, as expected, as the number of terms considered in the series expansion is increased, the residual decreases (i.e., the series is converging

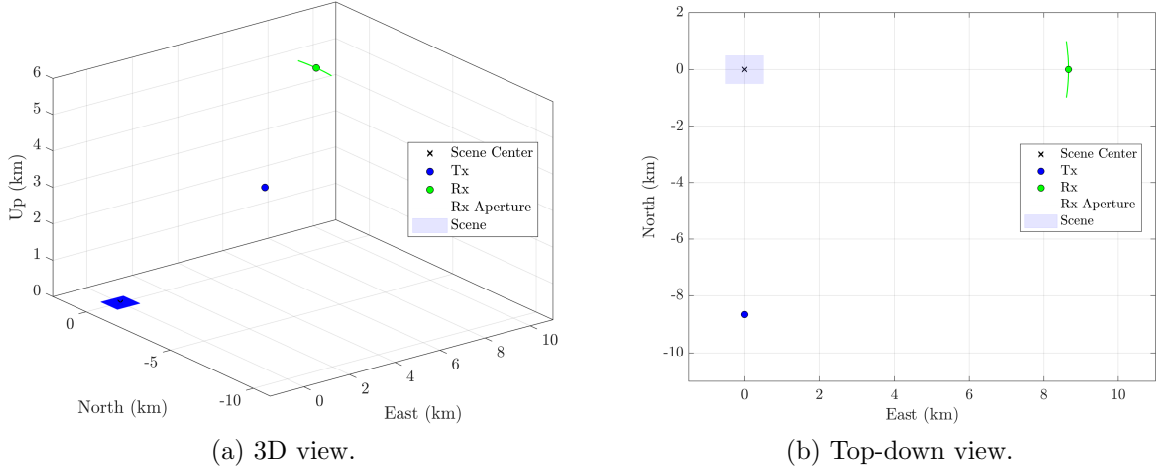


Figure 13. Far-field imaging scenario.

to the true value).

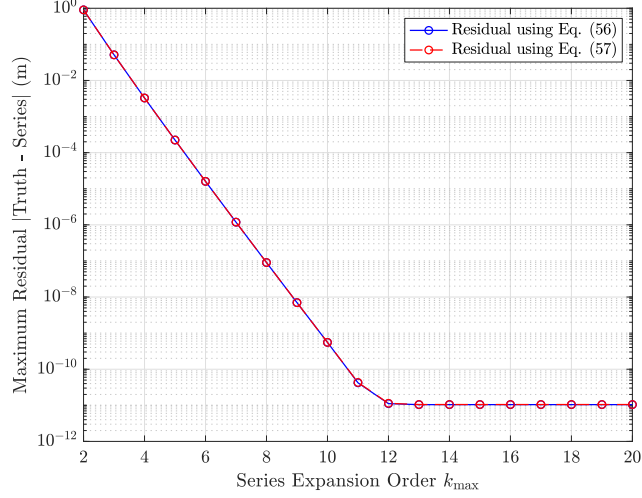


Figure 14. Convergence of (56) and (57).

Figure 15 depicts spatially invariant regions identified using (63) and (64), with $\beta = 1$, $k_{\max} = 4$, and $\|\delta \mathbf{p}_t\|, \|\delta \mathbf{p}_r\| = \frac{3^{3/2}}{2}\sigma$. Again, the value used for $\|\delta \mathbf{p}_t\|$ and $\|\delta \mathbf{p}_r\|$ roughly corresponds to the 90% confidence interval of the Euclidean norm of the Gaussian distributed position errors³. In Figure 15, the radius of the spatially invariant ball (the red contour) is ≈ 95 m. The blue contours in Figure 15 show the

maximum differential phase error at each point across the synthetic aperture. Mathematically, this is

$$\max_{n \in \{0, \dots, N-1\}} 2\pi f_o |\delta\tau_{d,i,n}| \quad (69)$$

at each point $\mathbf{p}_i$. Again, as expected, the predicted spatially invariant regions fall within the $\pi/2$ differential phase error contour.

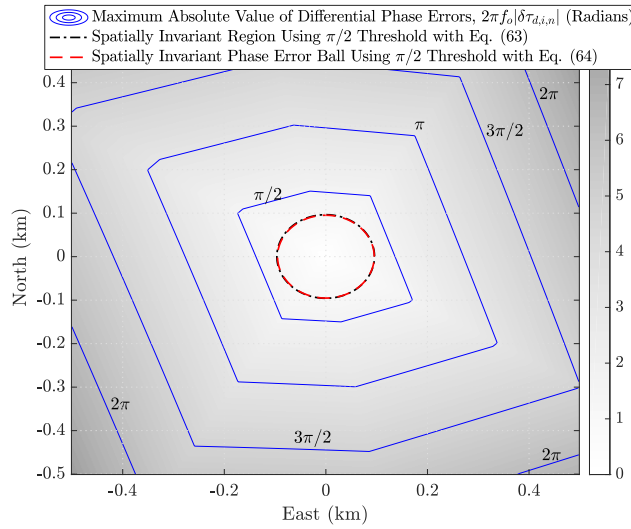


Figure 15. Spatially invariant regions identified using (63) and (64) with $\beta = 1$, $k_{\max} = 4$, and $\|\delta\mathbf{p}_t\|, \|\delta\mathbf{p}_r\| = \frac{3^{3/2}}{2}\sigma$.

Figure 16 depicts the minimum, maximum, and mean differential phase error for each point located within the identified spatially invariant region in Figure 15, for each pulse of the synthetic aperture. Mathematically, this is

$$\min / \text{mean} / \max_{i \in \mathcal{S}} 2\pi f_o \delta\tau_{d,i,n} \quad (70)$$

for each slow-time pulse across the synthetic aperture, where $\mathcal{S}$ is the set of points within the identified spatially invariant region in Figure 15. In addition, Figure 16

depicts the $\pi/2$ differential phase error bound, making it evident no points within the identified spatially invariant region exceed the specified threshold.

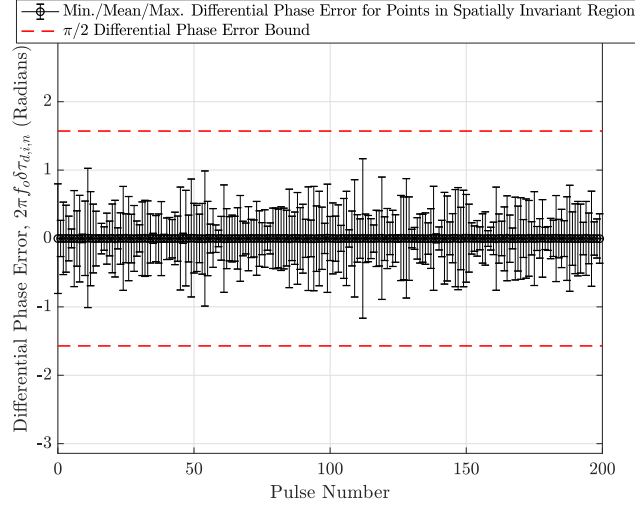


Figure 16. Minimum, maximum, and mean differential phase error for each point located within the identified spatially invariant region in Figure 15.

Synthetic Aperture Radar Imaging Results In this section, example autofocus SAR images formed using two different processing schemes are compared. With the first scheme, autofocus is performed on the full SAR image, ignoring the predicted spatially invariant phase error bound. In other words, a single phase error correction is estimated and applied for each point of the SAR image. With the second scheme, autofocus is performed on sub-images, where the sub-image scene size is determined by the predicted spatially invariant phase error bound. In other words, a separate phase error correction is estimated for each sub-image and applied for each point in the sub-image. Then, the full SAR image is formed by mosaicing together the individually autofocused sub-images. SAR images formed using these two schemes make evident the usefulness of the spatially invariant phase error bound derived and presented in Section 3.2.2.

For each of the two processing schemes the SAR image is formed using the BPA. In addition, the images are autofocused using the phase-only SDR autofocus algorithm presented in [60], where it is assumed that the true location of strong in-scene scatterers is known. Although the SDR autofocus algorithm is considered, any phase-only autofocus algorithm could be used to obtain comparable results to those presented herein. The collection geometry assumed for the two processing schemes consists of the near-field collection geometry described in Figure 9. Figure 17 depicts the collection geometry, this time with the SAR scene consisting of 25 stationary isotropic point scatterers spaced every 4 meters along the east and north dimensions. A 4 meter-by-4 meter sub-image scene size, which is approximately the spatially invariant phase error bound predicted in Figure 11, is considered for producing the mosaic SAR image.

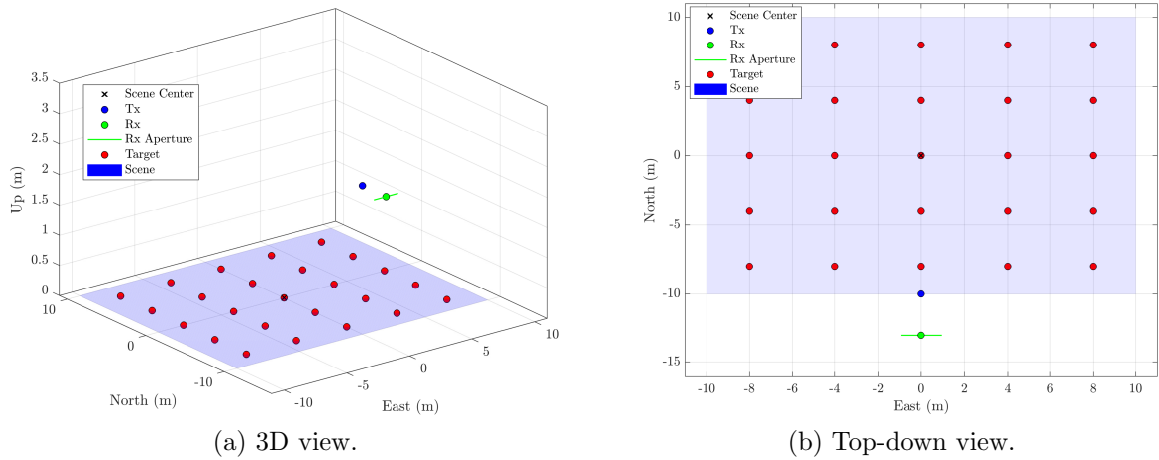


Figure 17. Near-field imaging scenario.

Figures 18a and 18b depict the ideal and defocused SAR images, respectively. Figure 18c depicts the autofocused image produced by performing autofocus on the full SAR image (i.e., by ignoring the predicted spatially invariant phase error bound). Figure 18d depicts the autofocused mosaic SAR image produced by performing auto-

focus on individual 4 meter-by-4 meter sub-images (i.e., by adhering to the predicted spatially invariant phase error bound).

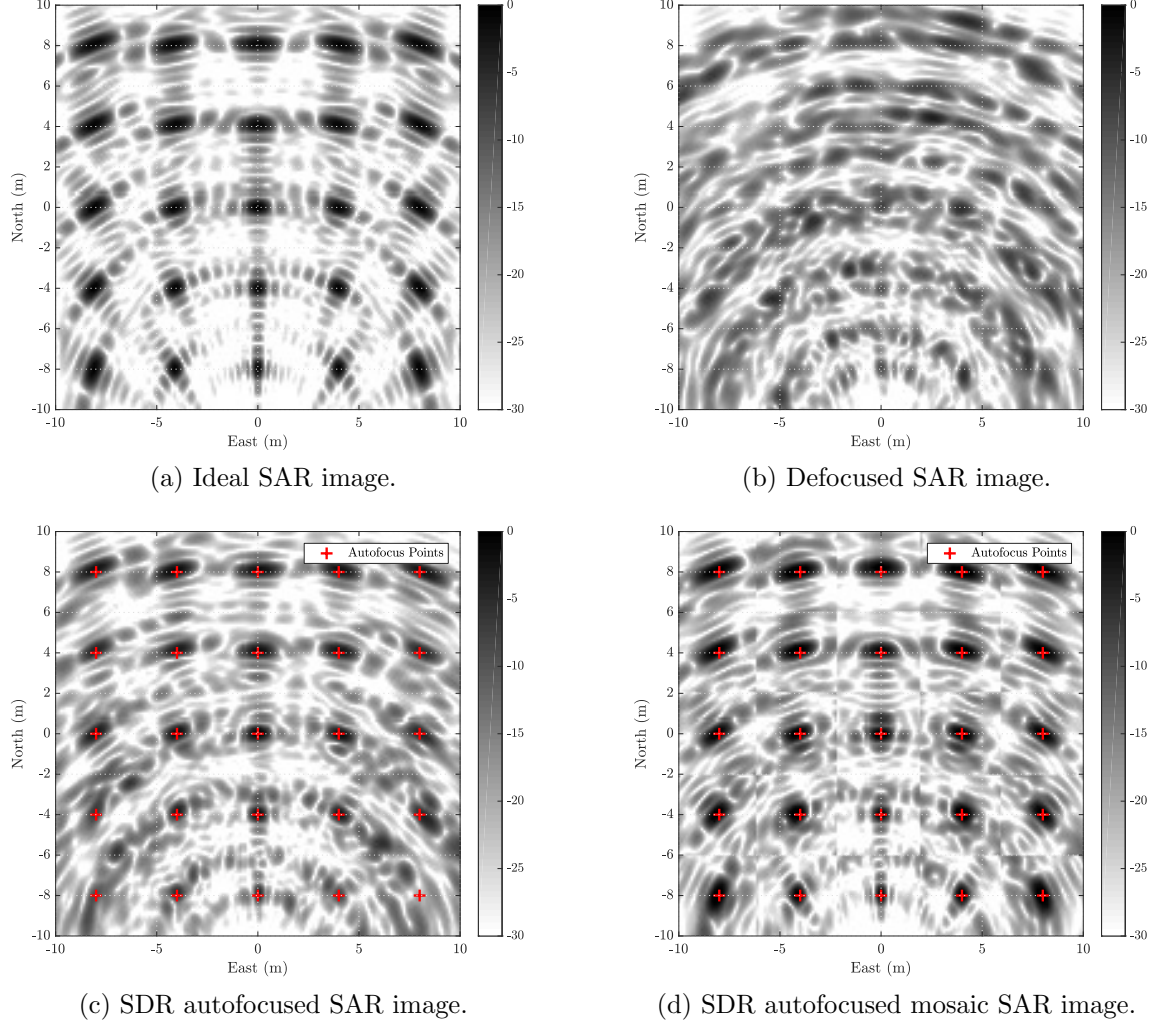


Figure 18. Comparison of phase-only autofocus SAR image quality as the spatially invariant phase error model breaks down.

As expected, the SAR image in Figure 18d, where individual 4 meter-by-4 meter sub-images are autofocused and mosaiced together, is better focused than the SAR image in Figure 18c, where a single autofocus solution is applied⁴. The added defocus

⁴Although the mosaic SAR image is better focused, it does contain discontinuities from separate autofocus solutions being applied across the image. Dealing with these discontinuities extends beyond the scope of this work.

observed in Figure 18c is a direct result of the spatially invariant phase error model breaking down. For instance, near the center of the SAR image in Figure 18c, the focus is comparable to the ideal image in Figure 18a. However, near the edges of the SAR image in Figure 18c, where the spatially invariant phase error model breaks down (see Figure 11), the scattering responses become less distinguishable. On the other hand, the defocus observed in the mosaic SAR image in Figure 18d is relatively constant throughout the SAR image. That is, the residual defocus is no worse near the edges of the SAR image as compared to the center of the SAR image. The consistency of the residual defocus observed in Figure 18d, as compared to that observed in Figure 18c, is a consequence of the spatially invariant phase error model not breaking down within each of the autofocused sub-images.

3.3 Conclusion

In this section, matched filter and image formation analysis was performed to identify bounds on motion measurement errors and SAR scene size limits such that the phase-only defocus model and spatially invariant phase error assumption are applicable. In the next section, the GPGA algorithm is developed and studied.

IV. Generalized Phase Gradient Autofocus Algorithm

In this section the GPGA algorithm is formulated. The formulation begins by first providing a side-by-side description of the PGA and GPGA algorithms, making evident the GPGA algorithm includes the PGA algorithm as a special case. Next, it is shown that the GPGA algorithm is applicable with the BPA, while preserving the four crucial signal processing steps comprising the PGA algorithm. Lastly, it is shown that with the generalized formulation, whether using the PFA or BPA, the GPGA algorithm provides an approximate MMLE of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers.

4.1 Using the Polar Format Algorithm

For all remaining discussion, unless explicitly stated, the PGA algorithm refers to the algorithm depicted in Figure 19, which consists of processing steps described in [19]. Before discussing each of these processing steps, it is first important to understand the assumed relationship between the range profile or range compressed data and the SAR image.

In [19], the PGA algorithm is formulated using the PFA, meaning the range profile data is related to the SAR image by a fast Fourier transform (FFT) across slow-time. As alluded to in Section 2.1 with (25), this implies $\tilde{\phi}_i$ is a Fourier basis vector, and ξ_i is a single row of range profile samples from the polar-to-Cartesian resampled data. As will be seen, understanding this relationship is important as well-known Fourier identities are used to develop the GPGA algorithm.

The PGA algorithm depicted in Figure 19 begins by first forming a PFA image.

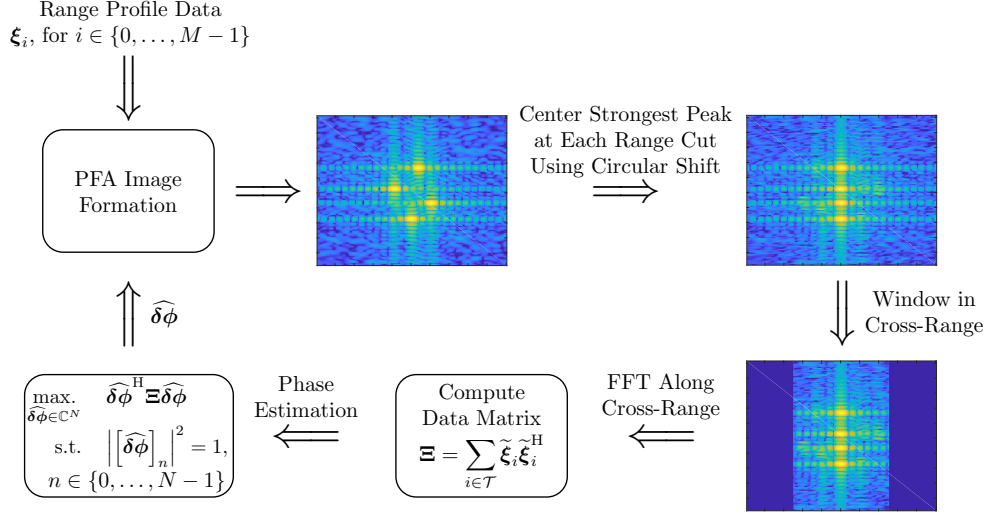


Figure 19. PGA algorithm [19].

Next, the strongest peaks from a subset of range cuts from the formed PFA image are centered in cross-range using a circular shift. The set of indices corresponding to the identified strongest peaks (i.e., selected scatterers) is denoted by $\mathcal{T}$. Upon selecting and centering scatterers, a window is applied along the cross-range dimension for each selected range cut. Then, an FFT is computed along cross-range to move from the image domain to the range profile domain. After moving to the range profile domain, what is referred to as the data matrix is computed as a sum of outer products of the circularly shifted and windowed range cuts. Lastly, phase estimation is performed by solving or finding an approximate solution to the NP-hard phase estimation problem depicted in Figure 19.

Figure 20 depicts what is termed the GPGA algorithm. Using Fourier identities, the GPGA algorithm performs each of the signal processing steps of the PGA algorithm directly in the range profile domain. For instance, centering the strongest peak in a single range cut using a circular shift in the image domain is equivalent to element-wise multiplying the range profile samples by the phase vector applied to produce the

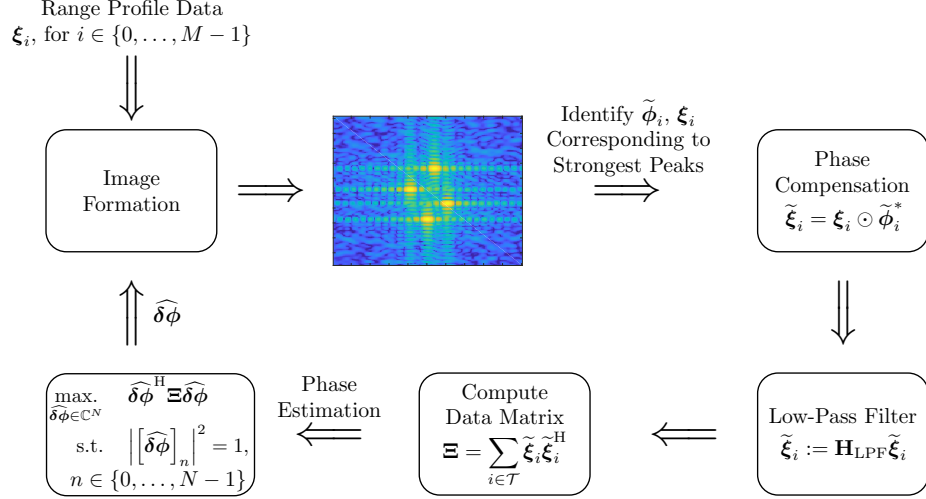


Figure 20. GPGA algorithm.

strongest peak. Mathematically, this equivalence is

$$\begin{array}{c}
 \text{Center } \zeta(\mathbf{p}_i) \\
 \text{Using Circular Shift}
 \end{array}
 \iff \tilde{\boldsymbol{\xi}}_i = \boldsymbol{\xi}_i \odot \tilde{\boldsymbol{\phi}}_i^*. \quad (71)$$

Applying a window along the cross-range dimension in the image domain is equivalent to applying a low-pass filter (LPF) to phase compensated range profile samples. Mathematically, this equivalence is

$$\begin{array}{c}
 \text{Window in} \\
 \text{Cross-Range}
 \end{array}
 \iff \tilde{\boldsymbol{\xi}}_i := \mathbf{H}_{\text{LPF}} \tilde{\boldsymbol{\xi}}_i, \quad (72)$$

where $\mathbf{H}_{\text{LPF}} \in \mathbb{C}^{N \times N}$ is an LPF matrix. The remaining steps of the PGA algorithm, which include computing the data matrix and performing phase estimation, are performed in the range profile domain and are identical for the GPGA algorithm. Thus, it is seen that each of the signal processing steps comprising the PGA algorithm may be performed directly in the range profile domain by way of the GPGA algorithm depicted in Figure 20.

At this point in the discussion, the GPGA algorithm has been described as being applied with the PFA to enable a side-by-side comparison to be made with the PGA algorithm. In the next section, the GPGA algorithm is described as being applied with the BPA.

4.2 Using the Backprojection Algorithm

As depicted in Figure 20, the GPGA algorithm using the BPA begins by first forming a BPA image. Next, the range profile samples and phase vector corresponding to the strongest peaks in the formed BPA image are identified. These quantities are ξ_i and $\tilde{\phi}_i$ for $i \in \mathcal{T}$, where again, $\mathcal{T}$ is the set of indices corresponding to the strongest peaks (i.e., selected scatterers) in the formed BPA image. For the BPA, the range profile samples ξ_i need not correspond to a single range cut (see Figure 5b), and the phase vector need not correspond to a Fourier basis vector (recall, the BPA phase vector is $[\tilde{\phi}_i]_n = e^{-j\omega_o(\tilde{\tau}_{i,n} - \tilde{\tau}_{0,n})}$). However, the centering step is still completed through phase compensation, where an element-wise multiply $\tilde{\xi}_i = \xi_i \odot \tilde{\phi}_i^*$ is computed. Upon performing phase compensation, the windowing step is performed by applying a LPF to the phase compensated range profile samples, as shown in the right hand side of (72). Next, what is referred to as the data matrix is computed as a sum of outer products of the phase compensated and low-pass filtered range profile samples $\tilde{\xi}_i$. Lastly, phase estimation is performed by solving or finding an approximate solution to the NP-hard phase estimation problem depicted in Figure 20.

Before discussing strategies for applying the GPGA algorithm, it is first important to point out some of the intuition behind the centering and windowing steps of the GPGA algorithm when used with the BPA (this same intuition for the centering and windowing steps of the PGA algorithm when used with the PFA is discussed in [3,20].)

First, phase compensation is performed to effectively center the point spread function of each selected scatterer at the scene center of the SAR image (i.e., at the spatial location $\mathbf{p}_0$ where the differential bistatic delay is zero). This centering of the point spread function is better evidenced by expanding (21)

$$\begin{aligned}\zeta(\mathbf{p}_i) &= \mathbf{1}^H \tilde{\boldsymbol{\xi}}_i \\ &= \sum_{n=0}^{N-1} \left[\tilde{\boldsymbol{\xi}}_i \right]_n e^{j\omega_o(\tilde{\tau}_{0,n} - \tilde{\tau}_{0,0})} \\ &= \tilde{\boldsymbol{\phi}}_0^H \tilde{\boldsymbol{\xi}}_i,\end{aligned}\tag{73}$$

which closely resembles the expression for the SAR image at the scene center: $\zeta(\mathbf{p}_0) = \tilde{\boldsymbol{\phi}}_0^H \tilde{\boldsymbol{\xi}}_0 = \mathbf{1}^H \tilde{\boldsymbol{\xi}}_0$. Upon centering/aligning the point spread function of each selected scatterer, windowing is performed. Windowing is performed to try to capture the majority of the blur or defocus of each point spread function, while filtering energy outside of the main defocus region to try to improve the effective SINR of each selected scatterer [3, 20].

From [3, 20] and the discussion points above, one may see that the GPGA algorithm when used with the BPA shares the same intuitions as the PGA and GPGA algorithms when used with the PFA. Albeit, the intuition for the steps performed with the BPA is not as readily seen as when the PFA is used, where well-known Fourier identities may be leveraged.

In Sections 4.3 to 4.5, strategies for performing three of the four crucial signal processing steps of the GPGA algorithm are discussed. Discussion on arguably the most important step of the GPGA algorithm, phase estimation, is withheld from this discussion as the entirety of Chapter V is dedicated to describing and discussing trade-offs between different GPGA phase estimators.

4.3 Scatterer Selection

As will be discussed and shown in Section 4.7, with the GPGA algorithm there is statistical motivation for selecting scatterers with either moderate or nearly constant SINR. Since these scattering locations are not generally known *a priori*, a strategy must be devised for identifying these scattering locations from formed (and potentially defocused) SAR images.

Arguably, the most straightforward approach for scatterer selection is to apply a relative threshold to the formed SAR image. That is, indices of image points with an image intensity within a specified threshold of the maximum image intensity are included in $\mathcal{T}$. The intuition behind this approach is that image points with higher intensity potentially have higher SINR.

A potentially more robust and complicated scheme may consist of using strategies from constant false alarm rate (CFAR) detection. For instance, the 2D CFAR windowing/filtering process could be applied to the formed SAR image to estimate the SINR at each point in the image. Then, image points with an estimated SINR exceeding a specified SINR threshold could be included in $\mathcal{T}$. For more details regarding the many 2D CFAR windowing/filtering strategies available, the reader is referred to [73–76].

4.4 Low-Pass Filter Design

For GPGA LPF design two strategies commonly employed with the PGA algorithm are proposed [3]. These strategies include: 1) setting the low-pass width to be the estimated blur width of the phase errors and 2) employing an LPF with a constant

shrinking low-pass width over GPGA iterations.

For estimating the blur width of the phase errors, a non-coherent sum of the spectrum of each set of phase compensated range profile samples is first computed as

$$\mathbf{x} = \sum_{i \in \mathcal{T}} \left(\mathbf{F} \tilde{\boldsymbol{\xi}}_i \right) \odot \left(\mathbf{F} \tilde{\boldsymbol{\xi}}_i \right)^* . \quad (74)$$

Substituting in (25) for $\tilde{\boldsymbol{\xi}}_i$, and assuming nearly constant reflectivity, high SINR scatterers are selected during the scatterer selection process (i.e., $\boldsymbol{\gamma}_i \approx \gamma_i \mathbf{1}$ and $\mathbf{w}_i$ is negligible), $\mathbf{x}$ from (74) is seen to be an estimate of the magnitude of the spectrum of the phase errors

$$\begin{aligned} \mathbf{x} &= \sum_{i \in \mathcal{T}} \left(\mathbf{F}(\text{Diag}(\boldsymbol{\delta}\boldsymbol{\phi}) \boldsymbol{\gamma}_i + \mathbf{w}_i) \right) \odot \left(\mathbf{F}(\text{Diag}(\boldsymbol{\delta}\boldsymbol{\phi}) \boldsymbol{\gamma}_i + \mathbf{w}_i) \right)^* \\ &\approx \sum_{i \in \mathcal{T}} |\gamma_i|^2 (\mathbf{F} \boldsymbol{\delta}\boldsymbol{\phi}) \odot (\mathbf{F} \boldsymbol{\delta}\boldsymbol{\phi})^* . \end{aligned} \quad (75)$$

Thus, from $\mathbf{x}$ the blur or low-pass width of the window in cross-range, let it be denoted by W in terms of the number of samples, may be estimated by thresholding $\mathbf{x}$ at a chosen level below the maximum value of $\mathbf{x}$ (similar to the approach suggested in [3]). Once the width has been estimated, the LPF matrix used in (72) may be computed by

$$\mathbf{H}_{\text{LPF}} = \mathbf{F}^H \text{Diag}(\mathbf{h}) \mathbf{F}, \quad (76)$$

where

$$[\mathbf{h}]_n = \begin{cases} 1, & |[n - N/2]| \leq \frac{W}{2}, \\ 0, & \text{else,} \end{cases} \quad (77)$$

for $n \in \{0, \dots, N-1\}$.

The second and simpler GPGA LPF design strategy is to employ an LPF with a constant shrinking low-pass width over GPGA iterations. The intuition here is that after each iteration the image becomes better focused, and the residual phase errors approach a constant. Hence, the spectrum of the residual phase errors has less blur width (i.e., as the residual phase errors approach a constant, the spectrum approaches a delta function with no blur width). For the constant shrinking low-pass width approach, the LPF matrix may be computed using (76) with $\mathbf{h}$ defined as

$$[\mathbf{h}]_n = \begin{cases} 1, & |[n - N/2]| \leq \alpha^k \frac{N}{2}, \\ 0, & \text{else,} \end{cases} \quad (78)$$

where $\alpha \in (0, 1]$ is the constant shrink factor and $k \in \{0, 1, \dots\}$ is the GPGA iteration number.

4.5 Iterations

As discussed in [19, 20], a crucial step of the PGA algorithm is iterating the algorithm. The intuition for this step is that with each iteration the SAR image becomes better focused enabling the strongest scattering locations to more easily be identified, and thereby leads to improved phase estimation. As depicted in Figure 20, the GPGA algorithm is designed to preserve the crucial step of iterating.

In [3, 20] strategies or stopping criteria are discussed for determining an appropriate number of PGA algorithm iterations. These strategies apply equally well with the GPGA algorithm. The simplest of these strategies, and the strategy used for results presented in Chapter VI, is running the GPGA algorithm for a fixed number of

iterations. Similar to findings in [19] with the PGA algorithm, the authors have found the GPGA algorithm provides adequate performance with a number of data sets using no more than three GPGA iterations. However, the system designer is free to change the iteration strategy or stopping criteria as seen fit, so both a robust and computationally efficient implementation of the GPGA algorithm may be realized.

4.6 Summary

At first glance, the GPGA algorithm may only seem useful for applying the PGA algorithm with the BPA. However, the generalization also has an implication which improves upon the PGA algorithm from [3, 16, 17, 19, 20] using the PFA. For instance, with the PGA algorithm there is a known limitation in that only one strong peak per range cut may be centered and used for phase error estimation [14, 45, 55]. This limitation is alleviated using the generalized implementation shown in Figure 20, as the phase compensation and low-pass filtering steps may be performed on any arbitrary set of image points $\mathcal{T}$.

In the next section, the GPGA algorithm, whether using the PFA or BPA, is shown to provide an approximate MMLE of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers.

4.7 Maximum Marginal Likelihood Estimate

Let the variables in (26) to (28) be distributed as

$$\omega_o \delta \tau_{i,n} \sim \mathcal{U} [-\pi, \pi], \quad (79)$$

$$\boldsymbol{\gamma}_i \sim \mathcal{CN} (\mathbf{0}, \sigma_{\gamma,i}^2 \mathbf{1}\mathbf{1}^H + \sigma_{v,i}^2 \mathbf{I}), \quad (80)$$

$$\mathbf{w}_i \sim \mathcal{CN} (\mathbf{0}, \sigma_w^2 \mathbf{I}), \quad (81)$$

where $\sigma_{\gamma,i}^2 \in \mathbb{R}$ is the expected matched filtered power per pulse for the i th scatterer and $\sigma_{v,i}^2, \sigma_w^2 \in \mathbb{R}$ are the expected matched filtered interference and noise powers per pulse at the i th scattering location, respectively. In (79) to (81) each of the variables are assumed to be independent with respect to one another. For the statistical model described by (79) to (81), the following is observed:

1. The complex-valued reflectivities for the i th scatterer, $\gamma_{i,n}$, are expected to be perfectly correlated from pulse-to-pulse (i.e., perfectly correlated across n).
2. Both the matched filtered interference and noise powers, $\sigma_{v,i}^2$ and σ_w^2 , are constant from pulse-to-pulse (i.e., constant across n).

From the observations stated above a physical implication follows. In particular, for an anisotropic scattering model a small enough receive aperture, in terms of angular extent, must be considered for performing autofocus such that the scatterer's complex-valued reflectivities and the interference power are expected to be constant from pulse-to-pulse. This physical implication does not represent a limitation or shortcoming of any autofocus algorithm leveraging the above statistical model. Rather, this physical implication dictates how an autofocus algorithm leveraging the above statistical model must be applied. For instance, if a wide aperture collection geometry is considered,

then the wide aperture data should be broken up into smaller sub-apertures (a step commonly done in practice) for performing autofocus so that the statistical model from (79) to (81) applies.

The SINR, given the above statistical model, for the i th scattering location is

$$\text{SINR}_i = \frac{\sigma_{\gamma,i}^2}{\sigma_{v,i}^2 + \sigma_w^2}. \quad (82)$$

Without loss of generality, let the interference-plus-noise term be one (i.e., $\sigma_{v,i}^2 + \sigma_w^2 = 1$), so that SINR is parameterized solely by the expected scatterer's matched filtered power $\sigma_{\gamma,i}^2$. Using this assumption and the derivation from Appendix A, the marginal likelihood for the phase compensated range profile samples from (25), having marginalized over the unknown complex-valued reflectivities γ_i , is

$$\tilde{\xi}_i | \boldsymbol{\delta\phi}, \boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\gamma,i}^2 \boldsymbol{\delta\phi} \boldsymbol{\delta\phi}^H + \mathbf{I}), \quad (83)$$

where $\boldsymbol{\eta} \in \mathbb{R}^P$ is a vector containing all $\sigma_{\gamma,i}^2$ for $i \in \mathcal{T}$ (i.e., vector containing all hyperparameters) and P is the number of scatterers selected during the scatterer selection process (i.e., the cardinality of $\mathcal{T}$).

Using the matrix determinant lemma and Woodbury matrix identity, the determinant and inverse of the covariance matrix from (83) are

$$|\sigma_{\gamma,i}^2 \boldsymbol{\delta\phi} \boldsymbol{\delta\phi}^H + \mathbf{I}| = 1 + \sigma_{\gamma,i}^2 N \quad (84)$$

and

$$(\sigma_{\gamma,i}^2 \boldsymbol{\delta\phi} \boldsymbol{\delta\phi}^H + \mathbf{I})^{-1} = \mathbf{I} - \frac{\boldsymbol{\delta\phi} \boldsymbol{\delta\phi}^H}{\frac{1}{\sigma_{\gamma,i}^2} + N}. \quad (85)$$

Substituting (84) and (85) in the log marginal likelihood function for (83) produces

$$\begin{aligned}\mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta}) &\propto - \sum_{i \in \mathcal{T}} \ln(1 + \sigma_{\gamma,i}^2 N) + \sum_{i \in \mathcal{T}} \frac{\left| \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} \right|^2}{\frac{1}{\sigma_{\gamma,i}^2} + N} \\ &\approx - \sum_{i \in \mathcal{T}} \ln(1 + \sigma_{\gamma,i}^2 N) + \alpha(\boldsymbol{\eta}) \sum_{i \in \mathcal{T}} \left| \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} \right|^2,\end{aligned}\quad (86)$$

where constants with respect to $\boldsymbol{\delta\phi}$ and $\boldsymbol{\eta}$ have been omitted and $\alpha : \mathbb{R}^P \rightarrow \mathbb{R}$, $\alpha(\boldsymbol{\eta}) > 0$ is an approximation constant. For the approximation in (86) to hold, it is assumed selected scatterers have moderate or nearly constant SINR (i.e., $N\sigma_{\gamma,i}^2 \gg 1$ or $\sigma_{\gamma,i}^2$ is nearly constant for all $i \in \mathcal{T}$). One of these two conditions will typically be satisfied using the GPGA algorithm. In fact, in many cases both conditions will be satisfied. For instance, the first step of the GPGA algorithm is the selection of strongest peaks from the formed SAR image (i.e., image points with potentially high SINR). Moreover, the system designer may enforce that the image intensities of selected scatterers lie within a given interval such that the image intensities are approximately constant (i.e., image points with potentially constant SINR).

From (86) the approximate log marginal likelihood is seen to be separable in terms of the phase errors and the hyperparameters

$$\mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta}) \approx \mathcal{L}_1(\boldsymbol{\eta}) + \alpha(\boldsymbol{\eta}) \mathcal{L}_2(\boldsymbol{\delta\phi}), \quad (87)$$

where $\mathcal{L}_1 : \mathbb{R}^P \rightarrow \mathbb{R}$ and $\mathcal{L}_2 : \mathbb{C}^N \rightarrow \mathbb{R}$ are defined as

$$\mathcal{L}_1(\boldsymbol{\eta}) = - \sum_{i \in \mathcal{T}} \ln(1 + \sigma_{\gamma,i}^2 N), \quad (88)$$

$$\mathcal{L}_2(\boldsymbol{\delta\phi}) = \sum_{i \in \mathcal{T}} \left| \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} \right|^2. \quad (89)$$

Thus, given selected scatterers with moderate or nearly constant SINR, an MMLE for the phase errors, $\widehat{\boldsymbol{\delta\phi}}^*$, is obtained by finding a global maximum point to

$$\begin{aligned} & \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} && \widehat{\boldsymbol{\delta\phi}}^H \boldsymbol{\Xi} \widehat{\boldsymbol{\delta\phi}}, \\ & \text{subject to} && \left| \left[\widehat{\boldsymbol{\delta\phi}} \right]_n \right|^2 = 1, \quad n \in \{0, \dots, N-1\}, \end{aligned} \quad (90)$$

where $\boldsymbol{\Xi} = \widetilde{\boldsymbol{\xi}} \widetilde{\boldsymbol{\xi}}^H \in \mathbb{H}^N$ is the GPGA data matrix (see Figure 20) and $\widetilde{\boldsymbol{\xi}} = \left\{ \widetilde{\boldsymbol{\xi}}_i \mid i \in \mathcal{T} \right\} \in \mathbb{C}^{N \times P}$ is a matrix consisting of column vectors of phase compensated range profile samples for each selected scatterer. The optimization problem in (90) is equivalent to the GPGA phase estimation problem stated in (32), recognizing

$$\begin{aligned} \widehat{\boldsymbol{\delta\phi}}^H \boldsymbol{\Xi} \widehat{\boldsymbol{\delta\phi}} &= \widehat{\boldsymbol{\delta\phi}}^H \widetilde{\boldsymbol{\xi}} \widetilde{\boldsymbol{\xi}}^H \widehat{\boldsymbol{\delta\phi}} \\ &= \widehat{\boldsymbol{\delta\phi}}^H \left(\sum_{i \in \mathcal{T}} \widetilde{\boldsymbol{\xi}}_i \widetilde{\boldsymbol{\xi}}_i^H \right) \widehat{\boldsymbol{\delta\phi}} \\ &= \sum_{i \in \mathcal{T}} \left| \widehat{\boldsymbol{\delta\phi}}^H \widetilde{\boldsymbol{\xi}}_i \right|^2. \end{aligned} \quad (91)$$

As mentioned in Section 2.2, the GPGA phase estimation problem is NP-hard. Thus, for SAR imaging applications, where the size of the GPGA data is typically large, it is not reasonable to expect to always solve (90) (i.e., find a global optimal point). Instead, heuristics are needed to obtain an approximate solution to (90) in a reasonably efficient manner. As previously mentioned, because heuristics are involved a number of phase estimators or approximate MMLEs exist, with each providing a trade-off in terms of performance/robustness, computational complexity, and suboptimality of the phase error estimate with respect to the solution set of MMLEs for (90). In the next section, three GPGA phase estimators, two of which are derived from existing PGA phase estimators, are described in detail and their trade-offs discussed. However, before continuing on to the next section, some additional comments on the

GPGA algorithm are made.

In light of the many comments made in the literature regarding assumptions and shortcomings of the PGA algorithm, the authors feel it is necessary to highlight assumptions made and address potential shortcomings with the GPGA algorithm:

1. In (11), the SAR scene is assumed to consist of a collection of stationary point scatterers — a common modeling assumption made with many SAR imaging applications. Moreover, (11) relies on stop-and-hop and narrowband signal approximations, and best models scenarios where scattering bodies are large compared to the carrier wavelength [77, 78].
2. Timing/position errors, which may result from propagation effects, hardware limitations, and/or motion measurement errors, are assumed to produce spatially invariant phase errors. For this assumption to hold, the SAR scene size is assumed to be chosen according to the scene size limits derived in Section 3.2.
3. For the statistical model described by (79) to (81), the GPGA algorithm provides an approximate MMLE of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers. For the presented formulation of the GPGA algorithm, no assumptions regarding specific scene types or scene content are made (e.g., scene containing strong isolated point scatterers and/or low/high contrast regions). This statement is not to be misinterpreted as reading, “the GPGA algorithm performs well with all different scene types,” just that the formulation does not rely upon any scene type or scene content assumptions.
4. Naturally, the performance of the GPGA algorithm (as well as the PGA algorithm) depends upon the SINR of selected scatterers. Thus, for SAR scenes

containing moderate to high SINR scatterers and/or regions, it is reasonable to assume the GPGA algorithm will provide an adequate autofocus solution (although, extreme edge cases may exist where this is not true). For scenes devoid of these convenient moderate or high SINR features, the GPGA algorithm phase error estimation accuracy is generally expected to degrade. Depending upon a number of factors, most notably SINR, this reduction in phase error estimation accuracy may or may not be suitable. In other words, while the estimation accuracy of the GPGA algorithm is generally expected to degrade with SAR scenes containing lower SINR scatterers and/or regions, it is difficult to say precisely for which scene types and scene content the GPGA algorithm phase error estimation accuracy will be inadequate.

4.8 Conclusion

In this section, the GPGA algorithm was formulated by first providing a side-by-side description of the PGA and GPGA algorithms, making evident the GPGA algorithm includes the PGA algorithm as a special case. Next, it was shown that the GPGA algorithm was applicable with the BPA, while preserving the four crucial signal processing steps comprising the PGA algorithm. Lastly, it was shown that with the generalized formulation, whether using the PFA or BPA, the GPGA algorithm provides an approximate MMLE of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers. In the next section, trade-offs between three GPGA phase estimators or approximate MMLEs are discussed.

V. Generalized Phase Gradient Autofocus Phase Estimation

In this section, discussion is provided on three GPGA phase estimators: 1) PD, 2) EVR, and 3) Max-SDR phase estimators. In this discussion, each of the three GPGA phase estimators are described in detail and their respective trade-offs are discussed. In addition, in this section a specialized IPM is presented, which exploits low-rank structure typically associated with the GPGA phase estimation problem, for more efficiently performing Max-SDR phase estimation.

5.1 Phase Differencing

The PD phase estimator from [20] is

$$\text{Arg} \left(\left[\widehat{\delta\phi} \right]_n \right) = \begin{cases} 0, & n = 0, \\ \text{Arg} \left(\sum_{n'=0}^{n-1} \sum_{i \in \mathcal{T}} \left[\tilde{\xi}_i \right]_{n'}^* \left[\tilde{\xi}_i \right]_{n'+1} \right), & \text{else,} \end{cases} \quad (92)$$

where without loss of generality the argument of the first element of the phase error estimate is assumed to be zero since the focus of a SAR image is invariant to a constant phase offset (i.e., one element of the phase error estimate may be arbitrarily assigned, with all other elements expressed relative to the arbitrarily assigned element).

As noted in [19], the PD phase estimator is derived assuming nearly constant reflectivity, high SINR scatterers are selected during the scatterer selection process (i.e., $\gamma_i \approx \gamma_i \mathbf{1}$ and $\mathbf{w}_i$ is negligible). For example, making this assumption, the matched filtered response from (15) for the i th selected scatterer is approximately

$$\tilde{\xi}_i = \text{Diag}(\delta\phi) \gamma_i + \mathbf{w}_i \approx \gamma_i \delta\phi. \quad (93)$$

Using (93) and rearranging summations, the PD phase estimator from (92) is approximately

$$\text{Arg}\left(\left[\widehat{\boldsymbol{\delta\phi}}\right]_n\right) \approx \begin{cases} 0, & n = 0, \\ \text{Arg}\left(\sum_{i \in \mathcal{T}} |\gamma_i|^2 \sum_{n'=0}^{n-1} [\boldsymbol{\delta\phi}]_{n'}^* [\boldsymbol{\delta\phi}]_{n'+1}\right), & \text{else.} \end{cases} \quad (94)$$

which consists of a weighted sum of the numerical integration of the measured phase difference or phase gradient, $[\boldsymbol{\delta\phi}]_{n'}^* [\boldsymbol{\delta\phi}]_{n'+1}$, for each selected scatterer (hence, it is termed the PD phase estimator).

From (92), one may see the PD phase estimator is easy to implement and has a relatively low computational complexity of $\mathcal{O}(NP)$. Although the PD phase estimator is computationally efficient, its performance suffers when high SINR scatterers do not exist, or are not located and selected during the scatterer selection process [19]. Moreover, little can be said regarding the suboptimality of the PD phase error estimate with respect to the solution set of (90). In the next section a more robust phase estimator, in the sense it does not rely upon a high SINR approximation, is discussed.

5.2 Eigenvector Relaxation

To formulate the EVR phase estimator, a change of variables is made with the GPGA phase estimation problem to equivalently express (90) as

$$\begin{aligned} & \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} \quad \widehat{\boldsymbol{\delta\phi}}^H \boldsymbol{\Xi} \widehat{\boldsymbol{\delta\phi}}, \\ & \text{subject to} \quad \left| \left[\widehat{\boldsymbol{\delta\phi}} \right]_n \right|^2 = \frac{1}{N}, \quad n \in \{0, \dots, N-1\}, \end{aligned} \quad (95)$$

where $\widehat{\boldsymbol{\delta\phi}} := \widehat{\boldsymbol{\delta\phi}}/\sqrt{N}$ is the change of variables. By relaxing the equality constraint in (95), what is referred to as the EVR problem is obtained

$$\begin{aligned} & \underset{\widehat{\boldsymbol{\delta\phi}} \in \mathbb{C}^N}{\text{maximize}} && \widehat{\boldsymbol{\delta\phi}}^H \boldsymbol{\Xi} \widehat{\boldsymbol{\delta\phi}}, \\ & \text{subject to} && \left\| \widehat{\boldsymbol{\delta\phi}} \right\|^2 = 1. \end{aligned} \tag{96}$$

A global optimal point to (96), let it be denoted by $\mathbf{v}^*$, is the leading eigenvector of the GPGA data matrix, $\boldsymbol{\Xi}$. In order for $\mathbf{v}^*$ to be directly useful for the phase estimation problem in (90), it is necessary for $\mathbf{v}^*$ to have unit-modulus elements (i.e., be feasible to (90)). Thus, a natural approach for ensuring a feasible point to (90) is obtained from the solution to (96), is to normalize the elements of $\mathbf{v}^*$ to be unit-modulus

$$\left[\widehat{\boldsymbol{\delta\phi}} \right]_n = \frac{[\mathbf{v}^*]_n}{|[\mathbf{v}^*]_n|}. \tag{97}$$

Applying this conversion process to obtain a feasible point to (90) is referred to as EVR, since the elements of the leading eigenvector of $\boldsymbol{\Xi}$ are normalized to be unit-modulus.

In [19, 20] the authors report the EVR phase estimator outperforms the PD phase estimator, especially at lower SINRs where the modeling assumption for the PD phase estimator breaks down. As will be seen in Chapter VI, this reported performance improvement obtained with the EVR phase estimator over the PD phase estimator is consistent with findings in this research. Additionally, in [19] the EVR phase estimator is shown to include the PD phase estimator as a special case when the number of pulses $N = 2$.

For performing EVR phase estimation, computing the leading eigenvector of $\boldsymbol{\Xi}$ is complexity $\mathcal{O}(N^2)$ [79]. However, when $P < N$, a condition which may be enforced by

the system designer of the GPGA algorithm, EVR phase estimation may be performed more efficiently. For instance, recognizing $\Xi = \tilde{\Xi}\tilde{\Xi}^H$ where $\tilde{\Xi} \in \mathbb{C}^{N \times P}$, EVR phase estimation may be performed with complexity $\mathcal{O}(NP)$ by computing the leading singular vector of $\tilde{\Xi}$ [79]. Although the computational complexity for the PD and EVR phase estimators is the same when $P < N$, in Chapter VI the EVR phase estimator is shown to have a slightly higher average run-time than the PD phase estimator.

Similar to the PD phase estimator, the EVR phase estimator generally provides little insight into the suboptimality of the phase error estimate with respect to the solution set of (90). The exception to this statement is the special case where the leading eigenvector of Ξ has constant modulus elements, in which case the EVR phase estimate is known to be a global optimal point to (90). In other words, only in special cases does the EVR phase estimator provide insight into the suboptimality of the phase error estimate, which is more than can be said for the PD phase estimator. In the next section a phase estimator which provides a worst-case suboptimality with respect to the solution set of (90) is discussed (i.e., provides insight into the suboptimality of the phase error estimate in all cases).

5.3 Max-Semidefinite Relaxation

To formulate the Max-SDR phase estimator, the GPGA phase estimation problem in (90) is first lifted to a higher dimension. Specifically, letting $\Phi = \widehat{\delta\phi}\widehat{\delta\phi}^H \in \mathbb{H}^N$, (90)

is equivalently expressed as

$$\begin{aligned}
& \underset{\Phi \in \mathbb{H}^N}{\text{maximize}} && \text{Tr}(\Phi \Xi), \\
& \text{subject to} && \text{diag}(\Phi) = \mathbf{1}, \\
& && \Phi \succeq \mathbf{0}, \\
& && \text{Rank}(\Phi) = 1.
\end{aligned} \tag{98}$$

In (98), the objective function and all constraints, with the exception of the rank-1 constraint, are convex. By relaxing the rank-1 constraint a semidefinite program is obtained

$$\begin{aligned}
& \underset{\Phi \in \mathbb{H}^N}{\text{maximize}} && \text{Tr}(\Phi \Xi), \\
& \text{subject to} && \text{diag}(\Phi) = \mathbf{1}, \\
& && \Phi \succeq \mathbf{0},
\end{aligned} \tag{99}$$

which is referred to as the SDR problem. The SDR problem in (99) is a convex optimization problem. Hence, a global maximum point to (99), let it be denoted by Φ^* , may be found to an arbitrary accuracy in polynomial time using an IPM [70].

In order for Φ^* to be directly useful for the phase estimation problem in (98), it is necessary Φ^* be rank-1 (i.e., be feasible to (98)). When Φ^* is rank-1, Φ^* is a global maximum point to (98). In this case, an MMLE of the phase errors, which is a global maximum point to (90), is obtained by normalizing the leading eigenvector of Φ^* to have unit-modulus elements. In the case Φ^* is not rank-1, a conversion process must be employed to map Φ^* to a feasible point of (90).

A natural approach to this conversion process is to apply EVR with the matrix Φ^* . More specifically, a feasible point to (90) is computed by normalizing the elements of the leading eigenvector of Φ^* to be unit-modulus. Albeit straightforward, little can be

said about the suboptimality of a feasible point to (90) obtained in this manner [80]. Hence, a more robust conversion process, in terms of the suboptimality of an obtained feasible point, may be desired.

Fortunately, this desired robustness is readily obtained by applying a process known as *randomization* [70]. The randomization process for obtaining a feasible point to (90) from Φ^* proceeds as follows:

1. Generate Q realizations distributed as $\varphi_q \sim \mathcal{CN}(\mathbf{0}, \Phi^*)$ for $q \in \{0, \dots, Q-1\}$.
2. Normalize the realizations to have unit-modulus elements: $[\varphi_q]_n := \frac{[\varphi_q]_n}{|[\varphi_q]_n|}$ for $q \in \{0, \dots, Q-1\}$.
3. Find the best feasible point to (90) from the set of normalized realizations:

$$\widehat{\delta\phi} = \arg \max_{\varphi_0, \dots, \varphi_{Q-1}} \varphi_q^H \Xi \varphi_q.$$

At first blush, the randomization process described above may seem to be anything but robust in terms of the suboptimality of the obtained feasible point. However, as stated in [70], “the intuitive ideas behind randomization are not difficult to see, yet the theoretical implications that follow are far from trivial.” In particular, in [69] the randomization process described above is shown to yield a feasible point to (98) with a worst-case approximation bound of $\pi/4$. Specifically, letting $f : \mathbb{C}^N \rightarrow \mathbb{R}$, $f(\varphi) = \varphi^H \Xi \varphi$ be the objective function from (90), the worst-case approximation bound derived in [69] states

$$f(\widehat{\delta\phi}^*) \geq \mathbb{E}[f(\varphi)] \geq \frac{\pi}{4} f(\widehat{\delta\phi}^*), \quad (100)$$

where $\widehat{\delta\phi}^*$ is an MMLE (i.e., global optimal point to (90)) and $\varphi \in \mathbb{C}^N$ is generated by the first two steps of the randomization process described above. The bound derived

in (100) not only has a profound theoretical implication, but practically speaking, it also has an implication. As noted in [70], the bound derived in (100) will tend to be tighter in practice as it is likely a single realization from the randomization process will yield an objective function value greater than the expected objective function value (assuming a reasonable number of realizations are generated).

To provide such a strong statement on the worst-case approximation bound for a phase error estimate, the Max-SDR phase estimator sacrifices computational complexity. In many SAR imaging applications the demand for increased computational complexity associated with the Max-SDR phase estimator using a generic IPM is too steep to make it practical. However, as noted in a number of works (most notably [81–83]), the complexity and run-time requirements of an IPM may be greatly reduced by exploiting structure associated with a given problem.

In the next section, a specialized IPM is presented which exploits low-rank structure typically associated with the GPGA phase estimation problem — thereby making Max-SDR phase estimation more viable for SAR imaging applications leveraging the GPGA algorithm.

5.4 Efficient Max-Semidefinite Relaxation

In this section, it is shown that the GPGA data matrix, Ξ , for the SDR problem in (99) may be constrained to be low-rank, while having negligible effect on the quality of the autofocus solution (i.e., on the quality of the focused SAR image). Then, by exploiting this low-rank structure, a specialized IPM is developed and presented for more efficiently solving the SDR problem in (99) as compared to using a generic IPM.

5.4.1 Low-Rank Data

A key signal processing step of the GPGA algorithm is scatterer selection. In Section 4.3, potential strategies for performing scatterer selection are discussed. Regardless of the chosen strategy, the objective of the scatterer selection process is the same — obtain as many high-quality observations of the phase errors as possible. However, computationally speaking, there is a benefit to selecting fewer scatterers. In the discussion to follow, a bound is derived on the number of selected scatterers needed, so both an efficient and robust implementation of the GPGA algorithm may be realized.

From the derivation in Appendix B, the inverse of the $(N - 1) \times (N - 1)$ Fisher information matrix (FIM) for the argument of the phase errors, $[\boldsymbol{\theta}]_n = \text{Arg}([\boldsymbol{\delta\phi}]_n)$ for $n \in \{0, \dots, N - 2\}$, is

$$\mathbf{J}_{\boldsymbol{\theta}, \boldsymbol{\theta}}^{-1} = \frac{1}{2N \sum_{i \in \mathcal{T}} \frac{(\sigma_{\gamma, i}^2)^2}{1 + N\sigma_{\gamma, i}^2}} (\mathbf{I} + \mathbf{1}\mathbf{1}^H), \quad (101)$$

where again, without loss of generality $\sigma_{\gamma, i}^2$ represents the SINR of the i th selected scatterer. Letting $\sigma_{\gamma, \min}^2$ denote the minimum SINR for all selected scatterers (i.e., $\sigma_{\gamma, \min}^2 \leq \sigma_{\gamma, i}^2$ for $i \in \mathcal{T}$), the diagonal elements of the inverse of the FIM are bounded above by

$$[\mathbf{J}_{\boldsymbol{\theta}, \boldsymbol{\theta}}^{-1}]_{n, n} \leq \frac{1 + N\sigma_{\gamma, \min}^2}{NP (\sigma_{\gamma, \min}^2)^2}. \quad (102)$$

For an efficient estimator, (102) may be used to determine the number of selected scatterers needed to produce a phase error estimate with a specified mean squared error (MSE), given a minimum selected scatterer SINR and number of slow-time pulses. For instance, using the generally accepted phase error threshold of $\pi/4$ for

negligible image defocus [2, 3], the minimum number of selected scatterers for an efficient estimator is determined from the bound on squared error by solving for P from

$$\frac{1 + N\sigma_{\gamma,\min}^2}{NP(\sigma_{\gamma,\min}^2)^2} \leq \left(\alpha \frac{\pi}{4}\right)^2, \quad (103)$$

where $\alpha \in (0, 1]$ is a proportionality constant set by the system designer for a phase error estimate with a specified MSE.

As will be seen in Chapter VI, Monte Carlo simulation results suggest the Max-SDR phase estimator is efficient for moderate SINRs (e.g., $\text{SINR} \geq 0$ dB) and a small number of slow-time pulses (e.g., $N \geq 10$). Thus, for most SAR imaging applications, where the number of slow-time pulses and SINR conditions are satisfied, the Max-SDR phase estimator may be considered to be an efficient estimator. Therefore, the number of selected scatterers needed for Max-SDR phase estimation, to obtain a phase error estimate with a given MSE, may be determined from (103). For example, using equality with (103) and considering a proportionality constant of $\alpha = 1/4$ to provide a margin of error for realizations that may have a larger squared error than the expectation (i.e., a margin of error of roughly 6 dB), the number of selected scatterers needed is

$$P = \left\lceil \left(\frac{16}{\pi}\right)^2 \frac{1 + N\sigma_{\gamma,\min}^2}{N(\sigma_{\gamma,\min}^2)^2} \right\rceil. \quad (104)$$

Figure 21 shows (104) evaluated over the parameter space of number of slow-time pulses and minimum selected scatterer SINR. From Figure 21 one may see that under moderate conditions the number of selected scatterers needed to obtain an adequate autofocus solution is no more than roughly 30. The moderate conditions under which

no more than roughly 30 selected scatterers are needed are as follows: 1) minimum selected scatterer SINR of roughly 0 dB and 2) number of slow-time pulses is greater than or equal to 10.

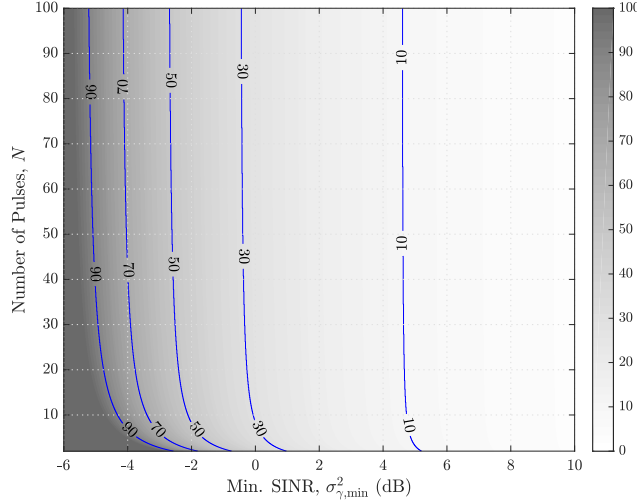


Figure 21. Number of scatterers, P , needed during the scatterer selection process for an efficient estimator to produce a phase error estimate with an MSE of $(\alpha\pi/4)^2$, with $\alpha = 1/4$.

In the next section, a specialized IPM is presented, which exploits the low-rank structure typically associated with the GPGA phase estimation problem.

5.4.2 Specialized Interior Point Method

In this section, a specialized IPM is developed for solving (99) using a similar strategy to that employed in [84]. In the context of Max-SDR phase estimation, the strategy employed in [84] proceeds as follows: 1) solve the Lagrange dual problem associated with the SDR problem in (99) using an IPM, then 2) use the computed Lagrange dual solution to calculate an optimal point to the SDR problem in (99). To improve upon the efficiency of the approach used in [84], the low-rank structure of the GPGA data matrix discussed in Section 5.4.1 is exploited.

The Lagrange dual problem for (99), referred to as the dual semidefinite relaxation (DSDR) problem, is derived in Appendix C and shown to be

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimize}} && -\mathbf{1}^T \mathbf{x}, \\ & \text{subject to} && \mathbf{G}(\mathbf{x}) \preceq \mathbf{0}, \end{aligned} \tag{105}$$

where $\mathbf{x} \in \mathbb{R}^N$ is the dual variable and $\mathbf{G} : \mathbb{R}^N \rightarrow \mathbb{H}^N$, $\mathbf{G}(\mathbf{x}) = \text{Diag}(\mathbf{x}) + \mathbf{\Xi}$ is a matrix involving the dual variable, $\mathbf{x}$, and the GPGA data matrix, $\mathbf{\Xi}$.

To solve (105) to an arbitrary accuracy, a logarithmic barrier function is used to enforce the linear matrix inequality (LMI), and a series of unconstrained optimization problems are solved using Newton's method. The series of unconstrained optimization problems, referred to as the dual barrier problems, are

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimize}} -\mu_k \mathbf{1}^T \mathbf{x} + \psi(\mathbf{x}), \tag{106}$$

where $\mu_{k+1} > \mu_k > 0$ is an increasing sequence of barrier parameters defining the series of dual barrier problems and the logarithmic barrier function $\psi : \mathbb{R}^N \rightarrow \mathbb{R}$ applied to the LMI in (105) is

$$\psi(\mathbf{x}) = \begin{cases} -\ln(|-\mathbf{G}(\mathbf{x})|), & |-\mathbf{G}(\mathbf{x})| > 0, \\ +\infty, & \text{else.} \end{cases} \tag{107}$$

Algorithm 1 is pseudocode describing how the series of dual barrier problems are used to obtain an ε -suboptimal point to the SDR problem in (99), where $\varepsilon > 0$ is an arbitrary suboptimality. From inspection of Algorithm 1, one may see the algorithm involves two key steps: 1) solving the unconstrained dual barrier problem in (106) using Newton's method, and 2) computing an ε -suboptimal point to (99) given an

ε -suboptimal point $\mathbf{x}_{k+1}^*$ to the dual barrier problem. In the proceeding discussion, additional details are provided on each of these key steps.

Algorithm 1: Specialized SDR Interior Point Method

inputs : GPGA data $\tilde{\boldsymbol{\xi}} \in \mathbb{C}^{N \times P}$ and suboptimality $\varepsilon > 0$

output: ε -suboptimal point $\boldsymbol{\Phi}^* \in \mathbb{H}^N$ to (99)

// Normalize data for numerical stability, since (99) is invariant to a positive scaling of the objective function.

```

1  $\tilde{\boldsymbol{\xi}} \leftarrow \tilde{\boldsymbol{\xi}} / \|\tilde{\boldsymbol{\xi}}\|_F$ 
2  $\mathbf{x}_0^* \leftarrow -(1 + 10^{-6})\mathbf{1}$  ; // strictly feasible point to (105)
3  $\mu_0 \leftarrow 1$  ; // initial barrier parameter
4  $a \leftarrow 1.1$  ; // barrier parameter update rate,  $a > 1$ 
5 for  $k \leftarrow 0$  to  $max. \text{ iter.} - 1$  do
6    $\mathbf{x}_{k+1}^*, \mathbf{G}^{-1}(\mathbf{x}_{k+1}^*) \leftarrow$  Solve (106) with Algorithm 2, given barrier parameter  $\mu_k$ ,
   initial strictly feasible point  $\mathbf{x}_k^*$ , suboptimality  $\varepsilon$ , and
   GPGA data  $\tilde{\boldsymbol{\xi}}$ 
7    $\mu_{k+1} \leftarrow a\mu_k$  ; // update the barrier parameter
8   if  $N/\mu_k \leq \varepsilon$  then // check suboptimality
9     // Approx.  $\varepsilon$ -suboptimal SDR point
9      $\boldsymbol{\Phi}^* \leftarrow N\mathbf{G}^{-1}(\mathbf{x}_{k+1}^*) / (\mathbf{1}^T \text{diag}(\mathbf{G}^{-1}(\mathbf{x}_{k+1}^*)))$ 
10    if  $-\mathbf{1}^T \mathbf{x}_{k+1}^* - \text{Tr}(\boldsymbol{\Xi} \boldsymbol{\Phi}^*) > \varepsilon$  then
11      | continue to the next iteration
12    else
13      | break from the loop
14    end
15  end
16 end
```

5.4.3 Newton's Method

In order to apply Newton's method to (106), both the gradient and Hessian of the dual barrier objective function are needed. Using Wirtinger calculus [85–89], the

gradient and Hessian are (see Appendix D)

$$\nabla_{\mathbf{x}} (-\mu_k \mathbf{1}^T \mathbf{x} + \psi(\mathbf{x})) = -\mu_k \mathbf{1} - \text{diag}(\mathbf{G}^{-1}(\mathbf{x})), \quad (108)$$

and

$$[\nabla_{\mathbf{x}}^2 (-\mu_k \mathbf{1}^T \mathbf{x} + \psi(\mathbf{x}))]_{m,n} = \left| [\mathbf{G}^{-1}(\mathbf{x})]_{m,n} \right|^2. \quad (109)$$

Given expressions for the gradient and Hessian, Algorithm 2 is pseudocode describing how the dual barrier problem in (106) is solved to an arbitrary accuracy using Newton's method. In Algorithm 2 the most computationally complex steps are the inversion of $\mathbf{G}(\mathbf{x})$ and the Hessian, as each of these inversions is complexity $\mathcal{O}(N^3)$ using a brute force solver. However, given the low-rank structure of the GPGA data matrix (e.g., $P < N$), the inverse of $\mathbf{G}(\mathbf{x})$ may be computed more efficiently with complexity $\mathcal{O}(NP^2)$ using the Woodbury matrix identity

$$\mathbf{G}^{-1}(\mathbf{x}) = \text{Diag}^{-1}(\mathbf{x}) - \mathbf{A}\mathbf{B}^{-1}\mathbf{A}^H, \quad (110)$$

where $\mathbf{A} = \text{Diag}^{-1}(\mathbf{x})\tilde{\boldsymbol{\xi}} \in \mathbb{C}^{N \times P}$ and $\mathbf{B} = \mathbf{I} + \tilde{\boldsymbol{\xi}}^H \text{Diag}^{-1}(\mathbf{x})\tilde{\boldsymbol{\xi}} \in \mathbb{H}^P$. Because Algorithm 2 is initialized with and only steps to strictly feasible points of the DSDR problem in (105), the inverse of $\mathbf{B}$ exists at each point. To see this, let $\mathbf{x}$ be a strictly feasible point to (105) satisfying the LMI $\mathbf{0} \preceq \tilde{\boldsymbol{\xi}}\tilde{\boldsymbol{\xi}}^H \prec \text{Diag}(-\mathbf{x})$. Then, using the matrix determinant lemma it is seen $0 < |-\mathbf{G}(\mathbf{x})| = |\text{Diag}(-\mathbf{x})| |\mathbf{B}|$, where $|\text{Diag}(-\mathbf{x})| > 0$ since $\mathbf{0} \prec \text{Diag}(-\mathbf{x})$. It then follows that $|\mathbf{B}| \neq 0$, which is true if and only if the inverse of $\mathbf{B}$ exists.

Next, a method is presented for computing the inverse Hessian with reduced com-

Algorithm 2: Newton's Method

input : Barrier parameter $\mu \in \mathbb{R}$, initial strictly feasible point $\mathbf{x}_0 \in \mathbb{R}^N$,
suboptimality $\varepsilon > 0$, and GPGA data $\boldsymbol{\xi} \in \mathbb{C}^{N \times P}$

output: ε -suboptimal point $\mathbf{x} \in \mathbb{R}^N$ to (106)

```
1 for  $k \leftarrow 0$  to  $\text{max. iter.} - 1$  do
2   if  $P < N$  then
3      $\mathbf{G}^{-1}(\mathbf{x}) \leftarrow$  inverse using (110)
4   else
5      $\mathbf{G}^{-1}(\mathbf{x}) \leftarrow$  brute force inverse
6   end
7    $\mathbf{g} \leftarrow -\mu \mathbf{1} - \text{diag}(\mathbf{G}^{-1}(\mathbf{x}))$  ; // gradient using (108)
8    $[\mathbf{H}]_{m,n} \leftarrow \left| [\mathbf{G}^{-1}(\mathbf{x})]_{m,n} \right|^2$  ; // Hessian using (109)
9   if  $P^2 < N/4$  then // See footnote10
10     $\mathbf{H}^{-1} \leftarrow$  inverse Hessian using (114)
11  else
12     $\mathbf{H}^{-1} \leftarrow$  brute force inverse Hessian
13  end
14   $\Delta \mathbf{x} \leftarrow -\mathbf{H}^{-1} \mathbf{g}$  ; // compute Newton step
15  if  $-\mathbf{g}^T \Delta \mathbf{x} \leq \varepsilon$  then // check suboptimality
16    break from the loop
17  end
18   $\beta \leftarrow$  backtracking line search step-size [82, pg. 464]
19  if  $\beta \|\Delta \mathbf{x}\| / \|\mathbf{x}\| \leq \varepsilon$  then // check progress
20    break from the loop
21  else // take Newton step
22     $\mathbf{x} \leftarrow \mathbf{x} + \beta \Delta \mathbf{x}$ 
23  end
24 end
```

plexity when¹ $P^2 < N$. First, let $\mathbf{B}^{-1} = \mathbf{C}_0 \mathbf{C}_0^H$ and $\mathbf{C} = \mathbf{A} \mathbf{C}_0 \in \mathbb{C}^{N \times P}$. Then, using (110) the Hessian from (109), let it be denoted by $\mathbf{H} \in \mathbb{S}^N$, is

$$\begin{aligned} \mathbf{H} &= (\text{Diag}^{-1}(\mathbf{x}) - \mathbf{C} \mathbf{C}^H) \odot (\text{Diag}^{-1}(\mathbf{x}) - \mathbf{C} \mathbf{C}^H)^* \\ &= \mathbf{D} + \tilde{\mathbf{C}} \tilde{\mathbf{C}}^H, \end{aligned} \quad (111)$$

where $\mathbf{D} \in \mathbb{S}^N$ is a diagonal matrix

$$\mathbf{D} = \text{Diag}^{-2}(\mathbf{x}) - 2\text{Diag}^{-1}(\mathbf{x}) \odot (\mathbf{C} \mathbf{C}^H), \quad (112)$$

and $\tilde{\mathbf{C}} \in \mathbb{C}^{N \times P^2}$ is

$$\tilde{\mathbf{C}} = \begin{bmatrix} \text{Diag}^H(\mathbf{c}_0) \mathbf{C}, & \dots, & \text{Diag}^H(\mathbf{c}_{P-1}) \mathbf{C} \end{bmatrix}, \quad (113)$$

with $\mathbf{c}_p$ denoting the p th column vector of $\mathbf{C}$. Using the Woodbury matrix identity, the inverse Hessian may be computed with complexity $\mathcal{O}(NP^4)$ by

$$\mathbf{H}^{-1} = \mathbf{D}^{-1} - \mathbf{D}^{-1} \tilde{\mathbf{C}} \left(\mathbf{I} + \tilde{\mathbf{C}}^H \mathbf{D}^{-1} \tilde{\mathbf{C}} \right)^{-1} \tilde{\mathbf{C}}^H \mathbf{D}^{-1}, \quad (114)$$

since $\mathbf{D}$ is diagonal and easily inverted.

5.4.4 Max-Semidefinite Relaxation Feasible Point

From (99) and (105) one may see Slater's condition is satisfied, since for each problem there exists a strictly feasible point [82]. For example, $\mathbf{\Phi} = \mathbf{I}$ and $\mathbf{x} = -(\|\mathbf{\Xi}\|_F^2 + \delta)$ are strictly feasible points to (99) and (105), respectively, where $\delta > 0$ is an arbitrary

¹Although the complexity is reduced when $P^2 < N$, the author has found through experimentation that due to added computational steps, the described method for computing the inverse Hessian tends to provide a run-time performance improvement when $P^2 < N/4$.

positive constant. Because Slater's condition is satisfied, the duality gap between the primal and dual problems in (99) and (105) is zero [82, 90, 91], meaning

$$\text{Tr}(\Xi\Phi^*) = -\mathbf{1}^T\mathbf{x}^*, \quad (115)$$

where $\Phi^* \in \mathbb{H}^N$ and $\mathbf{x}^* \in \mathbb{R}^N$ are global optimal points to (99) and (105), respectively.

As shown in [82, pp. 600], given an optimal point to (106), let it be denoted by $\mathbf{x}_{k+1}^*$, a feasible point to (99) is obtained by $\Phi^* = -\mathbf{G}^{-1}(\mathbf{x}_{k+1}^*)/\mu_k$, where μ_k is the barrier parameter used in (106) for obtaining $\mathbf{x}_{k+1}^*$. In Algorithm 2 this step is not performed explicitly, since only an ε -suboptimal point is obtained to the dual barrier problem in (106) (i.e., the ε -suboptimality potentially leads to the scale factor $-1/\mu_k$ not being precisely correct). Instead, to account for this potential imprecision, the inverse of $\mathbf{G}(\mathbf{x})$ is scaled by the mean of the diagonal elements, $\mathbf{1}^T \text{diag}(\mathbf{G}^{-1}(\mathbf{x}_{k+1}^*)) / N$. Since each of the diagonal elements consists of a potentially perturbed version of $-1/\mu_k$ due to having an ε -suboptimal point $\mathbf{x}_{k+1}^*$, the mean of the diagonal values captures the true perturbed scale factor needed to compute an approximate feasible point to (99). In addition, because Φ^* is computed using an ε -suboptimal point to (105), the suboptimality of Φ^* may not be precisely ε when $N/\mu_k \leq \varepsilon$. As a result, Algorithm 1 is designed to evaluate the true suboptimality and perform additional iterations of the specialized SDR IPM if necessary.

5.4.5 Computational Complexity

To conclude this section, the complexity of the presented specialized IPM is determined and compared against a generic IPM. From [82], the worst-case number of iterations needed for Algorithm 1, ignoring leading constant terms, is $\ln(N/\varepsilon)$. Ad-

ditionally, the worst-case number of Newton’s method iterations (i.e., Algorithm 2 iterations) is $\sqrt{N}$ [82]. Using these worst-case number of iterations, Table 3 is a tabulation of the computational complexity for the specialized and a generic IPM for solving (99). Although the computational complexity for the specialized IPM is only reduced when $P^2 < N$, the number of $\mathcal{O}(N^3)$ operations performed to solve (99) is reduced compared to a generic IPM when $P < N$. In Chapter VI run-time results are presented making this point evident.

Table 3. Computational complexity comparison between the specialized IPM in Algorithm 1 and a generic IPM for solving (99).

	Specialized IPM	Generic IPM
$P^2 < N$	$\mathcal{O}(N^{1.5}P^4 \ln(N/\varepsilon))$	$\mathcal{O}(N^{3.5} \ln(N/\varepsilon))$
$P^2 \geq N$	— $\mathcal{O}(N^{3.5} \ln(N/\varepsilon))$ —	

5.5 Conclusion

In this section, discussion was provided on three GPGA phase estimators: 1) PD, 2) EVR, and 3) Max-SDR phase estimators. Table 4 is a summary of the trade-offs discussed for each of the three GPGA phase estimators. In addition, in this section a specialized IPM was presented, which exploits low-rank structure typically associated with the GPGA phase estimation problem, for more efficiently performing Max-SDR phase estimation. In the next section, results produced using each of the discussed GPGA phase estimators are presented and compared.

Table 4. Summary of trade-offs between different GPGA phase estimators.

Estimator	Summary
PD	- Assumes high SINR scatterers are selected - Most efficient with complexity $\mathcal{O}(NP)$ - No insight into suboptimality of phase error estimate
EVR	- Assumes moderate or nearly constant SINR scatterers are selected - Second most efficient with complexity $\mathcal{O}(NP)$ - In general, no insight into suboptimality of phase error estimate
Max-SDR	- Assumes moderate or nearly constant SINR scatterers are selected - Least efficient with complexity $\mathcal{O}(N^{1.5}P^4 \ln(N/\varepsilon))$ when $P^2 \leq N$, and $\mathcal{O}(N^{3.5} \ln(N/\varepsilon))$ otherwise - Provides phase error estimate with worst-case suboptimality

VI. Results

In this section, four sets of results produced using the GPGA algorithm with different phase estimators are presented and discussed. The first set of results includes Monte Carlo simulation results examining the MSE of the argument of phase error estimates for different GPGA phase estimators versus SINR. The second set of results includes Monte Carlo simulation results examining the average run-time for different GPGA phase estimators versus the size of the GPGA data (i.e., number of slow-time pulses N and number of selected scatterers P). For the third set of results, the GPGA algorithm is used to autofocus a measured, near-field bistatic SAR data set collected using a small-scale experimental setup in the Radar Instrumentation Laboratory (RAIL) at the Air Force Institute of Technology (AFIT). For the fourth set of results, the GPGA algorithm is applied to measured, far-field SAR data from the GOTCHA Public Release Dataset [92], where synthetic phase errors are applied to the data to produce image defocus.

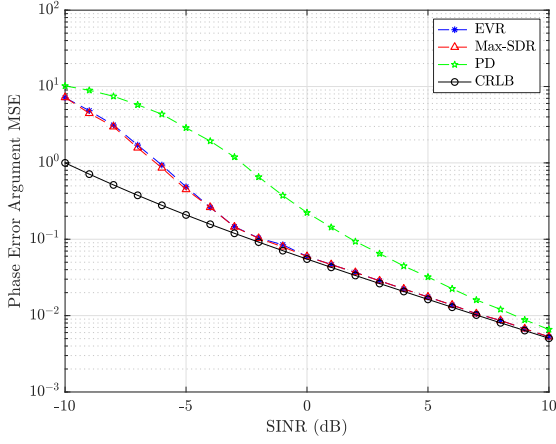
For each of the presented results, the GPGA algorithm is configured as follows: 1) to attempt to select either moderate or constant SINR scatterers, only image points with an image intensity within 10 dB of the maximum image intensity may be selected; 2) a maximum number of 30 scatterers may be selected per GPGA iteration to reduce computational complexity; 3) the GPGA algorithm is run for three iterations (except where explicitly stated). In addition, for performing Max-SDR phase estimation, Algorithm 1 is implemented in MATLAB with a suboptimality of $\varepsilon = 10^{-3}$ and number of randomization realizations $Q = 500$. Each of the aforementioned settings are free to be changed by the system designer as seen fit. However, the author has found the listed configuration works well on a number of data sets, all while reducing computational complexity.

6.1 Phase Error Estimate Mean Squared Error

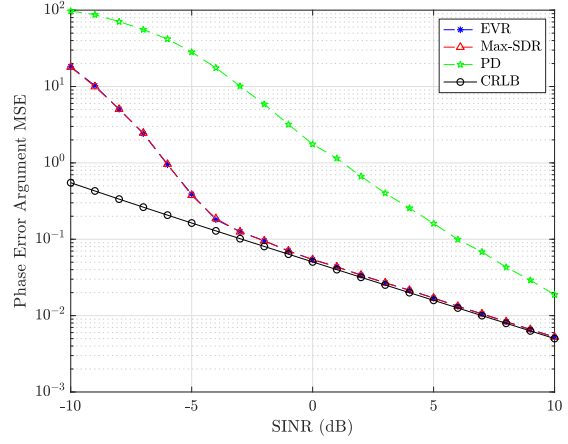
Figure 22 depicts Monte Carlo simulation results showing the MSE of the argument of phase error estimates versus SINR for different GPGA phase estimators, with $N \in \{10, 100\}$ slow-time pulses and $P \in \{20, 30\}$ selected scatterers. In addition, Figure 22 shows the CRLB for each SINR calculated from (101). For each trial of the Monte Carlo simulation, P vectors containing constant SINR range profile samples are drawn from the marginal distribution in (83), and used with the different GPGA phase estimators to produce a phase error estimate. Then, the MSE of the argument of each phase error estimate, as compared to the argument of the true phase errors, is computed. For each SINR shown in Figure 22, 5×10^3 Monte Carlo trials are run.

From Figure 22 it is clear the EVR and Max-SDR phase estimators drastically outperform the PD phase estimator at lower SINRs, which is consistent with discussion points in Section 5.1 and comments made in the literature regarding the bias of the PD phase estimator at lower SINRs [17, 19]. In addition, Figure 22 suggests that on average, the performance of the EVR and Max-SDR phase estimators is nearly equivalent, at least as simulated with the ideal model in (83). However, as mentioned previously, the EVR phase estimator does not share the same worst-case approximation bound provided with the Max-SDR phase estimator (i.e., little can be said regarding the suboptimality of the EVR phase error estimate with respect to the solution set of MMLEs).

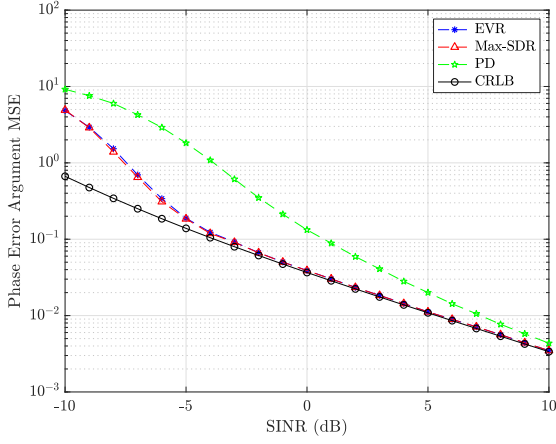
From the expression for the CRLB in (101), it is seen that for large pulse-SINR products (i.e., $N\sigma_{\gamma,i}^2 \gg 1$), the CRLB is inversely related to only the SINR and number of selected scatterers. In other words, for large pulse-SINR products the



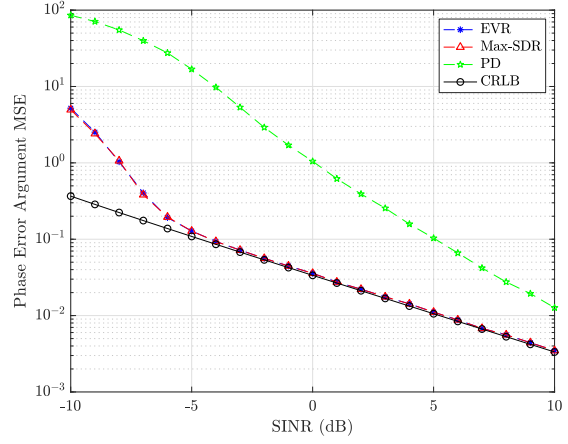
(a) $N = 10, P = 20$



(b) $N = 100, P = 20$



(c) $N = 10, P = 30$



(d) $N = 100, P = 30$

Figure 22. MSE of the argument of the phase error estimate for the PD, EVR, and Max-SDR phase estimators versus SINR and the size of the GPGA data (i.e., number of slow-time pulses N and number of selected scatterers P).

CRLB is approximately

$$[\mathbf{J}_{\theta, \theta}^{-1}]_{n,n} \approx \frac{1}{P\sigma_{\gamma,i}^2}. \quad (116)$$

Thus, for an efficient estimator with a large pulse-SINR product, the variance on phase error estimates is driven primarily by the SINR and number of selected scatterers. From Figure 22 this relationship is made evident. For instance, in Figures 22a and 22b, and Figures 22c and 22d, the CRLB approaches nearly the same value as SINR

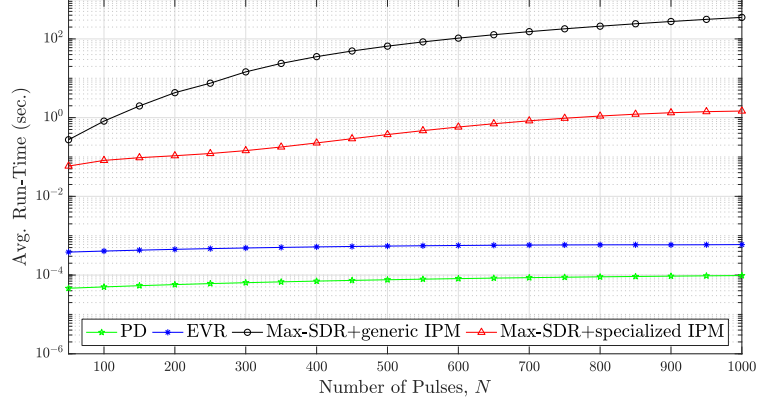
increases, whether using $N = 10$ or $N = 100$. In the next section, the average run-time for different GPGA phase estimators is presented and discussed.

6.2 Average Run-Time

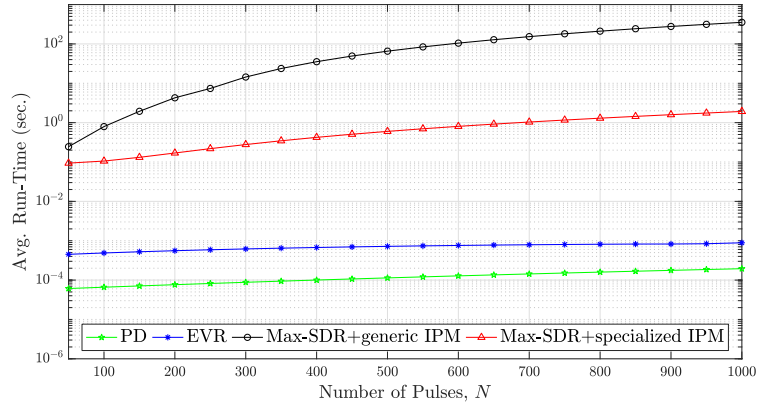
In this section, Monte Carlo simulations are run to compute the average run-time for different GPGA phase estimators versus the size of the GPGA data (i.e., number of slow-time pulses N and number of selected scatterers P). In addition, to make evident the efficiency improvement provided by the specialized IPM over a generic IPM in solving (99), two separate implementations of the Max-SDR phase estimator are considered using the following IPMs: 1) specialized IPM described in Algorithm 1 and 2) generic IPM from the SDPT3 solver shipped with CVX [93]. For the Monte Carlo simulation the average run-time for each GPGA phase estimator is recorded using 10^3 Monte Carlo trials.

The run-time results in Figure 23 make evident the efficiency improvement realized with the specialized IPM, as compared to a generic IPM, when the number of slow-time pulses is large compared to the number of selected scatterers. For example, for $N = 50$ and $P = 30$ the generic IPM is slightly faster than the specialized IPM. However, for $N = 10^3$, which is typical for SAR imaging applications, the specialized IPM is roughly two orders of magnitude faster than the generic IPM.

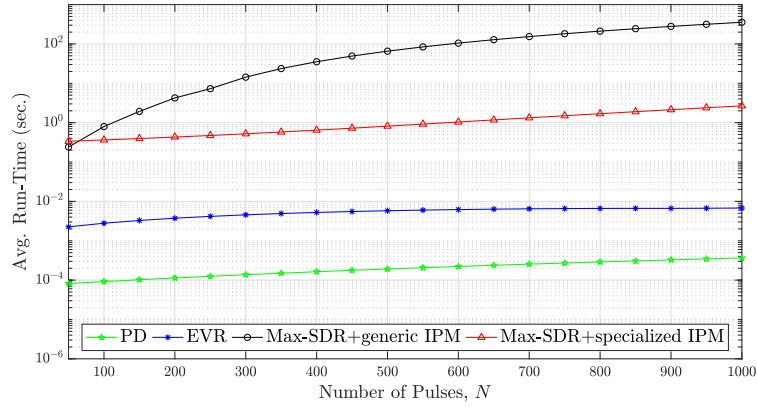
As expected, the average run-time for the PD and EVR phase estimators is significantly less than the Max-SDR phase estimator, as both of their computational complexities is $\mathcal{O}(NP)$. However, as previously discussed, the efficiency gained with the PD and EVR phase estimators over the Max-SDR phase estimator comes at the expense of robustness, in terms of having a worst-case approximate bound on a phase



(a) $P = 2$



(b) $P = 10$



(c) $P = 30$

Figure 23. Average run-time for different GPGA phase estimators versus the size of the GPGA data (i.e., number of slow-time pulses N and number of selected scatterers P).

error estimate as compared to the solution set of MMLEs (i.e., represents an efficiency versus robustness trade-off).

As an additional point of context regarding the run-time associated with SDR phase estimation, [55] reported in 2013 that SDR phase estimation, not exploiting low-rank structure and using a spectral bundle method, took on the order of 30 minutes and 24 hours to converge to a solution for $N \in \{1500, 2000\}$ with 8 GB of RAM and a clock rate of 2.66 GHz. On the other hand, the presented specialized IPM takes roughly 7 and 13 seconds for performing SDR phase estimation for $P = 30$ and $N \in \{1500, 2000\}$ with 16 GB of RAM and a clock rate of 2.6 GHz. In the next section the GPGA algorithm is used to autofocus a measured, near-field bistatic SAR data set.

6.3 Near-Field Imaging

Figure 24 depicts the small-scale experimental setup used for collecting bistatic SAR data in the RAIL at AFIT. The system and setup shown in Figure 24 has been used previously to generate test data for passive and bistatic SAR algorithm validation [60, 94–97]. On the transmit side, the RAIL system consists of a Tektronix arbitrary waveform generator (AWG) connected to an AirMax 2G-16-90 sector antenna. The AWG is controlled with a laptop through a MATLAB graphical user interface (GUI), and is capable of loading MATLAB data files containing arbitrary waveforms. On the receive side, the RAIL system consists of an AirMax 2G-16-90 sector antenna mounted on a linear motor driven rail that is controlled with a laptop through a MATLAB GUI. The receive antenna is fed into a Tektronix TDS6124 digital storage oscilloscope, which records and stores data for post-processing.

Figure 25a depicts the measured defocused SAR image, where the defocus is primarily a result of unknown, random timing delays in the software triggers of the system

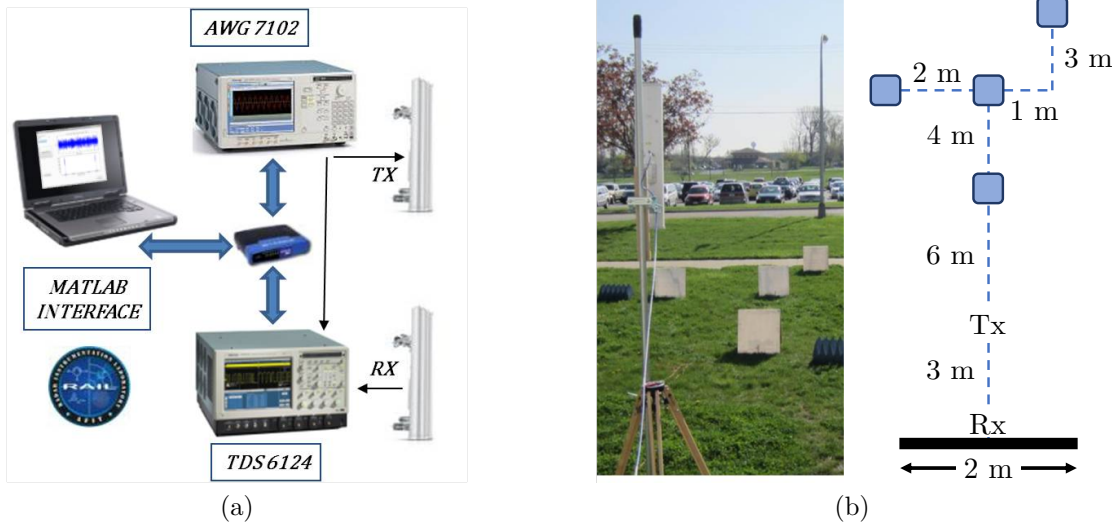


Figure 24. Experimental system and setup from [60, 94–97].

which contribute to receiver position uncertainty during the collection. From Figure 25a a near-field response is evidenced by the curvature of the energy in the image along the cross-range dimension. Hence, the PFA is not well-suited for this collection geometry as the far-field approximation does not hold. Figures 25b to 25d depict the autofocused BPA images and the scatterers selected during the third iteration of the GPGA algorithm using the PD, EVR, and Max-SDR phase estimators. For each of the autofocused images, the four plate reflectivities are more clearly discernible than in the measured defocused image, with only minor differences present between each of the results. For instance, with the PD phase estimator each of the autofocused plate responses seem to be rotated from the scene center, while appearing to have little effect on the overall focus of the image. With the EVR and Max-SDR phase estimation results, the autofocused images are nearly identical with only subtle differences arising in the sidelobes of the plate responses.

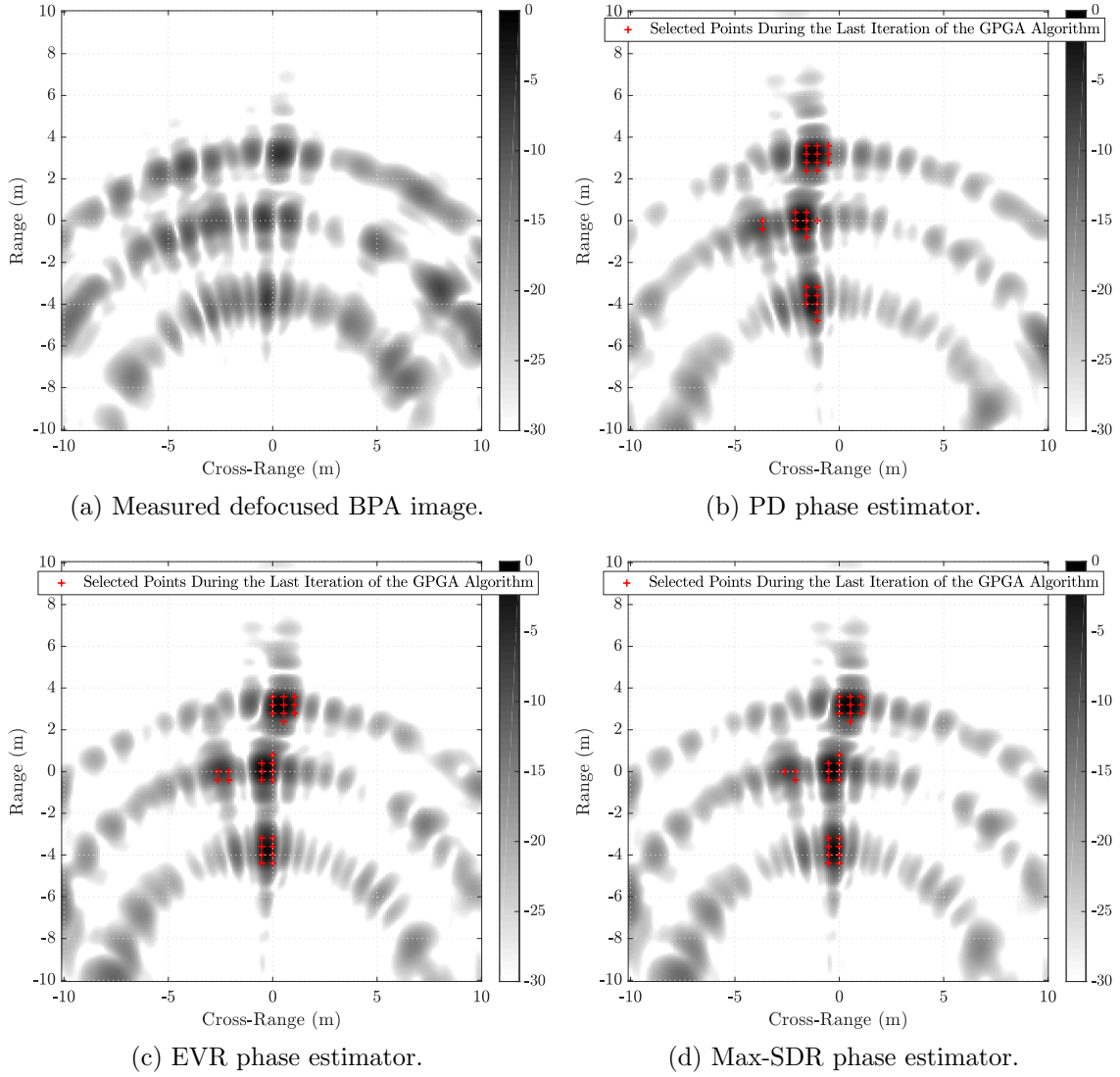


Figure 25. Measured and autofocused BPA images produced using the GPGA algorithm with different phase estimators, for the experimental setup in Figure 24.

6.4 Far-Field Imaging

In this section, the public release data set, referred to as “Gotcha Volumetric SAR Data Set, Version 1.0,” is used to produce example autofocus results using the GPGA algorithm [92]. The Gotcha data set consists of X-band SAR data with a bandwidth of 640 MHz [92]. For results shown herein, an aperture consisting of 469 slow-time pulses of data, ranging from 0° to 4° azimuth, is used.

6.4.1 Using the Polar Format Algorithm

For applying the GPGA algorithm with the PFA, the Gotcha SAR data is first resampled from a polar to a uniformly sampled Cartesian grid, resulting in 1024 slow-time pulses of resampled data. Then, synthetic phase errors are applied to the resampled data. The synthetic phase errors consist of white phase errors for each slow-time pulse chosen from a uniform distribution over the interval $[-\pi, \pi)$.

Figures 26a and 26b depict the ideal and defocused PFA SAR images. Figures 26c to 26e show autofocused PFA SAR images after applying the GPGA algorithm for three iterations using different GPGA phase estimators. Figure 26f shows the true and estimated phase errors. For the shown PFA autofocus results, quantities which the autofocus solution are invariant to are removed. These invariant quantities include: 1) constant phase offset, 2) 2π jumps in phase, and 3) linear phase ramp (which results in image translation) [98]. As evidenced by Figures 26c to 26e, the GPGA algorithm using the PFA with the PD, EVR, and Max-SDR phase estimators is able to restore the defocused image to closely resemble the ideal image. Moreover, with the shown autofocused results there appear to be no distinguishable differences between the autofocused images produced using the PD, EVR, and Max-SDR phase estimators.

To provide additional insight into how the performance of the GPGA algorithm may change based on its configuration, the MSE of the phase error estimates produced from the defocused SAR image in Figure 26b is evaluated versus different configuration parameters. In particular, Tables 5 to 7 are tabulations of the MSE of the phase error estimates and number of selected scatterers versus the applied relative threshold for scatterer selection and number of GPGA iterations run. As illustrated with the results in Tables 5 to 7, the MSE of the phase error estimates is generally expected

to decrease with the number of moderate or constant SINR scatterers selected and number of GPGA iterations run. Moreover, the Max-SDR phase estimator is generally expected to provide phase error estimates with the lowest MSE, followed in order by the EVR and PD phase estimators.

Table 5. MSE of PD phase error estimates in radians² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.

		Iteration				
		Iter. #1	Iter. #2	Iter. #3	Iter. #4	Iter. #5
Rel. Threshold	15 dB	3.251 30	0.796 30	0.092 17	0.075 8	0.053 7
	10 dB	3.251 30	0.330 3	0.168 6	0.055 6	0.045 3
	5 dB	3.251 30	2.485 1	0.308 1	0.308 1	0.308 1

Table 6. MSE of EVR phase error estimates in radians² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.

		Iteration				
		Iter. #1	Iter. #2	Iter. #3	Iter. #4	Iter. #5
Rel. Threshold	15 dB	3.311 30	0.115 30	0.045 7	0.045 7	0.045 7
	10 dB	3.311 30	2.798 10	0.045 3	0.045 3	0.045 3
	5 dB	3.311 30	3.248 3	0.097 1	0.087 1	0.087 1

6.4.2 Using the Backprojection Algorithm

For applying the GPGA algorithm with the BPA, synthetic phase errors are applied directly to the 469 pulses of data (i.e., there is no polar-to-Cartesian resampling of

Table 7. MSE of Max-SDR phase error estimates in radians² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.

		Iteration				
		Iter. #1	Iter. #2	Iter. #3	Iter. #4	Iter. #5
Rel. Threshold	15 dB	3.278 30	0.126 30	0.012 7	0.012 7	0.012 7
	10 dB	3.278 30	0.285 24	0.013 3	0.013 3	0.013 3
	5 dB	3.278 30	3.267 12	0.333 2	0.075 1	0.075 1

the data). Again, the synthetic phase errors consist of white phase errors for each slow-time pulse chosen from a uniform distribution over the interval $[-\pi, \pi)$.

Figures 27a and 27b depict the ideal and defocused BPA SAR images. Figures 27c to 27e show the autofocused BPA SAR images after applying the GPGA algorithm for three iterations using different GPGA phase estimators. Figure 27f shows the true and estimated phase errors. For the shown BPA autofocus results, quantities which the autofocus solution are invariant to are removed. These invariant quantities include: 1) constant phase offset and 2) 2π jumps in phase [98]. As evidenced by Figures 27c to 27e, the GPGA algorithm using the BPA with the PD, EVR, and Max-SDR phase estimators is able to restore the defocused image to closely resemble the ideal image. Moreover, with the shown autofocused results there appear to be no distinguishable differences between the autofocused images produced using the PD, EVR, and Max-SDR phase estimators.

Again, to provide additional insight into how the performance of the GPGA algorithm may change based on its configuration, the MSE of the phase error estimates produced from the defocused SAR image in Figure 27b is evaluated versus different configuration parameters. In particular, Tables 8 to 10 are tabulations of the MSE of

the phase error estimates and number of selected scatterers versus the applied relative threshold for scatterer selection and number of GPGA iterations run. As illustrated with the results in Tables 8 to 10, the MSE of the phase error estimates is generally expected to decrease with the number of moderate or constant SINR scatterers selected and number of GPGA iterations run. Moreover, the Max-SDR phase estimator is generally expected to provide phase error estimates with the lowest MSE, followed in order by the EVR and PD phase estimators.

Table 8. MSE of PD phase error estimates in radians² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.

		Iteration				
		Iter. #1	Iter. #2	Iter. #3	Iter. #4	Iter. #5
Rel. Threshold	15 dB	3.439 30	0.037 30	0.037 30	0.038 30	0.039 30
	10 dB	3.439 30	0.037 30	0.029 17	0.029 16	0.029 17
	5 dB	3.439 30	0.037 30	0.032 6	0.032 6	0.032 6

Table 9. MSE of EVR phase error estimates in radians² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.

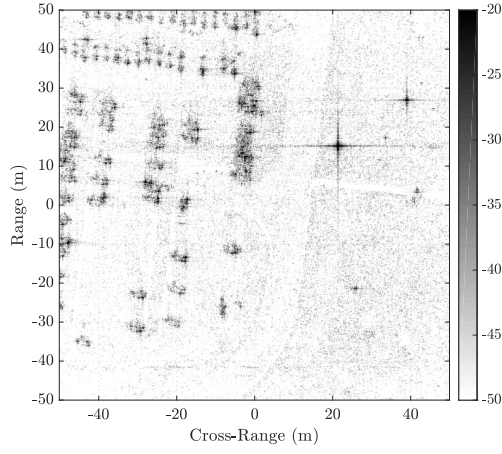
		Iteration				
		Iter. #1	Iter. #2	Iter. #3	Iter. #4	Iter. #5
Rel. Threshold	15 dB	0.069 30	0.031 30	0.031 30	0.031 30	0.031 30
	10 dB	0.069 30	0.029 19	0.029 19	0.029 18	0.029 18
	5 dB	0.069 30	0.030 6	0.030 6	0.030 6	0.030 6

Table 10. MSE of Max-SDR phase error estimates in radians² and number of selected scatterers versus the relative threshold for scatterer selection and the number of GPGA iterations run.

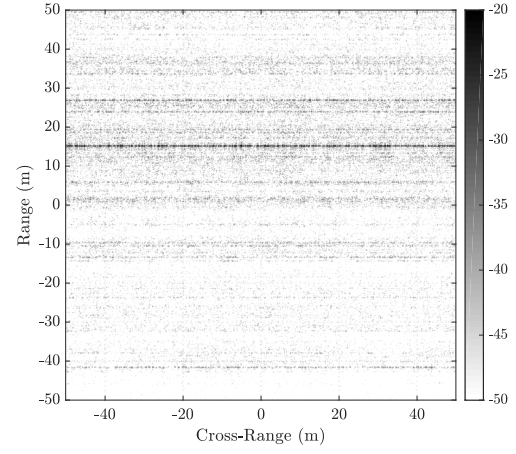
		Iteration				
		Iter. #1	Iter. #2	Iter. #3	Iter. #4	Iter. #5
Rel. Threshold	15 dB	0.034 30	0.029 30	0.029 30	0.029 30	0.029 30
	10 dB	0.034 30	0.029 19	0.029 19	0.029 18	0.029 17
	5 dB	0.034 30	0.029 6	0.029 6	0.029 6	0.029 6

6.5 Conclusion

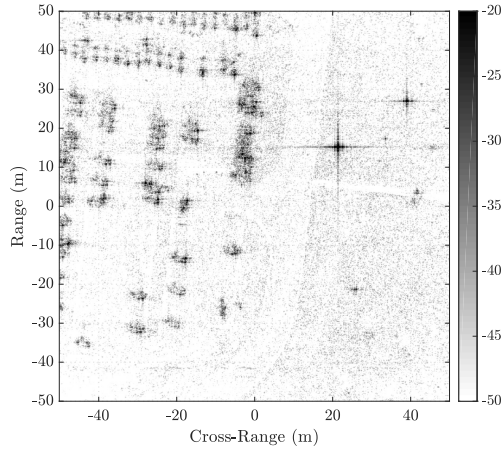
In this section, four sets of results produced using the GPGA algorithm with different phase estimators were presented and discussed. The presented results illustrated the effectiveness of the GPGA algorithm using either the PFA or BPA, and better made evident the performance versus efficiency trade-off that exists with the PD, EVR and Max-SDR phase estimators. In the next section concluding remarks are made.



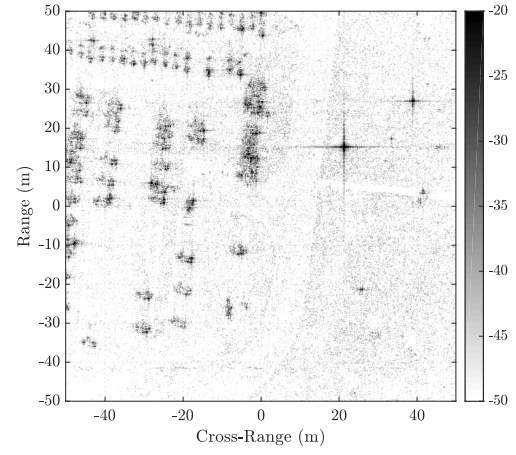
(a) Ideal PFA Image.



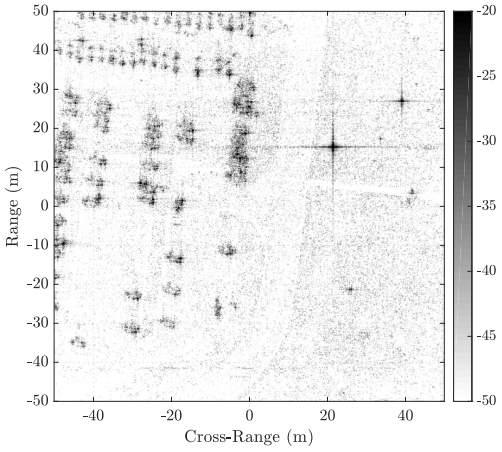
(b) Defocused PFA Image.



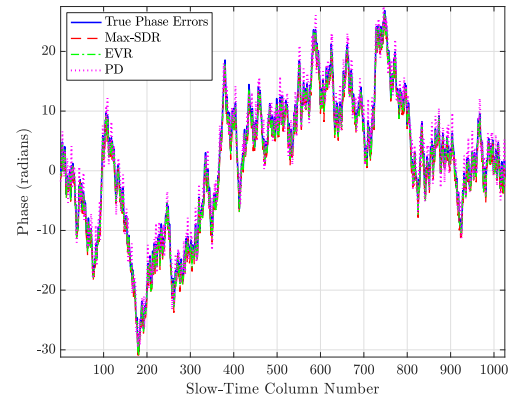
(c) PD phase estimator.



(d) EVR phase estimator.

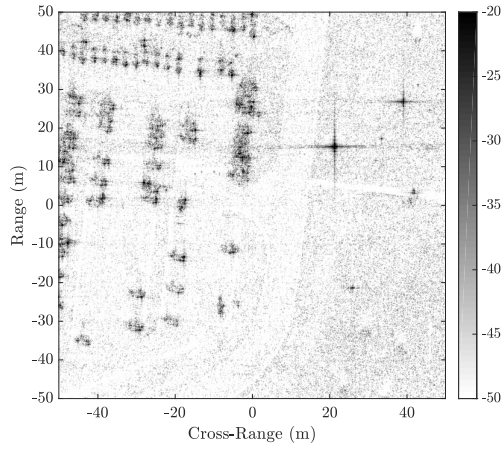


(e) Max-SDR phase estimator.

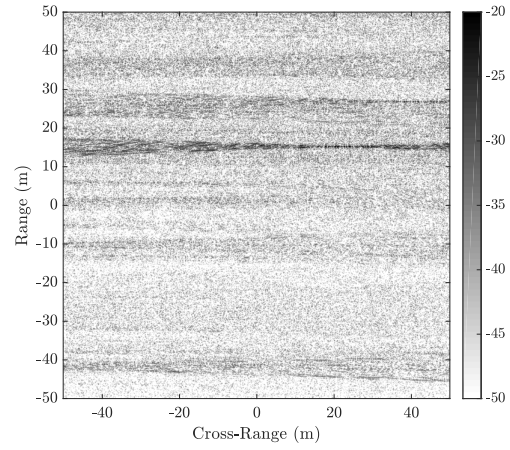


(f) True and estimated phase errors.

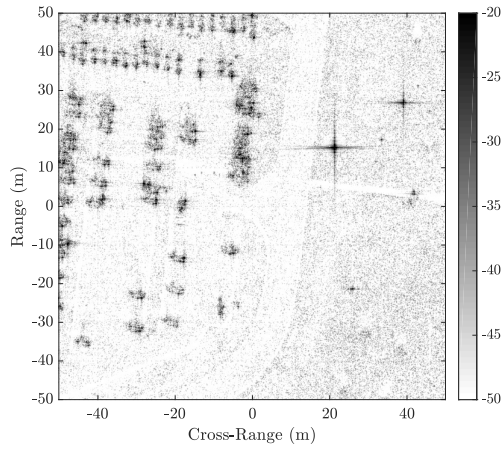
Figure 26. Autofocus results produced using the GPGA algorithm with different phase estimators and the PFA for SAR image formation.



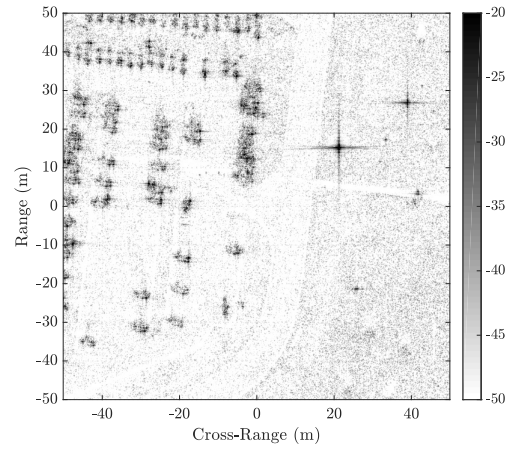
(a) Ideal BPA Image.



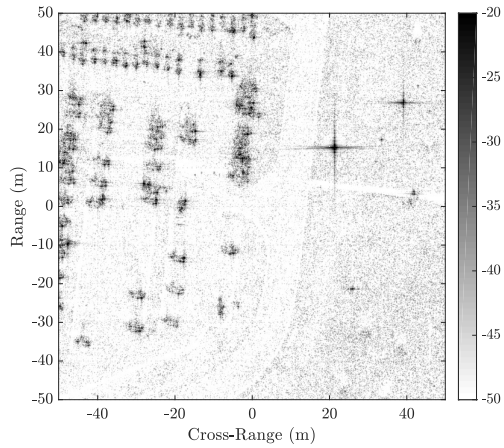
(b) Defocused BPA Image.



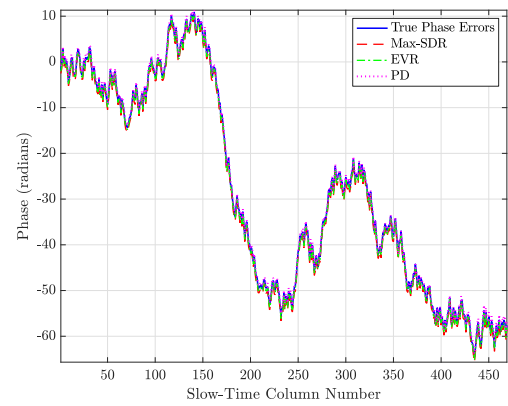
(c) PD phase estimator.



(d) EVR phase estimator.



(e) Max-SDR phase estimator.



(f) True and estimated phase errors.

Figure 27. Autofocus results produced using the GPGA algorithm with different phase estimators and the BPA for SAR image formation.

VII. Conclusion

In this research, advancements to the current state-of-the-art in SAR autofocus were made by extending or generalizing the well-known PGA algorithm, resulting in what was termed the GPGA algorithm. The material from this research consists of three primary contributions and a number of secondary contributions, which have culminated in four publications [60, 66–68]. Together, the primary and secondary contributions of this research provide greater insight into topics concerning: 1) how the GPGA algorithm relates to other existing SAR autofocus algorithms, 2) when and where the GPGA algorithm is applicable, 3) strategies for configuring/applying the GPGA algorithm, 4) statistical motivation for the GPGA algorithm, and 5) trade-space study for three different GPGA phase estimators. In the sections to follow each of the three primary contributions are summarized, with each section primarily concerned with material presented in each of the respective publications [66–68].

7.1 Spatially Invariant Defocus Model

Many existing autofocus algorithms, including the GPGA algorithm, rely on the assumption that phase errors, which cause SAR image defocus, are spatially invariant [1–3, 13, 15–17, 19, 20, 23, 26, 28, 30–33, 35, 38, 39, 42, 45, 46, 48, 50, 55, 59, 60]. In many cases this assumption holds, but for larger scene sizes or scenarios where the far-field approximation does not hold, the spatially invariant phase error approximation breaks down. In this work, an approximate upper bound was derived on the spatially variant phase error component of phase errors. The derived bound applies with near- and far-field bistatic geometries, with monostatic as a special case, and is expressed in terms of either known quantities or quantities which typically have a known upper

bound. The derived bound enables SAR scene size limits to be computed, such that the phase errors are approximately spatially invariant according to a specified phase threshold.

7.2 Generalized Phase Gradient Autofocus Algorithm

In this research, the GPGA algorithm was developed and shown to be applicable with both the PFA and BPA, thereby directly supporting a wide range of collection geometries and SAR imaging modalities. The GPGA algorithm preserves the four crucial signal processing steps comprising the PGA algorithm, while alleviating the constraint of using a single scatterer per range cut for phase error estimation which exists with the PGA algorithm. Moreover, the GPGA algorithm, whether using the PFA or BPA, yields an approximate MMLE of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers.

7.3 Max-Semidefinite Relaxation Phase Estimation

In this work, a new approximate MMLE, termed the Max-SDR phase estimator, was proposed for use with the GPGA algorithm. Leveraging recent work on SDR, the Max-SDR phase estimator provides a phase error estimate with a worst-case approximation bound compared to the solution set of MMLEs (i.e., worst-case suboptimality for a feasible point to the NP-hard GPGA phase estimation problem). Additionally, a specialized IPM was presented for more efficiently performing Max-SDR phase estimation by exploiting low-rank structure typically associated with the GPGA phase estimation problem. When the size of the data involved with the phase estimation problem is large, which is typical for SAR imaging applications, the computational

complexity and run-time are reduced by roughly N^2 and two orders of magnitude, respectively, as compared to a generic IPM. Hence, the presented specialized IPM makes Max-SDR phase estimation more practical for SAR imaging applications leveraging the GPGA algorithm.

Appendix A. Marginal Likelihood

Assume N complex-valued observations, $\mathbf{y} \in \mathbb{C}^N$, are made of an affine data model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}, \quad (117)$$

where $\mathbf{A} \in \mathbb{C}^{N \times N}$, $\mathbf{x}, \mathbf{b} \in \mathbb{C}^N$, and each of the parameters is assumed to be independent with respect to one another. Moreover, assume $\mathbf{A}$ is arbitrarily distributed and the remaining parameters in the data model are distributed as

$$\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\mathbf{x}}(\boldsymbol{\eta})), \quad (118)$$

$$\mathbf{b} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{b}}(\boldsymbol{\eta}), \mathbf{C}_{\mathbf{b}}(\boldsymbol{\eta})), \quad (119)$$

where $\boldsymbol{\mu}_{\mathbf{b}} : \mathbb{C}^P \rightarrow \mathbb{C}^N$ returns the mean vector for $\mathbf{b}$, $\mathbf{C}_{\mathbf{x}}, \mathbf{C}_{\mathbf{b}} : \mathbb{C}^P \rightarrow \mathbb{C}^{N \times N}$ return the covariance matrix for $\mathbf{x}$ and $\mathbf{b}$, respectively, $\boldsymbol{\eta} \in \mathbb{C}^P$ is a vector containing all hyperparameters (i.e., distribution parameters), and P is the number of hyperparameters. Given the descriptions above, the data observations, given all parameters except $\mathbf{b}$, are distributed as

$$\mathbf{y} | \mathbf{A}, \mathbf{x}, \boldsymbol{\eta} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{y}}(\mathbf{A}\mathbf{x}, \boldsymbol{\eta}), \mathbf{C}_{\mathbf{b}}(\boldsymbol{\eta})), \quad (120)$$

where $\boldsymbol{\mu}_{\mathbf{y}} : \mathbb{C}^N \times \mathbb{C}^P \rightarrow \mathbb{C}^N$, $\boldsymbol{\mu}_{\mathbf{y}}(\mathbf{A}\mathbf{x}, \boldsymbol{\eta}) = \mathbf{A}\mathbf{x} + \boldsymbol{\mu}_{\mathbf{b}}(\boldsymbol{\eta})$. In the proceeding it is shown that the marginal likelihood of $\mathbf{y}$, marginalizing over $\mathbf{x}$, is distributed as

$$\mathbf{y} | \mathbf{A}, \boldsymbol{\eta} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{b}}(\boldsymbol{\eta}), \mathbf{A}\mathbf{C}_{\mathbf{x}}(\boldsymbol{\eta})\mathbf{A}^H + \mathbf{C}_{\mathbf{b}}(\boldsymbol{\eta})), \quad (121)$$

assuming both $\mathbf{C}_{\mathbf{x}}(\boldsymbol{\eta})$ and $\mathbf{C}_{\mathbf{b}}(\boldsymbol{\eta})$ are full rank (hence, invertible).

The marginal likelihood, marginalizing over $\mathbf{x}$, is

$$\begin{aligned}\Pr(\mathbf{y} \mid \mathbf{A}, \boldsymbol{\eta}) &= \int \Pr(\mathbf{y} \mid \mathbf{A}, \mathbf{x}, \boldsymbol{\eta}) \Pr(\mathbf{x} \mid \mathbf{A}, \boldsymbol{\eta}) \, d\mathbf{x} \\ &= \int \Pr(\mathbf{y} \mid \mathbf{A}, \mathbf{x}, \boldsymbol{\eta}) \Pr(\mathbf{x}) \, d\mathbf{x},\end{aligned}\tag{122}$$

recalling each of the parameters are assumed to be independent with respect to one another. Using (118) and (120), and letting $\mathbf{z} : \mathbb{C}^N \times \mathbb{C}^P \rightarrow \mathbb{C}^N$, $\mathbf{z}(\mathbf{y}, \boldsymbol{\eta}) = \mathbf{y} - \boldsymbol{\mu}_{\mathbf{b}}(\boldsymbol{\eta})$, the marginal likelihood is (omitting arguments to functions to simplify notation)

$$\begin{aligned}\Pr(\mathbf{y} \mid \mathbf{A}, \boldsymbol{\eta}) &= \int \frac{1}{\pi^N |\mathbf{C}_{\mathbf{b}}|} e^{-(\mathbf{z} - \mathbf{A}\mathbf{x})^H \mathbf{C}_{\mathbf{b}}^{-1} (\mathbf{z} - \mathbf{A}\mathbf{x})} \frac{1}{\pi^N |\mathbf{C}_{\mathbf{x}}|} e^{-\mathbf{x}^H \mathbf{C}_{\mathbf{x}}^{-1} \mathbf{x}} \, d\mathbf{x} \\ &= \frac{e^{-\mathbf{z}^H \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{z}}}{\pi^N |\mathbf{C}_{\mathbf{b}}| |\mathbf{C}_{\mathbf{x}}|} \int \frac{1}{\pi^N} e^{-\mathbf{x}^H (\mathbf{C}_{\mathbf{x}}^{-1} + \mathbf{A}^H \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{A}) \mathbf{x}} e^{2\text{Re}\{\mathbf{z}^H \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{A} \mathbf{x}\}} \, d\mathbf{x}.\end{aligned}\tag{123}$$

Letting $\mathbf{C}_1^{-1}(\boldsymbol{\eta}) = \mathbf{C}_{\mathbf{x}}^{-1}(\boldsymbol{\eta}) + \mathbf{A}^H \mathbf{C}_{\mathbf{b}}^{-1}(\boldsymbol{\eta}) \mathbf{A}$ and $\tilde{\mathbf{y}}(\boldsymbol{\eta}) = \mathbf{A}^H \mathbf{C}_{\mathbf{b}}^{-H}(\boldsymbol{\eta}) \mathbf{z}$, the result in (123) becomes

$$\begin{aligned}\Pr(\mathbf{y} \mid \mathbf{A}, \boldsymbol{\eta}) &= \frac{|\mathbf{C}_1|}{\pi^N |\mathbf{C}_{\mathbf{b}}| |\mathbf{C}_{\mathbf{x}}|} e^{-\mathbf{z}^H \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{z}} e^{\tilde{\mathbf{y}}^H \mathbf{C}_1 \tilde{\mathbf{y}}} \int \frac{1}{\pi^N |\mathbf{C}_1|} e^{-(\mathbf{x} - \mathbf{C}_1 \tilde{\mathbf{y}})^H \mathbf{C}_1^{-1} (\mathbf{x} - \mathbf{C}_1 \tilde{\mathbf{y}})} \, d\mathbf{x} \\ &= \frac{|\mathbf{C}_1|}{\pi^N |\mathbf{C}_{\mathbf{b}}| |\mathbf{C}_{\mathbf{x}}|} e^{-\mathbf{z}^H \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{z}} e^{\tilde{\mathbf{y}}^H \mathbf{C}_1 \tilde{\mathbf{y}}}.\end{aligned}\tag{124}$$

Substituting back in for $\tilde{\mathbf{y}}$ and letting $\mathbf{C}_2^{-1}(\boldsymbol{\eta}) = \mathbf{C}_{\mathbf{b}}^{-1}(\boldsymbol{\eta}) - \mathbf{C}_{\mathbf{b}}^{-1}(\boldsymbol{\eta}) \mathbf{A} \mathbf{C}_1(\boldsymbol{\eta}) \mathbf{A}^H \mathbf{C}_{\mathbf{b}}^{-H}(\boldsymbol{\eta})$ produces

$$\begin{aligned}\Pr(\mathbf{y} \mid \mathbf{A}, \boldsymbol{\eta}) &= \frac{|\mathbf{C}_1|}{\pi^N |\mathbf{C}_{\mathbf{b}}| |\mathbf{C}_{\mathbf{x}}|} e^{-\mathbf{z}^H \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{z}} e^{\mathbf{z}^H \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{A} \mathbf{C}_1 \mathbf{A}^H \mathbf{C}_{\mathbf{b}}^{-H} \mathbf{z}} \\ &= \frac{1}{\pi^N |\mathbf{C}_{\mathbf{b}}| |\mathbf{C}_{\mathbf{x}}| |\mathbf{C}_1^{-1}|} e^{-\mathbf{z}^H \mathbf{C}_2^{-1} \mathbf{z}}.\end{aligned}\tag{125}$$

Using the Woodbury matrix identity it is seen that

$$\mathbf{C}_2^{-1}(\boldsymbol{\eta}) = \left(\mathbf{A}\mathbf{C}_x(\boldsymbol{\eta})\mathbf{A}^H + \mathbf{C}_b(\boldsymbol{\eta}) \right)^{-1}, \quad (126)$$

implying $\mathbf{C}_2(\boldsymbol{\eta}) = \mathbf{A}\mathbf{C}_x(\boldsymbol{\eta})\mathbf{A}^H + \mathbf{C}_b(\boldsymbol{\eta})$. Using the matrix determinant lemma it is seen that

$$|\mathbf{C}_2(\boldsymbol{\eta})| = |\mathbf{C}_1^{-1}(\boldsymbol{\eta})| |\mathbf{C}_x(\boldsymbol{\eta})| |\mathbf{C}_b(\boldsymbol{\eta})|. \quad (127)$$

Using (126) and (127), and substituting in the definition for $\mathbf{z}$ in (125) produces

$$\begin{aligned} \Pr(\mathbf{y} \mid \mathbf{A}, \boldsymbol{\eta}) &= \frac{1}{\pi^N |\mathbf{C}_2|} e^{-(\mathbf{y}-\boldsymbol{\mu}_b)^H \mathbf{C}_2^{-1} (\mathbf{y}-\boldsymbol{\mu}_b)} \\ &= \mathcal{CN}(\boldsymbol{\mu}_b, \mathbf{A}\mathbf{C}_x\mathbf{A}^H + \mathbf{C}_b), \end{aligned} \quad (128)$$

which is equivalent to the marginal likelihood in (121).

Appendix B. Fisher Information Matrix

From (83) in Section 4.7, the marginal likelihood for the phase compensated range profile samples for the i th selected scatterer, having marginalized over the unknown complex-valued reflectivities γ_i , is

$$\tilde{\xi}_i | \delta\phi, \eta \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_i(\delta\phi, \eta)), \quad (129)$$

where $\eta \in \mathbb{C}^P$ is the vector of hyperparameters (i.e., vector containing all $\sigma_{\gamma,i}^2$) and $\mathbf{C}_i : \mathbb{C}^N \times \mathbb{C}^P \rightarrow \mathbb{C}^{N \times N}$, $\mathbf{C}_i(\delta\phi, \eta) = \sigma_{\gamma,i}^2 \delta\phi \delta\phi^H + \mathbf{I}$ is the covariance matrix of the phase compensated range profile samples for the i th scatterer. Using the provided model, the FIM for the argument of the phase error vector and vector of hyperparameters, $\mathbf{J} \in \mathbb{C}^{(N-1)P \times (N-1)P}$, is

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\theta,\theta} & \mathbf{J}_{\theta,\eta} \\ \mathbf{J}_{\theta,\eta}^H & \mathbf{J}_{\eta,\eta} \end{bmatrix}, \quad (130)$$

where the phase error vector, in terms of the argument of the phase error vector $\theta \in \mathbb{C}^{N-1}$, is

$$[\delta\phi]_n = \begin{cases} e^{j[\theta]_n}, & n = 0, \dots, N-2, \\ 1, & n = N-1, \end{cases} \quad (131)$$

and

$$[\mathbf{J}_{\theta,\theta}]_{n',n} = -\mathbb{E} \left[\frac{\partial^2 \mathcal{L}(\delta\phi, \eta)}{\partial [\theta]_{n'} \partial [\theta]_n} \right] \in \mathbb{C}^{(N-1) \times (N-1)}, \quad (132)$$

$$[\mathbf{J}_{\theta,\eta}]_{n,p} = -\mathbb{E} \left[\frac{\partial^2 \mathcal{L}(\delta\phi, \eta)}{\partial [\theta]_n \partial [\eta]_p} \right] \in \mathbb{C}^{(N-1) \times P}, \quad (133)$$

$$[\mathbf{J}_{\boldsymbol{\eta}, \boldsymbol{\eta}}]_{p', p} = -\mathbb{E} \left[\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\eta}]_{p'} \partial [\boldsymbol{\eta}]_p} \right] \in \mathbb{C}^{P \times P}. \quad (134)$$

Recall, the estimated phase error vector (i.e., autofocus solution) is independent of a constant. Hence, one value of the phase error vector may be assumed to be known and arbitrarily assigned. For convenience, in (131) the last element of the phase error vector is assumed to be known and equal to one.

In the remaining sections, the blocks of the FIM are derived using the log marginal likelihood of the phase compensated range profile samples

$$\mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta}) \propto -\sum_{i \in \mathcal{T}} \ln(1 + N\sigma_{\gamma, i}^2) + \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma, i}^2}{1 + N\sigma_{\gamma, i}^2} \left| \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} \right|^2, \quad (135)$$

where constants with respect to $\boldsymbol{\delta\phi}$ and $\boldsymbol{\eta}$ have been omitted. To simplify the notation, arguments to the covariance matrix function are omitted for the derivations.

2.1 Argument of the Phase Error Vector

The partial derivative with respect to $[\boldsymbol{\theta}]_n$ is

$$\begin{aligned} \frac{\partial \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\theta}]_n} &= \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma, i}^2}{1 + N\sigma_{\gamma, i}^2} \left(\frac{\partial \boldsymbol{\delta\phi}^H \tilde{\boldsymbol{\xi}}_i}{\partial [\boldsymbol{\theta}]_n} \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} + \boldsymbol{\delta\phi}^H \tilde{\boldsymbol{\xi}}_i \frac{\partial \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi}}{\partial [\boldsymbol{\theta}]_n} \right) \\ &= 2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma, i}^2}{1 + N\sigma_{\gamma, i}^2} \operatorname{Re} \left\{ \frac{\partial \boldsymbol{\delta\phi}^H \tilde{\boldsymbol{\xi}}_i}{\partial [\boldsymbol{\theta}]_n} \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} \right\} \\ &= 2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma, i}^2}{1 + N\sigma_{\gamma, i}^2} \operatorname{Re} \left\{ -j [\boldsymbol{\delta\phi}]_n^* [\tilde{\boldsymbol{\xi}}_i]_n \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} \right\}. \end{aligned} \quad (136)$$

The second order partial derivative with respect to $[\boldsymbol{\theta}]_n$ is

$$\begin{aligned}\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\theta}]_n^2} &= 2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma,i}^2}{1 + N\sigma_{\gamma,i}^2} \text{Re} \left\{ -[\boldsymbol{\delta\phi}]_n^* [\tilde{\boldsymbol{\xi}}_i]_n \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} + [\boldsymbol{\delta\phi}]_n^* [\tilde{\boldsymbol{\xi}}_i]_n [\tilde{\boldsymbol{\xi}}_i]_n^* [\boldsymbol{\delta\phi}]_n \right\} \\ &= 2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma,i}^2}{1 + N\sigma_{\gamma,i}^2} \text{Re} \left\{ -[\boldsymbol{\delta\phi}]_n^* [\tilde{\boldsymbol{\xi}}_i]_n \tilde{\boldsymbol{\xi}}_i^H \boldsymbol{\delta\phi} + \left| [\tilde{\boldsymbol{\xi}}_i]_n \right|^2 \right\}.\end{aligned}\quad (137)$$

Taking the expectation and negating produces

$$\begin{aligned}-\mathbb{E} \left[\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\theta}]_n^2} \right] &= 2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma,i}^2}{1 + N\sigma_{\gamma,i}^2} \text{Re} \left\{ \sum_{p=0}^{N-1} [\boldsymbol{\delta\phi}]_n^* [\mathbf{C}_i]_{n,p} [\boldsymbol{\delta\phi}]_p - (\sigma_{\gamma,i}^2 + 1) \right\} \\ &= 2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma,i}^2}{1 + N\sigma_{\gamma,i}^2} \text{Re} \left\{ \sigma_{\gamma,i}^2 \sum_{\substack{p=0 \\ p \neq n}}^{N-1} [\boldsymbol{\delta\phi}]_n^* [\boldsymbol{\delta\phi}]_n [\boldsymbol{\delta\phi}]_p^* [\boldsymbol{\delta\phi}]_p \right\} \\ &= 2 \sum_{i \in \mathcal{T}} \frac{(\sigma_{\gamma,i}^2)^2}{1 + N\sigma_{\gamma,i}^2} (N-1).\end{aligned}\quad (138)$$

The second order partial derivative with respect to $[\boldsymbol{\theta}]_n$ and $[\boldsymbol{\theta}]_{n'}$, where $n \neq n'$, is

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\theta}]_{n'} \partial [\boldsymbol{\theta}]_n} = 2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma,i}^2}{1 + N\sigma_{\gamma,i}^2} \text{Re} \left\{ [\boldsymbol{\delta\phi}]_n^* [\tilde{\boldsymbol{\xi}}_i]_n [\tilde{\boldsymbol{\xi}}_i]_{n'}^* [\boldsymbol{\delta\phi}]_{n'} \right\}.\quad (139)$$

Taking the expectation and negating produces

$$\begin{aligned}-\mathbb{E} \left[\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\theta}]_{n'} \partial [\boldsymbol{\theta}]_n} \right] &= -2 \sum_{i \in \mathcal{T}} \frac{\sigma_{\gamma,i}^2}{1 + N\sigma_{\gamma,i}^2} \text{Re} \left\{ [\boldsymbol{\delta\phi}]_n^* [\mathbf{C}_i]_{n,n'} [\boldsymbol{\delta\phi}]_{n'} \right\} \\ &= -2 \sum_{i \in \mathcal{T}} \frac{(\sigma_{\gamma,i}^2)^2}{1 + N\sigma_{\gamma,i}^2}.\end{aligned}\quad (140)$$

Combining the above results one obtains

$$\mathbf{J}_{\boldsymbol{\theta}, \boldsymbol{\theta}} = 2 \sum_{i \in \mathcal{T}} \frac{(\sigma_{\gamma,i}^2)^2}{1 + N\sigma_{\gamma,i}^2} (N\mathbf{I} - \mathbf{1}\mathbf{1}^H).\quad (141)$$

2.2 Cross Terms

The second order partial derivative with respect to $[\boldsymbol{\theta}]_n$ and $[\boldsymbol{\eta}]_p$ is

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\eta}]_p \partial [\boldsymbol{\theta}]_n} = \frac{2}{(1 + N\sigma_{\gamma,p}^2)^2} \text{Re} \left\{ -j [\boldsymbol{\delta\phi}]_n^* [\tilde{\boldsymbol{\xi}}_p]_n \tilde{\boldsymbol{\xi}}_p^H \boldsymbol{\delta\phi} \right\}. \quad (142)$$

Taking the expectation and negating produces

$$\begin{aligned} -\mathbb{E} \left[\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\eta}]_p \partial [\boldsymbol{\theta}]_n} \right] &= \frac{2}{(1 + N\sigma_{\gamma,p}^2)^2} \text{Re} \left\{ j \sum_{q=0}^{N-1} [\boldsymbol{\delta\phi}]_n^* [\mathbf{C}_p]_{n,q} [\boldsymbol{\delta\phi}]_q \right\} \\ &= \frac{2}{(1 + N\sigma_{\gamma,p}^2)^2} \text{Re} \left\{ j \left((\sigma_{\gamma,p}^2 + 1) + \sum_{\substack{q=0 \\ q \neq p}}^{N-1} [\boldsymbol{\delta\phi}]_n^* [\mathbf{C}_p]_{n,q} [\boldsymbol{\delta\phi}]_q \right) \right\} \\ &= \frac{2}{(1 + N\sigma_{\gamma,p}^2)^2} \text{Re} \left\{ j ((\sigma_{\gamma,p}^2 + 1) + \sigma_{\gamma,p}^2 (N - 1)) \right\} \\ &= 0, \end{aligned} \quad (143)$$

meaning

$$\mathbf{J}_{\boldsymbol{\theta}, \boldsymbol{\eta}} = \mathbf{0}. \quad (144)$$

2.3 Hyperparameters

The partial derivative with respect to $[\boldsymbol{\eta}]_p$ is

$$\frac{\partial \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\eta}]_p} = -\frac{N}{1 + N\sigma_{\gamma,p}^2} + \frac{1}{(1 + N\sigma_{\gamma,p}^2)^2} \left| \tilde{\boldsymbol{\xi}}_p^H \boldsymbol{\delta\phi} \right|^2. \quad (145)$$

The second order partial derivative with respect to $[\boldsymbol{\eta}]_p$ and $[\boldsymbol{\eta}]_{p'}$ is

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\eta}]_{p'} \partial [\boldsymbol{\eta}]_p} = \begin{cases} \frac{N^2}{(1+N\sigma_{\gamma,p}^2)^2} - \frac{2N}{(1+N\sigma_{\gamma,p}^2)^3} \left| \tilde{\boldsymbol{\xi}}_p^H \boldsymbol{\delta\phi} \right|^2, & p = p', \\ 0, & p \neq p'. \end{cases} \quad (146)$$

Taking the expectation and negating produces

$$\begin{aligned} -\mathbb{E} \left[\frac{\partial^2 \mathcal{L}(\boldsymbol{\delta\phi}, \boldsymbol{\eta})}{\partial [\boldsymbol{\eta}]_{p'} \partial [\boldsymbol{\eta}]_p} \right] &= \begin{cases} -\frac{N^2}{(1+N\sigma_{\gamma,p}^2)^2} + \frac{2N}{(1+N\sigma_{\gamma,p}^2)^3} \boldsymbol{\delta\phi}^H \mathbf{C}_p \boldsymbol{\delta\phi}, & p = p', \\ 0, & p \neq p', \end{cases} \\ &= \begin{cases} \frac{N^2}{(1+N\sigma_{\gamma,p}^2)^2}, & p = p', \\ 0, & p \neq p'. \end{cases} \end{aligned} \quad (147)$$

Combining the above results one obtains

$$\mathbf{J}_{\boldsymbol{\eta}, \boldsymbol{\eta}} = \text{Diag} \left(\left[\frac{N^2}{(1+N\sigma_{\gamma,0}^2)^2}, \quad \dots, \quad \frac{N^2}{(1+N\sigma_{\gamma,P-1}^2)^2} \right] \right). \quad (148)$$

2.4 Conclusion

Using above derived results, the FIM from (130) simplifies to

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\boldsymbol{\theta}, \boldsymbol{\theta}} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\boldsymbol{\eta}, \boldsymbol{\eta}} \end{bmatrix}. \quad (149)$$

Thus, the inverse of the FIM, which is used for the CRLB, is obtained by inverting the blocks of the FIM. Using the Sherman-Morrison formula, the inverse of $\mathbf{J}_{\boldsymbol{\theta}, \boldsymbol{\theta}}$ is

$$\mathbf{J}_{\boldsymbol{\theta}, \boldsymbol{\theta}}^{-1} = \frac{1}{2N \sum_{i \in \mathcal{T}} \frac{(\sigma_{\gamma,i}^2)^2}{1+N\sigma_{\gamma,i}^2}} (\mathbf{I} + \mathbf{1}\mathbf{1}^H). \quad (150)$$

And, the inverse of $\mathbf{J}_{\boldsymbol{\eta},\boldsymbol{\eta}}$ is

$$\mathbf{J}_{\boldsymbol{\eta},\boldsymbol{\eta}}^{-1} = \text{Diag} \left(\left[\frac{(1+N\sigma_{\gamma,0}^2)^2}{N^2}, \quad \dots, \quad \frac{(1+N\sigma_{\gamma,P-1}^2)^2}{N^2} \right] \right). \quad (151)$$

Appendix C. Lagrange Dual Problem

Consider the convex optimization problem

$$\begin{aligned}
& \underset{\mathbf{\Phi} \in \mathbb{H}^N}{\text{maximize}} && \text{Tr}(\mathbf{\Phi}\mathbf{\Xi}), \\
& \text{subject to} && \text{diag}(\mathbf{\Phi}) = \mathbf{c}, \\
& && \mathbf{\Phi} \succeq \mathbf{0},
\end{aligned} \tag{152}$$

where $\mathbf{c} \in \mathbb{R}^N$ and $\mathbf{\Xi} \in \mathbb{H}^N$ are given data. The Lagrangian, $\mathcal{L} : \mathbb{R}^N \times \mathbb{H}^N \times \mathbb{H}^N \rightarrow \mathbb{R}$, for (152) is

$$\begin{aligned}
\mathcal{L}(\mathbf{x}, \mathbf{Y}, \mathbf{\Phi}) &= \text{Tr}(\mathbf{\Phi}\mathbf{\Xi}) + \mathbf{x}^T (\text{diag}(\mathbf{\Phi}) - \mathbf{c}) + \text{Tr}(\mathbf{\Phi}\mathbf{Y}) \\
&= \text{Tr}(\mathbf{\Phi} (\text{Diag}(\mathbf{x}) + \mathbf{\Xi} + \mathbf{Y})) - \mathbf{c}^T \mathbf{x},
\end{aligned} \tag{153}$$

where $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{Y} \in \mathbb{H}^N$, $\mathbf{Y} \succeq \mathbf{0}$ are dual variables introduced for the equality and inequality constraints in (152), respectively. The Lagrange dual function, $g : \mathbb{R}^N \times \mathbb{H}^N \rightarrow \mathbb{R}$, for (152) is

$$\begin{aligned}
g(\mathbf{x}, \mathbf{Y}) &= \sup_{\mathbf{\Phi} \in \mathbb{H}^N} \{\mathcal{L}(\mathbf{x}, \mathbf{Y}, \mathbf{\Phi})\} \\
&= \sup_{\mathbf{\Phi} \in \mathbb{H}^N} \{\text{Tr}(\mathbf{\Phi} (\text{Diag}(\mathbf{x}) + \mathbf{\Xi} + \mathbf{Y}))\} - \mathbf{c}^T \mathbf{x} \\
&= \begin{cases} +\infty, & \text{Diag}(\mathbf{x}) + \mathbf{\Xi} + \mathbf{Y} \neq \mathbf{0}, \\ -\mathbf{c}^T \mathbf{x}, & \text{Diag}(\mathbf{x}) + \mathbf{\Xi} + \mathbf{Y} = \mathbf{0}, \end{cases}
\end{aligned} \tag{154}$$

where the simplification in the last line follows from a linear function being bounded above only when the function is zero [82]. Using the Lagrange dual function from

(154), the Lagrange dual problem for (152) is

$$\begin{aligned}
& \underset{\mathbf{x} \in \mathbb{R}^N, \mathbf{Y} \in \mathbb{H}^N}{\text{minimize}} && -\mathbf{c}^T \mathbf{x}, \\
& \text{subject to} && \text{Diag}(\mathbf{x}) + \mathbf{\Xi} + \mathbf{Y} = \mathbf{0}, \\
& && \mathbf{Y} \succeq \mathbf{0}.
\end{aligned} \tag{155}$$

Eliminating the slack variable $\mathbf{Y}$, the Lagrange dual problem in (155) may be equivalently expressed as [82]

$$\begin{aligned}
& \underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimize}} && -\mathbf{c}^T \mathbf{x}, \\
& \text{subject to} && \text{Diag}(\mathbf{x}) + \mathbf{\Xi} \preceq \mathbf{0}.
\end{aligned} \tag{156}$$

Appendix D. Logarithmic Barrier Gradient and Hessian

The logarithmic barrier function, $\psi : \mathbb{R}^N \rightarrow \mathbb{R}$, for the DSDR problem in (105) is

$$\begin{aligned}\psi(\mathbf{x}) &= -\ln(|-\mathbf{G}(\mathbf{x})|) \\ &= -\ln(|-\text{Diag}(\mathbf{x}) - \mathbf{\Xi}|),\end{aligned}\tag{157}$$

where $\mathbf{G} : \mathbb{R}^N \rightarrow \mathbb{H}^N$, $\mathbf{G}(\mathbf{x}) = \text{Diag}(\mathbf{x}) + \mathbf{\Xi}$, $\mathbf{x} \in \mathbb{R}^N$ is the dual variable, and $\mathbf{\Xi} \in \mathbb{H}^N$ is a given data matrix. Using (157), Wirtinger calculus [85–89], and the chain rule, the gradient of the logarithmic barrier function is

$$\begin{aligned}[\nabla_{\mathbf{x}}(\psi(\mathbf{x}))]_{\ell} &= -\frac{\partial \ln(|-\mathbf{G}(\mathbf{x})|)}{\partial [\mathbf{x}]_{\ell}} \\ &= \frac{-1}{|-\mathbf{G}(\mathbf{x})|} \frac{\partial |-\mathbf{G}(\mathbf{x})|}{\partial [\mathbf{G}^*(\mathbf{x})]_{m,n}} \frac{\partial [\mathbf{G}^*(\mathbf{x})]_{m,n}}{\partial [\mathbf{x}]_{\ell}}.\end{aligned}\tag{158}$$

Recognizing

$$\frac{\partial [\mathbf{G}^*(\mathbf{x})]_{m,n}}{\partial [\mathbf{x}]_{\ell}} = \begin{cases} 1, & m, n = \ell, \\ 0, & \text{else,} \end{cases}\tag{159}$$

and that the diagonal elements of $\mathbf{G}(\mathbf{x})$ are purely real, the gradient of the logarithmic barrier function, using (42) from [99], is

$$\begin{aligned}[\nabla_{\mathbf{x}}(\psi(\mathbf{x}))]_{\ell} &= \frac{-1}{|-\mathbf{G}(\mathbf{x})|} \frac{\partial |-\mathbf{G}(\mathbf{x})|}{\partial [\mathbf{G}(\mathbf{x})]_{\ell,\ell}} \\ &= \frac{-|-\mathbf{G}(\mathbf{x})|}{|-\mathbf{G}(\mathbf{x})|} \text{Tr}(\mathbf{G}^{-1}(\mathbf{x}) \mathbf{e}_{\ell} \mathbf{e}_{\ell}^T) \\ &= -[\mathbf{G}^{-1}(\mathbf{x})]_{\ell,\ell},\end{aligned}\tag{160}$$

where $\mathbf{e}_\ell \in \mathbb{R}^N$ is the ℓ th standard basis vector (i.e., vector with one as the ℓ th element and zero elsewhere).

Using (160) and the chain rule, the Hessian of the logarithmic barrier function is

$$\begin{aligned} [\nabla_{\mathbf{x}}^2 (\psi(\mathbf{x}))]_{k,\ell} &= - \frac{\partial \ln (|-\mathbf{G}(\mathbf{x})|)}{\partial [\mathbf{x}]_k \partial [\mathbf{x}]_\ell} \\ &= - \frac{\partial [\mathbf{G}^{-1}(\mathbf{x})]_{\ell,\ell}}{\partial [\mathbf{G}^{-1}(\mathbf{x})]_{m,n}^*} \frac{\partial [\mathbf{G}^{-1}(\mathbf{x})]_{m,n}^*}{\partial [\mathbf{x}]_k}. \end{aligned} \quad (161)$$

Recognizing

$$\frac{\partial [\mathbf{G}^{-1}(\mathbf{x})]_{m,n}^*}{\partial [\mathbf{x}]_k} = \begin{cases} 1, & m, n = k, \\ 0, & \text{else,} \end{cases} \quad (162)$$

and that the inverse of $\mathbf{G}(\mathbf{x})$ is conjugate symmetric, the Hessian of the logarithmic barrier function, using (60) from [99], is

$$\begin{aligned} [\nabla_{\mathbf{x}}^2 (\psi(\mathbf{x}))]_{k,\ell} &= - \frac{\partial [\mathbf{G}^{-1}(\mathbf{x})]_{\ell,\ell}}{\partial [\mathbf{G}^{-1}(\mathbf{x})]_{k,k}} \\ &= [\mathbf{G}^{-1}(\mathbf{x})]_{\ell,k} [\mathbf{G}^{-1}(\mathbf{x})]_{k,\ell} \\ &= \left| [\mathbf{G}^{-1}(\mathbf{x})]_{k,\ell} \right|^2. \end{aligned} \quad (163)$$

Bibliography

- [1] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar Systems and Signal Processing*. New York, NY: John Wiley & Sons, 1991.
- [2] W. G. Carrara, R. S. Goodman, and R. M. Majewski, *Spotlight Synthetic Aperture Radar Signal Processing Algorithms*. Boston, MA: Artech House, 1995.
- [3] C. V. Jakowatz, Jr., D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. Boston, MA: Kluwer Academic Publishers, 1996.
- [4] G. G. Kuperman, “Subjective Assessment of SAR Imagery Enhancement Algorithms,” Crew Systems Directorate: Human Engineering Division, Tech. Rep. September, 1997.
- [5] M. L. Skolnik, *Introduction to Radar Systems*, 3rd ed. New York, NY: The McGraw-Hill Company, 2001.
- [6] N. J. Willis and H. D. Griffiths, *Advanced in Bistatic Radar*. Raleigh, NC: SciTech Publishing, Inc., 2007.
- [7] M. Cheney and B. Borden, *Fundamentals of Radar Imaging*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2008.
- [8] ———, “Problems in Synthetic-Aperture Radar Imaging,” in *Inverse Problems*, vol. 25, no. 12, 2009, p. 123005.
- [9] A. Moreira, P. Prats-Iraola, M. Younis, K. Gerhard, I. Hajnsek, and K. P. Papathanassiou, “A Tutorial on Synthetic Aperture Radar,” in *IEEE Geoscience*

and Remote Sensing Magazine, vol. 1, no. 1, 2013, pp. 6–43.

- [10] C. W. Sherwin, J. P. Ruina, and R. D. Rawcliffe, “Some Early Developments in Synthetic Aperture Radar Systems,” in *IRE Transactions on Military Electronics*, vol. MIL-6, no. 2, 1962, pp. 111–115.
- [11] C. A. Wiley, “Pulsed Doppler Radar Methods and Apparatus,” 1965.
- [12] M. Soumekh, *Synthetic Aperture Radar Signal Processing with MATLAB Algorithms*. New York, NY: John Wiley & Sons, 1999.
- [13] B. D. Rigling and R. L. Moses, “Motion Measurement Errors and Autofocus in Bistatic SAR,” in *IEEE Transactions on Image Processing*, vol. 15, no. 4, 2006, pp. 1008–1016.
- [14] R. M. Naething and R. D. West, “Analysis of SAR Autofocus Performance,” in *Proceedings of SPIE*, vol. 8714, 2013, p. 871414.
- [15] R. G. Paxman and J. C. Marron, “Aberration Correction of Speckled Imagery with an Image-Sharpness Criterion,” in *Proceedings of SPIE*, vol. 976, 1988, pp. 37 – 47.
- [16] P. H. Eichel, D. C. Ghiglia, and C. V. Jakowatz, Jr., “Speckle Processing Method for Synthetic-Aperture-Radar Phase Correction,” in *Optics Letters*, vol. 14, no. 1, 1989, pp. 1–3.
- [17] P. H. Eichel and C. V. Jakowatz, Jr., “Phase-Gradient Algorithm as an Optimal Estimator of the Phase Derivative,” in *Optics Letters*, vol. 14, no. 20, 1989, pp. 1101–1103.

- [18] D. C. Ghiglia and G. A. Mastin, “Two-Dimensional Phase Correction of Synthetic-Aperture-Radar Imagery,” in *Optics Letters*, vol. 14, no. 20, 1989, pp. 1104–1106.
- [19] C. V. Jakowatz, Jr. and D. E. Wahl, “Eigenvector Method for Maximum-Likelihood Estimation of Phase Errors in Synthetic-Aperture-Radar Imagery,” in *Journal of the Optical Society of America*, vol. 10, no. 12, 1993, pp. 2539–2546.
- [20] D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and C. V. Jakowatz, Jr., “Phase Gradient Autofocus – A Robust Tool for High Resolution SAR Phase Correction,” in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, 1994, pp. 827–835.
- [21] D. E. Wahl, C. V. Jakowatz, Jr., P. A. Thompson, and D. C. Ghiglia, “New Approach to Strip-Map SAR Autofocus,” in *Proceedings of IEEE 6th Digital Signal Processing Workshop*, 1994, pp. 53–56.
- [22] J. S. Bates, “The Phase Gradient Autofocus Algorithm with Range Dependent Stripmap SAR,” Ph.D. dissertation, Brigham Young University, 1998.
- [23] H. L. Chan and T. S. Yeo, “Noniterative Quality Phase-Gradient Autofocus (QPGA) Algorithm for Spotlight SAR Imagery,” in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, 1998, pp. 1531–1539.
- [24] D. G. Thompson, J. S. Bates, D. V. Arnold, D. G. Long, and A. Robertson, “Range Dependent Phase Gradient Autofocus,” in *IEEE International Geoscience and Remote Sensing Symposium*, vol. 5, 1998, pp. 2634–2636.
- [25] A. F. Yegulalp, “Minimum Entropy SAR Autofocus,” in *Adaptive Sensor Array Processing (ASAP) Workshop*. MIT Lincoln Laboratory, 1999.

- [26] J. R. Fienup, “Synthetic-Aperture Radar Autofocus by Maximizing Sharpness,” in *Optics Letters*, vol. 25, no. 4, 2000, pp. 221–223.
- [27] D. W. Warner, D. C. Ghiglia, A. Fitzgerald, and J. Beaver, “Two-Dimensional Phase Gradient Autofocus,” in *Proceedings of SPIE*, vol. 4123, 2000, pp. 162–173.
- [28] J. R. Fienup and J. J. Miller, “Generalized Image Sharpness Metrics for Correcting Phase Errors,” in *Journal of the Optical Society of America*, 2001, pp. 84–86.
- [29] H. J. Callow, M. P. Hayes, and P. T. Gough, “Stripmap Phase Gradient Autofocus,” in *Oceans 2003. Celebrating the Past ... Teaming Toward the Future*, vol. 5, 2003, pp. 2414–2421.
- [30] J. R. Fienup and J. J. Miller, “Aberration Correction by Maximizing Generalized Sharpness Metrics,” in *Journal of the Optical Society of America*, vol. 20, no. 4, 2003, pp. 609–620.
- [31] J. Kolman, “PACE: An Autofocus Algorithm for SAR,” in *IEEE International Radar Conference*, 2005, pp. 310–314.
- [32] R. L. Morrison and M. N. Do, “A Multichannel Approach to Metric-Based SAR Autofocus,” in *IEEE International Conference on Image Processing*, vol. 2, 2005, pp. 1067–1070.
- [33] T. J. Kragh, “Monotonic Iterative Algorithm for Minimum-Entropy Autofocus,” in *Adaptive Sensor Array Processing (ASAP) Workshop*. MIT Lincoln Laboratory, 2006.

- [34] —, “Presentation on Monotonic iterative Algorithm for Minimum-Entropy Autofocus,” 2006.
- [35] R. L. Morrison and M. N. Do, “Multichannel Autofocus Algorithm for Synthetic Aperture Radar,” in *IEEE International Conference on Image Processing*, 2006, pp. 2341–2344.
- [36] W. L. Van Rossum, M. P. G. Otten, and R. J. P. Van Bree, “Extended PGA for Range Migration Algorithms,” in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 2, 2006, pp. 478–488.
- [37] M. Martorella, M. Preiss, B. Haywood, and B. Bates, “Polarimetric Phase Gradient Autofocus,” in *International Geoscience and Remote Sensing Symposium*, 2007, pp. 133–136.
- [38] R. L. Morrison, M. N. Do, and D. C. Munson, “SAR Image Autofocus by Sharpness Optimization: A Theoretical Study,” in *IEEE Transactions on Image Processing*, vol. 16, no. 9, 2007, pp. 2309–2321.
- [39] T. J. Schulz, “Optimal Sharpness Function for SAR Autofocus,” in *IEEE Signal Processing Letters*, vol. 14, no. 1, 2007, pp. 27–30.
- [40] H. J. Cho and D. C. Munson, “Overcoming Polar-Format Issues in Multichannel SAR Autofocus,” in *Asilomar Conference on Signals, Systems and Computers*, 2008, pp. 523–527.
- [41] C. V. Jakowatz, Jr., D. E. Wahl, and D. A. Yocky, “Beamforming as a Foundation for Spotlight-Mode SAR Image Formation by Backprojection,” in *Proceedings of SPIE*, vol. 6970, 2008, p. 697015.

- [42] K.-H. Liu and D. C. Munson, “Fourier-Domain Multichannel Autofocus for Synthetic Aperture Radar,” in *Asilomar Conference on Signals, Systems and Computers*, 2008, pp. 848–852.
- [43] K.-H. Liu and D. C. Munson, Jr., “Autofocus in Multistatic Passive SAR Imaging,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 1277–1280.
- [44] C. V. Jakowatz, Jr. and D. E. Wahl, “Considerations for Autofocus of Spotlight-Mode SAR Imagery Created Using a Beamforming Algorithm,” in *Proceedings of SPIE*, vol. 7337, 2009, p. 73379.
- [45] R. L. Morrison, M. N. Do, and D. C. Munson, “MCA: A Multichannel Approach to SAR Autofocus,” in *IEEE Transactions on Image Processing*, vol. 18, no. 4, 2009, pp. 840–853.
- [46] K.-H. Liu, A. Wiesel, and D. C. Munson, “Synthetic Aperture Radar Autofocus via Semidefinite Relaxation,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 1342–1345.
- [47] H. J. Cho and D. C. Munson, Jr., “CBP-Based Multichannel Autofocus for Near-Field SAR Imaging,” in *Proceedings of SPIE*, vol. 8051, 2011, p. 80519.
- [48] K.-H. Liu and D. C. Munson, “Fourier-Domain Multichannel Autofocus for Synthetic Aperture Radar,” in *IEEE Transactions on Image Processing*, vol. 20, no. 12, 2011, pp. 3544–3552.
- [49] K.-H. Liu, A. Wiesel, and D. C. Munson, “Maximum Likelihood SAR Autofocus with Low-Return Region,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011, pp. 1377–1380.

- [50] J. N. Ash, “An Autofocus Method for Backprojection Imagery in Synthetic Aperture Radar,” in *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 1, 2012, pp. 104–108.
- [51] K.-H. Liu, A. Wiesel, and D. C. Munson, “Synthetic Aperture Radar Autofocus Based on a Bilinear Model,” in *IEEE Transactions on Image Processing*, vol. 21, no. 5, 2012, pp. 2735–2746.
- [52] R. D. West, J. Gunther, and T. Moon, “Convex Optimization of Subspace Fitting Autofocus for Stripmap SAR Images,” in *IEEE Radar Conference*, 2012, pp. 0731–0734.
- [53] S. Zhou, M. Xing, X. G. Xia, Y. Li, L. Zhang, and Z. Bao, “An Azimuth-Dependent Phase Gradient Autofocus (APGA) Algorithm for Airborne/Stationary BiSAR Imagery,” in *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 6, 2013, pp. 1290–1294.
- [54] M. I. Duersch, “Backprojection for Synthetic Aperture Radar,” Ph.D. dissertation, Brigham Young University, 2013.
- [55] K.-H. Liu, A. Wiesel, and D. C. Munson, “Synthetic Aperture Radar Autofocus via Semidefinite Relaxation,” in *IEEE Transactions on Image Processing*, vol. 22, no. 6, 2013, pp. 2317–2326.
- [56] K. Hu, X. Zhang, S. He, H. Zhao, and J. Shi, “A Less-Memory and High-Efficiency Autofocus Back Projection Algorithm for SAR Imaging,” in *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 4, 2015, pp. 890–894.
- [57] S. Wei, X. Zhang, K. Hu, and W. Wu, “LASAR Autofocus Imaging Using Maximum Sharpness Back Projection via Semidefinite Programming,” in *IEEE Radar*

Conference, 2015, pp. 1311–1315.

- [58] S. Chen, F. Lu, J. Wang, and M. Liu, “An Improved Phase Gradient Autofocus Method for One-Stationary Bistatic SAR,” in *IEEE International Conference on Signal Processing, Communications and Computing, Conference*, 2016.
- [59] H. J. Cho, “Autofocus and Back-Projection in Synthetic Aperture Radar Imaging,” Ph.D. dissertation, University of Michigan, 2016.
- [60] A. Evers, N. Zimmerman, and J. A. Jackson, “Semidefinite Relaxation Autofocus for Bistatic Backprojection SAR,” in *IEEE Radar Conference*, 2017, pp. 1332–1337.
- [61] Y. Li, J. Wu, W. Pu, J. Yang, Y. Huang, W. Li, H. Yang, and W. Huo, “An Autofocus Method Based on Maximum Image Sharpness for Fast Factorized Back-Projection,” in *IEEE Radar Conference*, 2017, pp. 1201–1204.
- [62] X. Wang, Q. Wu, and J. Yang, “Extended PGA Processing of High Resolution Airborne SAR Imagery Reconstructed via Backprojection Algorithm,” in *IEEE International Radar Conference*, 2016, pp. 1–3.
- [63] S. Wei, L. Zhou, X. Zhang, and J. Shi, “Fast Back-Projection Autofocus for Linear Array SAR 3-D Imaging via Maximum Sharpness,” in *IEEE Radar Conference*. IEEE, 2018, pp. 525–530.
- [64] A. Doerry, “Autofocus Correction of Excessive Migration in Synthetic Aperture Radar Images,” Sandia National Laboratories, Tech. Rep., 2004.
- [65] J. M. Kantor, “Minimum Entropy Autofocus Correction of Residual Range Cell Migration,” in *IEEE Radar Conference*, 2017, pp. 11–16.

- [66] A. Evers and J. A. Jackson, “SAR Scene Size Limits for Spatially Invariant Defocus with Near- and Far-Field Geometries,” in *IEEE Transactions on Computational Imaging*, vol. 4, no. 4, 2018, pp. 640–647.
- [67] —, “A Generalized Phase Gradient Autofocus Algorithm,” in *IEEE Transactions on Computational Imaging*, Accepted Jan. 2019.
- [68] —, “Generalized Phase Gradient Autofocus Using Semidefinite Relaxation Phase Estimation,” in *IEEE Transactions on Computational Imaging*, Submitted Feb. 2019.
- [69] S. Zhang and Y. Huang, “Complex Quadratic Optimization and Semidefinite Programming,” in *SIAM Journal on Optimization*, vol. 16, no. 3, 2006, pp. 871–890.
- [70] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, “Semidefinite Relaxation of Quadratic Optimization Problems,” in *IEEE Signal Processing Magazine*, vol. 27, no. 3, 2010, pp. 20–34.
- [71] J. M. Harris, J. L. Hirst, and M. J. Mossinghoff, *Combinatorics and Graph Theory*. New York, NY: Springer Science+Business Media, LLC, 2008.
- [72] S. D. Fisher, *Complex Variables*. Mineola, NY: Dover Publications, Inc., 1990.
- [73] M. Weiss, “Analysis of Some Modified Cell-Averaging CFAR Processors in Multiple-Target Situations,” in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 18, no. 1, 1982, pp. 102–114.
- [74] H. Rohling, “Radar CFAR Thresholding in Clutter and Multiple Target Situations,” in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 19,

no. 4, 1983, pp. 608–621.

- [75] P. P. Gandhi and S. A. Kassam, “Analysis of CFAR Processors in Nonhomogeneous Background,” in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 4, 1988, pp. 427–445.
- [76] N. Levanon, “Detection Loss Due to Interfering Targets in Ordered Statistics CFAR,” in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 6, 1988, pp. 678–681.
- [77] J. B. Keller, “Geometrical Theory of Diffraction,” in *Journal of the Optical Society of America*, vol. 52, no. 2, 1962, pp. 116–130.
- [78] L. C. Potter and R. L. Moses, “Attributed Scattering Centers for SAR ATR,” in *IEEE Transactions on Image Processing*, vol. 6, no. 1, 1997, pp. 79–91.
- [79] R. Zadeh, “Lecture Notes on Singular Value Decomposition,” 2017. [Online]. Available: [https://stanford.edu/{~}rezab/classes/cme323/S17/notes/lecture17/cme323_{_}lec17.pdf](https://stanford.edu/~rezab/classes/cme323/S17/notes/lecture17/cme323_{_}lec17.pdf)
- [80] A. Singer, “Angular Synchronization by Eigenvectors and Semidefinite Programming,” in *Applied and Computational Harmonic Analysis*, vol. 30. Elsevier Inc., 2011, pp. 20–36.
- [81] L. Vandenberghe and S. Boyd, “Semidefinite Programming,” in *Society for Industrial and Applied Mathematics*, vol. 38, no. 1, 1996, pp. 49–95.
- [82] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, NY: Cambridge University Press, 2004.

- [83] E. De Klerk, “Exploiting Special Structure in Semidefinite Programming: A Survey of Theory and Applications,” in *European Journal of Operational Research*, vol. 201, no. 1, 2009, pp. 1–10.
- [84] F. Jarre, “An Interior-Point Method for Minimizing the Maximum Eigenvalue of a Linear Combination of Matrices,” in *Society for Industrial and Applied Mathematics*, vol. 31, no. 5, 1993, pp. 1360–1377.
- [85] R. F. H. Fischer, “Wirtinger Calculus,” in *Precoding and Signal Shaping for Digital Transmission*. New York, NY: John Wiley & Sons, 2002, ch. Appendix A, pp. 405–413.
- [86] A. Hjørungnes and D. Gesbert, “Complex-Valued Matrix Differentiation: Techniques and Key Results,” in *IEEE Transactions on Signal Processing*, vol. 55, no. 6, 2007, pp. 2740–2746.
- [87] K. Kreutz-Delgado, “The Complex Gradient Operator and the CR-Calculus,” University of California, Tech. Rep., 2007.
- [88] J. Eriksson, E. Ollila, and V. Koivunen, “Essential Statistics and Tools for Complex Random Variables,” in *IEEE Transactions on Signal Processing*, vol. 58, no. 10, 2010, pp. 5400–5408.
- [89] T. Adah, P. J. Schreier, and L. L. Scharf, “Complex-Valued Signal Processing: The Proper Way to Deal With Impropriety,” in *IEEE Transactions on Signal Processing*, vol. 59, no. 11, 2011, pp. 5101–5125.
- [90] D. P. Bertsekas, A. Nedic, and A. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 2003.

- [91] D. P. Bertsekas, *Convex Optimization Algorithms*. Nashua, NH: Athena Scientific, 2015.
- [92] C. H. Casteel, Jr., L. A. Gorham, M. J. Minardi, S. M. Scarborough, K. D. Naidu, and U. K. Majumder, “A Challenge Problem for 2D/3D Imaging of Targets from a Volumetric Data Set in an Urban Environment,” in *Proceedings of SPIE*, vol. 6568, 2007, p. 65687.
- [93] M. Grant and S. Boyd, “CVX: Matlab Software for Disciplined Convex Programming, version 2.1,” 2014. [Online]. Available: <http://cvxr.com/cvx>
- [94] J. R. Gutierrez, “Passive Synthetic Aperture Radar Imaging Using Commercial OFDM Communication Networks,” Ph.D. dissertation, Air Force Institute of Technology, 2012.
- [95] J. R. Gutierrez and J. A. Jackson, “Collecting and Processing WiMAX Ground Returns for SAR Imaging,” in *IEEE Radar Conference*, 2013, pp. 1–6.
- [96] J. R. Gutierrez, J. A. Jackson, and M. A. Temple, “Receive Signal Processing for OFDM-Based Radar Imaging,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 2775–2779.
- [97] A. Evers and J. A. Jackson, “Experimental Passive SAR Imaging Exploiting LTE, DVB, and DAB Signals,” in *IEEE Radar Conference*, 2014, pp. 680–685.
- [98] J. R. Fienup, “Invariant Error Metrics for Image Reconstruction,” in *Applied Optics*, vol. 36, no. 32, 1997, pp. 8352–8357.
- [99] K. B. Petersen and M. S. Pedersen, “The Matrix Cookbook,” Technical University of Denmark, Tech. Rep., 2012. [Online]. Available: <https://www.math.u Copenhagen.dk/~madsen/MatrixCookbook/>

[//www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf](http://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf)

REPORT DOCUMENTATION PAGE					Form Approved OMB No. 0704-0188	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.						
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.						
1. REPORT DATE (DD-MM-YYYY) 21 Mar 2019		2. REPORT TYPE Dissertation			3. DATES COVERED (From - To) March 2015 - March 2019	
4. TITLE AND SUBTITLE A Generalized Phase Gradient Autofocus Algorithm				5a. CONTRACT NUMBER		
				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Evers, Aaron, S., Mr.				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way Wright-Patterson AFB OH 45433-7765					8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENG-DS-19-M-025	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Intentionally Left Blank					10. SPONSOR/MONITOR'S ACRONYM(S)	
					11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Distribution Statement A. Approved for Public Release; Distribution Unlimited.						
13. SUPPLEMENTARY NOTES This work is declared a work of the U.S. Government and is not subject to copyright protection in the United States.						
14. ABSTRACT In this work, a generalized phase gradient autofocus (GPGA) algorithm is developed which is applicable with both the polar format algorithm (PFA) and backprojection algorithm (BPA). The GPGA algorithm preserves the four crucial signal processing steps comprising the PGA algorithm, while alleviating the constraint of using a single scatterer per range cut for phase error estimation which exists with the PGA algorithm. The GPGA algorithm, whether using the PFA or BPA, yields an approximate maximum marginal likelihood estimate (MMLE) of phase errors having marginalized over unknown complex-valued reflectivities of selected scatterers. Also, in this work a new approximate MMLE, termed the max-semidefinite relaxation (Max-SDR) phase estimator, is proposed for use with the GPGA algorithm. The Max-SDR phase estimator provides a phase error estimate with a worst-case approximation bound compared to the solution set of MMLEs. Additionally, in this work a specialized interior-point method (IPM) is presented for more efficiently performing Max-SDR phase estimation. Lastly, simulation and experimental results are presented.						
15. SUBJECT TERMS Bistatic synthetic aperture radar, synthetic aperture radar autofocus, generalized phase gradient autofocus algorithm, semidefinite relaxation, interior point method.						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Julie A. Jackson, AFIT/ENG	
U	U	U	UU	146	19b. TELEPHONE NUMBER (Include area code) (937) 255-3636 x4678 julie.jackson@afit.edu	