



AFRL-RY-WP-TR-2020-0153

DESIGN AND EVALUATION OF STOCHASTIC PROCESSES AS PHYSICAL RADAR WAVEFORMS

Charles Mohr
University of Kansas

MAY 2020
Final Report

Approved for public release; distribution is unlimited.

See additional restrictions described on inside pages

STINFO COPY

**AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE**

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 this 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-YY) May 2020			2. REPORT TYPE Thesis		3. DATES COVERED (From - To) 16 April 2020 –16 April 2020		
4. TITLE AND SUBTITLE DESIGN AND EVALUATION OF STOCHASTIC PROCESSES AS PHYSICAL RADAR WAVEFORMS			5a. CONTRACT NUMBER FA8650-14-D-1722/N00014-16-C-2029				
			5b. GRANT NUMBER				
			5c. PROGRAM ELEMENT NUMBER N/A				
6. AUTHOR(S) Charles Mohr			5d. PROJECT NUMBER N/A				
			5e. TASK NUMBER N/A				
			5f. WORK UNIT NUMBER Y16A				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Kansas 450 Jayhawk Blvd. Lawrence, KS 66045			8. PERFORMING ORGANIZATION REPORT NUMBER				
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory Sensors Directorate (AFRL/RYMD) Wright-Patterson Air Force Base, OH 45433-7320 Air Force Materiel Command United States Air Force			10. SPONSORING/MONITORING AGENCY ACRONYM(S) AFRL/RYMD				
			11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S) AFRL-RY-WP-TR-2020-0153				
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.							
13. SUPPLEMENTARY NOTES PAO case number 88ABW-2020-1418, Clearance Date 16 April 2020. Submitted to the graduate degree program in Electrical Engineering and Computer Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctorate of Philosophy. This work was funded in whole or in part by Department of the Air Force. The U.S. Government has for itself and others acting on its behalf an unlimited, paid-up, nonexclusive, irrevocable worldwide license to use, modify, reproduce, release, perform, display, or disclose the work by or on behalf of the U.S. Government. Report contains color.							
14. ABSTRACT Recent advances in waveform generation and in computational power have enabled the design and implementation of new complex radar waveforms. Still, even with these advances in computation, in a pulse agile mode, where the radar transmits unique waveforms at every pulse, the requirement to design physically robust waveforms which achieve good autocorrelation sidelobes, are spectrally contained, and have a constant amplitude envelope for high power operation, can require expensive computation equipment and can impede real time operation. This work addresses this concern in the context of FM noise waveforms which have been demonstrated in recent years in both simulation and in experiments to achieve low autocorrelation sidelobes through the high dimensionality of coherent integration when operating in a pulse agile mode. However while they are effective, the approaches to design these waveforms requires the optimization of each individual waveform making them subject to the concern above. This dissertation takes a different approach. Since these FM noise waveforms are meant to be noise like in the first place, the waveforms here are instantiated as the sample functions of a stochastic process which has been specially designed to produce spectrally contained, constant amplitude waveforms with noise like cancellation of sidelobes. This makes the waveform creation process little more computationally expensive than pulling numbers from a random number generator (RNG) since the optimization designs a waveform generating function (WGF) itself rather than each waveform themselves. This goal is achieved by leveraging gradient descent optimization methods to reduce the expected frequency template error (EFTE) cost function for both the pulsed stochastic waveform generation (StoWGe) waveform model and a new CW version of StoWGe denoted CW-StoWGe. The effectiveness of these approaches and their ability to generate useful radar waveforms is analyzed using several stochastic waveform generation metrics developed here. The EFTE optimization is shown through simulation to produce WGFs which generate FM noise waveforms in both pulsed and CW modes which achieve good spectral containment and autocorrelation sidelobes. The resulting waveforms will be demonstrated in both loopback and in open-air experiments to be robust to physical implementation.							
15. SUBJECT TERMS FM noise waveforms, optimization, stochastic waveforms, waveform design, waveform diversity							
16. SECURITY CLASSIFICATION OF:		17. LIMITATION OF ABSTRACT: SAR		18. NUMBER OF PAGES 168		19a. NAME OF RESPONSIBLE PERSON (Monitor) Charles Mohr	
a. REPORT Unclassified		b. ABSTRACT Unclassified		c. THIS PAGE Unclassified		19b. TELEPHONE NUMBER (Include Area Code) N/A	

Design and Evaluation of Stochastic Processes as Physical Radar Waveforms

By

Charles Mohr

Submitted to the graduate degree program in Electrical Engineering and Computer Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctorate of Philosophy.

Dr. Shannon Blunt, Chairperson

Dr. Chris Allen

Committee members

Dr. Jim Stiles

Dr. Carl Leuschen

Dr. Zsolt Talata

Date defended: _____

The Dissertation Committee for Charles Mohr certifies
that this is the approved version of the following dissertation :

Design and Evaluation of Stochastic Processes as Physical Radar Waveforms

Dr. Shannon Blunt, Chairperson

Date approved: _____

Abstract

Recent advances in waveform generation and in computational power have enabled the design and implementation of new complex radar waveforms. Still, even with these advances in computation, in a pulse agile mode, where the radar transmits unique waveforms at every pulse, the requirement to design physically robust waveforms which achieve good autocorrelation sidelobes, are spectrally contained, and have a constant amplitude envelope for high power operation, can require expensive computation equipment and can impede real time operation. This work addresses this concern in the context of FM noise waveforms which have been demonstrated in recent years in both simulation and in experiments to achieve low autocorrelation sidelobes through the high dimensionality of coherent integration when operating in a pulse agile mode. However while they are effective, the approaches to design these waveforms requires the optimization of each individual waveform making them subject to the concern above.

This dissertation takes a different approach. Since these FM noise waveforms are meant to be noise like in the first place, the waveforms here are instantiated as the sample functions of a stochastic process which has been specially designed to produce spectrally contained, constant amplitude waveforms with noise like cancellation of sidelobes. This makes the waveform creation process little more computationally expensive than pulling numbers from a random number generator (RNG) since the optimization designs a waveform generating function (WGF) itself rather than each waveform themselves. This goal is achieved by leveraging gradient descent optimization methods to reduce the expected frequency template error (EFTE) cost function for

both the pulsed stochastic waveform generation (StoWGe) waveform model and a new CW version of StoWGe denoted CW-StoWGe. The effectiveness of these approaches and their ability to generate useful radar waveforms is analyzed using several stochastic waveform generation metrics developed here. The EFTE optimization is shown through simulation to produce WGFs which generate FM noise waveforms in both pulsed and CW modes which achieve good spectral containment and autocorrelation sidelobes. The resulting waveforms will be demonstrated in both loopback and in open-air experiments to be robust to physical implementation.

Acknowledgements

This work was supported by the Office of Naval Research under Contract #N00014-16-C-2029 and by a subcontract with Matrix Research, Inc. for the Air Force Research Laboratory under Prime Contract #FA8650-14-D-1722.

Contents

1	Introduction	1
2	Background	5
2.1	Basic Radar Processing	6
2.1.1	The Unmodulated Pulse and Basic Radar Operation	6
2.1.2	Linear Frequency Modulation and Pulse Compression	8
2.1.3	Pulse Integration and LFM Sidelobe Mitigation	17
2.1.3.1	Pulse Integration	17
2.1.3.2	LFM Sidelobe Mitigation	19
2.2	Pulse Compression Waveform Design	20
2.2.1	Non-linear FM	21
2.2.2	Phased-coded Waveforms	21
2.2.2.1	Phase-coded Waveform Examples	22
2.2.2.2	Phase-coded Waveform Spectral Characteristics	24
2.2.3	Polyphase-coded Frequency Modulation	26
2.2.4	FM Noise Waveforms	31
2.2.4.1	Psuedo-Random Optimized FM (PRO-FM)	32
2.2.4.2	Process Based FM Noise Waveforms	37
2.3	Continuous Wave Radar	37
2.4	The Ambiguity Function and Doppler Processing	39
2.5	Gradient Descent Techniques	39
3	Stochastic Waveform Analysis and Evaluation	41

3.1	Random Variables	41
3.1.1	Covariance and Independence	42
3.1.2	Continuous Random Variables and the Characteristic Function	44
3.2	Random Processes	45
3.2.1	Correlation Functions	46
3.2.2	Stationarity and the Power Spectral Density (PSD)	47
3.2.3	Estimation of the Autocorrelation and the PSD	49
3.2.3.1	Power Spectral Density Estimation	49
3.2.3.2	Autocorrelation Estimation	52
3.3	Stochastic Processes as Radar Signals	53
3.3.1	Pulsed Stochastic Processes as Radar Signals	53
3.3.2	Stochastic Processes as CW Radar Signals	56
3.4	Evaluating Stochastic Radar Waveforms	58
3.4.1	The Discrete Waveform Generating Process	58
3.4.2	Metrics for Stochastic Radar Waveforms	60
3.4.2.1	Aggregate Metrics	60
3.4.2.2	Individual Metrics	62
3.4.2.3	Summary of Stochastic Waveform Measures	63
3.4.3	Stochastic Waveform Moments	63
4	Pulsed Stochastic Waveform Generation (StoWGe)	66
4.1	The Pulsed StoWGe Signal Model	66
4.2	The Expected Frequency Template Error (EFTE) for Pulsed StoWGe	69
4.3	Optimization of the Pulsed StoWGe Expected Power Spectrum	71
4.3.1	The Pulsed StoWGe EFTE Gradient	72
4.3.2	Gradient Descent Implementation	73
4.3.3	Initializations	75
4.4	Optimization Results	82

4.4.1	Small N	83
4.4.1.1	Discrete Distributions	83
4.4.1.2	Continuous Distributions	84
4.4.2	Pulsed StoWGe Evaluation	86
4.4.2.1	Gaussian Template Results	89
4.4.2.2	Super-Gaussian Template Results	92
4.4.2.3	Rectangular Template Results	93
4.4.2.4	DU2 Distribution and S4G2 Spectral Template Case Study	98
4.4.2.5	Basis and Correlation Matrices	100
4.4.3	Pulsed StoWGe Optimization Conclusions	102
4.5	Experimental Results	102
4.5.1	Loop-back Testing	106
4.5.2	Open-Air Testing	106
4.6	Conclusions	106
5	CW Stochastic Waveform Generation (CW-StoWGe)	107
5.1	The CW-StoWGe Signal Model	107
5.2	The Expected Frequency Template Error (EFTE) for CW-StoWGe	112
5.3	Optimization of the CW-StoWGe Expected Power Spectrum	113
5.3.1	The CW-StoWGe EFTE Gradient	113
5.3.2	Gradient Descent Implementation	113
5.3.3	Initializations	114
5.4	Optimization Results	115
5.5	Experimental Results	116
5.5.1	Loopback Testing	116
5.5.2	Open-Air Testing	116
5.6	Conclusions	116

6 Conclusions & Future Work	117
A StoWGe Derivations	118
A.1 Pulsed StoWGe	119
A.1.1 The Pulsed StoWGe Expected Spectrum	119
A.1.2 The Pulsed StoWGe EFTE Gradient	120
A.1.2.1 Basis Function Matrix Gradient	120
A.1.2.2 Mean Phase Value Vector Gradient	122
A.1.3 The Pulsed StoWGE EFTE Gradient for Selected Distributions	123
A.1.3.1 Discrete uniform distribution with two states (DU2)	123
A.1.3.2 Continuous uniform distribution (CU)	124
A.1.3.3 Gaussian distribution (G)	125
A.2 CW-StoWGe	126
A.2.1 The CW-StoWGe Expected Spectrum	126
A.2.2 The CW-StoWGe EFTE Gradient	130
A.2.3 The CW-StoWGe EFTE Gradient for Selected Distributions	131
B Tabulated Optimization Results	132
B.1 Pulsed StoWGe	132
B.2 CW-StoWGe	132
B.3 StoWGe Comms	132

List of Figures

2.1	Ideal, simplistic radar data where the entire pulse envelopes are visible, undistorted, and are free of noise. The leading edges clearly indicate the positions of six distinct objects.	7
2.2	The same radar radar data as in Fig. 2.1, but with additive white Gaussian noise . . .	9
2.3	The autocorrelation response of a baseband tone	12
2.4	The noisy radar data from Fig. 2.2 after being matched filtered with a baseband tone	13
2.5	The autocorrelation of 10 us tone and a 10 us, 5 MHz LFM	15
2.6	The power spectrum of 10 us tone and a 10 us, 5 MHz LFM	16
2.7	The noisy radar data from Fig. 2.2 where the transmit signal was an LFM after being matched filtered	16
2.8	The noisy radar data from Fig. 2.2 where the transmit signal was an LFM after being matched filtered and coherently integrated over 100 pulses	19
2.9	Comparison of an LFM waveform with and without a Taylor tapering window . . .	20
2.10	The noisy radar data from Fig. 2.2 with either an LFM or a tapered LFM and coherent integration	21
2.11	The autocorrelation of a length-13 Barker code	23
2.12	The autocorrelation of a length-64 MPS code	24
2.13	The power spectral density of the Barker and MPS codes from Figs. 2.11 and 2.12 respectively	26
2.14	A P4 code compared to its implementation as a PCFM waveform	29
2.15	The noisy radar data from Fig. 2.2 with varying CPI sizes of 1, 10 and 100 corresponding to (a), (b), and (c) respectively	33

2.16	The noisy radar data from Fig. 2.2 but the transmit waveform is a single noise waveform with a 100 pulse CPI	34
2.17	RMS spectrum of 100 optimized PRO-FM waveforms and their desired template	36
2.18	Coherently integrated and RMS autocorrelation of 100 PRO-FM waveforms compared to the RMS autocorrelation of 100 generic noise waveforms	36
4.1	Desired spectra for the EFTE optimization with a 3 dB oversampling factor of $K = 2$	79
4.2	Corresponding autocorrelation response of the desired EFTE spectra	79
4.3	Desired spectra for the EFTE optimization with a 3 dB oversampling factor of $K = 4$	80
4.4	Structure of the "Identity" basis matrix initialization (\mathbf{B}_{ID})	81
4.5	Normalized power spectral deviation for the DU2 distribution, the G2 template, and various values of N	85
4.6	Normalized power spectral deviation for the G distribution, the G2 template, and various values of N	86
4.7	Optimized expected power spectrum for the G2 template (a). Optimized expected power spectrum for G4 template (b). Expected power spectral deviation for the G2 template after optimization (c). Expected power spectral deviation for the G4 template after optimization (d).	90
4.8	Expected autocorrelation for the G2 template after optimization (a). Expected autocorrelation for the G4 template after optimization (b). Expected RMS autocorrelation for the G2 template after optimization (c). Expected RMS autocorrelation for the G4 template after optimization (d). Expected RMS cross-correlation for the G2 template after optimization (e). Expected RMS cross-correlation for the G4 template after optimization (f).	91
4.9	Optimized expected power spectrum for the S4G2 template (a). Optimized expected power spectrum for S4G4 template (b). Expected power spectral deviation for the S4G2 template after optimization (c). Expected power spectral deviation for the S4G4 template after optimization (d).	93

4.10 Expected autocorrelation for the S4G2 template after optimization (a). Expected autocorrelation for the S4G4 template after optimization (b). Expected RMS autocorrelation for the S4G2 template after optimization (c). Expected RMS autocorrelation for the S4G4 template after optimization (d). Expected RMS cross-correlation for the S4G2 template after optimization (e). Expected RMS cross-correlation for the S4G4 template after optimization (f)	94
4.11 Optimized expected power spectrum for the R2 template (a). Optimized expected power spectrum for R4 template (b). Expected power spectral deviation for the R2 template after optimization (c). Expected power spectral deviation for the R4 template after optimization (d)	96
4.12 Expected autocorrelation for the R2 template after optimization (a). Expected autocorrelation for the R4 template after optimization (b). Expected RMS autocorrelation for the R2 template after optimization (c). Expected RMS autocorrelation for the R4 template after optimization (d). Expected RMS cross-correlation for the R2 template after optimization (e). Expected RMS cross-correlation for the R4 template after optimization (f)	97
4.13 All results for the DU2 distribution and the S4G2 template. $N = 32$ is plotted in blue. $N = 256$ is plotted in red. (a) expected power spectrum (b) power spectral deviation (c) expected autocorrelation (d) expected RMS autocorrelation (e) expected RMS cross-correlation	99
4.14 Correlation matrices and selected basis functions for each distribution and the G4 template. Correlation matrices are plotted in magnitude on a dB scale. (a) selected G basis functions (b) G correlation matrix (c) selected CU basis functions (d) CU correlation matrix (e) selected DU2 basis functions (f) DU2 correlation matrix . . .	103

4.15 Correlation matrices and selected basis functions for each distribution and the S4G4 template. Correlation matrices are plotted in magnitude on a dB scale. (a) selected G basis functions (b) G correlation matrix (c) selected CU basis functions (d) CU correlation matrix (e) selected DU2 basis functions (f) DU2 correlation matrix	104
4.16 Correlation matrices and selected basis functions for each distribution and the R4 template. Correlation matrices are plotted in magnitude on a dB scale. (a) selected G basis functions (b) G correlation matrix (c) selected CU basis functions (d) CU correlation matrix (e) selected DU2 basis functions (f) DU2 correlation matrix . . .	105
5.1 The CW-StoWGe phase structure at time m assuming is a multiple of nT_s for $T_s = 4$ and $L = 3$	110

List of Tables

3.1	Random Walk PMFs for $n = 0, 1, 2, 3, 4, 5$	46
3.2	Random walk partial sequences for $n = 0, 1, 2, 3, 4, 5, 6$	46
3.3	Summary of Stochastic Waveform Measures	63
4.1	Waveform generating process parameters which must be selected or initialized prior to optimization	75
4.2	Initializations for the StoWGe parameters. All combinations of every optimization are considered resulting in 288 total optimizations	82
4.3	The best resulting WGFs for each $p_X(x)$ and template as a function of N for $N \geq 32$ and \mathbf{B}_0	87
5.1	Waveform generating process parameters which must be selected or initialized prior to optimization	114
B.1	Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)	132
B.2	Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)	133
B.3	Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)	133

B.4 Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)	134
B.5 Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 2 with respect to its absolute bandwidth ($K = 2$)	134
B.6 Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 4 with respect to its absolute bandwidth ($K = 4$)	135
B.7 CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)	136
B.8 CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)	137
B.9 CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)	138
B.10 CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)	139
B.11 CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 2 with respect to its absolute bandwidth ($K = 2$)	140

B.12 CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 4 with respect to its absolute bandwidth ($K = 4$) 141

Chapter 1

Introduction

Over the years the generation of the radar waveforms has changed greatly from simple spark-gap generators in the early 20th century [1] to the high powered magnetrons [2] and the pulse compression waveforms of the mid 20th century [2–5]. More recently the trend is towards sophisticated arbitrary waveform generators (AWG). High fidelity AWGs along with the incredible computational abilities of modern computers have motivated a great deal of interest in the field of waveform diversity and design [6–10]. The goals behind designing new waveforms is of course to make a given radar system or radar mode more effective. What *more effective* means is application specific; however, some generalizations can be made.

In general, radar waveforms should produce *unambiguous* responses. In the context of the matched filter, a basic and effective processing tool, the autocorrelation sidelobes represent ambiguities which can hide or mask targets of interest. Consequently, a great deal of radar waveform design focuses on minimizing these ambiguous responses in both the range and Doppler domains [3, 5, 11].

Radar waveforms should be spectrally contained. The electromagnetic spectrum is a finite resource. If two users use the same bandwidth then they will interfere with each other and neither will be able to operate effectively. Given the proliferation of spectrum usage especially for commercial applications, it is more important than ever that Radar systems operate in a bandwidth efficient manner for their own sake and for the sake of other users [12–14].

Radar waveforms should be amenable to implementation on high powered equipment. Generally speaking, in order to combat the R^4 power loss, where R is the range to some object, incurred by the two way spherical spreading of the electromagnetic energy, radars often operate at very

high power levels. High power operation necessitates high power amplifiers which operate in the saturation region. If a radar signal or any signal for that matter with amplitude modulation (AM) is passed through an amplifier operating in saturation, it will invariably endure non-linear distortion effects which degrade radar performance and expand the signal spectrum into adjacent bands diminishing spectral containment [6, 15]. Though more sophisticated techniques exist such as predistortion [16, 17], the most straightforward way to mitigate this effect is to design constant modulus (constant amplitude) waveforms.

Numerous waveform implementations and design schemes have been proposed over the years to address these issues and others. This work however considers the design and implementation of what are known as frequency modulated (FM) noise waveforms.

Noise waveforms in general take advantage of the high dimensionality of noise and noise like signals to reduce ambiguous responses such as autocorrelation sidelobes [18–21]. To get an idea of how this works, consider an experiment. If the results are noisy one might run the experiment again and again to mitigate the noise and get a clean result. This is akin to transmitting the same pulse over and over again as with traditional LFM based radar. While the noise is reduced, the autocorrelation sidelobes remain the same. With noise radar or pulse agile radar in general where different pulses are transmitted every pulse repetition interval (PRI), this is like doing an entirely new experiment every pulse. The results can still be combined to lower the noise power, but since the sidelobe responses are different also, they too decrease when coherently combined. This diminishes the ambiguity due to range sidelobes with more and more unique pulses, though it could be argued the ambiguity appears elsewhere rather than being completely mitigated as will be shown with range sidelobe modulation (RSM) [6, 22].

The difficulty with noise radar is a high peak to average power ratio (PAPR) or a lot of AM which leads to the non-linear distortion discussed above. To mitigate this problem while retaining the benefits of noise radar, FM noise waveforms preserve the desirable high dimensionality sidelobe reduction properties of noise radar while achieving a constant amplitude temporal envelope (hence FM). The high dimensionality reduction of sidelobes and the constant amplitude character

of these waveforms address two of the aforementioned design goals. Often, the most difficult aspect of FM noise waveform design and constant amplitude waveform design as a whole is the spectral containment aspect. For FM noise waveforms this often entails an iterative, sometimes computationally expensive design process to make each individual FM noise waveform have a desirable spectrum and good autocorrelation sidelobes [23–26]. Other FM noise implementations are process based. Rather than optimizing each waveform, the phase of the waveform is the sample function of a random process, though in their current formulations these methods lack much design flexibility [27–29].

The goal of this work is to combine the optimization based and the process based FM noise radar implementations by defining a *waveform generating function* (WGF) according to the *stochastic waveform generation* (StoWGe) model [30] and optimizing the WGF to instantiate waveforms with a desired power spectral density (PSD). In this way, the entire optimization process only has to be performed once rather than on a per waveform level and unique FM noise waveforms are generated as sample functions of the WGF. Additionally, this model is extended to continuous-wave a (CW) form where rather than transmitting pulses, the radar emits a single, long-duration waveform. The next several paragraphs outline the structure of this dissertation.

In order to provide context to the waveform design implementations in the later chapters, Chapter 2 introduces various basic radar principles in the context of an admittedly simplistic but representative radar scenario where the role of the radar waveform and subsequent digital signal processing (DSP) are emphasized.

If the radar waveforms developed in this work are the sample functions of random processes, then it makes sense to define and analyze them in terms of the classical definitions and notation random signals and noise. Chapter 3, provides a brief overview of random variables and random processes before explaining how FM noise waveforms and noise waveforms as a whole can be analyzed as random processes. Then, various stochastic signal based metrics are defined in order to provide a means to evaluate the efficacy of a given WGF for producing useful radar waveforms.

In Chapter 4, the StoWGe model is introduced as a parameterized stochastic process. Then by

optimizing its parameters by minimizing the expected frequency template error (EFTE) cost function, various WGFs are optimized to generate waveforms with desirable spectral characteristics. These WGFs are then analyzed in other respects according to the stochastic waveform properties defined in Chapter 3. Waveforms generated by several of the WGFs are then implemented in both loop-back and open-air settings to evaluate their physical robustness.

In Chapter 5, the StoWGe model is modified to allow for the design of CW radar waveforms. The resulting model is very similar to the communications scheme continuous phase modulation (CPM). In a similar progression as Chapter 4, the CW-StoWGe parameters are optimized according to the EFTE cost function. The resulting optimized WGFs are then evaluated according to the stochastic waveform metrics defined in Chapter 3. Waveforms generated by several of the WGFs are then implemented in both loop-back and open-air settings to evaluate their physical robustness.

Chapter 6 provides conclusions as well as a discussion on future work while the appendices provide tabulated optimization results, the derivation of numerous equations relating to StoWGe, CW-StoWGe, and their respective optimizations, and finally a summary of terminology.

Chapter 2

Background

The goal of this chapter is not provide a comprehensive overview of radar principles. Rather, it is to give context to the radar waveform design schemes and objectives of the later chapters. To do so, this chapter introduces numerous basic radar concepts through the examination of an admittedly simplistic, but representative sensing scenario. Since this dissertation is concerned with the design of radar waveforms, throughout this chapter an emphasis is placed on the role of the radar waveform in enabling the radar to do its job.

In Section 2.1, it is first shown how noise complicates the detection process before introducing the matched filter as a means to maximize the signal to noise ratio (SNR) in presence of additive white Gaussian noise (AWGN). Then linear frequency modulated (LFM) pulse is then introduced as a superior pulse compression waveform as compared to the simple unmodulated pulse in that it leverages bandwidth to achieve finer range resolution. Finally, pulse integration is shown to improve SNR and tapering is shown to mitigate to an extent the deleterious autocorrelation side-lobes inherent to the LFM waveform. Section 2.2 introduce several waveform design schemes and topologies for achieving such goals as sidelobe mitigation and spectral containment. This section provides an introduction to FM noise waveforms which are the primary focus of this work. Section 2.3 introduces CW radar, Section 2.4 discusses the ambiguity function and doppler processing, while Section 2.5 provides a brief overview of the gradient descent techniques used in this work.

2.1 Basic Radar Processing

2.1.1 The Unmodulated Pulse and Basic Radar Operation

The simplest form of electromagnetic energy a radar can transmit is the unmodulated pulse. This signal has no amplitude modulation (AM) such that its amplitude is constant over the pulse duration. Additionally, it has no frequency modulation (FM) in that its frequency is constant over the pulse duration. The unmodulated pulse is defined as

$$s_{\text{pb}}(t) = \begin{cases} A \cos(2\pi f_c t) & 0 < t < T \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

where f_c is the carrier frequency of the pulse, A is the amplitude, and T is the duration. The unmodulated pulse definition in 2.1 is represented as a *passband* signal as implied by the subscript pb. Alternatively, it could be represented in *complex-baseband*. For this simple waveform its complex representation is simply a real-valued, time limited, DC pulse such that

$$s_{\text{bb}}(t) = \begin{cases} A & 0 < t < T \\ 0 & \text{otherwise} \end{cases} \quad (2.2)$$

where the subscript bb denotes baseband. More generally, any constant amplitude passband signal is defined as

$$s_{\text{pb}}(t) = A \cos(2\pi f_c + \phi(t)) \quad (2.3)$$

where $\phi(t)$ is some phase function. Alternatively, any constant amplitude baseband signal is defined as

$$s_{\text{bb}}(t) = A \exp(j\phi(t)). \quad (2.4)$$

For the unmodulated pulse of (2.1), $\phi(t)$ is zero.

Physically speaking, the baseband signal is up-converted to the carrier frequency f_c before being transmitted. On receive, the passband signal is down-converted to the baseband representation

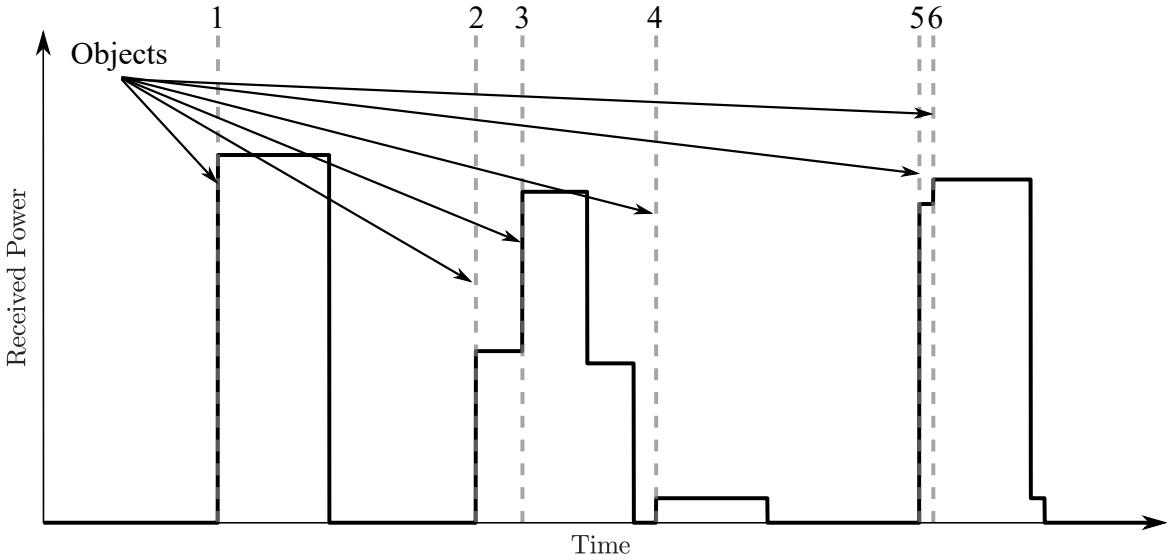


Figure 2.1: Ideal, simplistic radar data where the entire pulse envelopes are visible, undistorted, and are free of noise. The leading edges clearly indicate the positions of six distinct objects.

for processing. For this work, the details of this process are only important inasmuch as the up-conversion and down-conversion steps can be performed. For the purposes of design, analysis, and processing it is more convenient to utilize the baseband representation. From here on out, these tasks will be performed at complex-baseband. More information on the physical hardware implementation of radar signals can be found in any radar textbook such as [31].

Consider a radar operating in some environment. At time t_0 it transmits an unmodulated pulse and then listens for the echoes. In this simplistic case, its goal is to simply determine the distance to any objects in the scene. In an ideal world the received power envelopes may look like Fig. 2.1. where various, distinct reflections are clearly visible and their range can be easily evaluated by identifying the time delay of the leading edge of the pulse. Their amplitudes vary based on the distance to the objects, their particular reflectivity, and numerous other factors which are not necessary to consider here. Even when the reflections fall on top of each other, their position in Fig. 2.1 is obvious. Based on the propagation speed of electromagnetic radiation, the distance to each object is

$$R = \frac{\tau c}{2} \quad (2.5)$$

where τ is the delay of the leading edge of the reflection, c is the speed of light in a vacuum, and the factor of 2 indicates the delay to the object is doubled because the pulse has to travel to and from the object.

Unfortunately, reality is never as simple as in Fig. 2.1. Without any kind of additional processing a more reasonable range response is shown in Fig. 2.2. Filtering, distortion, and primarily the presence of noise has completely obscured the positions of the objects which are plainly visible in Fig. 2.1. The most straightforward solution to this problem would be to simply transmit as much power as possible to raise the signal well above the noise, but this is not as simple as it sounds. In (2.5) the factor of 1/2 accounts for the two way propagation of the radar waveform. This two way propagation also results in a two way spherical spreading loss such that the power returned by any given object is inversely proportional to the fourth power of its range. In other words, if two objects would otherwise reflect the same amount of energy but one is twice as far away from the radar, the further object will only return 1/16 the energy as the closer object. Transmitting more and more energy is a losing battle with respect to range losses. Consequently, minimizing noise, interference, and maximizing detectability through post-processing are essential to radar detection as discussed in the next section.

2.1.2 Linear Frequency Modulation and Pulse Compression

In Fig. 2.1, the energy returned from each pulse is spread out over the entire pulse duration. Previously, only the leading edge was considered. However, it is possible to realize a greater response by "compressing" the returns from each object by using the *matched filter* to implement *pulse compression*. To demonstrate this, consider the ideal response in Fig. 2.1. In this noise free environment, the returned signal is a convolution of the time reversed transmit signal with the impulse response of the environment such that

$$y_s(t) = \int_{-\infty}^{\infty} s(\tau - t)x(\tau)d\tau \quad (2.6)$$

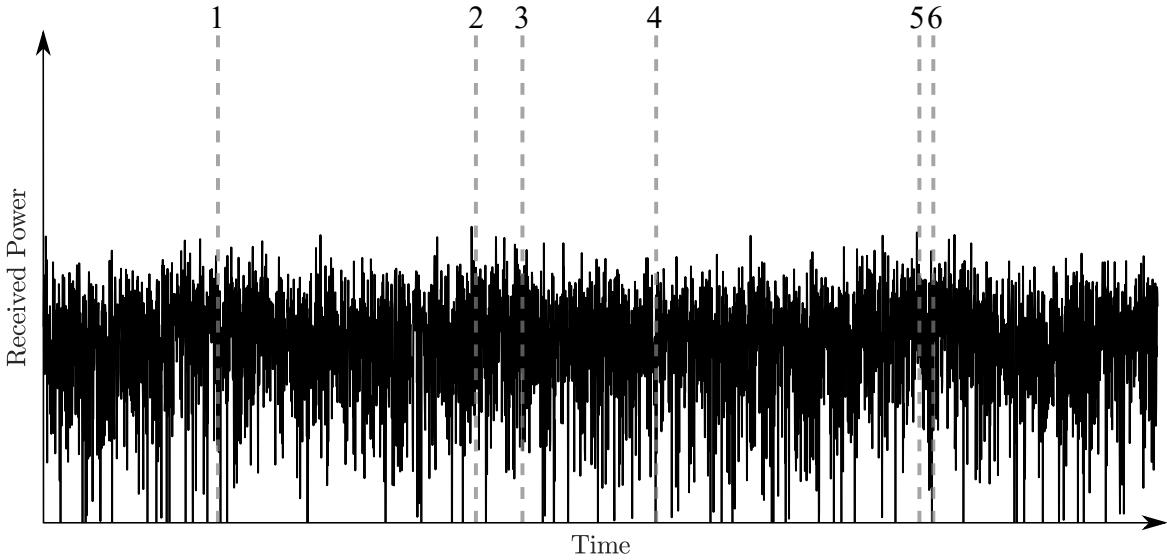


Figure 2.2: The same radar radar data as in Fig. 2.1, but with additive white Gaussian noise

where $s(t)$ is the transmit signal and $x(t)$ is the channel's impulse response. In this case, the sequence of six objects can be described via a linear combination of delta functions such that

$$x(r) = a_n \sum_{n=1}^6 \delta(r - r_n) \quad (2.7)$$

where r is range from the transceiver and a_n is some complex valued scalar which is proportional to the objects reflectivity. r and t are related by (2.5). In Fig. 2.2, the returns are distorted by the addition of white Gaussian noise (WGN) which is an excellent approximation for many natural sources of noise like the thermal noise which is inherent to all electronic devices. The addition of noise modifies (2.6) to

$$y(t) = \int_{-\infty}^{\infty} s(\tau - t)x(\tau)d\tau + v(t) \quad (2.8)$$

where $v(t)$ is the additive WGN. In Fig. 2.2, the noise power of $v(t)$ is strong enough to completely obscure the objects which were easily visible in Fig. 2.1. The relative power between the returned signal and the noise is the signal to noise ratio (SNR). Unless other factors cause further interference, SNR is an effective tool for determining whether something is detectable. In fact, the *radar*

range equation which typically measures the maximum range a radar can detect something, is a function of SNR. A detailed discussion of the radar range equation is not needed here, but can be found in [31].

The fact remains that, even without describing any specific values, some additional processing is necessary to make the returns in Fig. 2.2 useful. To do so, it is often useful to apply a linear filter. In the time domain, this operation is mathematically described via a convolution. Thus the filtered data becomes

$$y_f(t) = \int_{-\infty}^{\infty} f(t - \tau) y(\tau) d\tau \quad (2.9)$$

where $f(t)$ is some filter function and $y_f(t)$ is the returned data under the filtering operation. For the matched filter, $f(t)$ becomes

$$f(t) = a s^*(-t) \quad (2.10)$$

which is the complex conjugated, time reversed version of the baseband signal $s(t)$ with an arbitrary scale factor a . By replacing $f(t)$ with the matched filter and by replacing $y(t)$ with its signal component, $y_s(t)$, and its noise component, $v(t)$, (2.9) becomes

$$y_{mf}(t) = \int_{-\infty}^{\infty} a s^*(\tau - t) (y_s(\tau) + v(\tau)) d\tau \quad (2.11)$$

where $y_{mf}(t)$ is the returned data under the matched filtering operation. Recall, that $y_s(t)$ as defined in (2.6) is a linear combination of time shifted versions of the transmit signal. Under the matched filtering operation $y_s(t)$ instead becomes a linear combination of time shifted *autocorrelations*. The autocorrelation is defined as

$$r(t) = \int_{-\infty}^{\infty} s^*(t - \tau) s(t) d\tau. \quad (2.12)$$

The usefulness of the matched filter is described by the Schwartz inequality. For two arbitrary

functions, $f_1(x)$ and $f_2(x)$, a relationship exists such that [31]

$$\int_a^b f_1(x)f_2(x)dx \leq \int_a^b f_1(x)dx \int_a^b f_2(x)dx \quad (2.13)$$

which holds with equality iff $f_1(x) = af_2(x)$ where a is some constant. At $t = 0$ in (2.12), this is precisely the case. Consequently, whenever $t = t_0$ where t_0 is the location of an object in $x(t)$, (2.11) becomes

$$y(t_0) = \int_0^T |s(t)|^2 dt + \int_{t_0-T/2}^{t_0+T/2} s^*(t)v(t)t. \quad (2.14)$$

The first term in (2.14) is, by definition, the total energy of the transmit signal and also the form of the Schwartz inequality that realizes the equality condition. If $v(t)$ is exclusively WGN, the matched filter maximizes the SNR. This is the most important aspect of the matched filter. A more complete discussion of this can be found in [5, 7, 32, 33].

The autocorrelation of the baseband unmodulated pulse realizes a triangular function. This triangular function becomes somewhat like a downward facing parabola as shown in Fig. 2.3 when it is plotted on a dB scale. Applying the matched filter to the data in Fig. 2.2 realizes Fig. 2.4 where several responses are clearly visible however there are still some obvious issues. While object 1 is plainly visible, 2 and 3 are hard to tell apart, 4 looks like part of 2 and 3, while 5 and 6 look like one return. For the most part, the problem here is *resolution*, there just is not enough separation between many of the objects to reliably identify them given the wide autocorrelation response of the unmodulated pulse. The resolution of the tone is usually defined as the Rayleigh resolution which is [31]

$$\delta R = \frac{cT}{2} \quad (2.15)$$

From (2.15), one solution is to shorten the pulse. The time width of the autocorrelation response is 2 times the temporal length of the signal so shortening the signal shortens the autocorrelation. However, this means less energy on target leading to poorer SNRs. To compensate, the radar could transmit at a higher power, but this leads to a vicious cycle of transmitting ever shorter pulses at

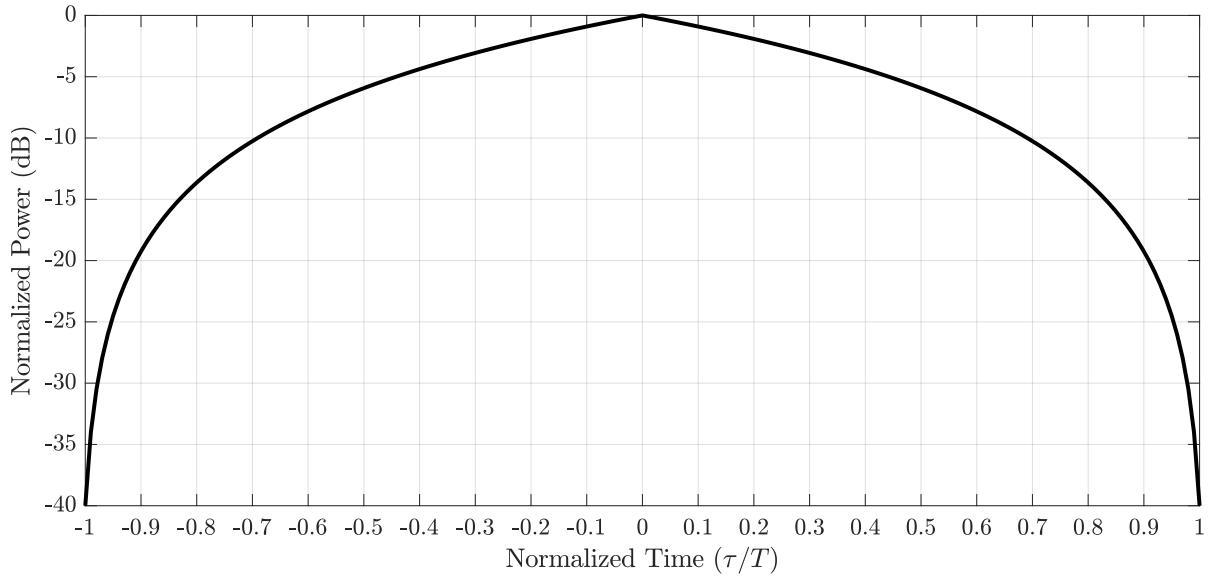


Figure 2.3: The autocorrelation response of a baseband tone

ever higher powers. Due to hardware constraints, there are practical limits to both.

To address the resolution problem, we can transmit a pulse with a more advantageous auto-correlation. For decades the linear frequency modulated (LFM) pulse has been the prototypical modulated radar waveform. As the name implies, the frequency function of an LFM is a linear function of time such that at passband the LFM is defined [7, 31, 32, 34]

$$s_{\text{LFM}}(t) = \begin{cases} A \cos(2\pi f_c t + \pi \frac{B}{T} t^2) & -T/2 < t < T/2 \\ 0 & \text{otherwise} \end{cases} \quad (2.16)$$

which has been centered at $t = 0$ for convenience. As described by (2.3), the phase function of (2.16) is

$$\phi(t) = 2\pi f_c t + \pi \frac{B}{T} t^2 \quad -T/2 < t < T/2 \quad (2.17)$$

and the radial frequency function is the derivative of the phase function yielding

$$\frac{d\phi(t)}{dt} = 2\pi f_c + 2\pi \frac{B}{T} t \quad -T/2 < t < T/2 \quad (2.18)$$

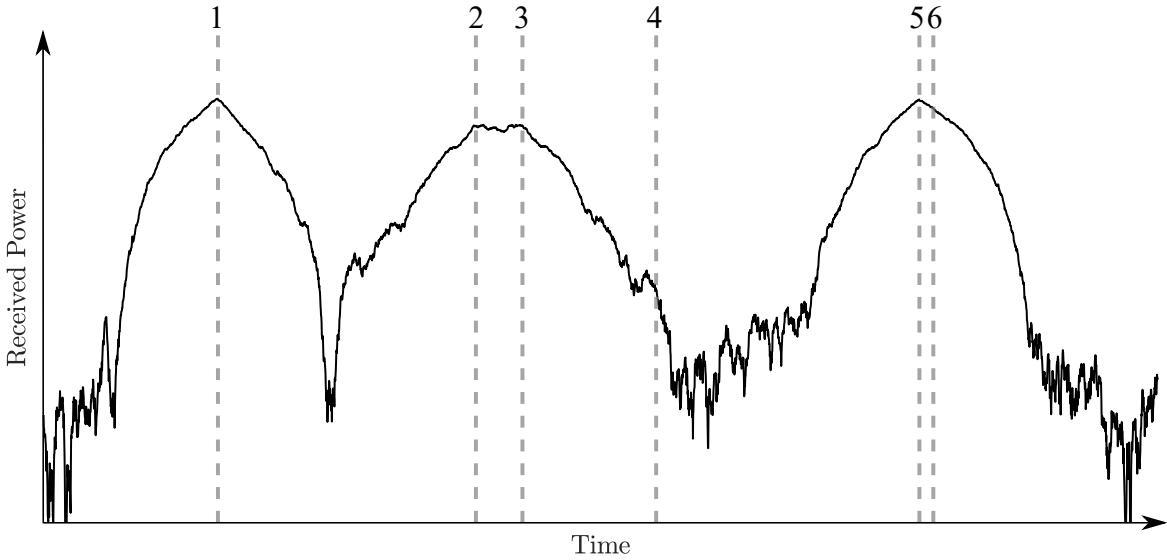


Figure 2.4: The noisy radar data from Fig. 2.2 after being matched filtered with a baseband tone

such that the frequency function in Hz is

$$f(t) = f_c + \frac{B}{T}t \quad -T/2 < t < T/2 \quad (2.19)$$

which is a linear function of time. Over the course of the pulse, the signal chirps through B Hz in T seconds. To understand what effect the modulation has on the autocorrelation function and further the scene in Fig. 2.4, it is useful to apply some arbitrary, but representative numbers to the modulated and unmodulated pulses.

Consider a modulated and an unmodulated pulse, both with $T = 10$ us and for the LFM $B = 10$ MHz. Each pulse has been normalized to unit energy such that $|s(t)|^2 = 1$. The autocorrelation response of each is shown in Fig. 2.7 where showing only half of the autocorrelations is necessary since they are symmetric about $t = 0$. The LFM autocorrelation clearly will do a much better job of resolving relatively close together objects as opposed to the tone owing to its well defined peak. Additionally, a lobing structure has been revealed where the lobe around the match point, $t = 0$, is called the *mainlobe* while all other lobes are known as *sidelobes*. As will be shown, these sidelobes are inherent to virtually all waveforms and are problematic in their tendency to hide

weaker reflections. For now though, the LFM autocorrelation is superior to the tone autocorrelation from its resolution improvement alone. To show why introducing a modulation has resulted in such a drastic autocorrelation improvement, it is useful to examine their power spectra.

The power spectra of either the LFM or the tone or any pulsed radar waveform for that matter is evaluated by taking the magnitude squared of the Fourier transform of the pulse. The result of these operations is shown in Fig. 2.6. The baseband tone is simply a DC pulse so its power is concentrated at 0 frequency and takes on a sinc squared envelope owing to its rectangular pulse shape. The LFM spectrum on the other hand is spread fairly evenly throughout its swept bandwidth ($|f| \leq 2.5\text{MHz}$). The vertical lines in Fig.2.6 represent the bandwidth of the tone and the LFM. For the tone the bandwidth is considered to be

$$B_{\text{tone}} \approx \frac{1}{T} \quad (2.20)$$

which in this case is 100kHz. This corresponds to approximately the 4 dB bandwidth of the tone. For the LFM, the bandwidth is normally considered to be its swept bandwidth which corresponds to approximately is 6 dB bandwidth which in this case is 5 MHz.

In general, the notion of bandwidth is defined for the application. For matters of resolution, the relevant bandwidth metrics are usually defined with respect to the spectral power falling below some relative power threshold like the ones just mentioned for the tone or the LFM. For non-LFM waveforms, the 3dB bandwidth (the point at which the spectral power falls below 1/2 (-3dB) of the peak power) is commonly used to estimate the resolution. In other applications it may be relevant to consider the XX% bandwidth or the bandwidth which contains XX% of the signal power. A common value may be the 99% bandwidth. The absolute bandwidth is the frequency beyond which there is no frequency content.

Regardless, the threshold bandwidth metrics (3dB, 4dB, 6dB), tie nicely into what is known as the time-bandwidth product (BT). For any tone, the BT is

$$BT = T \frac{1}{T} = 1. \quad (2.21)$$

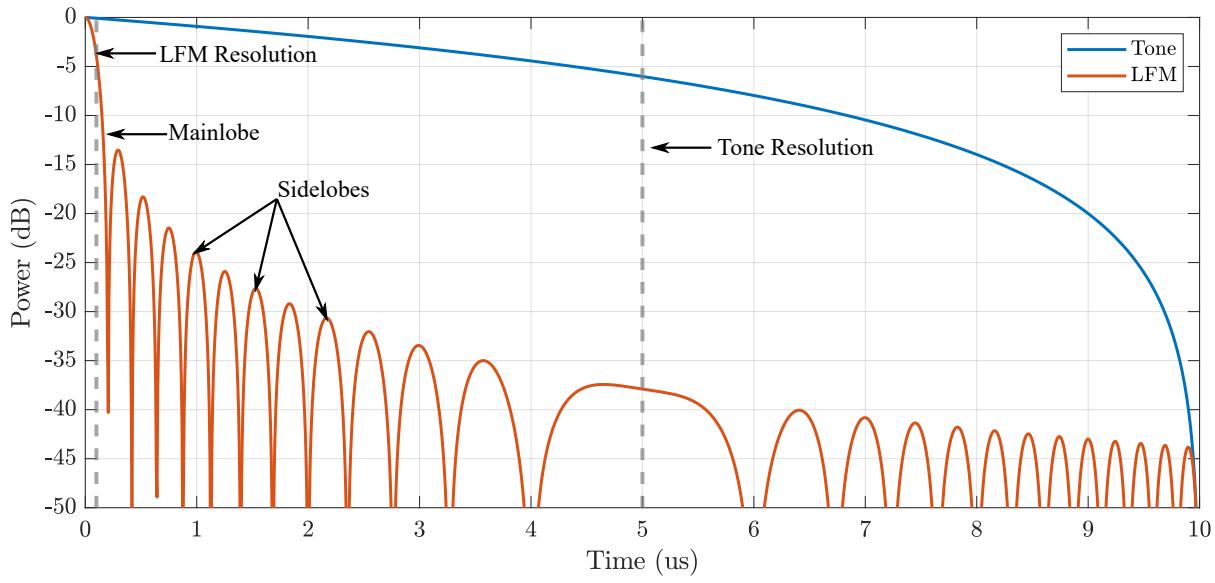


Figure 2.5: The autocorrelation of 10 us tone and a 10 us, 5 MHz LFM

For the LFM from above, using the swept bandwidth or the 6dB bandwidth which are synonymous in this case, the BT becomes

$$BT = (10^{-5} \text{ s})(5 \cdot 10^6 \text{ Hz}) = 50. \quad (2.22)$$

In general, a waveform with X times the BT of another waveform will likewise have an X times improved resolution. So, for a given pulse length increasing the bandwidth improves the range resolution. Finally, if an LFM were transmitted rather than a tone, after the matched filtering operation, the result in Fig. 2.2 becomes Fig. 2.7 where objects 2 and 3 are now clearly separable. Still, there are problems. Despite the processing gains of matched filtering and using the much higher resolution LFM, objects 4 and 6 are still buried beneath the noise since the SNR is too low to detect them.

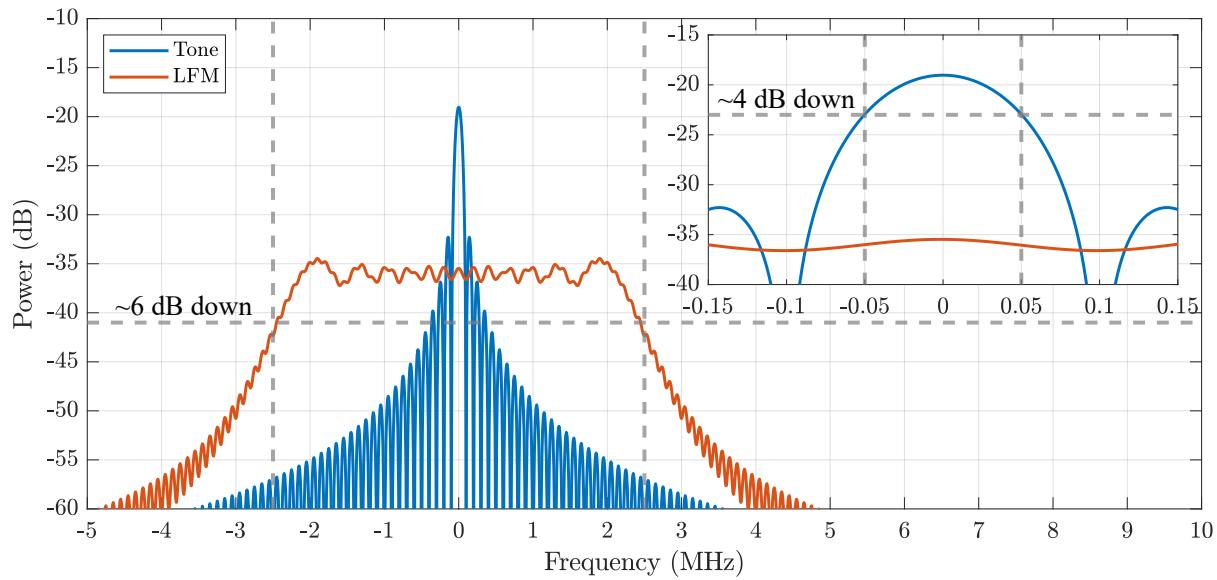


Figure 2.6: The power spectrum of 10 us tone and a 10 us, 5 MHz LFM

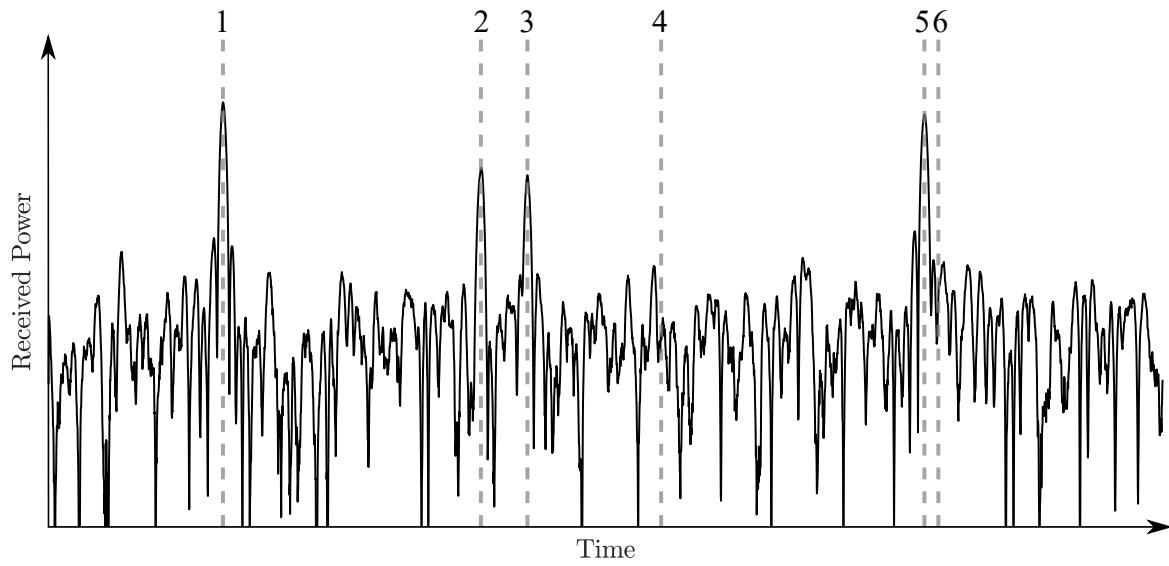


Figure 2.7: The noisy radar data from Fig. 2.2 where the transmit signal was an LFM after being matched filtered

2.1.3 Pulse Integration and LFM Sidelobe Mitigation

2.1.3.1 Pulse Integration

Radar systems rarely operate in a single pulse mode. Typically, whether it is attached to an aircraft, a spacecraft, or is stationary, the view of the radar will be changing as either it moves itself or the object it is attached to moves. Still, radars will often transmit with *pulse repetition frequencies* (PRF) in the kHz or even tens of kHz where thousands of pulses will be transmitted every second such that despite the motion of the radar or the platform, consecutive pulses will return data from largely the same scene, especially for objects that are stationary with respect to the radar. The inverse of the PRF is the *pulse repetition interval* (PRI) which says a pulse is transmitted every PRI seconds. Sets of consecutive pulses are often jointly processed as determined by the length of the *coherent processing interval* (CPI). Organizing pulses in this way is useful for several reasons.

Due to the Doppler effect, objects moving with respect to the radar induce a frequency shift on the signal which is very effective for identifying moving objects in the presence of stationary objects. By collecting multiple pulses in a CPI, these (usually) small frequency shifts can be measured to estimate velocity. More will be said on this later.

The more immediately relevant reason for transmitting multiple pulses is achieving more power on target. Assuming the scene changes minimally over the course of a CPI or if the changes due to object or platform motion can be compensated for, the energy from various pulses can be added together or *integrated* to achieve a better signal to noise ratio.

To demonstrate how this works consider the matched filter data from N different pulses in a CPI. For every pulse, the underlying scene is the same, but the noise is assumed to be independent, identically distributed (i.i.d.), and zero mean. At any arbitrary point in time, the summation (integration) of this data is

$$y_N(t_0) = \sum_{n=1}^N y_n(t_0) + v_n(t_0). \quad (2.23)$$

Since the signal data in each case is the same and in phase (2.23) becomes

$$y_N(t_0) = Ny(t_0) + \sum_{n=1}^N v_n(t_0). \quad (2.24)$$

where the n subscript has been dropped since all $y_n(t_0)$ are by definition equivalent. The signal power is then $(Ny(t_0))^2$. The noise power however is evaluated as the variance of the noise component. Since the noise is zero mean, its variance is defined

$$\mathbb{E} \left[\left| \sum_{n=1}^N v_n(t_0) \right|^2 \right] = \mathbb{E} \left[\left(\sum_{n=1}^N v_n(t_0) \right) \left(\sum_{n=1}^N v_n^*(t_0) \right) \right] \quad (2.25)$$

Since the noise is i.i.d., the cross terms cancel and (2.25) becomes [35]

$$\mathbb{E} \left[\left(\sum_{n=1}^N v_n(t_0) \right) \left(\sum_{n=1}^N v_n^*(t_0) \right) \right] = N\sigma_v^2 \quad (2.26)$$

where σ_v^2 is the noise variance of a single sample. From this result, the signal power increased by a factor of N^2 while the noise power only increased by a factor of N . The SNR is then improved by a factor N , the ratio between these values. Thus for the integration of N pulses, the relative noise power will decrease by $10\log_{10}(N)$. Much more will be said about random variables in Chapter 3.

Consider Fig. 2.7. Objects 5 and 6 are buried beneath the noise, but this is for a single pulse. With coherent pulse integration Fig. 2.7 becomes Fig. 2.8 where 100 pulses have been coherently integrated. In Fig. 2.8 object 4 is now plainly visible and the LFM sidelobes have been revealed from the noise. Still, object 6 is not identifiable. It appears to no longer be buried in the noise, but it instead is hidden beneath the sidelobes of object 5. Because of this, no degree of pulse integration will reveal object 6 necessitating other approaches. One option is to use a waveform other than the LFM which has lower autocorrelation sidelobes. In general, the mitigation of autocorrelation sidelobes is the driving force behind waveform design.

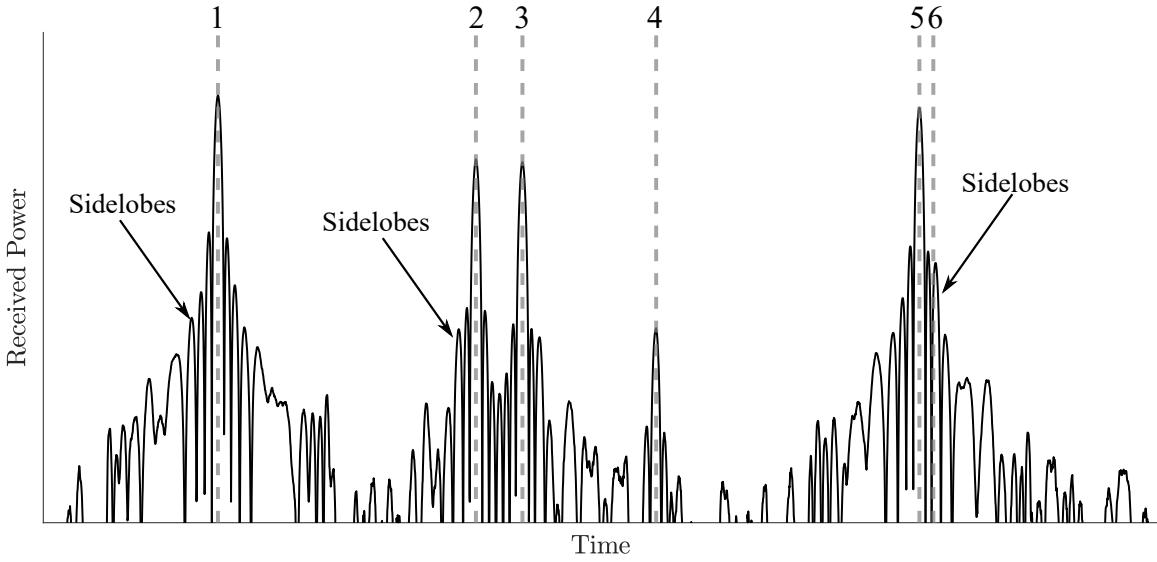


Figure 2.8: The noisy radar data from Fig. 2.2 where the transmit signal was an LFM after being matched filtered and coherently integrated over 100 pulses

2.1.3.2 LFM Sidelobe Mitigation

One of the best ways to mitigate the sidelobes of the LFM is through tapering the received data's spectrum such that

$$Y_{\text{taper}}(f) = Y(f)W(f) \quad (2.27)$$

where $Y(f)$ is the spectrum of the returned data and $W(f)$ is the tapering function. In general, these tapering functions will smooth the sharp corners of the LFM spectrum that are seen in Fig. 2.6. In the time domain, this has the effect of lowering the autocorrelation sidelobes significantly, but also inducing SNR and range resolution losses. In Fig. 2.9, the resulting autocorrelation and spectrum are shown after tapering an LFM with a Taylor window [31]. The autocorrelation response is calculated via the Weiner-Khinchine theorem by taking an inverse Fourier transform of the tapered spectrum. The resulting autocorrelation in Fig. 2.9 demonstrates a sidelobe level about 30 dB below the original peak, but the autocorrelation peak itself is almost 5 dB lower than its untapered counterpart and it has a slightly poorer resolution. Given a scenario with sufficient SNR, tapering would likely be desirable.

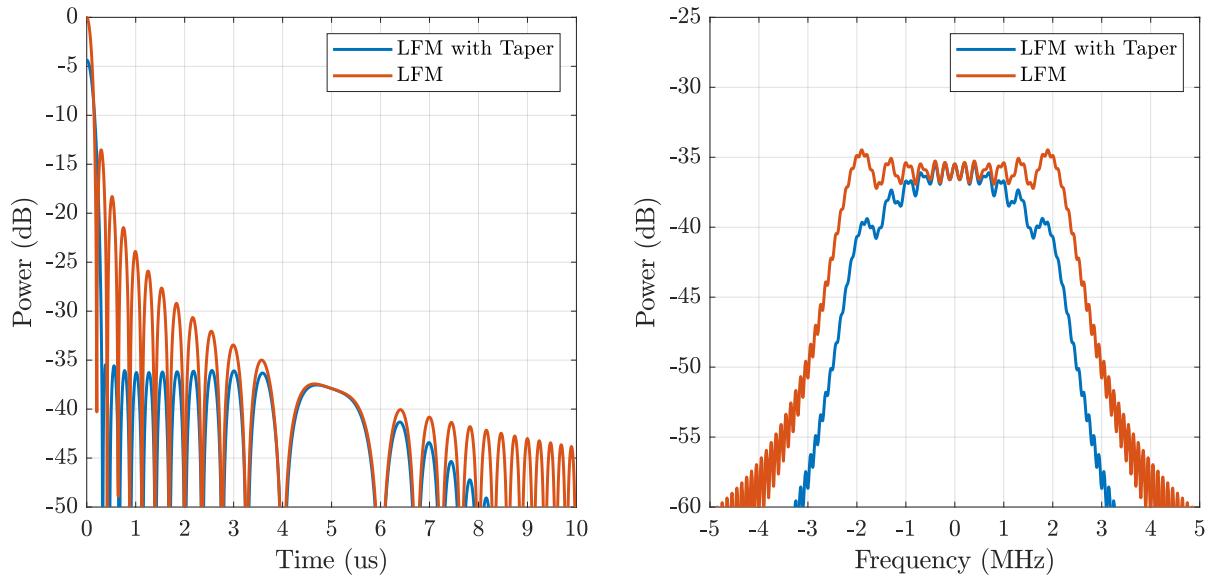


Figure 2.9: Comparison of an LFM waveform with and without a Taylor tapering window

Applying the taper to the LFMs in Fig. 2.8 yields Fig. 2.10 where the new tapered results are shown in red on top of the untapered LFM results. The clear lowering of sidelobes is apparent as well as a loss in SNR. Critically, despite the loss in SNR the improved sidelobe levels have revealed object 6 as just past object 5. This example is just representative as there are numerous window functions each of which have their own trade offs and can be applied in different ways. Further information on tapers can be found in [36, 37].

2.2 Pulse Compression Waveform Design

Section 2.1.2 introduced the concept of autocorrelation sidelobes and Section 2.1.3 showed through Fig. 2.8 where they become a problem. Section 2.1.3.2 introduced the means to suppress them and this section will discuss in detail the various waveforms and waveform design schemes which have been developed to lower autocorrelation sidelobes and achieve other goals.

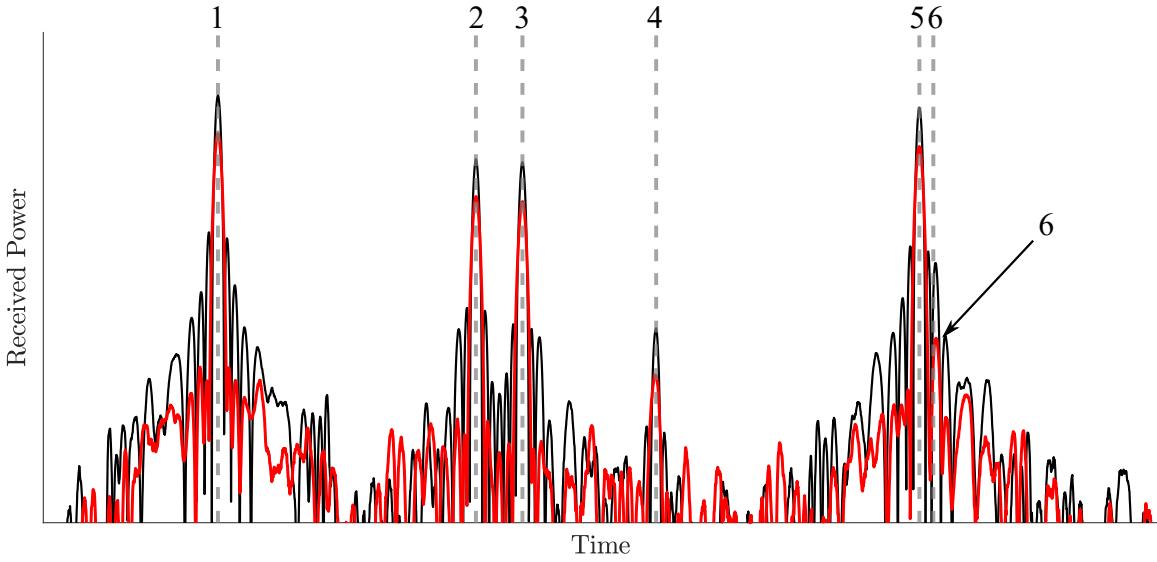


Figure 2.10: The noisy radar data from Fig. 2.2 with either an LFM or a tapered LFM and coherent integration

2.2.1 Non-linear FM

After discussing the LFM waveform and tapering techniques it is natural to first discuss what are typically considered non-linear FM waveforms (NLFM). In section 2.1.3.2, tapering was discussed as means to lower autocorrelation sidelobes at the expense of reduced range resolution and SNR loss. In the frequency domain, tapering LFM waveforms has the effect of making the power spectrum more Gaussian like. NLFM waveforms inherently posses this Gaussian like spectrum and consequently much lower autocorrelation sidelobes. They are often design via the principle of stationary phase (POSP) and have frequency function which resembles a "sideways-S" [38–40].

2.2.2 Phased-coded Waveforms

The phase-coded signal model considers radar waveforms as a sequence of discrete values. Practically speaking, such a model is intuitive. After all most electronic systems nowadays are digital systems which sample the input data into a sequence of discrete values anyway. Additionally, the phase-coded model is relatively simple making it mathematically tractable from a design stand-

point. Stated formally, a unit energy phase coded signal of duration T is defined as [7]

$$s(t) = \frac{1}{\sqrt{T}} \sum_{n=1}^N \exp(j\phi_n) \text{rect}\left(\frac{t - (m - 1/2)t_b}{t_b}\right) \quad (2.28)$$

where $\text{rect}(\cdot)$ is defined

$$\text{rect}(t) = \begin{cases} 1 & -1/2 < t < 1/2 \\ 0 & \text{otherwise} \end{cases} \quad (2.29)$$

and the N phase values, $\phi_1, \phi_2, \dots, \phi_N$ are collectively the phase code. t_b is the time width of each $\text{rect}(\cdot)$ and is known as the chip time. Since the model is based on a sequence of rectangular chips, the autocorrelation can be evaluated as the discrete correlation of the phase coded sequence with itself. The continuous time autocorrelation is then a linear interpolation of the discrete correlation as is seen in Figs. 2.11 and 2.12.

2.2.2.1 Phase-coded Waveform Examples

As a field, there has been a huge number of contributions to the study of phase codes, far too many to study in detail here. However, many excellent resources exist such as [7, 41]. As a representative example, Barker codes and Minimum Peak Sidelobe (MPS) codes are discussed in some detail here while some other design schemes are introduced.

Perhaps the most well known set of phase codes are the Barker codes. Originally developed in 1953, Barker codes realize a peak to sidelobe ratio (PSL) of $1/N$ where N is the length of the code [42]. Barker codes are *binary* codes in that the phase values only take on one of two antipodal states usually referred to as 1 and 0 which map to π and 0 respectively but any two relative states are acceptable so long as they are opposite each other on the unit circle. The unfortunate aspect of Barker codes is that they are only known to exist for values $N \leq 13$. The $N = 13$ Barker code autocorrelation is shown in Fig. 2.11.

With this limitation in mind, numerous authors have sought to find the next best thing which are termed minimum peak sidelobe (MPS) codes. As the name implies, for a length N binary

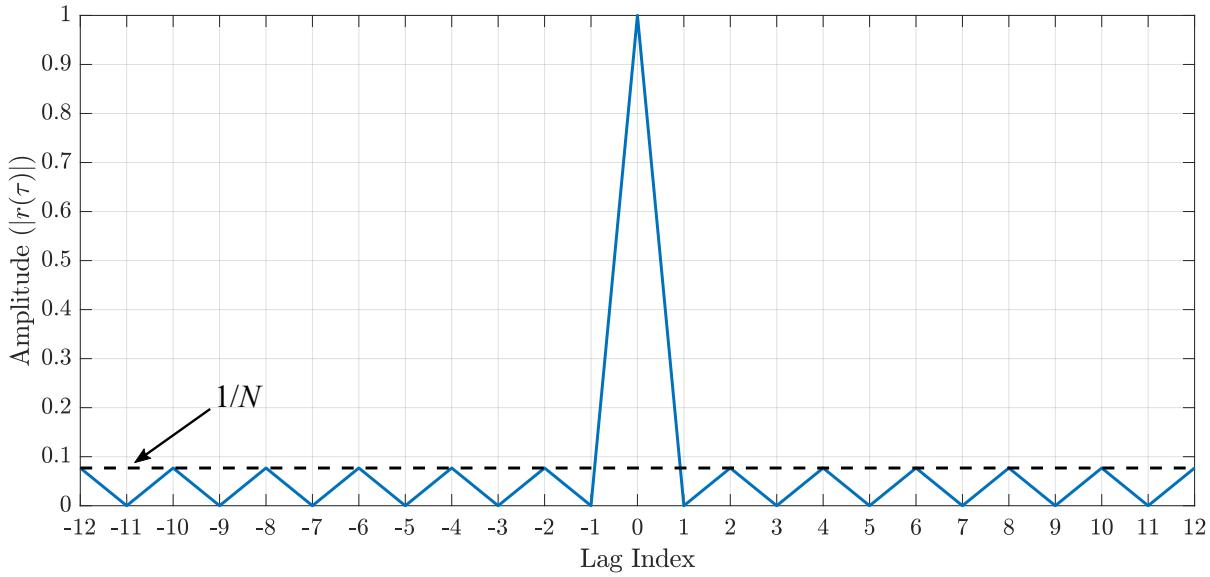


Figure 2.11: The autocorrelation of a length-13 Barker code

code, the MPS sequence achieves the lowest possible sidelobe level. Due to the binary nature of the code, numerical optimization methods are not suited to the problem. Consequently exhaustive searches have been employed to find MPS codes. The problem with an exhaustive search is how many sequences there are for a given value of N since $N_s = 2^N$ where N_s is the number of candidate binary sequences. In 1975, Lindner implemented such an exhaustive search for $N \leq 40$, but due to the sheer number of codes to check, the computerized search took 50 days of computation [43]! Since then, other more efficient approaches have extended the list of known MPS codes such as [44] which found codes all the way to $N = 70$. Fig. 2.12 shows the autocorrelation of an MPS code for $N = 64$ where the optimal PSL happens to be $4/N$ or $1/16$.

If one considers codes with larger alphabets such that ϕ_n is allowed to take on values beyond just π or 0, sequences which meet the Barker code performance of $1/N$ can be found for longer than $N = 13$ sequences. These are known as polyphase barker codes [45]. Beyond Barker, polyphase Barker and MPS codes, numerous other code designs have been proposed. These include Frank codes [46], Zadoff-Chu codes [47], P-codes [48, 49], Golomb codes [50], and many more. [11, 41] provide excellent overviews of this topic along with a more in depth analysis of particular codes

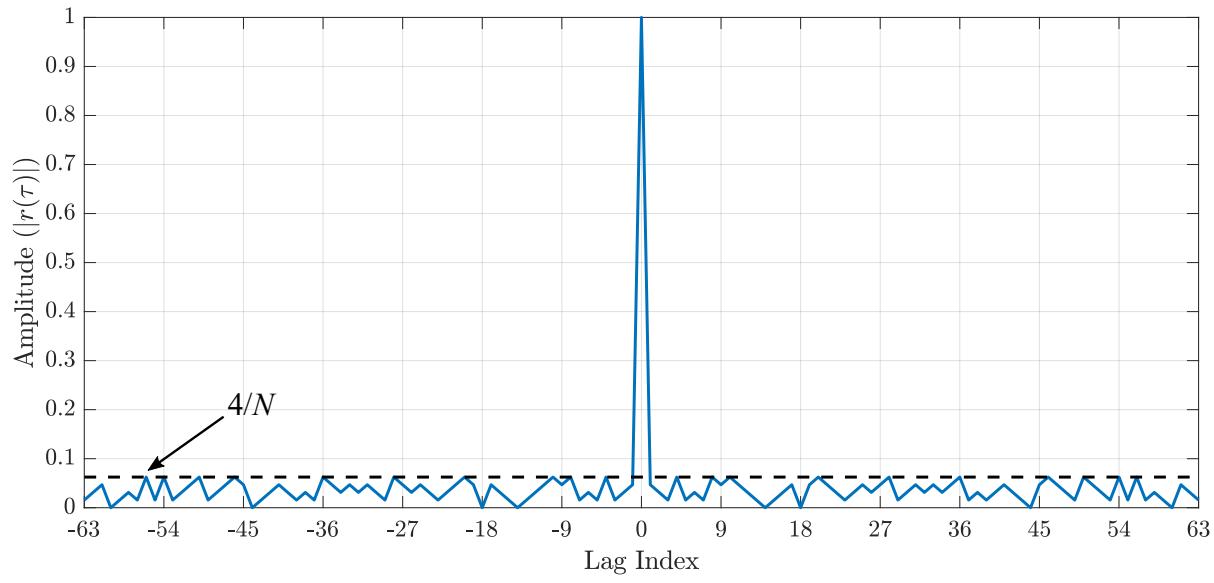


Figure 2.12: The autocorrelation of a length-64 MPS code

and coding schemes.

2.2.2.2 Phase-coded Waveform Spectral Characteristics

To this point, little has been said about the spectral content of radar signals. In Section 2.1.2, bandwidth was discussed with respect to the autocorrelation response and resolution, but for tones andLFMs bandwidth is not really an issue from a waveform generation standpoint since each have a spectra which decays or "roll-off" quite quickly beyond its main bandwidth. For phase codes this is not the case. The rectangular shape of the chips and the instantaneous phase changes between the chips tend towards waveforms with poor spectral roll-off.

Taking the Fourier transform of (2.28) yields

$$S(f) = \frac{1}{\sqrt{T}} \sum_{n=1}^N \frac{\sin(\pi f t_b)}{\pi f} \exp(j(\phi_n - 2\pi f(n-1/2)t_b)) \quad (2.30)$$

and the power spectral density is then

$$|S(f)|^2 = \frac{\sin^2(\pi f t_b)}{T(\pi f)^2} \sum_{n=1}^N \sum_{m=1}^N \exp(j(\phi_n - \phi_m - 2\pi f(n-m)t_b)) \quad (2.31)$$

which can be reduced to

$$|S(f)|^2 = \frac{\sin^2(\pi f t_b)}{T(\pi f)^2} \left(N + 2 \sum_{n=1}^{N-1} \sum_{m=n+1}^N \cos(\phi_n - \phi_m - 2\pi f(n-m)t_b) \right) \quad (2.32)$$

which is a super-position cosine modulated sinc squared functions which results in poor roll-off [11]. The cosine terms result in a repeating pattern of images every $1/t_b$ Hz centered at 0 Hz in this baseband representation. These images are then attenuated by the sinc squared function. For the Barker and MPS codes in figs. 2.11 and 2.12, the spectra are plotted in Fig. 2.13 where each spectrum has been normalized to the same bandwidth. Based on Fig. 2.13, the fundamental bandwidth, $|f| < 1/2$, is determined by the chip length t_b . In general, as with the LFM spectrum in Fig. 2.6, the goal is to contain the spectrum to within this interval to the degree possible. Otherwise, the transmit electronics will filter out the higher frequency components of the signal leading to linear signal distortion and AM. This AM can then result in further non-linear distortion in the high power amplifier. Even if the transmission system can handle the extended spectrum, then the signal may interfere with other users and the receiver then needs to be able to handle the wide bandwidth itself. Accepting such a wide bandwidth may result in additional interference to the radar from other users that would otherwise be attenuated out of band if the radar could just focus on the fundamental interval. Additionally, a wider bandwidth means accepting a higher noise power as additional higher frequency noise is accepted.

Clearly, there good reasons to keep the signal spectrally contained. Consequently, several methods have been proposed to aid in spectral roll-off. One method is to smooth the phase transitions by linearly changing the phase over a fraction of the chip width [7]. For bi-phase codes such as the ones shown in this section, the bi-phase to quadriphase (BTQ) transform [51] and derivative phase shift keying [52] have been shown to be affective at containing the bi-phase spectrum. Another approach is to use a different shaping filter for the chips. Rather than using rectangular chips Chen and Cantrell suggested using a Gaussian weighted sinc function [53]. However, this approach introduces AM to the signal which presents its own problems. Finally, a method that was originally

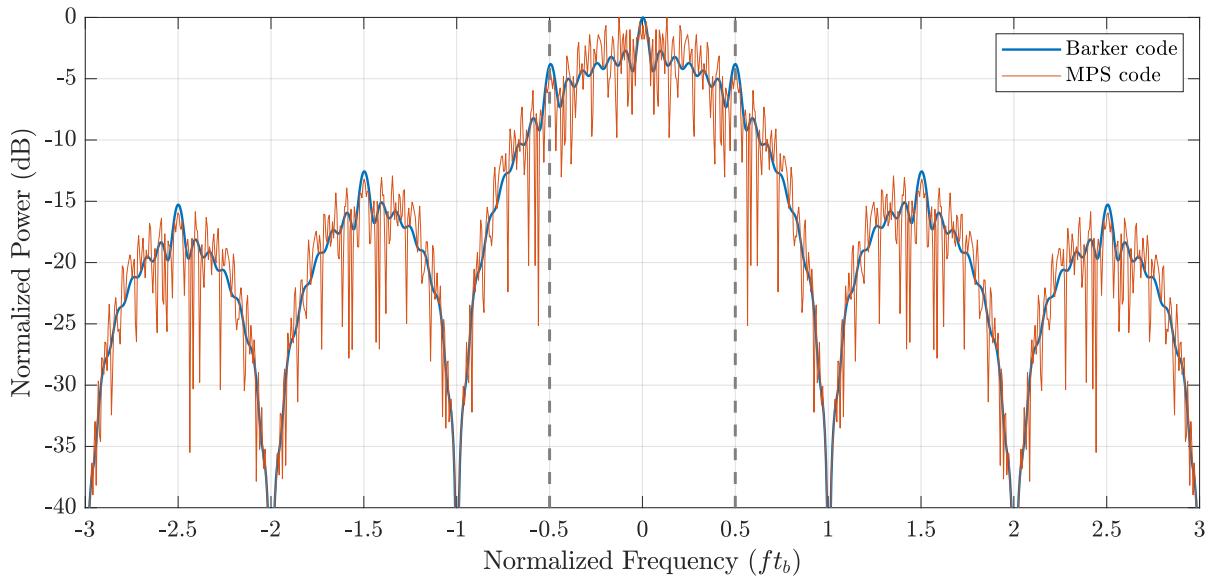


Figure 2.13: The power spectral density of the Barker and MPS codes from Figs. 2.11 and 2.12 respectively

developed for implementing poly-phase codes is discussed in the next Section.

2.2.3 Polyphase-coded Frequency Modulation

The last section ended with a discussion on the spectral challenges associated with implementing bi-phase and poly-phase coded waveforms. Polyphase-coded Frequency Modulation (PCFM) was developed as a means to address these issues [54]. The PCFM waveform model borrows heavily from the continuous phase modulation (CPM) communications scheme which is commonly used for power constrained applications where power efficiency is key such as the BluetoothTM wireless standard [55–58]. Being a phase modulation scheme CPM is constant amplitude like phase codes making it amenable to high power transmitters. However, its phase function is continuous as well making it actually a frequency modulation scheme with, in general, better spectral containment than a purely phase modulated scheme.

The CPM signal model is predicated on a continuous wave (CW) signal in that it has no math-

ematically defined beginning or end such that

$$s_{\text{CPM}}(t; \mathbf{I}) = \exp \left(j2\pi \sum_{k=-\infty}^m I_k h_k q(t - kT_s) \right), \quad mT \leq t \leq (m+1)T \quad (2.33)$$

where \mathbf{I} is an infinite length vector of information carrying symbols, T_s is the symbol time, h_k is a scalar known as the modulation index which can change with every symbol. $q(t)$ is the symbol phase response which itself is defined as

$$q(t) = \int_0^t g(\tau) d\tau \quad (2.34)$$

where $g(t)$ is the frequency shaping filter or frequency pulse. To adapt (2.33) to a pulsed Radar waveform model, the communications aspects were dropped and the signal was made time limited yielding

$$s_{\text{PCFM}}(t; \mathbf{x}) = \begin{cases} \exp(j \sum_{n=1}^N x_n q(t - nT_s)) & 0 \leq t \leq (N-1)T \\ 0 & \text{otherwise} \end{cases} \quad (2.35)$$

where the N -length vector $\mathbf{x} = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_N]^T$ is comprised of the PCFM parameters which have subsumed the 2π term [54,59]. (2.35) can alternatively be written in terms of the frequency shaping filter such that (2.35) becomes

$$s_{\text{PCFM}}(t; \mathbf{x}) = \begin{cases} \exp \left\{ j \left(\int_0^t g(\tau) * \left[\sum_{n=1}^N \alpha_n \delta(t - (n-1)T_s) \right] d\tau \right) \right\} & 0 \leq t \leq (N-1) \\ 0 & \text{otherwise} \end{cases} \quad (2.36)$$

where the integration stage shows explicitly that the PCFM phase (and the CPM phase function on which it is based) are continuous functions of times. For CPM, $g(t)$ can take on many different shapes yielding different advantages and disadvantages when it comes to spectral containment and demodulation. PCFM however, was constructed as a means to implement poly-phase codes which dictates $q(t)$ should be a rectangular function such that at the end of every subpulse (every T_s interval), the PCFM phase will match the phase of the poly-phase code it was meant to implement.

The shaping filter $g(t)$ is normalized to integrate to 1 such that the shaping filter is

$$g(t) = \begin{cases} \frac{1}{T_s} & 0 < t \leq T_s \\ 0 & 0 < \text{otherwise} \end{cases} \quad (2.37)$$

and the PCFM parameters are bounded on the interval $\alpha_n \in \{-\pi, \pi\}$. To then implement a poly-phase code as a PCFM waveform, the PCFM code is computed as the piecewise difference of the poly-phase code. Since the element by element difference of an N -length vector results in an $(N - 1)$ -length vector, the first element of the N -length PCFM code is set to the first element of the poly-phase code. In this way, the PCFM parameters are akin to instantaneous frequencies. Such a process results in the phase functions in Fig. 2.14(a) where a P4 code [49] has been adapted to the PCFM model to improve containment. 2.14(b) shows the dramatic improvement in spectral containment however 2.14(c) shows a degradation in the PSL level has likewise occurred.

Given the spectral containment capabilities of the PCFM model, it did not take long for its capabilities as a standalone waveform design scheme to be realized. Instead, waveforms optimized based on the PCFM model waveforms were demonstrated in [60] where greedy search methods were shown as an effective means find PCFM codes which result in good spectral containment autocorrelation sidelobes.

More recently, a new representation of the PCFM model has been used to implement highly efficient and effective PCFM optimization schemes which utilize gradient descent methods, the subject of Section 2.5. Consider the integration and convolution steps of (2.36). If these are evaluated, the PCFM phase takes on an exceedingly simple definition

$$\phi(t; \mathbf{x}) = \sum_{n=1}^N \alpha_n b_n(t), \quad (2.38)$$

where each basis function

$$b_n(t) = \int_0^t g(\tau - (n-1)T_s) d\tau \quad (2.39)$$

is the integral of $g(t)$ time shifted by an integer multiple of T_s . Given the rectangular shaping filter,

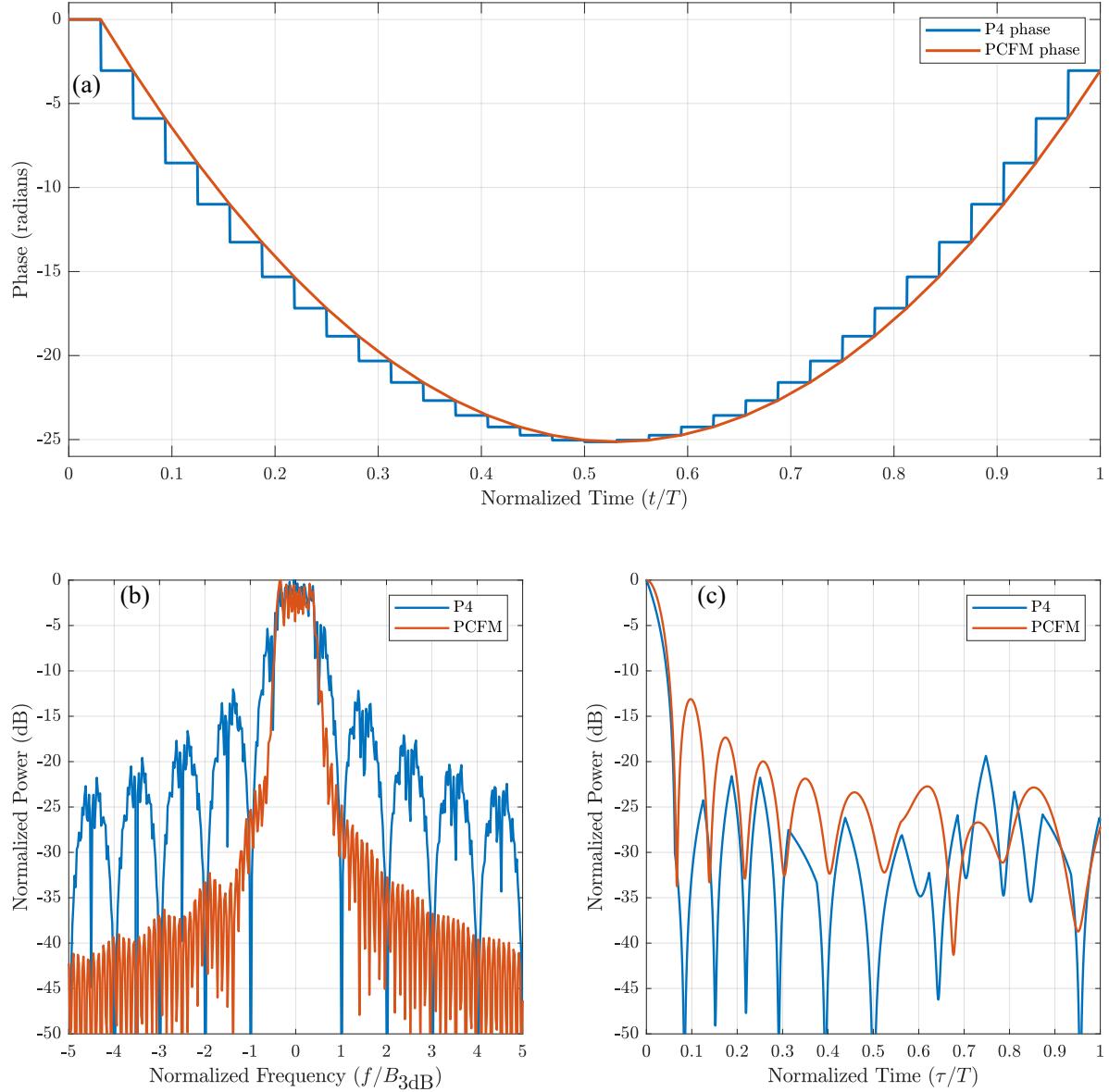


Figure 2.14: A P4 code compared to its implementation as a PCFM waveform

each $b_n(t)$ becomes a time shifted ramp function such that

$$g(t) = \begin{cases} 0 & 0 < t \leq (n-1)T_s \\ (t - (n-1)T_s)/T_s & (n-1)T_s < t \leq nT_s \\ 1 & nT_s < \leq NT_s \end{cases} \quad (2.40)$$

The definition of $b_n(t)$ results in what is known as first order PCFM. Additional integration stages can be incorporated into the PCFM definition to realize function with are continuous not only in phase, but also frequency, chirp-rate, etc, but for the sake of brevity these are not considered here [61]. Likewise, the basis functions in 2.40 could be generalized to any desirable function to realized coded FM (CFM) such as through the use of legendre polynomials [62].

Other options for $b_n(t)$ aside, the convenience of (2.38) is realized when one considers how the PCFM form has to be optimized and handled on a computer. It has to be sampled and the sampled version needs to somehow capture the continuous nature of the PCFM phase that achieves excellent spectral containment. Conveniently, the PCFM form can be easily discretized by sampling the N basis functions and collecting them into the basis function matrix \mathbf{B} . The discretized PCFM waveform is then realized as

$$\mathbf{s} = \exp(j\mathbf{Bx}) \quad (2.41)$$

where B is an $M \times N$ matrix. By bounding the PCFM parameters to the digital frequency space $[-\pi, \pi]$, the number of PCFM parameters, N , becomes approximately equivalent to BT such that $N \approx BT$. The ratio between M and N is then the *oversampling* factor K which represents the ratio between the maximum digital bandwidth allowed by the number of samples in the discrete waveform and the waveform's 3 dB bandwidth. In Fig. 2.14(b), $K = 10$ since the edges of the digital bandwidth extend to $|f| = 5$ and the 3 dB bandwidth occurs at approximately $|f| = 1/2$. The key aspect of the oversampling factor is that it allows for the unambiguous digital representation of the roll-off region. By ensuring the digital waveform exhibits good spectral roll-off in this region, the physical implementation of the waveform will likewise exhibit a good spectral roll-off enabling

a high fidelity, physical radar waveform.

Given these parameters, the original PCFM formulation guaranteed that for an oversampling factor of K , there were also K samples of the ramping portion of each $b_n(t)$. In [63], it was shown that by relaxing the condition $N \approx BT$ such that $N > BT$ further sidelobe suppression could be achieved. However, to guarantee a BT lower than N , the PCFM parameters had to be further restricted to the interval $[-\pi/L, \pi/L]$ where L is the termed the *over-coding* factor. Given [63] utilized a greedy search, to maintain the new interval the search just ignored values outside that interval.

In further optimization work, gradient descent methods were utilized. Since explicitly limiting the interval of the PCFM parameters would greatly complicate the gradient descent implementations if not make them impossible, the optimizations themselves were designed to seek spectrally contained solutions regardless of the degree of over-coding. In [23], this was achieved using a Frequency Template Error (FTE) metric, which had previously been examined in [60], to explicitly optimize for a good spectral roll-off. In [64], only initializations which tend towards spectrally contained solutions were considered. In [65], since the cost function did not tend towards contained solution, no over-coding was used and spectral containment was achieved explicitly through the PCFM form itself. In general, the PCFM form has been used in a variety of ways for a variety of purposes some of which will be discussed in the next section.

2.2.4 FM Noise Waveforms

The concept behind noise waveforms is relatively simple. Given a set of unique individual waveforms where the sidelobe response of one waveform is completely uncorrelated with the response of others then when coherently summed, the sidelobe responses of the different waveforms will add in a noise like manner. Thus, through coherent integration as discussed in section 2.1.3, the sidelobe response can be mitigated just like the noise response.

To demonstrate this, Fig. 2.15 shows the same scene examined in numerous figures thus far, but tested with varying numbers of noise waveforms under three different CPI length ranging from

1 to 10 to 100. Critically, this is a *pulse agile* setting where a unique pulse is transmitted at every PRI. In Fig. 2.15(a) with only one waveform to work with, objects 4 and 6 are obscured and the scene looks much like Fig. 2.7. In Fig. 2.15(b), object 4 has now been revealed from underneath the noise, but critically in Fig. 2.15(c) where 100 unique noise waveforms are used, object 6 has been revealed and there are no visible sidelobes. It is not that these waveforms do not have sidelobes, rather the sidelobes are already at or below the noise floor so they are indistinguishable from the noise. Since they are noise waveforms, the sidelobes then decay with the noise level as more and more waveforms are coherently integrated. This is in direct contrast to Fig. 2.16 where instead of transmitting and coherently integrated 100 unique noise waveforms, the same single noise waveform was transmitted at every PRI. The noise floor is lower, but this time the sidelobes are revealed, not suppressed, and object 6 is obscured.

Given the sidelobe suppression shown in Fig. 2.15, radar noise waveforms offer a unique performance advantage such that better detection can be achieved in the presence or large returns by simply transmitting more unique waveforms. The question then is how to make and design them.

In 1959, Horton answered this question simply for use on aircraft altimeters. Just transmit noise [21]. Surely if the radar waveform itself is literally noise then it will exhibit the desired noise like characteristics. However, just transmitting noise implies significant AM resulting in a high PAPR making it poorly suited for high power operation. To overcome this, one could consider the noise waveform implementations suggested by [27–29, 66] where the noise modulation is in the phase such that the signals are constant amplitude and thus amenable to high power transmission. This is the first mention of what are known as FM noise waveform which are the focus of this work. The rest of this section introduces some of the latest work on this topic.

2.2.4.1 Psuedo-Random Optimized FM (PRO-FM)

To achieve good autocorrelation sidelobes and spectral roll-off, PRO-FM utilizes an alternating projection approach where a candidate waveform is projected in an iterative fashion between the

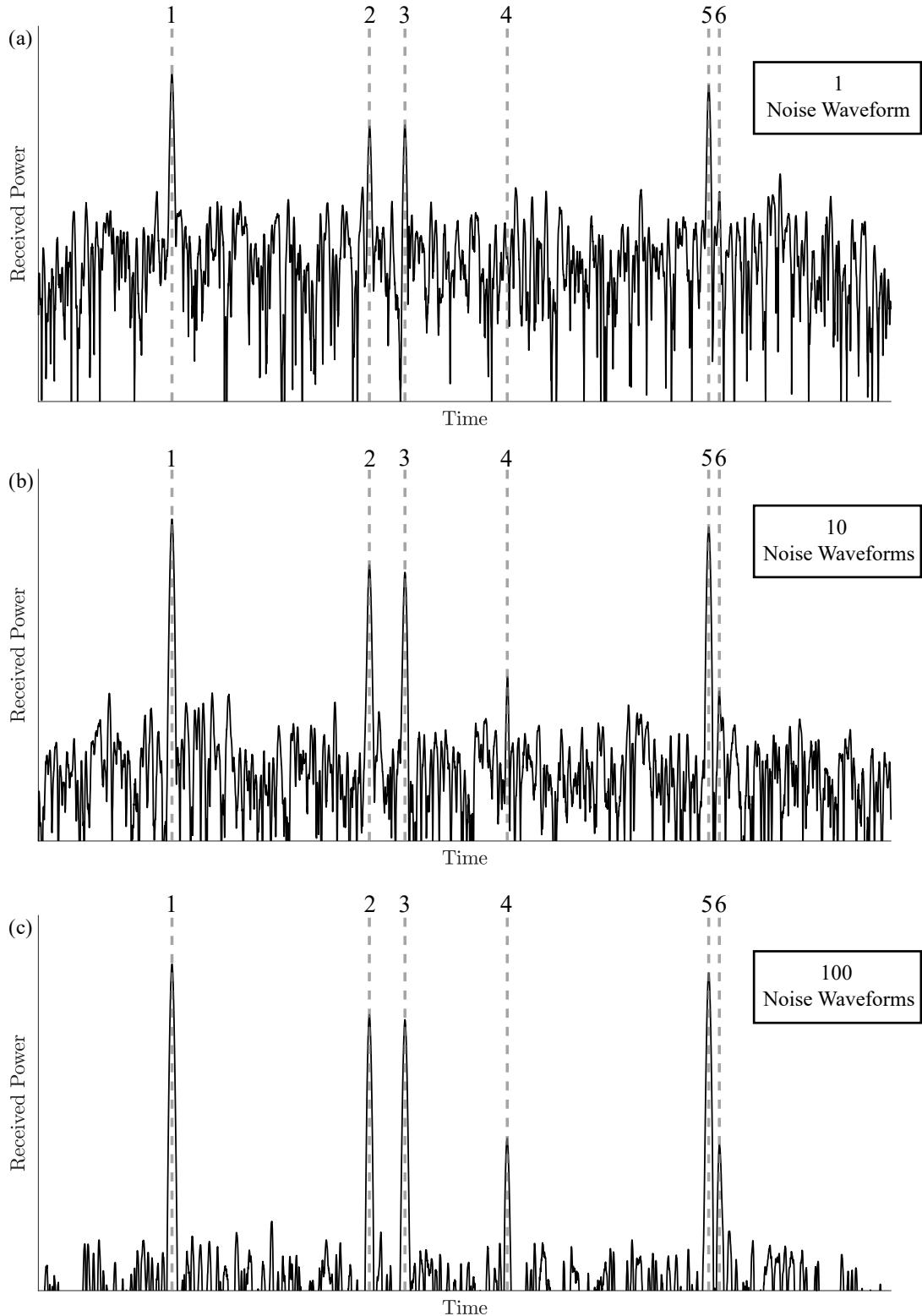


Figure 2.15: The noisy radar data from Fig. 2.2 with varying CPI sizes of 1, 10 and 100 corresponding to (a), (b), and (c) respectively

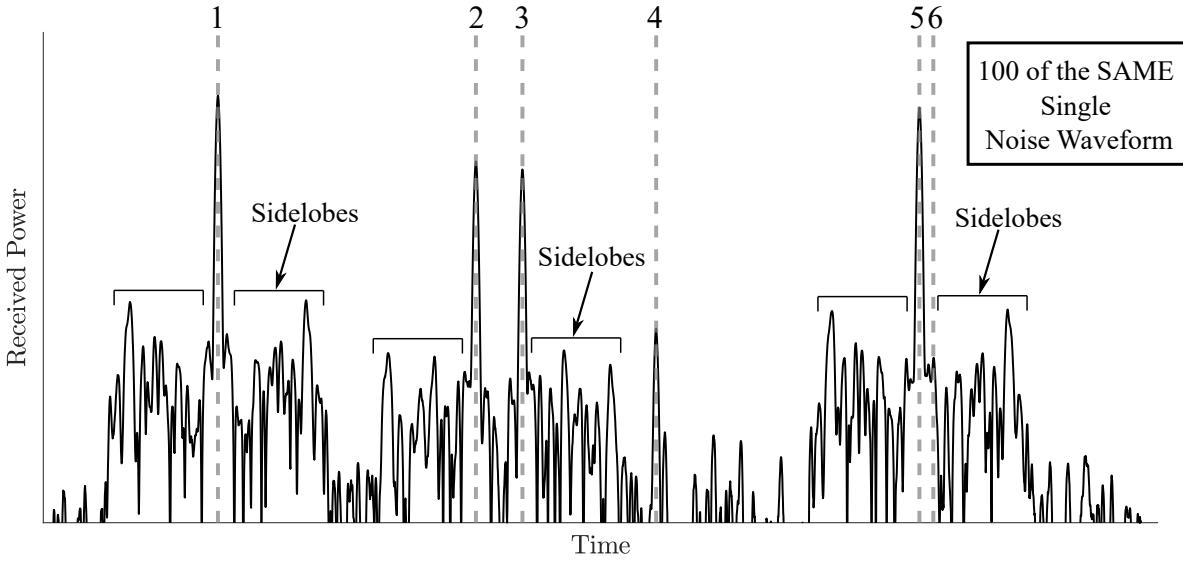


Figure 2.16: The noisy radar data from Fig. 2.2 but the transmit waveform is a single noise waveform with a 100 pulse CPI

time and frequency domains [25, 26]. For the k iteration, the process is defined by the alternating application of

$$r_{k+1}(t) = \mathbb{F}^{-1}\{|G(f)| \exp(j\angle \mathbb{F}\{p_k(t)\})\} \quad (2.42)$$

and

$$p_{k+1} = u(t) \exp(j\angle(r_{k+1}(t))) \quad (2.43)$$

where \mathbb{F} and \mathbb{F}^{-1} are the Fourier and inverse Fourier transforms respectively, $|G(f)|$ is some desired spectral envelope, $u(t)$ is a constant amplitude envelope, and \angle extracts the angle of the argument. The key to producing good FM noise waveforms from (2.42) and (2.43) is the choice of $|G(f)|$.

In [25, 26], $|G(f)|$ was chosen to be a Gaussian envelope. In terms of the spectrum a Gaussian envelope exhibits decent roll-off aiding in spectral containment. In terms of autocorrelation performance, since the inverse Fourier transform of a Gaussian function is likewise Gaussian, if the PSD of the waveform is Gaussian then the autocorrelation will likewise be Gaussian resulting in zero autocorrelation sidelobes. In practice, the optimization will result in a good spectral match, but not perfect.

To demonstrate the effectiveness of the PRO-FM approach. 100 FM noise waveforms were optimized according to (2.42) and (2.43) for a BT of 128. Adopting the PCFM notation for oversampling with respect to the 3 dB bandwidth, K was set to 4. These parameters and optimization results are shown in Figs. 2.17 and 2.18 which show the resulting Spectra and autocorrelations respectively.

In Fig. 2.17, the RMS spectrum of the 100 PRO-RM waveforms matches the spectral template out to about $|f| < 1$ (2 times the 3 dB bandwidth) before it deviates from the template due to the rectangular pulse shape. Thus the PRO-FM waveforms exhibit decent spectral roll-off. In 2.18, perhaps the most impressive aspect of PRO-FM is its RMS autocorrelation response. For a generic noise waveform such as white Gaussian noise, the RMS autocorrelation sidelobes should be $-10\log_{10}(BT)$. After optimization, the PRO-FM RMS sidelobes near the mainlobe are a full order of magnitude better at $-20\log_{10}(BT)$. Then to demonstrate the noise like coherent integration of FM noise waveform sidelobes, the coherent integration of the autocorrelations results in an approximately $10\log_{10}(100) = 20$ dB sidelobe level improvement versus the RMS sidelobes since 100 autocorrelation were summed.

These waveforms have been demonstrated experimentally and they have been utilized for numerous applications such as for spectral notching to limit the impact of in band interference [67–69], radar and communications spectrum sharing [70–72], simultaneous dual-polarized radar [73], non-linear harmonic radar with FM-noise waveforms [74], and even random movement radar which mimics the human eye [75].

Since the formulation of PRO-FM, a few other approaches have been formulated to optimize FM noise waveforms such as Temporal Template Error (TTE) waveforms which utilize a hybrid approach between a gradient based optimization and a projection [24], and Logarithmic frequency template error (Log-FTE) waveforms which utilize a gradient descent and the PCFM waveform model approach to achieve excellent spectral containment [23].

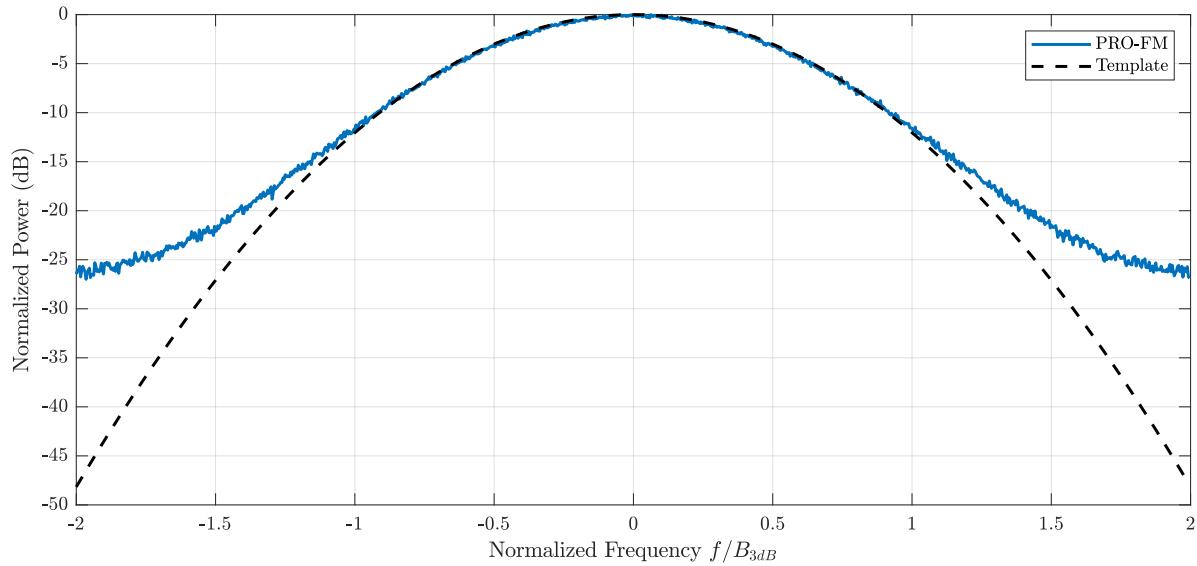


Figure 2.17: RMS spectrum of 100 optimized PRO-FM waveforms and their desired template

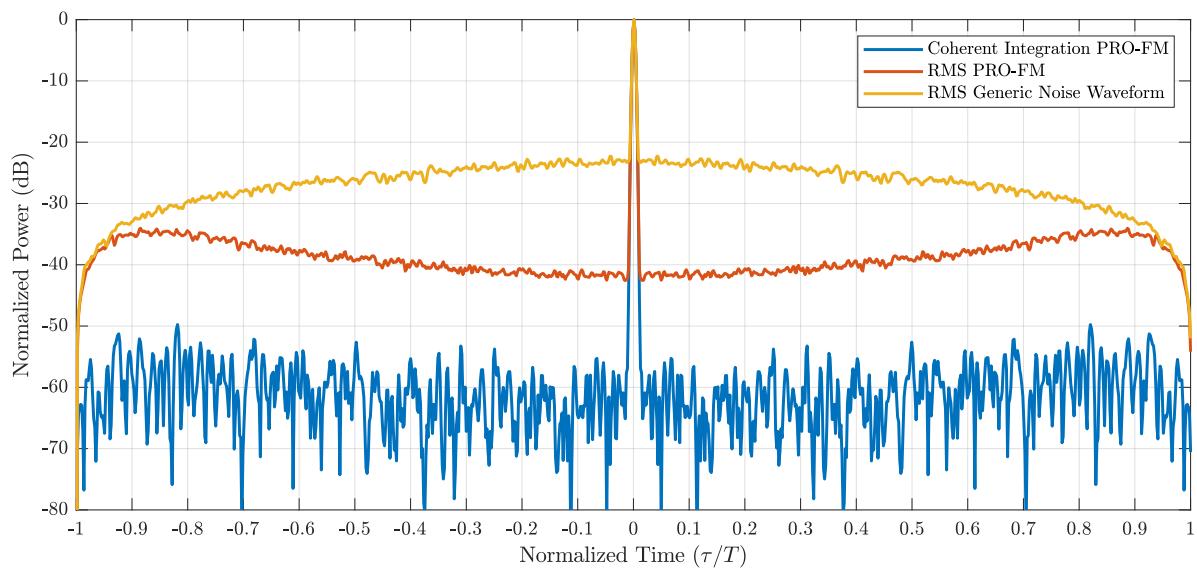


Figure 2.18: Coherently integrated and RMS autocorrelation of 100 PRO-FM waveforms compared to the RMS autocorrelation of 100 generic noise waveforms

2.2.4.2 Process Based FM Noise Waveforms

As opposed to the previously discussed FM noise waveform approaches which are optimized or designed on an individual level, process based FM noise waveforms are more similar to drawing a sample function from a random process. Some examples of this kind of implementation are [27–30,66,76], where the phase function of the waveforms are based on a Gaussian random process and instantiating the phase function is as simple as using an RNG to generate the phase and thus the waveforms. Using this kind of methodology, creating waveforms with a given set of predefined characteristics is very cheap computationally speaking.

The majority of this dissertation expands these types of waveforms and on the work in [30]. Chapter 3 discusses these kinds of waveforms within the classical terminology of random variables and processes and how they translate to a more radar centric perspective. Chapters 4-6 put these concepts to work in designing what are called waveform generating functions (WGF) to achieve desirable radar waveform characteristics.

2.3 Continuous Wave Radar

In a pulsed mode, over a PRI the energy on target will be proportional to the average power of the transmission. If the pulse is constant amplitude and it has power P_t , then the average power over a PRI is

$$P_{\text{avg}} = \frac{P_t t_d}{T_{\text{PRI}}} \quad (2.44)$$

where t_d is the duty factor and T_{PRI} is the length of the PRI in seconds. The duty factor arises from the fact that the radar is only transmitting a fraction of the time. If the transmitter is only active 5% of the time, then $t_d = 0.05$. The distinction between a pulsed mode and continuous wave (CW) radar lies in the duty factor. A CW radar transmits 100% of the time so $t_d = 1$. In the context of (2.44), this means a CW radar can transmit at a fraction of the power of a pulsed radar and achieve the same power on target. However, transmitting continuously presents other issues [31].

In a pulsed mode, the highly sensitive, delicate receive electronics have to be shut off during

transmission to protect them from the extremely high output power that would otherwise break them. This results in a blind range which is proportional to the pulse length. However after transmission, the receiver can listen without being drowned out by the transmitter. For a CW radar, this is no longer the case. The transmitter and receiver are necessarily simultaneously operating so the receiver has to contend with the direct path from the transmitter. Even though CW radars can get away with transmitting at lower powers overall the, the direct path signal, if unmitigated, would be orders of magnitude stronger than any returns. To combat this, a large part of CW radar design is devoted to mitigating direct path interference [31]. Still, this constraint leads to CW radars usually be low power systems with relatively short operating ranges.

In terms of the transmit waveform and processing, CW operation enables some interesting modes. In frequency modulated continuous wave (FMCW) radar, the transmission is frequency modulated such as with continuous repetitions of an LFM or an alternating sequence of up-chirps and down-chirps. With appropriate processing, range and velocity information can be obtained [77].

In terms of noise radar, the approaches proposed by [21, 66] were CW based. The original PRO-FM formulation was likewise for CW operation [25] where matched filtering was used by cutting the transmit waveform into contiguous segments. The advantage CW radar provides to noise waveforms is an extremely high BT . In general, for noise waveforms their sidelobe levels are proportional to their BT . In a pulse mode we can consider the concept of an aggregate BT which is the BT of an individual waveform times the number of waveforms in a CPI. This aggregate BT will be proportional to the final sidelobe level after coherent integration. However, if each waveform were longer or more were transmitted during a CPI, then the aggregate BT would increase for a given bandwidth. The limit of this process is CW radar which intuitively achieves the highest BT for a given bandwidth. For noise radar this is important because it mitigates the interference effect of the direct path by lowering its sidelobes. If the sidelobes of the direct path are low enough, they will not mask targets of interest. For the CW FM noise waveforms designed in Chapter 6, the same segmented matched filtering process will be used.

2.4 The Ambiguity Function and Doppler Processing

This section will provide a brief discussion of the ambiguity function and doppler processing. This is important for the open air results of Chapters 4-6 where the range Doppler plots are formed from the open-air testing results.

2.5 Gradient Descent Techniques

The different waveform design schemes employed in this work all utilize a the gradient descent class of optimization techniques. In mathematical optimization problems the goal is to minimize (or maximize) some cost function value which measures some attribute of a problem. Often times this is described as minimizing the error, but it could just as easily mean attempting to maximize profits. In general, the mathematical optimization problem can be defined as [78]

$$\begin{aligned} & \text{minimize} && J(\mathbf{x}) \\ & \text{such that} && f_i(\mathbf{x}) \leq c_i, \quad i = 1, 2, \dots, M \end{aligned} \tag{2.45}$$

where \mathbf{x} is a vector of parameters, each $f_i(\mathbf{x}) \leq c_i$ is a constraint on the parameters where each c_i is a constant. The goal of any optimization process is then to find \mathbf{x}^* such that

$$J(\mathbf{x}^*) \leq J(\mathbf{x}) \text{ for all } \mathbf{x}. \tag{2.46}$$

where the $*$ indicates \mathbf{x}^* minimizes the cost function value. As defined in 2.46, \mathbf{x}^* is a *global* minimizer of $J(\mathbf{x})$. That is there is no \mathbf{x} such that $J(\mathbf{x})$ is less than $J(\mathbf{x}^*)$. In practice however, it is often only possible to find an \mathbf{x} such that

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \text{ for all } \mathbf{x} \in \mathcal{N} \tag{2.47}$$

where \mathcal{N} is an infinitesimal open set (local neighborhood) containing \mathbf{x}^* [11].

In general, the gradient descent optimization process is iterative. At the k th iteration, the set of

optimizable parameters is updated as

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (2.48)$$

where \mathbf{x}_k is the vector of parameters, \mathbf{p}_k is the search direction and α_k is the step size or how far to move \mathbf{x}_k in the direction of \mathbf{p}_k . Generally speaking, the step direction \mathbf{p}_k is chosen as

$$\mathbf{p}_k = \begin{cases} -\nabla J_k & \text{when } k = 0 \\ -\nabla J_k + \beta_k \mathbf{p}_{k-1} & \text{otherwise} \end{cases} \quad (2.49)$$

such that the search direction is a linear combination of the negative of the current gradient and the previous step directions. Gradient descent techniques usually vary in their calculation of β . One popular family of techniques are non-linear conjugate gradient (NLCG) techniques such as was used to design PCFM waveforms in [64]. An excellent survey of NLCG methods can be found in [79].

For this work however, the relatively simple heavy ball gradient descent method is used where β is defined such that

$$0 < \beta < 1. \quad (2.50)$$

In this way, the search direction has a kind of inertia where the search direction can not change too quickly [80]. If the resulting search direction results in a direction of ascent, then it is reset to the negative of the current gradient which is guaranteed to decrease the cost function value to some degree. For the cost functions in this work, this method has been found to be a simple, efficient means of reaching a locally optimal solution.

For the final dissertation, this section will be expanded upon to provide a more comprehensive overview of optimization and gradient descent techniques.

Chapter 3

Stochastic Waveform Analysis and Evaluation

The purpose of this chapter is to develop a mathematical framework for analyzing and evaluating radar signals with stochastic characteristics and specifically the waveform generating function (WGFs) designed in the rest of the work. The concepts developed here build directly off of the fundamentals of random signals analysis covered in any random signals text such as [35, 81–83] to name a few. The usual progression in these texts is to begin with a basic overview of random variables and random processes before venturing into more advanced topics. The approach taken here is no different.

Sections 3.1 and 3.2 provide a brief overview of random variables and random processes respectively. Section 3.3 applies the topics in sections 3.1 and 3.2 to radar signals and radar signal processing. Section 3.4 defines various metrics for evaluating random processes as radar signal generating functions.

3.1 Random Variables

Random variables describe the possible outcomes of an experiment such as rolling a fair die or the toss of a fair coin. In the case of the fair die, there are six equally probable outcomes while for the fair coin there are two equally probable outcomes. In either case the experiment, that is the roll of the die or the flip of the coin, is entirely described by their probability mass function or PMF. In

the case of the fair die, this can be written as

$$P_X(x = x_i) = \begin{cases} 1/6 & x_i = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}. \quad (3.1)$$

(3.1) can read as, "The probability that the random variable X is equal to x_i is $1/6$ for $x_i = 1, 2, 3, 4, 5, 6$ and 0 otherwise". Additionally, (3.1) has several other properties which generalize to any probability mass function. Notice that for (3.1), and any PMF in general,

$$P_X(x = x_i) \geq 0 \quad (3.2)$$

and

$$\sum_{i=1}^n P_X(x = x_i) = 1. \quad (3.3)$$

In the case of (3.1), $n = 6$. To further characterize PMFs and random variables in general, it is useful to define what are called *moments* which are calculated via the expectation operator, $E[\cdot]$, to evaluate the *expected value*. The most commonly used moments are the mean and variance defined as

$$E[X] = \sum_{i=1}^n x_i P_X(x = x_i) = \mu_X \quad (3.4)$$

and

$$E[(X - \mu_X)^2] = \sum_{i=1}^n (x_i - \mu_X)^2 P_X(x = x_i) \quad (3.5)$$

respectively, where the mean is the first moment and the variance is the second moment. In plain language, (3.4) can be understood as the average value of X while (3.5) can be understood as the average squared deviation of X from its average value.

3.1.1 Covariance and Independence

Practically speaking, random experiments rarely exist in a vacuum. In fact it may be necessary to consider jointly two or more experiments at the same time. Instead of rolling one die what if

two fair die are rolled? In this case of random variables X_1 and X_2 for the first and second die respectively, the joint PMF is

$$P(x_1 = x_{1,i}; x_2 = x_{2,i}) = \begin{cases} 1/36 & x_{1,i} = 1, 2, 3, 4, 5, 6; \quad x_{2,i} = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}. \quad (3.6)$$

An important concept that comes from considering multiple random variables together is the concept of dependence (or independence). In (3.6) the two experiments are intuitively independent. That is, the result of one die roll does not effect the other. Consequently, in (3.6) the probability of

$$P(x_1 = x_{1,i}; x_2 = x_{2,i}) = P_{X_1}(x_1 = x_{1,i})P_{X_2}(x_2 = x_{2,i}). \quad (3.7)$$

Independence also indicates an important result for expectations. For independent random variables

$$\mathbb{E}[X_1 X_2] = \mathbb{E}[X_1]\mathbb{E}[X_2]. \quad (3.8)$$

In other words the expectation of their product can be written a product of their expectations. Now consider a case where the random variables are dependent. To do so, define a new random variable as

$$Y = X_1 + X_2 \quad (3.9)$$

where X_1 and X_2 are the die rolls from before and Y is the sum of their results. The relationship between Y and either X_1 or X_2 can be measured as the covariance between them defined as

$$\sigma_{XY} = \mathbb{E}[(Y - \mu_Y)(X - \mu_X)] \quad (3.10)$$

which is a measure of how Y and X move together. In words, how do they vary together around their respective means? If they both tend to either be more or less than the their means then σ_{XY} will be positive. If they both tend to have opposite signs about their means then σ_{XY} will be negative. If they have no tendency to be either positive or negative relative to their means and each

other, then σ_{XY} will be 0. The concept of covariance is crucially important to random variables and random processes and almost any subject where they are used.

3.1.2 Continuous Random Variables and the Characteristic Function

Thus far the discussion on random variables has focused exclusively on discrete random variables and their PMFs. However, many natural processes such as the thermal noise in a circuit are instead described by continuous random variables and the *probability density function* (PDF). The continuous analogue to the uniform (equal probability) PMF of the die roll in (3.1), is the continuous uniform distribution defined as

$$f_X(x) = \begin{cases} 1/(b-a) & a < x < b \\ 0 & \text{otherwise} \end{cases} \quad (3.11)$$

for $a < x < b$. The outcome of an experiment described by 3.11 is equally likely to be any value in the range $a < x < b$. Like the PMF, the PDF is also strictly positive and integrates to 1. The moments of a continuous random variable are defined by an integral rather than a sum as

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} xf_X(x)dx = \mu_X, \quad (3.12)$$

while the variance is defined as

$$\mathbb{E}[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx = \sigma_X^2, \quad (3.13)$$

and the covariance becomes

$$\mathbb{E}[(X - \mu_Y)(X - \mu_X)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{X,Y}(x,y) dx dy = \sigma_{XY}. \quad (3.14)$$

Another useful tool, and one that will be used extensively in the development of the waveform

design techniques in this work, is the characteristic function, which is defined

$$\Psi_X(\omega) = \text{E}[\exp(j\omega X)] = \int_{-\infty}^{\infty} \exp(j\omega x) f_X(x) dx. \quad (3.15)$$

This form applies to both continuous and discrete random variable, but for the discrete case the PMF has to be written on a continuous number line using dirac delta functions. The study of random variables is much more extensive and rigorous than what has been presented here, but this should be sufficient for the content of this work. More details can be found in any statistics or random processes text such as [35, 82].

3.2 Random Processes

One way to look at a random processes is as a sequence (continuous or discrete) of random variables. At any instant of time, the random process defines a random variable with its own PDF or PMF and perhaps a dependence on other parts of the random process (other random variables). To see this, consider the *random walk* denoted as $X[n]$ (square brackets indicate discrete time) and is defined as follows. At each instant of time, n , a fair coin is flipped. If it comes up heads, one is added to the counter. If it comes up tails, one is subtracted from the counter. At time zero, the counter is set to 0 such that $X[0] = 0$. The table below shows the PMFs of $X[n]$ for the first several times n . For clarity, the zero valued cells have been filled with '-'. At each time n the value of $X[n]$ is described by the PMF in each column of table 3.1. Something that is not immediately evident from the table however, is the dependence between the different random variables at each time n . Based on the definition of the random walk, if $X[n_0] = 4$, then $X[n_0 + 1]$ can only equal either 3 or 5 with equal probability, but there is really no way to see just from the content of Table 3.1.

This is much more apparent when viewing random processes from a more holistic perspective. Rather than looking at the process as a sequence of correlated random variables, it can be viewed as a collection of member function which each have their own probability of occurring. For the random walk some possible partial sequences are

	n					
	0	1	2	3	4	5
5	-	-	-	-	-	1/32
4	-	-	-	-	1/16	-
3	-	-	-	1/8	-	5/32
2	-	-	1/4	-	1/4	-
1	-	1/2	-	3/8	-	5/16
$X[n]$	0	1	-	1/2	-	3/8
	-1	-	1/2	-	3/8	-
	-2	-	-	1/4	-	1/4
	-3	-	-	-	1/8	-
	-4	-	-	-	-	1/16
	-5	-	-	-	-	-

Table 3.1: Random Walk PMFs for $n = 0, 1, 2, 3, 4, 5$

n	=	0,	1,	2,	3,	4,	5,	6,	...
$x_a[n]$	=	0,	-1,	-2,	-1,	0,	-1,	-2,	...
$x_b[n]$	=	0,	1,	2,	3,	4,	3,	4,	...
$x_c[n]$	=	0,	1,	0,	-1,	-2,	-3,	-2,	...

Table 3.2: Random walk partial sequences for $n = 0, 1, 2, 3, 4, 5, 6$

where x_a or any other letter subscript for that matter indicates a unique sample function of the random process $X[n]$. In this form, the dependence on previous values is more obvious. Since the counter is only able to change by one unit at every time step, the sample functions only change by one unit from time n to time $n + 1$. The next several subsections detail metrics, properties, and estimation techniques for stochastic processes.

3.2.1 Correlation Functions

Table 3.1 showed how a stochastic process at a given function of time is itself a random variable. Table 3.2 anecdotally showed how the random variables at different times can be correlated with each other. In other words, if for a given sample function, the value at time n_1 is known and n_1 and

n_2 are correlated, then something can be said about the sample function at time n_2 . For a random process this correlation between samples at different times is described by the two dimensional autocorrelation function defined as

$$R_{XX}[n_1, n_2] \triangleq E[X(n_1)X^*(n_2)]. \quad (3.16)$$

or in the continuous case with continuous time t

$$R_{XX}(t_1, t_2) \triangleq E[X(t_1)X^*(t_2)]. \quad (3.17)$$

where $(\cdot)^*$ indicates complex conjugation in the case the process is complex. The autocorrelation can be generalized to describe the correlation between difference random processes, in which case (3.16) and (3.17) become cross-correlation functions as

$$R_{XY}[n_1, n_2] \triangleq E[X(n_1)Y^*(n_2)]. \quad (3.18)$$

or in the continuous case with continuous time t

$$R_{XY}(t_1, t_2) \triangleq E[X(t_1)Y^*(t_2)]. \quad (3.19)$$

(3.16-3.19) are two dimensional functions of n_1 and n_2 or t_1 and t_2 .

3.2.2 Stationarity and the Power Spectral Density (PSD)

A powerful concept which can greatly reduce the complexity of analyzing a stochastic process is stationarity. In general, stationarity comes in a couple of different forms but they are based on the moments of a process being invariant in time. For the first order moment stationarity indicates the mean is a constant μ_X for all time,

$$E[X(t)] = \mu_X. \quad (3.20)$$

For the autocorrelation function (second moment) this indicates

$$R_{XX}(t_1, t_2) = R_{XX}(t_1 - t_2) \quad (3.21)$$

meaning the autocorrelation is only a function of the relative time between each sample rather than the absolute time of the samples. This makes the autocorrelation a simpler, one dimensional function. If this pattern, continuous for all higher order moments of $X(t)$, then $X(t)$ is said to be Strict Sense Stationary (SSS). If only the first and second moments are stationary, as in (3.21), then $X(t)$ is said to be Wide Sense Stationary (WSS).

In addition to simplifying the form of the autocorrelation, if a process is WSS then its PSD is likewise constant in that its spectral content is constant as a function of time. By the Weiner-Kinchine relation, the PSD is [35]

$$S_{XX}(f) = \int_{-\infty}^{\infty} R_{XX}(\tau) \exp(-j2\pi f\tau) d\tau. \quad (3.22)$$

If the autocorrelation were a function of multiple variables, t_1 and t_2 , different cuts of $R_{XX}(t_1, t_2)$ would yield unique PSDs indicating the spectral content of $X(t)$ changes as a function of time. These results can likewise be applied to cross-correlations and cross-power spectral densities as long as $X(t)$ and $Y(t)$ are jointly stationary. These take the form

$$R_{XY}(t_1, t_2) = R_{XY}(t_1 - t_2). \quad (3.23)$$

and

$$S_{XY}(f) = \int_{-\infty}^{\infty} R_{XY}(\tau) \exp(-j2\pi f\tau) d\tau. \quad (3.24)$$

respectively. Their discrete counterparts are

$$R_{XY}[n_1, n_2] = R_{XY}[n_1 - n_2] \quad (3.25)$$

and

$$S_{XY}(f) = \sum_{\tau=-\infty}^{\infty} R_{XY}[\tau] \exp(-j2\pi f\tau). \quad (3.26)$$

3.2.3 Estimation of the Autocorrelation and the PSD

Often times when analyzing a stochastic process its properties are unknown a priori so it is necessary to estimate them by observing the behavior of sample functions. Broadly speaking, this estimation process falls into one of two categories, either model-based or model-free estimation. In model-based estimation, the process being analyzed is assumed to have some sort of structure. If this assumption is correct, then parameters such as the autocorrelation and PSD can be effectively estimated with relatively little data. For example, model based estimation is used for the efficient digitization of human speech and adaptive filtering [81, 84]. For the purposes of this work and in the context of basic radar processing such as match filtering and spectral estimation, it is more important to consider model free estimation.

Before discussing the estimation of the PSD and the autocorrelation it is important to note that the estimation procedures presented here assume the underlying process being estimated is WSS. If only one sample function can be measured, then the process is also assumed to be ergodic. The estimation is carried out on the samples of a process so the development in this section is necessarily discrete.

3.2.3.1 Power Spectral Density Estimation

For either the autocorrelation or the PSD, the estimation process consists of estimating the second moments of numerous, individual random variables which are the sample data. For a stationary random process, the mean can be estimated as

$$\hat{\mu}_X = \frac{1}{N} \sum_{n=1}^N x[n] \quad (3.27)$$

where the hat of $\hat{\mu}_X$ indicates this is an estimate of μ_X or the sample mean and each $x[n]$ is a discrete sample of the random process $X(t)$ or the discrete process $X[n]$. (3.27) is *unbiased*, that is

$$E[\hat{\mu}_X] = \frac{E[\sum_{n=1}^N x[n]]}{N} = \frac{N\mu_X}{N} = \mu_X \quad (3.28)$$

In other words, (3.27) will approach exactly μ_X (assuming the process is stationary).

Likewise, for a WSS process, an estimator of the PSD can be formed as the DFT of a finite length vector of samples of some random process whether it be continuous or discrete. This DFT estimate of the spectrum is defined as

$$\hat{X}_{XX}(f) = \sum_{n=0}^{N-1} x[n] \exp(-j2\pi fn) \quad (3.29)$$

and the PSD estimate is then

$$\hat{S}_{XX}(f) = |\hat{X}_{XX}[f]|^2 = \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} x[n] x^*[k] \exp(-j2\pi f(n-k)) \quad (3.30)$$

This form of spectral estimate is called a periodogram. The bias of this estimator is evaluated by taking the expectation yielding

$$E[\hat{S}_{XX}(f)] = \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} E[x[n] x^*[k]] \exp(-j2\pi f(n-k)) \quad (3.31)$$

where by linearity, the expectation has been moved inside of the double summation. Since the random process is assumed to be stationary, the expectation in 3.31 is a function of only the difference between n and k such that

$$E[\hat{S}_{XX}(f)] = \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} R_{XX}[n-k] \exp(-j2\pi f(n-k)). \quad (3.32)$$

For each lag, $\ell = n - k$, there are $N - |\ell|$ relevant terms in the summation such that 3.32 can be

rewritten and normalized by N as

$$E[\hat{S}_{XX}(f)] = \sum_{\ell=-(N-1)}^{N-1} \left(1 - \frac{|\ell|}{N}\right) R_{XX}[\ell] \exp(-j2\pi f\ell) \quad (3.33)$$

From 3.33, the expectation of the DFT estimate of the PSD is the DFT of a windowed form of the process's autocorrelation function. This indicates the periodogram estimate is biased. The window takes the form

$$w[\ell] = \begin{cases} \left(1 - \frac{|\ell|}{N}\right) & |\ell| < N \\ 0 & \text{otherwise} \end{cases}. \quad (3.34)$$

This window is a discrete triangular function and is known as a Bartlett window [81–84]. If the DFT is evaluated in 3.33, the expectation of the periodogram results in a convolution of the processes PSD and the DFT of the Bartlett window such that

$$E[\hat{S}_{XX}(f)] = \frac{1}{N} \left(\frac{\sin(\pi f N)}{\sin(\pi f)} \right)^2 * S_{XX}(f) \quad (3.35)$$

where $*$ denotes convolution. In other words, the mean value of the estimate provided by the periodogram will never approach the true PSD of $X(t)$; however, it can be made arbitrarily close by increasing the number of DFT points.

In addition to the bias it is important to consider the variance of an estimator. That is, on average how far will the estimate deviate from the mean in a mean squared error sense. In general this is a difficult quantity to measure as the result will depend on the fourth order moments of the process, however it can be computed exactly in special cases such as with WGN [35, 81, 82]. In general the periodogram realizes poor performance in terms of variance and bias. To address this, various other spectral estimation techniques have been proposed [81]. With respect to periodogram methods, windowing techniques are commonly applied which work by applying windows to the

autocorrelation estimate such that

$$E[\hat{S}_{XX}(f)] = \sum_{\ell=-(N-1)}^{N-1} w[\ell] R_{XX}[\ell] \exp(-j2\pi f\ell) \quad (3.36)$$

where $w[\ell]$ is some real function which is zero for $|\ell| > N$. Through a judicious choice of $w[\ell]$ a trade off can be made between bias and variance. Any spectral estimation or random signals processing text will discuss various windows that have been proposed over the years [35, 81–83, 85].

3.2.3.2 Autocorrelation Estimation

Because of their Fourier transform relationship, it is difficult to discuss the PSD and the autocorrelation as separate topics. Case and point, from the form of PSD estimate in the previous section it should be intuitive how the autocorrelation can be estimated at least in a biased form. However, An unbiased estimate of the autocorrelation can be defined as

$$\hat{R}_{XX}[\ell] = \frac{1}{N - |\ell|} \sum_{k=1+\ell}^N x[k]x^*[k-\ell] \quad \ell < N \quad (3.37)$$

where by definition of the autocorrelation of a stationary process the negative lags are $R_{XX}[\ell] = R_{XX}^*[-\ell]$. The fractional term in (3.37) normalizes the sum by the number of terms in the sum for lag ℓ realizing an unbiased estimate. The issue with (3.37) lies in the variance. At values of ℓ near N , the number of samples used to estimate $\hat{R}_{XX}[\ell]$ is very small such that it will have a high variance. Alternatively, the autocorrelation can be estimated as

$$\hat{R}_{XX}[\ell] = \frac{1}{N} \sum_{k=1+\ell}^N x[k]x^*[k-\ell] \quad \ell < N \quad (3.38)$$

which is a biased estimator such that the expectation of (3.38) is

$$E[\hat{R}_{XX}[\ell]] = \left(1 - \frac{|\ell|}{N}\right) R_{XX}[\ell] \quad |\ell| < N \quad (3.39)$$

Interestingly, (3.39) is exactly the form of the scaled autocorrelation in (3.33). The windowing of the autocorrelation reduces the influence of the high variance lags of the autocorrelation estimate. Additionally, computing the periodogram from (3.38) guarantees the PSD estimate will be positive [81]. Finally, (3.38) is the resulting estimate from matched filtering a data vector. For these reasons, in the context of periodograms and radar signal processing (3.38) is the desired autocorrelation estimator.

3.3 Stochastic Processes as Radar Signals

At this point it is good to recall the reasoning behind using noise-like radar waveforms. As discussing in Section 2.1.2, the match filtering process results in autocorrelation sidelobes. These sidelobes can hide objects of interest and degrade radar performance. Noise-like waveforms address this through uniqueness. Since each waveform has unique, uncorrelated sidelobes with each other. Their coherent summation through pulse integration will, on average, result in lower side-lobe levels. The amount of sidelobe level improvement is a function of BT . In this pulsed, case this is the aggregate BT of all the pulses integrated, and in the CW case this is the BT of the entire radar *CPI*.

For the pulsed case this begs the question, what is the average autocorrelation performance of a single waveform and what improvement can be expected when integrating pulses? For the CW case the question is what autocorrelation level can be expected for a given BT ? In either case the answer depends on processes's autocorrelation function.

3.3.1 Pulsed Stochastic Processes as Radar Signals

A complex baseband pulsed stochastic process is time limited such that

$$S(t) = \begin{cases} S_p(t) & 0 < t < T \\ 0 & \text{otherwise} \end{cases} \quad (3.40)$$

During the interval $0 < t < T$, $S_p(t)$ is random, but outside of this interval it is deterministically 0. Because of this, the ensemble autocorrelation function of $S(t)$ is zero valued beyond the stochastic interval such that

$$R_{SS}(t_1, t_2) = \begin{cases} R_{S_p S_p}(t_1, t_2) & 0 < t_1 < T, 0 < t_2 < T \\ 0 & \text{otherwise} \end{cases} \quad (3.41)$$

For radar processing purposes, if a sample function of $S(t)$ were used as a radar waveform, the matched filter response of the function is computed as

$$r(\tau) = \frac{1}{T} \begin{cases} \int_0^{T+\tau} s(t)s^*(t-\tau)dt & -T < \tau \leq 0 \\ \int_\tau^T s(t)s^*(t-\tau)dt & 0 < \tau < T \\ 0 & \text{otherwise} \end{cases}. \quad (3.42)$$

where the bounds of the integrals have been carefully selected to omit the zero valued portions of the sample function $s(t)$, and $r(\tau)$ is the sample matched filter response¹. Interestingly, this is the exact form of the autocorrelation estimator of a stationary process in the previous section. With this in mind, the expectation of (3.42) is

$$\mathbb{E}[r(\tau)] = \frac{1}{T} \begin{cases} \int_0^{T+\tau} R_{SS}(t, t-\tau)dt & -T < \tau \leq 0 \\ \int_\tau^T R_{SS}(t, t-\tau)dt & 0 < \tau < T \\ 0 & \text{otherwise} \end{cases}. \quad (3.43)$$

The right hand side of (3.43) is the integral over the contours of $R_{SS}(t_1, t_2)$ defined by $t_1 - t_2 = \tau$ over the non-zero portions of $R_{SS}(t_1, t_2)$. Since $S(t)$ is not necessarily stationary the matched filtering operation results in an averaging operation over $R_{SS}(t_1, t_2)$, such that $\mathbb{E}[r_{SS}(\tau)]$ is the average expected correlation between samples in $S(t)$ separated by τ seconds. If $S(t)$ happens to be WSS over the course of the pulse, then (3.42) becomes exactly the autocorrelation estimation function

¹To avoid a conflict in terminology, $R_{SS}(t_1, t_2)$ will be referred to as the *autocorrelation* of the random process $S(t)$, while $r(\tau)$ is the *matched filter response* (in the context of radar processing) of a sample function of $S(t)$.

in the previous section (assuming the process is sampled as well).

Still, (3.43) represents the ensemble average of $S(t)$ under the matched filtering operation. That is, if every member function of $S(t)$ were matched filtered and averaged, the result would approach (3.43). Practically speaking, $S(t)$ will likely have an infinite number of sample functions while a radar CPI will likely, at most, consist of hundreds of waveforms. The question is then how does $r_{SS}(\tau)$ behave when only a few to a few hundred autocorrelation sample function are coherently integrated?

To determine this, consider the fact that $r(t)$ is itself a random variable whose mean is the right hand side of (3.43). It then also has a variance that is decreased with increasing coherent integration so the expected sidelobe level of a single sample function should be proportional to the second moment of $r(t)$ which is

$$E[|r(\tau)|^2]. \quad (3.44)$$

If N matched filter responses from N unique sample function of $S(t)$ are coherently integrated and normalized then the variance of this sum becomes

$$E \left[\left| \frac{1}{N} \sum_{n=1}^N r_n(\tau) \right|^2 \right] \quad (3.45)$$

which can be expanded to

$$\frac{1}{N^2} \sum_{n_1=1}^N \sum_{n_2=1}^N E [(r_{n_1 n_2}(\tau)) (r_{n_1 n_2}^*(\tau))] \quad (3.46)$$

The terms defined by $n_1 = n_2$ are the second moment of the matched filter response of $S(t)$ while the other terms are correlations between the matched filter responses of different sample functions. Since the different sample function are ideally independent of each other, their correlations are the product of their expectations. Since they were generated using the same waveform generating

function they have the same expected autocorrelation functions such that.

$$E \left[\left| \frac{1}{N} \sum_{n=1}^N r_n(\tau) \right|^2 \right] = \frac{1}{N} E[|r(\tau)|^2] + \frac{N-1}{N} |E[r(\tau)]|^2 \quad (3.47)$$

From (3.47) the contribution of second moment decreases at a rate of $1/N$ or $-10\log_{10}(N)$. Thus, the sidelobe levels will decrease from their nominal single waveform level of $E[|r(\tau)|^2]$ to $E[r(\tau)]^2$ at a rate of $-10\log_{10}(N)$. Assuming $E[r(\tau)]$ is ideal in that it has zero sidelobes, the sidelobes will likewise approach zero.

By the Weiner-Kinchine theorem, the power spectral density, in the context of radar processing, can be written as a function of the expected matched filter response such that

$$S(f) = \int_{-\infty}^{\infty} E[r(\tau)] \exp(-j2\pi f\tau) d\tau. \quad (3.48)$$

Here, it is important to note that $S(f)$, the power spectral density as measured by the matched filter response in the context of radar processing, is not equivalent to $S_{SS}(f)$ since the pulsed waveform generating process, $S(t)$, is not stationary. The PSD of $S(t)$ changes as a function of time. Just like $E[r(\tau)]$ is the average correlation between samples, $S(f)$ is the average PSD of $S(t)$.

3.3.2 Stochastic Processes as CW Radar Signals

For CW signals, $S(t)$ is now considered to be infinite in time. However, since CPI's are time limited only a portion of a sample function of $S(t)$ can ever be generated, transmitted, and recorded. Consequently, the matched filter will only ever be as large as some window of a sample function of $S(t)$. Assuming this time window is the same size as the CPI, the matched filter response becomes

$$r(\tau) = \frac{1}{T_{win}} \begin{cases} \int_0^{T_{win}+\tau} s(t)s^*(t-\tau) dt & -T_{win} < \tau \leq 0 \\ \int_{\tau}^{T_{win}} s(t)s^*(t-\tau) dt & 0 < \tau < T_{win} \\ 0 & \text{otherwise} \end{cases} \quad (3.49)$$

where T_{win} is the size of the matched filter window. In this case, since $S(t)$ is infinitely long and it needs to "look" the same at all times for the autocorrelation to be consistent at any reference time, in the CW case, $S(t)$ is assumed to be WSS. Under these assumptions, the expectation of the matched filter response of the windowed $S(t)$ is

$$E[r(\tau)] = \frac{1}{T_{win} - |\tau|} R_{SS}(\tau) \quad -T_{win} < \tau < T_{win}. \quad (3.50)$$

Since $S(t)$ is WSS, the autocorrelation is a function of only τ and (3.50) is exactly the form of the time averaging autocorrelation estimate. The final wrinkle in the CW case is that for a given CPI there is only one sample function to work with and thus no pulse integration. This means that, while the results of (3.44) - (3.47) apply here, they are not relevant without pulse integration. Alternatively, the sidelobe levels of a CW noise like waveform should be proportional to BT . For a given bandwidth, this means increasing T_{win} . For this decrease in sidelobe level to be guaranteed, $S(t)$ needs to also be ergodic such that the limit of time averages of a single sample function are equivalent to the ensemble averages. This realizes

$$R_{SS}(\tau) = \lim_{T_{win} \rightarrow \infty} \frac{1}{T_{win}} \begin{cases} \int_0^{T_{win}+\tau} s(t)s^*(t-\tau)dt & -T_{win} < \tau \leq 0 \\ \int_\tau^{T_{win}} s(t)s^*(t-\tau)dt & 0 < \tau < T_{win} \\ 0 & \text{otherwise} \end{cases}. \quad (3.51)$$

If $S(t)$ is ergodic such that (3.51) applies, then $r(\tau)$ will, as a function increasing BT , approach $R_{SS}(\tau)$ which ideally has zero autocorrelation sidelobes.

As in the pulsed case, the expected PSD can be calculated via the Weiner-Kinchine theorem such that

$$S(f) = \int_{-T_{win}}^{T_{win}} \frac{1}{T_{win} - |\tau|} R_{SS}(\tau) d\tau = \left(\frac{\sin(\pi f T_{win})}{\pi f} \right)^2 * S_{SS}(f). \quad (3.52)$$

where $*$ denotes convolution. In the CW case, the result in the context of radar processing, is exactly the windowed (and biased) PSD estimate of $S_{SS}(f)$ as defined in section 3.2.3 (assuming a sampled form of $s(t)$).

3.4 Evaluating Stochastic Radar Waveforms

Section 3.3 provided an overview of basic radar processing (i.e. matched filtering and Fourier transforms) looks like when using the sample functions of stochastic processes as radar signals. This was discussed for both the pulsed and CW cases. Throughout that analysis, the waveform generating function $S(t)$ and its member functions were continuous. However, in modern systems the matched filtering and spectral estimation steps are performed on computers and is necessarily discrete. Additionally, the design and optimization in chapters 4 - 6 are likewise performed on a computer so it is necessary to consider the content of section 3.3 in discrete terms.

3.4.1 The Discrete Waveform Generating Process

The discrete analogue to the pulsed waveform generating process defined in (3.40) is

$$S[m] = \begin{cases} S[m] & m = 1, 2, \dots, M \\ 0 & \text{otherwise} \end{cases}. \quad (3.53)$$

where the pulsed stochastic process $S[m]$ is stochastic for exactly M samples from $m = 1, 2, \dots, M$ and deterministically zero otherwise. In CW case, $S[m]$ is stochastic for all m and is denoted $S_{\text{cw}}[m]$. The discrete matched filter responses are

$$r[\ell] = \begin{cases} \sum_{m=1}^{M+\ell} s[m]s^*[m-\ell]dt & \ell = 1-M, 2-M, \dots, 0 \\ \sum_{\ell}^M s[m]s^*[m-\ell]dt & \ell = 1, 2, \dots, M-1 \\ 0 & \text{otherwise} \end{cases} \quad (3.54)$$

for the pulsed case, and

$$r[\ell] = \begin{cases} \sum_{m=1}^{M_{\text{win}}+\ell} s[m]s^*[m-\ell]dt & \ell = 1-M_{\text{win}}, 2-M_{\text{win}}, \dots, 0 \\ \sum_{\ell}^{M_{\text{win}}} s[m]s^*[m-\ell]dt & \ell = 1, 2, \dots, M_{\text{win}}-1 \\ 0 & \text{otherwise} \end{cases} \quad (3.55)$$

for the CW case, where M_{win} is some window length since only a portion of the infinitely long CW stochastic process can be processed.

Before discussing the discrete PSDs, it is convenient to write (3.53) as the M -length vector, \mathbf{s} , where the zero-valued portions of (3.53) are omitted². A particular sample function of the stochastic process, \mathbf{s} , is written as $\hat{\mathbf{s}}$.

Section 3.2.3 defined the periodogram estimate of a signal's power spectrum. For the pulsed stochastic signal, \mathbf{s} , the periodogram estimate can be written as a matrix multiply of a zero-padded version of a sample function as

$$\hat{\mathbf{s}}_f = \mathbf{A} \hat{\mathbf{s}} \quad (3.56)$$

where $\bar{\mathbf{s}}$ is a zero-padded version of \mathbf{s} such that

$$\bar{\mathbf{s}} = [\mathbf{s}^T \ \mathbf{0}_{W-M}^T]^T, \quad (3.57)$$

$\hat{\mathbf{s}}$ is a sample function of the zero-padded random process to account for the pulsed nature of \mathbf{s} , \mathbf{A} is the $W \times W$ DFT matrix, and W is $W \geq 2M - 1$ (such that (3.57) is at least the length of the sample matched filter response, (3.54)). (3.56) is defined in terms of a sample function of $\bar{\mathbf{s}}$. For future convenience, the transformation of the random process $\bar{\mathbf{s}}$, is defined as

$$\mathbf{s}_f = \mathbf{A} \bar{\mathbf{s}}. \quad (3.58)$$

For the CW case, $S_{cw}[m]$ can also be written as \mathbf{s} , but in this case it is an M_{win} -length vector indicating that it is a truncated version of an infinitely long process. Consequently, when evaluating the periodogram no zero-padding is necessary and the estimate becomes

$$\hat{\mathbf{s}}_f = \mathbf{A} \hat{\mathbf{s}} \quad (3.59)$$

²Previously, the stochastic processes and random variables themselves have been written as capital letters. This implies the vectorized form of the random process should be 'S' however this creates a conflict with typical vector notation where bold, capital letters denote matrices and lower case bold letters are vectors. To avoid confusion, typical vector notation is applied here and the distinction will be made in context

and biases apply as discussed in 3.2.3. The transformation of the CW process \mathbf{s} is similar to (3.58), but without the zero-padding. With these definitions, it is possible to start defining metrics and the tools needed to evaluate the usefulness of stochastic processes for analysis on a digital system.

3.4.2 Metrics for Stochastic Radar Waveforms

Thus far, an effort has been made to distinguish between terms which specifically refer to the fundamental characteristics of a random process such as the autocorrelation function, $R_{SS}(t_1, t_2)$, and terms which refer to these quantities in the context of radar processing. For example, in radar processing the matched filter response is interchangeably called the autocorrelation, but to avoid confusing this with $R_{SS}(t_1, t_2)$, it has been referred to exclusively as the matched filter response. At this point, this chapter pivots from a random signals focus, to a radar analysis and processing focus. Consequently, terms such as the autocorrelation and the PSD can be assumed to be in the context of radar signal processing rather than random signals analysis unless explicitly stated as such.

The following subsections consider two classes of stochastic radar waveform metrics which apply to both pulsed and CW radar waveforms. Section 3.4.2.1 defines *aggregate* metrics. These measure the average performance of the entire ensemble of functions defined by \mathbf{s} (i.e. the resulting autocorrelation after coherently integrating all member functions). Section 3.4.2.2 discusses the metrics used to evaluate the behavior of the stochastic waveforms defined by \mathbf{s} on an individual basis (i.e. what autocorrelation performance can be expected if only a single sample function is used). Both classes of metrics are vitally important for evaluating the usefulness of a particular stochastic process for the generation of radar waveforms.

3.4.2.1 Aggregate Metrics

Expected Power Spectral Density

$$E \left[|\mathbf{s}_f|^2 \right] \quad (3.60)$$

In (3.60), $E[\cdot]$ evaluates the element wise expectation and $|\cdot|^2$ is the element wise magnitude squared. It measures the average PSD over all the sample waveforms defined by the process $s[m]$. If the PSD of all the sample function in $s[m]$ were evaluated and averaged, the result would be (3.60).

Expected Autocorrelation

$$\mathbf{A}^H E \left[|\mathbf{s}_f|^2 \right] \quad (3.61)$$

Since the Fourier transform, and here the inverse discrete Fourier transform in the form of matrix multiplication, is a linear process. The expected autocorrelation can be calculated from an IFFT of the expected PSD by the Weiner-Khinchine theorem. If the autocorrelation of all the sample functions of $s[m]$ where evaluated and averaged, the result would be (3.61).

Expected Cross-Power Spectral Density

$$E [\mathbf{s}_{af} \odot \mathbf{s}_{bf}^*] \quad (3.62)$$

In (3.62), the subscripts 'a' and 'b' denote that \mathbf{s}_{af} and \mathbf{s}_{bf} are the spectra of unique sample functions of the waveform generating process $s[m]$ as opposed to the magnitude squared of a single sample function of $s[m]$ as in (3.60) and (3.62). In this way, (3.62) measures the mean pair-wise cross-power spectral density of $s[m]$. If the sample average were taken over all of the pair-wise cross-power spectral densities of the sample functions defined by $s[m]$, the result would approach (3.62).

Expected Cross-Correlation

$$\mathbf{A}^H E [\mathbf{s}_{af} \odot \mathbf{s}_{bf}^*] \quad (3.63)$$

Similarly to (3.62), (3.63) measures a pairwise relationship, but here it is the time-domain cross correlation. If the sample average were taken over all of the pair-wise cross-correlations of the sample functions defined by $S[m]$, the result would approach (3.63).

3.4.2.2 Individual Metrics

Power Spectral Deviation

$$\left(E \left[|s_f|^4 \right] - E \left[|s_f|^2 \right]^2 \right)^{1/2} \quad (3.64)$$

Whereas (3.60) measures the average PSD of the sample functions of $S[m]$, (3.64) measures the magnitude of the expected error between a single sample function's spectrum and of the expected PSD in (3.60).

Expected RMS Autocorrelation

$$\left(E \left[\left| A^H s_f \right|^2 \right] \right)^{1/2} \quad (3.65)$$

(3.65) is the square root of the second moment of the autocorrelation of the sample functions in $S[m]$. If the square root of the sample average of the magnitude square of the autocorrelations of all the sample functions in $S[m]$ were calculated, the result would approach (3.65). In other words, (3.65) represents the average magnitude of the autocorrelation of a single sample function of $S[m]$ as opposed to the autocorrelation after coherently integrating the autocorrelation response of every sample function of $S[m]$ as in (3.61).

Expected RMS Cross-Power Spectral Density

$$\left(E \left[|s_{af} \odot s_{bf}^*|^2 \right] \right)^{1/2} \quad (3.66)$$

(3.66) measures the average magnitude of the cross-power spectral density of two waveforms. In other words, if only one pair of waveforms is considered, on average, how much common spectral content do they have.

Expected RMS Cross-Correlation

$$\left(E \left[\left| A^H E [s_{af} \odot s_{bf}^*] \right|^2 \right] \right)^{1/2} \quad (3.67)$$

If only one pair of waveforms generated by $S[m]$ are correlated, (3.67) represents the average magnitude of this cross-correlation.

3.4.2.3 Summary of Stochastic Waveform Measures

Table 3.3: Summary of Stochastic Waveform Measures

Aggregate Measures	Power Spectrum	$E[s_f ^2]$
	Autocorrelation	$\mathbf{A}^H E[s_f ^2]$
	Cross-Power Spectral Density	$E[s_{af} \odot s_{bf}^*]$
	Cross-Correlation	$\mathbf{A}^H E[s_{af} \odot s_{bf}^*]$
Individual Measures	Power Spectral Deviation	$\left(E[s_f ^4] - E[s_f ^2]^2 \right)^{1/2}$
	RMS Autocorrelation	$\left(E[\mathbf{A}^H s_f ^2 ^2] \right)^{1/2}$
	RMS Cross-Power Spectral Density	$\left(E[s_{af} \odot s_{bf}^* ^2] \right)^{1/2}$
	RMS Cross-Correlation	$\left(E[\mathbf{A}^H [s_{af} \odot s_{bf}^*] ^2] \right)^{1/2}$

3.4.3 Stochastic Waveform Moments

The aforementioned equations, 3.60 through 3.67, are functions of the moments of $S[m]$. To see this, begin by expanding the expected PSD in 3.60 or 3.61 to

$$E[|s_f|^2] = E[\mathbf{A}\bar{s} \odot (\mathbf{A}\bar{s})^*] \quad (3.68)$$

where \odot is the Hadamard (element-wise) product and $(\cdot)^*$ is complex conjugation. The w^{th} sample of (3.68) is

$$E \left[|s_{f,w}|^2 \right] = E \left[\left(\sum_{m_1=1}^W a_{w,m_1} \bar{s}_{m_1} \right) \left(\sum_{m_2=1}^W a_{w,m_2} \bar{s}_{m_2} \right) \right] \quad (3.69)$$

This product of sums can be written as a sum of products such that

$$E \left[|s_{f,w}|^2 \right] = \sum_{m_1=1}^W \sum_{m_2=1}^W a_{w,m_1} a_{w,m_2}^* E \left[\bar{s}_{m_1} \bar{s}_{m_2}^* \right] \quad (3.70)$$

where the expectation has been moved inside the summation since it is a linear operator. In (3.64), and less obviously in (3.65), a similar sum of expectations appears as

$$E \left[|s_{f,w}|^4 \right] = \sum_{m_1=1}^W \sum_{m_2=1}^W \sum_{m_3=1}^W \sum_{m_4=1}^W a_{w,m_1} a_{w,m_2}^* a_{w,m_3}^* a_{w,m_4} E \left[\bar{s}_{m_1} \bar{s}_{m_2}^* \bar{s}_{m_3}^* \bar{s}_{m_4} \right]. \quad (3.71)$$

In (3.70) and (3.71), the terms

$$E \left[\bar{s}_{m_1} \bar{s}_{m_2}^* \right] \quad (3.72)$$

and

$$E \left[\bar{s}_{m_1} \bar{s}_{m_2}^* \bar{s}_{m_3}^* \bar{s}_{m_4} \right] \quad (3.73)$$

represent the second and fourth order moments of the samples in \bar{s} respectively. For convenience these terms can be collected into $\bar{\mathbf{C}}_{W \times W}$ and $\bar{\mathbf{K}}_{W \times W \times W \times W}$ where $\bar{\mathbf{C}}$ is the correlation matrix and $\bar{\mathbf{K}}$ is a fourth-order moment hyper-cube. Additionally, $\bar{\mathbf{C}}$ and $\bar{\mathbf{K}}$ include the zero-valued portions of \bar{s} meaning the majority of $\bar{\mathbf{C}}$ and $\bar{\mathbf{K}}$ are actually zero and the non-zero portions are the first M samples in each dimension is denoted as $\mathbf{C}_{M \times M}$ and $\mathbf{K}_{M \times M \times M \times M}$ respectively. In the CW case, there are no zero valued portions and the moment structures become $\mathbf{C}_{M_{win} \times M_{win}}$ and $\mathbf{K}_{M_{win} \times M_{win} \times M_{win} \times M_{win}}$.

For the expected cross terms, (3.62), (3.63), (3.66), and (3.67), two different samples functions "a" and "b" are considered such that their samples are completely independent of each other. Consider the expected cross-PSD (3.62) which can be expanded in a similar fashion to the expected

PSD (3.60) such that for a single sample, the equation becomes

$$E[s_{af,w} s_{bf,w}^*] = \sum_{m_1=1}^W \sum_{m_2=1}^W a_{w,m_1} a_{w,m_2}^* E[\bar{s}_{a,m_1}] E[\bar{s}_{b,m_2}^*] \quad (3.74)$$

In (3.74), the expectation of the product can be written as a product of the expectations. Since \bar{s}_a and \bar{s}_b were generated from the same random process the expectations of their individual samples are equivalent and (3.74) becomes

$$E[s_{af,w} s_{bf,w}^*] = \sum_{m_1=1}^W \sum_{m_2=1}^W a_{w,m_1} a_{w,m_2}^* E[\bar{s}_{m_1}] E[\bar{s}_{m_2}^*] \quad (3.75)$$

where $E[\bar{s}_{m_2}^*]$ is the first moment of the m_1^{th} sample of $s[m]$. Likewise, using the same reasoning, similar moments appear in (3.66) and (3.67) as

$$E[|\mathbf{s}_{af} \odot \mathbf{s}_{bf}^*|^2] = \sum_{m_1=1}^W \sum_{m_2=1}^W \sum_{m_3=1}^W \sum_{m_4=1}^W a_{w,m_1} a_{w,m_2}^* a_{w,m_3}^* a_{w,m_4} E[\bar{s}_{m_1} \bar{s}_{m_2}^*] E[\bar{s}_{m_3}^* \bar{s}_{m_4}] \quad (3.76)$$

such that the RMS cross-correlation and the RMS cross-PSD are functions of the second order moments of $S[m]$.

Chapter 4

Pulsed Stochastic Waveform Generation (StoWGe)

Chapter 3 introduced stochastic processes as radar waveform generating functions (WGF) where the sample functions of a given process are utilized as radar waveforms. In the same vein, various metrics were introduced as a means to assess the suitability of a given random process as a radar WGF. However, nothing was said about how to find or design a random process that has desirable properties. In [30], the stochastic waveform generation (StoWGe) signal model was introduced as a means to design and optimize stochastic processes which produce FM noise waveforms with desirable power spectra. This chapter reintroduces this topic and expands upon it greatly by formulating the model in more general terms, implementing it in a more flexible manner, and considering a much more comprehensive set of test cases.

This chapter begins with an overview of the StoWGe model in Section 4.1. Section 4.2 introduces the expected frequency template error (EFTE) for pulsed StoWGe, while Section 4.3 introduces the means to optimize the StoWGe expected power spectrum by minimization of the EFTE cost function along with a regimen of test cases. Section 4.4 discusses the results of the optimization and Section 4.5 discusses both loop-back and open-air experimental results.

4.1 The Pulsed StoWGe Signal Model

The StoWGe signal model produces waveforms of the form in 3.40. However, to facilitate optimization on a computer and since the waveforms produced by StoWGe are constant amplitude, the

most appropriate baseband signal mode is

$$S[m] = \begin{cases} \exp(j\Phi[m]) & m = 1, 2, \dots, M \\ 0 & \text{otherwise} \end{cases} \quad (4.1)$$

where the capital S and the capital Φ indicate that these are stochastic processes themselves, while individual sample functions would be lower case. In (4.1), $S[m]$ is a non-linear transformation of the stochastic process $\Phi[m]$. The goal then is to design $\Phi[m]$ such that the sample functions of $S[m]$ are useful pulsed radar waveforms.

For convenience, (4.1) can be rewritten using the same vector notation as in section 3.4.1 such that the non-zero portions of (4.1) are

$$\mathbf{s} = \exp(j\boldsymbol{\Phi}). \quad (4.2)$$

To facilitate the optimization of \mathbf{s} as a WGF, according to the pulsed StoWGe model $\boldsymbol{\Phi}$ is parameterized as

$$\boldsymbol{\Phi} = \mathbf{B}\mathbf{x} + \boldsymbol{\mu} \quad (4.3)$$

where \mathbf{B} is an $M \times N$, real-valued matrix of basis functions, \mathbf{x} is an $N \times 1$ vector of independent, identically distributed, zero-mean, random variables, and $\boldsymbol{\phi}$ is an $M \times 1$ vector of constant real values. In this form, \mathbf{x} provides $\boldsymbol{\Phi}$ with its stochastic character and by designing \mathbf{B} , the distributions of $\boldsymbol{\Phi}$ can be tailored such that the WGF, \mathbf{s} , has desirable characteristics.

$\boldsymbol{\mu}$ provides $\boldsymbol{\Phi}$ with a mean value such that

$$\mathbb{E}[\boldsymbol{\Phi}] = \boldsymbol{\mu}. \quad (4.4)$$

where the expectation operator is applied on an element wise basis. Alternatively, $\boldsymbol{\mu}$ could be omitted and the members of \mathbf{x} could be allowed to have non-zero means to achieve the same result

in which case (4.4) would become

$$E[\boldsymbol{\phi}] = \mathbf{B}\boldsymbol{\mu}_X \quad (4.5)$$

where $\boldsymbol{\mu}_x$ is a vector of the means of the random variables in \mathbf{x} . While this would be just as effective at providing any desired mean to the elements of $\boldsymbol{\phi}$ it unnecessarily obfuscates them by tying them into the elements of \mathbf{B} and the mean values of \mathbf{x} .

To gain a better idea of how the pulsed StoWGe structure functions, (4.3) can be rewritten such that

$$\boldsymbol{\phi} = \left(\sum_{n=1}^N X_n \mathbf{b}_n \right) + \boldsymbol{\mu} \quad (4.6)$$

where X_n is the n th random variable in the vector \mathbf{x} and \mathbf{b}_n is the n th column vector of \mathbf{B} . In this form, $\boldsymbol{\phi}$ is a linear combination of basis functions which are the columns of \mathbf{B} , and the weight of each basis function is randomly chosen as a sample value of each member of \mathbf{x} . Since the elements of \mathbf{x} are by definition zero-mean, $\boldsymbol{\mu}$ can provide any mean value to the elements of $\boldsymbol{\phi}$.

Section 3.4.3 established that the metrics defined in Section 3.4.2 are functions of the moments of the WGF. For the pulsed StoWGe model, these moments are likewise functions of \mathbf{B} and $\boldsymbol{\mu}$. The correlation between the m_1 th sample and the m_2 th is defined

$$E[s_{m_1} s_{m_2}^*] = E[\exp(j(\phi_{m_1} - \phi_{m_2}))] \quad (4.7)$$

where the phase values can be expanded based on the structure of $\boldsymbol{\mu}$ such that

$$E[s_{m_1} s_{m_2}^*] = E[\exp(j((\mathbf{b}_{m_1} - \mathbf{b}_{m_2})\mathbf{x} + \boldsymbol{\mu}_{m_1} - \boldsymbol{\mu}_{m_2}))] \quad (4.8)$$

where \mathbf{b}_{m_1} and \mathbf{b}_{m_2} are row vectors of \mathbf{B} . Since the values of $\boldsymbol{\mu}$ are constant, they can be factored out of the expectation and the equation can be rewritten as

$$E[s_{m_1} s_{m_2}^*] = \exp(j(\boldsymbol{\mu}_{m_1} - \boldsymbol{\mu}_{m_2})) E \left[\prod_{n=1}^N \exp(j(b_{m_1,n} - b_{m_2,n})X_n) \right]. \quad (4.9)$$

Since the elements of \mathbf{x} are assumed to be statistically independent, the expectation can be moved inside the product operator yielding

$$\mathbb{E}[s_{m_1}s_{m_2}^*] = \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{n=1}^N \mathbb{E}[\exp(j(b_{m_1,n} - b_{m_2,n})X_n)]. \quad (4.10)$$

Such that (4.9) becomes a product of expectations rather than the expectation of a product. Similar to the analysis in [27, 28, 76], note that the expectation in (4.10) is the form of the characteristic function of a random variable which itself is defined as

$$\psi_X(\omega) = \mathbb{E}[\exp(j\omega X)] \quad (4.11)$$

where $\psi_X(\omega)$ is the characteristic function of the random variable X as defined in Section 3.1.2.

Putting (4.11) into (4.10), the correlation of the signal samples become

$$\mathbb{E}[s_{m_1}s_{m_2}^*] = \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{n=1}^N \psi_{X_n}(b_{m_1,n} - b_{m_2,n}). \quad (4.12)$$

Likewise, the fourth order moment is

$$\mathbb{E}[s_{m_1}s_{m_2}s_{m_3}^*s_{m_4}^*] = \exp(j(\mu_{m_1} + \mu_{m_2} - \mu_{m_3} - \mu_{m_4})) \prod_{n=1}^N \psi_{X_n}(b_{m_1,n} + b_{m_2,n} - b_{m_3,n} - b_{m_4,n}). \quad (4.13)$$

In (4.12) and (4.13), the moments of \mathbf{s} are functions of the elements of \mathbf{B} and $\boldsymbol{\mu}$. Therefore, (4.3) can be optimized according to the metrics in Section 3.4 by adjusting the values of the elements of \mathbf{B} and $\boldsymbol{\mu}$.

4.2 The Expected Frequency Template Error (EFTE) for Pulsed StoWGe

In [30], the expected frequency template error (EFTE) cost function was introduced as a means to measure the squared error between the expected power spectrum of some waveform generating

process and some desired spectrum. This can be stated mathematically as

$$J_P = \left\| E \left[|\mathbf{s}_f|^2 \right] - \mathbf{u} \right\|_2^2 \quad (4.14)$$

where

$$E \left[|\mathbf{s}_f|^2 \right] \quad (4.15)$$

is the expected power spectrum of the WGF \mathbf{s} as defined in Section 3.4, $\|\cdot\|_2^2$ is the squared Euclidean norm, and \mathbf{u} is some desired power spectrum. Then, as (4.14) is minimized, the WGF will on average produce waveforms with a power spectrum that is more and more similar to \mathbf{u} . Overall, the EFTE cost function has important advantages relative to other stochastic waveform metrics.

For one, the expected power spectrum shares a Fourier transform relationship with the expected autocorrelation. Accordingly, by selecting a desired spectrum which results in an excellent expected autocorrelation, the expected autocorrelation can be optimized simultaneously. This concept has been used extensively in the design of FM noise waveforms [23, 24, 26] and in the design of single, low auto-correlation sidelobe waveforms [38, 39] as was discussed in Section 2.2. However, these waveform design implementations operated on a per waveform basis whereas here the notion is applied to an entire family of waveforms via the WGF.

Additionally, the expected power spectrum is only a second order function of the waveform generating process as apposed to a fourth order function in the individual metrics of Section 3.4. This represents a clear advantage in terms of the processing and complexity requirements of any optimization using this metric. Alternatively, one may consider optimizing the expected autocorrelation directly since it is also a second order function of the waveform generating process. The most straightforward choices are to optimize via minimizing either ISL, PSL, or the more flexible generalized integrated sidelobe level (GISL) [64]. However, these metrics largely disregard spectral containment making the EFTE metric more desirable at least for noise like waveforms [86].

One major drawback of the EFTE cost function is that it only considers the average behavior

of waveforms. Individual waveforms could have very poor properties such as wildly spread out spectra or high autocorrelation sidelobes, but as long as they average out to the desirable outcome the EFTE metric will not capture this behavior. Practically, speaking we will only ever be able to coherently integrate a finite number of waveforms so it is important to consider their individual behavior in addition to their aggregate characteristics. With this mind, it is helpful to at least examine the individual metrics of Section 3.4 to ensure the minimization of the EFTE cost function is producing useful WGFs.

4.3 Optimization of the Pulsed StoWGe Expected Power Spectrum

As alluded to in section 4.2, the pulsed StoWGe expected power spectrum can be optimized by minimizing the EFTE cost function as was performed in [30]. In order to determine the best optimization method, the 2-norm in (4.14) can be expanded such that it becomes

$$J = \left(E \left[|s_f|^2 \right] \right)^T \left(E \left[|s_f|^2 \right] \right) - 2\mathbf{u}^T \left(E \left[|s_f|^2 \right] \right) + \mathbf{u}^T \mathbf{u} \quad (4.16)$$

which is a fourth order function of s . Additionally, as shown in (4.12), the second order moments of the samples in s are already products of the characteristic functions of the random variables of \mathbf{x} . The characteristic functions themselves are dependent on \mathbf{B} and $\boldsymbol{\mu}$. Consequently, the EFTE cost function (4.14) is highly non-linear. This, along with it being unconstrained makes it unlikely that (4.14) is convex, in that it is unlikely to have a global minimum.

In addition to being non-convex and non-linear, (4.14) is also unconstrained. That is the elements of \mathbf{B} and $\boldsymbol{\mu}$ are allowed to take on any real value. As discussed in Section 2.5, non-linear cost functions can be difficult to optimize, but at the very least it is straightforward to apply gradient descent methods to find at least a local minimum so long as the function is differentiable as is the case of (4.14). With this in mind the most important part of gradient descent optimization is to calculate the gradient itself.

4.3.1 The Pulsed StoWGe EFTE Gradient

As the name implies gradient descent techniques require the evaluation of the cost function gradient. For the EFTE cost function this can be calculated for both the basis function matrix, \mathbf{B} , and the mean value vector, $\boldsymbol{\mu}$. In terms of a single element of \mathbf{B} , the derivative of (4.14) is

$$\frac{\partial J}{\partial b_{k,n}} = 2 \left(\frac{\partial E[|s_f|^2]}{\partial b_{k,n}} \right)^T (E[|s_f|^2] - \mathbf{u}) \quad (4.17)$$

where $b_{k,n}$ is the element of \mathbf{B} in the k th row and n th column, and the derivative of the expected power spectrum is evaluated on an element wise basis. The derivative of the w th element of $E[|s_{f,w}|^2]$ is

$$\frac{\partial E[|s_{f,w}|^2]}{\partial b_{k,n}} = 2\Re \left\{ \sum_{m=1}^M a_{w,m} a_{w,k}^* \exp(j(\mu_m - \mu_k)) \frac{\partial \Psi_{X_p}(b_{m,n} - b_{k,n})}{\partial b_{k,n}} \prod_{\substack{p=1 \\ p \neq k}}^N \Psi_{X_p}(b_{m,n} - b_{k,n}) \right\} \quad (4.18)$$

where $\Re\{\cdot\}$ extracts the real part. A full derivation of (4.18) can be found in appendix A.1.2.1. (4.17) can be evaluated for each element of \mathbf{B} and collected into the structure

$$\nabla_{\mathbf{B}} J = \begin{bmatrix} \frac{\partial J}{\partial b_{1,1}} & \frac{\partial J}{\partial b_{1,2}} & \cdots & \frac{\partial J}{\partial b_{1,N}} \\ \frac{\partial J}{\partial b_{2,1}} & \frac{\partial J}{\partial b_{2,2}} & \cdots & \frac{\partial J}{\partial b_{2,N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial b_{M,1}} & \frac{\partial J}{\partial b_{M,2}} & \cdots & \frac{\partial J}{\partial b_{M,N}} \end{bmatrix} \quad (4.19)$$

to realize the gradient with respect to \mathbf{B} .

Likewise, the derivative with respect to an single value of $\boldsymbol{\mu}$ is

$$\frac{\partial J}{\partial \mu_k} = 2 \left(\frac{\partial E[|s_f|^2]}{\partial \mu_k} \right)^T (E[|s_f|^2] - \mathbf{u}) \quad (4.20)$$

where μ_k is a single element of $\boldsymbol{\mu}$. Taking the derivative on an element wise basis, the derivative of

the w th element of $E[|s_{f,w}|^2]$ is

$$\frac{\partial E[|s_{f,w}|^2]}{\partial \mu_k} = 2\Im \left\{ a_{w,k}^* \exp(-j\mu_k) \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} \exp(j\mu_m) \prod_{p=1}^P \Psi_{X_p}(b_{k,p} - b_{m,p}) \right\} \quad (4.21)$$

where $\Im\{\cdot\}$ extracts the imaginary part. A complete derivation of (4.21) can be found in Section A.1.2.2. The results of (4.20) and (4.21) can be collected into the structure

$$\nabla_{\boldsymbol{\mu}} J = \begin{bmatrix} \frac{\partial J}{\partial \mu_1} \\ \frac{\partial J}{\partial \mu_2} \\ \vdots \\ \frac{\partial J}{\partial \mu_M} \end{bmatrix} \quad (4.22)$$

to realize the gradient of the EFTE cost function with respect to the mean value phase vector $\boldsymbol{\mu}$.

4.3.2 Gradient Descent Implementation

As discussed in Section 2.5, the local optimization of a non-linear problem can be more of an art than a science. The quality of the final result is often strongly dependent on the choice of initialization, as will be seen in the results of work, and the type of optimization and its parameters (e.g. β for gradient descent methods). Consequently, the choice of parameters is often a heuristic process of trial and error and educated guesses as was the case here.

In the course of evaluating the efficiency of various gradient descent techniques such as the ones in [79], it was found that one of the simplest varieties, heavy ball gradient descent, was uniquely robust and efficient at minimizing the pulsed StoWGe EFTE cost function. As was discussed in more detail in Section 2.5, for gradient descent methods a search direction is chosen such that

$$\mathbf{q}_i = \begin{cases} -\mathbf{g}_i & \text{when } i = 0 \\ -\mathbf{g}_i + \beta \mathbf{q}_{i-1} & \text{otherwise} \end{cases} \quad (4.23)$$

where for heavy ball gradient descent β takes on a value such that $0 < \beta < 1$ which makes the search direction at the i th iteration a linear combination of the current gradient and weighted versions of previous search directions [80]. The fact that heavy ball gradient descent would be more efficient (converging more quickly than other implementations) than other more sophisticated approaches such as non-linear conjugate gradient methods, is itself a surprising result considering they generally work very well.

Nevertheless, it does not really matter what type of gradient descent method is used so long as it is used properly. Given the same initialization and line search method, they should come to the same solutions. It is just matter of how quickly they do so. More impactful, to the final results are the stopping conditions. Stopping conditions which are too strict will stop the optimization while significant improvement could still be attained. Stopping conditions what are too lax will achieve a much better approximation of the local minimum than the former case, but a significant number of iterations may result in little improvement and wasted time. The stopping conditions chosen here were pragmatically chosen based on a process of trial and error. With this in mind, the goal was to provide the optimization with as many iterations as it needed to find a good solution but no more.

- No matter what, in each case the optimization was performed for at least 5000 iterations.

This supersedes the following stopping conditions

- The optimization was ended after 1000000 iterations.
- If the aggregate decrease in the cost function over the course of any consecutive 1000 iterations was less than .01 dB, the optimization was ended. Such small decreases in the cost function value indicate further meaningful improvement is unlikely even with an extreme number of iterations.
- If the cost function value fell below -200 dB, the optimization was ended. A cost function value below -200 dB indicates the total squared error between the expected spectrum and desired spectrum is $< 10^{-20}$. Practically speaking, the error is 0 at this point making further optimization unnecessary.

In a vacuum these parameters and stopping conditions seem completely arbitrary, but in practice they were found to provide a good compromise between optimization time and finding an excellent approximation of a local minimum.

4.3.3 Initializations

One difficulty in evaluating the optimization of StoWGe via the EFTE cost function is the number of parameters to consider. Table 4.1 highlights this issue by listing all the independent parameters. The parameters M , N , $p_X(x)$, and \mathbf{u} are fixed throughout the optimization, while the basis function

Fixed Parameters	M	Number of samples per waveform
	N	Number of random variables per waveform
	$p_X(x)$	Random Variable Distribution Function
	\mathbf{u}	The Desired expected spectrum
Optimizable Parameters	\mathbf{B}	Basis Function matrix
	$\boldsymbol{\mu}$	Waveform Sample mean value vector

Table 4.1: Waveform generating process parameters which must be selected or initialized prior to optimization

matrix \mathbf{B} and the mean phase value vector $\boldsymbol{\mu}$ are optimized. The fixed parameters have to be set before beginning the optimization while the optimizable parameters have to be initialized. With a total of six parameters to consider the number of optimizations to run can quickly become impractical. With this in mind, it is necessary to select a meaningful subset of test cases. The next several paragraphs discuss the selected initializations for the parameters in Table 4.1 along with the rationale for those selections.

Number of Waveform Samples: M – To understand how the EFTE minimization performs on StoWGe, the only requirement on M is that it is sufficiently large such that a meaningfully large BT with respect to the 3 dB bandwidth can be realized. For digital waveforms the maximum BT is equal to the number of digital samples. If the waveform is oversampled relative to its 3 dB bandwidth then the BT is reduced by that factor. To allow for a meaningfully large BT , M is

fixed such that $M = 512$ samples. This way, even if the expected spectrum has a relatively high oversampling factor relative to the 3 dB bandwidth such as 4, the waveforms produced by the optimized WGF will still possess a relatively high BT of 128.

Number of Random Variables: N – The smaller the value of N , the fewer the number of random variables that have to be instantiated with every waveform and the fewer the number of basis function that have to be summed to instantiate the phase function. From this, it is desirable that N is kept as small as possible. The questions are then how small can N be and still be able to match the desired spectrum and with a small value of N are the waveform instantiations sufficiently different to realize noise like sidelobe reductions with coherent integration. To answer these questions, N will be varied over the powers of 2 between 1 and 512 such that $N \in \{2, 4, 8, 16, 32, 64, 128, 256\}$.

Random Variable Distribution Function: $p_X(x)$ – The possible forms of $p_X(x)$ include all valid PDFs. However, here we only need to consider PDFs with zero mean. The phase mean value vector μ makes giving $p_X(x)$ a mean value unnecessary. Still the options for $p_X(x)$ are vast. The approach taken here is to consider a few common and relatively simple distributions which have mathematically tractable characteristic functions. With this in mind the first distribution to be considered is the binary uniform distribution less formally known as a fair coin toss. That is, the distribution has equal probability of taking one of two states such that

$$\text{DU2: } p_X(x) = \frac{1}{2} \delta(x + \pi) + \frac{1}{2} \delta(x - \pi) \quad (4.24)$$

where here the two states are $\pm\pi$. Alternatively, we also consider a continuous uniform distribution defined as

$$\text{CU: } p_X(x) = \frac{1}{2\pi} \begin{cases} 1 & -\pi \leq x \leq \pi \\ 0 & \text{otherwise} \end{cases} \quad (4.25)$$

where the distribution extends from $-\pi$ to π . In addition to these uniform distributions it is informative to compare them to the Gaussian distribution which has already been evaluated for the purposes of FM noise waveforms in other works such as [27–30, 66, 76]. Here, the Gaussian pro-

cess is defined to have unit variance such that

$$G: p_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (4.26)$$

For convenience the three different $p_X(x)$ will be referred to via shorthand as DU2, CU , and G respectively. These three distributions should provide insight into the behavior and usefulness of highly constrained distributions such as DU2 and less restrictive distributions such as CU and G which can take on a continuum of values.

Desired Spectrum: u – As discussed in Chapter 2, it is often desirable for a waveform to exhibit a Gaussian like spectrum since this spectrum ideally results in zero autocorrelation sidelobes and good spectral roll-off. Still, for low oversampling factors with respect to the 3 dB bandwidth the roll-off is fairly poor. In fact, for an oversampling factor of 2, a Gaussian spectrum only decays to about -12 dB in normalized power by the edge of the sample bandwidth. This may be enough motivation to move away from the Gaussian spectrum in certain cases. One alternative is the *Super-Gaussian* spectral shape defined as

$$u(f) = A \exp\left(\frac{|f|^n}{x_{3dB}}\right) \quad (4.27)$$

where A arbitrarily scales the power, x_{3dB} can be chosen to set B_{3dB} and n is an integer greater than two. If n is set to two, then (4.27) is simply a Gaussian function. Values of n greater than two result in a function with a similar shape as a Gaussian, but with a flatter passband and a much steeper roll-off. As n approaches infinity (4.27) approaches a rectangular function which is the third spectral shape considered here.

The baseband rectangular spectral template is defined such that

$$u(f) = \begin{cases} 1 & -f_B/2 \leq f \leq f_B/2 \\ 0 & \text{otherwise} \end{cases} \quad (4.28)$$

where the 3 dB bandwidth, f_B , is also the absolute bandwidth in this case. This template repre-

sents perfect spectral containment. In Fig. 4.1 these three functions, Gaussian, Super Gaussian with $n = 4$, and the rectangular function, are plotted with an oversampling factor with respect to the 3 dB bandwidth of $K = 2$. The Super-Gaussian decays much more quickly towards the sample bandwidth edges ($\pm B_{\text{samp}}/2$) as compared to Gaussian function and the rectangular functions resembles a "brick-wall" at the 3 dB bandwidth representing the best possible spectral roll-off.

The downside to the Super-Gaussian function as compared to the traditional Gaussian is that it will never achieve zero autocorrelation sidelobes. However, it will not have nearly as high of sidelobes as the rectangular function either. To examine this, the best case autocorrelation responses of these functions are plotted in Fig. 4.2 with an oversampling factor of $K = 8$ in order to provide good visibility of the sidelobe structure. The Gaussian autocorrelation function has no sidelobes shown and if it were zoomed out arbitrarily far it never would. The rectangular spectrum results in the familiar sinc like sidelobes of an LFM with peak levels at roughly -13.4 dB. The super Gaussian spectrum however has a peak sidelobe at roughly -20 dB, but subsequent peaks decay rapidly from this level with the next peaks at roughly -37 dB. The key takeaway from this plot is that even for noise or FM noise waveforms where coherent integration of multiple pulse can be expected to lower the sidelobe levels, if the average spectrum approaches the shapes in Fig. 4.1 their autocorrelation responses can be no better than the shapes in Fig. 4.2.

Historically, the super-Gaussian function has been a topic of much interest in optics [87], here it is suggested a compromise between the zero-autocorrelation sidelobe producing Gaussian spectrum that demonstrates modest spectral containment and the rectangular spectrum which achieves ideal spectral containment, but exhibits poor sidelobe performance. With this in mind, these three templates are used with two different oversampling factors of $K = 2, 4$ plotted in Figs. 4.1 and 4.3 respectively.

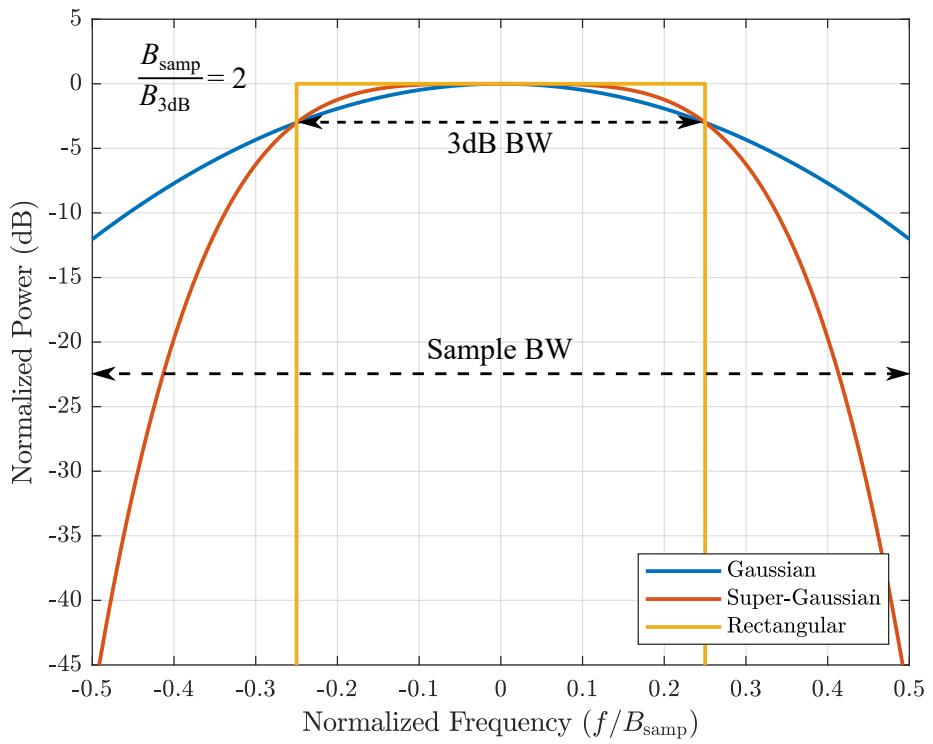


Figure 4.1: Desired spectra for the EFTE optimization with a 3 dB oversampling factor of $K = 2$

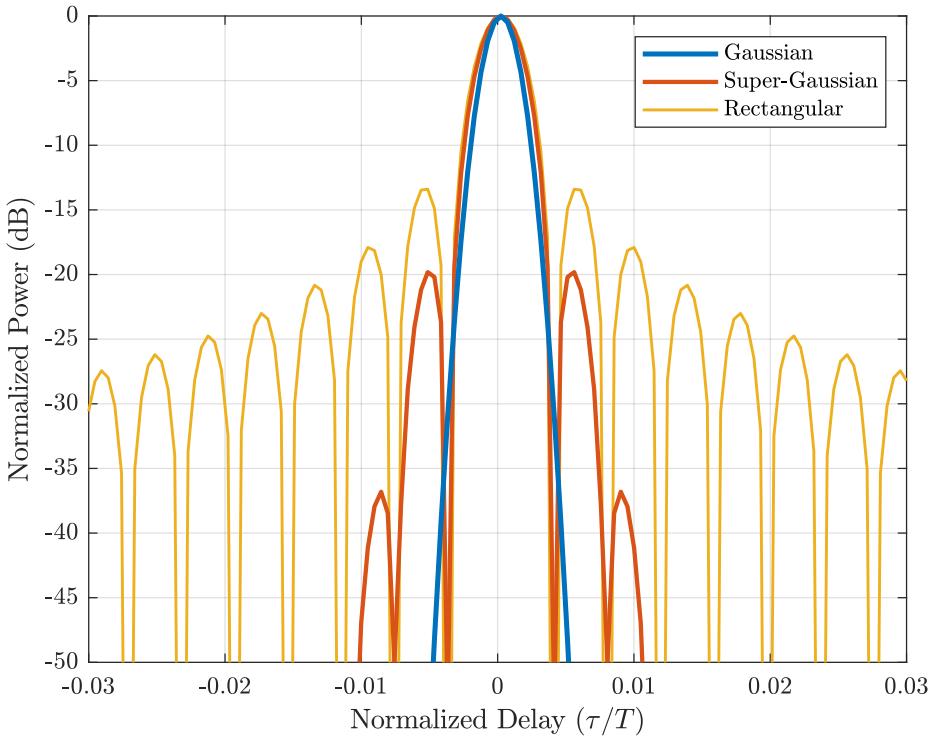


Figure 4.2: Corresponding autocorrelation response of the desired EFTE spectra

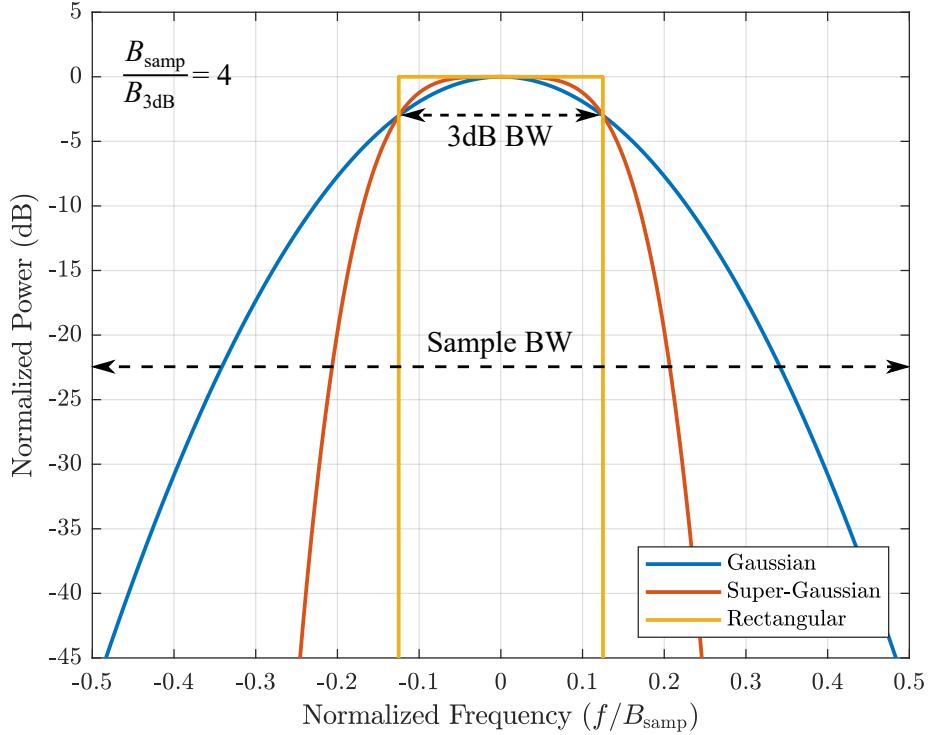


Figure 4.3: Desired spectra for the EFTE optimization with a 3 dB oversampling factor of $K = 4$

For future convenience, these spectral templates are referred to by shorthand. The Gaussian templates with either $K = 2$ or $K = 4$ are G2 and G4 respectively. The Super-Gaussian templates with $n = 4$ and either $K = 2$ or $K = 4$ are S4G2 and S4G4 respectively. The Rectangular templates with either $K = 2$ or $K = 4$ are R2 and R4 respectively.

Initial Basis Function Matrix: \mathbf{B}_0 — As has been shown in other works for non-convex optimization, the initialization can have a dramatic effect on the result of the optimization. In general, the challenge is to find an initialization that tends towards good solutions. For this work, two initializations are examined to demonstrate this behavior. The first one is based on the PCFM framework as discussed in Section 2.2.3. The initial \mathbf{B} has the exact form of the PCFM basis matrix for a given M , N , and oversampling factor of the desired spectrum with respect to the 3 dB bandwidth.

The second initialization is an identity matrix when the number of random variables matches the number of waveform samples, but when $N < M$ the matrix is modified as in Fig. 4.4

$$\frac{M}{N} = 4 \left\{ \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 1 & 0 & & 0 \\ 0 & 1 & & \\ 0 & 1 & & \vdots \\ 0 & 1 & & \vdots \\ 0 & 1 & & \\ & & \ddots & \\ & & & 1 & 0 \\ & & & & 1 & 0 \\ & & & & & 1 & 0 \\ & & & & & & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{array} \right\} M \times N$$

Figure 4.4: Structure of the "Identity" basis matrix initialization (\mathbf{B}_{ID})

For convenience, the PCFM basis matrix initialization is referred to as \mathbf{B}_{PC} and the modified identity matrix initialization is referred to as \mathbf{B}_{ID} .

Initial Phase Mean Value Vector: μ_0 – An asymmetric baseband expected spectrum can be achieved in only a few ways. Either the phase values have a non-zero mean value, the random variable distributions are asymmetric about zero, or the combination of the two. The result of these cases is a characteristic function with an imaginary component leading to an asymmetric baseband expected spectrum. Since this work, is only concerned with the symmetric spectral templates discussed above, it is not necessary for the phase of the StoWGe signals to have a non-zero mean value. Although, for future work this can be considered for such cases as spectral notching where asymmetric baseband spectra are required. For this work however, μ_0 will be initialized as a vector of zeros and will not be optimized.

Summary: – All the initializations and their shorthand notation from above are summarized in Table 4.2.

Variable	Definition	Test Cases	
M	Number of waveform Samples	$M =$	512
N	Number of random variables	$N =$	2, 4, 8, 16, 32, 64, 128, 256
$p_X(x)$	Random Variable Distribution	DU2	– Discrete Uniform, 2 states
		CU	– Continuous Uniform
		G	– Gaussian with unit variance
		G2	– Gaussian ($K = 2$)
		G4	– Gaussian ($K = 4$)
\mathbf{u}	Desired Spectrum	S4G2	– Super-Gaussian ($n = 4, K = 2$)
		S4G4	– Super-Gaussian ($n = 4, K = 4$)
		R2	– Rectangular ($K = 2$)
		R4	– Rectangular ($K = 4$)
\mathbf{B}_0	Basis Function Initialization	\mathbf{B}_{PC}	– PCFM basis matrix
		\mathbf{B}_{ID}	– Modified identity matrix
$\boldsymbol{\mu}_0$	Phase Mean Value Vector	$\mathbf{0}$	– $[0 \ 0 \ \dots \ 0]_{M \times 1}^T$

Table 4.2: Initializations for the StoWGe parameters. All combinations of every optimization are considered resulting in 288 total optimizations

4.4 Optimization Results

In one sense, determining which optimization was the most successful in terms of $p_X(x)$, N , and \mathbf{B}_0 is as simple looking for the most negative values (best) in the tables of appendix B where the results of the test cases outlined in Table 4.2 are tabulated. Objectively speaking, these results provided the closest expected power spectrum matches to the desired templates in the squared error sense. From this perspective the DU2 distribution almost always outperformed the CU and G distributions so at this cursory level is almost always the best choice for the spectral templates considered here. However, the problem with such a sweeping statement is that the goal is not simply to find a WGF that has the desired expected power spectrum, it is to find WGFs which are useful for generating FM noise waveforms. The expected power spectrum is just one aspect of that goal albeit an important one. An excellent example of where minimizing the EFTE cost function

has produced WGFs with near perfect matches to the desired template, but are not useful as FM noise waveforms are when N is small.

4.4.1 Small N

4.4.1.1 Discrete Distributions

According to the tables in Appendix B, for many of the templates the DU2 distribution realizes excellent matches to the desired power spectral template regardless of the value of N . In the case of the S4G2 template, the values of $N = 2, 4$ actually perform the best and even meet the -200 dB stopping point of the optimization, but these excellent matches hide a significant issue.

If $p_X(x)$ is a discrete distribution with a finite number of states then the WGF will likewise have a finite number of sample functions. For the DU2 distribution, two states per random variable, the WGF can only produce 2^N unique waveforms, so when $N = 2, 4$ the expected power spectrum measured by the EFTE is just the sample mean of the 4 or 16 sample functions respectively. With so few sample functions to choose from, the WGF in this case is clearly not useful for generating FM noise waveforms, since the last thing any noise waveforms should do is repeat themselves. In fact, when $N = 2$ there is a 25% chance that two of the same waveform will be transmitted consecutively assuming they are all equally probable as they would be for the DU2 case.

Instead, in a completely roundabout way, these small sets of waveforms are more like complementary waveforms [65, 88, 89], in that they combine to match a desired property. In the Gaussian template cases, they combine to produce zero autocorrelation sidelobes making them explicitly complementary waveforms. Still, what really sets the small N cases for discrete distributions apart is the very small number of member functions. Because of this, only WGFs with a sufficiently large sample space will be considered for FM noise purposes.

To a degree "sufficiently large" is somewhat arbitrary, but it can be thought of this way. Given a 100 pulse CPI, or some other reasonably sized CPI, what are the odds that the same waveform is repeated in those 100 pulses? For the DU2 case, all sample functions are equally probable so the answer is just $100(1/2^N)$. For $N = 2, 4, 8, 16$, the solutions are 100%, 100%, 39%, and 0.15%

respectively. With $N = 16$, in about 15 out of every 1000 CPIs there will be 2 of the same sample function in a randomly chosen set of 100. Although with $N = 16$ the chance of a repeat is small, out of an abundance of caution this work will only consider values of $N \geq 32$ as useful for FM noise purposes from here on out for the DU2 distribution.

4.4.1.2 Continuous Distributions

For the CU and G distributions, the small N cases tend to perform worse than the larger N cases according to the tables in Appendix B. Consequently, they are of less concern since they would not be considered "good" in the first place. Still, they sometimes do outperform the larger N cases so it is prudent to examine the behavior of the WGF with a continuous $p_X(x)$ and small N .

Being continuous, regardless of the value of N the WGF with a continuous $p_X(x)$ can produce an infinite number of sample functions so the small sample function set size will not be an issue here. However, they do present a less obvious issue that is revealed by examining the power spectral deviation as defined in (3.64).

In Table B.1, the \mathbf{B}_{PC} initialization results in a better match to the power spectral template for smaller values of N for $p_X(x) = \text{CU}$, and nearly equivalent results for any value of N for $p_X(x) = \text{G}$. Figs. 4.5 and 4.6 show the resulting power spectral deviation for the CU case and the G case respectively for values of N including 2, 8, 32, and 128. Additionally, the G2 template is shown for reference however the resulting expected power spectra for each case are not. For the CU case, the optimization was so effective that the expected power spectra are virtually indistinguishable from the desired template.

The most notable aspects of Figs. 4.5 and 4.6 is the tremendous spike in the power spectral deviation at 0 frequency for the smaller values of N . As N increases, the spike decreases and the power spectral deviation function becomes more smooth overall until in the $N = 128$ case it is indistinguishable from the template meaning the deviation is on the order of the expected power spectrum. This is a striking parallel to the periodogram of Gaussian processes where the variance of the power spectrum estimate is likewise on the order of the measurement itself [35].

For the smaller values of N , the spike is reasonable. With so few random variables to construct the phase for any given sample function the probability that they are all close to zero or all well away from 0 is non-negligible. This represents a problem from a pulse integration standpoint. Because of this, it will take more pulse integration for the sample power spectrum to resemble the expected power spectrum. To safely avoid this issue only WGFs with $N \geq 32$ will be considered for either the CU or G distributions same as the DU2 case.

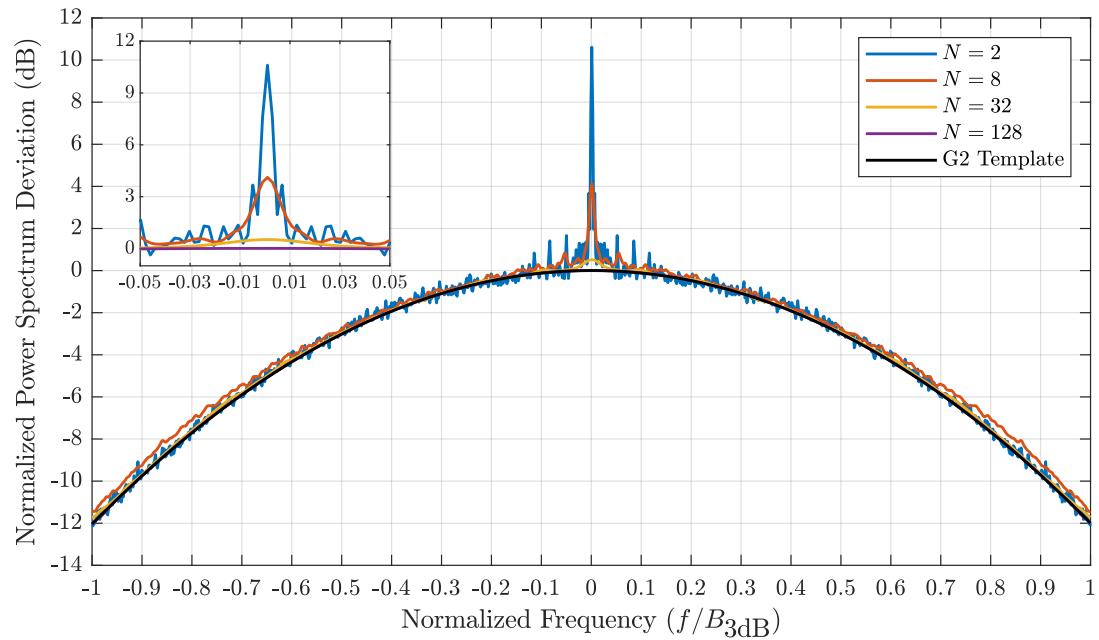


Figure 4.5: Normalized power spectral deviation for the DU2 distribution, the G2 template, and various values of N

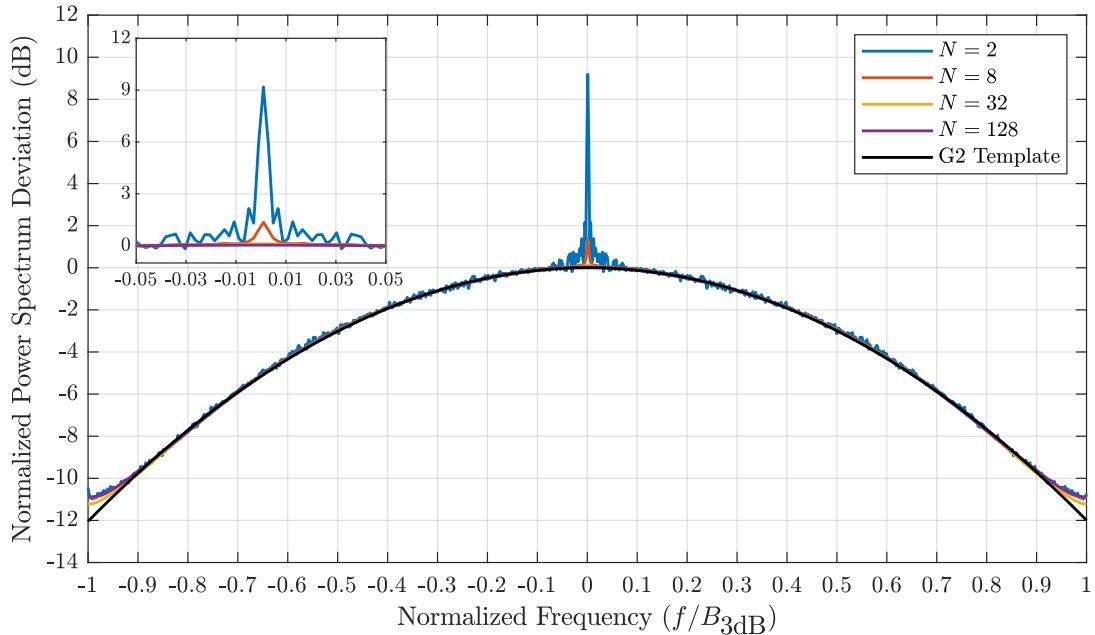


Figure 4.6: Normalized power spectral deviation for the G distribution, the G2 template, and various values of N

4.4.2 Pulsed StoWGe Evaluation

With the small N cases addressed, the question becomes how to evaluate the effectiveness of the rest of the WGFs for instantiating useful radar waveforms. At a first level, it is only necessary to examine the cases with the best resulting cost function values as a function of \mathbf{u} and $P_X(x)$ for $N \geq 32$. With six templates and three distributions this leaves 18 WGFs to examine in more detail. These cases are collected in Table 4.3.

Beyond the template match, it is important to understand how these WGFs perform in other ways, but especially in the time domain and on an individual basis. Such an analysis can be performed by utilizing several of the metrics discussed in 3.4. The metrics deemed helpful for evaluating the usefulness of the optimized WGFs are

- Expected power spectrum - This metric is directly optimized by the cost function. It is clearly relevant to the analysis.

Template (u)	Distribution ($p_X(x)$)											
	DU2				CU				DU2			
	J	N	\mathbf{B}_0		J	N	\mathbf{B}_0		J	N	\mathbf{B}_0	
Template (u)	G2	-154.1	128	\mathbf{B}_{PC}	-154.0	32	\mathbf{B}_{Id}		-74.2	32	\mathbf{B}_{PC}	
	G4	-200.0	128	\mathbf{B}_{PC}	-115.1	32	\mathbf{B}_{PC}		-88.5	256	\mathbf{B}_{PC}	
	S4G2	-157.3	32	\mathbf{B}_{PC}	-89.6	32	\mathbf{B}_{PC}		-46.2	64	\mathbf{B}_{PC}	
	S4G4	-88.8	64	\mathbf{B}_{PC}	-76.2	64	\mathbf{B}_{PC}		-42.0	32	\mathbf{B}_{PC}	
	R2	-57.1	128	\mathbf{B}_{PC}	-42.9	32	\mathbf{B}_{PC}		-37.1	32	\mathbf{B}_{PC}	
	R4	-48.1	64	\mathbf{B}_{PC}	-39.6	32	\mathbf{B}_{PC}		-33.8	32	\mathbf{B}_{PC}	

Table 4.3: The best resulting WGFs for each $p_X(x)$ and template as a function of N for $N \geq 32$ and \mathbf{B}_0 .

- Power spectral deviation - While it is essential that the waveforms on average have a good spectrum, it is also useful to understand how the power spectrum of a given waveform can be expected to deviate from the expected power spectrum. If each waveform deviates wildly from the expected spectrum, it may take much more coherent integration for the sample power spectrum to resemble the expected power spectrum as opposed to a WGF with much less power spectral deviation.
- Expected autocorrelation - For a WGF to be useful for creating noise like waveforms, at a minimum the autocorrelation sidelobes should decrease with coherent integration. The expected autocorrelation represents the lowest possible sidelobes given sufficient coherent integration. In general, the expected autocorrelation should have extremely low sidelobe levels, but ultimately this depends on the spectral template and the WGFs ability to realize that template.
- Expected RMS autocorrelation - While it is useful to know that the sidelobe levels will decrease with coherent integration, it is likewise useful to know at what level they will begin. The expected RMS autocorrelation represents the expected autocorrelation level of a single WGF generated waveform.

- Expected RMS cross-correlation - If any single pair of StoWGe waveforms are filtered with each other, the expected RMS cross-correlation represents the magnitude of this result. In general, it is desirable that unique waveforms have little cross-correlation to maximize separability and reduce ambiguity.

This of course is only five of the eight metrics listed in Table 3.3. The other three, are in general, unhelpful for this analysis. Consider the expected cross-correlation and the expected cross-power spectrum. In either case these measure the integration of pulse pairs under a cross-correlation or as a product of their power spectra. Since each waveform in each pair are different sample functions, zeroing out the expected cross-correlation and the expected cross-power spectrum is as simple as putting a random initial phase on each sample function according to an uniform distribution on $U [-\pi, \pi]$. In any radar setting, this can be done without any performance loss since matched filtering will remove this phase anyway allowing for coherent integration without issue.

The other unused metric is the RMS cross-power spectral density which measures the expected magnitude of the product of two different power spectra. The formulation of StoWGe guarantees that any waveform instantiation is independent of all others. For the RMS cross-power spectral density this means the unique waveform terms in 3.66 are separable such that

$$\left(E \left[|s_{af} \odot s_{bf}^*|^2 \right] \right)^{1/2} = \left(E \left[|s_{af}|^2 \right] E \left[|s_{bf}|^2 \right] \right)^{1/2} = E \left[|s_f|^2 \right] \quad (4.29)$$

which is simply the expected power spectrum again. Still, this could be useful in another context such as when there is coupling between different instantiations of the waveforms, but for StoWGe this simply is not the case. The next three sections discuss the results for each type of template (Gaussian, Super-Gaussian, Rectangular). For each template all of the five different metrics which were deemed useful are plotted for each $p_X(x)$ in order to provide an informative understanding of how each distribution is suited to each template.

4.4.2.1 Gaussian Template Results

In Figs. 4.7 and 4.8 the spectral and temporal metrics of each of the two Gaussian templates respectively. Judging from Fig. 4.7(a,b) each of the different distributions were able to achieve fairly good matches to the Gaussian templates where for the G2 template the expected spectra are indistinguishable from temple while for the G4 template there is some deviation from the template on the part of the CU and G distributions at the sample bandwidth edges.

The spectral deviation plots are a little more interesting. In 4.7(c), the DU2 distribution realizes some periodic, small spikes in deviation while in 4.7(d), the G distribution shows much more variance over its bandwidth compared to the others indicating its individual waveforms could on occasion have much more or less power in the roll-off regions. In Fig. 4.8(a,b), the expected autocorrelation (ACF) for each distribution either matches the templates resulting ACF or even outperforms it in case of the G distribution. Although, the CU distribution does somewhat worse for the G4 template, but not meaningfully so. The expected autocorrelation represents the resulting autocorrelation with infinite coherent integration. The autocorrelation levels of noise like waveforms decrease at a rate of $10\log_{10}(L)$ where L is the number of waveforms being coherently integrated. If a given waveform of a WGF realizes sidelobe level around -40 dB, it would still take the coherent integration of 1000000 waveforms to run into the sidelobe floor presented by the CU distribution in 4.8(b), so really the expected autocorrelation sidelobes only need to be as low as what can be reasonably obtained given the RMS autocorrelation sidelobes of a given WGF and the size of the CPI.

In Fig. 4.8(c), the RMS autocorrelations for the poorly contained spectrum of the G2 template all exhibit a fast oscillatory component with the DU2 distribution having the highest amplitude oscillations. For the G4 template in Fig. 4.8(d), the G distribution exhibits a broadening at the base of the mainlobe and a couple of fairly large oscillations thereafter. The CU distribution has some higher frequency oscillations while for the DU2 distribution, there is a small bump at the mainlobe but is then smooth thereafter. For noise and noise like waveforms without any optimization, the peak sidelobe level should be at about $-10\log_{10}(BT)$ which are -24.1 and -21.1 for the G2 and

G4 templates respectively. Even accounting for the oscillations, roughly speaking each choice of distribution was able to meet or exceed these values. Though, the broadening of the base of the mainlobe in the G distribution case for the G4 template makes it less obvious where the sidelobe region truly begins.

In Fig. 4.8(e), the oscillations present for the G2 template RMS autocorrelations are likewise present in the RMS cross-correlations. However in Fig. 4.8(f), the large scale oscillations disappear from the RMS autocorrelations. Overall however, the RMS cross-correlation results largely resemble the RMS autocorrelation results but without a mainlobe.

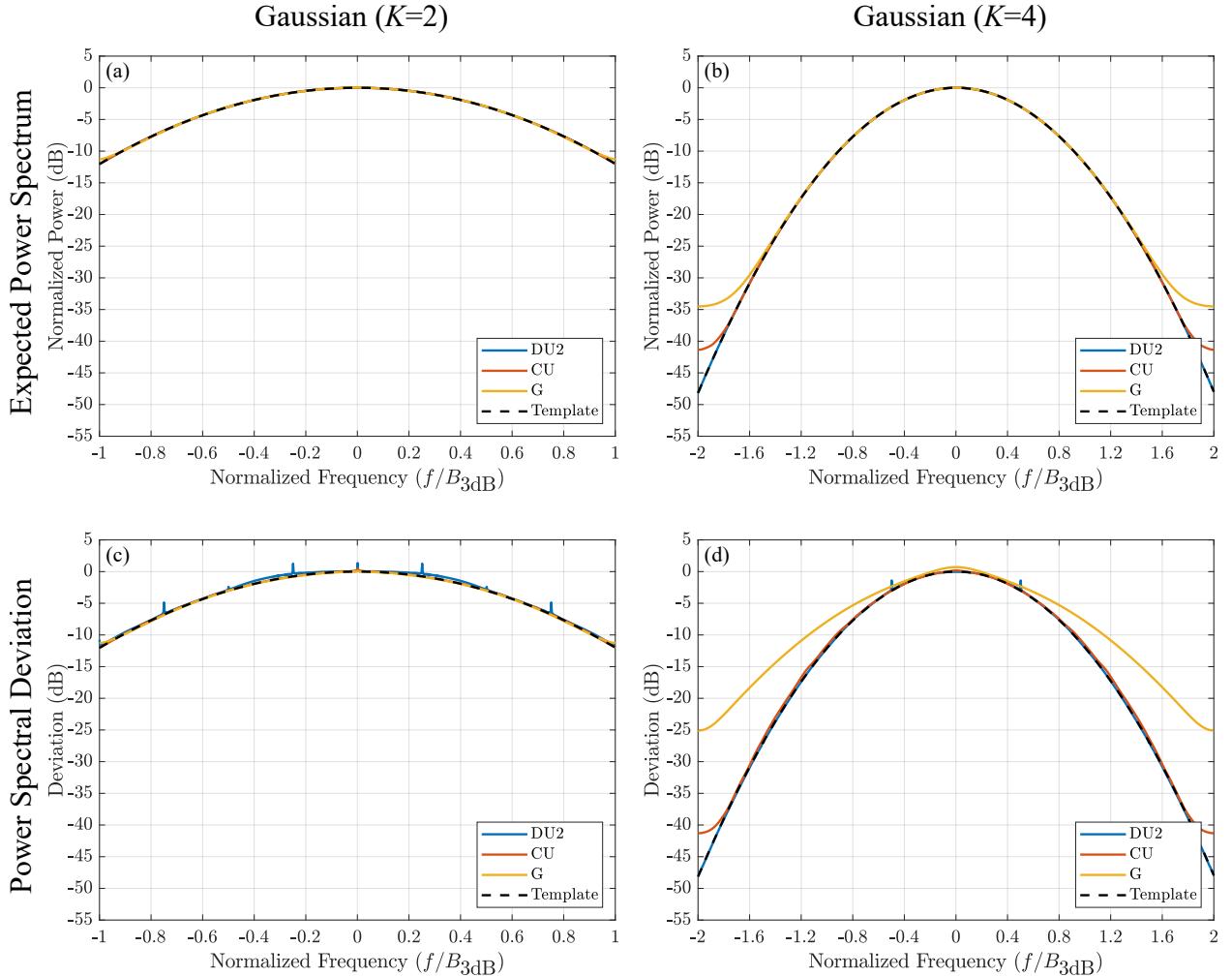


Figure 4.7: Optimized expected power spectrum for the G2 template (a). Optimized expected power spectrum for G4 template (b). Expected power spectral deviation for the G2 template after optimization (c). Expected power spectral deviation for the G4 template after optimization (d).

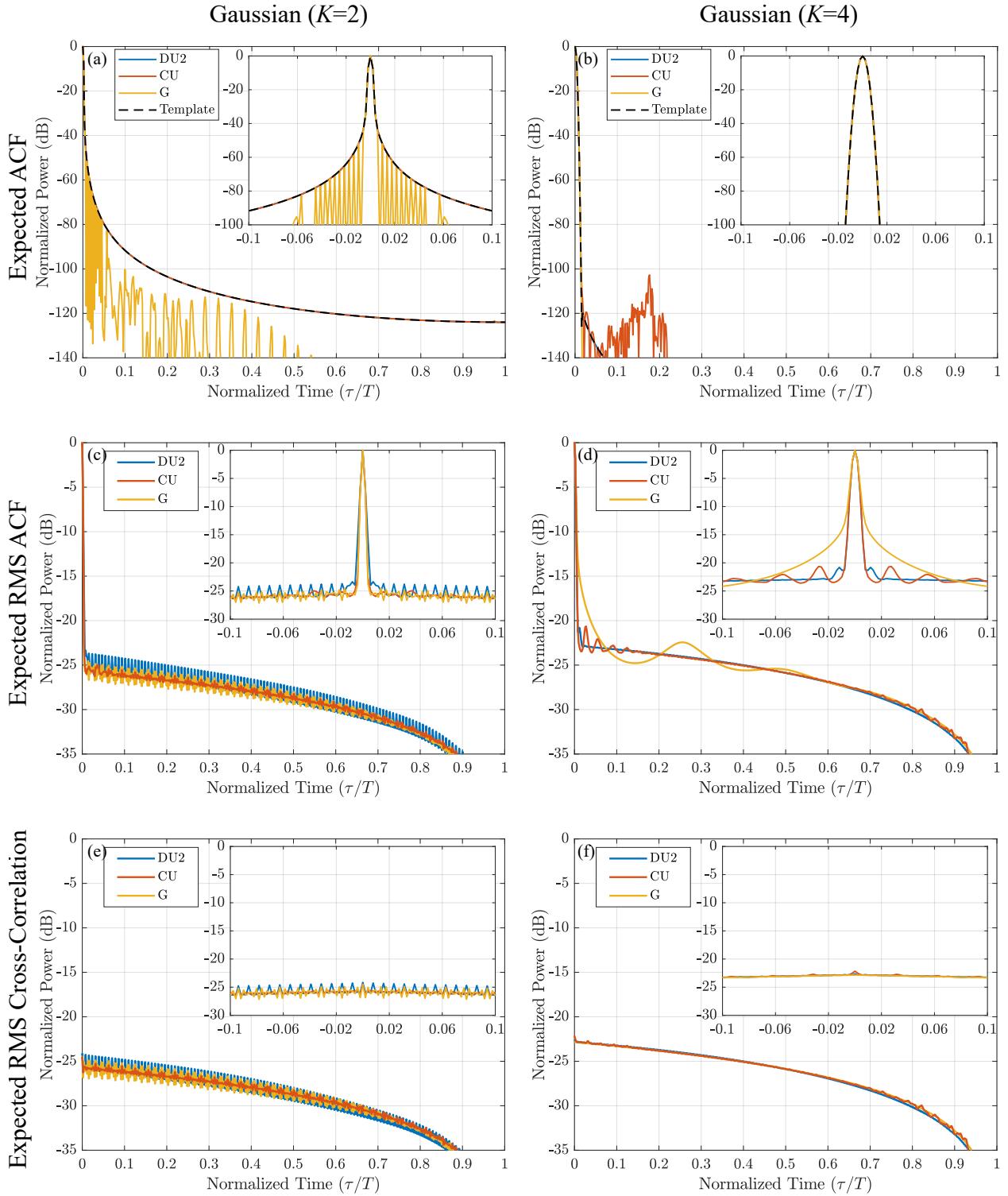


Figure 4.8: Expected autocorrelation for the G2 template after optimization (a). Expected autocorrelation for the G4 template after optimization (b). Expected RMS autocorrelation for the G2 template after optimization (c). Expected RMS autocorrelation for the G4 template after optimization (d). Expected RMS cross-correlation for the G2 template after optimization (e). Expected RMS cross-correlation for the G4 template after optimization (f).

4.4.2.2 Super-Gaussian Template Results

Compared to the G2 and G4 templates, there is much more disparity between the distributions when it comes to matching the S4G2 and S4G4 templates. In Fig. 4.9(a,b) the G distribution does a very poor job of matching to either template. The CU distribution does much better in comparison. For the S4G2 template in 4.9(a) the DU2 expected power spectrum is nearly indistinguishable from the template while for the S4G4 template of 4.9(b) it matches the template down to about -35 dB which clearly outperforms the other the other distributions. In 4.9(c,d), the spectral deviation plots are largely proportional to the expected power spectra. However, the G distribution experiences a somewhat larger degree of spectral deviation while the DU2 case for the S4G2 template has some small spikes similar to its behavior for the G2 template in 4.7(c).

As opposed to the G2 and G4 templates, the S4G2 and S4G4 templates have near in sidelobes. In 4.10(a,b), each distribution was largely able to match these sidelobes. However, beyond the first few sidelobes the distributions start to really deviate below -80 dB. However, these sidelobes are still low enough that they would only ever be seen with an exceptionally large amount of coherent integration.

The RMS autocorrelations for the S4G2 template in Fig. 4.10(c) exhibit some ringing as they did for the G2 template in 4.8(c), but the DU2 distribution now exhibits very prominent oscillations suggesting some kind of periodic component in the WGF. This will be revisited in 4.4.2.4. Interestingly, these oscillations almost completely disappear for the S4G4 template in 4.10(d), however the G distribution shows the large scale oscillations that were also present in the G4 template of 4.8(d).

The RMS cross correlations in Fig. 4.10(e,f) once again are very similar to the RMS autocorrelation, but with less dramatic oscillations in the S4G4 case and with the mainlobe removed.

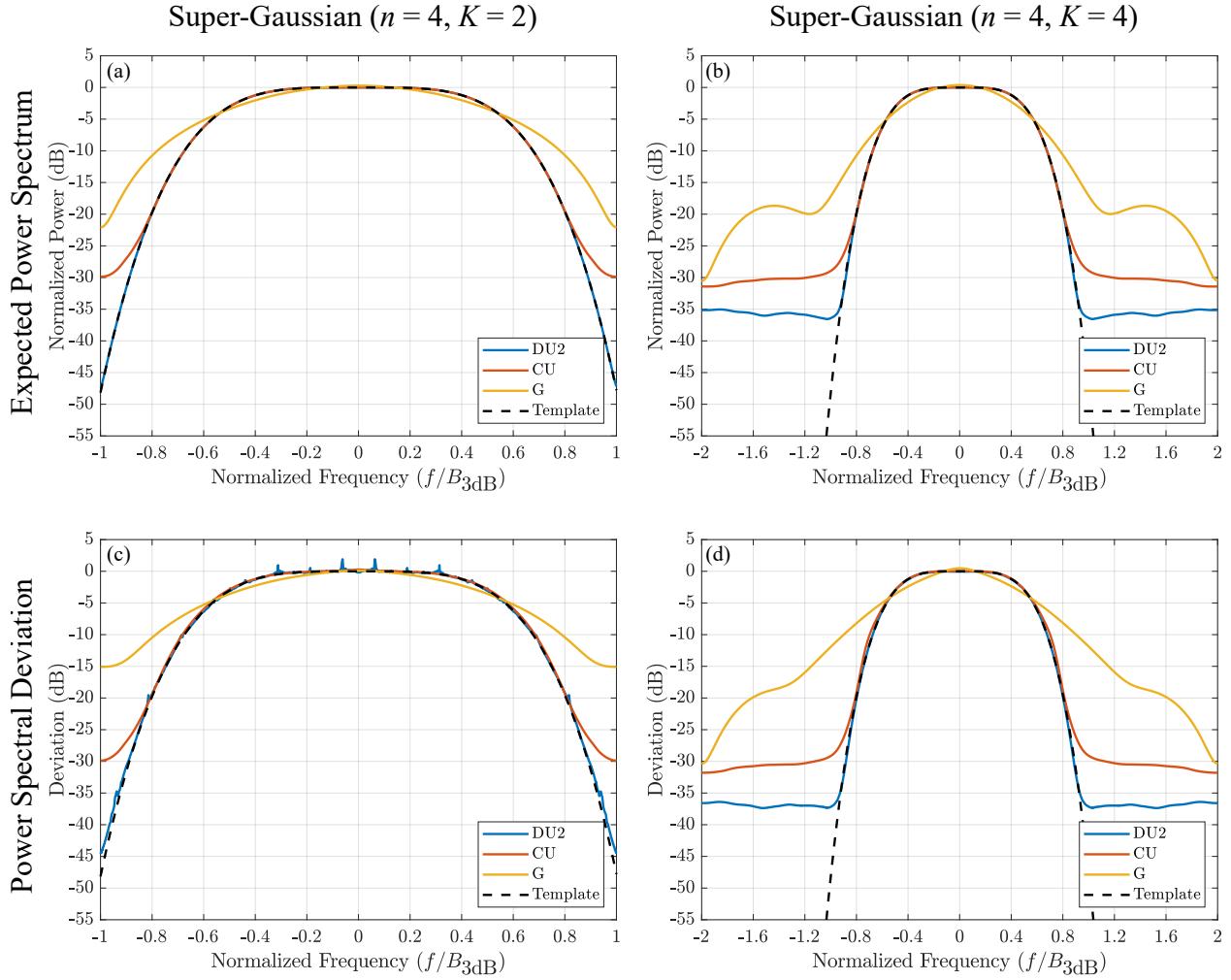


Figure 4.9: Optimized expected power spectrum for the S4G2 template (a). Optimized expected power spectrum for S4G4 template (b). Expected power spectral deviation for the S4G2 template after optimization (c). Expected power spectral deviation for the S4G4 template after optimization (d).

4.4.2.3 Rectangular Template Results

The rectangular template represents the most restrictive power spectrum and unsurprisingly it resulted in the worst optimized values of the EFTE cost function of the three template types. However, the DU2 distribution performed significantly better than either the G or CU distributions. In Fig. 4.11(a,b) the CU distribution was able to match to the template marginally better than the G distribution, but they both achieved poor matches the template. The DU2 distribution did a pretty good job implementing the brick wall at the 3 dB bandwidth, but even it bottomed out at about

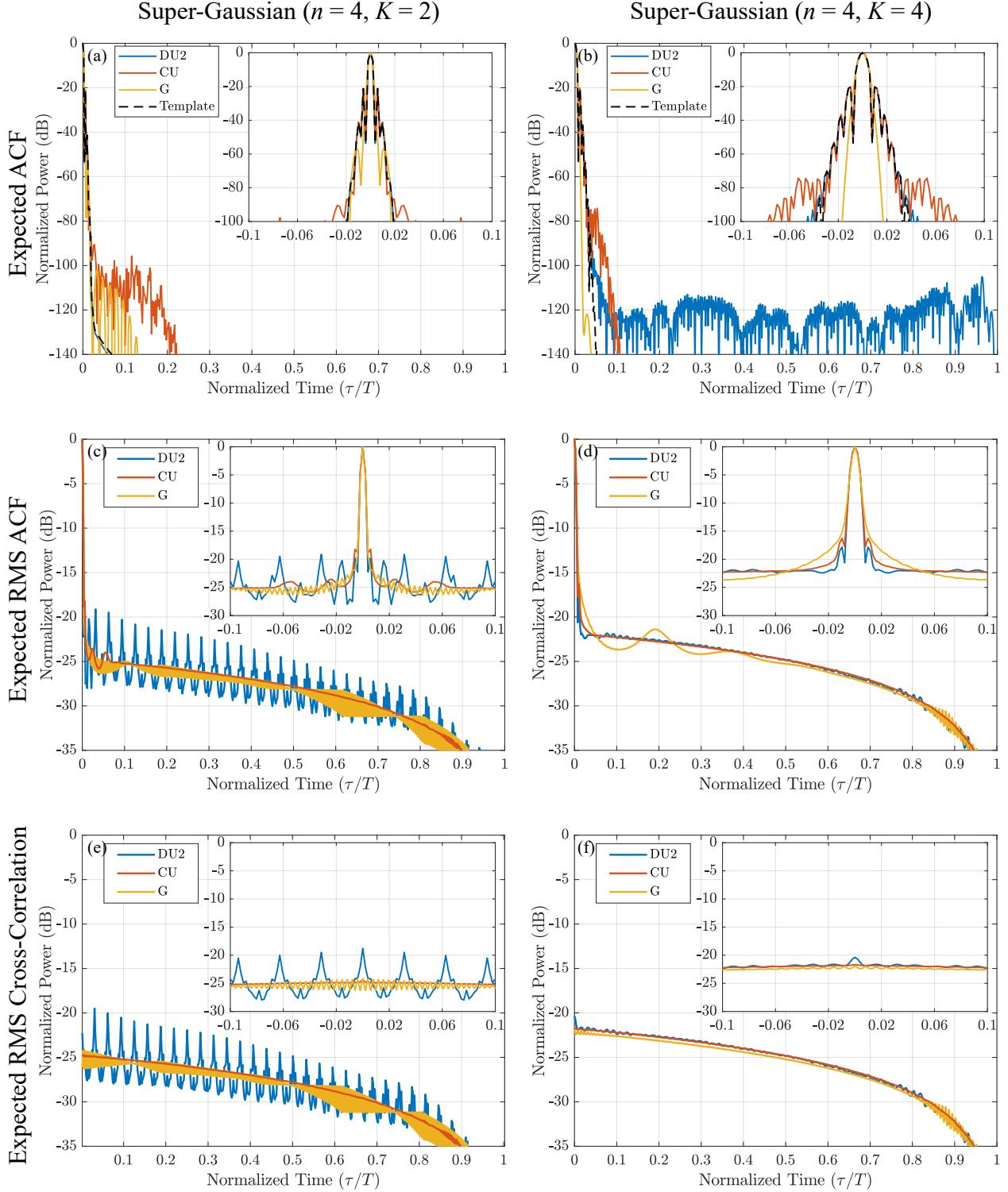


Figure 4.10: Expected autocorrelation for the S4G2 template after optimization (a). Expected autocorrelation for the S4G4 template after optimization (b). Expected RMS autocorrelation for the S4G2 template after optimization (c). Expected RMS autocorrelation for the S4G4 template after optimization (d). Expected RMS cross-correlation for the S4G2 template after optimization (e). Expected RMS cross-correlation for the S4G4 template after optimization (f).

-25 dB in the roll-off region in the R2 case. In the R4 case of 4.11(b) it realized a sort of spectral pedestal before dropping steeply again to -35 dB.

The spectral deviation plots in Fig.4.11(c-d) are roughly proportional to the spectral results, however in either the R2 or the R4 case the DU2 spectral deviation is slightly smoother and lower at the edges of the template bandwidth than what would be expected from the expected spectrum. This indicates that from sample function to sample function, these edges are a little more consistent than the rest expected spectrum.

Compared to the sidelobes of the other templates, the R2 and R4 templates exhibit dramatically higher sidelobe levels akin to the sinc like sidelobes of an LFM. In Fig. 4.12(a,b), the previously included template trace has been left off to improve clarity. However, the DU2 trace effectively follows what would be expected of the template owing to its decent spectral match. The poorly matched G and CU expected spectrum result in likewise poor matches to the template's expected autocorrelation.

Fig. 4.12c,d shows the expected RMS autocorrelation response where the first several sidelobes of the expected autocorrelation are clearly visible. No matter how much coherent integration is used these will remain. The rest of the expected RMS autocorrelation behaves similarly to the other templates. The DU2 distribution does have an interesting bump in the expected RMS autocorrelation towards the edges of the sidelobe regions. It is important to note, that as the DU2 waveforms are coherently integrated, their autocorrelation response will more and more resemble the sinc like response in 4.12.

The expected RMS cross-correlation also presents an interesting behavior in the DU2 distribution case. In Fig. 4.12(e,f), the DU2 trace has a correlation spike at 0 delay indicating that there is some degree of correlation between the samples functions of its corresponding WGF. It would appear this correlation was need to achieve the good spectral matches shown in 4.11(a-b).

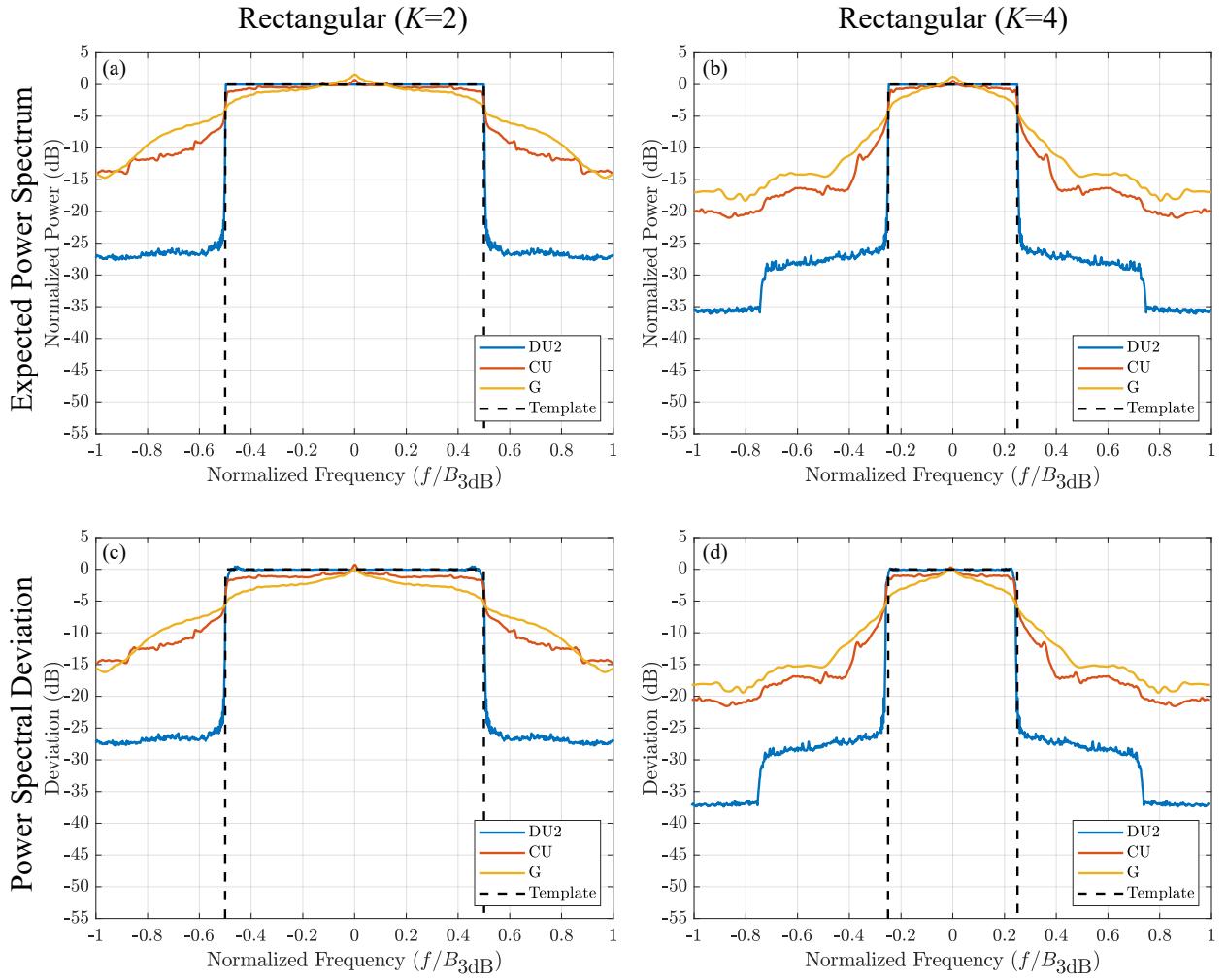


Figure 4.11: Optimized expected power spectrum for the R2 template (a). Optimized expected power spectrum for R4 template (b). Expected power spectral deviation for the R2 template after optimization (c). Expected power spectral deviation for the R4 template after optimization (d).

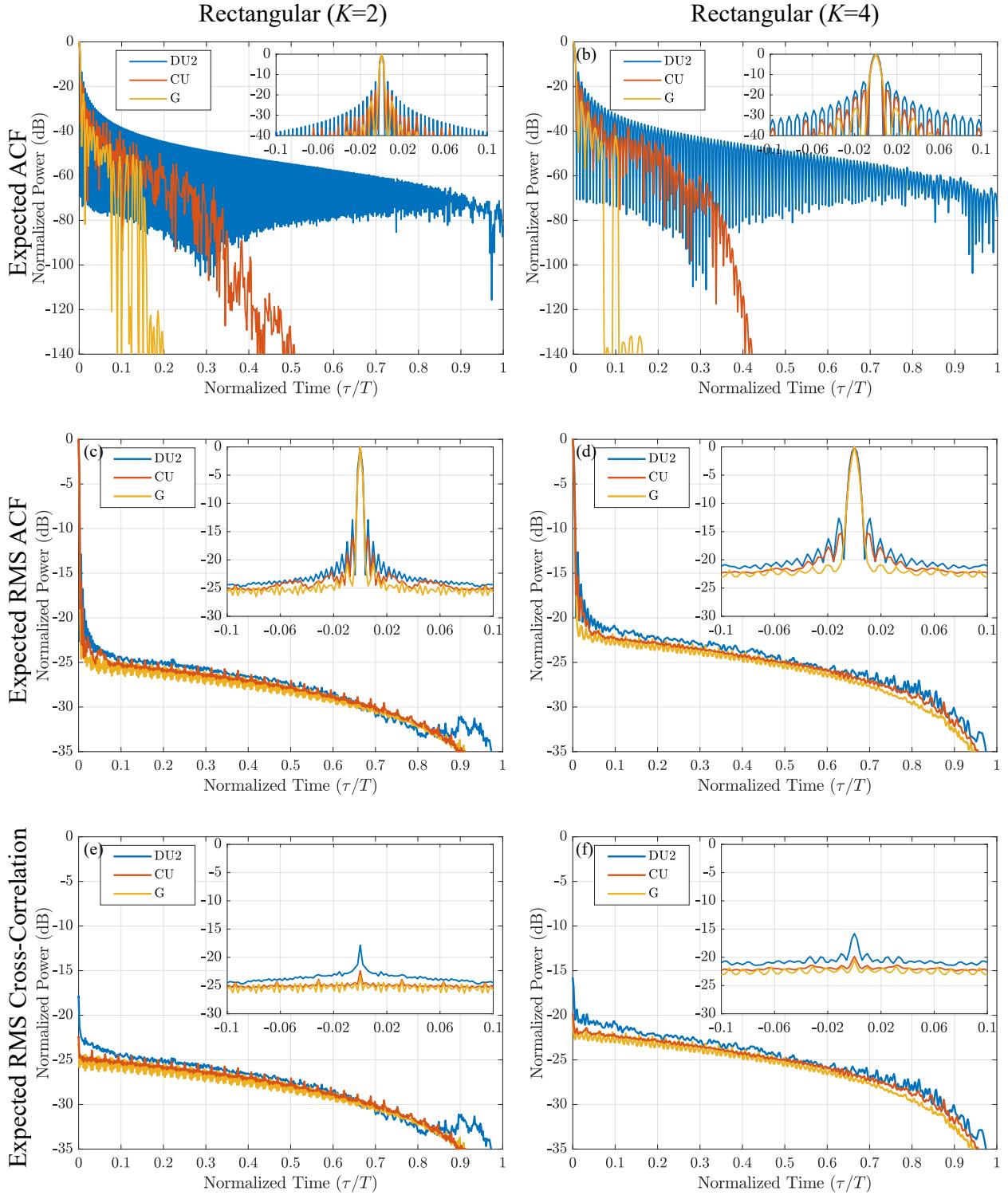


Figure 4.12: Expected autocorrelation for the R2 template after optimization (a). Expected autocorrelation for the R4 template after optimization (b). Expected RMS autocorrelation for the R2 template after optimization (c). Expected RMS autocorrelation for the R4 template after optimization (d). Expected RMS cross-correlation for the R2 template after optimization (e). Expected RMS cross-correlation for the R4 template after optimization (f).

4.4.2.4 DU2 Distribution and S4G2 Spectral Template Case Study

Section 4.4.2.2 commented on the large oscillations present in the expected RMS autocorrelations and cross correlations of the DU2 distribution for the S4G2 template in Fig. 4.10(c,e). While the $N = 32$ case provided the best spectral match it also produced those dramatic oscillations that result in higher autocorrelation sidelobes. This begs the question, is there another result from the DU2 results of Table B.3 that produces a better expected RMS autocorrelation? Although it would have a poorer spectral template match, if the expected RMS autocorrelation is much better the poorer spectral performance may be worth it.

To explore this idea, all the results for $N \geq 32$ for the DU2 distribution in Table B.3 are plotted in Fig. 4.13. Each sub figure has 8 traces corresponding to $N = 32, 64, 128, 256$ and $\mathbf{B}_0 = \mathbf{B}_{\text{PC}}, \mathbf{B}_{\text{ID}}$. The traces belonging to the $\mathbf{B}_0 = \mathbf{B}_{\text{PC}}, N = 32$ case from Figs. 4.9 and 4.10 are plotted in blue. While what is considered to be the best WGF here, the $\mathbf{B}_0 = \mathbf{B}_{\text{PC}}, N = 256$ case, is plotted in red. For convenience since they utilized the same \mathbf{B}_0 , these case will be referred to as $N = 32$ and $N = 256$. All other traces are plotted in gray for clarity.

Interestingly, despite achieving a cost function value that is more than 13 dB worse than the ideal case, the expected spectrum trace in Fig. 4.13(a) of the $N = 256$ case and the $N = 32$ are nearly indistinguishable. This is reasonable though. In linear terms the total difference between -157.3 dB and -142.9 dB is quite small even if the ratio between them is more than an order of magnitude.

Perhaps more notably, the spectral deviation of the $N = 256$ case in Fig. 4.13(b) does not possess the small peaks seen in the $N = 32$ case and on others. Further, the expected RMS autocorrelation and cross-correlation of the $N = 256$ case in Fig. 4.13(d,e) lack the large oscillations apparent in the $N = 32$ case indicating these features may be related. Since the $N = 256$ WGF provides largely the same expected power spectrum performance as the $N = 32$ case while achieving a better expected RMS autocorrelation, the $N = 256$ appears to be a better choice as a WGF.

All Results for the DU2 distribution and the S4G2 template

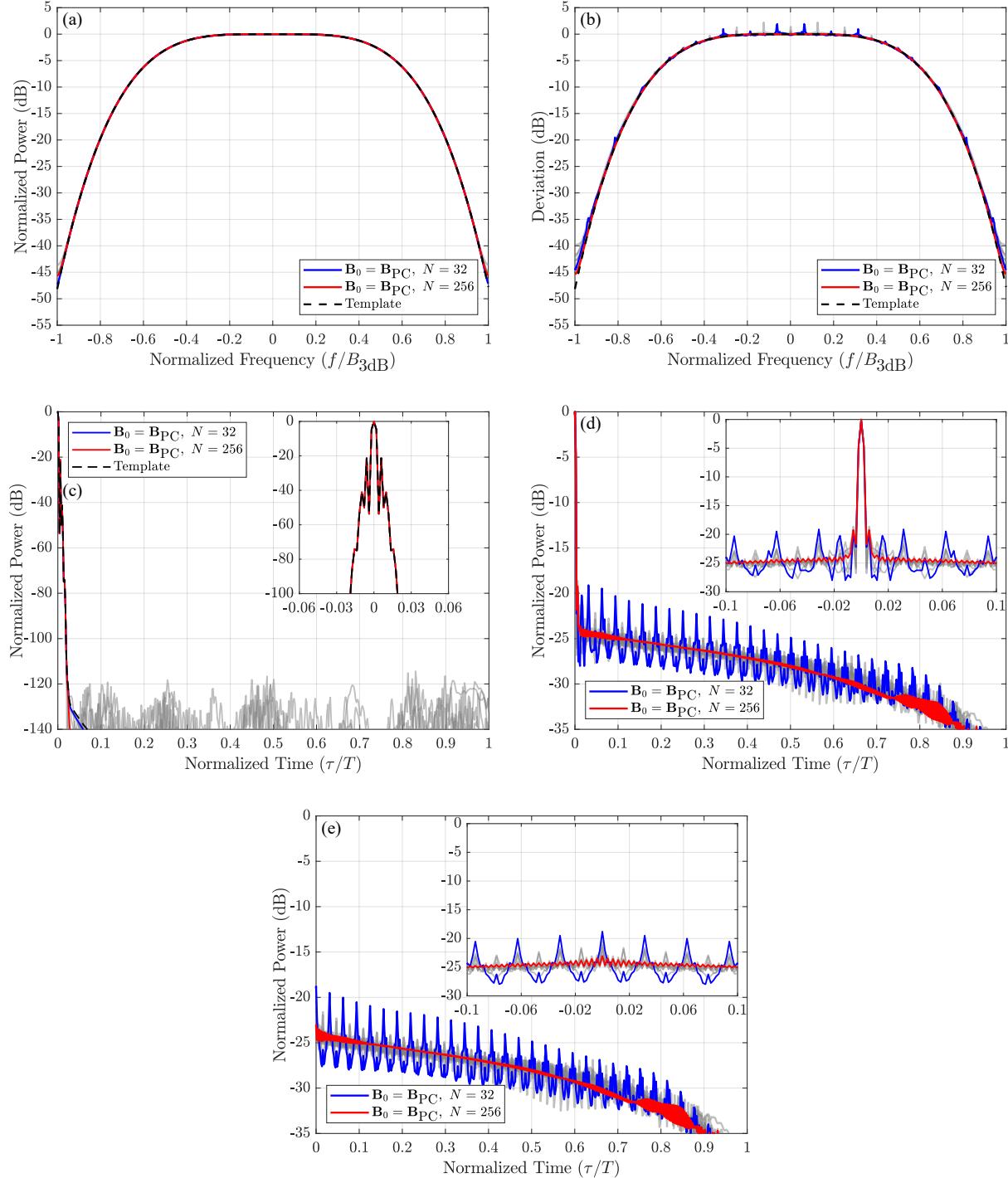


Figure 4.13: All results for the DU2 distribution and the S4G2 template. $N = 32$ is plotted in blue. $N = 256$ is plotted in red. (a) expected power spectrum (b) power spectral deviation (c) expected autocorrelation (d) expected RMS autocorrelation (e) expected RMS cross-correlation

4.4.2.5 Basis and Correlation Matrices

The previous sections looked at how well the StoWGe model and the EFTE optimization is able to optimize for a desired expected power spectrum. This section looks at how it did so. The correlation matrices and selected basis functions for the $K = 4$ templates (G4, S4G4, R4) are plotted in Figs. 4.14 - 4.16 respectively. Sixteen of the N basis functions are plotted in each case since plotting all of them makes the plots difficult to interpret. The correlation matrices are plotted in magnitude on a dB scale. Since the optimized WGFs have a 0 valued μ , the correlation matrices are real and with the exception of the Gaussian distributed RVs, they can be negative.

Beginning with the G4 template, in Fig. 4.14(a,c,e) each distribution has optimized towards unique basis function shapes from the smooth wavelet like functions of the G distribution to the step like functions of the DU2 distribution. While this in itself is interesting, for gaining intuition into how the distributions achieve their respective curves in Figs. 4.7 and 4.8, it is more important to examine the correlation matrices in 4.14(b,d,f).

Consider the G distribution correlation matrix in 4.14(b). Due to the nature of the characteristic function of Gaussian distributions and the fact that μ is set to zeros, the correlation matrix is strictly positive. Now recall that the expected autocorrelation is defined as the sum across the diagonals of the correlation matrix and the expected spectrum is the IFFT of that result. In 4.7(b), the G result nearly matches the template. Since the correlation matrix is strictly positive, the only way for this to happen is for the regions above and below the main diagonal to be extremely close to zero. Indeed, in 4.14(b) this is the case.

The CU and DU2 results are a little more nuanced. Their correlation matrix values can be negative, so while there may be meaningful correlation in the diagonals off from the main diagonal, it is possible the sum of those diagonals is small resulting in low autocorrelation sidelobes. This appears to be the case for the DU2 correlation matrix in 4.14(f) where significant correlation is observed at the bottom left and top right corners perhaps as large as -12 dB indicating the beginning and end of the DU2 waveforms are somewhat correlated. However, in 4.8(b) the resulting expected autocorrelation is a near perfect match to the templates ideal autocorrelation so the correlation must

cancel itself out when summed. In other words, since the waveforms are not necessarily stationary over the pulse duration, the matched filter result only depends on the average correlation at a given delay hence the summing operation over the diagonals. Intriguingly, the DU2 distribution resulted in the best template match in 4.7(b) meaning the correlation between the ends of the waveform may be helpful to further improve spectral containment.

For the S4G4 template results in 4.15 recall that the G distribution did a poor job of matching to the spectral template (4.9(b)). In an attempt to match to the template, the optimization resulted in sort of square shaped steps at the edges of the basis functions. The impact of these features is seen on the main diagonal of the resulting correlation matrix in 4.15(b) where the top left and bottom right regions of the diagonal become like an identity matrix indicating little if any correlation between samples. Despite its best attempts this was not enough and the G distribution appears ill-suited to achieve the S4G4 template.

The CU and DU2 distributions did significantly better, but the DU2 distribution once again provided the best performance in the roll-off region. In this case, some sidelobes are present in the template autocorrelation meaning some correlation should be expected off from the main diagonal as is the case for the CU and DU2 distributions in Fig. 4.15(d,f). The CU distribution resulted in smooth basis functions while the DU2 distribution resulted in relatively smooth basis function, but with periodic spikes. If Fig. 4.10(b), the DU2 distribution produced the best match to the template autocorrelation although it exhibits a sidelobe floor around -110 dB. This floor is attributable to the significant correlation in the upper and lower triangular regions of Fig. 4.15(f) where evidently the correlation does not cancel as well as it did in the G4 case (4.14(f)).

Finally, for the R4 template in 4.16, only the DU2 distribution was able to produce a decent match to the desired template (Fig. 4.11(b)). Like the with the G distribution for the S4G4 template, both the G and CU distributions optimized towards jagged basis functions and bizarre correlation matrices. This appears to be an indicator the optimization had trouble achieving a good template match. Considering the LFM like sidelobes of the R4 template it makes sense that the DU2 correlation matrix in 4.16(f) has significant correlation throughout the entire matrix. This was

necessary to achieve the expected autocorrelation in Fig. 4.12(b) and the corresponding spectrum in 4.11(b).

4.4.3 Pulsed StoWGe Optimization Conclusions

Overall, the StoWGe model was capable of minimizing the EFTE cost function to realize various desired expected power spectral densities. However, the choice of various underlying parameters had a significant impact on how well the gradient descent optimization was able to do so and how useful the StoWGe WGFs are for producing good radar waveforms.

In Section 4.4.1 it was shown that N should be kept larger than a few dozen. In Section 4.4.2, it was shown that the DU2 distribution largely outperforms either the G or the CU distributions for any of the templates considered here and for either of the basis function initializations. This in itself is an interesting result. The tendency with diverse radar waveform design is that more degrees of freedom results in better outcomes. Here, the G and CU distributions are far less restricted than the DU2 distribution since they are continuous, but it would appear their freedom is not useful. Rather, it is unwieldy. While the DU2 distribution can only take on two values the optimization then only has to account for those two outcomes rather than a continuum of them.

Consider the G distribution. however unlikely its possible for any one of the random variable to be extremely large and expand the spectrum of given waveform greatly. For the CU case or the G case, the optimization has to consider what happens when several of the parameters happen to be zero leading to a very small tight for that waveform. The DU2 distribution only has to consider two outcomes per random variable and at least for the cases considered here, that appears to be very advantageous.

4.5 Experimental Results

The goal of this Section is to verify the simulated results of Section 4.4.2 in both loop-back and open-air testing. This will consist of selecting several cases from Section 4.4.2 which were deemed

Gaussian ($\mathbf{u} = \mathbf{G4}$)

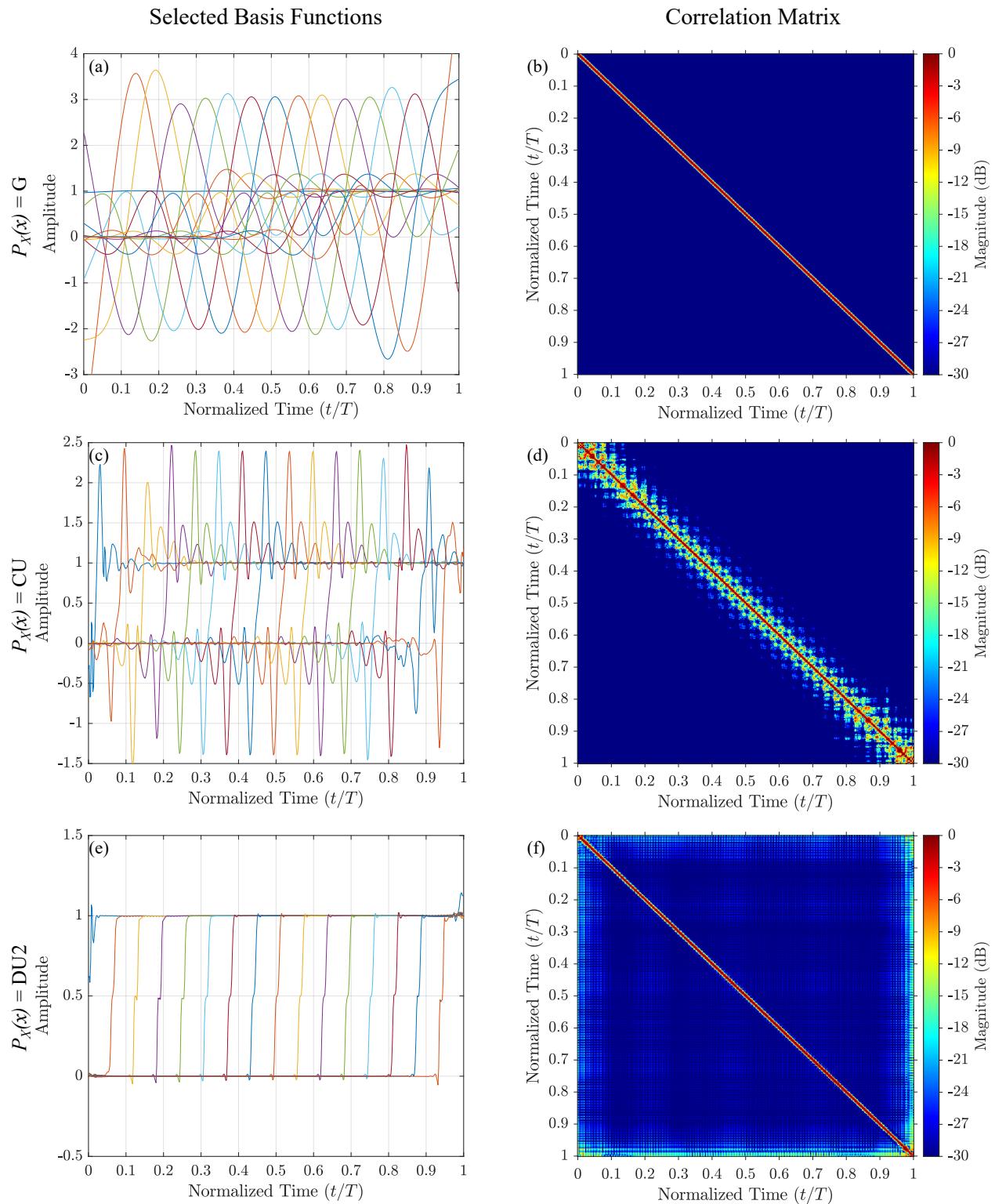


Figure 4.14: Correlation matrices and selected basis functions for each distribution and the G4 template. Correlation matrices are plotted in magnitude on a dB scale. (a) selected G basis functions (b) G correlation matrix (c) selected CU basis functions (d) CU correlation matrix (e) selected DU2 basis functions (f) DU2 correlation matrix

Super-Gaussian ($\mathbf{u} = \mathbf{S4G4}$)

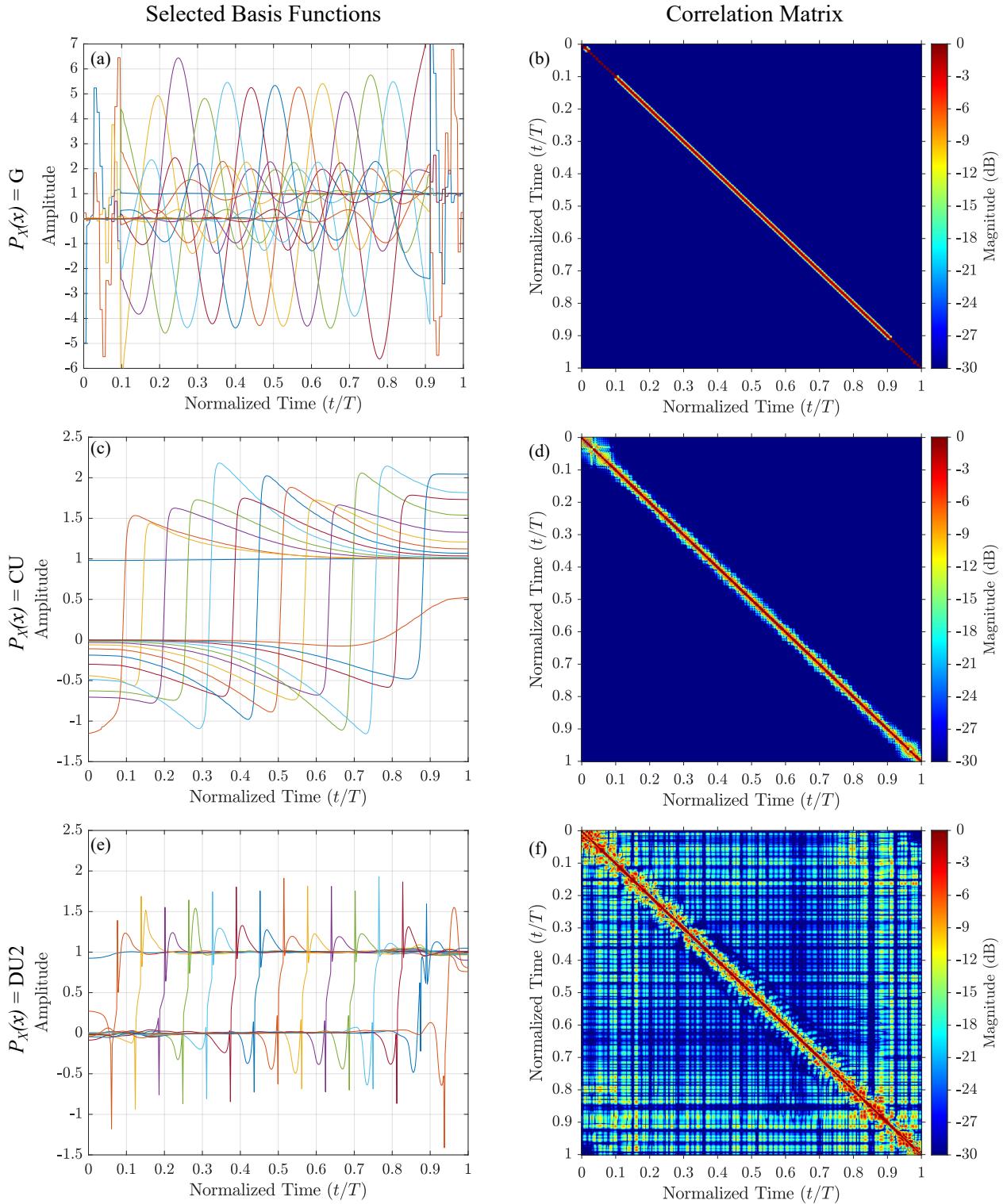


Figure 4.15: Correlation matrices and selected basis functions for each distribution and the S4G4 template. Correlation matrices are plotted in magnitude on a dB scale. (a) selected G basis functions (b) G correlation matrix (c) selected CU basis functions (d) CU correlation matrix (e) selected DU2 basis functions (f) DU2 correlation matrix

Rectangular ($\mathbf{u} = \mathbf{R4}$)

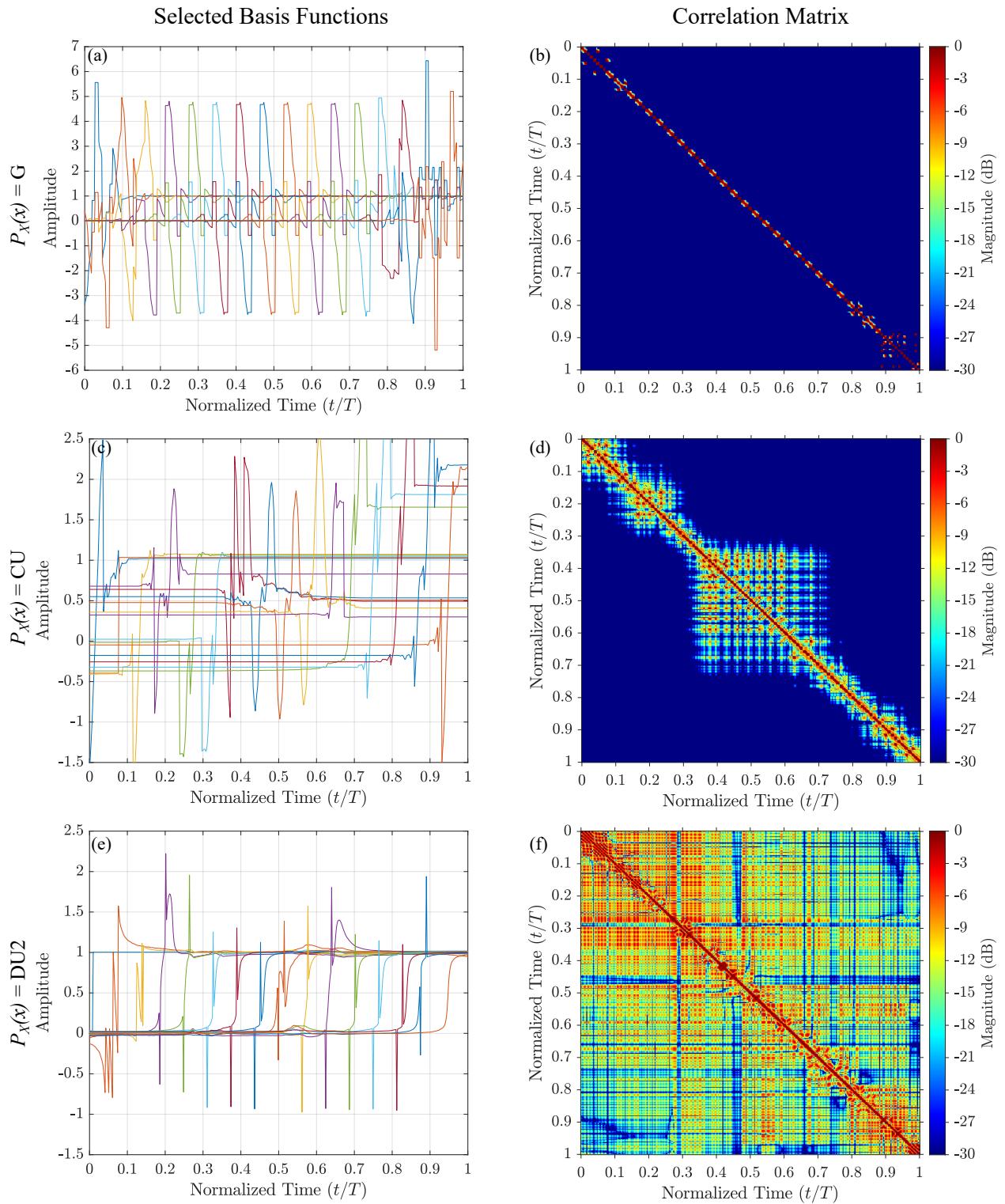


Figure 4.16: Correlation matrices and selected basis functions for each distribution and the R4 template. Correlation matrices are plotted in magnitude on a dB scale. (a) selected G basis functions (b) G correlation matrix (c) selected CU basis functions (d) CU correlation matrix (e) selected DU2 basis functions (f) DU2 correlation matrix

to be useful. For the G distribution this means implementing the G2 and G4 template results only since the G distribution was unable to achieve acceptable matches to the other spectral templates. For the CU distribution this means implementing the G2, G4, S4G2, and S4G4 for the same reasons. Finally, the all the DU2 cases will be implemented as it was able to acceptably match to every template.

4.5.1 Loop-back Testing

4.5.2 Open-Air Testing

4.6 Conclusions

Chapter 5

CW Stochastic Waveform Generation (CW-StoWGe)

As discussed in 2.3, CW modes maximize the BT for a given bandwidth. Since the sidelobe performance of noise waveforms and specifically the FM noise waveforms considered in this work is directly proportional to the BT it is logical to consider FM noise waveforms for a CW mode.

This chapter begins by introducing the CW-StoWGe model in Section 5.1 and then discusses the EFTE cost function for CW-StoWGe in Section 5.2. In Section 5.3, the means to optimize the CW-StoWGe power spectrum are discussed, Section 5.4 presents the optimization results, and finally the CW-StoWGe waveforms will be shown to be robust to both loop-back and open-air implementations experimentally in Section 5.5.

5.1 The CW-StoWGe Signal Model

For pulsed StoWGe, the basis matrix \mathbf{B} and the mean phase value vector $\boldsymbol{\mu}$ described the phase of the entire signal and while the signal could be arbitrarily long, it was always finite. In a CW mode however, the infinitely long signal would likewise require an infinitely large \mathbf{B} and $\boldsymbol{\mu}$. For this reason, to extend StoWGe to an infinitely length in time it is necessary to fundamentally alter the structure such that there are only a finite number of parameters to optimize. To do so, consider the CPM signal model defined as [57, 58, 90]

$$s(t; \mathbf{I}) = \exp \left(j2\pi \sum_{k=-\infty}^m I_k h_k q(t - kT_s) \right), \quad mT \leq t \leq (m+1)T \quad (5.1)$$

where \mathbf{I} is an infinite length vector of information carrying symbols, T_s is the symbol time, h_k is a scalar known as the modulation index which can change with every symbol, and $q(t)$ is the phase response defined as

$$q(t) = \int_0^t g(\tau) d\tau \quad (5.2)$$

where $g(t)$ is the *frequency shaping filter* or *frequency pulse* which is required to be time limited. For communications, CPM has been found to be extremely useful for its constant amplitude nature making it energy efficient and its spectral efficiency in terms of bits/hertz. CW-StoWGe likewise takes advantage of CPM's energy and spectral efficiency. From an optimization stand point, the finite length frequency shaping filter makes this possible.

Consider (5.1), the phase is an infinite superposition of time shifted versions of a single basis function while the pulsed StoWGe phase is a finite superposition of numerous basis functions. The fact that the frequency pulse is time limited means it can be modeled by a finite set of parameters which themselves can be optimized to achieve desired characteristics. Still, to put (5.1) completely into the StoWGe framework it is necessary to modify it to a degree such that the CW-StoWGe waveform model is

$$s(t; \mathbf{x}) = \exp \left(j \sum_{k=-\infty}^m X_k q(t - kT_s) \right), \quad mT \leq t \leq (m+1)T \quad (5.3)$$

where h_k has been removed entirely and \mathbf{I} has been replaced with \mathbf{x} to emphasize the random variables no longer carry information and can take on any zero-mean distribution. For communications purposes, the information symbols would almost certainly be distributed on a discrete uniform basis. Still, even with these changes the forms of (5.1) and (5.3) are very similar, but for (5.1) the design paradigm calls for a good communications signal while for (5.3) the design goal is for good radar waveforms that posses low autocorrelation sidelobes and good spectral containment. Chapter 6 will consider utilizing the CW-StoWGe optimization for communications purposes in addition to radar.

In chapter 4, the StoWGe signal model was introduced in discrete terms to facilitate its design

and optimization on a computer. Likewise here, the CW-StoWGe model can be discretized as

$$s[m] = \exp \left(j \left(\sum_{n=-\infty}^k X_n q[m - nT_s] \right) \right) \quad kT_s \leq m < (k+1)T_s \quad (5.4)$$

where $q(t)$ becomes

$$q[m] = \sum_{k=0}^m g[k]. \quad (5.5)$$

In (5.4), T_s is now a positive integer such that a new random variable contributes to the signal phase every T_s samples. The frequency pulse, $g[k]$, is defined to be non-zero for $0 \leq n < LT_s$ samples where L is an integer. In CPM, $L = 1$ one corresponds to *full-response* CPM [57] while $L > 1$ corresponds to *partial-response* CPM [58]. For CPM, the value of L has a significant impact on how easy or difficult it is to demodulate a signal, but it is useful for shaping the signal's power spectrum. For radar, L is a means to gain more design freedom with little down side other than the computational complexity of optimizing more parameters. The structure of (5.4) is visualized in Fig. 5.1.

Figure 5.1: The CW-StoWGe phase structure at time m assuming m is a multiple of nT_s for $T_s = 4$ and $L = 3$.

As with pulsed StoWGe, to evaluate the usefulness of a given CW-StoWGe implementation according to the metrics in Section 3.4 it is necessary to evaluate its second and fourth order moments. The correlation between two samples of (5.4) can be expressed as the expectation of the product between the two samples at times m_1 and m_2 . For convenience, it is assumed $m_1 > m_2$

such that the moment is

$$\text{E}[s[m_1]s^*[m_2]] = \text{E} \left[\exp \left(j \sum_{n_1=k_2+1}^{k_1} X_{n_1} q[m_1 - n_1 T_s] + j \sum_{n_2=m_2-L+1}^{k_2} X_{n_2} (q[m_1 - n_2 T_s] - q[m_2 - n_2 T_s]) \right) \right] \quad (5.6)$$

where the common components of the phase of $s[m_1]$ and $s^*[m_2]$ have been canceled. after a significant amount of manipulation and accounting for that fact that the CW-StoWGe process is cyclo-stationary, the correlation between samples of the CW-StoWGe process is

$$C[\ell] = \sum_{v=0}^{T_s-1} \left[\prod_{n_1=L}^{\lfloor \frac{\ell+v}{T_s} \rfloor + L-1} \Psi_{X_{n_1}} \left(\sum_{k=0}^{v+\ell+(L-1-n_1)T_s} g[k] \right) \prod_{n_2=0}^{L-1} \Psi_{X_{n_2}} \left(\sum_{k=v+1+(L-1-n_2)T_s}^{v+\ell+(L-1-n_2)T_s} g[k] \right) \right]. \quad (5.7)$$

where $\lfloor \cdot \rfloor$ is the floor operation and $\Psi_{X_{n_1}}$ and $\Psi_{X_{n_2}}$ are the characteristic functions of the n_1 th random variable and the n_2 th random variable respectively. The results of (5.7) can be organized into a $(2W - 1) \times 1$ correlation vector where W is arbitrary such that

$$\mathbf{c} = [C[-W+1] \ C[-W+2] \ \dots \ C[W-2] \ C[W-1]]^T \quad (5.8)$$

which in turn can be used to create a toeplitz $W \times W$ correlation matrix with $C[0]$ centered on the main diagonal such that

$$\mathbf{C} = \begin{bmatrix} C[0] & C[-1] & \dots & C[-W+1] \\ C[1] & C[0] & \dots & C[-W+2] \\ \vdots & \vdots & \ddots & \vdots \\ C[W-1] & C[W-2] & \dots & C[0] \end{bmatrix}. \quad (5.9)$$

5.2 The Expected Frequency Template Error (EFTE) for CW-StoWGe

On the surface, the EFTE for CW-StoWGe is very similar to that of the pulsed StoWGe case. In fact, without a deeper look it is exactly the same aside from a subscript. It is defined as

$$J_{\text{CW}} = \left\| \mathbb{E} \left[|\mathbf{s}_f|^2 \right] - \mathbf{u} \right\|_2^2. \quad (5.10)$$

where J_{CW} indicates the CW-StoWGe model. Consider the expected power spectrum of the pulsed StoWGe model where a single sample is defined as

$$\mathbb{E} \left[|s_{f,w}|^2 \right] = \mathbf{a}_w \bar{\mathbf{C}} \mathbf{a}_w^H \quad (5.11)$$

where \mathbf{a}_w is the w th row of the DFT matrix \mathbf{A} . For pulsed StoWGe, the expected power spectrum is defined using the zero-padded version of its correlation matrix, $\bar{\mathbf{C}}$. This is to account for its pulsed nature. Whereas for the CW case, the expected power spectrum is defined such that

$$\mathbb{E} \left[|s_{f,w}|^2 \right] = \mathbf{a}_w \mathbf{C} \mathbf{a}_w^H \quad (5.12)$$

where there is no zero-padding on the correlation matrix \mathbf{C} . Instead, it is by definition truncated since the full matrix would be infinite in size. This implication was discussed in Section 3.3.2 where the autocorrelation and the PSD of a CW signal were discussed. In that section it was shown that by estimating the autocorrelation or the PSD from a finite portion of a sample function, the resulting estimates are biased even if the process is ergodic. For the autocorrelation estimate this bias takes the form of a triangular window while the PSD is convolved with the Fourier transform of the triangular window, a sinc squared function as shown in (3.52).

This affects the EFTE cost function in meaningful ways. For the pulsed case, the EFTE measures the error between a desired template and the true expected power spectrum of the process. No bias is incurred since the sample functions of the process are inherently time limited. In the CW case, the EFTE measures the error between a desired template and a biased version of the

true expected power spectrum of the process. Consequently, minimizing the EFTE cost function for CW-StoWGe minimizes the squared error between a desired template and the expected power spectrum under the assumption the expected power spectrum is estimated via a DFT (periodogram) with W points. As will be shown, the value of W has a meaningful impact on the apparent roll-off and thus spectral containment of any CW-StoWGe generating function.

5.3 Optimization of the CW-StoWGe Expected Power Spectrum

Optimization of the CW-StoWGE expected power spectrum is very similar to that of the pulsed power spectrum. Since the same cost function is used, albeit with a different waveform model, many of the same conditions from the pulsed case apply. The cost function is still unconstrained, non-linear, and non-convex. However, in this case the shaping filter, \mathbf{g} , is optimized rather than either the basis function matrix \mathbf{B} or the mean phase value vector, $\boldsymbol{\mu}$. Since the cost function is differentiable with respect to the shaping filter, gradient descent methods will once again be utilized.

5.3.1 The CW-StoWGe EFTE Gradient

This section will briefly discuss the CW-StoWGe EFTE Gradient while a more detailed derivation will be found in Appendix A.2.2. The challenge right now is determining the best way to write out the indexing of (A.61). I was actually able to calculate this in MatLabTM to perform the optimization since the indexing was just a matter selecting certain values from matrices.

5.3.2 Gradient Descent Implementation

In the final version this section will include similar information to 4.3.2 such as the gradient descent variant used and stopping conditions.

5.3.3 Initializations

As in the pulsed case, there are numerous parameters that must be considered in optimizing the CW-StoWGe WGF. These parameters are listed in Table 5.1. In this case, the parameters T_s , L , $p_X(x)$, and \mathbf{u} must be initialized while the shaping filter function consists of the parameters to be optimized. The next section discuss each parameter, their selected initialization(s) and the rational

	T_s	Number of samples per waveform
Fixed Parameters	L	Number of random variables per waveform
	$p_X(x)$	Random Variable Distribution Function
	\mathbf{u}	The Desired expected spectrum
Optimizable Parameters	\mathbf{g}	Shaping Filter Function

Table 5.1: Waveform generating process parameters which must be selected or initialized prior to optimization

behind those selections.

Intervariable Spacing, T_s , and Partial Response, L : The intervariable spacing and the partial response parameter are discussed together since their product, LT_s , determines the length of \mathbf{g} and consequently the number of optimizable parameters. The intuition for the product LT_s is that it should be as large as possible since more degrees of freedom should result in more design freedom resulting in a better spectral match. However, it remains to be seen that for a given product of T_s and L is better to have more random variables tightly spaced in time but with long responses (small T_s , large L)? Alternatively, is it better to have fewer, more space out random variables with less overlap in time (large T_s , small L)? To examine these questions, various combinations of T_s and L will be considered, including an experiment where their product is kept constant, but their relative values are varied.

Random Variable Distribution Function $p_X(x)$: To maintain a consistent comparison to the pulsed results, the same set of random variable distribution functions will be examined as for the pulsed StoWGe model in Chapter 4.

Desired Spectrum \mathbf{u} : To maintain a consistent comparison to the pulsed results, the same set of spectral templates will be examined as for the pulsed StoWGe model in Chapter 4.

Initial Frequency Shaping Filter \mathbf{g}_0 : The choice of the initial shaping filter is somewhat arbitrary, but it is reasonable to use shaping filters which are already common to use in CPM implementations and variations thereof. With this in mind four initializations will be considered. They include a full rectangular function and a partial function which are defined as

$$\mathbf{g}_0 = \mathbf{g}_{\text{RECF}} = \mathbf{1}_{T_s L \times 1} \quad (5.13)$$

and

$$\mathbf{g}_0 = \mathbf{g}_{\text{PREC}} = [\mathbf{0}_{1 \times T_s \lfloor (L-1)/2 \rfloor} \quad \mathbf{1}_{1 \times T_s} \quad \mathbf{0}_{1 \times T_s \lceil (L-1)/2 \rceil}]^T \quad (5.14)$$

respectively, where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling operators respectively. The second two are based on the commonly used raised cosine shaping filter [90]. The full and partial initializations are defined

$$\mathbf{g}_0 = \mathbf{g}_{\text{RCF}} = (1 - \cos(2\pi[1/2 \ 3/2 \ \dots \ T_s L - 1/2]))/(T_s L) \quad (5.15)$$

and

$$\mathbf{g}_0 = \mathbf{g}_{\text{RCP}} = [\mathbf{0}_{1 \times T_s \lfloor (L-1)/2 \rfloor} \quad (1 - \cos(2\pi[1/2 \ 3/2 \ \dots \ T_s - 1/2]))/T_s \quad \mathbf{0}_{1 \times T_s \lceil (L-1)/2 \rceil}]^T \quad (5.16)$$

respectively where $\cos(\cdot)$ operates on an element wise basis.

5.4 Optimization Results

This section will perform a similar analysis to 4.4 where the results of Tables B.7 - B.12 are analyzed in detail

5.5 Experimental Results

This section will analyze the implementation of selected WGFs from Tables B.7 - B.12. They will be implemented in both loopback and open-air testing to evaluate their robustness to physical effects and usefulness as CW radar waveforms.

5.5.1 Loopback Testing

5.5.2 Open-Air Testing

5.6 Conclusions

Chapter 6

Conclusions & Future Work

Appendix A

StowGe Derivations

This appendix derives the expected power spectrum as defined in (3.60) in terms of the pulsed StoWGe waveform generating function of Chapter 4 and the CW StoWGe waveform generating function of Chapter 5. In doing so, the most important aspect of the process in either case is the calculation of the moments of the waveform generating process \mathbf{s} . If these can be calculated, the calculation of the expected power spectrum and the other metrics in Section 3.4 is straightforward extension of the moment calculation.

This appendix also derives the gradients with respect to the waveform generating process basis matrix \mathbf{B} and the phase mean value vector $\boldsymbol{\mu}$ for the pulsed StoWGe EFTE cost function. These gradients are defined as

$$\nabla_{\mathbf{B}} J = \begin{bmatrix} \frac{\partial J}{\partial b_{1,1}} & \frac{\partial J}{\partial b_{1,2}} & \cdots & \frac{\partial J}{\partial b_{1,N}} \\ \frac{\partial J}{\partial b_{2,1}} & \frac{\partial J}{\partial b_{2,2}} & \cdots & \frac{\partial J}{\partial b_{2,N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial b_{M,1}} & \frac{\partial J}{\partial b_{M,2}} & \cdots & \frac{\partial J}{\partial b_{M,N}} \end{bmatrix} \quad (\text{A.1})$$

and

$$\nabla_{\boldsymbol{\mu}} J = \begin{bmatrix} \frac{\partial J}{\partial \mu_1} \\ \frac{\partial J}{\partial \mu_2} \\ \vdots \\ \frac{\partial J}{\partial \mu_M} \end{bmatrix} \quad (\text{A.2})$$

respectively.

For the CW StoWGe-implementation, the gradient is evaluated with respect to the frequency

shaping filter vector g . This gradient is defined as

$$\nabla_g J = \begin{bmatrix} \frac{\partial J}{\partial g_1} \\ \frac{\partial J}{\partial g_2} \\ \vdots \\ \frac{\partial J}{\partial g_N} \end{bmatrix}. \quad (\text{A.3})$$

A.1 Pulsed StoWGe

A.1.1 The Pulsed StoWGe Expected Spectrum

As in (4.2) and (4.3) The pulsed StoWGe waveform generating process is defined as

$$\mathbf{s} = \exp(j\boldsymbol{\phi}) \quad (\text{A.4})$$

where

$$\boldsymbol{\phi} = \mathbf{B}\mathbf{x} + \boldsymbol{\mu} \quad (\text{A.5})$$

and \mathbf{B} is a $M \times N$ matrix of basis functions, \mathbf{x} is a $N \times 1$ vector of random variables, and $\boldsymbol{\mu}$ is a vector of constant values which are the expected value of the samples of $\boldsymbol{\phi}$. From 3.70, the expected spectrum can be written a function of the moments of \mathbf{s} such that

$$E\left[|s_{f,w}|^2\right] = \sum_{m_1=1}^W \sum_{m_2=1}^W a_{w,m_1} a_{w,m_2}^* E\left[\bar{s}_{m_1} \bar{s}_{m_2}^*\right] \quad (\text{A.6})$$

For pulsed StoWGe, the correlation between samples is

$$E\left[s_{m_1} s_{m_2}^*\right] = E\left[\exp(j((\mathbf{b}_{m_1} - \mathbf{b}_{m_2})\mathbf{x} + \boldsymbol{\mu}_{m_1} - \boldsymbol{\mu}_{m_2}))\right] \quad (\text{A.7})$$

such that the correlation is a function of the characteristic functions of the random variables in the vector \mathbf{x} as

$$\mathbb{E}[s_{m_1}s_{m_2}^*] = \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{n=1}^N \psi_{X_n}(b_{m_1,n} - b_{m_2,n}). \quad (\text{A.8})$$

Putting (A.6) and (A.8) together, a single sample of the expected spectrum of the pulsed StoWGe waveform generating function is

$$\mathbb{E}\left[|s_{f,w}|^2\right] = \sum_{m_1=1}^M \sum_{m_2=1}^M a_{w,m_1} a_{w,m_2}^* \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{n=1}^N \psi_{X_n}(b_{m_1,n} - b_{m_2,n}) \quad (\text{A.9})$$

In terms of the pulsed StoWGe second moment matrix this becomes

$$\mathbb{E}\left[|s_{f,w}|^2\right] = \mathbf{a}_w \bar{\mathbf{C}} \mathbf{a}_w^H \quad (\text{A.10})$$

where \mathbf{a}_w is the w th row of the $W \times W$ DFT matrix \mathbf{A} .

A.1.2 The Pulsed StoWGe EFTE Gradient

A.1.2.1 Basis Function Matrix Gradient

To calculate the gradient of the EFTE cost function, defined as

$$J = \left\| \mathbb{E}\left[|\mathbf{s}_f|^2\right] - \mathbf{u} \right\|_2^2, \quad (\text{A.11})$$

with respect to basis matrix \mathbf{B} , begin by evaluating the derivative of the cost function with respect to a single element of \mathbf{B} such that

$$\frac{\partial J}{\partial b_{k,n}} = \frac{\partial}{\partial b_{k,n}} \left\| \mathbb{E}\left[|\mathbf{s}_f|^2\right] - \mathbf{u} \right\|_2^2 \quad (\text{A.12})$$

where $b_{k,n}$ is the element of \mathbf{B} in the k th row and the n th column. According to the chain rule, (A.12) becomes

$$\frac{\partial J}{\partial b_{k,n}} = 2 \left(\frac{\partial \mathbb{E} [|\mathbf{s}_f|^2]}{\partial b_{k,n}} \right)^T \left(\mathbb{E} [|\mathbf{s}_f|^2] - \mathbf{u} \right) \quad (\text{A.13})$$

where the derivative is applied to $\mathbb{E} [|\mathbf{s}_f|^2]$ element wise. For the w th sample of the expected spectrum, the derivative is

$$\frac{\partial \mathbb{E} [|s_{f,w}|^2]}{\partial b_{k,n}} = \sum_{m_1=1}^M \sum_{m_2=1}^M a_{w,m_1} a_{w,m_2}^* \exp(j(\mu_{m_1} - \mu_{m_2})) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^P \Psi_{X_p}(b_{m_1,p} - b_{m_2,p}) \right). \quad (\text{A.14})$$

The only possibly non-zero terms of the derivative in (A.14) occur when exclusively either m_1 or m_2 are equal to k . Therefore, (A.14) can be rewritten as

$$\begin{aligned} \frac{\partial \mathbb{E} [|s_{f,w}|^2]}{\partial b_{k,n}} &= a_{w,k} \exp(j\mu_k) \sum_{\substack{m_2=1 \\ m_2 \neq k}}^M a_{w,m_2}^* \exp(-j\mu_{m_2}) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^P \Psi_{X_p}(b_{k,p} - b_{m_2,p}) \right) \\ &\quad + a_{w,k}^* \exp(-j\mu_k) \sum_{\substack{m_1=1 \\ m_1 \neq k}}^M a_{w,m_1} \exp(j\mu_{m_1}) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^P \Psi_{X_p}(b_{m_1,p} - b_{k,p}) \right) \end{aligned} \quad (\text{A.15})$$

The two summations are complex conjugates of each other such that

$$\frac{\partial \mathbb{E} [|s_{f,w}|^2]}{\partial b_{k,n}} = 2\Re \left\{ a_{w,k} \exp(j\mu_k) \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m}^* \exp(-j\mu_m) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^P \Psi_{X_p}(b_{k,p} - b_{m,p}) \right) \right\} \quad (\text{A.16})$$

where the subscript of the iterator m has been dropped since it is no longer necessary. Finally, there is exactly one term in the product where $b_{n,k}$ appears resulting in

$$\frac{\partial \mathbb{E} [|s_{f,w}|^2]}{\partial b_{k,n}} = 2\Re \left\{ \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} a_{w,k}^* \exp(j(\mu_m - \mu_k)) \frac{\partial \Psi_{X_p}(b_{m,n} - b_{k,n})}{\partial b_{k,n}} \prod_{\substack{p=1 \\ p \neq n}}^N \Psi_{X_p}(b_{m,p} - b_{k,p}) \right\} \quad (\text{A.17})$$

From (A.17) and (A.13), the entire structure of $\nabla_{\mathbf{B}} J$ can be calculated.

A.1.2.2 Mean Phase Value Vector Gradient

According to the chain rule, the derivative of the EFTE cost function with respect to a single element of the mean value vector μ is

$$\frac{\partial J}{\partial \mu_k} = 2 \left(\frac{\partial E[|s_f|^2]}{\partial \mu_k, n} \right)^T (E[|s_f|^2] - \mathbf{u}) \quad (\text{A.18})$$

where μ_k is the k th element of μ . The derivative of $E [|s_f|^2]$ can be evaluated in an element wise manner such that

$$\frac{\partial E[|s_{f,w}|^2]}{\partial \mu_k} = \sum_{m_1=1}^M \sum_{m_2=1}^M a_{w,m_1} a_{w,m_2}^* \frac{\partial}{\partial \mu_k} \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{p=1}^P \Psi_{X_p}(b_{m_1,p} - b_{m_2,p}). \quad (\text{A.19})$$

The non-zero terms of (A.19) occur when $k = m_1$ or $k = m_2$ but not both. Taking this into account, one of the sums can be removed yielding

$$\begin{aligned} \frac{\partial E[|s_{f,w}|^2]}{\partial \mu_k} = & ja_{w,k} \exp(j\mu_k) \sum_{\substack{m_2=1 \\ m_2 \neq k}}^M a_{w,m_2}^* \exp(-j\mu_{m_2}) \prod_{p=1}^P \Psi_{X_p}(b_{k,p} - b_{m_2,p}) \\ & - ja_{w,k}^* \exp(-j\mu_k) \sum_{\substack{m_1=1 \\ m_1 \neq k}}^M a_{w,m_1} \exp(j\mu_{m_1}) \prod_{p=1}^P \Psi_{X_p}(b_{m_1,p} - b_{k,p}) \end{aligned} \quad (\text{A.20})$$

Since this is a difference of conjugates, (A.20) can be written as the imaginary part of either term such that

$$\frac{\partial E[|s_{f,w}|^2]}{\partial \mu_k} = 2\Im \left\{ a_{w,k}^* \exp(-j\mu_k) \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} \exp(j\mu_m) \prod_{p=1}^P \Psi_{X_p}(b_{k,p} - b_{m,p}) \right\} \quad (\text{A.21})$$

where \Im extracts the imaginary part. From (A.21) and (A.18) the gradient defined by (A.2) can be calculated.

A.1.3 The Pulsed StoWGE EFTE Gradient for Selected Distributions

Evaluating the gradient with respect to specific distributions is a matter of evaluating the characteristic function for the distribution, its derivative with respect to the relevant parameter, and inserting these results into the generalized gradients which is (A.17) and A.21 in this case.

A.1.3.1 Discrete uniform distribution with two states (DU2)

Multiplying a DU2 RV by $b_{m,n} - b_{k,n}$ shifts the position of the deltas in its PDF such that for the distribution used here

$$\text{DU2: } p_X(x) = \frac{1}{2}\delta(x + \pi(b_{m,n} - b_{k,n})) + \frac{1}{2}\delta(x - \pi(b_{m,n} - b_{k,n})) \quad (\text{A.22})$$

The characteristic function is then

$$\text{DU2: } \Psi_X(b_{m,n} - b_{k,n}) = \cos(\pi(b_{m,n} - b_{k,n})) \quad (\text{A.23})$$

and its derivative with respect to $b_{k,n}$ is

$$\text{DU2: } \frac{\partial \Psi_X(b_{m,n} - b_{k,n})}{\partial b_{k,n}} = \pi \sin(\pi(b_{m,n} - b_{k,n})) \quad (\text{A.24})$$

Inserting A.23 and A.24 into A.17 realizes

$$\frac{\partial E[|s_{f,w}|^2]}{\partial b_{k,n}} = 2\Re \left\{ \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} a_{w,k}^* \exp(j(\mu_m - \mu_k)) \pi \sin(\pi(b_{m,n} - b_{k,n})) \prod_{\substack{p=1 \\ p \neq n}}^N \cos(\pi(b_{m,n} - b_{k,n})) \right\} \quad (\text{A.25})$$

Inserting (A.28) into (A.21) yields

$$\frac{\partial E[|s_{f,w}|^2]}{\partial \mu_k} = 2\Im \left\{ a_{w,k}^* \exp(-j\mu_k) \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} \exp(j\mu_m) \prod_{p=1}^P \cos(\pi(b_{m,n} - b_{k,n})) \right\} \quad (\text{A.26})$$

A.1.3.2 Continuous uniform distribution (CU)

Multiplying a continuous uniform RV with a scalar stretches the distribution. For the $U[-\pi, \pi]$ distribution used here, this realizes

$$\text{CU: } p_X(x) = \frac{1}{2\pi(b_{m,n} - b_{k,n})} \begin{cases} 1 & -\pi(b_{m,n} - b_{k,n}) \leq x \leq \pi(b_{m,n} - b_{k,n}) \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.27})$$

The characteristic function is then

$$\text{CU: } \Psi_X(b_{m,n} - b_{k,n}) = \frac{\sin(\pi(b_{m,n} - b_{k,n}))}{\pi(b_{m,n} - b_{k,n})} \quad (\text{A.28})$$

The derivative of A.33 with respect to $b_{k,n}$ is then

$$\text{CU: } \frac{\partial \Psi_X(b_{m,n} - b_{k,n})}{\partial b_{k,n}} = \frac{-\pi^2(b_{m,n} - b_{k,n}) \cos(\pi(b_{m,n} - b_{k,n})) + \pi \sin(\pi(b_{m,n} - b_{k,n}))}{\pi^2(b_{m,n} - b_{k,n})^2} \quad (\text{A.29})$$

Inserting A.28 and A.29 into A.17 realizes

$$\frac{\partial E[|s_{f,w}|^2]}{\partial b_{k,n}} = 2\Re \left\{ \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} a_{w,k}^* \exp(j(\mu_m - \mu_k)) \times \left(\frac{\sin(\pi(b_{m,n} - b_{k,n}))}{\pi(b_{m,n} - b_{k,n})^2} - \frac{\cos(\pi(b_{m,n} - b_{k,n}))}{b_{m,n} - b_{k,n}} \right) \prod_{\substack{p=1 \\ p \neq n}}^N \frac{\sin(\pi(b_{m,n} - b_{k,n}))}{\pi(b_{m,n} - b_{k,n})} \right\} \quad (\text{A.30})$$

where the derivative has been rearranged to make it more concise. Inserting (A.28) into (A.21) yields

$$\frac{\partial E[|s_{f,w}|^2]}{\partial \mu_k} = 2\Im \left\{ a_{w,k}^* \exp(-j\mu_k) \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} \exp(j\mu_m) \prod_{p=1}^P \frac{\sin(\pi(b_{m,n} - b_{k,n}))}{\pi(b_{m,n} - b_{k,n})} \right\} \quad (\text{A.31})$$

A.1.3.3 Gaussian distribution (G)

The zero mean Gaussian distribution with variance $(b_{m,n} - b_{k,n})^2$ is defined

$$G: p_X(x) = \frac{1}{(b_{m,n} - b_{k,n})\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{b_{m,n} - b_{k,n}}\right)^2} \quad (A.32)$$

The characteristic function is realized by taking the Fourier transform of (A.32) yielding

$$G: \Psi_X(b_{m,n} - b_{k,n}) = e^{-\frac{1}{2}(b_{m,n} - b_{k,n})^2 x^2} \quad (A.33)$$

The derivative of A.33 with respect to $b_{k,n}$ is then

$$G: \frac{\partial \Psi_X(b_{m,n} - b_{k,n})}{\partial b_{k,n}} = (b_{m,n} - b_{k,n}) e^{-\frac{1}{2}(b_{m,n} - b_{k,n})^2 x^2} \quad (A.34)$$

Inserting A.33 and A.34 into A.17 yields

$$\frac{\partial E[|s_{f,w}|^2]}{\partial b_{k,n}} = 2\Re \left\{ \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} a_{w,k}^* \exp(j(\mu_m - \mu_k)) (b_{m,n} - b_{k,n}) \prod_{p=1}^N \exp(-1/2(b_{m,n} - b_{k,n})^2 x_p^2) \right\} \quad (A.35)$$

where the $p \neq n$ has been removed from the product operator since it subsumed the exponential component of the derivative term. The derivative with respect to mean value vector becomes

$$\frac{\partial E[|s_{f,w}|^2]}{\partial \mu_k} = 2\Im \left\{ a_{w,k}^* \exp(-j\mu_k) \sum_{\substack{m=1 \\ m \neq k}}^M a_{w,m} \exp(j\mu_m) \prod_{p=1}^P \exp(-1/2(b_{m,n} - b_{k,n})^2 x_p^2) \right\} \quad (A.36)$$

A.2 CW-StoWGe

A.2.1 The CW-StoWGe Expected Spectrum

The discrete CW StoWGe waveform generating process is defined as

$$s[m] = \exp \left(j \left(\sum_{n=-\infty}^k X_k q[m - nT_s] \right) \right) \quad kT_s \leq m < (k+1)T_s \quad (\text{A.37})$$

where positive integer T_s is the inter-variable spacing such that a new X_k contributes to the phase content of (A.37) every T_s samples. The phase function $q[m]$ is the cumulative sum of the frequency shaping filter or frequency pulse defined as

$$q[m] = \sum_{k=0}^m g[k]. \quad (\text{A.38})$$

and the frequency shaping filter is time limited such that

$$g[m] = \begin{cases} g[m] & n = 0, 1, \dots, LT_s \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.39})$$

where positive integer L is the response length of the frequency shaping filter such that each random variable contributes to the frequency content of (A.37) for LT_s samples. The non-zero elements of $g[k]$ can be collected into the $(LT_s) \times 1$ vector \mathbf{g} . To begin evaluating the expected spectrum, the product of different samples of (A.37) is

$$s[m_1]s^*[m_2] = \exp \left(j \left(\sum_{n_1=-\infty}^{k_1} X_{n_1} q[m_1 - n_1 T_s] - \sum_{n_2=-\infty}^{k_2} X_{n_2} q[m_2 - n_2 T_s] \right) \right) \quad (\text{A.40})$$

then by assuming $m_1 > m_2$, the common components of the infinite sums can be canceled such that

$$s[m_1]s^*[m_2] = \exp\left(j\left(\sum_{n_1=k_2+1}^{k_1} X_{n_1} q[m_1 - n_1 T_s]\right) + j\left(\sum_{n_2=k_2-L+1}^{k_2} X_{n_2} (q[m_1 - n_2 T_s] - q[m_2 - n_2 T_s])\right)\right). \quad (\text{A.41})$$

Since these sums are the argument of a complex exponential, (A.41) can be written as a product of complex exponentials yielding

$$s[m_1]s^*[m_2] = \prod_{n_1=v_2+1}^{k_1} \exp(jX_{n_1} q[m_1 - n_1 T_s]) \times \prod_{n_2=k_2-L+1}^{k_2} \exp(jX_{n_2} (q[m_1 - n_2 T_s] - q[m_2 - n_2 T_s])) \quad (\text{A.42})$$

The correlation between these two samples is the expectation between them such that

$$\mathbb{E}[s[m_1]s^*[m_2]] = \mathbb{E}\left[\prod_{n_1=v_2+1}^{k_1} \exp(jX_{n_1} q[m_1 - n_1 T_s]) \times \prod_{n_2=k_2-L+1}^{k_2} \exp(jX_{n_2} (q[m_1 - n_2 T_s] - q[m_2 - n_2 T_s]))\right] \quad (\text{A.43})$$

Since each random variable is independent, the expectation of the products can be written as the product of the expectations such that

$$\mathbb{E}[s[m_1]s^*[m_2]] = \prod_{n_1=v_2+1}^{k_1} \mathbb{E}[\exp(jX_{n_1} q[m_1 - n_1 T_s])] \times \prod_{n_2=k_2-L+1}^{k_2} \mathbb{E}[\exp(jX_{n_2} (q[m_1 - n_2 T_s] - q[m_2 - n_2 T_s]))] \quad (\text{A.44})$$

In this form, the terms in the products are exactly the characteristic functions of their respective random variables where

$$\omega = q[m_1 - k_1 T_s] \quad (\text{A.45})$$

or

$$\omega = q[m_1 - k_1 T_s] - q[m_2 - k_2 T_s] \quad (\text{A.46})$$

given the characteristic function is defined

$$\Psi_X(\omega) = \mathbb{E}[\exp(j\omega X)] \quad (\text{A.47})$$

as was discussed in Section 3.1.2. Consequently, (A.44) can be rewritten in terms of characteristic functions such that

$$\begin{aligned} \mathbb{E}[s[m_1]s^*[m_2]] &= \prod_{n_1=k_2+1}^{k_1} \Psi_{X_{n_1}}(q[m_1 - n_1 T_s]) \times \\ &\quad \prod_{n_2=k_2-L+1}^{k_2} \Psi_{X_{n_2}}(q[m_1 - n_2 T_s] - q[m_2 - n_2 T_s]) \end{aligned} \quad (\text{A.48})$$

Unfortunately, the form of (A.48) is problematic in that it is not stationary. Consider the times (A.48), but increment each time index by 1 to realize

$$\begin{aligned} \mathbb{E}[s[m_1 + 1]s^*[m_2 + 1]] &= \prod_{n_1=k_2+1}^{k_1} \Psi_{X_{n_1}}(q[m_1 - n_1 T_s + 1]) \times \\ &\quad \prod_{n_2=k_2-L+1}^{k_2} \Psi_{X_{n_2}}(q[m_1 - n_2 T_s + 1] - q[m_2 - n_2 T_s + 1]) \end{aligned} \quad (\text{A.49})$$

Since there is no guarantee that the incremented $q[m + 1]$ equals $q[m]$ (otherwise it would be constant), equations (A.48) and (A.49) are not equal indicating the CW-StoWGe waveform generating process is not stationary. From the discussion in Section 3.3.1 this means the expected autocorrelation will differ depending on the reference time and consequently so will the expected power spectrum. However, CW-StoWGe is cyclo-stationary in that the autocorrelation is independent of certain time shifts. In the case of (A.48), incrementing the time indeces by multiples of T_s results in commensurate changes of the values of k_1 and k_2 such that

$$\mathbb{E}[s[m_1]s^*[m_2]] = \mathbb{E}[s[m_1 + nT_s]s^*[m_2 + nT_s]] \quad (\text{A.50})$$

where n is an integer. Averaging over the interval T_s realizes a stationary correlation function defined as

$$C[\ell] = E[s[m+\ell]s^*[m]] = \sum_{v=0}^{T_s-1} \left[\prod_{n_1=k_2+v}^{k_1} \Psi_{X_{n_1}}(q[m+v+\ell-n_1T_s]) \times \prod_{n_2=k_2-L+1}^{k_2} \Psi_{X_{n_2}}(q[m+v+\ell-n_2T_s] - q[m+v-n_2T_s]) \right] \quad (\text{A.51})$$

where ℓ is a positive integer. However since the autocorrelation is conjugate symmetric $C[\ell] = C^*[-\ell]$. Averaging a cyclo-stationary process in this way is a common way to evaluate the expected spectrum and autocorrelation for CW signals. [90].

For convenience, m can be set to 0 such that values for k_1 and k_2 are also set in terms of ℓ and v . Given the definition given by (A.37), k_1 and k_2 become

$$k_1 = \left\lfloor \frac{\ell+v}{T_s} \right\rfloor + L - 1 \quad (\text{A.52})$$

and

$$k_2 = L - 1 \quad (\text{A.53})$$

respectively, such that (A.51) becomes

$$C[\ell] = \sum_{v=0}^{T_s-1} \left[\prod_{n_1=L}^{\lfloor \frac{\ell+v}{T_s} \rfloor + L - 1} \Psi_{X_{n_1}}(q[v+\ell-n_1T_s]) \times \prod_{n_2=0}^{L-1} \Psi_{X_{n_2}}(q[v+\ell-n_2T_s] - q[v-n_2T_s]) \right]. \quad (\text{A.54})$$

(A.54) is written in terms of the phase function $q[m]$ but the optimization performed in Chapter 5 is performed on the frequency shaping filter. $q[m]$ can be replace by $g[m]$ by leveraging (A.39).

However, the indexing becomes onerous. Nevertheless, in terms of $g[m]$, (A.54) becomes

$$C[\ell] = \sum_{v=0}^{T_s-1} \left[\prod_{n_1=L}^{\lfloor \frac{\ell+v}{T_s} \rfloor + L-1} \Psi_{X_{n_1}} \left(\sum_{k=0}^{v+\ell+(L-1-n_1)T_s} g[k] \right) \prod_{n_2=0}^{L-1} \Psi_{X_{n_2}} \left(\sum_{k=v+1+(L-1-n_2)T_s}^{v+\ell+(L-1-n_2)T_s} g[k] \right) \right]. \quad (\text{A.55})$$

where the upper bounds of the sums can include some of the zero valued portions of $g[m]$, but this is allowable since they will not effect the value of $C[\ell]$. According to the discussion in section 3.4.3, the expected spectrum can be calculated as a function of the correlation matrix. With (A.55) an arbitrarily large ($W \times W$) correlation matrix can be created as a toeplitz construction of all lags for $[C_{-W+1} \ C_{-W+2} \ \cdots \ C_{W-2} \ C_{W-1}]^T$ by placing the lag $\ell = 0$ along the main diagonal. The w th term in W length expected power spectrum is then calculated as

$$\mathbb{E} [|s_{f,w}|^2] = \mathbf{a}_w \mathbf{C} \mathbf{a}_w^H. \quad (\text{A.56})$$

The key distinction between the expected power spectrum here and the pulsed StoWGe case is the evaluation of \mathbf{C} and the lack of zero-padding. Zero-padding implies a pulsed signal which is clearly not applicable in the CW case.

A.2.2 The CW-StoWGe EFTE Gradient

The calculation of the CW-StoWGe EFTE gradient begins identically to the Pulsed StoWGe gradient except the derivative is taken with respect to the frequency shaping filter instead a basis function matrix. The EFTE cost function is defined as

$$J_{CW} = \left\| \mathbb{E} [|\mathbf{s}_f|^2] - \mathbf{u} \right\|_2^2. \quad (\text{A.57})$$

The derivative with respect to the y th element of $g[m]$ is then

$$\frac{\partial J_{CW}}{\partial g_y} = \frac{\partial}{\partial g_y} \left\| \mathbb{E} [|\mathbf{s}_f|^2] - \mathbf{u} \right\|_2^2. \quad (\text{A.58})$$

According to the chain rule, (A.58) becomes

$$\frac{\partial J_{CW}}{\partial g_y} = 2 \left(\frac{\partial E[|s_f|^2]}{\partial g_y} \right)^T (E[|s_f|^2] - \mathbf{u}) \quad (\text{A.59})$$

where the derivative is applied to $E[|s_f|^2]$ element wise. For the w th sample of the expected power spectrum, the derivative is

$$E[|s_{f,w}|^2] = \mathbf{a}_w \frac{\partial \mathbf{C}}{\partial g_y} \mathbf{a}_w^H \quad (\text{A.60})$$

where the partial derivative is evaluated with respect to each element of \mathbf{C} . Since \mathbf{C} is toeplitz, there are only W unique derivatives to evaluate corresponding to $[C_0 \ C_1 \ \dots \ C_{W-1}]^T$. The derivative of the correlation function at lag ℓ is

$$\frac{\partial C[\ell]}{\partial g_y} = \frac{\partial}{\partial g_y} \left(\sum_{v=0}^{T_s-1} \left[\prod_{n_1=L}^{\lfloor \frac{\ell+v}{T_s} \rfloor + L-1} \Psi_{X_{n_1}} \left(\sum_{k=0}^{v+\ell+(L-1-n_1)T_s} g[k] \right) \prod_{n_2=0}^{L-1} \Psi_{X_{n_2}} \left(\sum_{k=v+1+(L-1-n_2)T_s}^{v+\ell+(L-1-n_2)T_s} g[k] \right) \right] \right) \quad (\text{A.61})$$

In a mathematical sense, applying the partial derivative in (A.61) is straightforward in that it just requires repeated application of the product rule. The difficultly lies in proper indexing.

A.2.3 The CW-StoWGe EFTE Gradient for Selected Distributions

This section will evaluate the CW-StoWGe EFTE Gradient for the DU2, CU, and G distributions.

Appendix B

Tabulated Optimization Results

B.1 Pulsed StoWGe

		Gaussian ($K = 2$)					
		DU2		CU		Gaussian	
# of Random Variables (N)	$\mathbf{B}_0 :$	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}
	2	-154.1	-154.1	-153.9	-57.4	-73.9	-59.0
	4	-154.1	-154.1	-152.4	-129.4	-74.2	-74.2
	8	-154.1	-149.1	-129.5	-143.4	-74.2	-74.2
	16	-154.1	-153.6	-121.0	-153.8	-74.2	-74.2
	32	-154.1	-154.0 *	-114.3	-154.0	-74.2	-74.2
	64	-142.8	-117.0	-107.6	-153.3	-70.5	-62.6
	128	-154.1	-116.7	-100.8	-153.8	-74.1	-74.0
	256	-154.1	-117.1	-97.7	-153.3	-74.1	-73.5

Table B.1: Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)

B.2 CW-StoWGe

B.3 StoWGe Comms

Gaussian ($K = 4$)

$P_X(x) :$	DU2		CU		Gaussian		
$\mathbf{B}_0 :$	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}	
# of Random Variables (N)	2	-135.6	-179.2	-64.5	-41.4	-56.0	-48.6
	4	-173.9	-169.7	-102.8	-55.7	-71.2	-60.2
	8	-158.6	-156.0	-115.5	-56.8	-76.4	-68.5
	16	-163.1	-139.5 *	-112.9	-60.0	-88.2	-65.7
	32	-138.0	000	-115.1	-59.2	-88.1	-60.1
	64	-170.6	-130.0	-114.6	-57.2	-88.2	-54.3
	128	-200.0	-129.8	-113.2	-84.4	-84.6	-50.9
	256	-188.2	-129.4	-110.6	-90.9	-88.5	-88.0

Table B.2: Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)

Super-Gaussian ($n = 4, K = 2$)

$P_X(x) :$	DU2		CU		Gaussian		
$\mathbf{B}_0 :$	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}	
# of Random Variables (N)	2	-200.1	-148.0	-61.5	-42.9	-45.7	-43.4
	4	-200.0	-182.1	-84.2	-46.2	-45.9	-45.9
	8	-151.9	-150.0	-90.7	-50.9	-45.9	-45.9
	16	-155.9	-137.0	-93.2	-52.1	-45.9	-45.7
	32	-157.3	-132.4	-89.6	-58.3	-45.4	-45.3
	64	-153.2	-129.9	-87.2	-57.0	-46.2	-44.1
	128	-154.7	-129.8	-84.9	-56.2	-46.2	-45.6
	256	-142.9	-129.0	-59.3	-59.5	-46.2	-45.5

Table B.3: Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)

Super-Gaussian ($n = 4, K = 4$)

$P_X(x) :$	DU2		CU		Gaussian	
$\mathbf{B}_0 :$	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}
# of Random Variables (N)	2	-86.1	-85.9	-58.2	-39.8	-41.5
	4	-88.1	-88.4	-73.0	-50.1	-41.8
	8	-88.4	-88.4	-78.3	-52.1	-41.8
	16	-88.1	000	-77.6	-50.7	-42.0
	32	-88.1	-88.3	-73.1	-47.1	-42.0
	64	-88.8	-88.1	-76.2	-28.2	-42.0
	128	-88.8	-79.0	-54.5	-49.3	-42.0
	256	-76.3	-88.6	-61.4	-59.1	-42.0

Table B.4: Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)

Rectangular ($K = 2$)

$P_X(x) :$	DU2		CU		Gaussian	
$\mathbf{B}_0 :$	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}	\mathbf{B}_{PC}	\mathbf{B}_{Id}
# of Random Variables (N)	2	-56.9	-56.8	-42.4	-38.5	-37.2
	4	-57.0	-57.0	-43.1	-35.3	-37.2
	8	-57.1	-57.0	-43.4	-33.0	-37.3
	16	-57.0	-57.1	-43.3	-34.4	-37.2
	32	-56.9	-56.7	-42.9	-37.5	-37.1
	64	-57.0	-57.0	-42.8	-42.4	-36.9
	128	-57.1	-56.7	-40.5	-42.5	-36.8
	256	-57.0	-54.4	-39.5	-41.7	-36.8

Table B.5: Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 2 with respect to its absolute bandwidth ($K = 2$)

Rectangular ($K = 4$)

$P_X(x) :$	DU2		CU		Gaussian		
	$\mathbf{B}_0 :$	\mathbf{B}_{PCFM}	\mathbf{B}_{Id}	\mathbf{B}_{PCFM}	\mathbf{B}_{Id}	\mathbf{B}_{PCFM}	\mathbf{B}_{Id}
# of Random Variables (N)	2	-51.0	-51.0	-38.8	-33.3	-33.9	-33.6
	4	-51.1	-51.1	-40.0	-31.5	-33.9	-33.9
	8	-51.0	-51.1	-40.0	-31.6	-34.0	-33.9
	16	-51.0	-51.0	-39.8	-33.3	-34.0	-33.9
	32	-50.9	-51.1	-39.6	-38.9	-33.8	-33.7
	64	-51.1	-50.4	-38.1	-39.4	-33.7	-33.6
	128	-51.0	-49.2	-36.5	-38.7	-33.8	-33.5
	256	-48.1	-49.7	-37.2	-39.3	-33.7	-33.6

Table B.6: Pulsed StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 4 with respect to its absolute bandwidth ($K = 4$)

Gaussian ($K = 2$)

\mathbf{g}_0	L	$P_X(x)$:	DU2			CU			G			
			T_s :	2	4	8	2	4	8	2	4	
RECF	1		-45.3	-42.5	-47.9		-53.5	-54.0	-76.7	-54.8	-59.0	-62.4
	2		-48.9	-83.0	-71.4		-68.1	-75.2	-84.1	-61.6	-64.2	-65.9
	4		-89.2	-88.9	-98.1		-78.6	-82.1	-87.1	-64.9	-66.3	-67.1
	8		-89.4	-99.5	-104.3		-79.0	-82.5	-87.4	-66.4	-67.2	-67.6
RECP	1		-45.3	-42.5	-47.9		-53.5	-54.0	-76.7	-54.8	-59.0	-62.4
	2		-85.6	-63.7	-54.0		-74.0	-85.2	-91.6	-61.3	-64.2	-66.0
	4		-81.4	-89.4	-95.6		-79.0	-83.5	-87.9	-65.1	-66.4	-67.2
	8		-89.2	-99.5	-102.0		-79.0	-82.6	-87.2	-66.3	-67.2	-67.6
RCF	1		-45.3	-42.5	-47.9		-53.5	-54.0	-76.7	-54.8	-57.9	-61.8
	2		-48.9	-83.1	-71.4		-68.1	-75.3	-90.4	-62.7	-64.8	-66.1
	4		-89.2	-88.9	-98.1		-78.6	-82.1	-87.1	-65.4	-66.5	-67.2
	8		-89.4	-99.5	-104.3		-79.0	-82.5	-87.4	-66.7	-67.3	-67.6
RCP	1		-45.3	-42.5	-47.9		-53.5	-54.0	-76.7	-54.8	-57.9	-61.8
	2		-85.6	-63.7	-54.0		-74.0	-83.7	-89.6	-61.3	-63.4	-65.6
	4		-81.4	-89.4	-95.6		-79.0	-85.6	-92.3	-65.1	-65.9	-67.0
	8		-89.2	-99.5	-102.0		-79.0	-86.3	-92.7	-66.3	-67.0	-67.5

Table B.7: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)

Gaussian ($K = 4$)

		$P_X(x)$:			DU2			CU			G		
g_0	L	T_s :	2	4	8		2	4	8		2	4	8
RECF	1		-57.1	-49.6	-52.3		-54.1	-53.2	-54.1		-50.3	-56.3	-61.3
	2		-76.5	-73.5	-60.9		-85.3	-69.7	-70.7		-58.3	-64.8	-70.3
	4		-77.5	-88.6	-116.2		-79.0	-84.7	-99.1		-65.8	-71.8	-77.5
	8		-105.5	-108.0	-113.7		-95.9	-97.7	-115.8		-72.3	-78.2	-84.0
RECP	1		-57.1	-50.0	-52.4		-54.1	-86.9	-54.1		-50.3	-56.3	-61.3
	2		-96.2	-73.5	-60.2		-85.1	-69.7	-75.9		-57.2	-64.4	-70.2
	4		-79.7	-126.8	-98.1		-87.0	-92.9	-104.2		-64.7	-72.4	-78.2
	8		-80.1	-112.2	-116.0		-86.9	-98.3	-113.8		-67.6	-77.2	-84.0
RCF	1		-57.1	-49.6	-52.3		-54.1	-53.2	-54.1		-50.3	-56.4	-59.9
	2		-76.5	-55.3	-79.0		-85.3	-69.7	-70.7		-58.3	-66.6	-71.7
	4		-80.0	-108.5	-117.4		-87.0	-84.7	-101.9		-67.7	-73.4	-78.4
	8		-102.4	-112.2	-112.5		-94.4	-98.4	-107.7		-74.5	-79.4	-84.6
RCP	1		-57.1	-49.6	-52.4		-54.1	-53.2	-54.1		-50.3	-56.4	-59.9
	2		-96.2	-73.5	-52.4		-85.1	-69.7	-80.0		-57.2	-64.0	-68.6
	4		-79.7	-118.9	-113.0		-87.0	-95.6	-108.1		-64.7	-71.0	-75.9
	8		-80.1	-111.9	-112.3		-86.9	-94.7	-113.5		-67.6	-74.9	-81.8

Table B.8: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Gaussian spectrum which is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)

Super-Gaussian ($n = 4, K = 2$)

g_0	L	$P_X(x)$:	DU2			CU			G			
			T_s :	2	4	8	2	4	8	2	4	
RECF	1		-53.1	-41.3	-46.1		-54.1	-59.8	-62.5	-46.8	-48.4	-49.4
	2		-54.7	-70.8	-67.0		-59.9	-74.6	-75.9	-49.1	-49.8	-50.2
	4		-64.0	-90.2	-111.8		-61.0	-80.4	-93.1	-50.0	-50.3	-50.5
	8		-75.5	-79.2	-114.5		-61.0	-82.7	-94.3	-50.3	-50.5	-50.6
RECP	1		-53.1	-41.3	-46.1		-54.1	-59.8	-62.5	-46.8	-48.4	-49.4
	2		-67.4	-61.3	-75.2		-64.4	-76.9	-80.5	-49.0	-49.8	-50.2
	4		-66.9	-105.8	-111.0		-64.5	-81.9	-93.0	-50.0	-50.3	-50.5
	8		-80.5	-115.2	-112.4		-64.5	-82.7	-94.3	-50.3	-50.5	-50.6
RCF	1		-53.1	-41.3	-46.1		-54.1	-59.8	-62.5	-46.8	-48.1	-49.2
	2		-54.7	-57.3	-57.6		-59.9	-74.6	-75.9	-49.4	-50.0	-50.3
	4		-89.1	-81.6	-113.7		-61.0	-80.4	-93.1	-50.1	-50.4	-50.5
	8		-79.7	-109.0	-111.4		-61.0	-82.7	-94.3	-50.4	-50.5	-50.6
RCP	1		-53.1	-41.3	-46.1		-54.1	-59.8	-62.5	-46.8	-50.3	-49.2
	2		-67.4	-61.3	-60.2		-64.4	-76.9	-80.5	-49.0	-49.7	-50.2
	4		-66.9	-106.3	-115.4		-64.5	-81.9	-93.0	-50.0	-50.2	-50.5
	8		-80.5	-109.7	-116.3		-64.5	-82.7	-94.3	-50.3	-50.4	-50.6

Table B.9: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)

Super-Gaussian ($n = 4, K = 4$)

		$P_X(x)$:	DU2			CU			G		
g_0	L	T_s :	2	4	8	2	4	8	2	4	8
RECF	1	-44.1	-58.6	-48.8		-43.6	-55.8	-58.6	-42.8	-45.1	-46.1
	2	-49.9	-60.1	-61.1		-47.7	-59.0	-72.7	-45.3	-46.7	-47.2
	4	-50.8	-59.4	-86.8		-48.5	-59.8	-77.9	-46.8	-47.3	-47.5
	8	-50.8	-74.3	-76.8		-48.5	-59.8	-80.4	-47.3	-47.6	-47.7
RECP	1	-44.1	-58.6	-48.8		-43.6	-55.8	-58.6	-42.8	-45.1	-46.1
	2	-49.9	-60.3	-57.3		-47.7	-59.5	-72.7	-45.0	-46.7	-47.2
	4	-50.8	-64.9	-65.6		-48.5	-59.8	-78.7	-46.5	-47.3	-47.6
	8	-50.8	-71.1	-85.0		-48.5	-59.8	-80.4	-46.9	-47.5	-47.7
RCF	1	-44.1	-58.6	-48.8		-43.6	-55.8	-58.6	-42.8	-45.2	-45.9
	2	-49.9	-60.1	-56.0		-47.7	-59.0	-72.7	-45.2	-46.8	-47.3
	4	-50.8	-60.3	-75.1		-48.5	-59.8	-77.9	-46.9	-47.4	-47.6
	8	-50.8	-71.4	-82.8		-48.5	-59.8	-80.4	-47.4	-47.6	-47.7
RCP	1	-44.1	-58.6	-48.8		-43.6	-55.8	-58.6	-42.8	-45.2	-45.9
	2	-49.9	-60.3	-57.3		-47.7	-59.5	-72.7	-45.0	-46.7	-47.1
	4	-50.8	-64.9	-74.9		-48.5	-59.8	-78.7	-46.5	-47.3	-47.5
	8	-50.8	-71.1	-69.9		-48.5	-59.8	-80.4	-46.9	-47.5	-47.6

Table B.10: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired Super-Gaussian spectrum with a roll-off factor of 4 ($n = 4$) and is oversampled by a factor of 4 with respect to its 3 dB bandwidth ($K = 4$)

Rectangular ($K = 2$)

		$P_X(x)$:	DU2			CU			G		
g_0	L	T_s :	2	4	8	2	4	8	2	4	8
RECF	1		-47.5	-38.3	-42.5	-43.9	-45.7	-47.9	-41.3	-41.9	-42.2
	2		-50.7	-48.5	-48.0	-44.4	-46.4	-48.4	-42.1	-42.3	-42.4
	4		-52.3	-54.4	-52.7	-44.6	-46.4	-48.4	-42.4	-42.4	-42.5
	8		-43.1	-49.6	-49.6	-44.6	-46.4	-48.4	-42.5	-42.5	-42.5
RECP	1		-47.5	-38.3	-42.5	-43.9	-45.7	-47.9	-41.3	-41.9	-42.2
	2		-52.1	-51.0	-49.6	-44.6	-46.3	-48.4	-42.1	-42.3	-42.4
	4		-52.4	-52.5	-55.8	-44.6	-46.4	-48.4	-42.4	-42.4	-42.5
	8		-52.9	-53.0	-52.9	-44.7	-46.4	-48.4	-42.4	-42.5	-42.5
RCF	1		-47.5	-38.3	-42.5	-43.9	-45.7	-47.9	-41.3	-41.8	-42.1
	2		-50.7	-47.1	-51.6	-44.4	-46.4	-48.4	-42.2	-42.4	-42.4
	4		-52.3	-53.2	-50.5	-44.6	-46.4	-48.4	-42.4	-42.5	-42.5
	8		-52.9	-55.8	-51.3	-44.6	-46.4	-48.4	-42.5	-42.5	-42.5
RCP	1		-47.5	-38.3	-42.5	-43.9	-45.7	-47.9	-41.3	-41.8	-42.1
	2		-52.1	-51.0	-49.3	-44.6	-46.3	-48.4	-42.1	-42.3	-42.4
	4		-52.4	-51.0	-55.8	-44.6	-46.4	-48.4	-42.4	-42.4	-42.5
	8		-52.9	-53.6	-54.6	-44.7	-46.4	-48.4	-42.4	-42.5	-42.5

Table B.11: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 2 with respect to its absolute bandwidth ($K = 2$)

Rectangular ($K = 2$)

		$P_X(x)$:	DU2			CU			G		
g_0	L	T_s :	2	4	8	2	4	8	2	4	8
RECF	1		-38.3	-46.2	-41.2	-38.1	-41.3	-43.0	-37.8	-38.7	-39.1
	2		-40.0	-48.8	-47.2	-39.5	-41.5	-43.3	-38.8	-39.3	-39.4
	4		-40.3	-49.7	-52.3	-39.7	-41.7	-43.3	-39.3	-39.4	-39.5
	8		-40.3	-44.1	-52.7	-39.7	-41.7	-43.3	-39.4	-39.5	-39.5
RECP	1		-38.3	-46.2	-41.2	-38.1	-41.3	-43.0	-37.8	-38.7	-39.1
	2		-40.0	-49.0	-44.4	-39.5	-41.6	-43.3	-38.7	-39.2	-39.4
	4		-40.3	-49.6	-50.1	-39.7	-41.7	-43.3	-39.2	-39.4	-39.5
	8		-40.4	-50.1	-49.9	-39.7	-41.7	-43.3	-39.3	-39.5	-39.5
RCF	1		-38.3	-46.2	-41.2	-38.1	-41.3	-43.0	-37.8	-38.8	-39.0
	2		-40.0	-48.8	-48.2	-39.5	-41.5	-43.3	-38.7	-39.3	-39.4
	4		-40.3	-49.7	-50.2	-39.7	-41.7	-43.3	-39.3	-39.4	-39.5
	8		-40.3	-50.1	-52.8	-39.7	-41.7	-43.3	-39.4	-39.5	-39.5
RCP	1		-38.3	-46.2	-41.2	-38.1	-41.3	-43.0	-37.8	-38.8	-39.0
	2		-40.0	-49.0	-48.2	-39.5	-41.6	-43.3	-38.7	-39.2	-39.4
	4		-40.3	-49.6	-50.1	-39.7	-41.7	-43.3	-39.2	-39.4	-39.5
	8		-40.4	-50.1	-50.3	-39.7	-41.7	-43.3	-39.3	-39.5	-39.5

Table B.12: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum is oversampled by a factor of 4 with respect to its absolute bandwidth ($K = 4$)

Bibliography

- [1] T. K. Sarkar, M. Salazar Palma, and E. L. Mokole, “Echoing across the years: A history of early radar evolution,” *IEEE Microwave Magazine*, vol. 17, pp. 46–60, Oct 2016.
- [2] S. S. Swords, “Technical history of the beginnings of radar,” Institution of Electrical Engineers, 1986.
- [3] D. O. North, “An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems,” *Proceedings of the IEEE*, vol. 51, pp. 1016–1027, July 1963.
- [4] J. R. Klauder, A. Price, S. Darlington, and W. J. Albersheim, “The theory and design of chirp radars,” *Bell System Technical Journal*, vol. 39, no. 4, pp. 745–808, 1960.
- [5] G. Turin, “An introduction to matched filters,” *IRE Transactions on Information Theory*, vol. 6, pp. 311–329, June 1960.
- [6] S. D. Blunt and E. L. Mokole, “Overview of radar waveform diversity,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, pp. 2–42, November 2016.
- [7] N. Levanon and E. Mozeson, *Radar signals*. John Wiley & Sons, 2004.
- [8] M. C. Wicks, E. Mokole, S. D. Blunt, R. S. Schneible, and V. J. Amuso, *Principles of waveform diversity and design*. SciTech, 2010.
- [9] F. Gini, A. De Maio, and L. Patton, *Waveform design and diversity for advanced radar systems*. Institution of engineering and technology London, 2012.
- [10] U. Pillai, K. Y. Li, I. Selesnick, and B. Himed, *Waveform diversity: theory & applications*. McGraw-Hill, 2011.

- [11] J. Nocedal and S. J. Wright, *Nonlinear Equations*. Springer, 2006.
- [12] H. Griffiths, L. Cohen, S. Watts, E. Mokole, C. Baker, M. Wicks, and S. Blunt, “Radar spectrum engineering and management: Technical and regulatory issues,” *Proceedings of the IEEE*, vol. 103, pp. 85–102, Jan 2015.
- [13] S. D. Blunt and E. S. Perrins, *Radar and Communication Spectrum Sharing*. SciTech Publishing, 2018.
- [14] H. Griffiths, S. Blunt, L. Cohen, and L. Savy, “Challenge problems in spectrum engineering and waveform diversity,” in *2013 IEEE Radar Conference (RadarCon13)*, pp. 1–5, April 2013.
- [15] J. C. Pedro and N. B. Carvalho, “Intermodulation distortion in microwave and wireless circuits, artech house,” *Inc., Norwood MA*, 2003.
- [16] F. H. Raab, P. Asbeck, S. Cripps, P. B. Kenington, Z. B. Popovic, N. Pothecary, J. F. Sevic, and N. O. Sokal, “Power amplifiers and transmitters for rf and microwave,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, pp. 814–826, March 2002.
- [17] S. C. Cripps, “RF power amplifiers for wireless communications,” 2006.
- [18] M. Malanowski and K. Kulpa, “Detection of moving targets with continuous-wave noise radar: Theory and measurements,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 3502–3509, Sep. 2012.
- [19] S. R. J. Axelsson, “Random noise radar/sodar with ultrawideband waveforms,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 1099–1114, May 2007.
- [20] Xiaojian Xu and R. M. Narayanan, “Range sidelobe suppression technique for coherent ultra wide-band random noise radar imaging,” *IEEE Transactions on Antennas and Propagation*, vol. 49, pp. 1836–1842, Dec 2001.

- [21] B. M. Horton, “Noise-modulated distance measuring systems,” *Proceedings of the IRE*, vol. 47, pp. 821–828, May 1959.
- [22] J. Jakabosky, S. D. Blunt, and B. Himed, “Spectral-shape optimized FM noise radar for pulse agility,” in *2016 IEEE Radar Conference (RadarConf)*, pp. 1–6, May 2016.
- [23] C. A. Mohr, P. M. McCormick, S. D. Blunt, and C. Mott, “Spectrally-efficient FM noise radar waveforms optimized in the logarithmic domain,” in *2018 IEEE Radar Conference (RadarConf18)*, pp. 0839–0844, April 2018.
- [24] C. A. Mohr and S. D. Blunt, “FM noise waveforms optimized according to a temporal template error (TTE) metric,” in *2019 IEEE Radar Conference (RadarConf)*, pp. 1–6, April 2019.
- [25] J. Jakabosky, S. D. Blunt, and B. Himed, “Waveform design and receive processing for nonrecurrent nonlinear fmcw radar,” in *2015 IEEE Radar Conference (RadarCon)*, pp. 1376–1381, May 2015.
- [26] J. Jakabosky, S. D. Blunt, and B. Himed, “Spectral-shape optimized FM noise radar for pulse agility,” in *2016 IEEE Radar Conference (RadarConf)*, pp. 1–6, May 2016.
- [27] S. R. J. Axelsson, “Noise radar using random phase and frequency modulation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 2370–2384, Nov 2004.
- [28] L. Pralon, B. Pompeo, and J. M. Fortes, “Stochastic analysis of random frequency modulated waveforms for noise radar systems,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, pp. 1447–1461, April 2015.
- [29] J. L. Stewart, “The power spectrum of a carrier frequency modulated by gaussian noise,” *Proceedings of the IRE*, vol. 42, pp. 1539–1542, Oct 1954.
- [30] C. A. Mohr and S. D. Blunt, “Design and generation of stochastically defined, pulsed FM noise waveforms,” in *Intl. Radar Conf*, 2019.

- [31] M. A. Richards, J. Scheer, W. A. Holm, and W. L. Melvin, *Principles of modern radar*. Citeseer, 2010.
- [32] M. A. Richards, *Fundamentals of radar signal processing*. Tata McGraw-Hill Education, 2005.
- [33] P. M. Woodward, *Probability and Information Theory, with Applications to Radar: International Series of Monographs on Electronics and Instrumentation*, vol. 3. Elsevier, 1953.
- [34] P. Peebles, “Radar principles john wiley & sons,” *Inc., Canada*, 1998.
- [35] K. S. Shanmugan and A. M. Breipohl, *Random signals, detection, estimation and data analysis*. Wiley, 1988.
- [36] F. J. Harris, “On the use of windows for harmonic analysis with the discrete fourier transform,” *Proceedings of the IEEE*, vol. 66, pp. 51–83, Jan 1978.
- [37] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Prentice Hall Press, 2009.
- [38] E. Fowle, “The design of FM pulse compression signals,” *IEEE Transactions on Information Theory*, vol. 10, pp. 61–67, January 1964.
- [39] C. E. Cook, “A class of nonlinear FM pulse compression signals,” *Proceedings of the IEEE*, vol. 52, pp. 1369–1371, Nov 1964.
- [40] I. Gladkova, “Design of frequency modulated waveforms via the zak transform,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, pp. 355–359, Jan 2004.
- [41] S. D. Blunt and E. L. Mokole, “Overview of radar waveform diversity,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, pp. 2–42, November 2016.
- [42] R. Barker, “Group synchronization of binary digital systems,” *Communication Theory, Willis and Jackson (Butterworth's Scientific Publications, 1953)*, 1953.

- [43] J. Lindner, “Binary sequences up to length 40 with best possible autocorrelation function,” *Electronics Letters*, vol. 11, pp. 507–507, October 1975.
- [44] G. Coxson and J. Russo, “Efficient exhaustive search for optimal-peak-sidelobe binary codes,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, pp. 302–308, Jan 2005.
- [45] L. Bomer and M. Antweiler, “Polyphase barker sequences,” *Electronics Letters*, vol. 25, pp. 1577–1579, Nov 1989.
- [46] R. Frank, “Polyphase codes with good nonperiodic correlation properties,” *IEEE Transactions on Information Theory*, vol. 9, pp. 43–45, January 1963.
- [47] D. Chu, “Polyphase codes with good periodic correlation properties (corresp.),” *IEEE Transactions on Information Theory*, vol. 18, pp. 531–532, July 1972.
- [48] B. L. Lewis and F. F. Kretschmer, “A new class of polyphase pulse compression codes and techniques,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-17, pp. 364–372, May 1981.
- [49] B. L. Lewis and F. F. Kretschmer, “Linear frequency modulation derived polyphase pulse compression codes,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-18, pp. 637–641, Sep. 1982.
- [50] S. W. Golomb, “Two-valued sequences with perfect periodic autocorrelation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, pp. 383–386, April 1992.
- [51] J. W. Taylor and H. J. Blinchikoff, “Quadriphase code-a radar pulse compression signal with unique characteristics,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, pp. 156–170, March 1988.

- [52] H. H. Faust, B. Connolly, T. M. Firestone, R. C. Chen, B. H. Cantrell, and E. L. Mokole, “A spectrally clean transmitting system for solid-state phased-array radars,” in *Proceedings of the 2004 IEEE Radar Conference (IEEE Cat. No.04CH37509)*, pp. 140–144, April 2004.
- [53] R. Chen and B. Cantrell, “Highly bandlimited radar signals,” in *Proceedings of the 2002 IEEE Radar Conference (IEEE Cat. No.02CH37322)*, pp. 220–226, April 2002.
- [54] S. D. Blunt, M. Cook, J. Jakabosky, J. D. Graaf, and E. Perrins, “Polyphase-coded FM (PCFM) radar waveforms, part I: implementation,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, pp. 2218–2229, July 2014.
- [55] J. B. Anderson, T. Aulin, and C.-E. Sundberg, *Digital phase modulation*. Springer Science & Business Media, 1986.
- [56] S. Bluetooth, “Specifications of the bluetooth system, version 5.2.,” *IB, 3317 Specification*, vol. 5, 2019.
- [57] T. Aulin and C. Sundberg, “Continuous phase modulation - part I: Full response signaling,” *IEEE Transactions on Communications*, vol. 29, pp. 196–209, March 1981.
- [58] T. Aulin, N. Rydbeck, and C. . Sundberg, “Continuous phase modulation - part II: Partial response signaling,” *IEEE Transactions on Communications*, vol. 29, pp. 210–225, March 1981.
- [59] S. Blunt, M. Cook, E. Perrins, and J. de Graaf, “CPM-based radar waveforms for efficiently bandlimiting a transmitted spectrum,” in *2009 IEEE Radar Conference*, pp. 1–6, May 2009.
- [60] S. D. Blunt, J. Jakabosky, M. Cook, J. Stiles, S. Seguin, and E. L. Mokole, “Polyphase-coded fm (pcfm) radar waveforms, part ii: optimization,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, pp. 2230–2241, July 2014.

- [61] P. S. Tan, J. Jakabosky, J. M. Stiles, and S. D. Blunt, “Higher-order implementations of polyphase-coded fm radar waveforms,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, pp. 2850–2870, Dec 2019.
- [62] P. M. McCormick and S. D. Blunt, “Gradient-based coded-FM waveform design using legendre polynomials,” in *International Conference on Radar Systems (Radar 2017)*, pp. 1–6, Oct 2017.
- [63] J. Jakabosky, S. D. Blunt, and B. Himed, “Optimization of “over-coded” radar waveforms,” in *2014 IEEE Radar Conference*, pp. 1460–1465, May 2014.
- [64] P. M. McCormick and S. D. Blunt, “Nonlinear conjugate gradient optimization of polyphase-coded fm radar waveforms,” in *2017 IEEE Radar Conference (RadarConf)*, pp. 1675–1680, May 2017.
- [65] C. A. Mohr, P. M. McCormick, and S. D. Blunt, “Optimized complementary waveform subsets within an FM noise radar CPI,” in *2018 IEEE Radar Conference (RadarConf18)*, pp. 0687–0692, April 2018.
- [66] N. Abramson, “Bandwidth and spectra of phase-and-frequency-modulated waves,” *IEEE Transactions on Communications Systems*, vol. 11, pp. 407–414, December 1963.
- [67] J. Jakabosky, S. D. Blunt, and A. Martone, “Incorporating hopped spectral gaps into non-recurrent nonlinear FMCW radar emissions,” in *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 281–284, Dec 2015.
- [68] J. W. Owen, B. Ravenscroft, B. H. Kirk, S. D. Blunt, C. T. Allen, A. F. Martone, K. D. Sherbondy, and R. M. Narayanan, “Experimental demonstration of cognitive spectrum sensing notching for radar,” in *2018 IEEE Radar Conference (RadarConf18)*, pp. 0957–0962, April 2018.

- [69] J. Jakabosky, B. Ravenscroft, S. D. Blunt, and A. Martone, “Gapped spectrum shaping for tandem-hopped radar/communications cognitive sensing,” in *2016 IEEE Radar Conference (RadarConf)*, pp. 1–6, May 2016.
- [70] B. Ravenscroft, P. M. McCormick, S. D. Blunt, J. Jakabosky, and J. G. Metcalf, “Tandem-hopped OFDM communications in spectral gaps of FM noise radar,” in *2017 IEEE Radar Conference (RadarConf)*, pp. 1262–1267, May 2017.
- [71] B. Ravenscroft, P. M. McCormick, S. D. Blunt, E. Perrins, and J. G. Metcalf, “A power-efficient formulation of tandem-hopped radar communications,” in *2018 IEEE Radar Conference (RadarConf18)*, pp. 1061–1066, April 2018.
- [72] B. Ravenscroft, P. McCormick, S. D. Blunt, E. Perrins, C. Sahin, and J. Metcalf, “Experimental assessment of tandem-hopped radar and communications (THoRaCs),” in *2019 International Conference on Radar (RADAR)*, pp. 1–6, Sep 2019.
- [73] G. Zook, P. M. McCormick, S. D. Blunt, C. Allen, and J. Jakabosky, “Dual-polarized FM noise radar,” in *International Conference on Radar Systems (Radar 2017)*, pp. 1–5, Oct 2017.
- [74] J. Owen, S. D. Blunt, K. Gallagher, P. McCormick, C. Allen, and K. Sherbondy, “Nonlinear radar via intermodulation of FM noise waveform pairs,” in *2018 IEEE Radar Conference (RadarConf18)*, pp. 0951–0956, April 2018.
- [75] G. Zook, P. McCormick, and S. D. Blunt, “Fixational eye movement radar: Random spatial modulation,” in *2018 IEEE Radar Conference (RadarConf18)*, pp. 0827–0832, April 2018.
- [76] L. Pralon, G. Beltrao, B. Pompeo, M. Pralon, and J. M. Fortes, “Near-thumbtack ambiguity function of random frequency modulated signals,” in *2017 IEEE Radar Conference (RadarConf)*, pp. 0352–0355, May 2017.
- [77] M. Jankiraman, *FMCW Radar Design*. Artech House, 2018.
- [78] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

- [79] W. W. Hager and H. Zhang, “A survey of nonlinear conjugate gradient methods,” *Pacific journal of Optimization*, vol. 2, no. 1, pp. 35–58, 2006.
- [80] E. Ghadimi, H. R. Feyzmahdavian, and M. Johansson, “Global convergence of the heavy-ball method for convex optimization,” in *2015 European Control Conference (ECC)*, pp. 310–315, IEEE, 2015.
- [81] P. Stoica, R. L. Moses, *et al.*, “Spectral analysis of signals,” 2005.
- [82] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- [83] M. B. Priestley, *Spectral analysis and time series*, vol. 1. Academic press London, 1981.
- [84] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2005.
- [85] A. Papoulis, “Minimum-bias windows for high-resolution spectral estimates,” *IEEE Transactions on Information Theory*, vol. 19, pp. 9–12, January 1973.
- [86] C. A. Mohr, P. M. McCormick, C. A. Topliff, S. D. Blunt, and J. M. Baden, “Gradient-based optimization of PCFM waveforms,” *Submitted to IEEE Transactions on Aerospace and Electronic Systems*, 2020.
- [87] A. Parent, M. Morin, and P. Lavigne, “Propagation of super-gaussian field distributions,” *Optical and quantum electronics*, vol. 24, no. 9, pp. S1071–S1079, 1992.
- [88] M. Golay, “Complementary series,” *IRE Transactions on Information Theory*, vol. 7, pp. 82–87, April 1961.
- [89] C. Tseng and C. Liu, “Complementary sets of sequences,” *IEEE Transactions on Information Theory*, vol. 18, pp. 644–652, Sep. 1972.
- [90] J. G. Proakis and M. Salehi, *Digital communications*, vol. 4. McGraw-hill New York, 2001.