AFRL-RY-WP-TR-2021-0159

DESIGN AND EVALUATION OF STOCHASTIC PROCESSES AS PHYSICAL RADAR WAVEFORMS (Preprint)

Charles Mohr
Distributed Radio Frequency Sensing Branch
Multispectral Sensing & Detection Division

JUNE 2021
Final Report

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

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
June 2021	Thesis/Dissertation	10 June 2021 –10 June 2021

4. TITLE AND SUBTITLE
DESIGN AND EVALUATION OF STOCHASTIC PROCESSES AS PHYSICAL RADAR WAVEFORMS (Preprint)

5a. CONTRACT NUMBER
FA8650-14-D-1722-0002/N00014-16-C-2029

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
65502F/62204F

6. AUTHOR(S)
Charles Mohr

5d. PROJECT NUMBER
6095

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
Y1AC

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Sensors Directorate (AFRL/RYMS)
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

Office of Naval Research (ONR/ACO)
230 S. Dearborn Street
Chicago, IL 60604

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RYMS

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RY-WP-TR-2021-0159

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
PAO case number AFRL-2021-1799, Clearance Date 10 June 2021. 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. The U.S. Government is joint author of this work and has the right to use, modify, reproduce, release, perform, display, or disclose the work. 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 despite these advances, in a waveform agile mode where the radar transmits unique waveforms for every pulse or a nonrepeating signal continuously, effective operation can be difficult due the waveform design requirements. In general, for radar waveforms to be both useful and physically robust they must achieve good autocorrelation sidelobes, be spectrally contained, and possess a constant amplitude envelope for high power operation. Meeting these design goals represents a tremendous computational overhead that can easily impede real-time operation and the overall effectiveness of the radar. This work addresses this concern in the context of random FM waveforms (RFM) that 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 waveform agile mode. However, while they are effective, the approaches to design these waveforms require optimization of each individual waveform, making them subject to costly computational requirements.
continued on back of page

15. SUBJECT TERMS
nonconvex optimization, radar waveform design, stochastic waveforms

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	b. ABSTRACT	c. THIS PAGE			Charles Mohr
Unclassified	Unclassified	Unclassified	SAR	290	19b. TELEPHONE NUMBER *(Include Area Code)* N/A

14. ABSTRACT, continued

This dissertation takes a different approach. Since RFM waveforms are meant to be noise like in the first place, the waveforms here are instantiated as the sample functions of an underlying stochastic process called a waveform generating function (WGF). This approach enables the convenient generation of spectrally contained RFM wave-forms for little more computational cost than pulling numbers from a random number generator (RNG). To do so, this work translates the traditional mathematical treatment of random variables and random processes to a more radar centric perspective such that the WGFs can be analytically evaluated as a function of the usefulness of the radar waveforms that they produce via metrics such as the expected matched filter response and the expected power spectral density (PSD). Further, two WGF models denoted as pulsed stochastic waveform generation (Pulsed StoWGe) and continuous-wave stochastic waveform generation (CW-StoWGe) are devised as means to optimize WGFs to produce RFM waveform with good spectral containment and design flexibility between the degree of spectral containment and autocorrelation sidelobe levels for both pulsed and CW modes. This goal is achieved by leveraging gradient descent optimization methods to reduce the expected frequency template error (EFTE) cost function. The EFTE optimization is shown analytically using the metrics above, as well as others defined in this work and through simulation, to produce WGFs whose sample functions achieve these goals and thus produce useful random FM waveforms. To complete the theory-modeling-experimentation design life cycle, the resultant StoWGe waveforms are implemented in a loop-back configuration and are shown to be amenable to physical implementation.

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.

Committee members

Dr. Shannon Blunt, Chairperson

Dr. Chris Allen

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 despite these advances, in a waveform agile mode where the radar transmits unique waveforms for every pulse or a nonrepeating signal continuously, effective operation can be difficult due the waveform design requirements. In general, for radar waveforms to be both useful and physically robust they must achieve good autocorrelation sidelobes, be spectrally contained, and possess a constant amplitude envelope for high power operation. Meeting these design goals represents a tremendous computational overhead that can easily impede real-time operation and the overall effectiveness of the radar. This work addresses this concern in the context of random FM waveforms (RFM) that 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 waveform agile mode. However, while they are effective, the approaches to design these waveforms require optimization of each individual waveform, making them subject to costly computational requirements.

This dissertation takes a different approach. Since RFM waveforms are meant to be noise like in the first place, the waveforms here are instantiated as the sample functions of an underlying stochastic process called a waveform generating function (WGF). This approach enables the convenient generation of spectrally contained RFM waveforms for little more computational cost than pulling numbers from a random number generator (RNG). To do so, this work translates the traditional mathematical treatment of random variables and random processes to a more radar centric perspective

such that the WGFs can be analytically evaluated as a function of the usefulness of the radar waveforms that they produce via metrics such as the expected matched filter response and the expected power spectral density (PSD). Further, two WGF models denoted as pulsed stochastic waveform generation (Pulsed StoWGe) and continuous-wave stochastic waveform generation (CW-StoWGe) are devised as means to optimize WGFs to produce RFM waveform with good spectral containment and design flexibility between the degree of spectral containment and autocorrelation sidelobe levels for both pulsed and CW modes. This goal is achieved by leveraging gradient descent optimization methods to reduce the expected frequency template error (EFTE) cost function. The EFTE optimization is shown analytically using the metrics above, as well as others defined in this work and through simulation, to produce WGFs whose sample functions achieve these goals and thus produce useful random FM waveforms. To complete the theory-modeling-experimentation design life cycle, the resultant StoWGe waveforms are implemented in a loop-back configuration and are shown to be amenable 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

List of Figures

List of Tables

Chapter 1

Introduction

Over the years the generation of radar waveforms has changed greatly from the 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 [6–11]. 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, 5–7, 7–10].

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 systems operate in a bandwidth efficient manner for their own sake and for the sake of other users [12–15].

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

high power levels. High power operation necessitates high power amplifiers which operate in the saturation region of the amplifier gain curve. 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, 14, 16]. Though more sophisticated techniques exist such as predistortion [17, 18], 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 random frequency modulated (RFM) waveforms.

Random or noise-like waveforms in general take advantage of the high dimensionality of noise and noise like signals to reduce ambiguous responses such as autocorrelation sidelobes [6, 19–22]. 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 linear frequency modulation (LFM) based radar. While the noise is reduced, the autocorrelation sidelobes remain the same. With noise or noise-like radar signals, where unique pulses are transmitted at 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 as well, they too decrease when coherently combined. This diminishes the ambiguity due to the range sidelobes with more and more unique pulses.

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, random FM 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 as-

pect of random FM waveform design and constant amplitude waveform design in general is the spectral containment aspect. For random FM waveforms this often entails an iterative, computationally expensive design process to make each individual random FM waveform which have a desirable spectrum and good autocorrelation sidelobes [23–26]. Other random FM 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 since the random process is based on Gaussian noise itself [27–30].

With this in mind, the goal of this work is to combine the optimization based and the process based random FM radar implementations by defining a new class of random processes called *waveform generating functions* (WGFs) whose sample functions are themselves either pulsed or continuous wave (CW) radar waveforms. In this way, unique random FM radar waveforms can be created by simply pulling numbers from a random number generator (RNG) followed by some simple mathematical transformation. Compared to running a full-blown optimization on each and every waveform, this implementation is extremely efficient from a computational standpoint. However, the overall challenge in this approach is guaranteeing that the randomly generated waveforms are actually useful from a radar perspective. The waveforms need to be FM and they need to possess a desired spectrum with sufficient spectral containment while achieving a noise like reduction in autocorrelation sidelobes with coherent integration. The structure of this work reflects a solution to this challenge.

In an effort 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. To demonstrate the need and motivation for the design of radar waveforms, the scene is interrogated with progressively more sophisticated radar signals and post-processing. A particular emphasis is paid to the roll of the autocorrelation sidelobes in obscuring and hiding targets of interest within the scene. Subsequent sections provide a brief overview of current radar waveform models and design schemes especially as they relate to spectral and autocorrelation performance. The Chapter concludes by examining some miscellaneous topics which

are pertinent to this work.

Chapter 3 begins the process of framing noise and noise-like waveforms within the context random variables and random processes. To do so, the first few sections provide the fundamentals behind these topics. The next section goes into significant detail regarding the estimation of the properties of random processes with a focus on the properties of stationarity and ergodicity as well the autocorrelation function and the power spectral density (PSD). These estimation methods are at the heart of what it means to define a WGF. Accordingly in the next sections, they are subtly altered and redefined such that they describe not just the autocorrelation function of the process but the *expected matched filter response*, not just the PSD of the random process but the expected PSD of the WGF *from a radar perspective*. Such an analysis is performed for both pulsed and CW WGFs which are shown to behave in profoundly different manners from a stochastic processes perspective. These properties and others defined in Chapter 3 describe how the waveforms produced by the WGF can be *expected* to behave. In other words, these metrics can be used to evaluate the usefulness of the waveforms of a particular WGF. The Chapter concludes by translating these metrics into a discretized framework that is suitable evaluation and manipulation on a computer.

To put the tools of Chapter 3 to work, Chapters 4 and 5 define the pulsed stochastic waveform generation (Pulsed StoWGe) and the CW stochastic waveform generation (CW-StoWGe) models respectively. In either case, the models define a constant amplitude (FM) WGF whose phase is a linear combination of stochastically weighted basis functions. These models posses two primary advantages which make them suitable analysis, optimization, and physical implementation. First, they have have been designed in such a way that evaluating their moments and subsequently their characteristics as defined in Chapter 3 is mathematically tractable and computationally reasonable. In other words for a WGF based on either model, it is practical to analytically determine the usefulness of the radar waveforms that can be generated by the WGF. Second, the basis functions consist of a finite number of adjustable parameters. Then, by tweaking these parameters, the WGF can be tailored to produce radar waveforms with desirable characteristics. In order to take advantage of this flexibility, both Chapters 4 and 5 define the expected frequency template error

(EFTE) cost function which measures the mean squared error between the expected PSD of the WGF and some desired PSD for their respective models. Using a gradient descent optimization method [31], the EFTE cost functions are minimized for a multitude of WGFs corresponding to unique desired spectra and numerous permutations of the parameters of either the Pulsed StoWGe or the CW-StoWGe models. In either case, the resulting WGFs were analytically evaluated and used to produce the individual radar waveforms. In simulation, these waveforms demonstrate the veracity of the models and in experimental loopback measurements the waveforms are shown to amenable to physical implementation.

Finally, Chapter 6 provides overall conclusions regarding the modeling and use of WGFs as well as a discussion on possible future work. The appendices provide the derivation of numerous equations relating to Pulsed StoWGe, CW-StoWGe, and their respective optimizations as well as tabulated results of the optimizations performed in Chapters 4 and 5.

Chapter 2

Background

The purpose of this chapter is not to provide a comprehensive overview of radar principles. Rather, it is to provide a sufficient background to give context to the radar waveform design schemes and objectives of 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, 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 the presence of additive white Gaussian noise (AWGN). Then the linear frequency modulated (LFM) pulse is introduced as a superior pulse compression waveform as compared to the simple unmodulated pulse in that it leverages an expanded bandwidth to achieve a finer range resolution. Finally, pulse integration is shown to improve SNR and tapering is shown to mitigate, to a degree, the deleterious autocorrelation sidelobes inherent to the LFM waveform. Section 2.2 introduces several waveform design schemes and topologies for achieving such goals as sidelobe mitigation and spectral containment. This section provides an introduction to random FM 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

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

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) represents a *passband* signal as implied by the subscript 'pb'. Alternatively, it could be represented in *complex-baseband*. For this simple waveform, its complex-baseband representation is simply a real-valued, time limited, DC pulse such that

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

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

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

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

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

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

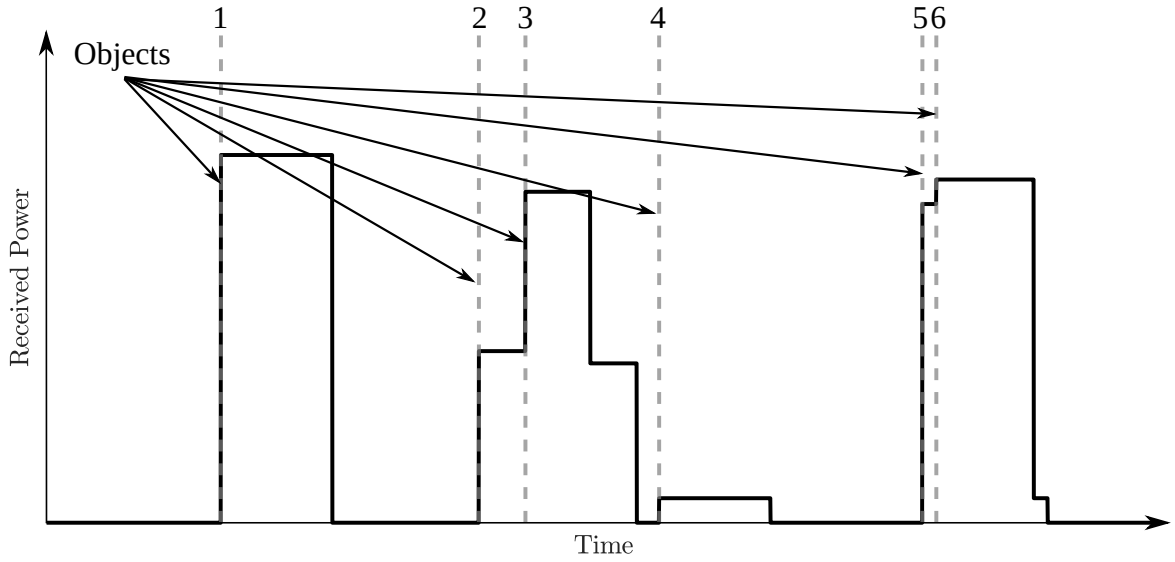

Figure 2.1: Ideal, simplistic radar data where the entire pulse envelopes are visible, undistorted, and are free of noise. The leading edges clearly indicate the positions of six distinct objects.

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. Consequently 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 [32–36] to name a few.

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

$$R = \frac{\tau c}{2} \tag{2.5}$$

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

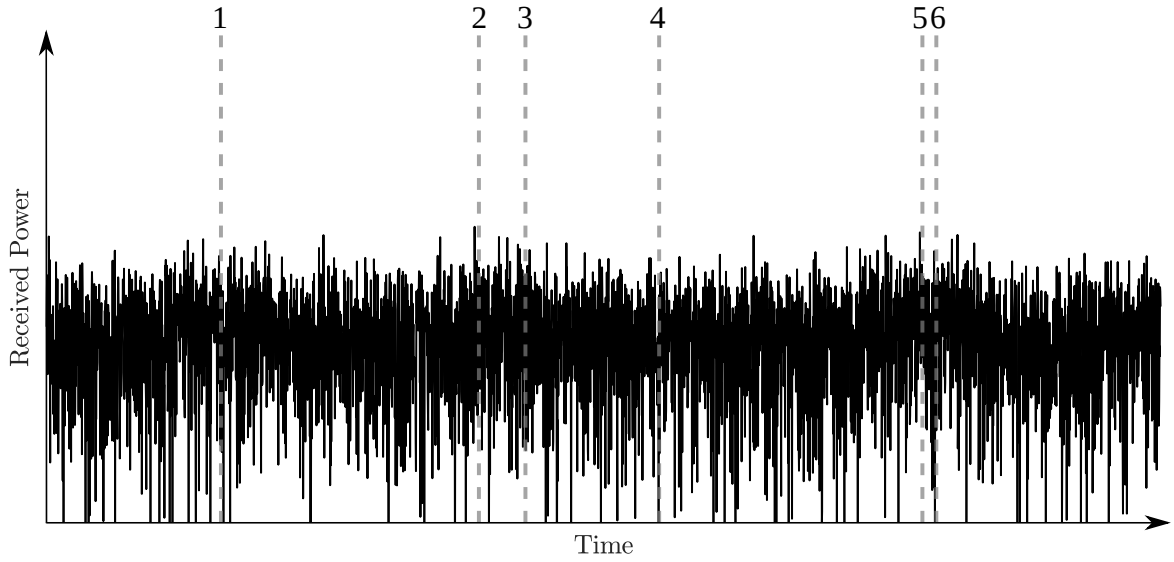

Figure 2.2: The same radar radar data as in Fig. 2.1, but with additive white Gaussian noise

impulse response of the environment such that

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

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

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

where r is range from the transceiver and a_n is some complex valued scaler 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

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

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 [32–36].

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, linear filters can be applied. In the time domain, this operation is mathematically described via a convolution. Thus the filtered data becomes

$$y_{\mathrm{f}}(\tau) = \int_{-\infty}^{\infty} f(\tau - t) y(t) dt \tag{2.9}$$

where $f(t)$ is some filter function and $y_{\mathrm{f}}(t)$ is the returned data under the filtering operation. For the matched filter, $f(t)$ becomes

$$f(t) = as^*(-t) \tag{2.10}$$

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, $y_{\mathrm{s}}(t)$, and its noise component, $v(t)$, (2.9) becomes

$$y_{\mathrm{mf}}(\tau) = \int_{-\infty}^{\infty} as^*(t - \tau)(y_s(t) + v(t)) dt \tag{2.11}$$

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

The autocorrelation is defined as

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

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 [32]

$$\int_a^b f_1(x)f_2(x)dx \leq \int_a^b f_1(x)dx \int_a^b f_2(x)dx \tag{2.13}$$

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

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

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, which is the most important aspect of the matched filter. A more complete discussion of this can be found in [5, 7, 37, 38].

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 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

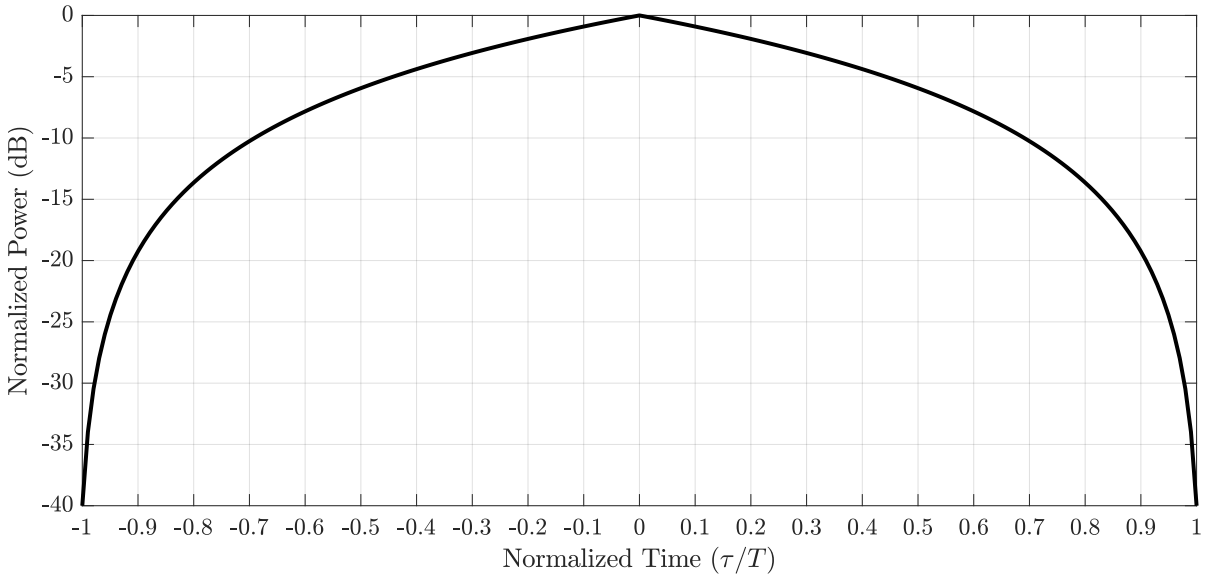

Figure 2.3: The autocorrelation response of a baseband tone

which is [32]

$$\delta R = \frac{cT}{2} \tag{2.15}$$

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 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 autocorrelation. 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, 32, 37, 39]

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

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

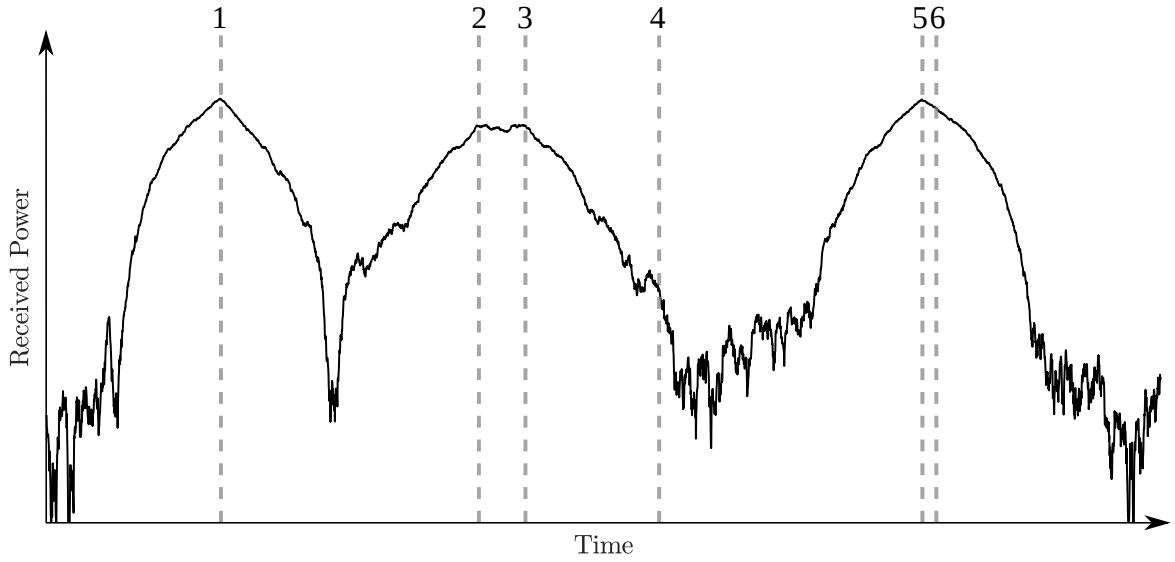

Figure 2.4: The noisy radar data from Fig. 2.2 after being matched filtered with a baseband tone

(2.16) is

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

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

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

such that the frequency function in Hz is

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

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 $B = 5$ MHz. Each pulse has been normalized to unit energy such that $\int_0^T |s(t)|^2 dt = 1$. The autocorre-

lation 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 evalutated 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 ($|f| \leq 2.5$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

$$B_{\text{tone}} \approx \frac{1}{T} \tag{2.20}$$

which in this case is 100 kHz. This corresponds to approximately the 4 dB bandwidth of the tone or where the spectral power falls below 4 dB of the peak. For the LFM, the bandwidth is normally considered to be its swept bandwidth which corresponds to approximately its 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

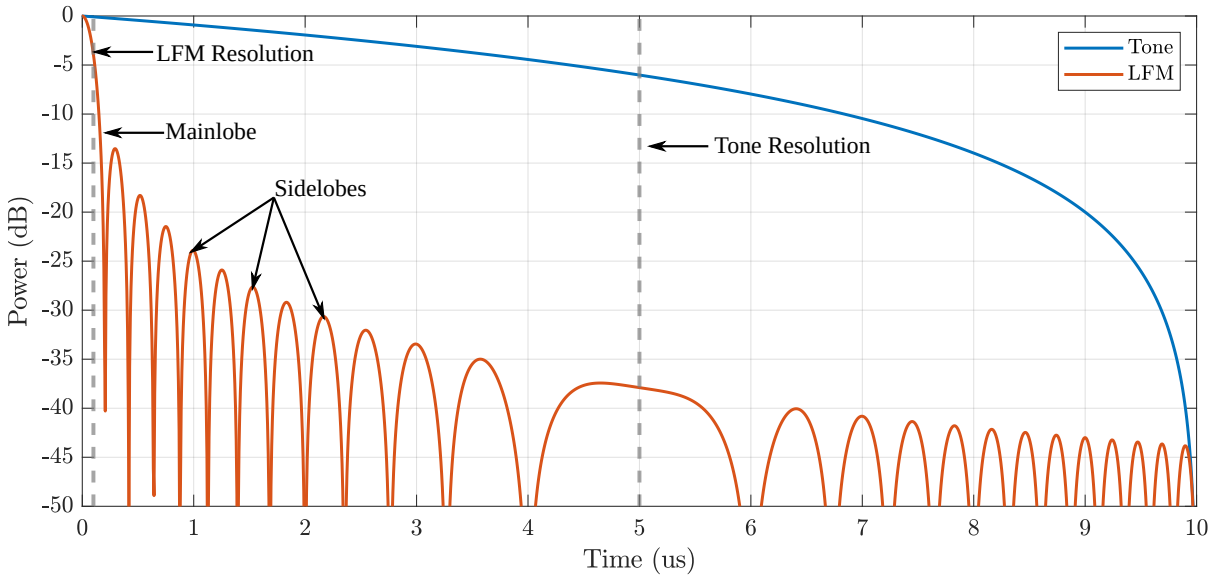

Figure 2.5: The autocorrelation of 10 us tone and a 10 us, 5 MHz LFM

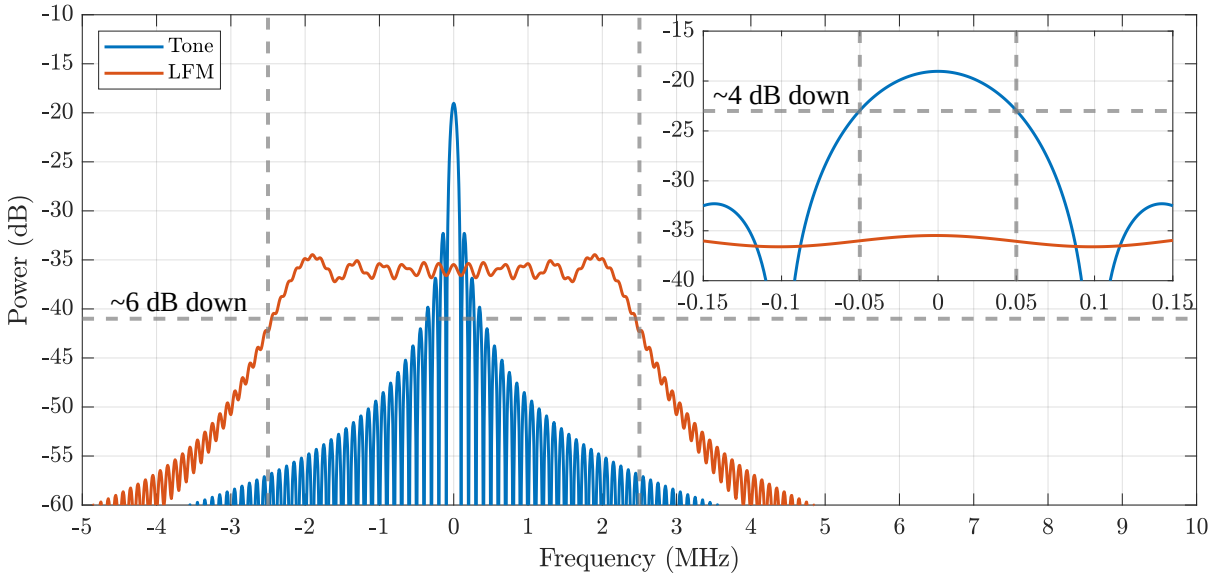

Figure 2.6: The power spectrum of 10 us tone and a 10 us, 5 MHz LFM

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 an unmodulated pulse, the BT is

$$BT = T\frac{1}{T} = 1. \tag{2.21}$$

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

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

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.

2.1.3 Pulse Integration and LFM Sidelobe Mitigation

2.1.3.1 Pulse Integration

Radars 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 a *pulse repetition frequency*

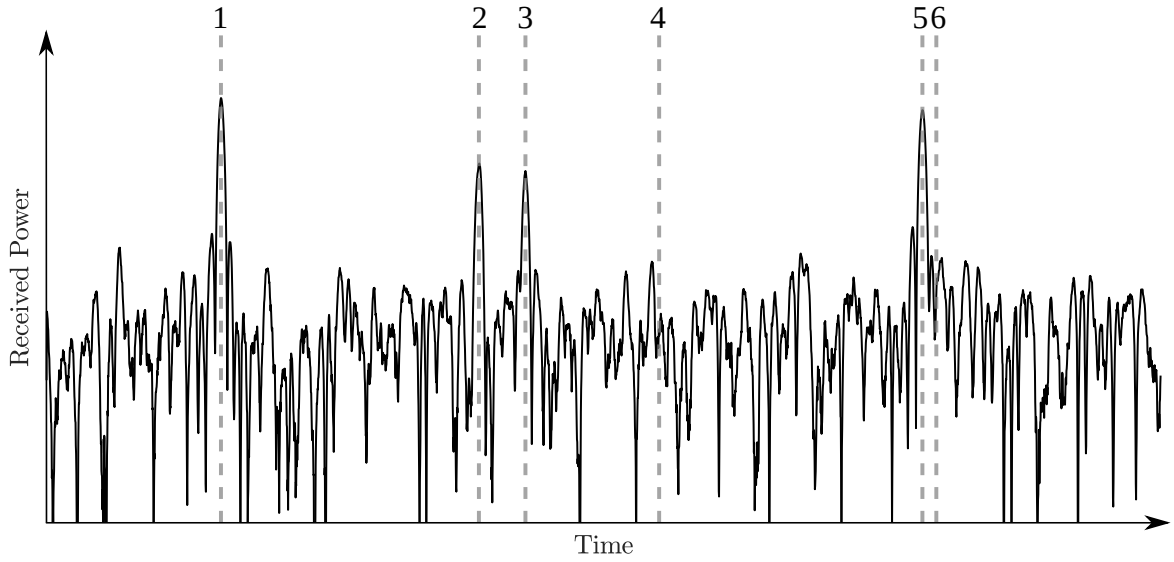

Figure 2.7: The noisy radar data from Fig. 2.2 where the transmit signal was an LFM after being matched filtered

(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 t_{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 immediate 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 SNR.

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

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

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

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

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

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

Since the noise is i.i.d., the cross terms cancel and (2.25) becomes [40]

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

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 be 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

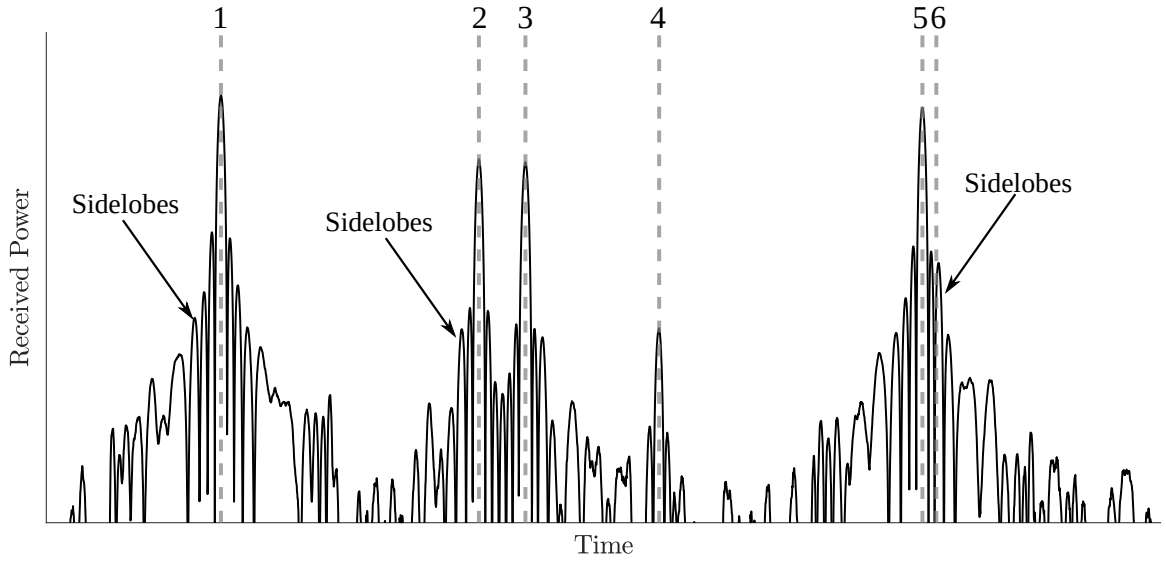

Figure 2.8: The noisy radar data from Fig. 2.2 where the transmit signal was an LFM after being matched filtered and coherently integrated over 100 pulses

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.

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

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

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 [32]. 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.

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 [41–43].

2.2 Pulse Compression Waveform Design

Given the SNR benefits provided by matched filtering, much of the work in designing radar waveforms has naturally focused on the mitigation of autocorrelation sidelobes in order to further improve radar performance. Likewise, this section expands upon the spectral challenges that result from attempting to lower those same sidelobes. To demonstrate these challenges and some of their proposed solutions this section is broken into several subsections which describe different waveform models and design schemes including, non-linear FM (NLFM), phase codes, poly-phase coded FM (PCFM), and random FM waveforms. These subsections are in now way exhaustive in the topic of waveform diversity [6, 8, 9]. However, they do provide a good starting point when it comes to designing waveforms specifically for lower autocorrelation sidelobes and spectral containment.

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

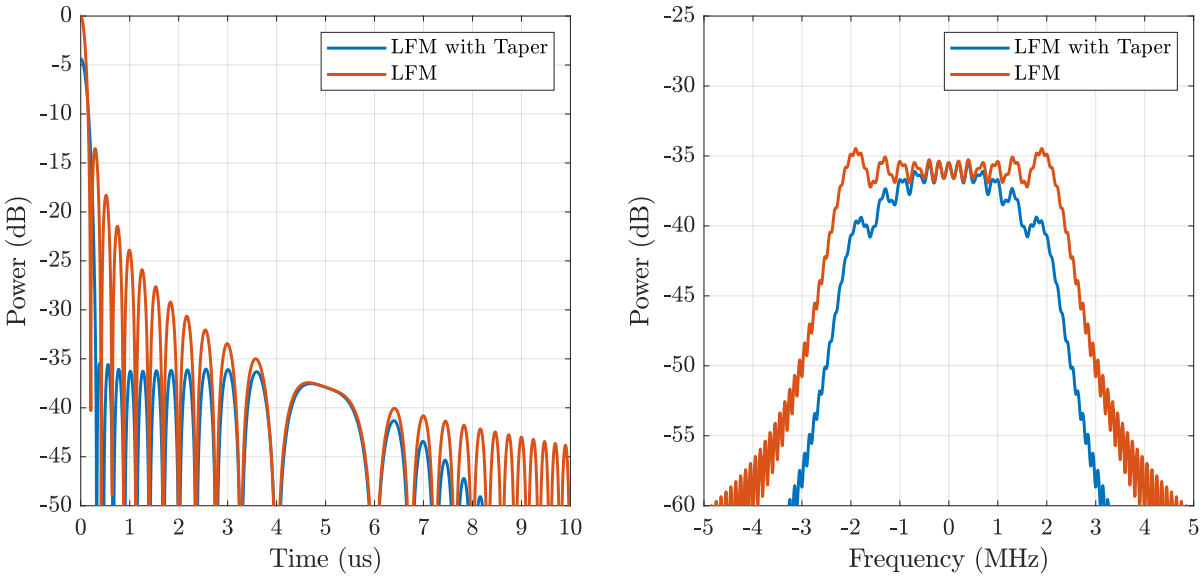

Figure 2.9: Comparison of an LFM waveform with and without a Taylor tapering window

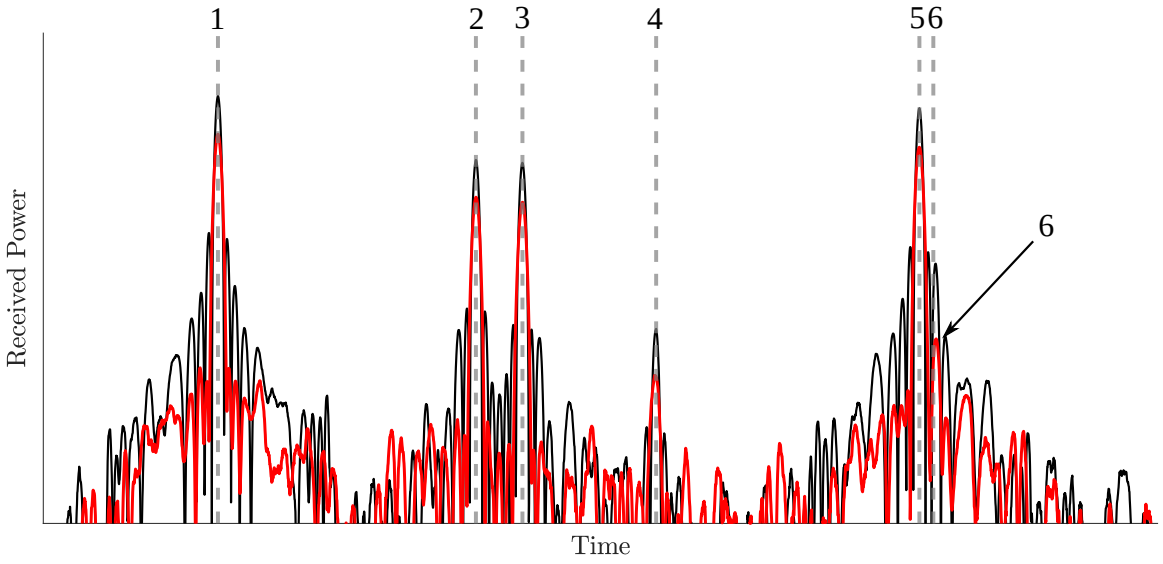

Figure 2.10: The noisy radar data from Fig. 2.2 with either an LFM or a tapered LFM and coherent integration

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 designed via the principle of stationary phase (POSP) and have time-frequency function which resembles a "sideways-S" [44–46].

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

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

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

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

and the N phase values, $\phi_1,\ \phi_2,\ \cdots,\ \phi_N$ 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 [6, 7]. 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 [47]. 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 that are opposite of 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 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 [48]! Since then, other more efficient approaches have extended the list of known MPS codes such as [49] 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 [50]. Beyond Barker, polyphase Barker, and MPS codes, numerous other code designs have been proposed. These include Frank

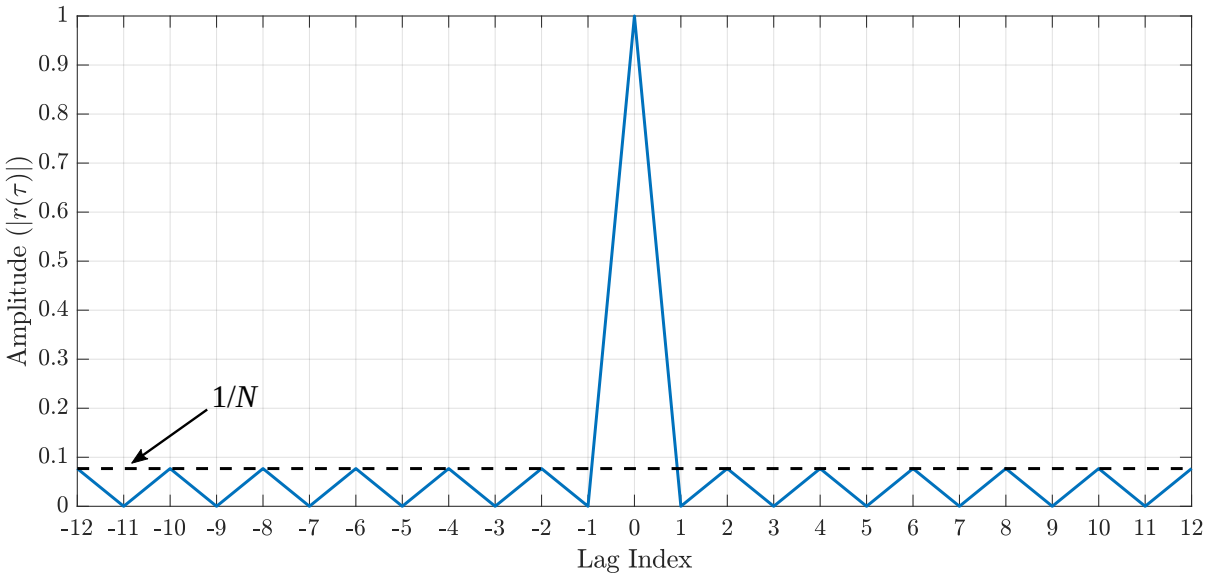

Figure 2.11: The autocorrelation of a length-13 Barker code

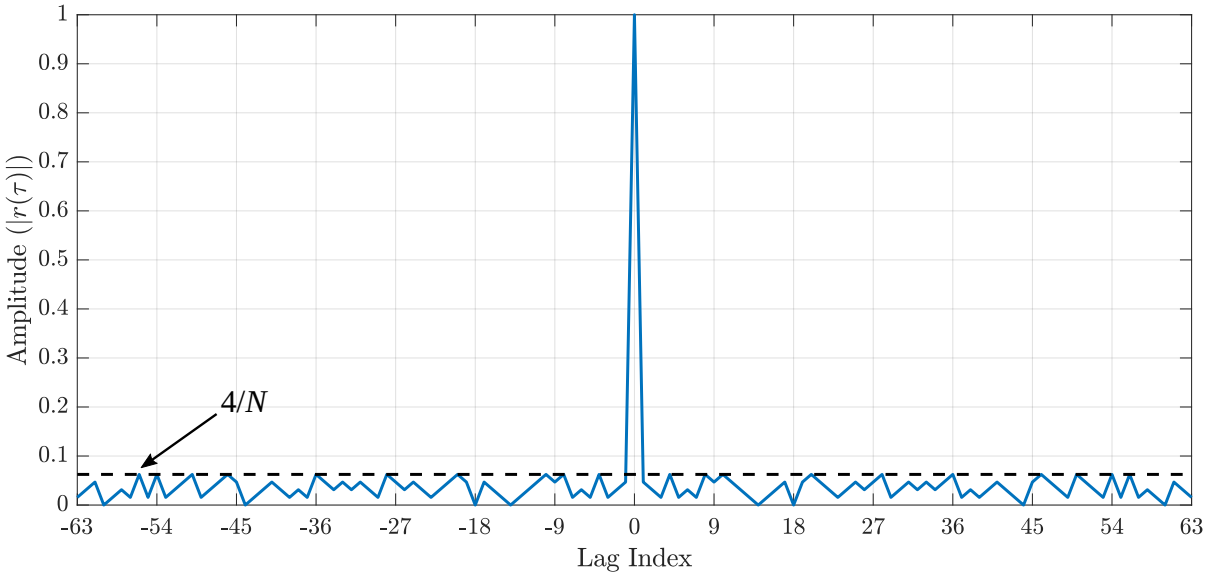

Figure 2.12: The autocorrelation of a length-64 MPS code

codes [51], Zadoff-Chu codes [52], P-codes [53, 54], Golomb codes [55], and many more. [6–8] provide excellent overviews of this topic along with a more in depth analysis of particular codes 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 and LFMs bandwidth is not really an issue from a waveform generation standpoint since each have a spectra which decays or "rolls-off" quite quickly beyond its primary 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

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

and the power spectral density is then

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

which can be reduced to

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

which is a super-position of cosine modulated sinc squared functions which result in poor spectral roll-off [56]. The cosine terms result in a repeating pattern of images every $1/t_{\mathrm{b}}$ 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

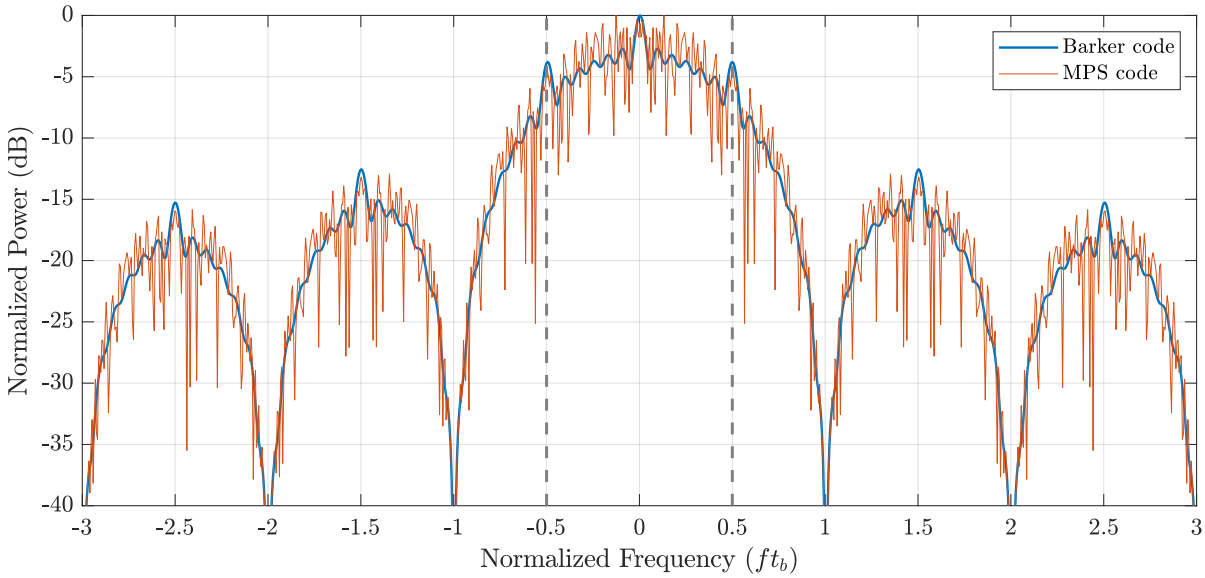

Figure 2.13: The power spectral density of the Barker and MPS codes from Figs. 2.11 and 2.12 respectively

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 [12].

Clearly, there are 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 [57] and derivative phase shift keying [58] have been shown to be affective at containing the bi-phase spectrum.

Still some increase in sidelobe levels is incured from these methods. 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 [59]. However, this approach introduces AM to the signal which presents its own problems. Finally, a method that was originally 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 [60]. 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 Bluetooth^{TM} wireless standard [61–64]. 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 mathematically defined beginning or end such that

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

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

$$q(t) = \int_0^t g(\tau)d\tau \tag{2.34}$$

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 to be time limited yielding

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

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 [60, 65] in (2.33). (2.35) can alternatively be written in terms of the frequency shaping filter such that (2.35) becomes

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

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

$$g(t) = \begin{cases} \frac{1}{T_{\text{s}}} & 0 < t \le T_{\text{s}} \\ 0 & 0 < \text{otherwise} \end{cases} \tag{2.37}$$

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 [54] has been adapted to the PCFM model to improve spectral 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 of simply implementing existing poly-phase codes, waveforms based on the PCFM model waveforms were optimized and demonstrated in [66] where greedy search methods were shown as an effective means to find PCFM codes which result in good spectral containment and low 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

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

where each basis function

$$b_n(t) = \int_0^t g(\tau - (n-1)T_\mathrm{s})d\tau \tag{2.39}$$

is the integral of $g(t)$ and is time shifted by an integer multiple of T_s. Given the rectangular shaping filter, each $b_n(t)$ becomes a time shifted ramp function such that

$$g(t) = \begin{cases} 0 & 0 < t \leq (n-1)T_\mathrm{s} \\ (t-(n-1)T_\mathrm{s})/T_\mathrm{s} & (n-1)T_\mathrm{s} < t \leq nT_\mathrm{s} \\ 1 & nT_\mathrm{s} \leq NT_\mathrm{s} \end{cases} \tag{2.40}$$

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 functions which are continuous not only in phase, but also frequency, chirp-rate, etc, but for the sake of brevity these are not considered here, but can be found in [67]. Likewise, the basis functions in (2.40) could be generalized to any

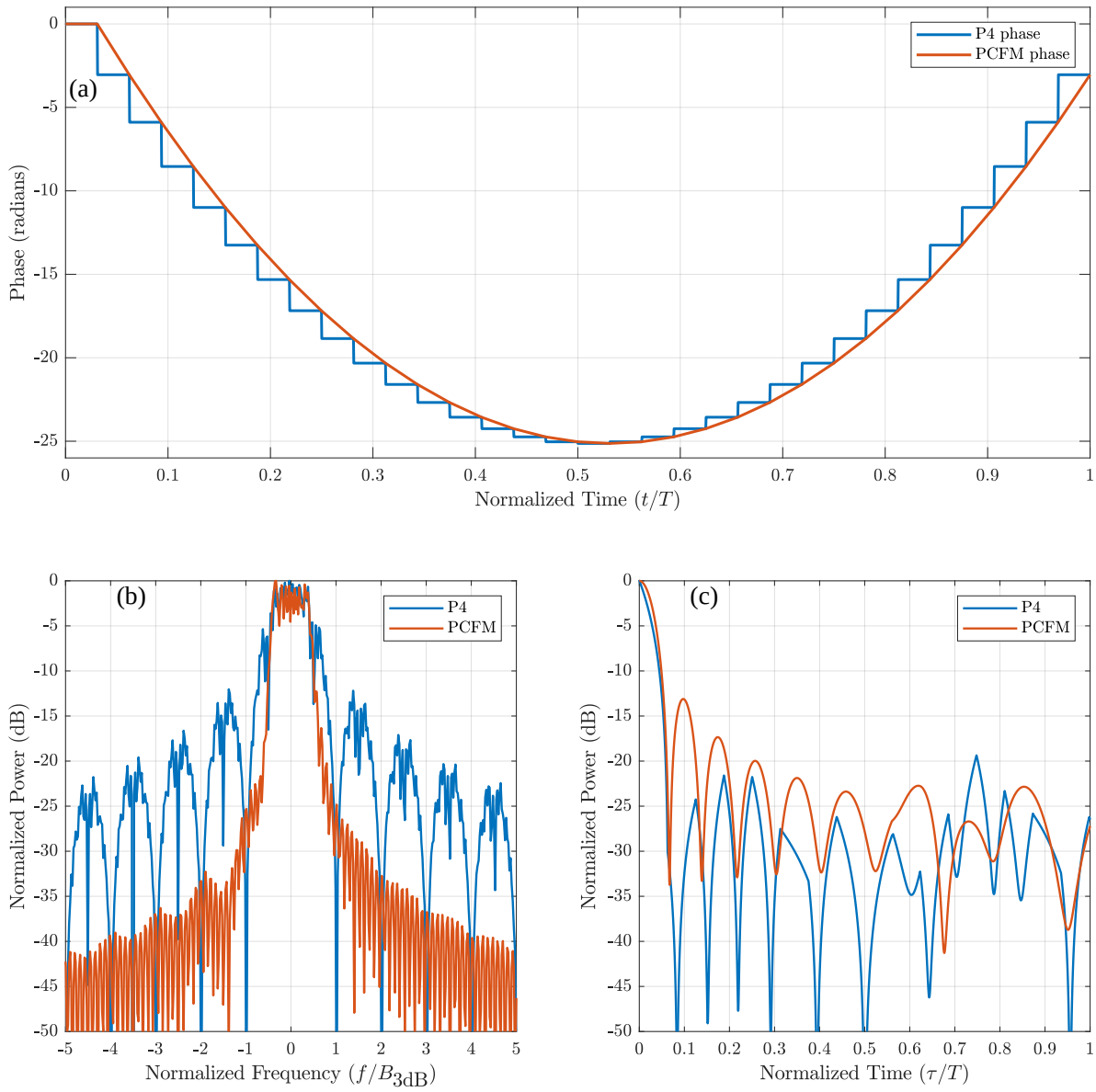

Figure 2.14: A P4 code compared to its implementation as a PCFM waveform

desirable function to realize coded FM (CFM) such as through the use of legendre polynomials [68].

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 this 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 directly and collecting these samples into the basis function matrix $\mathbf{B}$. The discretized PCFM waveform is then realized as

$$\mathbf{s} = \exp(j\mathbf{B}\mathbf{x}) \tag{2.41}$$

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 [69], 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 [69] 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 [66], to explicitly optimize for a good spectral roll-off. In [70], only initializations which tend towards spectrally contained solutions were considered. In [71], since the cost function did not tend towards spectrally contained solutions, 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 many of ways for a variety of design goals some of which will be discussed in the next section.

2.2.4 Random FM Waveforms

The previous sections largely considered the performance of single waveforms, but their performance can be considered in aggregate as well. To do so, 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 (not necessarily RFM) under three different CPI lengths 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 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

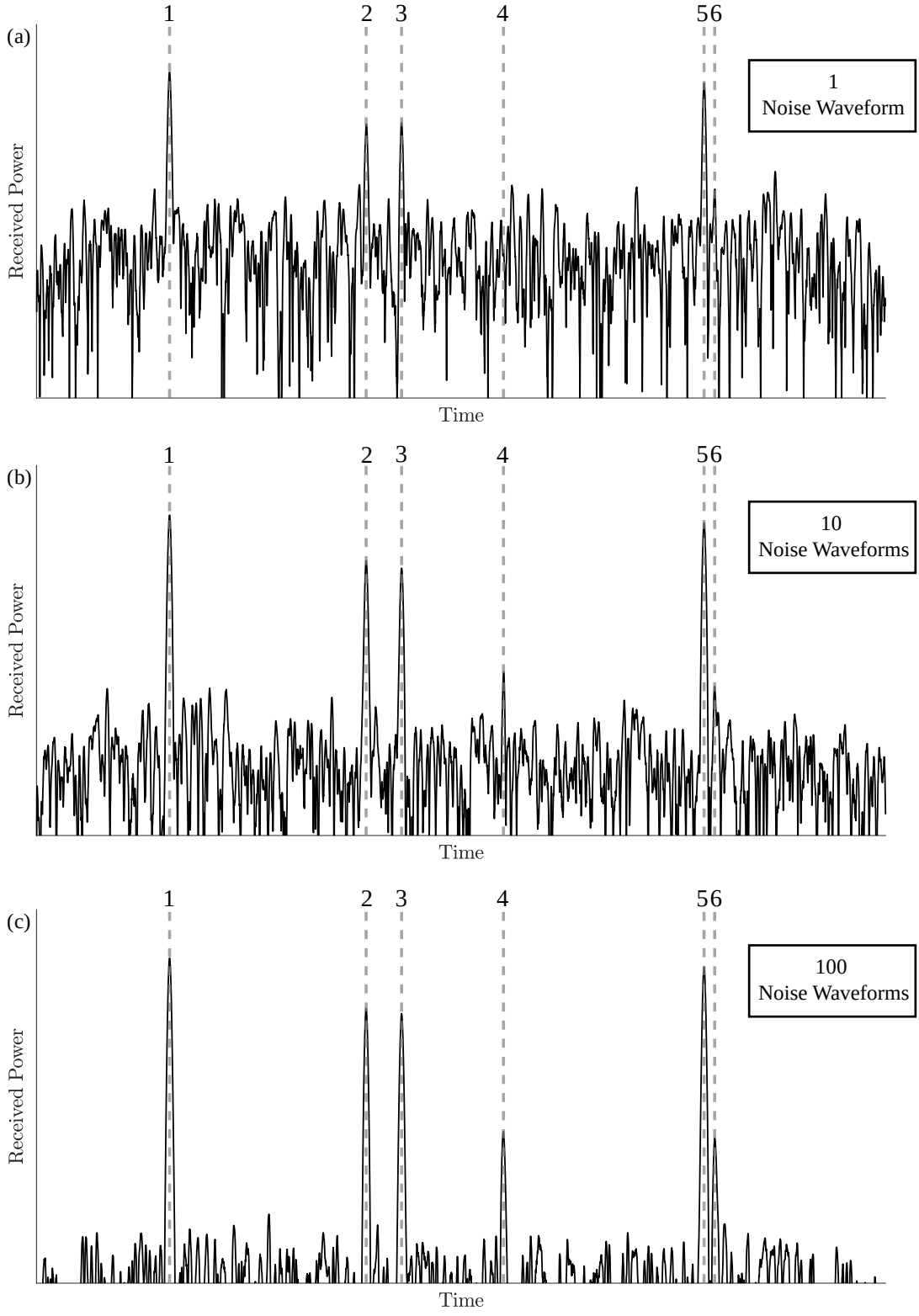

Figure 2.15: The noisy radar data from Fig. 2.2 with varying CPI sizes of 1, 10 and 100 corresponding to (a), (b), and (c) respectively

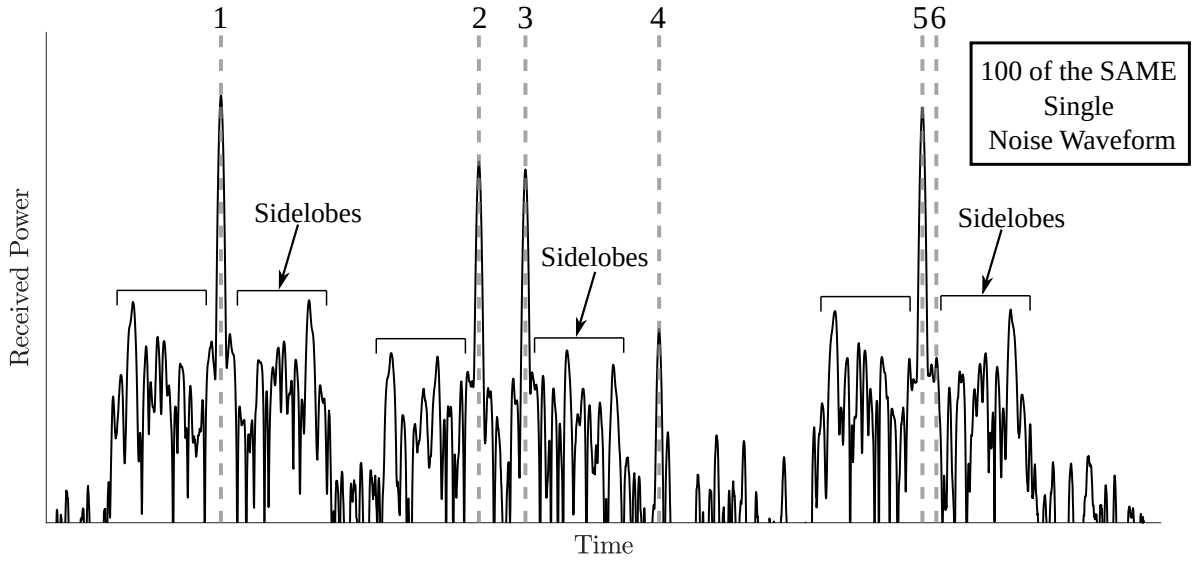

Figure 2.16: The noisy radar data from Fig. 2.2 but the transmit waveform is a single noise waveform with a 100 pulse CPI

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 presense 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 [22]. 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, 72] 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 random FM waveforms which are the focus of this work. The rest of this section introduces some the latest work on this topic.

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

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

and

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

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 RFM 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 ideally 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 random FM 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.

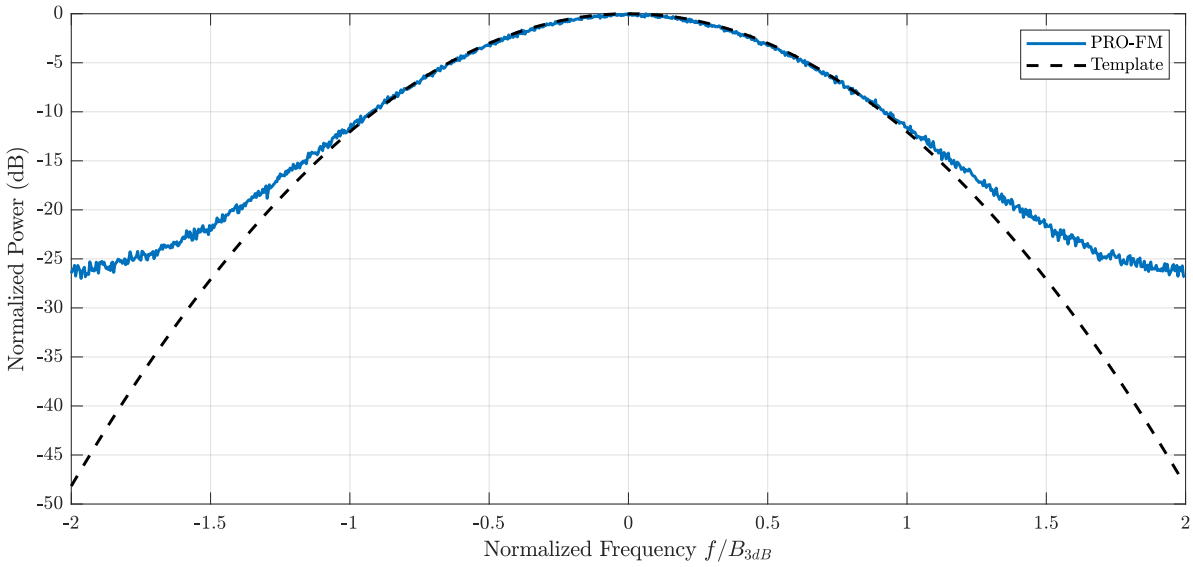

Figure 2.17: RMS spectrum of 100, optimized PRO-FM waveforms and their desired template

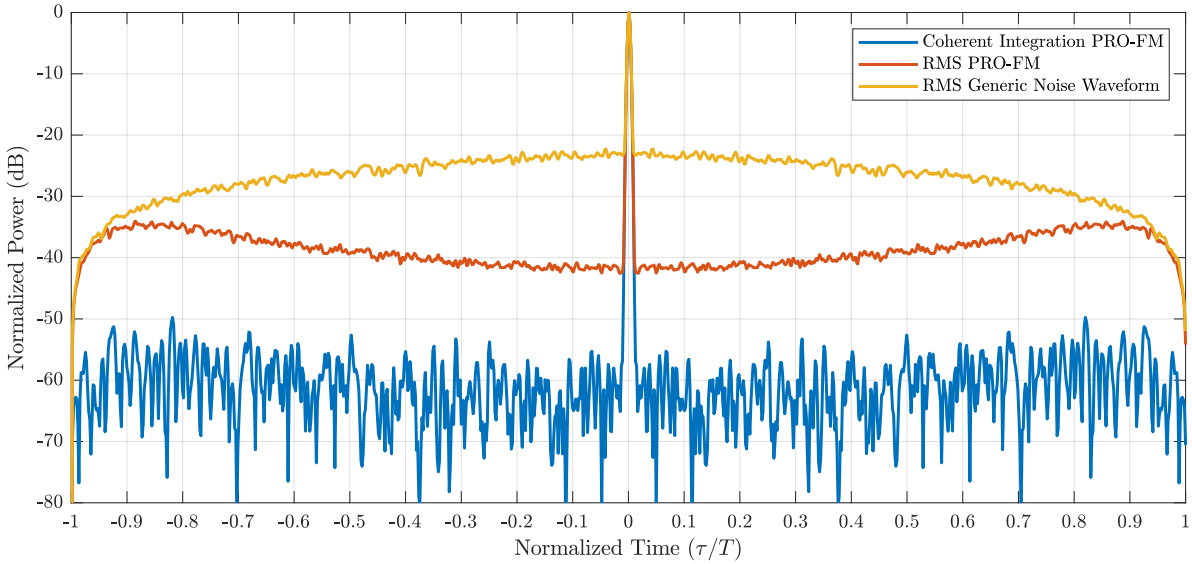

Figure 2.18: Coherently integrated and RMS autocorrelations of 100 PRO-FM waveforms compared to the RMS autocorrelation of 100 generic noise waveforms

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 at about $-10\log_{10}(BT)$. After optimization, the PRO-FM RMS sidelobes near the mainlobe are a full order of magnitude better at about $-20\log_{10}(BT)$. Then to demonstrate the noise like coherent integration of random FM 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 [73–75], radar and communications spectrum sharing [76–78], simultaneous dual-polarized radar [79], non-linear harmonic radar with FM-noise waveforms [80], and even random movement radar which mimics the human eye [81].

Since the formulation of PRO-FM, a few other approaches have been formulated to optimize random FM 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 optimization and the PCFM waveform model approach to achieve excellent spectral containment [23]. See [11] for an overview of several random FM waveform implementations.

2.2.4.2 Process Based Random FM Waveforms

As apposed to the previously discussed random FM waveform approaches which are optimized or designed on an individual level, process based random FM waveforms are more similar to drawing a sample function from a random process. Some examples of this kind of implementation are [27–29, 31, 72, 82, 83], 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. Further, [30] actually bases the phase function of constant amplitude orthogonal frequency division multiplexing (OFDM). 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 [31]. 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 4-5 put these concepts to work in designing what are called waveform generating functions (WGFs) 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

$$P_{\mathrm{avg}} = \frac{P_{\mathrm{t}} t_{\mathrm{d}}}{T_{\mathrm{PRI}}} \tag{2.44}$$

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 for a fraction of the time. If the transmitter is only active 5% of the time, then $t_{\mathrm{d}} = 0.05$. 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 $t_{\mathrm{d}} = 1$. 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 [32].

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 of the transmitter 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 [32, 84]. Still, this constraint leads to CW radars usually be low power systems with relatively short operating ranges. (Although, there is at least one major exception to this due to special circumstances [85]).

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 [86].

In terms of noise radar, the approaches proposed by [22, 72] 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 pulsed 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 it sidelobes. If the sidelobes of the direct path are low enough, they will not mask targets of interest. For the CW random FM waveforms designed in Chapter 5, the same segmented matched filtering process will be used.

2.4 Doppler Processing and the Ambiguity Function

2.4.1 The Doppler Effect

One the most useful capabilities of many radar systems is their ability to accurately discern the velocity of a given object, whose ability to do this is based on the Doppler effect. Just about everyone has experienced the Doppler effect in the context of a siren on an ambulance or firetruck becoming lower or higher pitched as it moves away or approaches respectively. A similar effect occurs with electromagnetic radiation.

To explain, consider an object moving radially away from a radar. When the radar transmits the pulse the object is at some range R_0 relative to the radar. However, by the time the leading edge of the pulse reaches the object, the object will have moved beyond R_0. Accounting for the velocity of the object v and the speed of propagation of the radar pulse c, the time for the leading edge of the pulse to come into contact with the object can be calculated by setting the position of the pulse's leading edge equal to the position of the object such that

$$ct = R_0 + vt \tag{2.45}$$

by solving for t. Doing so yields

$$t_0 = \frac{R_0}{c - v} \tag{2.46}$$

where t_0 represents the time when the leading edge of the pulse impinges on the moving object. The round trip time is then two times this value such that

$$t_{\mathrm{r1}} = \frac{2R_0}{c - v} \tag{2.47}$$

where t_{r0} is the total time from the transmission of the leading edge of the pulse to its returning to the receiver. A similar analysis can be performed for the trailing edge of the pulse. With respect to the leading edge of the pulse, the trailing edge is transmitted after a T second delay where T

is the duration of the pulse. During the pulses transmission, the object has continued moving so by the time the trailing edge leaves the transmitter it is at a new position R_1. The delay from the transmission of the leading edge of the pulse to the trailing edge of the pulse hitting the object is then

$$t_2 = \frac{R_1}{c-v} + T. \tag{2.48}$$

Based on the delay to transmitting the trailing edge of the pulse and the motion of the object, $R_1 = R_0 + vT$. This transforms (2.48) into

$$t_1 = \frac{R_0 + vT}{c-v} + T. \tag{2.49}$$

The fractional term in (2.49) represents the flight time of the trailing edge of the pulse. Doubling this term realizes the delay from the transmission of the leading pulse edge to the reception of the reflected trailing pulse edge such that

$$t_{\mathrm{r1}} = 2\frac{R_0 + vT}{c-v} + T. \tag{2.50}$$

The difference between (2.50) and (2.47) represents the duration of the reflected pulse where

$$T_{\mathrm{r}} = 2\frac{R_0 + vT}{c-v} + T - \frac{2R_0}{c-v} = T\frac{c+v}{c-v}. \tag{2.51}$$

In other words, the motion of the object has modulated the duration of the pulse such that the returned pulse is longer or shorter depending on whether the object is moving away or towards the radar respectively. (2.51) holds true so long as nothing is moving at a significant fraction of the speed of light which should be true for any terrestrial application. The ratio between the received duration and the transmit duration is then

$$\frac{T_{\mathrm{r}}}{T} = \frac{c+v}{c-v}. \tag{2.52}$$

To this point the discussion has focused on the difference in arrival time between the leading and trailing edges of a pulse, but (2.52) can be applied to any arbitrary points during a signal transmission CW or otherwise. For a sinusoidal (tone) transmission, defining these points as the beginning and end of a frequency period shows how this expansion or contraction of the pulse effects frequency. Assuming T is the transmit period and T_r is the received period, then the change in frequency is the reciprocal such that

$$\frac{f_{\mathrm{r}}}{f_{\mathrm{t}}} = \frac{c-v}{c+v}. \tag{2.53}$$

Further since typically $v \ll c$, (2.53) can be approximated and the new frequency is

$$f_{\mathrm{r}} \approx f_{\mathrm{t}} \left(1 - \frac{2v}{c}\right). \tag{2.54}$$

This can be rewritten in terms the transmit wavelength λ such that the Doppler frequency shift takes on the familiar form

$$f_{\mathrm{D}} = f_{\mathrm{r}} - f_{\mathrm{t}} \approx -\frac{2v}{\lambda}. \tag{2.55}$$

The question is then whether (2.55) is applicable to non-tonal waveforms such as the ones discussed in the previous sections. This introduces the concept of narrow-band versus wide-band signals and percent bandwidth. Percent bandwidth is defined such that

$$\%\mathrm{BW} = \frac{B}{f_{\mathrm{c}}} \times 100\% \tag{2.56}$$

where f_{c} is the carrier frequency of the signal and B is the bandwidth of the signal by some metric such as those mentioned previously. Generally speaking, if (2.56) is less than 10% then the signal is considered narrow-band and the dominant impact of Doppler on the signal is described by (2.55). Otherwise, the pulse duration modulation effect described by (2.52) needs to be used.

2.4.2 The Ambiguity Function

Understanding how the match filtering process is impacted by Doppler shifts of the waveforms is critical to evaluating the behavior of a given radar waveform. The prototypical means of assessing this is through what is known as a ambiguity function. The ambiguity function is defined such that

$$A(\tau, f_{\mathrm{d}}) = \int_{-\infty}^{+\infty} \exp(j2\pi f_{\mathrm{d}} t) s(t) s^*(t+\tau) dt \tag{2.57}$$

where $f_{(d)}$ is the Doppler frequency. Note that the only difference between the ambiguity function and the matched filter is the addition of the Doppler term. That is, if f_{d} is set to zero, the (2.57) is the autocorrelation response. More specifically, (2.57) is the narrow band ambiguity function where the Doppler frequency is approximated as only a function of the carrier frequency and the waveform undergoes a uniform frequency shift. More general forms of the ambiguity function take into account the pulse duration dilation to capture the Doppler shift for every frequency present in the signal. However, for narrow-band signals, (2.57) is perfectly sufficient.

To understand how Doppler shifts can impact the match filtering process, Figs. 2.19 and 2.20 show the ambiguity functions of an LFM (down-chirp) and a PRO-FM waveforms both with a BT of 128. In Fig. 2.19, the LFM ambiguity function conforms to a very regular structure sometimes referred to as a "sheared ridge". As the degree of Doppler increases, there is an obvious coupling between range and Doppler such that higher and higher amounts of Doppler will result in range registration errors. This behavior is a direct result of the LFM's linearly sweeping time-frequency behavior. The vertical line at zero Doppler is exactly the LFM matched filter response. In Fig. 2.20, there is no such regular structure. Off from zero Doppler, there is no meaningful mainlobe whatsoever. In this case it is not so much that a high degree of Doppler results in a range registration error so much that it would result in no registration at all. This is known as a "thumbtack" response. The LFM ambiguity function is said to be Doppler-tolerant in that detection is possible in the presence of significant Doppler, while this is clearly not the case for the PRO-FM waveform and any noise like waveform like it. This brief discussion on the ambiguity function is meant to

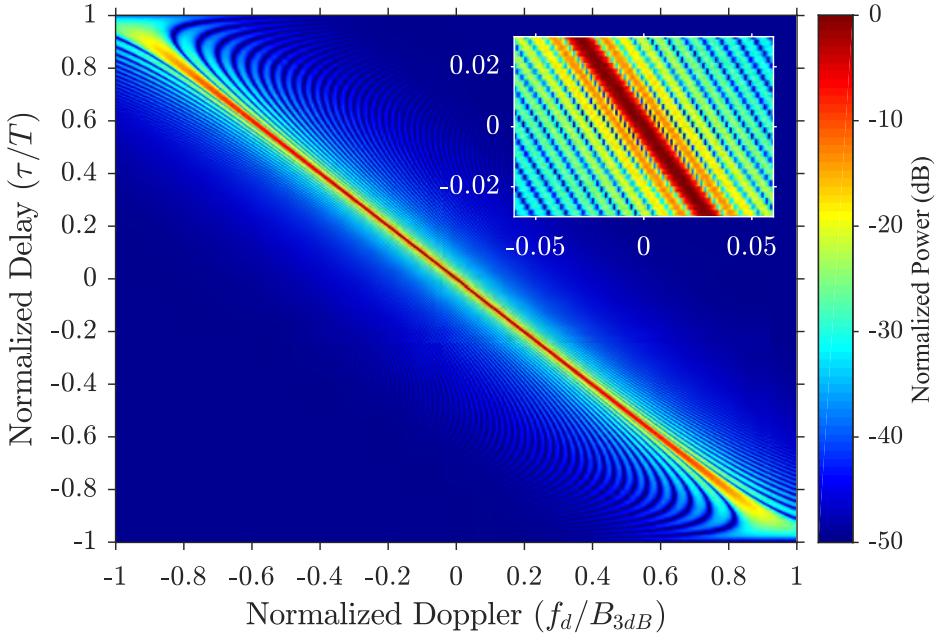

Figure 2.19: Range-Doppler ambiguity function of an LFM waveform

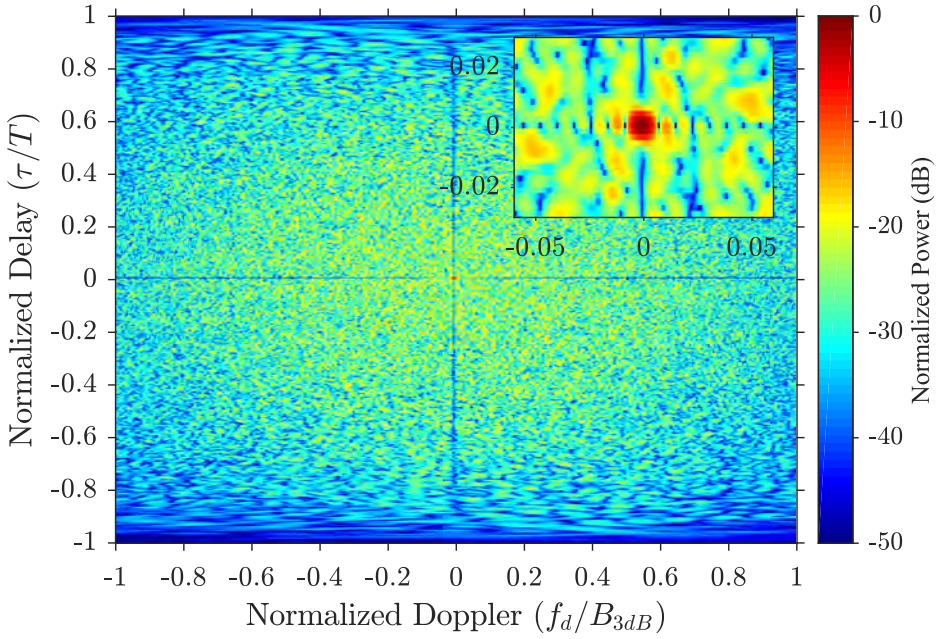

Figure 2.20: Range-Doppler ambiguity function of a single PRO-FM waveform

demonstrate the significant effects Doppler shifts can impart on the match filtering process. A more in depth discussion on the ambiguity function and Doppler in general can be found in numerous resources such as [7, 32].

2.5 Gradient Descent Techniques

The different waveform design schemes employed in this work all utilize 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 [87]

$$\begin{aligned} &\text{minimize} \quad J(\mathbf{x}) \\ &\text{such that} \quad f_i(\mathbf{x}) \leq c_i, \;\; i = 1, 2, \ldots, M \end{aligned} \tag{2.58}$$

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

$$J(\mathbf{x}^*) \leq J(\mathbf{x}) \text{ for all } \mathbf{x}. \tag{2.59}$$

where the $*$ indicates $\mathbf{x}^*$ minimizes the cost function value. As defined in 2.59, $\mathbf{x}^*$ 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

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \text{ for all } \mathbf{x} \in \mathscr{N} \tag{2.60}$$

where $\mathscr{N}$ is an infinitesimal open set (local neighborhood) containing $\mathbf{x}^*$ [56].

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

optimizable parameters is updated as

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \tag{2.61}$$

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

$$\mathbf{p}_k = \begin{cases} -\nabla J_k & \text{when } k = 0 \\ -\nabla J_k + \beta_k \mathbf{p}_{k-1} & \text{otherwise} \end{cases} \tag{2.62}$$

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 [70]. An excellent survey of NLCG methods can be found in [88].

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

$$0 < \beta < 1. \tag{2.63}$$

In this way, the search direction has a kind of inertia where the search direction can not change very quickly [89]. 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 core goal of this entire work is to design, evaluate, and implement what are known as waveform generating functions (WGFs). By definition, a WGF is a random process whose sample functions are intended to be radar signals. Of course, this definition is extremely general. It does not say anything about the stochastic process other than its intended use. Without designing the underlying random process of the, WGF the resulting "waveforms" could be extremely poor or even nonsensical from a radar perspective. Consequently, there is a need to be able to evaluate the WGF for its ability to create useful radar waveforms.

Since the WGF is a stochastic entity, it has to be evaluated using statistical tools and these tools have to be framed such that they reveal useful characteristics of the WGF from a radar perspective. For example, a simple statistical analysis of the WGF may answer the question, "what is the correlation between samples of the WGF separated by t seconds?" or "what is the mean value of the WGF over time?", but the answers to these kinds of questions reveal little about the usefulness of the WGF for producing good radar waveforms. Instead, from a radar perspective more useful questions are, "on average what spectrum will the sample functions of the WGF possess?" or "on average what will the matched filter response look like?". Answering these questions requires applying additional nuance to the typical analysis of stochastic processes. With this in mind, the analytical tools developed here build directly off of the fundamentals of random signals analysis covered in any random signals text such as [40, 43, 90–92], so it is prudent to first cover the basics of these topics as they apply to the discrete time random processes designed as WGFs in this work. 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.

To begin, Section 3.1 provides a brief overview of the fundamentals of random variables through examples, by introducing the concept of moments, and defining the relationships between different random variables. Section 3.2 extends these ideas to random processes and the relationships between them. Section 3.3 explains how the topics in apply to stochastically defined radar signals and stochastic radar signal processing. Finally, Section 3.4 aggregates all this information into a set of eight metrics for evaluating both CW and pulsed random processes as WGFs for their ability to produce useful radar waveforms.

3.1 Random Variables Review

Generally speaking, random variables describe the possible outcomes of an experiment such as rolling a die or the toss of a coin. For example, in the case of a fair die, there are six equally probable outcomes while for a 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

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

which 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. Note that for (3.1), and any PMF in general,

$$p_X(x = x_i) \geq 0 \tag{3.2}$$

and

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

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, $\mathrm{E}[\,\cdot\,]$ such that the *n*th moment is defined

$$\mathrm{E}[X^n] = \sum_{i=1}^{N} x_i p_X(x = x_i), \tag{3.4}$$

where N is the total number of possible outcomes of the random variable X. The simplest moment is naturally the first moment which is also known as the mean and is defined such that

$$\mathrm{E}[X] = \sum_{i=1}^{N} x_i p_X(x = x_i) = \mu_X, \tag{3.5}$$

which can be intuitively understood as the sample average of the random variable X. In other words, if more and more independent outcomes of the random variable X where averaged, the results would approach μ_X. Additionally, central moments are defined relative to the mean such that the *n*th central moment is defined

$$\mathrm{E}[(X - \mu_X)^n] = \sum_{i=1}^{N} (x_i - \mu_X)^n p_X(x = x_i). \tag{3.6}$$

The most commonly used central moment is the variance which is defined

$$\mathrm{E}[(X - \mu_X)^2] = \sum_{i=1}^{N} (x_i - \mu_X)^2 p_X(x = x_i) = \sigma_X^2, \tag{3.7}$$

where the variance can be understood as the degree to which the outcomes of the random variable vary around the mean. The positive valued square root of the variance is the standard deviation.

These definitions can be extended to continuous random variables which can describe many natural processes such as the thermal noise in a circuit. These continuous random variables as apposed to the example discrete random variable defined in (3.1) can take on a continuum of values and are instead described by the *probability density function* (PDF). The continuous analogue to

the discrete uniform (equal probability) PMF of the die roll in (3.1), is the continuous uniform distribution defined as

$$p_X(x) = \begin{cases} 1/(b-a) & a < x < b \\ 0 & \text{otherwise} \end{cases}. \tag{3.8}$$

The outcome of an experiment described by (3.8) is equally likely to be any value in the range $a < x < b$. Like the PMF, the PDF is also strictly positive although it integrates to 1 on the real number line rather than summing to one over its members. The moments of a continuous random variable are defined by an integral rather than a sum such that the nth moment becomes

$$\mathrm{E}[X^n] = \int_{-\infty}^{\infty} x^n p_X(x) dx, \tag{3.9}$$

where the first moment is the mean and is defined

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x p_X(x) dx = \mu_X. \tag{3.10}$$

The nth central moment becomes

$$\mathrm{E}[(X-\mu_X)^n] = \int_{-\infty}^{\infty} (x-\mu_X)^n f_X(x) dx. \tag{3.11}$$

and the variance becomes

$$\mathrm{E}[(X-\mu_X)^2] = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x) dx = \sigma_X^2, \tag{3.12}$$

Together, moments and central moments for both the discrete and continuous random variable are a useful means of evaluating their behavior, but often times two or more random variables will be related to each other. These relationships can likewise be mathematically characterized.

Consider the fair die roll from above. What if two fair dice are rolled simultaneously? In the

case of random variables X_1 and X_2 for the first and second die respectively, the joint PMF is

$$p_{X_1,X_2}(x_1 = x_{1,i}; x_2 = x_{2,i}) = \begin{cases} 1/36 & x_{1,i} = 1,2,3,4,5,6; \quad x_{2,i} = 1,2,3,4,5,6 \\ 0 & \text{otherwise} \end{cases} . \tag{3.13}$$

where the individual variables can take on any integer value from one to six and the probability of any pair of values is $1/36$. When rolling two fair dice it is intuitive that the roll of one die does not impact the probability the other. In this way, X_1 and X_2 are independent in the sense that the outcome of one does not impact the PMF of the other or vice versa. Consequently, in (3.13) the probability of

$$p_{X_1,X_2}(x_1 = x_{1,i}; x_2 = x_{2,i}) = p_{X_1}(x_1 = x_{1,i}) p_{X_2}(x_2 = x_{2,i}). \tag{3.14}$$

Independence also provides an important result for expectations. For independent random variables [40]

$$\mathrm{E}[\, X_1 X_2 \,] = \mathrm{E}[\, X_1 \,]\mathrm{E}[\, X_2 \,]. \tag{3.15}$$

Now consider a new random variable, Y, where Y is sum of the outcomes between X_1 and X_2 such that

$$Y = X_1 + X_2 \tag{3.16}$$

where the result of either X_1 or X_2 will clearly impact the value of Y. The relationship between Y and either X_1 or X_2 can be measured by their respective covariance where

$$\sigma_{XY} = \mathrm{E}[\, (Y - \mathrm{E}[\, Y \,])(X - \mathrm{E}[\, X \,]) \,] \tag{3.17}$$

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 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.

If σ_{XY} is something other than zero, then the variables are correlated to some degree. If the covariance is zero then the variables are uncorrelated. However, they are not necessarily independent. (3.17) is only a measure of the linear relationship between random variables. (3.17) can be zero, but the random variables may be related in a non-linear way and thus uncorrelated but not independent.

Finally, another useful tool for examining the behavior of random variables is the characteristic function which is defined

$$\psi_X(\omega) = \mathrm{E}[\exp(j\omega X)] = \int_{-\infty}^{\infty} \exp(j\omega x) f_X(x) dx. \tag{3.18}$$

where this form can apply to both continuous and discrete random variables; however, in the discrete case the PMF must be written in terms of Dirac delta functions. As a tool, the characteristic function is an effective means of calculating the moments of the random variable X, since it is often easier to calculate these moments from the characteristic function than it is to do so directly using (3.4) or (3.9). More details on how this can be done can be found in [40, 90, 93]. Interestingly, the characteristic function is extensively used in Chapters 4 and 5 of this work, but only indirectly and not for calculating moments. This is made clear in those chapters.

In summary, this section acts a very brief introduction to the fundamentals of random variables as they are used in this work. A much more comprehensive discussion on these topics can be found in any random variables and stochastic processes dedicated textbook such as [40, 90, 93] to name just a few.

3.2 Random Processes Review

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 $X[m]$ (square brackets indicate discrete time) and

is defined as follows. At each instant of time, m, 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 defined to be 0 such that $X[0] = 0$. The table below shows the PMFs of $X[m]$ for the first several times m. For clarity, the zero valued cells have been filled with '-'. At each time, m,

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

Table 3.1: Random walk PMFs for $m = 0, 1, 2, 3, 4, 5$

the value of $X[m]$ 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 m. Based on the definition of the random walk, if $X[m_0] = 4$, then $X[m_0 + 1]$ can only equal either 3 or 5 with equal probability, but there is really no way to see that 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 random variables, it can instead be viewed as a collection of member functions which each have their own probability of occurring. For the random walk some possible partial sequences are shown in Table 3.2. where x_a or any other subscript indicates a unique sample function of the random process $X[m]$. 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.

m	=	0,	1,	2,	3,	4,	5,	6,	$\cdots$
$x_a[m]$	=	0,	-1,	-2,	-1,	0,	-1,	-2,	$\cdots$
$x_b[m]$	=	0,	1,	2,	3,	4,	3,	4,	$\cdots$
$x_c[m]$	=	0,	1,	0,	-1,	-2,	-3,	-2,	$\cdots$

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

These observations made here regarding the random walk process and the characterization of random processes in general is formalized in the following sections. However, this analysis will focus only on discrete time stochastic processes. While the development in the following sections can be generalized to continuous random processes, doing so is not necessary since the WGF design of this work is performed using exclusively discrete time processes.

3.2.1 Correlation Functions

Table 3.1 demonstrates how a stochastic process at a given instant 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

$$R[m_1,m_2] \triangleq \mathrm{E}\left[X[m_1]X^*[m_2]\right]. \tag{3.19}$$

Alternatively, the autocovariance function measures the relationship at times m_1 and m_2 relative to the mean such that

$$V[m_1,m_2] \triangleq \mathrm{E}\left[\left(X[m_1]-\mathrm{E}[X[m_1]]\right)\left(X[m_2]-\mathrm{E}[X[m_2]]\right)^*\right]. \tag{3.20}$$

where the relationship between (3.20) and (3.17) is obvious. However, instead of considering the relationship between different random variables, (3.20) considers the relationships that exist within a random process at different times. If the random process $X[m]$ happens to be zero mean for all time, then (3.19) and (3.20) are equivalent.

This concept can also be extended across different random processes using the cross-correlation function where

$$R_{XY}[m_1, m_2] \triangleq \mathrm{E}[\, X[m_1] Y^*[m_2] \,], \tag{3.21}$$

or the cross-covariance function where

$$V_{XY}[m_1, m_2] \triangleq \mathrm{E}\left[\left(X[m_1] - \mathrm{E}[X[m_1]]\right) \left(Y[m_2] - \mathrm{E}[Y[m_2]]\right)^* \right], \tag{3.22}$$

If the random processes $X[m]$ and $Y[m]$ are both zero mean for all time, then (3.21) and (3.22) are equivalent.

Understanding the autocorrelation or the autocovariance function of a random process goes a long way in characterizing that process; however, there a few properties of (3.19) and (3.20) that can make them difficult to use, but primarily the issue is that they are a function of absolute time which means to use these functions the underlying process must have an absolute time line. Mathematically, this may not be an issue, but consider modeling the wave height on the ocean as a random process. What does absolute time even mean in this context? In many physical systems an absolute time reference makes little sense. Fortunately many of these same systems can be modeled such that they are not functions of absolute time at all. Instead, their autocorrelation function only depends on relative time.

For a stochastic process to be stationary in its autocorrelation function, two conditions must be met. First its mean must be constant as a function of time such that

$$\mathrm{E}[X[nm]] = \mu_X, \tag{3.23}$$

and the autocorrelation function must be invariant to a constant time shift such that

$$R[m_1, m_2] = R[m_1 - \ell, m_2 - \ell] \tag{3.24}$$

where ℓ is an integer. A random process which meets these conditions is said to be *wide sense stationary* (WSS). Stationarity can likewise be defined for higher order moments. A random process which is stationary for all of its moments is said to be *strict sense stationary* (SSS).

If a process is WSS, then its autocorrelation can be defined more simply such that

$$R[\ell] = \mathrm{E}[X[m]X^*[m-\ell]] \tag{3.25}$$

which is a function of only one variable, ℓ. Such a behavior is useful and common to many real processes even if only over short periods of time. Consider the wave height problem. The system may be WSS while the wind speed and direction are constant, but once those change so to may the properties of the wave height, but even stationarity over short time periods can be very useful such as for human speech modeling [43, 94].

Finally, a random process may not be a function of relative time for all time shifts, but it may be stationary for certain shifts. Such a random process is said to be cyclo-stationary where

$$R[m_1, m_2] \neq R[m_1 - \ell, m_2 - \ell] \tag{3.26}$$

but

$$R[m_1, m_2] = R[m_1 - K\ell, m_2 - K\ell] \tag{3.27}$$

and K is some positive valued integer. If the interval K is small enough, it is often practical to simply average the autocorrelation over that interval and to treat the random process as if it were WSS anyway realizing

$$R[\ell] = \frac{1}{K}\sum_{k=0}^{K-1} \mathrm{E}[X[m+k]X^*[m+k-\ell,]]\,. \tag{3.28}$$

Such an approach is used for the random processes in Chapter 5.

3.2.2 Power Spectral Density (PSD)

For any discrete function, whether it be something deterministic or the random processes considered here, spectral quantities are evaluated via the *discrete-time Fourier transform* (DTFT) which is defined

$$S(f) = \sum_{m=-\infty}^{\infty} x[m]\exp(-j2\pi fm) \quad -1/2 < f < 1/2 \tag{3.29}$$

where the frequency variable f is in units of cycles per sampling interval.

The PSD of a random process describes the average power of the process at a particular frequency. Since random processes are assumed to be infinite in time evaluating the PSD of a process given the definition of the DTFT requires taking a limit such that the PSD becomes

$$S(f) = \lim_{M\to\infty} \mathrm{E}\left[\frac{1}{M}\left|\sum_{n=1}^{M} x[m]\exp(-j2\pi fm)\right|^2\right] \tag{3.30}$$

However, hidden within this definition of the PSD is that the PSD is actually a transformation of the autocorrelation function. Expanding 3.30 yields

$$S(f) = \lim_{M\to\infty} \frac{1}{M} \sum_{m_1,m_2=1}^{M} \mathrm{E}\left[x[m_1]x^*[m_2]\right]\exp(-j2\pi f(m_1-m_2)) \tag{3.31}$$

where the expectation has been moved inside the summations by its linearity and the expectation term is the definition of the correlation function for a given m_1 and m_2. However, this prevents an issue. The frequency term is a function of only the difference between m_1 and m_2 meaning the process could have different spectral characteristics at different absolute times, but (3.31) would simply average them all. Consequently, the PSD is usually only considered for WSS processes, cyclo-stationary processes, or processes which are stationary for some meaningful amount of time

where 'meaningful' is context dependent. Thus, assuming stationarity (3.31) becomes

$$S(f) = \lim_{M\to\infty} \frac{1}{M} \sum_{m_1,m_2=1}^{M} R[\ell] \exp(-j2\pi f\ell) = \lim_{M\to\infty} \frac{1}{M} \sum_{\ell=-(M-1)}^{M-1} (M-|\ell|) R[\ell] \exp(-j2\pi f\ell) \tag{3.32}$$

where in the final part (3.32) like terms from the double summation were consolidated. Finally, (3.32) can be further rewritten such that

$$S(f) = \sum_{\ell=-\infty}^{\infty} R[\ell] \exp(-j2\pi f\ell) + \lim_{M\to\infty} \frac{1}{M} \sum_{\ell=-(M-1)}^{M-1} |\ell| R[\ell] \exp(-j2\pi f\ell) \tag{3.33}$$

where, under the condition that $|\ell|R[\ell]$ decays quickly enough, the limit term in (3.33) approaches zero such that

$$S(f) = \sum_{\ell=-\infty}^{\infty} R[\ell] \exp(-j2\pi f\ell) \tag{3.34}$$

which is simply the DTFT of the autocorrelation function. Likewise, a cross-spectral density between two jointly stationary random processes can be defined such that

$$S_{XY}(f) = \sum_{\ell=-\infty}^{\infty} R_{XY}[\ell] \exp(-j2\pi f\ell) \tag{3.35}$$

More details on this formulation of the PSD and the autocorrelation above can be found in [43, 91, 92, 95].

3.2.3 Estimation

The previous subsections defined the mathematical tools used to examine random processes, but practically speaking these quantities are typically unknown in advance and the entire goal is to determine these properties. Consequently, they must be estimated from a finite amount of recorded data. 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 autocor-

relation 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 in adaptive filtering [43, 94]. 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.

3.2.3.1 Random Variable Estimation

Estimation is easiest to understand from the viewpoint of estimating the properties of a single random variable. Consider some random variable, X, with some unknown distribution, but its mean (first moment) needs to be determined. To do so, several outcomes are pulled from the random variable and are denoted as, x_n, where n is the index of the outcome sample. Estimating the mean of this random variable can be performed by taking the sample average of the independent samples such that

$$\hat{\mu}_X = \frac{1}{N}\sum_{n=1}^{N} x_n \tag{3.36}$$

where the hat of $\hat{\mu}_X$ indicates this is an estimate of the expected value, μ_X, as defined in (3.23). At a basic level, an estimator is an empirical tool, while the expectation operator is an analytical tool. In (3.23), the expectation operator uses known properties of the random variable to mathematically determine the first moment. (3.36) however assumes nothing about the random variable and then uses actual, measured sample values to approximate the moment and so there will inevitably be some error in the estimate. Perhaps most confusingly, this implies the estimator itself is also a random variable with its own mean and variance. Consequently, the expectation operator can be used to analytically determine the properties of the estimator itself. Applying this to (3.36) realizes

$$\mathrm{E}[\,\hat{\mu}_X\,] = \frac{\mathrm{E}[\sum_{n=1}^{N} x_n]}{N} = \frac{N\mu_X}{N} = \mu_X. \tag{3.37}$$

Intuitively, (3.37) indicates that (3.36) should approach exactly the mean value of the random variable, but it does not reveal how quickly it will approach that value given some number of

samples, N. To do that, it is necessary to evaluate the variance of (3.36) which is by definition

$$\sigma^2_{\hat{\mu}_X} = \mathrm{E}\left[\left| \frac{1}{N} \sum_{n=1}^{N} x_n \right|^2 \right] - \left| \mathrm{E}\left[\frac{1}{N} \sum_{n=1}^{N} x_n \right] \right|^2 \tag{3.38}$$

Expanding the first term and utilizing the result of (3.37) to evaluate the second terms yields

$$\sigma^2_{\hat{\mu}_X} = \frac{1}{N^2} \mathrm{E}\left[\sum_{n_1,n_2=1}^{N} x_{n_1} x^*_{n_2} \right] - |\mu_X|^2. \tag{3.39}$$

For each term in the double summation where $n_1 = n_2$, the result is $\sigma^2_X + |\mu_X|^2$, but since the samples are independent of each other, where $n_1 \neq n_2$ the result is simply $|\mu_X|^2$. Thus, (3.39) simplifies to

$$\sigma^2_{\hat{\mu}_X} = \frac{1}{N} \sigma^2_X \tag{3.40}$$

where the variance of the estimator is a scaled version of the variance of the random variable itself. As N is increased, the variance of the estimator is decreased and better and better estimates of the first moment are realized.

However, how helpful is this if σ^2_X is also unknown? The variance can also be estimated although doing so is somewhat more complicated. A variance estimator can be defined such that

$$\hat{\sigma}^2_X = \frac{1}{N-1} \sum_{n=1}^{N} |x_n - \hat{\mu}_X|^2 \tag{3.41}$$

where $\hat{\mu}_X$ is the same sample mean defined in (3.36). Taking the expectation of this estimator realizes

$$\mathrm{E}[\,\hat{\sigma}^2_X\,] = \frac{1}{N-1} \sum_{n=1}^{N} \mathrm{E}[\,|x_n - \hat{\mu}_X|^2\,] = \frac{1}{N-1} \sum_{n=1}^{N} \mathrm{E}[x_n x^*_n - x_n \hat{\mu}^*_X - x^*_n \hat{\mu}_X + \hat{\mu}_X \hat{\mu}^*_X\,]. \tag{3.42}$$

Since the expectation is a linear operator it can be applied to each term individually such that (3.42)

becomes

$$\mathrm{E}[\ \hat{\sigma}_X^2\ ] = \frac{N}{N-1}\left((\sigma_X^2 + |\mu_X|^2) - 2(\frac{1}{N}\sigma_X^2 + |\mu_X|^2) + (\frac{1}{N}\sigma_X^2 + |\mu_X|^2)\right) = \sigma_X^2 \tag{3.43}$$

where the estimator will approach the random variable's variance. Unfortunately, determining the variance of (3.41) involves the fourth moment of the random variable. Nevertheless, the effectiveness of second order estimators is an important topic especially as it pertains to random processes and to the waveforms designed in this work.

3.2.3.2 Autocorrelation Estimation

Often time the focus of estimating autocorrelation functions is on WSS processes which is reasonable. WSS processes have some convenient properties and are found in many practical situations. However, for the purposes of the content of Chapter 4, it is important to specifically address how the correlation function of a non-stationary process can be estimated. To do so, two things are necessary: independent sample functions of the process and an absolute time reference. Consider the random walk sample functions shown in Table 3.2. Each sample function is independent of the others and has a well defined starting point. Thus, a given sample term in the correlation function can be evaluated such that

$$\hat{R}[m_1, m_2] = \frac{1}{N}\sum_{n=1}^{N} x_n[m_1]x_n^*[m_2] \tag{3.44}$$

where n references a given sample function and m_1 and m_2 refer to specific discrete times within those sample functions. The independence of the different sample functions makes this possible. It is important to note as well that for non-stationary process the absolute values of m_1 and m_2 are important since in general $R[m_1, m_2] \neq R[m_1 + k, m_2 + k]$ where k is an integer. In other words, for each sample function there is only one available sample for each m_1 and m_2. Effectively, the estimation in (3.44) is identical to that of a single random variable. The primary caveat is

maintaining a consistent reference time. For this estimator, the expectation is

$$\mathrm{E}[\hat{R}[m_1,m_2]] = \frac{1}{N}\sum_{n=1}^{N}\mathrm{E}[x_n[m_1]x_n^*[m_2]] = \frac{NR[m_1,m_2]}{N} = R[m_1,m_2]. \tag{3.45}$$

For a WSS process evaluating the autocorrelation can technically be performed in the same way. However, since for a WSS $R[m_1,m_2] = R_{XX}[m_1+k,m_2+k]$, each sample function technically provides an infinite number of samples for each relative time value. Consequently, an entire autocorrelation function can be estimated to arbitrary precision with a single sample function. Although, WSS does not guarantee that each sample function has the same autocorrelation. It only guarantees that the autocorrelations of each sample function are themselves a function of only relative time. This distinction is made by examining the difference between time averages and ensemble averages. First, consider the ensemble average autocorrelation of a WSS discrete stochastic process which is defined

$$R[\ell] = \mathrm{E}[X[m]X[m+\ell]] \tag{3.46}$$

The ensemble average as defined by the expectation operator is the average autocorrelation across all possible sample functions, but not every sample function will necessarily have the same autocorrelation. Alternatively, the time average estimate of the autocorrelation for a given sample function is defined

$$\hat{R}[\ell] = \lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}x[m]x^*[m+\ell] \tag{3.47}$$

Then, if the time average estimate approaches the ensemble average in a squared error sense as $M\to\infty$, then the process is said to be *ergodic* in the autocorrelation. Stated succinctly

$$R[\ell] = \mathrm{E}[X[m]X[m+\ell]] = \lim_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}x[m]x^*[m+\ell] \tag{3.48}$$

where $x[m]$ is some sample function of the ergodic process $X[m]$ [40]. A similar statement can be made for the mean.

Thus, given M samples of a discrete ergodic process, the lags of the autocorrelation from $1-M$

to $M-1$ can be estimated from a time average such that

$$\hat{R}[\ell] = \frac{1}{M-|\ell|} \begin{cases} \sum_{k=1+\ell}^{M} x[k]x^*[k-\ell] & 0 \leq \ell < M \\ \sum_{k=1}^{M+\ell} x[k]x^*[k-\ell] & -M < \ell < 0 \end{cases} \tag{3.49}$$

where the factor out front normalizes each lag to the number of samples being averaged. (3.49) is known as the standard unbiased autocorrelation sequence (ACS) estimate where ACS is another name for the discrete autocorrelation function. However, the normalization term in (3.49) introduces some issues. First of all for a total of M samples, for a lag of $|\ell| = 1$ there are $M-1$ unique products to average. For $|\ell| = 2$, there are $M-2$ unique products to average. For $|\ell| = M-1$ there is only one product to work will. Consequently, the larger lags will possess a much greater degree of variance compared to small lags and vary erratically. Secondly, (3.49) is not guaranteed to be postive semidefinite such that a PSD estimate derived from (3.49) could be negative valued at certain frequencies [40, 43, 91, 92, 95].

These concerns can be addressed by only normalizing by the total number of samples such that

$$\hat{R}[\ell] = \frac{1}{M} \begin{cases} \sum_{k=1+\ell}^{M} x[k]x^*[k-\ell] & 0 \leq \ell < M \\ \sum_{k=1}^{M+\ell} x[k]x^*[k-\ell] & -M < \ell < 0 \end{cases}. \tag{3.50}$$

which is known as the standard biased ACS estimate. Normalizing in this fashion diminishes the impact of the erratically varying higher lag terms and is guaranteed to be postive semi-definite such that any PSD estimate obtained from (3.50) is guaranteed to be positive. For these reasons, (3.50) is more commonly used [43].

However, as the name implies, (3.50) introduces what is known as a bias. Consider the expectation of (3.50). Each autocorrelation lag becomes scaled by its own number of terms in the summation such that

$$\mathrm{E}[\hat{R}[\ell]] = \left(1 - \frac{|\ell|}{M}\right) R[\ell] \quad -M < \ell < M. \tag{3.51}$$

the estimator converges to a windowed version of the autocorrelation function rather than simply

the autocorrelation itself. This windowing effect is perhaps more important when considering the spectral estimation techniques described in the next section.

3.2.3.3 Power Spectral Density Estimation

The most straightforward way to estimate the PSD is to simply use the definition. Assuming $X[m]$ is WSS and given M data samples, the definition of the PSD given in (3.30) becomes

$$\hat{S}(f) = \frac{1}{M}\left|\sum_{m=1}^{M} x[m]\exp(-j2\pi fm)\right|^2 \tag{3.52}$$

which is known as the *periodogram spectral estimator*. As a method the periodogram has been used to find periodicity in time series data for well over 100 years at this point [96, 97].

The behavior of the periodogram can be examined by expanding the magnitude squared term such that (3.52) becomes

$$\hat{S}(f) = \frac{1}{M}\sum_{m_1,m_2=1}^{M} x[m_1]x^*[m_2]\exp(-j2\pi f(m_1-m_2)) \tag{3.53}$$

Since the process is assumed to be stationary only the difference between m_1 and m_2 is relevant to the estimation process combining like terms in (3.53) allows for a single summation described by the variable ℓ such that

$$\hat{S}(f) = \frac{1}{M}\sum_{\ell=-(M-1)}^{M-1}\exp(-j2\pi f\ell)\begin{cases}\sum_{k=1+\ell}^{M} x[k]x^*[k-\ell] & 0\le\ell<M\\ \sum_{k=1}^{M+\ell} x[k]x^*[k-\ell] & -M<\ell<0\end{cases}, \tag{3.54}$$

which is simply the DTFT of the standard biased autocorrelation estimate. Written more succinctly, (3.54) becomes

$$\hat{S}(f) = \sum_{\ell=-(M-1)}^{M-1}\hat{R}[\ell]\exp(-j2\pi f\ell). \tag{3.55}$$

However, the PSD estimate based on the autocorrelation estimate is known as a correlogram even though it is equivalent to the periodogram in its result [98]. Unfortunately, without modification

the periodogram (or correlogram) estimate actually does a poor job of estimating the spectra for because of bias and variance.

First, the expectation of (3.55) yields

$$\mathrm{E}[\hat{S}(f)] = \sum_{\ell=-(M-1)}^{M-1} \mathrm{E}[\hat{R}[\ell]] \exp(-j2\pi f\ell) = \sum_{\ell=-(M-1)}^{M-1} \left(1 - \frac{|\ell|}{M}\right) R[\ell] \exp(-j2\pi f\ell) \tag{3.56}$$

where the windowing of the standard biased autocorrelation estimate will now impact the PSD estimate. This biasing window is notated such that

$$w_{\mathrm{B}}[m] = \begin{cases} \left(1 - \frac{|m|}{M}\right) & |m| < M \\ 0 & \text{otherwise} \end{cases}, \tag{3.57}$$

where m is an integer. (3.57) is known as a *triangular* or *Bartlett* window. It should be noted that the periodogram and the correlogram are both based on the DTFT which sums over $(-\infty, \infty)$, but with only a finite amount of data from which to estimate, these sums are necessarily truncated to the relevant intervals in (3.52) and (3.55). Regardless of this truncation, the properties of the DTFT still apply. Consequently, the product of the autocorrelation and its bias is equivalent to a convolution in the frequency domain such that.

$$\mathrm{E}[\hat{S}(f)] = \int_{-\frac{1}{2}}^{\frac{1}{2}} S(\theta) W_{\mathrm{B}}(f - \theta) d\theta \tag{3.58}$$

For the Bartlett window, the DTFT can be calculated such that

$$W_{\mathrm{B}}(f) = \frac{1}{M} \left(\frac{\sin(\pi f M/2)}{\sin(\pi f/2)} \right)^2. \tag{3.59}$$

Thus, even with a perfect, zero variance estimate of the periodogram, the resulting PSD estimate itself would still be biased by a convolution with the window, (3.59), as pictured in Fig. 3.1. This convolution primarily results in two effects on the spectral estimate. First, the mainlobe of the window has a smoothing effect on the estimate such that the frequency resolution is limited. The

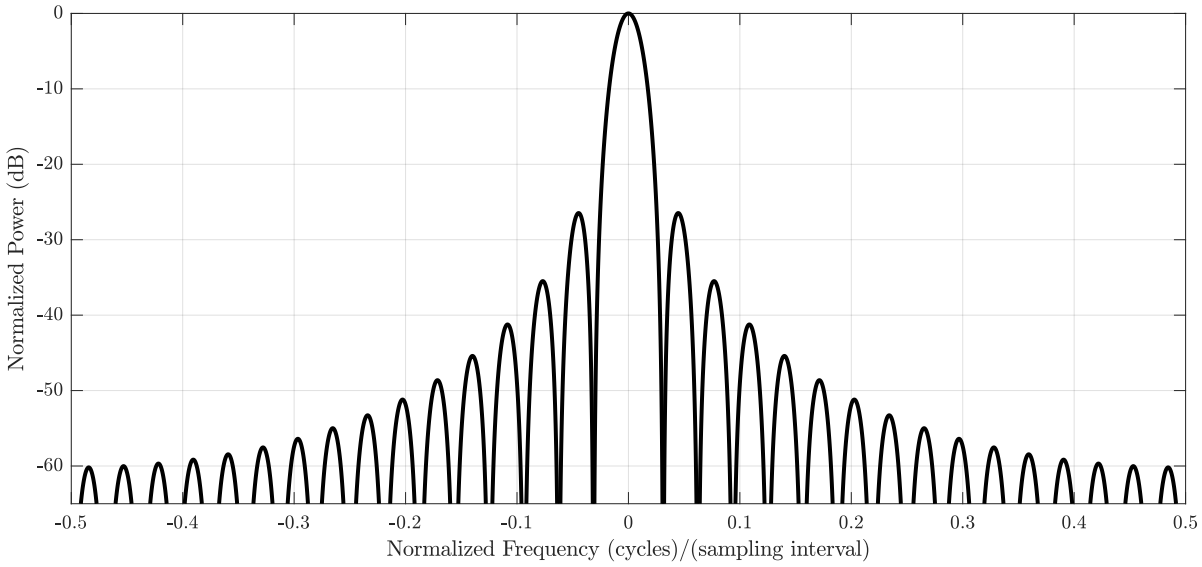

Figure 3.1: The normalized Bartlett spectrum on a dB scale with $M = 32$ as defined by (3.59)

sidelobes on the other hand result in what is known as spectral leakage where power from higher portions of the spectrum are smeared into the lower power portions of the roll-off region. The second periodogram problem, variance, is more difficult to evaluate in general since it depends on the fourth order moments of a given random process. (Although, to foreshadow, this can be evaluated for the WGFs optimized in Chapters 4 and 5.) Still, although such a proof is beyond the scope of this work, it suffices to say that the periodogram is an *inconsistent estimator*. In other words, even as the size of the periodogram, M, is allowed to approach infinity the variance of the periodogram will not approach zero. Intuitively, this effect is evident from the biased autocorrelation estimate. Regardless, of how large M becomes, the higher lags as ℓ approaches M in the autocorrelation estimate will always posses a high degree of variance due to the limited number of terms being averaged. A more complete discussion on this topic can be found in [43, 99].

Due to the bias and variance issues inherent to the periodogram, over the decades various techniques have been introduced to create trade-offs between these two effects. One such method involves the use of windows. Rather than simply accepting the Bartlett window, windowing techniques introduce their own windows with more desirable characteristics to the data. To do so, the

Blackman-Tukey spectral estimator is defined such that [98]

$$\hat{S}_{\mathrm{BT}}(f) = \sum_{\ell=-(K-1)}^{K-1} w[\ell]\hat{r}[\ell]\exp(-j2\pi f\ell) \tag{3.60}$$

where $w[\ell]$ is an even function where usually $w[0] = 1$ and $w[\ell] = 0$ for $|\ell| \geq K$ and $w[\ell]$ decays smoothly to zero with increasing $|\ell|$ and $K < M$. Since this window weights the lags of the auto-correlation estimate, it is known as a *lag window*. While this is a time domain window, its impact can also be written in frequency domain such that

$$\hat{S}_{\mathrm{BT}}(f) = \hat{S}(f) * W(f) \tag{3.61}$$

which shows that the Blackman-Tukey spectral estimate is itself a biased form of the periodogram estimate and $*$ denotes convolution. The key point however is that $w[\ell]$ can be designed to achieve a useful trade-off in the estimate between variance, resolution, and spectral leakage. For example defining $w[\ell]$ to be a rectangular window of length K where $K < M$ will produce a lower variance estimate since such a window would remove the highest variance members of $\hat{r}[\ell]$. However, this comes at the cost of resolution. A full discussion on this topic is beyond the scope of this work, but it can be shown that in general if the spectral resolution of the periodogram is on the order of $1/M$, then the spectral resolution of the Blackman-Tukey spectral estimate will be on the order of $1/K$. Likewise if the variance of the periodogram is normalized to 1 then the Blackman-Tukey spectral estimate variance will be on the order of K/M. Still, this only addresses the windows length. The choice of the window's shape represents a trade-off between smearing and leakage effects [43]. With these goals in mind, over the years a multitude of windows have been proposed some of which are likewise used in radar processing to mitigate autocorrelation sidelobes for LFM waveforms. Some of the more common windows include Hanning, Hamming, Kaiser, and Blackman windows. Much more information on these and other windows can be found in [41, 43, 91, 92, 100, 101].

Fundamentally, the Blackman-Tukey spectral estimate treats the M available samples as a single record to be evaluated all at once. In contrast, there are methods which break the data into

multiple chunks, process these chucks individually, and then average those results to obtain lower variance, poorer resolution estimates than the periodogram estimate alone.

The simplest of these approaches is the Bartlett method [102, 103] where a data set of M samples is split into N data sets of $N = M/W$ such that each set is W samples long. A periodogram (or correlogram) is then evaluated for each subset and then these results are averaged. The Bartlett method can be written such that the spectral estimate is

$$\hat{S}_{\mathrm{B}}(f) = \frac{1}{W} \sum_{n=1}^{N} \hat{S}_n(f) \tag{3.62}$$

where the subscript n denotes each member of the N data sets and $\hat{S}_n(f)$ is the periodogram estimate of the nth member. The frequency resolution of this method is on the order of the resolution of the individual periodograms which is $1/W$ while the variance is reduced by a factor of N. This method is essentially how spectral analysis is performed in Chapter 5. In a similar way the Welch method divides the data record into segments, but it differs in that it also allows these segments to overlap each other and for these segments to be windowed [104].

Overall, the estimation of the PSD of a given random process is a deceptively difficult task. On the surface it seems like it should be as simple as taking a partial DTFT of the data and moving on but as has been discussed, this periodogram estimator is a poor estimator in both its bias a variance. This problem has motivated the development of numerous refined spectral estimation techniques many of which were at least introduced here. For this work, the most relevant of which is the Bartlett method which will be used to process noise-like, CW radar signals.

3.3 Stochastic Processes as Radar Signals

The beginning of this chapter introduced the admittedly vague concept of the WGF, but with the tools of Sections 3.1 and 3.2, it is now possible to characterize WGFs in the context of radar processing. For example, if a WGF is meant to produce pulsed waveforms, then the underlying stochastic process must be time limited. If it produces CW waveforms, it must be unlimited in time.

If it is meant to produce FM waveforms, then it must be constant in amplitude. Whether or not these characteristics are true for a given WGF will have an impact on how the WGF is evaluated. This is especially important when considering the difference between a pulsed WGF versus a CW one. In any case however, evaluating the WGF requires adapting the general stochastic process autocorrelation and power spectral density concepts to their specific meanings in a radar context.

3.3.1 Pulsed Stochastic Processes as Radar Signals

Thus far stochastic processes have been defined as infinite in time, but for a WGF to produce pulsed radar waveforms, it must be time-limited. One solution to this problem is to apply a window function such as a rectangular window to an infinite in duration random process to force it to be time limited. However, the approach taken here is to define the stochastic process to be time limited by definition. Defining a stochastic process in this way is unusual and it necessarily means that these WGFs are not stationary. Clearly, if the process is strictly 0 valued for some time and then allowed to vary, its autocorrelation function will also vary as a function of absolute time. Consequently, when examining the pulsed WGFs defined here, the focus is on ensemble averages.

With this in mind, consider a complex, time-limited stochastic process which is meant to be used as a WGF for pulsed radar waveform applications such that

$$S_{\mathrm{p}}[m] = \begin{cases} S[m] & m = 1,2,\cdots,M \\ 0 & \text{otherwise} \end{cases} . \tag{3.63}$$

where the subscript 'p' indicates the pulsed nature of the process. During the interval $[1,2,\ldots,M]$ the process is stochastic, but outside of this interval it is deterministically zero valued. Because of this, the autocorrelation function outside of this interval is also zero valued such that

$$R_{\mathrm{p}}[m_1,m_2] = \begin{cases} R[m_1,m_2] & 1 \leq m_1 \leq M,\ 1 \leq m_2 \leq M \\ 0 & \text{otherwise} \end{cases} , \tag{3.64}$$

Since the process is not stationary, each sample function only provides one sample for each pair of m_1 and m_2 such that (3.64) can only be estimated through ensemble averages over numerous sample functions in the manner discussed in Section 3.2.3.2. However, in the context of radar signal processing, we do not necessarily care about finding (3.64) specifically. Instead, we care about finding the matched filter response.

3.3.1.1 Pulsed Matched Filter Estimation

Recall the matched filter defined in 2.1.2. Confusingly, in the context of radar the term "autocorrelation function" is often used interchangeably with the "matched filter response". However, this radar defined autocorrelation is not equivalent to the autocorrelation function in the random process sense. To address this ambiguity, the autocorrelation in the context of radar will be referred to specifically as the matched filter response although this terminology is relaxed in later chapters. With this in mind, in discrete terms the matched filter is defined such that

$$r[\ell] = \sum_{m=-\infty}^{\infty} x[m]x^*[m-\ell] \tag{3.65}$$

which is simply the convolution of the signal $x[m]$ with $x^*[-m]$. For a given sample function of the pulsed random process defined in (3.63) specifically, the matched filter response can be written as

$$\hat{r}_{\mathrm{p}}[\ell, N=1] = \frac{1}{M}\begin{cases} \sum_{m=1+\ell}^{M} s_n[m]s_n^*[m-\ell] & 0 \le \ell < M \\ \sum_{m=1}^{M+\ell} s_n[m]s_n^*[m-\ell] & -M < \ell < 0 \end{cases} \tag{3.66}$$

which is zero valued elsewhere. For completeness, (3.66) has been written for all non-zero values of the matched filter estimate, but for future convenience the matched filter estimate will be written only in terms of positive ℓ since the matched filter estimate is conjugate symmetric such that $\hat{r}_{\mathrm{p}}[\ell, N] = \hat{r}_{\mathrm{p}}^*[-\ell, N]$.

(3.66) emphasizes the time limited nature of the signal and the subscript n indicates the nth sample function of (3.63) where n is arbitrary. The estimator is written as a function of N to

emphasize the concept of coherent integration across unique sample functions or waveforms from a noise-like radar perspective. Specifically in (3.66), $N = 1$ shows the estimate is being made with a single sample function, but generally the ensemble average matched filter response can be estimated using an arbitrary number of sample functions such that (3.66) becomes

$$\hat{r}_{\mathrm{p}}[\ell, N] = \frac{1}{N}\sum_{n=1}^{N}\hat{r}_{\mathrm{p},n}[\ell, 1] = \frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1+\ell}^{M} s_n[m]s_n^*[m-\ell] \quad 0 \le \ell < M \tag{3.67}$$

where the N sample function estimate represents an average over N single sample function matched filter estimates.

Interestingly, even though the process is clearly not stationary or ergodic due to its pulsed definition, the matched filter response treats it as if it were in a way. To see this, the expectation of (3.67) yields.

$$r_{\mathrm{p}}[\ell] = \mathrm{E}[\hat{r}_{\mathrm{p}}[\ell, N]] = \frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1+\ell}^{M}\mathrm{E}[s_n[m]s_n^*[m-\ell]] \quad 0 \le \ell < M \tag{3.68}$$

where evaluating the expectation realizes

$$r_{\mathrm{p}}[\ell] = \mathrm{E}[\hat{r}_{\mathrm{p}}[\ell, N]] = \frac{1}{M}\begin{cases} \sum\limits_{m=1+\ell}^{M} R_{\mathrm{p}}[m, m-\ell] & 0 \le \ell < M \\ \sum\limits_{m=1}^{M+\ell} R_{\mathrm{p}}[m, m-\ell] & -M < \ell < 0 \end{cases}. \tag{3.69}$$

such that the matched filter response given a lag of ℓ is actually the average correlation of the process's autocorrelation function for the same lag, ℓ.

To put $r_{\mathrm{p}}[\ell]$ into perspective, as discussed in Section 2.2.4 for noise-like waveforms, coherent integration is meant to lower the sidelobe response. In (3.67), increasing amounts of coherent integration is used to improve the estimate of the process's matched filter response and (3.69) shows that the limit of this coherent integration is exactly the expected matched filter response of the process itself, $r_{\mathrm{p}}[\ell]$. Consequently, $r_{\mathrm{p}}[\ell]$ represents just how low those sidelobes can be driven via coherent integration for a given $S_{\mathrm{p}}[m]$. Thus, if $r_p[\ell]$ is already achieved, then further coherent

integration via (3.67) (by increasing N) will result in no further lowering of the sidelobe level. In general, this ideal matched filter response, $r_p[\ell]$, is a goal that can never be truly reached as doing so would require coherently integrating an impractical or even infinite number of waveforms. The question then becomes how quickly does coherent integration lower the sidelobes and approach $r_{\mathrm{p}}[\ell]$. What if there are only 100 sample functions to coherently integrate? How the matched filter response can be expected to behave given exactly N coherent integrations is described by the second moment of the estimator,

$$\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,N]\right|^2\right], \tag{3.70}$$

which in a radar context can be described as the *Expected RMS Autocorrelation Function* or the *Expected RMS Matched Filter Response*. These names describe how (3.70) itself can be estimated, where instead of coherently averaging the matched filter estimates, they are instead averaged via RMS. (Although, it is technically just a mean square method since no square root operation has been applied.)

In general evaluating the second moment of an estimator can be difficult, however since this estimator looks across different sample functions rather than as a time averaging operation over a single sample function, evaluating (3.70) is fairly straightforward as the different sample functions are independent of each other. With this in mind, (3.70) becomes

$$\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,N]\right|^2\right] = \frac{1}{N^2}\mathrm{E}\left[\left|\sum_{n=1}^{N}\hat{r}_{\mathrm{p},n}[\ell,1]\right|^2\right]. \tag{3.71}$$

The magnitude squared operation over the sum can be expanded such that (3.71) is now

$$\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,N]\right|^2\right] = \frac{1}{N^2}\sum_{n_1,n_2=1}^{N}\mathrm{E}\left[\hat{r}_{\mathrm{p},n_1}[\ell,1]\hat{r}^{*}_{\mathrm{p},n_2}[\ell,1]\right]. \tag{3.72}$$

Since the matched filter responses of different sample functions are independent of each other, the expectation of their product can be written as the product of their expectations. Thus, where $n_1 \neq n_2$ the terms in the double summation become the magnitude squared of the expected matched filter

response $\left|r_{\mathrm{p}}[\ell]\right|^2$. There are $N(N-1)$ of these terms. Otherwise when $n_1 = n_2$, the expectation represents the second moment of the matched filter estimate where $N = 1$. There are N of these terms. Consequently, (3.72) can be simplified such that

$$\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,N]\right|^2\right] = \frac{1}{N}\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] + \frac{N-1}{N}\left|r_{\mathrm{p}}[\ell]\right|^2. \tag{3.73}$$

In (3.73) As N increases, the contribution of the second moment of the single sample function expected matched filter decreases while the contribution of the magnitude squared of the ideal matched filter response approaches one. Thus, as expected, the second moment of the estimator approaches the ideal response with increasing N.

However, this moment is not yet in terms of its variance plus the magnitude squared of its mean. To find this form, the variance of the estimator given a single sample function is by definition.

$$\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1] - r_{\mathrm{p}}[\ell]\right|^2\right] = \mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] - \left|r_{\mathrm{p}}[\ell]\right|^2 \qquad 0 \leq \ell < M \tag{3.74}$$

Inserting this result into (3.73) realizes

$$\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,N]\right|^2\right] = \frac{1}{N}\left(\mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] - \left|r_{\mathrm{p}}[\ell]\right|^2\right) + \left|r_{\mathrm{p}}[\ell]\right|^2 \tag{3.75}$$

where the first term on the right hand side is the variance and the second term is the magnitude squared of the mean. Clearly, as N increases the variance decreases.

To demonstrate the contribution of the variance and mean terms in (3.75) individually, consider some pulsed WGF whose sample functions are constant in energy. For this WGF, Fig. 3.2 plots the analytically calculated RMS matched filter result itself as well as its mean and variance separately for $N = 10^4$ coherent integrations. From a radar perspective, the black trace represents the average matched filter response given 10^4 coherent integrations, while the blue trace represents the ideal result with further coherent integration. The orange trace represents the variance term. Where the variance dominates the response, $\ell > 120$, the RMS response largely follows the variance trace.

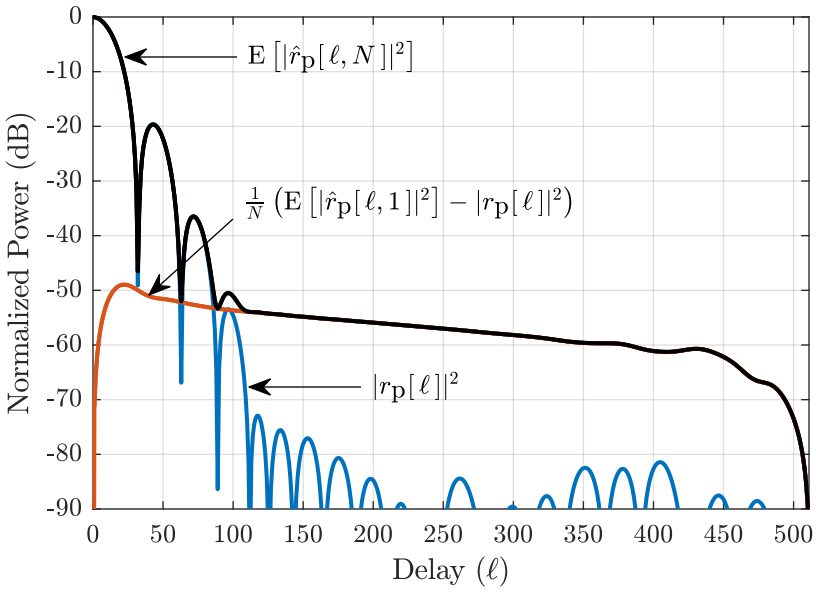

Figure 3.2: The analytically calculated expected RMS matched filter response of an example WGF for $N = 10^4$ sample functions and its mean and variance as defined in (3.75)

Where the mean term dominates the response, $\ell < 90$, the RMS result largely follows the mean trace. Where the variance is roughly equivalent to the mean $90 \leq \ell \leq 120$, the RMS response deviates from either of them as it is the sum of both. Near $\ell = 0$ the variance races towards zero. Since the energy in each sample function is constant, the peak of the matched filter estimate is identical for every sample function such that the peak experiences no variance.

To demonstrate the impact of coherent integration, Fig. 3.3 plots the expected RMS matched filter response, (3.70), for $N = [10^0,\ 10^2,\ 10^4,\ 10^7]$ along with the expected matched filter response, $r_{\mathrm{p}}[\ell]$ which is now plotted in black. Away from the mainlobe for lags greater than about 120, $|r_{\mathrm{p}}[\ell]|^2$ is fairly small on the order of -80 dB. Because of this, for $N = [10^0,\ 10^2,\ 10^4]$ the matched filter estimate is dominated by the variance term in (3.73). Consequently, when going from $N = 10^0$ to $N = 10^2$ to $N = 10^4$ there is roughly a 20 dB decrease in the expected sidelobe level of the estimate for each step. It is not until $N = 10^7$ that the impact of the expected matched filter response starts to becomes apparent in these regions. For the near in sidelobes however for roughly $\ell < 120$, the expected matched filter response is already evident in the estimate at $N = 10^2$

owing to its relatively high power.

The behavior apparent in Fig. 3.3 allows for a distinction to be made between the variance component and the mean component of (3.73). The sidelobes contributed by the mean component which can also be called the the expected matched filter response or the ideal matched filter response are *persistent*. They cannot be mitigated through coherent integration as they are a fundamental property of a given WGF as defined by its respective $r_{\mathrm{p}}[\ell]$. However, The variance term in (3.73) represents the *transient* component of the sidelobes. These are those sidelobes which can be mitigated through coherent integration.

Before moving on, it is important to recognize that the traces for $N = [10^0,\ 10^2,\ 10^4,\ 10^7]$ represent expected values or averages given N coherent integrations. In practice, the empirical sidelobe response from actually integrating instantiated unique matched filter responses will not resemble the smooth lines shown in Fig. 3.3. Instead, the transient sidelobes will vary a fair degree around the same line as can be see in Fig. 3.4 where 10^5 sample functions of the pulsed WGF used in Fig. 3.3 have been, instantiated, match filtered, and coherently integrated. This result is compared to the analytically calculated RMS matched filter response given $N = 10^5$ coherent integrations. In the region where the transient sidelobes are dominant ($\ell > 120$), the empirical estimate of $r_{\mathrm{p}}[\ell]$ varies greatly around the analytical expectation. Which makes sense as these sidelobes are a manifestation of the estimate's variance. However, in the persistent sidelobes dominated region ($\ell < 120$), the analytical expectation and the empirical estimate appear very close on the dB scale since every single sample function possesses the same persistent sidelobe component as determined by the underlying WGF itself. This behavior can alternatively be examined in the frequency domain.

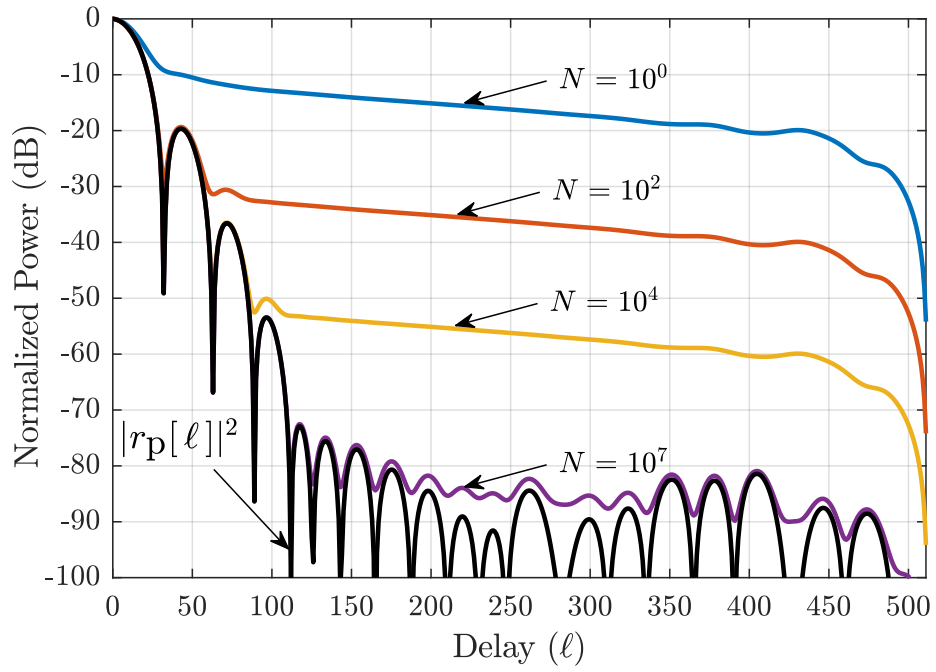

Figure 3.3: The analytically calculated expected RMS matched filter response for an example WGF given varying degrees of coherent integration

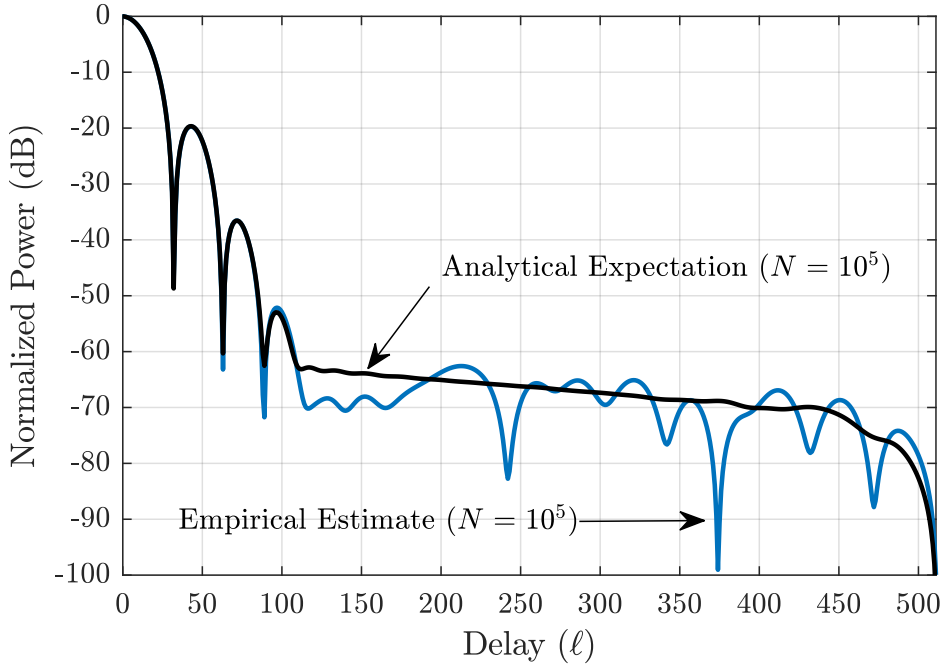

Figure 3.4: The analytically calculated expected RMS matched filter response of an example WGF compared to an empirical estimate of the matched filter given $N = 10^5$ coherent integrations

3.3.1.2 Pulsed PSD Estimation

The estimation of the PSD of (3.63) presents similar issues as with the estimation of the matched filter response. Since (3.63) is not stationary, the typical stochastic process definition of the PSD does not apply here. Instead, for (3.63) as a WGF the PSD represents the average spectral content of the process over all possible sample functions. The most straightforward way to show this is to invoke the autocorrelation definition of the PSD where for a discrete radar signal the PSD is simply the DTFT of the matched filter response. Since the matched filter response has been shown to be an average over the correlation function $R_{\mathrm{p}}[m_1, m_2]$, then so to is the PSD from a WGF perspective.

With this in mind the PSD can be defined relative to the matched filter response such that

$$S_{\mathrm{p}}(f) = \sum_{\ell=-(M-1)}^{M-1} r_{\mathrm{p}}[\ell] \exp(-j2\pi f \ell). \tag{3.76}$$

An estimate of the PSD can then be obtained as the DTFT of the matched filter estimate such that

$$\begin{aligned} \hat{S}_{\mathrm{p}}(f,N) &= \frac{1}{N} \sum_{n=1}^{N} \hat{S}_{\mathrm{p},n}[\ell, 1] \\ &= \frac{1}{N} \sum_{\ell=-(M-1)}^{M-1} \sum_{n=1}^{N} \hat{r}_{\mathrm{p},n}[\ell, 1] \exp(-j2\pi f \ell) \\ &= \sum_{\ell=-(M-1)}^{M-1} \hat{r}_{\mathrm{p}}[\ell, N] \exp(-j2\pi f \ell). \end{aligned} \tag{3.77}$$

where the expectation of (3.77) will clearly result in (3.76). Likewise, the second moment of (3.77) can be evaluated in a similar manner to the matched filter case. Using the same principle of independence across different sample functions, the second moment of the PSD estimator is

$$\mathrm{E}\left[\left| \hat{S}_{\mathrm{p}}(f,N) \right|^2 \right] = \frac{1}{N} \mathrm{E}\left[\left| \hat{S}_{\mathrm{p}}(f,1) \right|^2 \right] + \frac{N-1}{N} \left| S_{\mathrm{p}}(f) \right|^2. \tag{3.78}$$

Again as in the matched filter case, (3.78) can be written in terms of the variance plus a mean term

such that

$$\mathrm{E}\left[\left|\hat{S}_{\mathrm{p}}(f,N)\right|^2\right] = \frac{1}{N}\left(\mathrm{E}\left[\left|\hat{S}_{\mathrm{p}}(f,1)\right|^2\right] - \left|S_{\mathrm{p}}(f)\right|^2\right) + \left|S_{\mathrm{p}}(f)\right|^2. \tag{3.79}$$

As N approaches infinity, the variance term dissipates and (3.79) approaches the magnitude squared of the PSD as expected. (3.78) can be directly related to the mean and variance of the matched filter estimate such that (3.79) becomes

$$\begin{aligned}\mathrm{E}\left[\left|\hat{S}_{\mathrm{p}}(f,N)\right|^2\right] = \frac{1}{N}\left(\mathrm{E}\left[\left|\sum_{\ell=-(M-1)}^{M-1}\hat{r}_{\mathrm{p}}[\ell,1]\exp(-j2\pi f\ell)\right|^2\right] - \left|\sum_{\ell=-(M-1)}^{M-1}r_{\mathrm{p}}[\ell]\exp(-j2\pi f\ell)\right|^2\right) \\ + \left|\sum_{\ell=-(M-1)}^{M-1}r_{\mathrm{p}}[\ell]\exp(-j2\pi f\ell)\right|^2.\end{aligned} \tag{3.80}$$

However, the variance term is somewhat peculiar in this case as it appears to contain the correlation between different correlation estimates. To see this, the variance term can be expanded such that

$$\left(\sum_{\ell_1,\ell_2=-(M-1)}^{M-1}\mathrm{E}\left[\hat{r}_{\mathrm{p}}[\ell_1,1]\hat{r}_{\mathrm{p}}^*[\ell_2,1]\right]\exp(-j2\pi f(\ell_1-\ell_2))\right) - \left|S_{\mathrm{p}}(f)\right|^2. \tag{3.81}$$

where the double summations take the same form as those in (3.53). In (3.53) it was shown that the double summation could be written as a single summation over the elements of the process's autocorrelation function. Here, the initial summations are already over the matched filter response which is akin to the random process autocorrelation function. Using the result in (3.80), another conjugate symmetric function can be defined such that

$$k[\eta] = \begin{cases} \sum\limits_{\ell=\eta-(M-1)}^{M-1}\mathrm{E}[\hat{r}_{\mathrm{p}}[\ell,1]\hat{r}_{\mathrm{p}}^*[\ell-\eta,1]] & 0 \le \eta < 2M-1 \\ \sum\limits_{\ell=-(M-1)}^{M-1+\eta}\mathrm{E}[\hat{r}_{\mathrm{p}}[\ell,1]\hat{r}_{\mathrm{p}}^*[\ell-\eta,1]] & -(2M-1) \le \eta < 0 \end{cases} \tag{3.82}$$

which confusingly, is the $4M-3$ length expected matched filter response of the matched filter

response of $S_{\mathrm{p}}[m]$. Putting (3.82) into (3.79) realizes

$$\mathrm{E}\left[\left|\hat{S}_{\mathrm{p}}(f,N)\right|^2\right] = \frac{1}{N}\left(\sum_{\eta=-(2M-2)}^{2M-2} k[\eta]\exp(-j2\pi f\eta) - \left|S_{\mathrm{p}}(f)\right|^2\right) + \left|S_{\mathrm{p}}(f)\right|^2, \tag{3.83}$$

such that the variance of the PSD estimator depends on the DTFT of the expected matched filter response of the matched filter response.

Finally, the total variance of the PSD estimate can be compared to the total variance of the matched filter estimate such that

$$\sum_{\ell=-(M-1)}^{M-1} \mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] - |r_{\mathrm{p}}[\ell]|^2 = \int_{-1/2}^{1/2} \sum_{\eta=-(2M-2)}^{2M-2} k[\eta]\exp(-j2\pi f\eta) - |S_{\mathrm{p}}(f)|^2 df. \tag{3.84}$$

Beginning with the term containing $k[\eta]$, the summation can be moved outside of the integral such that the integral extends only over the complex exponential term. For $\eta \neq 0$, this integral extends over an integer number of cycles of the exponential such that it is equal to zero realizing

$$\sum_{\ell=-(M-1)}^{M-1} \mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] - |r_{\mathrm{p}}[\ell]|^2 = k[0] - \int_{-1/2}^{1/2} |S_{\mathrm{p}}(f)|^2 df. \tag{3.85}$$

The second term on the right hand side can be expanded such that

$$\sum_{\ell=-(M-1)}^{M-1} \mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] - |r_{\mathrm{p}}[\ell]|^2 = k[0] - \int_{-1/2}^{1/2} \left|\sum_{\ell=-(M-1)}^{M-1} r_{\mathrm{p}}[\ell]\exp(-j2\pi f\ell)\right|^2 df. \tag{3.86}$$

Next, distributing the magnitude squared operation realizes

$$\begin{aligned}\sum_{\ell=-(M-1)}^{M-1} \mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] - |r_{\mathrm{p}}[\ell]|^2 &= k[0] \\ &- \sum_{\ell_1,\ell_2=-(M-1)}^{M-1} r_{\mathrm{p}}[\ell_1] r_{\mathrm{p}}^*[\ell_2] \int_{-1/2}^{1/2} \exp(-j2\pi f(\ell_1-\ell_2)) df.\end{aligned} \tag{3.87}$$

where again the integral is zero except for when $\ell_1 = \ell_2$. Thus, (3.87) becomes

$$\sum_{\ell=-(M-1)}^{M-1} \mathrm{E}\left[\left|\hat{r}_{\mathrm{p}}[\ell,1]\right|^2\right] - |r_{\mathrm{p}}[\ell]|^2 = k[0] - \sum_{\ell=-(M-1)}^{M-1} |r_{\mathrm{p}}[\ell]|^2. \tag{3.88}$$

Based on (3.82), the left most term of (3.88) is equivalent to $k[0]$, which demonstrates that the variance of the matched filter estimate is equivalent to the variance of the spectral estimate. The implications of the variance of spectral estimate can be visualized as a function of N.

As with the matched filter estimation, it is helpful to examine the behavior of the estimator on an example WGF. To do so, consider the same WGF used for Figs. 3.2-3.4. In Fig. 3.5, the three components of (3.79) have been plotted for $N = 1$ where the each term has been square rooted such that they are on the same order as the expected PSD $S_{\mathrm{p}}(f)$. Consequently, the standard deviation is plotted here rather than the variance. In the spectral case, the impact of variance/deviation is much less dramatic than for the expected matched filter. Regardless, as N increases the black trace will approach the blue trace while the orange trace reduces at a rate of $10\log_{10}(N)$.

For the particular WGF function used here each sample function possess the same total energy. In the expected matched filter case this meant that the variance approached zero at the peak of the mainlobe. In the frequency domain however, this means the total power in any individual spectral estimate must be equivalent to the power in the expected PSD such that

$$\int_{-1/2}^{1/2} \hat{S}_{\mathrm{p}}(f,N)df = \int_{-1/2}^{1/2} S_{\mathrm{p}}(f)df \tag{3.89}$$

for any value of N. Consequently, the spectral estimate can only vary above and below $S_{\mathrm{p}}(f)$ such that (3.89) is maintained. Visualizing the impact of the variance in the spectral case, is easier to do by examining empirical estimates of the spectrum as opposed to examining the analytical RMS matched filter response as in Fig. 3.3. In Fig. 3.6, the PSD of the WGF was estimated using $N = [10^0\ 10^2\ 10^4]$. With each increase in the number of sample functions/waveforms used to estimate the PSD, the variance clearly decreases and the estimate approaches the expected PSD, $S_{\mathrm{p}}(f)$.

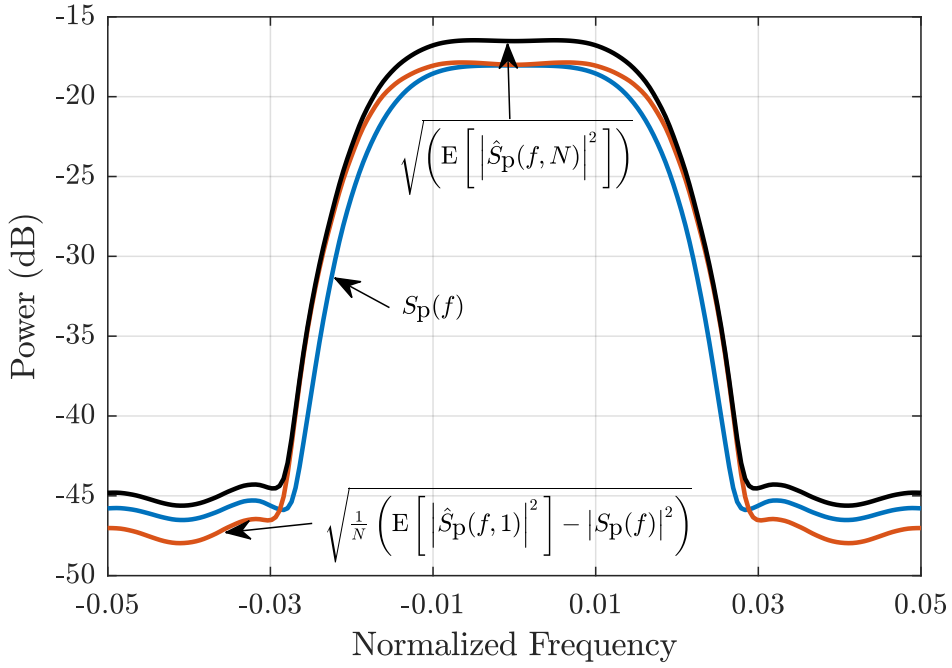

Figure 3.5: The analytically calculated RMS PSD of an example WGF and its mean and variance for $N = 1$ coherent integrations

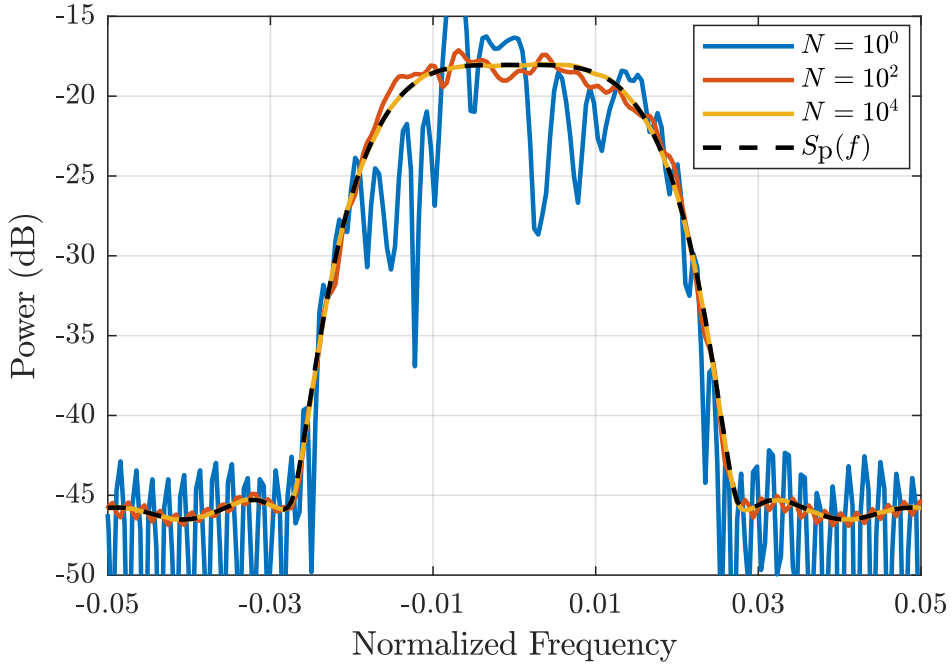

Figure 3.6: Empirical estimates of an example WGF's PSD given varying degrees of coherent integration

Perhaps the most important equation in this entire discussion is simply (3.76) since it demonstrates how the expected matched filter response, $r_{\mathrm{p}}(\ell)$, and the expected PSD, $S_{\mathrm{p}}(f)$, are fundamentally connected. In estimating the matched filter the ideal result is simply $r_{\mathrm{p}}(\ell)$ where further coherent integration will lead to no further decrease in the sidelobes. The sidelobes of $r_{\mathrm{p}}(\ell)$ in turn are a fundamental property of the (3.76). Thus, the expected matched filter response can only ever be as good as the expected PSD allows it to be and vice versa.

3.3.1.3 Pulsed Cross-correlation and Cross-spectral Density Estimation

In addition to understanding the spectral and autocorrelation characteristics of the waveforms produced by (3.63), it is important to understand how the waveforms interact with each other through their cross-correlation and cross-spectral density characteristics. However, in a radar context these terms differ from how they would be used in a random process context where cross-correlation and cross-spectral density refer to the interaction between completely different random processes. Here, these terms refer to the interaction between different sample functions (waveforms) of the same random process (pulsed WGF).

For this analysis it is actually easier to begin with the expected cross-spectral characteristics where it is more straightforward to take advantage of the independence of different sample functions. To do so, the expected cross-spectral density is defined here as the expected product of one sample functions spectrum with the conjugate of another such that

$$C_{\mathrm{p}}(f) = \mathrm{E}\left[\left(\sum_{m=1}^{M} s_a[m]\exp(-j2\pi fm)\right)\left(\sum_{m=1}^{M} s_b[m]\exp(-j2\pi fm)\right)^{*}\right] \tag{3.90}$$

where the subscripts a and b indicate that these are unique sample functions of the process $S_{\mathrm{p}}[f]$. Since the sample functions are independent, the expectation operation can be applied to each spectrum individually such that (3.90) becomes

$$C_{\mathrm{p}}(f) = \left(\sum_{m=1}^{M} \mathrm{E}[s_a[m]]\exp(-j2\pi fm)\right)\left(\sum_{m=1}^{M} \mathrm{E}[s_b[m]]\exp(-j2\pi fm)\right)^{*}. \tag{3.91}$$

Distributing the summations realizes

$$C_{\mathrm{p}}(f) = \sum_{m_1,m_2=1}^{M} \mathrm{E}[s_a[m_1]]\mathrm{E}[s_b^*[m_2]]\exp(-j2\pi f(m_1-m_2)). \tag{3.92}$$

which is similar in form to (3.31), but there is of course no limit operation and the expectation is applied to each term individually rather than as a pair. Regardless, the summation can be rewritten over a single variable, ℓ, in a similar manner to (3.31) such that the expected cross-spectral density is equivalent to the DTFT over the cross-correlation function,

$$C_{\mathrm{p}}(f) = \sum_{\ell=-(M-1)}^{M-1} c_{\mathrm{p}}[\ell]\exp(-j2\pi f\ell). \tag{3.93}$$

where the cross correlation function $c_{\mathrm{p}}[\ell]$ is defined,

$$c_{\mathrm{p}}[\ell] = \frac{1}{M}\begin{cases} \sum_{m=1+\ell}^{M} \mathrm{E}[s_a[m]]\mathrm{E}[s_b^*[m-\ell]] & 0 \le \ell < M \\ \sum_{m=1}^{M+\ell} \mathrm{E}[s_a[m]]\mathrm{E}[s_b^*[m-\ell]] & -M < \ell < 0 \end{cases} \tag{3.94}$$

In terms of measuring these quantities, the cross-spectral density can be estimated via

$$\hat{C}_{\mathrm{p}}(f,N) = \frac{1}{N}\sum_{n=1}^{N}\left(\sum_{m=1}^{M} s_{n,a}[m]\exp(-j2\pi fm)\right)\left(\sum_{m=1}^{M} s_{n,b}[m]\exp(-j2\pi fm)\right)^* \tag{3.95}$$

where n indexes over different pairs of unique sample functions. As for the expected autocorrelation, and the expected PSD, the second moment of this estimator can be shown to be

$$\mathrm{E}\left[|\hat{C}_{\mathrm{p}}(f,N)|^2\right] = \frac{1}{N}\left(\mathrm{E}\left[|\hat{C}_{\mathrm{p}}(f,1)|^2\right] - |C_{\mathrm{p}}(f)|^2\right) + |C_{\mathrm{p}}(f)|^2. \tag{3.96}$$

The term $\mathrm{E}\left[|\hat{C}_{\mathrm{p}}(f,1)|^2\right]$ can then be written as

$$\mathrm{E}\left[|\hat{C}_{\mathrm{p}}(f,1)|^2\right] = \mathrm{E}\left[\left|\sum_{m=1}^{M} s_{n,a}[m]\exp(-j2\pi fm)\right|^2\right]\mathrm{E}\left[\left|\sum_{m=1}^{M} s_{n,b}[m]\exp(-j2\pi fm)\right|^2\right] \tag{3.97}$$

where the expectation has been distributed to each independent term. The expectations in (3.97) should be recognized as a definition of the processes expected PSD, $S_{\mathrm{p}}(f)$, such that (3.96) becomes

$$\mathrm{E}\left[\,|\hat{C}_{\mathrm{p}}(f,N)|^2\,\right] = \frac{1}{N}\left(S_{\mathrm{p}}^2(f) - |C_{\mathrm{p}}(f)|^2\right) + |C_{\mathrm{p}}(f)|^2. \tag{3.98}$$

In other words, the expected envelope of a single cross-spectral density pair should largely resemble the PSD of the underlying pulsed WGF. However, this only applies to the envelope since in general each cross-correlation pair will have a random phase while $S_{\mathrm{p}}(f)$ by definition has zero phase for all f. Consequently, with repeated coherent integration these phases cancel each other out and the variance term in (3.98) diminishes.

Similarly, the expected cross-correlation function can be estimated via

$$\hat{c}_{\mathrm{p}}[\ell,N] = \frac{1}{NM}\sum_{n=1}^{N}\begin{cases} \sum\limits_{m=1+\ell}^{M} s_{n,a}[m]s_{n,b}^*[m-\ell] & 0 \le \ell < M \\ \sum\limits_{m=1}^{M+\ell} s_{n,a}[m]s_{n,b}^*[m-\ell] & -M < \ell < 0 \end{cases} \tag{3.99}$$

where the second moment of the estimator is

$$\mathrm{E}\left[\left|\hat{c}_{\mathrm{p}}[\ell,N]\right|^2\right] = \frac{1}{N}\left(\mathrm{E}\left[\left|\hat{c}_{\mathrm{p}}[\ell,1]\right|^2\right] - \left|c_{\mathrm{p}}[\ell]\right|^2\right) + \left|c_{\mathrm{p}}[\ell]\right|^2. \tag{3.100}$$

For only the positive lags, $\mathrm{E}\left[\left|\hat{c}_{\mathrm{p}}[\ell,1]\right|^2\right]$ becomes

$$\mathrm{E}\left[\left|\hat{c}_{\mathrm{p}}[\ell,1]\right|^2\right] = \frac{1}{M^2}\sum_{m_1,m_2=1+\ell}^{M}\mathrm{E}[s_{n,a}[m_1]s_{n,a}^*[m_2]]\mathrm{E}[s_{n,b}^*[m_1-\ell]s_{n,b}[m_2-\ell]] \quad 0 \le \ell < M \tag{3.101}$$

where the expectation can be applied separately to the independent sample functions. This results in a product of correlations such that (3.101) becomes

$$\mathrm{E}\left[\left|\hat{c}_{\mathrm{p}}[\ell,1]\right|^2\right] = \frac{1}{M^2}\sum_{m_1,m_2=1+\ell}^{M} R_{\mathrm{p}}[m_1,m_2]R_{\mathrm{p}}^*[m_1-\ell,m_2-\ell] \quad 0 \le \ell < M. \tag{3.102}$$

Perhaps the most interesting result from (3.102) occurs when $\ell = 0$ where the second moment is the

sum over the magnitude squared of every element of the WGFs correlation function. Consequently, when comparing two WGFs which produce waveforms of the same 3 dB bandwidth, but one has higher expected matched filter sidelobes, then the one with the higher sidelobes will experience a higher peak RMS cross-correlation.

From a practical standpoint, the expected cross-spectral density and the expected autocorrelation function are unlikely to be very useful since it is trivial to force them to be zero for any given WGF. To see this, consider a contrived WGF whose sample functions are all the exact same LFM, but with a uniformly random phase. Such a phase is canceled by the matched filtering process, but would remain in the cross-correlation and cross-spectral density estimates such that their expected values are zero. Because of this, it is not unreasonable to assume that these terms will be zero valued or near zero for a more useful WGF.

3.3.2 Stochastic Processes as CW Radar Signals

Compared to pulsed WGFs, CW WGFs fit much more neatly within the traditional understanding of stochastic processes simply because these WGFs are defined to be infinite in time meaning much of the estimation development of Section 3.2.3 applies. Although, these operations take on a more specific meaning in the context of radar. In the pulsed case, how the WGF was measured was driven by how its sample functions are processed as radar waveforms. Similarly, it is necessary to establish how CW WGF sample functions can be utilized and processed as nonrepeating CW signals.

In the pulsed case, this is fairly intuitive were the processing is focused on each individual pulse, but it the CW case the processing is not as straightforward. Consider a CW scenario. A portion of a single sample function is instantiated from the WGF and then transmitted for some CPI length. The receiver records at the same time. Is it best to treat the entire transmission as a single pulse and match filter the result? Or, is it better to subdivide the data into subsets and then process those results?

3.3.2.1 Autocorrelation Estimation

Since the WGFs of this section are infinite in time and only a single sample function will be used for each radar CPI, it now makes sense to use the time averages of Section 3.2.3 rather than the ensemble averages as in the pulsed case. Consequently, the CW WGFs considered in this section are assumed to be not only WSS but also ergodic in the autocorrelation, which conveniently means the match filtering estimate is actually equivalent to the biased autocorrelation estimate of the process itself as defined in (3.50) such that

$$\hat{r}_{\mathrm{CW}}[\ell] = \hat{R}[\ell] = \frac{1}{M} \begin{cases} \sum_{k=1+\ell}^{M} s[k]s^*[k-\ell] & 0 \leq \ell < M \\ \sum_{k=1}^{M+\ell} s[k]s^*[k-\ell] & -M < \ell < 0 \end{cases}. \tag{3.103}$$

where M is the total number of samples at hand such as those recorded in a CPI and $\hat{r}_{\mathrm{CW}}[\ell]$ is an estimate of the WGF's matched filter response. Likewise the expectation of (3.103), is then

$$\mathrm{E}[\hat{r}_{\mathrm{CW}}[\ell]] = \left(1 - \frac{|\ell|}{M}\right) r_{\mathrm{CW}}[\ell] = \mathrm{E}[\hat{R}[\ell]] = \left(1 - \frac{|\ell|}{M}\right) R[\ell] \quad -M < \ell < M. \tag{3.104}$$

However, while estimating the matched filter via (3.103) is a perfectly reasonable way to estimate the matched filter response of the WGF, it is not necessarily a good way to process noise-like CW radar data due to Doppler effects.

Another way to think about the CW CPI is as a single, very long pulse which means the mainlobe will be highly sensitive to Doppler shifts. Further, the waveforms considered here are noise like and will exhibit ambiguity functions similar to the Doppler intolerant one shown in Fig. 2.20 where the only meaningful response exists near zero Doppler as apposed to the Doppler tolerant LFM ambiguity function in Fig. 2.19. The most straightforward way to accommodate this issue is to filter the data with different Doppler shifted versions of the transmit signal in order to cancel out the Doppler and measure velocity. However, this is computationally very costly since in general for a CW scenario M will be very large and each Doppler filter requires its own matched filtering operation.

Fortunately, the Doppler problem can be largely mitigated by using pseudo-pulse processing where the single, very long pulse is broken into numerous segments where each segment is treated as its own pulse. How big these pseudo-pulses should be is a matter of unambiguous Doppler considerations versus processing requirements. In principal, pseudo-pulse processing is equivalent to the matched filter estimation in (3.103) for zero Doppler. Away from zero-Doppler, the Doppler shift is effectively quantized across the different pseudo-pulses such that it can be mostly canceled via standard FFT based Doppler processing across slow-time.

Written in terms of pseudo-pulse processing where the CW signal has been split into N, W-length pseudo-pulses, 3.103 becomes

$$\hat{r}_{\mathrm{CW}}[\ell, M] = \frac{1}{M}\sum_{n=1}^{N}\begin{cases} \sum_{m=1+W(n-1)+\ell}^{Wn} s[m]s^*[m-\ell] & 0 \le \ell < nW-1 \\ \sum_{m=1+W(n-1)}^{Wn+\ell} s[m]s^*[m-\ell] & -(n+1-N)W+1 < \ell < 0 \end{cases} . \quad (3.105)$$

where (3.105) is actually equivalent to (3.103) for the same value of M, if M is assumed to be evenly divisible by W such that $N = M/W$.

From a practical standpoint, the analysis of (3.105) can be simplified significantly. Since, in general, W will be much less than M such that the impact of the triangular biasing window in (3.104) will be very small within the vicinity of the pseudo-pulse matched filter peak. Thus, the pseudo-pulse estimator can be redefined in terms of the correlation between one of these pseudo-pulses and its infinite length sample function such that

$$\hat{r}_{\mathrm{CW}}[\ell, W, 1] = \frac{1}{W}\sum_{m=1}^{W} s[m]s^*[m-\ell] \quad (3.106)$$

where the pseudo-pulse begins at the arbitrary time $m = 1$. Technically as defined in (3.106), ℓ extends from minus infinity to positive infinity, but in general only perhaps the first few tens of thousands of $|\ell|$ will matter in any practical situation.

Further, in the interest of modeling coherent integration across consecutive pseudo-pulses

(3.106) can be written in terms of N, W-length pulses such that

$$\hat{r}_{\mathrm{CW}}[\ell,W,N] = \frac{1}{NW}\sum_{n=1}^{N}\left(\sum_{m=W(n-1)+1}^{nW} s[m]s^*[m-\ell]\right). \tag{3.107}$$

Since $S_{\mathrm{CW}}[m]$ is assumed to be ergodic, (3.107) will approach $r_{\mathrm{CW}}[\ell]$ as either N or W approach infinity. In (3.107) from the way the pseudo-pulses are directly adjacent to each other in time, the autocorrelation of one pulse is mixed with the cross-correlations with other pulses. Beyond $\ell = W-1$, the correlation function is described entirely by the cross-correlations between a given pseudo-pulse and others. The key assumption made about the pseudo-pulses is that they are effectively uncorrelated each other such that each one represents a unique sample set of the process. Effectively, this means the expected matched filter response, $r_{\mathrm{CW}}[\ell]$, of a good CW WGF should decay quickly. From a radar standpoint, this is desirable anyway since a quickly decaying expected matched filter function implies better autocorrelation sidelobe performance with increasing coherent integration.

The variance of (3.107) is defined

$$\mathrm{E}[|\hat{r}_{\mathrm{CW}}[\ell,W,N]|^2] = \frac{1}{M^2}\mathrm{E}\left[\left|\sum_{m=1}^{M} s[m]s^*[m-\ell]\right|^2\right] \tag{3.108}$$

where for convenience, the double summation over N and W has been replaced with a single summation over M where $M = NW$. Distributing the magnitude squared operation over the summation yields

$$\mathrm{E}[|\hat{r}_{\mathrm{CW}}[\ell,W,N]|^2] = \frac{1}{M^2}\sum_{m_1,m_2=1}^{M}\mathrm{E}\left[s[m_1]s^*[m_1-\ell]s^*[m_2]s[m_2-\ell]\right] \tag{3.109}$$

where the variance is shown to depend on the fourth order moments of the process $S_{\mathrm{CW}}[m]$. Unfortunately, unlike in the pulsed case which took advantage of the independence across different sample functions to simplify the second moment, no such quality can be used here. However, the anticipation is that different pseudo-pulses will be largely uncorrelated with each other such that

they are effectively independent samples of the matched filter function. In fact, for a stationary, ergodic process this must be true in the limit. Otherwise, the time average would never approach the ensemble average.

Because of this, the CW matched filter response, $r_{\mathrm{CW}}[\ell]$ and its second moment in (3.109) can be viewed in a similar way as the pulsed case. $r_{\mathrm{CW}}[\ell]$ represents the ideal matched filter response when an infinite number of pseudo-pulses are coherently integrated while the second moment of its estimator represents the expected matched filter response given the integration of a finite number of pseudo-pulses of length W.

To see this behavior, in Fig. 3.7 the RMS matched filter response for various values of M or combinations of N and W are plotted in comparison to the expected matched filter response. This figure is analogous to Fig. 3.3 for the pulsed case. Where the RMS response is much larger than the expected autocorrelation, the RMS matched filter sidelobes decrease at a rate of roughly $10\log_{10}(M)$ while where the expected matched filter sidelobes are dominate the response stays the same with increasing M. Perhaps the most notable difference between 3.3 and 3.7 is the relatively flat response of the the RMS mathced filter sidelobes in the CW case compared to the pulse case where they approach zero at the edge of the pulse. This is because in the CW case the plotted response is necessarily truncated as it technically extends infinitely in either direction. Consequently, the triangular window that realizes the sloping matched filter response in the pulsed case is inconsequential here. Finally, as in the pulsed case, the expected RMS autocorrelation responses are only averages. Any empirical estimate made from actual sample functions of the process will vary from the expected RMS response to a fair degree as can be seen in Fig. 3.8.

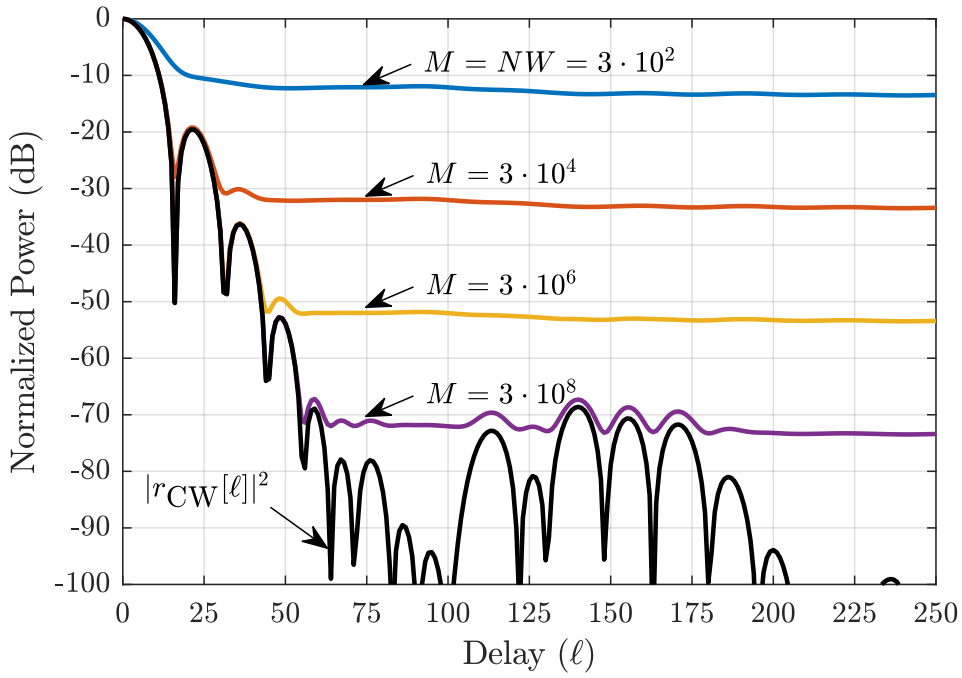

Figure 3.7: The analytically calculated RMS matched filter response of an example WGF given varying total numbers of samples, M

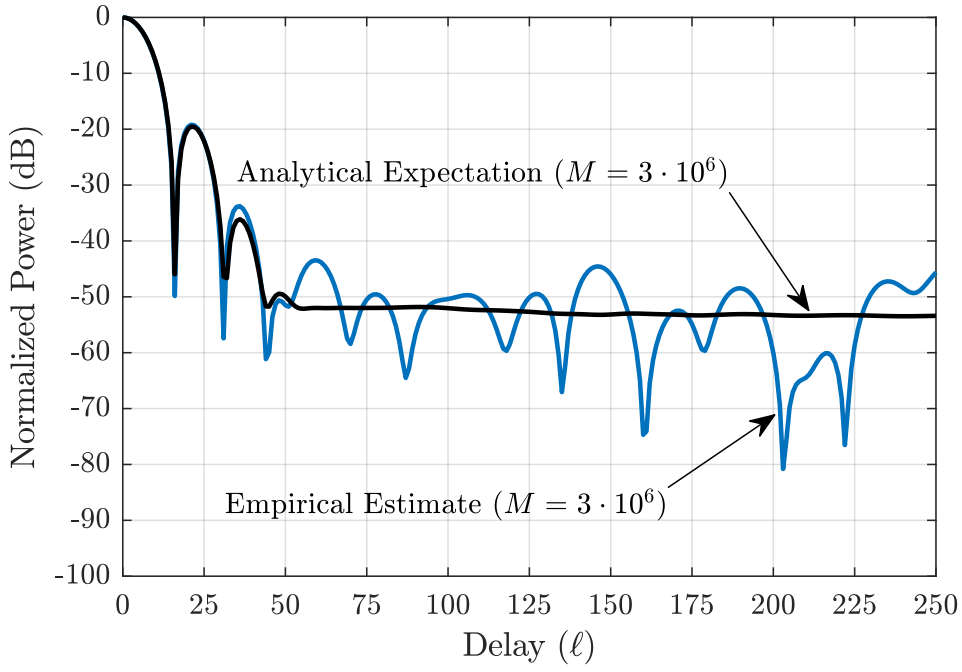

Figure 3.8: Comparison between the analytically calculated RMS matched filter response of an example WGF and an empirical estimate of the same WGF's matched filter response for $M = 3 \cdot 10^6$

3.3.2.2 CW WGF Power Spectral Density Estimation

One convenient aspect of pseudo-pulse processing is that it directly lends itself to a particular form of spectral estimation were the received data is cut into subsets anyway. Because of this, the specific approach used in this work is the Bartlett method where the individual DTFT estimates of the process's spectrum over a number of subsets (pseudo-pulses) are averaged to reduce the variance of the estimate overall. This estimation method is described in Section 3.2.3.3 and is defined for a CW WGF such that

$$\hat{S}_{\mathrm{CW}}(f,W,N) = \frac{1}{N}\sum_{n=1}^{N}\left|\sum_{m=W(n-1)+1}^{nW} s[m]\exp(-j2\pi f(m-W(n-1)))\right|^2 \tag{3.110}$$

were the overall PSD estimate is the average over the individual PSD estimates of each pseudo-pulse. As with any estimator of the PSD of an infinite random process, the trade-offs between different values of N and W involve compromises between bias, variance, and resolution. To see this, (3.110) can be written as the mean of corellogram estimates such that it becomes

$$\hat{S}_{\mathrm{CW}}(f,W,N) = \frac{1}{NW}\sum_{n=1}^{N}\sum_{\ell=-(W-1)}^{W-1}\hat{r}_n[\ell]\exp(-j2\pi f\ell) \tag{3.111}$$

where $\hat{r}_n[\ell]$ represents the matched filter estimate of the nth pseudo-pulse. Note that this differs from the pseudo-pulse estimation of the autocorrelation function where the matched filter estimates were mixed with the cross-correlation terms between a given pseudo-pulse and its immediate neighbors. Here, the spectrum of each pseudo-pulse is considered individually. Because of this the, the expectation of (3.111) yields

$$\mathrm{E}\left[\hat{S}_{\mathrm{CW}}(f,W,N)\right] = \sum_{\ell=-(W-1)}^{W-1}\left(1-\frac{|\ell|}{W}\right)R_{\mathrm{CW}}[\ell]\exp(-j2\pi f\ell) = \frac{1}{W}\left(\frac{\sin(\pi fW)}{\sin(\pi f)}\right)^2 * S_{\mathrm{CW}}(f) \tag{3.112}$$

which contains the Bartlett window biasing term. According to (3.112) it is preferable to make W as large as possible such that the biasing term is minimized. However, (3.112) says nothing about

the variance, which can be examined through the second moment of (3.111)

$$\begin{aligned} \mathrm{E}\left[\left|\hat{S}_{\mathrm{CW}}(f,W,N)\right|^2\right] &= \frac{1}{(NW)^2}\mathrm{E}\left[\left|\sum_{n=1}^{N}\sum_{\ell=-(W-1)}^{W-1}\hat{r}_n[\ell]\exp(-j2\pi f\ell)\right|^2\right] \\ &= \frac{1}{(NW)^2}\sum_{n_1,n_2=1}^{N}\sum_{\ell_1,\ell_2=-(W-1)}^{W-1}\mathrm{E}\left[\hat{r}_{n_1}[\ell_1]\hat{r}^*_{n_2}[\ell_2]\right]\exp(-j2\pi f(\ell_1-\ell_2)) \end{aligned} \tag{3.113}$$

which is once again, a fourth order function of the WGF. The assumption behind the Bartlett method is that the different windows are largely independent of each other such that the variance is minimized with the coherent integration of the different spectral estimates.

To demonstrate this spectral estimation scheme, the expected PSDs for the previously considered WGF for varying values of W are plotted in Fig. 3.9. In the passband region, the value of W has little effect, but it clearly has a significant impact on the apparent spectral roll-off where higher values of W (larger Barlett window sizes) clearly result in better spectral roll-off. This effect is directly attributable to biasing term present in the Bartlett spectral estimation scheme as defined in (3.112). In other words, how the spectrum is estimated has a significant impact on the apparent spectral roll-off of the WGFs PSD.

By itself, Fig. 3.9 seems to indicate that large windows are strictly better than smaller ones as they result in better spectral roll-off. However, the traces in Fig. 3.9 assume an infinite amount of data is available. In Fig. 3.10, $M = 2^{20}$ samples of the process have been instantiated and the spectrum of the process has been empirically estimated using the Bartlett method for varying values of N and W such that $M = NW$. Here the trade-off between roll-off meaning a less biased estimate of the processes PSD, and variance is obvious. Greater degrees of coherent integration through increasing N result in a smoother spectral estimate at the expense of a more biased PSD with more spectral leakage that results in a poorer spectral roll-off.

This result raises an interesting question. The pseudo-pulse processing matched filter estimation leads in nicely to the Bartlett spectral estimation scheme, but there is no reason another spectral estimation scheme could not be used. Although, exploring this in detail is beyond the

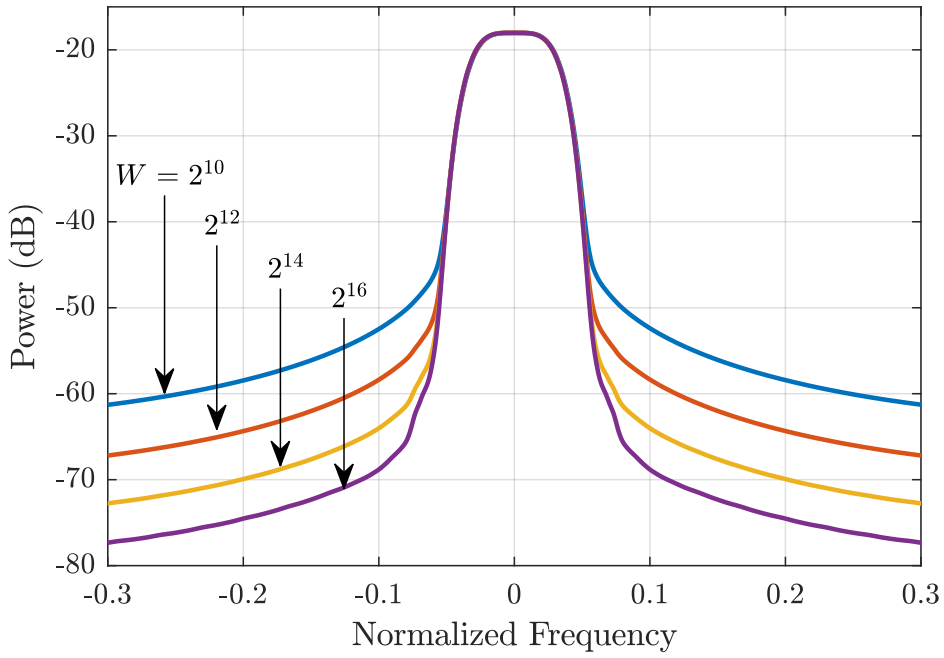

Figure 3.9: The analytically calculated expected PSD of an example WGF given varying Bartlett window sizes.

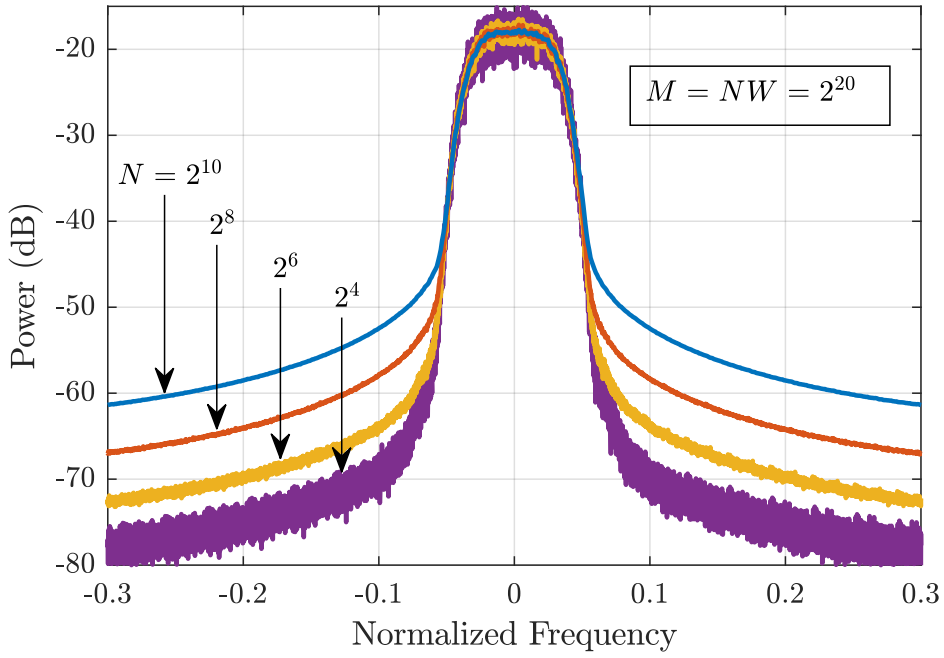

Figure 3.10: Empirical estimates of an example WGF's PSD given varying numbers and sizes of Bartlett windows for $M = 2^{20}$ total samples

scope of this work it is worth noting that any other windowing spectral estimation scheme will result in some SNR loss do the window application. In radar, where SNR is often the biggest problem in performing adequate detection, such a trade-off could prove undesirable.

3.4 Evaluating Waveform Generating Processes

Thus far, WGFs and their analysis have been discussed in purely mathematical terms using the DTFT, but the DTFT is continuous in frequency which clearly cannot be represented on a digital system. Consequently, it is necessary to redefine the topics discussed in 3.3 in terms of the discrete Fourier transform (DFT) using matrix vector notation.

3.4.1 Pulsed WGFs

Beginning with the pulsed WGF defined in (3.63), this can be represented such that

$$\mathbf{s} = \begin{bmatrix} S_{\mathrm{p}}[1] \; S_{\mathrm{p}}[2] \; \cdots \; S_{\mathrm{p}}[M] \end{bmatrix}^T \tag{3.114}$$

where $\mathbf{s}$ is an $M \times 1$ column vector containing the non-zero valued portions of the pulsed waveform generating function. $\mathbf{s}$ is used interchangeably to stand for the WGF itself as well as an arbitrary sample function of the WGF as in the following equations of this section.

The non-zero portion of the WGF's correlation function, $R_{\mathrm{p}}[m_1, m_2]$ can be conveniently represented via a matrix such that

$$\mathbf{C} = R_{\mathrm{p}}[m_1, m_2] = \mathrm{E}\left[\mathbf{s}\mathbf{s}^H\right] \tag{3.115}$$

where $\mathbf{C}$ is an $M \times M$ correlation matrix. According to (3.69), the expected matched filter response is the sum across the diagonals of (3.115) such the ℓth element of the expected matched filter vector $\mathbf{r}$ is

$$r_\ell = \mathrm{E}\left[\hat{r}_\ell\right] = \sum_{m=1+\ell}^{M} C_{m,m-\ell} = \sum_{m=1+\ell}^{M} \mathrm{E}\left[s_m s_{m-\ell}^*\right] \quad 0 \le \ell \le M-1 \tag{3.116}$$

where $C_{m,m-\ell}$ is the element in the mth row and the $(m-\ell)$th column of $\mathbf{C}$ and s_m is the m element of $\mathbf{s}$. Taken together for $-(M-1) \leq \ell \leq (M-1)$ where $r_\ell = r^*_{-\ell}$, the expected matched filter vector, $\mathbf{r}$, is $2M-1$ in length.

Alternatively, $\mathbf{r}$ can be represented as the inverse discrete Fourier transform (IDFT) of the WGF's discrete PSD. However, to do so the WGF's PSD must be evaluated with a sufficient number of points such that

$$\mathbf{s}_\mathrm{f} = \mathrm{E}\left[\hat{\mathbf{s}}_\mathrm{f}\right] = \mathrm{E}\left[\mathbf{A}\bar{\mathbf{s}} \odot \mathbf{A}\bar{\mathbf{s}}^*\right] \tag{3.117}$$

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

$$\bar{\mathbf{s}} = \left[\mathbf{s}^T \ \mathbf{0}^T_{M-1\times 1}\right]^T \tag{3.118}$$

and $\bar{\mathbf{s}}$ is $2M-1$ in length. Here, $\mathbf{A}$ is a $(2M-1)\times(2M-1)$ DFT matrix and $\odot$ is the hadamard product which outputs a vector that is the element by element product of its arguments. The vector $\mathbf{s}_\mathrm{f}$ is equivalent to a sampled form of $S_\mathrm{p}(f)$. The expected matched filter response is then realized as the IDFT of (3.117) realizing

$$\mathbf{r} = \mathbf{A}^H \mathrm{E}\left[\mathbf{A}\bar{\mathbf{s}} \odot \mathbf{A}\bar{\mathbf{s}}^*\right]. \tag{3.119}$$

Although, (3.119) will likely need to be scaled and have its samples rearranged in order to exactly match (3.116) depending on exactly how $\mathbf{A}$ is defined. Likewise, $\mathbf{A}^H$ is the IDFT matrix.

Similarly, the second moments of these functions can be evaluated to realize the expected RMS matched filter response and the power spectral deviation. Using the definition in (3.116), the expected RMS matched filter response for a single sample function can be written as the expectation of the estimate of the expected matched filter such that

$$\mathrm{E}\left[|\hat{r}_\ell|^2\right] = \mathrm{E}\left[\left|\sum_{m=1+\ell}^{M} s_m s^*_{m-\ell}\right|^2\right] = \sum_{m_1,m_2=1+\ell}^{M} \mathrm{E}\left[s_{m_1} s^*_{m_1-\ell} s^*_{m_2} s_{m_2-\ell}\right] \quad 0 \leq \ell \leq M-1 \tag{3.120}$$

which is symmetric about $\ell = 0$ and is a function of the fourth moment of $\mathbf{s}$. For convenience, these

fourth order moments can be collected into the structure $\mathbf{K}_{M\times M\times M\times M}$ where

$$K_{m_1,m_2,m_3,m_4} = \mathrm{E}\left[s_{m_1} s^*_{m_2} s^*_{m_3} s_{m_4} \right]. \tag{3.121}$$

Alternatively, the expected RMS matched filter for a single sample function estimate can be defined in terms of DFTs using (3.119) such that

$$\mathrm{E}[\,|\hat{\mathbf{r}}_\ell|^2\,] = \mathrm{E}\left[\left| \mathbf{A}^H \left(\mathbf{A}\bar{\mathbf{s}} \odot (\mathbf{A}\bar{\mathbf{s}})^* \right) \right|^2 \right]. \tag{3.122}$$

The power spectral deviation for a single sample function estimate can then be defined as a function of the estimate of the power spectral density defined in (3.117) realizing

$$\left(\mathrm{E}\left[|\hat{\mathbf{s}}_\mathrm{f}|^2 \right] - \mathrm{E}\left[\hat{\mathbf{s}}_\mathrm{f} \right]^2 \right)^{1/2}. \tag{3.123}$$

Defining the cross metrics requires a similar process, but with notation to show that there are two unique sample functions. Using the DFT notation, the expected cross-spectral density becomes

$$\mathrm{E}\left[\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^* \right], \tag{3.124}$$

where a and b indicate unique sample functions, and the expected cross-correlation becomes

$$\mathbf{A}^H \mathrm{E}\left[\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^* \right]. \tag{3.125}$$

Similarly, for a single sample function pair estimate, the expected RMS cross-correlation is

$$\mathrm{E}\left[\left| \mathbf{A}^H \left(\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^* \right) \right|^2 \right], \tag{3.126}$$

and the single sample function pair estimate of the RMS cross-spectral density is

$$\left(\mathrm{E}\left[\left|\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^*\right|^2\right]\right)^{1/2}. \tag{3.127}$$

Generally, there are two types of functions defined here. There are those which measure the behavior of the sample functions in aggregate such as the expected matched filter response, and those which measure the behavior of the individual sample functions. These terms are collected into Table 3.3 for convenience.

Table 3.3: Summary of stochastic waveform measures for pulsed WGFs

Aggregate Measures	Power Spectrum	$\mathbf{s}_{\mathrm{f}} = \mathrm{E}[\hat{\mathbf{s}}_{\mathrm{f}}] = \mathrm{E}[\mathbf{A}\bar{\mathbf{s}} \odot \mathbf{A}\bar{\mathbf{s}}^*]$
	Autocorrelation	$\mathbf{A}^H \mathrm{E}[\mathbf{A}\bar{\mathbf{s}} \odot \mathbf{A}\bar{\mathbf{s}}^*]$
	Cross-Spectral Density	$\mathrm{E}\left[\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^*\right]$
	Cross-Correlation	$\mathbf{A}^H \mathrm{E}\left[\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^*\right]$
Individual Measures	Power Spectral Deviation	$\left(\mathrm{E}\left[\left\|\hat{\mathbf{s}}_{\mathrm{f}}\right\|^2\right] - \mathrm{E}[\hat{\mathbf{s}}_{\mathrm{f}}]^2\right)^{1/2}$
	RMS Autocorrelation	$\mathrm{E}[\,\|\hat{\mathbf{r}}_\ell\|^2] = \mathrm{E}\left[\left|\mathbf{A}^H (\mathbf{A}\bar{\mathbf{s}} \odot (\mathbf{A}\bar{\mathbf{s}})^*)\right|^2\right]$
	RMS Cross-Spectral Density	$\mathrm{E}\left[\,\|\mathbf{A}^H \left(\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^*\right)\|^2\right]$
	RMS Cross-Correlation	$\left(\mathrm{E}\left[\left|\mathbf{A}\bar{\mathbf{s}}_a \odot (\mathbf{A}\bar{\mathbf{s}}_b)^*\right|^2\right]\right)^{1/2}$

3.4.2 CW WGFs

CW WGFs can be treated in the same manner. The length M section of a single sample function can be represented as

$$\mathbf{s} = [S_{\mathrm{CW}}[m+1]\; S_{\mathrm{CW}}[m+2] \cdots S_{\mathrm{CW}}[M]] \tag{3.128}$$

where m is arbitrary. Since CW WGFs are assumed to be stationary, their metrics are somewhat simplified. The expected matched filter response is identical to the random processes autocorrelation function such that for the first $|\ell| \leq M-1$ lags, the expected matched filter vector is defined

$$\mathbf{r} = R_{\mathrm{CW}}[\ell] = \mathrm{E}[s_m s^*_{m-\ell}] \tag{3.129}$$

where again m is arbitrary. The $2M-1$ length PSD given Bartlett method estimation can be defined in terms of the DFT matrix and the vector $\mathbf{r}$ such that

$$\mathbf{s}_{\mathrm{f}} = \mathbf{A}\left(\mathbf{w}_{\mathrm{B}} \odot \mathbf{r}\right) = \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] = \mathrm{E}\left[\mathbf{A}\bar{\mathbf{s}} \odot \left(\mathbf{A}\bar{\mathbf{s}}^*\right)\right] \tag{3.130}$$

where $\mathbf{w}_{\mathrm{B}}$ is the Bartlett window in vector form and $\bar{\mathbf{s}}$ is the $2M-1$ length zero-padded version of (3.128). Again, depending on the specific form of $\mathbf{A}$, sample shifting and scaling may be required to guarantee all the terms in (3.130) are exactly equivalent.

Evaluating the RMS matched filter response is unfortunately more complicated without making further assumptions about the WGF. According to (3.109), given an N and W, this can be calculated as

$$\mathrm{E}[\,|\hat{\mathbf{r}}_\ell|^2\,] = \frac{1}{M^2} \sum_{m_1,m_2=1}^{M} \mathrm{E}\left[\, s_{m_1} s^*_{m_1-\ell} s^*_{m_2} s_{m_2-\ell} \,\right]. \tag{3.131}$$

Finally, the power spectral deviation given M samples can be evaluated from (3.130) such that

$$\left(\mathrm{E}\left[\left|\hat{\mathbf{s}}_{\mathrm{f}}\right|^2\right] - \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]^2\right)^{1/2}. \tag{3.132}$$

3.4.3 Summary

This chapter introduced the basic concepts behind the estimation of random variables and processes and then applied them to a special class of random processes called waveform generating functions (WGFs) for both pulsed and CW radar modes. In doing so, the basics behind the estimation of random signals were re-framed in the context of random radar signals processing in great detail. In fact, estimators were introduced for fourth order terms such as the RMS matched filter response and the power spectral deviation. In general higher order estimators such as these simply are not that useful other than for analysis purposes. In a practical scenario where samples from some random process have been collected, estimating these fourth order terms is simply too computationally expensive and requires to much data to reach a good estimate. For instance, if M samples have been collected then there are M^4 possible fourth order moments to estimate!

However, the purpose of random radar signals compared to the traditional purpose of random signals estimation is different. Consider a communications signal. The goal is to estimate the parameters of the process in order to reveal the information hidden within the signal, but in a traditional monostatic radar scenario the transmitted signals are already known. Instead, the goal is to estimate the environment's impact on the signal usually through Doppler shifts and time delays. Consequently, the development of this chapter has been entirely in the anticipation that the WGF has some known, analytical model such that the tools of this chapter can applied analytically to generally characterize the usefulness of the random waveforms generated by a given WGF. In short, there is no anticipation that the structure $\mathbf{K}$ for example will be experimentally determined. It will only ever be analytically calculated from the known model.

With this in mind, the next two chapters introduce the pulsed Stochastic Waveform Generation (StoWGe) and the CW Stochastic Waveform Generation (CW-StoWGe) models whose parameterized WGFs have been specifically formulated such that the metrics defined in Section 3.4 can be analytically calculated and optimized according to radar signals performance goals.

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 random 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, little was said about how to find or design a random process that has desirable properties. In [31], the stochastic waveform generation (StoWGe) signal model was introduced as a means to design and optimize stochastic processes which produce random FM 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 with more flexibility, and considering a much more comprehensive set of test cases.

This chapter begins with an overview of the pulsed 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 pulsed 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 as well as comparisons to previous random FM waveforms and, Section 4.5 presents experimental results which demonstrate both the soundness of the pulsed StoWGe model and design scheme as well as the suitability of the waveforms for implementation on physical hardware. Finally, Section 4.6 provides conclusions regarding the pulsed StoWGe optimization scheme and results.

4.1 The Pulsed StoWGe Signal Model

The StoWGe signal model produces waveforms of the form shown in (3.63). However, since the waveforms produced by the pulsed StoWGe model are constant modulus, the most appropriate baseband signal mode is

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

where the capital S and the capital Φ indicate that these are stochastic processes. In (4.1), $S[m]$ is a non-linear transformation of the stochastic process $\Phi[m]$. The goal is then 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 vector notation such that the zero valued portions of (4.1) are left off and the non-zero valued portion is

$$\mathbf{s} = \exp(j\boldsymbol{\phi}). \tag{4.2}$$

To facilitate the optimization of $\mathbf{s}$ as a WGF, $\boldsymbol{\phi}$ is, by definition of the pulsed StoWGe model, parameterized such that

$$\boldsymbol{\phi} = \mathbf{B}\mathbf{x} + \boldsymbol{\mu} \tag{4.3}$$

where $\mathbf{B}$ is an $M \times N$, real-valued matrix, $\mathbf{x}$ 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 each element of $\boldsymbol{\phi}$ can be tailored such that the WGF, $\mathbf{s}$, has desirable characteristics.

Additionally, $\boldsymbol{\mu}$ provides each element of $\boldsymbol{\phi}$ with a mean value such that

$$\mathrm{E}[\,\boldsymbol{\phi}\,] = \boldsymbol{\mu}. \tag{4.4}$$

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

$$\mathrm{E}[\,\boldsymbol{\phi}\,] = \mathbf{B}\boldsymbol{\mu}_X \tag{4.5}$$

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

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

where X_n is the nth random variable in the vector $\mathbf{x}$ and $\mathbf{b}_n$ is the nth 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 $\boldsymbol{\phi}$.

As was discussed in much detail in Section 3.4, the WGF metrics are functions of the WGFs moments. For the pulsed StoWGe model, these moments are likewise functions of $\mathbf{B}$ and $\boldsymbol{\mu}$. The correlation between the m_1th sample and the m_2th is defined

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

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

$$\mathrm{E}\left[s_{m_1} s_{m_2}^*\right] = \mathrm{E}\left[\exp(j((\mathbf{b}_{m_1} - \mathbf{b}_{m_2})\mathbf{x} + \mu_{m_1} - \mu_{m_2}))\right] \tag{4.8}$$

where $\mathbf{b}_{m_1}$ and $\mathbf{b}_{m_2}$ are the m_1th and m_2 row vectors of $\mathbf{B}$ respectively. Since the values of $\boldsymbol{\mu}$ are constant, they can be factored out of the expectation and the equation can be rewritten as

$$\mathrm{E}\left[\, s_{m_1} s_{m_2}^* \,\right] = \exp(j(\mu_{m_1} - \mu_{m_2}))\mathrm{E}\left[\prod_{n=1}^{N} \exp(j(b_{m_1,n} - b_{m_2,n})X_n)\right], \tag{4.9}$$

where the inner product argument of the complex exponential, $(\mathbf{b}_{m_1} - \mathbf{b}_{m_2})\mathbf{x}$, can be rewritten using a product operator since it is an exponent and $b_{m_1,n}$ is the nth member of the m_1th row of $\mathbf{B}$. Since the elements of $\mathbf{x}$ are by definition statistically independent, the expectation can be moved inside the product operator yielding

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

Such that (4.9) becomes a product of expectations rather than the expectation of a product. Similar to the analysis in [27, 28, 82], note that the expectation in (4.10) is the form of the characteristic function of a random variable which itself is defined as

$$\psi_X(\omega) = \mathrm{E}\left[\exp(j\omega X)\right] \tag{4.11}$$

where $\psi_X(\omega)$ is the characteristic function of the random variable X as defined in Section 3.1. Putting (4.11) into (4.10), the correlation of the signal samples becomes

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

where $\psi_{X_n}\omega$ is the characteristic function of the nth random variable and $\omega = b_{m_1,n} - b_{m_2,n}$. Further, the fourth order moment can be derived similarly such that

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

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 [31], 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 spectral shape. This can be stated mathematically as

$$J_{\mathrm{p}} = \left|\left| \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] - \mathbf{u} \right|\right|_2^2 \tag{4.14}$$

where

$$\mathbf{s}_{\mathrm{f}} = \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] \tag{4.15}$$

is the expected power spectrum of the WGF, $\mathbf{s}$, as defined in Section 3.4, $||\cdot||_2^2$ is the squared Euclidean norm, $\mathbf{u}$ is some desired power spectrum, and the subscript, p, indicates the *pulsed* StoWGe model. Then, as (4.14) is minimized, the WGF will *on average* produce waveforms with a power spectrum that is more and more similar to $\mathbf{u}$ in a mean squared error sense. Potentially, a cost function could be defined for any of the metrics discussed in Section 3.4, but the EFTE cost function offers advantages relative to other stochastic waveform metrics.

For one, the expected power spectrum shares a Fourier transform pair 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 random FM waveforms [23, 24, 26] and in the design of single, low auto-correlation sidelobe waveforms [44, 45, 105] 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 defined via a WGF.

Additionally, the expected power spectrum is only a second order function of the WGF as apposed to a fourth order function like the power spectral deviation or the RMS matched filter response. This represents a clear advantage in terms of the processing and complexity requirements. Simply put, for an M length pulsed WGF there are M^2 second order terms, but there are M^4 fourth

order terms.

Alternatively, one may consider optimizing the expected autocorrelation directly since it is also a second order function. The most straightforward choices are to optimize via minimizing either ISL, PSL, or the more flexible generalized integrated sidelobe level (GISL) metric [70, 105]. However, these metrics largely disregard spectral containment making the EFTE metric more desirable at least for random FM waveforms [105].

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 so 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 crucial to at least examine the individual metrics of Section 3.4 to ensure the minimization of the EFTE cost function is producing useful WGFs. Such an examination is provided in the following sections.

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 in a similar manner [31]. In order to determine the best optimization method, the 2-norm in (4.14) can be expanded such that it becomes

$$J_{\mathrm{p}} = \left(\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]\right)^{T}\left(\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]\right) - 2\mathbf{u}^{T}\left(\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]\right) + \mathbf{u}^{T}\mathbf{u} \tag{4.16}$$

which is a fourth order function of $\mathbf{s}$. Additionally, as shown in (4.12), the second order moments of the samples in $\mathbf{s}$ are already products of the characteristic functions of the random variables of $\mathbf{x}$. The characteristic functions themselves are dependent on $\mathbf{B}$. Consequently, the EFTE cost function is highly non-linear. This makes it unlikely that (4.14) is convex such that finding its global minimum is extremely difficult if not impossible.

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 which (4.14) is. 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

$$\frac{\partial J_{\mathrm{p}}}{\partial b_{k,n}} = 2\left(\frac{\partial \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]}{\partial b_{k,n}}\right)^{T}\left(\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]-\mathbf{u}\right) \tag{4.17}$$

where $b_{k,n}$ is the element of $\mathbf{B}$ in the kth row and nth column, and the derivative of the expected power spectrum is evaluated on an element wise basis. The derivative of the wth element of $\mathrm{E}[\hat{s}_{\mathrm{f},w}]$ is

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = 2\Re\left\{\sum_{m=1}^{M} a_{w,m}a_{w,k}^{*}\exp\left(j(\mu_m-\mu_k)\right)\frac{\partial \psi_{X_p}(b_{m,n}-b_{k,n})}{\partial b_k,n}\prod_{\substack{p=1\\p\neq k}}^{N}\psi_{X_p}(b_{m,n}-b_{k,n})\right\} \tag{4.18}$$

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

$$\nabla_{\mathbf{B}} J_{\mathrm{p}} = \begin{bmatrix} \frac{\partial J_{\mathrm{p}}}{\partial b_{1,1}} & \frac{\partial J_{\mathrm{p}}}{\partial b_{1,2}} & \cdots & \frac{\partial J_{\mathrm{p}}}{\partial b_{1,N}} \\ \frac{\partial J_{\mathrm{p}}}{\partial b_{2,1}} & \frac{\partial J_{\mathrm{p}}}{\partial b_{2,2}} & \cdots & \frac{\partial J_{\mathrm{p}}}{\partial b_{2,N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J_{\mathrm{p}}}{\partial b_{M,1}} & \frac{\partial J_{\mathrm{p}}}{\partial b_{M,2}} & \cdots & \frac{\partial J_{\mathrm{p}}}{\partial b_{M,N}} \end{bmatrix} \tag{4.19}$$

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

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

$$\frac{\partial J}{\partial \mu_k} = 2\left(\frac{\partial \mathrm{E}[\hat{\mathbf{s}}_\mathrm{f}]}{\partial \mu_k}\right)^T (\mathrm{E}[\hat{\mathbf{s}}_\mathrm{f}] - \mathbf{u}) \tag{4.20}$$

where μ_k is the kth element of $\boldsymbol{\mu}$. Taking the derivative on an element wise basis, the derivative of the wth element of $\mathrm{E}[|s_{\mathrm{f},w}|^2]$ is

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial \mu_k} = 2\Im\left\{ a^*_{w,k} \exp(-j\mu_k) \sum_{\substack{m=1 \\ m\neq k}}^{M} a_{w,m} \exp(j\mu_m) \prod_{p=1}^{P} \psi_{X_p}(b_{k,p} - b_{m,p}) \right\} \tag{4.21}$$

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

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

to realize the gradient of the EFTE cost function with respect to the mean phase value 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 seemingly arbitrary initialization decisions 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, error and educated guesses.

In the course of evaluating the efficiency of various gradient descent techniques such as the ones in [88], 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

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

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 [89]. 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 better than the simplistic heavy ball method.

The particulars of the heavy ball gradient descent method used in this work are detailed in Table 4.1. Going from the first steps to the last, the following is an overview of heavy ball gradient descent minimization of the EFTE cost function for pulsed StoWGe WGFs. First, in step 1 various parameters are initialized. The parameters M, N, $p_X(x)$, $\mathbf{u}$, $\mathbf{B}_0$, and $\boldsymbol{\mu}_0$ pertain directly to the pulsed StoWGe model and are discussed in detail in the next section. The parameters β, ρ_{up}, ρ_{down}, c, and γ are used exclusively for the gradient descent optimization. As step 2 indicates, the optimization is run iteratively until a set of condition(s) is met. The optimization begins in step 3 were the current cost function value and the gradients with respect to both $\mathbf{B}$ and $\boldsymbol{\mu}$ are evaluated. In step 4, the search direction components corresponding to the both $\mathbf{B}$ and $\boldsymbol{\mu}$ are calculated as a function of the heavy ball parameter β and their respective previous search direction components.

At this point it is important to remember what these two different gradients mean in the context of the gradient descent scheme. As presented, they appear as if they are two distinct gradients and consequently independent search directions, but this is just a matter of notational convenience. Instead, as they are treated here, they are really just two different portions of the same total gradient with respect to all the parameters in both $\mathbf{B}$ and $\boldsymbol{\mu}$.

Table 4.1: Pseudo-code for the gradient descent optimization of pulsed StoWGe WGFs through the minimization of the EFTE

EFTE Gradient Descent Optimization of the Pulsed StoWGe WGF

1: **Initialize:** $M,\ N,\ p_X(x),\ \mathbf{u},\ \mathbf{B}_0,\ \boldsymbol{\mu}_0,\ \beta,\ \rho_{\text{up}},\ \rho_{\text{down}},\ c,\ \gamma$
set $i = 0, \mathbf{Q}_0 = \mathbf{0}_{M\times N}, \mathbf{q}_0 = \mathbf{0}_{M\times 1}$

2: **Repeat**

3: Evaluate: $J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i)$, $\nabla_{\mathbf{B}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i)$, and $\nabla_{\boldsymbol{\mu}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i)$

4: $\mathbf{Q}_{i+1} = -\nabla_{\mathbf{B}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i) + \beta \mathbf{Q}_i, \quad \mathbf{q}_{i+1} = -\nabla_{\boldsymbol{\mu}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i) + \beta \mathbf{q}_i$

5: **If** $\left(\langle \mathbf{Q}_{i+1}, \nabla_{\mathbf{B}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i)\rangle + \mathbf{q}_{i+1}^T (\nabla_{\boldsymbol{\mu}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i))\right) \geq 0$

6: $\mathbf{Q}_{i+1} = -\nabla_{\mathbf{B}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i), \quad \mathbf{q}_{i+1} = -\nabla_{\boldsymbol{\mu}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i)$

7: **End** (If)

8: **While** $J_{\text{p}}(\mathbf{B}_i + \gamma \mathbf{Q}_{i+1},\ \boldsymbol{\mu}_i + \gamma \mathbf{q}_{i+1})$
$> J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i) + c\gamma\left(\langle \mathbf{Q}_{i+1}, \nabla_{\mathbf{B}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i)\rangle_F + \mathbf{q}_{i+1}^T (\nabla_{\boldsymbol{\mu}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i))\right)$

9: $\gamma = \rho_{\text{down}} \gamma$

10: **End** (While)

11: $\mathbf{B}_{i+1} = \mathbf{B}_i + \gamma \mathbf{Q}_{i+1}, \quad \boldsymbol{\mu}_{i+1} = \boldsymbol{\mu}_i + \gamma \mathbf{q}_{i+1}, \quad \gamma = \rho_{\text{up}} \gamma$

12: $i = i + 1$

13: **If** C_{Stop}

14: **Stop**

15: **End** (If)

16: **End** (Repeat)

In steps 4 through 6, the new search direction (both $\mathbf{Q}_{i+1}$ and $\mathbf{q}_{i+1}$) is tested to make sure it actually is a descent direction. As noted in Section 2.5, if the search direction and the gradient are vectors, then if the inner product between them is positive then the search direction is actually an ascent direction and it needs to be reset to the negative of the current gradient. Here, the search direction consists of both a matrix and a vector, but the principle is the same. In step 5, the Frobenius product is evaluated with respect to $\mathbf{Q}_{i+1}$ and $\nabla_{\mathbf{B}} J_{\text{p}}(\mathbf{B}_i, \boldsymbol{\mu}_i)$ which is defined such that

$$\langle \mathbf{T}, \mathbf{V} \rangle := \sum_{n,m} t_{nm} v_{nm} \tag{4.24}$$

where **T** and **V** are arbitrary matrices of the same dimensionality. The Frobenius product itself is equivalent to stringing the columns of each matrix into a single, long vector and calculating the resulting inner product. If the resulting sum of the Frobenius product and the inner product in step 5 is greater than zero then the search direction is actually a direction of ascent and it is reset to the negative of gradient.

Next, in steps 8 through 10, a simple line search is performed to find an acceptable step size, γ. Section 2.5 introduces the strong Wolfe conditions, but here only the Armijo condition (the first of the strong Wolfe conditions), corresponding to the term multiplied by c, is used as apposed to using both strong Wolfe conditions. The second Wolfe condition, the curvature condition, is left off here since, while it would result in a more efficient optimization on a per iteration basis, it would also require repeated evaluations of the gradient which would significantly increase the per iteration computational cost. Consequently, the trade off here is to sacrifice some per iteration improvement for a meaningful reduction in computational expense to enable a much greater overall number of iterations in the same amount of time as compared to using both conditions. In step 9 if the Armijo condition is not met, the step size is reduced by a factor of $\rho_{\text{down}} < 1$ until an acceptable step size is found.

After the line search, in step 11 the parameters are updated and the step size in increased by a factor of $\rho_{\text{up}} > 1$. Increasing γ here is a natural consequence of the line search. Without this step, the step size would likely get smaller and smaller and the per iteration improvement would dwindle. In affect, this is a rudimentary compensation for leaving out the curvature condition which keeps the step size from getting too small. In other words, the Armijo condition keeps the step size from getting too large while step 11 keeps the step size from getting to small. In step 12 the iteration index is increased. Finally, in steps 13 through 15 the stopping conditions are checked where in step 13 C_{Stop} is a generic place holder for the stopping condition(s). If they are met, then the optimization is finished. The particular values and conditions used for the optimization in Table 4.1 are discussed in detail in Sections 4.3.3 and 4.3.4.

4.3.3 Pulsed StoWGe Model Initializations

One difficulty in evaluating the optimization of pulsed StoWGe via the EFTE cost function is the number of parameters to consider. Table 4.2 highlights this issue by listing all the independent parameters.

Table 4.2: WGF parameters which must be selected or initialized prior to optimization

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

The parameters M, N, $p_X(x)$, and $\mathbf{u}$ are fixed throughout the optimization, while the basis function 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 unique 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 4.2 along with the rationale for those selections.

Number of Waveform Samples: M – To understand how the EFTE minimization performs on pulsed 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 a noise like sidelobe reduction 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/PMFs, although we can disregard PDFs/PMFs with a non-zero mean. The phase mean value vector $\boldsymbol{\mu}$ 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. 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 an probability of taking one of two states such that

$$\text{DU2:} \quad p_X(x) = \frac{1}{2}\delta(x+\pi) + \frac{1}{2}\delta(x-\pi) \tag{4.25}$$

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

$$\text{CU:} \quad p_X(x) = \frac{1}{2\pi} \begin{cases} 1 & -\pi \leq x \leq \pi \\ 0 & \text{otherwise} \end{cases} \tag{4.26}$$

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 random FM waveforms in other works such as [27–29, 31, 72, 82]. Here, the Gaussian random variable is defined to have unit variance such that

$$\text{G:} \quad p_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{4.27}$$

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 as this spectrum ideally results in zero autocorrelation sidelobes and good spectral roll-off. Consequently, this spectrum will be examined here as well.

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 peak normalized power by the edge of the sample bandwidth. This may be enough motivation in some cases to move away from the Gaussian spectrum. One alternative is the *Super-Gaussian* spectral shape defined as

$$u(f) = A \exp\left(\frac{|f - f_c|^n}{x_{3dB}}\right) \tag{4.28}$$

where A arbitrarily scales the power, $x_{3\text{dB}}$ can be chosen to set $B_{3\text{dB}}$, n is an integer greater than two, and f_c is the center frequency. If n is set to two, then (4.28) 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 (4.28) approaches a rectangular function which is the third spectral shape considered here.

The baseband rectangular spectral template is defined such that

$$u(f) = \begin{cases} 1 & -f_B/2 \leq f \leq f_B/2 \\ 0 & \text{otherwise} \end{cases} \tag{4.29}$$

where the 3 dB bandwidth, f_B, is also the absolute bandwidth in this case. This template represents 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 the Gaussian function while the rectangular functions resembles a "brick-wall" at the 3 dB bandwidth representing the ideal 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 where the oversampling factor has been set to $K = 8$ 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 random FM waveforms where coherent integration of multiple pulses 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. In this way, the choice of spectra determines what kind of *persistent* sidelobes the WGF will have.

Historically, the super-Gaussian function has been a topic of much interest in optics [106], here it is suggested as a compromise between the zero persistent autocorrelation sidelobe Gaussian spectrum that demonstrates modest spectral containment and the rectangular spectrum which achieves ideal spectral containment, but with extensive persistent sidelobes. With this in mind, these three templates are used with two different oversampling factors of $K = 2$ or $K = 4$ as plotted in Figs. 4.1 and 4.3 respectively.

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

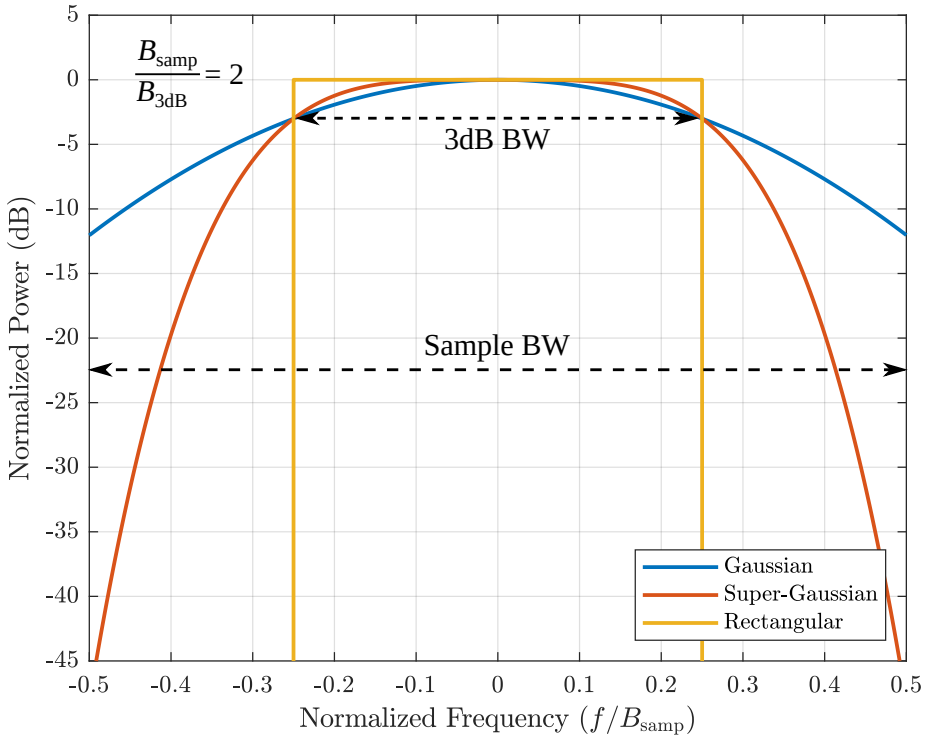

Figure 4.1: Desired spectra for the EFTE optimization with a 3 dB oversampling factor of $K = 2$

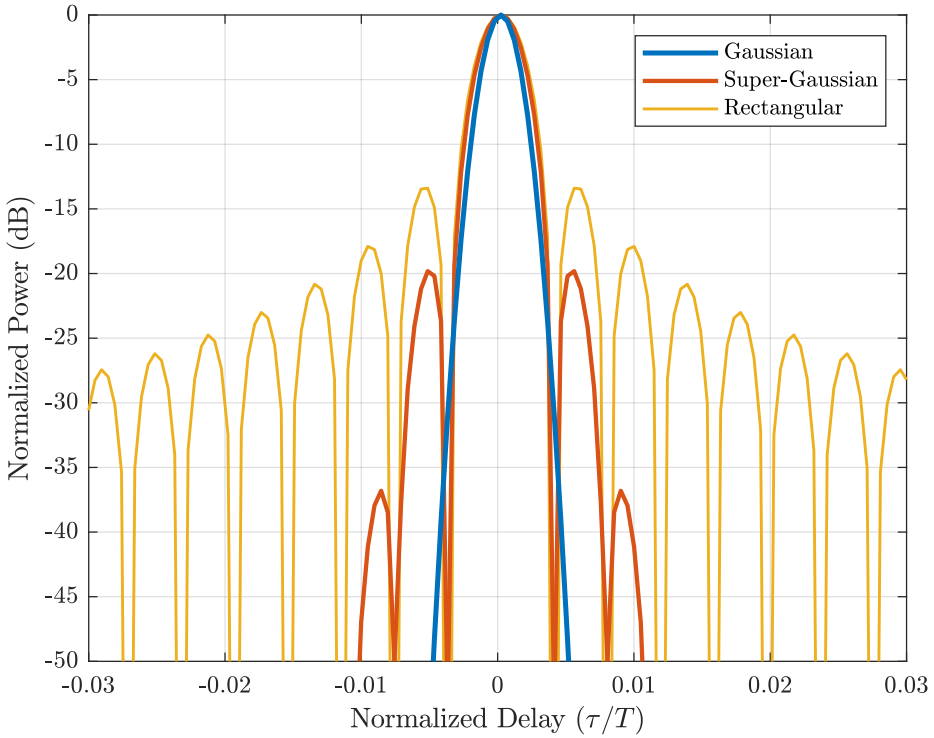

Figure 4.2: Corresponding autocorrelation response of the desired EFTE spectra

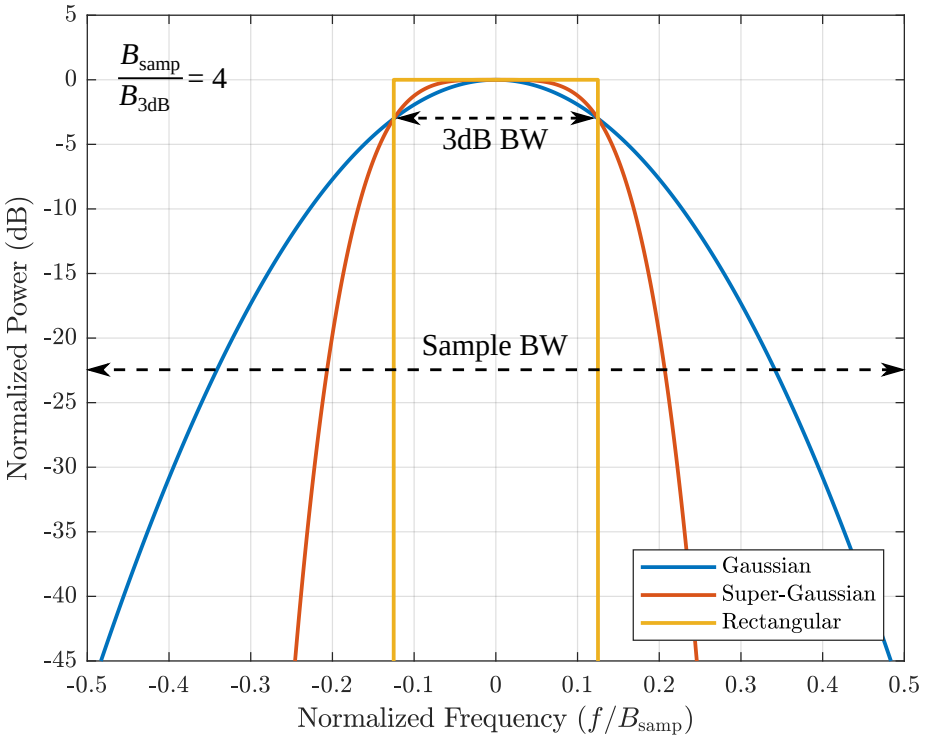

Figure 4.3: Desired spectra for the EFTE optimization with a 3 dB oversampling factor of $K = 4$

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 toward 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, K, 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

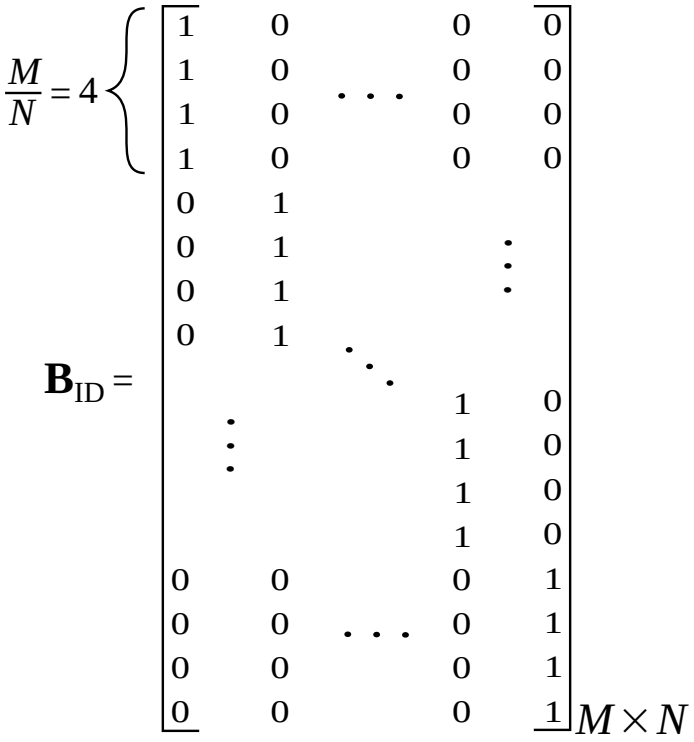

Figure 4.4: Structure of the "Identity" basis matrix initialization ($\mathbf{B}_{\text{ID}}$)

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

Initial Phase Mean Value Vector: $\boldsymbol{\mu}_0$ – An asymmetric baseband expected spectrum can only be achieved if the phase values have a non-zero mean value. However, since this work is only concerned with the symmetric spectral templates discussed above, it is not necessary for the phase of the pulsed StoWGe signals to have a non-zero mean value. In fact it would be detrimental to matching to the templates. For future work the optimization of $\boldsymbol{\mu}$ can be considered for such cases as spectral notching where asymmetric baseband spectra are required. For this work however, $\boldsymbol{\mu}_0$ will be initialized as a vector of zeros and will not be optimized. For completeness and to show how it can be done, the algorithm in Table 4.1 include the vector $\boldsymbol{\mu}$ as a set of optimizable parameters, but for the optimizations in this work these terms are simply removed.

Summary: – All the initializations and their shorthand notation from above are summarized in Table 4.3.

Table 4.3: Optimization initializations for the pulsed StoWGe parameters. All combinations are considered resulting in 288 total optimizations

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
$\mathbf{u}$	Desired Spectrum	G2	– Gaussian $(K = 2)$
		G4	– Guassian $(K = 4)$
		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	$\mathbf{B}_{\text{PC}}$	– PCFM basis matrix
		$\mathbf{B}_{\text{ID}}$	– Modified identity matrix
$\boldsymbol{\mu}_0$	Phase Mean Value Vector	$\mathbf{0}$	$-\left[0 \;\; 0 \;\; \cdots \;\; 0\right]^T_{M\times 1}$

4.3.4 Pulsed StoWGe Gradient Descent Parameter Initializations

The specific values of the gradient descent parameters, the choice of line search, and the choice of heavy ball gradient descent as apposed to others were made with the intention of making the optimization reasonably fast but more importantly robust. It is certainly possible if not likely that there is a better set of parameters, line search, and gradient descent method that could minimize the EFTE cost function more quickly than what it presented here, but the goal of this work is to show that the minimization can produce useful WGFs and not necessarily how quickly it can do so. After all, one of the biggest advantages of StoWGe is that the optimization does not have to be performed in a just in time manner since entire stochastic families of waveforms are optimized rather than individual waveforms.

With this mindset, the parameters, $\beta,\ \rho_{\text{up}},\ \rho_{\text{down}},\ c,$ and γ where chosen via trial and error.

For every one of the test cases outlined in Table 4.3, $\beta,\ \rho_{\text{up}},\ \rho_{\text{down}},\ c$ where set to $.95, 1.01, .9, .01$ respectively and γ was initialized as 10^{-6}.

In Table 4.1, the stopping conditions were represented by a place holder since the chosen conditions really only make sense in the context of the values chosen for the other parameters. For the pulsed StoWGe optimization, three stopping conditions were used. The first is simply that the optimization will run for at least 5000 iterations. This is a consequence of the initial step size being very small, which in turn was a decision made to prevent the optimization from jumping into a poor solution space. In general, the initial basis matrix, $\mathbf{B}_0$, results in a very poor match to the expected spectrum, which in turn means basically any step in any direction of any size stands a good chance of lowering the cost function value and being "acceptable" as far as the algorithm is concerned. For example, if adding 1 billion to each parameter in the initial basis matrix results in a decrease in the cost function value, then the algorithm will accept that change. The problem is that it is unlikely that an "excellent" local minimum exists in the neighborhood of such an exotic region of the solution space. Making the initial step size so small is intended to prevent this from happening and to guarantee that whatever optimization solution is found is in or near the neighborhood of the initialization. The down side is that since it starts out so small such that it can take a fair number if iterations for it to become large enough such that the cost function is meaningfully decreased with each iteration. The minimum of 5000 iterations prevents the other stopping conditions from prematurely ending the optimization.

Second, if the total decrease of the cost function in any consecutive set of 1000 iterations is less than .01 dB, the optimization is stopped. The rationale is simply that if the optimization is progressing so slowly (outside of the first 5000 iterations), then it is unlikely that further iterations will realize meaningful improvement.

Finally, the optimization will stop at either 10^6 iterations or if the cost function value goes below -200 dB. These are both pragmatic conditions. If the iteration count makes it to 10^6, then the cost function value is likely already very small and is barely not meeting beating the previous condition so the optimization is stopped in the interest of time. The -200 dB condition is somewhat

arbitrary, but the idea is that -200 dB is effectively 0 in this context. In the sense that is there really a difference between -200 dB and say -210 dB in a way that can be meaningfully evaluated when producing the stochastic waveforms? In linear terms the difference is minuscule.

The key takeaway in how these stopping conditions were chosen is not how arbitrary they seem to be, but that they were intentionally chosen to be conservative (designed not to stop too early) through a heuristic process of trial and error. After all, non-linear, local minimization itself is as much an art as it is a science [87].

4.4 Pulsed StoWGe Optimizations Results

By the definition of the EFTE cost function, the best pulsed StoWGe WGFs are those which result in the smallest cost function values as a function of $\mathbf{B}_0$, N, and $p_X(x)$. Every result of each of the optimizations detailed in Table 4.3 is tabulated in Tables B.7-B.12 on a dB scale where more negative values (smaller) represent a better optimization result. If the goal of this work were only to produce WGFs which achieve a desired expected power spectrum, then all the relevant analysis could be performed by simply examining the contents of those tables. However, the goal behind the WGFs optimized here is that they produce useful random FM waveforms for radar purposes and having a good expected spectrum is only a part of that goal. Still, it is expected that these two goals are well correlated and that if a WGF matches well to its desired expected spectrum then it is more likely to produce useful radar waveforms than otherwise. Because of this, Tables B.1-B.6 represent an excellent starting point.

The analysis begins in Section 4.4.1 by examining a problematic corner case that occurs for the WGFs with a small value of N. Then, using many of the metrics outlined in Table 3.3, Section 4.4.2 evaluates the WGFs with the best cost function values for their ability to produce useful radar waveforms. Section 4.4.3 looks at how the optimizations achieved their results by qualitatively examining the optimized basis and correlation matrices. Finally, Section 4.4.4 compares the pulsed StoWGe waveforms produced here to previous pulsed random FM waveforms.

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 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 random FM waveforms, since the last thing noise-like waveforms should do is repeat themselves. In fact, when $N = 2$ there is a 25% chance that two of the same waveform will be instantiated 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 [71, 107, 108], 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 apart for discrete distributions 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 given 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 random FM 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 WGFs 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.

In Table B.1, the $\mathbf{B}_{\text{PC}}$ 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.

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.

For the smaller values of N, the spike is intuitive. 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, which represents a problem from a pulse integration standpoint. It will take more pulse integration for the sample power spectrum to resemble the expected power spectrum. To avoid this issue, only WGFs with $N \geq 32$ will be considered for either the CU or G

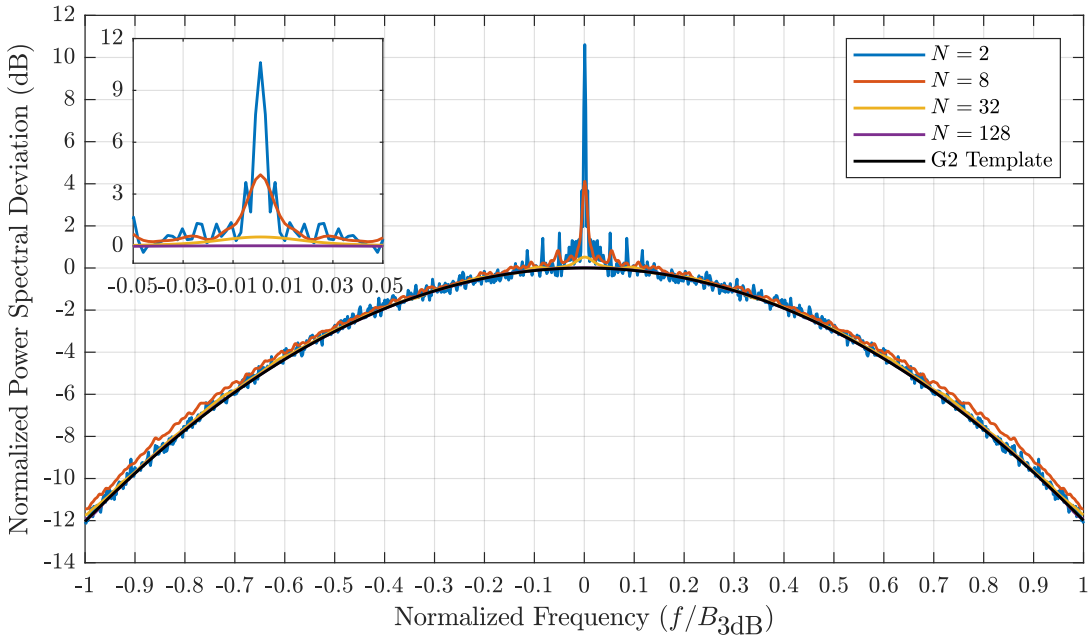

Figure 4.5: Normalized power spectral deviation for the CU distribution, the G2 template, and various values of N

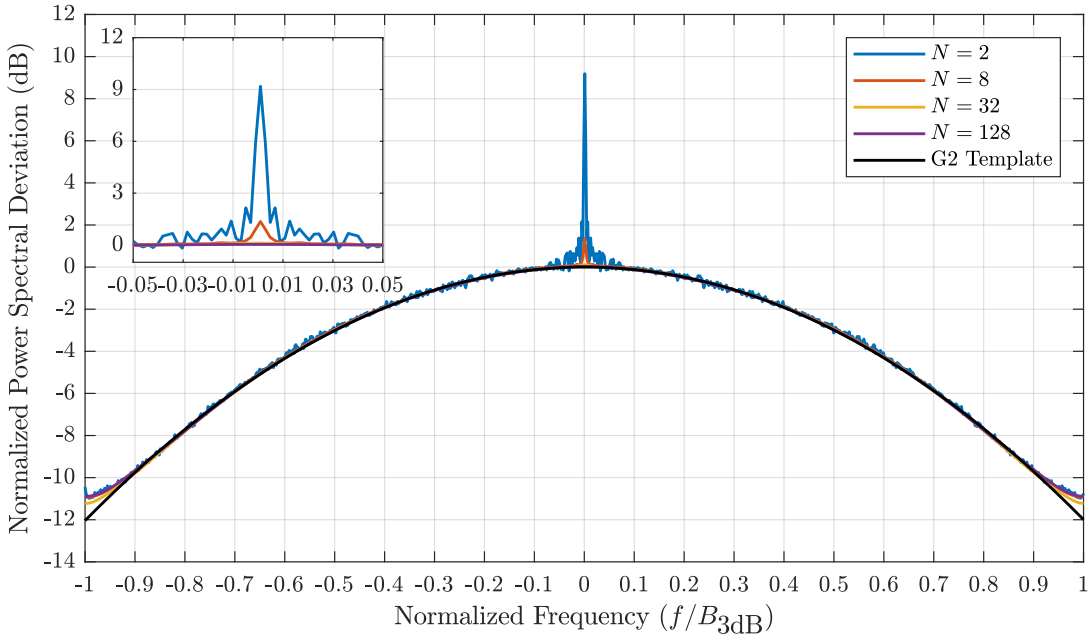

Figure 4.6: Normalized power spectral deviation for the G distribution, the G2 template, and various values of N

distributions in the same way as for the DU2 case.

4.4.2 Pulsed StoWGe Optimized WGF Analysis

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. The most straightforward place to begin, is to examine the cases with the best 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 4.4. Beyond the template match, it is important to understand

Table 4.4: Optimized WGFs which resulted in the lowest value of J_{p} for each $p_X(x)$ and $\mathbf{u}$ as a function of N for $N \geq 32$ and $\mathbf{B}_0$.

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

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.
- 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 is only five of the eight metrics listed in Table 3.3. The other three metrics as discussed in 3.3.1.3 generally are not useful and were included for completeness.

4.4.2.1 Gaussian Template Results

In Figs. 4.7 and 4.8 the analytically calculated spectral and temporal metrics of each of the two Gaussian templates are plotted 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 the template 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 its deviation while in 4.7(d), the G distribution shows much more variability 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. Such behavior indicates it would likely take more coherent integration for the spectral estimates of the Gaussian random variable WGF sample function to approach their expected power spectrum as opposed to the other distributions.

In Fig. 4.8(a,b), the expected autocorrelation function (ACF) for each distribution either matches the templates corresponding ACF or even outperforms it in the 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 an individual sidelobe level of 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 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 average sidelobe level (near the mainlobe) should be around $-10\log_{10}(BT)$ 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, the RMS cross-correlation results largely resemble the RMS autocorrelation results but without a mainlobe.

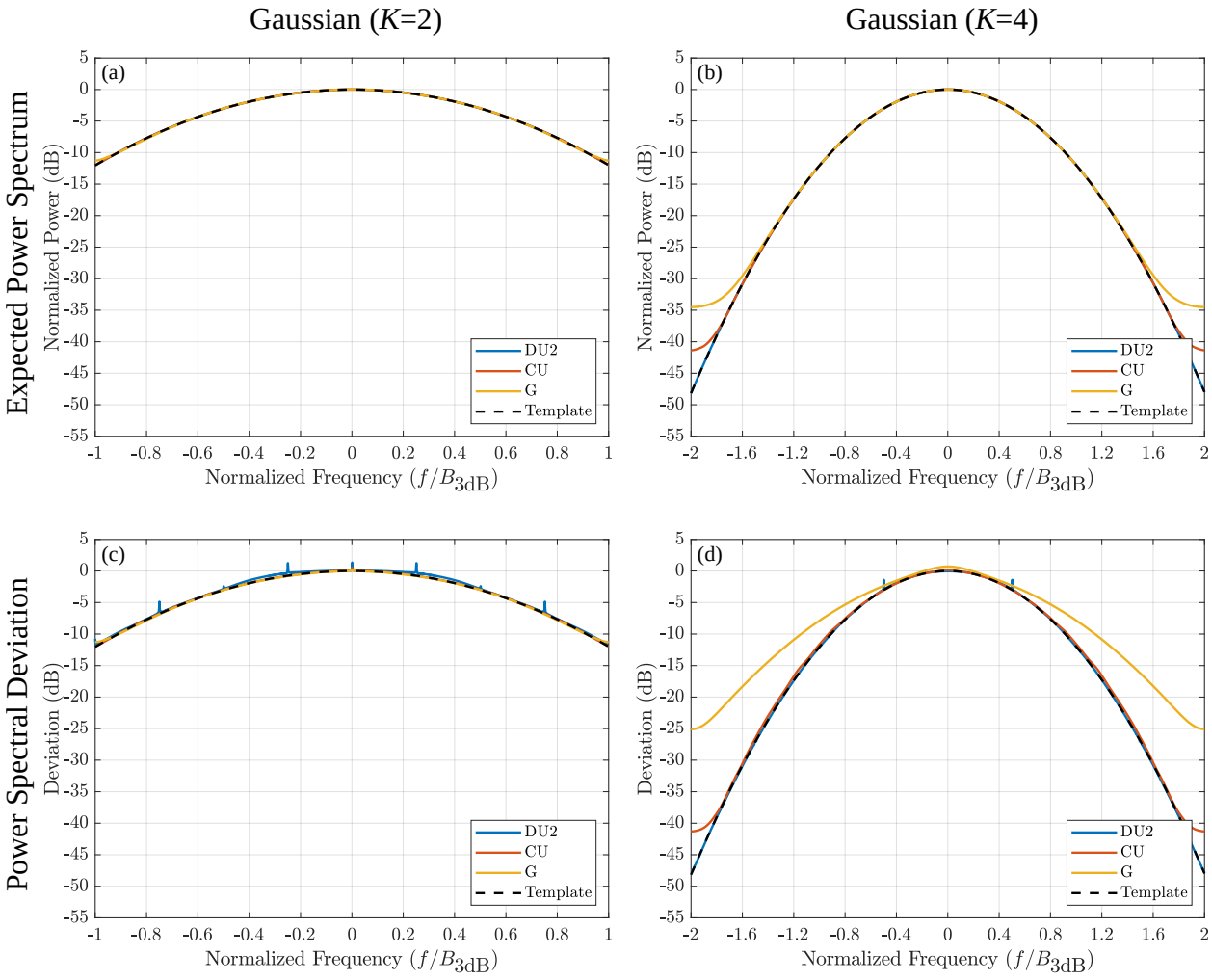

Figure 4.7: Pulsed StoWGe spectral optimization results for Gaussian templates: expected power spectrum for the G2 template (a), expected power spectrum for the G4 template (b), expected power spectral deviation for the G2 template (c), expected power spectral deviation for the G4 template (d)

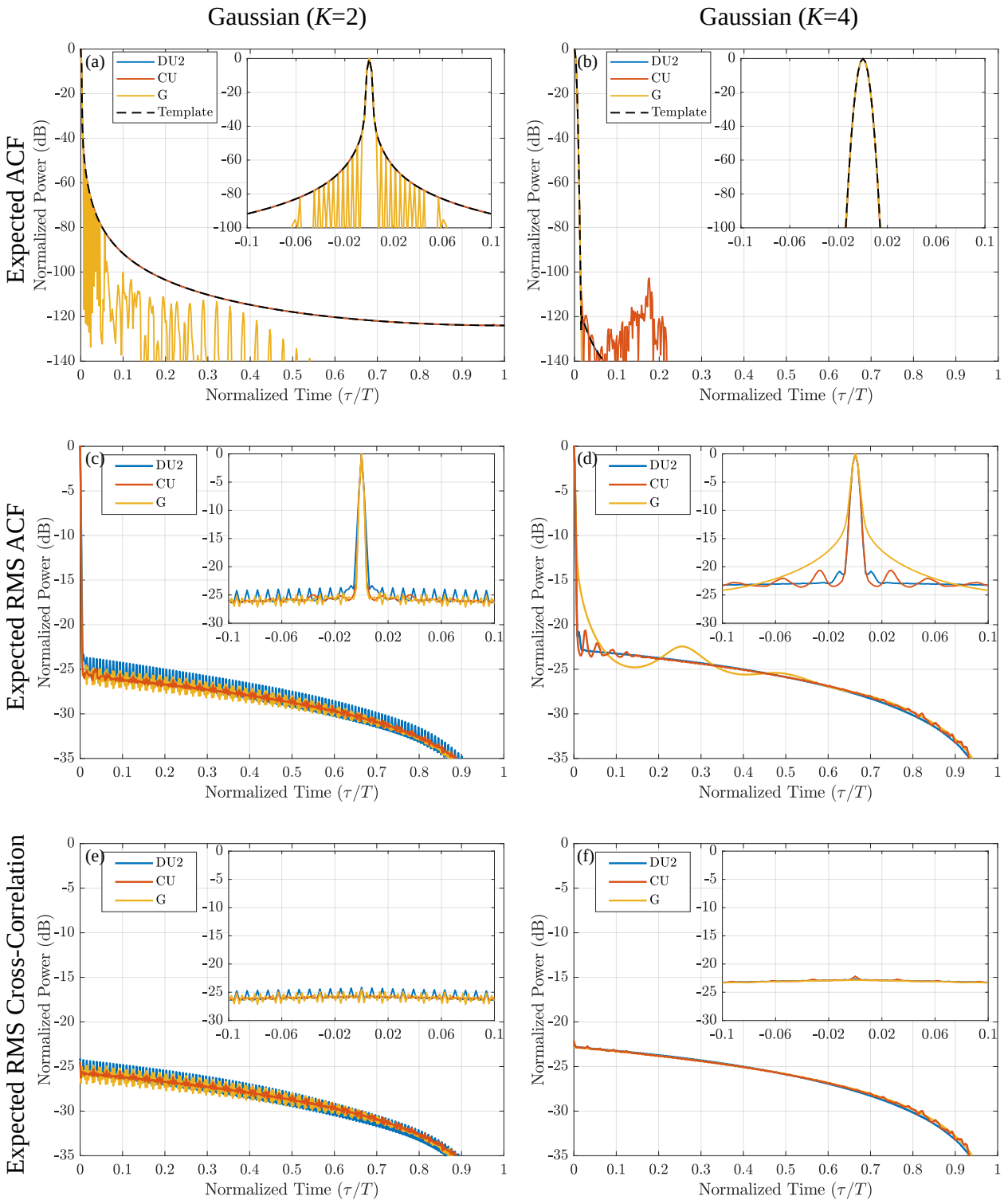

Figure 4.8: Pulsed StoWGe temporal optimization results for Gaussian templates: expected autocorrelation for the G2 template (a), expected autocorrelation for the G4 template (b), expected RMS autocorrelation for the G2 template (c), expected RMS autocorrelation for the G4 template (d), expected RMS cross-correlation for the G2 template (e), expected RMS cross-correlation for the G4 template (f)

4.4.2.2 Super-Gaussian Template Results

Compared to the G2 and G4 templates, there is much more disparity between the distributions for 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, while the CU distribution did 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) the DU2 distributed random variable enabled a much better match to the template down to about -35 dB which clearly outperforms 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 Section 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 auto-correlation, but with less dramatic oscillations in the S4G4 case and with the mainlobe removed.

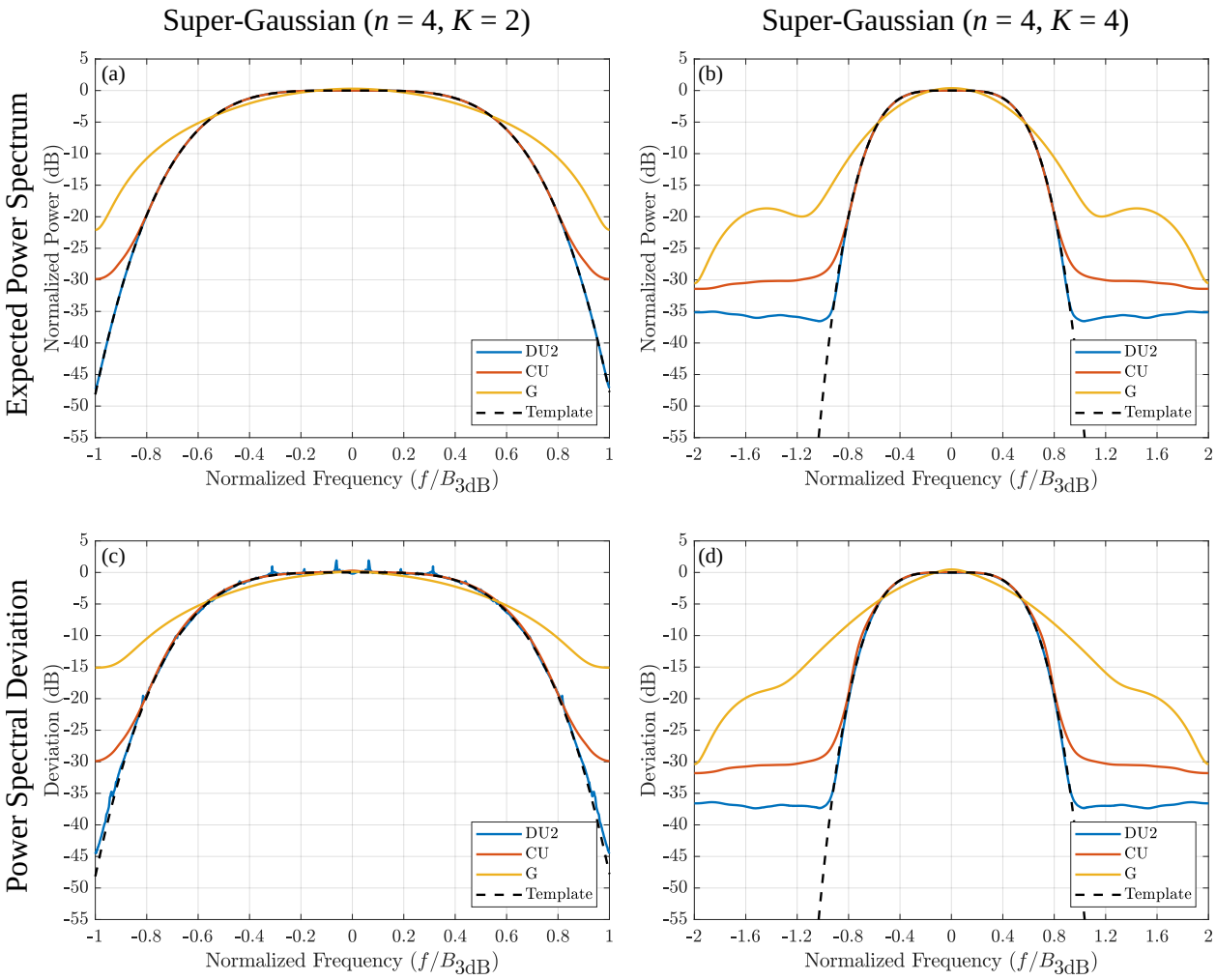

Figure 4.9: Pulsed StoWGe spectral optimization results for super-Gaussian templates: expected power spectrum for the S4G2 template (a), expected power spectrum for the S4G4 template (b), expected power spectral deviation for the S4G2 template (c), expected power spectral deviation for the S4G4 template (d)

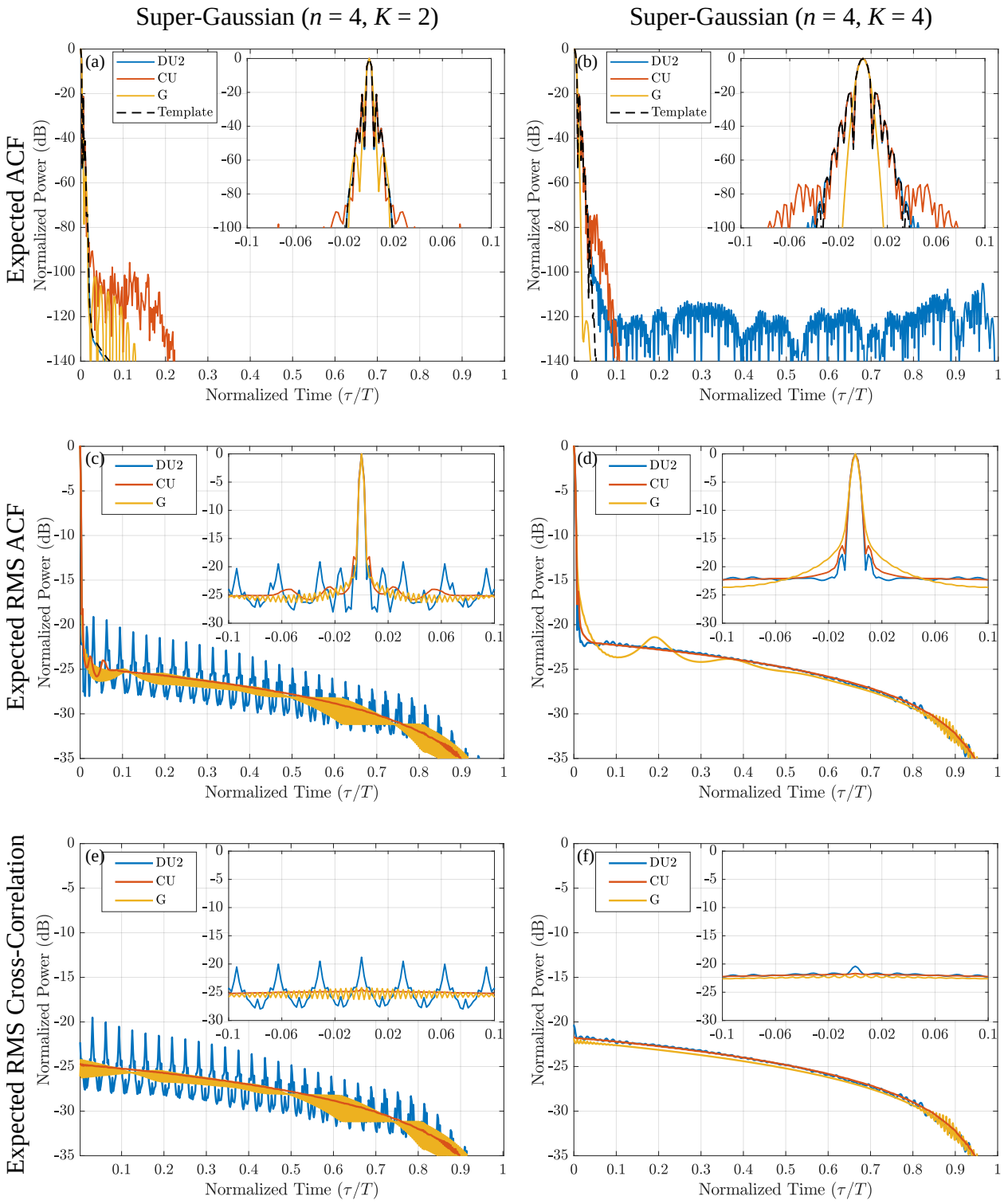

Figure 4.10: Pulsed StoWGe temporal optimization results for super-Gaussian templates: expected autocorrelation for the S4G2 template (a), expected autocorrelation for the S4G4 template (b), expected RMS autocorrelation for the S4G2 template (c), expected RMS autocorrelation for the S4G4 template (d), expected RMS cross-correlation for the S4G2 template (e), expected RMS cross-correlation for the S4G4 template (f)

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 to 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 -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 and leveling off at -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 implied from the expected spectrum. This indicates that from sample function to sample function, these edges are a little more consistent than the rest of the 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 spectra result in likewise poor matches to the template's expected autocorrelation.

Fig. 4.12(c,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 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(a,b).

The expected RMS cross-correlation also presents an interesting behavior in the DU2 distribu-

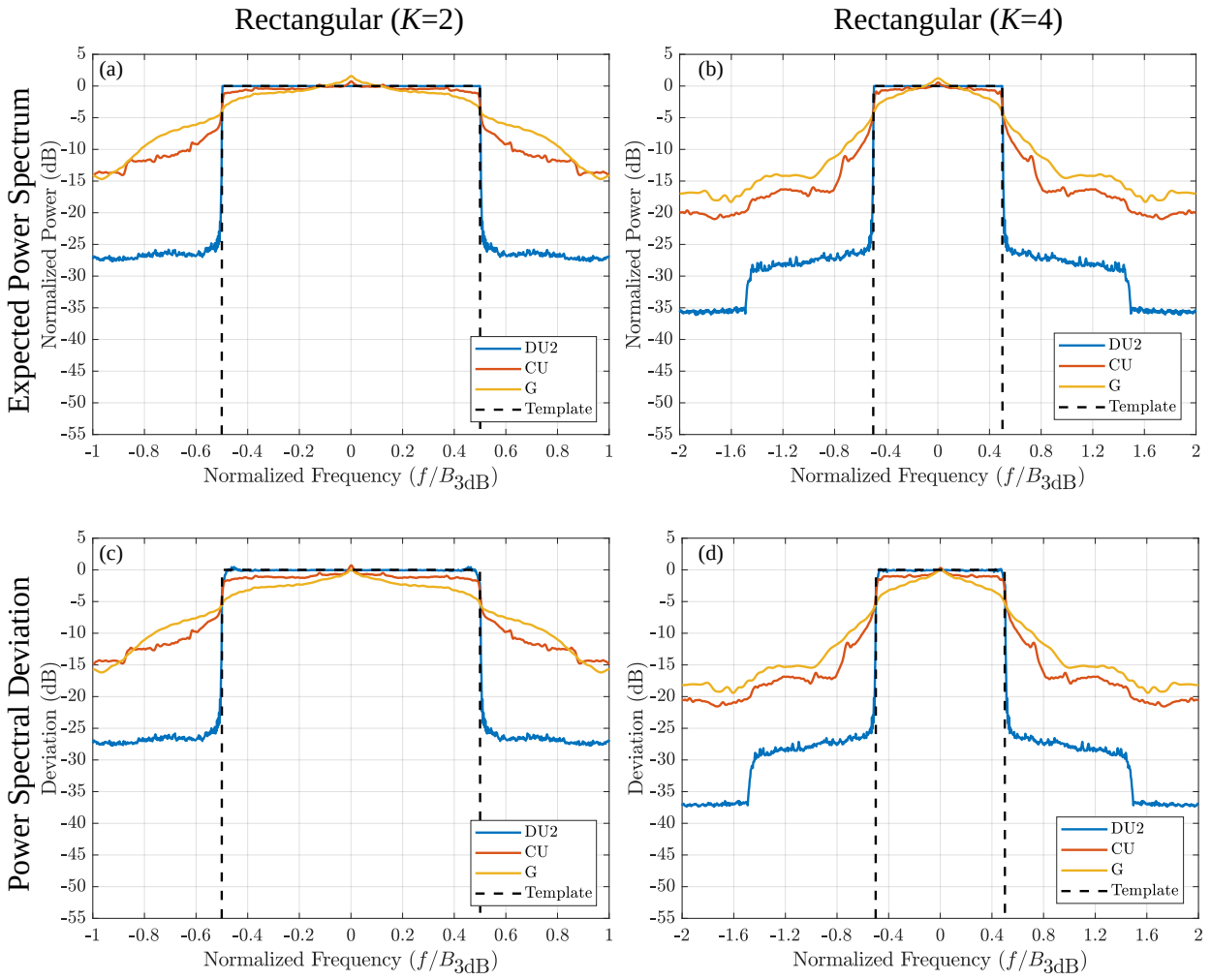

Figure 4.11: Pulsed StoWGe spectral optimization results for rectangular templates: expected power spectrum for the R2 template (a), expected power spectrum for the R4 template (b), expected power spectral deviation for the R2 template (c), expected power spectral deviation for the R4 template (d)

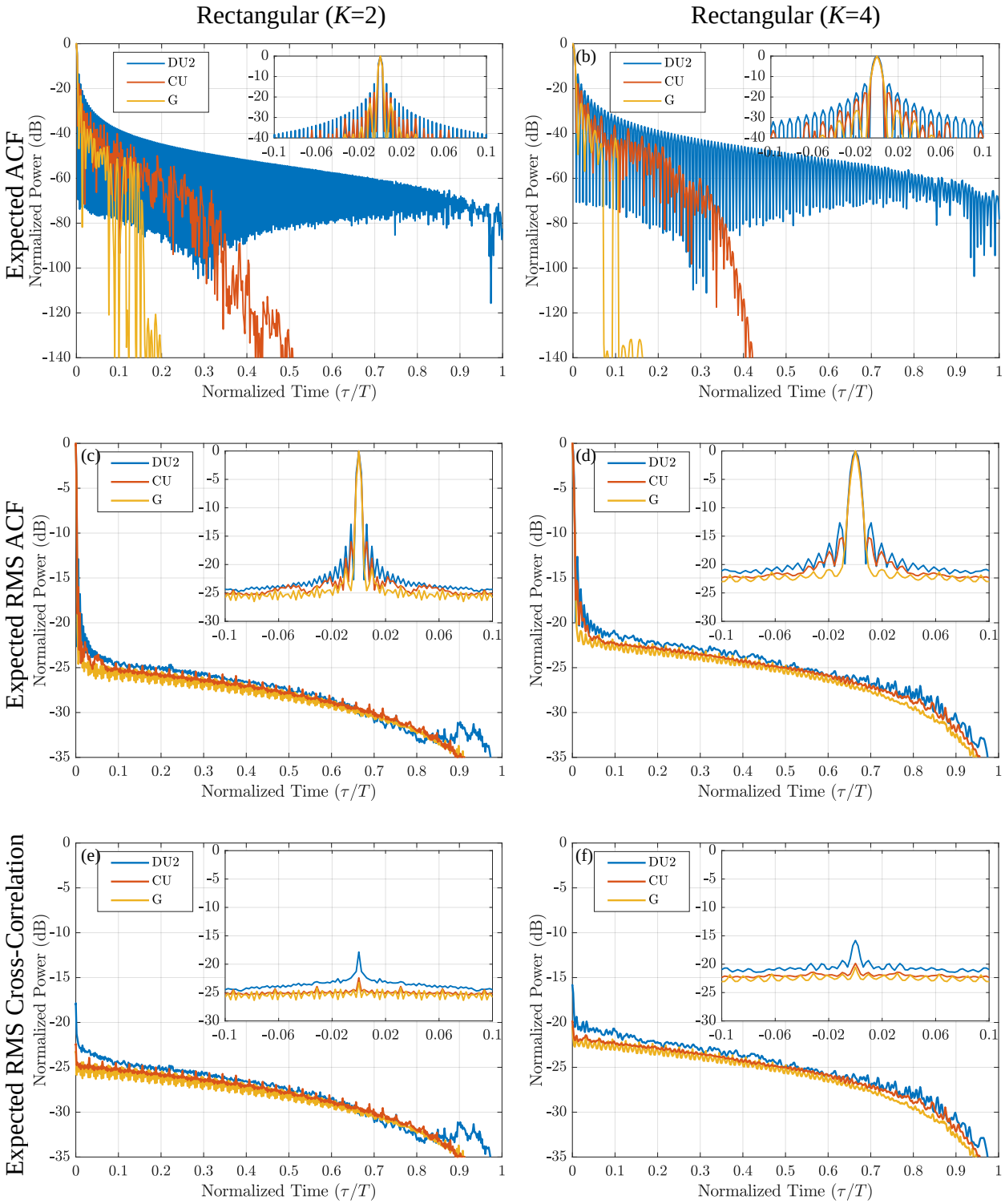

Figure 4.12: Pulsed StoWGe temporal optimization results for rectangular templates: expected autocorrelation for the R2 template (a), expected autocorrelation for the R4 template (b), expected RMS autocorrelation for the R2 template (c), expected RMS autocorrelation for the R4 template (d), expected RMS cross-correlation for the R2 template (e), expected RMS cross-correlation for the R4 template (f)

tion case. In Fig. 4.12(e,f), the DU2 trace has a correlation spike at 0 delay. It would appear this correlation might have been needed to achieve the good spectral matches shown in 4.11(a-b).

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 having 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 posses 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.

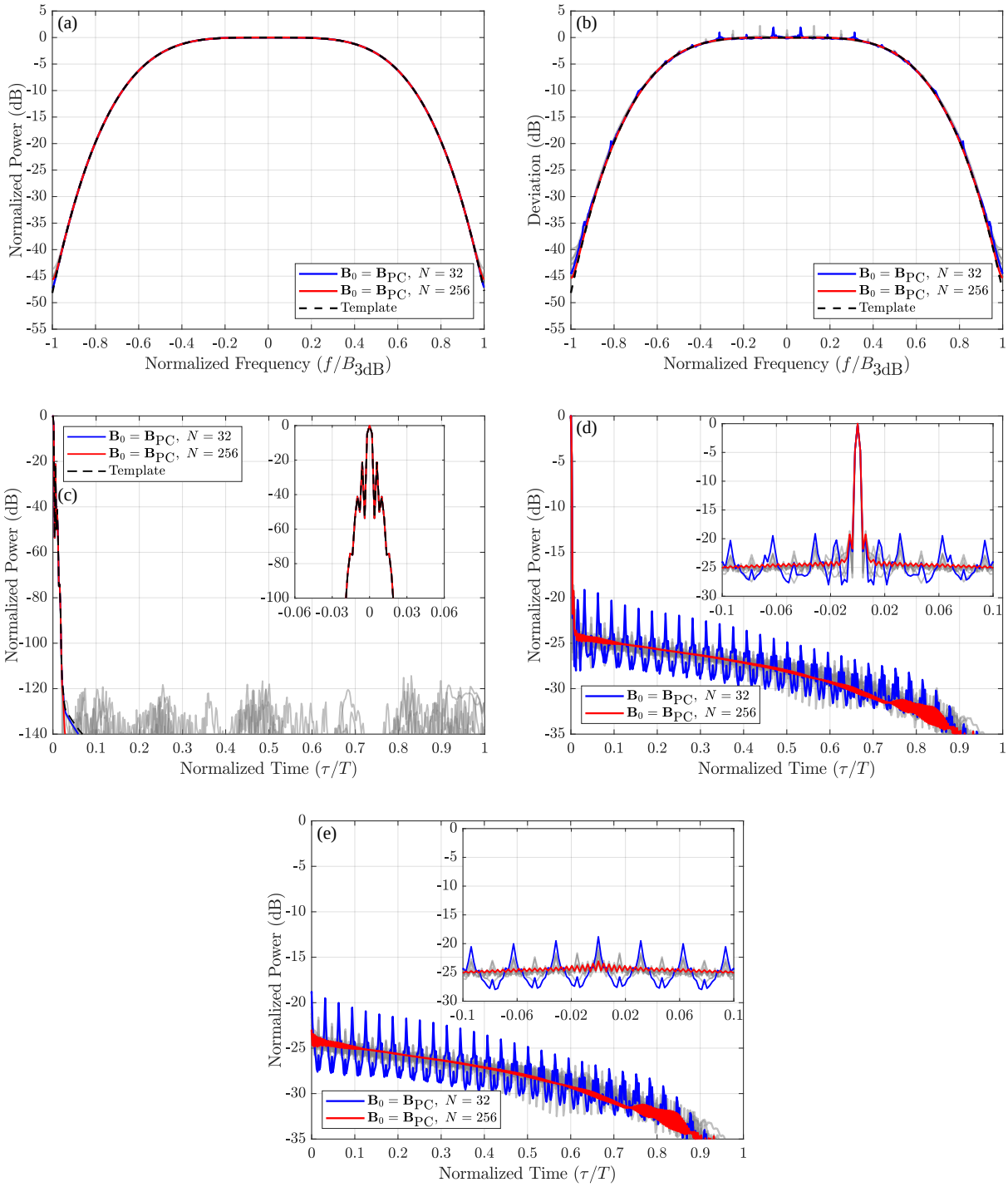

Figure 4.13: All results for the DU2 distribution and the S4G2 template: $N = 32$ in blue, $N = 256$ in red, expected power spectrum (a), expected power spectral deviation (b), expected autocorrelation (c), expected RMS autocorrelation (d), expected RMS cross-correlation (e)

4.4.3 Optimized Basis and Correlation Matrices

The previous section looked at how well the pulsed 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 would make the plots difficult to interpret. The correlation matrices are plotted in magnitude on a dB scale. Since the optimized WGFs have a 0 valued $\boldsymbol{\mu}$, the correlation matrices are real.

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 characteristic function of Gaussian distributions and the fact that $\boldsymbol{\mu}$ 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 still possible for the sum to result 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 template's ideal autocorrelation so the correlation must cancel itself out

when summed 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 in Fig. 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 attempt, this was not enough and the Gaussian distributed random variables 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 smooth basis functions, but with periodic spikes. If Fig. 4.10(b), the DU2 distribution produced the best match to the template autocorrelation though exhibiting 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)).

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).

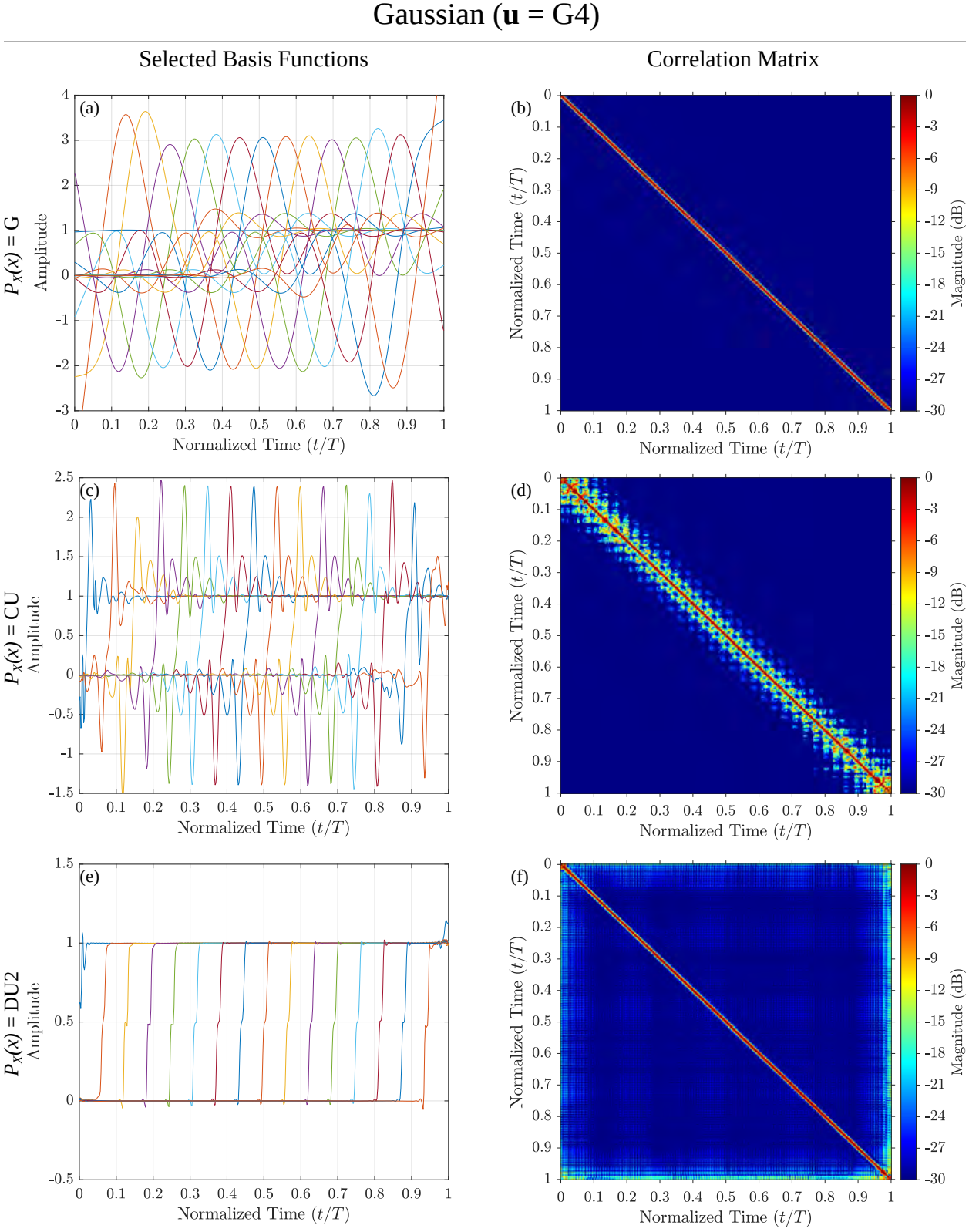

Figure 4.14: Optimized correlation matrices and selected basis functions for each distribution for the G4 template: Correlation matrices are plotted in magnitude on a dB scale. selected G basis functions (a), G correlation matrix (b), selected CU basis functions (c), CU correlation matrix (d), selected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

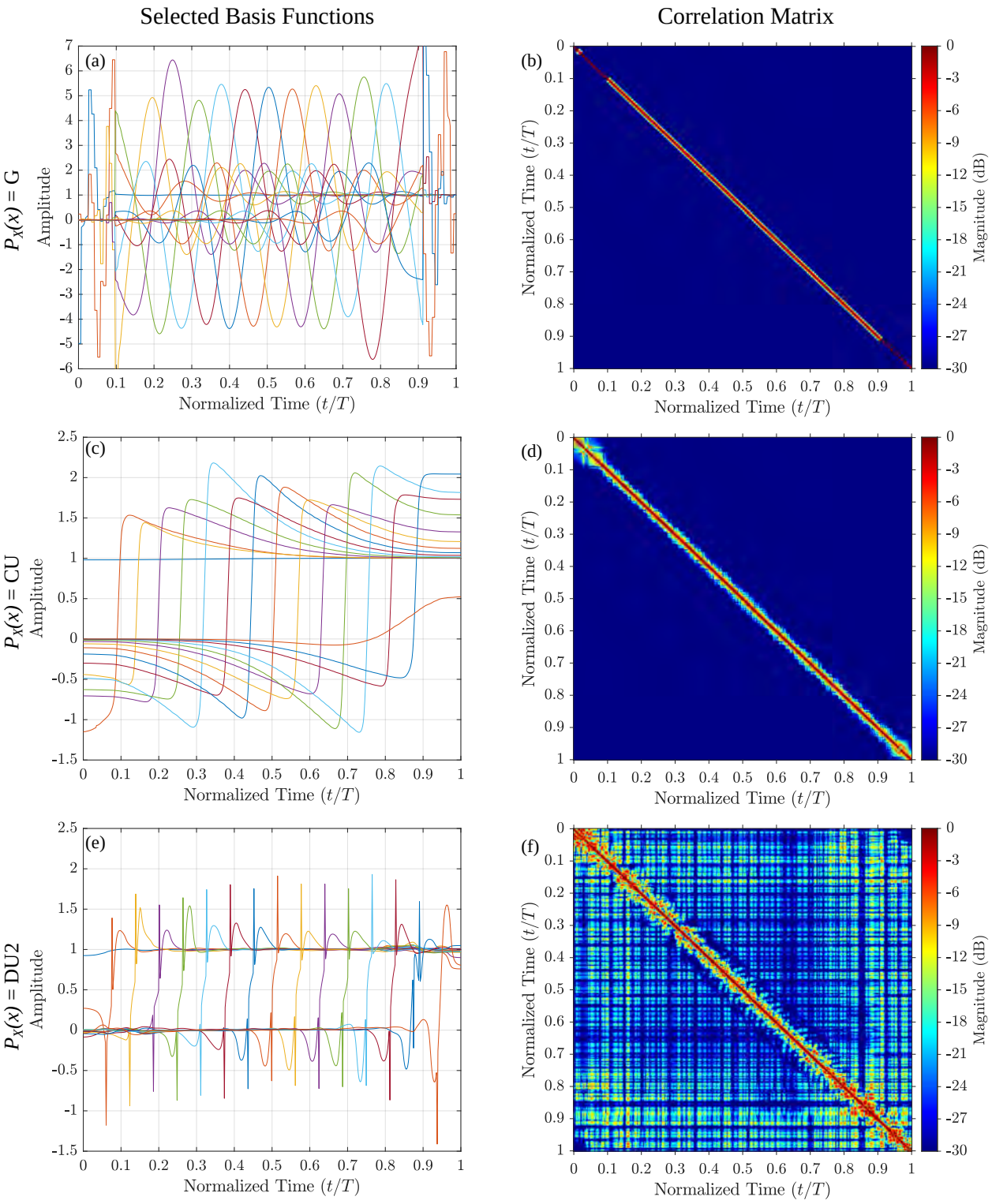

Figure 4.15: Optimized correlation matrices and selected basis functions for each distribution for the S4G4 template: Correlation matrices are plotted in magnitude on a dB scale. selected G basis functions (a), G correlation matrix (b), selected CU basis functions (c), CU correlation matrix (d), selected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

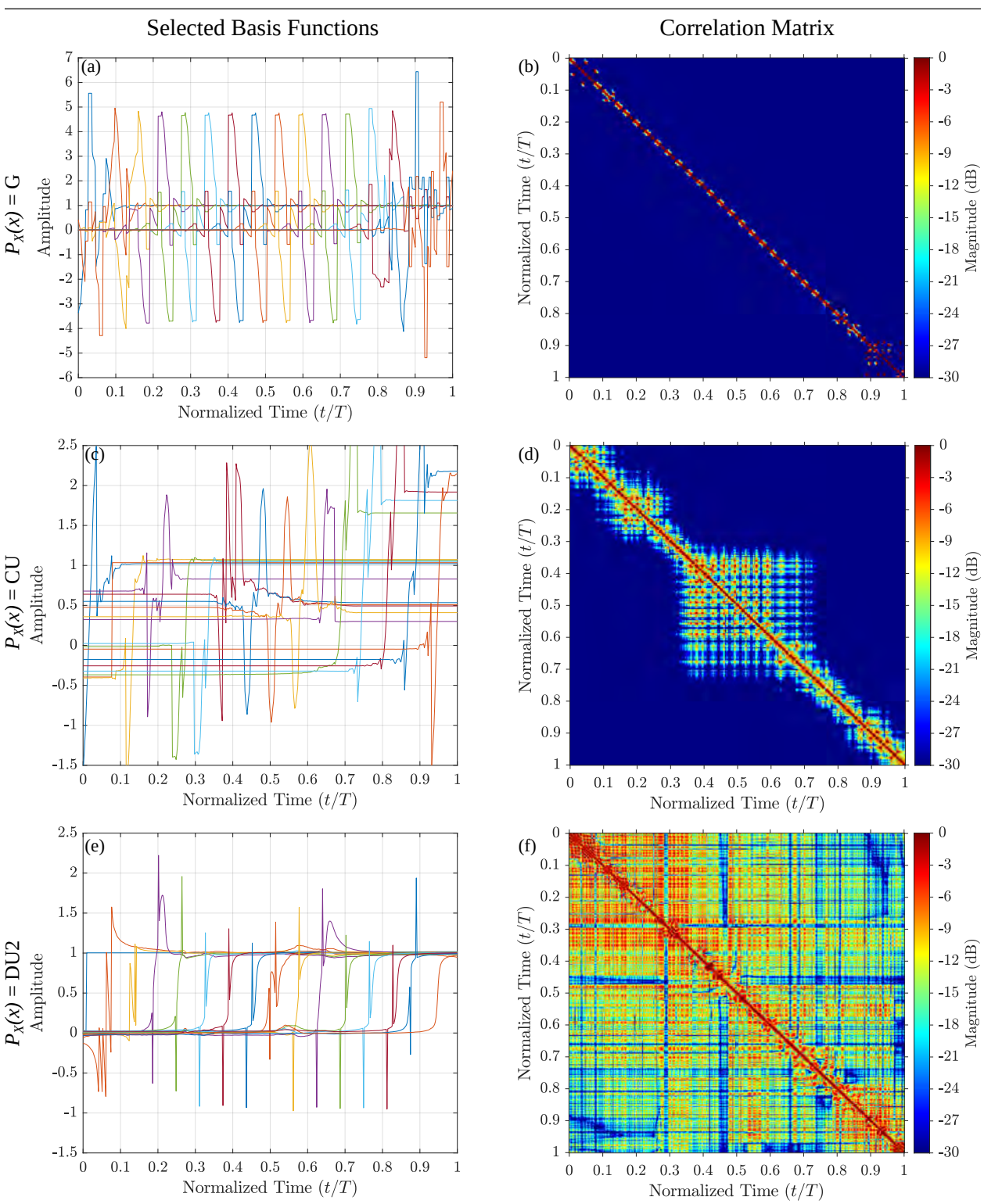

Figure 4.16: Optimized correlation matrices and selected basis functions for each distribution for the R4 template: Correlation matrices are plotted in magnitude on a dB scale. selected G basis functions (a), G correlation matrix (b), selected CU basis functions (c), CU correlation matrix (d), selected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

4.4.4 Pulsed StoWGe Comparisons to Previous RFM Waveforms

The previous sections established that the optimized DU2 random variable pulsed StoWGe WGFs generally produce the best radar waveforms as determined by various metrics. The goal of this section is to provide further context to those results by comparing those same DU2 based WGFs to previous random FM waveforms according to the same metrics. Thus far in the literature [23, 24, 71], the PRO-FM waveforms discussed in Section 2.2.4.1 have been used as a standard of comparison. They are used in the same way here. With this in mind, the goal of this section is not to establish whether or not these pulsed StoWGe waveforms are better than the PRO-FM waveforms specifically. Rather, the goal is to compare more generally the behavior of the pulsed StoWGe waveforms to that of other random FM waveforms. The comparison to the PRO-FM waveforms is especially useful since the PRO-FM have much more in common with other developed random FM waveforms than they do to the pulsed StoWGe waveforms. In this way, the general comparisons to the PRO-FM waveforms should be generally true when compared to other classes of random FM waveforms as well such as those developed in [23, 24, 71].

To perform the comparison, the DU2 pulsed StoWGe WGFs are compared to the PRO-FM waveforms in the same manner the DU2 pulsed StoWGe WGFs were compared to the G and CU pulsed StoWGe variants for each of the six different spectral templates. In Figs. 4.17-4.22, the DU2 traces were analytically calculated and are the same as those in Figs. 4.7-4.12. However, since there is no known way to analytically calculate the expected spectrum, the expected autocorrelation, and so on for the PRO-FM waveforms, these have to be estimated in a monte carlo fashion. To do so, 10000 PRO-FM waveforms were optimized for each of the desired spectral templates for a total of 60000 unique waveforms. These waveforms were also designed to have the same BT as the equivalent pulsed StoWGe waveforms. The metrics were then estimated via these waveforms.

4.4.4.1 Pulsed StoWGe and PRO-FM with zero Doppler

In 4.7-4.12, a few broad trends can be identified. First, looking at the spectral results in Figs. 4.7, 4.9, and 4.11 In general, the StoWGe waveforms achieve better spectral containment than the PRO-

FM waveforms. In each case, the expected spectrum of the StoWGe waveforms drops off more precipitously than that of the PRO-FM waveforms. However, there are other random FM variants that achieve a better spectral roll-off than the PRO-FM waveforms anyway [71]. With this in mind, a perhaps more consequential difference between them can be seen in the spectral deviation plots. In every case, for every spectral template, the PRO-FM waveforms have far less deviation across their passband. This means, on average, individual PRO-FM waveforms will be much closer to their expected spectrum than a given StoWGe waveform would be. The consequence of this behavior is evident in the RMS autocorrelation traces of Figs. 4.8, 4.10, and 4.12 where in every case, the PRO-FM waveforms can be expected to have lower autocorrelation sidelobes on a per waveforms basis than can be expected of the StoWGe waveforms. This makes intuitive sense. If an individual waveform is expected to be closer to its expected spectrum than otherwise, then its autocorrelation will likewise be closer to its expected autocorrelation than otherwise.

This is perhaps the most important difference between the StoWGe waveforms developed here and previous random FM waveforms. Simply put, the StoWGe waveforms were designed to have a good *average* spectrum. No constraint was placed on the spectrum of each individual waveform. The PRO-FM waveforms were *individually* designed each to have the desired spectrum. If that desired spectrum results in low autocorrelation sidelobes, then the key takeaway is that random FM waveforms which are designed on a per waveform basis will likewise have lower autocorrelation sidelobes on a per waveform basis as compared to random FM waveforms which were designed in aggregate. While this looks like a significant shortcoming of the StoWGe waveforms, this work never claims that the StoWGe waveforms achieve an exceptional per waveform autocorrelation performance. Instead, the key advantages of StoWGe is that the waveforms can be spectrally shaped, they are constant amplitude, but critically they are also extremely computationally cheap to instantiate. All the optimization is moved up front, whereas for PRO-FM and many other random FM instantiations, the optimization is both computationally expensive and must be performed for every single waveform.

Still, it would be objectively better if the StoWGe waveforms had an improved per waveform

autocorrelation performance. To do so, it may be possible to optimize directly the expected RMS autocorrelation or the spectral deviation, but for now such an approach is beyond the scope of this work.

As a final note, it is important to recall that the PRO-FM expected autocorrelation traces are simply estimates. The PRO-FM traces in Figs. 4.8(a,b), 4.10(a,b), and 4.12(a,b) would likely be even lower and perhaps match or exceed the StoWGe traces with a larger number of waveforms.

4.4.4.2 Pulsed StoWGe and PRO-FM Ambiguity Functions

Section 4.4 evaluated the optimization of the pulsed StoWGe waveforms for three different random variable distributions, but it only did so for zero Doppler metrics. In the interest of addressing the Doppler question, this section presents pulsed StoWGe expected and expected RMS ambiguity functions, but for the sake of conciseness only for the DU2 WGFs since these were generally shown in Section 4.4 to be superior to the G and CU distribution based WGFs. For clarity, only the ambiguity functions corresponding to the 4 times oversampled spectra are shown. Additionally, in the interest of comparing to previous random FM waveforms the pulsed StoWGe results are shown alongside equivalent PRO-FM results.

Similar to previous comparisons, the PRO-FM ambiguity functions are estimates based on the 10000 waveforms instantiated for each of the 4 times oversampled spectra. Since the ambiguity function is a measure of fast time Doppler, for each Doppler bin the corresponding doppler shift was applied to each waveform and was then filtered with the zero Doppler conjugate of itself as defined by the ambiguity function.

For the StoWGe waveforms, the ambiguity functions were calculated analytically by modifying their expected autocorrelation and expected RMS autocorrelations to include a Doppler shift. To do so, for a given doppler frequency, f_{d}, a Doppler shift vector, $\mathbf{d}$, can be formed such that

$$\mathbf{d} = \exp(j2\pi f_{\mathrm{d}}\mathbf{t}) \tag{4.30}$$

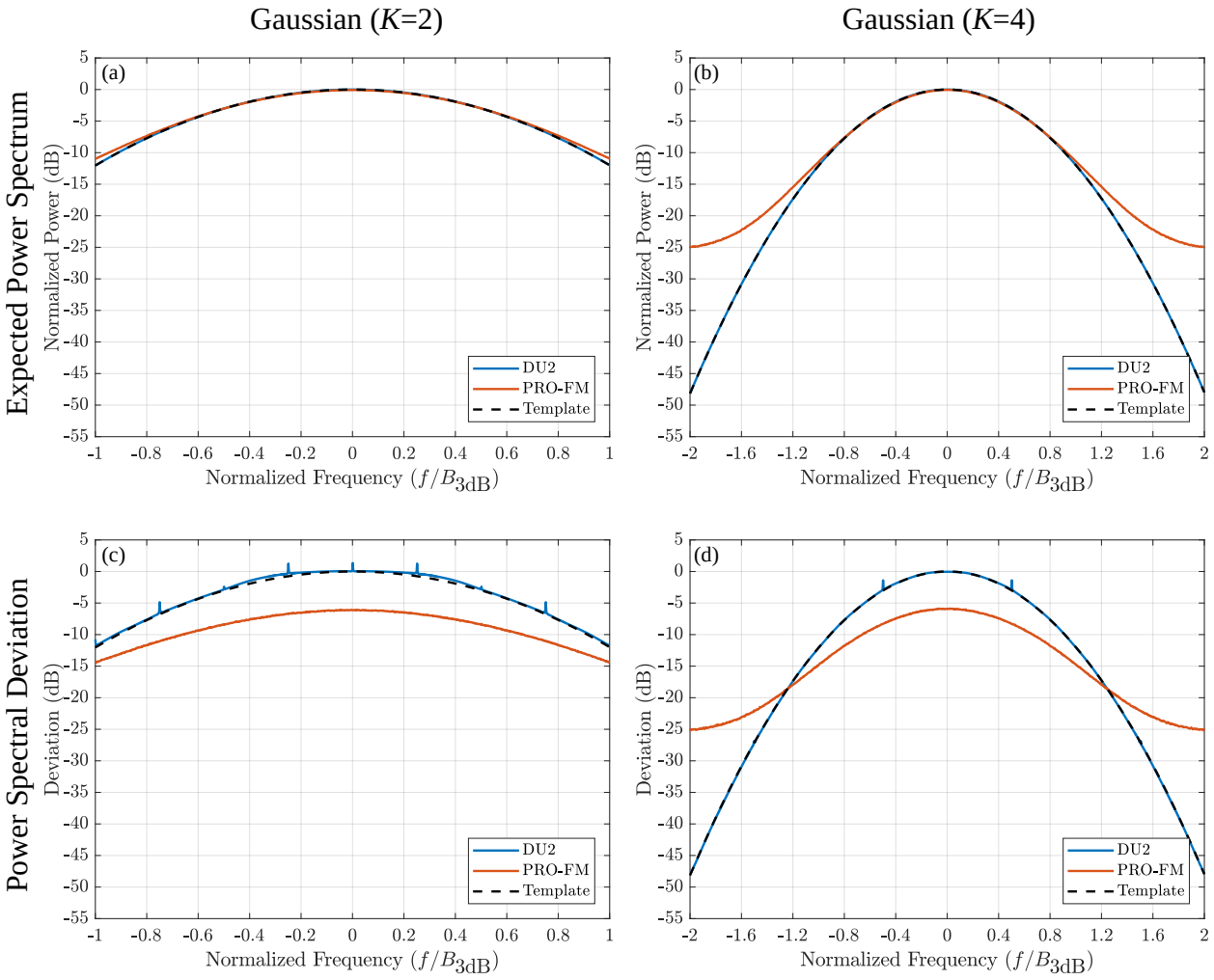

Figure 4.17: Comparison between the analytical spectral characteristics of optimized DU2 based pulsed StoWGe and estimated spectral PRO-FM characteristics, expected power spectrum for the G2 template (a), expected power spectrum for the G4 template (b), expected power spectral deviation for the G2 template (c), expected power spectral deviation for the G4 template (d)

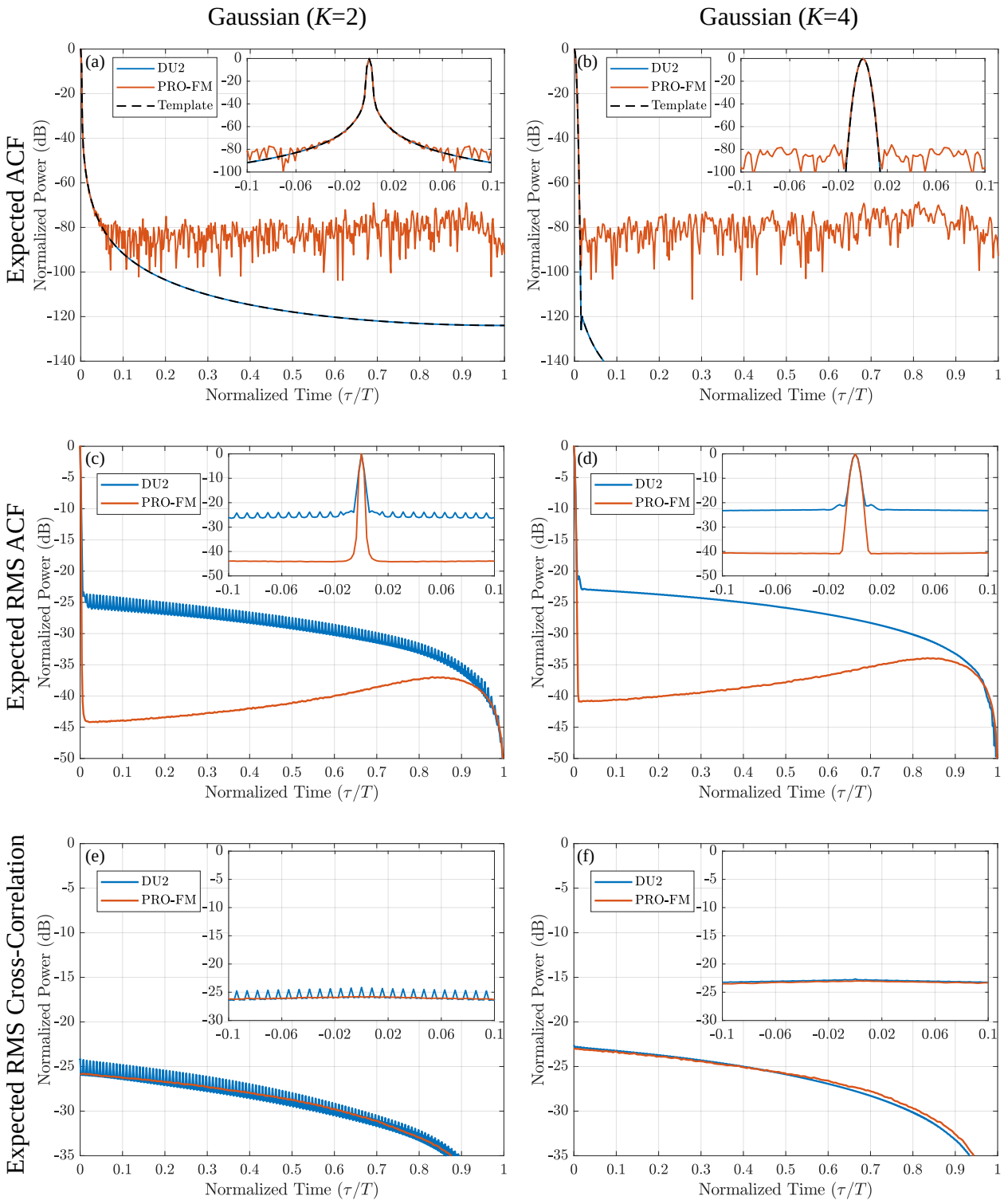

Figure 4.18: Comparison between the analytical temporal characteristics of optimized DU2 based pulsed StoWGe and estimated temporal PRO-FM characteristics, expected autocorrelation for the G2 template (a), expected autocorrelation for the G4 template (b), expected RMS autocorrelation for the G2 template (c), expected RMS autocorrelation for the G4 template (d), expected RMS cross-correlation for the G2 template (e), expected RMS cross-correlation for the G4 template (f)

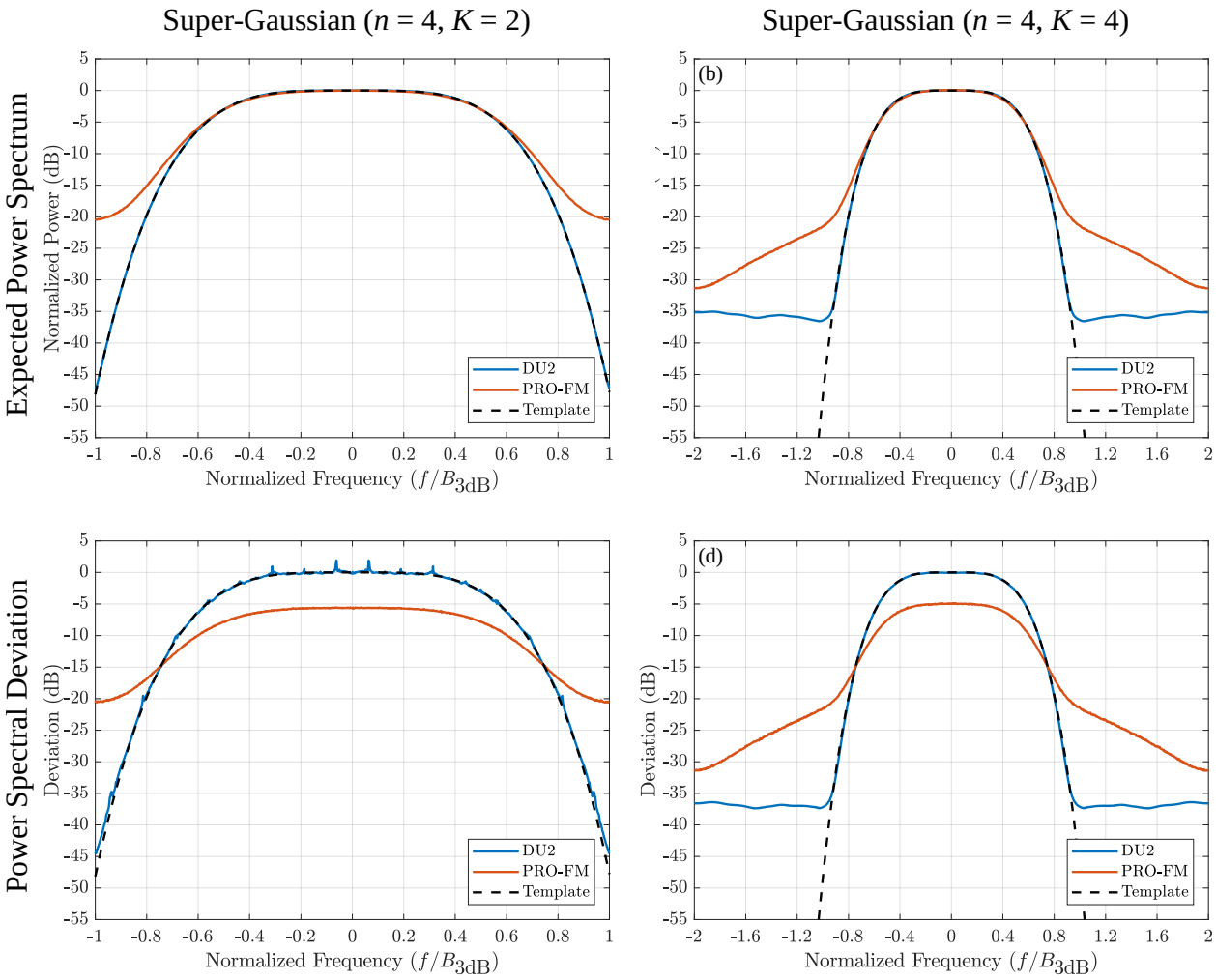

Figure 4.19: Comparison between the analytical spectral characteristics of optimized DU2 based pulsed StoWGe and estimated spectral PRO-FM characteristics, expected power spectrum for the S4G2 template (a), expected power spectrum for the S4G4 template (b), expected power spectral deviation for the S4G2 template (c), expected power spectral deviation for the S4G4 template (d)

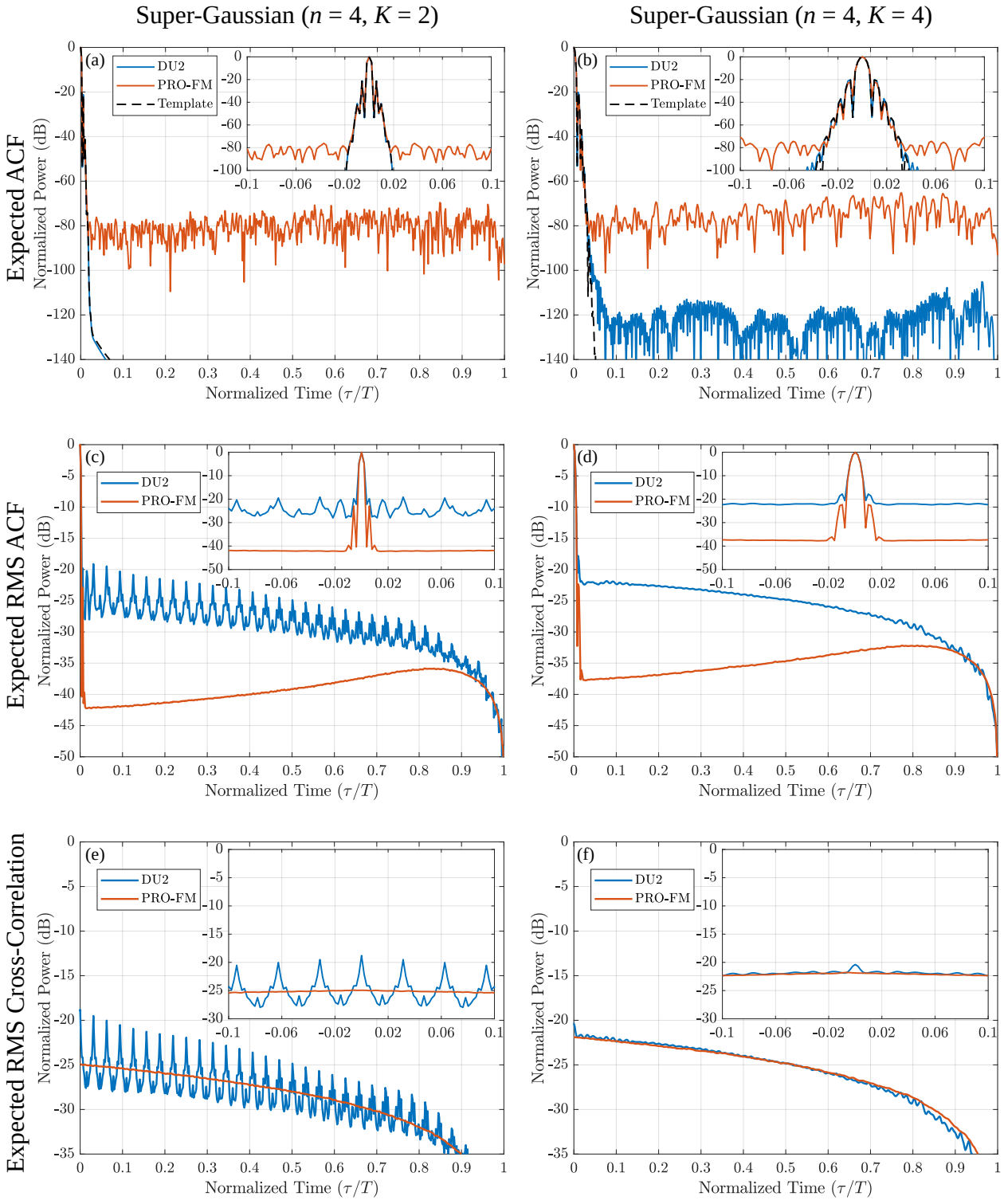

Figure 4.20: Comparison between the analytical temporal characteristics of optimized DU2 based pulsed StoWGe and estimated temporal PRO-FM characteristics, expected autocorrelation for the S4G2 template (a), expected autocorrelation for the S4G4 template (b), expected RMS autocorrelation for the S4G2 template (c), expected RMS autocorrelation for the S4G4 template (d), expected RMS cross-correlation for the S4G2 template (e), expected RMS cross-correlation for the S4G4 template (f)

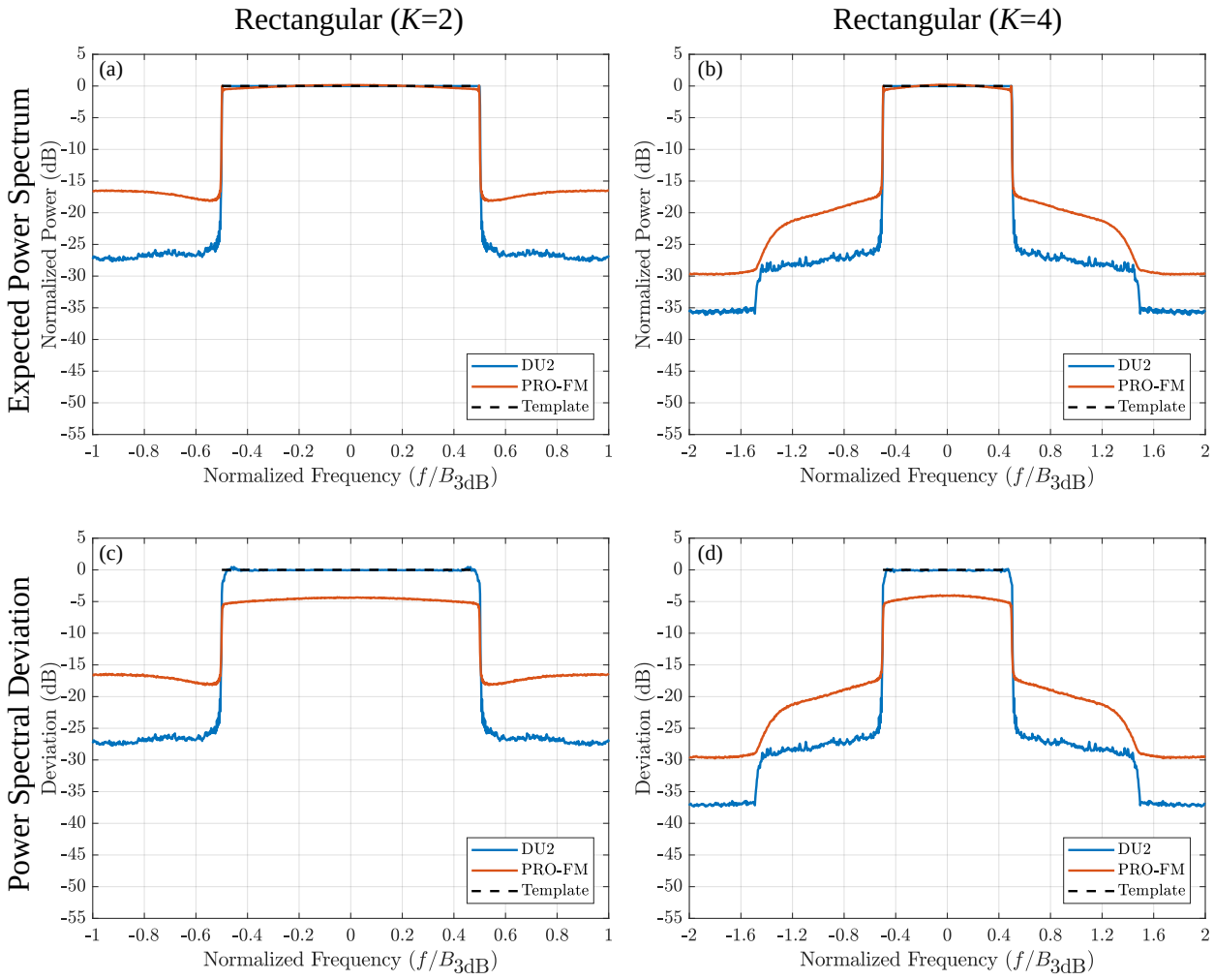

Figure 4.21: Comparison between the analytical spectral characteristics of optimized DU2 based pulsed StoWGe and estimated spectral PRO-FM characteristics, expected power spectrum for the R2 template (a), expected power spectrum for the R4 template (b), expected power spectral deviation for the R2 template (c), expected power spectral deviation for the R4 template (d)

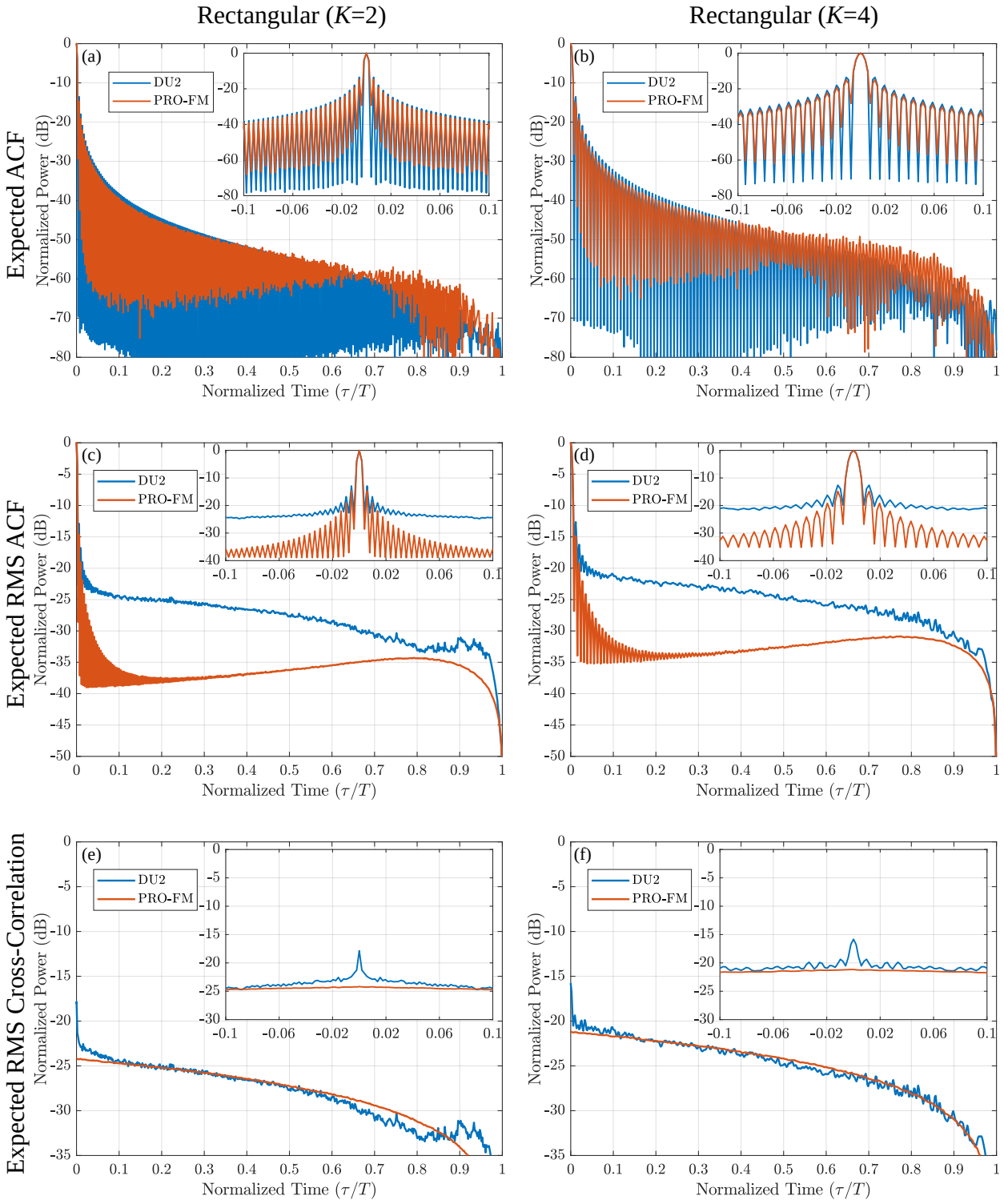

Figure 4.22: Comparison between the analytical temporal characteristics of optimized DU2 based pulsed StoWGe and estimated temporal PRO-FM characteristics, expected autocorrelation for the R2 template (a), expected autocorrelation for the R4 template (b), expected RMS autocorrelation for the R2 template (c), expected RMS autocorrelation for the R4 template (d), expected RMS cross-correlation for the R2 template (e), expected RMS cross-correlation for the R4 template (f)

where $\mathbf{t}$ is an $M \times 1$ normalized time vector. Despite a Doppler shift, the expected autocorrelation for a pulsed StoWGe can be calculated in a similar manner as before where an IFFT is taken over the power spectral density. However, with a Doppler shift it is not longer a power spectral density. Instead, the IFFT is taken with respect to a cross spectral density such that

$$\mathrm{E}\left[\hat{\mathbf{r}}_{f_\mathrm{d}}\right] = \mathbf{A}^H \mathrm{E}\left[\mathbf{A}\mathbf{s} \odot \left(\mathbf{A}\left(\mathbf{s} \odot \mathbf{d}\right)\right)^*\right] \tag{4.31}$$

With respect to the expectation operator, the Doppler shift can simply be pulled out and the Doppler shifted second moment simply becomes

$$d_{m_2}^* C_{m_1,m_2} = d_{m_2}^* \mathrm{E}\left[s_{m_1} s_{m_2}^*\right] = d_{m_2}^* \prod_{n=1}^{N} \psi_{X_n} \exp\left(j\left(\mu_{m_1} - \mu_{m_2}\right)\right)\left(b_{m_1,n} - b_{m_2,n}\right) \tag{4.32}$$

which allows for the calculation of the expected ambiguity function.

Similarly, the Doppler shift can also be pulled out of the Doppler shifted RMS autocorrelation such that

$$\mathrm{E}\left[\left|\hat{r}_{\ell,f_\mathrm{d}}\right|^2\right] = d_{m_1}^* d_{m_2} \sum_{m_1,m_2=1}^{\ell} \mathrm{E}\left[s_{M-\ell+m_1} s_{m_1}^* s_{M-\ell+m_2}^* s_{m_2}\right] \tag{4.33}$$

Taken together, these equations allow for the analytical calculation of the plots on the left sides of Figs. 4.23 and 4.24 respectively.

As was shown in Fig. 2.20, RFM waveforms should be expected to exhibit a thumbtack like ambiguity function with little Doppler tolerance. In both Figs. 4.23 and 4.24 this is clearly the case. Fig. 4.23 shows the analytical expected autocorrelations of the StoWGe waveforms in Fig. 4.23(a,c,e) while the estimated PRO-FM results are shown in 4.23(b,d,f). In the PRO-FM case it cannot be definitively said what the ambiguity functions will do with further coherent integration, but since they are noise like waveforms it is reasonable to expect the further minimization of energy off from the mainlobe making the response even more thumbtack like.

For the pulsed StoWGe waveforms, there is no such question. Since the plots in Fig. 4.23(a,c,e) represent the analytical expectation of the ambiguity function, there is clearly a floor to the ambi-

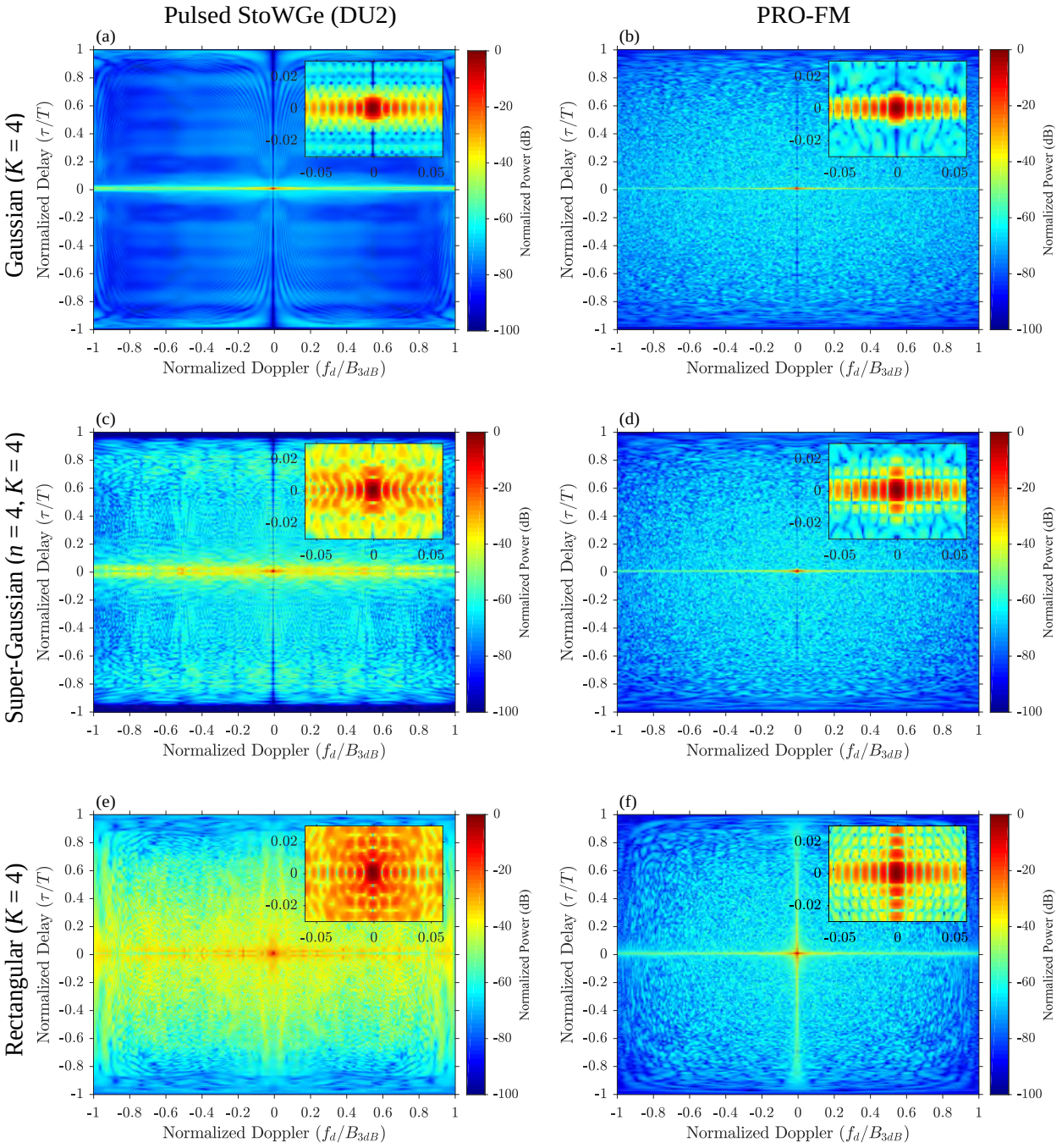

Figure 4.23: Comparison between the analytical ambiguity functions of optimized DU2 based pulsed StoWGe and estimated ambiguity functions for PRO-FM: pulsed StoWGe for the G4 template (a), PRO-FM for the G4 template (b), pulsed StoWGe for the S4G4 template (c), PRO-FM for the S4G4 template (d), pulsed StoWGe for the R4 template (e), PRO-FM for the R4 template (f)

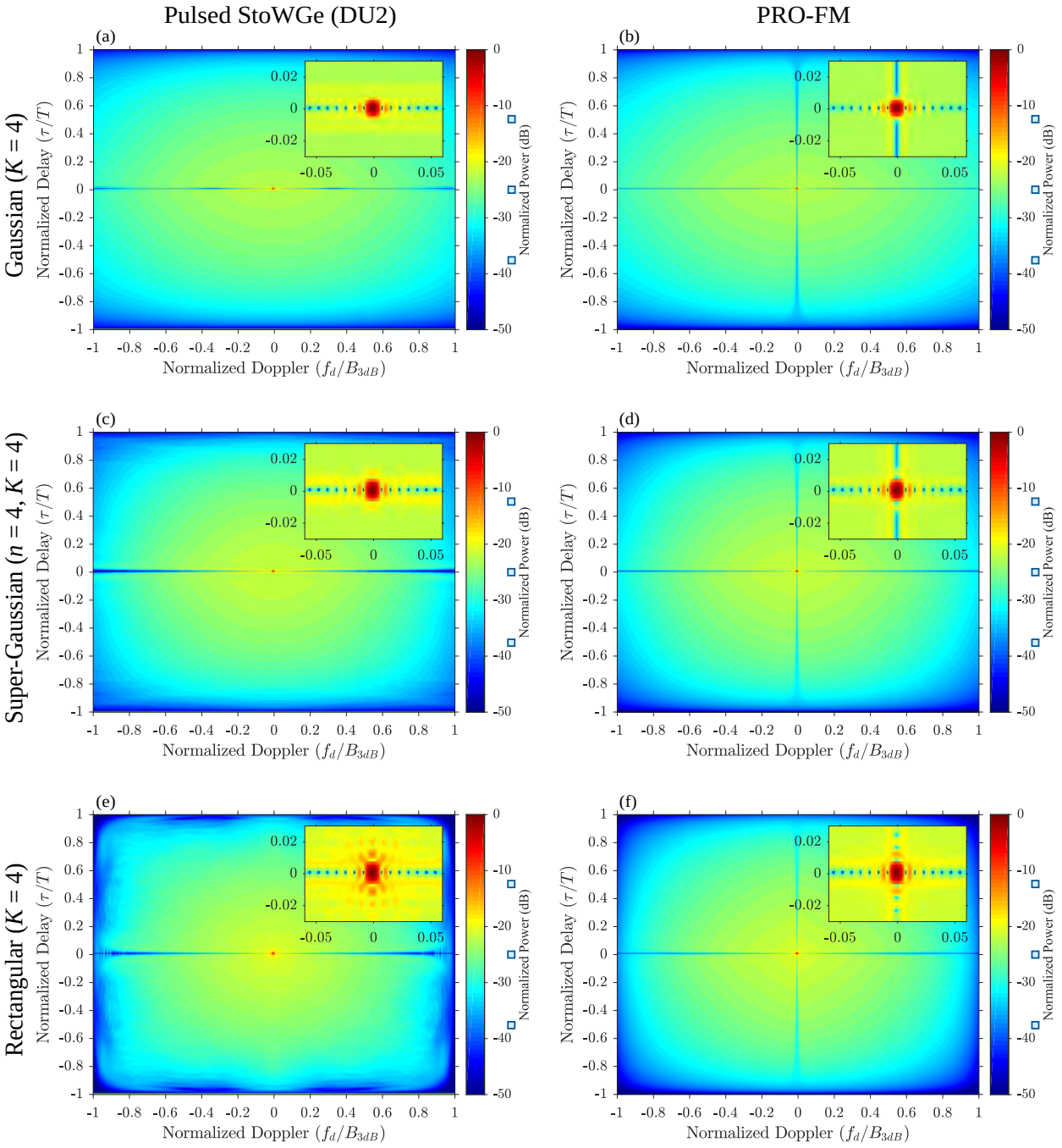

Figure 4.24: Comparison between the analytical RMS ambiguity functions of optimized DU2 based pulsed StoWGe and estimated RMS ambiguity functions for PRO-FM: pulsed StoWGe for the G4 template (a), PRO-FM for the G4 template (b), pulsed StoWGe for the S4G4 template (c), PRO-FM for the S4G4 template (d), pulsed StoWGe for the R4 template (e), PRO-FM for the R4 template (f)

guity surface that is lower for the G4 template, then higher for the S4G4 template, and higher yet for the R4 template. These floors will not decrease with coherent integration. Importantly, as can be seen from the insets of the mainlobe in each plot, these floor only exist off of the zero-Doppler axis.

The expected RMS ambiguity plots in Fig. 4.24 are perhaps easier to interpret since there is a much higher decree of homogeneity between the subplots. The most obvious difference between the StoWGe results and the PRO-FM results is the presence of a zero-Doppler valley of sorts in the case of the PRO-FM waveforms, which is a direct consequence of the lower RMS autocorrelation sidelobes of the PRO-FM waveforms shown in Figs. 4.18(d), 4.20(d), and 4.22(d). Other than that, the StoWGe and PRO-FM results are remarkably similar for each of the three desired spectra.

4.5 Experimental Results

As important as modeling and simulation are to any design problem there is simply no substitute for experimental demonstration. In this section, the analytical models of the optimized pulsed StoWGe WGFs shown in Section 4.4.2 are compared directly to both simulated estimates of the WGF properties as well as physically measured signals which were implemented on hardware in a loopback configuration. Since the DU2 based WGFs have been shown to be the best generally speaking, they are the ones considered here.

In each of Figs. 4.25-4.30 there are three traces. The first is the "analytical" trace. These traces are identical to the ones in 4.4.2 and they were calculated directly using the definition of the pulsed StoWGe model and the optimized WGFs for the metrics defined in Section 3.4.

The second is the "simulation" trace. For each of the DU2 based pulsed StoWGe WGFs shown in detail in Section 4.4.2, 1000 waveforms were instantiated for each of the desired spectra for a total of 6000 waveforms. In the same manner as for the PRO-FM waveforms in Section 4.4.4, the different random FM metrics were estimated on a computer.

Finally, the third is the "loopback" trace. Each of the 6000 waveforms used to create the simulated results were up-sampled by a factor of 12.5 using sinc interpolated, and then they were

projected onto a constant modulus envelope. They were then digitally up-converted to a center frequency of -$f_s/4$. Given the system DAC rate of 2.5 GHz, the signals have a 3 dB bandwidth of either 100 MHz or 50 MHz for the 2 times oversampled or the 4 times oversampled spectrum respectively and a center frequency of 625 MHz in either case. The signals were then implemented on an arbitrary waveform generator, linearly amplified, and subsequently attenuated and recorded at passband at a rate of 2.5 GHz. The recorded signals were then digitally down converted and re-sampled down to the baseband sample rate of 200 MHz. The different random FM metrics were then estimated from this data.

Speaking generally about Figs. 4.25-4.30, the analytical results and the simulated results match very well. This is especially true in spectral results where the blue trace, the analytical results, are almost completely hidden by the orange trace, the simulation results, but they are not quite the same since the orange traces only represents an estimate. This discrepancy is most apparent in the expected autocorrelation plots of Figs. 4.26(a,b), 4.28(a,b), and 4.30(a,b) where there is a clear divergence in the sidelobe regions between the analytical and the simulated traces, but this difference is expected. The analytical expected autocorrelation shows what happens when an infinite number of waveform autocorrelations are coherently integrated, but only 1000 waveforms were actually integrated in each simulated case. Consequently, on a dB scale, the difference is readily apparent where the estimation error is large compared to the magnitude of the analytical value such as in the sidelobe regions of the expected autocorrelation. With more and more integration it can be expected that the simulated traces will approach the analytical results. In contrast to the expected autocorrelation plots, the RMS autocorrelation simulated traces provide an excellent match to the analytical results as in the spectral cases. The key takeaway in comparing the analytical to the simulated results is that the mathematical characterization of the WGF accurately describes the statistics of the random FM waveforms that it produces.

In order to determine whether or not the simulated results reflect physically realizable waveforms, those same waveforms have been implemented in a loopback configuration. In contrast to the simulated results, there is a more notable difference between the loopback and analytical spec-

tral traces, but these differences can likely be attributed to the upsampling process and filtering effects.

First, the upsampling distortion is most obvious in the out of band regions of 4.25(b) and 4.25(a,b), where the normalized power of the loopback traces exceed that of the analytical or the simulated traces at least briefly. The initial sinc interpolation is ideal in that the baseband sample bandwidth is perfectly preserved at the higher sample rate, but projecting the signal back on a constant amplitude envelope will inevitably distort the spectrum to some degree. Nevertheless, this step is necessary to maintain an ideally constant amplitude signal.

Second, filtering effects are most obviously present in 4.29(a,b) where there is a clear ripple in the spectra across the passband. These affects are present in the other spectra as well, but the flat response of the rectangular spectral shapes makes it especially conspicuous. Without other information, it is ambiguous whether this ripple is a filtering effect or an artifact of the implementation of the StoWGe waveforms; however, the PRO-FM waveforms were also implemented in loopback in the same manner. As can be see in Fig. 4.31 the loopback PRO-FM waveforms exhibit similar passband distortion as in the pulsed StoWGe case. Previous experimentation with PRO-FM waveforms has not shown this effect indicating the distortion is hardware related and not a fundamental characteristic of the waveforms under test [26]. Apart from these effects, the loopback traces largely reflect the analytical and simulated results.

Interestingly, the spectral distortion has had little effect on the temporal results. In Figs. 4.26, 4.28, and 4.30, there is very little difference between the simulated and loopback traces further validating the analytical description of the waveforms as well as their suitability for physical implementation.

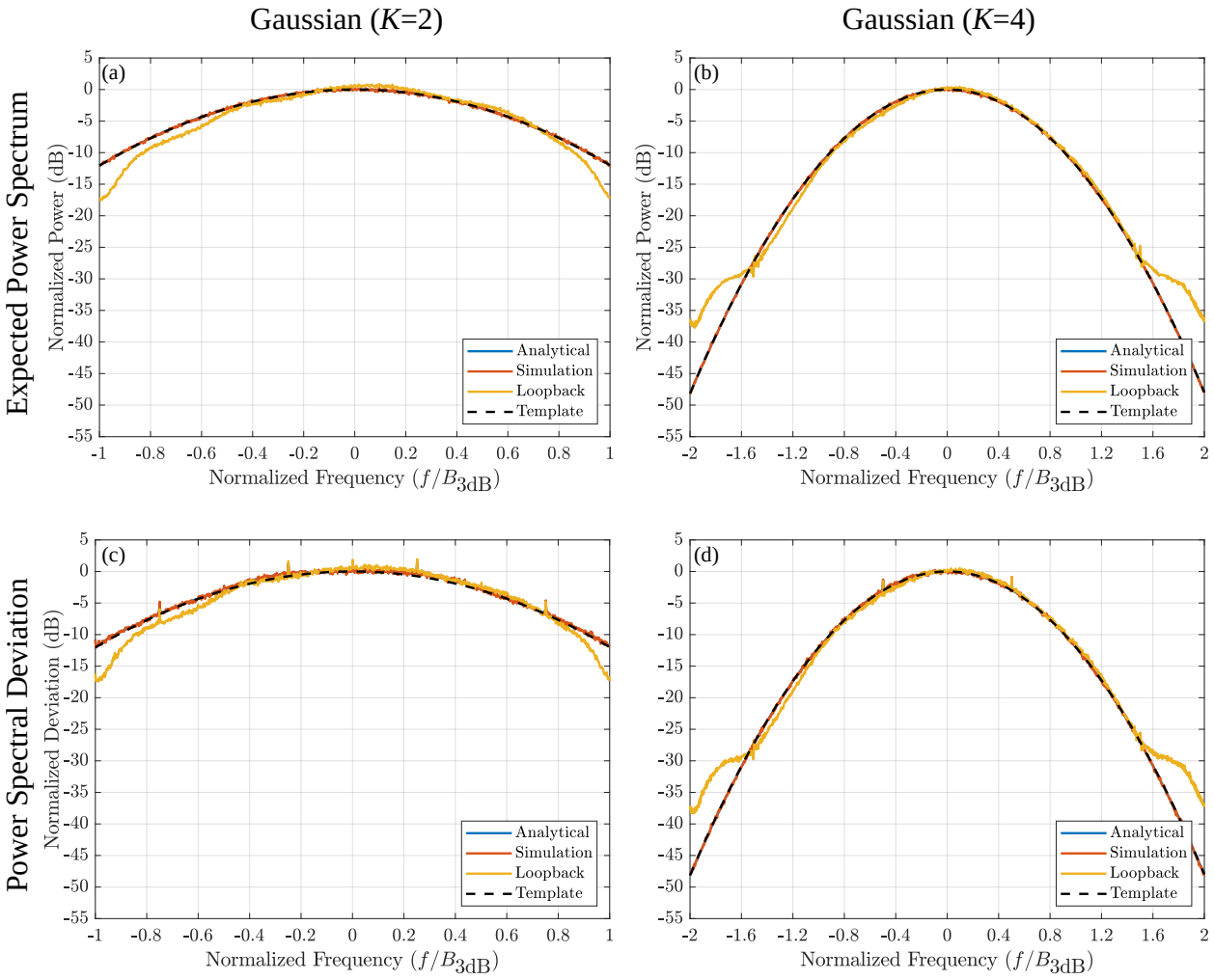

Figure 4.25: Optimized, DU2 based, pulsed StoWGe analytical, simulated, and loopback spectral results for Gaussian templates: expected power spectrum for the G2 template (a), expected power spectrum for the G4 template (b), expected power spectral deviation for the G2 template (c), expected power spectral deviation for the G4 template (d)

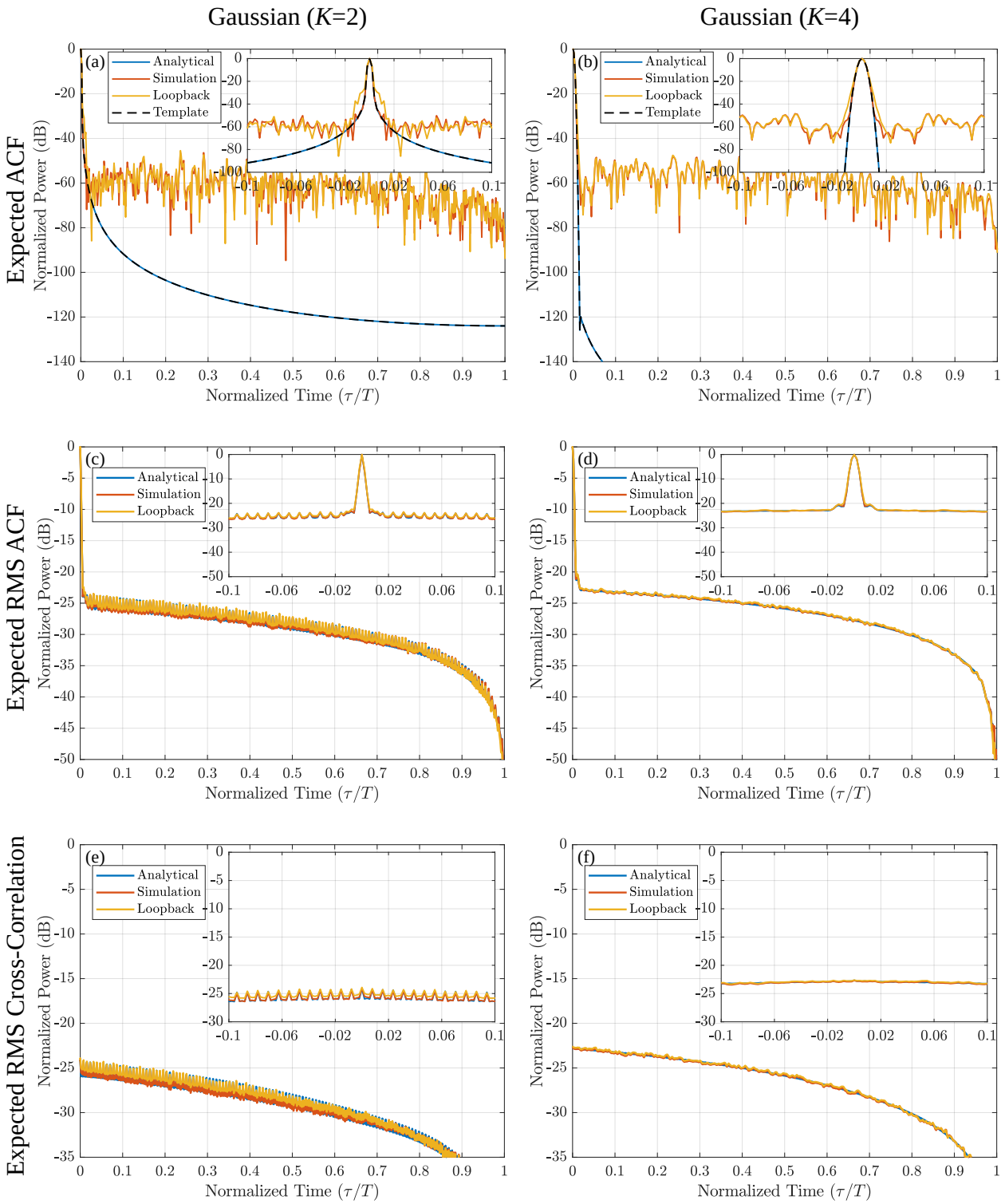

Figure 4.26: Optimized, DU2 based, pulsed StoWGe analytical, simulated, and loopback temporal results for Gaussian templates: expected autocorrelation for the G2 template (a), expected autocorrelation for the G4 template (b), expected RMS autocorrelation for the G2 template (c), expected RMS autocorrelation for the G4 template (d), expected RMS cross-autocorrelation for the G2 template (e), expected RMS cross-autocorrelation for the G4 template (f)

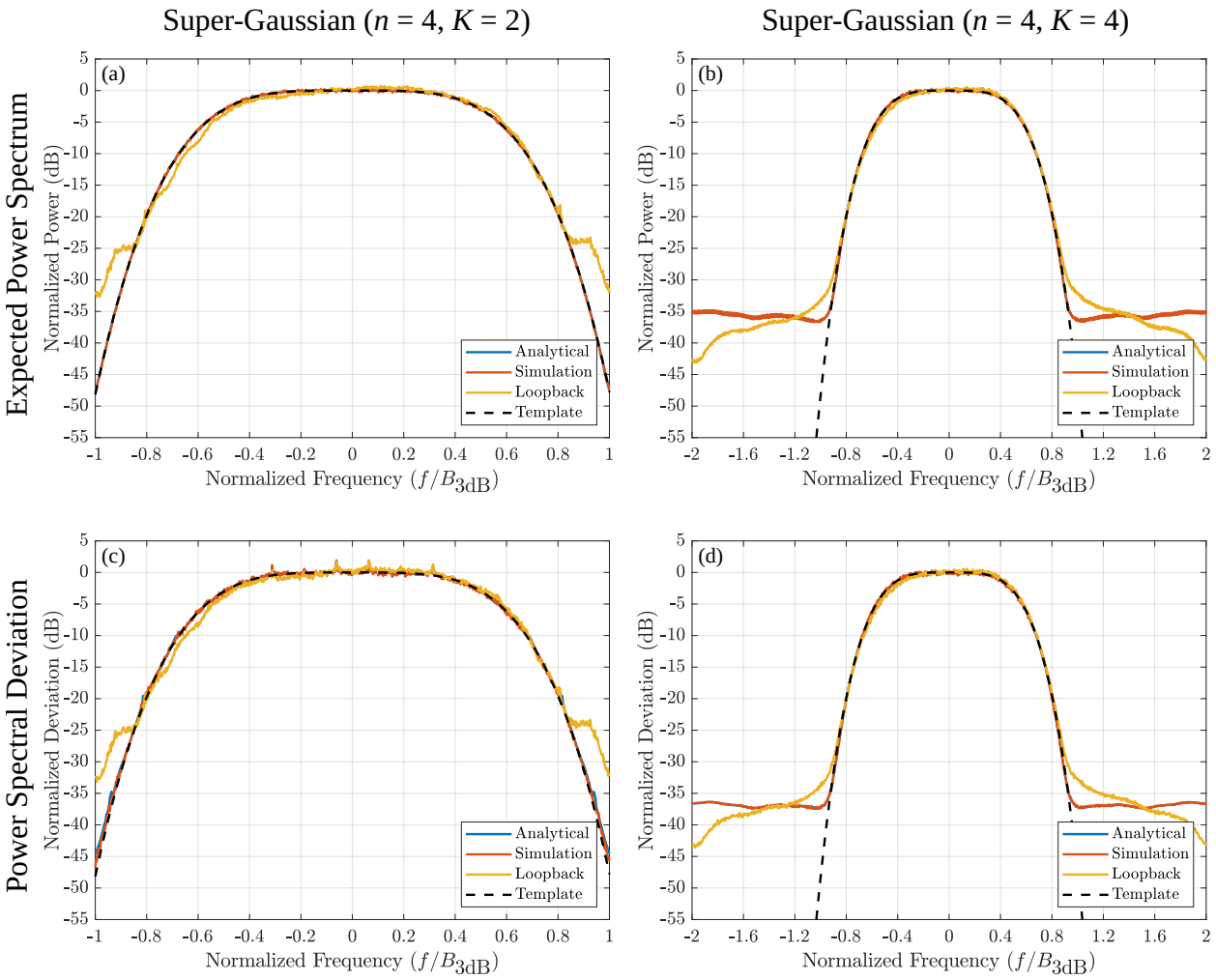

Figure 4.27: Optimized, DU2 based, pulsed StoWGe analytical, simulated, and loopback spectral results for super-Gaussian templates: expected power spectrum for the S4G2 template (a), expected power spectrum for the S4G4 template (b), expected power spectral deviation for the S4G2 template (c), expected power spectral deviation for the S4G4 template (d)

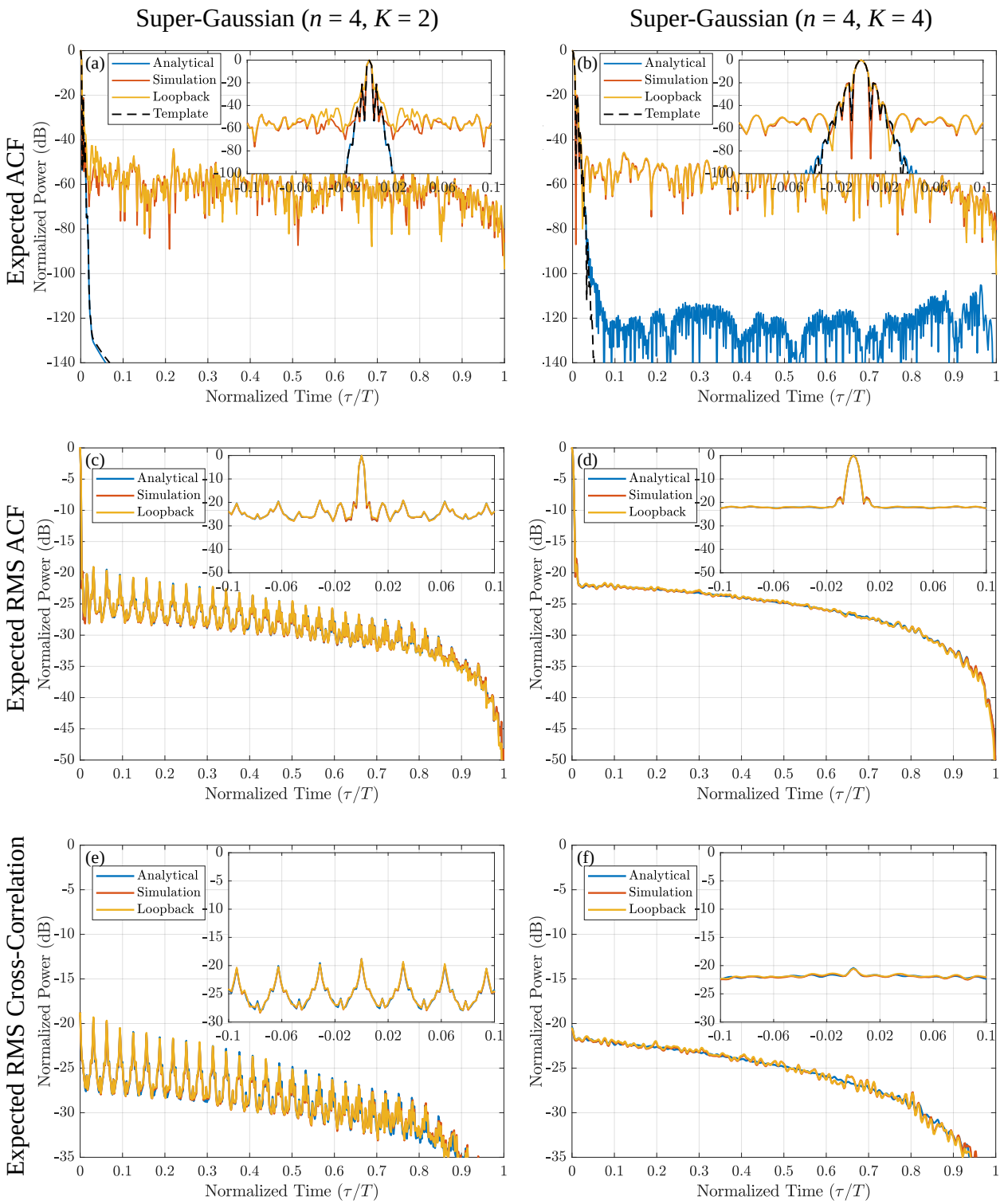

Figure 4.28: Optimized, DU2 based, pulsed StoWGe analytical, simulated, and loopback temporal results for super-Gaussian templates: expected autocorrelation for the S4G2 template (a), expected autocorrelation for the S4G4 template (b), expected RMS autocorrelation for the S4G2 template (c), expected RMS autocorrelation for the S4G4 template (d), expected RMS cross-autocorrelation for the S4G2 template (e), expected RMS cross-autocorrelation for the S4G4 template (f)

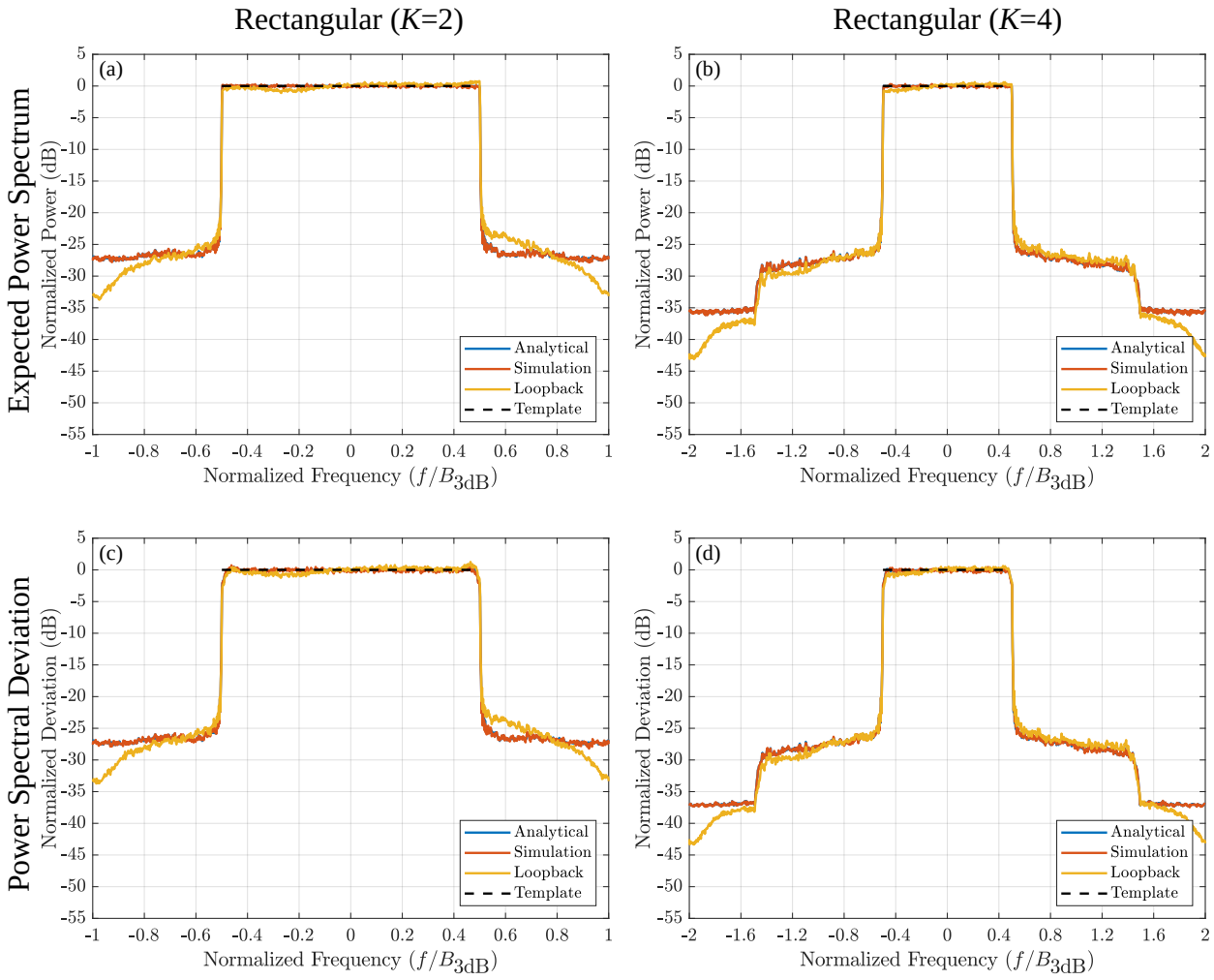

Figure 4.29: Optimized, DU2 based, pulsed StoWGe analytical, simulated, and loopback spectral results for rectangular templates: expected power spectrum for the R2 template (a), expected power spectrum for the R4 template (b), expected power spectral deviation for the R2 template (c), expected power spectral deviation for the R4 template (d)

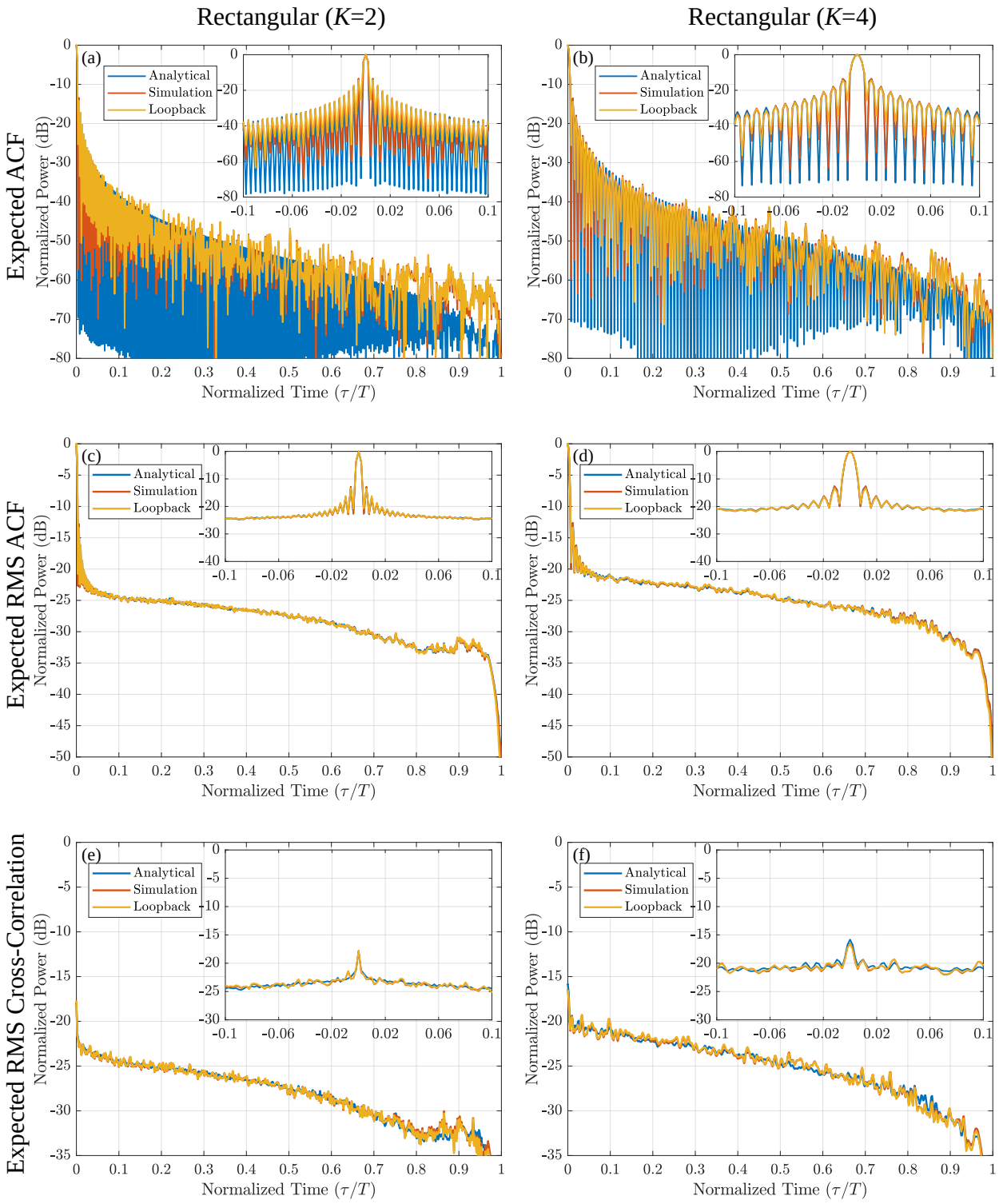

Figure 4.30: Optimized, DU2 based, pulsed StoWGe analytical, simulated, and loopback temporal results for rectangular templates: expected autocorrelation for the R2 template (a), expected autocorrelation for the R4 template (b), expected RMS autocorrelation for the R2 template (c), expected RMS autocorrelation for the R4 template (d), expected RMS cross-autocorrelation for the R2 template (e), expected RMS cross-autocorrelation for the R4 template (f)

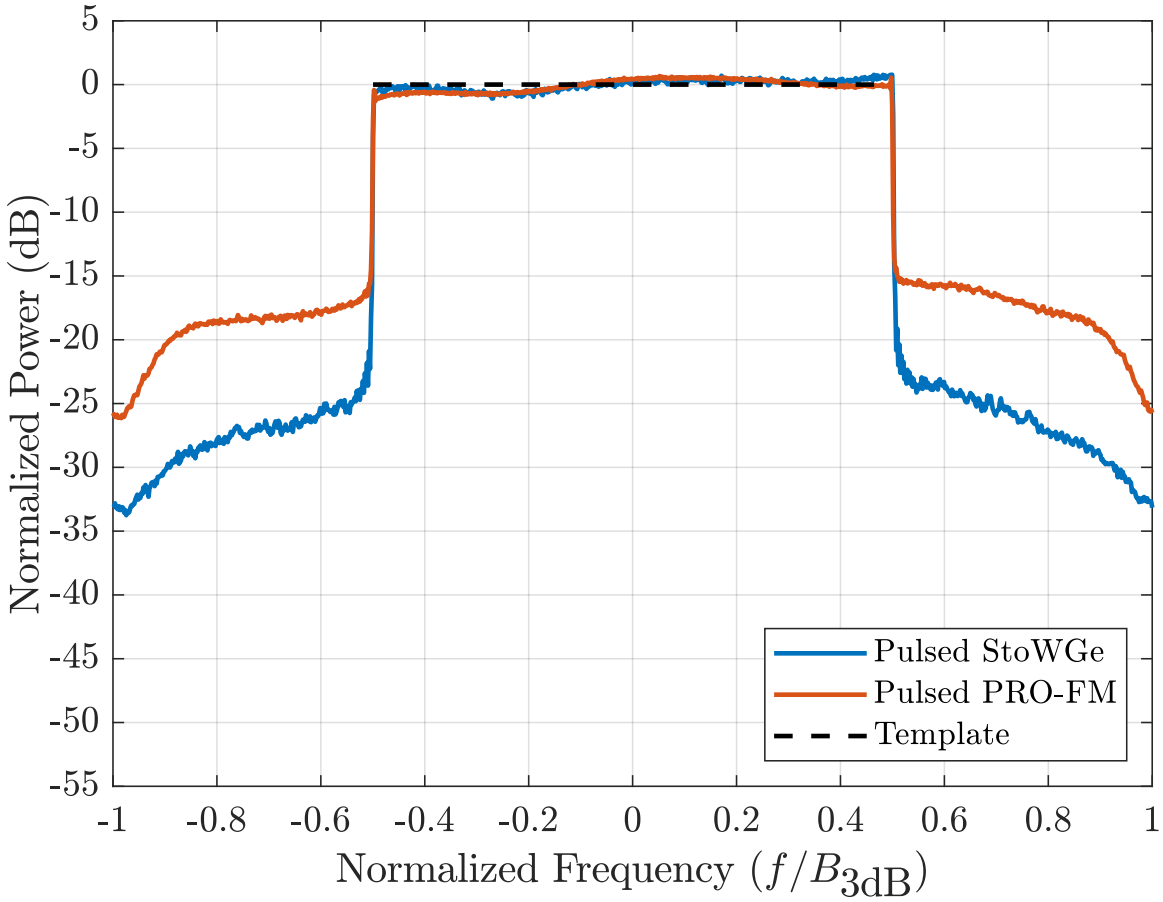

Figure 4.31: Comparison between the estimated expected spectrum of the loopback measured pulsed StoWGe DU2 waveforms and the loopback measured PRO-FM waveforms for the R2 template. Both sets exhibit similar passband spectrum distortion due to hardware effects.

4.6 Pulsed StoWGe Summary

The overall goal in defining the pulsed StoWGe WGF was to enable the computationally cheap instantiation of spectrally contained, constant modulus, pulsed radar waveforms. The efficient instantiation is achieved through the model by definition since the signal phase is created through a matrix vector multiply and by pulling values from an RNG. However, guaranteeing the resultant signals are useful in a radar context requires an involved optimization of the underlying parameters of the signal model.

Nevertheless, the gradient based minimization of the expected frequency template error (EFTE)

cost function defined in Section 4.2 was show in Section 4.4 to be an effective means of shaping the expected power spectrum of the waveforms instantiated by a particular stochastic waveform generating function (WGF). In doing so, the resultant waveforms were shown to be constant modulus and to adhere generally to desired spectral shapes while still being noise-like in that their autocorrelation sidelobes decrease with coherent integration and having noise-like cross-correlation characteristics. In other words, they were shown to be desirable for radar purposes.

Further, to gain more insight into the behavior of the pulsed StoWGe model, various spectral templates were considered along with varying numbers of basis functions and random variable distributions. While proving out the flexibility of the model, this testing and analysis likewise strongly indicated that random number distributions with *fewer* degrees of freedom results in better WGFs, but most especially when it comes to matching the expected power spectrum to more restrictive shapes like a rectangular spectrum. This result will require further research.

In Section 4.4.4, these waveforms were compared directly to previous random FM waveforms where it was shown that on an individual basis previous pulsed, random FM waveforms can achieve superior autocorrelation performance, but each individual waveform has to be optimized which creates a trade space between computational cost and the autocorrelation levels of the individual waveforms. In a real-time scenario, optimizing each individual waveform can easily be computationally prohibitive, but for pulsed StoWGe a WGF with a desired spectrum and BT can be optimized ahead of time and then be used to create an effectively infinite number or radar waveforms at minimal real-time computational cost.

Finally, in Section 4.5 the pulsed StoWGe waveforms were examined in both simulation and loopback tests where the analytically derived performance of the pulsed StoWGe waveforms was verified demonstrating the soundness of the pulsed StoWGe model as well as the real world physicality of the resultant waveforms.

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 results of the optimization and compares the CW-StoWGe waveforms to previously designed and implemented random FM waveforms. Section 5.5 demonstrates the suitability of the CW-StoWGe waveforms for physical implementation by experimentally verifying them in a loopback scenario. Finally, Section 5.6 provides a summary and some conclusions regarding the optimization and implementation of the CW-StoWGe signal model.

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 pulsed StoWGe to an infinite length in time, it is necessary to fundamentally alter its structure such that there are only a finite number of parameters to optimize. To do so, consider

the continuous phase modulation (CPM) signal model defined as [63, 64, 109]

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

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

$$q(t) = \int_0^t g(\tau) d\tau \tag{5.2}$$

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 by being energy efficient since it is constant amplitude and for 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

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

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.

In Chapter 4, the StoWGe signal model was introduced in discrete terms to facilitate its design and optimization on a computer. Likewise, the CW-StoWGe model can be discretized as

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

where $q(t)$ becomes

$$q[m] = \sum_{p=0}^{m} g[p]. \tag{5.5}$$

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 [63] while $L > 1$ corresponds to *partial-response* CPM [64]. 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 where the rows correspond to individual time domain samples of the CW-StoWGe model while the columns depict how a given random variable contributes the phase of a given sample. All q_m for $m \geq LT_s$ are the same value. This is indicated by the stair-stepping solid black line. A new random variable contributes to the phase every T_s samples and then contributes uniquely for LT_s samples, after which its contribution is just a constant phase offset.

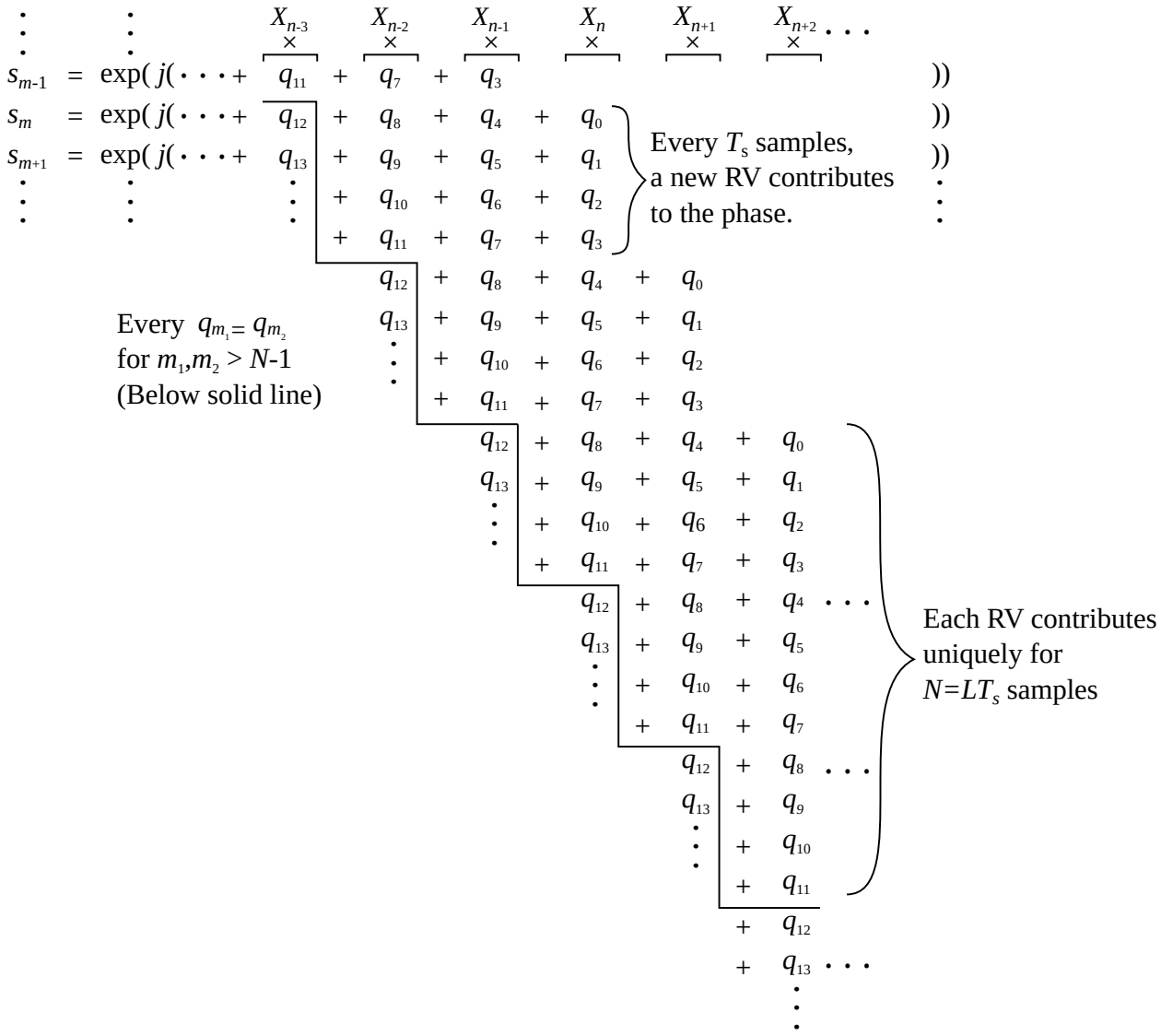

Figure 5.1: The CW-StoWGe phase structure at time m for $T_s = 4$ and $L = 3$.

As with pulsed StoWGe, to evaluate the characteristics of a particular CW-StoWGe implementation, it is necessary to evaluate the second and fourth order moments of the model. However, an apparent issue with CW-StoWGe is that it is not actually stationary. Instead, it is cyclo-stationary (defined in Section 3.2.1) over the interval T_s such that

$$\mathrm{E}\Big[s[m]s[m+\ell]\Big] \neq \mathrm{E}\Big[s[m+k]s[m+k+\ell]\Big] \tag{5.6}$$

where k is an integer, but it *is* true that

$$\mathrm{E}\Big[s[m]s[m+\ell]\Big] = \mathrm{E}\Big[s[m+kT_s]s[m+kT_s+\ell]\Big] \tag{5.7}$$

Consequently, to acheive a correlation vector which is a function of only one variable it is necessary to average the correlation over this relatively small interval. In doing so, the evaluation of these moments is carried out in detail in Appendix A.2.1, but the final results are reproduced here for convenience. The second moment (correlation) at delay ℓ is

$$C[\ell] = \sum_{v=0}^{T_s-1} \left[\prod_{n=\left\lceil \frac{v-\ell-LT_s+2}{T_s} \right\rceil}^{0} \psi_{X_n} \left(\sum_{k=0}^{v-nT_s} g[k] - \sum_{k=0}^{v-\ell-nT_s} g[k] \right) \right] \tag{5.8}$$

where ψ_{X_n} is characteristic function of the random variable X_n. The indexes of the summation of $g[k]$ will often go beyond it non-zero values ($k \geq LT_s$), but including these does not impact the value of $C[\ell]$. (5.8) could be modified to omit these terms, but doing so would significantly complicate the expression and so these terms are left. Additionally, the upper limits of the sums could be less than the lower limit. In these cases, the summations are by definition equal to zero.

Treating the CW-StoWGe model as a stationary process allows the second moments to be organized into a convenient structure. First a $(2W-1)\times 1$ correlation vector can be constructed as

$$\mathbf{c} = [\, C[-W+1] \quad C[-W+2] \quad \cdots \quad C[W-2] \quad C[W-1] \,]^T. \tag{5.9}$$

This vector can then be used to create a toeplitz correlation matrix such that

$$\mathbf{C} = \begin{bmatrix} C[0] & C[-1] & \cdots & C[-W+1] \\ C[1] & C[0] & \cdots & C[-W+2] \\ \vdots & \vdots & \ddots & \vdots \\ C[W-1] & C[W-2] & \cdots & C[0] \end{bmatrix}. \tag{5.10}$$

which will be useful for evaluating the expected power spectrum.

Likewise, a derivation of the fourth order moment is provided in Appendix A.2.1, but is reproduced here for convenience as

$$K[\ell_1,\ell_2,\ell_3] = \sum_{v=0}^{T_s-1}\left[\prod_{n=n_{\min}}^{n_{\max}} \psi_{X_n}\left(\sum_{k=0}^{v-nT_s} g[k] - \sum_{k=0}^{v-\ell_1-nT_s} g[k] - \sum_{k=0}^{v-\ell_2-nT_s} g[k] + \sum_{k=0}^{v-\ell_3-nT_s} g[k]\right)\right] \quad (5.11)$$

where

$$n_{\min} = \left\lceil \frac{v - \max\{v,\ell_1,\ell_2,\ell_3\} - LT_s + 2}{T_s} \right\rceil \quad (5.12)$$

and

$$n_{\max} = \left\lfloor \frac{\min\{v,\ell_1,\ell_2,\ell_3\}}{T_s} \right\rfloor. \quad (5.13)$$

Once again, the upper limits of the summations may exceed the non-zero portions of $g[k]$ and the summations may be 0. Eliminating these terms from (5.11) is possible, but doing so would significantly complicate the expression. The impact of cyclo-stationarity is also seen in the fourth moment. Consequently $K[\ell_1,\ell_2,\ell_3]$ represents the average fourth moments at relative delays ℓ_1,ℓ_2,ℓ_3. This is a consequence of the fact that

$$\mathrm{E}\Big[s[m]s[m-\ell_1]s[m-\ell_2]s[m-\ell_3]\Big] \neq \mathrm{E}\Big[s[m+k]s[m-\ell_1+k]s[m-\ell_2+k]s[m-\ell_3+k]\Big] \quad (5.14)$$

where k is an integer, but it *is* true that

$$\mathrm{E}\Big[s[m]s[m-\ell_1]s[m-\ell_2]s[m-\ell_3]\Big] = \mathrm{E}\Big[s[m+kT_s]s[m-\ell_1+kT_s]s[m-\ell_2+kT_s]s[m-\ell_3+kT_s]\Big]. \quad (5.15)$$

The fourth moment is used in the calculation of the RMS autocorrelation and the power spectral deviation and the RMS autocorrelation function.

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

The previous Chapter demonstrated how shaping the expected spectrum of pulsed StoWGe waveforms is an effective means of creating WGFs which generate useful radar waveforms. The same

EFTE metric can be used for the CW-StoWGe model as well. On the surface, the EFTE for the CW-StoWGe model is virtually identical to that of the pulsed StoWGe model with the only visual difference being a subscript such that

$$J_{\mathrm{CW}} = \left|\left| \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] - \mathbf{u} \right|\right|_2^2. \tag{5.16}$$

where the subscript 'CW' indicates the CW-StoWGe model and $\hat{\mathbf{s}}_{\mathrm{f}}$ represents the spectral estimate using the Bartlett spectral estimation scheme such that the CW-StoWGe EFTE measures the mean squared error between some desired spectral shape and the expected power spectral density for the WGF *given the Bartlett spectral estimation method is used.* In other words, the CW-StoWGe EFTE does not optimize the "true" WGF PSD directly. Instead it optimizes it with respect to how it is measured.

Further, this has an impact on how $\mathbf{s}_{\mathrm{f}}$ can be evaluated. For pulsed StoWGe, the expected autocorrelation function is realized as a sum over the diagonals of the correlation matrix. For CW-StoWGe, the correlation matrix is Hermitian and the expected autocorrelation function is evaluated directly via (5.8). However, the Bartlett estimate of the PSD is based on a windowed version of this correlation matrix such that

$$\mathbf{s}_{\mathrm{f}} = \mathbf{A}\left(\mathbf{w}_{\mathrm{B}} \odot \mathbf{c}\right) \tag{5.17}$$

where $\mathbf{w}_{\mathrm{B}}$ is the Bartlett window in vector form.

In general, the Bartlett method and the use of the standard unbiased autocorrelation estimate are used as a means to reduce the impact of the variance of the higher lag estimates of the correlation vector. However, given the CW-StoWGe model, these correlation terms can be evaluated analytically such that removing the window term from (5.17) should yield an exceptionally close estimate to the "true" PSD of the CW-StoWGe WGF assuming enough samples of the correlation vector are included such that it rolls-off to nearly zero value. At the same time, how useful is knowing this from a practical standpoint? The spectrum can never be experimentally measured in this way. Consequently, it is examined and optimized here as it would be observed in a physical

experiment.

It is expected however that as the number of DFT points in (5.17) approaches infinity, then (5.17) would approach the true PSD of the WGF. This tendency can be seen in Fig. 3.9 where increasing the DFT points results in better and better spectral roll-off as the expectation of the estimate approaches the true PSD of the WGF.

5.3 Optimization of the CW-StoWGe Expected Power Spectrum

Since the CW-StoWGe EFTE cost function is similarly non-linear and unconstrained as with the pulse StoWGe EFTE cost function, it can likewise be locally optimized via gradient descent methods. However, given the very different waveform model the particulars of the gradient descent method will differ as well. Here, instead of optimizing the basis function matrix, $\mathbf{B}$, or the mean phase value vector, $\boldsymbol{\mu}$, the shaping filter, $\mathbf{g}$, is optimized in order to make the expected PSD of a given WGF resemble some desirable spectral shape. As with any gradient descent method, the most important calculation is that of the gradient itself.

5.3.1 The CW-StoWGe EFTE Gradient

Compared to the pulsed StoWGe EFTE gradient, the difficulty is not so much in the evaluation of any given derivitive, it is instead a problem of indexing. A detailed derivation of the CW-StoWGe EFTE gradient is provided in Sections A.2.2 and A.2.3; however, a brief overview is provided here for convenience.

In Section A.2.2, the derivative of the CW-StoWGe EFTE cost function with respect to the yth element of $\mathbf{g}$ is shown to be

$$\frac{\partial J_{\mathrm{CW}}}{\partial g_y} = 2\left(\frac{\partial \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]}{\partial g_y}\right)^{T}\left(\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]-\mathbf{u}\right). \tag{5.18}$$

The derivative of the wth element of the CW-StoWGe expected spectrum is then

$$\frac{\partial \mathrm{E}\left[\hat{s}_{\mathrm{f},w}\right]}{\partial g_y} = \mathbf{a}_w \frac{\partial \mathbf{C}}{\partial g_y} \mathbf{a}_w^H. \tag{5.19}$$

where $\mathbf{a}_w$ is the wth row of the DFT matrix $\mathbf{A}$ and the derivative is applied in an element wise manner. (5.19) shows how the gradient is a function of the derivative with respect to the CW-StoWGe correlation matrix. Since this correlation matrix is Hermitian, there are actually only W unique derivatives to evaluate despite the $W \times W$ sized correlation matrix. Further steps are included in Section A.2.2, but the derivative of the ℓ element of the CW-StoWGe correlation vector is

$$\frac{\partial c[\ell]}{\partial g_y} = \frac{1}{T_\mathrm{s}} \sum_{v=0}^{T_\mathrm{s}-1} \sum_{z=n_{\min}}^{0} \left(\frac{\partial \psi_{X_z}\left(\sum_{k=v-|\ell|-zT_\mathrm{s}+1}^{v-zT_\mathrm{s}} g[k]\right)}{\partial g_y} \prod_{\substack{n=n_{\min} \\ n\neq z}}^{0} \psi_{X_n}\left(\sum_{k=v-|\ell|-nT_\mathrm{s}+1}^{v-nT_\mathrm{s}} g[k]\right)\right). \tag{5.20}$$

It is important to note that the derivative in (5.20) will often be zero since in many cases the argument of the characteristic function may not depend on g_y. However, rewriting (5.20) such that these terms are excluded would make it significantly more complicated and would not add any clarity to the problem. The derivatives with respect to particular characteristic function account for this ambiguity in Section A.2.3.

Regardless, the derivative in (5.20) can be used to evaluate the derivative in (5.19) and subsequently the derivative in (5.18). Putting the results for each term in $\mathbf{g}$ together, the gradient of J_{CW} with respect the $\mathbf{g}$ can be written such that

$$\nabla_{\mathbf{g}} J_{\mathrm{CW}} = \begin{bmatrix} \frac{\partial J_{\mathrm{CW}}}{\partial g_1} \\ \frac{\partial J_{\mathrm{CW}}}{\partial g_2} \\ \vdots \\ \frac{\partial J_{\mathrm{CW}}}{\partial g_{LT_\mathrm{s}-1}} \\ \frac{\partial J_{\mathrm{CW}}}{\partial g_{LT_\mathrm{s}}} \end{bmatrix} \tag{5.21}$$

5.3.2 CW-StoWGe Gradient Descent Implementation

As with pulsed StoWGe, the heavy ball gradient descent method described in Sections 2.5 and 4.3.2 was found to be an effective means of optimization. Consequently, Tables 4.1 and 5.1 mirror each other in significant ways other than the explicit differences in the models themselves.

In step 1 the various parameters are initialized. The initialization of the CW-StoWGe model parameters, $T_s, L, p_X(x), \mathbf{u}$, and $\mathbf{g}_0$, is described in detail in 5.3.3, while the initialization of the heavy ball gradient descent parameters $\beta, \rho_{\text{up}}, \rho_{\text{down}}$, and γ, is described in Section 5.3.4.

Beginning in step 3, one of the more prominent differences between the pulsed and CW models can be seen in that the gradient in this case consists of a single vector ($\nabla_{\mathbf{g}} J_{\text{CW}}(\mathbf{g}_i)$) rather than the combination of a vector and a matrix as in Table 4.1. Consequently, Table 5.1 more obviously follows a gradient descent framework with no need for special matrix operations such as the Frobenious product.

In step 4, the new search direction is defined as a combination of the negative of the current gradient and a scaled version of the previous search direction where $0 < \beta < 1$. In step 5, the inner product is taken of the new search direction and the current gradient. If the resulting quantity is positive the new search direction is actually a direction of ascent so it must be reset to the negative of the current gradient in step 6.

Steps 8 through 10 represent a simple backtracking algorithm in order to find an appropriate step size, γ. In step 8, the current shaping filter, $\mathbf{g}_i$ is incremented by $\gamma \mathbf{q}_{i+1}$. The cost function is then evaluated at this point. If the new cost function value represents a decrease by at least $c\gamma \mathbf{q}_{i+1}^T \nabla_{\mathbf{g}} J_{\text{CW}}(\mathbf{g}_i)$, then the candidate step size has met the Armijo condition which is the first of the Strong Wolfe conditions. As in the pulsed case, the second Wolfe condition, the curvature condition is ignored since it requires an additional evaluation of the gradient with the candidate step size. In this way, each iteration is somewhat less efficient than what could be achieved with both Wolfe conditions, but each iteration is also more efficient making more iterations possible in a given amount of time. If the Armijo condition is not met, the step size is decreased by the factor ρ_{down}. This process is repeated until an acceptable step size is found.

After completing the line search, in step 11 the shaping filter is updated according to the new step size and search direction. Then, the step size is increased by a factor of ρ_{up} which keeps the step size from ever getting too small due to the line search. As in the pulsed case, this is a simple means of compensating for not implementing the curvature condition.

In step 12, the iteration variable is incremented. Finally, in step 13 C_{Stop} is a place holder for the stopping conditions which are explained in Section 5.3.4. Step 16 indicates that the optimization is repeated until the stopping conditions are met.

EFTE Gradient Descent Optimization of the CW-StoWGe WGF

1:	**Initialize:** $T_{\text{s}},\ L,\ p_X(x),\ \mathbf{u},\ \mathbf{g}_0,\ \beta,\ \rho_{\text{up}},\ \rho_{\text{down}},\ c,\ \gamma$
	set $i=0, \mathbf{q}_0 = \mathbf{0}_{M\times 1}$
2:	**Repeat**
3:	Evaluate: $J_{\text{CW}}(\mathbf{g}_i),\ \nabla_{\mathbf{g}} J_{\text{CW}}(\mathbf{g}_i)$
4:	$\mathbf{q}_{i+1} = -\nabla_{\mathbf{g}} J_{\text{CW}}(\mathbf{g}_i) + \beta \mathbf{q}_i$
5:	**If** $\mathbf{q}_{i+1}^T (\nabla_{\mathbf{g}} J_{\text{CW}}(\mathbf{g}_i)) \geq 0$
6:	$\mathbf{q}_{i+1} = -\nabla_{\mathbf{g}} J_{\text{CW}}(\mathbf{g}_i)$
7:	**End** (If)
8:	**While** $J_{\text{CW}}(\mathbf{g}_i + \gamma \mathbf{q}_{i+1}) > J_{\text{CW}}(\mathbf{g}_i) + c\gamma \mathbf{q}_{i+1}^T (\nabla_{\mathbf{g}} J_{\text{CW}}(\mathbf{g}_i)$
9:	$\gamma = \rho_{\text{down}} \gamma$
10:	**End** (While)
11:	$\mathbf{g}_{i+1} = \mathbf{g}_i + \gamma \mathbf{q}_{i+1}, \quad \gamma = \rho_{\text{up}} \gamma$
12:	$i = i+1$
13:	**If** C_{Stop}
14:	**Stop**
15:	**End** (If)
16:	**End** (Repeat)

Table 5.1: Pseudo-code for the gradient descent optimization of CW-StoWGe WGFs through the minimization of the EFTE

5.3.3 CW-StoWGe Model 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 5.2. In this case, the parameters T_s, L, $p_X(x)$, and $\mathbf{u}$ must be initialized while the shaping filter, $\mathbf{g}$ consists of the parameters to be optimized. The next several paragraphs discuss each parameter, their selected initialization(s) and the rational behind those selections.

Table 5.2: WGF parameters which must be selected or initialized prior to optimization

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

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. Intuitively, the product LT_s should be as large as possible since more degrees of freedom should result in more design freedom and consequently better spectral matches. However, it remains to be seen if it is better to have more random variables tightly spaced in time but with longer responses (small T_s, large L), or is it better to have fewer, more spaced 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. For T_s these include $\{2, 4, 6, 8\}$. For L these include $\{1, 2, 3, 4, 5, 6, 7, 8\}$.

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. These include the DU2, CU, and G distributions discussed in detail in Section 4.3.3.

Desired Spectrum 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. These include the G2, G4, S4G2, S4G4, R2, and R4 spectral templates discussed in detail in Section 4.3.3.

Initial Frequency Shaping Filter $\mathbf{g}_0$: The two initial shaping filters were chosen with two goals in mind. The initializations should demonstrate the non-convexity of the cost function, and they should generally result in "good" matches to the PSD.

To demonstrate the non-convexity of the cost function, the initializations need to be sufficiently unique such that when all other parameters are held constant, the different initializations realize unique optimized shaping filters. The second goal of a good match to the template was achieved through trial and error. Numerous initializations were examined, and the two that are used here are not claimed to be the best initialization in any given case, but in general they have been found to tend towards good matches to the spectral templates as will be shown in Section 5.4.

The first initialization is based on the rectangular shaping used for CPM where $L = 1$ [109]. In continuous terms, the rectangular shaping filter is defined

$$\text{RECT}: \quad g_0(t) = \begin{cases} \frac{1}{2T_\text{s}} & 0 < t < T_\text{s} \\ 0 & \text{otherwise} \end{cases} \tag{5.22}$$

To optimize, (5.22) can be discretized such that

$$\text{RECT}: \quad \mathbf{g}_0 = \left[\mathbf{0}_{1\times T_\text{s}\lfloor (L-1)/2 \rfloor} \quad \frac{1}{2T_\text{s}} \mathbf{1}_{1\times T_\text{s}} \quad \mathbf{0}_{1\times T_\text{s}\lceil (L-1)/2 \rceil} \right]^T, \tag{5.23}$$

where the floor operator, $\lfloor \cdot \rfloor$, returns the largest integer smaller than the argument, and the ceiling operator, $\lceil \cdot \rceil$, returns the smallest integer larger than the argument. In (5.23), the initial shaping filter is defined such that $L = 1$ since the non-zero portions of $\mathbf{g}_0$ only extend over one T_s length interval. However, the total number of optimizable parameters is defined such that $N = LT_\text{s}$. In this way, the initial shaping filter has no partial response($L = 1$), but the optimization is allowed to create a partial response in order to minimize the cost function assuming it has been allotted the

degrees of freedom to do so.

The second initialization tested here is not based on any existing shaping filter to the author's knowledge rather it is based on a process of trial and error. In continuous terms it is defined

$$\text{EXP}: \quad g_0(t) = \begin{cases} \frac{1}{A}\exp(-|t-L/2|) & 0 < t < L \\ 0 & \text{otherwise} \end{cases} \tag{5.24}$$

where division by A normalizes the initialization to integrate to a total of $1/2$. (5.24) can be discretized for optimization such that

$$\text{EXP}: \quad \mathbf{g}_0 = \frac{1}{A}\exp(-|\mathbf{t}-L/2|) \tag{5.25}$$

where $\mathbf{t}$ is an LT_s length vector whose values extend uniformly over the range $[0, L]$.

These two initializations will be referred to as $\mathbf{g}_{\text{RECT}}$ and $\mathbf{g}_{\text{EXP}}$ respectively.

Summary: – All the initializations and their shorthand notation from above are summarized in Table 5.3.

5.3.4 CW-StoWGe Gradient Descent Parameter Initializations

In the course of performing the optimizations for this work, the pulsed StoWGe optimizations were carried out first. As a starting point for the CW-StoWGe optimizations, the same values for the gradient descent parameters, β, ρ_{up}, ρ_{down}, c and the stopping conditions were carried over from the pulsed StoWGe optimization, and these values were found to be sufficient for this optimization as well. The good suitability of these parameters and stopping conditions for both optimizations is likely a consequence of them being relatively conservative. (i.e. the sufficient decrease parameter, c, is fairly small in the scope of gradient methods.) The values for β, ρ_{up}, ρ_{down}, c are .95, 1.01, .9, and .01 respectively and a complete description of the stopping conditions can be found in 4.3.4.

It bears repeating that there are likely better sets of parameters and more efficient means of optimizing the CW-StoWGe shaping filters in terms of computational time. However, the goal of

Table 5.3: Optimization initializations for the pulsed StoWGe parameters. All combinations are considered resulting in 1152 total optimizations

Variable	Definition	Test Cases	
T_s	Intervariable Spacing	$T_s =$	2, 4, 6, 8
L	Partial Response	$N =$	1, 2, 3, 4, 5, 6, 7, 8
$p_X(x)$	Random Variable Distribution	DU2	– Discrete Uniform, 2 states
		CU	– Continuous Uniform
		G	– Gaussian with unit variance
$\mathbf{u}$	Desired Spectrum	G2	– Gaussian $(K = 2)$
		G4	– Guassian $(K = 4)$
		S4G2	– Super-Gaussian $(n = 4, K = 2)$
		S4G4	– Super-Gaussian $(n = 4, K = 4)$
		R2	– Rectangular $(K = 2)$
		R4	– Rectangular $(K = 4)$
$\mathbf{g}_0$	Basis Function Initialization	$\mathbf{g}_{\text{RECT}}$	– Rectangular vector
		$\mathbf{g}_{\text{EXP}}$	– Exponential vector

this work is not demonstrate the most efficient means of optimization for these problems, rather it is to show that they can be optimized such that the resulting WGFs produce useful random FM waveforms. Computational speed was only a concern inasmuch as the optimizations could be performed in a practical time frame such as hours instead of weeks.

5.4 CW-StoWGe Optimization Results

The best WGFs according to the EFTE optimization are those which achieved the smallest cost function value as a function of $\mathbf{g}_0$, T_s, L, and $p_X(x)$. Every result of each of the optimizations detailed in Table 5.4 is tabulated in Tables B.7-B.12 on a dB scale where more negative values (smaller) represent a better optimization result. However, as in the pulsed case, there is more to producing good radar waveforms than simply achieving a desired expected power spectrum. Still,

the goals of producing good radar waveforms and achieving a desired expected power spectrum are closely related meaning a WGF which results in a good expected spectral match is more likely to produce useful radar waveforms than otherwise. Consequently, this section as a whole examines the WGFs in Tables B.7-B.12 for their ability to produce useful random FM radar waveforms in terms of the expected spectrum and other metrics.

Table 5.4: Optimized WGFs which resulted in the lowest value of J_{CW} for each $p_X(x)$ and $\mathbf{u}$ as a function of T_{s}, L, and $\mathbf{g}_0$.

		Distribution $(p_X(x))$											
		DU2				CU				G			
		J	T_s	L	$\mathbf{g}_0$	J	T_s	L	$\mathbf{g}_0$	J	T_s	L	$\mathbf{g}_0$
Template ($\mathbf{u}$)	G2	-115.4	4	8	$\mathbf{g}_{\mathrm{EXP}}$	-96.4	8	5	$\mathbf{g}_{\mathrm{RECT}}$	-78.5	2	4	$\mathbf{g}_{\mathrm{RECT}}$
	G4	-132.1	4	4	$\mathbf{g}_{\mathrm{EXP}}$	-124.7	8	4	$\mathbf{g}_{\mathrm{RECT}}$	-100.4	8	8	$\mathbf{g}_{\mathrm{RECT}}$
	S4G2	-127.3	8	5	$\mathbf{g}_{\mathrm{RECT}}$	-95.0	8	7	$\mathbf{g}_{\mathrm{RECT}}$	-51.6	8	5	$\mathbf{g}_{\mathrm{RECT}}$
	S4G4	-105.3	8	8	$\mathbf{g}_{\mathrm{EXP}}$	-80.4	8	8	$\mathbf{g}_{\mathrm{RECT}}$	-47.9	6	5	$\mathbf{g}_{\mathrm{RECT}}$
	R2	-56.0	8	6	$\mathbf{g}_{\mathrm{RECT}}$	-48.4	8	2	$\mathbf{g}_{\mathrm{RECT}}$	-42.6	4	3	$\mathbf{g}_{\mathrm{RECT}}$
	R4	-52.8	8	8	$\mathbf{g}_{\mathrm{EXP}}$	-43.3	8	2	$\mathbf{g}_{\mathrm{RECT}}$	-39.6	4	5	$\mathbf{g}_{\mathrm{RECT}}$

To begin, Section 5.4.1 examines the WGFs which produced the best matches found in Tables B.7-B.12 as a function of the template and the random variable distribution. This analysis is performed by using several of the metrics in Table 3.3, albeit with the important CW processing caveats covered in Section 5.4.1. Section 5.4.2 takes a look at how the shaping filters were optimized to achieve expected power spectrum template matches. Finally, 5.4.3 compares the optimized CW-StoWGe waveforms to previous random FMCW waveforms.

5.4.1 CW-StoWGe Optimized WGF Analysis

As with the pulsed case, the most obvious place to start in evaluating the quality of a given optimized CW-StoWGe WGF is to look at how well its expected spectrum matches the desired template. Table 5.4 lists the cases with the best optimized cost function values as a function of T_s, L,

and $\mathbf{g}_0$. Given six different templates and 3 different random variable distributions there are a total of 18 CW-StoWGe waveform generating functions to examine in greater detail.

As with the pulse case, examining the performance of a given WGF for producing random FM waveforms requires examining them beyond their expected power spectrum. To do so, the several of the metrics 3.3 are used to determine how effective the optimized CW-StoWGe WGF are for producing useful random FMCW waveforms. The metrics deemed useful for evaluating the optimized CW-StoWGe WGFs are

- Expected power spectrum - Since this metric it directly optimized it is clearly relevant to the analysis.

- Power spectral deviation - In the pulsed case, the power spectral deviation described how each individual pulse can be expected to deviate from the expected power spectrum. In the CW-StoWGe case, the power spectral deviation describes how the estimated spectrum will vary with respect to the DFT of a single window. With pseudo-pulse processing, this effectively serves the same purpose as for the pulse case. It is desirable that the power spectral deviation is minimized such that the DFT of a given window is as close as possible to the desired power spectrum.

- Expected Autocorrelation - This metric demonstrates the best autocorrelation performance for a given WGF which occurs for CW-StoWGe when there are an infinite number of pseudo-pulses, or with one infinitely long pulse. Ideally, the expected autocorrelation sidelobes should be extremely low such that with increasing coherent integration the sidelobes are lowered further and further. Any difference between this metric and the expected RMS autocorrelation represents a potential for lowering the autocorrelation sidelobes with coherent integration.

- Expected RMS autocorrelation - Since any given observation scenario will invariably be time limited, it is important to understand the expected autocorrelation level given finite time

support. The expected RMS autocorrelation represents the average autocorrelation level for a given delay and filter length.

Notably absent from the aforementioned metrics are the cross-correlation metrics. Since in a CW setting only one sample function of the WGF will ever be transmitted at a time, there will never be cross-correlation terms to consider making them irrelevant to this analysis. However, there could be a scenario where multiple transmitters send out unique sample functions of the WGF in a multi-static application, but that is beyond the scope of this work.

5.4.1.1 Gaussian Template Results

If Figs. 5.2 and 5.3 the analytically calculated spectral and temporal plots of the results in Table 5.4 are plotted respectively. As evidenced by 5.2(a,b), each distribution was able to make a good match to the desired template for either of the G2 or the G4 templates. However, in the G4 case the expected spectra do deviate from the desired template beginning below about -30 dB in normalized power. Without further analysis, it is impossible to say whether this is a result of the optimization fundamentally being unable to match to the template at such low power levels or if this is the impact of spectral estimation bias from too few DFT point (W=1024 in this case).

In 5.2(c) for the G2 template, each distribution achieved a spectral deviation curve that is on the order of the spectral template itself which mirrors the results in the pulsed case. Likewise, the spectral deviation for the G4 template in Fig. 5.2(d) mirrors the pulsed case in that the DU2 distribution spectral deviation largely matches the spectral template power but contains small, periodic spikes, the CU distribution spectral deviation largely matches the spectral template power, and the G distribution spectral deviation results in a higher degree of deviation over almost the whole sample bandwidth but primarily in the roll-off regions. This higher degree of deviation indicates it will take a larger amount of coherent integration for the sample function spectral estimates to approach the expected power spectrum as compared to the other distributions.

In Fig. 5.3(a,b) for either the G2 or the G4 templates, the expected autocorrelations achieve lower autocorrelation sidelobes than what would be expected given a perfect template matched

as evidenced by the template trace. The template trace is simply the IFFT of the power spectral template. The expected autocorrelation traces indicate that as the window length W is increased or the number of coherently integrated pseudo-pulses is increased, then the autocorrelation level will approach that of their respective trace. In each case, the expected autocorrelation quickly falls below -140 dB. This result demonstrates that the waveforms produced by the WGFs are truly noise like in that their autocorrelation sidelobe decrease with an increase in coherent integration.

The RMS autocorrelation are perhaps more useful from a practical standpoint since the expected ACF plots assume infinite coherent integration. Whereas, Figs. 5.3(c,d) represent the expected autocorrelation levels when processing with a single pseudo-pulse of length T_{filt}. An interesting trend in either of 5.3(c,d) is the pronounced sinusoidal pattern in the RMS autocorrelation as a function of normalized time which is especially prevalent for the G distribution for the G2 template. This oscillatory structure is due to the fact that the CW-StoWGe WGFs are cyclostationary rather than simply stationary.

Additionally, as discussed early, for these traces $W = 1024$. For a 3dB oversampling factor of 2 for the G2 template and 4 for the G4 template, this results in an approximate $W/2 = 512$ and a $W/4 = 256$ BT for each case respectively. For noise and noise-like CW waveforms, the average autocorrelation level should be around $10\log_{10}(BT)$. For the $G2$ template results in 5.3(c), this level is about -27.1 dB and for 5.3(d) the level is about -24.1. In either case the traces level off or average to within a couple dB of these levels further indicating the noise like character of waveforms.

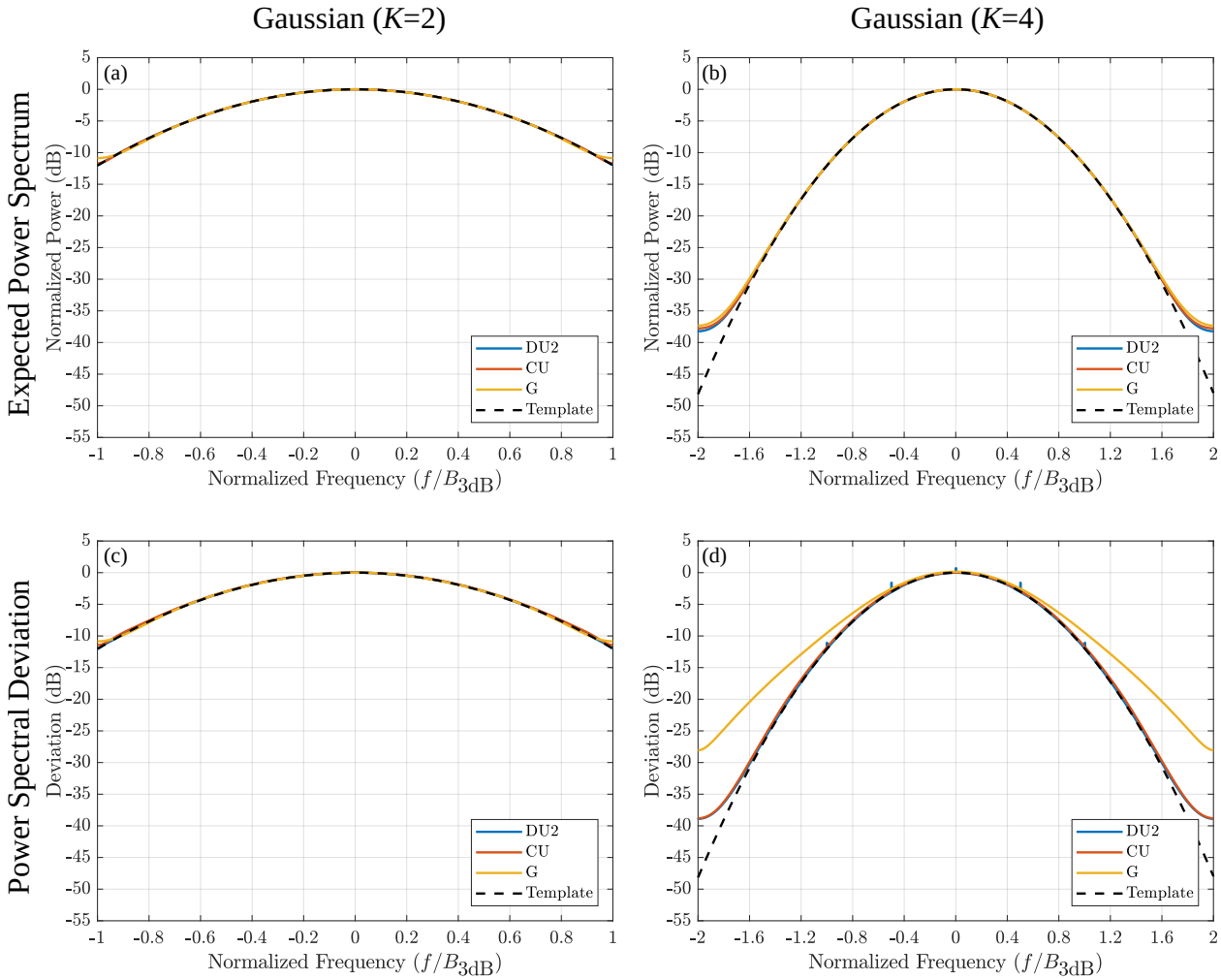

Figure 5.2: CW-StoWGe spectral optimization results for Gaussian templates: expected power spectrum for the G2 template (a), expected power spectrum for the G4 template (b), expected power spectral deviation for the G2 template (c), expected power spectral deviation for the G4 template (d)

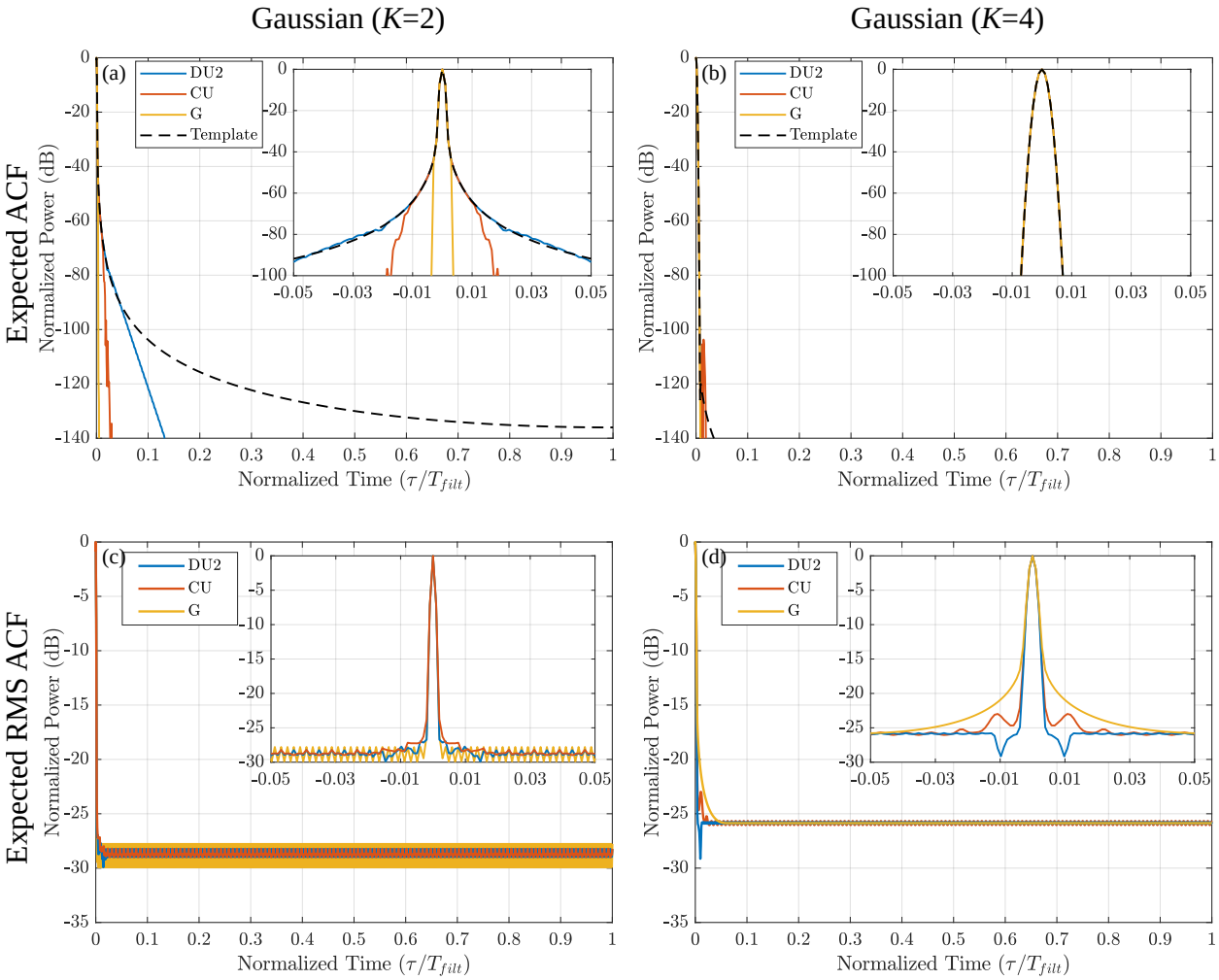

Figure 5.3: CW-StoWGe temporal optimization results for Gaussian templates: expected autocorrelation for the G2 template (a), expected autocorrelation for the G4 template (b), expected RMS autocorrelation for the G2 template (c), expected RMS autocorrelation for the G4 template (d)

5.4.1.2 Super-Gaussian Template Results

As with pulsed-StoWGe, there is once again much more disparity between the different random variable distributions for the Super-Gaussian templates as compared to the Gaussian templates. In Fig. 5.4(a,b), the G distribution performed the worst while the CU distribution performed markedly better, but the DU2 distribution clearly outperformed either. Interestingly in 5.4(b) (the S4G4 template), the CU and DU2 distribution hit a sort of spectral containment floor just above -40 dB in normalized power. However, this floor can be attributed directly to measurement bias.

In Fig. 5.4(c,d), a similar pattern is seen compared to the Gaussian spectral templates. The G

distribution experiences a higher degree of spectral deviation while the CU and DU2 distributions have a deviation level that is largely proportional to their respective expected power spectra. However, for these templates, the DU2 distributed random variables no longer result in small, periodic spikes in the spectral deviation plots.

In Fig. 5.5)(a,b), the inevitable sidelobes of the Super-Gaussian spectral shape are clearly present as the dashed black lines in the inset plots. The poor spectral match of the G distribution leads to a likewise poor match to the templates' respective autocorrelations. However, decent spectral matches achieved by the CU and DU2 distributions lead to descent matches to the spectral template autocorrelations. This is at least the case for the first several significant sidelobes. Beyond those sidelobes, the DU2 distribution exhibits some ringing, but otherwise quickly fades away as is desired. Thus, the sidelobes in any case can be arbitrarily lowered through coherent integration.

The RMS autocorrelation plots in Fig. 5.5(c,d) once again average to about the $10\log_{10}(BT)$ which are the same levels as before. Interestingly, as in the pulsed case, for the S4G2 template, the DU2 distributed random variables result in an RMS autocorrelation with a significant oscillatory component, while the G distribution has a similar but less pronounced behavior, and the CU distribution trace is nearly flat. For the S4G4 template in Fig. 5.5(d) each distribution resulted in a relatively flat RMS autocorrelation response. However, the DU2 distribution results in a small dip near the mainlobe.

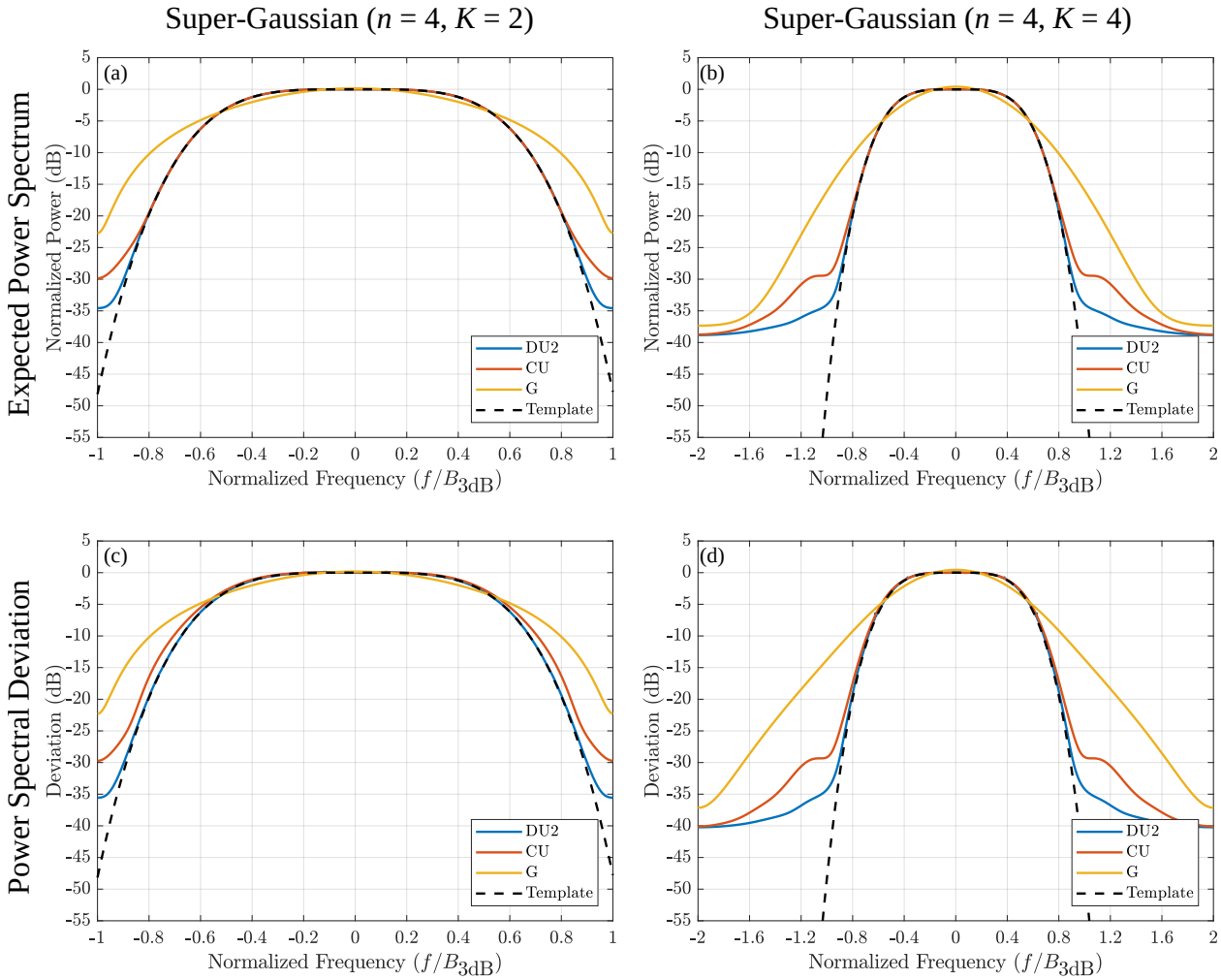

Figure 5.4: CW-StoWGe spectral optimization results for super-Gaussian templates: expected power spectrum for the S4G2 template (a), expected power spectrum for the S4G4 template (b), expected power spectral deviation for the S4G2 template (c), expected power spectral deviation for the S4G4 template (d)

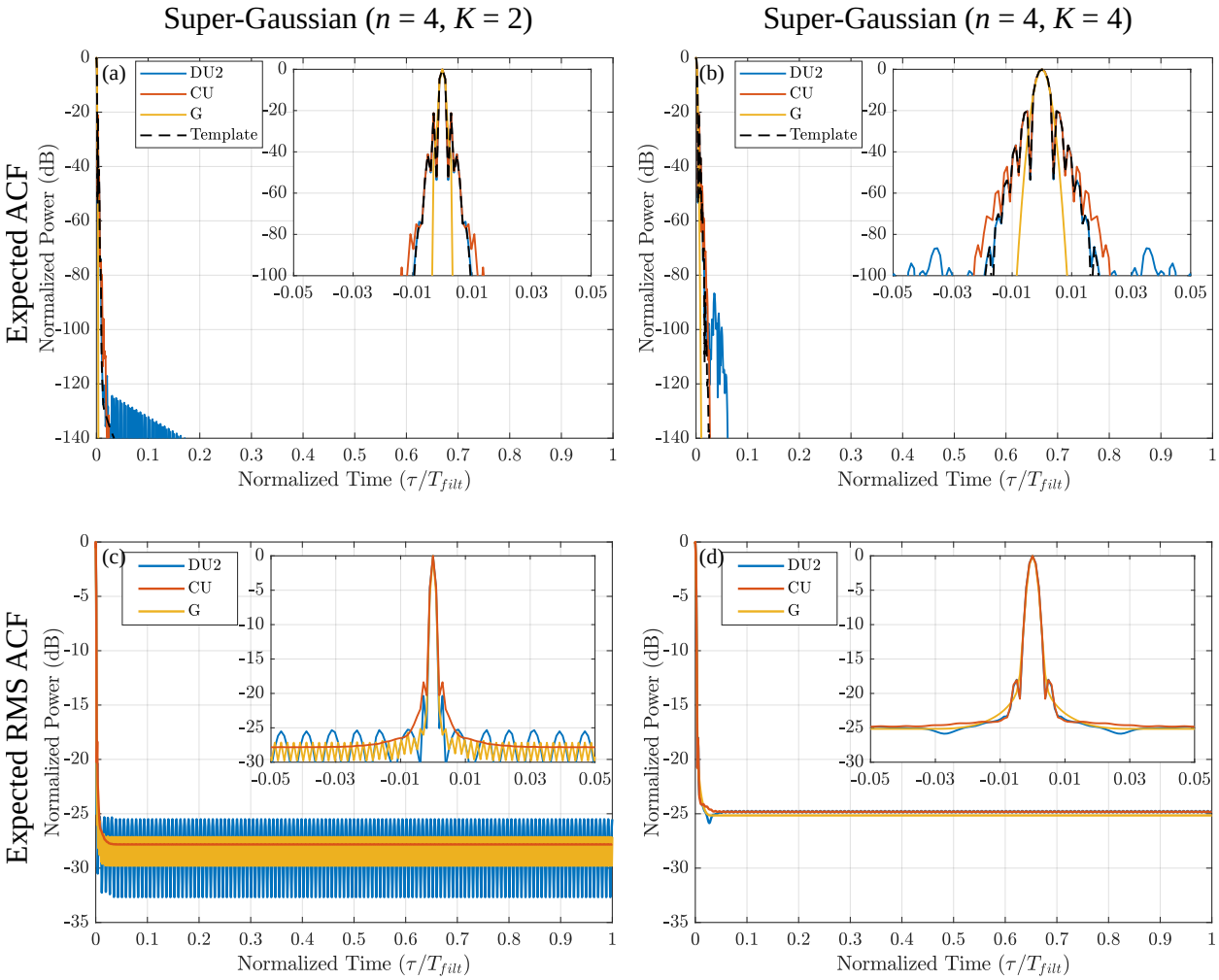

Figure 5.5: CW-StoWGe temporal optimization results for super-Gaussian templates: expected autocorrelation for the S4G2 template (a), expected autocorrelation for the S4G4 template (b), expected RMS autocorrelation for the S4G2 template (c), expected RMS autocorrelation for the S4G4 template (d)

5.4.1.3 Rectangular Template Results

For pulsed StoWGe the rectangular spectral templates results in the poorest power spectral matches where only the DU2 distributed random variables were really able to produce a decent match. This is likewise true in the CW case, but to a greater extreme where none of the distributions were able to achieve an excellent spectral roll-off. However, as evidenced by Fig. 5.6(a,b) the DU2 distribution still did the best job of matching to the pass band of the spectrum and rolling-off to some degree.

The spectral deviation traces in Fig. 5.6(c,d) are largely proportional to the expected spectral traces. However, once again, the G distributed random variable result in a slightly higher degree of deviation as compared to the other distributions, but this is of little consequence in this case since the G distribution resulted in such a poor spectral match anyway.

In the expected autocorrelation plots of Fig. 5.7(a,b), since only the DU2 distribution was able to meaningfully match to the passband of the rectangular templates, only the DU2 traces exhibit the expected sinc like autocorrelation sidelobes like those of an LFM. As the amount of coherent integration is increased, the autocorrelation sidelobes will decrease no further than those that are shown in Fig. 5.7(a,b).

For the expected RMS autocorrelation plots in 5.7(c,d), the sinc like sidelobes are high enough that they are evident in these traces. The effect is especially pronounced for the DU2 random variable distribution results. Otherwise, the RMS autocorrelation sidelobe levels once again decay to about $10\log_{10}(BT)$

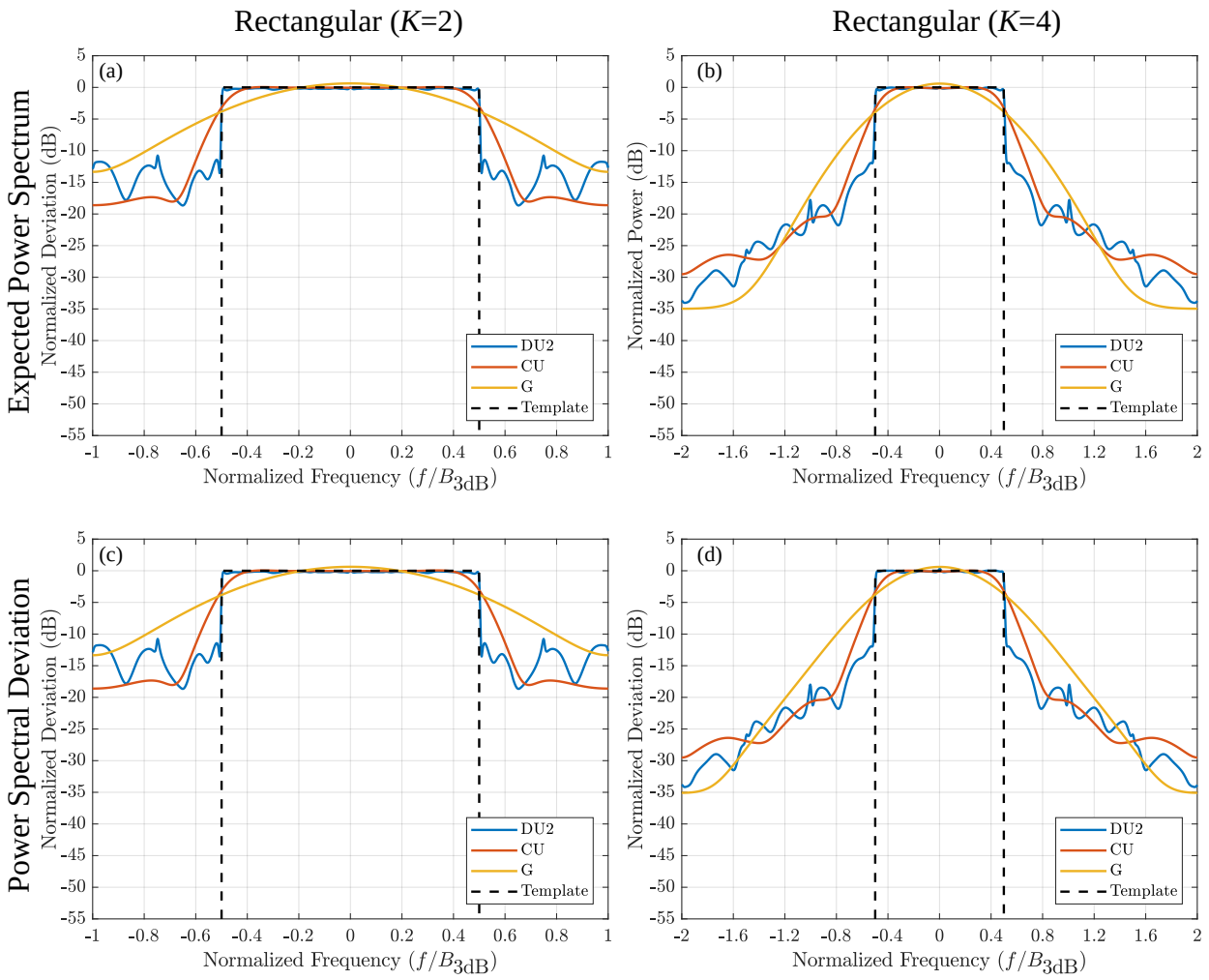

Figure 5.6: CW-StoWGe spectral optimization results for rectangular templates: expected power spectrum for the R2 template (a), expected power spectrum for the R4 template (b), expected power spectral deviation for the R2 template (c), expected power spectral deviation for the R4 template (d)

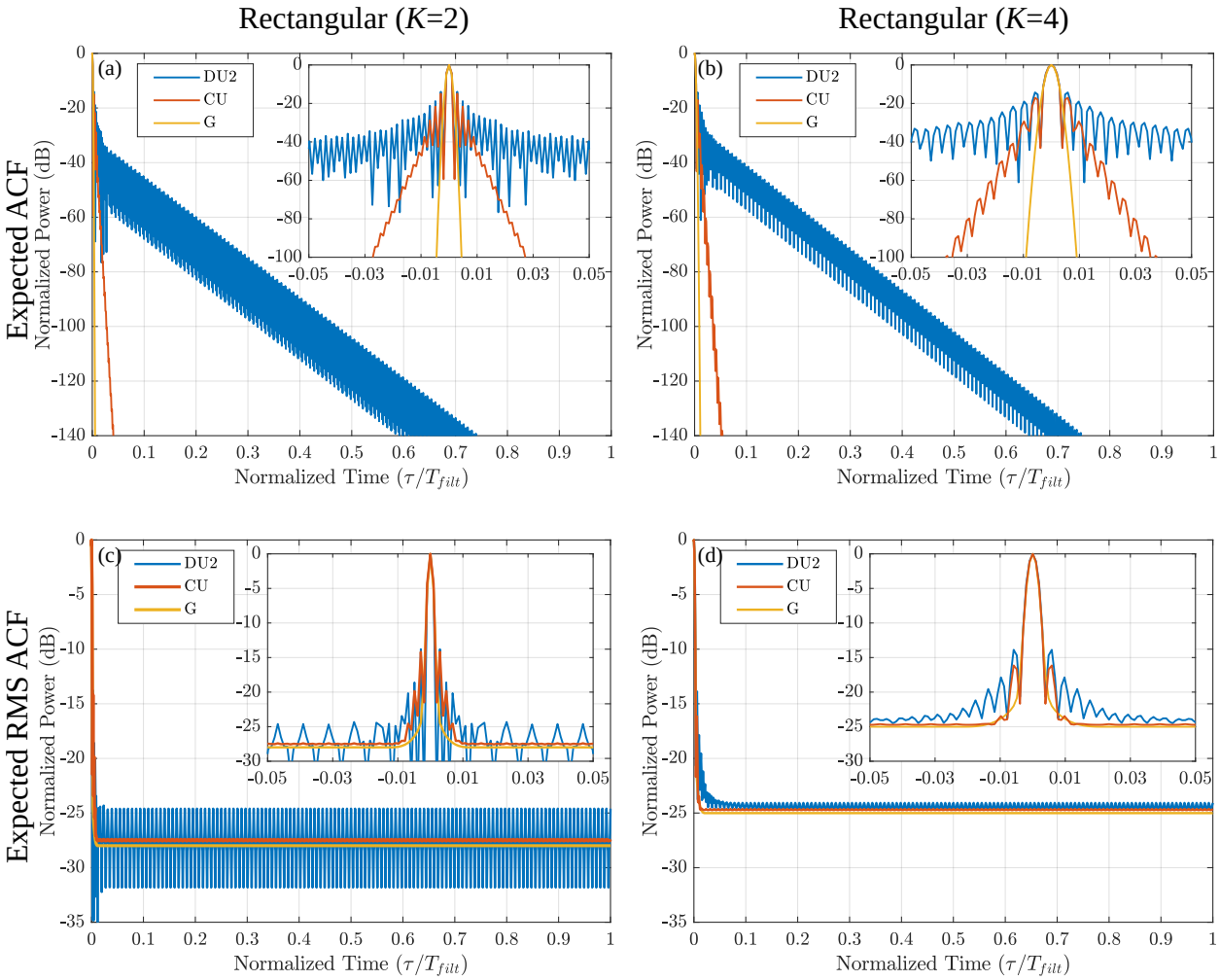

Figure 5.7: CW-StoWGe temporal optimization results for rectangular templates: expected autocorrelation for the R2 template (a), expected autocorrelation for the R4 template (b), expected RMS autocorrelation for the R2 template (c), expected RMS autocorrelation for the R4 template (d)

5.4.2 Optimized Frequency Shaping Filters

For pulsed StoWGe, determining how the optimization minimized the cost function required examining both the basis function matrix and the correlation matrix. With CW-StoWGe, the only optimized structure is the frequency shaping filter and the since the CW-StoWGe model is cylcostationary the correlation matrix is largely described by the autocorrelation functions of the previous section.

With this in mind, Figs. 5.8 and 5.9 show the optimized shaping filters for all of the 2 and 4

times oversampled desired spectra respectively for all the cases examined in the previous section. These cases are listed in Table 5.4.

In Fig. 5.8, the x-axis extends from 1 to 64 where 64 is the maximum possible length of the shaping filters optimized in this work. Such a length corresponds to $T_s = 8$ and $L = 8$. Shaping filters shorter than this were centered in the plot to make for an easier comparison to the other filters. The range of the y-axis was chosen such that the entirety of the shaping filters are visible. Although, in the case of the Fig. 5.8(b) the Gaussian (G) random variable based shaping filter was allowed to go off the plot since showing it in its entirety would have poorly scaled the other filters. Likewise, the range of the y-axis was adjusted in 5.9 to best show the shaping filters.

Perhaps the most interesting aspect of the shaping filters is their inconsistency. For example compare the DU2 traces in 5.8(b) and 5.9(b). In both cases the optimization was trying to match to a super-Gaussian expected spectrum, but with a different oversampling factor. Despite the similarity in spectral shape, the optimization came to two very different solutions. On the other hand, there does seem to be a pattern to the CU traces for the same templates. For the rectangular templates in 5.8(c) and 5.9(c) it is really only worth looking at the DU2 filters since these are the only ones that produced any kind of reasonable match to the desired spectra. In these cases, there seems to be little in common between the different shaping filters. Perhaps these results are a testament to the usefulness of the gradient descent optimization of a non-convex function, since there is nothing intuitive about the solutions shown in Figs. 5.8 and 5.9, but as evidenced by the previous section, they often achieve good matches to the desired expected spectrum. This is especially true of the DU2 distributed random variables which generally achieved the best matches compared to the other random variable distributions.

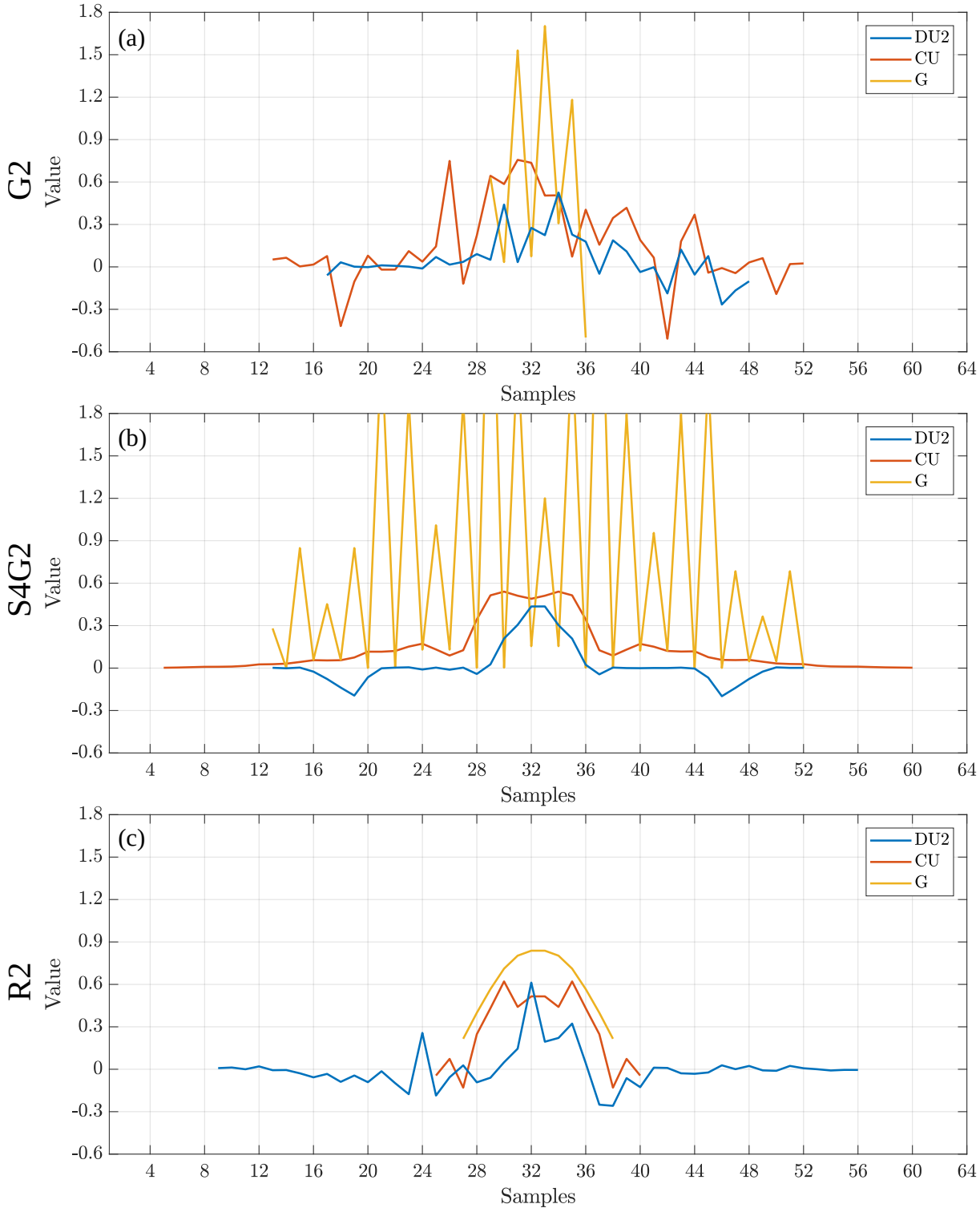

Figure 5.8: Optimized frequency shaping filters for each distribution for the 2 times oversampled templates: G2 template (a), S4G2 template (b), R2 template (c)

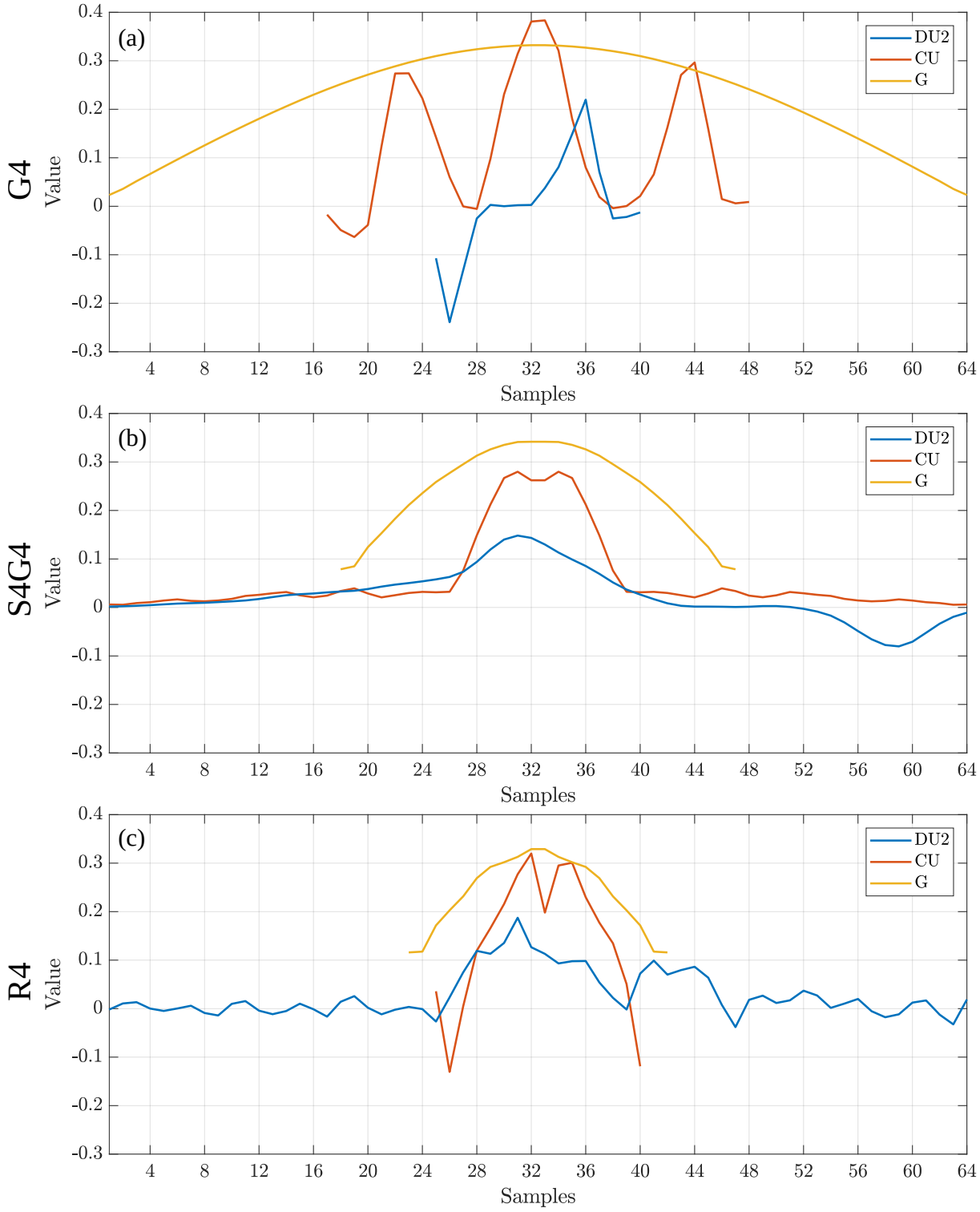

Figure 5.9: Optimized frequency shaping filters for each distribution for the 4 times oversampled templates: G4 template (a), S4G4 template (b), R4 template (c)

5.4.3 CW-StoWGe Comparisons to Previous Random FM Waveforms

As in the pulsed case, the DU2 distributed random variable WGFs generally produce the best radar waveforms. Since the PRO-FM waveforms have been demonstrated in a CW mode (PRO-FMCW), they once again make sense as a standard of comparison [25]. In a similar manner to 4.4.4, the results of the DU2 CW-StoWGe waveforms will be compared directly to the PRO-FM waveforms using the same metrics as in 5.4.1. The goal of this comparison is not to determine which of either the CW-StoWGe or the PRO-FM waveforms are better. Instead, the goal is to establish what the CW-StoWGe waveforms do and do not have in common with previous random FMCW waveforms.

5.4.3.1 CW-StoWGe and PRO-FM with zero Doppler

To do so, the CW-StoWGe waveforms traces in Figs. 5.10 to 5.15 are pulled directly from section 5.4.1, while the PRO-FM traces are monte carlo estimates based on instantiated PRO-FMCW waveforms. In contrast to the pulsed waveforms, getting a good monte carlo estimate of the PRO-FMCW is a matter of creating one, very long, high BT waveform rather than many smaller ones. Using the PRO-FMCW instantiation and optimization method defined in [25], for the 4 times oversampled templates, 3 waveforms with a BT of approximately 2560000 where instantiated. While for the 2 times oversampled templates, 3 waveforms with a BT of approximately 5120000 where instantiated. There resulting expected spectra, spectral deviations, autocorrelations, and RMS autocorrelations where then estimated via the FMCW estimation methods defined in 3.4. The DU2 results and the PRO-FMCW estimates are presented in Figs. 5.10.

Starting the with expected spectra in Figs. 5.10, 5.12, and 5.14, the CW-StoWGe results generally achieve better spectral roll-off and containment. Although in the case of the 1‘rectangular templates of Fig. 5.14(a,b), the PRO-FMCW waveforms achieve somewhat better spectral containment. The bigger contrast between these waveform generation methods is seen in the power spectral deviation results. For the CW-StoWGe waveforms, the deviation is almost directly proportional to the expected spectrum. For the PRO-FMCW, like their closely related pulsed variant, the spectral deviation is significantly less than the expected spectrum level. Although, the PRO-

FMCW waveforms are of a different structure than the pulsed PRO-FM waveforms, they are still optimized in a similar manner. The PRO-FMCW waveforms are instantiated by optimizing *individual* segments to achieve the desired power spectrum. These segments are then phase aligned and concatenated to create the overall CW waveform. The CW-StoWGe signals however are only tasked with matching the expected spectrum as the length of the signal approaches infinity. While the mechanics are different from the pulsed case, the result is similar. The power spectrum of a PRO-FMCW waveform will be much closer to its expected spectrum in a mean squared error sense over a given time interval than a commensurate CW-StoWGe waveform will.

The impact of this fact is evident in the RMS autocorrelation plots of Figs. 5.11, 5.13, and 5.15. Regardless of the template, for the given filter length the PRO-FM waveforms achieve a lower autocorrelation sidelobe level on a per pseudo-pulse level. This effect is mostly concentrated near the mainlobe and the advantage decreases as the delay increases. As the normalized time approaches 1, the sidelobe level approaches that of the CW-StoWGe waveforms which happens to be approximately $10\log_{10}(BT)$. For the 4 times oversampled template with the 1024 sample length filter this comes out to approximately -24.1 dB while for the 2 time oversampled templates and the same length filter this comes out to about -27.1 dB. In contrast to the pulsed waveforms where the autocorrelation decays due to the time limited nature of the pulses. In pseudo-pulse processing, the filtered data vector extends well beyond the filter length. In this case, the filter length is also the length of each PRO-FMCW segment. As the delay increases, the filtering segments overlaps more and more with the adjacent segment rather than with the matched segment. Since the adjacent segments were optimized and instantiated separately from the filtering segment, then there should be no structure between them resulting in a level that is described by the BT of the filtering segment. All of this is to say that for delays beyond the length of the pseudo-pulse filter, there is no difference in the RMS autocorrelation performance for the CW-StoWGe and the PRO-FMCW waveforms for the same BT and spectrum. Because of this, the lower autocorrelation level advantage achieved by the pulsed PRO-FM waveforms and other random FM waveform types is actually diminished somewhat in the CW case since the sidelobe levels level off to about $10\log_{10}(BT)$ anyway. One

solution would be to make longer and longer filters and segments, but this would be increasingly computationally expensive. Regardless, as in the pulsed case the major advantages afforded by the CW-StoWGe waveforms are good spectral containment and an extremely computationally cheap instantiation.

5.4.3.2 CW-StoWGe and PRO-FM Ambiguity Functions

In order to examine the Doppler characteristics of the CW-StoWGe waveforms, the analytical expected RMS ambiguity function was evaluated for each of the 4 times oversampled spectral templates and for only the DU2 distributed random variable WGFs used in the previous section since these tended to produce the best random FM waveforms.

As a standard of comparison, the PRO-FM waveforms designed to match the corresponding spectra where also used to estimate the RMS ambiguity functions for th PRO-FMCW case. Each of the three PRO-FMCW waveforms have a BT of approximately 256000.

As in the pulsed case, examining the impact of a Doppler shift of the RMS autocorrelation requires introducing that shift to the RMS autocorrelation equation. For a given Doppler shift, f_{d}, a Doppler shift function, $d[m]$ can be formed such that

$$d[m] = \exp(j2\pi f_{\mathrm{d}} t[m]) \tag{5.26}$$

where $t[m]$ is a normalized time axis. This Doppler shift can be introduced to (3.108) such that the Doppler shifted RMS matched filter response becomes

$$\mathrm{E}\left[\ |\hat{r}[\ell, W, f_{\mathrm{d}}]|^2\ \right] = \mathrm{E}\left[\left|\frac{1}{W}\sum_{m=1}^{W} s^*[m]d^*[m]s[m+\ell]\right|^2\right]. \tag{5.27}$$

where m starts arbitrarily at 1. Expanding the magnitude squared realizes

$$\mathrm{E}\left[\ |\hat{r}[\ell, W, f_{\mathrm{d}}]|^2\ \right] = \frac{1}{W^2}\sum_{m_1,m_2=1}^{W} \mathrm{E}\left[s^*[m_1]d^*[m_1]s[m_1+\ell]s^*[m_2]d^*[m_2]s[m_2+\ell]\right]. \tag{5.28}$$

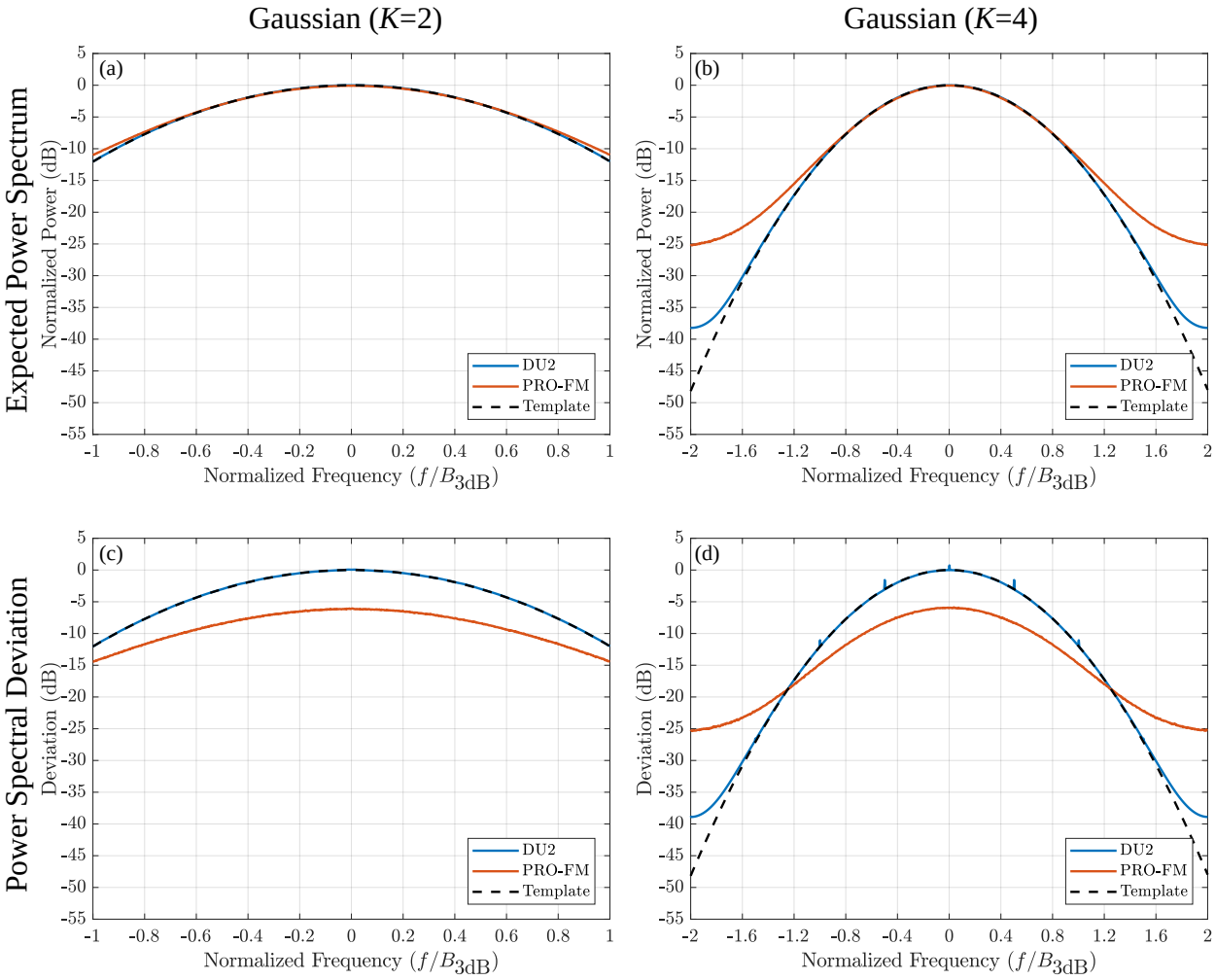

Figure 5.10: Comparison between the analytical spectral characteristics of optimized DU2 based CW-StoWGe and estimated spectral PRO-FMCW characteristics, expected power spectrum for the G2 template (a), expected power spectrum for the G4 template (b), expected power spectral deviation for the G2 template (c), expected power spectral deviation for the G4 template (d)

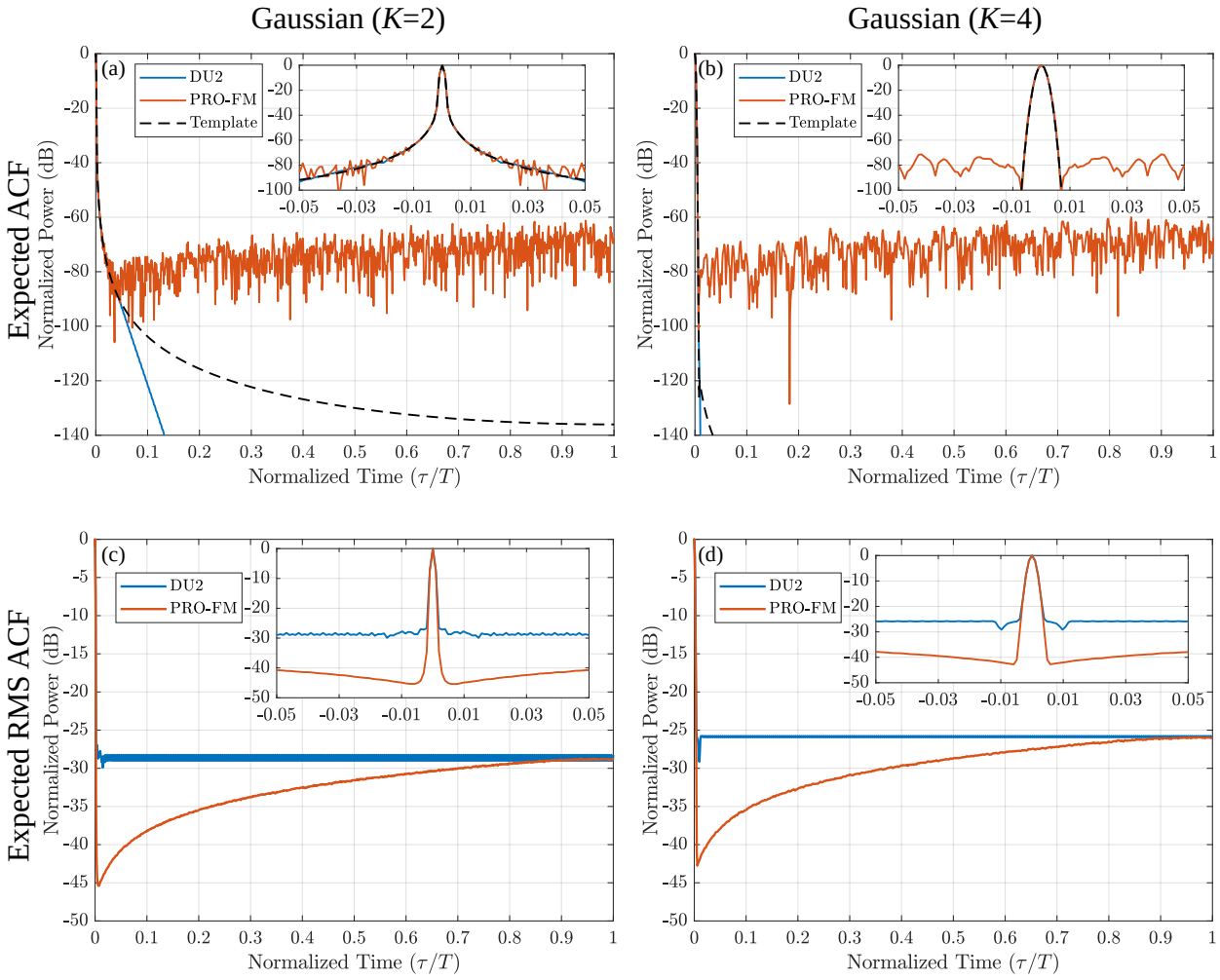

Figure 5.11: Comparison between the analytical temporal characteristics of optimized DU2 based CW-StoWGe and estimated temporal PRO-FMCW characteristics, expected autocorrelation for the G2 template (a), expected autocorrelation for the G4 template (b), expected RMS autocorrelation for the G2 template (c), expected RMS autocorrelation for the G4 template (d)

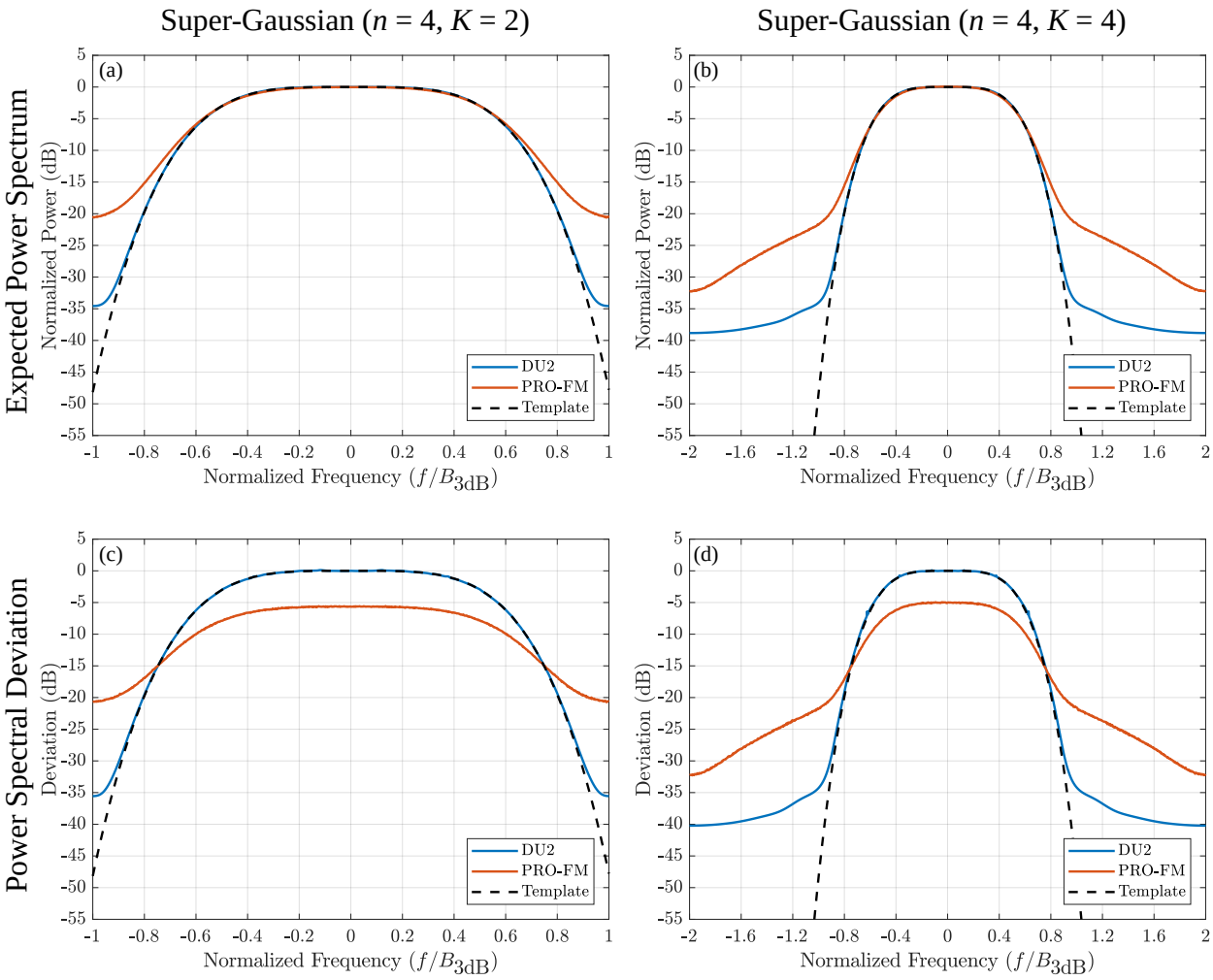

Figure 5.12: Comparison between the analytical spectral characteristics of optimized DU2 based CW-StoWGe and estimated spectral PRO-FMCW characteristics, expected power spectrum for the S4G2 template (a), expected power spectrum for the S4G4 template (b), expected power spectral deviation for the S4G2 template (c), expected power spectral deviation for the S4G4 template (d)

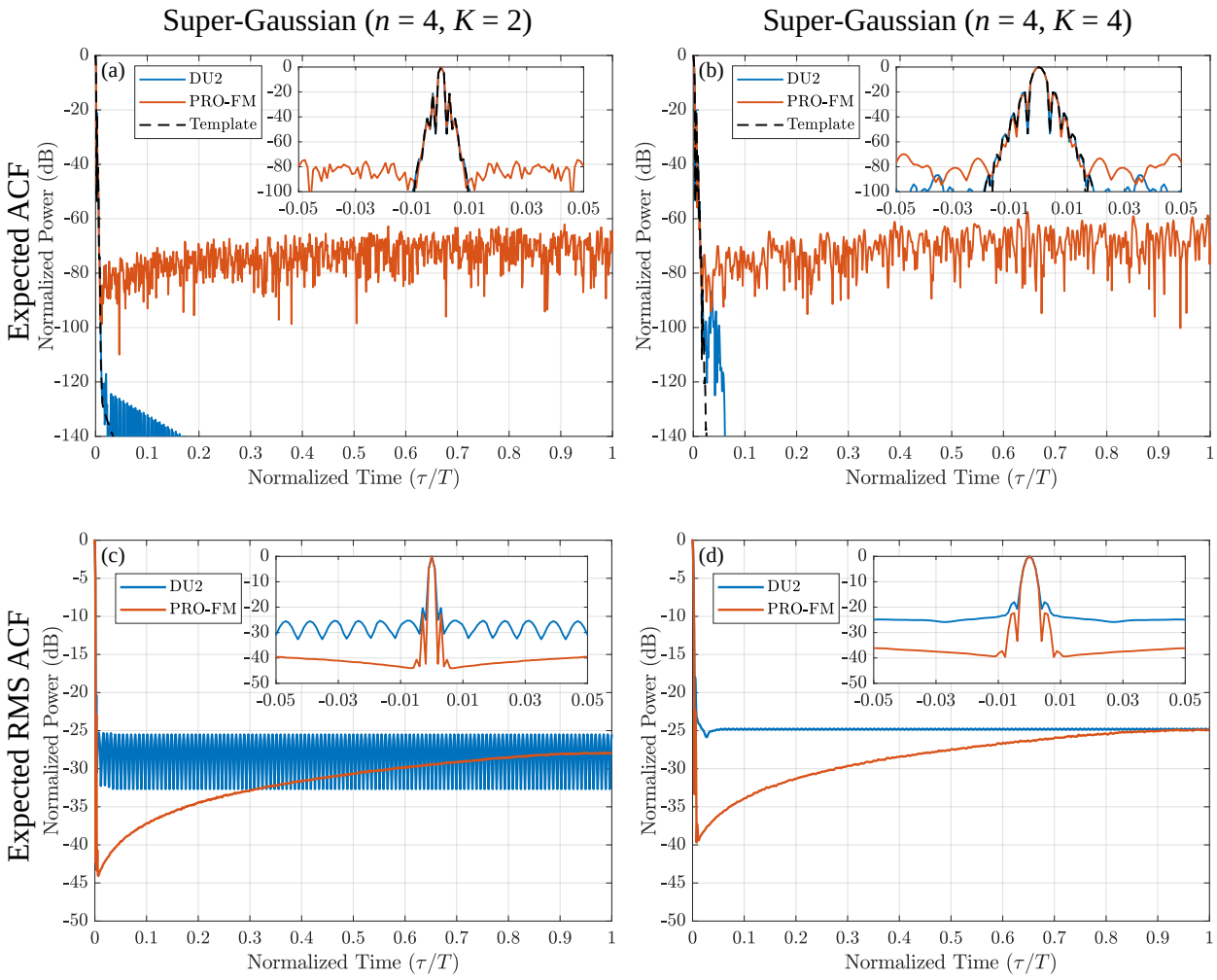

Figure 5.13: Comparison between the analytical temporal characteristics of optimized DU2 based CW-StoWGe and estimated temporal PRO-FMCW characteristics, expected autocorrelation for the S4G2 template (a), expected autocorrelation for the S4G4 template (b), expected RMS autocorrelation for the S4G2 template (c), expected RMS autocorrelation for the S4G4 template (d)

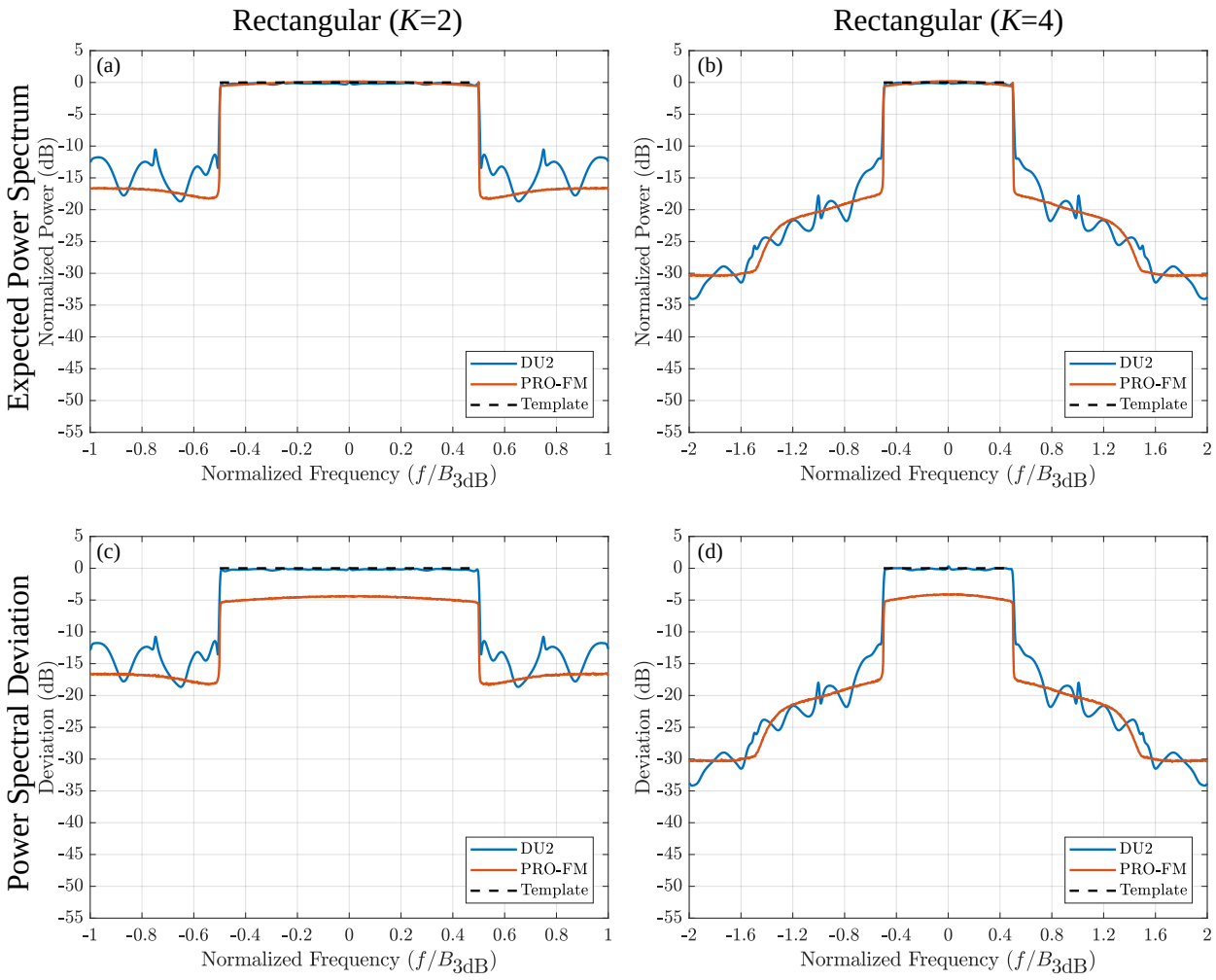

Figure 5.14: Comparison between the analytical spectral characteristics of optimized DU2 based CW-StoWGe and estimated spectral PRO-FMCW characteristics, expected power spectrum for the R2 template (a), expected power spectrum for the R4 template (b), expected power spectral deviation for the R2 template (c), expected power spectral deviation for the R4 template (d)

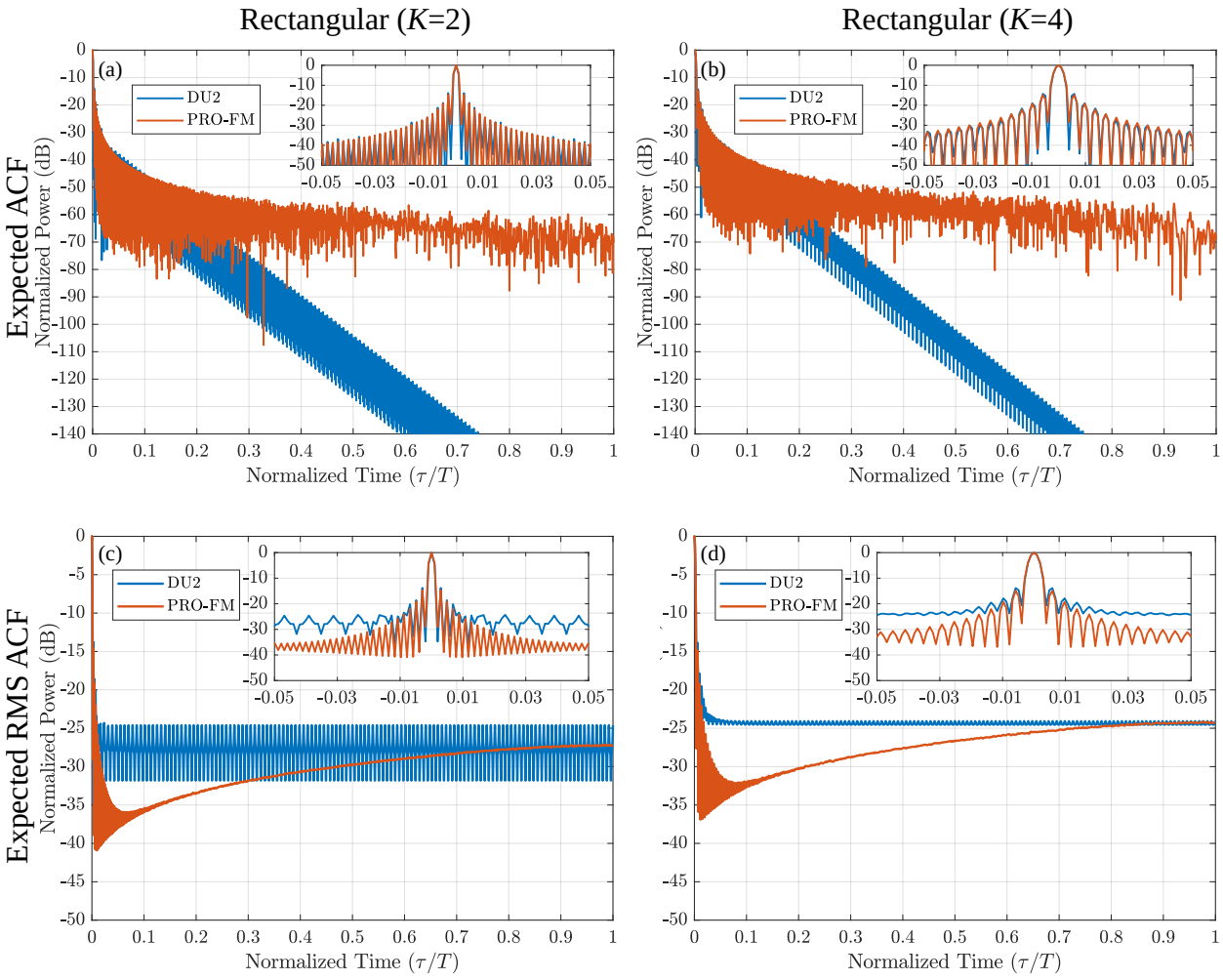

Figure 5.15: Comparison between the analytical temporal characteristics of optimized DU2 based CW-StoWGe and estimated temporal PRO-FMCW characteristics, expected autocorrelation for the R2 template (a), expected autocorrelation for the R4 template (b), expected RMS autocorrelation for the R2 template (c), expected RMS autocorrelation for the R4 template (d)

where the Doppler terms relative to the expectation are constant and can be pulled out of the expectation such that

$$\mathrm{E}\left[\ |\hat{r}[\ell, W, f_{\mathrm{d}}]|^2\ \right] = \frac{1}{W^2} \sum_{m_1, m_2=1}^{W} d^*[m_1] d[m_2] \mathrm{E}\left[s^*[m_1] s[m_1+\ell] s[m_2] s^*[m_2+\ell]\right]. \tag{5.29}$$

where the Doppler shift represents a phase shift on each term in the double summation.

In contrast to the pulsed case where both the expected ambiguity functions and the expected RMS ambiguity functions were shown, only the RMS ambiguity functions are shown here. This is a consequence of the form of CW-StoWGe expected autocorrelation function as defined in Section 3.2.3 which is defined

$$r[\ell] = \lim_{W\to\infty} \frac{1}{W} \sum_{m=1}^{W} s^*[m] s[m+\ell]. \tag{5.30}$$

as a consequence of ergodicity. Multiplying this by a Doppler shift realizes

$$r_{\mathrm{d}}[\ell, f_{\mathrm{d}}] = \lim_{W\to\infty} \frac{1}{W} \sum_{km=1}^{W} d^*[m] s^*[m] s[m+\ell]. \tag{5.31}$$

In general, the Doppler terms will uniformly distribute the phase of each term in (5.31) around the unit circle such that the sum approaches 0 for all non-zero doppler. However, there is an exception when the period of the Doppler frequency matches the period of the cyclo-stationarity. (Recall that (5.30) represents the average correlation at the ℓth delay.)

For the CW-StoWGe cyclo-stationarity period of T_{s} samples, this means the average correlation with ℓ delay can be written as

$$r[\ell] = \frac{1}{T_{\mathrm{s}}} \sum_{m=1}^{T_{\mathrm{s}}} \mathrm{E}[s^*[m] s[m+\ell]] \tag{5.32}$$

Given a period matching Doppler shift such that

$$d[m] = d[m+T_{\mathrm{s}}], \tag{5.33}$$

then (5.31) becomes

$$r[\ell, f_{\mathrm{d}}] = \frac{1}{T_{\mathrm{s}}} \sum_{m=1}^{T_{\mathrm{s}}} d^*[m] \mathrm{E}[s^*[m] s[m+\ell]]. \tag{5.34}$$

Doppler vectors are defined such that they sum to zero of their period, but in (5.34) they are in general each multiplied by a unique constant such that this is not guaranteed. However, if the Doppler shift does not satisfy (5.33) then (5.34) is not longer valid since in general each consecutive set of T_{s} terms would be multiplied by different sets of Doppler terms forcing the correlation to approach zero.

All of this is to say, that the expected CW-StoWGe ambiguity function is not very useful since it is zero valued almost everywhere.

In Fig. 5.16, The RMS ambiguity functions for the DU2 based CW-StoWGe WGFs for the 4 times oversampled templates are plotted in 5.16(a,c,e). Their estimated PRO-FM equivalents are plotted in 5.16(b,d,f). As in the pulsed case, the most notable difference between them is the zero-Doppler valley present in the PRO-FM results which corresponds to the lower near-in autocorrelation sidelobes seen in the PRO-FM results of Figs. 5.11, 5.13, and 5.15. Off from the zero-Doppler axis, the ambiguity functions of the CW-StoWGe and the PRO-FM waveforms are much more similar in that they resemble a thumbtack response that should be expected for random FM waveforms. In contrast to the pulsed waveforms the ambiguity functions extend beyond the y-axis as a consequence of their CW nature. Regardless, beyond the zero-Doppler axis there is little difference between the CW-StoWGe waveforms and previous random FM waveforms.

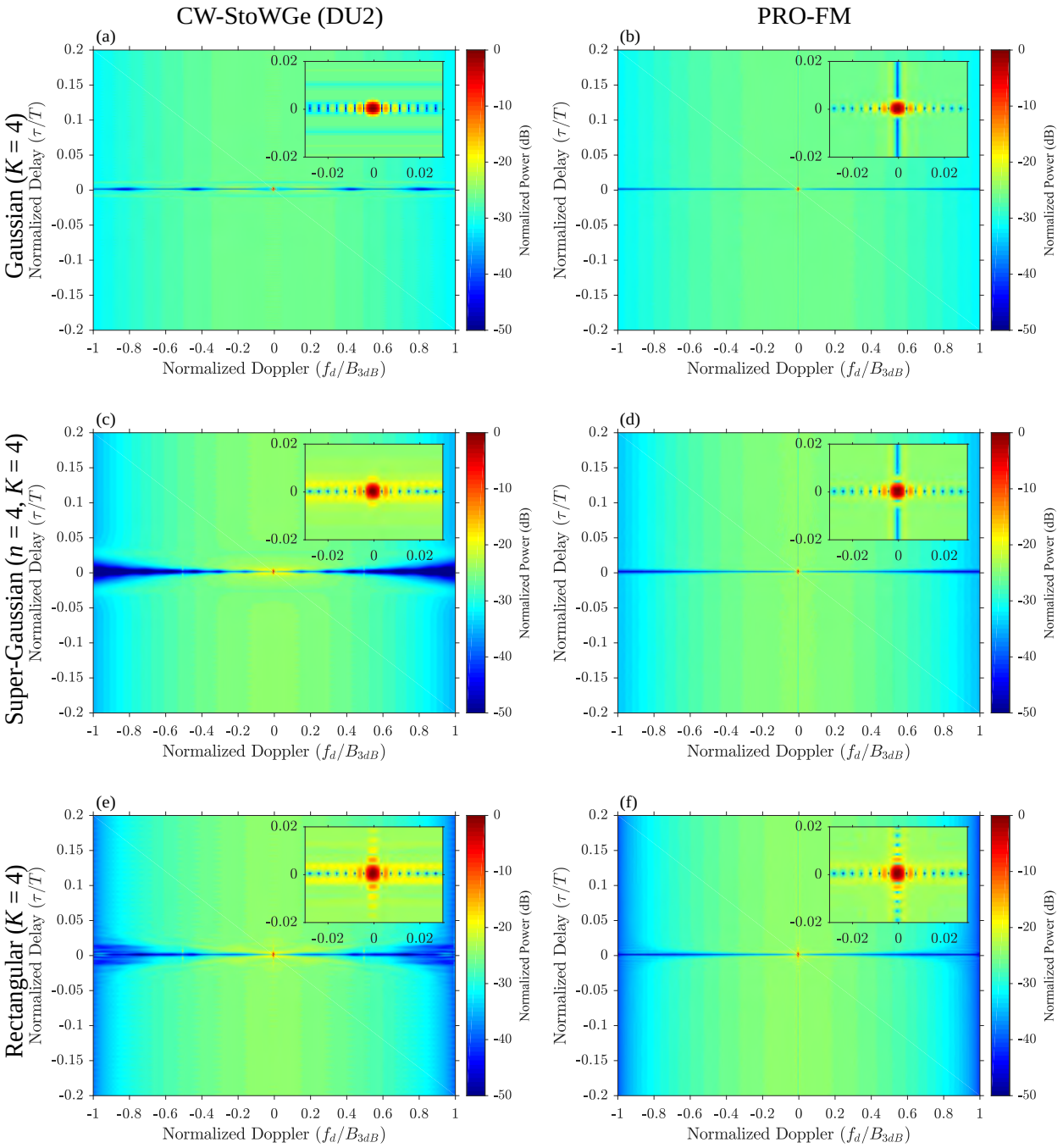

Figure 5.16: Comparison between the analytical RMS ambiguity functions of optimized DU2 based CW-StoWGe and estimated RMS ambiguity functions for PRO-FMCW: CW-StoWGe for the G4 template (a), PRO-FMCW for the G4 template (b), CW-StoWGe for the S4G4 template (c), PRO-FMCW for the S4G4 template (d), CW-StoWGe for the R4 template (e), PRO-FMCW for the R4 template (f)

5.5 CW-StoWGe Experimental Results

In Section 4.5, the analytical pulsed StoWGe model was verified via both simulation and loopback experimentation. Since the CW-StoWGe model significantly differs from its pulsed counterpart it is likewise prudent to verify through similar means. To do so, the analytical models of the optimized CW-StoWGe WGFs from Section 5.4 are compared directly to both simulated estimates as well as physically implemented and measured signals in a loopback configuration. Since the DU2 based WGFs tended to produce the best radar waveforms, these are the ones examined here.

In Figs. 5.17 -5.22 there are three traces. The analytical traces were calculated directly using the definition of the CW-StoWGe model and were lifted directly from the figures in Section 5.4.

The simulation traces were calculated by estimating in a similar manner as the PRO-FMCW results in Section 5.4.3. For each of the DU2 WGFs for each of the spectral templates, a sample function of 1024000 samples was instantiated resulting in a BT of approximately 512000 for the 2 times oversampled templates and about 256000 for the 4 times oversampled templates respectively and a total of 6 quasi-CW radar waveforms. These waveforms were then segmented and evaluated for the various metrics via the methods described in Section 3.3.2.

Finally, the CW-StoWGe signals were implemented on the same loopback system as the pulsed StoWGe waveforms. Each of the 6 CW-StoWGe waveforms were sinc interpolated by a factor of 12.5 and then pojected onto a constant modulus envelope. They were then digitally up-converted to a center frequency of -$f_s/4$. Given a DAC rate of 2.5 GHz, the signals have a 3 dB bandwidth of either 100 MHz or 50 MHz for the 2 times oversamples or the 4 times oversampled spectra respectively and a center frequency of 625 MHz. The signals were implemented on an arbitrary waveform generator at a passband rate of 2.5 GHz, before being linearly amplified, then attenuated, and finally recorded at passband at a rate of 2.5 GHz. The received signals were then digitally down converted and re-sampled to a baseband sample rate of 200 MHz. The various metrics were then calculated using this data via the same methods as for the simulated waveforms. The PRO-FMCW waveforms were likewise implemented in the same manner as a standard of comparison.

In examining Figs. 5.17-5.22, the analytical traces are almost entirely hidden by the simulated

traces especially for the the spectral results. Such a match indicates that the analytical analysis of the model accurately reflects the statistics of the actual sample functions. However, this excellent match is notably absent in the sidelobe regions of the expected autocorrelations of 5.18(a,b), 5.20(a,b), and 5.22(a,b) where the analytical and simulation traces clearly diverge. This behavior was seen in the pulsed StoWGe case as well and for a similar reason. Here, the expected autocorrelation traces represents the autocorrelation of an infinitely long sample function, but of course the simulated result is finite thus the traces diverge. In linear terms, the error between the simulated trace and the analytical traces is very small, but on a dB scale, the difference is several orders of magnitude where the value of the analytical trace is also very small. It can be expected that with longer and longer sample functions (waveforms), the simulated traces would approach the analytical traces.

On the other hand, there are more meaningful differences between the simulated traces and the loopback traces in a similar manner to the pulsed StoWGe results, but this is to be expected. In that section the loopback spectral distortion relative to the ideal simulated results was attributed to the signal upsampling and hardware filtering effects. Since the CW-StoWGe waveforms were upsampled in exactly the same way and implemented on the same equipment, it is reasonable that they should exhibit similar forms of distortion. In fact, in comparing the loopback traces of Figs. 4.25(a) and 5.17(a), where both the pulsed and the CW waveforms achieved nearly the same spectral result, the distortion from the ideal simulated and analytical results is nearly identical. This effect is likely attributable to filtering effects within the loopback system. To further verify this conclusion, PRO-FMCW waveforms of the same BT and spectral shapes were also implemented in loopback. In Fig. 5.23, the R2 expected spectrum loopback results of both the CW-StoWGe and the PRO-FMCW waveforms are plotted together where they both exhibit a similar distortion effect over the passband region. In previous work, PRO-FMCW waveforms have not shown this behavior [25]. Consequently, it is reasonable to conclude this is a filtering effect and not a fundamental issue with the implementation of the CW-StoWGe waveforms.

Despite these spectral effects, there is relatively little distortion in any of the corresponding

temporal results. Overall after accounting for the upsampling process and the loopback systems filtering effects, there is an excellent match between the simulated results and loopback results indicating the suitability of the CW-StoWGe waveforms for physical implementation.

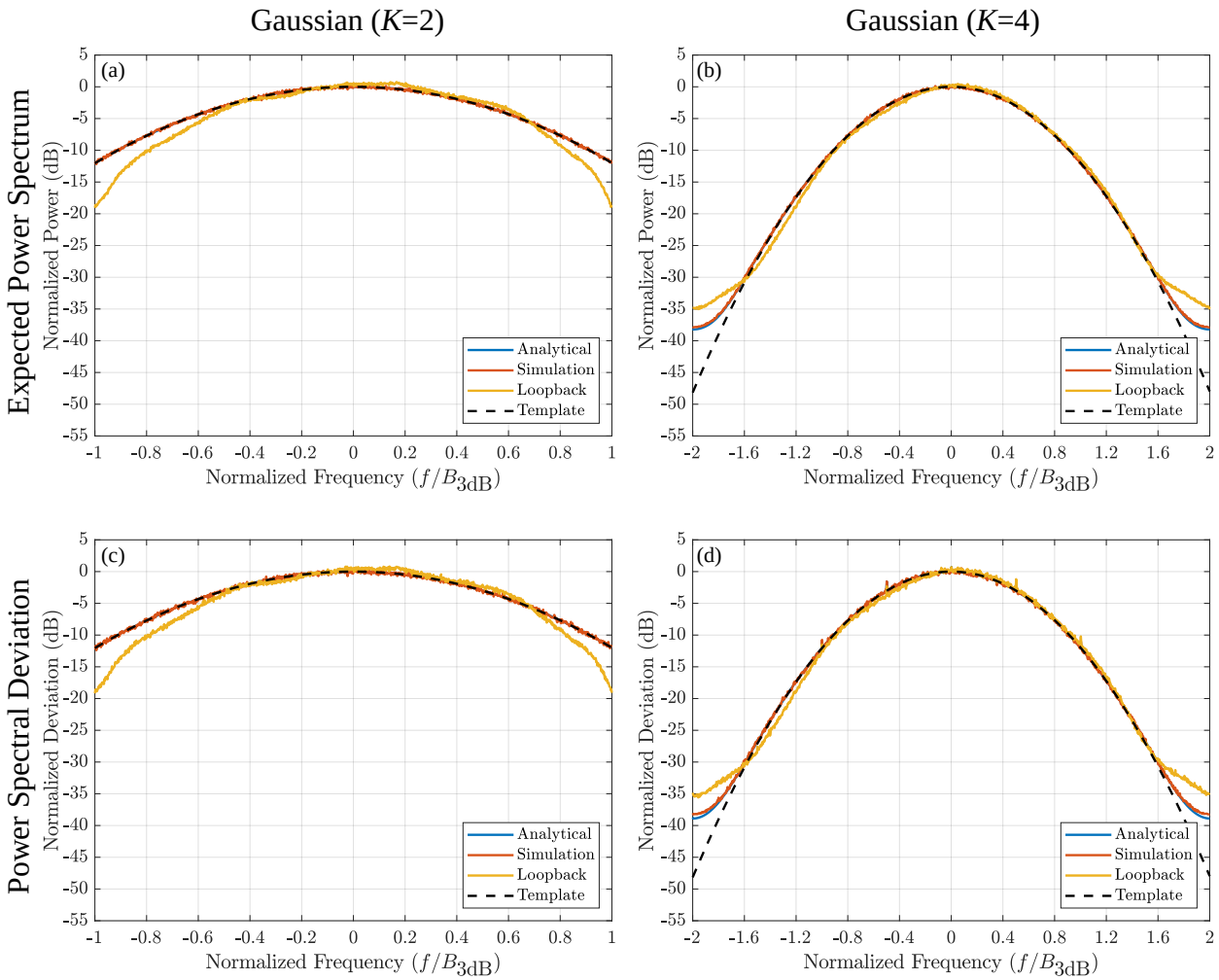

Figure 5.17: Optimized, DU2 based, CW-StoWGe analytical, simulated, and loopback spectral results for Gaussian templates: expected power spectrum for the G2 template (a), expected power spectrum for the G4 template (b), expected power spectral deviation for the G2 template (c), expected power spectral deviation for the G4 template (d)

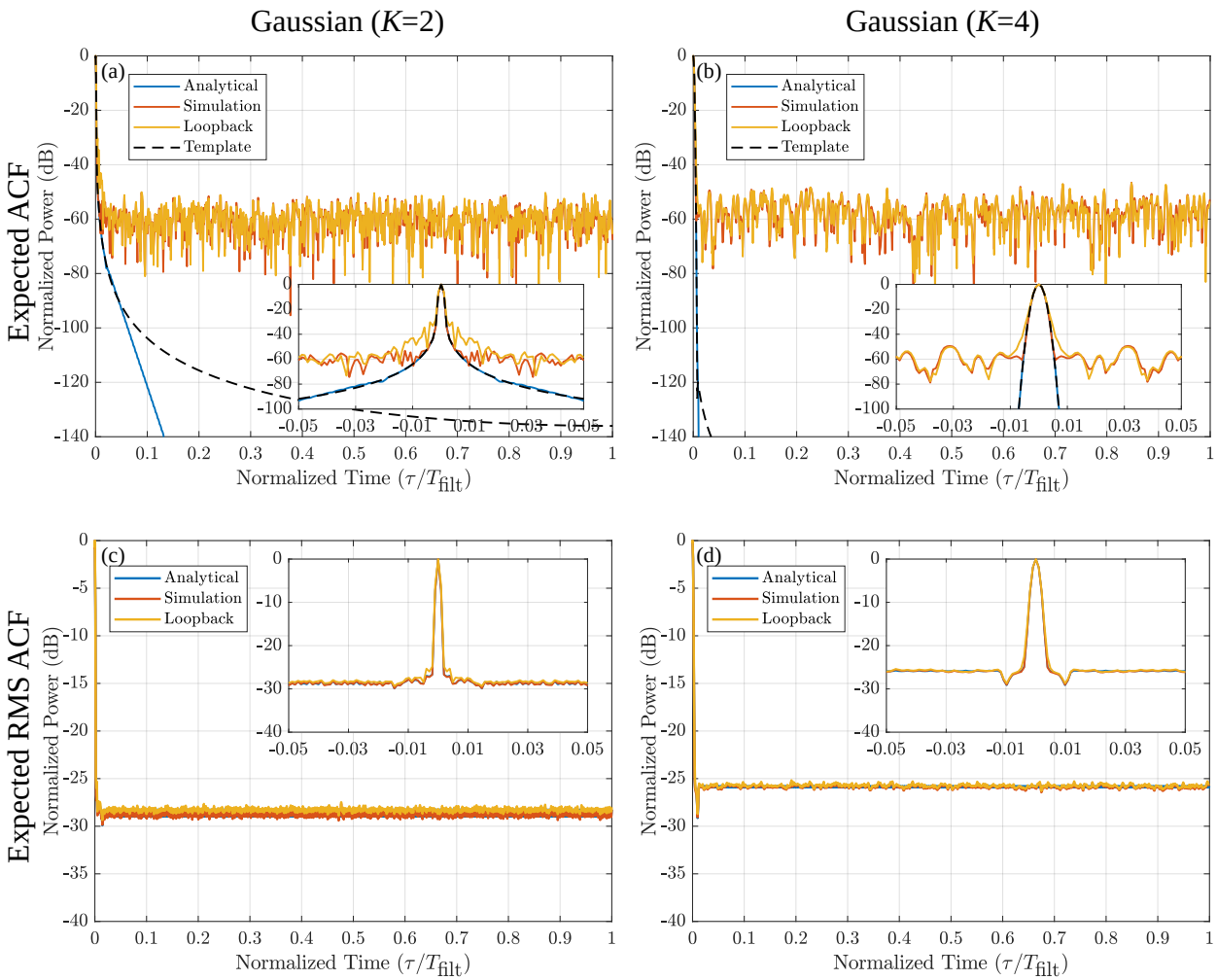

Figure 5.18: Optimized, DU2 based, CW-StoWGe analytical, simulated, and loopback temporal results for Gaussian templates: expected autocorrelation for the G2 template (a), expected autocorrelation for the G4 template (b), expected RMS autocorrelation for the G2 template (c), expected RMS autocorrelation for the G4 template (d)

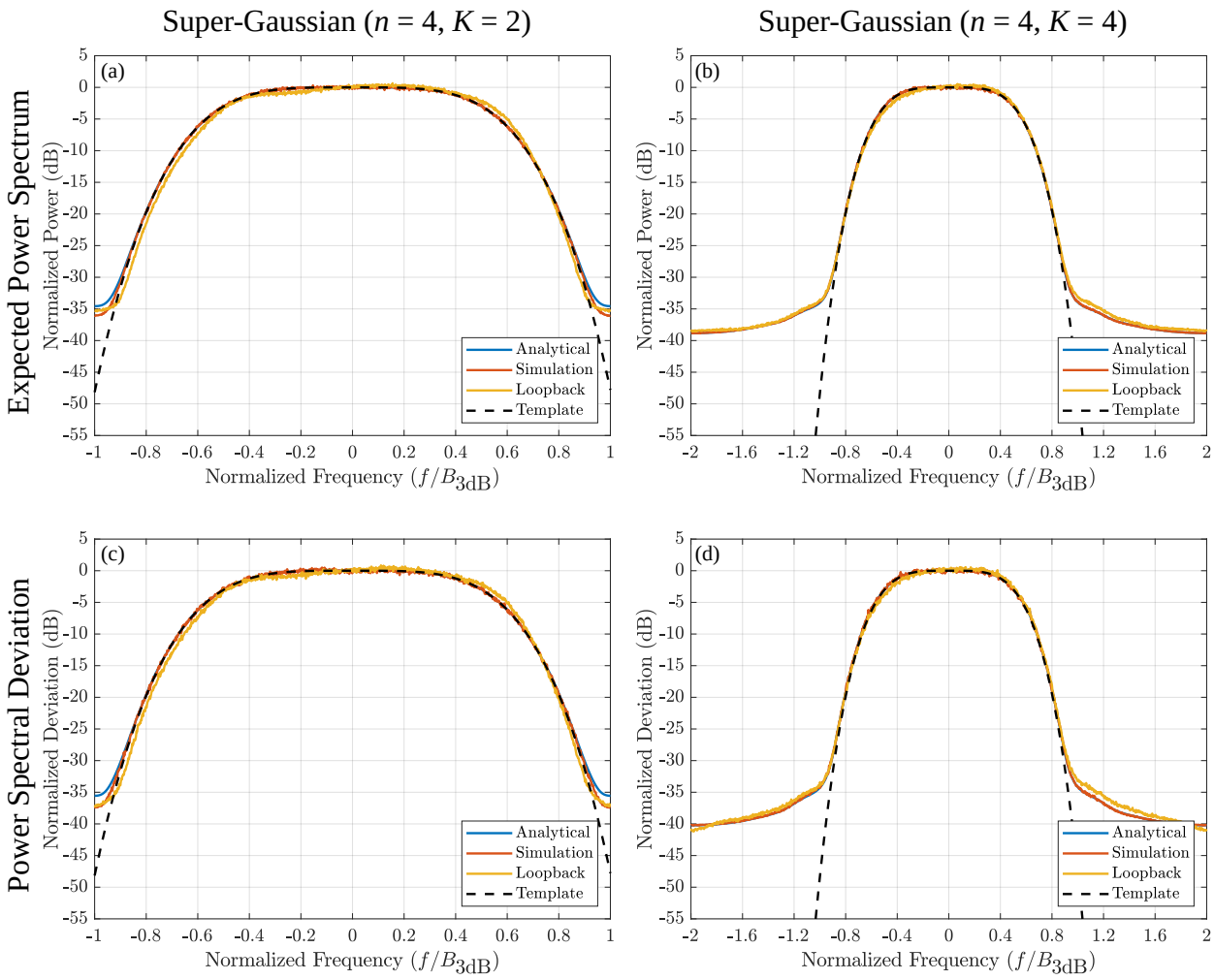

Figure 5.19: Optimized, DU2 based, CW-StoWGe analytical, simulated, and loopback spectral results for super-Gaussian templates: expected power spectrum for the S4G2 template (a), expected power spectrum for the S4G4 template (b), expected power spectral deviation for the S4G2 template (c), expected power spectral deviation for the S4G4 template (d)

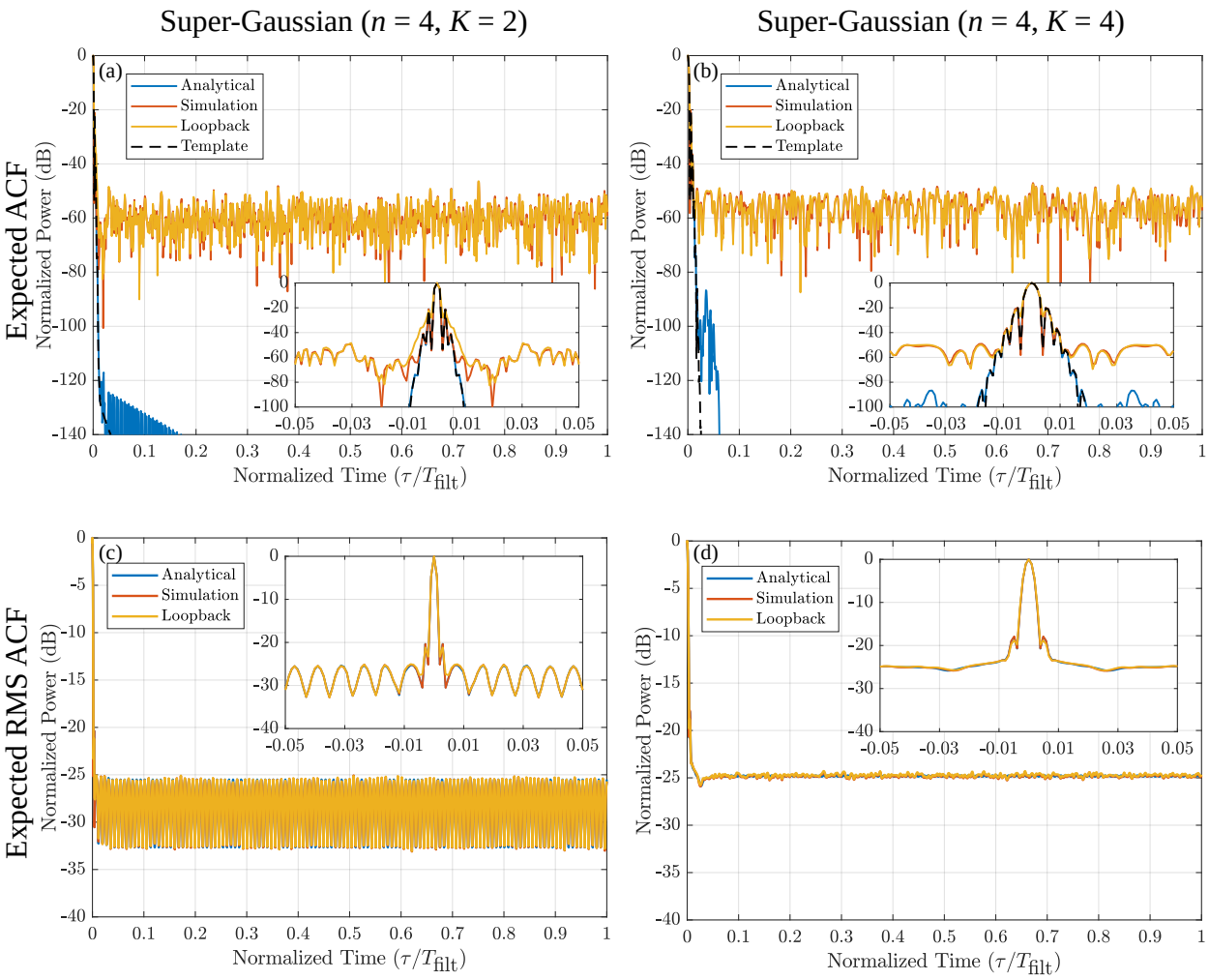

Figure 5.20: Optimized, DU2 based, CW-StoWGe analytical, simulated, and loopback temporal results for super-Gaussian templates: expected autocorrelation for the S4G2 template (a), expected autocorrelation for the S4G4 template (b), expected RMS autocorrelation for the S4G2 template (c), expected RMS autocorrelation for the S4G4 template (d)

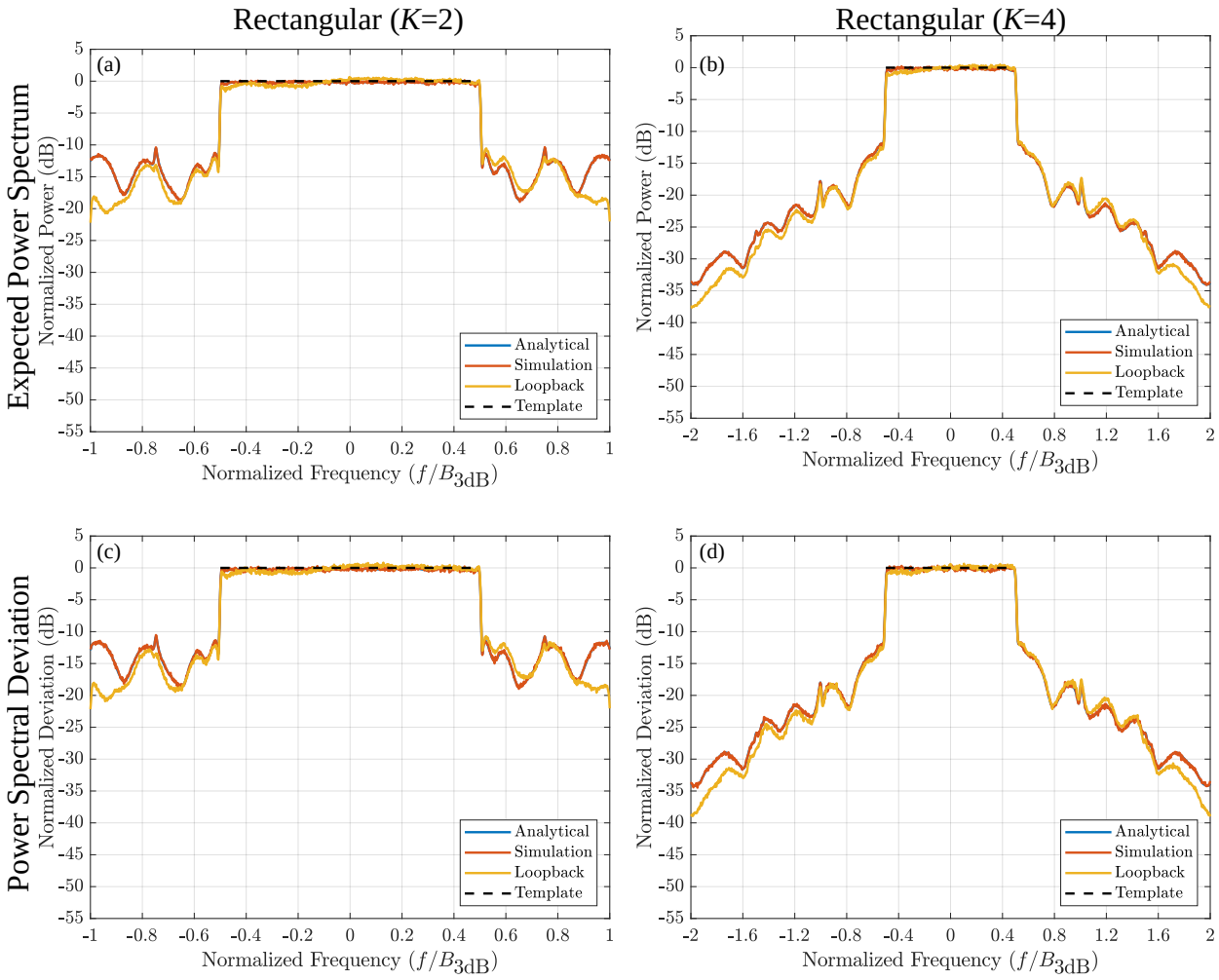

Figure 5.21: Optimized, DU2 based, CW-StoWGe analytical, simulated, and loopback spectral results for rectangular templates: expected power spectrum for the R2 template (a), expected power spectrum for the R4 template (b), expected power spectral deviation for the R2 template (c), expected power spectral deviation for the R4 template (d)

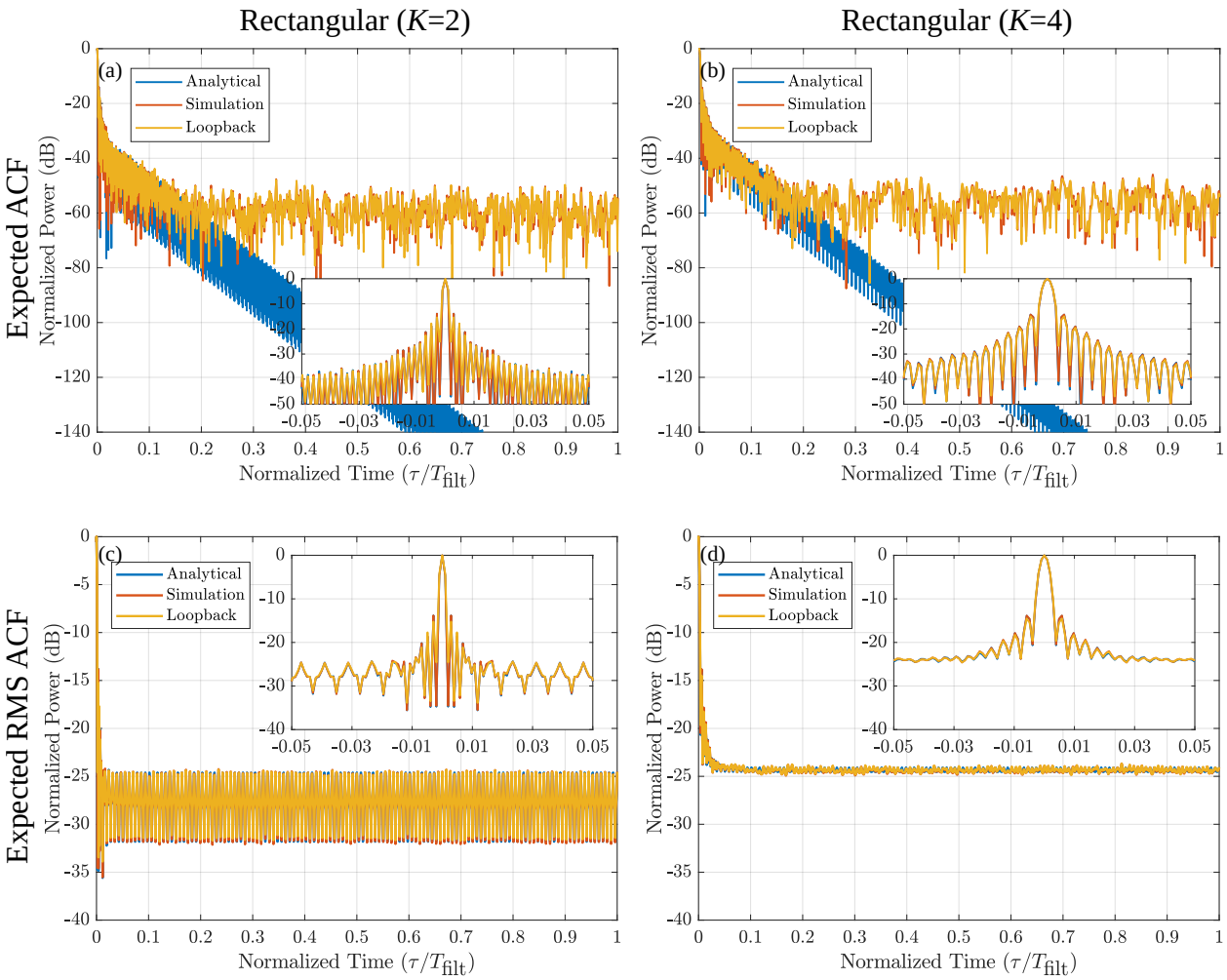

Figure 5.22: Optimized, DU2 based, CW-StoWGe analytical, simulated, and loopback temporal results for rectangular templates: expected autocorrelation for the R2 template (a), expected autocorrelation for the R4 template (b), expected RMS autocorrelation for the R2 template (c), expected RMS autocorrelation for the R4 template (d)

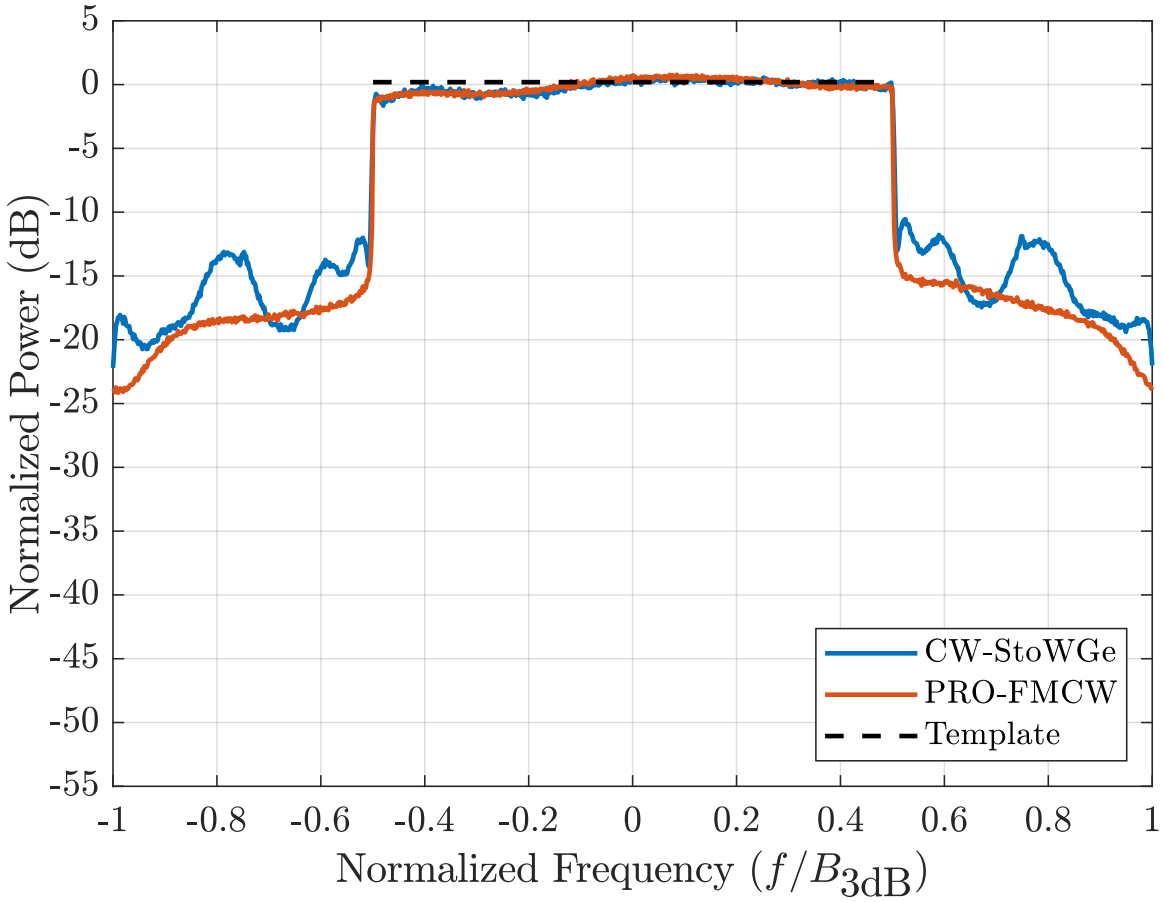

Figure 5.23: Comparison between the estimated expected spectrum of the loopback measured CW-StoWGe DU2 waveforms and the loopback measured PRO-FMCW waveforms for the R2 template. Both sets exhibit similar passband spectrum distortion due to hardware effects.

5.6 CW-StoWGe Summary

Similar the pulsed case, the goal behind CW-StoWGe was to enable the computationally cheap instantiation of useful radar waveforms, but here for continuous wave (CW) purposes. Likewise to pulsed StoWGe, the cheap efficient instantiation is a fundamental characteristic of the CW-StoWGe model where the phase is created through the convolution of a random variables with a frequency shaping filter, but guaranteeing their usefulness requires an optimization of the signal model's underlying parameters.

To further complicate matters, in the CW case it is necessary to select a means of estimating

the power spectrum and autocorrelation since there are numerous methods to choose from. For the sake of tractability, periodograms and pseudo-pulse processing were chosen for then. Under these conditions, the gradient based minimization of the expected frequency template error (EFTE) cost function defined in Section 5.2, was shown in Section 5.4 to be an effective means of shaping the CW-StoWGe models expected power spectrum while remaining constant modulus and retaining noise-like characteristics such as lower autocorrelation sidelobes with greater integration time.

The CW-StoWGE model was optimized for the same set of desired spectra and random variable distributions as in the pulsed case, but with varying degrees of the partial response parameter, L, and the inter-variable spacing, T_s. Unsurprisingly, it was shown that higher values of T_s and L generally result in better, optimized WGFs. (more design degrees of freedom.) However, it was also shown, as in the pulsed case, that more restrictive random variable distributions tend to result in better matches to desired expected spectra, a surprising results that warrants further investigation.

Then to evaluate the CW-StoWGe waveforms against previous research, the CW-StoWGe waveforms were compared directly against previously designed FMCW noise waveforms. Although, a comparison was more difficult in the CW case since existing noise-like FMCW waveforms are instantiated on segment by segment basis, while the CW-StoWGe waveforms instantiated continuously. This difference in design resulted in the previous FMCW waveforms achieved better (lower) near-in sidelobe while in comparison to CW-StoWGe. However, this advantage only extended as far as the design segment length. Beyond this this time frame, the previous waveforms and the CW-StoWGe waveforms achieve similar sidelobe levels.

Finally, CW-StoWGe radar waveforms were instantiated based on the best WGFs as determined by the optimization and the following analysis, and then implemented on hardware in a loopback configuration. These loopback results were then compared to both the analytical and the simulated results where it was found that the physically implemented signals were nearly identical to the analytical and simulated results demonstrating the suitability of the CW-StoWGe waveforms for physical implementation. Overall this chapter has shown that the design and optimization of

the CW-StoWGe model enables the computationally cheap instantiation of random FMCW waveforms with desirable spectral and autocorrelation characteristics.

Chapter 6

Conclusions & Future Work

Previous work and analysis has demonstrated the performance advantages of random FM waveforms. Like general noise waveforms, they are unambiguous such that there are no unambiguous range considerations since each waveform is unique. Additionally, their autocorrelation sidelobe performance can be improved arbitrarily through increasing their time-bandwidth (BT) product. Unlike general noise waveforms, random FM waveforms are constant in amplitude meaning they are amenable to implementation on high-powered amplifiers (HPA). However, maintaining a constant modulus envelope while also achieving sufficient spectral containment and good autocorrelation properties has proven difficult in practice. One general approach is to optimize individual waveforms such that they achieve these goals. While this has been shown to be an effective means of meeting these goals, it is also a computationally expensive process making real-time operation difficult. Alternatively, random FM waveforms have been defined as a stochastic process such that they can be produced with minimal computational expense via a random number generator (RNG). But in past work, the waveforms produced from these approaches have been static in that their average spectra is Gaussian with no means of optimizing them to suit other design goals.

The primary motivation of this work was to achieve the performance level of the optimization approach while circumventing its computational cost by implementing the waveforms in a manner consistent with the second approach. To do so, instead of thinking of random FM waveforms as singular entities to be examined and optimized on an individual basis, they were considered as equivalent members of a waveform family as defined by some underlying waveform generating function (WGF). In this way, an entire family of waveforms can be optimized simultaneously by optimizing their underlying WGF thus eliminating the need to perform a computationally expen-

sive optimization of each and every waveform. The waveforms themselves are then produced by instantiating sample functions from the underlying WGF.

To implement this framework, Chapter 3 introduced the mathematical basis behind random variables and stochastic processes, and then explained how these same concepts and analytical tools can be used to address the design and analysis of random FM waveforms. Most importantly, Section 3.4 introduced a series of eight metrics which translate the fundamentals of stochastic processes to a radar centered perspective. Then, by leveraging these ideas, Chapters 4 and 5 defined the pulsed stochastic waveform generation (Pulsed StoWGe) and the continuous-wave stochastic waveform generation (CW-StoWGe) models for pulsed and CW random FM waveforms respectively.

In summary, these models define a parameterized WGF. Then, by optimizing these parameters the WGFs can be designed to produce radar waveforms with desirable characteristics. The approach taken here was to optimize these parameters according to the expected frequency template error (EFTE) cost function which measures the average error between the expected power spectral density of the waveforms produced by their respective WGF and some desired spectrum. By minimizing the EFTE, the instantiated waveforms would on average have a power spectrum which resembles the desired spectrum. The question in performing the gradient descent optimization was to what degree can the EFTE be minimized given a desired spectrum and the pulsed or CW StoWGe models. To answer this question an extensive set of combinations of desired spectra and model parameters were chosen and systematically tested by minimizing the EFTE cost function via gradient descent techniques. In chapters 4 and 5 it was clearly shown that for certain combinations of pulsed or CW-StoWGe parameters the EFTE cost function can be effectively minimized such that the waveforms instantiated via the optimized WGFs achieve an expected spectrum that is very similar to the desired spectrum in a squared error sense. In particular, using discrete uniform random variables with two states (DU2) and maximizing the model's degrees of freedom were found to result in good spectral matches.

To further prove out the process and to evaluate the resulting waveforms, random FM wave-

forms were then instantiated from these WGFs and then verified to posses the expected characteristics based on their respective WGFs. But perhaps most importantly, their expected spectral, autocorrelation, and cross-correlation characteristics were all verified in an experimental demonstration on physical equipment showing that the mathematical theory accurately describes reality. These waveforms were shown in monte carlo comparisons to be largely equivalent to some previously designed random FM waveforms in terms of both spectral containment and cross-correlation, but perhaps worse in autocorrelation performance on a per waveform basis. Despite this relative shortcoming, the StoWGe waveforms do enjoy a massive advantage in terms of there instantiation since there is no need to optimize them on a per waveform basis. Consequently, the pulsed and CW StoWGe waveforms can easily be implemented in a real-time through an RNG once the underlying WGF itself is optimized.

Even so, there is likely further room for improvement within the StoWGe framework. This work optimized the expected spectrum of the WGFs. Considering the importance of the spectrum for radar waveforms, this was a sensible place to start. However, there could be advantages in optimizing other metrics as well. For example the expected spectrum is an average measure, so optimizing it does not directly control how the StoWGe waveforms behave on and individual level. Consequently, it can take a fair amount of coherent integration before the sample average spectrum strongly resembles the expected spectrum. To address this in future work, it may be possible to optimize the individual measures of Section 3.4 such as the power spectral deviation or the RMS autocorrelation to improve the per waveform spectral and autocorrelation properties of the StoWGe waveforms. Although such an optimization would likely be much more complicated on account of the per waveform metrics being fourth order functions of the WGF rather than second order functions as in the expected spectrum. Regardless, such an approach is worth exploring in future work.

More broadly, this work applied optimization to theory to the previously conceived idea of explicitly treating random FM waveforms as stochastic processes rather than as individual items. The pulsed and CW StoWGe models are simply a means of enabling the optimization of stochastic

processes for radar purposes, but the pulsed and CW-StoWGe models are not necessarily the only way to do so. In fact, there is a very real possibility that may not be the best method in first place, but until other models are formulated this of course cannot be known. Even so, StoWGe has now been shown to produce useful, physical random FM waveforms, but perhaps from a more forward looking perspective StoWGe's greatest contribution is simply showing that designing stochastic processes in such a dynamic way is even possible. With this in mind, future work should explore other possible WGF models and parameterizations in an effort to improve upon the per-waveform and aggregate performance characteristics achieved by the pulsed and CW-StoWGe models themselves.

Appendix A

StowGe Derivations

This appendix derives the second and fourth order moments of the pulsed StoWGe waveform generating function of Chapter 4 and the CW-StoWGe waveform generating function of Chapter 5 such that the metrics in Section 3.4 can be analytically evaluated for these models. Additionally, the derivatives with respect to the pulsed and CW EFTE cost functions are also derived.

A.1 Pulsed StoWGe

Evaluation of the Pulsed StoWGe EFTE and its gradient requires the calculation of its second and fourth order moments. Section A.1.1 derives these moments, Section A.1.2 uses the second moment to derive the gradients of the EFTE cost function, and Section A.1.3 derives the gradient with respect to selected random variable distributions.

A.1.1 Pulsed StoWGe Moments

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

$$\mathbf{s} = \exp(j\boldsymbol{\phi}) \tag{A.1}$$

where

$$\boldsymbol{\phi} = \mathbf{B}\mathbf{x} + \boldsymbol{\mu} \tag{A.2}$$

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 $\boldsymbol{\phi}$. By definition, the

second moment or correlation between two samples in (A.1) is

$$\mathrm{E}\left[s_{m_1} s_{m_2}^*\right] = \mathrm{E}\left[\exp(j((\mathbf{b}_{m_1} - \mathbf{b}_{m_2})\mathbf{x} + \mu_{m_1} - \mu_{m_2}))\right] \tag{A.3}$$

where m_1 and m_2 are arbitrary. Since the mean terms are constant they can be pulled out of the expectation such that

$$\mathrm{E}\left[s_{m_1} s_{m_2}^*\right] = \exp(j(\mu_{m_1} - \mu_{m_2}))\mathrm{E}\left[\exp(j(\mathbf{b}_{m_1} - \mathbf{b}_{m_2})\mathbf{x})\right]. \tag{A.4}$$

Next, the inner product inside the argument of the complex exponential can be written as a summation. Since this summation exists inside and exponential, the whole quantity can be written as a product realizing

$$\mathrm{E}\left[s_{m_1} s_{m_2}^*\right] = \exp(j(\mu_{m_1} - \mu_{m_2}))\mathrm{E}\left[\prod_{n=1}^{N} \exp(j(b_{m_1,n} - b_{m_2,n})X_n)\right], \tag{A.5}$$

where $b_{m_1,n}$ is the element in the m_1th row and the nth column of $\mathbf{B}$ and x_n is the nth element of $\mathbf{x}$. Next and perhaps most importantly, each member of $\mathbf{x}$ is independent such that the expectation of their product can be written as product of their expectations. Using this fact allows for the expectation to be pulled inside of the product such that

$$\mathrm{E}\left[s_{m_1} s_{m_2}^*\right] = \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{n=1}^{N} \mathrm{E}\left[\exp(j(b_{m_1,n} - b_{m_2,n})X_n)\right]. \tag{A.6}$$

The individual terms can be recognized as the form of the characteristic function for the random variable X_n as defined in Section 3.1 such that

$$\psi_X(\omega) = \mathrm{E}[\exp(j\omega X)] = \int_{-\infty}^{\infty} \exp(j\omega x) f_X(x) dx. \tag{A.7}$$

where $\omega = (b_{m_1,n} - b_{m_2,n})$. (A.6) becomes

$$\mathrm{E}\left[s_{m_1} s^*_{m_2}\right] = \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{n=1}^{N} \psi_{X_n}(b_{m_1,n} - b_{m_2,n}). \tag{A.8}$$

Then by selecting a distribution for each X_n, the correlation between the samples can be evaluated via (A.8)

Similarly, the fourth moment between samples of (A.1) is by definition

$$\mathrm{E}\left[s_{m_1} s^*_{m_2} s^*_{m_3} s_{m_4}\right] = \mathrm{E}\left[\exp(j((\mathbf{b}_{m_1} - \mathbf{b}_{m_2} - \mathbf{b}_{m_3} + \mathbf{b}_{m_4})\mathbf{x} + \mu_{m_1} - \mu_{m_2} - \mu_{m_3} + \mu_{m_4}))\right]. \tag{A.9}$$

Following an identical set of steps as above, the fourth moment can be shown to be

$$\mathrm{E}\left[s_{m_1} s^*_{m_2} s^*_{m_3} s_{m_4}\right] = \exp(j(\mu_{m_1} - \mu_{m_2} - \mu_{m_3} + \mu_{m_4})) \prod_{n=1}^{N} \psi_{X_n}(b_{m_1,n} - b_{m_2,n} - b_{m_3,n} + b_{m_4,n}), \tag{A.10}$$

which can be evaluated after choosing a distribution for each X_n.

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

$$J_{\mathrm{p}} = \left|\left|\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] - \mathbf{u}\right|\right|_2^2, \tag{A.11}$$

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

$$\frac{\partial J_{\mathrm{p}}}{\partial b_{k,n}} = \frac{\partial}{\partial b_{k,n}} \left|\left|\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] - \mathbf{u}\right|\right|_2^2 \tag{A.12}$$

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

$$\frac{\partial J}{\partial b_{k,n}} = 2\left(\frac{\partial \mathrm{E}[\hat{\mathbf{s}}_\mathrm{f}]}{\partial b_{k,n}}\right)^T (\mathrm{E}[\hat{\mathbf{s}}_\mathrm{f}] - \mathbf{u}) \tag{A.13}$$

where the derivative is applied to $\mathrm{E}[\hat{\mathbf{s}}_\mathrm{f}]$ on an element wise basis. For the wth sample of the expected spectrum, the derivative is

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = \sum_{m_1,m_2=1}^{M} a_{w,m_1} a^*_{w,m_2} \exp(j(\mu_{m_1} - \mu_{m_2})) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^{N} \psi_{X_p}(b_{m_1,p} - b_{m_2,p})\right). \tag{A.14}$$

The only possible 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

$$\begin{aligned}\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = \quad & a_{w,k} \exp(j\mu_k) \sum_{\substack{m_2=1 \\ m_2 \neq k}}^{M} a^*_{w,m_2} \exp(-j\mu_{m_2}) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^{N} \psi_{X_p}(b_{k,p} - b_{m_2,p})\right) \\ & + a^*_{w,k} \exp(-j\mu_k) \sum_{\substack{m_1=1 \\ m_1 \neq k}}^{M} a_{w,m_1} \exp(j\mu_{m_1}) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^{N} \psi_{X_p}(b_{m_1,p} - b_{k,p})\right)\end{aligned} \tag{A.15}$$

The two summations are complex conjugates of each other such that

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = 2\Re\left\{ a_{w,k} \exp(j\mu_k) \sum_{\substack{m=1 \\ m \neq k}}^{M} a^*_{w,m} \exp(-j\mu_m) \frac{\partial}{\partial b_{k,n}} \left(\prod_{p=1}^{N} \psi_{X_p}(b_{k,p} - b_{m,p})\right)\right\} \tag{A.16}$$

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

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = 2\Re\left\{ \sum_{\substack{m=1 \\ m \neq k}}^{M} a_{w,m} a^*_{w,k} \exp\left(j(\mu_m - \mu_k)\right) \frac{\partial \psi_{X_p}(b_{m,n} - b_{k,n})}{\partial b_{k,n}} \prod_{\substack{p=1 \\ p \neq n}}^{N} \psi_{X_p}(b_{m,p} - b_{k,p})\right\} \tag{A.17}$$

From (A.17) and (A.13), each partial derivative with respect to the elements of $\mathbf{B}$ can be placed into the structure $\nabla_{\mathbf{B}} J_{\mathrm{p}}$ such that

$$\nabla_{\mathbf{B}} J_{\mathrm{p}} = \begin{bmatrix} \frac{\partial J_{\mathrm{p}}}{\partial b_{1,1}} & \frac{\partial J_{\mathrm{p}}}{\partial b_{1,2}} & \cdots & \frac{\partial J_{\mathrm{p}}}{\partial b_{1,N}} \\ \frac{\partial J_{\mathrm{p}}}{\partial b_{2,1}} & \frac{\partial J_{\mathrm{p}}}{\partial b_{2,2}} & \cdots & \frac{\partial J_{\mathrm{p}}}{\partial b_{2,N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J_{\mathrm{p}}}{\partial b_{M,1}} & \frac{\partial J_{\mathrm{p}}}{\partial b_{M,2}} & \cdots & \frac{\partial J_{\mathrm{p}}}{\partial b_{M,N}} \end{bmatrix} \tag{A.18}$$

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 $\boldsymbol{\mu}$ is

$$\frac{\partial J_{\mathrm{p}}}{\partial \mu_k} = 2 \left(\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f}}]}{\partial \mu_k, n} \right)^T (\mathrm{E}[\hat{s}_{\mathrm{f}}] - \mathbf{u}) \tag{A.19}$$

where μ_k is the kth element of $\boldsymbol{\mu}$. The derivative of $\mathrm{E}[\hat{s}_{\mathrm{f}}]$ can be evaluated in an element wise manner such that

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial \mu_k} = \sum_{m_1, m_2 = 1}^{M} a_{w,m_1} a^*_{w,m_2} \frac{\partial}{\partial \mu_k} \exp(j(\mu_{m_1} - \mu_{m_2})) \prod_{p=1}^{N} \psi_{X_p}(b_{m_1,p} - b_{m_2,p}). \tag{A.20}$$

The non-zero terms of (A.20) occur when $k = m_1$ or $k = m_2$ but not both. Taking this into account, one of the sums can be removed yielding

$$\begin{aligned} \frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial \mu_k} = \quad & j a_{w,k} \exp(j\mu_k) \sum_{\substack{m_2 = 1 \\ m_2 \neq k}}^{M} a^*_{w,m_2} \exp(-j\mu_{m_2}) \prod_{p=1}^{N} \psi_{X_p}(b_{k,p} - b_{m_2,p}) \\ & - j a^*_{w,k} \exp(-j\mu_k) \sum_{\substack{m_1 = 1 \\ m_1 \neq k}}^{M} a_{w,m_1} \exp(j\mu_{m_1}) \prod_{p=1}^{N} \psi_{X_p}(b_{m_1,p} - b_{k,p}) \end{aligned} \tag{A.21}$$

Since this is a difference of conjugates, (A.21) can be written as the imaginary part of either term such that

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial \mu_k} = 2\Im\left\{ a_{w,k}^{*} \exp(-j\mu_k) \sum_{\substack{m=1 \\ m\neq k}}^{M} a_{w,m} \exp(j\mu_m) \prod_{p=1}^{N} \psi_{X_p}(b_{k,p} - b_{m,p}) \right\} \tag{A.22}$$

where $\Im[\cdot]$ extracts the imaginary part. From (A.22) and (A.19) the gradient with respect to the mean value phase vector $\boldsymbol{\mu}$ can be evaluated such that

$$\nabla_{\boldsymbol{\mu}} J_{\mathrm{p}} = \begin{bmatrix} \frac{\partial J_{\mathrm{p}}}{\partial \mu_1} \\ \frac{\partial J_{\mathrm{p}}}{\partial \mu_2} \\ \vdots \\ \frac{\partial J_{\mathrm{p}}}{\partial \mu_N} \end{bmatrix}. \tag{A.23}$$

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 are (A.17) and (A.22) in this case.

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

Multiplying a discrete, uniform random variable with two state (denoted DU2 here) by $b_{m,n} - b_{k,n}$ shifts the position of the deltas in its PDF such that for the distribution used here becomes

$$\text{DU2:} \quad p_X(x) = \frac{1}{2}\delta\left(x + \pi(b_{m,n} - b_{k,n})\right) + \frac{1}{2}\delta\left(x - \pi(b_{m,n} - b_{k,n})\right) \tag{A.24}$$

The characteristic function is then

$$\text{DU2:} \quad \psi_X(b_{m,n} - b_{k,n}) = \cos(\pi(b_{m,n} - b_{k,n})) \tag{A.25}$$

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

$$\text{DU2:} \quad \frac{\partial \psi_X(b_{m,n}-b_{k,n})}{\partial b_{k,n}} = \pi \sin(\pi(b_{m,n}-b_{k,n})) \tag{A.26}$$

Inserting (A.25) and (A.26) into (A.17) realizes

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = 2\Re\left\{ \sum_{\substack{m=1 \\ m\neq k}}^{M} a_{w,m} a_{w,k}^{*} \exp\left(j(\mu_m-\mu_k)\right) \pi \sin(\pi(b_{m,n}-b_{k,n})) \prod_{\substack{p=1 \\ p\neq n}}^{N} \cos(\pi(b_{m,n}-b_{k,n})) \right\} \tag{A.27}$$

Inserting (A.30) into (A.22) yields

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial \mu_k} = 2\Im\left\{ a_{w,k}^{*} \exp(-j\mu_k) \sum_{\substack{m=1 \\ m\neq k}}^{M} a_{w,m} \exp(j\mu_m) \prod_{p=1}^{N} \cos(\pi(b_{m,n}-b_{k,n})) \right\} \tag{A.28}$$

A.1.3.2 Continuous uniform distribution (CU)

Multiplying a continuous uniform distribution (denoted here as CU) with a scaler stretches the distribution. For the $U[-\pi,\pi]$ distribution used here, this realizes

$$\text{CU:} \quad p_X(x) = \frac{1}{2\pi(b_{m,n}-b_{k,n})} \begin{cases} 1 & -\pi(b_{m,n}-b_{k,n}) \leq x \leq \pi(b_{m,n}-b_{k,n}) \\ 0 & \text{otherwise} \end{cases} \tag{A.29}$$

The characteristic function is then

$$\text{CU:} \quad \psi_X(b_{m,n}-b_{k,n}) = \frac{\sin(\pi(b_{m,n}-b_{k,n}))}{\pi(b_{m,n}-b_{k,n})} \tag{A.30}$$

The derivative of (A.35) with respect to $b_{k,n}$ is then

$$\text{CU:} \quad \frac{\partial \psi_X(b_{m,n}-b_{k,n})}{\partial b_{k,n}} = \frac{-\pi^2(b_{m,n}-b_{k,n})\cos(\pi(b_{m,n}-b_{k,n})) + \pi\sin(\pi(b_{m,n}-b_{k,n}))}{\pi^2(b_{m,n}-b_{k,n})^2} \tag{A.31}$$

Inserting (A.30) and (A.31) into (A.17) realizes

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = 2\Re\left\{\sum_{\substack{m=1\\m\neq k}}^{M} a_{w,m}a_{w,k}^{*}\exp\left(j(\mu_m-\mu_k)\right)\times \left(\frac{\sin(\pi(b_{m,n}-b_{k,n}))}{\pi(b_{m,n}-b_{k,n})^2}-\frac{\cos(\pi(b_{m,n}-b_{k,n}))}{b_{m,n}-b_{k,n}}\right)\prod_{\substack{p=1\\p\neq n}}^{N}\frac{\sin(\pi(b_{m,n}-b_{k,n}))}{\pi(b_{m,n}-b_{k,n})}\right\} \tag{A.32}$$

where the derivative has been rearranged to make it more concise. Inserting (A.30) into (A.22) yields

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial \mu_k} = 2\Im\left\{a_{w,k}^{*}\exp(-j\mu_k)\sum_{\substack{m=1\\m\neq k}}^{M} a_{w,m}\exp(j\mu_m)\prod_{p=1}^{N}\frac{\sin(\pi(b_{m,n}-b_{k,n}))}{\pi(b_{m,n}-b_{k,n})}\right\} \tag{A.33}$$

A.1.3.3 Gaussian distribution (G)

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

$$\text{G:}\quad p_X(x) = \frac{1}{(b_{m,n}-b_{k,n})\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x}{b_{m,n}-b_{k,n}}\right)^2} \tag{A.34}$$

The characteristic function is realized by taking the Fourier transform of (A.34) yielding

$$\text{G:}\quad \psi_X(b_{m,n}-b_{k,n}) = e^{-\frac{1}{2}(b_{m,n}-b_{k,n})^2} \tag{A.35}$$

The derivative of A.35 with respect to $b_{k,n}$ is then

$$\text{G:}\quad \frac{\partial \psi_X(b_{m,n}-b_{k,n})}{\partial b_{k,n}} = (b_{m,n}-b_{k,n})e^{-\frac{1}{2}(b_{m,n}-b_{k,n})^2} \tag{A.36}$$

Inserting A.35 and A.36 into A.17 yields

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial b_{k,n}} = 2\Re\left\{\sum_{\substack{m=1\\m\neq k}}^{M} a_{w,m}a_{w,k}^{*}\exp\left(j(\mu_m-\mu_k)\right)(b_{m,n}-b_{k,n})\prod_{p=1}^{N}\exp\left(-1/2(b_{m,n}-b_{k,n})^2x^2\right)\right\} \tag{A.37}$$

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

$$\frac{\partial \mathrm{E}[\hat{s}_{\mathrm{f},w}]}{\partial \mu_k} = 2\Im \left\{ a^*_{w,k} \exp(-j\mu_k) \sum_{\substack{m=1 \\ m\neq k}}^{M} a_{w,m} \exp(j\mu_m) \prod_{p=1}^{N} \exp\left(-1/2(b_{m,n} - b_{k,n})^2 x^2\right) \right\} \tag{A.38}$$

A.2 CW-StoWGe Derivations

Evaluation of the CW-StoWGe EFTE and its gradient requires the calculation of its second and fourth order moments. Section A.2.1 derives these moments, Section A.2.2 uses the second moment to derive the gradient of the EFTE cost function, and Section A.2.3 derives the gradient with respect to selected random variable distributions.

A.2.1 CW-StoWGe Moments

The discrete, CW-StoWGe waveform generating process is defined as

$$s[m] = \exp\left(j \left(\sum_{n=-\infty}^{k} X_n q[m - nT_\mathrm{s}] \right) \right) \quad kT_\mathrm{s} \leq m < (k+1)T_\mathrm{s} \tag{A.39}$$

where positive integer T_s is the inter-variable spacing such that a new X_n contributes to the phase content of (A.39) every T_s samples. The phase function $q[m]$ is the cumulative sum of the frequency shaping filter or frequency pulse defined as

$$q[m] = \sum_{k=0}^{m} g[k]. \tag{A.40}$$

and the frequency shaping filter is time limited such that

$$g[m] = \begin{cases} g_0[m] & m = 0, 1, \cdots, LT_\mathrm{s} - 1 \\ 0 & \text{otherwise} \end{cases} \tag{A.41}$$

where $g[m]$ can take on any real value in the interval $0 \leq n < LT_s$. Accordingly, the positive integer LT_s is the response length of the frequency shaping filter such that each random variable contributes dynamically (changes from sample to sample) to the frequency content of (A.39) for LT_s samples.

The form of (A.39) and the value of k emphasize the causal nature of the CW-StoWGe form. As the time index m increases, more random variables contribute to the phase, but future random variables $(X_n; n > k)$ do not contribute to the phase. However, this is already guaranteed by the form of $g[m]$ (and subsequently $q[m]$) since $g[m < 0] = 0$. With in mind, (A.39) can be rewritten as

$$s[m] = \exp\left(j \sum_{n=-\infty}^{\infty} X_n q[m - nT_s] \right) \tag{A.42}$$

where k has been dropped and replaced with ∞.

Using this form, the correlation between two samples of (A.42) is by definition

$$\mathrm{E}\left[s[m_1] s^*[m_2] \right] = \mathrm{E}\left[\exp\left(j \left(\sum_{n_1=-\infty}^{\infty} X_{n_1} q[m_1 - n_1 T_s] - \sum_{n_2=-\infty}^{\infty} X_{n_2} q[m_2 - n_2 T_s] \right) \right) \right] \tag{A.43}$$

which can be simplified such that

$$\mathrm{E}\left[s[m_1] s^*[m_2] \right] = \mathrm{E}\left[\exp\left(j \sum_{n=-\infty}^{\infty} X_n \Big(q[m_1 - nT_s] - q[m_2 - nT_s] \Big) \right) \right]. \tag{A.44}$$

Since $q[m < 0] = 0$ and $q[m \geq LT_s] = b$ where $b = \sum_{m=0}^{LT_s - 1} g[m]$ is a constant almost every term in the summation is zero meaning only a finite number of random variables contribute to the phase argument of (A.44) for any given m_1 and m_2.

With this in mind, the last (largest or most positive n) random variable to have a non-zero contribution is determined by the larger of either m_1 or m_2. Using m as a stand-in for the larger of the two reference times, the largest relevant value of n can be found by solving

$$m - nT_s < 0 \tag{A.45}$$

for n which results in

$$n > \frac{m}{T_s}. \tag{A.46}$$

Since n must be and integer and new random variables only contribute every T_s samples, n is rounded down such that

$$n_{\max} = \left\lfloor \frac{\max\{m_1, m_2\}}{T_s} \right\rfloor \tag{A.47}$$

where $\max\{\cdot\}$ returns the maximum over its argument, $\lfloor \cdot \rfloor$ is the floor function which returns the greatest integer less than its argument, and $n_{\max}$ represents the largest value of n corresponding to a random variable with a non-zero contribution to the phase.

Likewise, the first (smallest or most negative n) random variable to have a non-zero contribution is determined by the smaller of either m_1 or m_2. Using m as a stand-in for the smaller of the two reference times, the smallest relevant value of n can be found by solving

$$m - nT_s \leq LT_s - 2 \tag{A.48}$$

for n which results in

$$n \geq \frac{m - LT_s + 2}{T_s}. \tag{A.49}$$

Since n must be an integer and new random variables only contribute every T_s samples, n needs to be rounded up such that

$$n_{\min} = \left\lceil \frac{\min\{m_1, m_2\} - LT_s + 2}{T_s} \right\rceil \tag{A.50}$$

where $\min\{\cdot\}$ returns the minimum over its argument, $\lceil \cdot \rceil$ is the ceiling function which returns the smallest integer greater than its, argument and $n_{\min}$ represents the smallest value of n corresponding with a random variable with a non-zero contribution to the phase.

Taken together, these allow (A.44) to be equivalently rewritten as

$$\mathrm{E}\left[s[m_1]s^*[m_2]\right] = \mathrm{E}\left[\exp\left(j \sum_{n=n_{\min}}^{n_{\max}} X_n\Big(q[m_1 - nT_s] - q[m_2 - nT_s]\Big)\right)\right] \tag{A.51}$$

where the derived bounds replace the summation limits. Recall now from Section 3.1 the definition of the characteristic function which is repeated here for convenience

$$\psi_X(\omega) = \mathrm{E}[\exp(j\omega X)] = \int_{-\infty}^{\infty} \exp(j\omega x) f_X(x) dx. \tag{A.52}$$

Additionally, if X_1 and X_2 are independent random variables then

$$\psi_X(\omega) = \mathrm{E}[\exp(j\omega(X_1 + X_2))] = \mathrm{E}[\exp(j\omega X_1)]\mathrm{E}[\exp(j\omega X_2)]. \tag{A.53}$$

Conveniently, since the random variables in (A.51) are by definition independent, (A.53) can be applied to (A.51) to realize

$$\mathrm{E}\left[s[m_1]s^*[m_2]\right] = \prod_{n=n_{\min}}^{n_{\max}} \psi_{X_n}\Big(q[m_1 - nT_\mathrm{s}] - q[m_2 - nT_\mathrm{s}]\Big) \tag{A.54}$$

where the summation has been replaced with the product of the random variables' individual characteristic functions. (A.54) adequately represents the correlation between the two samples at times m_1 and m_2, but what is needed for evaluating the expected spectrum is a function where it does not matter what m_1 and m_2 are in absolute terms. Instead, what matters is the difference between them. Otherwise, the expected spectrum could change as a function of time. In other words, the CW-StoWGe model needs to be stationary, but by examining (A.54) this is not the case. Based on the definitions of $q[m]$, $n_{\min}$, and $n_{\max}$, it is found that in general

$$\mathrm{E}\left[s[m_1]s^*[m_2]\right] \neq \mathrm{E}\left[s[m_1 + k]s^*[m_2 + k]\right] \tag{A.55}$$

where k is an integer. However, if instead of allowing any joint shift in m_1 and m_2 and instead only shifts which are a multiple of T_s are allowed then

$$\mathrm{E}\left[s[m_1]s^*[m_2]\right] = \mathrm{E}\left[s[m_1 + kT_\mathrm{s}]s^*[m_2 + kT_\mathrm{s}]\right] \tag{A.56}$$

which is a property referred to as cyclo-stationarity. The same behavior is observed with CPM [109]. In general, $T_{\rm s}$ will be much smaller than the DFT size ($T_{\rm s} << W$) and also much smaller than a receive data record. With this in mind it is reasonable to consider the average correlation over the interval $T_{\rm s}$ which realizes

$$C[m_1-m_2]=\frac{1}{T_{\rm s}}\sum_{v=0}^{T_{\rm s}-1}\mathrm{E}\left[s[m_1+v]s^*[m_2+v]\right]=\frac{1}{T_{\rm s}}\sum_{v=0}^{T_{\rm s}-1}\prod_{n=n_{\min}}^{n_{\max}}\psi_{X_n}\Big(q[m_1+v-nT_{\rm s}]-q[m_2+v-nT_{\rm s}]\Big) \tag{A.57}$$

which will realize the same result given any joint shift in m_1 and m_2. Thus, m_1 and m_2 can be chosen arbitrarily and (A.57) can be described entirely by the difference between them. To simplify (A.57), set

$$m_1 = v \tag{A.58}$$

and

$$m_2 = v-\ell \tag{A.59}$$

such that

$$C[\ell]=\frac{1}{T_{\rm s}}\sum_{v=0}^{T_{\rm s}-1}\mathrm{E}\left[s[v]s^*[v-\ell]\right]=\frac{1}{T_{\rm s}}\sum_{v=0}^{T_{\rm s}-1}\prod_{n=n_{\min}}^{n_{\max}}\psi_{X_n}\Big(q[v-nT_{\rm s}]-q[v-\ell-nT_{\rm s}]\Big) \tag{A.60}$$

where

$$n_{\max}=\left\lfloor\frac{\max\{v,v-\ell\}}{T_s}\right\rfloor \tag{A.61}$$

and

$$n_{\min}=\left\lceil\frac{\min\{v,v-\ell\}-LT_{\rm s}+2}{T_{\rm s}}\right\rceil. \tag{A.62}$$

and ℓ represents the difference between m_1 and m_2. $C[\ell]$ represents the average correlation given a spacing of ℓ.

It is more useful to define (A.60) in terms $g[m]$ rather than $q[m]$ since $g[m]$ is the actual

function modified to minimize the EFTE. Using (A.41), this realizes

$$C[\ell] = \frac{1}{T_\mathrm{s}} \sum_{v=0}^{T_\mathrm{s}-1} \prod_{n=n_\mathrm{min}}^{n_\mathrm{max}} \psi_{X_n} \left(\sum_{k_1=0}^{v-nT_\mathrm{s}} g[k_1] - \sum_{k_2=0}^{v-\ell-nT_\mathrm{s}} g[k_2] \right) \tag{A.63}$$

In (A.63) for $\ell \geq 0$,the difference of sums results in the canceling of the terms of $g[k]$ from 0 to $v-\ell-nT_\mathrm{s}$. This simplifies the form of (A.63) such that

$$C[\ell] = \frac{1}{T_\mathrm{s}} \sum_{v=0}^{T_\mathrm{s}-1} \prod_{n=n_\mathrm{min}}^{n_\mathrm{max}} \psi_{X_n} \left(\sum_{k=v-\ell-nT_\mathrm{s}+1}^{v-nT_\mathrm{s}} g[k] \right) \qquad \ell \geq 0 \tag{A.64}$$

However, since the CW-StoWGe model is cyclo-stationary and real, $C[\ell] = C[-\ell]$. Consequently, (A.64) can be written in terms of the magnitude of ℓ which realizes

$$C[\ell] = \frac{1}{T_\mathrm{s}} \sum_{v=0}^{T_\mathrm{s}-1} \prod_{n=n_\mathrm{min}}^{0} \psi_{X_n} \left(\sum_{k=v-|\ell|-nT_\mathrm{s}+1}^{v-nT_\mathrm{s}} g[k] \right) \tag{A.65}$$

for all ℓ. Likewise, n_max and n_min are simplified such that

$$n_\mathrm{max} = \left\lfloor \frac{v}{T_s} \right\rfloor = 0 \tag{A.66}$$

and

$$n_\mathrm{min} = \left\lceil \frac{v-|\ell|-LT_\mathrm{s}+2}{T_\mathrm{s}} \right\rceil. \tag{A.67}$$

In (A.65) the argument inside the summation will likely include 0 valued portions of $g[k]$. However, since these terms do not affect the value of (A.65) and since excluding them would significantly complicate the form of (A.65), they are left.

To enable the calculation of the expected spectrum and the EFTE cost function, it is convenient to organize the results of (A.63) into vector and matrix structures. First, a $(2W-1)\times 1$ correlation vector can be constructed such that

$$\mathbf{c} = [\, C[-W+1] \quad C[-W+2] \quad \cdots \quad C[W-2] \quad C[W-1] \,]^T. \tag{A.68}$$

This vector can then be used to create a Hermitian correlation matrix such that.

$$\mathbf{C} = \begin{bmatrix} C[0] & C[-1] & \cdots & C[-W+1] \\ C[1] & C[0] & \cdots & C[-W+2] \\ \vdots & \vdots & \ddots & \vdots \\ C[W-1] & C[W-2] & \cdots & C[0] \end{bmatrix}. \tag{A.69}$$

Like the second moment, the fourth moment can be used to evaluate useful waveform generating function metrics such as the RMS autocorrelation and the power spectral deviation. Fortunately, evaluating the fourth moment is similar to evaluating the second moment only with more terms. Even the cyclo-stationarity property still applies although, this takes it down from four independent variables down to three rather than from two to one as for the second moment. Accordingly, the fourth moment is a function of ℓ_1, ℓ_2, and ℓ_3 such that

$$K[\ell_1,\ell_2,\ell_3] = \frac{1}{T_\mathrm{s}} \sum_{v=0}^{T_\mathrm{s}-1} \left[\prod_{n=n_\mathrm{min}}^{n_\mathrm{max}} \psi_{X_n} \left(\sum_{k_1=0}^{v-nT_\mathrm{s}} g[k_1] - \sum_{k_2=0}^{v-\ell_1-nT_\mathrm{s}} g[k_2] - \sum_{k_3=0}^{v-\ell_2-nT_\mathrm{s}} g[k_3] + \sum_{k_4=0}^{v-\ell_3-nT_\mathrm{s}} g[k_4] \right) \right] \tag{A.70}$$

where

$$n_\mathrm{min} = \left\lceil \frac{\min\{v, v-\ell_1, v-\ell_2, v-\ell_3\} - LT_\mathrm{s} + 2}{T_\mathrm{s}} \right\rceil \tag{A.71}$$

and

$$n_\mathrm{max} = \left\lfloor \frac{\max\{v, v-\ell_1, v-\ell_2, v-\ell_3\}}{T_\mathrm{s}} \right\rfloor. \tag{A.72}$$

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 gradient is taken with respect to the frequency shaping filter instead of the basis function matrix. The CW-StoWGe EFTE cost function is defined as

$$J_\mathrm{CW} = \left|\left| \mathrm{E}[\hat{\mathbf{s}}_\mathrm{f}] - \mathbf{u} \right|\right|_2^2. \tag{A.73}$$

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

$$\frac{\partial J_{\mathrm{CW}}}{\partial g_y} = \frac{\partial}{\partial g_y}\left|\left|\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] - \mathbf{u}\right|\right|_2^2. \tag{A.74}$$

According to the chain rule, (A.74) becomes

$$\frac{\partial J_{\mathrm{CW}}}{\partial g_y} = 2\left(\frac{\partial \mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]}{\partial g_y}\right)^T \left(\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right] - \mathbf{u}\right) \tag{A.75}$$

where the derivative is applied to $\mathrm{E}\left[\hat{\mathbf{s}}_{\mathrm{f}}\right]$ on an element wise basis. For the wth sample of the expected power spectrum, the derivative is

$$\frac{\partial \mathrm{E}\left[\hat{s}_{\mathrm{f},w}\right]}{\partial g_y} = \mathbf{a}_w \frac{\partial \mathbf{C}}{\partial g_y} \mathbf{a}_w^H. \tag{A.76}$$

where $\mathbf{a}_w$ is the wth row of the DFT matrix and the partial derivative is evaluated with respect to each element of $\mathbf{C}$. Since $\mathbf{C}$ is Hermitian, there are only W unique derivatives to evaluate corresponding to $\left[\, C_0 \quad C_1 \quad \cdots \quad C_{W-1} \,\right]^T$. The derivative of the correlation function at lag ℓ is

$$\frac{\partial C[\ell]}{\partial g_y} = \frac{\partial}{g_y}\left(\frac{1}{T_{\mathrm{s}}} \sum_{v=0}^{T_{\mathrm{s}}-1} \prod_{n=n_{\min}}^{0} \psi_{X_n}\left(\sum_{k=v-|\ell|-nT_{\mathrm{s}}+1}^{v-nT_{\mathrm{s}}} g[k]\right)\right) \tag{A.77}$$

The derivative can be moved inside of the summation such that

$$\frac{\partial C[\ell]}{\partial g_y} = \frac{1}{T_{\mathrm{s}}} \sum_{v=0}^{T_{\mathrm{s}}-1} \frac{\partial}{g_y} \prod_{n=n_{\min}}^{0} \psi_{X_n}\left(\sum_{k=v-|\ell|-nT_{\mathrm{s}}+1}^{v-nT_{\mathrm{s}}} g[k]\right). \tag{A.78}$$

Then by using the product rule, (A.78) becomes

$$\frac{\partial C[\ell]}{\partial g_y} = \frac{1}{T_{\mathrm{s}}} \sum_{v=0}^{T_{\mathrm{s}}-1} \sum_{z=n_{\min}}^{0} \left(\frac{\partial \psi_{X_z}\left(\sum_{k=v-|\ell|-zT_{\mathrm{s}}+1}^{v-zT_{\mathrm{s}}} g[k]\right)}{\partial g_y} \prod_{\substack{n=n_{\min} \\ n\neq z}}^{0} \psi_{X_n}\left(\sum_{k=v-|\ell|-nT_{\mathrm{s}}+1}^{v-nT_{\mathrm{s}}} g[k]\right)\right). \tag{A.79}$$

In general, many of the terms inside of the inner summation will actually be zero, since in many cases g_y will not appear within the derivative making the partial derivative zero valued. Unfortunately, accounting for these cases and eliminating them from (A.79) is tedious as it depends on T_s, L, ℓ such that the indexing would become exceptionally onerous and would not really add any clarity to equation. If anything, doing so would make (A.79) much more complicated and less intuitive. These partial derivatives can be collected into the gradient structure $\nabla_{\mathbf{g}} J_{\mathrm{CW}}$ such that

$$\nabla_{\mathbf{g}} J_{\mathrm{CW}} = \begin{bmatrix} \frac{\partial J_{\mathrm{CW}}}{\partial g_1} \\ \frac{\partial J_{\mathrm{CW}}}{\partial g_2} \\ \vdots \\ \frac{\partial J_{\mathrm{CW}}}{\partial g_N} \end{bmatrix}. \tag{A.80}$$

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

Unlike in the pulsed StoWGe case, the derivative term in A.79 does not necessarily contain the variable g_y and so could actually be zero valued. Fortunately, this ambiguity can be addressed using substitution and the chain rule. To do so, define a function

$$f(z,\ell) = \sum_{k=v-|\ell|-zT_{\mathrm{s}}+1}^{v-hT_{\mathrm{s}}} g[k] \tag{A.81}$$

Then the derivative with respect to g_y is

$$\frac{\partial f(z,\ell)}{\partial g_y} = \begin{cases} 1 & (v-|\ell|-zT_{\mathrm{s}}+1) \leq \quad y \quad \leq (v-zT_{\mathrm{s}}) \\ 0 & \text{otherwise} \end{cases} \tag{A.82}$$

Putting (A.81) into (A.79) realizes

$$\frac{\partial C[\ell]}{\partial g_y} = \frac{1}{T_{\mathrm{s}}} \sum_{v=0}^{T_{\mathrm{s}}-1} \sum_{z=n_{\min}}^{0} \left(\frac{\partial \psi_{X_z}(f(z,\ell))}{\partial g_y} \prod_{\substack{n=n_{\min} \\ n\neq z}}^{0} \psi_{X_n}(f(n,\ell)) \right), \tag{A.83}$$

which by the chain rule becomes

$$\frac{\partial C[\ell]}{\partial g_y} = \frac{1}{T_\mathrm{s}} \sum_{\nu=0}^{T_\mathrm{s}-1} \sum_{z=n_\mathrm{min}}^{0} \left(\frac{\partial f(z,\ell)}{\partial g_y} \frac{\partial \psi_{X_z}(f(z,\ell))}{\partial f(z,\ell)} \prod_{\substack{n=n_\mathrm{min} \\ n\neq z}}^{0} \psi_{X_n}(f(n,\ell)) \right). \tag{A.84}$$

Since the partial with respect g_y addresses the zero or non-zero ambiguity, the gradients for each distribution can be evaluated in terms of the place holder function (A.81).

A.2.3.1 Discrete uniform distribution with two states (DU2)

Multiplying a discrete, uniform random variable with two states (denoted DU2 here) by $f(z,\ell)$ shifts the position of the deltas in its PDF such that for the distribution used here

$$\text{DU2:} \quad p_X(x) = \frac{1}{2}\delta(x+\pi f(z,\ell)) + \frac{1}{2}\delta(x-\pi f(z,\ell)) \tag{A.85}$$

The characteristic function is then

$$\text{DU2:} \quad \psi_X(f(z,\ell)) = \cos(\pi f(z,\ell)) \tag{A.86}$$

and its derivative with respect to $f(z,\ell)$ is

$$\text{DU2:} \quad \frac{\partial \psi_X(f(z,\ell))}{\partial f(z,\ell)} = -\pi \sin(\pi(f(z,\ell))) \tag{A.87}$$

Inserting (A.86) and (A.87) into (A.84) realizes

$$\frac{\partial C[\ell]}{\partial g_y} = -\frac{\pi}{T_\mathrm{s}} \sum_{\nu=0}^{T_\mathrm{s}-1} \sum_{z=n_\mathrm{min}}^{0} \left(\frac{\partial f(z,\ell)}{\partial g_y} \sin(\pi(f(z,\ell))) \prod_{\substack{n=n_\mathrm{min} \\ n\neq z}}^{0} \cos(\pi f(z,\ell)) \right). \tag{A.88}$$

A.2.3.2 Continuous uniform distribution (CU)

Multiplying a continuous uniform distribution (denoted here as CU) with a scaler stretches the distribution. For the $U[-\pi,\pi]$ distribution used here, this realizes

$$\text{CU:}\quad p_X(x) = \frac{1}{2\pi(f(z,\ell))}\begin{cases} 1 & -\pi(f(z,\ell)) \le x \le \pi(f(z,\ell)) \\ 0 & \text{otherwise} \end{cases} \tag{A.89}$$

The characteristic function is then

$$\text{CU:}\quad \psi_X(f(z,\ell)) = \frac{\sin(\pi(f(z,\ell)))}{\pi(f(z,\ell))} \tag{A.90}$$

The derivative of (A.90) with respect to $f(z,\ell)$ is then

$$\text{CU:}\quad \frac{\partial \psi_X(f(z,\ell))}{\partial f(z,\ell)} = \frac{\pi^2(f(z,\ell))\cos(\pi(f(z,\ell))) + \pi\sin(\pi(f(z,\ell)))}{\pi^2 f^2(z,\ell)} \tag{A.91}$$

Inserting (A.90) and (A.91) into (A.84) realizes

$$\frac{\partial C[\ell]}{\partial g_y} = \frac{1}{T_\text{s}}\sum_{v=0}^{T_\text{s}-1}\sum_{z=n_\text{min}}^{0}\left(\frac{\partial f(z,\ell)}{\partial g_y}\left(\frac{\sin(\pi f(z,\ell))}{\pi f^2(z,\ell)} + \frac{\cos(\pi f(z,\ell))}{f(z,\ell)}\right)\prod_{\substack{n=n_\text{min}\\ n\neq z}}^{0}\psi_{X_n}\frac{\sin(\pi(f(z,\ell)))}{\pi(f(z,\ell))}\right), \tag{A.92}$$

where the derivative has been rearranged to make it more concise.

A.2.3.3 Gaussian distribution (G)

The zero mean Gaussian distribution with variance $f^2(z,\ell)$ is defined

$$\text{G:}\quad p_X(x) = \frac{1}{f(z,\ell)\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x}{f(z,\ell)}\right)^2} \tag{A.93}$$

The characteristic function is realized by taking the Fourier transform of (A.34) yielding

$$\text{G:}\ \ \psi_X(f(z,\ell)) = e^{-\frac{1}{2}f^2(z,\ell)} \tag{A.94}$$

The derivative of A.35 with respect to $f(z,\ell)$ is then

$$\text{G:}\ \ \frac{\partial \psi_X(f(z,\ell))}{\partial f(z,\ell)} = -f(z,\ell)e^{-\frac{1}{2}f^2(z,\ell)} \tag{A.95}$$

Inserting A.94 and A.95 into (A.84) yields

$$\frac{\partial C[\ell]}{\partial g_y} = -\frac{\pi}{T_\text{s}} \sum_{v=0}^{T_\text{s}-1} \sum_{z=n_\text{min}}^{0} \left(\frac{\partial f(z,\ell)}{\partial g_y} f(z,\ell) e^{-\frac{1}{2}f^2(z,\ell)} \prod_{\substack{n=n_\text{min} \\ n\neq z}}^{0} e^{-\frac{1}{2}f^2(z,\ell)} \right). \tag{A.96}$$

where the $p \neq n$ has been removed from the product operator since it subsumed the exponential component of the derivative term.

Appendix B

Tabulated Optimization Results

B.1 Pulsed StoWGe

Gaussian ($K = 2$)

$P_X(x)$:		DU2		CU		Gaussian	
	$\mathbf{B}_0$:	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$
# of Random Variables (N)	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

Gaussian ($K = 4$)

$P_X(x)$:		DU2		CU		Gaussian	
	$\mathbf{B}_0$:	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$
# 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	-132.7	-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	
	$\mathbf{B}_0$:	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$
# 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	
	$\mathbf{B}_0$:	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$
# of Random Variables (N)	2	-86.1	-85.9	-58.2	-39.8	-41.5	-39.8
	4	-88.1	-88.4	-73.0	-50.1	-41.8	-41.7
	8	-88.4	-88.4	-78.3	-52.1	-41.8	-41.8
	16	-88.1	-88.4	-77.6	-50.7	-42.0	-41.7
	32	-88.1	-88.3	-73.1	-47.1	-42.0	-41.4
	64	-88.8	-88.1	-76.2	-28.2	-42.0	-40.8
	128	-88.8	-79.0	-54.5	-49.3	-42.0	-40.8
	256	-76.3	-88.6	-61.4	-59.1	-42.0	-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	
	$\mathbf{B}_0$:	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PC}}$	$\mathbf{B}_{\text{Id}}$
# of Random Variables (N)	2	-56.9	-56.8	-42.4	-38.5	-37.2	-36.8
	4	-57.0	-57.0	-43.1	-35.3	-37.2	-37.1
	8	-57.1	-57.0	-43.4	-33.0	-37.3	-37.1
	16	-57.0	-57.1	-43.3	-34.4	-37.2	-37.3
	32	-56.9	-56.7	-42.9	-37.5	-37.1	-37.1
	64	-57.0	-57.0	-42.8	-42.4	-36.9	-36.9
	128	-57.1	-56.7	-40.5	-42.5	-36.8	-36.8
	256	-57.0	-54.4	-39.5	-41.7	-36.8	-36.7

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

Rectangular ($K = 4$)

# of Random Variables (N)	$P_X(x)$:	DU2		CU		Gaussian	
	$\mathbf{B}_0$:	$\mathbf{B}_{\text{PCFM}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PCFM}}$	$\mathbf{B}_{\text{Id}}$	$\mathbf{B}_{\text{PCFM}}$	$\mathbf{B}_{\text{Id}}$
	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 which is oversampled by a factor of 4 with respect to its absolute bandwidth ($K = 4$)

Gaussian ($K = 2$)

$p_X(x)$	L	$\mathbf{g}_0$:	RECT				EXP			
		T_s:	2	4	6	8	2	4	6	8
DU2	1		-45.3	-42.5	-48.6	-47.9	-45.3	-42.5	-48.6	-47.9
	2		-48.9	-45.1	-95.2	-71.4	-48.9	-45.1	-95.2	-71.4
	3		-49.9	-70.5	-101.3	-80.7	-78.5	-95.0	-95.8	-97.5
	4		-55.2	-93.3	-99.0	-103.3	-59.3	-97.5	-100.3	-101.3
	5		-62.2	-91.2	-110.1	-107.9	-67.6	-94.3	-99.3	-106.1
	6		-88.3	-97.4	-103.2	-107.7	-95.5	-99.0	-104.4	-108.0
	7		-110.2	-99.4	-107.1	-107.8	-99.6	-99.7	-106.3	-108.8
	8		-93.8	-101.2	-107.0	-109.1	-96.0	-115.4	-106.5	-109.7
CU	1		-53.5	-54.0	-62.3	-61.6	-53.5	-54.0	-62.3	-61.6
	2		-68.1	-84.7	-88.3	-92.6	-68.1	-75.3	-90.1	-84.1
	3		-78.6	-86.8	-89.3	-93.1	-78.5	-86.8	-85.7	-87.4
	4		-78.8	-86.3	-89.5	-96.2	-78.7	-85.3	-85.8	-87.5
	5		-79.1	-86.5	-85.9	-96.4	-78.8	-82.6	-85.8	-87.5
	6		-79.3	-86.5	-85.9	-96.5	-79.0	-86.5	-85.9	-87.5
	7		-79.3	-86.5	-85.9	-88.1	-79.0	-86.5	-85.9	-87.5
	8		-79.3	-86.5	-85.9	-88.1	-79.0	-86.5	-85.9	-87.5
G	1		-54.9	-60.6	-63.5	-68.6	-54.9	-60.6	-63.5	-68.6
	2		-63.8	-66.7	-67.4	-67.7	-63.8	-66.7	-67.4	-78.5
	3		-65.8	-67.4	-78.5	-78.5	-65.8	-67.4	-78.5	-78.5
	4		-78.5	-78.5	-78.5	-78.5	-66.7	-78.5	-78.5	-78.5
	5		-67.1	-78.5	-78.5	-78.5	-67.1	-78.5	-78.5	-78.5
	6		-78.5	-78.5	-78.5	-78.5	-67.4	-78.5	-78.5	-78.5
	7		-78.5	-78.5	-78.5	-78.5	-78.5	-78.5	-78.5	-78.5
	8		-78.5	-78.5	-78.5	-78.5	-78.5	-78.5	-78.5	-78.5

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 ($K = 4$)

$p_X(x)$	L	$\mathbf{g}_0$:	RECT				EXP			
		T_s:	2	4	6	8	2	4	6	8
DU2	1		-57.1	-49.6	-45.1	-52.4	-57.1	-49.6	-54.6	-52.4
	2		-76.5	-51.1	-61.1	-73.6	-76.5	-55.3	-61.1	-73.6
	3		-78.7	-59.5	-75.5	-119.3	-78.7	-84.7	-67.3	-70.9
	4		-80.0	-118.9	-130.3	-127.8	-80.0	-132.1	-128.6	-129.2
	5		-80.3	-130.7	-126.9	-128.7	-80.3	-76.0	-121.9	-130.7
	6		-82.9	-127.5	-128.8	-129.5	-86.2	-124.3	-123.1	-130.3
	7		-82.9	-127.0	-128.0	-127.2	-87.0	-127.3	-128.7	-132.1
	8		-128.4	-123.7	-125.2	-127.1	-130.1	-116.1	-126.9	-129.3
CU	1		-54.1	-53.2	-53.6	-54.1	-54.1	-53.2	-53.6	-54.1
	2		-85.3	-69.7	-69.3	-76.0	-85.3	-69.7	-69.2	-70.7
	3		-87.0	-88.0	-82.4	-107.3	-87.0	-87.3	-82.2	-97.8
	4		-87.7	-95.8	-95.5	-124.7	-87.0	-84.7	-108.4	-89.2
	5		-87.1	-96.6	-96.9	-120.6	-87.1	-91.0	-105.2	-116.4
	6		-87.8	-97.3	-117.0	-121.2	-87.2	-98.8	-112.4	-119.5
	7		-87.8	-98.2	-120.7	-120.5	-94.0	-99.4	-121.9	-119.6
	8		-87.8	-99.3	-123.2	-121.3	-97.4	-102.5	-116.6	-120.4
G	1		-50.5	-56.7	-60.3	-62.8	-50.5	-56.7	-60.3	-62.8
	2		-58.9	-68.8	-75.0	-79.5	-58.9	-68.8	-75.0	-79.5
	3		-64.6	-75.0	-81.4	-85.9	-64.6	-75.0	-81.4	-85.9
	4		-68.9	-79.5	-85.9	-90.3	-68.9	-79.5	-85.9	-90.3
	5		-72.3	-83.0	-89.4	-93.7	-72.3	-83.0	-89.4	-93.7
	6		-75.1	-85.9	-92.1	-96.4	-75.1	-85.9	-92.1	-96.4
	7		-77.5	-88.3	-94.5	-98.6	-77.5	-88.3	-94.5	-98.6
	8		-79.5	-90.3	-96.4	-100.4	-79.5	-90.3	-96.4	-100.4

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 ($K = 2$)

$p_X(x)$	L	$\mathbf{g}_0$:	RECT				EXP			
		T_s:	2	4	6	8	2	4	6	8
DU2	1		-53.1	-41.3	-48.5	-46.1	-53.1	-41.3	-48.5	-46.1
	2		-54.7	-60.8	-55.0	-57.6	-54.7	-60.8	-55.0	-57.6
	3		-55.2	-76.5	-95.9	-96.6	-55.2	-47.6	-101.0	-77.7
	4		-57.3	-95.9	-113.5	-126.5	-57.3	-81.6	-99.3	-126.8
	5		-57.6	-97.3	-122.3	-127.3	-77.7	-94.9	-119.3	-91.0
	6		-59.4	-113.7	-124.6	-125.4	-78.0	-112.9	-115.0	-124.6
	7		-77.6	-109.5	-109.0	-125.9	-78.2	-121.8	-123.4	-125.5
	8		-74.7	-122.0	-116.3	-124.3	-66.1	-110.9	-120.5	-125.4
CU	1		-54.1	-59.8	-60.7	-62.5	-54.1	-59.8	-60.7	-62.5
	2		-59.9	-76.9	-77.3	-80.5	-59.9	-74.6	-72.3	-75.9
	3		-64.5	-79.2	-83.7	-87.9	-61.0	-76.8	-79.1	-86.2
	4		-64.5	-80.7	-86.3	-93.7	-61.0	-80.4	-84.6	-93.2
	5		-64.5	-82.6	-88.5	-94.8	-61.0	-81.8	-85.6	-94.2
	6		-64.5	-83.2	-88.8	-94.3	-61.0	-82.4	-85.8	-94.3
	7		-64.5	-83.5	-89.0	-95.0	-61.0	-82.6	-85.9	-94.4
	8		-64.5	-83.6	-89.0	-95.0	-61.0	-82.7	-85.9	-94.4
G	1		-46.8	-48.8	-49.5	-49.8	-46.8	-48.8	-49.5	-49.8
	2		-49.7	-50.5	-50.7	-50.7	-49.7	-50.5	-50.7	-50.7
	3		-50.3	-50.7	-50.8	-50.8	-50.3	-50.7	-50.8	-50.8
	4		-50.5	-50.7	-50.8	-51.6	-50.5	-50.7	-50.8	-50.8
	5		-50.6	-50.8	-51.5	-51.6	-50.6	-50.8	-50.8	-51.6
	6		-50.7	-50.8	-51.5	-51.6	-50.7	-50.8	-51.5	-51.6
	7		-50.7	-51.5	-51.5	-51.6	-50.7	-51.5	-51.5	-51.6
	8		-50.7	-50.8	-51.5	-51.6	-50.7	-51.5	-51.5	-51.6

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 ($K = 4$)

$p_X(x)$	L	$\mathbf{g}_0$:	RECT				EXP			
		T_s:	2	4	6	8	2	4	6	8
DU2	1		-44.1	-58.6	-44.4	-48.8	-44.1	-58.6	-44.4	-48.8
	2		-49.9	-60.1	-49.6	-56.0	-49.9	-60.1	-49.6	-61.4
	3		-50.1	-60.6	-54.0	-63.8	-50.1	-60.6	-72.4	-63.8
	4		-50.8	-60.8	-72.6	-75.1	-50.8	-60.3	-72.6	-75.1
	5		-50.8	-65.6	-81.1	-89.1	-50.8	-61.0	-81.1	-89.1
	6		-50.8	-61.3	-86.4	-88.7	-50.8	-64.1	-86.4	-88.7
	7		-50.8	-70.3	-86.4	-94.2	-50.8	-72.3	-86.4	-93.8
	8		-50.8	-71.0	-91.8	-94.5	-50.8	-72.3	-91.8	-105.3
CU	1		-43.6	-55.8	-57.1	-58.6	-43.6	-55.8	-57.1	-58.6
	2		-47.7	-59.5	-69.9	-72.7	-47.7	-59.0	-69.9	-72.7
	3		-48.4	-59.8	-71.4	-76.4	-48.4	-59.8	-71.4	-74.6
	4		-48.5	-59.8	-71.4	-77.9	-48.5	-59.8	-71.4	-77.9
	5		-48.5	-59.8	-71.4	-79.5	-48.5	-59.8	-71.4	-79.5
	6		-48.5	-59.8	-71.4	-80.0	-48.5	-59.8	-71.4	-80.0
	7		-48.5	-59.8	-71.4	-80.3	-48.5	-59.8	-71.4	-80.3
	8		-48.5	-59.8	-71.4	-80.4	-48.5	-59.8	-71.4	-80.4
G	1		-42.8	-45.5	-46.0	-46.6	-33.3	-29.6	-31.2	-31.8
	2		-45.5	-47.1	-47.5	-47.7	-33.3	-34.4	-36.5	-35.3
	3		-46.6	-47.5	-47.7	-47.8	-34.6	-40.9	-38.3	-36.2
	4		-47.1	-47.7	-47.8	-47.9	-40.4	-43.2	-39.2	-36.2
	5		-47.3	-47.8	-47.9	-47.9	-42.5	-44.5	-39.2	-36.2
	6		-47.5	-47.8	-47.9	-47.9	-46.4	-44.8	-39.2	-36.2
	7		-47.6	-47.9	-47.9	-47.9	-46.0	-44.8	-39.2	-36.2
	8		-47.7	-47.9	-47.9	-47.9	-45.1	-44.8	-39.2	-36.2

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 ($K = 2$)

$p_X(x)$	L	$\mathbf{g}_0$:	RECT				EXP			
		T_s:	2	4	6	8	2	4	6	8
DU2	1		-47.5	-38.3	-39.7	-44.4	-47.5	-38.3	-44.6	-44.4
	2		-50.7	-45.9	-50.9	-51.6	-50.7	-45.9	-50.9	-51.6
	3		-51.7	-52.3	-49.7	-53.2	-51.7	-47.3	-52.3	-53.7
	4		-52.3	-52.4	-50.7	-53.2	-52.3	-53.2	-52.5	-50.6
	5		-52.4	-52.6	-50.7	-55.9	-52.4	-48.0	-51.2	-54.1
	6		-52.7	-53.6	-53.4	-56.0	-52.7	-48.0	-52.4	-53.5
	7		-52.7	-53.6	-53.7	-54.7	-52.7	-51.8	-53.6	-54.3
	8		-52.9	-53.6	-54.0	-55.4	-52.9	-55.9	-49.8	-51.3
CU	1		-43.9	-45.7	-47.0	-47.9	-43.9	-45.7	-47.0	-47.9
	2		-44.4	-46.4	-47.5	-48.4	-44.4	-46.4	-47.5	-48.4
	3		-44.6	-46.4	-47.6	-48.4	-44.6	-46.4	-47.6	-48.4
	4		-44.7	-46.4	-47.6	-48.4	-44.6	-46.4	-47.6	-48.4
	5		-44.7	-46.4	-47.6	-48.4	-44.6	-46.4	-47.6	-48.4
	6		-44.7	-46.4	-47.6	-48.4	-44.6	-46.4	-47.6	-48.4
	7		-44.7	-46.4	-47.6	-48.4	-44.6	-46.4	-47.6	-48.4
	8		-44.7	-46.4	-47.6	-48.4	-44.6	-46.4	-47.6	-48.4
G	1		-41.3	-42.0	-42.2	-42.3	-41.3	-42.0	-42.2	-42.3
	2		-42.3	-42.5	-42.6	-42.6	-42.3	-42.5	-42.6	-42.6
	3		-42.5	-42.6	-42.6	-42.6	-42.5	-42.6	-42.6	-42.6
	4		-42.5	-42.6	-42.6	-42.6	-42.5	-42.6	-42.6	-42.6
	5		-42.5	-42.6	-42.6	-42.6	-42.5	-42.6	-42.6	-42.6
	6		-42.6	-42.6	-42.6	-42.6	-42.6	-42.6	-42.6	-42.6
	7		-42.6	-42.6	-42.6	-42.6	-42.6	-42.6	-42.6	-42.6
	8		-42.6	-42.6	-42.6	-42.6	-42.6	-42.6	-42.6	-42.6

Table B.11: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum which is oversampled by a factor of 2 with respect to its 3 dB bandwidth ($K = 2$)

Rectangular ($K = 4$)

$p_X(x)$	L	$\mathbf{g}_0$:	RECT				EXP			
		T_s:	2	4	6	8	2	4	6	8
DU2	1		-38.3	-46.2	-40.5	-41.2	-38.3	-46.2	-40.5	-41.2
	2		-40.0	-48.8	-41.6	-39.7	-40.0	-48.8	-41.6	-39.7
	3		-40.1	-49.1	-44.9	-49.0	-40.1	-49.1	-44.9	-44.3
	4		-40.3	-49.7	-44.2	-49.2	-40.3	-49.7	-44.2	-50.2
	5		-40.3	-49.7	-46.6	-50.2	-40.3	-49.7	-46.6	-49.2
	6		-40.3	-50.0	-45.4	-50.3	-40.3	-50.0	-45.4	-45.4
	7		-40.3	-50.0	-45.4	-50.3	-40.3	-50.0	-45.4	-49.4
	8		-40.3	-50.1	-45.6	-50.3	-40.3	-50.1	-45.6	-52.8
CU	1		-38.1	-41.3	-41.8	-43.0	-38.1	-41.3	-41.8	-43.0
	2		-39.5	-41.6	-42.5	-43.3	-39.5	-41.5	-42.5	-43.3
	3		-39.7	-41.6	-42.7	-43.3	-39.7	-41.6	-42.6	-43.3
	4		-39.7	-41.7	-42.7	-43.3	-39.7	-41.7	-42.6	-43.3
	5		-39.7	-41.7	-42.7	-43.3	-39.7	-41.7	-42.6	-43.3
	6		-39.7	-41.7	-42.7	-43.3	-39.7	-41.7	-42.6	-43.3
	7		-39.7	-41.7	-42.7	-43.3	-39.7	-41.7	-42.6	-43.3
	8		-39.7	-41.7	-42.7	-43.3	-39.7	-41.7	-42.6	-43.3
G	1		-37.8	-38.9	-39.0	-39.2	-37.8	-38.9	-39.0	-39.2
	2		-38.8	-39.4	-39.5	-39.5	-38.8	-39.4	-39.5	-39.5
	3		-39.2	-39.5	-39.5	-39.6	-39.2	-39.5	-39.5	-39.6
	4		-39.4	-39.5	-39.6	-39.6	-39.4	-39.5	-39.6	-39.6
	5		-39.4	-39.6	-39.6	-39.6	-39.4	-39.6	-39.6	-39.6
	6		-39.5	-39.6	-39.6	-39.6	-39.5	-39.6	-39.6	-39.6
	7		-39.5	-39.6	-39.6	-39.6	-39.5	-39.6	-39.6	-39.6
	8		-39.5	-39.6	-39.6	-39.6	-39.5	-39.6	-39.6	-39.6

Table B.12: CW-StoWGe EFTE optimized cost function values for various combinations of parameters and initializations for a desired rectangular spectrum which is oversampled by a factor of 2 with respect to its 3 dB 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] S. D. Blunt, J. K. Jakabosky, C. A. Mohr, P. M. McCormick, J. W. Owen, B. Ravenscroft, C. Sahin, G. D. Zook, C. C. Jones, J. G. Metcalf, and T. Higgins, "Principles and applications of random fm radar waveform design," *IEEE Aerospace and Electronic Systems Magazine*, vol. 35, no. 10, pp. 20–28, 2020.

[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] H. Deng and B. Himed, "Interference mitigation processing for spectrum-sharing between radar and wireless communications systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 3, pp. 1911–1919, 2013.

[16] J. C. Pedro and N. B. Carvalho, "Intermodulation distortion in microwave and wireless circuits, artech house," *Inc., Norwood MA*, 2003.

[17] 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.

[18] S. C. Cripps, "RF power amplifiers for wireless communications," 2006.

[19] 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.

[20] 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.

[21] 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.

[22] B. M. Horton, "Noise-modulated distance measuring systems," *Proceedings of the IRE*, vol. 47, pp. 821–828, May 1959.

[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 non-recurrent 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] E. R. Biehl, C. A. Mohr, B. Ravenscroft, and S. D. Blunt, "Assessment of constant envelope ofdm as a class of random fm radar waveforms," in *2020 IEEE Radar Conference (RadarConf20)*, pp. 1–6, 2020.

[31] C. A. Mohr and S. D. Blunt, "Design and generation of stochastically defined, pulsed FM noise waveforms," in *Intl. Radar Conf*, 2019.

[32] M. A. Richards, J. Scheer, W. A. Holm, and W. L. Melvin, *Principles of modern radar*. Citeseer, 2010.

[33] M. I. Skolnik *et al.*, *Introduction to radar systems*, vol. 3. McGraw-hill New York, 1962.

[34] M. I. Skolnik, *Radar handbook*. McGraw-Hill Education, 2008.

[35] J. Eaves and E. Reedy, *Principles of modern radar*. Springer Science & Business Media, 2012.

[36] D. K. Barton, *Radar system analysis and modeling*. Artech House, 2004.

[37] M. A. Richards, *Fundamentals of radar signal processing*. Tata McGraw-Hill Education, 2005.

[38] P. M. Woodward, *Probability and Information Theory, with Applications to Radar: International Series of Monographs on Electronics and Instrumentation*, vol. 3. Elsevier, 1953.

[39] P. Peebles, "Radar principles john wiley & sons," *Inc., Canada*, 1998.

[40] K. S. Shanmugan and A. M. Breipohl, *Random signals, detection, estimation and data analysis*. Wiley, 1988.

[41] 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.

[42] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Prentice Hall Press, 2009.

[43] P. Stoica, R. L. Moses, *et al.*, "Spectral analysis of signals," 2005.

[44] E. Fowle, "The design of FM pulse compression signals," *IEEE Transactions on Information Theory*, vol. 10, pp. 61–67, January 1964.

[45] C. E. Cook, "A class of nonlinear FM pulse compression signals," *Proceedings of the IEEE*, vol. 52, pp. 1369–1371, Nov 1964.

[46] 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.

[47] R. Barker, "Group synchronization of binary digital systems," *Communication Theory, Willis and Jackson (Butterworth's Scientific Publications, 1953)*, 1953.

[48] J. Lindner, "Binary sequences up to length 40 with best possible autocorrelation function," *Electronics Letters*, vol. 11, pp. 507–507, October 1975.

[49] 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.

[50] L. Bomer and M. Antweiler, "Polyphase barker sequences," *Electronics Letters*, vol. 25, pp. 1577–1579, Nov 1989.

[51] R. Frank, "Polyphase codes with good nonperiodic correlation properties," *IEEE Transactions on Information Theory*, vol. 9, pp. 43–45, January 1963.

[52] D. Chu, "Polyphase codes with good periodic correlation properties (corresp.)," *IEEE Transactions on Information Theory*, vol. 18, pp. 531–532, July 1972.

[53] 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.

[54] 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.

[55] S. W. Golomb, "Two-valued sequences with perfect periodic autocorrelation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, pp. 383–386, April 1992.

[56] J. Nocedal and S. J. Wright, *Nonlinear Equations*. Springer, 2006.

[57] 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.

[58] 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.

[59] 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.

[60] 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.

[61] J. B. Anderson, T. Aulin, and C.-E. Sundberg, *Digital phase modulation*. Springer Science & Business Media, 1986.

[62] S. Bluetooth, "Specifications of the bluetooth system, version 5.2.," *IB, 33l7 Specification*, vol. 5, 2019.

[63] T. Aulin and C. Sundberg, "Continuous phase modulation - part I: Full response signaling," *IEEE Transactions on Communications*, vol. 29, pp. 196–209, March 1981.

[64] 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.

[65] 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.

[66] 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.

[67] 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.

[68] 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.

[69] J. Jakabosky, S. D. Blunt, and B. Himed, "Optimization of "over-coded" radar waveforms," in *2014 IEEE Radar Conference*, pp. 1460–1465, May 2014.

[70] 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.

[71] 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.

[72] N. Abramson, "Bandwidth and spectra of phase-and-frequency-modulated waves," *IEEE Transactions on Communications Systems*, vol. 11, pp. 407–414, December 1963.

[73] 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.

[74] 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.

[75] B. Ravenscroft, J. W. Owen, J. Jakabosky, S. D. Blunt, A. F. Martone, and K. D. Sherbondy, "Experimental demonstration and analysis of cognitive spectrum sensing and notching for radar," *IET Radar, Sonar & Navigation*, vol. 12, no. 12, pp. 1466–1475, 2018.

[76] 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.

[77] 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.

[78] 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.

[79] 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.

[80] 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.

[81] 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.

[82] 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.

[83] L. Pralon, G. Beltrao, A. Barreto, and B. Cosenza, "On the analysis of pm/fm noise radar waveforms considering modulating signals with varied stochastic properties," *Sensors*, vol. 21, no. 5, p. 1727, 2021.

[84] W. L. Melvin and J. A. Scheer, *Principles of Modern Radar: Radar Applications, Volume 3*, vol. 2. IET, 2013.

[85] G. A. Fabrizio, *High frequency over-the-horizon radar: fundamental principles, signal processing, and practical applications*. McGraw-Hill Education, 2013.

[86] M. Jankiraman, *FMCW Radar Design*. Artech House, 2018.

[87] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[88] 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.

[89] 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.

[90] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[91] M. B. Priestley, *Spectral analysis and time series*, vol. 1. Academic press London, 1981.

[92] S. L. Marple Jr and W. M. Carey, "Digital spectral analysis with applications," 1989.

[93] W. B. Davenport, W. L. Root, *et al.*, *An introduction to the theory of random signals and noise*, vol. 159. McGraw-Hill New York, 1958.

[94] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2005.

[95] L. H. Koopmans, *The spectral analysis of time series*. Elsevier, 1995.

[96] A. Schuster, "On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena," *Terrestrial Magnetism*, vol. 3, no. 1, pp. 13–41, 1898.

[97] A. Schuster, *The Periodogram of Magnetic Declination as Obtained from the Records of the Greenwich Observatory During the Years 1871-1895*. Cambridge philos. soc., 1900.

[98] R. B. Blackman and J. W. Tukey, "The measurement of power spectra from the point of view of communications engineering," 1959.

[99] P. H. M. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued gaussian random variables," *IEEE Transactions on Automatic Control*, vol. 33, no. 9, pp. 867–870, 1988.

[100] S. M. Kay, *Modern spectral estimation: theory and application*. Pearson Education India, 1988.

[101] B. Porat, *A course in digital signal processing*. Wiley, 1997.

[102] M. S. Bartlett, "Smoothing periodograms from time-series with continuous spectra," *Nature*, vol. 161, no. 4096, pp. 686–687, 1948.

[103] M. S. Bartlett, "Periodogram analysis and continuous spectra," *Biometrika*, vol. 37, no. 1/2, pp. 1–16, 1950.

[104] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[105] C. A. Mohr, P. M. McCormick, C. A. Topliff, S. D. Blunt, and J. M. Baden, "Gradient-based optimization of PCFM waveforms," *IEEE Transactions on Aerospace and Electronic Systems*, 2021.

[106] 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.

[107] M. Golay, "Complementary series," *IRE Transactions on Information Theory*, vol. 7, pp. 82–87, April 1961.

[108] C. Tseng and C. Liu, "Complementary sets of sequences," *IEEE Transactions on Information Theory*, vol. 18, pp. 644–652, Sep. 1972.

[109] J. G. Proakis and M. Salehi, *Digital communications*, vol. 4. McGraw-hill New York, 2001.