Waveform Optimization for Multistatic Radar Imaging using Mutual Information

Zacharie Idriss,*†*Student Member*, Raghu G. Raj,†*Senior Member*, Ram M. Narayanan,**Life Fellow*,
†U.S. Naval Research Laboratory, Radar Division, Washington, DC 20375
*School of Electrical and Computer Engineering, The Pennsylvania State University, University Park, PA 16802

***Abstract*—This paper addresses the problem of radar waveform design for imaging targets in a cluttered environment. A multistatic radar scenario is considered for sparse and random sensor positions. The target impulse response (TIR) is modeled using the simulated frequency response of a buried metallic mine, and the clutter is modeled using the compound Gaussian (CG) distribution. We explore novel waveform design techniques with respect to the Mutual Information (MI) criterion based on CG clutter. The first method presents a waveform that exploits the CG distribution of the scene reflectivity function when projected onto a sparse basis. This is compared to the second method, called the target-specific approach that uses knowledge of the target and clutter frequency response to optimize a matched illumination waveform. In both cases, the Taguchi and particle swarm optimization (PSO) solvers are employed for MI based waveform design optimization. To validate and compare the effectiveness of the optimized waveforms, the resulting scene reflectivity function is estimated using the sparsity-driven regularization radar imaging method. Our experimental results demonstrate that both waveform optimization techniques result in significantly better image reconstruction performance than the traditional LFM waveform; and that the target-specific approach additionally suppresses clutter information in the scene.**

***Index Terms*—Waveform Optimization, Mutual Information, Multistatic Imaging, Compound Gaussian, Taguchi algorithm, Particle Swarm Optimization, Sparsity-Driven Regularized Radar Imaging**

NOMENCLATURE

TIR, CIR	Target/Clutter Impulse Response
CG	Compound Gaussian
MI	Mutual Information
PSO	Particle Swarm Optimization
LFM	Linearly Frequency Modulated
PSD	Power Spectral Density
TS, non-Ts	Target Specific/non-Target Specific
TFR, CFR	Target/Clutter Frequency Response
$\mathbf{g}$	Vectorized Scene
$\mathbf{c}$	Wavelet Coefficient Vector
Σ_w	Gaussian Covariance Matrix
$\mathbf{\Phi}$	Wavelet Dictionary
$\mathbf{\Psi}$	Radon transform of Wavelet Dictionary
$\mathbf{x}$	Transmit Waveform
$p_{\theta_i}(t)$	Impulse Response of Scene
z	Compound Gaussian Multiplier
$\mathbf{X}$	Transmit Convolution Matrix
$\mathbf{y}$	Received Signal Vector
$S_{g/c}(f)$	Target/Clutter Power Spectral Density

I. INTRODUCTION

High range resolution (HRR) imaging techniques such as strip-mode/spotlight-mode synthetic aperture radar (SAR), inverse SAR (ISAR), interferometric SAR (IFSAR), multistatic imaging and others provide inherent advantages over optical methods. The advantages are evident in heavily cluttered environments such as heavy vegetation (forests, or long grass) and ground penetrating applications such as nondestructive subsurface imaging [1]. In this paper, waveform optimization techniques for *multistatic* radar imaging are presented. For scene surveillance, multistatic radar provides illumination from multiple different transmit receive-pairs, allowing for improved coverage of potential targets and potentially useful information for target recognition.

In mine detection applications, radar returns often contain unwanted and waveform dependent reflections from the environment, in turn restricting the radars ability to detect and image targets. Waveform design seeks to optimally probe the environment to maximize target return information at the receiver to mitigate these effects. This channel can be modeled as a linear time-invariant (LTI) system as the effects of clutter on the waveform are linear [2]. Therefore, by viewing the surrounding environment as an LTI system, the problem is viewed as solving for the upper bound of the channel capacity and the change in information between the input and output is measured using information-theoretic metrics. Therefore, extracting target information is posed as maximizing the mutual information between the quantity of interest and received waveform conditioned upon the input; a transmit signal.

Mutual information (MI) for radar waveform design and imaging has received a lot of attention [3]–[9]. The authors of [10] use MI for multiple-input multiple-output (MIMO) radar waveform design. They show that maximizing the MI at the receiver for a random TIR is equivalent to minimizing the minimum mean square error (MMSE) in estimating the TIR. In [11], MI is applied to waveform optimization where the target ensemble is modeled as a Gaussian mixture model (GMM). The concept of dual-MI is used in [12] to maximize and minimize the TIR and clutter impulse response (CIR) returns respectively. The optimization of mutual information has also seen use in the coexistence of radar and communication systems. More recently, the mutual information is derived for a target ensemble modeled as a Gaussian mixture in the context of matched illumination [13]. However, it is still unclear whether maximizing mutual information at the receiver is an

optimal solution to waveform design when compared to other metrics. The authors of [14] aim to answer this question by providing a relationship between mutual information, signal-to-noise ratio (SNR), and the Kullback-Leibler divergence.

In this paper, two methods of waveform optimization for HRR imaging using MI are presented and named the non-target specific (non-TS), and target specific (TS) waveform optimization procedures.

The first method presents a procedure that exploits the CG distribution of the radar image as a prior for modelling the scene response. The MI between the received signal and the CG image vector is derived to form the optimal transmit waveform. The first source of novelty in this method is the analytical characterization of the mutual information (MI) for a vector described by the compound Gaussian distribution (CG). In [15], [16], it was shown empirically that with a sparse wavelet representation of natural scenes, a CG distribution is a good fit to model the image vector. Therefore, by defining the scene reflectivity function as the product of the inverse of a 2-dimensional (2D) wavelet basis and corresponding wavelet coefficients, the CG distribution models these coefficients. The second source of novelty lies in using the CG statistics of the scene to the problem of waveform optimization. Here, the work makes a systematic application of the developed tools to imaging targets buried underground.

A third source of novelty is the developed theory for both target specific (TS) and non-TS waveform optimization for radar imaging. The authors provide a rigorous framework by which the waveforms are optimized for the case of multistatic imaging. These algorithms are then tested against a linearly frequency modulated waveform.

The second waveform optimization procedure extends the method to specifically maximize target information at the receiver using the theory of matched illumination. In the theory of matched illumination, the PSD of the transmit waveform is matched to the target response to maximize target returns at the receiver. This is achieved by weighting the available power toward the target frequencies. The reduced power at the remaining frequencies aims to suppress clutter effects on the channel. In [17]–[19], the concept of matched illumination for the maximization of the signal-to-interference-and-noise-ratio (SINR) is considered for colored and waveform dependent clutter/interference. In [20], the theory of matched illumination is used to optimize a full-polarization waveform and corresponding receiver impulse response for non-zero clutter.

The power spectral density (PSD) is used to differentiate between clutter and target returns due to its ability to describe how power is distributed over frequency. Therefore, the frequency response extracts distinguishing frequency information to characterize the each response.

Since the scattering effects of targets and clutter operate linearly on the channel, designing a PSD that maximizes the target information at the receiver is synonymous with maximizing the Shannon-Hartley capacity. We use the target frequency response (TFR) and the clutter frequency response (CFR) to do this. The CG distribution provides a model for the overall measured scene reflectivity function and serves as a rich model for high resolution clutter. The coherent detection of targets in CG clutter has been explored in [21], which models the scattering of electromagnetic waves interacting with the scene. In [22], [23], it is shown that clutter scenes generally possess this structure. Therefore, the clutter is modeled as being CG. The mine is modeled using the simulated frequency response of a metallic mine buried 2 cm beneath the surface.

The assumption of a known TFR implies that an estimation of the target impulse response (TIR) exists. Target estimation is often done using a Kalman filter [24]. Should the assumption fail, signal mismatch may occur. The authors of [25] aim to tackle this issue and provide a method for estimating the target impulse response.

A novel method of applying spectrally matched illumination to radar imaging is presented. To accomplish this, the MI cost functions are optimized using Taguchi [26] and particle swarm optimization (PSO) techniques [27]. The MI is conditioned upon the unknown transmit waveform which is found using the optimization algorithms. A total power constraint is placed upon the waveforms. This ensures that the power of the optimized waveform coincides with the transmit pulse design specifications and is fairly compared. Once the waveforms are found, a constant modulus is enforced using the method of projections onto convex sets (POCS) [28]. With a constant modulus, the system operates at peak power efficiency at all times.

To validate our analysis, the optimized waveforms are transmitted using the multistatic radar scenario. We estimate the sparse image vector using sparsity-driven regularization radar imaging, and the reconstruction performance is measured using the following metrics: structural similarity (SSIM) index, SNR, peak-SNR (PSNR), and mean square error (MSE).

In Section II, we present the multistatic imaging scenario. In Section III, we present the scene reflectivity function. Section IV presents the mutual information. Section V presents the optimal waveforms. Section VI presents the algorithms. Section VII presents the simulation parameters. Section VIII presents the results. Our conclusions are in Section IX.

II. System Setup

The radar system model is first defined for the case of a single monostatic radar transmitting in the environment. The defined model is then extended to the multistatic case of M sensors for radar imaging.

A. Radar Model

A single monostatic radar is modeled as transmitting a signal, $x(t) \in \mathbb{R}^N$, along the range direction, $\mathbf{r}$, that cuts through the center of the scene at a positive angle, θ, from the cross-range axis, $\mathbf{u}$.

An energy constraint $\mathbb{E} = T\rho$ is placed on the transmit waveform such that

$$\|\mathbf{x}\|^2 \leq T\rho, \tag{1}$$

where ρ is the average power per sample and T is the duration of the signal. The complex target reflectivity function, is defined as $g(u, v) \in \mathbb{C}$. With this formulation, the return at

time, t, from an image point (u_0, v_0) at a range, r_0, is defined as

$$y_{r_0}(t) = g(u_0, v_0)x\left(t - \frac{2r_0}{\nu_p}\right), \tag{2}$$

where $2r_0/\nu_p$ is the two-way time delay of the waveform at a range, r_0 [29], and ν_p is the speed of wave propagation.

From (2), the reflectivity $g(u_0, v_0)$ is defined as the integral over the radar returns from the set of all points along the line perpendicular to the range location, r_0, and written as

$$y_{r_0}(t) = R_{\theta_i, r_0}\left\{g(u, v)\right\} x\left(t - \frac{2r_0}{\nu_p}\right). \tag{3}$$

$R_{\theta_i, r_0}\big\{g(u, v)\big\} : \mathbb{R}^N \rightarrow \mathbb{R}$ is the Radon transform [30, Chapter 2] of $g(u, v)$ at angle, θ_i, and range location, r_0, which corresponds to the integral over all entries of $g(u, v)$ that lie along the line, $r = r_0$, (also known as an *iso-range contour*). To form the full receive vector, the returns from all L ranges at angle, θ_i, within the image is written as

$$y(t) = \int_r R_{\theta_i, r}\left\{g(u, v)\right\} x\left(t - \frac{2r}{\nu_p}\right) dr \tag{4}$$

which is recast as

$$y(t) = p_{\theta_i}(t) * x(t) \tag{5}$$

where $*$ is the convolution operator and $p_{\theta_i}(t) = R_{\theta_i, k}\big\{g(u, v)\big\}$ to shorten notation. This system equation is interpreted as the response to an LTI system with input, $x(t)$, impulse response, $p_{\theta_i}(t)$, and output, $y(t)$. The received signal $y(t)$ is then sampled at the receiver.

The impulse response of the scene, $p_{\theta_i}(t)$, provides a model for the interaction of the transmit waveform with the sensing environment. Given, $y(t)$, the goal of this paper is to estimate the scene impulse response, $p_{\theta_i}(t)$, by optimizing the sensing waveform, $x(t)$. The information contained in the received signal, $y(k)$, that describes the scene impulse response, $p_{\theta_i}(t)$, is obtained by measuring the mutual information between them. Equation (5) underlies the system model.

B. Multistatic Radar Model

The model is extended to a scenario in which two separate radars, Radar i and Radar j, are present as shown in Figure 1. The radars transmit above the ground at an elevation angle of 45 degrees. Figure 1 portrays a setup known as multistatic radar in which one radar transmits a waveform and both receive the returns, which are jointly processed to construct the scene reflectivity function. This structure holds when the model is extended to the multi-static case with M radars in which one radar transmits and all receive.

Given that Radar i transmits, the received signal at Radar i, $\mathbf{y}_i \in \mathbb{R}^{N+L-1}$, is written using (5) as

$$\mathbf{y}_i = A_{i,i}\mathbf{p}_{\theta_i} * \mathbf{x} + \mathbf{n}_i, \tag{6}$$

where $\mathbf{x} = [x(1), \ldots, x(N)]^T$, $A_{i,i}$ is a path loss coefficient, and $\mathbf{n}_i$ is additive Gaussian noise such that $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, \sigma_n^2\mathbf{I})$ where σ_n^2 is the variance.

To describe the return at Radar j, the response at Radar j due to Radar i is approximated as the response of an imaginary radar placed at the bistatic bisector of the two radars [31], [32]. This phenomenon is also studied in [33]. The bistatic bisector is the angle that divides the bistatic angle in half—this is denoted as $\theta_b = \frac{1}{2}(\theta_i - \theta_j)$. With this approximation, the return at Radar j is written as

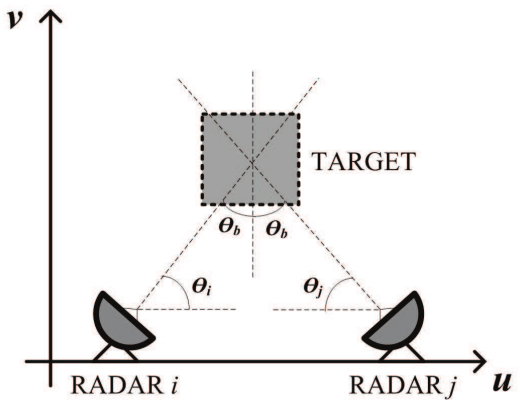

Fig. 1. Bistatic system in which two radars are available. A multistatic radar system is assumed in which one transmits and both receive. This is extended to M radars.

$$\mathbf{y}_j = A_{i,j}\mathbf{p}_{\theta_b} * \mathbf{x} + \mathbf{n}_j. \tag{7}$$

To exploit the CG distribution of the scene reflectivity function, the formulation of the system model will now incorporate the sparse structure of the image. First, the matrix version of the image $\mathbf{G}$ is defined as

$$\mathbf{G} \in \mathbb{R}^{I_r \times I_c} : g_{i,j} = g(i, j), \tag{8}$$

where I_r is the number of rows in the image and I_c is the number of columns, then we write $\mathbf{G}$ as

$$\tilde{\mathbf{g}} = \text{vec}(\mathbf{G}) = \mathbf{\Phi}\mathbf{c} \tag{9}$$

where $\mathbf{\Phi} \in \mathbb{R}^{L \times D}$ is a dictionary constructed of D vectorized wavelet *atoms* used as a basis and $\mathbf{c} \in \mathbb{R}^D$ is a random vector modeling the scene reflectivity function. As mentioned, the wavelet coefficient image vector, $\mathbf{c}$, is seen to be distributed as a CG, which implies

$$\mathbf{c} = z\mathbf{w}, \tag{10}$$

where $\mathbf{w}$ is a Gaussian vector such that $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_w)$ where $\mathbf{\Sigma}_w$ is the covariance matrix. Because the Gaussian scale mixture distribution is defined as an integral over, z, it is therefore continuous with respect to z. For the special case of a finite mixture of Gaussian distributions shown in (10), the multiplier, z, is a non-negative, scalar, and discrete random variable with probability mass function (PMF) $p_Z(z)$.

The overall set of wavelet coefficients is partitioned into subsets commonly called neighbourhoods. We assumed that each neighborhood has a common structure, as is common in image processing literature, and thus are similarly distributed. Therefore, the covariance matrix is constructed as a block diagonal matrix with copies of a neighborhood covariance matrix, $\mathbf{\Sigma}_{Nbr}$, on the diagonal [34], [35].

The neighborhood covariance matrix $\mathbf{\Sigma}_{Nbr}$ is explicitly computed from training data or is known *a priori*. The PMF

of the scalar mixing random variable, z, is computed explicitly from training data using maximum-likelihood techniques [15].

Using the sparse CG model (6) is rewritten as

$$\mathbf{y}_i = (A_{i,i}\mathbf{R}_{\theta_i}\mathbf{\Phi}\mathbf{c}) * \mathbf{x} + \mathbf{n}_i, \tag{11}$$

where, $\mathbf{R}_{\theta_i}$, is the Radon transform at angle θ_i written as a matrix operation because it is a linear operator. Defining the Radon transform of the sparse wavelet dictionary as $\mathbf{\Psi}_i = A_{i,i}\mathbf{R}_{\theta_i}\mathbf{\Phi}$, and form the convolution matrix $\mathbf{X} \in \mathbb{R}^{(N+L-1)\times L}$ as

$$\mathbf{X} = \begin{bmatrix} x(1) & 0 & & \cdots & 0 \\ x(2) & x(1) & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & x(N) & x(N-1) \\ 0 & \cdots & & 0 & x(N) \end{bmatrix}, \tag{12}$$

the returns at Radars i and j are written as

$$\mathbf{y}_i = \mathbf{X}\mathbf{\Psi}_i\mathbf{c} + \mathbf{n}_i, \tag{13}$$

$$\mathbf{y}_j = \mathbf{X}\mathbf{\Psi}_j\mathbf{c} + \mathbf{n}_j. \tag{14}$$

The convolution matrix in (12) is sparse, where the non-zero entries are contained within an uppermost and lowermost diagonal. We take advantage of these properties using sparse coding which reduces the computational complexity and enables us to perform compressive imaging. This is detailed in Subsection VII-A.

The model in (14) is extended to other radar systems in the multistatic array. For a total of M radar systems with only Radar i transmitting, the total responses are stated as a system of equations such as

$$\mathbf{y} = \overline{\mathbf{X}}\mathbf{\Psi}\mathbf{c} + \mathbf{n}, \tag{15}$$

where $\overline{\mathbf{X}}$ is a block diagonal matrix with M copies of $\mathbf{X}$ down the diagonal, and

$$\mathbf{\Psi} = [\mathbf{\Psi}_1^T, \mathbf{\Psi}_2^T, \ldots, \mathbf{\Psi}_M^T]^T,$$

$$\mathbf{n} = [\mathbf{n}_1^T, \mathbf{n}_2^T, \ldots, \mathbf{n}_M^T]^T,$$

where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2\mathbf{I})$.

Upon reception, estimating the image vector $\mathbf{c}$, from $\mathbf{y}$, is well-posed by minimizing a regularized least squares cost functional defined as

$$J(\mathbf{c}) = \|\mathbf{y} - \overline{\mathbf{X}}\mathbf{\Psi}\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_p^p \tag{16}$$

where the sparsity of $\mathbf{c}$ is achieved by setting $p = 1$ and the parameter $\lambda \geq 0$. The estimated image vector $\hat{\mathbf{c}}$ that minimizes (16) is projected onto the sparse wavelet basis to form the estimated scene reflectivity function, $\tilde{\mathbf{g}}_{\text{est}} = \mathbf{\Phi}\hat{\mathbf{c}}$.

III. Scene Reflectivity Modeling

The image impulse response $\mathbf{p}_{\theta_i}$, of the scene is modeled by assuming a metallic mine buried 2 cm beneath a rough and dry surface. The frequency dependent radar cross section (RCS) was simulated using the Army finite-difference time-domain (AFDTD) algorithm [36]. The in-phase and quadrature-phase frequency responses are shown in Figure 2.

The rough and dry surface is modeled as a CG distribution. This model is well suited for large sporadic variations in the reflected signal from rough surface scatterers [22]. To model the CFR, the interference model from [37] is borrowed; however the Gaussian interference is replaced with the more accurate CG distribution after the parameters of the distribution are estimated from the scene.

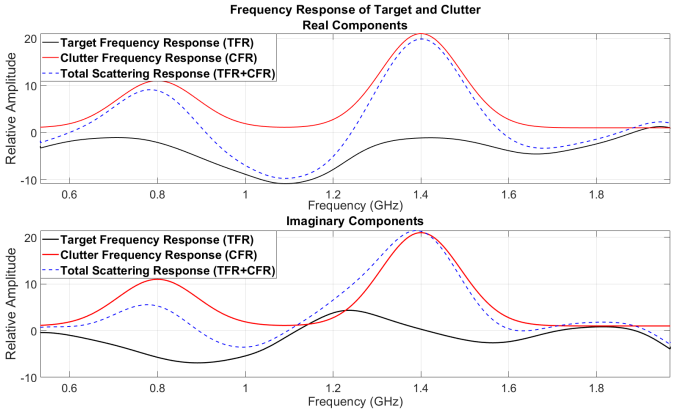

Fig. 2. Shown are the target and clutter frequency responses including their linear combination as the overall scattering response of the scene. The CFR shows the CG nature of the clutter. The TFR and CFR are plotted over the 3-dB bandwidth of the transmit waveform.

A common technique in the literature is to assume that the impulse response of a target is given by a complex Gaussian random variable. In [2, p. 48], the authors state in the case that an extended target is not known *a priori*, it makes sense to model the TIR using a finite-energy random process, as the distribution of scattering centers for an extended target are well modeled using a Gaussian random process. However, it can be shown that with natural scenes, the statistics (when represented on a wavelet basis) do not behave in a Gaussian manner, but is in fact characterized closely as a CG [15], [16]. In [16] it was also shown that radar images follow the CG distribution more closely than the Gaussian distribution. This fact is exploited to characterize the scene reflectivity function when deriving the MI.

IV. Waveform Design Using Mutual Information

A. Waveform 1: Non-Target Specific

This section derives the non-target specific waveform using knowledge of the CG distribution as a prior to model the scene reflectivity function. The MI between the complete set of received signals, $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_M^T]^T$, and the image vector, $\mathbf{c}$, given the transmit waveform, $\mathbf{x}$ is derived. The conditional MI is denoted by $I(\mathbf{y}; \mathbf{c}|\mathbf{x})$. Using the signal model in Section II, the conditional mutual information can be written as

$$I(\mathbf{y}; \mathbf{c}|\mathbf{x}) = h(\mathbf{y}|\mathbf{x}) - h(\mathbf{y}|\mathbf{c}, \mathbf{x}), \tag{17}$$

where $h(\cdot)$ denotes differential entropy [38, Chapter 9]. The second term on the right-hand side, $h(\mathbf{y}|\mathbf{c}, \mathbf{x})$, is reduced to the well-known Gaussian differential entropy

$$h(\mathbf{y}|\mathbf{c}, \mathbf{x}) = h(\mathbf{n}) = \frac{1}{2}K \log\left(2\pi e\sigma_n^2\right) \tag{18}$$

where the constant $K = M(N + L - 1)$. As (18) has no dependence on the transmit waveform, it has no effect on the solution. Thus, for optimization purposes this term is ignored and the focus shifts to the first term on the right hand side of (17).

The solution to the differential entropy, $h(\mathbf{y}|\mathbf{x})$, that remains waveform dependent (in (15)) is derived. Using the CG model of (10) for $\mathbf{c}$ in (17) yields

$$\mathbf{y} = z\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{w} + \mathbf{n}. \tag{19}$$

Since the multipliers in the first term on the right hand side are not random, $\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{w}$ is still a Gaussian vector. Denoting this quantity as $\mathbf{Q} = \overline{\mathbf{X}}\mathbf{\Psi}\mathbf{w}$ gives a new random vector with covariance matrix

$$\text{cov}(\mathbf{Q}, \mathbf{Q}) = \text{E}\big[\mathbf{Q}\mathbf{Q}^T\big] = \overline{\mathbf{X}}\mathbf{\Psi}\mathbf{\Sigma}_w\mathbf{\Psi}^T\overline{\mathbf{X}}^T \tag{20}$$

and the Gaussian vector is distributed as

$$\mathbf{Q} \sim \mathcal{N}\big(\mathbf{0}, \overline{\mathbf{X}}\mathbf{\Psi}\mathbf{\Sigma}_w\mathbf{\Psi}^T\overline{\mathbf{X}}^T\big). \tag{21}$$

The system model is then rewritten as

$$\mathbf{y} = z\mathbf{Q} + \mathbf{n}, \tag{22}$$

which is the sum of a CG and a Gaussian random variable. Thus, rewriting the expression for MI results in

$$I(\mathbf{y}; \mathbf{c}|\mathbf{x}) = h(z\mathbf{Q} + \mathbf{n}|\mathbf{x}) - \frac{1}{2}K \log\left(2\pi e\sigma_n^2\right). \tag{23}$$

Although Gaussian functions have a closed form differential entropy as shown in (18), there is no closed form solution for the differential entropy of the sum of a Gaussian and a CG.

This is due to the fact that differential entropy is defined as (in the discrete case) $h(\mathbf{x}) = -\sum_{-\infty}^{\infty} f(\mathbf{x}) \log f(\mathbf{x})$. Therefore the differential entropy of this vector is the logarithm of the sum of exponentials [39].

From [39], upper and lower bounds

$$h_l(\mathbf{x}) \leq h(z\mathbf{Q} + \mathbf{n}|\mathbf{x}) \leq h_u(\mathbf{x}) \tag{24}$$

for the $h(z\mathbf{Q} + \mathbf{n}|\mathbf{x})$ are found and are used to bound the MI, (23). The upper bound h_u, is defined (from [39]) for $h(z\mathbf{Q} + \mathbf{n}|\mathbf{x})$ as

$$h_u(\mathbf{x}) \triangleq \sum_i w_i \cdot \left(-\log w_i + \frac{1}{2}\log\left((2\pi e)^D|\mathbf{C}_i|\right)\right) \tag{25}$$

where $w_i = p_z(z_i)$ and $\mathbf{C}_i = z_i^2\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{\Sigma}_w\mathbf{\Psi}^T\overline{\mathbf{X}}^T + \sigma_n^2\mathbf{I}$ is the covariance matrix of $z\mathbf{Q} + \mathbf{n}$.

Similarly, the lower bound, $h_l(\mathbf{x})$ is defined as

$$-\sum_i w_i \cdot \log\left(\sum_j w_j\left[(2\pi)^{D/2}|\mathbf{C}_i + \mathbf{C}_j|\right]^{-1/2}\right) \tag{26}$$

where $\mathbf{C}_i$ and $\mathbf{C}_j$ are defined as in (25) but summed over different indices. Appendix A shows the derivation of (26) using [39] as it is less obvious.

The upper and lower bounds on the differential entropy shown in (25) and (26), respectively, then bound the conditional mutual information (17) between the received signal vector and the image vector. Therefore, the optimization of the waveform optimized depends on the optimization of the bounds as shown in in Section V-A and Section V-B for the upper and lower bounds, respectively. The resulting target reconstruction using these waveforms is shown in Section VIII-A.

B. Waveform 2: Target Specific

This section presents the matched illumination techniques used to form a *target specific* (TS) waveform. The TS waveform takes into account the TFR and CFR that make up the scene reflectivity function.

Since maximizing the MI is posed as a channel capacity problem, matched illumination waveform design makes use of the Shannon-Hartley theorem. From this theorem, maximizing the mutual information is reformulated in terms of the signal-to-interference-plus-noise ratio (SINR). The conditional mutual information in the frequency domain is written as [5]

$$I(\mathbf{y}; \mathbf{g}|\mathbf{x}) = T\int_B \log\left(1 + \frac{2|X(f)|^2 S_g(f)}{TP_n(f) + 2|X(f)|^2 S_c(f)}\right)df. \tag{27}$$

Here, S_g and S_c denote the power spectral density of the target and clutter respectively. This is viewed as optimizing the signal-to-interference-plus-noise ratio (SINR) where our degree of freedom is the power spectral density of the transmit waveform, $|X(f)|^2$.

V. Waveform Optimization

The optimization of the MI defined in Section IV is presented. The bounds on the mutual information based cost function are explicitly given in Section V-A and Section V-B respectively. The convexity of each bound is proven by decomposing it as a convex combination of elementary convex functions; and by exploiting the positive definiteness of the relevant Hessian matrix.

A. Optimization of Upper Bound, $h_u(\mathbf{x})$

The transmit waveform, $\mathbf{x}_u$, that maximizes the upper bound on the MI (25), is presented. The objective is defined as

$$\mathbf{x}_u = \underset{\mathbf{x}:\|\mathbf{x}\|^2 \leq T\rho}{\text{argmax}}\ h_u(\mathbf{x}). \tag{28}$$

Inserting the expression $h_u(\mathbf{x})$ and ignoring constants, (28) is rewritten as

$$\mathbf{x}_{up} = \underset{\mathbf{x}:\|\mathbf{x}\|^2 \leq T\rho}{\text{argmax}} \sum_i w_i \log\det\left(\mathbf{C}_i\right). \tag{29}$$

Since there is a one-to-one correspondence of the waveform, $\mathbf{x}$, and the convolution matrix, $\overline{\mathbf{X}}$, (29) is expanded as

$$\overline{\mathbf{X}}_{up} = \underset{\overline{\mathbf{X}}:\|\overline{\mathbf{X}}\|_F^2 \leq LMN\rho}{\text{argmax}} \sum_i w_i \log\det\left(z_i^2\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{\Sigma}_w\mathbf{\Psi}^T\overline{\mathbf{X}}^T + \sigma_n^2\mathbf{I}\right) \tag{30}$$

This upper bound is concave in $\overline{\mathbf{X}}$. This is shown by applying the Cholesky decomposition to $\mathbf{\Sigma}_w = \mathbf{D}_w\mathbf{D}_w^T$ since covariance matrices are at least positive semi-definite. Using

this and Sylvester's determinant theorem [40], the objective function (30) is rearranged (shown in Appendix B) to get

$$\sum_i w_i \log\det\left(\frac{z_i^2}{\sigma_n^2}\mathbf{D}_w^T\mathbf{\Psi}^T\overline{\mathbf{X}}^T\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{D}_w+\mathbf{I}\right). \tag{31}$$

Therefore, the log-determinant is concave as it depends on the positive quantity $\overline{\mathbf{X}}^T\overline{\mathbf{X}}$, [41]. Defining $\overline{\mathbf{R}}_{\overline{\mathbf{X}}}=\overline{\mathbf{X}}^T\overline{\mathbf{X}}$, as the main variable for optimization the upper bound is concave in $\overline{\mathbf{R}}_{\overline{\mathbf{X}}}$ and (31) is rewritten as

$$\underset{\overline{\mathbf{R}}_{\overline{\mathbf{X}}}:\mathrm{tr}(\overline{\mathbf{R}}_{\overline{\mathbf{X}}})\le LMN\rho}{\operatorname{argmax}} \sum_i w_i \log\det\left(\frac{z_i^2}{\sigma_n^2}\mathbf{D}_w^T\mathbf{\Psi}^T\overline{\mathbf{R}}_{\overline{\mathbf{X}}}\mathbf{\Psi}\mathbf{D}_w+\mathbf{I}\right). \tag{32}$$

$\overline{\mathbf{R}}_{\overline{\mathbf{X}}}$ is a block diagonal matrix where each of the M identical blocks are $\mathbf{R}_{\mathbf{X}}=\mathbf{X}^T\mathbf{X}$. Therefore, to find the optimal waveform we show that $\mathbf{R}_{\mathbf{X}}$ is a matrix of autocorrelation values in observation 1.

Observation 1. *The transmit convolution matrix as a function of length N sequence $\mathbf{x}$, and a length L impulse response is defined as $\mathbf{C}(\mathbf{x},L)\in\mathbb{R}^{(N+L-1)\times L}$. Therefore, the Hermitian-Toeplitz autocorrelation matrix $\mathbf{R}_{\mathbf{X}}=\mathbf{C}(\mathbf{x},L)^T\mathbf{C}(\mathbf{x},L)\in\mathbb{R}^{L\times L}$ is written as*

$$\mathbf{R}_{\mathbf{X}}=\begin{bmatrix} r_0 & r_1 & \cdots & r_{L-1}\\ r_1 & r_0 & \cdots & r_{L-2}\\ \vdots & & \ddots & \vdots\\ r_{L-1} & r_{L-2} & \cdots & r_0\end{bmatrix}. \tag{33}$$

Since the autocorrelation function of $\mathbf{x}$ is defined as

$$r(\tau)=\int_0^T x(t)\overline{x(t-\tau)}dt, \tag{34}$$

where $x(t)$ is nonzero only on the interval $0\le t<T$, the entries of, $\mathbf{R}_{\mathbf{X}}\in\mathbb{R}^{L\times L}$, are expressed as $r_{ij}=r(\tau)$ where $\tau=i-j$.

The proof is shown in Sub-appendix C.A. Therefore, building on observation 1, extracting $\mathbf{x}$ from $\mathbf{\Gamma}$ is given in corollary 1.

Corollary 1. *Given a Hermitian, Toeplitz, and positive semidefinite matrix $\mathbf{R}_{\mathbf{X}}\in\mathbb{R}^{L\times L}$, a convolution matrix $\mathbf{C}(\mathbf{x},L)$ that satisfies $\mathbf{R}_{\mathbf{X}}=\mathbf{C}(\mathbf{x},L)^T\mathbf{C}(\mathbf{x},L)$ is given by $\mathbf{C}(\mathbf{x}_0,L)$ where $\mathbf{x}_0$ is defined as*

$$\mathbf{x}_0=\mathcal{F}^{-1}\left\{\sqrt{\mathcal{F}\{\mathbf{r}\}}\right\}, \tag{35}$$

where $\mathbf{r}\in\mathbb{R}^N$ is an autocorrelation sequence built from $\mathbf{R}_{\mathbf{X}}$ and $\mathcal{F}\{\cdot\}$ is the discrete Fourier transform. Note that it is assumed that $N>L$, and $\mathbf{r}$ is zero-padded to be of length $2(N-1)+1$.

The proof is shown in Sub-appendix C.B. Therefore, the overall optimization problem is written as in Lemma 1.

Lemma 1. *The block diagonal transmit waveform autocorrelation matrix, $\overline{\mathbf{R}}_{\overline{\mathbf{X}}}=\overline{\mathbf{X}}^T\overline{\mathbf{X}}$, that maximizes the upper bound on mutual information, $h_{up}(\mathbf{x})$, is found using the following convex program,*

$$\begin{aligned}\operatorname{argmax}\quad & \sum_i p_z(z_i)\log\det\left(\frac{z_i^2}{\sigma_n^2}\mathbf{D}_w^T\mathbf{\Psi}^T\overline{\mathbf{R}}_{\overline{\mathbf{X}}}\mathbf{\Psi}\mathbf{D}_w+\mathbf{I}\right).\\ \text{such that}\quad & \mathrm{tr}(\overline{\mathbf{R}}_{\overline{\mathbf{X}}})\le LMN\rho,\\ & \mathbf{R}_{\mathbf{X}}^T=\mathbf{R}_{\mathbf{X}},\\ & \mathbf{R}_{\mathbf{X}}\ \text{Toeplitz},\\ & \mathbf{R}_{\mathbf{X}}\succeq 0.\end{aligned} \tag{36}$$

Proof. Follows from Observation 1 and Corollary 1. □

B. Optimization of Lower Bound, $f_{low}(\mathbf{x})$

For completeness, the lower bound on mutual information is be presented. Due to a similar structure, most of this optimization will follow the same pattern as optimization of the upper bound in the preceding section. Rewriting the lower bound on mutual information, (26), while ignoring constants gives

$$-\sum_i w_i \log\left(\sum_j w_j\left[\det\left(\mathbf{C}_i+\mathbf{C}_j\right)\right]^{-1/2}\right). \tag{37}$$

Expanding out the inside of the $\log(\cdot)$ using $\mathbf{C}_i$ defined in (25) gives

$$\sum_j w_j\left[\det\left(\left(z_i^2+z_j^2\right)\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{\Sigma}_w\mathbf{\Psi}^T\overline{\mathbf{X}}^T+2\sigma_n^2\mathbf{I}\right)\right]^{-1/2}. \tag{38}$$

Using the Cholesky decomposition and Sylvester's determinant theorem as in (30), (38) is equivalent to

$$\sum_j w_j\left[\det\left(\frac{z_i^2+z_j^2}{2\sigma_n^2}\mathbf{D}_w^T\mathbf{\Psi}^T\overline{\mathbf{R}}_{\overline{\mathbf{X}}}\mathbf{\Psi}\mathbf{D}_w+\mathbf{I}\right)\right]^{-1/2} \tag{39}$$

Since the convexity of the lower bound (39) is less obvious, the following Lemmas are presented.

Lemma 2. *Let $f(t)$ be a twice-differentiable, real function that takes only non-negative values (i.e., $f(t)\ge 0$) and whose derivatives satisfy the following inequality,*

$$\left(1+\lambda t+f\left(t\right)\right)f''(t)-f'(t)-\lambda\le 0, \tag{40}$$

where $\lambda>0$. Then, the following function, $g(t)$, is concave in $t\ge 0$,

$$g(t)=\log(1+\lambda t+f(t)). \tag{41}$$

The proof is shown in Sub-appendix D.A. Using the preceding lemma, the following result is shown.

Lemma 3. *Given a discrete non-negative random variable z with associated pmf $p_z(z)$ and constants $d_{\ell,j}\ge 0$, $\lambda_\ell\ge 0$ for $\ell=1,2,\ldots,D$, the following function, $h(t)$, is convex in $t\ge 0$,*

$$h(t)=\log\left(\sum_j p_z(z_j)\left[z_j^{2D}\prod_{\ell=1}^{D}[1+d_{\ell,j}+\lambda_\ell t]\right]^{-1}\right). \tag{42}$$

The proof is shown in Sub-appendix D.B.

Lemma 4. *The function,* $f(\mathbf{X}) : \mathbb{R}^{D\times D} \rightarrow \mathbb{R}$

$$f(\mathbf{X}) = -\sum_i p_z(z_i)\cdot \log\left(\sum_j p_z(z_j)\left[\det(\mathbf{I}+(z_i^2+z_j^2)\mathbf{X})\right]^{-1}\right) \tag{43}$$

is concave for a positive semidefinite $\mathbf{X}$.

The proof is shown in Sub-appendix D.C. Therefore, as per Lemma 4 it can be seen that the lower bound is convex.

C. Optimization of Target Specific Waveform

To form the TS waveform, we use the matched illumination techniques presented in Section IV-B. The transmit waveform that maximizes the TS mutual information of (27) is presented. The TS waveform is optimized for a known target response, and the clutter is described as being of the CG nature. Using this, the cost function (27) is optimized to find the optimal waveform. The TS matched illumination technique optimizes the PSD, $|X(\omega)|^2$, of the transmit waveform. Therefore, a time domain waveform with the desired PSD must be found.

By optimizing the power spectrum of the waveform,

$$X(\omega) = |X(\omega)|\exp\left(j\angle X\left(\omega\right)\right), \tag{44}$$

the phase function $\phi = \angle X(\omega)$ remains arbitrary. Therefore, designing a waveform that fulfils both the constant envelope and power spectrum requirements is turned into one of embedding the optimized spectrum into the phase of a time series signal. This is done using the 'Projections onto Convex Sets' (POCS) method [42].

This method finds a point that is either closest to, or within the intersection of, two or more sets by alternating projection operators between the sets. Convergence is guaranteed if all sets are convex, and their intersections are non-empty. However, this is not the case for waveform design [43].

Both the set of all signals whose Fourier transform has the same magnitude, C, and the set of constant envelope signals, D, do not form convex sets. However, using projection operators, a point, $\mathbf{x}_k$, is projected to a *nearest neighbor* within each set. The best point is the one with the smallest distance from all sets. The method is defined for an initialization signal, $\mathbf{x}_0$, as

$$\mathbf{x}_{k+1} = P_C(\mathbf{y}_k), \quad \mathbf{y}_k = P_D(\mathbf{x}_k), \quad k = 0, 1, 2, ... \tag{45}$$

where the projection operators, P_C, and, P_D, define the projections onto the set of Fourier transform magnitudes and constant envelope signals, respectively, over the iterations, k. The projection operators are defined as

$$\begin{aligned} P_D(\mathbf{x}_k) &= \mathcal{F}^{-1}\left[\mathbf{S}_{\text{opt}}\exp j(\angle\mathcal{F}\{\mathbf{x}_k\})\right] \\ P_C(\mathbf{y}_k) &= \mathbf{u}\exp j(\angle\mathbf{y}_k). \end{aligned} \tag{46}$$

where $\mathbf{S}_{\text{opt}}$ is the given PSD. This method is seen as a phase-retrieval technique and produces a phase-coded waveform. The waveform, $\mathbf{x}$, generated using the POCS is thus optimized by finding the optimal phase coding, $\mathbf{y}$. This algorithm can be initialized by defining, $\mathbf{x}_0$ to be a phase-coded waveform with each sample chosen to have a random phase in $[-\pi, \pi]$.

As both C and D are not convex sets, points that minimize the distance may not be unique. Therefore, it is necessary to develop tolerance and convergence criteria to ensure that the algorithm converges asymptotically as $k \rightarrow \infty$. The MSE is defined as

$$\text{MSE} = \frac{1}{N}\sum_{n=1}^{N}\left\|S_{\text{opt}} - \mathcal{F}\{\mathbf{R}_{\mathbf{x}}\}\right\|_2^2 \tag{47}$$

where $\mathbf{R}_{\mathbf{x}}$ is the autocorrelation of $\mathbf{x}_{k+1}$ in (45), and $\mathcal{F}$ is the Fourier transform operator. The distance between the two sets is said to be been minimized based upon this MSE. If

$$MSE < \epsilon \tag{48}$$

for a certain $\hat{K}$ number of iterations where ϵ is a set tolerance, a point, $\mathbf{x}^*$, is said to minimize the distance between the two sets. For a proof on the convergence of projections onto convex sets see [44].

The resulting point, $\mathbf{x}^*$, is a constant envelope signal, which approximates the optimized power spectrum. Although the initialization signal, $\mathbf{x}_0$, can be random with phase in $[-\pi, \pi]$, it was seen that the lowest MSE at convergence occurred when the initial point was an amplitude modulated LFM waveform.

VI. Optimization Algorithms

The complete optimization procedure from transmit to receive is shown for both the CG and TS based approaches described in the previous section. Algorithm 1 shows the closed loop procedure for the optimization of the CG waveform.

Algorithm 1 CG Waveform Optimization

Inputs: $\mathbf{\Psi}, \tilde{\mathbf{g}}, \lambda, M, N$

1. Estimate Clutter Statistics
   - $\mathbf{c} = \mathbf{\Psi}^T\tilde{\mathbf{g}} \leftarrow$ project scene onto wavelet basis
   - $\hat{\mathbf{\Sigma}}_w = \mathbb{E}\left[\mathbf{c}_N\mathbf{c}_N^T\right] \leftarrow$ estimate covariance
   - $z = \sqrt{\mathbf{c}_N^T\hat{\mathbf{\Sigma}}_w^{-1}\mathbf{c}_N/N} \leftarrow$ find CG multipliers
   - $\mathbf{\Sigma}_w = \mathbb{E}\left[\mathbf{c}_N\mathbf{c}_N^T/z^2\right]$
2. Optimize Waveform
   - Optimize (36) using Algorithm 3 or 4
   - $\overline{\mathbf{X}} \leftarrow$ form transmit matrix using (35) and (12)
3. Transmit Optimized Waveform
   - $\mathbf{y} = \overline{\mathbf{X}}\mathbf{\Psi}\mathbf{c} + \mathbf{n}$
4. Received Waveform
   - $\hat{\mathbf{c}} = \text{argmin}_{\mathbf{c}\in\mathbb{R}}\,||\mathbf{y} - \mathbf{X}\mathbf{\Psi}\mathbf{c}||_2^2 + \lambda||\mathbf{c}||_p^p$
   - $\tilde{\mathbf{g}}_{\text{est}} = \mathbf{\Psi}\hat{\mathbf{c}} \leftarrow$ form new scene

Algorithm 2 shows the closed loop procedure for the optimization of the TS waveform.

To solve step 2 in Algorithms 1 and 2, the Taguchi and PSO solvers are presented. Both Algorithm 1 and Algorithm 2 are slowed down by solving the sparsely regularized least squares minimization problem and are therefore, $O(N^k)$, complex.

A. Taguchi Algorithm

In order to optimize the MI, Taguchi optimization techniques are applied to the waveform design process. An orthogonal array (OA) is used as a search grid for weighting power

Algorithm 2 TS Waveform Optimization

Inputs: $\mathbf{\Psi}, \tilde{\mathbf{g}}, S_g(f), S_c(f), \epsilon, \lambda, M, N$

1. Estimate Clutter Statistics
   $\mathbf{c} = \mathbf{\Psi}^T \tilde{\mathbf{g}} \leftarrow$ project scene onto wavelet basis
   $\hat{\mathbf{\Sigma}}_w = \mathbb{E}\left[\mathbf{c}_N \mathbf{c}_N^T\right] \leftarrow$ estimate covariance
   $z = \sqrt{\mathbf{c}_N^T \hat{\mathbf{\Sigma}}_w^{-1} \mathbf{c}_N / N} \leftarrow$ find CG multipliers
   $\mathbf{\Sigma}_w = \mathbb{E}\left[\mathbf{c}_N \mathbf{c}_N^T / z^2\right]$
2. Optimize Waveform
   $\mathbf{S}_{\text{opt}} \leftarrow$ Optimize (27) using Algorithm 3 or 4
   $\mathbf{x} \leftarrow$ find time domain waveform using (46)
   $\overline{\mathbf{X}} \leftarrow$ form transmit matrix using (12)
3. Transmit Optimized Waveform
   $\mathbf{y} = \overline{\mathbf{X}} \mathbf{\Psi} \mathbf{c} + \mathbf{n}$
4. Received Waveform
   $\hat{\mathbf{c}} = \text{argmin}_{\mathbf{c} \in \mathbb{R}} ||\mathbf{y} - \mathbf{X}\mathbf{\Psi}\mathbf{c}||_2^2 + \lambda ||\mathbf{c}||_p^p$
   $\tilde{\mathbf{g}}_{\text{est}} = \mathbf{\Psi}\hat{\mathbf{c}} \leftarrow$ form new scene

at each frequency. By using the OA, the Taguchi algorithm is both efficient and versatile. Rather than an exhaustive test of all possible power level combinations for optimizing the weights, the OA provides the most informative subset of these combinations. The versatility is due to the Taguchi algorithm's ability to solve non-convex problems. A comprehensive reference for OAs is found at [45]. Pseudo code for the Taguchi algorithm is given in Algorithm 3.

Bold, lower case letters denote row vectors, upper case letters denote matrices, subscripts denote elements, and superscripts denote the iterations.

Algorithm 3 Taguchi Algorithm

```
Initialize ŵ^0 = 1_n ∈ R^n, s^0 = 0_l ∈ R^l
Bin width b = [l/n]
Define A_ij ∈ R^{m×n} and step size μ
for k = 1, 2, ... do ← iterations
    for i = 1, ..., M do ← rows
        for j = 1, ..., N do ← columns
            if (A_ij = −1) then W_ij^k = ŵ_j^{k−1} − μ;
            else if (A_ij = 0) then W_ij^k = ŵ_j^{k−1};
            else if (A_ij = 1) then W_ij^k = ŵ_j^{k−1} + μ;
        end for
        p = 1; q = b;
        for g = 1, ..., L do ← update input
            s_{p:q}^k = W_ig^k ∗ 1_b
            p = p + b;  q = q + b;
        end for
        c_i^k = f(s^k) ← compute MI
    end for
    r_best = max(c^k);  ŵ^k = W^k(r_best, :);
end for
```

B. Particle Swarm Optimization (PSO)

The PSO algorithm is also used for waveform optimization. The PSO algorithm randomly initializes points ('particles') in the waveform space defined by the cost function. Since the PSO algorithm is also non-convex optimizer, this method can be generalized for any objective in waveform optimization. Convergence for the PSO algorithm is defined such that the improvement in MI does not improve by a certain tolerance over a number of iterations. The algorithm is summarized in Algorithm 4.

Algorithm 4 PSO Algorithm

```
p ← define number of particles
X^0 ∈ R^{p×l} ← initialize particle positions over l dimensions
K ← define maximum iterations
while Convergence criteria not met and k < K do
    for i = 1, ..., p do
        c_i = f(x_i); ← compute MI
        ĉ_i = f(x̂_i) ← find personal best solution
        if c_i > ĉ_i then ĉ_i = c_i;
    end for
    g_best = max(ĉ) ← find global best solution
    for j = 1, ...p do
        v(x_j, x̂_j, g_best); ← find velocity
        x_j; ← find solution
    end for
end while
```

VII. Simulations

A. Sparsity-Driven Imaging

To test the optimized waveforms, the signals are transmit onto the scene with a metallic mine buried 2 cm beneath the surface. Sparsity-driven regularization radar imaging is used to estimate the sparse vector of coefficients that make up the scene reflectivity function [46], [47].

As is shown in (9), the vectorized scene reflectively function can be decomposed into a dictionary and a vector of coefficients, and received signal is written as (15). The transmit waveform convolution matrix $\mathbf{X}$ is formed using the optimized waveform. Since the receive vector (15), the sparse basis, and the transmit waveform are known, the sparse image vector is the only unknown and is estimated using l_1 regularized least squares of the form

$$\hat{\mathbf{c}} = \underset{\mathbf{c} \in \mathbb{R}}{\text{argmin}} ||\mathbf{y} - \mathbf{X}\mathbf{\Psi}\mathbf{c}||_2^2 + \lambda ||\mathbf{c}||_p^p \tag{49}$$

where $\lambda \geq 0$ is the regularization parameter, and $p = 1$. The truncated Newton interior-point method is used to solve (49) [47]. Once the coefficients $\hat{\mathbf{c}}$ are estimated, the received scene reflectivity function is $\tilde{\mathbf{g}}_{\text{est}} = \mathbf{\Phi}^T \hat{\mathbf{c}}$.

The regularization parameter is 0.5. The number of sensor positions is 5. Note that the regularization term, $\lambda ||\mathbf{c}||_1$, forces the scene reconstruction to be sparse. Although a geometrical explanation is intuitive, it suffices to note that the l_1-norm is simply the sum of absolute values. Therefore, minimizing this term forces all but some of the values in $\mathbf{c}$ to approach zero enforcing sparsity on the solution.

A note on the dimensionality: It was found that representing a scene $G \in \mathbb{R}^{M \times N}$, as a vector requires a dictionary of dimensionality $MN \times MN$ and thus the problem grows quadratically. Explicitly computing such a basis for a large M

and N requires a huge amount of Random Access Memory (RAM). In the wavelet dictionary, it was observed that the number of nonzero entries was $4MN$, whereas the total number of entries grew as $(MN)^2$. This enabled the use of sparse coding throughout, by storing only the nonzero elements and their locations.

Figure 3 shows the difference between the nonzero elements and the total number of elements for a dictionary.

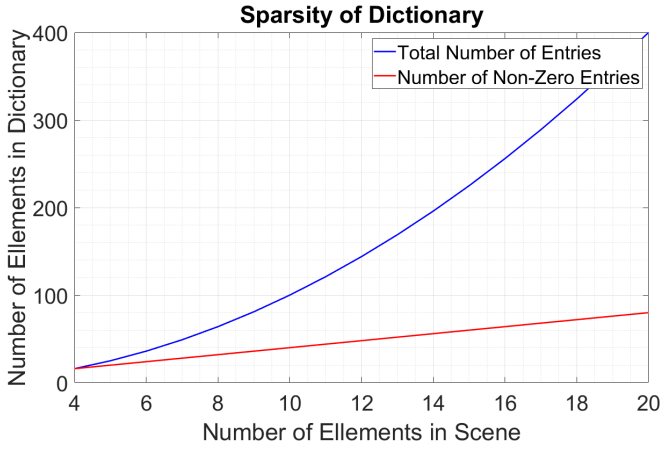

Fig. 3. The sparsity of the scene wavelet dictionary is shown.

B. Metrics

Image comparison metrics are used to assess the quality of waveforms in the reconstruction of the scene reflectivity function. The metrics are the SSIM index [48], the PSNR, the pixel to pixel MSE, and the SNR. The PSNR is defined as

$$\text{PSNR} = 20\log_{10}\left(\frac{\text{MAX}_\text{I}}{\sqrt{\text{MSE}}}\right). \tag{50}$$

where MAX_I is the maximum pixel value in the image. For example, when represented in 8-bit, $\text{MAX}_\text{I} = 255$. The MSE is the mean squared error between the ground truth image and reconstructed scene.

The SSIM is a widely used robust image comparison metric and encapsulates the spatial dependency of pixels on one-another rather than looking at absolute error. The SSIM is defined as

$$\text{SSIM}(\mathbf{x},\mathbf{y}) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{51}$$

where μ_x and σ_x are the mean and variance of a pixel patch $\mathbf{x} \in \mathbb{R}^K$, whose size takes into account a subset of pixels in the image. Both $\mathbf{x}$, and $\mathbf{y}$, are the same size and denote patches from the reconstructed and ground truth scene respectively. The constants are defined as $c_1 = (k_1 L)^2$, and $c_2 = (k_2 L)^2$, where L is the dynamic range of the pixel values, and $(k_1, k_2) = (0.01, 0.03)$ by default.

VIII. Results

In this section the estimated scene reflectivity functions are presented using the optimized non-matched illumination waveform defined in Section IV-A and the matched illumination waveform of Section IV-B. The phase history for each waveform is collected for a scene in which the response of a target buried 2 cm beneath the surface is phase and magnitude distorted by a randomly generated CG clutter response.

The optimization procedure for both the non-matched illumination waveform and the matched illumination waveform are defined in Algorithm 1 and Algorithm 2 respectively. Using the sparse imaging framework defined in Section VII-A, the scene reflectivity function for each waveform is estimated from the phase history. Each waveform is compared to the scene reflectivity function estimated from the phase history of an LFM waveform for the same bandwidth and energy.

The quality of each waveform in estimating the target reflectivity function is measured by the target reconstruction performance. Target reconstruction performance is defined as the similarity between the estimated scene and the target response in the presence of no clutter, shown in Figure 4. A high degree of similarity thus measures the degree to which the waveform is able to suppress clutter and therefore, Figure 4 will be considered the ground truth. This is quantified using the image metrics defined in Section VII-B. For each estimated scene reflectivity function, the SSIM values are given. A value of 0 indicates no similarity between two scenes, and a value of 1 indicates that the scenes are the same. The higher the SSIM value, the higher the similarity between the estimated scene and the ground truth target response thus quantifying the waveform's quality at reconstructing the target.

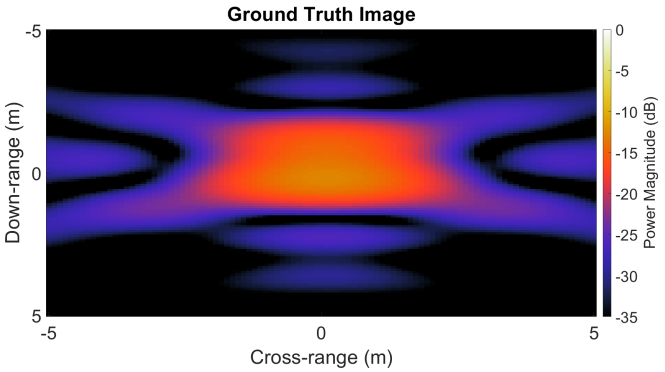

Fig. 4. Ground truth radar scene showing a metallic mine. When computing reconstruction performance for each waveform, this image is the reference.

Once the transmit waveform is optimized, the block diagonal convolution matrix as defined in (15) is formed. This is done by first forming a single convolution matrix, $\mathbf{X}$, using (12). Using the convolution matrix a block diagonal matrix, $\overline{\mathbf{X}} \in \mathbb{R}^{M(N+L-1)\times ML}$, is formed where $M = 5$ copies of $\mathbf{X}$ are placed, one for each sensor in the multistatic radar. The number of sensors equates to the number of Radon transform projections of the scene reflectivity function. Therefore, it is clear $M = 5$ is a sparse number of sensors as the scene is fully reconstructed using only five 'cross-range' locations.

A. Non-Target Specific Waveform

The results of optimizing the upper bound on the Non-TS waveform of Section V-A are presented first. The PSO and

Taguchi optimization procedures of Algorithm 3 and Algorithm 4 respectively are used to optimize the waveform. Table I and Table II show the target reconstruction performance using the PSO and Taguchi algorithms respectively.

The reconstructed scene reflectivity function from transmitting the PSO optimized waveform is presented in Figure 5. The results of the scene in Figure 5 to the ground truth scene

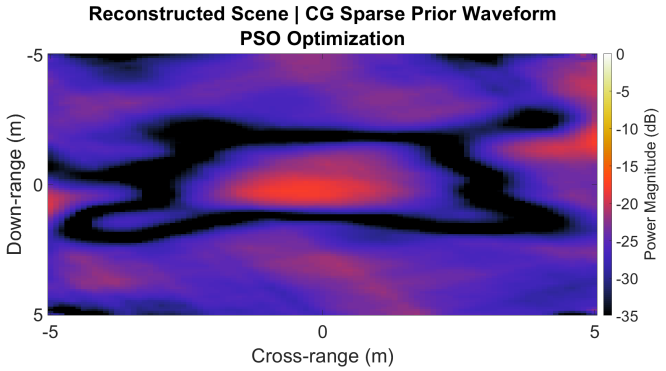

Fig. 5. Reconstructed scene using the sparse prior non-target specific waveform (32) optimized using the PSO algorithm. SSIM reconstruction value is 0.3634.

of Figure 4 are presented in column one of Table I.

The reconstructed scene reflectivity function from transmitting the Taguchi optimized waveform is presented in Figure 6. The results of comparing the scene in Figure 6 to the ground

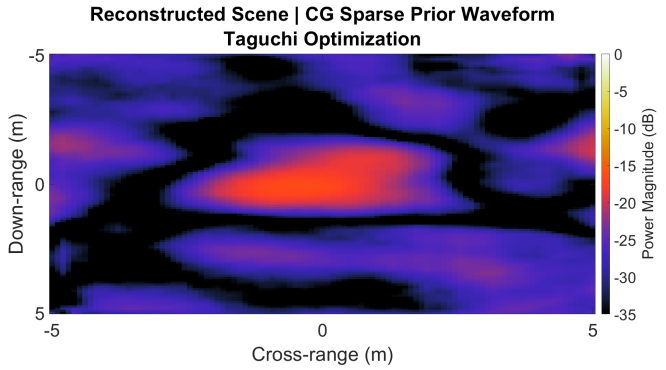

Fig. 6. Reconstructed scene using the sparse prior non-target specific waveform (32) optimized using the Taguchi algorithm. SSIM reconstruction value is 0.3859.

truth scene of Figure 4 are presented in column one of Table II.

The waveform is further compared to a constant modulus LFM waveform to show the improvement in target reconstruction that an optimized waveform offers over a conventional waveform. Each waveform is normalized to have the same energy. The resulting target reconstruction using the LFM waveform is shown in Figure 7. The resulting image metrics for the LFM waveform target reconstruction performance are given in Table III.

B. Target Specific Waveform

Secondly, the results of the TS waveforms as defined in Section IV-B are presented. The PSO and Taguchi optimization

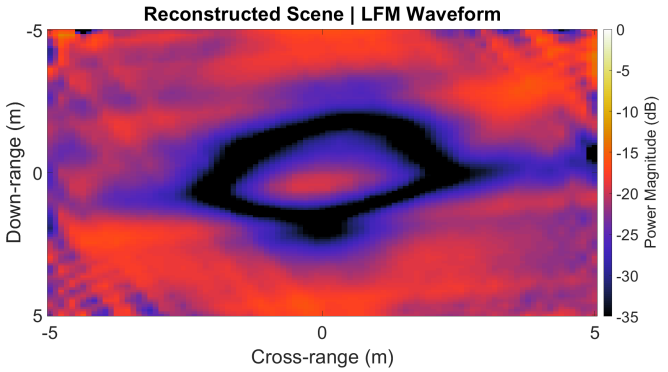

Fig. 7. Reconstructed scene using an LFM waveform. SSIM reconstruction value is 0.2928.

procedures of Algorithm 3 and Algorithm 4 respectively are used to optimize the waveform. Table I and Table II show the target reconstruction performance using the PSO and Taguchi algorithms respectively.

As the PSD of the TS waveform is optimized, a time domain waveform is formed using the POCS method outlined in Section V-C. The results are compared to the ground truth image of Figure 4 using the image metrics outlined in Section VII-A.

The reconstructed scene using the PSO algorithm to optimize the TS waveform is shown in Figure 8. The results of

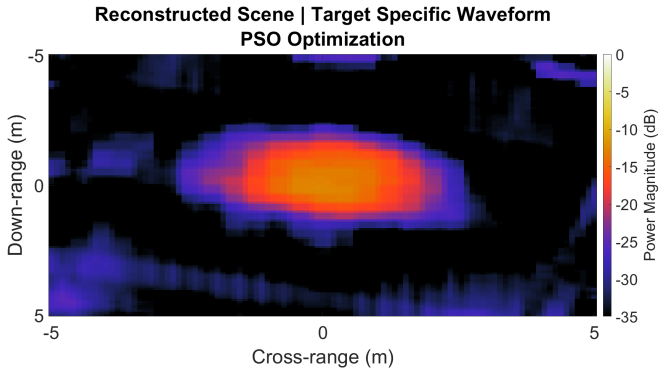

Fig. 8. Reconstructed scene using the PSO matched illumination PSD of (27). SSIM reconstruction value is 0.4467.

comparing the scene in Figure 8 to the ground truth scene of Figure 4 are presented in column two of Table I.

The reconstructed scene using the Taguchi algorithm to optimize the TS waveform is shown in Figure 9. The results of comparing the scene in Figure 9 to the ground truth scene of Figure 4 are presented in column two of Table II.

Comparing the results of the LFM waveform in Table III, both the non-target specific and target specific waveforms are seen to reconstruct the scene reflectivity function with higher similarity to the ground truth image.

Since the clutter is generated using the random CG distribution, the performance very much depends on the generation of the random numbers. Therefore, Monte Carlo methods provided bounds on the performance of each waveform.

Table I shows the results from the Monte Carlo experiments using the PSO algorithm to optimize the waveforms.

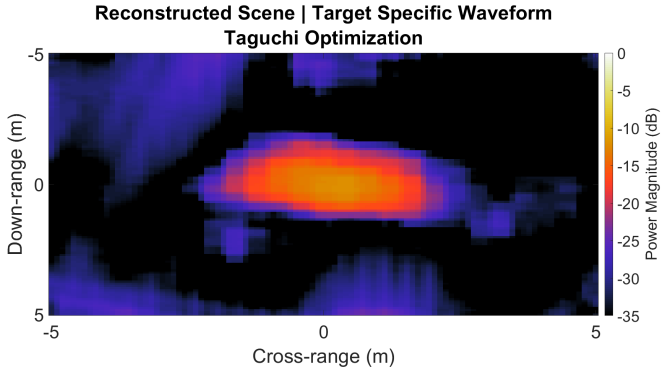

Fig. 9. Reconstructed scene using the Taguchi optimized matched illumination PSD of (27). SSIM reconstruction value is 0.4305.

TABLE I
PSO RECONSTRUCTION PERFORMANCE

Waveform	Non-TS	TS
SSIM	0.3612 ± 0.01981	0.4386 ± 0.02748
PSNR (dB)	14.28 ± 0.9800	16.56 ± 1.420
SNR (dB)	2.618 ± 0.9800	4.897 ± 1.420
MSE	0.03834 ± 0.009173	0.02332 ± 0.007963

Table II shows the results from the Monte Carlo experiments using the Taguchi algorithm to optimize the waveforms.

TABLE II
TAGUCHI RECONSTRUCTION PERFORMANCE

Waveform	Non-TS	TS
SSIM	0.3867 ± 0.02058	0.4333 ± 0.02583
PSNR (dB)	15.63 ± 1.203	20.09 ± 1.134
SNR (dB)	3.969 ± 1.203	8.435 ± 1.134
MSE	0.02843 ± 0.008060	0.01014 ± 0.003091

Table III shows the reconstruction performance bounds for the LFM waveform.

TABLE III
BENCHMARK RECONSTRUCTION PERFORMANCE

Waveform	LFM
SSIM	0.3102 ± 0.01803
PSNR (dB)	11.59 ± 0.7492
SNR (dB)	-0.06517 ± 0.7492
MSE	0.07031 ± 0.01206

These results demonstrate that our optimized waveforms show a significant improvement over the LFM waveform for all image reconstruction metrics, suggesting that the waveform is not algorithm or metric dependent and generally enhances the scene reflectivity function.

IX. Discussion

This paper presents two methods of waveform design for Sparsity-driven regularization imaging. The first method derives the mutual information when the scene is represented using a sparse basis. Therefore, the distribution is used as a prior in forming the waveform. Since the differential entropy of a CG distribution has no closed form solution, upper and lower bounds are derived for the MI. The maximization of the lower and upper bounds is over the transmit waveforms, and therefore the solutions provide optimal transmit signals for imaging.

The second, TS, waveform method takes full advantage of a cognitive radar's ability to estimate a TIR. Therefore, the matched illumination MI aims to maximize target information at the receiver. This is achieved by using the known power spectra of the target and clutter, and thus tailoring the spectrum of the transmit waveform by allocation power to the frequencies in which the clutter is least dominating. Therefore, the MI optimizes the PSD of the transmit waveform that maximizes the MI. To form a time domain signal with the optimal PSD from matched illumination the method of POCS is used. This allows the envelope of the transmit signal to be constant, while transmitting a signal with the optimized transmit spectrum.

The MI is optimized using both Taguchi optimization and PSO. The Taguchi algorithm is efficient and versatile as it uses OAs for the search space, and is able to solve non-convex problems. The PSO algorithm is extended to waveform design and compares favorably to the Taguchi algorithm.

The results show that by utilizing the scene prior distribution, and then taking full advantage of the TIR, the scene is reconstructed with a higher degree of accuracy as measured using the SSIM, PSNR, SNR, and MSE for both the PSO and Taguchi optimization algorithms.

The authors are currently extending this work by working from first principles. That is to say, starting with Maxwell's equations and numerically simulating the scenario by tracking the propagation of the EM waves as they are scattered. This includes the nature of the complex target, the type of soil, the loss in power experienced within the soil, and the second order scattering effects that occur when the EM signal penetrates boundary one (the air-surface boundary), is scattered from the second boundary (the target buried within the soil), and scattered again from the soil/air boundary.

In future work, we propose to conduct systematic studies into expanding the robustness of TS specific waveforms to encompass a class of target profile models. In particular, if the assumed target profile used to optimize the waveform does not match the true target profile in the scene, potential performance degradation may occur due to a mismatch between the assumed target profile and the true target profile. We provide initial results to explore the robustness of our TS waveform optimization framework with respect to model mismatch. This is achieved by optimizing the waveform based on a pristine (noiseless) target model and applying it to a scene where varying levels of noise are added to the pristine target profile; the results are shown Figure 10. Figure 10 shows the TS specific waveform is robust with respect to a reasonable variation in model mismatch, and eventually degrades, as expected, in performance as the model mismatch becomes larger.

Acknowledgements

This work was supported by the US Office of Naval Research (ONR) Grant Number N00014-16-1-2354 and the Naval Research Laboratory (NRL) Base Program. We thank

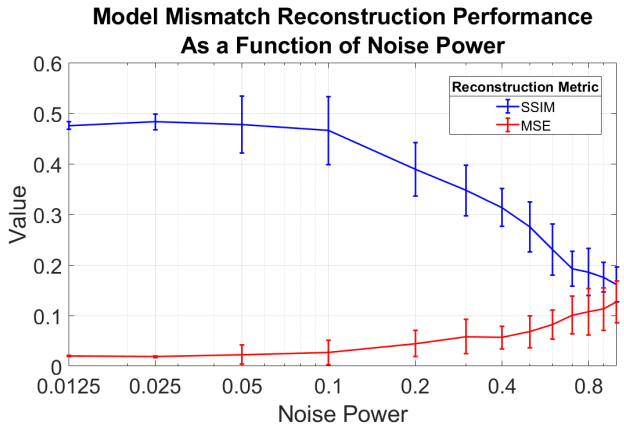

Fig. 10. Reconstruction performance of the scene as a function of increasing noise power level.

Dr. Traian Dogaru of the US Army Research Lab (ARL) for providing us with the simulated mine data used in our model development. We thank Dr. Zachary Chance of the MIT Lincoln Labs for his assistance in proving aspects of the convexity of the mutual information bounds. Finally, the authors thank the anonymous reviewers whose comments greatly helped improve the paper.

Appendix A
Lower Bound

From [39] the lower bound is defined as

$$h(\mathbf{x}) = -\sum_{i=1}^{L} w_i \log\left(\sum_{j=1}^{L} w_j z_{i,j}\right)$$
$$z_{i,j} = \int_{\mathbb{R}} \mathcal{N}(\mathbf{x}; 0, \mathbf{C}_i)\mathcal{N}(\mathbf{x}; 0, \mathbf{C}_j)d\mathbf{x} \tag{52}$$

The integral of the product of two zero mean Gaussian's, $z_{i,j}$, has a closed form solution as [49]

$$z_{i,j} = \frac{1}{\sqrt{(2\pi)^D |\mathbf{C}_i||\mathbf{C}_j|\left|\mathbf{C}_i^{-1} + \mathbf{C}_j^{-1}\right|}}. \tag{53}$$

Therefore, (52) simplifies to

$$h(\mathbf{x}) = -\sum_{i} w_i \cdot \log\left(\sum_{j} w_j \left[(2\pi)^D |\mathbf{C}_i + \mathbf{C}_j|\right]^{\frac{-1}{2}}\right). \tag{54}$$

Appendix B
Sylvester's Determinant

The determinant in (30) with the Cholesky decomposition for $\mathbf{\Sigma}_w = \mathbf{D}_w\mathbf{D}_w^T$ is

$$\left|\frac{z_i^2}{\sigma_n^2}\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{D}_w\mathbf{D}_w^T\mathbf{\Psi}^T\overline{\mathbf{X}}^T + \mathbf{I}\right| \tag{55}$$

Using Sylvester's determinant identity:

$$\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA}), \tag{56}$$
$$\mathbf{A} = \overline{\mathbf{X}}\mathbf{\Psi}\mathbf{D}_w$$
$$\mathbf{B} = \mathbf{D}_w^T\mathbf{\Psi}^T\overline{\mathbf{X}}^T. \tag{57}$$

Therefore, (55) can be rewritten as

$$\left|\frac{z_i^2}{\sigma_n^2}\mathbf{D}_w^T\mathbf{\Psi}^T\overline{\mathbf{X}}^T\overline{\mathbf{X}}\mathbf{\Psi}\mathbf{D}_w + \mathbf{I}\right| \tag{58}$$

and (31) follows.

Appendix C
Transmit Convolution Matrix

A. Observation 1

Proof. Since $r[-k] = r[k]$, the lower triangular portion of $\mathbf{R_X}$ is positive and the matrix is Hermitian. Finally, as, $\mathbf{R_X}$, is a covariance matrix, it follows that it is also positive semi-definite. □

B. Corollary 1

Proof. Since $\mathbf{R_X}$ and L are given, this problem becomes equivalent to finding $\mathbf{x}$ to satisfy the equality $\mathbf{R_X} = \mathbf{C}(\mathbf{x}, L)^T\mathbf{C}(\mathbf{x}, L)$. As was described in Observation 1, the entries of $\mathbf{R_X}$ correspond to a length $2(L-1)+1$ autocorrelation function, $\mathbf{r}$, of $\mathbf{x}$. Therefore, the equality will be satisfied if it is possible to find an $\mathbf{x}$ with the given autocorrelation sequence. Since

$$\mathcal{F}\{\mathbf{r}\} = |X_0(\omega)|^2, \tag{59}$$

where $X_0(\omega)$ is the discrete Fourier transform for $\mathbf{x}_0$. Therefore by taking the square root and the inverse discrete Fourier transform the expression $\mathbf{x}_0$ is produced. □

Appendix D
Convexity of Lower Bound

A. Lemma 2

Proof. Taking the second derivative of $g(t)$, we have

$$g''(t) = \frac{(1+\lambda t + f(t))f''(t) - f'(t) - \lambda}{(1+\lambda t + f(t))^2}. \tag{60}$$

For $g(t)$ to be concave, (60) must non-positive. Noting first that the denominator of (60) is strictly positive, the condition $g''(t) \leq 0$ is sufficient to produce the inequality (40) and thus the lemma. □

B. Lemma 3

Proof. To begin the proof, a simpler version of the equation in question is studies. Define $\tilde{h}(t)$ as $h(t)$ with $d_{\ell,j}$ changed to be d_ℓ, therefore

$$\tilde{h}(t) = \log\left(\sum_j p_z(z_j) z_j^{-2D}\right) - \sum_{\ell=1}^{D} \log\left(1 + d_\ell + \lambda_\ell t\right). \tag{61}$$

This rearrangement is valid as the product no longer depends on j and can be separated from the sum. It can be show that (61) is convex in t [41]. Thus, by mapping (42) to (61) the convexity would follow. This can be accomplished using the following definition of d_ℓ that depends on t

$$d_\ell(t) = \frac{1}{\sum_j \beta_j [1 + d_{\ell,j} + \lambda_\ell t]^{-1}} - (1 + \lambda_\ell t), \tag{62}$$

where $\beta_j = \frac{p_z(z_j)}{z_j^{2D}} \left[\sum_j \frac{p_z(z_j)}{z_j^{2D}} \right]^{-1}$. It can be shown that substituting $d_\ell(t)$ for d_ℓ in (61) gives $h(t) = \tilde{h}(t)$; thus, the two expressions are equivalent. Therefore, $\tilde{h}(t)$ can be rewritten as

$$\tilde{h}(t) = \log \left(\sum_j p_z(z_j) z_j^{-2D} \right) - \sum_{\ell=1}^{D} \log \left(1 + d_\ell(t) + \lambda_\ell t\right). \tag{63}$$

Since it can be verified that $d_\ell(t) \geq 0$, $d'_\ell(t) \geq 0$ and $d''_\ell(t) \leq 0$, by looking at the second term of (63) and Lemma 2 the full expression can be said to be convex. □

Using Lemma 2, and Lemma 3, the lower bound is shown to be convex in Lemma 4.

C. Lemma 4

Proof. To prove concavity, the convexity of $-f(\mathbf{X})$ is proven. Since a non-negative weighted sum of convex functions is convex it remains to show the convexity of

$$w(\mathbf{X}) = \log \left(\sum_j p_z(z_j) \left[\det(\mathbf{I} + (z_i^2 + z_j^2)\mathbf{X}) \right]^{-1} \right). \tag{64}$$

For simplicity, we let $\tilde{z}_j^2 = z_i^2 + z_j^2$ as it is now assumed that the index i is fixed. To prove the convexity of this expression, a commonly used method of checking the convexity using one variable is employed. This is done by seeing if $w(\mathbf{X} + t\mathbf{V})$ is convex in t, for any positive semidefinite matrices $\mathbf{X}$ and $\mathbf{V}$. Expanding out $w(\mathbf{X} + t\mathbf{V})$ results in

$$\begin{aligned}
&w(\mathbf{X} + t\mathbf{V}) = \\
&\log \left(\sum_j p_z(z_j) \left[\det(\mathbf{I} + \tilde{z}_j^2(\mathbf{X} + t\mathbf{V})) \right]^{-1} \right) \\
&= -\log\det(\mathbf{X}) + \\
&\log \left(\sum_j p_z(z_j) \left[\det\left(\mathbf{X} + \tilde{z}_j^2(\mathbf{I} + t\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}) \right) \right]^{-1} \right) \\
&= -\log\det(\mathbf{X}) + \\
&\log \left(\sum_j p_z(z_j) \left[\prod_{\ell=1}^{D} \left(\tilde{z}_j^2(1 + \lambda_l t) + \lambda_{x,\ell} \right) \right]^{-1} \right)
\end{aligned} \tag{65}$$

where $\lambda_l \geq 0$ and $\lambda_{x,\ell} \geq 0$ are the eigenvalues of $\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}$ and $\mathbf{X}$. Pulling out the $\tilde{z}_j^2$ from the product results in

$$\begin{aligned}
&= -\log\det(\mathbf{X}) \\
&+ \log \left(\sum_j p_z(z_j) \tilde{z}_j^{-2D} \left[\prod_{\ell=1}^{D} \left(1 + \lambda_l t + \frac{\lambda_{x,\ell}}{\tilde{z}_j^2} \right) \right]^{-1} \right).
\end{aligned} \tag{66}$$

where $d_{\ell,j} = \frac{\lambda_{x,\ell}}{\tilde{z}_j^2}$. Therefore, the second term is of the form in Lemma 3 and is convex in t. □

REFERENCES

[1] K. Gu, G. Wang, and J. Li, "Migration based SAR imaging for ground penetrating radar systems," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 151, no. 5, pp. 317–325, October 2004.

[2] F. Gini, A. D. Maio, and L. Patton, Eds., *Waveform Design and Diversity for Advanced Radar Systems*. Stevenage, UK: The Institution of Engineering and Technology, 2012.

[3] M. R. Bell, "Information theory and radar waveform design," *IEEE Transactions on Information Theory*, vol. 39, no. 5, pp. 1578–1597, September 1993.

[4] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[5] R. A. Romero, J. Bae, and N. A. Goodman, "Theory and application of SNR and mutual information matched illumination waveforms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 912–927, April 2011.

[6] A. Leshem, O. Naparstek, and A. Nehorai, "Information theoretic adaptive radar waveform design for multiple extended targets," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 42–55, June 2007.

[7] B. Tang and J. Li, "Spectrally constrained MIMO radar waveform design based on mutual information," *IEEE Transactions on Signal Processing*, vol. 67, no. 3, pp. 821–834, February 2019.

[8] L. Wang, L. Wang, Y. Zeng, and M. Wang, "Jamming power allocation strategy for MIMO radar based on MMSE and mutual information," *IET Radar, Sonar and Navigation*, vol. 11, no. 7, pp. 1081–1089, 2017.

[9] A. Farina and F. A. Studer, "Detection with high resolution radar - Great promise, big challenge," *Microwave Journal*, vol. 34, pp. 263–273, May 1991.

[10] Y. Yang and R. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 330–343, January 2007.

[11] Y. Gu and N. A. Goodman, "Information-theoretic waveform design for Gaussian mixture radar target profiling," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 3, pp. 1528–1536, June 2019.

[12] F. Xin, B. Wang, Z. Xu, and X. Chen, "A novel algorithm for optimal waveform design based on dual mutual information for radar systems," *Journal of Physics: Conference Series*, vol. 1324, p. 012022, October 2019.

[13] Y. Gu and N. A. Goodman, "Information-theoretic waveform design for gaussian mixture radar target profiling," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 3, pp. 1528–1536, 2019.

[14] Z. Zhu, S. Kay, and R. S. Raghavan, "Information-theoretic optimal radar waveform design," *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 274–278, March 2017.

[15] M. Wainwright and E. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," *Advances in Neural Information Processing Systems*, vol. 12, pp. 855–861, 1999.

[16] Z. Chance, R. G. Raj, and D. J. Love, "Information-theoretic structure of multistatic radar imaging," in *Proc. 2011 IEEE Radar Conference (RADAR)*, Kansas City, MO, 2011, pp. 853–858.

[17] S. U. Pillai, D. C. Youla, H. S. Oh, and J. R. Guerci, "Optimum transmit-receiver design in the presence of signal-dependent interference and channel noise," in *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 1999, pp. 870–875 vol.2.

[18] S. U. Pillai, H. S. Oh, and J. R. Guerci, "Multichannel matched transmit-receiver design in presence of signal-dependent interference and noise," in *Proceedings of the 2000 IEEE Sensor Array and Multichannel Signal Processing Workshop*, Cambridge, MA, March 2000, pp. 385–389.

[19] D. A. Garren, M. K. Osborn, A. C. Odom, J. S. Goldstein, S. U. Pillai, and J. R. Guerci, "Enhanced target detection and identification via optimised radar transmission pulse shape," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 148, no. 3, pp. 130–138, June 2001.

[20] D. A. Garren, A. C. Odom, M. K. Osborn, J. S. Goldstein, S. U. Pillai, and J. R. Guerci, "Full-polarization matched-illumination for target detection and identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 824–837, July 2002.

[21] K. J. Sangston, F. Gini, M. V. Greco, and A. Farina, "Structures for radar detection in compound Gaussian clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 445–458, April 1999.

[22] K. J. Sangston and A. Farina, "Coherent radar detection in compound-Gaussian clutter: Clairvoyant detectors," *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 11, pp. 42–63, November 2016.

[23] Z. Idriss, R. G. Raj, and R. M. Narayanan, "A compound Gaussian-based waveform design approach for enhanced target detection in multistatic radar imaging," in *Proc. SPIE Conference on Radar Sensor Technology XXIII*, vol. 11003, Baltimore, MD, April 2019.

[24] Y. Yao, J. Zhao, and L. Wu, "Waveform optimization for target estimation by cognitive radar with multiple antennas," *Sensors*, vol. 18, no. 6, p. 1743, 2018.

[25] A. Farina and F. A. Studer, "Detection with high resolution radar: Advanced topics potential applications," *Journal of Systems Engineering and Electronics*, vol. 3, no. 1, pp. 21–34, 1992.

[26] W. C. Weng, F. Yang, and A. Elsherbeni, *Electromagnetics and Antenna Optimization using Taguchi's Method*. Morgan Claypool, 2008.

[27] R. C. Eberhart and Y. Shi, "Particle swarm optimization: developments, applications and resources," in *Proceedings of the 2001 Congress on Evolutionary Computation*, vol. 1, Seoul, South Korea, May 2001, pp. 81–86 vol. 1.

[28] S. Boyd and J. Dattorro, "Alternating projections," Stanford University, Stanford, CA, Tech. Rep. EE392o, 2003. [Online]. Available: https://web.stanford.edu/class/ee392o/alt_proj.pdf

[29] D. C. Munson Jr., J. O'Brien, and W. K. Jenkins, "A tomographic formulation of spotlight-mode synthetic aperture radar," *Proc. of the IEEE*, vol. 71, no. 8, pp. 917–925, August 1983.

[30] C. V. Jakowatz, Jr., D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. New York, NY: Kluwer Academic Publishers, 1996.

[31] J. W. Crispin, R. F. Goodrich, and K. M. Siegel, "A theoretical method for the calculation of radar cross section of aircraft and missiles," University of Michigan Radiation Laboratory, Ann Arbor, MI, Rep. 2591-1-H, July 1959.

[32] N. J. Willis, *Bistatic Radar*. Raleigh, North Carolina: SciTech Publishing, 2005.

[33] R. E. Kell, "On the derivation of bistatic RCS from monostatic measurements," *Proc. of the IEEE*, vol. 53, no. 8, pp. 983–988, August 1965.

[34] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338–1351, November 2003.

[35] M. Wainwright, E. Simoncelli, and A. Willsky, "Random cascades on wavelet trees and their use in analyzing and modeling natural images," *Applied and Computational Harmonic Analysis*, vol. 11, no. 1, pp. 89–123, July 2001.

[36] J. R. Cornelius and T. Dogaru, "Modeling the radar signature of thin metallic objects with the AFDTD software," US Army Research Laboratory, Adelphi, MD, Tech. Rep. ARL-TR-7063, September 2014.

[37] S. Kay, "Optimal signal design for detection of Gaussian point targets in stationary Gaussian clutter/reverberation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 31–41, June 2007.

[38] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Hoboken, NJ: Wiley-Interscience, 2006.

[39] M. F. Huber, T. Bailey, H. Durrant-Whyte, and U. D. Hanebeck, "On entropy approximation for Gaussian mixture random vectors," in *Proc. IEEE International Conf. on Multisensor Fusion and Integration for Intelligent Systems*, Seoul, South Korea, August 2008, pp. 181–188.

[40] A. G. Akritas, E. K. Akritas, and G. I. Malaschonok, "Various proofs of Sylvester's (determinant) identity," *Mathematics and Computers in Simulation*, vol. 42, no. 4-6, pp. 585 – 593, November 1996.

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

[42] S. Boyd, "Alternating projections," class notes for EE592o, Department of Electrical Engineering, Stanford University, Stanford, CA, Autumn 2003.

[43] S. U. Pillai, K. Y. Li, and H. Beyer, "Reconstruction of constant envelope signals with given Fourier transform magnitude," in *Proc. 2009 IEEE Radar Conference*, Pasadena, CA, May 2009.

[44] H. H. Bauschke and J. M. Borwein, "On the convergence of von Neumann's alternating projection algorithm for two sets," *Set-Valued Analysis*, vol. 1, no. 2, pp. 185–212, Jun 1993.

[45] A. S. Hedayat, N. J. A. Sloane, and J. Stufken, *Orthogonal Arrays: Theory and Applications*. New York, NY: Springer-Verlag, 1999.

[46] Y. Fang, M. Cheney, and S. Roecker, "Multistatic radar imaging from sparse measurements," in *Proc. 2010 IEEE Radar Conference*, Washington, DC, May 2010, pp. 485–490.

[47] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1-regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.

[48] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.

[49] J. Duda, "Gaussian auto-encoder," arXiv 1811.04751 [cs], v4, January 2019.