Gradient-Based Optimization of PCFM Radar Waveforms

Charles A. Mohr, *Graduate Student Member, IEEE*, Patrick M. McCormick, *Member, IEEE*, Charles A. Topliff, *Graduate Student Member, IEEE*, Shannon D. Blunt, *Fellow, IEEE*, J. Michael Baden, *Senior Member, IEEE*

***Abstract*—While a number of signal structures have been proposed for radar, frequency modulation (FM) remains the most common in practice because it is well-suited to high-power transmitters, which tend to introduce significant distortion to other waveform classes. That said, various forms of coding provide useful parameterizations for which a variety of optimization methods can be readily applied to accomplish different operational goals. To that end, the polyphase-coded FM (PCFM) implementation was previously devised as a means to bridge this gap between optimizable parameters and physically realizable waveforms.**

However, the original method employed to optimize PCFM waveforms involved a piece-wise greedy search that, while relatively effective, was rather slow and cumbersome. Here the continuous nature of this framework is leveraged to formulate a gradient-based optimization approach that updates all parameters simultaneously and can be efficiently performed using fast Fourier transforms (FFTs), thus facilitating a general design methodology for practical waveforms that is directly extensible to myriad waveform-diverse arrangements. Results include a large number of optimization assessments to discern performance trends in aggregate and detailed analysis of specific cases, as well as both loopback and free-space experimental measurements to demonstrate practical efficacy.

***Index Terms*—frequency modulation, radar waveform design, gradient-based optimization, waveform diversity**

This work was supported in part 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. DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

C.A. Mohr is with the Radar Systems & Remote Sensing Lab (RSL), University of Kansas, Lawrence, KS, 66045 USA. He is also now with the Air Force Research Laboratory – Sensors Directorate, Wright-Patterson AFB, OH 45433 USA (e-mail: charles.mohr.1@us.af.mil).

P.M. McCormick was with the Radar Systems & Remote Sensing Lab (RSL), University of Kansas, Lawrence, KS, 66045 USA. He is now with the Air Force Research Laboratory – Sensors Directorate, Wright-Patterson AFB, OH 45433 USA (e-mail: patrick.mccormick.9@us. af.mil).

C.A. Topliff was with the Radar Systems & Remote Sensing Lab (RSL), University of Kansas, Lawrence, KS, 66045 USA. He is now with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30313 USA (e-mail: ctopliff0@gatech.edu).

S.D. Blunt is with the Radar Systems & Remote Sensing Lab (RSL) and Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045 USA (e-mail: sdblunt@ittc.ku.edu).

J.M. Baden is with the Sensors and Electromagnetic Applications Laboratory, Georgia Tech Research Institute, Smyrna, GA 30080 USA (e-mail: mike.baden@gtri.gatech.edu).

I. INTRODUCTION

The litany of radar waveform classes, signal structures, and design methodologies continues to grow (e.g. see [1-5] and references therein), and yet the very first waveform class based on frequency modulation (FM) [6-9] continues to be the workhorse of a majority of radar systems fielded today. Indeed, the simple and well-known linear FM (LFM) chirp is the standard performance benchmark and is still widely used due to its inherent Doppler tolerance and ease of implementation, which permits the generation of extremely wide bandwidths and the use of stretch processing [10, 11] on receive. Moreover, its primary limitation of high sidelobes can be largely addressed using receive tapering to shape the spectrum [1, 5], assuming the associated trade-offs in signal-to-noise ratio (SNR) loss and range resolution degradation are acceptable.

This spectral shaping concept has also been examined through various forms of nonlinear FM (NLFM) [1] that avoid the SNR loss, where the primary design goal is to determine a monotonic frequency function of time that is antisymmetric about its midpoint (see [12] for a recent summary and analysis). It was also recently shown that a compensated form of stretch processing can likewise be applied to these forms of NLFM [13]. Thus, in general, FM waveforms remain attractive because they provide relatively good spectral containment and possess a constant amplitude, both of which are attributes necessary for high-power transmitters.

For more benign environments where the spectrum is relatively quiet and operational time-scales are relatively long (e.g. scientific remote sensing applications), the LFM chirp will likely remain in widespread use for some time to come (e.g. [14, 15]). However, most other radar applications continue to experience a dynamic "complexification" of the radio frequency (RF) spectrum while concurrently having a conflicting need to achieve ever greater sensitivity, maneuverability, and adaptability in real-time. Consequently, there has been a tremendous amount of research under the heading of "waveform diversity" (e.g. [2-5]) to identify signals and design methodologies that provide further sophistication and flexibility.

At the hardware level this demand is being addressed by the enabling technologies of high-fidelity arbitrary waveform generation (AWG) capabilities [16], as well as emerging software defined radio/radar (SDR) systems [17] and RF system-on-a-chip (RF-SoC) modules [18]. However, realizing waveforms that achieve the above goals while likewise being suitable for these waveform generation systems, along with the more strenuous high-power transmitter effects that follow, remains a topic of ongoing investigation.

To that end the polyphase-coded FM (PCFM) waveform implementation [19, 20], a conceptual off-shoot of the digital FM communication structure known as continuous phase modulation (CPM) [21], was proposed as a means to make the linkage between goal-oriented parameter optimization and physically-realizable radar waveforms that are transmitter-amenable [22, 23]. It was subsequently demonstrated experimentally [23, 24] that transmitter distortion effects could be incorporated into the waveform design process, thereby further facilitating the prospect of joint transmitter/waveform optimization (e.g. [25-27]), a consequence of which could potentially be the effective mimicry (and the associated advanced sensing/discrimination proficiency) of biosonar capabilities found in nature [28-30].

The PCFM framework has likewise been recently leveraged to realize complementary FM waveforms [31], different forms of radar/communication spectrum sharing [32, 33], and nonlinear intermodulation radar [34]. Owing to their physically-realizable construction as FM waveforms, all of these applications of PCFM have been experimentally demonstrated in hardware as well.

While some amount of structure is needed to ensure that a given signal conforms to the FM form, the PCFM framework can arguably be viewed as a "maximally parameterized" structure relative to other NLFM arrangements. The justification for this statement relies on the fact that PCFM can employ multiple "orders" [35] as well as "over-coding" [36] that greatly exceeds the waveform time-bandwidth product (BT), where B is bandwidth (usually 3-dB) and T is pulsewidth. Moreover, PCFM is not constrained to frequency functions that are monotonic or symmetric (though the latter is found to still be roughly approximated for individual optimized FM waveforms that achieve a particular performance bound).

The piece-wise, greedy search over the PCFM parameter space described in [22-24] has been shown to yield effective results. However, it is also rather slow, thereby hindering the optimization of waveforms with high BT or applications requiring real-time waveform design/modification. Moreover, it is not readily extensible to multi-waveform modalities. To address these limitations, here we exploit the continuous nature of the PCFM waveform design problem to formulate a gradient-based optimization approach. Gradient methods have been used before to design waveforms belonging to the class of discrete sequences (e.g. [37-44]). The key difference here is that we are considering parameterized FM waveforms, whose constant amplitude and continuous structure are readily amenable to high-power transmitters. Consequently, experimental measurements will also be used to demonstrate the resulting optimized FM waveforms.

A particular benefit of gradient-based FM waveform optimization, relative to the greedy search of [22-24], is that all parameters are updated in parallel at each iteration. Thus this formulation is naturally extensible to the joint design of multiple waveforms, such as arises for waveform-diverse arrangements like multiple-input multiple-output (MIMO) or pulse agility. Moreover, as posited in [38, 43, 44], the gradient of some cost functions permits simplified computation through the use of FFTs. Consequently, optimization of these FM waveforms can be performed quite efficiently, thus enabling application to sensing modes involving large BT or the need for real-time, on-the-fly design.

Note that some applications of gradient-based FM waveform optimization were recently presented by the authors in [31, 33, 34, 45-47] for the particular class of random FM waveforms, which involves the generation of a stream of unique, nonrepeating waveforms (see [48] for a survey). In contrast, here we expand upon the approach of [49] by examining the underlying mathematical structure of this general design approach for individual waveforms (i.e. <u>not</u> part of a nonrepeating set) and assess the achievable performance when operating in the traditional mode of repeating the same waveform throughout the coherent processing interval (CPI). In other words, we are asking the question: "for a given initialization and (possibly quite large) set of parameters, what is the single best FM waveform that can be designed?" Of course, global optimality cannot be guaranteed due to the size of the solution space and its nonconvex nature. However, there does exist a judicious initialization selection that provides rather good results after the application of gradient descent, in many cases attaining a performance bound that exists when the continuous waveform is discretized for optimization and/or digital receive filtering.

The rest of the paper is organized as follows. Section II introduces and discusses the salient aspects of the PCFM implementation used for optimization of physically realizable waveforms, followed by a description of the Generalized Integrated Sidelobe Level (GISL) metric in Sect. III and its efficient gradient-descent solution in Sect. IV. In Sect. V the results of several optimization runs are examined to assess the complicated interplay between various parameters in the PCFM waveform implementation. Finally, loopback and open-air experimental measurements using selected optimized FM waveforms are presented in Sect. VI to demonstrate practical efficacy.

II. PCFM Radar Waveform Implementation

The PCFM signal structure [20] (Fig. 1) was adapted from the CPM communication framework that is used for aeronautical telemetry, deep-space communications, and is the basis for the Bluetooth™ wireless standard [21]. Being a form of digital FM, CPM (and thus also PCFM) has the desirable attributes of being both continuous and constant amplitude. The former facilitates good spectral containment (given appropriate coding) and the combination of both mitigates most of the distortion that would otherwise be imparted by a high-power transmitter (particularly the high-power amplifier (HPA) [50]).

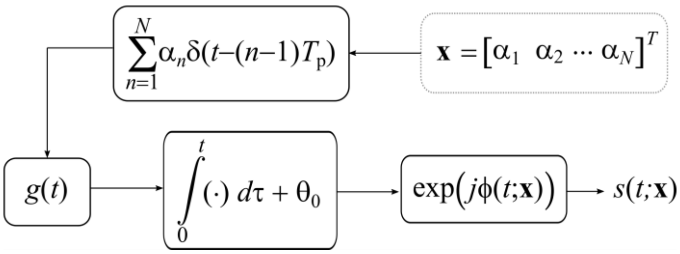

Fig. 1. Polyphase-coded FM (PCFM) radar waveform implementation [20]

In brief, the PCFM waveform implementation takes a vector of N parameters $\mathbf{x} = [\alpha_1\ \alpha_2 \cdots \alpha_N]^T$, for $(\bullet)^T$ the vector transpose operation, and uses them to weight a train of N impulses separated in time by T_p. This weighted impulse train is convolved with (frequency) shaping filter $g(t)$, which has time support on $[0,\ T_p]$, and is then integrated to obtain the continuous phase function of time $\phi(t\,;\mathbf{x})$. This phase function is subsequently exponentiated to produce the complex, baseband waveform of pulsewidth $T = NT_p$. The process in Fig. 1 can be written succinctly as [20]

$$\begin{aligned} s(t\,;\mathbf{x}) &= \exp\left\{ j\left(\int_0^t g(\tau) * \left[\sum_{n=1}^{N} \alpha_n \delta(t-(n-1)T_p) \right] d\tau \right) \right\} \\ &= \exp(j\phi(t\,;\mathbf{x})) \end{aligned} \quad (1)$$

where $*$ denotes convolution. For this "first-order" PCFM implementation (since there is a single integration stage) the values in $\mathbf{x}$ can be viewed as piece-wise, instantaneous digital frequencies that lie within $[-\pi, +\pi]$, though the notion of "over-coding" [36] provides a way to expand those limits for greater design freedom. While beyond the scope of this paper, "second-order" and higher PCFM implementations have also been examined [35] in which the coding is in terms of instantaneous digital chirp-rate (or higher). It remains to be seen how gradient-based optimization can likewise be applied to these higher-order implementations. For the remainder of the paper we shall only consider the first-order case. The reader should note, however, that Matlab™ code for the various PCFM implementation orders can be found in the appendix of [35].

Besides the power efficiency (constant amplitude), spectral containment (continuous phase), and transmitter-amenable implementation advantages of PCFM, the discrete set of underlying parameters also provides a convenient mechanism with which to perform waveform optimization, such as was first explored in [22-24]. More recent work by the authors [31, 34, 35, 45] takes advantage of the linearity of the phase construction in (1). By evaluating the convolution and integral, the phase component of (1) can be equivalently expressed as

$$\phi(t\,;\mathbf{x}) = \sum_{n=1}^{N} \alpha_n\, b_n(t)\,, \quad (2)$$

where each basis function

$$b_n(t) = \int_0^t g(\tau-(n-1)T_p)\, d\tau \quad (3)$$

is the integral of the shaping filter and delay-shifted by an integer multiple of T_p. For example, for first-order PCFM $g(t)$ is a rectangular shaping filter with amplitude $1/T_p$, and thus the nth basis function is the delay-shifted ramp function

$$b_n(t) = \begin{cases} 0, & 0 \le t \le (n-1)T_p \\ (t-(n-1)T_p)/T_p, & (n-1)T_p \le t \le nT_p \\ 1, & nT_p \le t \le NT_p\,. \end{cases} \quad (4)$$

The PCFM phase is therefore constructed from a linear combination of these N continuous basis functions, with each weighted by the corresponding PCFM parameter α_n.

Additionally, (2) can be viewed in a more general sense where the continuous phase function of all manner of coded FM waveforms is comprised of a weighted sum of continuous basis functions. One such example is the construction in [51], where orthogonal basis functions derived from Legendre polynomials were used to design FM waveforms. Moreover, this perspective is a variant of that previously taken by the seminal work of Wilcox [**52-54**] in which a formulation similar to (2) was used to represent the entire waveform $s(t)$ through a combination of orthonormal basis functions. In contrast, the construction in (2) is specific to the waveform's continuous phase component $\phi(t)$ and the constituent basis functions thereof are continuous but otherwise arbitrary. While the Legendre polynomial approach of [**51**] does rely on orthogonality of phase basis functions for compactness, the PCFM framework in (4) using a rectangular shaping filter clearly does not.

For the purposes of this paper, the representation in (2) is also important because it facilitates computationally efficient gradient-descent optimization. By casting the continuous PCFM phase in this way, it can easily be discretized by sampling the N continuous-time basis functions in (3). From there, the discretized PCFM form can be evaluated via a single matrix/vector multiply. Thus the length-M ($> N$) discretized PCFM waveform can be expressed as

$$\mathbf{s} = \exp(j\mathbf{B}\mathbf{x})\,, \quad (5)$$

where $\mathbf{B}$ is an $M \times N$ matrix consisting of the length-M sampled versions of the N basis functions. Given that a time-limited pulse cannot be bandlimited, thereby preventing true Nyquist sampling, the appropriate value for M bears closer examination. Further, while we are focusing on optimization of the PCFM implementation, note that the gradient-descent approach developed here is likewise applicable to any coded FM waveform that can be parameterized in the form of (2) and discretized with sufficient fidelity via (5).

In the original PCFM formulation [20], M took on a rigid definition with respect to the number of PCFM parameters N and the "oversampling" factor K relative to 3-dB bandwidth B, where $M = KN$ such that K was also the number of samples per T_p interval. When the PCFM parameters in $\mathbf{x}$ are allowed to span the digital frequency space of $[-\pi, +\pi]$, then $(1/T_\text{p}) \approx B$ and subsequently $BT \approx (1/T_\text{p})NT_\text{p} = N$. Consequently, the number of waveform samples can also be determined by setting $M = K(BT)$. The authors have observed that selecting K as low as 2 or 3 is generally sufficient for the discretized representation to possess a high enough dimensionality to adequately limit the amount of unavoidable aliasing (though in general the necessary value for K depends on the rate of spectral roll-off for a given waveform).

If much higher fidelity (i.e. much less aliasing) is required, such as in the case of transmit spectral notching [**55**], then a higher value of K may be necessary. Alternatively, discretization (for the purpose of optimization) could be performed in the frequency domain based on the analytical spectral representation of PCFM derived in [**33**] (further investigation of aliasing can be found in this reference).

In [36] it was shown that even better waveforms (based on sidelobe metrics) can be obtained by subdividing each T_p interval by a factor L, thereby increasing the number of PCFM parameters N and consequently the available degrees of freedom. Crucially though, while the number of PCFM parameters is increased, the relationship $M = K(BT)$ is maintained to preserve the same general spectral content. This subdivision of the code intervals, denoted as "over-coding" [36], therefore partially decouples the number of PCFM parameters N from BT, such that $BT \le N \le M$ and we can now set $N = ML/K = L(BT)$. Consequently, K is no longer the number of samples per T_p interval.

Of course, decoupling N from BT introduces a requirement to in some way constrain the PCFM values in $\mathbf{x}$, or to at least pay special attention to how $\mathbf{x}$ is optimized, such that the relationship $M = K(BT)$ is maintained. In [36], this requirement was met explicitly by limiting the span of each (over-coded) α_n to $[-\pi/L, +\pi/L]$. Alternatively, in [45] the optimization metric implicitly constrained the span of each α_n by shaping the waveform spectrum directly. This latter approach worked well enough that "maximal over-coding" (i.e. setting $M = N$, or $L = K$) could be used to exploit the maximum available degrees of freedom for the given discretized representation, while still preserving sufficient spectral containment. Table I provides a summary of these various PCFM terms and their relationships.

Finally, here we shall intentionally relax strict control on spectral containment and rely instead upon 1) setting the autocorrelation mainlobe width to effectively dictate the optimized waveform's 3-dB bandwidth and 2) the choice of initialization prior to optimization. Clearly the former will only impact the passband of the waveform and have little/no influence on spectral roll-off. The latter will only affect spectral containment insofar that the highly non-convex optimization problem <u>may</u> converge to local minima whose spectral content is similar to the initialization. Specifically, it has been observed through multiple trials that initializations exhibiting "chirping" behavior similar to LFM or many of the good NLFM waveforms do tend to retain much of their good spectral containment after optimization. This result is a consequence of the "conservation of ambiguity" phenomenon as it is instantiated in the prominent delay-Doppler sheared ridge for chirped waveforms. In contrast, random initializations generally do not enjoy this natural absorber of ambiguity and thus exhibit quite different optimization behavior (and would thereby otherwise require more explicit spectral control, e.g. [45]).

TABLE I
SUMMARY OF PCFM TERMS

Term	Description	Relationships
M	Number of PCFM waveform samples	$M = \dfrac{NK}{L} = K(BT)$
N	Number of optimizable PCFM parameters	$N = \dfrac{ML}{K} = L(BT)$
K	PCFM oversampling factor with respect to 3-dB bandwidth	$K = \dfrac{ML}{N} = \dfrac{M}{(BT)}$
L	PCFM over-coding factor	$L = \dfrac{NK}{M} = \dfrac{N}{(BT)}$

III. GENERALIZED INTEGRATED SIDELOBE LEVEL

Fast-time Doppler effects notwithstanding, metrics involving evaluation of the waveform autocorrelation

$$r(\tau) = \int_{-\infty}^{\infty} s(t)\, s^*(t+\tau)\, dt\,, \tag{6}$$

where $(\bullet)^*$ denotes complex conjugation, are the litmus tests by which waveform goodness is generally assessed. The two most well-known metrics are the integrated sidelobe level (ISL) ratio, which compares the total energy of the autocorrelation sidelobe region to the total energy in the mainlobe region, and the peak sidelobe level (PSL) ratio, which compares the peak value of the largest autocorrelation sidelobe to the value at the peak of the mainlobe. Mathematically, these metrics can be expressed as

$$\text{ISL} = \frac{\int_{\Delta t}^{T} |\, r(\tau)\,|^2 \; d\tau}{\int_{0}^{\Delta t} |\, r(\tau)\,|^2 \; d\tau} \tag{7}$$

and

$$\text{PSL} = \left(\frac{\max\{|\, r(\tau)\,|^2\}_{\Delta t}^{T}}{\max\{|\, r(\tau)\,|^2\}_{0}^{\Delta t}} \right) = \left(\frac{\max\{|\, r(\tau)\,|^2\}_{\Delta t}^{T}}{|\, r(0)\,|^2} \right), \tag{8}$$

where the mainlobe extends over $-\Delta t < 0 < +\Delta t$ and making use of $r(-\tau) = r^*(\tau)$. The latter form in (8) is the version of PSL most commonly used where, with proper normalization, the denominator can be set to unity and thereby removed altogether.

Lower values of ISL and PSL generally correspond to better radar waveforms. By lowering ISL the average energy in the sidelobes is diminished, thus reducing the overall sidelobe response of the waveform. In contrast, lowering PSL tends to have the effect of flattening the sidelobes, thus helping to reducing the probability of a false alarm due to a sidelobe, though the average sidelobe level in this case may be higher than for an ISL-minimized waveform.

These two metrics assess the problem of sidelobe minimization in different ways, yet whether one is better than the other depends on the application (and transmitter effects as demonstrated in Sect. VI). The question remains, however, of how best to reduce ISL and/or PSL values through waveform optimization. To answer this question it is helpful to examine (7) and (8) in further detail.

Generally speaking, it is easier to optimize (minimize in this case) "well-behaved" cost functions; i.e. those having attributes of convexity, linearity, and continuity. Neither (7) nor (8) have all of these properties, with (8) being especially problematic due to the $\max\{\cdot\}$ operator that prevents differentiation in a conventional manner. Consequently, we can rewrite (8) as

$$\mathrm{PSL} = \lim_{p \to \infty} \left(\frac{\int_{\Delta t}^{T} |r(\tau)|^p \, d\tau}{\int_0^{\Delta t} |r(\tau)|^p \, d\tau} \right)^{2/p}, \tag{9}$$

which takes advantage of the infinity norm that is equivalent to evaluating the $\max\{\cdot\}$ operator. Consequently, (7) and (9) can (effectively) be subsumed into [49]

$$\mathrm{GISL} = \left(\frac{\int_{\Delta t}^{T} |r(\tau)|^p \, d\tau}{\int_0^{\Delta t} |r(\tau)|^p \, d\tau} \right)^{2/p} \tag{10}$$

that realizes the generalized integrated sidelobe level (GISL) ratio for $2 \leq p < \infty$. Specifically, when setting $p = 2$ the GISL metric of (10) becomes ISL from (7). Likewise, if $p \rightarrow \infty$, (10) approaches the PSL metric of (8) or (9). Similar metrics have likewise been considered for discrete sequence design (see [41, 42, **56**, **57**]).

Values in between the extremes of $p = 2$ and ∞ provide a measure in between ISL and PSL. Further, while the GISL metric does not actually permit optimization according to PSL ($p = \infty$) in a strict sense, large (yet finite) values of p serve this intent for all practical purposes. In fact, anecdotal results thus far seem to indicate that there is little discernible benefit to setting p greater than 15 to 20.

Note that the autocorrelation $r(\tau)$ is already a second-order function of the waveform $s(t)$. The exponent p in (10) therefore realizes an optimization metric that is a $2p$-order function of $s(t)$. Given that, at the very least (for ISL), this metric could be a fourth-order function or higher, it is clearly nonlinear and non-convex with respect to waveform optimization. Hence the local minima achieved will depend on the choice of initialization. Fortunately, it is already known that (generally speaking, at least) good waveforms tend to possess a "sideways-S" [1, 5] frequency function of time, and thus classical chirped waveforms provide a good source of initialization. Moreover, since the PCFM representation in (1) is a continuous function of the parameters in $\mathbf{x}$, the subsequent autocorrelation via (6) and GISL evaluation via (10) likewise remain continuous, thus permitting the use of gradient-based optimization that can be performed efficiently due to the discretized form of (2) and (5), and FFT operations that arise from the gradient derivation.

IV. GISL Optimization of PCFM

As noted above, the GISL cost function is nonlinear and non-convex, which precludes global optimality, particularly as the number of optimizable parameters N grows large. Where for many types of problems this situation would subsequently necessitate the investigation of approaches/modifications whereby convexity could be imposed [**58, 59**], a useful attribute of waveform diversity [2-5] is in fact the wide assortment of options that could be employed as a consequence of all these local minima. Thus, while we are presently seeking to determine the single best FM waveform that can be achieved (for a given initialization and set of parameters), it is beneficial for the methodology to be extensible to the generation of more diverse waveform sets as well (e.g. see [48]).

In [31, 33, 34, 37-47, 49, 51] a variety of gradient-descent methods were used to optimize different types of radar signals. Specifically, in [49] the authors applied nonlinear conjugate gradient (NLCG) methods, of which numerous variants exist [60], to the particular problem of parameterized FM waveform optimization. The simpler, yet quite effective, "heavy ball" gradient descent technique was subsequently employed for many of these waveform design applications [31, 33, 34, 45-47]. It is this latter formulation that shall be used here as well.

A. *Discretizing the GISL Cost Function*

To numerically optimize the GISL cost function with respect to the parameter vector $\mathbf{x}$ for a particular integer value of p, it is first necessary to discretize (10), and therefore (6), in a manner consistent with (5) as discussed in Sect. II. In so doing the continuous autocorrelation of (6) becomes the discretized version

$$r[\ell] = \sum_{m=1}^{M} s[m]\, s^*[m+\ell], \tag{11}$$

with $s[m] = 0$ for $m \leq 0$ and $m > M$, and retaining the relationship $r[-\ell] = r^*[\ell]$ for delay index ℓ. Taking advantage of the Fourier relationship between the autocorrelation and spectral density, the vectorized form of (11) can be expressed as

$$\mathbf{r} = \mathbf{A}^H [(\mathbf{A}\bar{\mathbf{s}}) \odot (\mathbf{A}\bar{\mathbf{s}})^*], \tag{12}$$

where the zero-padded version of the length-M discretized waveform

$$\bar{\mathbf{s}} = [\mathbf{s}^T \; \mathbf{0}_{1\times(M-1)}]^T \tag{13}$$

has the same $(2M-1)$ length as the corresponding autocorrelation. Further, the $(2M-1) \times (2M-1)$ matrices $\mathbf{A}$ and $\mathbf{A}^H$ perform the discrete Fourier transform (DFT) and inverse DFT (IDFT), respectively, and $\odot$ is the Hadamard product.

Using (12), the discretized GISL cost function can then be written as

$$J_p = \frac{\|\mathbf{w}_{\mathrm{SL}} \odot \mathbf{r}\|_p^2}{\|\mathbf{w}_{\mathrm{ML}} \odot \mathbf{r}\|_p^2} \tag{14}$$

where $\|\cdot\|_p$ is the discrete p-norm and $\mathbf{w}_{\mathrm{SL}}$ and $\mathbf{w}_{\mathrm{ML}}$ are length $(2M-1)$ vectors comprised of zeros and ones that select the sidelobe and mainlobe regions of $\mathbf{r}$, respectively. The mainlobe portion of $\mathbf{r}$, as extracted by the non-zero central portion of $\mathbf{w}_{\mathrm{ML}}$, is set explicitly by the 3-dB bandwidth oversampling factor K, such that the central $(2K-1)$ samples of $\mathbf{r}$ are taken as the mainlobe. It was shown in [49, 51] that this approach is an effective way to control the desired bandwidth during optimization, though it does not regulate the degree of containment in the spectral roll-off region.

Finally, discretization of the autocorrelation imposes a lower bound on the achievable PSL value for constant amplitude waveforms. Specifically, at the outer delay indices of $\ell = \pm(M-1)$, the corresponding $r[\ell]$ in (11) consists of a single term, meaning the result can only be as small as the magnitude of this value. If the discretized waveform is normalized such that $r[0] = 1$, then the PSL lower bound is [49]

$$\mathrm{PSL}_{\text{sampled bound}} = 20\log_{10}\left(\frac{1}{M}\right) = -20\log_{10}(M) \qquad (15)$$

since the $\ell = \pm(M-1)$ terms in $r[\ell]$ will therefore be equal to $1/M$. This bound serves as a useful benchmark by assessing how close a given discretized waveform is to reaching this PSL value.

B. GISL Gradient Descent

For the discretized representation of the GISL cost function in (14) we now wish to compute the gradient with respect to the elements of $\mathbf{x}$, the operator for which is the length-N vector

$$\nabla_{\mathbf{x}} = \left[\frac{\partial}{\partial\alpha_1}\ \frac{\partial}{\partial\alpha_2}\ \cdots\ \frac{\partial}{\partial\alpha_N}\right]^T. \qquad (16)$$

When (16) is applied to (14), it is shown in Appendix A that the gradient ultimately becomes

$$\nabla_{\mathbf{x}} J_p = 4 J_p\, \bar{\mathbf{B}}^T \times \Im\left\{\bar{\mathbf{s}}^* \odot \left(\mathbf{A}^H\left[\mathbf{A}\left(\left[\frac{\mathbf{w}_{\mathrm{SL}}}{\mathbf{w}_{\mathrm{SL}}^T|\mathbf{r}|^p} - \frac{\mathbf{w}_{\mathrm{ML}}}{\mathbf{w}_{\mathrm{ML}}^T|\mathbf{r}|^p}\right]\odot|\mathbf{r}|^{(p-2)}\odot\mathbf{r}\right)\odot(\mathbf{A}\bar{\mathbf{s}})\right]\right)\right\} \qquad (17)$$

where we have expanded

$$\bar{\mathbf{B}} = \left[\mathbf{B}^T\ \mathbf{0}_{N\times(M-1)}\right]^T, \qquad (18)$$

the operator $\Im\{\cdot\}$ extracts the imaginary part of the argument, and $|\bullet|$ realizes the magnitude of each element in the vector argument. The matrix $\mathbf{B}$, containing the N discretized basis functions for the waveform construction in (5), is zero-padded in (18) in a manner consistent with the discretized waveform in (13). Critically, the gradient in (17) can be calculated using only FFTs/IFFTs and matrix/vector multiplies, thus permitting efficient computation for fast optimization. Specifically, the per-iteration cost is $O(M^3 N)$, which can also be expressed as $O(N^4 (K^3/L^3))$ via the relationships in Table I.

In general, gradient-based methods operate by iteratively updating the parameters being optimized such that the cost function is reduced at each iteration. At the ith iteration, this arrangement can be represented as

$$\mathbf{x}_i = \mathbf{x}_{i-1} + \mu_i\,\mathbf{q}_i\,, \qquad (19)$$

where $\mathbf{q}_i$ is the current search direction and μ_i is the current step-size. The search direction $\mathbf{q}_i$ is explicitly a function of the current gradient $\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1})$, and can also incorporate past search directions such that

$$\mathbf{q}_i = \begin{cases} -\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1}) & \text{when } i = 0 \\ -\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1}) + \beta\,\mathbf{q}_{i-1} & \text{otherwise} \end{cases}, \qquad (20)$$

where the gradient itself is a function of the current waveform $\mathbf{s}_{i-1}$ (and thus $\mathbf{x}_{i-1}$) obtained in the previous iteration.

While any of the various gradient-descent methods could be used, in the results presented we employ the heavy ball method [**61**] for which $0 < \beta < 1$. Where other implementations may be more sophisticated, and perhaps realize faster convergence, the simplicity of heavy ball (which relies on gradient "inertia" to dampen abrupt changes) makes it attractive from a computational perspective. As denoted in Table II, which shows the individual steps of the optimization process employed here, if the current search direction $\mathbf{q}_i$ is <u>not</u> actually a descent direction relative to the current gradient $\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1})$ (step 5), then the search direction is reset to the current gradient (step 6), thereby completely forgetting the past gradients (β can also be thought of as a "forgetting factor").

TABLE II
GISL GRADIENT OPTIMIZATION OF PCFM WAVEFORMS

1:	**Initialize**: $M, BT, L, \mathbf{B}, \mathbf{x}_0, p$ (norm), $N = L(BT)$, $\mathbf{q}_0 = \mathbf{0}_{N\times 1},\ \beta,\ \mu,\ \rho_{\mathrm{up}},\ \rho_{\mathrm{down}},\ c,$ and set $i = 1$
2:	**Repeat**
3:	Evaluate: $J_p(\mathbf{x}_{i-1})$ and $\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1})$ via (14) and (17)
4:	$\mathbf{q}_i = -\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1}) + \beta\mathbf{q}_{i-1}$
5:	**If** $\left(\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1})\right)^T \mathbf{q}_i \geq \mathbf{0}$
6:	$\mathbf{q}_i = -\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1})$
7:	**End** (If)
8:	**While** $J_p(\mathbf{x}_{i-1} + \mu\mathbf{q}_i) > J_p(\mathbf{x}_{i-1}) + c\mu\left(\nabla_{\mathbf{x}} J_p(\mathbf{x}_{i-1})\right)^T \mathbf{q}_i$
9:	$\mu = \rho_{\mathrm{down}}\,\mu$
10:	**End** (While)
11:	$\mathbf{x}_i = \mathbf{x}_{i-1} + \mu\mathbf{q}_i,\quad \mu = \rho_{\mathrm{up}}\,\mu$
12:	$i = i + 1$
13:	**Until** $i = I$ **or** $\left\|\nabla_{\mathbf{x}} J_p(\mathbf{x}_i)\right\| < g_{\min}$

The current step-size μ is determined via a simple backtracking approach that satisfies the Armijo condition [**62**] (otherwise known as the first Wolfe condition) with the "sufficient decrease" parameter c that is applied in step 8. In addition, a "step-size increase" parameter ρ_{up} (slightly > 1) is applied in step 11 and "backtracking" parameter ρ_{down} (slightly < 1) is applied in step 9 when the step-size is too large. The combination of ensuring descent (steps 5 and 6) and sufficient decrease (step 8), which triggers backtracking when necessary (step 9), serve the purpose of keeping the optimization within the local minimum region of the waveform initialization (perhaps not important in general but useful to assess the impact of the initialization).

The overall process is then run until either some maximum number of iterations I is reached or the magnitude of the gradient goes below some prescribed minimum $g_{\min}$. Finally, note that we do not consider the second Wolfe condition of "curvature" to avoid the higher computation cost it would incur as a trade-off for faster convergence.

V. Assessment of Gradient-Based PCFM Optimization

There are a large number of parameter combinations that could be considered when performing an evaluation of gradient-based optimization of the GISL cost function for PCFM waveforms. Specifically, there are the parameters of time-bandwidth product BT, the 3-dB oversampling factor K, and the over-coding factor L for the waveforms themselves. As described in Table I, these values collectively determine the number of optimizable PCFM parameters N, as well as the number of samples M in the discretized representation. There is likewise the waveform initialization, which we shall examine throughout this section.

From the cost function perspective, there is the particular norm value p to consider, along with the number of iterations of (19) to perform. As noted in the previous section, there is likewise the choice of gradient-descent formulation, though we restrict consideration here to the heavy ball approach for simplicity.

Clearly, exploring the complicated interplay between all these different parameters, implementation approaches, and initializations in an exhaustive manner is infeasible. Instead, we shall evaluate portions of the design space to ascertain meaningful trends, while showing an even smaller subsequent sampling of illustrative results in greater detail. Specifically, we shall fix M to be 1024 so that a variety of combinations of BT, K, and L can be assessed with a common basis for comparison. We also consider integer values of p ranging from 2 up to 20 to provide adequate breadth for appraisal.

Because no explicit constraint is being placed on spectral containment here (e.g. by imposing limits on α_n [23, 36] or shaping of the spectral roll-off [23, 45-47]), it is appropriate to set a limit on the allowable degree of over-coding. Specifically, to ensure *at least two* samples per T_p interval, the value of L is constrained such that $(BT)(L) \leq M/2$. Of course, it is important to note that the generation of these discretized PCFM waveforms in hardware may necessitate resampling relative to the digital-to-analog converter (DAC) rate of the system, where phase interpolation (instead of standard sinc interpolation or the like) must be used to avoid unnecessary amplitude distortion.

Table III summarizes 228 different combinations of parameters that have been examined and are presented here. In each case the optimization was initialized either with a discretized LFM waveform or a discretized random FM waveform, with the initial coding values for either case bounded on $[-\pi/L, +\pi/L]$ for the given value of L.

TABLE III
Optimization Parameters

M	P	BT	L
1024	2 to 20	64	1 to 8
1024	2 to 20	128	1 to 4

In the results that follow we set $\beta = 0.95$, which has been found empirically to provide a good trade-off between gradient "inertia" (smoothing) and responsiveness for the given cost function. The starting step-size is set smaller than is expected to be necessary for this problem (to $\mu = 10^{-4}$) to support the goal of staying within the local region of the waveform initialization, which matters for the LFM cases (attempting to preserve the chirp-like structure) but not so much for the random FM cases. The step-size increase (ρ_{up}) and backtracking (ρ_{down}) parameters are set to 1.01 and 0.9, respectively, as a crude balance between seeking the largest feasible step-size (for faster convergence) while keeping the number of backtracking steps low (maintain lower computation due to re-evaluations of the cost-function). Finally, empirical observation was also used to set the sufficient decrease parameter to $c = 10^{-2}$, the minimum gradient magnitude to $g_{\min} = 10^{-5}$, and the maximum number of iterations to $I = 10^6$.

A. General Results

The final PSL and ISL values of the 228 individual optimization runs using LFM initialization are tabulated in Appendix B, where Tables V and VI correspond to BT = 64 and Tables VII and VIII correspond to BT = 128. The tables have been shaded such that the lower values (better) are darker. Additionally for the PSL tables, the entries that meet the PSL bound of (15) are bordered in black, which for $M = 1024$ is a PSL value of –60.2 dB.

A couple trends are immediately apparent from these results. First, a higher degree of over-coding (larger L), which corresponds to an increase in the number of independent parameters, tends to realize improved ISL and PSL regardless of BT (aside from a few modest outlier cases for p > 17 in Tables VII and VIII, which may be a result of finite precision effects in the gradient calculation at higher values of p). From an optimization standpoint this general trend is not surprising since greater design degrees of freedom tend to provide better performance. From a waveform design standpoint, however, this outcome indicates that PSL and ISL can be decoupled from the value of BT for a given degree of oversampling. A similar result was observed in [49, 51].

Second, lower values of p tend to provide better final values of ISL, while higher values of p tend to provide better final values of PSL. For example, in Tables V and VII the best ISL values (for each N) correspond to p = 2, 3, or 4. In fact, once L exceeds 3 or 4 we find that p = 2 provides the best ISL results for all higher values of L. Conversely, Tables VI and VIII show that, in nearly all cases, PSL improves as p increases, though the incremental benefit of higher p is diminished as L increases. Indeed, Table VI demonstrates that when N = 512 (L = 8), the PSL bound is achieved for all cases where $p \geq 3$.

In light of these observed trends one can surmise that, to the degree that it is possible for a given sampling rate, higher L (more over-coding) is always desirable, with the selection of p then determined by the preferred emphasis on ISL or PSL. If the latter achieves the bound in (15), the lowest value of p to do so would therefore benefit both criteria. Given the clear benefit to higher L we shall henceforth select the largest value of L possible for the remaining cases examined.

B. Specific Cases

The tables in Appendix B survey the results of performing gradient-descent optimization for a PCFM waveform via the

GISL cost function for an LFM initialization. While results are for specific values of $BT = 64$ (or 128) and $M = 1024$ samples, the observed performance trends also provide indications of what one could expect in general for different BT, albeit with similar relational values of N, K, and L, per Table I.

Having considered the large-scale trends we now examine some specific cases in greater detail to discern particular aspects such as general autocorrelation attributes, spectral shape and containment, optimization convergence behavior, and the frequency function of time. Specifically, from Tables VII and VIII (where $BT = 128$), and focusing on $N = 512$ ($L = 4$), we select the cases $p = 2$ (best ISL), $p = 20$ (meets the PSL bound and largest value of p), and $p = 8$ (best ISL while also meeting the PSL bound).

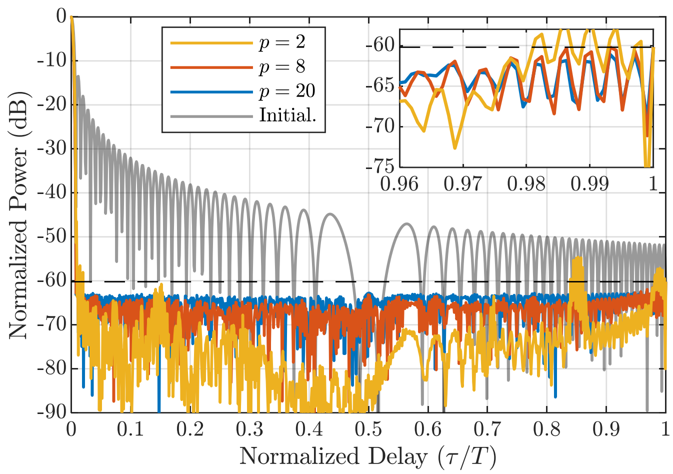

Fig. 2. Autocorrelations of optimized PCFM waveforms using LFM initialization with $BT = 128$, for $M = 1024$, $L = 4$, and $p = 2$, 8, and 20

Figure 2 depicts the one-sided autocorrelation for these three waveforms along with the initializing LFM, with the inset providing a close-up of the response near the outer edge where the PSL bound at –60.2 dB (dashed horizontal line) is applicable for every constant amplitude waveform discretized by M (= 1024 in this case). As the results in Table VIII indicate, the bound value is indeed the PSL value attained for the $p = 8$ (red trace) and $p = 20$ (blue trace) optimized waveforms since neither surpass the bound anywhere else. In contrast, the $p = 2$ optimized waveform (yellow trace) does exceed the bound in a few localized places (around normalized delays of 0.85 and 0.99, and the shoulder-lobes visible in Fig. 3), though it also has a much lower response between normalized delays of roughly 0.15 and 0.85, which serves to facilitate the lower ISL value (average sidelobe level).

Thus, as expected, PSL-focused (or at least larger p-norm) optimization leads to FM waveforms with flatter autocorrelation responses than those optimized for ISL. Moreover, the value of p does not need to be all that large (e.g. $p = 8$) to achieve this flattened result that is beneficial from the perspective of mitigating sidelobe-induced false alarms at the detection stage of the radar receiver.

Figure 3 likewise provides a detailed view of the autocorrelation mainlobe for the LFM initialization and the three optimization results. Here we observe a modest narrowing at the top, though the peak-to-null width is the same for all four cases due to specification of the mainlobe and sidelobe selection vectors from (14), which likewise conforms to the LFM mainlobe.

In Fig. 4 the delay-Doppler ambiguity function for the $p = 8$ optimized waveform is shown (the $p = 2$ and 20 cases are negligibly different). The presence of low range sidelobes is visible via to the narrow vertical notches at zero-Doppler above/below the mainlobe at zero-delay. The prominent sheared ridge of chirped waveforms is also clearly evident, along with the Fresnel lobes around it. Thus, this waveform (along with the other LFM-initialized optimization cases) retains some degree of Doppler tolerance from LFM, albeit with some increase in sidelobes as Doppler increases, as is the case with all NLFM waveforms.

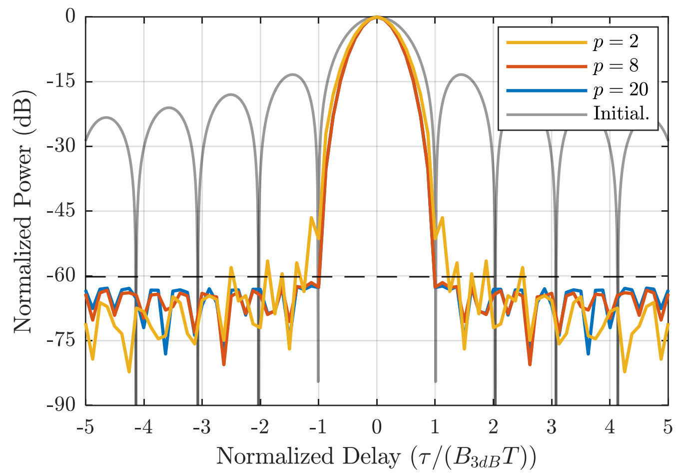

Fig. 3. Mainlobe autocorrelation detail of optimized PCFM waveforms using LFM initialization with $BT = 128$, for $M = 1024$, $L = 4$, and $p = 2$, 8, and 20

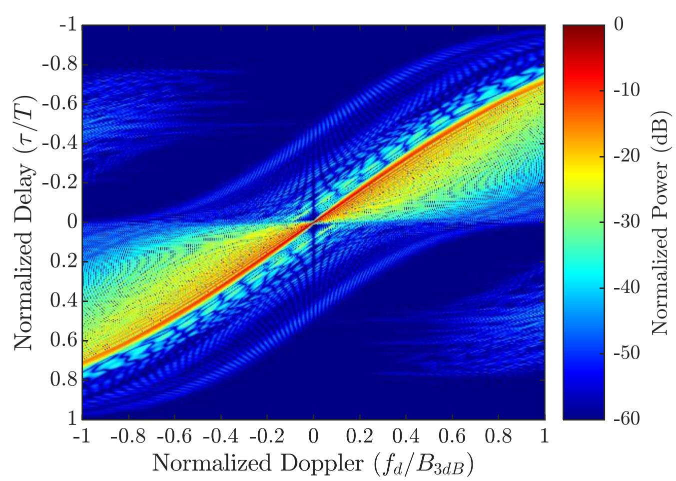

Fig. 4. Delay-Doppler ambiguity function of optimized PCFM waveform using LFM initialization with $BT = 128$, for $M = 1024$, $L = 4$, and $p = 8$

The degree of Doppler tolerance can be examined further by plotting the delay-Doppler ridge (or more generally, the ambiguity function peak value) as a function of normalized Doppler. Figure 5 illustrates this response for LFM, these three optimized waveforms based on LFM initialization, and another optimized waveform based on random initialization (discussed further in Sect. V.D). With LFM the benchmark for Doppler tolerance (see Chap. 20 of [**63**]), we observe that the three optimized waveforms above do incur a trade-off penalty for the far lower sidelobes attained, though compared to the thumbtack-like response of the optimized random FM waveform these three do still preserve at least some of the Doppler tolerance.

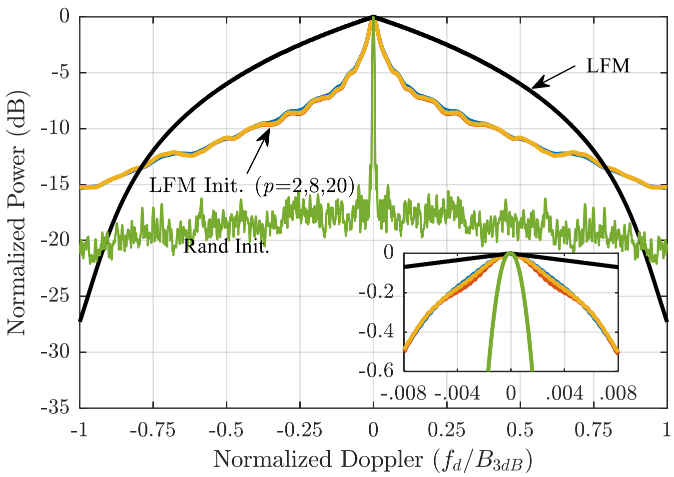

Fig. 5. Ambiguity function peak value as a function of normalized Doppler

A useful way in which to quantify the degree of Doppler tolerance is to measure the rate ψ of roll-off in Fig. 5 near the peak, since the amount of normalized Doppler encountered in practice tends to be small. Because the ambiguity function's Doppler response at zero delay is a $\text{sinc}^2(\bullet)$ function for constant-amplitude waveforms like FM, we can choose the first null of this response as a convenient standard. Specifically, at the boundary established by the first Doppler null of the zero-delay cut, determine the maximum normalized ambiguity function response as

$$\psi = \frac{\max_{\tau} |A(\tau, f_{\text{d}} = 1/T)|}{|A(0,0)|}, \qquad (21)$$

where

$$A(\tau, f_{\text{d}}) = \int_{-\infty}^{\infty} e^{j2\pi f_{\text{d}} t} s(t)\, s^*(t+\tau)\, dt \qquad (22)$$

is the ambiguity function.

For the normalized Doppler depiction below, and $BT = 128$ for these waveforms, the first Doppler null occurs at $f_d = \pm 0.0078B$, which is shown in the detail view of Fig. 5. Consequently, the values of ψ from (21) are –0.07 dB for LFM, –0.49 dB for the $p = 20$ optimized waveform, –0.50 dB for the $p = 2$ and 8 optimized waveforms, and a whopping –14.00 dB for the optimized random waveform (that is clearly not Doppler tolerant). Because the roll-off is quasi-linear in this small Doppler regime (see detail inset of Fig. 5), these values can be used to approximate the rate of loss one incurs for a given waveform as a linear function of normalized Doppler frequency.

Now consider the power spectra of these three optimized waveforms (see Fig. 6) relative to the power spectrum of the initializing LFM, along with a normalized Gaussian power spectrum having the same 3-dB bandwidth for comparison. Unlike in related work [23, 45-48], the latter was not used here, though it is interesting to observe that all three of the optimized waveforms realize power spectra that nonetheless closely conform to the upper 20 dB of the Gaussian. Anecdotally at least, this outcome reaffirms the use of a Gaussian power spectral template [**64**] for these other forms of waveform design that instead rely on spectrum matching. While not explored here, one could also relate this result to the spectral concentration problem that ultimately realizes the Slepian window [**65**], the mainlobe of which is likewise quite similar to the observed upper portion of the power spectrum in Fig. 6.

The 3-dB bandwidth is achieved indirectly by setting the autocorrelation mainlobe width via $\mathbf{w}_{\text{ML}}$ and $\mathbf{w}_{\text{SL}}$ in (14), thus establishing the range resolution. Specifically, the null-to-null autocorrelation width is preserved in this manner (relative to the LFM initialization), though the spectral shape changes somewhat. Indeed, it is clear from Fig. 6 that to achieve significant sidelobe reduction relative to LFM necessitates accepting more gradual spectral roll-off compared to LFM, which provides quite good spectral containment. Outside the Gaussian-like roll-off (about 20 dB below the peak), each optimized spectrum expands further, a direct consequence of using over-coding. Indeed, this "spectral fuzz" (as it was called in [36]) only appears when $L > 1$.

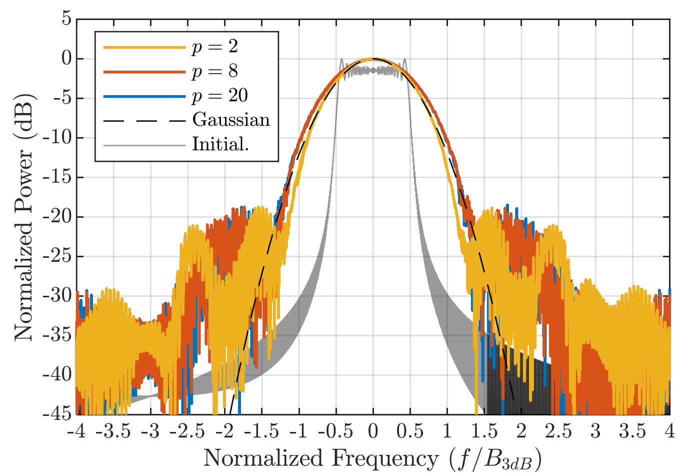

Fig. 6. Power spectra of optimized PCFM waveforms using LFM initialization with $BT = 128$, for $M = 1024$, $L = 4$, and $p = 2$, 8, and 20

Figure 7 depicts a discretization of the instantaneous frequency function of time (i.e. the values in $\mathbf{x}$) for each waveform, and in so doing illustrates the reason behind the expanded spectral content of Fig. 6. While the general chirping trend relative to LFM (in black) is maintained, along with the accelerated chirping at the pulse edges like in classical NLFM (the "sideways S" [5]), these optimized waveforms exhibit unique behavior by way of non-monotonic ringing at localized regions throughout each waveform (i.e. a "perturbed sideways S"). The $p = 2$ (yellow trace) ringing is the most localized, yet also has the most significant deviation from the classical monotonic NFLM form. The $p = 20$ case (blue trace) has a marginally greater distribution of ringing than the $p = 8$ case (red trace), yet both exhibit a slightly lower degree of deviation from classical NLFM than $p = 2$. All three waveforms also have small variations throughout (most visible at the pulse edges) that are likewise non-monotonic.

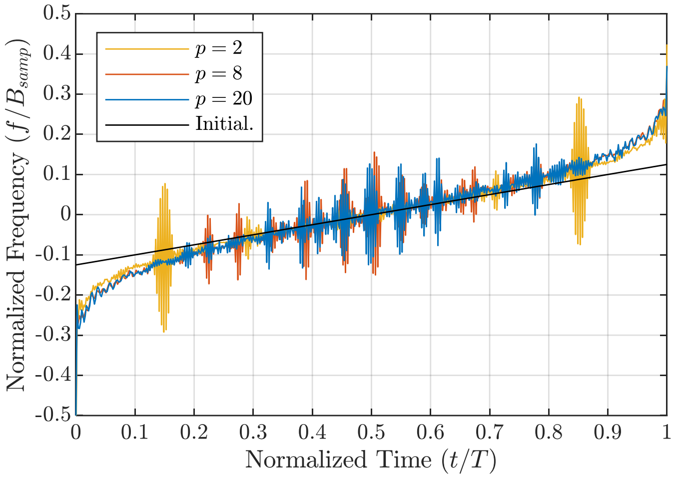

Fig. 7. Normalized instantaneous frequency function for the p = 2, 8, and 20 optimized waveforms using LFM initialization

It is also interesting to note that the $p = 2$ case appears to be closer to the typical anti-symmetry arrangement than the $p = 8$ or $p = 20$ cases, though the structure is at least approximately maintained for all of them. Clearly one take-away here is that monotonicity and symmetry are not necessarily requirements for good NLFM waveforms, though some degradation of spectral containment may be the trade-off, per Fig. 6.

Figure 8 shows the convergence in terms of ISL (dashed traces) and PSL (solid traces) when these three LFM-initialized waveforms were optimized using $p = 2$, 8, and 20. An interesting observation is that all these cases realize logarithmically increasing numbers of iterations in which the performance enhancement is quite small, interspersed by shorter intervals of significant improvement, which might indicate proximity to saddle points in the cost function surface.

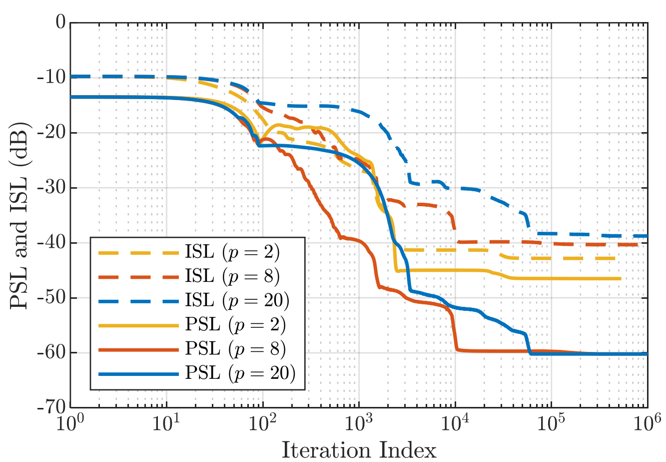

Fig. 8. PSL and ISL convergence when optimizing for p = 2, 8, and 20, with LFM initialization

It is instructive to examine the PSL convergence in Fig. 6 for the $p = 2$ case (solid yellow trace). Unlike the rest of the results in Fig. 8, this particular trace does not decrease monotonically, though overall it does clearly reduce significantly from its initial value. The reason for this distinction is that, while ISL and PSL both measure sidelobe performance in general, their cost functions also possess different local minima. It is this similar, yet different, relationship that was exploited by the "performance diversity" paradigm in [23] for a greedy search. It remains to be seen how this concept could also be applied to gradient-based optimization, though the similar prospect of cycling between different values of p [**57**] could certainly be considered.

Finally, to gain insight into how the instantaneous frequency features in Fig. 7 provide PSL/ISL improvement, Fig. 9 plots several intermediate instantaneous frequency functions for the $p = 8$ case taken at different points along the optimization process. The specific points were chosen to be the iteration indices immediately following significant reductions in either PSL or ISL (the solid/dashed orange traces in Fig. 8). For instance, after 2×10^2 iterations the frequency function (blue trace in Fig. 9) has begun to deviate somewhat from LFM. After 2×10^3 iterations (orange trace) we see a clear sideways-S trend, along with small perturbations. Then 2×10^4 roughly marks the point on Fig. 8 at which the last significant decrease in both PSL and ISL has recently occurred, and we observe in Fig. 9 (purple trace) that the frequency function has only changed a small amount further, now possessing even smaller perturbations. Interestingly, it is the rather miniscule remaining improvement (about 0.54 dB), during which the optimization reaches the PSL bound of (15), that introduces the large perturbations that appear almost speech-like (yellow trace). Consequently, it is not surprising that this attribute of optimal FM waveforms (in a PSL sense when discretized) has not been seen before.

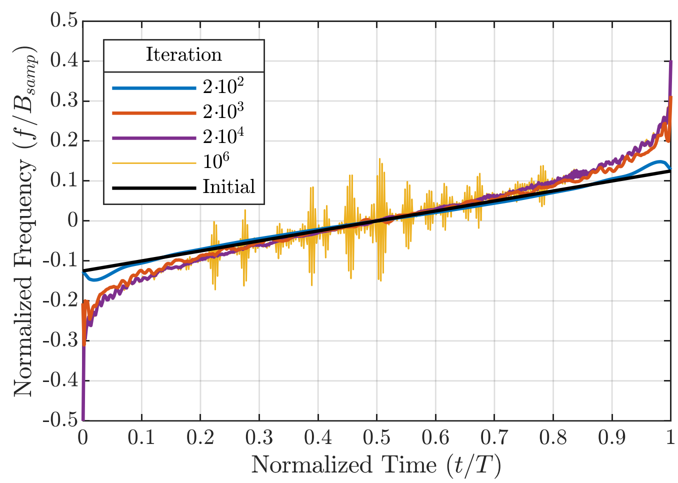

Fig. 9. Normalized instantaneous frequency function for the p = 8 optimized waveform using LFM initialization at several points during optimization

C. Other Chirped Initializations

For non-convex problems choosing the particular optimization parameters is often a heuristic process of trial and error. This concept likewise extends to the choice of initialization, where one set of optimized results may provide insight into better initialization choices.

For instance, consider the optimized frequency functions in Fig. 7. Ignoring the flared tails and small perturbations, one can perceive a general quasi-linear central frequency slope that is steeper than the initial LFM. This result begs the question: "would it be better to initialize with an LFM that has a wider bandwidth than that which is desired after optimization?" After all, it is the autocorrelation mainlobe selection vector $\mathbf{w}_{\mathrm{ML}}$ that ultimately determines the 3-dB bandwidth, and presumably setting the chirp-slope initialization closer to the expected final result should provide faster convergence.

In fact, we can take this line of thought a step further. Perturbations notwithstanding, these optimized frequency functions bear a strong resemblance to the inverse error function. Would this initialization yield even faster convergence?

To evaluate this comparison of chirp initializations, Fig. 10 illustrates the frequency function of the $p = 8$ optimization (gray trace) from Fig. 7 that relied on what we shall refer to as the "Old LFM" initialization (black trace). A best line fit to this optimization result is likewise shown (blue trace) and denoted as "New LFM". Finally, an inverse error function (orange trace) is included that is scaled to also follow the $p = 8$ optimized frequency function. These three initializations are subsequently employed to optimize using $p = 8$ (obviously yielding the same result in the Old LFM case).

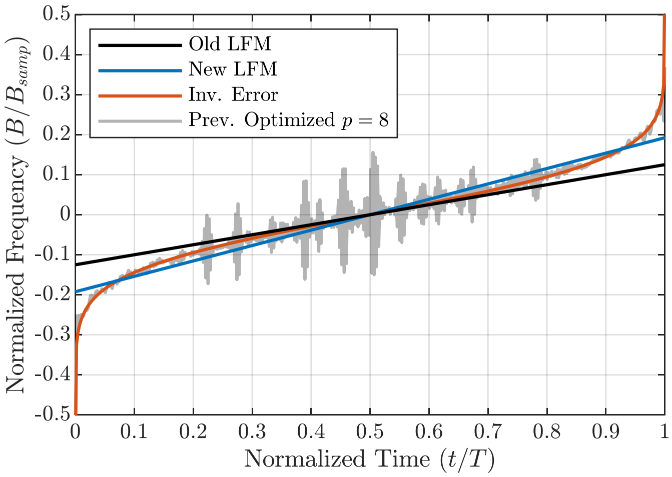

Fig. 10. Normalized instantaneous frequency functions used for initialization comparison, along with that for the *p* = 8 optimized waveform from Fig. 5

Figure 11 illustrates the PSL/ISL convergence for each of these initializations. All cases realize nearly the same final PSL/ISL values, with differences of less than 1 dB, and essentially identical power spectra and ambiguity functions. Relative to these converged final values, where the Old LFM initialization required roughly 10^4 iterations, the New LFM and inverse error function initializations need only about 3×10^3 iterations, or roughly a 3-fold improvement. Interestingly, while the inverse error function clearly starts with the best ISL/PSL, it still takes the same number of iterations to converge as New LFM. It should also be noted that the final waveforms obtained are quite similar, yet not identical, due to the non-convex cost function and the continuum of possible waveform structures.

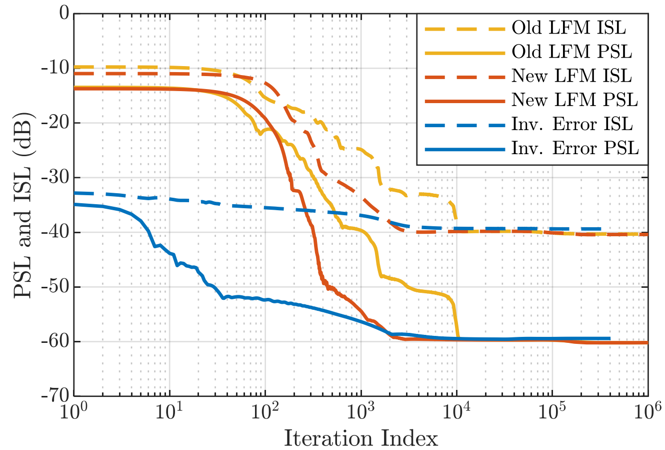

Fig. 11. PSL and ISL convergence when optimizing for *p* = 8 with different chirped initializations

D. Randomized Initialization

Up to this point we have focused only on chirped initializations because doing so leverages the sheared delay/Doppler ridge in a "conservation of ambiguity" sense [5]. We now consider various random initializations to determine how gradient-based optimization performs in general and to serve as a point of comparison with the results above.

Since introducing Monte Carlo trials over a set of random waveforms introduces even more complexity into the analysis, we shall focus on cases similar to the previous examples. Specifically, 2000 random FM waveforms with $BT = 128$ and $L = 4$ (so $N = 512$) were generated with $M = 1024$ samples. To prevent spectral expansion, each contiguous sequence of L PCFM parameters in $\mathbf{x}$ is identical (initially) and drawn from a uniform distribution on $[-\pi, +\pi]$, then divided by L, thus providing a reasonable (but in no means optimized) random FM waveform for the given BT.

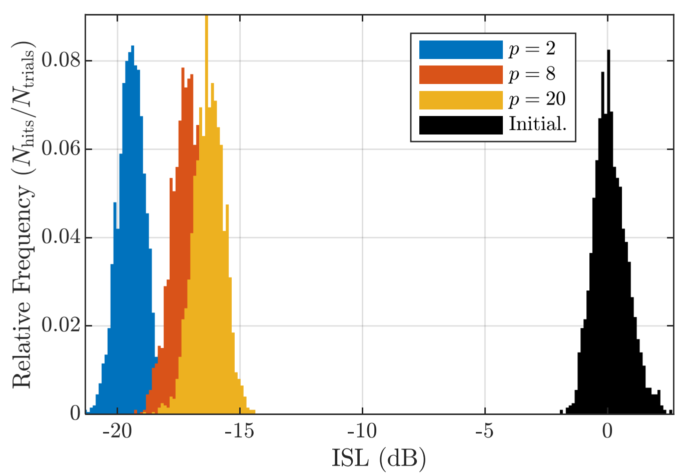

Fig. 12. ISL histograms for 2000 waveforms optimized using *p* = 2, 8, or 20 based on independent random initializations

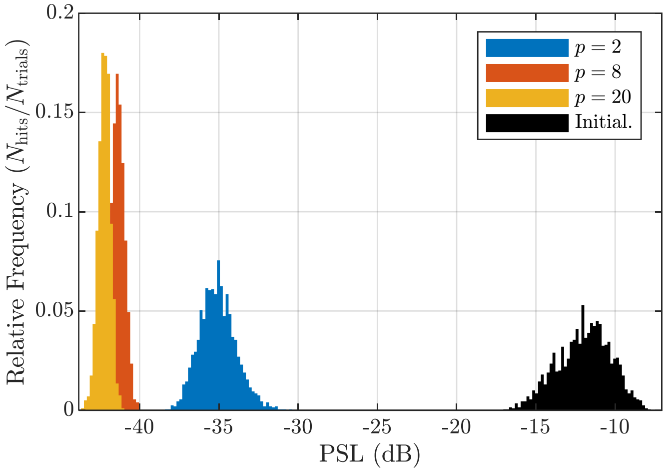

Fig. 13. PSL histograms for 2000 waveforms optimized using *p* = 2, 8, or 20 based on independent random initializations

Each of the 2000 random initializations were optimized using $p = 2$, 8, and 20 according to the approach in Table II (same as before). The resulting ISL and PSL values are plotted as histograms in Figs. 12 and 13, respectively. Generally speaking, the ISL values (Fig. 12) improve by roughly 17 to 20 dB over the initial values. More significantly, the PSL values improve by roughly 23 to 30 dB. However, comparing these results to what was achieved using a chirped initialization above, given the same parameters, shows that random initialization is clearly inferior on a per-waveform basis.

Comparisons to chirped initialization aside, it is interesting to note that similar trends are observed in the random initialization cases. Specifically, smaller values of p tend to provide better final ISL values (Fig. 12) while larger values of p tend to realize better final PSL values (Fig. 13).

An additional observation can be made with regard to distribution variances. Figure 13 reveals that random initializations tend to produce a relative wide PSL variance that is tightened considerably following optimization using higher p values. This behavior is linked to the flattening of sidelobes that occurs for PSL-based optimization.

Moreover, the useful metric $-20\log_{10}(BT)$ dB, which is related (by a further -3 dB) to a PSL bound specific to hyperbolic FM waveforms [**66**], is found to be -42.1 dB here (for $BT = 128$). In Fig. 13 we find this value to be quite close to the center of the PSL distribution for $p = 20$ optimized waveforms. Thus it can be inferred that the sidelobe performance for random FM waveforms may at best approach a limit determined by BT, for B the 3-dB bandwidth. In other words, the additional dimensionality provided by over-coding ($L > 1$), which had such a significant impact for LFM initialization in the previous section, may have a less discernible effect on random waveforms.

Like with the LFM initializations, it is also instructive to examine the characteristics of individual waveforms arising from optimization of random initializations. Figure 14 depicts the autocorrelation of an arbitrarily selected initialization from the Monte Carlo trials when optimized for $p = 2$, 8, and 20. As expected based on Figs. 12 and 13, the initial random waveform (gray trace) has a rather mediocre autocorrelation in terms of sidelobe level. All three optimization cases significantly improve upon this initialization, though the sidelobes are still considerably higher than what was achieved previously using LFM initialization. Interestingly, the $p = 8$ and 20 cases again do a good job of flattening the sidelobe response, even if the level is ~20 dB higher than that demonstrated in Fig. 2. Note that the mainlobe responses here (not shown) are indistinguishable from those in Fig. 3 (same null-to-null width).

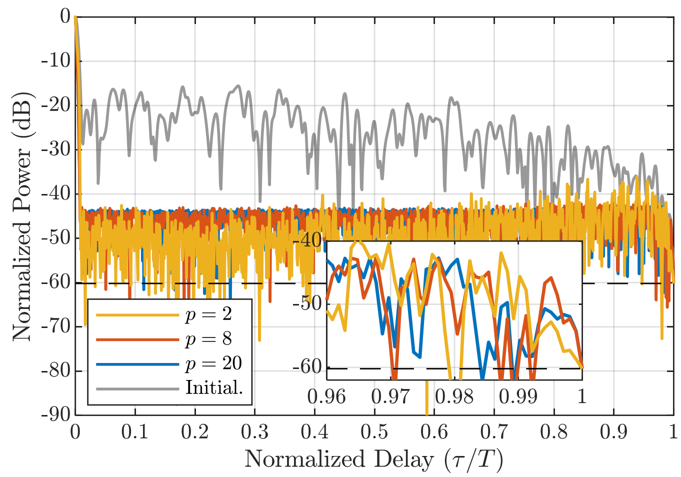

Fig. 14. Autocorrelations of optimized PCFM waveforms using random initialization with *BT* = 128, for *M* = 1024, *L* = 4, and *p* = 2, 8, and 20

Figure 15 depicts the delay-Doppler ambiguity function for the $p = 8$ case. Compared to the result in Fig. 4, which retains the sheared ridge structure of the LFM chirp, the response here has very little structure. The presence of lower range sidelobes (relatively speaking) on the vertical axis at zero Doppler is clearly visible in this case due to the diffusion of ambiguity over the delay-Doppler surface, again emphasizing the benefit of the sheared ridge for Doppler tolerant waveforms to serve as an absorber of ambiguity.

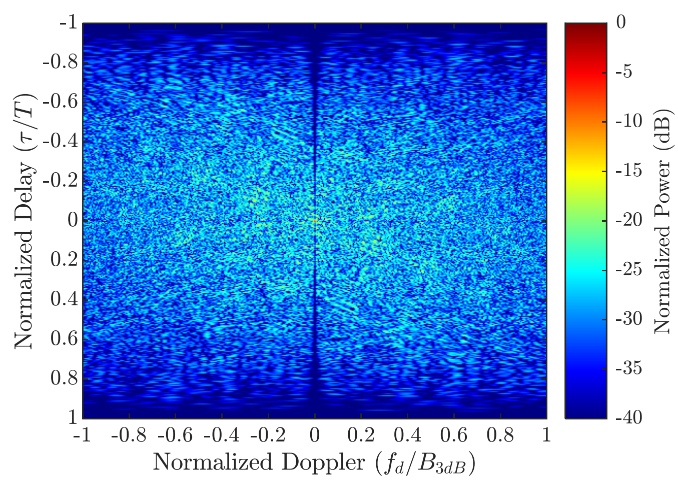

Fig. 15. Delay-Doppler ambiguity function of optimized PCFM waveform using random initialization with *BT* = 128, for *M* = 1024, *L* = 4, and *p* = 8

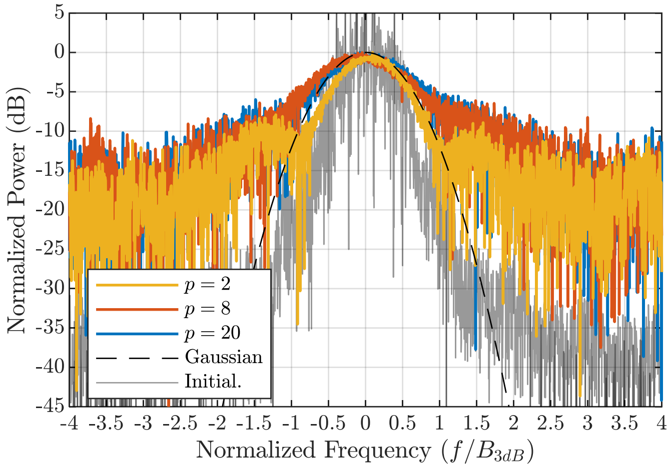

Fig. 16. Power spectra of optimized PCFM waveforms using random initialization with BT = 128, for M = 1024, L = 4, and p = 2, 8, and 20

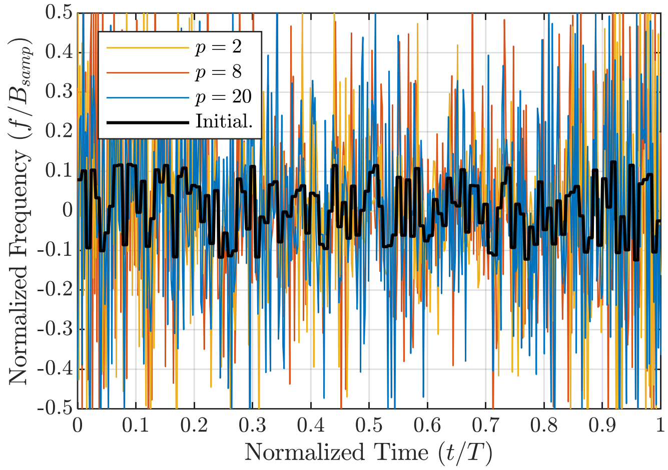

Fig. 17. Normalized instantaneous frequency function for the p = 2, 8, and 20 optimized waveforms using random initialization

Figures 16 and 17 show the power spectra and frequency functions of time, respectively, of the random initialization and optimized waveforms. Per Fig. 16, it is interesting to note that the random initialization has relatively good containment in the spectral roll-off region (based on how the initialization was performed) yet is clearly rather jagged in the passband. In contrast, the three optimized random spectra possess (again) a Gaussian-like response in the passband but exhibit greatly expanded roll-off. The latter is directly related to the rather wild instantaneous frequency excursions depicted in Fig. 17, which in contrast to Fig. 6, have no obvious anti-symmetry.

To quantify this observation, we can apply a normalized bilateral anti-symmetry (BAS) metric to the baseband waveform as

$$\mathrm{BAS} = \frac{2}{T}\left|\int_{0}^{T/2} s^{*}(t)\, s(T-t)\, dt\right|, \tag{23}$$

where values near 1 indicate the waveform is almost perfectly anti-symmetric (integration of a near constant in (23)), while values near 0 indicate a lack of anti-symmetry. In Fig. 17, the BAS metric was calculated for all of the randomly initialized optimized waveforms and plotted as a histogram. The BAS values for the waveforms optimized from LFM initialization (Sect. V.B) have likewise been assessed and included.

Clearly the LFM-initialized waveforms retain a high degree of anti-symmetric structure, which was also qualitatively observed in Fig. 8. In contrast, the randomly initialized waveforms possess almost no BAS structure since they are concentrated between 0 and 0.2. This result further confirms previous observations regarding the importance of anti-symmetric structure, though it need not be perfect to achieve the best PSL results.

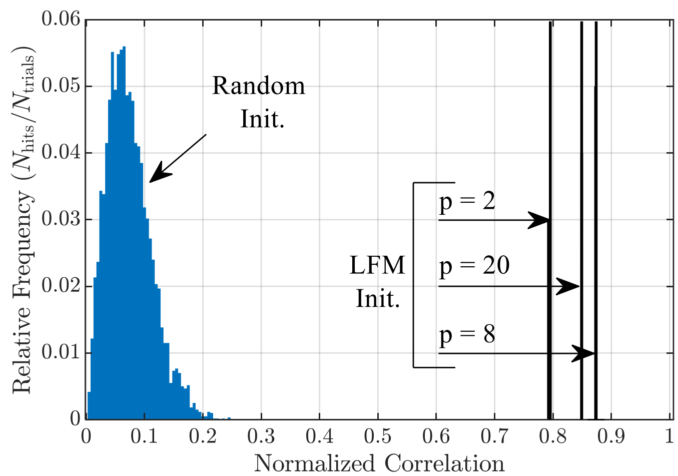

Fig 18. Normalized bilateral anti-symmetry (BAS) correlation via (23) of optimized PCFM waveforms based on random or LFM initializations

These random initialization results rather strongly indicate the need for spectral shaping when performing optimization of arbitrary FM waveforms. Specifically, the poorly contained spectra in Fig. 16 would be distorted by the (linear) transfer function of most transmitters, followed by further nonlinear distortion due the associated loss of constant amplitude. Methods to perform spectral shaping of FM waveforms have been examined in [23, 45-48].

It is also worth noting that, while these individual optimized FM waveforms obtained via random initialization are clearly inferior to those produced by LFM initialization, the single-waveform perspective is only one way to look at the sensing problem. It was discussed in [48] (and references therein) that generating a nonrepeating sequence of random FM waveforms has the benefit of substantially increasing design degrees of freedom and signal dimensionality (translates into an "aggregate BT " over the coherent processing interval instead of the BT for a single waveform). Consequently, this trade-off in performance for individual waveforms can actually facilitate new sensing capabilities when part of a larger set of different waveforms. Further investigation of that topic is, however, beyond the scope of this paper.

VI. Experimental Results

As described in [20, 23], the PCFM waveform implementation was developed as a means to leverage the parameter optimization benefits of coding while preserving a physical structure that is amenable to the rigors of a high-power radar transmitter. Thus we now examine how these gradient-descent optimized waveforms perform when implemented in hardware, which is necessary to understand their behavior in practice. Specifically, hardware implementation introduces unavoidable distortion due to the imposition of component transfer functions (i.e. filtering), nonlinear effects, and range straddling effects (non-ideal receive sampling). See [23, 51] for further details on the influence of distortion on optimized waveforms.

Here the impact of hardware is assessed in two ways. First, the three optimized waveforms using LFM initialization from Sect. V.B. are evaluated in a loopback configuration to determine how well their low sidelobe attributes are maintained after transmitter distortion. Based on this evaluation, the $p = 20$ waveform from this group was selected to illuminate a free-space scene that contains traffic traversing an intersection, thereby demonstrating practical utility for radar operation.

A. Loopback Assessment

To implement them in hardware the three LFM-initialized waveforms were up-sampled to 10 GSamples/sec using phase interpolation and then loaded onto a Tektronix AWG70000A arbitrary waveform generator (AWG), with each having a pulse width of $T = 5.12$ μs and a 3-dB bandwidth of 25 MHz (so range resolution of ~6 m). The signal produced by the AWG was passed through a Mini-Circuits TVA-82-213 RF amplifier (operating in saturation) and a subsequent Mini-Circuits BW-S40W2+ 40 dB attenuator to emulate an open-air environment before being captured in loopback (wired connection) and recorded on a Rohde & Schwarz real-time spectrum analyzer (RSA) at a receive sampling rate of 200 MHz.

Figure 19 presents autocorrelations of the loopback captured versions of the three optimized waveforms. Comparing against the ideal responses in Fig. 2, some modest degradation can be observed, with the most significant seeming to be with the $p = 2$ case, for which the lower regions from Fig. 2 are now filled in to roughly the same level as the $p = 8$ and 20 cases.

Table IV provides a quantitative comparison between the ideal responses obtained by optimization and the resulting loopback-measured responses. In terms of ISL, the $p = 2$, 8, and 20 cases respectively realize degradations of 7.0 dB, 4.4 dB, and 2.7 dB. Moreover, the 2-3 dB ISL advantage of the $p = 2$ case under ideal conditions has now become a 0.1-0.3 dB disadvantage, which is not necessarily surprising given the "filling in" observation above.

TABLE IV
IDEAL VS. LOOPBACK COMPARISON

p	ISL		PSL	
	Ideal	Measured	Ideal	Measured
2	–42.8	–35.8	–46.5	–41.7
8	–40.3	–35.9	–60.2	–53.8
20	–38.8	–36.1	–60.2	–55.1

With regard to PSL, first note that the value for the $p = 2$ case is dictated by a shoulder lobe just outside the mainlobe that is present both before and after transmitter distortion. While the degradation trend for PSL is a bit more even across the three cases (4.8 dB, 6.4 dB, and 5.1 dB, respectively) the clear winner is still $p = 20$, which retains a PSL value of −55.1 dB. These results therefore imply (at least anecdotally) that flattened autocorrelation responses realized by PSL-optimized waveforms (or least via sufficiently large p) may be more robust to transmitter distortion. Consequently, we shall use the $p = 20$ waveform for subsequent free-space measurements.

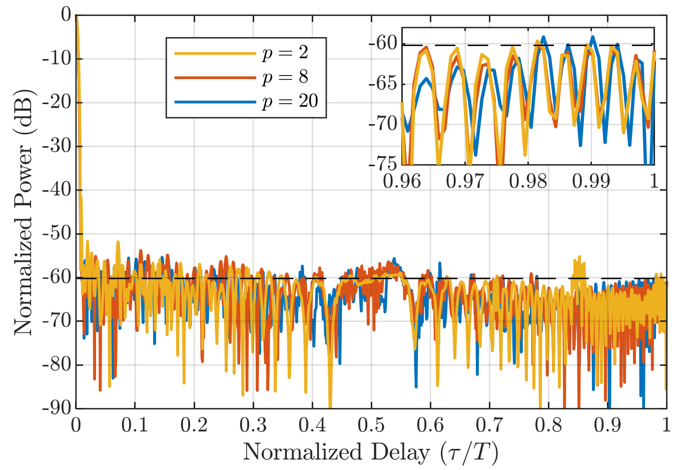

Fig. 19. Autocorrelations of loopback captured versions of optimized PCFM waveforms using LFM initialization with *BT* = 128, for *M* = 1024, *L* = 4, and *p* = 2, 8, and 20

It is also instructive to examine the power spectra of the loopback captured waveforms, as shown in Fig. 20. Comparing with the ideal power spectra in Fig. 6 it is observed that the three traces appear to be identical between normalized frequency values of roughly ±3. Beyond this point the loopback measurements experience a sharper roll-off, though this result is due to the anti-aliasing filter of our RSA "receiver" as opposed to any significant effects from the "transmitter" (the AWG in this context). Thus, one can conclude that this bandlimiting effect upon the high-frequency components may be the primary reason behind the modest autocorrelation degradation observed in Fig. 19 and Table IV.

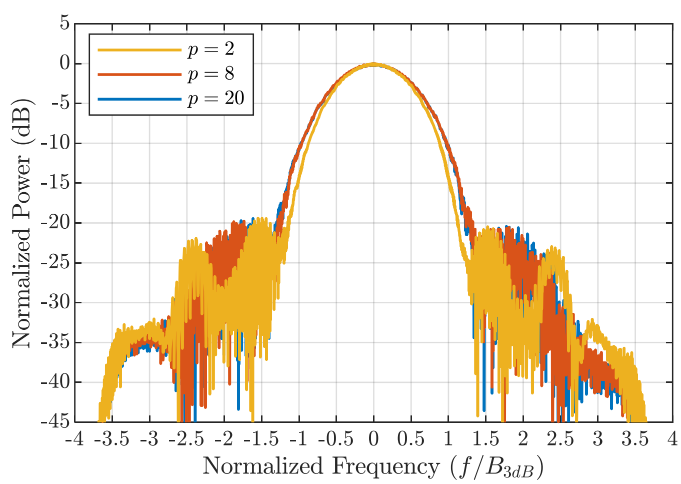

Fig. 20. Power spectra of loopback captured versions of optimized PCFM waveforms using LFM initialization with *BT* = 128, for *M* = 1024, *L* = 4, and *p* = 2, 8, and 20

For high-power operation the impact of the power amplifier (Class C or higher) is generally the main source of transmitter distortion. Thus, further work is needed to explore how the incorporation of predistortion [50] and/or hardware-in-the-loop optimization [23] can be incorporated into gradient-descent based waveform design. Moreover, there exist a variety of distortion models such as memory polynomial, Volterra, Hammerstein, etc. (see [**67**]) that, with sufficient determination of the transmitter model parameters, could potentially be incorporated into this gradient formulation to design waveforms tuned to a particular transmitter.

B. Open-Air Assessment

Each of the three waveforms implemented on the AWG was also transmitted in free-space at a center frequency of 3.55 GHz from the roof of a building on the University of Kansas campus toward the intersection of 23rd and Iowa streets in Lawrence, KS (a distance of about 1 km). Two parabolic dish antennas were used to facilitate simultaneous transmit and receive. Consequently, the direct path leakage between these antennas becomes the dominant received signal component (normalized to 0 dB), as shown in Fig. 21. The sidelobes induced by this direct path establish the floor at around –60 dB that extends out to almost 800 m, which agrees with the loopback-measured autocorrelation responses observed in Fig. 19.

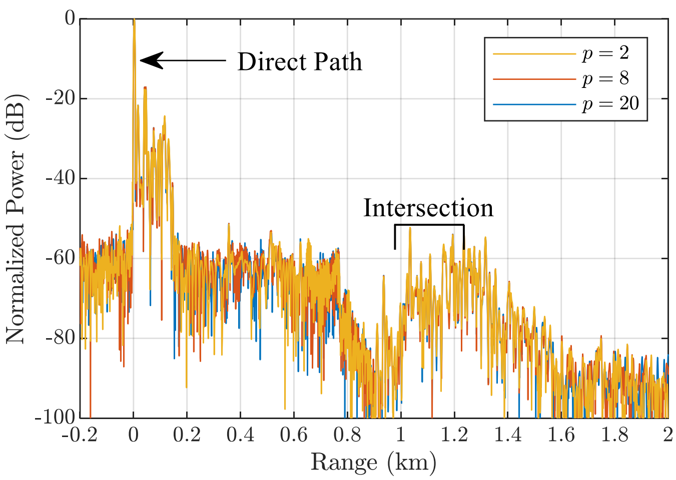

Fig. 21. Range profile (zero-Doppler cut) obtained from matched filtering the open-air emission of the p = 2, 8, and 20 optimized waveforms obtained using LFM initialization with BT = 128, M = 1024, and L = 4 (normalized to direct path response)

At the interval from about 1.0 to 1.4 km several scatterers are visible in Fig. 21. The traffic intersection residing at roughly 1.0 to 1.2 km contains multiple moving vehicles, with the rest being buildings on the other side of the intersection. Figure 22 then shows the range-Doppler response to a CPI of 1000 pulses modulated with the p = 20 waveform after standard two-pulse clutter cancellation and application of a Tukey taper to reduce Doppler sidelobes. The responses for the $p = 2$ and 8 waveforms are negligibly different and are therefore omitted. The pulse repetition frequency (PRF) here is 8.33 kHz and the dwell time is 120 ms, yielding an unambiguous velocity of 176.1 m/s and a Doppler resolution of (8.33 Hz) 0.35 m/s.

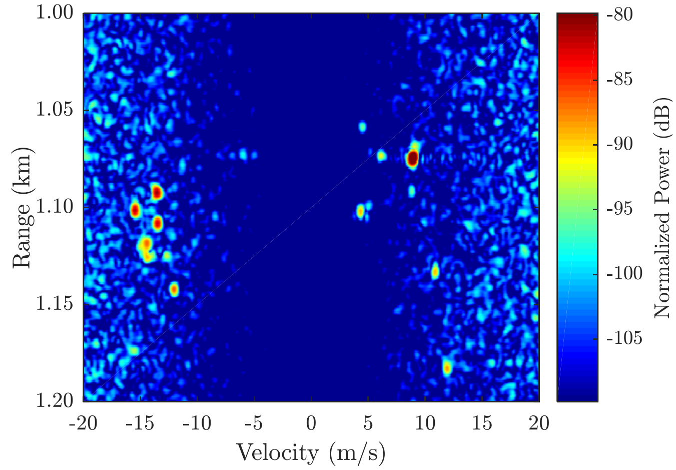

Fig. 22. Open-air range-Doppler response (direct path normalized) after standard two-pulse clutter cancellation when illuminating a traffic intersection with 1000 pulses modulated by the p = 20 optimized waveform obtained using LFM initialization with BT = 128, M = 1024, and L = 4

Multiple approaching and receding moving targets (cars and trucks) appear to be visible in Fig. 22. In fact, the dynamic range of about 30 dB between the scattering from the largest mover and the noise floor is not high enough for the range sidelobes to be observed because they all reside well below the noise. Thus the key take-away from this result is that these optimized FM waveforms can indeed be physically produced with high fidelity to enable practical radar operation with low sidelobes, and thus high sensitivity.

VII. CONCLUSIONS

A gradient-descent approach to optimizing physically realizable PCFM waveforms has been developed and demonstrated, both in terms of ideal performance and through experimental measurements. The approach optimizes all waveform parameters simultaneously in a computationally efficient manner that can be performed using FFT operations and matrix/vector multiplications. It is therefore extensible to a variety of higher dimensional, waveform-diverse applications such as MIMO and pulse agility.

It was observed that optimization using higher (but still relatively modest) values of the p-norm provide a practical surrogate to PSL-based optimization. Moreover, the flattened autocorrelation that results from using a higher p-norm appears to be more robust to transmitter distortion effects than when lower values of p are employed. Ongoing research involves the compensation of transmitter distortion effects (e.g. predistortion and/or hardware-in-the-loop design) and the appropriate inclusion of spectral shaping so that gradient descent can also be used to design arbitrary, random FM waveforms.

APPENDIX A. DERIVATION OF GISL GRADIENT

The gradient of the discretized GISL cost function in (14) can be evaluated in a number of ways. The approach taken here involves differentiating with respect to a single PCFM parameter α_n and then arranging the result in such a way that the derivative with respect to all N PCFM parameters can be computed in parallel.

Begin by rewriting (14) as

$$J_p = \frac{\|\mathbf{w}_{\text{SL}} \odot \mathbf{r}\|_p^2}{\|\mathbf{w}_{\text{ML}} \odot \mathbf{r}\|_p^2} = \left(\frac{\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right)^{(2/p)}, \tag{24}$$

making use of the fact that the elements of $\mathbf{w}_{\text{SL}}$ and $\mathbf{w}_{\text{ML}}$ contain only ones and zeroes. By the chain rule, the partial derivative of (24) with respect to α_n is

$$\begin{aligned}\frac{\partial J_p}{\partial \alpha_n} &= \left(\frac{2}{p}\right)\left(\frac{\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right)^{(2/p-1)} \left[\frac{\partial}{\partial \alpha_n}\left(\frac{\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right)\right] \\ &= \left(\frac{2}{p}\right) J_p^{(1-p/2)} \left[\frac{\partial}{\partial \alpha_n}\left(\frac{\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right)\right]\end{aligned}, \tag{25}$$

where we have employed the relationship

$$\left(\frac{\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right)^{(2/p-1)} = \left(\left(\frac{\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right)^{(2/p)}\right)^{(1-p/2)} = J_p^{(1-p/2)} \tag{26}$$

to maintain compact notation. Subsequent application of the quotient rule to (25) then yields

$$\begin{aligned}\frac{\partial J_p}{\partial \alpha_n} &= \left(\frac{2}{p}\right) J_p^{(1-p/2)} \left(\frac{1}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right)^2 \times \\ &\quad \left[\left(\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p\right)\left(\mathbf{w}_{\text{SL}}^T \frac{\partial |\mathbf{r}|^p}{\partial \alpha_n}\right) - \left(\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p\right)\left(\mathbf{w}_{\text{ML}}^T \frac{\partial |\mathbf{r}|^p}{\partial \alpha_n}\right)\right] \\ &= \left(\frac{2}{p}\right) J_p^{(1-p/2)} \left(\frac{\left(\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p\right)\left(\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p\right)}{\left(\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p\right)^2}\right) \times \\ &\quad \left[\frac{1}{\left(\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p\right)}\left(\mathbf{w}_{\text{SL}}^T \frac{\partial |\mathbf{r}|^p}{\partial \alpha_n}\right) - \frac{1}{\left(\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p\right)}\left(\mathbf{w}_{\text{ML}}^T \frac{\partial |\mathbf{r}|^p}{\partial \alpha_n}\right)\right] \\ &= \left(\frac{2}{p}\right) J_p \times \\ &\quad \left[\frac{1}{\left(\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p\right)}\left(\mathbf{w}_{\text{SL}}^T \frac{\partial |\mathbf{r}|^p}{\partial \alpha_n}\right) - \frac{1}{\left(\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p\right)}\left(\mathbf{w}_{\text{ML}}^T \frac{\partial |\mathbf{r}|^p}{\partial \alpha_n}\right)\right]\end{aligned} \tag{27}$$

since

$$\left(\frac{\mathbf{w}_{\text{SL}}^T |\mathbf{r}|^p}{\mathbf{w}_{\text{ML}}^T |\mathbf{r}|^p}\right) = J_p^{p/2}. \tag{28}$$

The partial derivative of $|\mathbf{r}|^p$ in (27) can be rewritten and evaluated using both the chain and product rules as

$$\begin{aligned}\frac{\partial |\mathbf{r}|^p}{\partial \alpha_n} &= \frac{\partial}{\partial \alpha_n}\left(\mathbf{r} \odot \mathbf{r}^*\right)^{(p/2)} \\ &= \left(\frac{p}{2}\right)\left(\mathbf{r} \odot \mathbf{r}^*\right)^{(p/2-1)} \odot \left[\mathbf{r} \odot \left(\frac{\partial}{\partial \alpha_n}\mathbf{r}^*\right) + \left(\frac{\partial}{\partial \alpha_n}\mathbf{r}\right) \odot \mathbf{r}^*\right] \\ &= p|\mathbf{r}|^{(p-2)} \odot \Re\left\{\left(\frac{\partial}{\partial \alpha_n}\mathbf{r}\right) \odot \mathbf{r}^*\right\}\end{aligned} \tag{29}$$

where the imaginary parts in the summation cancel and $\Re\{\cdot\}$ extracts the real part of the argument.

Now recall the definition of $\mathbf{r}$ in (12) based on the spectral density. Since the DFT and IDFT are linear operations, the derivative in (29) becomes

$$\begin{aligned}\frac{\partial}{\partial \alpha_n}\mathbf{r} &= \mathbf{A}^H \left[\frac{\partial}{\partial \alpha_n}\left((\mathbf{A}\bar{\mathbf{s}}) \odot (\mathbf{A}\bar{\mathbf{s}})^*\right)\right] \\ &= 2\mathbf{A}^H \Re\left\{(\mathbf{A}\bar{\mathbf{s}}) \odot \left(\mathbf{A}\frac{\partial \bar{\mathbf{s}}}{\partial \alpha_n}\right)^*\right\}\end{aligned}, \tag{30}$$

after once again invoking the product rule. Thus the final derivative involves the PCFM waveform itself. Using the compact representation in (5), the non-zero portion of the derivative in (30) is

$$\frac{\partial \mathbf{s}}{\partial \alpha_n} = \frac{\partial}{\partial \alpha_n}\exp\left(j\mathbf{B}\mathbf{x}\right) = j\mathbf{b}_n \odot \mathbf{s}, \tag{31}$$

where $\mathbf{b}_n$ is the n^{th} column of $\mathbf{B}$. The zero-padded form of the derivative in (30) can then be easily obtained via

$$\frac{\partial \bar{\mathbf{s}}}{\partial \alpha_n} = j\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}}, \tag{32}$$

in which

$$\bar{\mathbf{b}}_n = [\mathbf{b}_n^T \ \mathbf{0}_{1\times(M-1)}]^T, \tag{33}$$

and the collection of all N of these zero-padded basis vectors produces the zero-padded matrix in (18). Thus (30) becomes

$$\begin{aligned}\frac{\partial}{\partial \alpha_n}\mathbf{r} &= 2\mathbf{A}^H \Re\left\{-j(\mathbf{A}\bar{\mathbf{s}}) \odot \left(\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}})\right)^*\right\} \\ &= 2\mathbf{A}^H \Im\left\{(\mathbf{A}\bar{\mathbf{s}}) \odot \left(\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}})\right)^*\right\}.\end{aligned} \tag{34}$$

At this point we can substitute (34) back into (29), and then substitute that result into (27). Beginning with the sidelobe derivative term of (27) this substitution process yields

$$\begin{aligned}\mathbf{w}_{\text{SL}}^T \left(\frac{\partial |\mathbf{r}|^p}{\partial \alpha_n}\right) &= p\,\mathbf{w}_{\text{SL}}^T \left(|\mathbf{r}|^{(p-2)} \odot \Re\left\{\left(\frac{\partial}{\partial \alpha_n}\mathbf{r}\right) \odot \mathbf{r}^*\right\}\right) \\ &= 2p\,\mathbf{w}_{\text{SL}}^T \left(|\mathbf{r}|^{(p-2)} \odot \Re\left\{\left(\mathbf{A}^H \Im\left\{(\mathbf{A}\bar{\mathbf{s}}) \odot \left(\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}})\right)^*\right\}\right) \odot \mathbf{r}^*\right\}\right) \\ &= 2p\,\Re\left\{\left(\mathbf{w}_{\text{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r}^*\right)^T \mathbf{A}^H \Im\left\{(\mathbf{A}\bar{\mathbf{s}}) \odot \left(\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}})\right)^*\right\}\right\}\end{aligned} \tag{35}$$

with the last step taking advantage of the fact that $\mathbf{w}_{\text{SL}}$ and $|\mathbf{r}|^{(p-2)}$ are already real-valued and so can be subsumed inside the real operator. Because the term $(\mathbf{w}_{\text{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r}^*)$ in (35) is conjugate-symmetric about zero-delay, the DFT response resulting from $(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r}^*)^T \mathbf{A}^H$ is also necessarily real-valued. Moreover, with the vector $\Im\{(\mathbf{A}\bar{\mathbf{s}}) \odot (\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}}))^*\}$ likewise being real-valued, the subsequent inner product of these terms is itself real-valued, thereby making application of the real operation in (35) unnecessary. In fact, since the DFT response term above is real-valued, the imaginary operator on the latter term of (35) can subsume the entire equation. Thus we can write

$$\begin{aligned}
&\mathbf{w}_{\mathrm{SL}}^T \left(\frac{\partial |\mathbf{r}|^p}{\partial \alpha_n} \right) \\
&\quad = 2p\,\Im\left\{ \left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r}^* \right)^T \mathbf{A}^H \left((\mathbf{A}\bar{\mathbf{s}}) \odot \left(\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}}) \right)^* \right) \right\} \\
&\quad = 2p\,\Im\left\{ \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \right]^T \left((\mathbf{A}\bar{\mathbf{s}}) \odot \left(\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}}) \right)^* \right) \right\} \\
&\quad = 2p\,\Im\left\{ \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right]^T \left(\left(\mathbf{A}(\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}}) \right)^* \right) \right\} \\
&\quad = 2p\,\Im\left\{ \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right]^T \mathbf{A}^* (\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}})^* \right\},
\end{aligned} \tag{36}$$

sequentially making use of the real-valued nature of this DFT response to rearrange its terms (and $\mathbf{A}^H\,\mathbf{r}^* = \mathbf{A}\,\mathbf{r}$), associating the $\mathbf{A}\bar{\mathbf{s}}$ component with the result, and then bringing the complex conjugation inside the final term. The outcome of (36) can be further manipulated as

$$\begin{aligned}
&\mathbf{w}_{\mathrm{SL}}^T \left(\frac{\partial |\mathbf{r}|^p}{\partial \alpha_n} \right) \\
&\quad = 2p\,\Im\left\{ \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right]^T \mathbf{A}^* (\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}})^* \right\} \\
&\quad = 2p\,\Im\left\{ \left(\mathbf{A}^H \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right)^T (\bar{\mathbf{b}}_n \odot \bar{\mathbf{s}}^*) \right\} \\
&\quad = 2p\,\Im\left\{ \left(\left(\mathbf{A}^H \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right) \odot \bar{\mathbf{s}}^* \right)^T \bar{\mathbf{b}}_n \right\} \\
&\quad = 2p\,\bar{\mathbf{b}}_n^T\,\Im\left\{ \bar{\mathbf{s}}^* \odot \left(\mathbf{A}^H \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{SL}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right) \right\}
\end{aligned} \tag{37}$$

based on attributes of the transpose operation. Because they are identical aside from the selection weighting, the mainlobe derivative term from (27) can likewise employ (37) to write

$$\begin{aligned}
&\mathbf{w}_{\mathrm{ML}}^T \left(\frac{\partial |\mathbf{r}|^p}{\partial \alpha_n} \right) \\
&\quad = 2p\,\bar{\mathbf{b}}_n^T\,\Im\left\{ \bar{\mathbf{s}}^* \odot \left(\mathbf{A}^H \left[\mathbf{A}\left(\mathbf{w}_{\mathrm{ML}} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right) \right\}.
\end{aligned} \tag{38}$$

Inserting (37) and (38) into (27) therefore realizes the derivative

$$\begin{aligned}
&\frac{\partial J_p}{\partial \alpha_n} = 4J_p \times \\
&\left[\left(\bar{\mathbf{b}}_n^T\,\Im\left\{ \bar{\mathbf{s}}^* \odot \left(\mathbf{A}^H \left[\mathbf{A}\left(\frac{\mathbf{w}_{\mathrm{SL}}}{\mathbf{w}_{\mathrm{SL}}^T |\mathbf{r}|^p} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right) \right\} \right) \right. \\
&\left. - \left(\bar{\mathbf{b}}_n^T\,\Im\left\{ \bar{\mathbf{s}}^* \odot \left(\mathbf{A}^H \left[\mathbf{A}\left(\frac{\mathbf{w}_{\mathrm{ML}}}{\mathbf{w}_{\mathrm{ML}}^T |\mathbf{r}|^p} \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right) \right\} \right) \right] \\
&\quad = 4J_p\,\bar{\mathbf{b}}_n^T \times \\
&\Im\left\{ \bar{\mathbf{s}}^* \odot \left(\mathbf{A}^H \left[\mathbf{A}\left(\left[\frac{\mathbf{w}_{\mathrm{SL}}}{\mathbf{w}_{\mathrm{SL}}^T |\mathbf{r}|^p} - \frac{\mathbf{w}_{\mathrm{ML}}}{\mathbf{w}_{\mathrm{ML}}^T |\mathbf{r}|^p} \right] \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right) \right\}
\end{aligned} \tag{39}$$

by linearity. Note that, aside from the particular basis function $\mathbf{b}_n$, the components of (39) are the same regardless of PCFM code index n. Consequently, the scalar partial derivative of (39) can be easily generalized by collecting the basis functions into $\bar{\mathbf{B}}$ per (18) to form the final N-length gradient vector

$$\begin{aligned}
&\nabla_{\mathbf{x}} J_p = 4J_p\,\bar{\mathbf{B}}^T \times \\
&\Im\left\{ \bar{\mathbf{s}}^* \odot \left(\mathbf{A}^H \left[\mathbf{A}\left(\left[\frac{\mathbf{w}_{\mathrm{SL}}}{\mathbf{w}_{\mathrm{SL}}^T |\mathbf{r}|^p} - \frac{\mathbf{w}_{\mathrm{ML}}}{\mathbf{w}_{\mathrm{ML}}^T |\mathbf{r}|^p} \right] \odot |\mathbf{r}|^{(p-2)} \odot \mathbf{r} \right) \odot (\mathbf{A}\bar{\mathbf{s}}) \right] \right) \right\}
\end{aligned} \tag{40}$$

that is reproduced in (17).

TABLE V: ISL FOR M = 1024, BT = 64

Optimized GISL p-Norm	Number of Independent Parameters ($N = BT \times L$) 64	128	192	256	320	384	448	512
2	–24.0	–31.6	–35.8	–42.7	–47.6	–49.8	–51.3	–52.0
3	–27.8	–31.8	–37.3	–42.2	–46.8	–48.8	–50.2	–51.0
4	–27.2	–32.7	–36.8	–41.4	–45.6	–47.6	–49.3	–50.0
5	–26.7	–32.4	–36.2	–40.8	–44.6	–46.7	–48.2	–49.1
6	–26.3	–32.1	–35.8	–40.3	–44.0	–46.0	–47.7	–48.5
7	–26.9	–31.8	–35.4	–39.9	–43.4	–45.4	–47.1	–47.9
8	–27.6	–31.6	–35.1	–39.5	–42.9	–45.0	–46.7	–47.6
9	–27.4	–31.4	–34.8	–39.3	–42.7	–44.6	–46.5	–47.2
10	–27.3	–31.3	–34.6	–39.1	–42.5	–44.3	–46.3	–47.0
11	–27.2	–31.1	–34.4	–38.9	–42.3	–44.1	–45.9	–46.8
12	–27.0	–31.0	–34.2	–38.8	–42.1	–43.8	–45.6	–46.5
13	–26.9	–30.9	–34.1	–38.7	–41.8	–43.5	–45.4	–46.0
14	–26.8	–30.8	–34.0	–38.5	–41.7	–43.4	–45.0	–45.6
15	–26.7	–30.7	–33.9	–38.4	–41.6	–43.2	–44.8	–45.2
16	–26.5	–30.7	–33.8	–38.5	–41.5	–43.1	–44.6	–44.9
17	–26.5	–30.6	–33.7	–38.3	–41.4	–43.0	–44.4	–44.7
18	–26.4	–30.5	–33.6	–38.3	–41.3	–42.8	–44.2	–44.4
19	–26.4	–30.5	–33.5	–38.2	–41.3	–42.9	–43.9	–44.1
20	–26.3	–30.4	–33.5	–38.2	–41.2	–42.7	–43.8	–44.1

TABLE VI: PSL FOR M = 1024, BT = 64

Optimized GISL p-Norm	Number of Independent Parameters ($N = BT \times L$) 64	128	192	256	320	384	448	512
2	–20.8	–32.2	–35.8	–51.8	–52.0	–52.1	–52.9	–53.9
3	–33.6	–37.1	–48.9	–53.7	–55.8	–57.4	–59.3	–60.2
4	–36.1	–46.1	–50.8	–54.2	–57.0	–58.3	–60.2	–60.2
5	–37.3	–47.1	–51.7	–55.4	–57.2	–58.7	–60.2	–60.2
6	–38.1	–47.7	–52.0	–55.7	–57.5	–59.1	–60.2	–60.2
7	–40.5	–48.2	–52.1	–55.8	–58.2	–59.6	–60.2	–60.2
8	–42.7	–48.5	–52.3	–56.2	–58.6	–60.0	–60.2	–60.2
9	–43.0	–48.8	–52.4	–56.3	–58.6	–60.2	–60.2	–60.2
10	–43.2	–49.0	–52.5	–56.5	–58.6	–60.2	–60.2	–60.2
11	–43.5	–49.1	–52.6	–56.5	–58.6	–60.2	–60.2	–60.2
12	–43.6	–49.3	–52.7	–56.5	–58.6	–60.2	–60.2	–60.2
13	–43.8	–49.4	–52.7	–56.5	–58.6	–60.2	–60.2	–60.2
14	–44.0	–49.5	–52.7	–56.5	–58.6	–60.2	–60.2	–60.2
15	–44.1	–49.6	–52.7	–56.4	–58.7	–60.2	–60.2	–60.2
16	–44.2	–49.7	–52.7	–56.5	–58.7	–60.2	–60.2	–60.2
17	–44.3	–49.8	–52.7	–56.5	–58.7	–60.2	–60.2	–60.2
18	–44.4	–49.8	–52.8	–56.5	–58.7	–60.2	–60.2	–60.2
19	–44.5	–49.9	–52.8	–56.5	–58.8	–60.2	–60.2	–60.2
20	–44.5	–49.9	–52.8	–56.5	–58.8	–60.2	–60.2	–60.2

TABLE VII:
ISL FOR *M* = 1024, *BT* = 128

Optimized GISL *p*-Norm	# Independent Parameters ($N = BT \times L$): 128	256	384	512
2	–24.3	–32.4	–38.6	–42.8
3	–29.5	–34.8	–38.0	–41.1
4	–31.2	–34.2	–36.5	–41.9
5	–30.2	–33.5	–37.6	–41.4
6	–30.7	–33.2	–37.3	–41.0
7	–30.4	–33.8	–37.1	–40.6
8	–30.2	–33.7	–36.9	–40.3
9	–30.0	–34.1	–36.8	–40.3
10	–29.9	–34.0	–36.6	–39.5
11	–29.7	–33.9	–36.5	–39.3
12	–29.5	–33.8	–36.4	–39.7
13	–29.3	–33.7	–36.3	–39.5
14	–29.2	–33.7	–36.2	–-37.4
15	–29.0	–33.6	–36.1	–39.2
16	–28.9	–33.5	–36.0	–39.1
17	–28.8	–33.5	–35.9	–38.9
18	–28.7	–33.4	–33.4	–38.8
19	–28.6	–33.4	–33.2	–38.8
20	–28.5	–33.3	–33.2	–38.8

TABLE VIII:
PSL FOR *M* = 1024, *BT* = 128

Optimized GISL *p*-Norm	# Independent Parameters ($N = BT \times L$): 128	256	384	512
2	–20.9	–33.0	–45.2	–46.5
3	–35.7	–45.9	–46.8	–50.8
4	–43.2	–46.7	–49.2	–58.6
5	–44.7	–47.3	–57.7	–59.1
6	–46.8	–49.1	–57.7	–59.6
7	–47.3	–52.9	–57.7	–60.0
8	–47.6	–53.3	–57.7	–60.2
9	–47.8	–55.7	–57.8	–60.2
10	–48.0	–55.8	–57.8	–59.4
11	–48.2	–55.9	–57.9	–59.5
12	–48.3	–55.9	–57.9	–60.2
13	–48.4	–55.9	–58.0	–60.2
14	–48.5	–55.9	–58.0	–59.1
15	–48.6	–55.9	–58.1	–60.2
16	–48.6	–56.0	–58.1	–60.2
17	–48.7	–56.0	–58.2	–60.2
18	–48.7	–56.0	–55.3	–60.2
19	–48.8	–56.0	–55.4	–60.2
20	–48.8	–56.0	–55.5	–60.2

REFERENCES

[1] Levanon, N., Mozeson, E., *Radar Signals*, John Wiley & Sons, Inc., Hoboken, NJ, 2004.

[2] Wicks, M., Mokole, E., Blunt, S.D., Amuso, V., Schneible, R., eds., *Principles of Waveform Diversity & Design*, SciTech Publishing, 2010.

[3] Pillai, S., Li, K.Y., Selesnick, I., Himed, B., eds., *Waveform Diversity: Theory & Applications*, McGraw-Hill, 2011.

[4] Gini, F., De Maio, A., Patton, L.K., *Waveform Design and Diversity for Advanced Radar Systems*, IET Press, 2012.

[5] Blunt, S.D., Mokole, E.L., "An overview of radar waveform diversity," *IEEE AESS Systems Mag.*, vol. 31, no. 11, pp. 2-42, Nov. 2016.

[6] Klauder, J.R., Price, A.C., Darlington, S., Albersheim, W.J., "The theory and design of chirp radars," *The Bell System Technical Journal*, vol. XXXIX, no. 4, pp. 745-808, July 1960.

[7] Cook, C.E., "Pulse Compression -key to more efficient radar transmission," Proc. IRE, vol. 48, no. 3, pp. 310–316, Mar 1960.

[8] Fowle, E., "The design of FM pulse compression signals," *IEEE Trans. Information Theory*, vol. 10, no. 1, pp. 61-67, Jan. 1964.

[9] Cook, C.E., "A class of nonlinear FM pulse compression signals," *Proceedings of the IEEE*, vol. 52, no. 11, pp. 1369-1371, Nov. 1964.

[10] Caputi, W.J., "Stretch: a time-transformation technique," *IEEE Trans. Aerospace & Electronic Systems*, vol. AES-7, no. 2, pp. 269-278, Mar. 1971.

[11] L. Harnett, D. Hemmingsen, P. McCormick, S.D. Blunt, C. Allen, A. Martone, K. Sherbondy, D. Wikner, "Optimal and adaptive mismatch filtering for stretch processing," *IEEE Radar Conf.*, Oklahoma City, OK, Apr. 2018.

[12] S. Alphonse, G.A. Williamson, "Evaluation of a class of NLFM radar signals," *EURASIP J. Advances in Signal Processing*, vol. 2019, no. 62, Dec. 2019.

[13] D.M. Hemmingsen, P.M. McCormick, S.D. Blunt, C. Allen, A. Martone, K. Sherbondy, D. Wikner, "Waveform-diverse stretch processing," *IEEE Radar Conf.*, Oklahoma City, OK, Apr. 2018.

[14] F. Rodriguez-Morales, S. Gogineni, C.J. Leuschen, J.D. Paden, J. Li, C.S. Lewis, B. Panzer, D. Gomez-Garcia Alvestegui, A. Patel, K. Byers, R. Crowe, K. Player, R.D. Hale, E.J. Arnold, L. Smith, C.M. Gifford, D. Braaten, C. Panton, "Advanced multifrequency radar instrumentation for polar research," *IEEE Trans. Geoscience & Remote Sensing*, vol. 52, no. 5, pp. 2824-2842, May 2014.

[15] J.-B. Yan, D. Gomez-Garcia Alvestegui, J.W. McDaniel, Y. Li, S. Gogineni, F. Rodriguez-Morales, J. Brozena, C.J. Leuschen, "Ultrawideband FMCW radar for airborne measurements of snow over sea ice and land," *IEEE Trans. Geoscience & Remote Sensing*, vol. 55, no. 2, pp. 834-843, Feb. 2017.

[16] "Arbitrary waveform generator, AWG", *Electronics Notes*, www.electronics-notes.com/articles/test-methods/signal-generators/arbitrary- waveform-generator-awg.php

[17] W. Tuttlebee, eds., *Software Defined Radio: Enabling Technologies*, John Wiley & Sons, West Sussex, England, 2002.

[18] Xilinx, "Zynq Ultrascale+ RFSoC data sheet: overview," https://www.xilinx.com/support/documentation/data_sheets/ds889-zynq-usp-rfsoc-overview.pdf, Aug. 2019.

[19] Blunt, S.D., Cook, M., Perrins, E., de Graaf, J., "CPM-based radar waveforms for efficiently bandlimiting a transmitted spectrum," *IEEE Radar Conf.*, Pasadena, CA, May 2009.

[20] Blunt, S.D., Cook, M., Jakabosky, J., Graaf, J.D., Perrins, E., "Polyphase-coded FM (PCFM) radar waveforms, part I: implementation," *IEEE Trans. Aerospace & Electronic Systems*, vol. 50, no. 3, pp. 2218–2229, July 2014.

[21] Anderson, J.B., Aulin, T., Sunberg, C.-E. *Digital Phase Modulation*. Plenum Press: New York, 1986.

[22] Jakabosky, J., Anglin, P., Cook, M., Blunt, S.D., Stiles, J., "Non-linear FM waveform design using marginal Fisher's information within the CPM framework," *IEEE Radar Conf.*, Kansas City, MO, May 2011.

[23] Blunt, S.D., Jakabosky, J., Cook, M., Stiles, J., Seguin, S., Mokole, E.L., "Polyphase-coded FM (PCFM) radar waveforms, part II: optimization," *IEEE Trans. Aerospace & Electronic Systems*, vol. 50, no. 3, pp. 2230–2241, July 2014.

[24] Jakabosky, J., Blunt, S.D., Cook, M.R., Stiles, J., Seguin, S.A., “Transmitter-in-the-loop optimization of physical radar emissions,” *IEEE Radar Conf.*, Atlanta, GA, May 2012.

[25] Griffiths, H., Blunt, S., Cohen, L., Savy, L., “Challenge problems in spectrum engineering and waveform diversity,” *IEEE Radar Conf.*, Ottawa, Canada, Apr./May 2013.

[26] Griffiths, H., Cohen, L., Watts, S., Mokole, E., Baker, C., Wicks, M., Blunt, S., “Radar spectrum engineering and management: technical and regulatory issues,” *Proceedings of the IEEE*, vol. 103, no. 1, pp. 85–102, Jan. 2015.

[27] Latham, C., Egbert, A., Baylis, C., Cohen, L., Marks II, R.J., “Joint radar amplifier circuit and waveform optimization for ambiguity function, power-added efficiency, and spectral compliance,” *IEEE Trans. Aerospace & Electronic Systems*, vol. 55, no. 3, pp. 1190-1199, June 2019.

[28] Simmons, J.A., Gaudette, J.E., “Biosonar echo processing by frequency-modulated bats,” *IET Radar, Sonar & Navigation*, vol. 6, no. 6, pp. 556-565, July 2012.

[29] Au, W.W.L, Martin, S.W., “Why dolphin biosonar performs so well in spite of mediocre ‘equipment’,” *IET Radar, Sonar & Navigation*, vol. 6, no. 6, pp. 566-575, July 2012.

[30] Balleri, A., Griffiths, H., Baker, C., “Biologically-inspired radar and sonar: lessons from nature,” *SciTech Publishing*, 2017.

[31] Mohr, C.A., McCormick, P.M., Blunt, S.D., “Optimized complementary waveform subsets within an FM noise radar CPI,” *IEEE Radar Conf.*, Oklahoma City, OK, Apr. 2018.

[32] Sahin, C., Jakabosky, J., McCormick, P., Metcalf, J., Blunt, S., "A novel approach for embedding communication symbols into physical radar waveforms," *IEEE Radar Conf.*, Seattle, WA, May 2017.

[33] Mohr, C.A., Blunt, S.D., “Analytical spectrum representation for physical waveform optimization requiring extreme fidelity,” *IEEE Radar Conf.*, Boston, MA, Apr. 2019.

[34] Owen, J., Mohr, C., Blunt, S.D., Gallagher, K., “Nonlinear radar via intermodulation of jointly optimized FM noise waveform pairs,” *IEEE Radar Conf.*, Boston, MA, Apr. 2019.

[35] Tan, P.S., Jakabosky, J., Stiles, J.M., Blunt, S.D., "Higher-order implementations of polyphase-coded FM radar waveforms," *IEEE Trans. Aerospace & Electronic Systems*, vol. 55, no. 6, pp. 2850-2870, Dec. 2019.

[36] Jakabosky, J., Blunt, S.D., Himed, B., “Optimization of ‘over-coded’ radar waveforms,” *IEEE Radar Conf.*, Cincinnati, OH, May 2014.

[37] Skinner, B.J., Donohoe, J.P., Ingels, F.M., “Gradient descent method for designing optimum radar signals,” *IEEE Intl. Radar Conf.*, Alexandria, VA, May 1995.

[38] J.M. Baden, “Efficient optimization of the merit factor of long binary sequences,” *IEEE Trans. Information Theory*, vol. 57, no. 12, pp. 8084-8094, Dec. 2011.

[39] Patton, L.K., Rigling, B.D., “Phase retrieval for radar waveform optimization,” *IEEE Trans. Aerospace & Electronic Systems*, vol. 48, no. 4, pp. 3287-3302, Oct. 2012.

[40] Li, L., Yang, H., Cui, G., Kong, L., Yang, X., “Efficient phase-modulated waveform design for active sensing systems,” *IEEE Radar Conf.*, Cincinnati, OH, May 2014.

[41] Baden, J.M., Davis, M.S., Schmeider, L., “Energy efficient gradient calculations for binary and polyphase sequences,” *IEEE Intl. Radar Conf.*, Washington, DC, May 2015.

[42] Tan, U., Adnet, C., Rabaste, O., Arlery, F., Ovarlez, J.-P., Guyvarch, J.-P., “Phase code optimization for coherent MIMO radar via a gradient descent,” *IEEE Radar Conf.*, Philadelphia, PA, May 2016.

[43] Zhao, D., Wei, Y., Liu, Y., “Spectrum optimization via FFT-based conjugate gradient method for unimodular sequence design,” *Elsevier Signal Processing*, vol. 142, pp. 354-365, Jan. 2018.

[44] Baden, J.M., O’Donnell B., Schmieder, L., “Multiobjective sequence design via gradient descent methods,” *IEEE Trans. Aerospace & Electronic Systems*, vol. 54, no. 3, pp. 1237-1252, June 2018.

[45] Mohr, C.A., McCormick, P.M., Blunt, S.D., Mott, C., “Spectrally-efficient FM noise radar waveforms optimized in the logarithmic domain,” *IEEE Radar Conf.*, Oklahoma City, OK, Apr. 2018.

[46] Mohr, C.A., Blunt, S.D., “FM noise waveforms optimized according to a temporal template error (TTE) metric,” *IEEE Radar Conf.*, Boston, MA, Apr. 2019.

[47] Mohr, C.A., Blunt, S.D., “Design and generation of stochastically defined, pulsed FM noise waveforms,” *Intl. Radar Conf.*, Toulon, France, Sept. 2019.

[48] Blunt, S.D., Jakabosky, J.K., Mohr, C.A., McCormick, P.M., Owen, J.W., Ravenscroft, B., Sahin, C., Zook, G.D., Jones, C.C., Metcalf, J.G., Higgins, T., "Principles & applications of random FM radar waveform design," to appear in *IEEE AES Systems Magazine*.

[49] McCormick, P.M., Blunt, S.D., “Nonlinear conjugate gradient optimization of polyphase-coded FM radar waveforms,” *IEEE Radar Conf.*, Seattle, WA, May 2017.

[50] Raab, F.H., Asbeck, P., Cripps, S., Kenington, P.B., Popovic, Z.B., Pothecary, N., Sevic, J.F., Sokal, N.O., “Power amplifiers and transmitters for RF and microwave,” *IEEE Trans. Microwave Theory & Techniques*, vol. 50, no. 3, pp. 814-826, Mar. 2002.

[51] McCormick, P.M., Blunt, S.D., “Gradient-based coded-FM waveform design using Legendre polynomials,” *IET Intl. Conf. Radar Systems*, Belfast, UK, Oct. 2017.

[52] C. W. Wilcox, "The synthesis problem for radar ambiguity functions," Math. Res. Center, U.S. Army, Univ. of Wisconsin Rept. 157, April 1960.

[53] S. Sussman, "Least-square synthesis of radar ambiguity functions," Information Theory, IRE Transactions on, vol. 8, no. 3, pp. 246-254, April 1962.

[54] I. Gladkova, D. Chebanov, “On a new extension of Wilcox’s method,” *WSEAS Intl. Conf. Applied Mathematics*, Apr. 2004.

[55] B. Ravenscroft, J.W. Owen, J. Jakabosky, S.D. Blunt, A.F. Martone, K.D. Sherbondy, "Experimental demonstration and analysis of cognitive spectrum sensing & notching," *IET Radar, Sonar & Navigation*, vol. 12, no. 12, pp. 1466-1475, Dec. 2018.

[56] Song, J., Babu, P., Palomar, D.P., “Sequence design to minimize the weighted integrated and peak sidelobe levels,” *IEEE Trans. Signal Process*ing, vol. 64, no. 8, pp. 2051–2064, April 2016.

[57] M.A. Kerahroodi, A. Aubry, A. De Maio, M.M. Naghsh, M. Modarres-Hashemi, “A coordinate-descent framework to design low PSL/ISL sequences,” *IEEE Trans. Signal Processing*, vol. 65, no. 22, pp. 5942-5956, Nov. 2017.

[58] Boyd, S., Vandenberghe, L., *Convex Optimization*, Cambridge University Press, 2004.

[59] X. Du, A. Aubry, A. De Maio, G. Cui, “Hidden convexity in robust waveform and receive filter bank optimization under range unambiguous clutter,” *IEEE Signal Processing Letters*, vol. 27, pp. 885-889, May 2020.

[60] Hager, W.W., Zhang, H., “A survey of nonlinear conjugate gradient methods,” *Pacific Journal of Optimization*, vol. 2, no. 1, pp. 35-58, 2006.

[61] Ghadimi, E., Feyzmahdavian, R., Johansson, M., “Global convergence of the heavy-ball method for convex optimization,” *European Control Conf.*, Linz, Austria, July 2015.

[62] Nocedal, J., Wright, S., *Numerical Optimization*, Springer Science & Business Media, 2006.

[63] M.A. Richards, J.A. Scheer, W.A. Holm, *Principles of Modern Radar: Basic Principles*, SciTech Publishing, 2010.

[64] Johnston, J.A., Fairhead, A.C., “Waveform design and doppler sensitivity analysis for nonlinear FM chirp pulses,” *IEE Proc. F - Communications, Radar and Signal Processing*, vol. 133, no. 2, pp. 163-175, Apr. 1986.

[65] L. Wei, R.A. Kennedy, T.A. Lamahewa, “An optimal basis of band-limited functions for signal analysis and design,” *IEEE Trans. Signal Processing*, vol. 58, no. 11, pp. 5744-5755, Nov. 2010.

[66] Collins, T., Atkins, P., “Nonlinear frequency modulation chirps for active sonar,” *IEE Proc. Radar, Sonar & Navigation*, vol. 146, no. 6, pp. 312-316, Dec. 1999.

[67] F.M. Ghannouchi, O. Hammi, “Behavioral modeling and predistortion,” *IEEE Microwave Magazine*, vol. 10, no. 7, pp. 52-64, Dec. 2009.