
Article:
Zhang, J. orcid.org/0000-0002-6439-8834, Xu, H. orcid.org/0000-0002-9559-3691, Liu, W. orcid.org/0000-0003-2968-2888 et al. (3 more authors) (2022) Joint design of SAR waveform and imaging filters based on target information maximization. IEEE Journal of Selected Topics in Signal Processing. ISSN 1932-4553

https://doi.org/10.1109/jstsp.2022.3217895

Joint Design of SAR Waveform and Imaging Filters Based on Target Information Maximization

Jiawei Zhang, Huaping Xu, *Member, IEEE*, Wei Liu, *Senior Member, IEEE*, Chunsheng Li, Liang Feng and Yifan Chen, *Senior Member, IEEE*

***Abstract*—In this paper, a joint design of both transmit waveform and imaging filters for synthetic aperture radar (SAR) is proposed to improve its information acquisition capability. First, the mutual information between SAR image and target scattering characteristics is considered as the performance metric, and its equivalent analytical version is derived in the 2-D frequency domain. The design is formulated as an optimization problem with an energy condition and a similarity constraint. Then, to tackle the resultant non-convex problem, by referring to the Dinkelbach's method, an algorithm is derived to find the desired solution via a cyclic maximization procedure alternating between three subproblems. Based on minorization-maximization, a unified optimization method with an increasing penalty on constraint violation is proposed to solve all subproblems. Convergence of the developed algorithm is analytically proved. Finally, numerical examples are presented to demonstrate the effectiveness of the proposed design.**

***Index Terms*—SAR, transmit waveform, imaging filter, information acquisition, constrained optimization.**

I. Introduction

BY transmitting a wideband waveform periodically and receiving backscatter echoes of the observed scene, synthetic aperture radar (SAR) can generate high-resolution 2-D images, which are widely used for surveillance and monitoring of Earth surface [1–3]. With the improvement of resolution, even small and closely spaced targets become visible and identifiable, which opens up a gateway to a number of advanced applications, including aircraft detection [4], ship classification [5], vehicle recognition [6], and so on. The use scenarios for SAR are no longer limited to traditional large-scale topographic mapping and point target discovery, but also developed to description and perception of interested targets, whose performance depends on the amount of acquired information about the observed target. As a result, information acquisition for SAR has become an increasingly important problem.

Since the principle of synthetic aperture was introduced, one emphasis in the application of SAR has been the detection of point-like targets, where the signal-to-noise ratio (SNR) has been the primary metric used in SAR waveform selection and imaging [7]. The most common waveform employed in SAR is the linear frequency modulated (LFM) signal, whose range profile obtained through matched filtering is the sinc function. To reduce its sidelobes, weighted windowing, adaptive filtering, and other optimization techniques have been introduced [8]. Moreover, another SAR waveform is generated by a nonlinear frequency modulated (NLFM) synthesis method [9–11]. A proof-of-principle experiment was performed to demonstrate the potential of NLFM pulses based on a real SAR platform on the ground [9]. Using multi-objective optimization, a piecewise linear frequency modulated signal was generated for NLFM design, leading to a higher SNR and lower sidelobes compared with the LFM signal. Based on the principle of stationary phase, an NLFM waveform was obtained to suppress the undesired sidelobes, and this scheme was demonstrated by an airborne SAR experiment [10]. An advanced NLFM waveform optimization framework was reported in [11], with higher SNR gain achieved.

As can be seen, the SAR waveform design has always been focused on higher SNR and lower sidelobes, for better point target detection capability. However, for extended targets, information acquisition is the key factor [12]. Meanwhile, with the improvement in wideband signal generation and wideband processing, resolution of SAR has experienced a substantial increase, which transforms many originally point-like targets into extended ones. As early as in 1964, Woodward pointed out that pursuing SNR blindly can mislead radar design and data processing because there is no theory implying that maximizing SNR can ensure maximal information acquisition [13]. A similar issue was also addressed by Bell in [14]. Target detection, classification and recognition with a high confidence level, are still a difficult problem, even with high or ultra-high resolution SAR images [15]. Although this problem has been widely studied over the years, it is still far from being fully solved [16]. It is clear that the present SAR waveforms are not optimal for target information acquisition, and the associated SAR imaging algorithms employing a matched filter associated with the waveform are also not oriented for target information. Under the narrowband SAR system where the point scatter model is usually sufficient, maximizing SNR is desired in a general sense, and information acquisition of target details does not present much importance. However, the contradiction between the SNR criterion and information acquisition requirement for extended targets is more significant in the wideband case. The SAR optimization criterion for a

H. Xu is the corresponding author.

J. Zhang, H. Xu, and C. Li are with the School of Electronic and Information Engineering, Beihang University, Beijing 100191, China (e-mail: jwzhang@buaa.edu.cn, xuhuaping@buaa.edu.cn, lichunsheng@buaa.edu.cn).

W. Liu is with the Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield S1 3JD, U.K.(e-mail: w.liu@sheffield.ac.uk).

L. Feng is with the Bejing Institute of Radio Measurement, Beijing 100854, China (e-mail: liangf0325@163.com).

Y. Chen is with the School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China. He is also with the School of Engineering, University of Waikato, Hamilton 3240, New Zealand (e-mail: yifan.chen@uestc.edu.cn).

larger amount of information should be considered, especially for a wideband system.

SAR waveform design based on the maximizing target information criterion (MTIC) is rarely studied, while similar activities have already been carried out for various non-imaging radars, such as single waveform radar [17], multi-input multi-output radar [18] and cognitive radar [19]. Meanwhile, the MTIC has been involved in various application scenarios. For example, a MTIC-based waveform was implemented to perform target recognition from a set of possible alternatives [19]. It was shown that waveform design by maximizing mutual information (MI) can provide more desired estimation and classification capabilities [20]. Such a criterion was also employed in target class discrimination applications, leading to better classification results [21]. It demonstrates that the information acquisition capability of an active sensor can be improved by judiciously designing its transmit waveform [22, 23]. It is not easy to extend these works to SAR, since it is a 2-D imaging radar with some unique requirements, such as desired imaging filters, high resolution (wideband property) and low sidelobe level. In [24], only a suboptimal wideband waveform was designed to improve information acquisition of high-resolution radar, while the receive filter was still based on the matched filter derived by maximizing SNR. As a result, significant performance degradation may occur, as the receive filter is not optimized for information acquisition. Although this scheme was improved by a joint design of waveform and receive filter, the optimization in azimuth filter was ignored [25].

In this article, a novel method is developed to jointly design the SAR transmit waveform and imaging filters under the MTIC. To our best knowledge, this is the first time to address the joint design problem for information acquisition maximization of SAR. The MI between the SAR image and the target scattering characteristic is employed as the performance metric, which is transformed into an equivalent version of the Pearson correlation coefficient. An analytical expression of the objective function is then established in the 2-D frequency domain with detailed derivations. Combined with a similarity requirement in both directions, the constrained optimization problem is finally formulated.

The resultant problem is non-convex, with a fractional multivariate objective function and multiple constraints. To tackle the problem, referring to the Dinkelbach's method, the numerator and denominator of the objective function are decoupled by an auxiliary function. By using the cyclic maximization procedure, the multivariate problem is divided into three subproblems with respect to waveform, range filter and azimuth filter, respectively. Based on minorization-maximization (MM), the non-convex subproblems are converted to convex forms, with an increasing penalty on constraint violations. They are then solved alternately, and their convergence is analyzed. Numerical examples show that the joint design result synthesized by the proposed algorithms outperform their counterparts.

The rest of this paper is organized as follows. The SAR information acquisition model and the joint design problem are formulated in Section II. The proposed solution with detailed analysis is presented in Section III. Numerical results are provided in Section IV, and conclusions are drawn in Section V.

Notation: Throughout this article, matrices are denoted by bold uppercase letters, and vectors by bold lowercase letters. $\mathbb{N}$ and $\mathbb{R}$ denote the set of natural numbers and real numbers, respectively. $\mathrm{E}[\cdot]$ is the expectation operation. $\lambda_{\min}[\mathbf{A}]$ is the minimum eigenvalue of matrix $\mathbf{A}$. $\mathscr{F}_\tau$ and $\mathscr{F}_\eta$ are the Fourier transform in the range and azimuth directions, respectively. The symbol $\mathrm{tr}[\mathbf{A}]$ indicates the trace of $\mathbf{A}$. The superscript $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ represent the vector/matrix transpose, complex conjugate and the Hermitian transpose, separately. The notation $|\cdot|$ indicates the modulus of a complex scalar. $||\cdot||$ is the Euclidean norm of a vector. $\mathrm{Diag}(\mathbf{a})$ represents the diagonal matrix formed by components of the vector $\mathbf{a}$. The notation $\otimes$ represents the 2-D convolution operator, and $\odot$ is the Hadamard (element-wise) product of vectors/matrices.

II. Signal Model and Problem Formulation

A. SAR Signal Model

A SAR mounted on a moving platform is similar to a conventional radar, whose waveform is sequentially transmitted and backscattered echoes are collected by the radar antenna. A SAR waveform is generally written as

$$s(\tau) = \omega_r(\tau)\cos\left[2\pi f_0 \tau - \varphi(\tau)\right] \tag{1}$$

where τ, f_0 and $\varphi(\tau)$ are the range time, center frequency and phase code. $\omega_r(\tau)$ is a window function, and it varies slowly with respect to τ. As the platform advances, a pulse is transmitted at regular intervals. Meanwhile, the resultant SAR echo is written into successive rows, which is then processed by quadrature demodulation to generate 2-D baseband echo. The 2-D SAR echo of a point target with unity amplitude is the system impulse response [26]

$$\begin{aligned} h_{\mathrm{E}}(\tau,\eta) \approx\ & \omega_a(\eta)\exp\left(-j4\pi R_0/\lambda\right) \\ & \cdot\exp\left(-j\pi K_a\eta^2\right) x\left(\tau - 2R(\eta)/c\right) \end{aligned} \tag{2}$$

where η, λ, c, $\omega_a(\eta)$, R_0, K_a, $R(\eta)$ and $x(\tau)$ represent the azimuth time, wavelength, speed of light, azimuth beam pattern, slant range of scene center, azimuth frequency modulation, instantaneous slant range and complex baseband waveform $x(\tau) = \omega_r(\tau)\exp(j\varphi(\tau))$, respectively.

A general observation scene provides the baseband SAR signal data:

$$s_0(\tau,\eta) = [\sigma(\tau,\eta) + b(\tau,\eta)] \otimes h_{\mathrm{E}}(\tau,\eta) + n(\tau,\eta) \tag{3}$$

where the scene reflectivity is formed by superimposition between target $\sigma(\tau,\eta)$ and background $b(\tau,\eta)$. $n(\tau,\eta)$ is additive noise. Note that background $b(\tau,\eta)$ is defined as the part of the scene that the observer is not interested in. So, the observed scene is the superposition of target and background.

With this baseband 2-D echo, SAR imaging processing is viewed as a deconvolution process. Differing from low-resolution algorithm with high efficiency, *e.g.* SPECtral ANalysis (SPECAN) [27], more sophisticated imaging processing is considered. Referring to the Range Doppler (RD) algorithm, the processing $r(\tau,\eta)$ involves four steps: range filtering $m(\tau)$,

range cell migration correction (RCMC), phase compensating $h_{\mathrm{P}}(\eta)$, and azimuth filtering $w(\eta)$, so $r(\tau,\ \eta)$ is formed by cascading all these steps. Detailed derivations about $r(\tau,\ \eta)$ is presented in Section A of Appendix, while $r(\tau,\ \eta)$ is given as

$$r(\tau,\eta) = m(\tau) \otimes g_{\mathrm{RCMC}}(\tau,\eta) \otimes h_{\mathrm{P}}(\eta) \otimes w(\eta) \quad (4)$$

where $g_{\mathrm{RCMC}}(\tau,\ \eta)$ and $h_{\mathrm{P}}(\eta)$ are the inverse Fourier transform of the RCMC operation $G_{\mathrm{RCMC}}(f_\tau\ ,\ f_\eta)$ and phase compensation $H_{\mathrm{P}}(f_\eta)$ in (31) and (32), respectively, with f_τ and f_η denoting range and azimuth frequency, respectively. Then, $h_{\mathrm{E}}(\tau,\ \eta)$ and $r(\tau,\ \eta)$ are cascaded to obtain the whole SAR impulse response, given as

$$\begin{aligned} h(\tau,\eta) &= h_{\mathrm{E}}(\tau,\eta) \otimes r(\tau,\eta) \\ &= \exp\left(-j4\pi R_0/\lambda\right) x\left(\tau\right) \otimes m\left(\tau\right) \otimes \delta\left(\tau - 2R_0/c\right) \otimes w(\eta) \end{aligned}$$

where $\delta(\tau)$ is the Dirac delta function. A SAR image is obtained as

$$s_{\mathrm{I}}(\tau,\eta) = [\sigma(\tau,\eta)+b(\tau,\eta)] \otimes h(\tau,\eta) + n(\tau,\eta) \otimes r(\tau,\eta) \quad (5)$$

Similar to wireless communication systems, the above SAR information acquisition process can be viewed as a memoryless channel as described in Fig. 1 [28, 29]. The target scattering characteristics $\sigma(\tau,\ \eta)$ determine the information source, and the channel is determined by the medium through which information is transferred from the source to the destination which is the resultant SAR image $s_{\mathrm{I}}(\tau,\ \eta)$. It is expected that the joint design of $x(\tau)$, $m(\tau)$ and $w(\eta)$ can provide reliable communication, which means that target information is extracted by SAR image with arbitrarily small distortion.

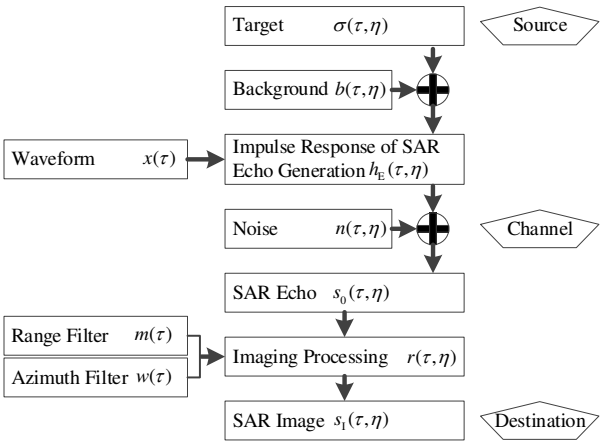

Figure 1. Information acquisition for SAR.

B. Problem Formulation

A SAR system makes measurements of a target in order to determine its unknown characteristics. In other words, SAR observes a target aiming at acquiring information about it. From Shannon theory, MI can tell us the amount of information a SAR image provides about the target, and we can use it to evaluate information acquisition quantitatively. The MI I_σ between the target scattering coefficient $\sigma(\tau,\eta)$ and SAR image $s_{\mathrm{I}}(\tau,\eta)$ is considered as performance index to be improved. To this end, taking waveform, range filter and azimuth filter as decision variables, the joint design can be written as a maximization problem $\max I_\sigma$. It is worth noting that the optimization in range/azimuth direction cannot be performed separately. Firstly, the range and azimuth coordinates are coupled in the raw signal domain, so it is not easy to derive a closed-form objective for the decision variable in one single direction. Moreover, even if an approximate closed-form optimization problem is given in each direction, separately, the resultant imaging result is inferior to the joint design scheme where information acquisition in the SAR image domain is maximized.

It is usually assumed that $\sigma(\tau,\eta)$, $b(\tau,\eta)$ and $n(\tau,\eta)$ are statistically independent, and follow a zero-mean complex Gaussian distribution [18, 19, 21, 29, 30]. Based on the MI definition, I_σ can be simplified into [31]

$$I_\sigma = -\ln\left(1 - |\rho_\sigma|^2\right) \quad (6)$$

where ρ_σ is defined as the Pearson correlation coefficient. It can be seen that the MI is only related to $|\rho_\sigma|^2$ and is a monotonically increasing function of the modulus. Therefore, maximizing I_σ is equivalent to maximizing $|\rho_\sigma|^2$ in the associated optimization problem, with

$$|\rho_\sigma|^2 = |\rho_1|^2/(\rho_2\rho_3)\ , \quad (7)$$

where

$$\begin{aligned} &\rho_1 = \mathrm{E}\left\{[\sigma(\tau,\eta) \otimes h(\tau,\eta)]\,\sigma^*(\tau,\eta)\right\}, \rho_2 = \mathrm{E}\left[\sigma(\tau,\eta)\sigma^*(\tau,\eta)\right], \\ &\rho_3 = \mathrm{E}\left\{[\sigma(\tau,\eta) \otimes h(\tau,\eta)]\,[\sigma^*(\tau,\eta) \otimes h^*(\tau,\eta)]\right. \\ &\quad + [b(\tau,\eta) \otimes h(\tau,\eta)][b^*(\tau,\eta) \otimes h^*(\tau,\eta)] \\ &\quad \left. + [n(\tau,\eta) \otimes r(\tau,\eta)][n^*(\tau,\eta) \otimes r^*(\tau,\eta)]\right\}. \end{aligned}$$

With detailed derivations in Section B of Appendix, in the frequency domain, we have

$$\begin{aligned} \rho_1 &= \iint P_\sigma(f_\tau,f_\eta)H(f_\tau,f_\eta)df_\tau df_\eta \\ \rho_2 &= \iint P_\sigma(f_\tau,f_\eta)df_\tau df_\eta \\ \rho_3 &= \iint \Big\{[P_\sigma(f_\tau,f_\eta) + P_b(f_\tau,f_\eta)]\,|H(f_\tau,f_\eta)|^2 \\ &\quad + P_n(f_\tau,f_\eta)|R(f_\tau,f_\eta)|^2\Big\} df_\tau df_\eta, \end{aligned} \quad (8)$$

where $P_\sigma(f_\tau,f_\eta)$, $P_b(f_\tau,f_\eta)$ and $P_n(f_\tau,f_\eta)$ denote the power spectrum density (PSD) of $\sigma(\tau,\eta)$, $b(\tau,\eta)$ and $n(\tau,\eta)$, respectively. ρ_2 is the variance of $\sigma(\tau,\eta)$, and thus a constant. To simplify the presentation, we choose $\rho_2 = 1$, which does not affect the following analysis. Meanwhile, $H(f_\tau,f_\eta)$ and $R(f_\tau,f_\eta)$ are the 2-D Fourier transform of $h(\tau,\eta)$ and $r(\tau,\eta)$, respectively, given as

$$\begin{aligned} H(f_\tau,f_\eta) &= \exp\left(-j4\pi R_0/\lambda\right) \\ &\quad \cdot \exp\left(-j4\pi R_0/c\ f_\tau\right) X(f_\tau)M(f_\tau)W(f_\eta) \end{aligned} \quad (9)$$

and

$$R(f_\tau,f_\eta) = M(f_\tau)G_{\mathrm{RCMC}}(f_\tau,f_\eta)H_{\mathrm{P}}(f_\eta)W(f_\eta),$$

with $X(f_\tau) = \mathscr{F}_\tau[x(\tau)]$, $M(f_\tau) = \mathscr{F}_\tau[m(\tau)]$ and $W(f_\eta) = \mathscr{F}_\eta[W(\eta)]$.

(8) is still very complex, so further simplification is needed. The integral in (8) is firstly approximated by summation. Then, $H(f_\tau, f_\eta)$ and $R(f_\tau, f_\eta)$ are simplified, w here s everal terms which have no effect on the optimization problem are ignored, such as the constant, time delay and unimodular phase-compensation functions. Finally, the summation is expressed in a matrix form. For detailed derivations, please refer to Section C of Appendix. The objective function $\rho(\mathbf{x}, \mathbf{m}, \mathbf{w})$ is then derived as

$$|\rho_\sigma|^2 \approx \rho(\mathbf{x}, \mathbf{m}, \mathbf{w}) = \frac{\alpha(\mathbf{x}, \mathbf{m}, \mathbf{w})}{\beta(\mathbf{x}, \mathbf{m}, \mathbf{w})} \tag{10}$$

where

$$\alpha(\mathbf{x}, \mathbf{m}, \mathbf{w}) = (\mathbf{x} \odot \mathbf{m})^H \mathbf{p}_{\sigma\tau} \mathbf{p}_{\sigma\tau}^T (\mathbf{x} \odot \mathbf{m}) \mathbf{w}^H \mathbf{p}_{\sigma\eta} \mathbf{p}_{\sigma\eta}^T \mathbf{w},$$

$$\begin{aligned}\beta(\mathbf{x}, \mathbf{m}, \mathbf{w}) = {} & (\mathbf{x} \odot \mathbf{m})^H \mathrm{Diag}(\mathbf{p}_{\sigma\tau}) (\mathbf{x} \odot \mathbf{m}) \mathbf{w}^H \mathrm{Diag}(\mathbf{p}_{\sigma\eta}) \mathbf{w} \\ & + (\mathbf{x} \odot \mathbf{m})^H \mathrm{Diag}(\mathbf{p}_{b\tau}) (\mathbf{x} \odot \mathbf{m}) \mathbf{w}^H \mathrm{Diag}(\mathbf{p}_{b\eta}) \mathbf{w} \\ & + \mathbf{m}^H \mathrm{Diag}(\mathbf{p}_{n\tau}) \mathbf{m} \mathbf{w}^H \mathrm{Diag}(\mathbf{p}_{n\eta}) \mathbf{w}.\end{aligned}$$

To maximize $\rho(\mathbf{x}, \mathbf{m}, \mathbf{w})$ with respect to the frequency-domain waveform $\mathbf{x}$, range filter $\mathbf{m}$ and azimuth filter $\mathbf{w}$, the joint design under the MTIC is then established as $\max\ \rho(\mathbf{x}, \mathbf{m}, \mathbf{w})$. Without constraints it may lead to a waveform-filter pair with poor resolution or unrealizability [32], so it is necessary to impose some reasonable constraints on $\mathbf{x}$, $\mathbf{m}$ and $\mathbf{w}$.

In general, SAR has several basic performance metrics or requirements, such as high resolution (wideband property) and low sidelobe [33]. Although maximizing MI is considered as the main objective, these common metrics cannot be compromised too much, because they are related to SAR image quality and specific applications. It is expected that MI is improved as much as possible, while these requirements are also maintained to some degree. To this end, by referring to the similarity constraints used in pulse radar and MIMO radar [34–37], the similarity constraints in both range and azimuth directions are introduced to SAR waveform-filter optimization. Firstly, it is imposed on waveform $\mathbf{x}$ with the classic LFM signal $\mathbf{c}$, i.e.

$$\|\mathbf{x} - \mathbf{F}\mathbf{c}\|^2 \le \varepsilon_1$$

where $\mathbf{F}$ and ε_1 denote the Fourier transform matrix and the similarity parameter, respectively. It guarantees that the resultant waveform will maintain some outstanding characteristics of $\mathbf{c}$, such as insignificant modulus variation, high resolution and Doppler tolerance. Then, a similarity constraint also needs to be imposed on the range filter. Since resolution and sidelobe level rely on range profile which involves both waveform and range filter, the newly designed range filter should help the filtered signal provide the desired resolution and sidelobe level, which can also be achieved by a similarity constraint with the matched filter of $\mathbf{c}$.

$$\|\mathbf{m} - \mathbf{F}^* \mathbf{c}^*\|^2 \le \varepsilon_2$$

where ε_2 is another similarity parameter. In the same way, let ε_3 be the similarity parameter in the azimuth direction, and the similarity requirement is also employed for $\mathbf{w}$ to obtain a similar azimuth profile:

$$\|\mathbf{w} - \mathbf{r}\|^2 \le \varepsilon_3$$

where $\mathbf{r}$ is the azimuth beam profile. These constraints in both directions will result in a similar range-azimuth profile with the classic SAR, so that performance in terms of these common metrics can be maintained as well. Besides, it can also relax the required prior information. A common waveform optimization method usually requires some prior information, while in practice the exact prior knowledge may not be available [18]. This mismatch will degrade the performance of designed waveform, but a similar imaging result can still be achieved even with totally inaccurate prior knowledge, because a similar range-azimuth profile is always guaranteed.

Finally, combining the objective function and constraints, the joint design can be cast as the following optimization problem $\mathcal{P}_1$:

$$\mathcal{P}_1 \begin{cases} \max\limits_{\mathbf{x},\mathbf{m},\mathbf{w}} \ \rho(\mathbf{x}, \mathbf{m}, \mathbf{w}) \\ \text{s.t.} \begin{cases} \mathbf{x}^H \mathbf{x} = E \\ \|\mathbf{x} - \mathbf{F}\mathbf{c}\|^2 \le \varepsilon_1 \\ \|\mathbf{m} - \mathbf{F}^* \mathbf{c}^*\|^2 \le \varepsilon_2 \\ \|\mathbf{w} - \mathbf{r}\|^2 \le \varepsilon_3 \end{cases} \end{cases} \tag{11}$$

where E is energy of the waveform.

III. Solving the Optimization Problem

The optimization problem $\mathcal{P}_1$ is non-convex with a fractional multivariate objective function and various constraints, and there is no easy solution from existing methods [38, 39]. In the following, a method is developed to tackle $\mathcal{P}_1$ effectively. Firstly, the fractional objective function is substituted by a quadratic function, and the multivariate problem is divided into three subproblems based on a cyclic maximizer. Then, a unified optimization method is proposed to solve all subproblems: by employing the MM algorithm, a surrogate function is found to minorize the objective function of subproblems, and the non-convex equality constraint is transformed into two convex constraints with an increasing penalty.

A. Algorithmic Procedure

Due to the fractional nature of the objective function whose numerator $\alpha(\mathbf{x}, \mathbf{m}, \mathbf{w})$ and denominator $\beta(\mathbf{x}, \mathbf{m}, \mathbf{w})$ are functions of variables to be optimized, we introduce an auxiliary function to decouple the numerator and denominator by referring to the Dinkelbach's algorithm [40, 41], i.e.

$$\begin{aligned}&\Psi\left(\mathbf{x}, \mathbf{m}, \mathbf{w}; \tilde{\mathbf{x}}, \tilde{\mathbf{m}}, \tilde{\mathbf{w}}\right) \\ &= \alpha(\mathbf{x}, \mathbf{m}, \mathbf{w}) - \beta(\mathbf{x}, \mathbf{m}, \mathbf{w}) \rho\left(\tilde{\mathbf{x}}, \tilde{\mathbf{m}}, \tilde{\mathbf{w}}\right).\end{aligned} \tag{12}$$

which is now considered as a new objective function for reformulating the problem $\mathcal{P}_1$. However, it is still a multivariate problem, and difficult to be solved jointly with respect to $\mathbf{x}$, $\mathbf{m}$ and $\mathbf{w}$. Based on the sequential optimization procedure, a multivariate optimization problem can be split into a number of suboptimization problems with only one variable, and they are alternately solved to generate a monotonically increasing objective function sequence [42].

1) Waveform Optimization : At the k-th step of iteration, given $\mathbf{m}=\tilde{\mathbf{m}}=\mathbf{m}^{(k-1)}$, $\mathbf{w}=\tilde{\mathbf{w}}=\mathbf{w}^{(k-1)}$ and $\tilde{\mathbf{x}}=\mathbf{x}^{(k-1)}$, $\mathbf{x}^{(k)}$ is expected to be found first. Without loss of generality, let $k=1$, so that the auxiliary function with respect to $\mathbf{x}$ is formed as

$$\Psi\left(\mathbf{x},\mathbf{m}^{(0)},\mathbf{w}^{(0)};\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)=\mathbf{x}^H\mathbf{A}_1\mathbf{x}+\mathbf{x}^H\mathbf{A}_2\mathbf{x}+\delta \tag{13}$$

where

$$\mathbf{A}_1=\mathbf{w}^{(0)H}\mathbf{p}_{\sigma\eta}\mathbf{p}_{\sigma\eta}^T\mathbf{w}^{(0)}\cdot\mathbf{p}_{\sigma\tau}\mathbf{p}_{\sigma\tau}^T\odot\left(\mathbf{m}^{(0)H}\mathbf{m}^{(0)}\right)^*,$$

$$\mathbf{A}_2=-\rho\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\cdot$$
$$\Big[\mathbf{w}^{(0)H}\mathrm{Diag}(\mathbf{p}_{\sigma\eta})\mathbf{w}^{(0)}\cdot\mathrm{Diag}(\mathbf{p}_{\sigma\tau})\odot\left(\mathbf{m}^{(0)}\mathbf{m}^{(0)H}\right)^*$$
$$+\mathbf{w}^{(0)H}\mathrm{Diag}(\mathbf{p}_{b\eta})\mathbf{w}^{(0)}\cdot\mathrm{Diag}(\mathbf{p}_{b\tau})\odot\left(\mathbf{m}^{(0)}\mathbf{m}^{(0)H}\right)^*\Big]$$

δ is a constant, and

$$\delta=-\rho\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\mathbf{m}^{(0)H}\mathrm{Diag}(\mathbf{p}_{n\tau})\mathbf{m}^{(0)}$$
$$\cdot\,\mathbf{w}^{(0)H}\mathrm{Diag}(\mathbf{p}_{n\tau})\mathbf{w}^{(0)}$$

Because adding or subtracting a constant to the objective function does not affect the solution to the optimization problem, δ has been ignored. Taking the constraints of $\mathbf{x}$ from $\mathcal{P}_1$, the first suboptimization problem is constructed as

$$\mathcal{P}_2\begin{cases}\max\limits_{\mathbf{x}} & \mathbf{x}^H\mathbf{A}_1\mathbf{x}+\mathbf{x}^H\mathbf{A}_2\mathbf{x}\\ \text{s.t.} & \begin{cases}\mathbf{x}^H\mathbf{x}=E\\ \|\mathbf{x}-\mathbf{F}\mathbf{c}\|^2\le\varepsilon_1\end{cases}\end{cases} \tag{14}$$

The first part of the objective function in $\mathcal{P}_2$ is convex, and the second one is concave. If the two are added together, it is not easy to determine whether the objective function is convex or concave. Moreover, the first constraint is non-convex, so both of them should be reformulated. Although all constraints of $\mathbf{x}$ become convex if $\mathbf{x}^H\mathbf{x}=E$ is relaxed to $\mathbf{x}^H\mathbf{x}\le E$, $\mathbf{x}^H\mathbf{x}=E$ is usually imposed on waveform optimization to make sure the energy always reaches the maximum for engineering considerations [32, 35, 43].

A general purpose semi-definite relaxation (SDR) solver is capable of converting $\mathcal{P}_2$ to a convex problem, but the worst-case complexity can be as high as $\mathcal{O}(N^{6.5})$ [44–46]. The MM algorithmic framework provides guidance in deriving problem-driven algorithms with low computational cost [47]. In the following, an MM-based algorithm is derived to handle the subproblem.

Starting from a feasible point, the MM procedure involves two steps. Assuming that we hope to tackle the maximization problem $\max f(\theta)$ over the feasibility set $\mathcal{Q}$. In the first minorization step, a surrogate function $g(\theta;\theta_{(t)})$ is derived at the t-th iteration, which locally approximates the objective function $f(\theta)$. $g(\theta;\theta_{(t)})$ is called a minorizer of $f(\theta)$, and it satisfies

$$g(\theta;\theta_{(t)})\le f(\theta),\ \text{for all }\theta\in\mathcal{Q};\quad g(\theta_{(t)};\theta_{(t)})=f(\theta_{(t)})$$

Then, in the maximization step, the surrogate function is maximized [48].

Using the MM optimization framework, one can find a surrogate function to replace the objective function in $\mathcal{P}_2$, so that a concave objective function is obtained for the maximization problem. Since a convex function is minorized by its supporting hyperplanes [49], we have

$$\mathbf{x}^H\mathbf{A}_1\mathbf{x}\ge\mathbf{x}_{(t)}^H\mathbf{A}_1\mathbf{x}_{(t)}+2\,\mathrm{Re}\left[\mathbf{x}_{(t)}^H\mathbf{A}_1(\mathbf{x}-\mathbf{x}_{(t)})\right] \tag{15}$$

where $\mathbf{x}_{(t)}$ is a feasible point. Thus, a minorizer of $f(\mathbf{x})$ is written as

$$g_1(\mathbf{x};\mathbf{x}_{(t)})=\mathbf{x}^H\mathbf{A}_2\mathbf{x}+2\,\mathrm{Re}\left(\mathbf{x}_{(t)}^H\mathbf{A}_1\mathbf{x}\right)-\mathbf{x}_{(t)}^H\mathbf{A}_1\mathbf{x}_{(t)} \tag{16}$$

The remaining step is how to reform the first non-convex constraint of $\mathcal{P}_2$. Firstly, it can be expressed as the pair of inequality constraints

$$\mathbf{x}^H\mathbf{x}-E\le0,\ E-\mathbf{x}^H\mathbf{x}\le0 \tag{17}$$

The second part is a concave constraint, which is approximated by the first order Taylor expansion at $\mathbf{x}_{(t)}$ [50], i.e.

$$E-\mathbf{x}^H\mathbf{x}\le E-2\,\mathrm{Re}\left(\mathbf{x}_{(t)}^H\mathbf{x}\right)+\mathbf{x}_{(t)}^H\mathbf{x}_{(t)}=2E-2\,\mathrm{Re}\left(\mathbf{x}_{(t)}^H\mathbf{x}\right)$$

Then, (17) is changed to

$$\mathbf{x}^H\mathbf{x}-E\le0,\ E-\mathrm{Re}\left(\mathbf{x}_{(t)}^H\mathbf{x}\right)\le0 \tag{18}$$

Even though all constraints are convex, the value of $\mathbf{x}$ cannot be updated, because $\mathbf{x}$ has only one feasible point $\mathbf{x}_{(t)}$ due to (18). This issue can be handled by an increasing penalty on constraint violation [51], so that $\mathcal{P}_2$ is reformulated as

$$\mathcal{P}_3\begin{cases}\max\limits_{\mathbf{x},s} & g_1(\mathbf{x};\mathbf{x}_{(t)})-\bar{\tau}_{(t)}s\\ \text{s.t.} & \begin{cases}\mathbf{x}^H\mathbf{x}\le E\\ E-\mathrm{Re}\left(\mathbf{x}_{(t)}^H\mathbf{x}\right)\le s\ ,\\ \|\mathbf{x}-\mathbf{F}\mathbf{c}\|^2\le\varepsilon_1\\ s\ge0\end{cases}\end{cases} \tag{19}$$

where s is a slack variable, and $\bar{\tau}_{(t)}$ is the penalty parameter. The above problem is convex, and can be solved by general solvers with lower worst-case complexity $\mathcal{O}(N^{3.5})$ than SDR [45]. The non-convex problem $\mathcal{P}_2$ is tackled by iterating $\mathcal{P}_3$, and $\mathbf{x}^{(1)}$ is obtained. The optimization procedure is given as Algorithm 1 below.

Algorithm 1: Optimization algorithm for subproblem $\mathcal{P}_3$

Input decision variable $\boldsymbol{\theta}$ and slack variable s
given initial point $\boldsymbol{\theta}_{(0)}=\boldsymbol{\theta}^{(0)}$, $\bar{\tau}_{(0)}>0$, $\bar{\tau}_{\max}>0$, $\mu>1$.
If the slack variable is not at the input, set $\bar{\tau}_{(0)}=0$.
$t:=0$
repeat
1. Form the surrogate function $g(\boldsymbol{\theta};\boldsymbol{\theta}_{(t)})$.
2. Solve the resultant problem by the CVX tool [52].
3. Update $\bar{\tau}$. $\bar{\tau}_{(t+1)}:=\min(\mu\bar{\tau}_{(t)},\bar{\tau}_{\max})$.
4. Update iteration. $t:=t+1$.

until stopping criterion is satisfied.
output $\boldsymbol{\theta}^{(1)}=\boldsymbol{\theta}_{(t^*)}$ with maximum iteration step $t=t^*$.

One reasonable stopping condition is that the variation of objection function is less than a minimum value e, i.e.

$$g_1\left(\mathbf{x}_{(t+1)};\mathbf{x}_{(t+1)}\right)-\bar{\tau}_{(t+1)}s_{(t+1)}-g_1\left(\mathbf{x}_{(t)};\mathbf{x}_{(t)}\right)+\bar{\tau}_{(t)}s_{(t)}\leq e$$

With $\mathbf{x}^{(1)}$ substituted in (13), it can be confirmed that

$$\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\geq\rho\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right) \tag{20}$$

The proof is presented in Section C of Appendix.

2) Range Filter Optimization : Given $\mathbf{x}=\tilde{\mathbf{x}}=\mathbf{x}^{(1)}$, $\mathbf{w}=\tilde{\mathbf{w}}=\mathbf{w}^{(0)}$ and $\tilde{\mathbf{m}}=\mathbf{m}^{(0)}$, the auxiliary function for $\mathbf{m}$ is written as

$$\Psi\left(\mathbf{x}^{(1)},\mathbf{m},\mathbf{w}^{(0)};\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)=\mathbf{m}^H\mathbf{B}_1\mathbf{m}+\mathbf{m}^H\mathbf{B}_2\mathbf{m}$$

where

$$\mathbf{B}_1=\mathbf{w}^{(0)H}\mathbf{p}_{\sigma\eta}\mathbf{p}_{\sigma\eta}^T\mathbf{w}^{(0)}\cdot\mathbf{p}_{\sigma\tau}\mathbf{p}_{\sigma\tau}^T\odot\left(\mathbf{x}^{(1)}\mathbf{x}^{(1)H}\right)^*,$$

$$\mathbf{B}_2=-\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\Big[\mathbf{w}^{(0)H}\mathrm{Diag}(\mathbf{p}_{\sigma\eta})\mathbf{w}^{(0)}\cdot\mathrm{Diag}(\mathbf{p}_{\sigma\tau})$$
$$\odot\left(\mathbf{x}^{(1)}\mathbf{x}^{(1)H}\right)^*+\mathbf{w}^{(0)H}\mathrm{Diag}(\mathbf{p}_{b\eta})\mathbf{w}^{(0)}\cdot\mathrm{Diag}(\mathbf{p}_{b\tau})$$
$$\odot\left(\mathbf{x}^{(1)}\mathbf{x}^{(1)H}\right)^*+\mathbf{w}^{(0)H}\mathrm{Diag}(\mathbf{p}_{n\eta})\mathbf{w}^{(0)}\cdot\mathrm{Diag}(\mathbf{p}_{n\tau})\Big]$$

Similarly, based on the MM algorithm, a minorizer of the objective function is given as

$$g_2(\mathbf{m};\mathbf{m}_{(t)})=\mathbf{m}^H\mathbf{B}_2\mathbf{m}+2\,\mathrm{Re}\left(\mathbf{m}_{(t)}^H\mathbf{B}_1\mathbf{m}\right)-\mathbf{m}_{(t)}^H\mathbf{B}_1\mathbf{m}_{(t)}$$

Combined with all constraints about $\mathbf{m}$, the associated problem with respect to $\mathbf{m}$, is cast as

$$\mathcal{P}_4\begin{cases}\max\limits_{\mathbf{m}}\, g_2(\mathbf{m};\mathbf{m}_{(t)})\\ \text{s.t. }\|\mathbf{m}-\mathbf{F}^*\mathbf{c}^*\|^2\leq\varepsilon_2\end{cases} \tag{21}$$

$\mathcal{P}_4$ is convex, and it is iterated through Algorithm 1, without penalty parameter. After the stopping criterion is satisfied , we have the solution $\mathbf{m}^{(1)}=\mathbf{m}_{(t^*)}$, and it can also be proved that

$$\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(1)},\mathbf{w}^{(0)}\right)\geq\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right). \tag{22}$$

The proof is similar with Section C of Appendix.

3) Azimuth Filter Optimization : Given $\mathbf{x}=\tilde{\mathbf{x}}=\mathbf{x}^{(1)}$, $\mathbf{m}=\tilde{\mathbf{m}}=\mathbf{m}^{(1)}$ and $\tilde{\mathbf{w}}=\mathbf{w}^{(0)}$, the auxiliary function for $\mathbf{w}$ is written as

$$\Psi\left(\mathbf{x}^{(1)},\mathbf{m}^{(1)},\mathbf{w};\mathbf{x}^{(1)},\mathbf{m}^{(1)},\mathbf{w}^{(0)}\right)=\mathbf{w}^H\mathbf{C}_1\mathbf{w}+\mathbf{w}^H\mathbf{C}_2\mathbf{w}$$

where

$$\mathbf{C}_1=\left(\mathbf{x}^{(1)}\odot\mathbf{m}^{(1)}\right)^H\mathbf{p}_{\sigma\tau}\mathbf{p}_{\sigma\tau}^T\left(\mathbf{x}^{(1)}\odot\mathbf{m}^{(1)}\right)\cdot\mathbf{p}_{\sigma\eta}\mathbf{p}_{\sigma\eta}^T,$$

$$\mathbf{C}_2=-\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(1)},\mathbf{w}^{(0)}\right)$$
$$\Big[\left(\mathbf{x}^{(1)}\odot\mathbf{m}^{(1)}\right)^H\mathrm{Diag}(\mathbf{p}_{\sigma\tau})\left(\mathbf{x}^{(1)}\odot\mathbf{m}^{(1)}\right)\cdot\mathrm{Diag}(\mathbf{p}_{\sigma\eta})$$
$$+\left(\mathbf{x}^{(1)}\odot\mathbf{m}^{(1)}\right)^H\mathrm{Diag}(\mathbf{p}_{b\tau})\left(\mathbf{x}^{(1)}\odot\mathbf{m}^{(1)}\right)$$
$$\cdot\mathrm{Diag}(\mathbf{p}_{b\eta})+\mathbf{m}^{(1)H}\mathrm{Diag}(\mathbf{p}_{n\tau})\mathbf{m}^{(1)}\cdot\mathrm{Diag}(\mathbf{p}_{n\eta})\Big]$$

A minorizer of the objective function corresponding to $\mathbf{w}$ is derived as

$$g_3(\mathbf{w};\mathbf{w}_{(t)})=\mathbf{w}^H\mathbf{C}_2\mathbf{w}+2\,\mathrm{Re}\left(\mathbf{w}_{(t)}^H\mathbf{C}_1\mathbf{w}\right)-\mathbf{w}_{(t)}^H\mathbf{C}_1\mathbf{w}_{(t)}$$

Combined with all constraints about $\mathbf{w}$, the associated problem with respect to $\mathbf{w}$, is cast as

$$\mathcal{P}_5\begin{cases}\max\limits_{\mathbf{w}}\, g_3(\mathbf{w};\mathbf{w}_{(t)})\\ \text{s.t. }\|\mathbf{w}-\mathbf{r}\|^2\leq\varepsilon_3\end{cases} \tag{23}$$

Similar with $\mathcal{P}_4$, $\mathcal{P}_5$ is convex. $\mathbf{w}^{(1)}$ is found when the iteration procedure of $\mathcal{P}_5$ converges. It can be confirmed that

$$\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(1)},\mathbf{w}^{(1)}\right)\geq\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(1)},\mathbf{w}^{(0)}\right). \tag{24}$$

Finally, a non-decreasing objective function sequence is generated through alternately iterating $\mathcal{P}_3$, $\mathcal{P}_4$ and $\mathcal{P}_5$, namely

$$\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(1)},\mathbf{w}^{(1)}\right)\geq\rho\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right). \tag{25}$$

Consequently, the algorithm derived above is effective in tackling the joint optimization problem $\mathcal{P}_1$, as it leads to a monotonically increasing sequence of the objective function during the iterative process. Its implementation is summarized in Algorithm 2.

Algorithm 2: Optimization algorithm for $\mathcal{P}_1$

initialize $\mathbf{x}^{(0)}$, $\mathbf{m}^{(0)}$, $\mathbf{w}^{(0)}$
$k:=1$
repeat
1. Form $g_1(\mathbf{x};\mathbf{x}_{(t)})$ with $\mathbf{m}^{(k-1)}$ and $\mathbf{w}^{(k-1)}$.
2. Solve $\mathcal{P}_3$ by Algorithm 1 to obtain $\mathbf{x}^{(k)}$.
3. Form $g_2(\mathbf{m};\mathbf{m}_{(t)})$ with $\mathbf{x}^{(k)}$ and $\mathbf{w}^{(k-1)}$.
4. Solve $\mathcal{P}_4$ by Algorithm 1 to obtain $\mathbf{m}^{(k)}$.
5. Form $g_3(\mathbf{m};\mathbf{m}_{(t)})$ with $\mathbf{x}^{(k)}$ and $\mathbf{m}^{(k)}$.
6. Solve $\mathcal{P}_5$ by Algorithm 1 to obtain $\mathbf{w}^{(k)}$.
7. $k:=k+1$.

until stopping criterion is satisfied.
output $\mathbf{x}^{(k)}$, $\mathbf{m}^{(k)}$ and $\mathbf{w}^{(k)}$.

By setting up a proper stopping condition

$$\rho\left(\mathbf{x}^{(k)},\mathbf{m}^{(k)},\mathbf{w}^{(k)}\right)-\rho\left(\mathbf{x}^{(k-1)},\mathbf{m}^{(k-1)},\mathbf{w}^{(k-1)}\right)\leq e,$$

the iterative process can be terminated to output final solutions for the waveform, range filter and azimuth filter.

B. Convergence Verification

Although the stopping condition is set, it still needs to be verified whether the iterative process can be terminated, especially when the minimum variation e is very small. To this end, the convergence of iterative process is analyzed in the following by showing that the objective function is monotonically increasing and converges to a finite value. Firstly, the convergence of the MM-based algorithm for suboptimization problem $\mathcal{P}_2$ is analyzed. For sufficiently large $\bar{\tau}_{\max}$, the slack variable is forced toward zero. Observe that once $\bar{\tau}=\bar{\tau}_{\max}$ in Algorithm 1, we have

$$f(\mathbf{x}_{(t+1)})\geq g(\mathbf{x}_{(t+1)};\mathbf{x}_{(t)})\geq g(\mathbf{x}_{(t)};\mathbf{x}_{(t)})=f(\mathbf{x}_{(t)}) \tag{26}$$

which indicates that $f(\mathbf{x}_{(t)})$ is not descending during iterations of $\mathcal{P}_3$. It is easy to prove that $f(\mathbf{x})$ has a upper bound:

$$f(\mathbf{x}) \leq \lambda_{\max}[\mathbf{A}_1 + \mathbf{A}_2]E \tag{27}$$

Thus, the solving method of suboptimization problem converges.

The convergence of alternate iteration procedure over three suboptimization problems should also be verified. As the monotonic increasing property has been verified in (25), further proof is needed to show that $\rho(\mathbf{x}, \mathbf{m}, \mathbf{w})$ is upper bounded.

Since the energy of target scattering characteristic function is limited, there must exist two positive numbers E_τ and E_η with $\|\mathbf{P}_{\sigma\tau}\|^2 \leq E_\tau$ and $\|\mathbf{P}_{\sigma\eta}\|^2 \leq E_\eta$. Then, the numerator of $\rho(\mathbf{x}, \mathbf{m}, \mathbf{w})$ has an upper bound, i.e.

$$\begin{aligned}\alpha(\mathbf{x}, \mathbf{m}, \mathbf{w}) &= (\mathbf{x} \odot \mathbf{m})^H \mathbf{p}_{\sigma\tau}\mathbf{p}_{\sigma\tau}^T(\mathbf{x} \odot \mathbf{m})\mathbf{w}^H \mathbf{p}_{\sigma\eta}\mathbf{p}_{\sigma\eta}^T\mathbf{w} \\ &\leq \|\mathbf{x} \odot \mathbf{m}\|^2 \|\mathbf{p}_{\sigma\tau}\|^2 \|\mathbf{w}\|^2 \|\mathbf{p}_{\sigma\eta}\|^2 \\ &= E_{xm}E_\tau E_w E_\eta\end{aligned}$$

where $\|\mathbf{x} \odot \mathbf{m}\|^2 = E_{xm}$. Meanwhile, the denominator has a lower bound, which is

$$\begin{aligned}&\beta(\mathbf{x}, \mathbf{m}, \mathbf{w}) \\ &= \text{tr}\left[\text{Diag}(\mathbf{p}_{\sigma\tau})(\mathbf{x} \odot \mathbf{m})(\mathbf{x} \odot \mathbf{m})^H\right] \text{tr}\left[\text{Diag}(\mathbf{p}_{\sigma\eta})\mathbf{w}\mathbf{w}^H\right] \\ &+ \text{tr}\left[\text{Diag}(\mathbf{p}_{b\tau})(\mathbf{x} \odot \mathbf{m})(\mathbf{x} \odot \mathbf{m})^H\right] \text{tr}\left[\text{Diag}(\mathbf{p}_{b\eta})\mathbf{w}\mathbf{w}^H\right] \\ &+ \text{tr}\left[\text{Diag}(\mathbf{p}_{n\tau})\mathbf{m}\mathbf{m}^H\right] \text{tr}\left[\text{Diag}(\mathbf{p}_{n\eta})\mathbf{w}\mathbf{w}^H\right] \\ &\geq \lambda_{\min}[\text{Diag}(\mathbf{p}_{\sigma\tau})]\text{tr}\left[(\mathbf{x} \odot \mathbf{m})(\mathbf{x} \odot \mathbf{m})^H\right] \\ &\cdot \lambda_{\min}[\text{Diag}(\mathbf{p}_{\sigma\eta})]\text{tr}\left[\mathbf{w}\mathbf{w}^H\right] \\ &+ \lambda_{\min}[\text{Diag}(\mathbf{p}_{b\tau})]\text{tr}\left[(\mathbf{x} \odot \mathbf{m})(\mathbf{x} \odot \mathbf{m})^H\right] \\ &\cdot \lambda_{\min}[\text{Diag}(\mathbf{p}_{b\eta})]\text{tr}\left[\mathbf{w}\mathbf{w}^H\right] \\ &+ \lambda_{\min}[\text{Diag}(\mathbf{p}_{n\tau})]\text{tr}\left[\mathbf{m}\mathbf{m}^H\right] \lambda_{\min}[\text{Diag}(\mathbf{p}_{n\eta})]\text{tr}\left[\mathbf{w}\mathbf{w}^H\right] \\ &= p_{\sigma 0}E_{xm}p_{\sigma 1}E_w + p_{b0}E_{xm}p_{b1}E_w + p_{n0}E_m p_{n1}E_w\end{aligned}$$

where $p_{\sigma 0}$, $p_{\sigma 1}$, p_{b0}, p_{b1}, p_{n0} and p_{n1} are minimum values of $\mathbf{p}_{\sigma\tau}$, $\mathbf{p}_{\sigma\eta}$, $\mathbf{p}_{b\tau}$, $\mathbf{p}_{b\eta}$, $\mathbf{p}_{n\tau}$ and $\mathbf{p}_{n\eta}$, respectively. For the first inequality, we have used the fact that, given $\mathbf{A}$, $\mathbf{B} \succcurlyeq \mathbf{0}$, $\text{tr}[\mathbf{AB}] \geq \lambda_{\min}[\mathbf{A}]\text{tr}[\mathbf{B}]$ [43]. Therefore,

$$\begin{aligned}\rho(\mathbf{x}, \mathbf{m}, \mathbf{w}) &= \frac{\alpha(\mathbf{x}, \mathbf{m}, \mathbf{w})}{\beta(\mathbf{x}, \mathbf{m}, \mathbf{w})} \\ &\leq \frac{E_{xm}E_\tau E_\eta}{E_{xm}p_{\sigma 0}p_{\sigma 1} + E_{xm}p_{b0}p_{b1} + E_m p_{n0}p_{n1}}\end{aligned} \tag{28}$$

Finally, the above bounded objective function in combination with the monotonically increasing property verifies the convergence.

IV. Numerical Examples

In this section, numerical simulations are performed to show the effectiveness of the problem formation and its solving procedure, where five SARs are compared in terms of information acquisition, SAR image quality, target detection and 2-D resolution performance, including the proposed scheme, the weighted SAR with weighted LFM signal and newly designed receive filter [25], the suboptimal SAR with suboptimal waveform [24], the classic SAR with the LFM signal and NLFM SAR with the NLFM signal generated by the POSP approach [10, 11]. The main parameters are listed in Table I.

Table I
List of Simulation Parameters

Parameter	Symbol	Value
Platform height	H_0	6km
Antenna length	D	2m
Effective radar velocity	v	150m/s
Look angle	θ	30°
Beam squint angle	ϕ	0°
Center frequency	f_0	5.3GHz
Pulse duration	T	1μs
Range Bandwidth	B_r	100MHz
Range sampling rate	F_r	120MHz
Azimuth sampling rate	F_a	180Hz
Number of range lines	N_a	256
Samples per range line	N_r	256

Details of the simulation process are given as follows:

Step 1 Joint design of waveform and imaging filters: $\mathbf{p}_{\sigma\tau}$ and $\mathbf{p}_{\sigma\eta}$ are set according to the target scattering characteristics PSD. Without loss of generality, suppose that $n(\tau, \eta)$ and $b(\tau, \eta)$ have a constant PSD. SNR and SCR are defined as $\text{SNR} = \text{E}\left[|\sigma(\tau, \eta)|^2\right]/\text{E}\left[|n(\tau, \eta)|^2\right]$ and $\text{SCR} = \text{E}\left[|\sigma(\tau, \eta)|^2\right]/\text{E}\left[|b(\tau, \eta)|^2\right]$, respectively. Initialize $\mathbf{x}^{(0)}$, $\mathbf{m}^{(0)}$ and $\mathbf{w}^{(0)}$ with $\mathbf{Fc}$, $\mathbf{F}^*\mathbf{c}^*$ and $\mathbf{r}$, and then Algorithm 2 is applied step by step until the stop condition is satisfied, where $e = 10^{-3}$, $\tau_0 = 1.1$, $\mu = 1.5$ and $\bar{\tau}_{\max} = 10^4$. Finally, the time-domain waveform $\mathbf{x}^\dagger$, range filter $\mathbf{m}^\dagger$ and azimuth filter $\mathbf{w}^\dagger$ based on MTIC are obtained.

Step 2 SAR echo generation: The observed scene is formed by an extended target embedded in background clutters, based on the system model in Fig. 1. $\mathbf{x}^\dagger$ is used as the SAR transmit waveform to illuminate the scene based on the parameters in Table I. As the platform advances, the resultant echoes are generated. For performance comparison, waveforms of the weighted SAR, suboptimal SAR, classic SAR, and NLFM SAR are generated, and their echos are also collected. The five waveforms have the same energy, pulse duration and bandwidth.

Step 3 SAR imaging: With the five groups of echo signals from Step 2, the SAR images are obtained, through range filtering, RCMC, phase compensation, and azimuth filtering. The weighted SAR uses newly designed receive filter to replace the matched filter in the range direction, and the classic SAR and the NLFM SAR apply the matched filter to achieve pulse compression. The resultant SAR images are used for performance analysis and comparison.

A. Theoretical Analysis Verification

1) Convergence of Algorithm 1: To show the convergence process of Algorithm 1, the value of $\Psi(\mathbf{x}, \mathbf{m}^{(0)}, \mathbf{w}^{(0)}; \mathbf{x}^{(0)}, \mathbf{m}^{(0)}, \mathbf{w}^{(0)}) = f_1(\mathbf{x})$ and $g_1(\mathbf{x}; \mathbf{x}_{(t)})$ versus iteration steps is given in Fig. 2, which also facilitates

the verification o f e ffectiveness f or t he s urrogate function. Similarly, the objective function and surrogate function corresponding to two other subproblems of $\mathbf{m}$ and $\mathbf{w}$ are also shown below.

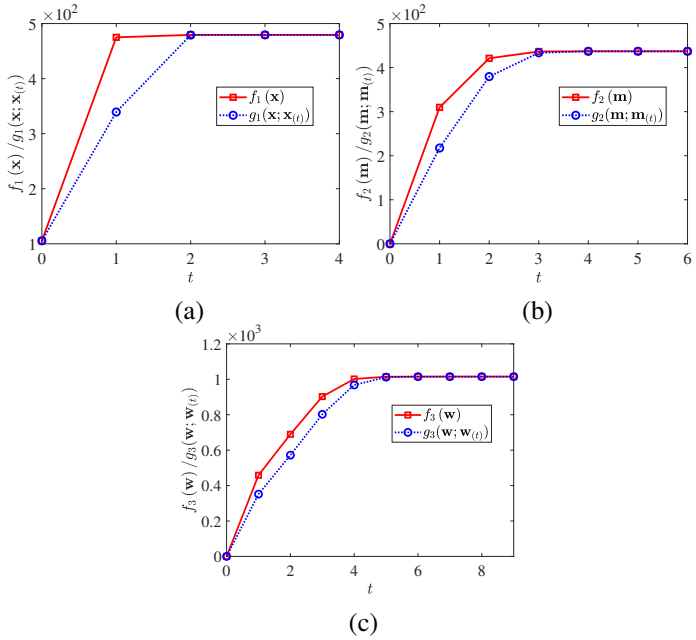

Figure 2. Objective function and surrogate function of subproblem versus iteration step: (a) $f_1(\mathbf{x})$ and $g_1(\mathbf{x};\mathbf{x}_t)$; (b) $f_2(\mathbf{m})$ and $g_2(\mathbf{m};\mathbf{m}_{(t)})$; (c) $f_3(\mathbf{w})$ and $g_3(\mathbf{w};\mathbf{w}_{(t)})$, where $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 6$ and SNR=SCR=0dB.

As can be seen, Algorithm 1 is effective in tackling the three subproblems, and it converges rapidly. Taking the first plot as an example, the objective function $f_1(\mathbf{x})$ is monotonically increasing with the iterative process proceeding, enforced by the improving minorizer or surrogate function $g_1(\mathbf{x};\mathbf{x}_{(t)})$. It is confirmed that $g_1(\mathbf{x};\mathbf{x}_{(t)})$ is effective to locally approximate the objective function.

2) Convergence of Algorithm 2: Examples of the convergence process for the optimization problem with different feasible regions are now provided. Algorithm 2 is performed, and the objective function $f(\mathbf{x},\mathbf{m},\mathbf{w})$ versus the iteration number is shown in Fig. 3, with respect to different similarity constraint parameters ε_1, ε_2 and ε_3.

It can be seen that the objective function sequence converges rapidly for all cases. After three iterations, the value of $\rho(\mathbf{x},\mathbf{m},\mathbf{w})$ tends to stay stable, and the stop condition is satisfied. A large value for the objective function is achieved even after one iteration. Moreover, as demonstrated in (25), the proposed algorithm to tackle $\mathcal{P}_1$ leads to a monotonically increasing objective function, with respect to the iteration number.

Compared to the bottom line associated with $\varepsilon_1 = 0.1$ in Fig. 3(a), increasing ε_x from a small value to ∞ results in a substantial growth of $\rho(\mathbf{x},\mathbf{m},\mathbf{w})$, as the feasible region of $\mathcal{P}_1$ becomes larger. Fig. 3(b) and (c) also show a similar pattern which means that a continuous increase can be seen in $\rho(\mathbf{x},\mathbf{m},\mathbf{w})$, resulting from a set of increasing ε_2 and ε_3. It is worth noting that different similarity parameters leads to different iteration number. Because the stop condition remains

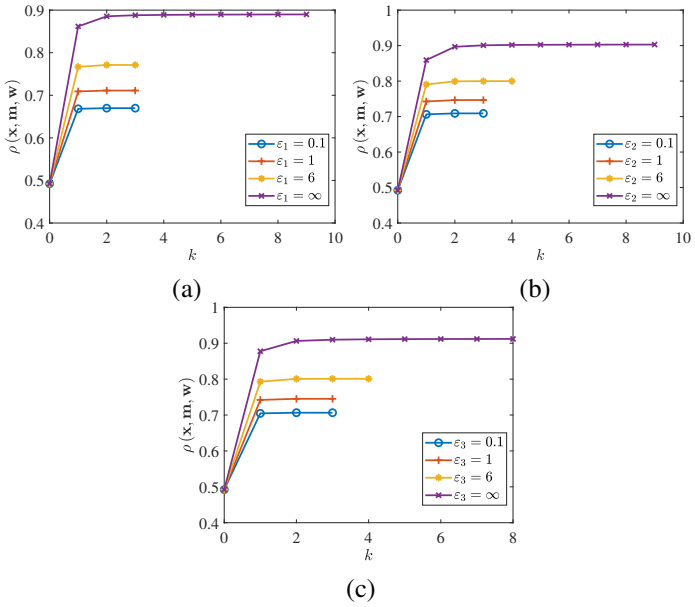

Figure 3. Objective function for different feasible regions: (a) $\rho(\mathbf{x},\mathbf{m},\mathbf{w})$ with varying ε_1; (b) $\rho(\mathbf{x},\mathbf{m},\mathbf{w})$ with varying ε_2; (c) $\rho(\mathbf{x},\mathbf{m},\mathbf{w})$ with varying ε_3, where the symbol ∞ means that the associated constraint is relaxed.

unchanged, the necessary iterations to meet the stop condition are naturally different.

3) Similarity Constrains: As mentioned before, a resultant waveform with high resolution, sidelobe level and easy generation can be achieved by the similarity constraint with an existing SAR waveform. In order to verify its effectiveness, the newly designed waveform with respect to different similarity parameters is depicted in Fig. 4.

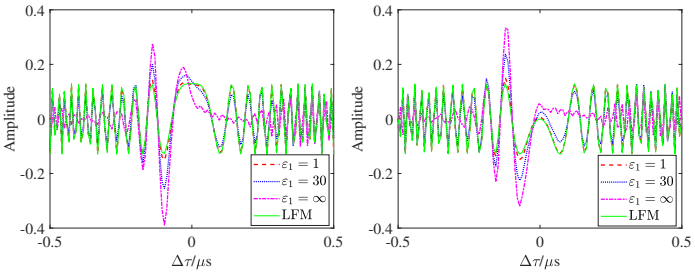

Figure 4. Resultant waveform and the LFM signal in the time domain: (a) real part; (b) imaginary part.

As expected, we can observe that a smaller ε_1 results in a waveform which is more similar to the LFM signal, which indicates that the similarity constraint works well in helping $\mathbf{x}^\dagger$ maintain some desired merits of $\mathbf{c}$. As the modulus of $\mathbf{x}^\dagger$ is close to $\mathbf{c}$ when ε_1 is relatively small, it facilitates hardware configuration in practice. Moreover, as mentioned before, a similar range-azimuth profile is desired, which is achieved by adjusting ε_1, ε_2 and ε_3 simultaneously. In order to investigate its effectiveness, the range-azimuth profile with different similarity parameters is shown in Fig. 5.

As expected, the range profile relies on the similarity parameters when the reference waveform $\mathbf{c}$ is fixed, which

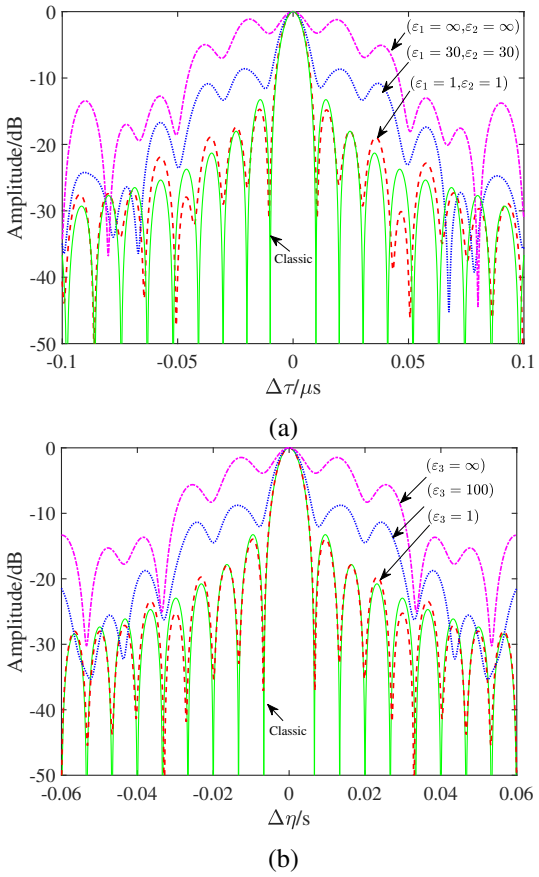

Figure 5. Range profile and azimuth profile with different similarity parameters: (a) range profile with fixed ε_3; (b) azimuth profile with fixed ε_1 and ε_2.

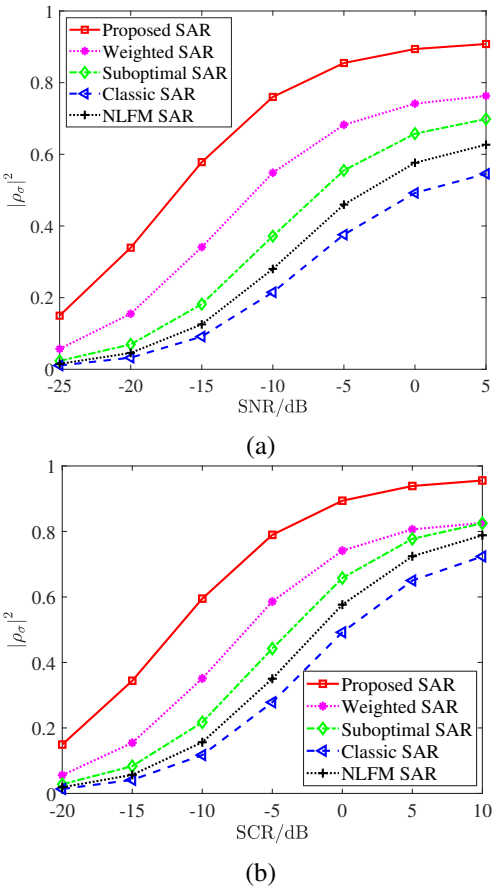

Figure 6. Information acquisition: (a) $|\rho_\sigma|^2$ with respect to different SNRs under SCR=0dB; (b) $|\rho_\sigma|^2$ with respect to different SCRs under SNR=0dB.

is demonstrated by more similar mainlobe width and sidelobe lever with $\mathbf{c}$ resulting from a smaller ε_1 and ε_2. A similar pattern is also shown in the azimuth profile. Referring to Fig. 3, with an increase of ε_1, ε_2 and ε_3, the value of the objective function also increases due to a larger feasibility region. It implies that a desired result for both the objective function and range-azimuth profile can be obtained by adjusting the similarity parameters.

B. Performance Assessment

1) Information Acquisition: In the following analysis, setting $\varepsilon_1 = 10$ and $\varepsilon_2 = \varepsilon_3 = 12$, the waveform-filter pair is generated through Algorithm 2, which is applied for performance evaluation. In order to compare the information acquisition ability, the $|\rho_\sigma|^2$ results versus different SNRs and SCRs are shown in Fig. 6.

The proposed SAR always achieves the largest $|\rho_\sigma|^2$ over all the considered SNR or SCR ranges compared with the other four. As can be seen, although the weighted SAR results in a larger amount of acquired information than the classic SAR and NLFM SAR, it is still lower than the proposed one. The reason is that SAR has 2-D range-azimuth coupling, whose information extraction capability relies on transmit-receive range pair and azimuth imaging filter. The weighted SAR without joint range-azimuth consideration cannot ensure maximal information acquisition in the 2-D SAR image. Similarly, the performance of suboptimal SAR is even worse than the weighted SAR, because the optimization in range filter is ignored. With the same SNR and SCR, the maximum gap between the proposed SAR and the classic one is about 0.54. It indicates that the proposed scheme can significantly improve the target information acquisition capability of SAR.

2) SAR Image Quality and Target Detection: With designed waveform and imaging filters, it is expected that a better visual quality of SAR images can be achieved. To illustrate this, an observed scene with five airliners is simulated, and the associated SAR images are formed. The five airliners are considered as the observed targets, and their reflectivity coefficients are different samples of one random process. Firstly, the $N_a \times N_r$ data matrices are generated according to the statistics of target scattering, background clutter, and noise. Then, delimiting the area for airplane, the associated cells of target data matrix are kept, and the observed scene is formed with the associated clutter components. Through Steps 2 and 3, four images resulting from the above-mentioned SARs, are shown in Fig. 7.

It can be found that the SAR image based on the MTIC has achieved a better image quality with more explicit airliner contour and shape than the other four SARs. Meanwhile, greater contrast between target and background cells is shown in Fig 7(a), and more detailed information of the target is provided

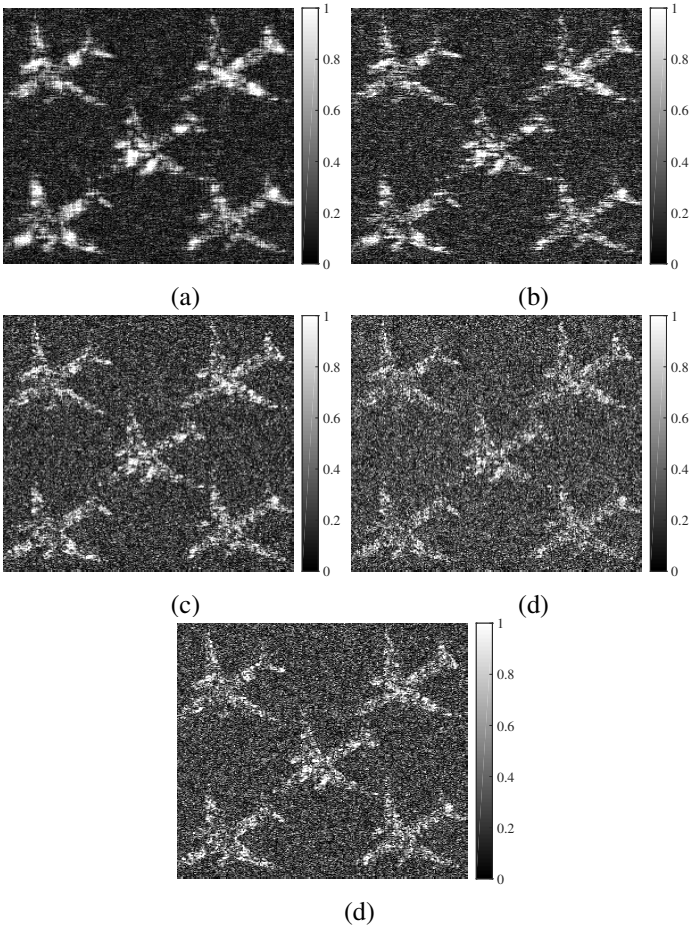

Figure 7. SAR images: (a) proposed SAR; (b) weighted SAR; (c) suboptimal SAR; (d) classic SAR; (e) NLFM SAR, where SNR=SCR=-5dB.

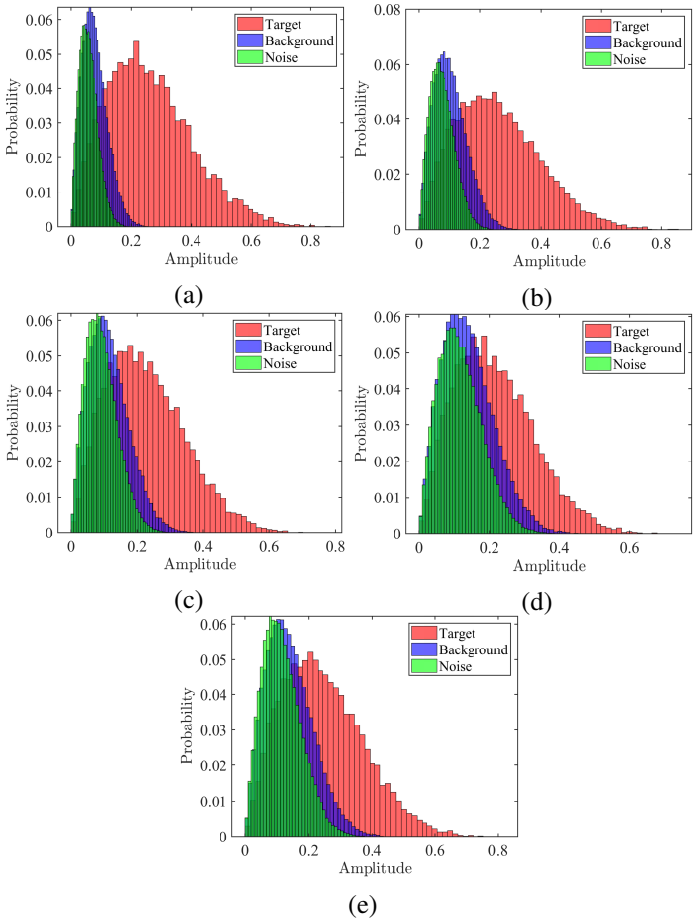

Figure 8. Probability histogram of SAR images: (a) proposed SAR; (b) weighted SAR; (c) suboptimal SAR; (d) classic SAR; (e) NLFM SAR.

by the proposed SAR. This has further demonstrated that the proposed method is effective in improving the information acquisition ability of SAR.

In order to further compare their ability in discriminating between the useful part and other signals, the probability histogram of each signal component in SAR images is plotted in Fig. 8, where it is clear that the proportion of target scattering corrupted by background clutter and noise in the proposed SAR is smaller than the other four. In Fig. 8(a), the discrimination between target and remaining part is better, which is verified by a longer distance between the peak of target and the peak of background or noise. The joint design based on the MTIC enhances target scattering while noise and clutter are suppressed, which benefits the process of target information extraction. Differing from conventional speckle noise suppression, the proposed scheme improves the discrimination by enlarging the difference between the mean values of target and other parts during the acquisition stage of SAR. Certainly, existing speckle noise suppression methods can also be applied to the SAR image generated through the proposed scheme to further improve the image quality.

The information acquisition ability can also be evaluated in terms of target detection performance [53, 54]. Based on Figs. 7 and 8, a target detection experiment is performed. Setting the threshold to traverse the entire signal value range in turn, the detection results (binary images) are obtained. Combined with the binary target image, detection probability P_D and false alarm probability P_{FA} are calculated, as shown in Fig. 9.

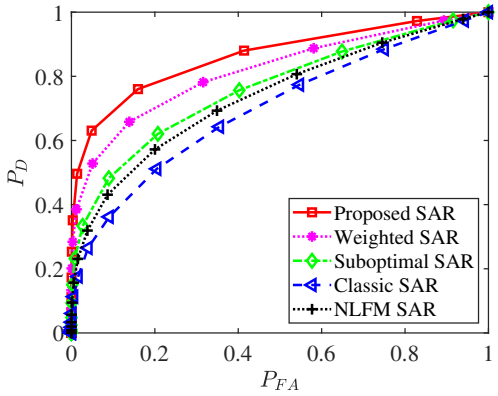

Figure 9. Detection probability and false alarm probability.

It can be seen that the proposed SAR results in better target detection result than other SARs, as a higher P_D is always held by the proposed scheme with any values of P_{FA}, and in particular, the improvement in P_D is more significant, when P_{FA} is smaller than 0.2. This further verifies that the

joint design can significantly i mprove t he S AR information acquisition capability.

3) Range-Azimuth Profile: In order to evaluate the range-azimuth profile, a corner reflector is placed in the scene. After Steps 2 and 3, the resultant SAR images and a point target 2-D profile are shown in Fig. 10.

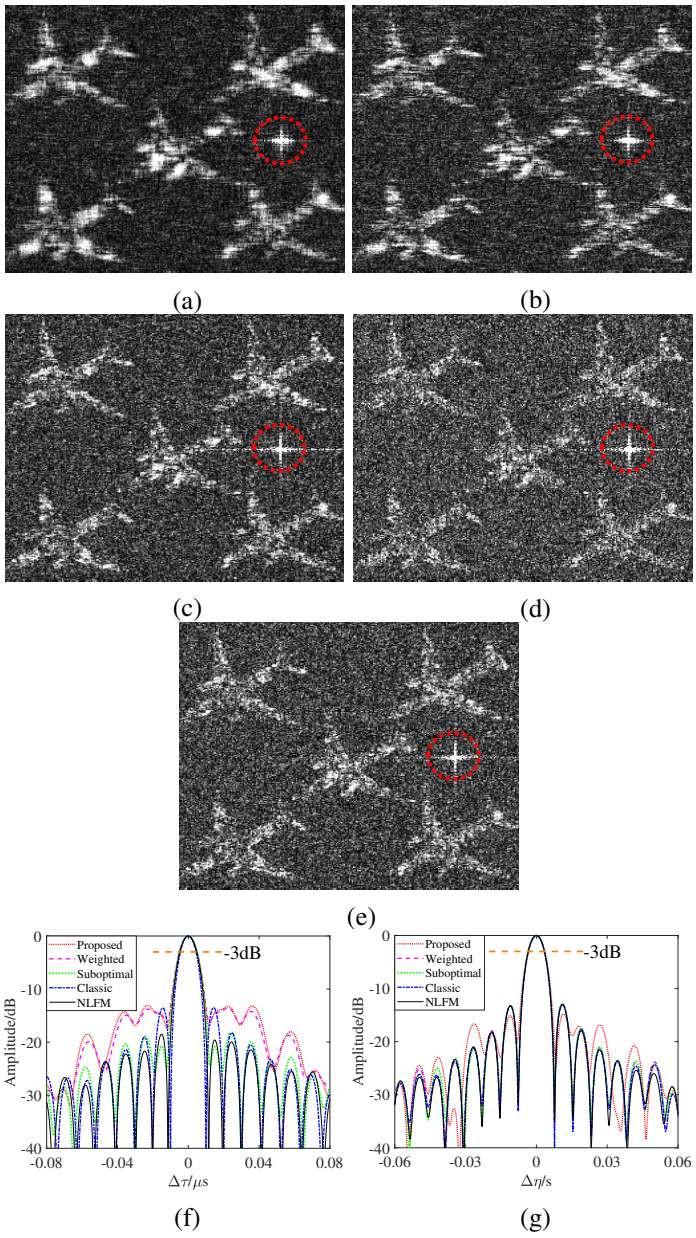

Figure 10. SAR images and a point target profile: (a) proposed SAR; (b) weighted SAR; (c) suboptimal SAR; (d) classic SAR; (e) NLFM SAR; (f) Range profile; (g) Azimuth profile.

As observed, the proposed scheme and the classic SAR have similar range-azimuth profile, which demonstrates that they can achieve similar 2-D resolution and sidelobe level. Therefore, the 2-D similarity constraint is effective, and the joint design has improved the target information acquisition capability when prior knowledge is accurate, while a similar imaging result can always be guaranteed even with inaccurate prior information.

V. Conclusion

A joint design of transmit waveform and imaging filters has been presented for more effective target information acquisition in SAR, based on the MTIC. Firstly, the SAR image extraction process with the designed waveform and imaging filters is derived. The MI between SAR image and target scattering characteristics is considered as the objective function whose analytical expression is given in the 2-D frequency domain. Combined with an energy limit and similarity constraints, the joint design is cast as a constrained optimization problem. To solve the problem, the numerator and denominator of the objective function are decoupled through an auxiliary function, and a unified MM-based optimization algorithm with an increasing penalty on constraint violation is proposed to tackle the three subproblems. They are solved in an alternate way and its convergence is proved. Finally, combining the designed transmit waveform, range filter and azimuth filter, the new SAR has achieved more effective information acquisition, better image quality and target detection performance than existing SAR designs. Moreover, one possible future research topic may be the extension of the proposed method to tackle multiple categories of targets which are sometimes embedded in observation scene at the same time.

VI. Appendix

A. Derivation of $r(\tau,\eta)$ and $h(\tau,\eta)$

The first imaging step is applying the range filter on the raw echo. After passing through the range filter $m(\tau)$, it follows that

$$\begin{aligned} s_r(\tau,\eta) &= h_{\mathrm{E}}(\tau,\eta)\otimes m(\tau) \\ &= p_r\left(\tau-2R(\eta)/c\right)\omega_a(\eta)\exp\left(-j4\pi R_0/\lambda\right)\exp\left(-j\pi K_a\eta^2\right) \end{aligned}$$

where $p_r = x(\tau)\otimes m(\tau)$ denotes the range filter output. Then, the azimuth Fourier transform is performed on each range gate to transform $s_r(\tau,\eta)$ into the range Doppler domain. Applying the POSP to the above equation [26], the relationship between azimuth frequency and time is given by $f_\eta = -K_a\eta$. Thus, the azimuth spectrum can be expressed as

$$\begin{aligned} S_r(\tau,f_\eta) &= p_r\left(\tau-2R_{rd}(f_\eta)/c\right)\mathcal{W}_a(f_\eta) \\ &\quad\cdot\exp\left(-j4\pi R_0/\lambda\right)\exp\left(j\pi f_\eta^2/K_a\right) \end{aligned} \tag{29}$$

where $\mathcal{W}_a(f_\eta)$ is the Fourier transform of antenna weighting function. Since the range equation can be approximated by

$$R(\eta)=\sqrt{R_0^2+v^2\eta^2}\approx R_0^2+\frac{v^2\eta^2}{2R_0},$$

$R_{rd}(f_\eta)$ is written as

$$R_{rd}(f_\eta)\approx R_0+\frac{\lambda^2R_0}{8v^2}f_\eta^2 = R_0+\Delta R(f_\eta).$$

The RCMC is then applied, and the azimuth filter $W(f_\eta)$ in frequency domain is introduced, so that (29) is transformed into

$$\begin{aligned} S_a(f_\tau,f_\eta) &= \mathscr{F}_\tau\left[S_r(\tau,f_\eta)\right]G_{\mathrm{RCMC}}(f_\tau,f_\eta)H_{\mathrm{P}}(f_\eta)W(f_\eta) \\ &\approx X(f_\tau)M(f_\tau)\exp\left(-j4\pi R_0f_\tau/c\right) \\ &\quad\cdot W(f_\eta)\mathcal{W}_a(f_\eta)\exp\left(-j4\pi R_0/\lambda\right). \end{aligned} \tag{30}$$

where $X(f_\tau)=\mathscr{F}_\tau[x(\tau)]$, $M(f_\tau)=\mathscr{F}_\tau[m(\tau)]$ and $\mathcal{W}_a(f_\eta)$ is assumed to be a rectangular window function with constant envelope 1 and $G_{\text{RCMC}}(f_\tau,f_\eta)$ denotes the RCMC operation:

$$G_{\text{RCMC}}(f_\tau,f_\eta)=\exp\{j4\pi f_\tau\Delta R(f_\eta)/c\} \tag{31}$$

where small range blocks are used, and the range invariant range cell migration is assumed within each block [26]. Moreover, the residual phase in the azimuth direction is removed by phase compensation given by

$$\mathscr{F}_\eta[h_{\text{P}}(\eta)]=H_{\text{P}}(f_\eta)=\exp\left(-j\pi f_\eta^2/K_a\right). \tag{32}$$

It can be seen from (30) that the RCMC can achieve decoupling between range frequency and azimuth frequency. Then, (30) is transformed into the 2-D time domain:

$$\begin{aligned}s_a(\tau,\eta)\approx&\exp\left(-j4\pi R_0/\lambda\right)x(\tau)\otimes m(\tau)\\&\otimes\delta\left(\tau-2R_0/c\right)\otimes w(\eta).\end{aligned} \tag{33}$$

Cascading all imaging steps including range filter $m(\tau)$, RCMC $g_{\text{RCMC}}(\tau,\eta)$, phase compensation $h_{\text{P}}(\eta)$ and azimuth filter $w(\eta)$, we have

$$r(\tau,\eta)=m(\tau)\otimes g_{\text{RCMC}}(\tau,\eta)\otimes h_{\text{P}}(\eta)\otimes w(\eta) \tag{34}$$

where $g_{\text{RCMC}}(\tau,\eta)$ is the inverse Fourier transform of $G_{\text{RCMC}}(f_\tau,f_\eta)$. It generates a direct filter from target scattering characteristics to SAR image, namely

$$\begin{aligned}&h(\tau,\eta)=h_{\text{E}}(\tau,\eta)\otimes r(\tau,\eta)\\&=\exp\left(-j4\pi R_0/\lambda\right)x(\tau)\otimes m(\tau)\otimes\delta\left(\tau-2R_0/c\right)\otimes w(\eta).\end{aligned}$$

B. Proof of Equation (8)

Taking ρ_1 as an example, ρ_2 and ρ_3 are derived in the same way.

$$\begin{aligned}\rho_1&=\text{E}\left[\sigma(\tau,\eta)\otimes h(\tau,\eta)\sigma^*(\tau,\eta)\right]\\&=\text{E}\left[\iint\sigma(\tau-x,\eta-y)h(x,y)dxdy\sigma^*(\tau,\eta)\right]\\&=\iint\text{E}\left[\sigma(\tau-x,\eta-y)\sigma^*(\tau,\eta)\right]h(x,y)dxdy\\&=\iint R_\sigma(-x,-y)h(x,y)dxdy\\&=R_\sigma(\Delta\tau,\Delta\eta)\otimes h(\Delta\tau,\Delta\eta)|_{\Delta\tau=0,\Delta\eta=0}\end{aligned}$$

where $R_\sigma(\Delta\tau,\Delta\eta)$ is the autocorrelation function of $\sigma(\tau,\eta)$. It is usually assumed that $\sigma(\tau,\eta)$ is a generalized random process [19, 29]. According to the Wiener-Khinchin theorem, the autocorrelation function can be defined as the inverse Fourier transform of the PSD, and then

$$\begin{aligned}&\rho_1=\\&\iint P_\sigma(f_\tau,f_\eta)h(f_\tau,f_\eta)e^{j2\pi f_\tau\Delta\tau}e^{j2\pi f_\eta\Delta\eta}df_\tau df_\eta\,|_{\Delta\tau=\Delta\eta=0}\\&=\iint P_\sigma(f_\tau,f_\eta)h(f_\tau,f_\eta)df_\tau df_\eta\end{aligned}$$

C. Proof of Equation (10)

To obtain (10), ρ_1, ρ_2 and ρ_3 need to be represented in a discrete form. In the following analysis, some terms can be ignored, because they have no impact on the optimization problem. As can be seen in (9), the first term $\exp(-j4\pi R_0/\lambda)$ is a constant, so it can be ignored in the optimization of $|\rho_\sigma|^2$. Meanwhile, the time delay in range direction does not influence the waveform and range filter design, and thus it can be considered as zero during the optimization of waveform-filter pair. The second term in (9) can also be ignored.

Then, ρ_1 is approximately transformed into a discrete form, i.e.

$$\begin{aligned}\rho_1\approx&\sum\nolimits_{k_1=1}^{N}P_{\sigma\tau}(f_{\tau k_1})X(f_{\tau k_1})M(f_{\tau k_1})\Delta f_\tau\\&\cdot\sum\nolimits_{k_2=1}^{M}P_{\sigma\eta}(f_{\eta k_2})W(f_{\eta k_2})\Delta f_\eta,\end{aligned}$$

where $N,\ M\in\mathbb{N}$, $f_{\tau k_1}=-B_\tau/2+k_1\Delta f_\tau$ and $f_{\eta k_2}=-B_\eta/2+k_2\Delta f_\eta$. $P_{\sigma\tau}(f_\tau)$ and $P_{\sigma\eta}(f_\eta)$ are PSDs in the range and azimuth directions of $\sigma(\tau,\eta)$, respectively. Because $G_{\text{RCMC}}(f_\tau,f_\eta)$ and $H_{\text{P}}(f_\eta)$ are both unimodular exponential functions in (31) and (32), respectively, we have

$$|R(f_\tau,f_\eta)|^2=|M(f_\tau)|^2|W(f_\eta)|^2. \tag{35}$$

ρ_3 is approximately written as

$$\begin{aligned}\rho_3\approx&\sum\nolimits_{k_1=1}^{N}P_{\sigma\tau}(f_{\tau k_1})\left|X(f_{\tau k_1})\right|^2|M(f_{\tau k_1})|^2\Delta f_\tau\\&\cdot\sum\nolimits_{k_2=1}^{M}P_{\sigma\eta}(f_{\tau k_2})\left|W(f_{\eta k_2})\right|^2\Delta f_\eta\\&+\sum\nolimits_{k_1=1}^{N}P_{b\tau}(f_{\tau k_1})\left|X(f_{\tau k_1})\right|^2|M(f_{\tau k_1})|^2\Delta f_\tau\\&\cdot\sum\nolimits_{k_2=1}^{M}P_{b\eta}(f_{\tau k_2})\left|W(f_{\eta k_2})\right|^2\Delta f_\eta\\&+\sum\nolimits_{k_1=1}^{N}P_{n\tau}(f_{\tau k_1})|M(f_{\tau k_1})|^2\Delta f_\tau\\&\cdot\sum\nolimits_{k_2=1}^{M}P_{n\eta}(f_{\tau k_2})\left|W(f_{\eta k_2})\right|^2\Delta f_\eta\end{aligned}$$

where $P_{b\tau}(f_\tau)$, $P_{b\eta}(f_\eta)$, $P_{n\tau}(f_\tau)$ and $P_{n\eta}(f_\eta)$ are the PSDs in the range and azimuth directions of $b(\tau,\eta)$ and $n(\tau,\eta)$, respectively. Use the notations $\mathbf{x}$, $\mathbf{m}$, $\mathbf{w}$, $\mathbf{p}_{\sigma\tau}$, $\mathbf{p}_{\sigma\eta}$, $\mathbf{p}_{b\tau}$, $\mathbf{p}_{b\eta}$, $\mathbf{p}_{n\tau}$ and $\mathbf{p}_{n\eta}$ to represent the vectors formed by $X(f_{\tau k_1})$, $M(f_{\tau k_1})$, $W(f_{\eta k_2})$, $P_{\sigma\tau}(f_{\tau k_1})$, $P_{\sigma\eta}(f_{\eta k_2})$, $P_{b\tau}(f_{\tau k_1})$, $P_{b\eta}(f_{\eta k_2})$, $P_{n\tau}(f_{\tau k_1})$, and $P_{n\eta}(f_{\eta k_2})$, respectively. As an example, for $\mathbf{x}=[X(f_{\tau 1}),X(f_{\tau 2}),\cdots,X(f_{\tau N})]^T$, we have

$$\rho_1\approx\mathbf{p}_{\sigma\tau}^T(\mathbf{x}\odot\mathbf{m})\cdot\mathbf{p}_{\sigma\eta}^T\mathbf{w}$$

Similarly, ρ_3 is rewritten as

$$\begin{aligned}\rho_3\approx&(\mathbf{x}\odot\mathbf{m})^H\text{Diag}(\mathbf{p}_{\sigma\tau})(\mathbf{x}\odot\mathbf{m})\mathbf{w}^H\text{Diag}(\mathbf{p}_{\sigma\eta})\mathbf{w}\\&+(\mathbf{x}\odot\mathbf{m})^H\text{Diag}(\mathbf{p}_{b\tau})(\mathbf{x}\odot\mathbf{m})\mathbf{w}^H\text{Diag}(\mathbf{p}_{b\eta})\mathbf{w}\\&+\mathbf{m}^H\text{Diag}(\mathbf{p}_{n\tau})\mathbf{m}\mathbf{w}^H\text{Diag}(\mathbf{p}_{n\eta})\mathbf{w}.\end{aligned}$$

Substitute ρ_1, ρ_2 and ρ_3 into (7), (10) is finally obtained.

D. Proof of Equation (20)

Suppose that $\mathbf{x}^{(1)}$ is the solution of $\mathcal{P}_2$ at the first step. We have

$$\Psi\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)};\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right) \geq \Psi\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)};\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right).$$

Based on the definition of auxiliary function, there must exist

$$\Psi\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)};\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)=\alpha\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right) - \beta\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\cdot\rho\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)=0.$$

It follows that

$$\Psi\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)};\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)=\alpha\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right) - \beta\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\cdot\rho\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\geq 0.$$

Meanwhile,

$$\Psi\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)};\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)=\alpha\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right) - \beta\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\cdot\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)=0.$$

With $\beta(\mathbf{x},\mathbf{m},\mathbf{w})\geq 0$, we have

$$\rho\left(\mathbf{x}^{(1)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right)\geq\rho\left(\mathbf{x}^{(0)},\mathbf{m}^{(0)},\mathbf{w}^{(0)}\right).$$

REFERENCES

[1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, 2013.

[2] C. Wang, J. Xu, G. Liao, X. Xu, and Y. Zhang, "A range ambiguity resolution approach for high-resolution and wide-swath SAR imaging using frequency diverse array," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 2, pp. 336–346, 2016.

[3] D. Ciuonzo, V. Carotenuto, and A. De Maio, "On multiple covariance equality testing with application to sar change detection," *IEEE Transactions on Signal Processing*, vol. 65, no. 19, pp. 5078–5091, 2017.

[4] Y. Kang, Z. Wang, J. Fu, X. Sun, and K. Fu, "SFR-Net: Scattering feature relation network for aircraft detection in complex sar images," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[5] H. Zhu, "Ship classification based on sidelobe elimination of SAR images supervised by visual model," in *Radar Conference*. IEEE, 2021, pp. 1–6.

[6] L. Zhang, X. Leng, S. Feng, X. Ma, K. Ji, G. Kuang, and L. Liu, "Domain knowledge powered two-stream deep network for few-shot SAR vehicle recognition," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[7] D. Cerutti-Maori, I. Sikaneta, and C. H. Gierull, "Optimum SAR/GMTI processing and its application to the radar satellite RADARSAT-2 for traffic monitoring," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3868–3881, 2012.

[8] J. Saeedi and K. Faez, "Synthetic aperture radar imaging using nonlinear frequency modulation signal," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 1, pp. 99–110, 2016.

[9] W. Wang, R. Wang, Y. Deng, Z. Zhang, X. Wu, and Z. Xu, "Demonstration of NLFM waveforms with experiments and Doppler shift compensation for SAR application," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1999–2003, 2016.

[10] W. Wang, R. Wang, Z. Zhang, Y. Deng, N. Li, L. Hou, and Z. Xu, "First demonstration of airborne SAR with nonlinear FM chirp waveforms," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 2, pp. 247–251, 2016.

[11] G. Jin, Y. Deng, R. Wang, W. Wang, P. Wang, Y. Long, Z. M. Zhang, and Y. Zhang, "An advanced nonlinear frequency modulation waveform for radar imaging with low sidelobe," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 6155–6168, 2019.

[12] A. Jain, P. Moulin, M. I. Miller, and K. Ramchandran, "Information-theoretic bounds on target recognition performance based on degraded image data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1153–1166, 2002.

[13] P. M. Woodward, *Probability and Information Theory With Applications to Radar.*, 2nd ed. Pergamon, 1964, pp. 62–80.

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

[15] G. Dong and G. Kuang, "Classification on the monogenic scale space: application to target recognition in SAR image," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2527–2539, 2015.

[16] G. Dong, H. Liu, and J. Chanussot, "Keypoint-based local descriptors for target recognition in SAR images: A comparative analysis," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 1, pp. 139–166, 2020.

[17] 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, 2007.

[18] Y. Yang and R. S. Blum, "Minimax robust mimo radar waveform design," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, p. 147, 2007.

[19] N. A. Goodman, P. R. Venkata, and M. A. Neifeld, "Adaptive waveform design and sequential hypothesis testing for target recognition with active sensors," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 105–113, 2007.

[20] B. Tang, J. Tang, and Y. Peng, "Mimo radar waveform design in colored noise based on information theory," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4684–4697, 2010.

[21] 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, 2011.

[22] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng,

L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE Journal of Selected Topics in Signal Processing*, 2021.

[23] J. Liang, H. C. So, C. S. Leung, J. Li, and A. Farina, "Waveform design with unit modulus and spectral shape constraints via Lagrange programming neural network," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1377–1386, 2015.

[24] H. Xu, J. Zhang, W. Liu, S. Wang, and C. Li, "High-resolution radar waveform design based on target information maximization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 5, pp. 3577–3587, 2020.

[25] J. Zhang, H. Xu, W. Liu, C. Li, and Y. Chen, "Joint design of transmit weight sequence and receive filter for improved target information acquisition in high-resolution radar," *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

[26] I. G. Cumming and F. H.-c. Wong, *Digital processing of synthetic aperture radar data: algorithms and implementation*. Artech House Norwood, MA, 2005, pp. 83–189.

[27] A. Vidal-Panteleoni and M. Ferrando, "A new spectral analysis algorithm for sar data processing of scansar data and medium resolution data without interpolation," in *IEEE International Geoscience and Remote Sensing. Symposium Proceedings*, vol. 2. IEEE, 1998, pp. 639–641.

[28] J. A. Malas, J. A. Cortese, and P. Ryan, "Uncertainty propagation and the Fano based information theoretic method," in *IEEE Radar Conference*, 2015, pp. 1638–1643.

[29] V. Frost and K. Shanmugan, "The information content of synthetic aperture radar images of terrain," *IEEE Transactions on Aerospace and Electronic Systems*, no. 5, pp. 768–774, 1983.

[30] S. Z. Alshirah, S. Gishkori, and B. Mulgrew, "Frequency-based optimal radar waveform design for classification performance maximization using multiclass fisher analysis," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[31] X. Zeng and T. Durrani, "Estimation of mutual information using copula density function," *Electronics Letters*, vol. 47, no. 8, pp. 493–494, 2011.

[32] S. M. Karbasi, A. Aubry, A. De Maio, and M. H. Bastani, "Robust transmit code and receive filter design for extended targets in clutter," *IEEE Transactions on Signal Processing*, vol. 63, no. 8, pp. 1965–1976, 2015.

[33] S. R. DeGraaf, "SAR imaging via modern 2-D spectral estimation methods," *IEEE Transactions on Image Processing*, vol. 7, no. 5, pp. 729–761, 1998.

[34] A. De Maio, S. De Nicola, Y. Huang, Z.-Q. Luo, and S. Zhang, "Design of phase codes for radar performance optimization with a similarity constraint," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 610–621, 2008.

[35] A. Aubry, A. DeMaio, A. Farina, and M. Wicks, "Knowledge-aided (potentially cognitive) transmit signal and receive filter design in signal-dependent clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 1, pp. 93–117, 2013.

[36] G. Cui, H. Li, and M. Rangaswamy, "Mimo radar waveform design with constant modulus and similarity constraints," *IEEE Transactions on signal processing*, vol. 62, no. 2, pp. 343–353, 2013.

[37] X. Yu, K. Alhujaili, G. Cui, and V. Monga, "Mimo radar waveform design in the presence of multiple targets and practical constraints," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1974–1989, 2020.

[38] I. M. Stancu-Minasian, *Fractional programming: theory, methods and applications*. Kluwer Academic Publishers, 1992, pp. 133–159.

[39] A. De Maio, Y. Huang, D. P. Palomar, S. Zhang, and A. Farina, "Fractional QCQP with applications in ML steering direction estimation for radar detection," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 172–185, 2010.

[40] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, 1967.

[41] A. Aubry, A. De Maio, and M. M. Naghsh, "Optimizing radar waveform and doppler filter bank via generalized fractional programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1387–1399, 2015.

[42] P. Stoica and Y. Selen, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 112–114, 2004.

[43] M. M. Naghsh, M. Soltanalian, P. Stoica, M. Modarres-Hashemi, A. De Maio, and A. Aubry, "A Doppler robust design of transmit sequence and receive filter in the presence of signal-dependent interference," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 772–785, 2013.

[44] B. Tang, Y. Zhang, and J. Tang, "An efficient minorization maximization approach for MIMO radar waveform optimization via relative entropy," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 400–411, 2018.

[45] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, and N. D. Sidiropoulos, "Feasible point pursuit and successive approximation of non-convex QCQPs," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 804–808, 2015.

[46] K. Huang and N. D. Sidiropoulos, "Consensus-ADMM for general quadratically constrained quadratic programming," *IEEE Transactions on Signal Processing*, vol. 64, no. 20, pp. 5297–5310, 2016.

[47] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, 2017.

[48] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "Dual-functional radar-communication waveform design: A symbol-level precoding approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1316–1331, 2021.

[49] B. Tang, J. Liu, H. Wang, and Y. Hu, "Constrained radar waveform design for range profiling," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1924–1937, 2021.

[50] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural computation*, vol. 15, no. 4, pp. 915–936, 2003.

[51] H. A. L. Thi, V. N. Huynh, and T. P. Dinh, "DC programming and DCA for general dc programs," in *Advanced Computational Methods for Knowledge Engineering*. Springer, 2014, pp. 15–35.

[52] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," 2014.

[53] 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, 2018.

[54] Y. Chen, Y. Nijsure, C. Yuen, Y. H. Chew, Z. Ding, and S. Boussakta, "Adaptive distributed MIMO radar waveform optimization based on mutual information," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 2, pp. 1374–1385, 2013.