

L-Hypersurface Based Parameters Selection in Composite Regularization Models With Application to SAR and TomoSAR Imaging

Yizhe Fan , Kun Wang , Jie Li , Guoru Zhou , Bingchen Zhang, and Yirong Wu

Abstract—Composite regularization models are widely used in sparse signal processing, making multiple regularization parameters selection a significant problem to be solved. Variety kinds of composite regularization models are used in sparse microwave imaging, including ℓ_1 and ℓ_2 penalty, nonconvex and total variation penalty, combined dictionary, etc. In this article, a new adaptive multiple regularization parameters selection method named L-hypersurface is proposed. The effectiveness of the proposed method is verified by experiments. Simulation experiments indicate that the selected optimal regularization parameters have satisfied reconstruction results, both visually and numerically. Furthermore, experiments on Gaofen-3 synthetic aperture radar satellite data are also exploited to show the performance of the proposed method.

Index Terms—Composite regularization, l-hypersurface, regularization parameter selection, sparse signal processing, synthetic aperture radar (SAR), tomographic SAR (TomoSAR).

I. INTRODUCTION

SPARSE signal processing focuses on representing the signal in a sparse way so as to make the processing faster and simpler with useful information stored in few coefficients [1]. Regularization is a basic method of sparse signal processing.

Synthetic aperture radar (SAR) is the major technology of modern microwave imaging in remote sensing and has the all-time and all-weather observing ability. The combination of sparse signal processing and SAR imaging can potentially improve the performance and reduce calculation complexity. With the development of sparse signal processing from point targets [1] to phase targets, regularization models are no longer

limited in single penalty term, but combined penalty terms, including ℓ_1 and total variation (TV) norm [2], nonconvex (NC) and TV-norm [3], NC and nonlocal TV norm [4], ℓ_1 and ℓ_2 norm [5], NC and $\ell_{2,1}$ norm [6], combined dictionaries [7], and morphology regularization [8]. Each penalty term of these composite regularization models is multiplied by a regularization parameter. The selection of regularization parameters directly controls the quality of reconstruction results.

Tomographic SAR (TomoSAR) is an essential technique for retrieving spatial information from multibaseline interferometric SAR images acquired with different view angles, which has been intensively developed in the past two decades and shows promising results [9], [10], [11], [12]. Recently, sparse signal processing has been widely used in SAR tomography reconstruction because of the advantages of superresolution and limited number of baselines. Introducing Spatial regularization, which is a composite regularization model, to TomoSAR [11] can retain the sparsity of targets and, at the same time, enhance spatial smoothness. Regularization parameters also have a great effect on the reconstruction quality of this method.

Composite regularization models have an extensive use not only in sparse signal processing, but also in variety fields using computational imaging [13], [14], [15]. In all application scenarios, regularization parameters control the effect of corresponding penalty terms and can greatly influence the performance of regularization models. Past studies have provided us a series of parameter selection methods for regularization models with single penalty term, such as the L-curve [16], [17], [18], Stein's unbiased risk estimate (SURE) [17], [18], [19], [20], generalized crossvalidation (GCV) [17], [18], [20], [21], [22], etc. However, the existing multiparameter selection methods summarized by Grasmair and Naumova [23] mainly deal with penalty terms in seminorm ($\|\mathbf{F}\mathbf{x}\|_p^p$) form, little attention had been paid to parameters selection for composite regularization models with arbitrary penalty terms in recent years. Thus, an adaptive multiparameter selection method is needed, regarding the gradually widely use of combined penalty terms. Although deep learning approaches [24], [25], [26], [27] can deal with multiple regularization parameters selection, the access to the ground truth image cannot be granted in real-life applications, such as in SAR imaging. In addition, deep learning approaches also can be very computationally demanding, but the training results sometimes only apply to limited situations, such as a fixed signal-to-noise ratio.

Manuscript received 31 May 2023; revised 22 July 2023 and 21 August 2023; accepted 2 September 2023. Date of publication 6 September 2023; date of current version 14 September 2023. This work was supported by the National Natural Science Foundation of China under Grant 61991421. (Corresponding author: Yizhe Fan.)

Yizhe Fan, Kun Wang, Jie Li, and Guoru Zhou are with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, also with the Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 101408, China (e-mail: fanyizhe18@mailsucas.ac.cn; wangkun181@mailsucas.ac.cn; lijie195@mailsucas.ac.cn; zhouguoru20@mailsucas.ac.cn).

Bingchen Zhang and Yirong Wu are with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 101408, China (e-mail: zhangbc@aircas.ac.cn; wyr@mails.ie.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3312510

The main contributions of this article are listed as follows.

- 1) We extend the L-curve method to multiparameter selection in composite regularization models and propose an L-hypersurface method, which can adaptively select multiple regularization parameters without limitation to the form of penalty terms.
- 2) We establish a corner location method for hypersurface based on inner product and orient area. Such method can locate the point of largest curvature on non-monotonous L-hypersurface.
- 3) We apply the L-hypersurface method to composite regularization model for SAR and TomoSAR imaging. The effectiveness of the proposed method is verified via simulation and real data experiments.

The rest of this article is organized as follows. Section II introduces the regularization models in SAR and TomoSAR image reconstruction. The fundamental theories of L-hypersurface method, as well as the specific algorithm are elaborated in Section III. Section IV analyzes the experimental results and evaluates the performance of the proposed method. Finally, Section V concludes this article.

II. REGULARIZATION MODEL IN SAR AND TOMOSAR IMAGE RECONSTRUCTION

The sparse microwave imaging model can be expressed as [1]

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \quad (1)$$

where $\mathbf{y} \in \mathbb{C}^{M \times 1}$ denotes the vector form of echo wave, $\mathbf{x} \in \mathbb{C}^{N \times 1}$ denotes the vector form of image, $\mathbf{A} \in \mathbb{C}^{M \times N}$ denotes the measurement matrix of the imaging system and $\mathbf{n} \in \mathbb{C}^{M \times 1}$ denotes the additive Gaussian white noise vector.

The sparse microwave imaging process is to solve a regularization problem, that is, to minimize the objective function

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \sum_{i=1}^K \lambda_i p_i(\mathbf{x}) \quad (2)$$

where $\hat{\mathbf{x}}$ denotes the imaging result, $\|\cdot\|_2$ denotes the ℓ_2 -norm, $\lambda_1, \lambda_2, \dots, \lambda_K$ denote the regularization parameters controlling penalty terms $p_1(\cdot), p_2(\cdot), \dots, p_K(\cdot)$, respectively.

A. SAR Regularization Model

Variety kinds of composite regularization models exist for SAR image reconstruction [2], [3], [4], [5], [6], [7]. Here, we introduce the regular-used NC and TV regularization model [3]

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda_1 p_{\text{NC}}(\mathbf{x}) + \lambda_2 p_{\text{TV}}(\mathbf{x}). \quad (3)$$

The NC penalty and TV penalty are defined as

$$p_{\text{NC}}(\mathbf{x}) = \text{MC}(\mathbf{x}) = \sum_{k=1}^N \begin{cases} |x_k| - \frac{|x_k|^2}{2\theta}, & |x_k| \leq \theta \\ \theta/2, & |x_k| > \theta \end{cases} \quad (4)$$

$$p_{\text{TV}}(\mathbf{x}) = \text{TV}(|\mathbf{x}|) = \sum_{i,j} \|\nabla(|\mathbf{X}|)_{i,j}\|_2 \quad (5)$$

where x_i denotes the i th element of image vector \mathbf{x} , \mathbf{X} is the 2-D complex-valued matrix of vector \mathbf{x} , the operator $|\cdot|$ represents magnitude calculation and $\nabla(|\mathbf{X}|)_{i,j}$ is the gradient vector of a pixel in the i th row and j th column, which is defined as

$$\nabla(|\mathbf{X}|)_{i,j} = \left(D_h |\mathbf{X}|_{i,j}, D_v |\mathbf{X}|_{i,j} \right) \quad (6)$$

$$D_h |\mathbf{X}|_{i,j} = |\mathbf{X}_{i+1,j}| - |\mathbf{X}_{i,j}| \quad (7)$$

$$D_v |\mathbf{X}|_{i,j} = |\mathbf{X}_{i,j+1}| - |\mathbf{X}_{i,j}| \quad (8)$$

λ_1 controls the effect of NC penalty and λ_2 controls the effect of TV penalty. In the NC-TV regularization-based sparse SAR imaging model, the NC-norm penalty, playing the role of sparsity inducing regularizer, can enhance point-based features, and the TV-norm penalty will enhance region-based features, maintaining the continuity of the backscattering coefficient of distributed targets within a certain area. Therefore, the NC-TV model can maintain the reconstruction accuracy of targets as well as protecting the reconstruction result from speckles. The distinguish functions of NC and TV penalties are instrumental in showing the tradeoff between them.

When λ_1 and λ_2 are relatively small, the function of penalty terms cannot be shown. When λ_1 is relatively large, the sparsity of image will be overly enhanced, which could lead to target loss. When λ_2 is relatively large, there will also be issues caused by oversmoothing, including the targets might be covered by noise and the texture details might loss. In general, the quality of image is controlled by λ_1 and λ_2 . Thus, it is significant to select parameters in NC-TV regularization.

B. TomoSAR Regularization Model

The spatial regularization model in TomoSAR reconstruction [11] is given as

$$\begin{aligned} \hat{\mathbf{x}} &= \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda_1 \|\mathbf{x}\|_1 + p_s(\mathbf{x}) \\ &= \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda_1 \|\mathbf{x}\|_1 \\ &\quad + \lambda_2 \|D_x |\mathbf{x}|\|_2^2 + \lambda_3 \|D_y |\mathbf{x}|\|_2^2 + \lambda_4 \|D_z |\mathbf{x}|\|_2^2 \end{aligned} \quad (9)$$

where $p_s(\mathbf{x})$ is the spatial smoothness penalty

$$p_s(\mathbf{x}) = \lambda_2 \|D_x |\mathbf{x}|\|_2^2 + \lambda_3 \|D_y |\mathbf{x}|\|_2^2 + \lambda_4 \|D_z |\mathbf{x}|\|_2^2. \quad (10)$$

$\|\cdot\|_1$ denotes the ℓ_1 norm, matrices D_x, D_y , and D_z stand for the finite difference operators in the x -, y - and z -directions. $\lambda_1, \lambda_2, \lambda_3$, and λ_4 control the weight of each term. λ_1 , associated with the sparsity constraint, has greatest influence on imaging result. When λ_1 is too large, there are some holes in the reconstruction, while a value of λ_1 that is too small leaves sidelobes and outliers. The effect of a too large value of λ_2, λ_3 , or λ_4 is the extension of structures in the corresponding direction of the spatial smoothing. When the spatial regularization is too weak, outliers located far from the actual surfaces can be observed, according to [11].

III. L-HYPERSURFACE METHOD

In this section, a multiparameter selection method named L-hypersurface is proposed. The kernel idea of L-curve method involves plotting the residual term $\|\mathbf{Ax} - \mathbf{y}\|_2^2$ and the penalty term $p_k(\mathbf{x})$, with the optimization parameter selected based on the corner of the L-shape curve. This allows for the best possible parameter selection result. We introduce the idea to multiple regularization parameters selection.

A. L-Curve Method

The L-curve is a classic method for selecting regularization parameter λ in regularization model with single penalty term, as (11) shows

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{y}\|_2^2 + \lambda p(\mathbf{x}). \quad (11)$$

The basic idea of L-curve method is to plot the residual term $\|\mathbf{Ax} - \mathbf{y}\|_2^2$ and the penalty term $p(\mathbf{x})$. The relationship curve is called an L-curve, because its shape contains a steep part and a horizontal part. The junction point of the steep and horizontal part on the L-curve is commonly named as corner, which represents the optimal parameter for the regularization model.

Compared with statistical methods, such as GCV and SURE, the L-curve is a graphical method, which can be more intuitive and efficient for multiparameter regularization problems. The L-curve method also has its ability to select parameters according to the needs of practical situation through defining corner in different ways. Regarding the low complexity and high applicability of L-curve, we decided to extend its ideas to multiparameter selection and establish an *L-hypersurface* method. Although Belge et al. [28] has the similar idea in establishing their own multiparameter selection method named by L-hypersurface, the performance of their method still depends on the form of penalty terms. In this article, we are going to establish a specific method for locating the corner on L-hypersurface regardless of the form of penalty terms.

B. Generation of L-Hypersurface

First, we generate an L-hypersurface in light of L-curve, expanding the 2-D curve to a high-dimensional hypersurface. The L-hypersurface for composite regularization model is given as

$$\mathcal{L} := \left\{ \left[\lg \|\mathbf{Ax}_\lambda - \mathbf{y}\|_2^2, \lg p_1(\hat{\mathbf{x}}_\lambda), \dots, \lg p_K(\hat{\mathbf{x}}_\lambda) \right] \mid \lambda \in \mathbb{R}^K \right\} \quad (12)$$

where $\hat{\mathbf{x}}_\lambda$ denotes the solution to (2) with $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_K)$ as regularization parameters. Each of the residual term and penalty terms associates with one dimension of the coordinate system, all of which formulate a K dimensional hypersurface on $(K + 1)$ dimensional hyperspace. The $\lg(\cdot)$ form is used to enhance the turning point of the hypersurface [16]. Each point on the hypersurface \mathcal{L} represents a solution to optimization problem (2) with a specific set of parameters. For parameters λ_0 , the

corresponding point on \mathcal{L} can be described as

$$\begin{aligned} \mathcal{L}(\lambda_0) &= [\lg(\delta_0), \lg(\delta_1), \dots, \lg(\delta_K)] \\ \delta_0 &= \|\mathbf{Ax}_{\lambda_0} - \mathbf{y}\|_2^2, \\ \delta_k &= p_k(\hat{\mathbf{x}}_{\lambda_0}), \quad (k = 1, 2, \dots, K). \end{aligned} \quad (13)$$

Next, we make the following statement: Each point on the hypersurface \mathcal{L} is a solution to the following set of equations:

$$\begin{aligned} \delta_0 &= \min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{y}\|_2^2 \text{ subject to} \\ p_k(\mathbf{x}) &\leq \delta_i \quad (i = 1, 2, \dots, K) \\ \delta_1 &= \min_{\mathbf{x}} p_1(\mathbf{x}) \text{ subject to} \\ \|\mathbf{Ax} - \mathbf{y}\|_2^2 &\leq \delta_0, p_k(\mathbf{x}) \leq \delta_k \quad (i = 2, 3, \dots, K) \\ \delta_2 &= \min_{\mathbf{x}} p_2(\mathbf{x}) \text{ subject to} \\ \|\mathbf{Ax} - \mathbf{y}\|_2^2 &\leq \delta_0, p_k(\mathbf{x}) \leq \delta_k \quad (i = 1, 3, \dots, K) \\ &\vdots \\ \delta_K &= \min_{\mathbf{x}} p_K(\mathbf{x}) \text{ subject to} \\ \|\mathbf{Ax} - \mathbf{y}\|_2^2 &\leq \delta_0, p_k(\mathbf{x}) \leq \delta_k \quad (i = 1, 2, \dots, K - 1). \end{aligned} \quad (14)$$

The statement can be proved by contradiction. Suppose that $\delta_0 \neq \min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{y}\|_2^2$ s.t. $p_k(\mathbf{x}) \leq \delta_k$ ($k = 1, 2, \dots, K$), which means there exists $\hat{\mathbf{x}}'$ such that $p_i(\hat{\mathbf{x}}') \leq \delta_k$ ($i = 1, 2, \dots, K$) and $\|\mathbf{Ax}' - \mathbf{y}\|_2^2 < \|\mathbf{Ax}_{\lambda_0} - \mathbf{y}\|_2^2 = \delta_0$. Then, we can get $\|\mathbf{Ax}' - \mathbf{y}\|_2^2 + \sum_{k=1}^K [\lambda_0]_k p_k(\hat{\mathbf{x}}') < \delta_0 + \sum_{k=1}^K [\lambda_0]_k \delta_k$, which contradicts to the fact that $\hat{\mathbf{x}}_{\lambda_0}$ is the optimized solution to (2). Similarly, each equation in (14) can be proved.

The statement guarantees that the L-hypersurface divides the hyperspace into two parts and any reconstruction result $\hat{\mathbf{x}}$ must correspond to a point above or on the hypersurface. (Strictly speaking, there is no above/below in high dimensional space, but it can be easily imagined and understood.) Once a set of parameters is given, the possible result can never correspond to a point beneath the L-hypersurface. Hence, the tradeoff must be made at a certain point on the hypersurface, which is regarded as the corner. Determining the location of the corner has a significant impact on the tradeoff, therefore a comprehensive discussion on this subject will follow in next section.

C. Determination of Corner

Locating the corner is the most significant issue for L-hypersurface method. Variety of methods exists on locating the corner of a single-parameter L-curve [18], [24], [29], [30], [31]. As the number of parameters increases, however, the complexity and uncertainty in locating the corner on the L-hypersurface also escalate, which leads to challenges including the definition of corner, the loss of monotonicity between residual and penalty terms, the intricate patterns of L-hypersurface. In order to dealing with these obstacles, we

establish a new corner determination method based on inner product and orient area. Detailed analyses are as follows.

1) *Definition of Corner*: There are many ideas to define the corner of a single-parameter L-curve, which are given as follows:

- 1) The point of maximum curvature.
- 2) The point closest to a reference point.
- 3) The point of tangency with a decided negative slope.

Here, we take 1) as the definition. Since there exists variety kinds of curvature for hypersurface, such as normal curvature, principal curvature, draw curvature, and Gaussian curvature, the meaning of curvature should be clarified. Thus, giving a reasonable interpretation of curvature is one of the main difficulties for corner determination. Here, we propose a definition of corner based on the inner product.

In application, an L-hypersurface can be generated by discrete sampling points. Let J_k be the number of sampling value of λ_k , then we have

$$\lambda_k \in \left\{ \lambda_k^1, \lambda_k^2, \dots, \lambda_k^{J_k} \right\}, 0 < \lambda_k^1 < \lambda_k^2 < \dots < \lambda_k^{J_k} \quad (15)$$

where λ_k^j ($j = 1, 2, \dots, J_k$) denotes the j th sampling value of parameter λ_k , arranging from smallest to largest. Through interpolating and smoothing operations, we can generate an approximate L-hypersurface of the testing regularization model from $J_1 \times J_2 \times \dots \times J_K$ sampling points.

We define the curvature of L-hypersurface \mathcal{L} at point A as the average of curvature on K propagation planes

$$\rho(A) = \frac{1}{K} \sum_{k=1}^K \rho_k(A) \quad (16)$$

where $\rho_k(A)$ denotes the curvature of propagated curve on the $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$ plane at point A . Specifically, given a certain point on the L-hypersurface

$$A = \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2}, \dots, \lambda_K^{j_K}) \quad (17)$$

the curvature ρ_k on the $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$ plane is defined with the assistance of another two points B_k and C_k

$$\begin{aligned} B_k &= \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2}, \dots, \lambda_{j_k}^{j_k}, \dots, \lambda_K^{j_K}) \\ C_k &= \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2}, \dots, \lambda_2^1, \dots, \lambda_K^{j_K}) \end{aligned} \quad (18)$$

where j'_k , which will be decided later, satisfies $1 < j_k < j'_k < J_k$. B_k and C_k correspond to the same j_1, \dots, j_K (except j_k) as A . C_k corresponds to the minimum value of parameter λ_k , which means the minimum value of $\|\mathbf{Ax} - \mathbf{y}\|_2^2$ on the subcurve is achieved at point C . Fig. 1(a) shows the projection of a subcurve, which can be described as $\mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2}, \dots, \lambda_{j_k}^{\hat{j}_k}, \dots, \lambda_K^{j_K})$ (\hat{j}_k is a variable), of L-hypersurface on plane $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$, where A , B_k , and C_k are marked on the curve. The curvature of point A on the $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$ plane is defined by inner product

$$\rho_k(A, B_k, C_k) = \pi - \arccos \frac{\langle \overrightarrow{A'B'_k}, \overrightarrow{A'C'_k} \rangle}{\left| \overrightarrow{A'B'_k} \right| \left| \overrightarrow{A'C'_k} \right|} \quad (19)$$

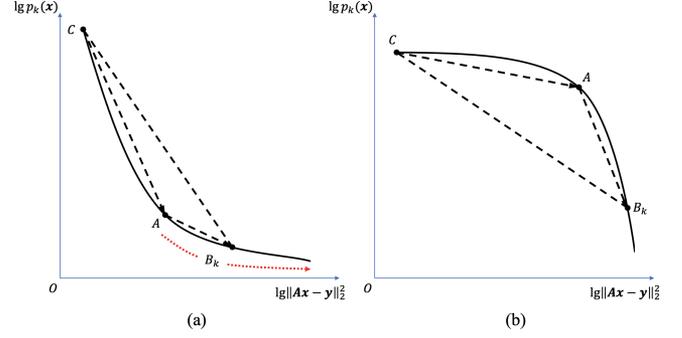


Fig. 1. Determination of corner: A subcurve of L-hypersurface projected on plane $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$. (a) Convex pattern of L-shape. (b) Concave pattern of L-shape.

where Q' denotes the projected point of Q on the $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$ plane, $\langle \cdot, \cdot \rangle$ is the inner product operator, $|\cdot|$ calculates the length of vector.

2) *Loss of Monotonicity Between Regularization Terms and Residual Term*: Back to the L-curve for single parameter, the residual term monotonically decreases as the regularization term increasing. For composite regularization problems, the expectation that regularization terms would be a decreasing function of residual term is usually failed, especially when the penalty term is not a Tikhonov form ($\|\mathbf{T}\mathbf{x}\|_2$), but, for example, the TV-norm. In addition, the solution algorithms, such as ADMM [32], ISTA [33], and FISTA [34], also might lead to nonmonotonicity of results. Most prior algorithms for locating the corner rely on monotonicity of the curve and discard those points where monotonicity is not fulfilled. For L-hypersurface, however, there might be so much discarded points that the true corner of the curve will be missed if we ignore all the nonmonotonic points. Thus, we should deal with those nonmonotonic points not by discarding them.

Here, we resolve the interference of nonmonotonicity through the selection of point B_k . More specifically, for all possible B_k s, we calculate the curvature $\rho_k(A, B_k)$ through (19) and set the maximum value as $\rho_k(A)$. Geometrically, while B_k moves from A to the end of the curve, the maximum curvature decided by $\triangle AB_kC_k$ gives the projected curvature of A on plane $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$

$$\begin{aligned} \rho_k(A) &= \max_{B_k} \rho_k(A, B_k, C_k) \\ &= \max_{(1 < j_k < j'_k < J_k)} \pi - \arccos \frac{\langle \overrightarrow{A'B'_k}, \overrightarrow{A'C'_k} \rangle}{\left| \overrightarrow{A'B'_k} \right| \left| \overrightarrow{A'C'_k} \right|} \\ &= \min_{(1 < j_k < j'_k < J_k)} \arccos \frac{\langle \overrightarrow{A'B'_k}, \overrightarrow{A'C'_k} \rangle}{\left| \overrightarrow{A'B'_k} \right| \left| \overrightarrow{A'C'_k} \right|}. \end{aligned} \quad (20)$$

Fig. 2 shows a nonmonotonic situation, the value of $p_k(\mathbf{x})$ increases as $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_k(\mathbf{x})$ increases in the circled part. For points A^1 and A^2 , we suppose B_k^1 and B_k^2 are the points corresponding to maximum curvature, respectively, so that the circled part has no effect on calculating $\rho_k(A^1)$ and $\rho_k(A^2)$. For

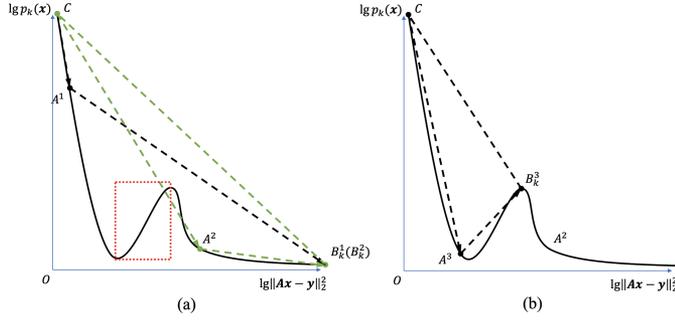


Fig. 2. Determination of corner: The nonmonotonicity of curve.

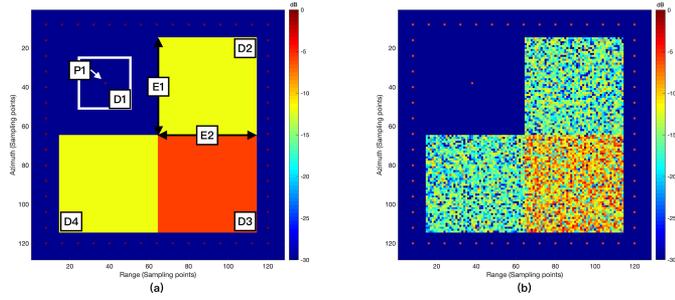


Fig. 3. SAR simulation experiment scene. (a) Real value. (b) Simulated value and selected point target (P1) with neighborhood (D1), distributed targets (D2,D3,D4), edge (E1,E2).

point A^3 , B_k^3 is the point corresponding to maximum curvature, which is on the nonmonotonically decreasing part. According to our method, the corner would be located near A^3 , since $\triangle A^3 B_k^3 C_k$ gives the maximum curvature on the projected curve. If we ignore the circled part, however, B_k^3 is ignored and the curvature of A^3 will be calculated differently, and the corner might move to A^2 . This is called the miss matching of corner and is avoided by our method.

3) *Convex and Concave Pattern of L-Shape*: No matter the curve is an L-shape [see convex shape in Fig. 1(a)] or an inverted L-shape [see concave shape in Fig. 1(b)], the formula (19) gives the same result. The shape of curve is absolutely decided by the form of regularizers and the solution algorithm. Therefore, we need to distinguish the convex or concave pattern of curves and select the maximum curvature point with L-shape curve. We introduce the concept of oriented area to separate the two patterns [30].

The oriented area of triangle $\triangle AB_k C_k$ is defined as

$$\text{area}(AB_k C_k) = \frac{1}{2} \det(\overrightarrow{C_k' A_k'}, \overrightarrow{A_k' B_k'}). \quad (21)$$

For certain triangle with vertex A_k , the sign of area identifies the shape of this part of the curve. The curve is an L-shape if $\text{area}(AB_k C_k) > 0$, and inverted L-shape, if $\text{area}(AB_k C_k) < 0$. Finally, we determine the corner of the L-hypersurface as the point with maximum curvature and corresponding triangles

Algorithm 1: L-Hypersurface method.

Input: $A, y, \{\lambda_k^1, \lambda_k^2, \dots, \lambda_k^{j_k}\} (k = 1, 2, \dots, K)$,

$\rho_{\max} = 0, \text{area}_k = 0, A_{\text{corner}} = \mathbf{0}$

1: **for** all possible λ **do**

2: $A = \mathcal{L}(\lambda)$

3: $C_k = \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2}, \dots, \lambda_2^1, \dots, \lambda_K^{j_K})$

4: **for** all possible j_1^1, \dots, j_K^1 **do**

5: $B_k = \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2}, \dots, \lambda_2^{j_k}, \dots, \lambda_K^{j_K})$

6: $\rho_k = \rho_k(A, B_k, C_k)$

7: $\text{area}_k = \text{area}(A, B_k, C_k)$

8: $\rho = \frac{1}{K} \sum_{k=1}^K \rho_k$

9: **if** $\text{area}_k > 0$ & $\rho > \rho_{\max}$ **then**

10: $\rho_{\max} = \rho, A_{\text{corner}} = A$

11: **end if**

12: **end for**

13: **end for**

Output: A_{corner}

which have positive area

$$A_{\text{corner}} := \arg \min_A \frac{1}{K} \sum_{k=1}^K \min_{B_k} \arccos \frac{\langle \overrightarrow{A' B_k'}, \overrightarrow{A' C_k'} \rangle}{\|\overrightarrow{A' B_k'}\| \|\overrightarrow{A' C_k'}\|}$$

subject to $\text{area}(AB_k C_k) > 0. \quad (22)$

The result can be reformulated as

$$A_{\text{corner}} = \arg \min_{A, j_k} \frac{1}{K} \sum_{k=1}^K \arccos \frac{\langle \overrightarrow{A' B_k'}, \overrightarrow{A' C_k'} \rangle}{\|\overrightarrow{A' B_k'}\| \|\overrightarrow{A' C_k'}\|}$$

subject to $1 < j_k < j_k' < J_k$

$\text{area}(AB_k C_k) > 0$

$(k = 1, \dots, K). \quad (23)$

D. Algorithm

Algorithm 1 gives the process of L-hypersurface method. Point $\mathcal{L}(\lambda)$ on the L-hypersurface is calculated by the regularization result \hat{x}_λ .

Algorithm 2 gives a detailed process of L-hypersurface for dual regularization parameters selection ($K = 2$). For NC-TV regularization SAR model of image reconstruction, use the modified ADMM algorithm given in [35] to calculate $\hat{x}_{\lambda_1, \lambda_2}$. For spatial regularization model of TomoSAR image reconstruction, use the algorithm given in [11] to calculate $\hat{x}_{\lambda_1, \lambda_2, \lambda_3, \lambda_4}$. To simplify the calculation, we assume $\lambda_2 = \lambda_3 = \lambda_4$, so that the model degenerates into $K = 2$ situation.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. SAR Imaging

In this section, we apply the proposed L-hypersurface method to the NC-TV regularization model (3) for SAR image reconstruction.

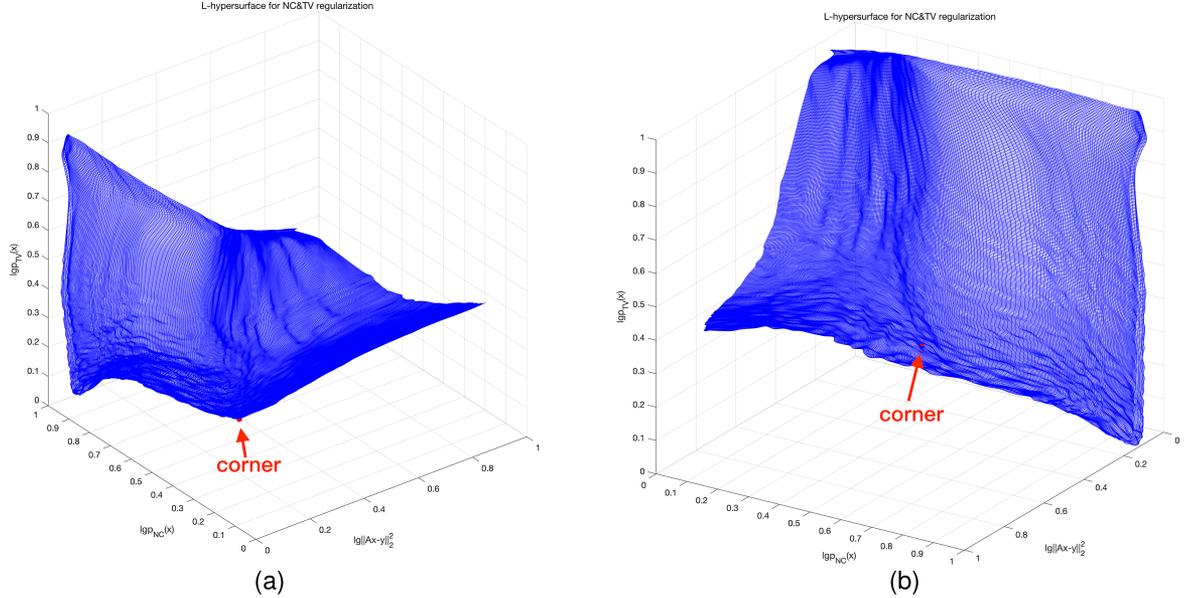


Fig. 4. L-hypersurface of simulation experiments. (a) Main view. (b) Another view.

Algorithm 2: L-Hypersurface Method for Dual Regularization Parameters Selection.

Input: \mathbf{A} , \mathbf{y} , $\{\lambda_1^1, \lambda_1^2, \dots, \lambda_1^{J_1}\}$, $\{\lambda_2^1, \lambda_2^2, \dots, \lambda_2^{J_2}\}$, $\rho_{\max} = 0$, $area_1 = area_2 = 0$, $A_{\text{corner}} = (0, 0, 0)$

- 1: **for** $j_1 = 2 : J_1 - 1$, $j_2 = 2 : J_2 - 1$ **do**
- 2: $A = \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2})$
- 3: $C_1 = \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2})$, $C_2 = \mathcal{L}(\lambda_1^{j_1}, \lambda_2^{j_2})$
- 4: **for** $j'_1 = j_1 + 1 : J_1$, $j'_2 = j_2 + 1 : J_2$ **do**
- 5: $B_1 = \mathcal{L}(\lambda_1^{j'_1}, \lambda_2^{j'_2})$, $B_2 = \mathcal{L}(\lambda_1^{j'_1}, \lambda_2^{j'_2})$
- 6: $\rho_1 = \pi - \arccos \frac{\langle \overrightarrow{A'B'_1}, \overrightarrow{A'C'_1} \rangle}{|\overrightarrow{A'B'_1}| |\overrightarrow{A'C'_1}|}$
 $\rho_2 = \pi - \arccos \frac{\langle \overrightarrow{A'B'_2}, \overrightarrow{A'C'_2} \rangle}{|\overrightarrow{A'B'_2}| |\overrightarrow{A'C'_2}|}$
- 7: $area_1 = -\det(\overrightarrow{A'C'_1}, \overrightarrow{A'B'_1})/2$
 $area_2 = -\det(\overrightarrow{A'C'_2}, \overrightarrow{A'B'_2})/2$
- 8: $\rho = (\rho_1 + \rho_2)/2$
- 9: **if** $area_1 > 0$ & $area_2 > 0$ & $\rho > \rho_{\max}$ **then**
- 10: $\rho_{\max} = \rho$, $A_{\text{corner}} = A$
- 11: **end if**
- 12: **end for**
- 13: **end for**

Output: A_{corner}

1) *Simulation Experiments:* In the simulation experiments, we generate a 128×128 pixels scene with distributed targets that occupy 50×50 pixels each and point targets in the blank, as Fig. 3(a) shows. According to [36], the complex distributed targets should be Rayleigh distributed in amplitude and uniformly distributed in phase for simulation, as Fig. 3(b) shows. A 1024×1024 Fourier matrix is chosen as the measurement

TABLE I
OBJECTIVE EVALUATION OF SIMULATION RECONSTRUCTION RESULTS

$\lambda_1 \backslash \lambda_2$		3.00	6.00	12.00
0.89	TBR of P1(dB)	26.831	20.803	1.088
	ENL of D3	3.0173	15.255	337.34
	ENL of D4	10.601	60.535	15.433
	EPI of E1	0.9408	1.2558	2.0543
	EPI of E2	0.8630	1.0393	2.1656
1.77	TBR of P1(dB)	29.773	28.933	10.660
	ENL of D3	2.9944	12.896	263.63
	ENL of D4	9.4203	24.270	8.2940
	EPI of E1	0.9693	1.2258	2.0581
	EPI of E2	0.8772	1.0873	2.0923
3.54	TBR of P1(dB)	29.781	28.961	20.205
	ENL of D3	2.9846	11.769	40.312
	ENL of D4	7.5648	8.8133	4.7898
	EPI of E1	0.9312	1.1040	1.8986
	EPI of E2	0.9021	1.1317	2.4466

matrix. A 20 dB Gaussian white noise is added to the raw echo wave, according to (1). The L-hypersurface is shown in Fig. 4 and the selected corner A_{corner} is marked on the figure. Fig. 5(a) shows the projection of L-hypersurface on $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_{\text{NC}}(\mathbf{x})$ space, which means each curve in (a) corresponds to the same λ_2 . Fig. 5(b) shows the projection of L-hypersurface on $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_{\text{TV}}(\mathbf{x})$ space, which means each curve in (b) corresponds to the same λ_1 . The projection points of the corner A_{corner} are also marked on the curves. From visual results, the corner is located at the turning points of L-shape curves on each projection space and at the point of relative large curvature on the L-hypersurface. Thus, our proposed corner determination method works. Next, the reconstruction quality of the selected corner is tested to demonstrate the effectiveness of our L-hypersurface method.

The performance of selected parameters is evaluated by visual results and objective indicators. The optimized regularization

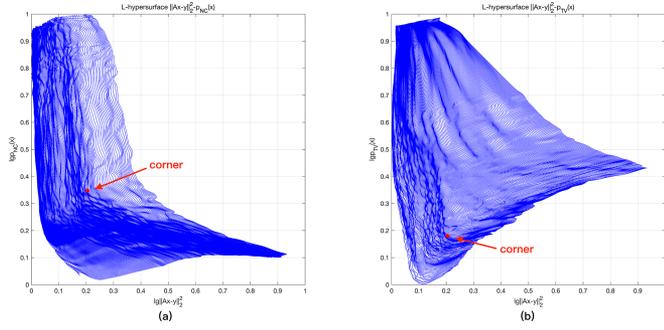


Fig. 5. Projection of L-hypersurface. (a) $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_{NC}(\mathbf{x})$ plane. (b) $\|\mathbf{Ax} - \mathbf{y}\|_2^2 - p_{TV}(\mathbf{x})$ plane.

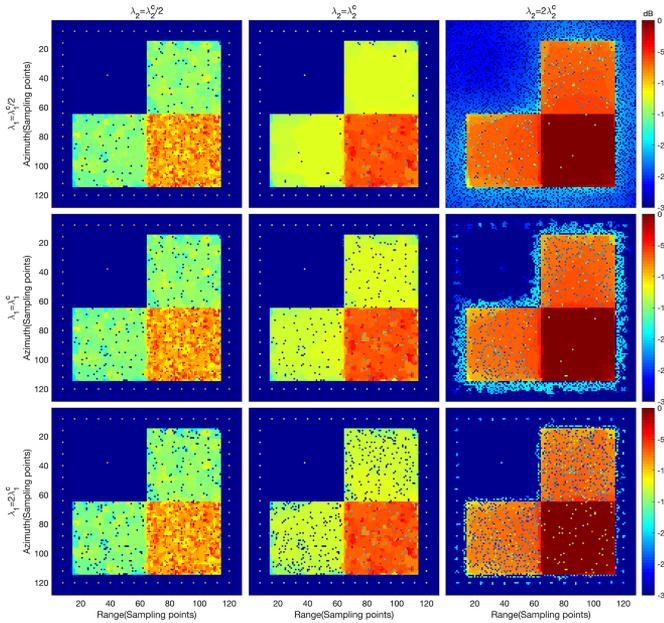


Fig. 6. SAR reconstruction images of simulation data via different pairs of regularization parameters.

parameters are decided by the corner

$$A_{\text{corner}} = \mathcal{L}(\lambda_1^c, \lambda_2^c) = \mathcal{L}(1.77, 6.00). \quad (24)$$

Fig. 6 presents reconstruction images via different pairs of regularization parameters $(\lambda_1, \lambda_2) \in \{\lambda_1^c/2, \lambda_1^c, 2\lambda_1^c\} \times \{\lambda_2^c/2, \lambda_2^c, 2\lambda_2^c\}$. The results indicate that $(\lambda_1^c, \lambda_2^c)$ give the best visual impression. Since the NC-TV regularization model can simultaneously enhance the point-based and region-based features, the speckle suppression and texture preserving performance should both be taken into consideration. In objective evaluation, amplitude of target background ratio (TBR) is used to evaluate point targets. A higher TBR value indicates a lower background noise. Equivalent number of looks (ENL) is used to evaluate distributed targets. When the ENL value is bigger, it indicates the image is smoothed well. Edge preservation index (EPI) are used to evaluate the texture preserving feature. The

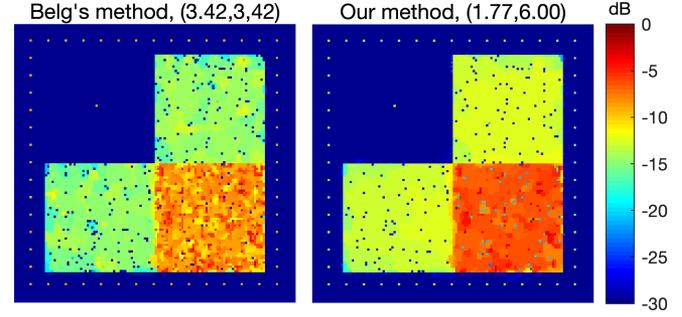


Fig. 7. Comparison between Belge's and our L-hypersurface method.

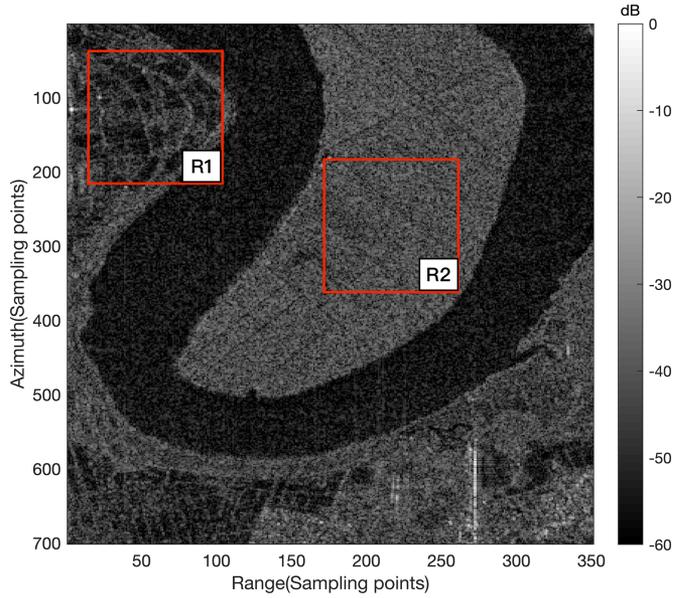


Fig. 8. Experimental scene (CSA) and selected targets.

definitions of TBR, ENL, and EPI are as follows:

$$\text{TBR} = 10 \lg \left(\frac{I_{\text{target}}}{\left(\sum_{(i,j) \in \mathcal{B}} I_{i,j} / N_{\mathcal{B}} \right)} \right) \quad (25)$$

where $\mathbf{I} = \{I_{i,j}\} = \{|\mathbf{X}|_{i,j}^2\}$ denotes the intensity matrix of image, I_{target} denotes the intensity of the evaluating point target, \mathcal{B} denotes the neighborhood of the target which has totally $N_{\mathcal{B}}$ pixels

$$\text{ENL} = \frac{\mu(\mathbf{I})^2}{\sigma(\mathbf{I})} \quad (26)$$

where $\mu(\cdot)$ calculates the mean and $\sigma(\cdot)$ calculates the standard deviation

$$\text{EPI} = \frac{\sum_{l=1}^L |\hat{\mathbf{x}}|_{l,1} - |\hat{\mathbf{x}}|_{l,2}}{\sum_{l=1}^L |\mathbf{x}|_{l,1} - |\mathbf{x}|_{l,2}} \quad (27)$$

where $|\hat{\mathbf{x}}|_{l,1}$ and $|\hat{\mathbf{x}}|_{l,2}$ denote the amplitude of reconstructed image on both sides of the edge, $|\mathbf{x}|_{l,1}$ and $|\mathbf{x}|_{l,2}$ denote the amplitude of reference image on both sides of the edge. Let Fig. 3 be the reference image in simulation experiment.

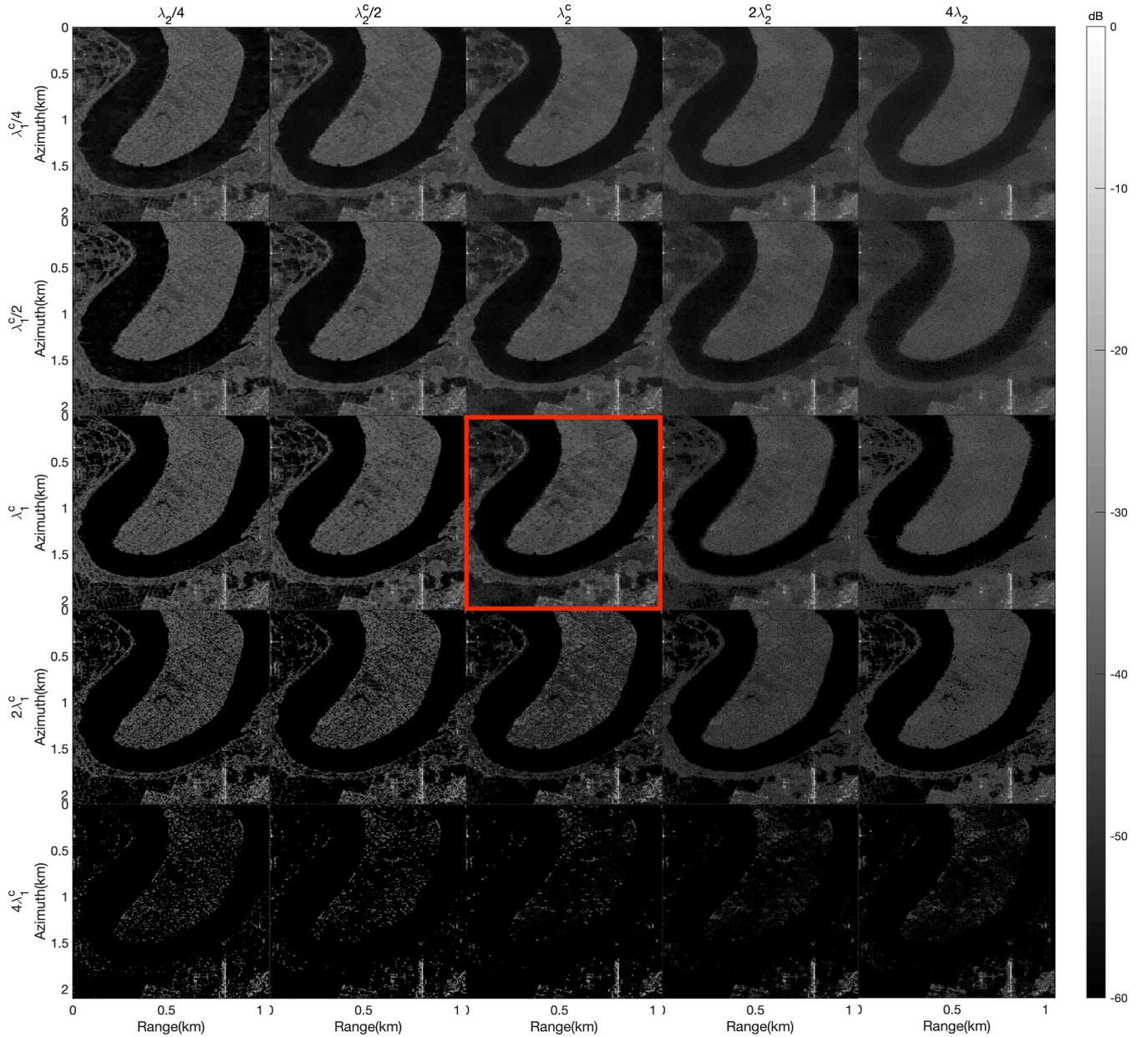


Fig. 9. SAR reconstruction images of real data via different pairs of regularization parameters.

As Fig. 3(a) shows, the TBR of point target P1 in neighborhood D1, the ENL of distributed targets D3, D4, the EPI of edges E1 and E2 are listed in Table I. The results corresponding to $(\lambda_1^c, \lambda_2^c)$ are in bold. According to the quantitative results, we draw the following conclusions.

- 1) The TBR of our selected parameters' reconstructed image is relatively high, which means a satisfactory denoising result. When $\lambda_1 = \lambda_1^c/2, \lambda_2 = \lambda_2^c$, the TBR becomes lower than the selected parameters, so that the image quality gets worse because of higher background noise. When $\lambda_2 = 2\lambda_2^c$ the point targets might be lost by reconstruction results.
- 2) The theoretical ENL of targets D3 and D4 in the original experiment scene, the amplitude of which is Rayleigh

distributed, is 1. Since the ENL of $(\lambda_1^c, \lambda_2^c)$'s reconstructed image is significantly improved, the selected parameters have a good speckle suppression performance. When $\lambda_2 = \lambda_2^c/2$, the ENL of targets D3 and D4 are still low. Although $\lambda_2 = 2\lambda_2^c$ has a better smoothing performance than λ_2^c , the noise is also excessively smoothed, so that the detailed feature might loss. More holes appear in the distributed target when $\lambda_1 = 2\lambda_1^c, \lambda_2 = \lambda_2^c$.

- 3) The EPI of $(\lambda_1^c, \lambda_2^c)$'s reconstructed image is close to 1, which means a superb texture preserving ability.

We also apply the Belge's L-hypersurface method [28] to the simulation data. Selected parameters of Belge's method are (3.42,3.42) and our method are (1.77,6.00). According to Fig. 7 and the analysis above, apparently, the optimum

TABLE II
ENL OF REAL RECONSTRUCTION RESULTS

$\lambda_1 \backslash \lambda_2$	37.5	75	150	300	600
17.5	2.1104	5.5098	22.8356	23.0734	14.8541
35	2.0826	4.9873	16.4872	12.5367	10.4246
70	1.8674	4.2174	11.1966	7.3961	7.4436
140	1.1579	1.4130	1.9288	3.8662	4.8727
280	0.2641	0.2287	0.1484	0.4102	0.5929

parameters of our method have better reconstruction performance. Less holes and smoother region can be observed in (1.77,6.00)'s reconstruction result.

To sum up, the simulation results successfully demonstrate the effectiveness of our L-hypersurface method.

2) *Real Data Experiments*: In order to show the performance of our parameter selection method, we conduct the experiment using the complex-value image of Gaofen-3 C-band SAR satellite. The main parameters of this data are as follows: the bandwidth of the signal is 60 MHz, the sampling frequency in the range direction is 66.66 MHz, the azimuth resolution is 3 m and the pulse repetition rate is 1484.64 Hz.

First, we use chirp scaling algorithm (CSA) to reconstruct the target scene. As shown in Fig. 8, there are many speckles in the image which cause the land to no longer be continuous and uniform. Then, we use the NC-TV regularization model for reconstruction and apply our L-hypersurface multiparameter selection method. The optimized regularization parameters are $(\lambda_1^c, \lambda_2^c) = (70, 150)$. Fig. 9 presents reconstruction images via different pairs of regularization parameters $(\lambda_1, \lambda_2) \in \{\lambda_1^c/4, \lambda_1^c/2, \lambda_1^c, 2\lambda_1^c, 4\lambda_1^c\} \times \{\lambda_2^c/4, \lambda_2^c/2, \lambda_2^c, 2\lambda_2^c, 4\lambda_2^c\}$. The highlighted reconstruction result generated by $(\lambda_1^c, \lambda_2^c)$, which has the best visual performance, achieves an ideal tradeoff between NC and TV penalties. When $\lambda_1 = 2\lambda_1^c$, a detail features lost happens in the field region R1 (marked on Fig. 8). When $\lambda_1 = \lambda_1^c/2$, the strong noise remains in the scene, resulting a weak edge preservation performance. When $\lambda_2 = 2\lambda_2^c$, the scene might be excessively smoothed, while $\lambda_2 = \lambda_2^c/2$ has weak speckle suppression ability and leaves distributed targets unsmoothed. When the value of λ_1 or λ_2 become even larger or smaller, the effect of corresponding penalty term is strengthened or weakened, leading to even worse reconstruction results. Table II gives the ENL of distributed targets marked as region R2 in Fig. 8. It can be inferred from the results that the selected parameters can simultaneously enhance region-based features while suppressing noise, fully demonstrating the effect of NC-TV regularization model.

B. TomoSAR Imaging

In this section, we apply the L-hypersurface method to the spatial regularization model (9) for SAR tomography reconstruction.

1) *Simulation Experiments*: The simulated data is generated corresponding to 40-channel complex SAR images and the system parameters of Yuncheng real data. A 0 dB Gaussian white noise is added to the raw echo wave and a uniformly distributed

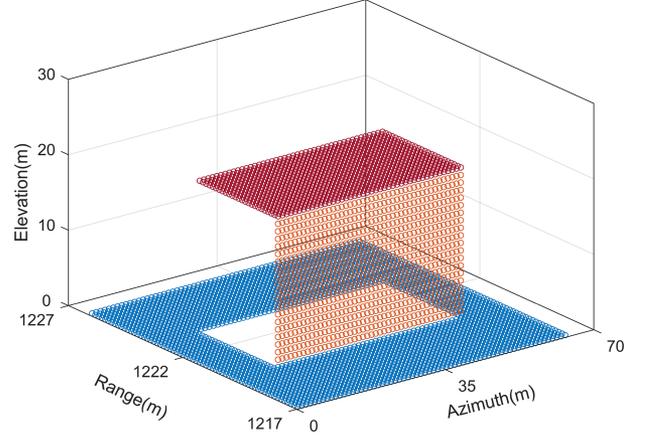


Fig. 10. TomoSAR simulation experiment scene.

random phase noise between $-\pi$ and π is added to the scene according to [10]. Fig. 10 shows the theoretical distribution of the scatters (ground truth). The scene are composed of a 64×64 pixels ground at altitude $h = 0$ m, a wall, and a 44×24 pixels roof at $h = 20$ m. The simulated building is higher than the estimated elevation resolution (as given by the Fourier inversion).

The optimized regularization parameters are $(\lambda_1^c, \lambda_2^c) = (2.4, 0.05)$. $(\lambda_1^c, \lambda_2^c)$ represent the optimized regularization parameters in model (9). Fig. 11 presents reconstruction results via different pairs of regularization parameters $(\lambda_1, \lambda_2) \in \{\lambda_1^c/4, \lambda_1^c, 4\lambda_1^c\} \times \{\lambda_2^c/4, \lambda_2^c, 4\lambda_2^c\}$, where Fig. 11(a) shows the 3-D scatters and Fig. 11(b) shows the elevation-range section of reconstruction results. The performance of selected parameters is objectively evaluated by the *accuracy* and *completeness* criteria introduced in [11] and [37]. For a discrete 3-D reconstruction result $\hat{\mathbf{P}}$, the accuracy and completeness are respectively defined as

$$A(\hat{\mathbf{P}}, \mathbf{P}) = \frac{1}{N_{\hat{\mathbf{P}}}} \sum_{j=1}^{N_{\hat{\mathbf{P}}}} \min_k \|\hat{p}_j - p_k\|_2 \quad (28)$$

$$C(\hat{\mathbf{P}}, \mathbf{P}) = \frac{1}{N_{\mathbf{P}}} \sum_{j=1}^{N_{\mathbf{P}}} \min_j \|\hat{p}_j - p_k\|_2 \quad (29)$$

where $N_{\hat{\mathbf{P}}}$ denotes the number of points in the reconstruction $\hat{\mathbf{P}}$, $N_{\mathbf{P}}$ denotes the number of points in the ground truth \mathbf{P} , \hat{p}_j denotes the j th point of $\hat{\mathbf{P}}$ and p_k denotes the k th point of \mathbf{P} . *Accuracy* represents the mean distance from each point in $\hat{\mathbf{P}}$ to the ground truth \mathbf{P} , indicating whether reconstructed points are correctly located. *Completeness* represents the mean distance from each point in \mathbf{P} to the points in reconstruction result $\hat{\mathbf{P}}$, indicating whether the ground truth is well represented by the set of points in the reconstruction. In Figs. 12 and 13, we plot accuracy as a function of completeness for different pairs of parameters, where Fig. 12 compares the performance of different λ_1 in model (9) as $\lambda_2 = \lambda_2^c (= \lambda_3 = \lambda_4)$ and Fig. 13 compares the performance of different λ_2 as $\lambda_1 = \lambda_1^c$. Based on the quantitative results, we draw to the following conclusions.

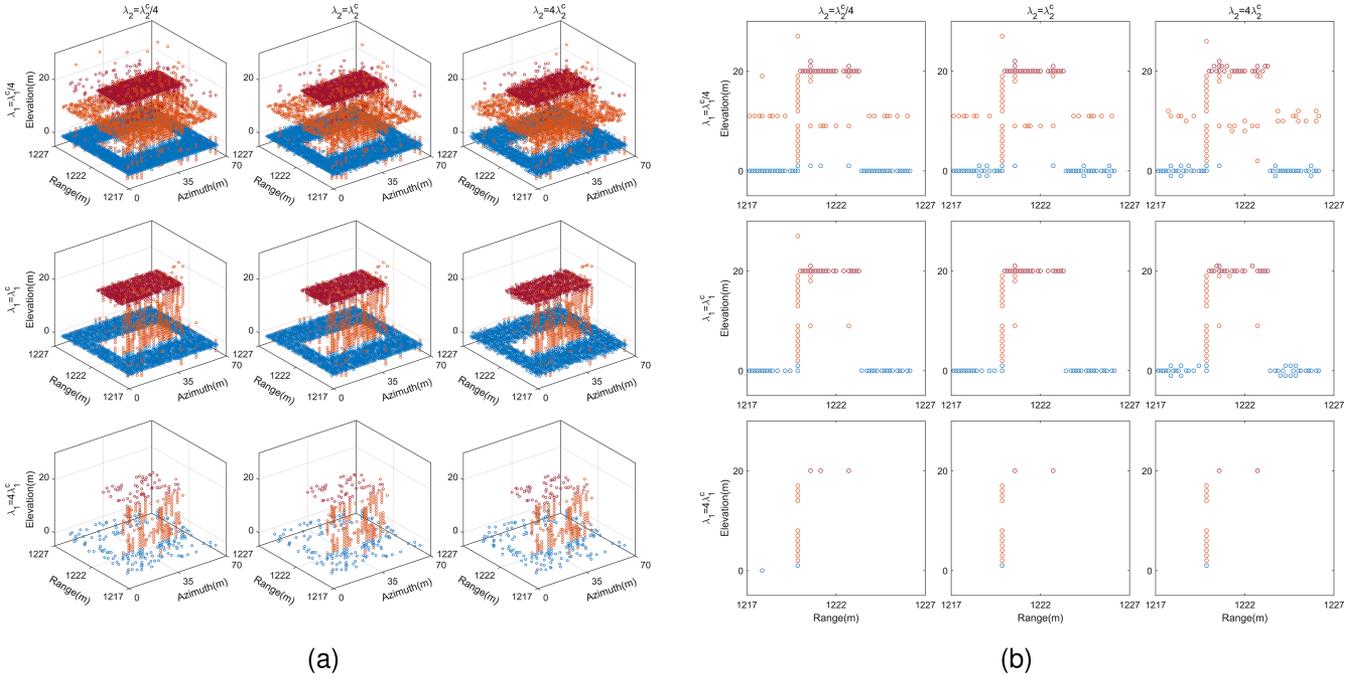


Fig. 11. TomoSAR reconstruction results of simulation data via different pairs of regularization parameters. (a) 3-D scatters. (b) Elevation-range sections.

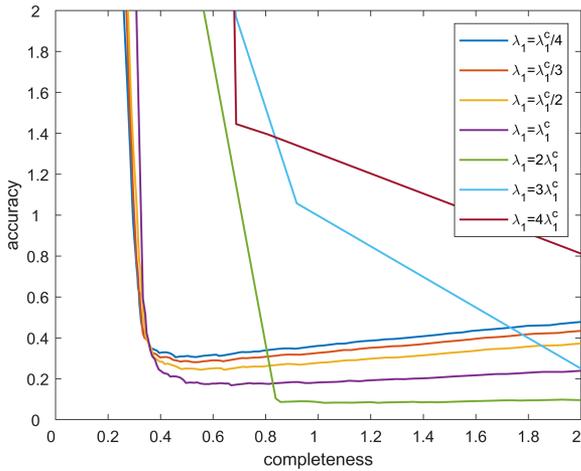


Fig. 12. Accuracy versus completeness as $\lambda_2 = \lambda_2^c = 2.4$.

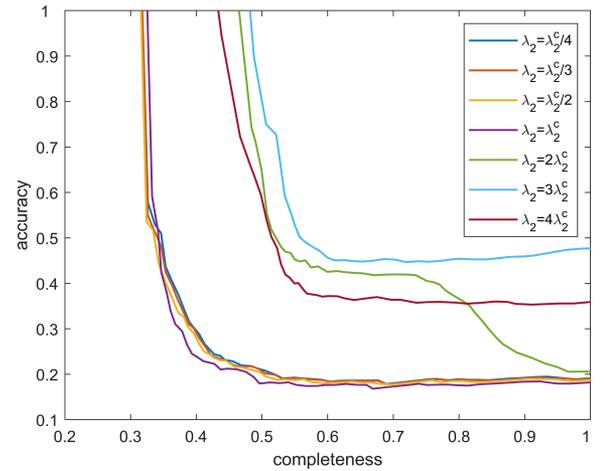


Fig. 13. Accuracy versus completeness as $\lambda_1 = \lambda_1^c = 0.05$.

- 1) When $\lambda_1 = \lambda_1^c/4$, numerous sidelobes and outliers are left in the scene. When $\lambda_1 = 4\lambda_1^c$, the reconstruction of roof and ground fails and the wall is overly sparsified. The accuracy versus completeness curve gradually moves away from the axes' origin while λ_1 is getting increasingly lower or higher than the optimized value λ_1^c .
- 2) When $\lambda_2 = \lambda_2^c/4$, the roof and ground are insufficiently smoothed and more holes appear in the wall. Relatively more outliers located far from the actual surfaces can be observed when the spatial regularization is too weak. When $\lambda_2 = 4\lambda_2^c$, the extension of structures in the elevation direction can be observed. The accuracy versus completeness curve shows the same trend as λ_1 's case,

while λ_2 is getting increasingly lower or higher than the optimized value λ_2^c .

By comparison, the best tradeoff between accuracy and completeness (point of the curve closest to the origin of the axes), is reached by selected parameters, indicating the effectiveness of proposed parameters selection method.

2) *Real Data Experiments:* Furthermore, we present the 3-D reconstruction results on one building. We conduct the experiment using the Aerospace Information Research Institute, Chinese Academy of Sciences array data acquired on the city of YunCheng in Shanxi province. The radar system operates at 15 GHz and has 8 channels in the crosstrack direction. The distance between the adjacent channels is 0.08 m. The height

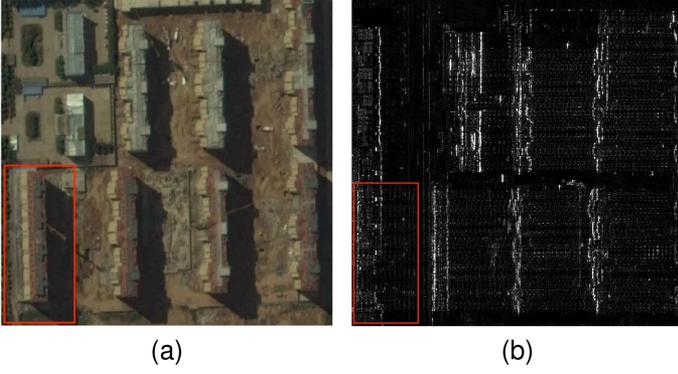


Fig. 14. Experimental area. (a) Optical image (copyright Google). (b) Intensity map.

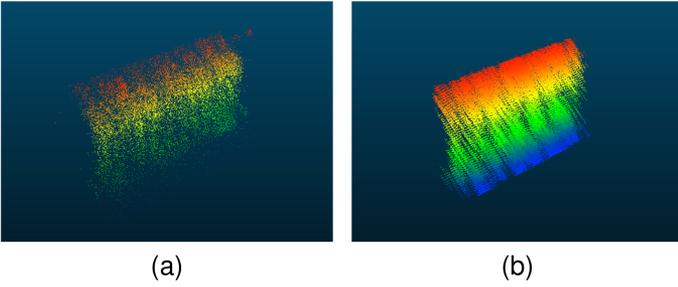


Fig. 15. TomoSAR reconstruction results of real data. (a) ℓ_1 regularization. (b) Spatial regularization.

of the radar platform is 972 m and the local incidence angle is 30° . The corresponding intensity maps of the areas are shown in Fig. 14. As we know, the height of highlighted building is 50 m.

For this scene, the regularization parameters are computed from the simulation experiments. Fig. 15(a) presents the reconstruction result of ℓ_1 regularization model. Fig. 15(b) presents the result of spatial regularization model, where the regularization parameters are selected by the L-hypersurface method. Compared with the ℓ_1 regularization, most of the outliers are suppressed and the building is reconstructed with higher completeness when the spatial regularization is used. The decreased in number of outliers and the increase in number of reconstruction scatters of the building show the effect of spatial smoothing. The results reflect a satisfied performance of the L-hypersurface multiparameter selection method in practical applications.

C. Computational Efficiency

Since the result of our L-hypersurface method should be obtained under different combinations of regularization parameters, the process can be very time consuming when high-precision optimal parameters are required. To improve the computational efficiency, we divide the corner determination process into two steps, rough calculation and precise calculation. Here are the following two strategies.

TABLE III
COMPARISON OF DIFFERENT STRATEGIES ($J_1 = J_2 = 500$)

Strategy	Sampling interval	$(\lambda_1^c, \lambda_2^c)$	Time Cost	
			Actual	Theoretical
No. 1	-	(1.77, 6.00)	646 min	250000T
No. 2	$U = 4$	(1.77, 6.00)	42.0 min	15986T
No. 2	$U = 10$	(1.77, 6.00)	13.2 min	4349T
No. 2	$U = 25$	(2.07, 4.79)	30.4 min	11009T

- 1) *Strategy No. 1*: Directly calculate the combinations of regularization parameters in sets

$$\{\lambda_1^1, \lambda_1^2, \lambda_1^3, \dots, \lambda_1^{J_1}\} \times \{\lambda_2^1, \lambda_2^2, \lambda_2^3, \dots, \lambda_2^{J_2}\}. \quad (30)$$

- 2) *Strategy No. 2*: First, rough calculation step. Calculate the combinations of regularization parameters in sets

$$\begin{aligned} & \{\lambda_1^1, \lambda_1^U, \lambda_1^{2U}, \lambda_1^{3U}, \dots, \lambda_1^{J_1}\} \\ & \times \{\lambda_2^1, \lambda_2^U, \lambda_2^{2U}, \lambda_2^{3U}, \dots, \lambda_2^{J_2}\} \end{aligned} \quad (31)$$

where U is a user selectable parameter that controls the sampling interval of rough calculation step. Second, precise calculation step. Let $(\lambda_1^{t_1 U}, \lambda_2^{t_2 U})$ denote the selected regularization parameters of the rough calculation step. Then, calculate the combinations of regularization parameters in sets

$$\begin{aligned} & \{\lambda_1^1, \lambda_1^{(t_1-2)U}, \lambda_1^{(t_1-2)U+1}, \dots, \lambda_1^{(t_1+2)U}, \lambda_1^{J_1}\} \\ & \times \{\lambda_2^1, \lambda_2^{(t_2-2)U}, \lambda_2^{(t_2-2)U+1}, \dots, \lambda_2^{(t_2+2)U}, \lambda_2^{J_2}\}. \end{aligned} \quad (32)$$

Suppose that the calculation time of each pair (λ_1, λ_2) is the same, denoted as T , the total time cost is approximately $J_1 \times J_2 \times T$ for strategy No. 1 and $\frac{J_1}{U} \times \frac{J_2}{U} \times T + (4U + 3)^2 \times T$ for strategy No. 2. We take the SAR simulation data in Section IV-A1 as an example. Table III gives the comparison of different strategies. The results indicate that a proper sampling interval U can significantly accelerate parameters selection process. An excessively large value of U will not only increase the time cost, but also lead to positioning deviation of the corner. We empirically suggest selecting U around $\lfloor \sqrt[4]{J_1 J_2} / 2 \rfloor$, where $\lfloor \cdot \rfloor$ denotes the floor operator, due to

$$\begin{aligned} & \arg \min_U \frac{J_1}{U} \times \frac{J_2}{U} \times T + (4U + 3)^2 \times T \\ & \approx \arg \min_U \frac{J_1}{U} \times \frac{J_2}{U} + 16U^2 \\ & = \frac{\sqrt[4]{J_1 J_2}}{2}. \end{aligned} \quad (33)$$

Based on the rough-and-precise strategy, the time cost of our L-hypersurface method is affordable.

V. CONCLUSION

An L-hypersurface parameters selection method for composite regularization is proposed. The proposed method is no longer limited to penalty terms of seminorm form, but can be

applied to regularization models with arbitrary penalty terms. Since reconstruction results based on the selected corner are better than other test points, both visually and numerically, the effectiveness of the proposed method has been verified by the simulation experiments. Experiments based on Gaofen-3 SAR satellites real data show that the optimized parameters have satisfactory reconstruction results, which implies the value of proposed method in practical application. The parameters selection performance of L-hypersurface in NC-TV regularization and spatial regularization indicates the strong potential of applying the L-hypersurface method to other composite regularization models.

The L-hypersurface method can be directly expanded to more than two regularization parameters, however, this will lead to more serious time consuming issue. Although an accelerating approach has been provided in this article, further researches focusing on improving the computational efficiency of locating the corner are considered by us. A strict mathematical analysis on the rationality of the proposed method is also expected.

REFERENCES

- [1] B. Zhang, W. Hong, and Y. Wu, "Sparse microwave imaging: Principles and applications," *Sci. China Inf. Sci.*, vol. 55, no. 8, pp. 1722–1754, 2012.
- [2] M. Çetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Trans. Image Process.*, vol. 10, no. 4, pp. 623–631, Apr. 2001.
- [3] Z. Xu, M. Liu, G. Zhou, Z. Wei, B. Zhang, and Y. Wu, "An accurate sparse SAR imaging method for enhancing region-based features via nonconvex and TV regularization," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 350–363, 2021, doi: [10.1109/JSTARS.2020.3034431](https://doi.org/10.1109/JSTARS.2020.3034431).
- [4] Z. Xu, B. Zhang, Z. Zhang, M. Wang, and Y. Wu, "Nonconvex-nonlocal total variation regularization-based joint feature-enhanced sparse SAR imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4515705, doi: [10.1109/LGRS.2022.3222185](https://doi.org/10.1109/LGRS.2022.3222185).
- [5] L. Prunte, "Elastic net regularization for SAR imaging from incomplete data," in *Proc. 12th Eur. Conf. Synthetic Aperture Radar*, 2018, pp. 1–6.
- [6] G. Zhou, M. Liu, Z. Xu, M. Wang, B. Zhang, and Y. Wu, "Azimuth ambiguities suppression using group sparsity and nonconvex regularization for sliding spotlight mode: Results on Qilu-1 SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 1660–1663.
- [7] S. Samadi, M. Çetin, and M. A. Masnadi-Shirazi, "Multiple feature-enhanced SAR imaging using sparsity in combined dictionaries," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 821–825, Jul. 2013.
- [8] L. Yang, S. Chen, S. Huan, H. Li, and M. Xing, "Structure-guaranteed SAR imagery via spatially-variant morphology regularization in ADMM manner," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5234414, doi: [10.1109/TGRS.2022.3197439](https://doi.org/10.1109/TGRS.2022.3197439).
- [9] A. Budillon, A. Evangelista, and G. Schirizzi, "Three-dimensional SAR focusing from multipass signals using compressive sampling," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 488–499, Jan. 2011.
- [10] X. X. Zhu and R. Bamler, "Tomographic SAR inversion by L_1 -norm regularization—The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, Oct. 2010.
- [11] C. Rambour, L. Denis, F. Tupin, and H. M. Oriot, "Introducing spatial regularization in SAR tomography reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8600–8617, Nov. 2019.
- [12] C. Rambour, A. Budillon, A. C. Johny, L. Denis, F. Tupin, and G. Schirizzi, "From interferometric to tomographic SAR: A review of synthetic aperture radar tomography-processing techniques for scatterer unmixing in urban areas," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 2, pp. 6–29, Jun. 2020.
- [13] A. Li, D. Chen, K. Lin, and G. Sun, "Hyperspectral image denoising with composite regularization models," *J. Sensors*, vol. 2016, 2016, Art. no. 6586032.
- [14] Z. Zhang, J. Zhang, Z. Wei, and L. Xiao, "Cartoon-texture composite regularization based non-blind deblurring method for partly-textured blurred images with poisson noise," *Signal Process.*, vol. 116, pp. 127–140, 2015.
- [15] M. Pham, X. Lin, A. Ruszczyński, and Y. Du, "An outer–inner linearization method for non-convex and nondifferentiable composite regularization problems," *J. Glob. Optim.*, vol. 81, pp. 179–202, 2021.
- [16] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the l-curve," *SIAM Rev.*, vol. 34, no. 4, pp. 561–580, 1992.
- [17] O. Batu and M. Cetin, "Parameter selection in sparsity-driven SAR imaging," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 4, pp. 3040–3050, Oct. 2011.
- [18] G. D. Martín-del Campo-Becerra, S. A. Serafín-García, A. Reigber, and S. Ortega-Cisneros, "Parameter selection criteria for tomo-SAR focusing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1580–1602, 2020, doi: [10.1109/JSTARS.2020.3042661](https://doi.org/10.1109/JSTARS.2020.3042661).
- [19] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Statist.*, vol. 9, no. 6, pp. 1135–1151, 1981.
- [20] S. Ramani, Z. Liu, J. Rosen, J.-F. Nielsen, and J. A. Fessler, "Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and sure-based methods," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3659–3672, Aug. 2012.
- [21] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.
- [22] M. Liu, B. Zhang, Z. Xu, and Y. Wu, "Efficient parameter estimation for sparse SAR imaging based on complex image and azimuth-range decouple," *Sensors*, vol. 19, no. 20, 2019, Art. no. 4549.
- [23] M. Grasmair and V. Naumova, "Multi-parameter approaches in image processing," in *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging: Mathematical Imaging and Vision*, Germany: Springer, 2023, pp. 943–967.
- [24] Y. Xu, Y. Pei, and F. Dong, "An extended l-curve method for choosing a regularization parameter in electrical resistance tomography," *Meas. Sci. Technol.*, vol. 27, no. 11, 2016, Art. no. 114002.
- [25] K. Kunisch and T. Pock, "A bilevel optimization approach for parameter learning in variational models," *SIAM J. Imag. Sci.*, vol. 6, no. 2, pp. 938–983, 2013.
- [26] E. De Vito, Z. Kereta, and V. Naumova, "Unsupervised parameter selection for denoising with the elastic net," 2018, *arXiv:1809.08696*.
- [27] T. T. Toma and D. S. Weller, "Fast automatic parameter selection for MRI reconstruction," in *Proc. IEEE 17th Int. Symp. Biomed. Imag.*, 2020, pp. 1078–1081.
- [28] M. Belge, M. E. Kilmer, and E. L. Miller, "Simultaneous multiple regularization parameter selection by means of the l-hypersurface with applications to linear inverse problems posed in the wavelet transform domain," in *Proc. SPIE Conf. Bayesian Inference Inverse Problems*, 1998, pp. 328–336.
- [29] S. Orintara, W. C. Karl, D. A. Castanon, and T. Q. Nguyen, "A method for choosing the regularization parameter in generalized Tikhonov regularized linear inverse problems," in *Proc. Int. Conf. Image Process.*, 2000, pp. 93–96.
- [30] J. L. Castellanos, S. Gómez, and V. Guerra, "The triangle method for finding the corner of the l-curve," *Appl. Numer. Math.*, vol. 43, no. 4, pp. 359–373, 2002.
- [31] A. Cultrera and L. Callegaro, "A simple algorithm to find the l-curve corner in the regularisation of ill-posed inverse problems," *IOP SciNotes*, vol. 1, no. 2, 2020, Art. no. 025004.
- [32] S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
- [33] I. Daubechies, M. DeFRise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.: A. J. Issued Courant Inst. Math. Sci.*, vol. 57, no. 11, pp. 1413–1457, 2004.
- [34] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [35] Z. Xu, B. Zhang, G. Zhou, L. Zhong, and Y. Wu, "Sparse SAR imaging and quantitative evaluation based on nonconvex and TV regularization," *Remote Sens.*, vol. 13, no. 9, 2021, Art. no. 1643.
- [36] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. Rijeka, Croatia: SciTech, 2004.
- [37] O. D'Hondt, C. López-Martínez, S. Guillaso, and O. Hellwich, "Non-local filtering applied to 3-D reconstruction of tomographic SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 272–285, Jan. 2018.



Yizhe Fan received the bachelor's degree in electronic information engineering in 2022 from the University of Chinese Academy of Sciences, Beijing, China, where he is currently working toward the Ph.D. degree in signal and information processing.

His research interests include radar signal processing, compressed sensing, sparse SAR imaging, and deep learning.



Guoru Zhou received the bachelor's degree in electronic information engineering from the Beijing Institute of Technology, Beijing, China, in 2020. She is currently working toward the Ph.D. degree in signal and information processing with the University of Chinese Academy of Sciences, Beijing, China.

Her research interests include radar signal processing, compressed sensing, and sparse SAR imaging.



Kun Wang received the bachelor's degree in electronic information engineering in 2022 from the University of Chinese Academy of Sciences, Beijing, China, where he is currently working toward the M.S. degree in signal and information processing.

His research interests include TomoSAR imaging and deep learning.



Bingchen Zhang received the bachelor's degree in electronic engineering and information science from the University of Science and Technology of China, Hefei, China, in 1996 and the M.S. degree and the Ph.D. degree in signal and information processing from the Institute of Electronics, Chinese Academy of Sciences (IECAS), Beijing, China, in 1999 and 2017, respectively.

Since 1999, he has been a Scientist with IECAS. His research interests include synthetic aperture radar (SAR) signal processing and airborne SAR system design, implementation, and data processing.



Jie Li received the bachelor's degree in electronic information engineering from the Beijing Institute of Technology, Beijing, China, in 2019. She is currently working toward the Ph.D. degree in signal and information processing with the University of Chinese Academy of Sciences, Beijing, China.

Her research interests include TomoSAR imaging and deep learning.



Yirong Wu received the Ms.D. degree in microwave electromagnetic field from the Beijing Institute of Technology, Beijing, China, in 1988 and the Ph.D. degree in signal and information processing from the Institute of Electronics, Chinese Academy of Sciences (IECAS), Beijing, China, in 2001.

Since 1988, he has been with IECAS, where he currently serves as the Director. He has over 20 years of experience in remote-sensing processing system design. His research interests include microwave imaging, signal and information processing, and related

applications.