

Article

Sub-Nyquist SAR Imaging and Error Correction Via an Optimization-Based Algorithm

Wenjiao Chen ¹ , Li Zhang ^{2,*}, Xiaocen Xing ¹, Xin Wen ¹ and Qiuxuan Zhang ¹

¹ The Department of Space Control and Communications, Space Engineering University, Beijing 102249, China; chenwenjiao008@163.com (W.C.); xxc012345@163.com (X.X.); wx19951017@126.com (X.W.); zhangqiuxuan@yeah.net (Q.Z.)

² The 15th Research Institute of China Electronics Technology Corporation, Beijing 100083, China

* Correspondence: zhangli35@cetc.com.cn

Abstract: Sub-Nyquist synthetic aperture radar (SAR) based on pseudo-random time–space modulation has been proposed to increase the swath width while preserving the azimuthal resolution. Due to the sub-Nyquist sampling, the scene can be recovered by an optimization-based algorithm. However, these methods suffer from some issues, e.g., manually tuning difficulty and the pre-definition of optimization parameters, and a low signal–noise ratio (SNR) resistance. To address these issues, a reweighted optimization algorithm, named pseudo- \mathcal{L}_0 -norm optimization algorithm, is proposed for the sub-Nyquist SAR system in this paper. A modified regularization model is first built by applying the scene prior information to nearly acquire the number of nonzero elements based on Bayesian estimation, and then this model is solved by the Cauchy–Newton method. Additionally, an error correction method combined with our proposed pseudo- \mathcal{L}_0 -norm optimization algorithm is also present to eliminate defocusing in the motion-induced model. Finally, experiments with simulated signals and strip-map TerraSAR-X images are carried out to demonstrate the effectiveness and superiority of our proposed algorithm.

Keywords: sub-Nyquist synthetic aperture radar (SAR) imaging; error correction; optimization-based algorithm; Bayesian estimation



Citation: Chen, W.; Zhang, L.; Xing, X.; Wen, X.; Zhang, Q. Sub-Nyquist SAR Imaging and Error Correction Via an Optimization-Based Algorithm. *Sensors* **2024**, *24*, 2840. <https://doi.org/10.3390/s24092840>

Academic Editor: Guangcai Feng

Received: 5 March 2024

Revised: 26 April 2024

Accepted: 28 April 2024

Published: 29 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

High-resolution wide-swath (HRWS) synthetic aperture radar (SAR) provides a short repeat cycle so that it has a high-efficiency acquisition capability [1,2]. Although the azimuthal multi-channel SAR [3] and multi-input multi-output (MIMO) SAR [4] can achieve a high azimuthal resolution and wide-range swath, it has a large amount of data and a large antenna. As the compressive sensing (CS) theorem posits [5–7], an innovative system concept called sub-Nyquist SAR based on pseudo-random time–space modulation breaks the limitation of the Nyquist sampling theorem with a single channel [8]. Simultaneously, the spatial and temporal phase modulation based on the optimization-based algorithm guarantees azimuthal resolution and wide-swath coverage mosaicked by several range sub-swaths. It adopts sub-Nyquist sampling along the azimuthal dimension, and the observed scene is recovered by sub-Nyquist SAR imaging [8].

In spite of the sub-Nyquist SAR achieving the above merits, sub-Nyquist SAR imaging still has two limitations, as follows:

- Under the assumption of satisfying restricted isometry property (RIP) [9], CS algorithms are vital to sub-Nyquist SAR imaging and include greedy algorithms [10,11], the \mathcal{L}_1 -norm optimization algorithms [12,13], and Bayesian-based methods [14–16], where the \mathcal{L}_1 -norm optimization algorithm has a better performance in terms of the recovered error evaluated by the mean square error (MSE) [17–19]. Although the \mathcal{L}_1 -norm optimization algorithm has achieved a better-recovered performance, these

methods still suffer from some problems, i.e., manually tuning difficulty and the pre-definition of optimization parameters (e.g., regularization parameter and thresholding parameter), and a low signal–noise ratio (SNR) resistance. However, these CS methods do not take full advantage of the scene prior information that we may hold, and sparse property is imposed uniformly and independently on each variable. Some low signal–noise ratio (SNR) targets in the sparse scene cannot be accurately recovered and it often yields false targets by the \mathcal{L}_1 -based method. Although some reweighted optimization-based algorithm has already been proposed [20,21], there is still no knowledge of how and why to select an approximately fair rule in sub-Nyquist SAR imaging to further mitigate the impact of empirical parameter setting on reconstructed performance.

- In addition, the imaging process requires the knowledge of motion parameters, e.g., radar position and radar equivalent velocity [5]. However, the radar platform may deviate from the pre-defined track, and the equivalent velocity is estimated by the curve-fitting method or approximate expression in a practical application [22]; uncertainties and errors may be introduced into the motion-induced model so that the recovered scene may defocus to decrease the image quality [23–26]. A technology called auto-focusing removes these phase errors [27]. In recent years, many sparsity-driven algorithms [28–34] have been proposed to solve the defocusing problem and achieve an effective performance. However, the references [28–33] do not fully formulate the motion error and adopt an approximate expression so that the error is not removed. The reference [34] integrated the deep SAR imaging algorithm to remove the motion error.

In this paper, we propose a pseudo- \mathcal{L}_0 -norm optimization algorithm based on Bayesian estimation to further improve the sub-Nyquist SAR imaging performance. The proposed algorithm penalizes the regularization item with the scene prior information, i.e., the reciprocal of its previous solution, to nearly acquire the number of nonzero values. Since this method approximates a \mathcal{L}_0 -based algorithm, which needs to be minimized in the sparse recovery, we named it a pseudo- \mathcal{L}_0 -norm optimization algorithm. Sub-Nyquist SAR imaging includes three steps: range compression, range cell migration correction (RCMC), and azimuth compression [5]. The traditional matched filtering (MF) method is adopted for range compression. After RCMC, a pseudo- \mathcal{L}_0 -norm optimization algorithm is adapted to achieve azimuth compression. This method not only takes advantage of the scene prior information but also establishes an approximately fair penalized rule so that it can recover low SNR targets and remove false targets compared to the \mathcal{L}_1 -norm optimization algorithm. In addition, an error correction method integrated with a pseudo- \mathcal{L}_0 -norm optimization algorithm eliminates the influence of phase error and improves the image quality. This method includes two steps: scene recovery based on a pseudo- \mathcal{L}_0 -norm optimization algorithm and error estimation by minimizing the least-squares target function. The two steps are successively iterative and the recovered matrix is updated according to the estimated error. For the exact expression of error is not fully formulated, we use the random phase error regardless of the error expression. The numerical simulation results are detailed in the following sections to make an evident advantage of the proposed algorithm.

The rest of this paper is organized as follows. In Section 2, the observation model for the sub-Nyquist SAR based on the pseudo-random time–space modulation is first described. Then, we build a pseudo- \mathcal{L}_0 -norm regularization model based on Bayesian estimation, and this regularization model can be solved by the Cauchy–Newton method. Additionally, an error correction method integrated with a pseudo- \mathcal{L}_0 -norm optimization algorithm is proposed to remove errors and eliminate defocusing. Simulation experiments and data experiments with real TerraSAR-X images confirm the effectiveness and superiority of our proposed method, presented in Section 3. In Section 4, the discussion shows the performance and advantages of our proposed algorithm. Section 5 concludes this paper.

2. Materials and Methods

In this section, the observation model of the sub-Nyquist SAR based on the pseudo-random time–space modulation is described. To further improve the sub-Nyquist SAR imaging performance, a pseudo- \mathcal{L}_0 -norm regularization model is built and solved by the Cauchy–Newton method to obtain a pseudo- \mathcal{L}_0 -norm optimization algorithm. In addition, an error correction method integrated with our proposed pseudo- \mathcal{L}_0 -norm optimization algorithm is proposed to eliminate the effect of the phase error and achieve autofocus.

2.1. Sub-Nyquist SAR Imaging and Error Correction Signal Models Based on Pseudo-Random Time–Space Modulation

2.1.1. Sub-Nyquist SAR Imaging Model

For traditional HRWS systems, e.g., the azimuthal multi-channel SAR [3] and MIMO SAR [4], the equivalent sampling still satisfies the Nyquist theorem. To lower the amount of data and break the conflict between high resolution and wide swath, the sub-Nyquist SAR based on pseudo-random time–space modulation has been proposed [8]. The imaging geometry of the sub-Nyquist SAR is shown in Figure 1, where η is slow time along the azimuth and $R_i(\eta)$ represents the range between the radar and the point target located at the coordinate $(x_i, y_i, 0)$ at the azimuth time η . x_i and y_i denote the azimuth and range coordinates, respectively. Because a two-dimensional image, i.e., azimuth and range, is considered, the coordinate $(x_i, y_i, 0)$ is simplified as (x_i, y_i) .

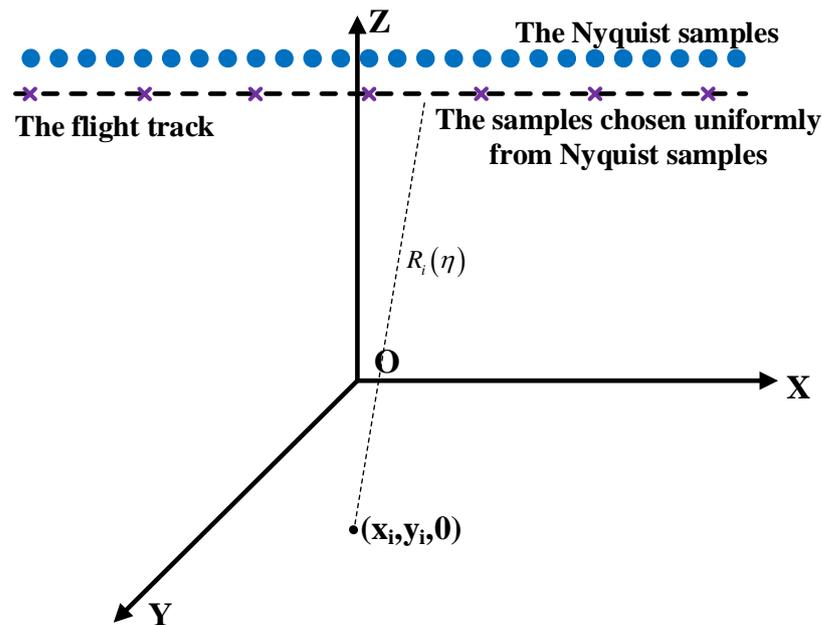


Figure 1. The imaging geometry of the sub-Nyquist SAR. ● denotes the Nyquist samples. × demonstrates the real azimuthal samples chosen uniformly from Nyquist samples in the sub-Nyquist SAR system based on the pseudo-random time–space modulation.

Similarly to the traditional SAR system, the raw data in the sub-Nyquist SAR for many point targets can be written as:

$$s_c(\tau, \eta) = \sum_{x_i, y_i} \sigma_i W_i(\tau, \eta) \exp\left\{j\pi K_r [\tau - 2R_i(\eta)/c]^2\right\} \cdot \exp\{-j4\pi R_i(\eta)/\lambda\} \cdot \exp\{j\varphi_i(\eta)\} + n(\tau, \eta) \quad (1)$$

where τ is the fast time along the range dimension, i.e., the sampling moment during one pulse width. η is the slow time along the azimuth dimension, i.e., the moment of the transmitting pulse. σ_i and $W_i(\tau, \eta)$ are the backscattering coefficient and the weighting pattern corresponding to the i -th target at (x_i, y_i) , respectively. $\varphi_i(\eta)$ is the random phase based on the pseudo-random time–space modulation corresponding to the target (x_i, y_i) .

K_r denotes the chirp rate of the linear frequency-modulated (LFM) signal, c is the light speed, λ is the wavelength, and $n(\tau, \eta)$ is the system noise.

To relieve the inherent contradiction in the HRWS system, sub-Nyquist sampling in the sub-Nyquist SAR is implied along the azimuthal dimension, and the sampling method is as shown in Figure 1. The sub-Nyquist SAR system only adopts the traditional MF method to not recover the scene exactly [5]. Accordingly, sub-Nyquist SAR imaging includes three steps: (1) range compression based on MF; (2) range cell migration correction (RCMC): interpolate by zero padding every azimuthal signal after range compression and then sum up all the range-compressed signal of grids on the same range cell to the pre-defined grid; and (3) azimuth reconstruction with sub-Nyquist samples based on the CS algorithm. After implementing range compression and RCMC to Equation (1), the signal at a certain range cell is represented by:

$$s_c(\tau_0, \eta) = \sum_{x_i, y_i} \sigma_i W_i(\tau_0, \eta) T_r \sin c \left\{ K_r T_r \left(\tau_0 - 2 \frac{R_i(\eta - \eta_{ci})}{c} \right) \right\} \cdot \exp \left\{ -j4\pi \frac{R_i(\eta - \eta_{ci})}{\lambda} \right\} \cdot \exp \{ j\varphi_i(\eta) \} + n(\tau_0, \eta) \quad (2)$$

where η_{ci} is the beam center crossing time for the target (x_i, y_i) . T_r is the pulse width and $\sin c(\cdot)$ denotes the *sinc* function.

Let $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_M]^T$ be the vectorized backscattering cross-sections of targets on the same range cell, and $\mathbf{s}_{N \times 1} = [s_c(\tau_0, \eta_1), s_c(\tau_0, \eta_2), \dots, s_c(\tau_0, \eta_N)]^T$ be the vectored signal after range compression and RCMC; then,

$$\mathbf{s}_{N \times 1} = \mathbf{D}_{N \times M} \sigma_{M \times 1} + \mathbf{n}_{N \times 1} \quad (3)$$

where N is the sampling number on the azimuthal dimension and M is the number of resolution cells at the certain range cell in the observed scene. $\mathbf{D}_{N \times M} = \{D_i(\tau_0, \eta_n)\}_{n=1, i=1}^{N, M}$ denotes the mapping relation between the received signal and the scene, and $D_i(\tau_0, \eta_n) = W_i(\tau_0, \eta_n) \cdot T_r \cdot \exp \{ -j4\pi R_i(\eta_n - \eta_{ci}) / \lambda \} \cdot \exp \{ j\varphi_i(\eta) \}$ $\mathbf{n}_{N \times 1} = [n(\tau_0, \eta_1), n(\tau_0, \eta_2), \dots, n(\tau_0, \eta_N)]^T$ is the noise.

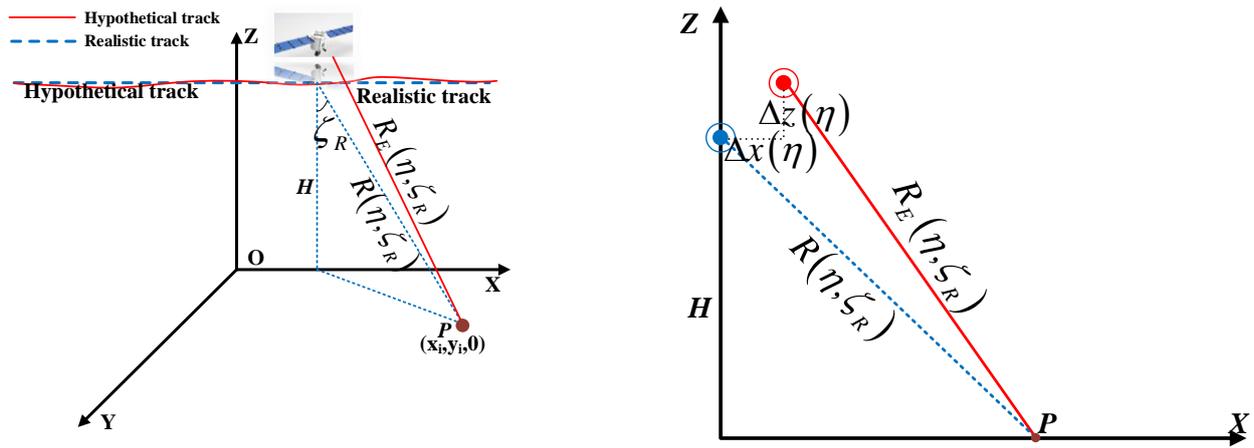
2.1.2. Motion Error Model

The recovered scene $\sigma_{M \times 1}$ can be estimated by Equation (3), while the recovered matrix $\mathbf{D}_{N \times M}$ can exactly reflect the relation between the raw data and the recovered scene. However, uncertainties and errors exist in the matrix $\mathbf{D}_{N \times M}$, and the inaccuracy of the motion-induced model leads to the phase error [22]. If the preset matrix $\mathbf{D}_{N \times M}$ in Equation (3) is still used to recover the scene $\sigma_{M \times 1}$ without extra processing steps, it may cause the defocusing of the reconstructed scene. A technology called autofocusing removes these phase errors.

To solve the problem of error correction, it is important to establish the exact mapping model so that we need to discern the uncertain factors. We mainly considered the position error and velocity error leading to the phase error in this paper. The SAR satellite is affected by different perturbations during on-orbit flying; so, it may cause the position error, as demonstrated in the following Figure 2, and the velocity error of the satellite. Firstly, we analyzed the position errors. The position error is mainly the deviation between the realistic position and the hypothetical position on the x-axis and z-axis. In Figure 2a, the solid line denotes the realistic track, and the dashed line denotes the hypothetical track. The hypothetical slant range between the radar and target without position error is:

$$R(\eta, \zeta_R) = \sqrt{\left(\frac{H}{\cos \zeta_R} \right)^2 + (V_e \eta - V_e \eta_0)^2} \approx \frac{H}{\cos \zeta_R} + \frac{V_e^2 \cos \zeta_R}{2H} (\eta - \eta_0)^2 \quad (4)$$

where V_e is the radar velocity. η_0 , the Doppler center moment, is the azimuthal moment when the radar is the nearest to the target.



(a) The imaging geometry with the position error.

(b) The imaging geometry in the XOZ plane.

Figure 2. The imaging geometry with the position error. The solid line denotes the realistic track, and the dashed line denotes the hypothetical track. $R(\eta, \zeta_R)$ and $R_E(\eta, \zeta_R)$ are the hypothetical slant range and real slant range with the position error, respectively. ζ_R is the pitch angle. P is the target point. H is the orbital height. Figure (b) is the projection of target point P on the XOZ plane. $\Delta x(\eta)$ and $\Delta z(\eta)$ are the range between the realistic position and the hypothetical position on the x-axis and z-axis, respectively.

When a position error exists, the real slant range $R_\zeta(\eta, \zeta_R)$ between the radar and target is:

$$\begin{aligned} R_\zeta(\eta, \zeta_R) &= \sqrt{(H \tan \zeta_R - \Delta x(\eta))^2 + (V_e \eta - V_e \eta_0)^2 + (H + \Delta z(\eta))^2} \\ &\approx \frac{H}{\cos \zeta_R} + \frac{V_e^2 \cos \zeta_R}{2H} (\eta - \eta_0)^2 + \Delta z(\eta) \cdot \cos \zeta_R - \Delta x(\eta) \cdot \sin \zeta_R \\ &\approx R(\eta, \zeta_R) + \Delta R_\zeta(\eta, \zeta_R) \end{aligned} \quad (5)$$

$\Delta R_\zeta(\eta, \zeta_R) = \Delta z(\eta) \cdot \cos \zeta_R - \Delta x(\eta) \cdot \sin \zeta_R$ is the position error. Regardless of the weight $W_i(\tau_0, \eta_n) \cdot T_r$ in Equation (2), the realistic echo signal $s_{error}(\eta, \zeta_R)$ of target point P at a certain range cell after range compression and RCMC is denoted as

$$\begin{aligned} s_{error}(\eta, \zeta_R) &= \sigma_P \cdot \exp\left\{-j\frac{4\pi}{\lambda} R_\zeta(\eta, \zeta_R)\right\} \\ &= \sigma_P \cdot \exp\left\{-j\frac{4\pi}{\lambda} (R(\eta, \zeta_R) + \Delta R_\zeta(\eta, \zeta_R))\right\} \\ &= s_0(\eta, \zeta_R) \cdot H_{\Delta R}(\eta, \zeta_R) \end{aligned} \quad (6)$$

where $s_0(\eta, \zeta_R) = \sigma_P \cdot \exp\{-j4\pi R(\eta, \zeta_R)/\lambda\}$ is the raw signal without the position error. $H_{\Delta R}(\eta, \zeta_R)$ is the position error signal.

Secondly, we analyzed the mathematical model for the velocity error. Regardless of the weight $W_i(\tau_0, \eta_n) \cdot T_r$ in Equation (2), the realistic echo signal $s_{error}(\eta, V_{Ee})$ of target point P at a certain range cell after range compression and RCMC is:

$$s_{error}(\eta, V_{Ee}) = \sigma_P \cdot \exp\left\{-j\frac{4\pi R_\zeta(\eta, V_{Ee})}{\lambda}\right\} \quad (7)$$

where V_{Ee} is the equivalent velocity. The raw signal is approximate to the linear frequency modulation signal and is written as:

$$s_{error}(\eta, V_{Ee}) = \exp\left\{f_d \eta + f_r (V_{Ee}) \eta^2\right\} \quad (8)$$

where f_d is the Doppler center frequency and the Doppler modulated rate is $f_r(V_{Ee}) = 2V_{Ee}^2 \cos^3 \zeta_A / \lambda R$, and ζ_A is the squint angle on the azimuthal dimension. When the hypothetical equivalent velocity V_E exists, $f_r(V_{Ee})$ is unfolded at $V_{Ee} = V_E$:

$$f_r(V_{Ee}) = f_r(V_E) + f_r'(V_E)(V_{Ee} - V_E) + \frac{1}{2!}f_r''(V_E)(V_{Ee} - V_E)^2 + \cdots + \frac{1}{n!}f_r^{(n)}(V_E)(V_{Ee} - V_E)^n + \cdots \quad (9)$$

If the error caused by the equivalent velocity is small, the above equation besides the first item $f_r(V_E)$ is also small and is denoted as $o_r(V_{Ee})$:

$$o_r(V_{Ee}) = f_r'(V_E)(V_{Ee} - V_E) + \frac{1}{2!}f_r''(V_E)(V_{Ee} - V_E)^2 + \cdots + \frac{1}{n!}f_r^{(n)}(V_E)(V_{Ee} - V_E)^n + \cdots \quad (10)$$

Substituting (24) and (25) into (23),

$$\begin{aligned} s_{error}(\eta, V_{Ee}) &= \exp\{f_d\eta + f_r(V_{Ee})\eta^2\} = \exp\{f_d\eta + f_r(V_E)\eta^2 + o_r(V_{Ee})\eta^2\} \\ &= \exp\{f_d\eta + f_r(V_E)\eta^2\} \cdot \exp\{o_r(V_{Ee})\eta^2\} \\ &= s_0(\eta, V_E) \cdot \exp\{o_r(V_{Ee})\eta^2\} = s_0(\eta, V_E) \cdot H_{V_{Ee}}(\eta) \end{aligned} \quad (11)$$

where $s_0(\eta, V_E) = \exp(f_d\eta + f_r(V_E)\eta^2) = \exp\{-j4\pi R(\eta, \zeta_R)/\lambda\}$ is the azimuthal signal without error and $H_{V_{Ee}}(\eta) = \exp(o_r(V_{Ee})\eta^2)$ is the error signal caused by the equivalent velocity V_{Ee} .

According to Equations (21) and (26), the azimuthal signal with the error can be written as the multiplication between the azimuthal signal without the error and the phase error. Considering the above phase error, the received raw data without the weight $W_i(\tau_0, \eta_n) \cdot T_r$ in Equation (2) after range compression and RCMC at a certain range cell is approximately denoted as:

$$s_{error}(\tau_0, \eta) = \sum_{i=1}^M \sigma_i \exp\left\{-j4\pi \frac{R_i(\eta)}{\lambda}\right\} \cdot \exp\{j\vartheta\} + n(\tau_0, \eta) \quad (12)$$

where ϑ is the error phase variation with the azimuthal sampling moment. Similarly to Equation (3), the vector matrix form of the raw signal (27) with an error along the azimuthal sampling moment η is:

$$\mathbf{s}_{N \times 1}^{error} = \mathbf{E}_{N \times N} \mathbf{D}_{N \times M} \boldsymbol{\sigma}_{M \times 1} + \mathbf{n}_{N \times 1} \quad (13)$$

where $\mathbf{E}_{N \times N} = \text{diag}(\exp(j\vartheta_1), \exp(j\vartheta_2), \dots, \exp(j\vartheta_N))$. A different error has a different expression form of the error matrix $\mathbf{E}_{N \times N}$, and the motion error has been not fully formulated yet due to some approximations in the models. Therefore, we assigned an error matrix $\mathbf{E}_{N \times N}$ random phase in the following simulation.

2.2. Sub-Nyquist SAR Imaging and Error Correction Based on the Pseudo- \mathcal{L}_0 -Norm Optimization Algorithm

In this subsection, the CS theorem is first briefly introduced. Then, a pseudo- \mathcal{L}_0 -norm regularization model is presented and solved by the Cauchy–Newton method in detail. Finally, sub-Nyquist SAR imaging and error correction based on pseudo-random space-time modulation are proposed based on our proposed pseudo- \mathcal{L}_0 -norm optimization algorithm.

2.2.1. CS Theorem

The key to the CS theorem is the effective recovered algorithm. While RIP is satisfied, there are three recovered algorithms, i.e., greedy algorithm, \mathcal{L}_1 -norm optimization algorithm, and sparse Bayesian learning method. The \mathcal{L}_1 -norm optimization algorithm is robust in the sense that it can effectively recover nearly sparse signals with/without measurement noise from remarkably few measurements, and its application is so wide that it broadly could be considered the modern least squares [12,13]. The \mathcal{L}_1 -norm optimization algorithm has a better-recovered performance in terms of the recovered error evaluated by

MSE [17]. Its optimization equation is to solve the underdetermined problem (3) without a subscript:

$$\hat{\sigma} = \underset{\sigma}{\operatorname{argmin}} \left\{ \|\mathbf{s} - \mathbf{D}\sigma\|_2^2 + \alpha \|\sigma\|_1 \right\} \quad (14)$$

where α is the regularization parameter. The first item $\|\mathbf{s} - \mathbf{D}\sigma\|_2^2$ ensures the recovery error, and the second item $\|\sigma\|_1$ guarantees the sparsity of the recovered scene. The parameter α balances between the recovered error and the sparsity, and is empirically chosen by minimizing the recovered error of the whole scene.

Even though the \mathcal{L}_1 -norm optimization algorithm achieved a good recovered performance, some low SNR targets cannot be accurately recovered, and it often yields false targets for an unfair penalization rule. The further optimization of the regularization model contributes to the recovered scene so that false targets can be removed, and then we carried out research about the optimization model. Is there an alternative to the \mathcal{L}_1 -norm regularization model to achieve a better-recovered performance? The problem of further improving the optimization algorithm is thus brought up. To solve this problem, a reweighted \mathcal{L}_1 -based algorithm has been already proposed to improve the recovered performance [20,21]. However, how to select an approximately fair regularization rule still does not have a theoretical analysis.

2.2.2. Sub-Nyquist SAR Imaging Based on the Pseudo- \mathcal{L}_0 -Norm Optimization Algorithm

For a sparse undetermined equation, the \mathcal{L}_0 -norm optimization algorithm should have the best recovered performance, but this optimization equation is a non-polynomial hard (NP-hard) problem [12]. Based on the idea of adaptive least absolute shrinkage and selection operator (Lasso) technique that is a popular technique for simultaneous estimation and variable selection [21], this paper penalized the regularization item to be close to the \mathcal{L}_0 -norm to achieve a better performance of the \mathcal{L}_0 -based algorithm, and the regularization model can be solved by the Cauchy–Newton method. Since it is closer to the \mathcal{L}_0 -based method in some sense, we named it a pseudo- \mathcal{L}_0 -norm optimization algorithm. In the following, we adopted Bayesian estimation to analyze and deduct a pseudo- \mathcal{L}_0 -norm regularization model by making full use of the scene prior information. The rule was deducted as follows: Usually, the noise $\mathbf{n}_{N \times 1}$ is assumed to be a Gaussian distribution with zero mean and variance σ_n^2 [14]. For simplicity, all the following matrixes/vectors omit subscripts.

$$p_n(n) = p_{\mathbf{s}/\sigma}(\mathbf{s}/\sigma) = \left(\frac{1}{\sqrt{2\pi}\sigma_n} \right)^N \exp \left\{ -\frac{\|\mathbf{s} - \mathbf{D}\sigma\|_2^2}{2\sigma_n^2} \right\} \quad (15)$$

Laplace distribution forces most coefficients to be small so that it can describe the sparse scene [35]. We assume:

$$p(\sigma_i) = \left(\frac{\xi_i}{2} \right) \exp \{ -\xi_i \|\sigma_i\|_1 \} \quad (16)$$

where ξ_i is the scale parameter of Laplace distribution and $\xi_i > 0$. Then, the probability distribution of the vector $\sigma_{M \times 1}$ is:

$$p_{\sigma}(\sigma) = \prod_{i=1}^M \left\{ \left(\frac{\xi_i}{2} \right) \exp \{ -\xi_i \|\sigma_i\|_1 \} \right\} = \prod_{i=1}^M \left\{ \left(\frac{\xi_i}{2} \right) \right\} \exp \left\{ -\sum_{i=1}^M \xi_i \|\sigma_i\|_1 \right\} \quad (17)$$

Based on the Bayesian rule in the information theory, the maximum posterior (MAP) probability of the vector $\sigma_{M \times 1}$ is:

$$\hat{\sigma} = \operatorname{argmax} [p_{\sigma/\mathbf{s}}(\sigma/\mathbf{s})] = \operatorname{argmax} [p_{\mathbf{s}/\sigma}(\mathbf{s}/\sigma) \cdot p_{\sigma}(\sigma)] \quad (18)$$

We took the logarithm of the above Formula (18):

$$\begin{aligned}\hat{\sigma} &= \operatorname{argmax}_{\sigma} [\log[p_{s/\sigma}(s/\sigma)] + \log[p_{\sigma}(\sigma)]] \\ &= \operatorname{argmax} \left[-\frac{1}{2\sigma_n^2} \|s - D\sigma\|_2^2 - \|\xi \cdot \sigma\|_1 \right] \\ &= \operatorname{argmin} \left[\|s - D\sigma\|_2^2 + \beta \|\xi \cdot \sigma\|_1 \right]\end{aligned}\quad (19)$$

where the reweighting matrix ξ is the diagonal matrix with $\{\xi_1, \xi_2, \dots, \xi_M\}$ on the diagonal and zeros elsewhere. Similarly to α in Equation (4), β is the regularization parameter. After the above deduction, the regularization model (19) is the modification to the \mathcal{L}_1 -norm penalization rule.

Additionally, the matrix ξ is first calculated to solve Equation (19). The logarithm function of the vector $\sigma_{M \times 1}$ is:

$$\begin{aligned}\log\{p_{\sigma}(\sigma)\} &= \log \left\{ \prod_{i=1}^M \left\{ \left(\frac{\xi_i}{2} \right) \exp\{-\xi_i |\sigma_i|_1\} \right\} \right\} \\ &= \log \left\{ \prod_{i=1}^M \left(\frac{\xi_i}{2} \right) \right\} + \log \left\{ \exp \left\{ -\sum_{i=1}^M \xi_i |\sigma_i|_1 \right\} \right\} = \sum_{i=1}^M \log \left(\frac{\xi_i}{2} \right) - \sum_{i=1}^M \xi_i |\sigma_i|_1\end{aligned}\quad (20)$$

Let the partial derivative function of Formula (10) with respect to ξ_i be equal to zero, and the estimation of the scale parameter ξ_i is:

$$\frac{\partial \log[p_{\sigma}(\sigma)]}{\partial \xi_i} = \frac{1}{\xi_i} - |\sigma_i|_1 = 0 \quad (21)$$

$$\xi_i = \frac{1}{|\sigma_i|} \quad (22)$$

In the case that $\sigma_i = 0$, Formula (12) makes no sense. Formula (12) is modified as:

$$\xi_i = \frac{1}{|\sigma_i| + \iota} \quad (23)$$

where $\iota > 0$ is a very small positive constant.

The above deduction explains the prior distribution, e.g., Laplace distribution in the sub-Nyquist SAR confirms the penalization rule and this method makes full use of the data prior information to achieve a good performance. Certainly, different probability distributions confirm different penalization rules. In the scenario of not knowing the variables themselves, using the iteratively updated method to establish an approximately fair penalization rule allows for the successively better estimation of nonzero variables. According to the optimization Equations (19) and (23), we know that the regularized item is iteratively penalized by itself to nearly acquire the number of nonzero values, and the large coefficients are more heavily penalized to discourage their effects compared to that with a small coefficient and is more likely to be identified as nonzero. Once the nonzero locations are identified, their influence attenuates to allow more sensitivity to identify the remaining small but nonzero elements. It means this algorithm can more accurately recover low SNR targets.

To solve the pseudo- \mathcal{L}_0 -norm regularization equation, the algorithm in Equation (19) is essentially a \mathcal{L}_1 penalization method so that it is convex and can be solved by the Cauchy–Newton method [36,37]. The solving method is as follows: \mathcal{L}_1 norm is not differentiable, while $\sigma_i = 0$. Firstly, the smoothing approximation is introduced:

$$\|\sigma\|_1 = \sum_{i=1}^N \sqrt{(|\sigma_i|^2 + \zeta)} \quad (24)$$

where $\varsigma > 0$ is a very small positive constant. Equation (19) can be expressed as:

$$\hat{\sigma} = \underset{\sigma}{\operatorname{argmin}} \left\{ \|\mathbf{s} - \mathbf{D}\sigma\|_2^2 + \beta \sum_{i=1}^M \xi_i \sqrt{(|\sigma_i|^2 + \varsigma)} \right\} \quad (25)$$

where $f(\hat{\sigma}) = \|\mathbf{s} - \mathbf{D}\sigma\|_2^2 + \beta \sum_{i=1}^M \xi_i \sqrt{(|\sigma_i|^2 + \varsigma)}$. The conjugate gradient function is written as:

$$\nabla_{\hat{\sigma}} f(\hat{\sigma}) = 2\mathbf{D}^H \mathbf{D}\sigma + \beta \mathbf{U}(\hat{\sigma}) \cdot \xi \sigma - 2\mathbf{D}^H \mathbf{s} = H(\hat{\sigma})\sigma - 2\mathbf{D}^H \mathbf{s} \quad (26)$$

where $H(\hat{\sigma}) = 2\mathbf{D}^H \mathbf{D} + \beta \mathbf{U}(\hat{\sigma}) \xi$, $\mathbf{U}(\hat{\sigma}) = \operatorname{diag} \left\{ 1 / \left(\sqrt{|\sigma_i|^2 + \varsigma} \right) \right\}$, $i = 1, 2, \dots, M$. \mathbf{D}^H is the conjugate transpose of the observed matrix \mathbf{D} .

According to the Newton method [36,37], the iterative equation of the reconstructed result σ is:

$$H(\hat{\sigma}^{(g)}) \hat{\sigma}^{(g+1)} = (1 - \gamma) H(\hat{\sigma}^{(g)}) \hat{\sigma}^{(g)} + 2\gamma \mathbf{D}^H \mathbf{s} \quad (27)$$

where γ is the step size of the iteration. When $\gamma = 1$, Equation (27) is simplified as:

$$\hat{\sigma}^{(g+1)} = 2\mathbf{H}(\hat{\sigma}^{(g)})^{-1} \mathbf{D}^H \mathbf{s} \quad (28)$$

Until the iteration terminates, $\hat{\sigma}$ is the recovered result. Normally, the MSE, the iterative number, etc., can be taken as the iterative criterion [23]. After the above deduction, the flow chart is concluded as follows:

- (1) Initialization: the iterative step $g = 1$, $\hat{\sigma}^1 = \mathbf{D}^H \mathbf{s}$;
- (2) Updating of the weighting matrix ξ and the matrix $\mathbf{H}(\hat{\sigma})$: $\xi^{(g)} = \operatorname{diag} \left(\frac{1}{|\sigma_1| + \iota}, \frac{1}{|\sigma_2| + \iota}, \dots, \frac{1}{|\sigma_M| + \iota} \right)$, $H(\hat{\sigma})^{(g)} = 2\mathbf{D}^H \mathbf{D} + \beta \mathbf{U}(\hat{\sigma}) \xi^{(g)}$;
- (3) Calculation: $\hat{\sigma}^{(g+1)} = 2 \left(\mathbf{H}(\hat{\sigma})^{(g)} \right)^{-1} \mathbf{D}^H \mathbf{s}$;
- (4) $g = g + 1$;
- (5) Loop;
- (6) Stopping iteration according to the iterative criterion.

During the deduction of the pseudo- \mathcal{L}_0 -norm optimization algorithm, there are four undetermined parameters, i.e., the regularization factor β , the parameter ι , the parameter ς , and the iterative criterion. Similarly to the regularization α , β is also an empirical value and is chosen by minimizing the recovered error [12]. In the simulation section, the iterative criterion selects the preset iterative number and this number is chosen empirically. Although the reference [23] demonstrates how to choose the parameter ι and ς , the pseudo- \mathcal{L}_0 -norm optimization algorithm has a good robustness and the reconstructed scene does not have the strictly sparse property; so, the proposed algorithm sets the parameters ι and ς as 10^{-3} and 10^{-6} , respectively.

2.2.3. Error Correction Based on the Pseudo- \mathcal{L}_0 -Norm Optimization Algorithm

Based on the observation model of Equation (13), we propose an error correction method integrated with the pseudo- \mathcal{L}_0 -norm optimization algorithm. This method considers the phase error as the model error and removes it during the scene reconstruction, and it enables our method to correct more artifacts due to the robustness of the pseudo- \mathcal{L}_0 -norm optimization algorithm. This error correction method is an update and iteration algorithm, and it includes the scene reconstruction and the error estimation during each iteration. In the first step of every iteration, the cost function is minimized with the scene based on the reweighted \mathcal{L}_1 -norm optimization algorithm, and in the second step, the phase error is estimated given the scene estimate. According to the estimated error, the error matrix $\mathbf{E}_{N \times N}$ in (13) is updated and delivered to the next iteration. The procedure is as follows:

- (1) The scene reconstruction

Based on the pseudo- \mathcal{L}_0 -norm optimization algorithm, the optimization equation of (13) without subscripts is:

$$\hat{\sigma} = \underset{\sigma}{\operatorname{argmin}} \left[\|\mathbf{s}^{error} - \mathbf{E}\mathbf{D}\sigma\|_2^2 + \beta \|\xi\sigma\|_1 \right] \quad (29)$$

The flow chart in Section 2.2.2 can resolve this optimization equation.

(2) The error estimation

To estimate the phase error, the cost function should minimize the *MSE* of the reconstructed result since the error exists at each sampling moment, and each element can be handled separately. The cost function at each sampling moment is:

$$\hat{\mathbf{E}}_n(g+1) = \underset{\mathbf{E}_n}{\operatorname{argmin}} \left\| \mathbf{s}_n^{error} - \mathbf{E}_n \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right\|_2^2, n \in \{1, 2, \dots, N\} \quad (30)$$

where $\hat{\mathbf{E}}_n(g+1)$ is the n -th diagonal element of the error matrix $\hat{\mathbf{E}}^{(g+1)}$ at the $(g+1)$ -th iteration. In the following, the cost function is unfolded for analysis:

$$\begin{aligned} \left\| \mathbf{s}_n^{error} - \mathbf{E}_n \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right\|_2^2 &= \left(\mathbf{s}_n^{error} - \mathbf{E}_n \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right)^H \cdot \left(\mathbf{s}_n^{error} - \mathbf{E}_n \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right) \\ &= (\mathbf{s}_n^{error})^H \cdot \mathbf{s}_n^{error} + \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n^H \cdot \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n - 2 \cos(\vartheta_n) \cdot \operatorname{Re} \left\{ (\mathbf{s}_n^{error})^H \cdot \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right\} + 2 \sin(\vartheta_n) \cdot \operatorname{Im} \left\{ (\mathbf{s}_n^{error})^H \cdot \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right\} \end{aligned} \quad (31)$$

where $\operatorname{Re}\{\cdot\}$ and $\operatorname{Im}\{\cdot\}$ are the real part and the imaginary part of a complex signal, respectively. Assuming that $\rho = \operatorname{Re} \left\{ (\mathbf{s}_n^{error})^H \cdot \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right\}$ and $v = \operatorname{Im} \left\{ (\mathbf{s}_n^{error})^H \cdot \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right\}$, the above equation can be written as:

$$\left\| \mathbf{s}_n^{error} - \mathbf{E}_n \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right\|_2^2 = (\mathbf{s}_n^{error})^H \cdot \mathbf{s}_n^{error} + \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n^H \cdot \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n - 2\sqrt{\rho^2 + v^2} \cdot \cos\left(\vartheta_n - \arctan\left(\frac{v}{\rho}\right)\right) \quad (32)$$

Estimating ϑ_n is equal to minimizing the above cost function (32); so, $\hat{\vartheta}_n$ satisfies:

$$\hat{\vartheta}_n = \angle \left((\mathbf{s}_n^{error})^H \cdot \left(\mathbf{D}^{(g)} \sigma^{(g+1)} \right)_n \right) \quad (33)$$

After the above deduction, the detailed flow chart is as follows:

- (1) Initialization: $g = 0, \mathbf{E} = \mathbf{I}$;
- (2) Recovering the scene: $\hat{\sigma}^{(g)} = \underset{\sigma}{\operatorname{argmin}} \left\| \mathbf{s}^{error} - \mathbf{E}^{(g)} \mathbf{D} \sigma^{(g)} \right\|_2 + \beta \left\| \xi^{(g)} \sigma^{(g)} \right\|_1$;
- (3) Estimating the error matrix: $\hat{\mathbf{E}}^{(g)} = \underset{\mathbf{E}}{\operatorname{argmin}} \left\| \mathbf{s}^{error} - \mathbf{E}^{(g)} \mathbf{D} \sigma^{(g)} \right\|_2$;
- (4) Updating $\mathbf{E}^{(g)}$;
- (5) $g = g + 1$;
- (6) Loop;
- (7) Stopping iteration according to the iterative criterion.

where $\hat{\sigma}^{(g)}$ and $\hat{\mathbf{E}}^{(g)}$ denote the recovered result and the estimated error matrix in the g -th iteration, respectively. During each iteration, the cost function focuses part of the defocused scene to generate a relatively accurate error matrix so that a relatively accurate model can lead to a better reconstruction result. The iterative criterion was mentioned in Section 2.2.2 and is not repeated in this section.

2.2.4. Analysis of the Computational Complexity

In order to obtain the total computational complexity of the proposed method, we first analyzed the time complexity per iteration quantitatively. As the notations used above, the total computational cost of the proposed method is in the order of $o(G \cdot NM \log(NM))$, where G is the required number of iterations to recover the result in one range cell. The value of G is difficult to confirm accurately through theoretical analysis, but in practice,

the running time of the proposed method is affordable. As for the memory cost of the proposed method, we only need to store the input, the output, and the parameter matrices. In summary, the memory cost is in the order of $o(NM)$.

3. Experiment

In this section, simulation experiments and data experiments verify the effectiveness and superiority of sub-Nyquist SAR imaging and error correction integrated with a pseudo- \mathcal{L}_0 -norm optimization algorithm.

3.1. Data Description

To verify the validity and effectiveness of sub-Nyquist SAR imaging and error correction based on the pseudo- \mathcal{L}_0 -norm optimization algorithm, we selected real strip-map TerraSAR-X images for the experiments. The reflectivity functions of images were used to simulate raw data according to the simulated parameters, and then the raw data were uniformly received, as shown in Figure 1. During the error correction, the error is added at each sampling moment and the error adopts the random distribution.

Figure 3 features a sea–land interface scene. This scene is more complex and used to verify the performance of our proposed algorithm in the sub-Nyquist SAR imaging compared with the \mathcal{L}_1 -norm optimization algorithm. Figure 4 is a sea containing several boats to verify the availability of error corrections based on the pseudo- \mathcal{L}_0 -norm optimization algorithm.

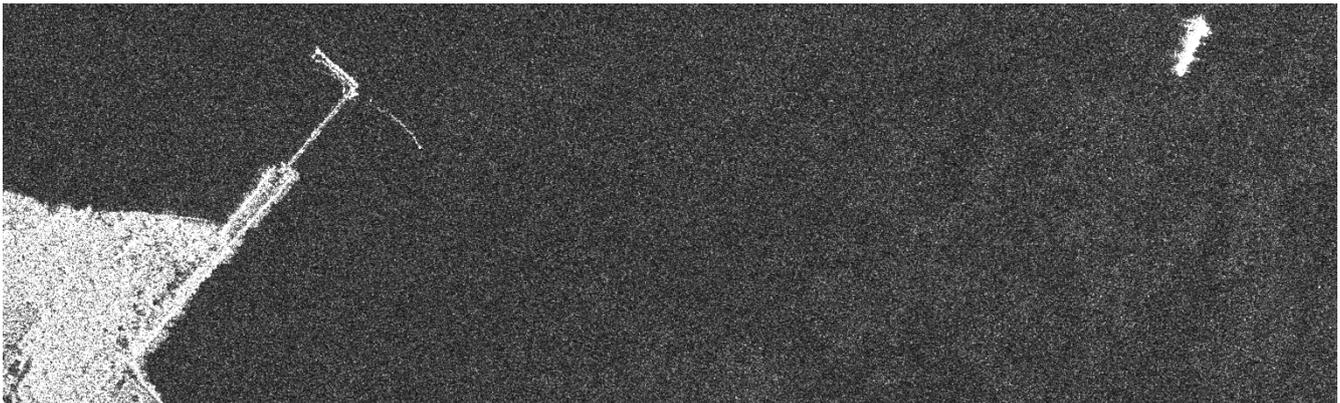


Figure 3. Sea–land interface scene in the SAR image.

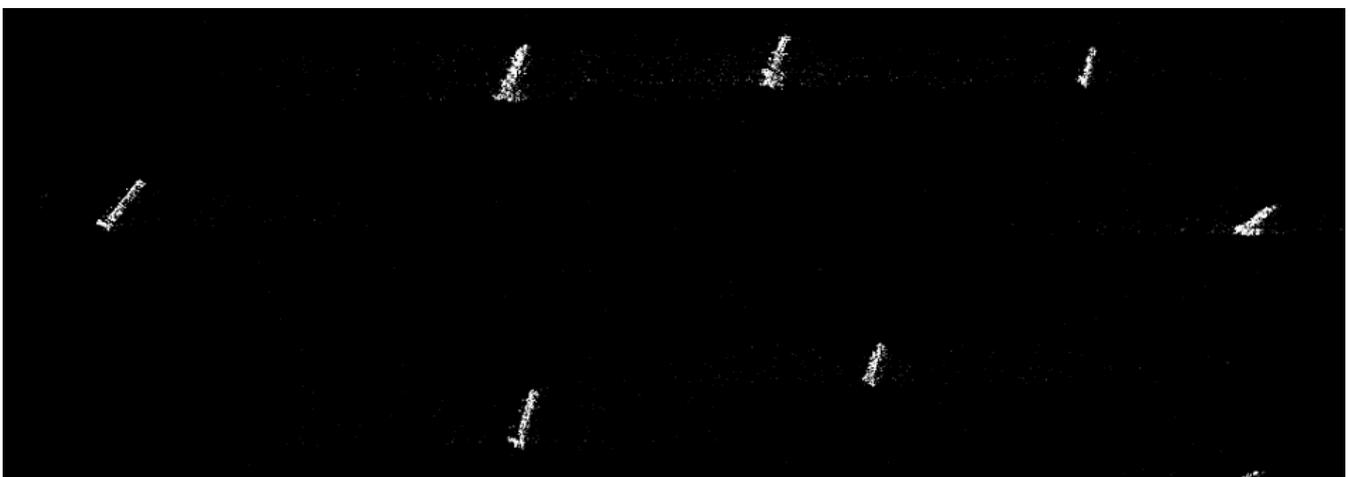


Figure 4. Sea containing several boats in the SAR image.

3.2. The Simulation of the Pseudo- \mathcal{L}_0 -Norm Optimization Algorithm

We used a one-dimensional simulation of point targets to verify the superiority of our proposed algorithm compared with the \mathcal{L}_1 -norm optimization algorithm. For a given under-recovered signal, σ , the recovered performance can be evaluated by the normalized MSE (NMSE):

$$NMSE = \frac{\|\hat{\sigma} - \sigma\|_2^2}{\|\sigma\|_2^2} \quad (34)$$

where $\hat{\sigma}$ is the recovered result.

The following one-dimensional simulation illustrates that the pseudo- \mathcal{L}_0 -norm algorithm has a better performance compared with the \mathcal{L}_1 -norm optimization algorithm. The simulated parameters are as follows. It selects a sparse signal, σ , of length $M = 256$, with $\|\sigma\|_0 = S$. The S nonzero spike positions are selected randomly, and the amplitude of nonzero elements obeys a zero-mean unit-variance Gaussian distribution. This simulation selects the measured number and a $N \times M$ random matrix, \mathbf{D} , with independent identically distributed (i.i.d.) Gaussian elements. The noise vector, \mathbf{n} , is drawn from the i.i.d. zero-mean Gaussian function with $\|\mathbf{n}\|_2 = 0.5$, so that $SNR = 10 \log_{10}(\|\sigma\|_2^2 / \|\mathbf{n}\|_2^2) = 21 \text{ dB}$. To recover the signal, σ , it adopts two algorithms, the \mathcal{L}_1 -norm optimization algorithm and pseudo- \mathcal{L}_0 -norm algorithm, to compare the recovered performance. The NMSE of the two algorithms and the recovered results are compared in Figures 5 and 6.

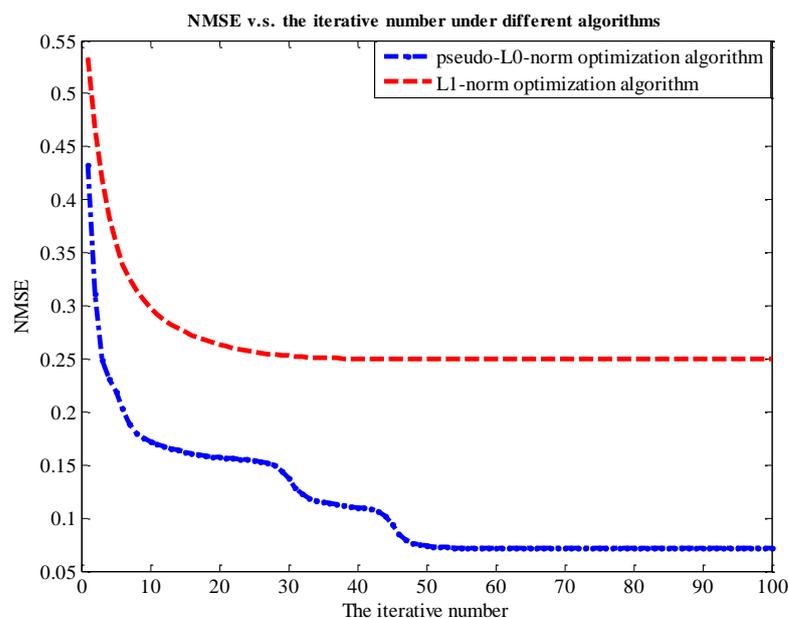


Figure 5. NMSE vs. the iterative number under different algorithms.

Figure 5 demonstrates that the pseudo- \mathcal{L}_0 -norm optimization algorithm can achieve the smallest recovered error, although this proposed algorithm has a slower rate of convergence. Our proposed method is sensitive to low SNR targets that are not easy to identify, so the convergence rate is slow. Figure 6 shows the recovered results under different algorithms. In the red circle, ①, the results for the \mathcal{L}_1 -norm optimization algorithm have weak false targets, but two other algorithms do not. In circles ② and ④, the weak target can be recovered under the optimization algorithm, while the \mathcal{L}_1 -norm optimization algorithm cannot recover it. In circle ③, the amplitude of targets can be exactly recovered by the proposed algorithm. It demonstrates that our proposed algorithm is more friendly to the low SNR target and removes false targets to achieve a good recovered performance.

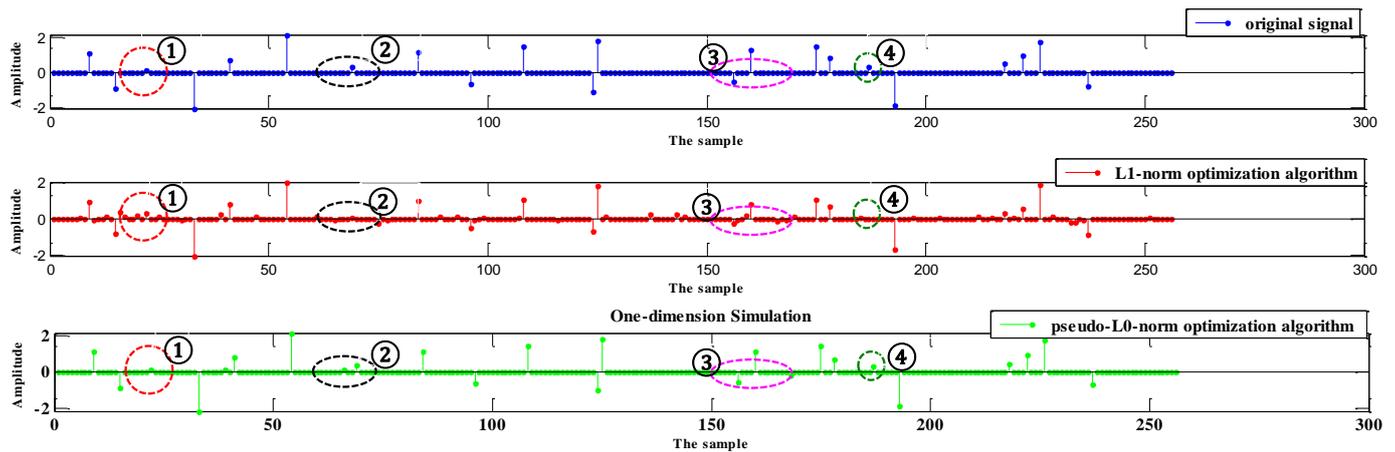


Figure 6. The recovered result under different algorithms.

3.3. The Simulation of Sub-Nyquist SAR Imaging

We used a simulation based on the real strip-map TerraSAR-X image in Figure 3 to verify the validity of our proposed algorithm for sub-Nyquist SAR imaging. The SAR image is over the sea–land interface scene. The simulated parameters are shown in Table 1, and the raw data were randomly received, as shown in Figure 1. In the simulation, the interfaced land is recovered without the loss of details under the proposed algorithm. According to the simulated result in Figure 7 and MSE in Table 2, this simulation also demonstrates that the pseudo- \mathcal{L}_0 -norm optimization algorithm can achieve a better recovered performance compared with the \mathcal{L}_1 -norm optimization algorithm.

Table 1. Simulated parameters.

Parameter	Data
Average PRF (Hz)	893
Range sampling frequency (MHz)	55
Referred slant range (km)	870
Chirp rate (Hz/s)	10^{12}
Doppler bandwidth (Hz)	2438
Wavelength (mm)	5.55
Velocity (m/s)	7513
Height (km)	693
Squint angle ($^{\circ}$)	0



(a) The reconstructed result based on the pseudo- \mathcal{L}_0 -norm optimization algorithm

Figure 7. Cont.



(b) The reconstructed result based on the \mathcal{L}_1 -norm optimization algorithm

Figure 7. The recovered result based on the pseudo- \mathcal{L}_0 -norm optimization algorithm and \mathcal{L}_1 -norm optimization algorithm.

Table 2. MSE under different algorithms.

	pseudo- \mathcal{L}_0 -norm optimization algorithm	\mathcal{L}_1 -norm optimization algorithm
MSE	0.114	0.874

3.4. Simulation of the Error Correction

The error correction mainly solves image defocusing for the phase error. As we all know, the better the focusing performance is, the smaller the image entropy is. We adopted image entropy $H(\sigma)$ to evaluate the reconstruction result.

$$H(\sigma) = -\sum_i \frac{|\sigma_i|^2}{\sum_i |\sigma_i|^2} \log_2 \frac{|\sigma_i|^2}{\sum_i |\sigma_i|^2} \quad (35)$$

Firstly, we simulated point targets to illustrate the effectiveness of our method. The simulated parameters are shown in Table 3. When the elements of the error matrix are uniformly distributed $[0, 17/18\pi]$, the simulated results are presented in Figure 8. Figure 8a is the original image, (b) is the reconstructed scene without an error correction in the traditional SAR, (c) is the reconstructed scene with an error correction in the sub-Nyquist SAR based on the \mathcal{L}_1 -norm optimization algorithm, and (d) is the reconstructed scene with the error correction in the sub-Nyquist SAR based on the pseudo- \mathcal{L}_0 -norm optimization algorithm. When the phase error exists, the reconstructed scene without an error correction is defocused (Figure 8b). Although it can also remove errors and focus the scene based on the \mathcal{L}_1 -norm optimization algorithm, there are false targets and lost targets in the red circle (Figure 8c). Our method almost removes errors and recovers the scene based on the pseudo- \mathcal{L}_0 -norm optimization algorithm (Figure 8d). From the quantitative analysis, Table 4 illustrates that the image entropy based on our method achieves the nearly same image entropy as the original image.

The simulation in Figure 9 is based on a strip-map TerraSAR-X image, i.e., sea containing several boats in Figure 4. The reflectivity function of the image was used to simulate raw data, and the simulated parameters are shown in Table 3. The elements of the error matrix are uniformly distributed, $[0, \pi/2]$. Figure 9a is the recovered result in the traditional SAR, but is defocused for the phase error. The error is partly removed from the Sub-Nyquist SAR with an error correction method based on the \mathcal{L}_1 -norm optimization algorithm in Figure 9b. Based on the pseudo- \mathcal{L}_0 -norm optimization algorithm, it almost removes the error and focuses the scene in Figure 9c. The image entropy in Table 5 illustrates that the image recovered by our method can achieve a similar performance to the original image.

Table 3. Simulated parameters.

Parameter	Data
Average PRF (Hz) in the sub-Nyquist SAR	155
PRF (Hz) in the traditional SAR	1907
Range sampling frequency (MHz)	120
Referred slant range (km)	888
Pulse width (us)	50
Doppler bandwidth (Hz)	1401
Wavelength (mm)	5.55
Velocity (m/s)	7513
Height (km)	693
Squint angle (°)	0

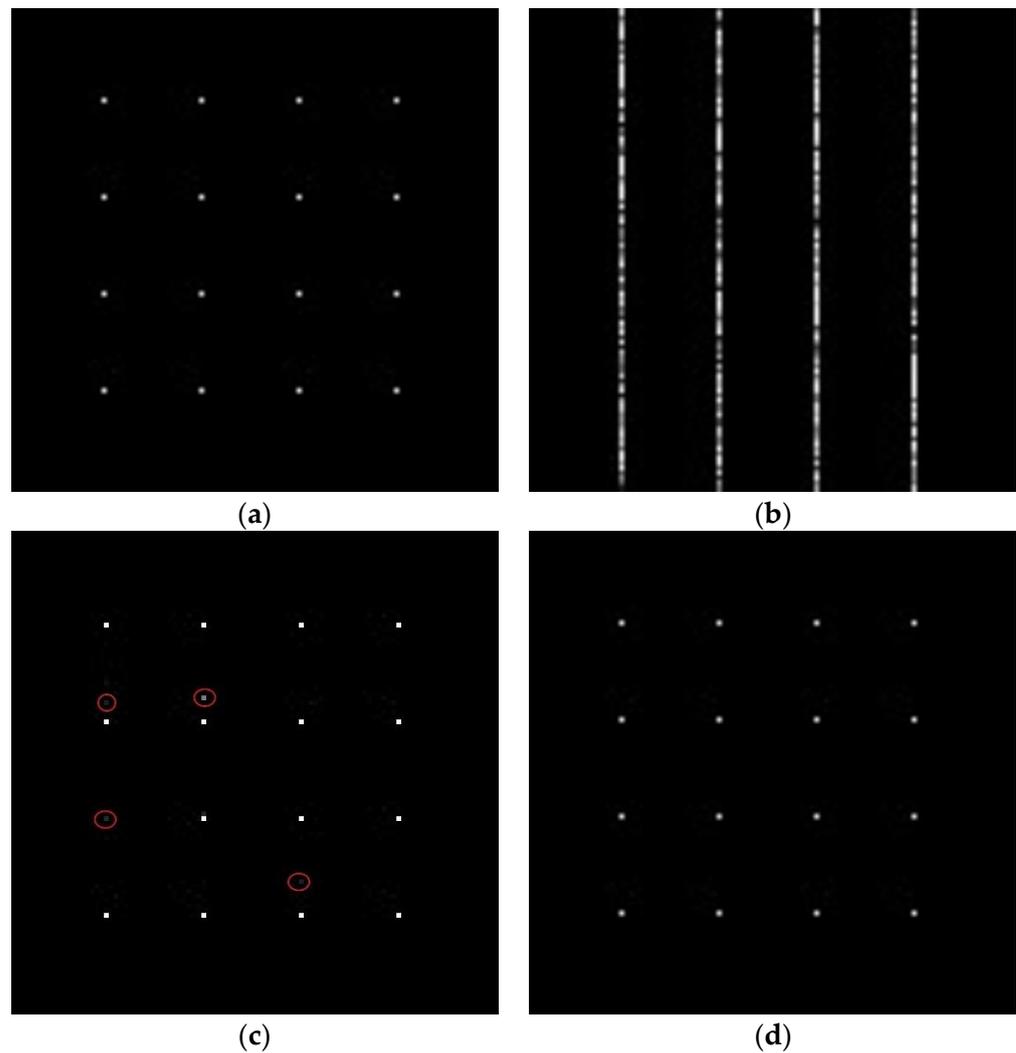


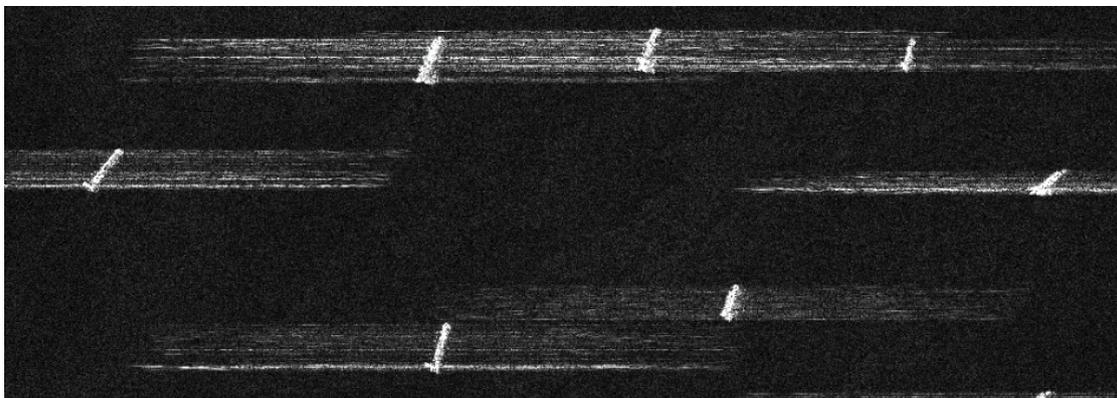
Figure 8. The reconstructed results. (a) The original image; (b) the reconstructed scene without error correction; (c) the reconstructed scene with error correction based on the \mathcal{L}_1 -norm optimization algorithm; and (d) the reconstructed scene with error correction based on the pseudo- \mathcal{L}_0 -norm optimization algorithm.

Table 4. Image entropy under different algorithms.

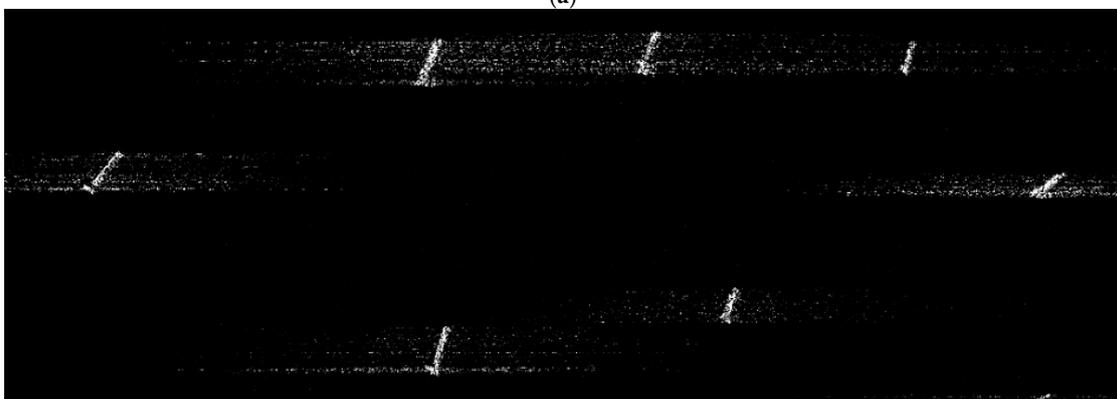
	The original image	The reconstructed scene without an error correction	The reconstructed scene with an error correction based on the \mathcal{L}_1 -norm optimization algorithm	The reconstructed scene with an error correction based on the pseudo- \mathcal{L}_0 -norm optimization algorithm
Image entropy (bit)	4.00	5.90	3.08	3.99

Table 5. Image entropy under different algorithms.

	The original image	The reconstructed scene without an error correction	The reconstructed scene with an error correction based on the \mathcal{L}_1 -norm optimization algorithm	The reconstructed scene with an error correction based on the pseudo- \mathcal{L}_0 -norm optimization algorithm
Image entropy (bit)	7.81	12.16	8.35	7.80

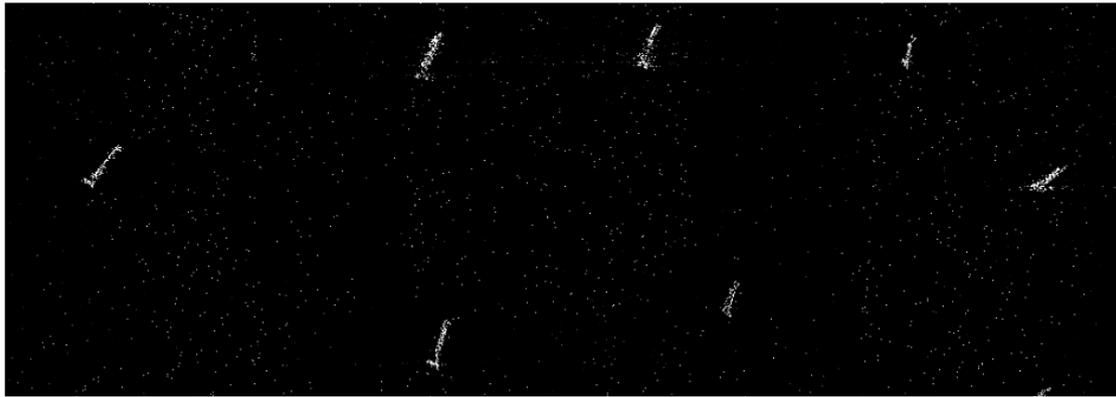


(a)



(b)

Figure 9. Cont.



(c)

Figure 9. The reconstructed result. (a) The scene reconstruction without error correction; (b) the scene reconstruction with error correction based on the \mathcal{L}_1 -norm optimization algorithm; and (c) the scene reconstruction with error correction based on the pseudo- \mathcal{L}_0 -norm optimization algorithm.

4. Discussion

HRWS SAR imaging has always been the goal of spaceborne SAR systems in remote sensing applications [2]. Since a high resolution and wide swath are inherently conflicting requirements, the HRWS SAR system creates new difficulties and challenges. These requirements are simultaneously satisfied by advanced observing modes, e.g., azimuthal multi-channel SAR [3] and MIMO SAR [4], and these observing modes both have the characteristics of large amounts of data and long antenna. As the CS theorem develops, a novel imaging mode named sub-Nyquist SAR based on the pseudo-random space–time modulation has been proposed without large data and long antenna, and sub-Nyquist SAR imaging based on the CS algorithms can recover the scene [8].

Although the \mathcal{L}_1 -norm optimization algorithm has achieved a quite good performance, some low SNR targets are not accurately recovered and it yields false targets. To further improve the recovered performance of sub-Nyquist SAR, we present a pseudo- \mathcal{L}_0 -norm optimization algorithm based on the Bayesian estimation. This algorithm penalizes the regularization item with the reciprocal of its previous solution to acquire nonzero values so that the rule is fair and makes full use of the data prior information. This proposed algorithm essentially adopts the idea of the adaptive Lasso technique [20,21], and the \mathcal{L}_1 -norm optimization algorithm is the Lasso variable selection method [13]. The Lasso variable selection is consistent with satisfying a necessary condition; sometimes Lasso is not consistent and does not have predictive properties. Adaptive Lasso is the modified version of the Lasso system with important differences. The modification is data-dependent and reasonably selected so that adaptive Lasso has predictive properties [20,21]. Our proposed algorithm is better in terms of consistency. For the recovered error, we deduced the expression of MSE in the following way. Assume that $\mathbf{F}_{N \times S}$ is the submatrix constructed by taking S columns from the recovered matrix, $\mathbf{D}_{N \times M}$, which are specified by the index vector, Λ , and each element in Λ satisfies the following:

$$\sigma_{\Lambda_l} \neq 0, (l = 1, 2, \dots, S) \quad (36)$$

The estimated covariance matrix for the nonzero components, σ_l , is

$$\mathbf{C} = \left(\mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} + \gamma_2 \Sigma(\sigma_{\Lambda_l}) \right)^{-1} \mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} \left(\mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} + \gamma_2 \Sigma(\sigma_{\Lambda_l}) \right)^{-1} \quad (37)$$

where $\Sigma(\sigma_{\Lambda_l}) = \text{diag}(\xi_{\Lambda_1}/\sigma_{\Lambda_1}, \xi_{\Lambda_2}/\sigma_{\Lambda_2}, \dots, \xi_{\Lambda_S}/\sigma_{\Lambda_S})$ is a diagonal matrix. The estimated error is the trace of the covariance matrix, \mathbf{C} :

$$\begin{aligned}
& \|\hat{\boldsymbol{\sigma}}_{\Lambda} - \boldsymbol{\sigma}_{\Lambda}\|_2^2 = \sigma_n^2 \text{trace}(\mathbf{C}) \\
& = \sigma_n^2 \text{trace} \left((\mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} + \gamma_2 \boldsymbol{\Sigma}(\boldsymbol{\sigma}_{\Lambda_l}))^{-1} \mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} (\mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} + \gamma_2 \boldsymbol{\Sigma}(\boldsymbol{\sigma}_{\Lambda_l}))^{-1} \right) \\
& = \sigma_n^2 \text{trace} \left(\left((\mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} + \gamma_2 \boldsymbol{\Sigma}(\boldsymbol{\sigma}_{\Lambda_l}))^{-1} \right)^2 \mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} \right) \\
& \geq \sigma_n^2 \left[2 \text{trace} \left((\mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} + \gamma_2 \boldsymbol{\Sigma}(\boldsymbol{\sigma}_{\Lambda_l}))^{-1} \right) - \text{trace} \left((\mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S})^{-1} \right) \right]
\end{aligned} \tag{38}$$

Assuming that $\mathbf{A} = \mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S}$, $\mathbf{B} = \mathbf{F}_{N \times S}^H \mathbf{F}_{N \times S} + \gamma_2 \boldsymbol{\Sigma}(\boldsymbol{\sigma}_{\Lambda_l})$, and their eigenvalues are $\lambda_1, \lambda_2, \dots, \lambda_S$ and $\alpha_1, \alpha_2, \dots, \alpha_S$, respectively. Based on the structure form of the recovered matrix, $\mathbf{D}_{N \times M}$, the \mathcal{L}_2 -norm of each column is nearly equal and assumes $\|\mathbf{D}_l\|_2^2 \approx L$, where L is the number of samples in one aperture time [8]. Formula (38) is simplified as:

$$\|\hat{\boldsymbol{\sigma}}_{\Lambda} - \boldsymbol{\sigma}_{\Lambda}\|_2^2 \geq \sigma_n^2 \left[2 \text{trace}(\mathbf{B}^{-1}) - \text{trace}(\mathbf{A}^{-1}) \right] \tag{39}$$

Gail's circle theorem [38] indicates that

$$\begin{aligned}
& |\lambda_l - L| \leq (S-1) \cdot u \cdot L \\
& (1 - (S-1) \cdot u) \cdot L \leq \lambda_l \leq (1 + (S-1) \cdot u) \cdot L \\
& \frac{1}{(1+(S-1) \cdot u) \cdot L} \leq \lambda_l^{-1} \leq \frac{1}{(1-(S-1) \cdot u) \cdot L} \\
& \frac{S}{(1+(S-1) \cdot u) \cdot L} \leq \sum_{l=1}^S \lambda_l^{-1} \leq \frac{S}{(1-(S-1) \cdot u) \cdot L} \\
& \frac{S}{(1+(S-1) \cdot u) \cdot L} \leq \text{trace}(\mathbf{A}^{-1}) \leq \frac{S}{(1-(S-1) \cdot u) \cdot L}
\end{aligned} \tag{40}$$

$$\begin{aligned}
& \left| \alpha_l - \left(L + \frac{\gamma_2 \cdot \xi_{\Lambda_l}}{\sigma_{\Lambda_l}} \right) \right| \leq (S-1) \cdot u \cdot L \\
& \frac{\sigma_{\Lambda_l} \cdot [1 - (S-1) \cdot u] \cdot L + \gamma_2 \cdot \xi_{\Lambda_l}}{\sigma_{\Lambda_l}} \leq \alpha_l \leq \frac{\sigma_{\Lambda_l} \cdot [1 + (S-1) \cdot u] \cdot L + \gamma_2 \cdot \xi_{\Lambda_l}}{\sigma_{\Lambda_l}} \\
& \frac{\sigma_{\Lambda_l}}{\sigma_{\Lambda_l} \cdot [1 + (S-1) \cdot u] \cdot L + \gamma_2 \cdot \xi_{\Lambda_l}} \leq \alpha_l^{-1} \leq \frac{\sigma_{\Lambda_l}}{\sigma_{\Lambda_l} \cdot [1 - (S-1) \cdot u] \cdot L + \gamma_2 \cdot \xi_{\Lambda_l}} \\
& \frac{S}{[1 + (S-1) \cdot u] \cdot L + \gamma_2 \cdot \max(\xi_{\Lambda_l} / \sigma_{\Lambda_l})} \leq \sum_{l=1}^S \alpha_l^{-1} \leq \frac{S}{[1 - (S-1) \cdot u] \cdot L + \gamma_2 \cdot \min(\xi_{\Lambda_l} / \sigma_{\Lambda_l})} \\
& \frac{S}{[1 + (S-1) \cdot u] \cdot L + \gamma_2 \cdot \max(\xi_{\Lambda_l} / \sigma_{\Lambda_l})} \leq \text{trace}(\mathbf{B}^{-1}) \leq \frac{S}{[1 - (S-1) \cdot u] \cdot L + \gamma_2 \cdot \min(\xi_{\Lambda_l} / \sigma_{\Lambda_l})}
\end{aligned} \tag{41}$$

where $u = \max_{1 \leq m_1 \neq m_2 \leq M} |\langle \mathbf{D}_{m_1}, \mathbf{D}_{m_2} \rangle| / (\|\mathbf{D}_{m_1}\|_2 \cdot \|\mathbf{D}_{m_2}\|_2)$ is the mutual coherence coefficient that reflects the maximum similarity between any two different columns, m_1, m_2 , in the recovered matrix, $\mathbf{D}_{N \times M}$. Combine Formulas (40) and (41):

$$\frac{2S}{[1 + (S-1) \cdot u] \cdot L + \gamma_2 \cdot \max(\xi_{\Lambda_l} / \sigma_{\Lambda_l})} - \frac{S}{(1 - (S-1) \cdot u) \cdot L} \leq 2 \text{trace}(\mathbf{B}^{-1}) - \text{trace}(\mathbf{A}^{-1}) \leq \frac{2S}{[1 - (S-1) \cdot u] \cdot L + \gamma_2 \cdot \min(\xi_{\Lambda_l} / \sigma_{\Lambda_l})} - \frac{S}{(1 + (S-1) \cdot u) \cdot L} \tag{42}$$

So, we have the MSE of the proposed algorithm:

$$\|\hat{\boldsymbol{\sigma}}_{\Lambda} - \boldsymbol{\sigma}_{\Lambda}\|_2^2 \geq \sigma_n^2 \left[2 \text{trace}(\mathbf{B}^{-1}) - \text{trace}(\mathbf{A}^{-1}) \right] \geq \sigma_n^2 \left[\frac{2S}{[1 + (S-1) \cdot u] \cdot L + \gamma_2 \cdot \max(\xi_{\Lambda_l} / \sigma_{\Lambda_l})} - \frac{S}{(1 - (S-1) \cdot u) \cdot L} \right] \tag{43}$$

Similarly, the MSE of the \mathcal{L}_1 -norm optimization algorithm is:

$$\|\hat{\boldsymbol{\sigma}}_{l_1} - \boldsymbol{\sigma}_{\Lambda}\|_2^2 \geq \sigma_n^2 \left[2 \text{trace}(\mathbf{B}^{-1}) - \text{trace}(\mathbf{A}^{-1}) \right] \geq \sigma_n^2 \left[\frac{2S}{[1 + (S-1) \cdot u] \cdot L + \gamma_1 \cdot \max(1 / \sigma_{\Lambda_l})} - \frac{S}{[1 - (S-1) \cdot u] \cdot L} \right] \tag{44}$$

From the above deduction, we can see the optimal performance is achieved by the two above algorithms. By comparing (43) and (44), it can be observed that the difference is in the dashed block. The pseudo- \mathcal{L}_0 -norm optimization algorithm is more sensitive to small

coefficients. For the low SNR targets, i.e., $\xi_{\Lambda_i} = 1/(\sigma_{\Lambda_i} + \iota) \gg 1$, our proposed algorithm performs better and it can more accurately recover low SNR targets. For example, in the one-dimensional simulation, our algorithm recovers the weak target, and the amplitude of the target is more accurately recovered in Figure 6.

Sub-Nyquist SAR imaging includes three steps: range compression, range cell migration correction (RCMC), and azimuth compression. After range compression and RCMC, a pseudo- \mathcal{L}_0 -norm optimization algorithm is employed to achieve the azimuth compression. Simulation experiments in Figures 5 and 6, and data experiments based on a real TerraSAR-X image in Figure 7, demonstrate that the algorithm can achieve smaller recovered errors and remove false targets compared to that of the \mathcal{L}_1 -norm optimization algorithm. Additionally, considering that CS algorithms themselves have a certain capacity to remove phase errors, an error correction method integrated with a pseudo- \mathcal{L}_0 -norm optimization algorithm eliminates the influence of phase errors and removes defocusing. Different from autofocusing technology as a post-processing method in a traditional SAR system, this method includes two steps: scene reconstruction based on a pseudo- \mathcal{L}_0 -norm optimization algorithm and error estimation by minimizing the least-square target function. These two steps are successively iterative and the recovered matrix is updated to the estimated error. Since the exact expression of the error is not fully formulated, we propose the error matrix with a random phase, whatever the expression for the error. The simulations in the Figures 8 and 9 explain the advantages of our proposed algorithm.

The potential applications of the pseudo- \mathcal{L}_0 -norm optimization algorithm can recover more low SNR targets, e.g., the sea-land interface in the above experiment, so that it is not limited to high SNR targets, e.g., the boats on the sea. This algorithm can also be applied to the traditional HRWS system, e.g., the azimuthal multi-channel SAR and MIMO SAR, to lower the amount of data to relieve the pressure on data storage.

5. Conclusions

In this paper, we propose a signal processing algorithm for an innovative single-channel HRWS system called sub-Nyquist SAR based on a pseudo-random space-time modulation. To further improve sub-Nyquist SAR imaging performance, a pseudo- \mathcal{L}_0 -norm optimization algorithm is proposed. Firstly, the Bayesian estimation explains how to take a more democratic approach with data prior information to acquire nonzero variables, so that it can more accurately recover low SNR targets and remove false targets compared with the prevalent \mathcal{L}_1 -norm optimization algorithm. Then we present an error correction method integrated with a pseudo- \mathcal{L}_0 -norm optimization algorithm to eliminate the effect of phase errors and achieve autofocusing. Finally, the simulated experiments demonstrate the effectiveness of the proposed algorithm.

Author Contributions: Methodology, W.C., L.Z., X.X., X.W. and Q.Z.; validation, W.C. and L.Z.; writing—original draft, W.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data were bought from Deutsches Zentrum für Luft- und Raumfahrt (DLR) and are available from the author Wenjiao Chen with the permission of DLR.

Acknowledgments: The authors would like to thank the editors and the reviewers for their help and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Krieger, G.; Moreira, A.; Fiedler, H.; Hajnsek, I.; Werner, M.; Younis, M.; Zink, M. TanDEM-X: A satellite formation for high-resolution SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3317–3341. [[CrossRef](#)]
2. Krieger, G.; Younis, M.; Gebert, N.; Bordoni, F.; Patyuchenko, A.; Moreira, A. Advanced concepts for high-resolution wide-swath SAR imaging. In Proceedings of the 8th European Conference on Synthetic Aperture Radar, Aachen, Germany, 7–10 June 2010.
3. Sikaneta, I.; Gierull, C.H.; Cerutti-Maori, D. Optimum signal processing for multichannel SAR: With application to high-resolution wide-swath imaging. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6095–6109. [[CrossRef](#)]
4. Krieger, G. MIMO-SAR: Opportunities and pitfalls. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2628–2645. [[CrossRef](#)]
5. Candes, E.J.; Tao, T. Decoding by linear programming. *IEEE Trans. Inf. Theory* **2005**, *51*, 4203–4215. [[CrossRef](#)]
6. Donoho, D.L.; Elad, M.; Temlyakov, V.N. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Trans. Inf. Theory* **2006**, *52*, 6–18. [[CrossRef](#)]
7. Baraniuk, R.G. Compressive sensing. *IEEE Signal Process. Mag.* **2007**, *24*, 118–121. [[CrossRef](#)]
8. Chen, W.; Li, C.; Yu, Z.; Xiao, P. Sub-Nyquist SAR Based on Pseudo-Random Time-Space Modulation. *Sensors* **2019**, *18*, 4343. [[CrossRef](#)] [[PubMed](#)]
9. Tillmann, A.M.; Pfetsch, M.E. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Trans. Inf. Theory* **2014**, *60*, 1248–1259. [[CrossRef](#)]
10. Tropp, J.A.; Gilbert, A.C. Signal recovery from partial information via orthogonal matching pursuit. *IEEE Trans. Inf. Theory* **2007**, *53*, 4655–4666. [[CrossRef](#)]
11. Cerrone, C.; Cerulli, R.; Golden, B. Carousel greedy: A generalized greedy algorithm with applications in optimization. *Comput. Oper. Res.* **2017**, *85*, 97–112. [[CrossRef](#)]
12. Candes, E.; Tao, T. The Dantzig selector: Statistical estimation when p is much larger than n . *Ann. Stat.* **2007**, *35*, 2313–2351.
13. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. B* **1996**, *58*, 267–288. [[CrossRef](#)]
14. Ji, S.; Xue, Y.; Carin, L. Bayesian Compressive Sensing. *IEEE Trans. Signal Process.* **2008**, *56*, 2346–2356. [[CrossRef](#)]
15. Tipping, M.E.; Faul, A.C. Fast marginal likelihood maximization for sparse Bayesian models. In Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, Key West, FL, USA, 3–6 January 2006.
16. Babacan, S.D.; Molina, R.; Katsaggelos, A.K. Bayesian compressive sensing using laplace priors. *IEEE Trans. Image Process.* **2010**, *19*, 53–63. [[CrossRef](#)] [[PubMed](#)]
17. Arjouni, Y.; Kaabouch, N.; Ghazi, H.E.; Tamtaoui, A. Compressive sensing: Performance comparison of sparse recovery algorithms. In Proceedings of the 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 9–11 January 2017.
18. Joshi, S.; Siddamal, K.V.; Saroja, V.S. Performance analysis of compressive sensing reconstruction. In Proceedings of the 2nd International Conference on Electronics and Communication Systems (ICECS), Coimbatore, India, 26–27 February 2015.
19. Celik, S.; Basaran, M.; Erkucuk, S.; Cirpan, H. Comparison of compressed sensing based algorithms for sparse signal reconstruction. In Proceedings of the 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, Turkey, 16–19 May 2016.
20. Hui, Z. The adaptive lasso and its oracle properties. *J. Am. Stat. Assoc.* **2006**, *101*, 1418–1429.
21. Candès, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing Sparsity by Reweighted L1 Minimization. *J. Fourier Anal. Appl.* **2007**, *14*, 877–905. [[CrossRef](#)]
22. Cumming, I.G.; Wong, F.H. *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*; Artech House: Norwood, MA, USA, 2004.
23. Pu, W.; Wu, J.; Huang, Y.; Yang, J.; Yang, H. Fast factorized backprojection imaging algorithm integrated with motion trajectory estimation for bistatic forward-looking SAR. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3949–3965. [[CrossRef](#)]
24. Pu, W.; Wu, J.; Huang, Y.; Li, W.; Sun, Z.; Yang, J.; Yang, H. Motion errors and compensation for bistatic forward-looking SAR with cubic-order processing. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6940–6957. [[CrossRef](#)]
25. Zhang, L.; Qiao, Z.; Xing, M.; Yang, L.; Bao, Z. A robust motion compensation approach for UAV SAR imagery. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3202–3218. [[CrossRef](#)]
26. Pu, W.; Wu, J.; Huang, Y.; Du, K.; Li, W.; Yang, J.; Yang, H. A rise-dimensional modeling and estimation method for flight trajectory error in bistatic forward-looking SAR. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 5001–5015. [[CrossRef](#)]
27. Wahl, D.E.; Eichel, P.H.; Ghiglia, D.C.; Jakowatz, C.V. Phase gradient autofocus—a robust tool for high resolution SAR phase correction. *IEEE Trans. Aerosp. Electron. Syst.* **1994**, *30*, 827–835. [[CrossRef](#)]
28. Onhon, N.; Cetin, M. A Sparsity-Driven Approach for Joint SAR Imaging and Phase Error Correction. *IEEE Trans. Image Process.* **2012**, *21*, 2075–2088. [[CrossRef](#)]
29. Chen, Y.; Li, G.; Zhang, Q.; Zhang, Q.; Xia, X. Motion compensation for airborne SAR via parametric sparse representation. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 551–562. [[CrossRef](#)]
30. Kelly, S.; Yaghoobi, M.; Davies, M. Sparsity-based autofocus for undersampled synthetic aperture radar. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 972–986. [[CrossRef](#)]
31. Yang, L.; Li, P.; Zhang, S.; Zhao, L.; Zhou, S.; Xing, M. Cooperative multitask learning for sparsity-driven SAR imagery and nonsystematic error autocalibration. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 5132–5147. [[CrossRef](#)]

32. Kantor, J.M. Polar format-based compressive SAR image reconstruction with integrated autofocus. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3458–3468. [[CrossRef](#)]
33. Gungor, A.; Cetin, M.; Guven, H.E. Autofocused compressive SAR imaging based on the alternating direction method of multipliers. In Proceedings of the IEEE Radar Conference, Seattle, WA, USA, 8–12 May 2017.
34. Pu, W. Deep SAR Imaging and Motion Compensation. *IEEE Trans. Geosci. Remote Sens.* **2021**, *30*, 2232–2247. [[CrossRef](#)]
35. Seeger, M.W.; Nickisch, H. Compressed sensing and Bayesian experimental design. In Proceedings of the 25th international conference on Machine Learning, Helsinki, Finland, 5–9 July 2008.
36. Birgin, E.G.; Martinez, J.M. A spectral conjugate gradient method for unconstrained optimization. *Appl. Math. Optim* **2001**, *43*, 117–128. [[CrossRef](#)]
37. Nazareth, J.L. Conjugate gradient method. *WIREs Comp. Stat.* **2009**, *1*, 348–353. [[CrossRef](#)]
38. Golub, G.H.; Loan, C.F.V. *Matrix Computations*, 4th ed.; Johns Hopkins University Press: Baltimore, MD, USA, 2013.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.