remote sensing

MDPI

Article

Recovery-Error-Bound-Guided Frequency Selection for Moving-Target Microwave Coincidence Imaging

Shu-Yang Wan, Chen Miao *, Shi-Shan Qi, Yue Ma, Yang-Ying Zhao and Wen Wu

Key Laboratory of Near-Range RF Sensing ICs and Microsystems, Nanjing University of Science and Technology, Nanjing 210094, China; wanshuyang@njust.edu.cn (S.-Y.W.); qiss@njust.edu.cn (S.-S.Q.); mayue@njust.edu.cn (Y.M.); zhaoyy@njust.edu.cn (Y.-Y.Z.); wuwen@njust.edu.cn (W.W.)
* Correspondence: miaochen78@njust.edu.cn

Highlights

What are the main findings?

- A conditional recovery-error bound links observation-matrix sensitivity and sparse-scatterer distinguishability to scattering-coefficient recovery accuracy. The bound converts a prescribed recovery-error threshold into a state-prediction-error tolerance, enabling frequency-set selection by maximizing the minimum tolerance over prescribed observation geometries.
- Simulations show that, compared with uniform frequency sampling, the selected frequency set raises the success rate for resolving two targets separated by 1.83° from 20% to 80% at a signal-to-noise ratio (SNR) of 10 dB and an offset of 0.08 m. When used for an extended target without reoptimization, the selected frequency set improves the peak signal-to-noise ratio of the reconstructed image by 1.5 dB and reduces scatterer-localization root-mean-square error for matched detections by 71.8% at an SNR of 0 dB.

What are the implications of the main findings?

- Recovery-accuracy requirements can guide frequency-set design to improve imaging robustness to residual state-prediction errors. This provides a complementary approach to motion compensation by configuring observations to tolerate the errors that remain after compensation.
- The results indicate that moving-target imaging robustness can be improved by reconfiguring the operating frequencies within a fixed metasurface system, without increasing the frequency count or transmitted energy per frame. These findings highlight the value of frequency-set design for improving imaging and localization performance under hardware and acquisition constraints.

Check for updates

Academic Editor: Massimiliano Pieraccini

Received: 28 October 2026
Revised: 10 November 2026
Accepted: 14 November 2026
Published: 16 November 2026

Copyright: © 2026 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.

Abstract

In microwave coincidence imaging (MCI), radiation patterns with low mutual correlation, generated by an appropriately selected frequency set, provide spatial diversity for high-resolution imaging. For moving targets, however, residual state-prediction errors after motion compensation can still cause a mismatch between the echoes and the observation matrix, degrading imaging performance despite this spatial diversity. To improve imaging robustness under this mismatch, frequency-set selection is guided by a conditional upper bound on the scattering-coefficient recovery error. This bound accounts for both the sensitivity of the observation matrix to state-prediction errors and the distinguishability of sparse scatterers. The state-prediction-error tolerance is chosen such that this bound does not exceed a prescribed recovery-error threshold. The frequency set is then selected by maximizing the minimum tolerance over the prescribed observation conditions. The

resulting frequency set is used for closed-loop tracking and imaging with a frequency-diverse metasurface. During closed-loop operation, the predicted state is used to form the observation matrix, and the image centroid is fed back to update the state estimate. This update incorporates the correlation between state-prediction errors and image-derived observation errors. Simulation results show that, at a signal-to-noise ratio (SNR) of 10 dB and an x-direction offset of 0.08 m, the success rate for resolving two targets separated by 1.83° increases from 20% with uniform frequency sampling to 80% with the selected frequency set. For an extended target at 0 dB SNR, the peak signal-to-noise ratio (PSNR), evaluated against a stationary-target reference reconstruction, improves by 1.5 dB, and the scatterer-localization root-mean-square error (RMSE) for matched detections decreases by 71.8% compared with uniform frequency sampling. These results demonstrate that the proposed criterion improves resolution, image quality, and localization accuracy in the presence of state-prediction errors, without increasing the number of operating frequencies.

Keywords: microwave coincidence imaging; frequency-diverse metasurface; frequency selection; moving-target imaging; closed-loop tracking

1. Introduction

Unmanned aerial vehicles (UAVs) have been widely used for disaster monitoring, emergency search and rescue, and environmental monitoring owing to their flexible deployment and observation capabilities [1,2]. Under low-visibility conditions, such as dense smoke and fog, the sensing capabilities of onboard visible-light cameras and light detection and ranging (LiDAR) sensors may be impaired, and microwave radar can provide an alternative means of perception [3–6]. In microwave coincidence imaging (MCI), spatial diversity is provided by coded radiation fields, and the target scattering distribution is reconstructed from measured echoes using an observation matrix formed from the known radiation fields and propagation responses [7]. Unlike conventional synthetic aperture radar (SAR) and inverse synthetic aperture radar (ISAR), MCI achieves spatial diversity through frequency-dependent radiation patterns without requiring platform motion or target rotation [8,9]. For moving targets, however, errors in the estimated target state can introduce a mismatch between the observation model and the measured echoes, thereby degrading reconstruction performance [10].

To mitigate this mismatch and the resulting imaging degradation, motion estimation, autofocus, and joint reconstruction methods have been employed to improve consistency between the observation model and the measured echoes. Li et al. [11] alternated sparse reconstruction with motion-error estimation to jointly reconstruct the target and compensate for multiple-input multiple-output (MIMO) array motion errors. Chen et al. [12] combined spectrum modification with two-dimensional autofocus to correct motion errors with spatial and frequency dependence in ultrahigh-resolution squint spotlight SAR. For Ka-band carrier-frequency-agility SAR, Zhou et al. [13] used coherent imaging and autofocus to address phase discontinuities caused by frequency hopping and residual motion errors. However, residual errors may remain after compensation, and their impact on reconstruction depends on both their magnitude and the sensitivity of the observation matrix to errors in the estimated target state. Reducing this sensitivity can therefore complement motion compensation.

For UAV-based MCI, the design of the sensing hardware is subject to onboard size, weight, and power constraints [14,15]. Frequency-diverse metasurfaces provide a compact means of spatial encoding, with frequency-dependent field distributions serving

as candidate measurement modes for target reconstruction. Zhao et al. [16] developed a frequency-diverse imager using a highly dispersive random metasurface and a dual-polarized aperture, and experimentally demonstrated computational imaging using measurement modes with low mutual correlation. Beyond mode generation, Li et al. [17] evaluated measurement modes based on regional near-field correlation and selected observations for imaging through contribution matrix sorting, reducing both the number of modes and the computational burden while maintaining reconstruction quality.

However, low radiation-field correlation alone does not ensure that the observation matrix is insensitive to state-prediction errors. Since the observation-matrix entries and their derivatives with respect to the target state vary with frequency, frequency selection can alter the matrix perturbation induced by a given state-prediction error. Maintaining sufficient distinguishability among sparse scattering distributions helps limit the amplification of perturbations during reconstruction. These considerations motivate a frequency-selection criterion that jointly accounts for mismatch sensitivity and sparse-distribution distinguishability under a prescribed recovery-error threshold.

To address this need, a frequency-set design criterion guided by a scattering-coefficient recovery-error bound is proposed and applied within a closed-loop tracking and imaging system based on a frequency-diverse metasurface. The main contributions are summarized as follows:

1. A conditional upper bound on the scattering-coefficient recovery error is derived in the presence of residual state-prediction errors. The resulting observation-matrix perturbation is characterized through first- and second-order variations of the observation model, while the distinguishability of sparse scattering distributions is quantified by a restricted minimum singular value. These two factors are incorporated into the bound, establishing a quantitative relationship between state-prediction errors and scattering-coefficient recovery error.
2. Using the recovery-error upper bound, a prescribed recovery-error threshold is translated into a state-prediction error tolerance. An offline frequency-set design criterion is then formulated to maximize the minimum tolerance over prescribed observation conditions.
3. The selected frequency set is used for closed-loop tracking and imaging, where the state update accounts for the correlation between state-prediction errors and image-derived observation errors. Two-target and extended-target simulations, together with comparisons against alternative frequency-selection criteria, are used to evaluate imaging robustness under residual state-prediction errors.

The remainder of this paper is organized as follows. Section 2 presents the moving-target observation model and closed-loop imaging framework, followed by the recovery-error analysis and offline frequency-set design. Section 3 presents the simulation setup and results. Section 4 discusses the findings and limitations. Section 5 concludes the paper.

2. Materials and Methods

2.1. Moving-Target Observation Model and Closed-Loop Imaging

2.1.1. Observation Scenario and Target Motion Prediction

Figure 1a shows the UAV-based active imaging system for moving targets. The transmitter and a frequency-diverse receiving metasurface with a fixed aperture layout are mounted on the UAV. The fields scattered by the target pass through the metasurface and are collected by a single receiving antenna connected to a single radio-frequency (RF) receive chain. Target translation in a two-dimensional imaging plane is predicted using a constant-acceleration model with process noise.

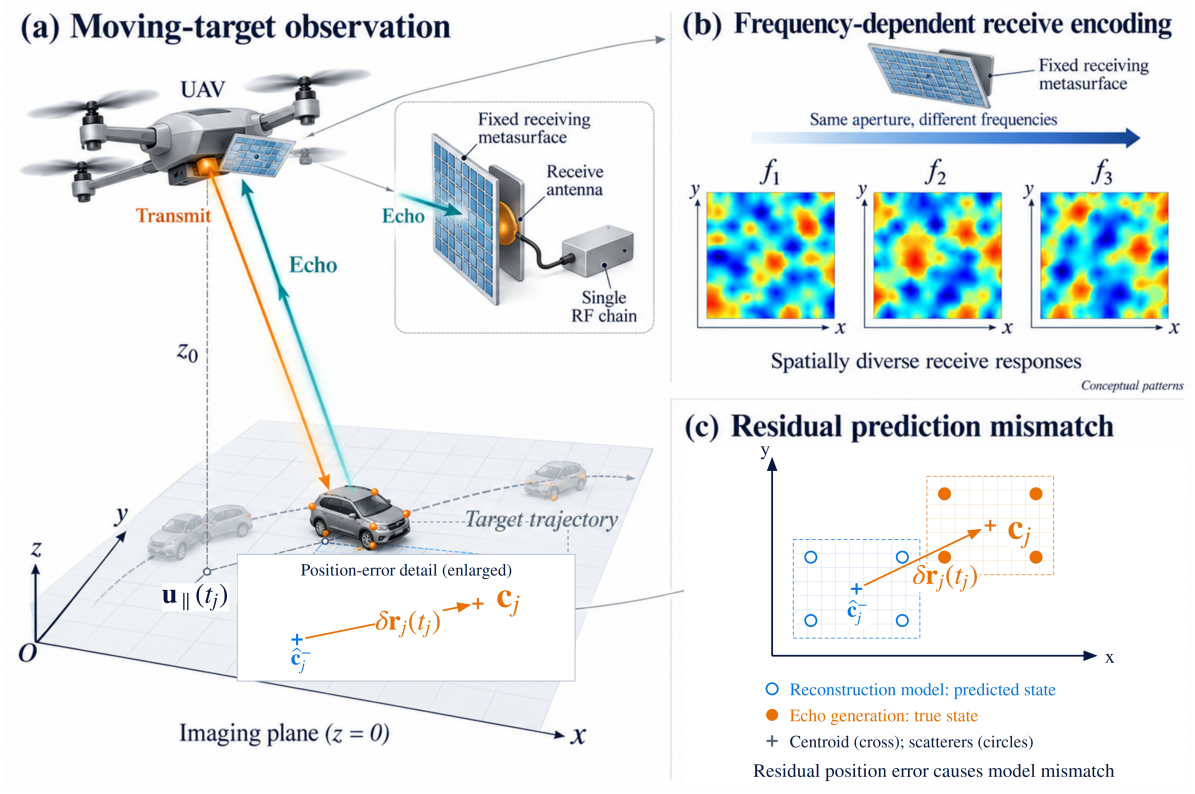


Figure 1. Moving-target microwave coincidence imaging (MCI) using an unmanned aerial vehicle (UAV): (**a**) observation geometry and single-chain reception; (**b**) frequency-dependent receive encoding; and (**c**) residual position-prediction mismatch. In (**b**), the color distributions are schematic and indicate spatial variations in the relative receive-response magnitude.

The target undergoes translation with a fixed spatial structure, and its position is represented by the energy centroid. At the reference time t_j of the jth imaging observation, the motion state is defined as $\mathbf{s}_j = [\mathbf{c}_j^{\mathsf{T}}, \mathbf{v}_j^{\mathsf{T}}, \mathbf{a}_j^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{R}^6$, where $\mathbf{c}_j$, $\mathbf{v}_j$, and $\mathbf{a}_j$ are the two-dimensional centroid position, velocity, and acceleration, respectively. Their initial values are denoted by $\mathbf{c}_0$, $\mathbf{v}_0$, and $\mathbf{a}_0$.

Let $\delta = t - t_j$ denote the time offset from the reference time t_j. Within each imaging acquisition, acceleration is modeled as constant, giving $\mathbf{c}(t) = \mathbf{T}(\delta)\mathbf{s}_j$, where $\mathbf{T}(\delta) = [\mathbf{I}_2, \delta\mathbf{I}_2, \frac{1}{2}\delta^2\mathbf{I}_2]$ maps the motion state at the reference time to the two-dimensional centroid position at t, and $\mathbf{I}_2$ is the 2×2 identity matrix.

For the interval $\Delta t_j = t_j - t_{j-1}$ between consecutive imaging observations, which is set to 10 ms in the simulations, the state is propagated as $\mathbf{s}_j = \mathbf{F}_j\mathbf{s}_{j-1} + \mathbf{w}_j$, where $\mathbf{w}_j \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_j)$ is process noise with covariance $\mathbf{Q}_j$, and the constant-acceleration state-transition matrix is $\mathbf{F}_j = \begin{bmatrix} \mathbf{I}_2 & \Delta t_j\mathbf{I}_2 & \frac{1}{2}\Delta t_j^2\mathbf{I}_2 \\ \mathbf{0} & \mathbf{I}_2 & \Delta t_j\mathbf{I}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \end{bmatrix}$. Process noise represents departures from constant-acceleration motion between consecutive observations, with $\mathbf{Q}_j \succeq \mathbf{0}$.

The Kalman filter is initialized with the state estimate $\hat{\mathbf{s}}_0^+$ and error covariance $\mathbf{P}_0^+$. Before the jth observation, the predicted state estimate and its error covariance are calculated as $\hat{\mathbf{s}}_j^- = \mathbf{F}_j\hat{\mathbf{s}}_{j-1}^+$ and $\mathbf{P}_j^- = \mathbf{F}_j\mathbf{P}_{j-1}^+\mathbf{F}_j^{\mathsf{T}} + \mathbf{Q}_j$, respectively [18]. The superscripts $-$ and $+$ indicate quantities before and after the current observation update, respectively.

The position-extraction matrix is defined as $\mathbf{H} = [\mathbf{I}_2, \mathbf{0}_{2\times2}, \mathbf{0}_{2\times2}]$, yielding the predicted centroid $\hat{\mathbf{c}}_j^- = \mathbf{H}\hat{\mathbf{s}}_j^-$ at the reference time. During this observation, the predicted centroid position is given by $\hat{\mathbf{c}}_j^-(t) = \mathbf{T}(t - t_j)\hat{\mathbf{s}}_j^-$. The UAV horizontal position is then prescribed as $\mathbf{u}_{\parallel}(t) = \hat{\mathbf{c}}_j^-(t) + \mathbf{d}_{\mathrm{U}}$, where $\mathbf{d}_{\mathrm{U}} = d_h[\cos\theta_{\mathrm{U}}, \sin\theta_{\mathrm{U}}]^{\mathsf{T}}$ is a fixed horizontal offset. Together

with the fixed height z_0, this maintains a constant nominal range $\sqrt{d_h^2 + z_0^2}$ and viewing direction relative to the predicted target position. The UAV position at each sampling time is used to calculate the observation geometry. For tracking comparisons conducted along a common prescribed UAV trajectory, all methods use the same trajectory rather than following their respective predicted target positions.

2.1.2. Multifrequency Echo Model and Image Reconstruction

Based on the motion prediction above, the observation matrix is constructed using the frequency-dependent metasurface responses and the corresponding propagation model, and the target scattering-coefficient distribution is recovered from the echoes. Each imaging observation employs N_F carrier frequencies f_p, with $p = 1, \ldots, N_F$. The transmitted signal at the pth carrier frequency in the jth observation is expressed as $S_{j,p}(t) = A_{j,p}(t)e^{i2\pi f_p t}$, where i is the imaginary unit and $A_{j,p}(t)$ is the known complex envelope of the pulse beginning at $t_j + (p-1)T_{\text{step}}$. The complex transmission responses of the metasurface units are obtained from full-wave electromagnetic simulations. Frequency-dependent metasurface responses provide spatial diversity for computational imaging [19], as shown in Figure 1b. Unit screening and aperture-layout optimization are performed offline, and the resulting geometries and layout remain fixed during acquisition. The receiving response at position $\mathbf{r}$ in the imaging plane is expressed as $\Psi_j(\mathbf{r}, f_p; t) = A_{\Psi,j}(\mathbf{r}, f_p; t)e^{i\varphi_{\Psi,j}(\mathbf{r}, f_p; t)}$, where $A_{\Psi,j}$ and $\varphi_{\Psi,j}$ are its amplitude and phase, respectively. The unit-design procedure, fixed aperture layout, and receiving-response model are described in Appendix A.

The imaging region is discretized into $N_x \times N_y$ grid points, with a total of $M = N_x N_y$. The two-dimensional offset of the mth grid point from the grid center is denoted by $\boldsymbol{\xi}_m$, where $m = 1, \ldots, M$. For echo modeling, the grid is centered at the actual centroid, whereas for image reconstruction, it is centered at the predicted centroid. Accordingly, the actual position of the mth grid point at time t is $\mathbf{r}_m(t; \mathbf{s}_j) = \mathbf{T}(t - t_j)\mathbf{s}_j + \boldsymbol{\xi}_m$, whereas the position used to form the observation matrix is $\widehat{\mathbf{r}}_{j,m}(t) = \mathbf{T}(t - t_j)\widehat{\mathbf{s}}_j^{-} + \boldsymbol{\xi}_m$. The resulting position-prediction error at the reference time is $\delta\mathbf{r}_j(t_j) = \mathbf{c}_j - \widehat{\mathbf{c}}_j^{-}$, as shown in Figure 1c. The target scattering distribution is represented by a common scattering-coefficient vector $\mathbf{x}_j = [x_{j,1}, \ldots, x_{j,M}]^{\mathsf{T}} \in \mathbb{C}^M$, which is assumed constant across carrier frequencies and throughout each imaging observation.

Within each imaging observation, the carriers are transmitted sequentially, starting at t_j, with an interval T_{step} between the start times of successive carrier transmissions. In the simulations, $T_{\text{step}} = 100\ \mu$s. Following a receive delay τ_{gate} from the start of each carrier transmission, L samples are collected at

$$t_{j,p,l} = t_j + (p-1)T_{\text{step}} + \tau_{\text{gate}} + \frac{l-1}{f_s},$$

where $p = 1, \ldots, N_F$, $l = 1, \ldots, L$, and f_s is the sampling rate. With approximately coincident transmit and receive phase centers and negligible target displacement during propagation, the round-trip delay for the mth grid point is

$$\tau_{j,p,l,m}(\mathbf{s}_j) = \frac{2}{c}\sqrt{\left\|\mathbf{r}_m(t_{j,p,l}; \mathbf{s}_j) - \mathbf{u}_{\parallel}(t_{j,p,l})\right\|_2^2 + z_0^2}. \tag{1}$$

Here, c is the speed of light, $\mathbf{r}_m(t_{j,p,l}; \mathbf{s}_j)$ is the two-dimensional position of the mth moving grid point at the reception time $t_{j,p,l}$, $\mathbf{u}_{\parallel}(t_{j,p,l})$ is the horizontal projection of the actual UAV position onto the imaging plane, and z_0 is the vertical distance between the UAV and the imaging plane.

The complex transmit response from the transmitter to planar position $\mathbf{r}$ and the corresponding receive response are denoted by $g_T(\mathbf{r}, f_p; t)$ and Ψ_j, respectively. Both include

propagation amplitudes and carrier phases, while propagation delays are modeled as time shifts of the transmitted complex envelope. Using $\mathbf{r}_m = \mathbf{r}_m(t_{j,p,l}; \mathbf{s}_j)$ in the following expression, the entries of the observation matrix $\mathbf{\Phi}_{j,p}(\mathbf{s}_j) \in \mathbb{C}^{L\times M}$ at the pth carrier frequency are given by

$$[\mathbf{\Phi}_{j,p}(\mathbf{s}_j)]_{l,m} = g_T(\mathbf{r}_m, f_p; t_{j,p,l})\Psi_j(\mathbf{r}_m, f_p; t_{j,p,l})A_{j,p}\big(t_{j,p,l} - \tau_{j,p,l,m}(\mathbf{s}_j)\big). \tag{2}$$

The echo at the pth carrier frequency satisfies $\mathbf{y}_{j,p} = \mathbf{\Phi}_{j,p}(\mathbf{s}_j)\mathbf{x}_j + \mathbf{n}_{j,p}$, where $\mathbf{y}_{j,p} \in \mathbb{C}^L$ is the echo vector and $\mathbf{n}_{j,p} \in \mathbb{C}^L$ is the corresponding zero-mean noise vector. By vertically stacking the echo vectors, noise vectors, and observation matrices in carrier-frequency order, the multifrequency echo model is obtained as [20]

$$\mathbf{y}_j = \mathbf{\Phi}_j(\mathbf{s}_j)\mathbf{x}_j + \mathbf{n}_j. \tag{3}$$

Here, $\mathbf{y}_j = \mathrm{col}_{p=1}^{N_F}(\mathbf{y}_{j,p})$, $\mathbf{n}_j = \mathrm{col}_{p=1}^{N_F}(\mathbf{n}_{j,p})$, and $\mathbf{\Phi}_j(\mathbf{s}_j) = \mathrm{col}_{p=1}^{N_F}[\mathbf{\Phi}_{j,p}(\mathbf{s}_j)]$, where col denotes vertical stacking, $\mathbf{y}_j, \mathbf{n}_j \in \mathbb{C}^{LN_F}$, and $\mathbf{\Phi}_j \in \mathbb{C}^{LN_F\times M}$.

For image reconstruction, the motion-compensated observation matrix $\widehat{\mathbf{\Phi}}_j = \mathbf{\Phi}_j(\widehat{\mathbf{s}}_j^-)$ is formed using the predicted state. Residual state-prediction errors can therefore cause a mismatch between the true observation matrix and the matrix used for reconstruction. The effect of this mismatch on recovery is analyzed in Section 2.2. The multifrequency echo-noise covariance is denoted by $\mathbf{C}_{n,j} = \mathbb{E}[\mathbf{n}_j\mathbf{n}_j^{\mathsf{H}}]$, where the superscript H denotes the conjugate transpose. Sparse Bayesian learning (SBL) reconstruction is performed using $\mathbf{y}_j$, $\widehat{\mathbf{\Phi}}_j$, and $\mathbf{C}_{n,j}$ [21], yielding the scattering-coefficient estimate $\widehat{\mathbf{x}}_j$ on the predicted grid. The image intensity at grid point m is $I_{j,m} = |\widehat{x}_{j,m}|^2$.

2.1.3. Centroid Observation and Closed-Loop Tracking Update

The intensity-weighted target centroid is extracted from the reconstructed image and used as the position observation for the Kalman-filter update. For an intensity threshold $\eta_j > 0$, the selected grid-point set is defined as $\mathcal{M}_j = \{m : I_{j,m} > \eta_j\}$. When $\mathcal{M}_j$ is nonempty, the normalized intensity weight is $\omega_{j,m} = I_{j,m}/\sum_{k\in\mathcal{M}_j} I_{j,k}$. Since the reconstruction grid is centered at the predicted centroid $\widehat{\mathbf{c}}_j^-$ at t_j, the resulting observation is

$$\mathbf{z}_j = \widehat{\mathbf{c}}_j^- + \sum_{m\in\mathcal{M}_j} \omega_{j,m}\boldsymbol{\xi}_m. \tag{4}$$

The image-derived position observation is used to update the predicted state. Because this observation is extracted from an image reconstructed using the predicted state, its error can be statistically correlated with the state-prediction error. The position-observation error and state-prediction error are defined as $\boldsymbol{\epsilon}_j = \mathbf{z}_j - \mathbf{H}\mathbf{s}_j$ and $\delta\mathbf{s}_j = \mathbf{s}_j - \widehat{\mathbf{s}}_j^-$, respectively. Let $\bar{\boldsymbol{\epsilon}}_j$ denote the mean observation error estimated from a separate calibration set. The covariance $\mathbf{R}_j = \mathrm{Cov}(\boldsymbol{\epsilon}_j)$ and cross-covariance $\mathbf{C}_j = \mathrm{Cov}(\delta\mathbf{s}_j, \boldsymbol{\epsilon}_j)$ are estimated from paired, mean-centered error samples in the same set. After correcting for the estimated observation bias, the innovation is $\boldsymbol{\nu}_j = \mathbf{z}_j - \bar{\boldsymbol{\epsilon}}_j - \mathbf{H}\widehat{\mathbf{s}}_j^-$, with covariance

$$\mathbf{V}_j = \mathbf{H}\mathbf{P}_j^-\mathbf{H}^{\mathsf{T}} + \mathbf{R}_j + \mathbf{H}\mathbf{C}_j + \mathbf{C}_j^{\mathsf{T}}\mathbf{H}^{\mathsf{T}}. \tag{5}$$

The corresponding Kalman gain is

$$\mathbf{K}_j = (\mathbf{P}_j^-\mathbf{H}^{\mathsf{T}} + \mathbf{C}_j)\mathbf{V}_j^{-1}. \tag{6}$$

The updated state estimate and error covariance are $\hat{\mathbf{s}}_j^+ = \hat{\mathbf{s}}_j^- + \mathbf{K}_j \boldsymbol{\nu}_j$ and $\mathbf{P}_j^+ = \mathbf{P}_j^- - \mathbf{K}_j \mathbf{V}_j \mathbf{K}_j^\mathsf{T}$, respectively.

The covariance update requires the joint covariance $\boldsymbol{\Sigma}_j$ to be positive semidefinite and the innovation covariance $\mathbf{V}_j$ to be positive definite. These conditions were verified numerically for all evaluated updates. The effect of accounting for this cross-covariance is evaluated in Appendix G.

When $\mathcal{M}_j$ is empty, the observation update is skipped, with $\hat{\mathbf{s}}_j^+ = \hat{\mathbf{s}}_j^-$ and $\mathbf{P}_j^+ = \mathbf{P}_j^-$. The resulting estimates are propagated to predict the target motion for the next imaging observation.

2.2. Recovery-Error Analysis and Robust Frequency-Set Design

The frequency-set design principle is outlined in Figure 2. Observation-matrix sensitivity and the distinguishability of sparse scattering distributions are jointly evaluated through a conditional recovery-error bound, from which a state-prediction error tolerance is derived for a prescribed recovery-error threshold.

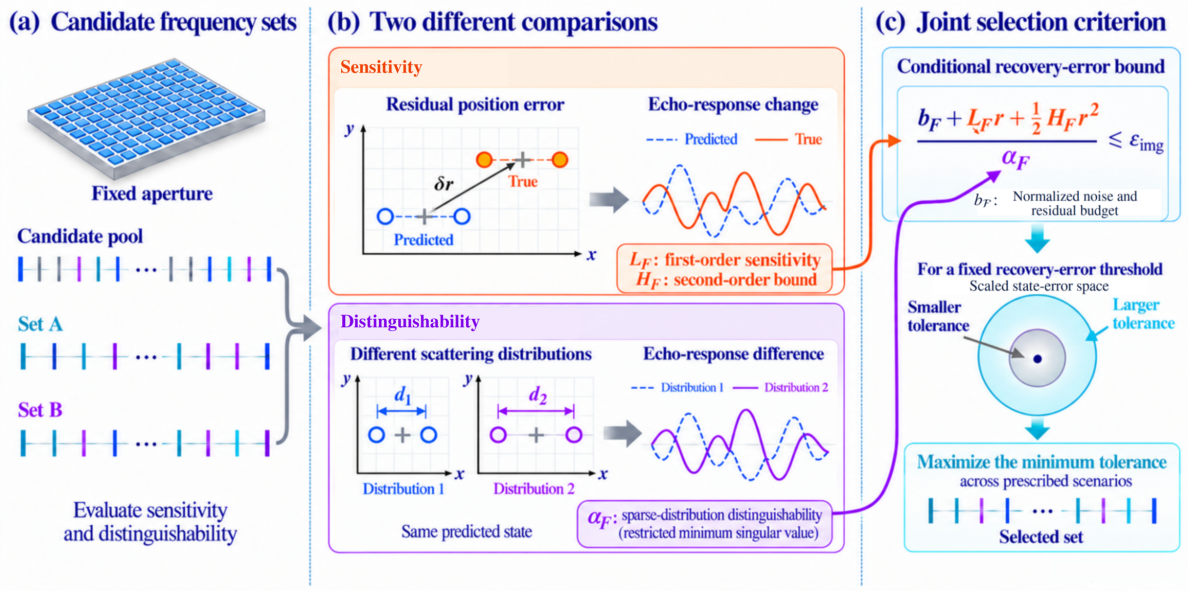


Figure 2. Recovery-error-bound-guided frequency-set design: (**a**) candidate frequency sets; (**b**) mismatch sensitivity and sparse-distribution distinguishability; and (**c**) joint evaluation and tolerance-based selection.

2.2.1. State-Error-Induced Observation-Matrix Mismatch

The state-prediction error $\delta\mathbf{s}_j$, defined in Section 2.1.3, induces a position offset at each sampling time. Using the motion mapping in Section 2.1.1, the offset at the lth sample of the pth carrier frequency is $\delta\mathbf{r}_{j,p,l} = \mathbf{T}(t_{j,p,l} - t_j)\delta\mathbf{s}_j$.

The carrier frequencies used in the current imaging observation are collected into the frequency set $\mathcal{F}$. For the frequency set $\mathcal{F}$, the predicted and actual observation matrices are given by $\widehat{\boldsymbol{\Phi}}_{j,\mathcal{F}} = \boldsymbol{\Phi}_{j,\mathcal{F}}(\hat{\mathbf{s}}_j^-)$ and $\boldsymbol{\Phi}_{j,\mathcal{F}}(\hat{\mathbf{s}}_j^- + \delta\mathbf{s}_j)$, respectively. Defining the observation-matrix perturbation as $\Delta\boldsymbol{\Phi}_{j,\mathcal{F}} = \boldsymbol{\Phi}_{j,\mathcal{F}}(\hat{\mathbf{s}}_j^- + \delta\mathbf{s}_j) - \widehat{\boldsymbol{\Phi}}_{j,\mathcal{F}}$, the echo model can be expressed as

$$\mathbf{y}_{j,\mathcal{F}} = \widehat{\boldsymbol{\Phi}}_{j,\mathcal{F}}\mathbf{x}_j + \Delta\boldsymbol{\Phi}_{j,\mathcal{F}}\mathbf{x}_j + \mathbf{n}_{j,\mathcal{F}}. \tag{7}$$

Here, $\mathbf{y}_{j,\mathcal{F}}$ and $\mathbf{n}_{j,\mathcal{F}}$ are the measured echo vector and noise vector for $\mathcal{F}$, respectively, and $\mathbf{x}_j$ is the target scattering-coefficient vector. When the predicted observation matrix is used for

scattering-coefficient recovery, the reconstruction is affected by the state-prediction error through the additional echo term $\Delta\boldsymbol{\Phi}_{j,\mathcal{F}}\mathbf{x}_j$.

To characterize this observation-matrix mismatch, each matrix entry is assumed to be twice continuously differentiable with respect to the state vector in the neighborhood under consideration. A Taylor expansion about the predicted state $\hat{\mathbf{s}}_j^-$ then gives [22] $\Delta\boldsymbol{\Phi}_{j,\mathcal{F}} = \sum_{a=1}^{6} \delta s_{j,a}\mathbf{D}_{j,\mathcal{F}}^{(a)} + \mathbf{R}_{j,\mathcal{F}}^{(2)}$. Here, a indexes the state components, $\delta s_{j,a}$ is the ath component of the state-prediction error, $\mathbf{D}_{j,\mathcal{F}}^{(a)} = \partial\boldsymbol{\Phi}_{j,\mathcal{F}}(\mathbf{s})/\partial s_a\big|_{\mathbf{s}=\hat{\mathbf{s}}_j^-}$ is the observation-matrix derivative with respect to that component, and $\mathbf{R}_{j,\mathcal{F}}^{(2)}$ is the second-order remainder.

To evaluate the state derivatives, let $h_{j,p,l}(\mathbf{r})$ denote the noiseless received signal from a point scatterer with unit scattering coefficient at position $\mathbf{r}$, evaluated at carrier frequency f_p and sampling time $t = t_{j,p,l}$: $h_{j,p,l}(\mathbf{r}) = g_T(\mathbf{r}, f_p; t)\Psi_j(\mathbf{r}, f_p; t)A_{j,p}\big(t - \tau(\mathbf{r}, t)\big)$. Here, $\tau(\mathbf{r}, t)$ is the round-trip delay in (1) evaluated at position $\mathbf{r}$. Evaluating $h_{j,p,l}$ at the moving grid point $\mathbf{r}_m(t; \mathbf{s}_j)$ gives the corresponding entry of $\boldsymbol{\Phi}_{j,p}(\mathbf{s}_j)$. With time and UAV geometry held fixed, the chain rule gives

$$\frac{\partial h_{j,p,l}}{\partial \mathbf{s}_j^{\mathsf{T}}} = (\nabla_{\mathbf{r}} h_{j,p,l})^{\mathsf{T}}\mathbf{T}(\delta) = \left[h_x, h_y, \delta h_x, \delta h_y, \frac{\delta^2}{2}h_x, \frac{\delta^2}{2}h_y\right]. \tag{8}$$

Here, $\delta = t_{j,p,l} - t_j$, $h_x = \partial h_{j,p,l}/\partial x$, and $h_y = \partial h_{j,p,l}/\partial y$, with $\nabla_{\mathbf{r}} h_{j,p,l} = [h_x, h_y]^{\mathsf{T}}$. All derivatives are evaluated at the predicted grid position; their expressions are given in Appendix A. Position, velocity, and acceleration errors affect the response through these spatial derivatives, weighted by 1, δ, and $\delta^2/2$, respectively.

At a fixed sampling time and observation geometry, the nonzero point-scatterer response can be expressed in amplitude–phase form as $h(\mathbf{r}, f) = a(\mathbf{r}, f)e^{i\varphi(\mathbf{r},f)}$, where $a > 0$ and φ denote its amplitude and phase, respectively. For a small position offset $\delta\mathbf{r} = \ell\mathbf{d}$, where $\mathbf{d} \in \mathbb{R}^2$ is a unit direction vector and ℓ is the signed displacement, the relative response deviation is

$$\frac{\delta h}{h} = \ell\frac{\mathbf{d}^{\mathsf{T}}\nabla_{\mathbf{r}} h}{h} + \mathcal{O}(\ell^2) = \ell\Big[\mathbf{d}^{\mathsf{T}}\nabla_{\mathbf{r}} \ln a + i\mathbf{d}^{\mathsf{T}}\nabla_{\mathbf{r}}\varphi\Big] + \mathcal{O}(\ell^2). \tag{9}$$

Here, $\delta h = h(\mathbf{r} + \ell\mathbf{d}, f) - h(\mathbf{r}, f)$, and $\mathcal{O}(\ell^2)$ denotes the second-order remainder. The terms $\mathbf{d}^{\mathsf{T}}\nabla_{\mathbf{r}} \ln a$ and $\mathbf{d}^{\mathsf{T}}\nabla_{\mathbf{r}}\varphi$ describe the directional rates of relative amplitude change and phase change, respectively.

These local spatial derivatives are not directly constrained by the correlation between radiation fields at different frequencies. Let $\boldsymbol{\psi}_{j,p} \in \mathbb{C}^M$ contain the complex receiving-response values at the predicted grid points at the reference time t_j, with $[\boldsymbol{\psi}_{j,p}]_m = \Psi_j(\hat{\mathbf{r}}_{j,m}(t_j), f_p; t_j)$. For nonzero field vectors, the inter-frequency correlation is $\rho_{p,q}^{\Psi} = |\boldsymbol{\psi}_{j,p}^{\mathsf{H}}\boldsymbol{\psi}_{j,q}|/(\|\boldsymbol{\psi}_{j,p}\|_2\|\boldsymbol{\psi}_{j,q}\|_2)$. A small $\rho_{p,q}^{\Psi}$ therefore does not by itself ensure low sensitivity to residual state-prediction errors. The proposed frequency-set design accounts for this sensitivity together with the distinguishability of sparse scattering distributions.

2.2.2. Conditional Recovery-Error Bound and State-Error Tolerance

Let $\mathbf{L}_j = (\mathbf{P}_j^-)^{1/2}$ denote the symmetric positive semidefinite square root of the prediction covariance computed with the prescribed process-noise covariance $\mathbf{Q}_j$ in Section 2.1.1. The state-prediction-error set is defined by $\delta\mathbf{s}_j = \mathbf{L}_j\mathbf{u}$, with $\mathbf{u} \in \mathbb{R}^6$, $\|\mathbf{u}\|_2 \le r$, and $0 \le r \le r_0$.

Here, $\mathbf{u}$ is the dimensionless error parameter in covariance-scaled coordinates, and r is a dimensionless radius bounding $\|\mathbf{u}\|_2$. The physical position error is obtained from $\delta\mathbf{r}_j(t) = \mathbf{T}(t - t_j)\mathbf{L}_j\mathbf{u}$. The corresponding directional position-error bounds are given in

Equation (14). The parameter $r_0 > 0$ specifies the neighborhood over which the second-derivative bound holds.

For a frequency set $\mathcal{F}$, the observation-matrix mismatch is bounded by substituting $\delta \mathbf{s}_j = \mathbf{L}_j \mathbf{u}$ into the Taylor expansion of the observation matrix associated with $\mathcal{F}$. Let $L_\mathcal{F}$ be the maximum norm of its first directional derivative with respect to $\mathbf{u}$ at $\mathbf{u} = \mathbf{0}$, taken over real unit directions. Let $H_\mathcal{F}$ be an upper bound on the spectral norm of the second directional derivative, uniformly over all real unit directions and all $\mathbf{u}$ satisfying $\|\mathbf{u}\|_2 \le r_0$. For $\|\mathbf{u}\|_2 \le r \le r_0$, the norms of the first-order term and second-order remainder are bounded by $L_\mathcal{F} r$ and $\frac{1}{2} H_\mathcal{F} r^2$, respectively, giving

$$\|\Delta \mathbf{\Phi}_{j,\mathcal{F}}\|_2 \le L_\mathcal{F} r + \frac{1}{2} H_\mathcal{F} r^2. \tag{10}$$

The expressions for $L_\mathcal{F}$ and $H_\mathcal{F}$ and the derivation of the observation-matrix mismatch bound are given in Appendix B. For $\|\mathbf{u}\|_2 \le r \le r_0$, (10) implies $\|\Delta \mathbf{\Phi}_{j,\mathcal{F}} \mathbf{x}_j\|_2 \le (L_\mathcal{F} r + \frac{1}{2} H_\mathcal{F} r^2) \|\mathbf{x}_j\|_2$.

Consider an estimate $\widehat{\mathbf{x}}_j$ obtained using $\widehat{\mathbf{\Phi}}_{j,\mathcal{F}}$. The recovery-error analysis assumes $\|\mathbf{n}_{j,\mathcal{F}}\|_2 \le \varepsilon_{n,\mathcal{F}}$ and $\|\widehat{\mathbf{\Phi}}_{j,\mathcal{F}} \widehat{\mathbf{x}}_j - \mathbf{y}_{j,\mathcal{F}}\|_2 \le \varepsilon_{\mathrm{rec},\mathcal{F}}$, where $\varepsilon_{n,\mathcal{F}}$ and $\varepsilon_{\mathrm{rec},\mathcal{F}}$ are the prescribed noise and residual bounds. For a design scenario indexed by ω, both $\mathbf{x}_j$ and $\widehat{\mathbf{x}}_j$ have at most K nonzero entries, with supports contained in $\mathcal{I}_\omega \subseteq \{1, \ldots, M\}$. Their difference is therefore supported on $\mathcal{I}_\omega$ and has at most $2K$ nonzero entries. The distinguishability of these sparse differences is characterized by the restricted minimum singular value

$$\alpha_\mathcal{F} = \inf_{\substack{\mathbf{v} \in \mathbb{C}^M,\ \mathbf{v} \neq \mathbf{0} \\ \|\mathbf{v}\|_0 \le 2K,\ v_m = 0\ (m \notin \mathcal{I}_\omega)}} \frac{\|\widehat{\mathbf{\Phi}}_{j,\mathcal{F}} \mathbf{v}\|_2}{\|\mathbf{v}\|_2}. \tag{11}$$

Here, $\mathbf{v}$ is a nonzero sparse difference vector, v_m is its mth entry, and $\|\mathbf{v}\|_0$ counts its nonzero entries. For $\alpha_\mathcal{F} > 0$ and a prescribed signal-norm lower bound $\|\mathbf{x}_j\|_2 \ge x_{\min} > 0$, the mismatch bound and the above recovery conditions yield

$$\frac{\|\widehat{\mathbf{x}}_j - \mathbf{x}_j\|_2}{\|\mathbf{x}_j\|_2} \le \frac{b_\mathcal{F} + L_\mathcal{F} r + \frac{1}{2} H_\mathcal{F} r^2}{\alpha_\mathcal{F}}. \tag{12}$$

Here, $b_\mathcal{F} = (\varepsilon_{\mathrm{rec},\mathcal{F}} + \varepsilon_{n,\mathcal{F}})/x_{\min}$ is the sum of the reconstruction-residual and noise bounds normalized by the prescribed lower bound on $\|\mathbf{x}_j\|_2$. A detailed derivation is provided in Appendix B.

Equation (12) links state-prediction errors to scattering-coefficient recovery accuracy. The coefficients $L_\mathcal{F}$ and $H_\mathcal{F}$ quantify mismatch sensitivity, while $\alpha_\mathcal{F}$ characterizes the distinguishability of sparse scattering distributions and appears in the denominator of the recovery-error bound. The bound therefore combines mismatch sensitivity and sparse-distribution distinguishability in a common recovery-error measure, as shown in Figure 2b,c. The bound applies to complex-valued scattering coefficients with arbitrary phases, provided that the stated sparsity, support, signal-norm, noise, and reconstruction-residual conditions are satisfied.

To compare frequency sets under the same recovery-accuracy requirement, a normalized recovery-error threshold $\varepsilon_{\mathrm{img}} > 0$ is specified. After the noise and reconstruction-residual term is accounted for, the remaining allowance for observation-matrix mismatch is $m_\mathcal{F} = \varepsilon_{\mathrm{img}} \alpha_\mathcal{F} - b_\mathcal{F}$. For $m_\mathcal{F} > 0$, the recovery-error bound does not exceed $\varepsilon_{\mathrm{img}}$ provided

that $L_{\mathcal{F}} r + \frac{1}{2} H_{\mathcal{F}} r^2 \leq m_{\mathcal{F}}$. For $H_{\mathcal{F}} > 0$, the largest $r \in [0, r_0]$ for which the recovery-error bound does not exceed ε_{img} is

$$r_{\text{safe}}(\mathcal{F}) = \min\left\{ r_0, \frac{\sqrt{L_{\mathcal{F}}^2 + 2H_{\mathcal{F}} m_{\mathcal{F}}} - L_{\mathcal{F}}}{H_{\mathcal{F}}} \right\}. \tag{13}$$

For a fixed design scenario, prediction covariance, and recovery-error threshold, a larger $r_{\text{safe}}(\mathcal{F})$ provides a wider state-error set over which the prescribed recovery accuracy is guaranteed under the stated sparsity, support, noise, and residual conditions. The resulting r_{safe} is a sufficient state-error tolerance under the stated conditions. The case $H_{\mathcal{F}} = 0$ is given in Appendix B. For a fixed frequency set and observation condition, the tolerance is nondecreasing in the permitted recovery error and nonincreasing in the noise and reconstruction-residual budgets.

To express this tolerance in position coordinates, use $\delta \mathbf{r}_j(t) = \mathbf{T}(t - t_j)\mathbf{L}_j \mathbf{u}$. For a two-dimensional unit direction vector $\mathbf{d}$, the maximum absolute position-error component along $\mathbf{d}$ over the state-error set $\|\mathbf{u}\|_2 \leq r_{\text{safe}}(\mathcal{F})$ is

$$e_{\text{safe},\mathbf{d}}(t) = r_{\text{safe}}(\mathcal{F}) \sqrt{\mathbf{d}^{\mathsf{T}} \mathbf{T}(t - t_j) \mathbf{P}_j^{-} \mathbf{T}^{\mathsf{T}}(t - t_j) \mathbf{d}}. \tag{14}$$

Thus, every state error in this set satisfies $|\mathbf{d}^{\mathsf{T}} \delta \mathbf{r}_j(t)| \leq e_{\text{safe},\mathbf{d}}(t)$. The x- and y-direction projection limits are obtained with $\mathbf{d} = [1, 0]^{\mathsf{T}}$ and $\mathbf{d} = [0, 1]^{\mathsf{T}}$, respectively. The derivation is given in Appendix B.

2.2.3. Offline Frequency-Set Design

The state-error tolerance derived above is employed to design the frequency set, while the metasurface geometry and aperture layout are kept fixed. For each design scenario, the prescribed prediction covariance defines the uncertainty model used to compare candidate frequency sets. The frequency set is selected offline and then retained throughout the subsequent closed-loop tracking and imaging process. During online operation, the state estimate and covariance are updated recursively, and the observation matrix is constructed using the current predicted state, whereas the frequency set is not reoptimized.

The candidate frequency set is denoted by $\mathcal{C} = \{f_1^{\text{c}}, \ldots, f_{N_c}^{\text{c}}\}$, where N_c is the number of candidate frequencies. A predetermined number $N_F < N_c$ of distinct frequencies is selected from $\mathcal{C}$ to form $\mathcal{F}$. A prescribed set of imaging scenarios, denoted by Ω, is used to evaluate each candidate frequency set. Each design scenario specifies a predicted target state, prediction covariance, UAV position, sampling times, and an admissible sparse-support set, together with the prescribed noise and reconstruction-residual limits. These inputs are fixed before frequency selection and shared by all candidate frequency sets. The scenario set Ω specifies the observation geometries and target-support priors considered in the design.

The criterion is evaluated on prescribed sparse-support regions for offline frequency selection; the implementation settings are given in Appendix C.

For each scenario $\omega \in \Omega$, the theoretical tolerance $r_{\text{safe}}(\mathcal{F}; \omega)$ is calculated from (13) using the corresponding values of $\alpha_{\mathcal{F}}$, $L_{\mathcal{F}}$, $H_{\mathcal{F}}$, and $b_{\mathcal{F}}$. For a prescribed normalized recovery-error threshold ε_{img}, the minimum tolerance across the prescribed scenarios is defined as $\underline{r}_{\text{safe}}(\mathcal{F}) = \min_{\omega \in \Omega} r_{\text{safe}}(\mathcal{F}; \omega)$. Under the conditions stated in Section 2.2.2, this value gives the largest r for which the recovery-error bound does not exceed the

threshold in every scenario. Maximizing $\underline{r}_{\text{safe}}(\mathcal{F})$, subject to constraints on frequency count, frequency span, radiation-pattern correlation, and positivity, yields

$$\begin{aligned} \max_{\mathcal{F}} \quad & \underline{r}_{\text{safe}}(\mathcal{F}) = \min_{\omega\in\Omega} r_{\text{safe}}(\mathcal{F};\omega), \\ \text{s.t.} \quad & \mathcal{F} \subseteq \mathcal{C}, \quad |\mathcal{F}| = N_F, \quad \max_{f\in\mathcal{F}} f - \min_{f\in\mathcal{F}} f = B, \\ & \rho^{\Psi}_{p,q}(\omega) \le \rho_{\max}, \quad \forall\, f_p, f_q \in \mathcal{F},\ p \ne q, \quad \forall\, \omega \in \Omega, \\ & \alpha_{\mathcal{F}}(\omega) > 0, \quad m_{\mathcal{F}}(\omega) > 0, \quad \forall \omega \in \Omega. \end{aligned} \tag{15}$$

Here, $|\mathcal{F}|$ is the number of selected frequencies and B is the prescribed frequency span. The quantity $\rho^{\Psi}_{p,q}(\omega)$ denotes the correlation coefficient between the radiation patterns at the pth and qth frequencies in scenario ω, and $\rho_{\max}$ is the prescribed upper bound for all distinct pairs of selected frequencies across all design scenarios. In that scenario, $\alpha_{\mathcal{F}}(\omega)$ and $m_{\mathcal{F}}(\omega)$ denote the restricted minimum singular value and the allowance for observation-matrix mismatch, respectively. The constraints $\alpha_{\mathcal{F}}(\omega) > 0$ and $m_{\mathcal{F}}(\omega) > 0$ ensure a positive theoretical tolerance. The correlation constraint limits inter-frequency redundancy, whereas the objective evaluates recovery robustness under state-prediction errors. All candidate frequency sets are evaluated using the same prescribed scenarios, sampling duration, and transmit energy.

The discrete optimization problem is solved by a constrained search based on the Fedorov exchange framework [23], as shown in Figure 3a. Starting from an initial feasible frequency set $\mathcal{F}_0$, one selected frequency is exchanged with an unselected frequency. An exchange is retained only if the resulting frequency set satisfies all constraints and increases $\underline{r}_{\text{safe}}(\mathcal{F})$. The search terminates when no feasible single-frequency exchange can further increase this objective, yielding $\mathcal{F}_{\text{sel}}$, which is locally optimal with respect to such exchanges.

The relative increase in the minimum tolerance is $G_{\text{safe}} = \underline{r}_{\text{safe}}(\mathcal{F}_{\text{sel}})/\underline{r}_{\text{safe}}(\mathcal{F}_0) - 1$. Because every accepted exchange increases the objective, $G_{\text{safe}} \ge 0$.

Figure 3 summarizes the offline frequency-selection and online imaging workflow. The selected frequency set is fixed during online operation, while image-derived position observations are fed back for state updates. The corresponding steps are listed in Algorithm 1.

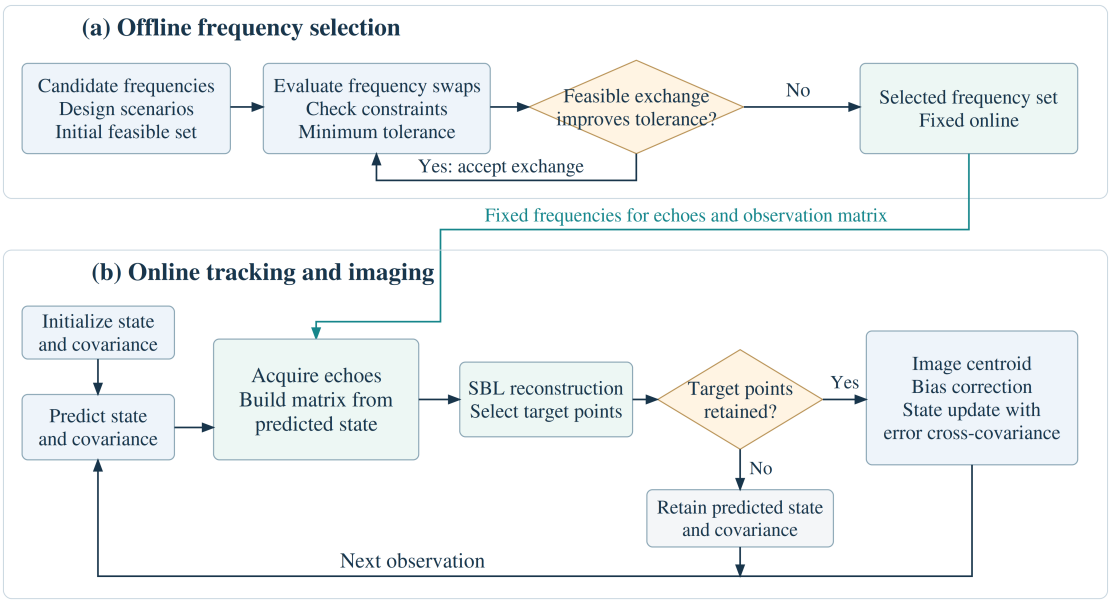


Figure 3. Workflow of (**a**) offline frequency selection and (**b**) online closed-loop tracking and imaging.

Frequency selection is performed offline. With identical matrix dimensions, reconstruction implementations, and iteration counts, the selected and uniform frequency sets have the same asymptotic online computational complexity.

Algorithm 1: Frequency-Set Optimization and Closed-Loop Imaging

Input:

Candidate set $\mathcal{C}$, initial feasible set $\mathcal{F}_0$, scenario set Ω, initial state estimate $\widehat{\mathbf{s}}_0^+$, and initial error covariance $\mathbf{P}_0^+$.

Output:

$\mathcal{F}_{\text{sel}}$, reconstructed images, and updated target-state estimates.

Offline frequency-set optimization

1 Initialize $\mathcal{F} \leftarrow \mathcal{F}_0$ and evaluate $\underline{r}_{\text{safe}}(\mathcal{F})$.

2 Exchange one selected frequency with one unselected frequency, retaining only feasible exchanges that increase $\underline{r}_{\text{safe}}(\mathcal{F})$.

3 Repeat Step 2 until no feasible exchange improves the objective; set $\mathcal{F}_{\text{sel}} \leftarrow \mathcal{F}$.

Online closed-loop imaging

4 For observation j, predict $\widehat{\mathbf{s}}_j^-$ and $\mathbf{P}_j^-$ and construct $\widehat{\mathbf{\Phi}}_j$ using $\mathcal{F}_{\text{sel}}$.

5 Recover $\widehat{\mathbf{x}}_j$ using SBL and extract the image-derived position observation $\mathbf{z}_j$ when $\mathcal{M}_j$ is nonempty.

6 If a valid observation is obtained, update $\widehat{\mathbf{s}}_j^+$ and $\mathbf{P}_j^+$; otherwise retain the predicted state estimate and covariance. Repeat Steps 4–6 for the next observation.

3. Results

The simulation parameters are summarized in Table 1. Frequency-set selection, echo generation, image reconstruction, closed-loop tracking, and performance evaluation were implemented in MATLAB R2022a (MathWorks). Three configurations were compared: open-loop MCI with the uniform frequency set $\mathcal{F}_0$ (OL), closed-loop MCI with $\mathcal{F}_0$ (CL), and closed-loop MCI with the selected set $\mathcal{F}_{\text{sel}}$ (CL-FS). OL used simulated position observations without image feedback, whereas CL and CL-FS used reconstructed-image centroids. Unless otherwise stated, each configuration used its own predicted state for UAV following and observation-matrix formation.

The selected frequency set and the reconstruction settings are specified in Appendix C.

The theoretical evaluation and position-offset simulations used two scatterers at $(-0.08, 0)$ and $(0.08, 0)$ m relative to the centroid, with fixed complex scattering coefficients $x_1 = x_2 = 1$, corresponding to unit magnitudes and zero phases. The same coefficients were used for all simulated position offsets and all frequency sets being compared. Propagation phases and the frequency-dependent complex metasurface responses were retained in the observation matrix defined in Equation (2). The uniform set $\mathcal{F}_0$ and the selected set $\mathcal{F}_{\text{sel}}$ used equal transmitted energy per frame. Circularly symmetric complex Gaussian noise was added with a common variance for all methods and both target configurations. The common noise variance was set to the mean power of the noiseless reference echo divided by $10^{\text{SNR}/10}$, where SNR is expressed in decibels. The reference echo was obtained using $\mathcal{F}_0$ for the two-scatterer target at the first frame with zero position offset. Thus, the reported SNR values are defined relative to the common two-scatterer reference echo rather than the actual echo power of each target configuration and frequency set.

Table 2 compares the conditional state-error tolerances and their positional equivalents at an SNR of 10 dB and $\varepsilon_{\text{img}} = 0.85$. For both frequency sets, the minimum tolerance over the prescribed geometries occurred at the zero-offset geometry ω_{rep}. The design geometries and numerical settings are specified in Appendix C.

Under the conditions in Section 2.2.2, the minimum tolerance increased from 0.000560600 to 0.000843224, a gain of approximately 50.4%. The positional equivalents were calculated using $\mathbf{P}_1^-$, for which $e_{\text{safe},x}(t_1) = e_{\text{safe},y}(t_1)$. These positional values are projections of the conditional state-error tolerance onto the position coordinates. Both frequency sets satisfy the prescribed field-correlation constraint.

Table 1. Simulation parameters.

Parameter	Value
Metasurface aperture	25×25 elements with a 5 mm pitch; 125 mm $\times$ 125 mm
Candidate frequencies	32–36 GHz with a 0.05 GHz spacing; 81 points
Uniform frequency set	$(32 + 0.2k)$ GHz, $k = 0, \ldots, 20$
Frequencies per set	21
Transmit pulse	Gaussian; full width at half maximum of the power envelope: 2 ns
Receive delay	33 ns
Inter-carrier transmission interval, T_{step}	100 µs
Inter-frame interval, Δt_j	10 ms
Sampling	8 GHz; 64 samples per frequency
Simulation length	150 frames
UAV height	5 m above the imaging plane
Reconstruction grid	0.8 m $\times$ 0.8 m with a 0.02 m spacing; 41 $\times$ 41 points
Initial target state	Position at the origin; velocity 8 m/s and acceleration 3 m/s^2, both along $+x$

UAV—unmanned aerial vehicle.

Table 2. Conditional state-error tolerance and its positional equivalent in the zero-offset geometry ω_{rep}.

Set	$\alpha_{\mathcal{F}}$	$\overline{L}_{\mathcal{F}}$	$H_{\mathcal{F}}$	$b_{\mathcal{F}}$	$m_{\mathcal{F}}$	r_{safe}	$e_{\text{safe},x}(t_1)$ (m)	$\rho^{\Psi}_{\text{peak}}$
$\mathcal{F}_0$	0.002823023	1.067628	6.728319	0.001800	0.000599570	0.000560600	$2.24520218 \times 10^{-5}$	0.437205
$\mathcal{F}_{\text{sel}}$	0.003145944	1.034017	6.032220	0.001800	0.000874052	0.000843224	$3.37710943 \times 10^{-5}$	0.040756

Figure 4 compares the simulated observation-matrix perturbations and normalized recovery errors with their corresponding theoretical bounds. The selected frequency set yields smaller matrix perturbations and lower recovery errors than the uniform frequency set over the evaluated range.

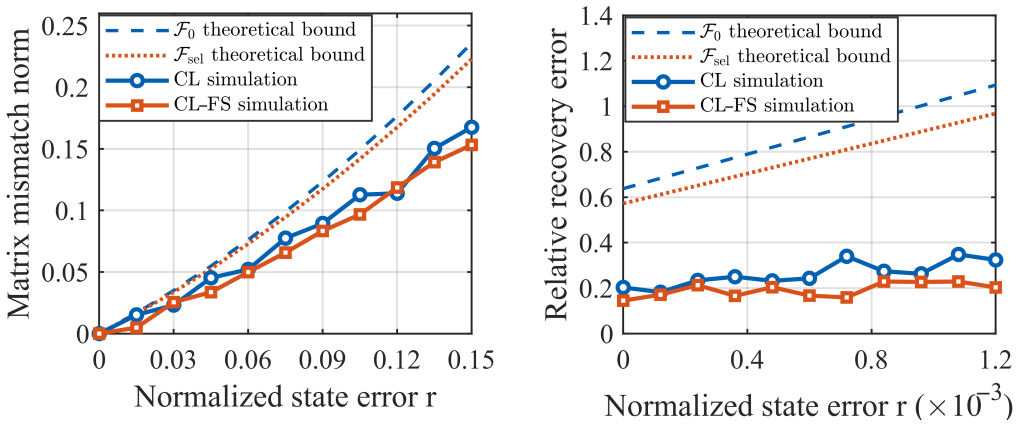


Figure 4. Simulated results for the uniform and selected frequency sets, labeled CL and CL-FS, respectively, together with their corresponding theoretical bounds: (**a**) observation-matrix perturbation and (**b**) normalized recovery error. The horizontal variable r is a dimensionless, covariance-scaled state-error radius. Its relation to physical position-error limits is given in Equation (14). Different ranges of the covariance-scaled state-error radius r are used: $r \in [0, 0.15]$ in (**a**) and $r \in [0, 1.2 \times 10^{-3}]$ in (**b**). The narrower range in (**b**) is used to evaluate recovery errors near the conditional tolerances reported in Table 2.

Figure 5 compares CL and CL-FS at x-direction position-prediction errors of 0, 0.04, and 0.08 m. Echoes were generated using the true target state, whereas the predicted state and UAV position used for reconstruction were held fixed across the position-offset tests, without tracking updates. The images show the central 0.6 m × 0.6 m region, with true positions marked by red crosses. At a range of approximately 5 m, the angular separation between the scatterers is 1.83°. As the offset increases, the CL reconstructions show broader scattering responses and more artifacts, whereas the two responses remain distinguishable in the CL-FS reconstruction at an offset of 0.08 m. The simulation evaluates the separability of the 1.83° target pair on the 0.02 m grid.

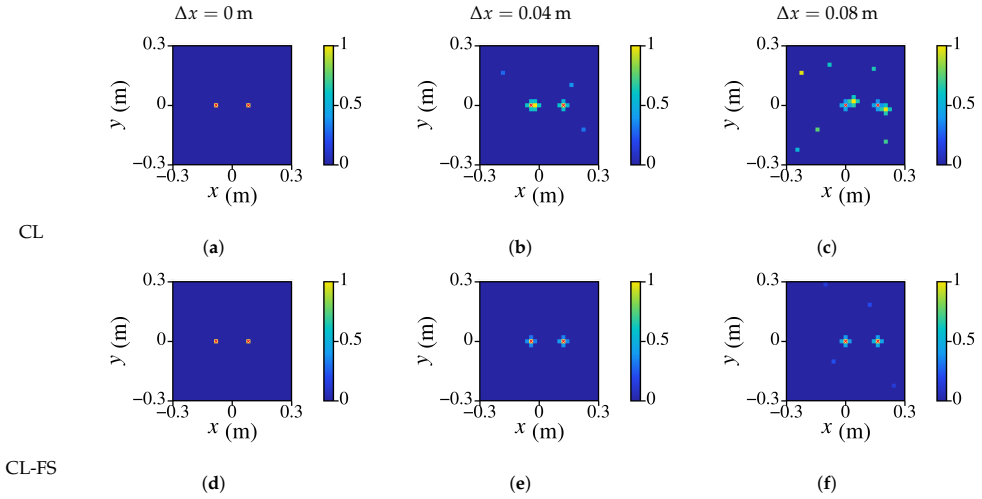


Figure 5. Two-target imaging results at a signal-to-noise ratio (SNR) of 10 dB under different x-direction position offsets, with $\Delta y = 0$: (**a**) CL with $\Delta x = 0$ m; (**b**) CL with $\Delta x = 0.04$ m; (**c**) CL with $\Delta x = 0.08$ m; (**d**) CL-FS with $\Delta x = 0$ m; (**e**) CL-FS with $\Delta x = 0.04$ m; and (**f**) CL-FS with $\Delta x = 0.08$ m. The red crosses indicate the true target positions.

To quantify two-target separability and isolate the effects of the selection criteria, Figure 6 compares five frequency sets under identical design and reconstruction settings. Three additional frequency sets were selected using alternative criteria: minimizing the worst-case radiation-field correlation (CL-Corr), maximizing the worst-case restricted minimum singular value $\alpha_{\mathcal{F}}$ (CL-Dist), and minimizing the worst-case mismatch bound $\overline{L}_{\mathcal{F}} r_* + \frac{1}{2} H_{\mathcal{F}} r_*^2$ with $r_* = 1$ (CL-Sens). The same feasibility constraints were applied to the frequency sets selected under all criteria. A trial is deemed successful if the two dominant peaks are matched one-to-one to the true targets within one grid spacing (0.02 m), and the valley intensity between the peaks is at least 3 dB below the weaker peak. Success rates were calculated from 100 independent noise realizations per offset at an SNR of 10 dB.

As the position offset increases, the success rate decreases more slowly for CL-FS than for the other methods. At an x-direction offset of 0.08 m, the success rate is 80% for CL-FS, compared with 20% for CL and 35–50% for the alternative selection criteria. The corresponding y-direction results are shown in Figure A2. The selected frequency set was then evaluated in closed-loop tracking using the extended-target configuration specified in Appendix C.2. Figure 7 compares five configurations using a common prescribed UAV trajectory. Each configuration uses its own predicted state for image reconstruction. OL-FS-M used independent position observations whose error covariance matched that of the image-derived centroid observations in CL-FS before the state update. CL-FS-C0 retained the calibrated observation bias and observation-error covariance of CL-FS but set $\mathbf{C}_j = \mathbf{0}$ in the state update. Calibration statistics were fixed before independent evaluation. Centroid root-mean-square error (RMSE) was calculated from updated estimates over all 150 frames

of 1000 evaluation trajectories, including skipped-update frames. Further settings are given in Appendix C.

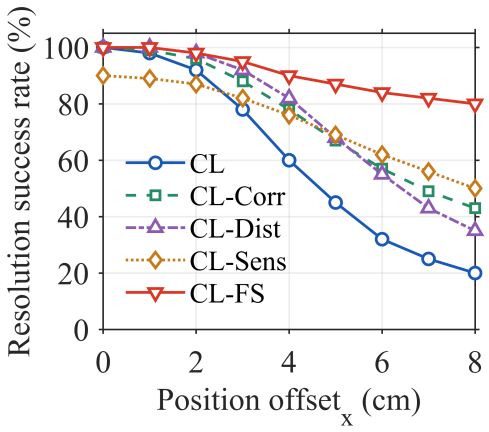


Figure 6. Two-target resolution success rates under x-direction offsets. CL and CL-FS use the uniform and proposed frequency sets, respectively. CL-Corr, CL-Dist, and CL-Sens use frequency sets selected by minimizing the worst-case field correlation, maximizing the worst-case restricted minimum singular value, and minimizing the worst-case mismatch bound, respectively. Success rates were evaluated at an SNR of 10 dB using 100 independent noise realizations for each offset.

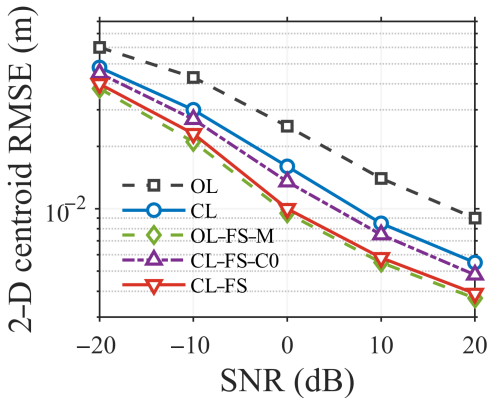


Figure 7. Two-dimensional centroid-estimation root-mean-square error (RMSE) versus signal-to-noise ratio (SNR) for five tracking configurations under a common prescribed unmanned aerial vehicle (UAV) trajectory. OL-FS-M denotes the open-loop reference using the selected frequency set and independent position observations with covariance matched to the pre-update image-centroid observation-error covariance of CL-FS. CL-FS-C0 retains the CL-FS calibration statistics but sets the cross-covariance to zero in the state update.

CL-FS yields lower centroid RMSE than CL and CL-FS-C0, while its RMSE remains slightly higher than that of OL-FS-M. Results with random acceleration disturbances are reported in Appendix G. The corresponding extended-target images obtained with OL, CL, and CL-FS at SNRs of 20, 0, and −20 dB are compared in Figure 8. As SNR decreases, the CL reconstructions become more fragmented, while background artifacts become more pronounced in the OL reconstructions. The main scattering region remains discernible in the CL-FS reconstruction at −20 dB, although fine details are degraded.

The comparison methods in Figures 9 and 10 were the consensus alternating direction method of multipliers (CADMM) [9] and efficient sparse Bayesian learning (ESBL) [24]. Both used $\mathcal{F}_{\text{sel}}$ and, where required, the CL-FS state predictions, so the comparisons primarily assess reconstruction performance under common conditions. Their adaptations are described in Appendix D.

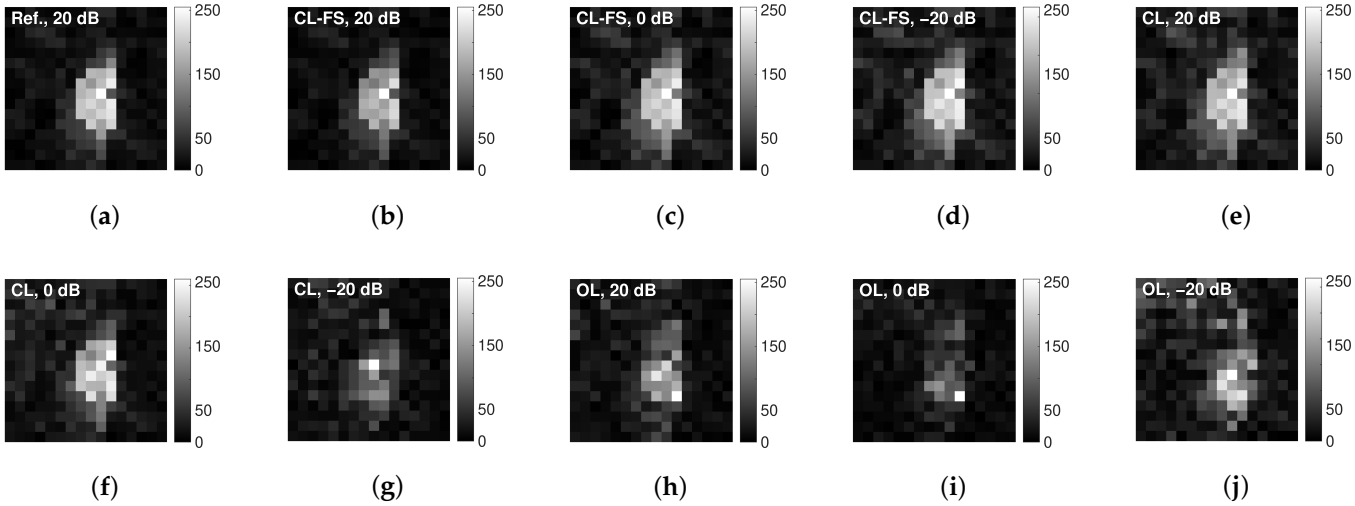


Figure 8. Stationary-target reference and moving-target reconstructions at different SNRs. (**a**) Stationary-target reference (labeled "Ref.") reconstructed with the uniform frequency set $\mathcal{F}_0$ at 20 dB SNR. Moving-target reconstructions at frame 75 are shown for CL-FS at (**b**) 20 dB, (**c**) 0 dB, and (**d**) −20 dB; CL at (**e**) 20 dB, (**f**) 0 dB, and (**g**) −20 dB; and OL at (**h**) 20 dB, (**i**) 0 dB, and (**j**) −20 dB. All panels display the central 0.32 m × 0.32 m region, whereas the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) reported in Figure 9 are calculated over the full 0.8 m × 0.8 m evaluation region. Gray levels indicate image intensity on a common 0–255 scale.

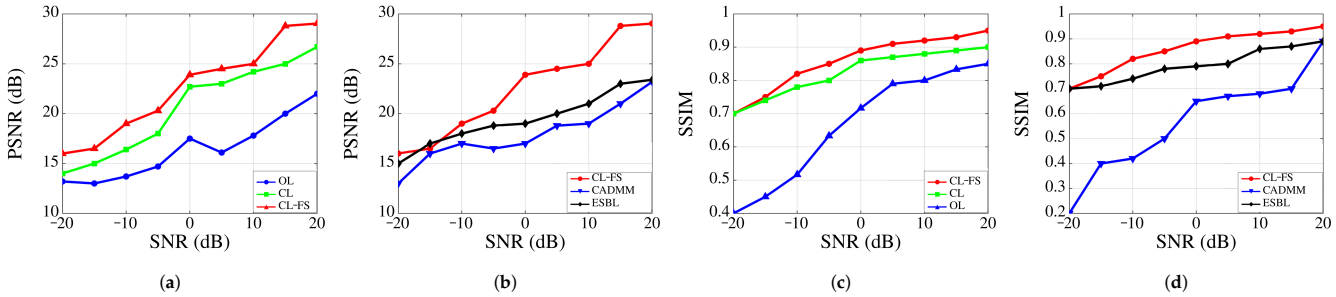


Figure 9. Peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) of moving-target reconstructions relative to the stationary-target reference reconstructed with $\mathcal{F}_0$ at an SNR of 20 dB and shown in Figure 8a. Metrics are evaluated over the full 0.8 m × 0.8 m region and averaged over all frames of 1000 independent simulations at each SNR: (**a**) PSNR for OL, CL, and CL-FS; (**b**) PSNR for CL-FS, CADMM, and ESBL; (**c**) SSIM for OL, CL, and CL-FS; and (**d**) SSIM for CL-FS, CADMM, and ESBL. CADMM [9] denotes the consensus alternating direction method of multipliers, and ESBL [24] denotes efficient sparse Bayesian learning. In (**b**) and (**d**), CL-FS, CADMM, and ESBL use the same frequency set $\mathcal{F}_{\text{sel}}$ and, where required, the CL-FS state predictions. These panels primarily compare reconstruction performance under these common conditions.

Figure 9 evaluates the agreement between the moving-target reconstructions and the stationary-target reference reconstruction in terms of the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). Figure 8 displays only the central 0.32 m × 0.32 m region for visualization, whereas both metrics in Figure 9 are calculated over the full 0.8 m × 0.8 m evaluation region specified in Appendix D. At each SNR, the metrics are averaged over all frames from 1000 independent simulations. The stationary-target image reconstructed using the uniform frequency set $\mathcal{F}_0$ at an SNR of 20 dB, as shown in Figure 8a, is used as the reference. Therefore, these metrics quantify agreement with the reference reconstruction rather than directly measuring the error relative to the underlying true scattering-coefficient distribution. The alignment procedure and metric settings are detailed in Appendix D. At 0 dB, CL-FS achieves a PSNR of approximately 24 dB and an SSIM of 0.89, compared with 22.5 dB and 0.86 for CL, respectively, indicating closer agreement

with the stationary-target reference. Compared with CADMM and ESBL, CL-FS generally achieves a higher PSNR, although ESBL achieves a slightly higher PSNR at −15 dB. CL-FS also achieves a higher SSIM at most tested SNRs, with an SSIM comparable to that of ESBL at −20 dB.

Figure 10 compares scatterer-localization RMSE calculated from matched detections across all frames of 100 independent simulations at each SNR. The RMSE is expressed in decibels relative to 0.1 m. Its definition and matching procedure are given in Appendix D. For matched detections, CL-FS yields lower localization RMSE than CL, OL, CADMM, and ESBL at moderate and high SNRs. Its RMSE decreases from approximately 0.0282 m at 0 dB to 0.00891 m at 20 dB.

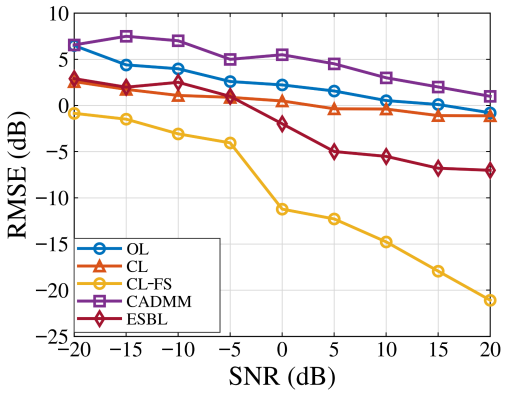


Figure 10. Scatterer-localization RMSE for matched detections versus SNR. The plotted values are expressed in decibels relative to 0.1 m as $\mathrm{RMSE}_{\mathrm{loc,dB}} = 20\log_{10}[\mathrm{RMSE}_{\mathrm{loc}}/(0.1\ \mathrm{m})]$. CADMM [9] denotes the consensus alternating direction method of multipliers, and ESBL [24] denotes efficient sparse Bayesian learning. CADMM and ESBL use $\mathcal{F}_{\mathrm{sel}}$ and, where required, the CL-FS state predictions. Their comparisons with CL-FS primarily assess reconstruction performance under these common conditions.

Table 3 compares uniform, correlation-based, and proposed frequency selection for the extended target at 0 dB. All frequency sets were determined using the two-scatterer design and held fixed during evaluation. The CL and CL-FS values correspond to the 0 dB results in Figures 9 and 10; CL-Corr was evaluated under the same settings. Compared with uniform sampling, correlation-based selection increases PSNR, SSIM, and detection coverage and reduces localization RMSE. The proposed criterion provides further improvements in all four metrics.

Table 3. Extended-target performance at an SNR of 0 dB.

Method	Frequency-Selection Criterion	PSNR (dB)	SSIM	Localization RMSE (m)	Detection Coverage (%)
CL	Uniform spacing	22.5	0.860	0.1000	72
CL-Corr	Radiation-field correlation	23.3	0.875	0.0501	82
CL-FS	Recovery-error bound	24.0	0.890	0.0282	90

PSNR—peak signal-to-noise ratio; SSIM—structural similarity index; RMSE—root-mean-square error. CL, CL-Corr, and CL-FS denote closed-loop configurations using uniform, correlation-based, and proposed frequency selection, respectively.

4. Discussion

The simulation results show that the selected frequency set improves imaging performance in the presence of residual state-prediction errors, consistent with a design that jointly accounts for observation-matrix sensitivity and sparse-distribution distinguishability. The conditional recovery-error bound in (12) quantitatively links these two factors to scattering-coefficient recovery accuracy. The resulting state-error tolerance in (13) gives

the largest covariance-scaled radius certified by the recovery-error bound for a prescribed recovery-error threshold. The frequency set is then selected by maximizing the minimum tolerance over the prescribed observation geometries, as formulated in (15), thereby directly incorporating the recovery-accuracy requirement into the design while accounting for geometry-dependent variations in error tolerance.

The benefit of jointly accounting for mismatch sensitivity and sparse-distribution distinguishability is supported by the comparisons among frequency-selection criteria. As shown in Figure 6, a two-target resolution success rate of 80% is obtained with the proposed criterion at an SNR of 10 dB and an x-direction offset of 0.08 m, compared with 35–50% for the three alternative criteria. Further evidence is provided by the additional frequency-set comparison in Appendix F. Although the random frequency set $\mathcal{F}_{\mathrm{R1}}$ has a lower peak field-correlation coefficient than the selected set, a resolution success rate of only 35% is obtained under the same conditions (Table A1). This comparison indicates that lower field correlation alone does not ensure better resolution in the presence of state-prediction errors. The observed improvement can be interpreted through the recovery-error analysis: relative to the uniform frequency set, the selected set has smaller mismatch-bound coefficients and a larger restricted minimum singular value (Table 2), both of which contribute to a lower recovery-error bound. This interpretation is consistent with the matrix-perturbation and recovery-error results in Figure 4. Together, these findings support the use of the recovery-error bound as a common basis for jointly evaluating mismatch sensitivity and sparse-distribution distinguishability in frequency-set selection.

Under the sparsity, support, noise, residual, and local perturbation conditions stated in Section 2.2.2, the derived state-error tolerance defines a range within which the normalized scattering-coefficient recovery error does not exceed the prescribed threshold. At position offsets exceeding the corresponding position-error limits reported in Table 2, the selected frequency set still yields higher two-target resolution success rates than uniform sampling, as shown in Figure 6. These results provide numerical evidence that tolerance-based frequency selection remains effective at the evaluated larger offsets, while the theoretical recovery-error guarantee remains subject to its stated conditions. Further evidence is provided by the extended-target results in Table 3, where improvements in image quality and localization accuracy are obtained using the frequency set designed for the two-scatterer configuration without reoptimization. These results support the applicability of the selected frequency set to the evaluated extended target, which was not used in the frequency-set design.

The benefits of frequency-set design are also observed in closed-loop state estimation using image feedback. Under a common prescribed UAV trajectory, lower centroid RMSE is obtained with CL-FS than with CL (Figure 7), indicating that closed-loop localization accuracy can be improved through tolerance-based frequency selection. In this feedback process, the observation matrix used for image reconstruction is formed from the predicted state, so image-derived observation errors may be statistically correlated with state-prediction errors. The effect of accounting for this correlation is examined by comparing CL-FS with CL-FS-C0, in which the same calibration statistics are retained but the cross-covariance is set to zero in the state update. The lower centroid RMSE of CL-FS relative to CL-FS-C0 supports accounting for this cross-covariance in the state update. Lower centroid RMSE is also obtained with CL-FS than with both CL and CL-FS-C0 under the random acceleration disturbances in Figure A3, providing further evidence of the benefits of frequency selection and cross-covariance correction under the evaluated motion disturbances.

These improvements are obtained while the metasurface geometry, aperture layout, number of operating frequencies, and transmitted energy per frame are held fixed, indicating that imaging and localization performance can be improved through frequency-set

selection under the same hardware and acquisition constraints. The frequency set is selected offline and remains fixed throughout tracking and imaging, eliminating the need for a per-frame frequency search. As shown by the complexity analysis in Appendix E, CL-FS and CL have the same asymptotic online computational complexity when identical matrix dimensions, implementations, and iteration counts are used. The additional computational cost is therefore incurred during offline frequency-set evaluation and search. In particular, support combinations must be enumerated to evaluate the restricted minimum singular value exactly, increasing the offline computational burden as the admissible support set or assumed sparsity level grows. Experimental assessment using measured aperture responses and moving-target echoes is still needed to establish performance in a physical system and to examine how aperture-response deviations and transmit–receive chain errors affect frequency selection and closed-loop imaging.

5. Conclusions

Low radiation-field correlation alone does not ensure robust reconstruction in the presence of residual state-prediction errors. Observation-matrix sensitivity and the distinguishability of sparse scattering distributions jointly influence the recovery-error bound. Their combined assessment through a conditional recovery-error bound provides a basis for selecting frequency sets that better tolerate prediction errors. In the evaluated scenarios, improved two-target resolution, image quality, and localization accuracy were obtained relative to uniform frequency sampling, with benefits also observed for an extended target not used in frequency selection. These results indicate that frequency-set design can complement motion compensation by reducing the imaging degradation caused by residual state-prediction errors.

Author Contributions: Conceptualization, S.-Y.W. and C.M.; methodology, S.-Y.W. and C.M.; software, S.-Y.W.; validation, S.-Y.W.; formal analysis, S.-Y.W.; investigation, S.-Y.W.; resources, C.M., S.-S.Q., Y.M., Y.-Y.Z. and W.W.; writing—original draft preparation, S.-Y.W.; writing—review and editing, C.M., S.-S.Q., Y.M., Y.-Y.Z. and W.W.; visualization, S.-Y.W.; supervision, C.M. and W.W.; funding acquisition, C.M., S.-S.Q. and W.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant numbers 62271257, 62301254, and 62171464; the Natural Science Foundation of Jiangsu Province, grant number BK20230916; and the Fundamental Research Funds for the Central Universities, grant number 30925010602.

Data Availability Statement: The simulation data and code supporting the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: During the preparation of this work, the authors used Google Gemini 1.5 Pro to improve the English language and refine sentence structure. The authors subsequently reviewed and edited the material and take full responsibility for the final content.

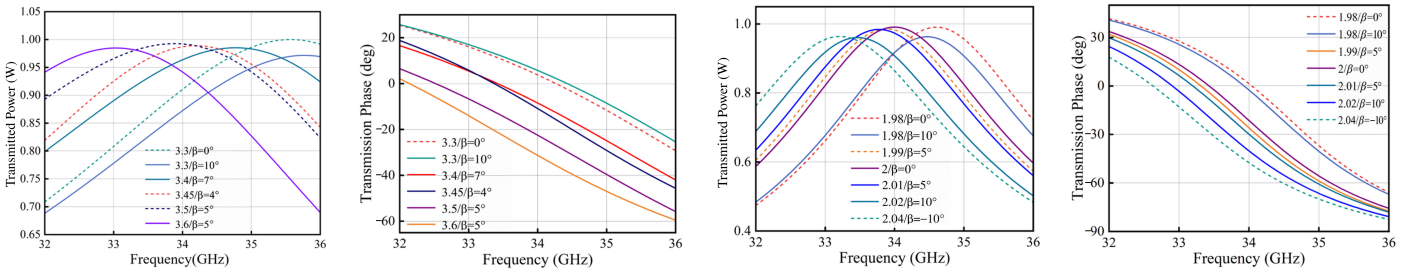
Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Metasurface Response and Position Derivatives

Appendix A.1. Unit Responses and Fixed Aperture

Four metal patterns were scaled and rotated within a 5×5 mm^2 footprint on a 0.787-mm-thick Rogers RT/duroid 5880 substrate (0.5-oz copper; minimum gap of 0.2 mm). Following neural network screening for high transmission and phase diversity, the transmission responses of 24 selected geometries were evaluated using Ansys High-Frequency Structure Simulator

(HFSS), version 2021 R1, with periodic boundary conditions over 31.5–36.5 GHz. Their complex responses $\mathfrak{t}_\chi(f) = S_{21,\chi}(f)$, $\chi = 1, \ldots, 24$, were used at all 81 candidate frequencies.



(**a**) (**b**) (**c**) (**d**)

Figure A1. Ansys High-Frequency Structure Simulator (HFSS) transmission responses: (**a**,**b**) unit type 1; (**c**,**d**) unit type 2. (**a**,**c**) Transmitted power; (**b**,**d**) transmission phase. Legend entries indicate the unit size in millimeters and the in-plane rotation angle β.

Figure A1 shows the unit transmission responses. For power-normalized ports and $P_{\text{inc}} = 1$ W, $P_{\text{tr},\chi}(f) = P_{\text{inc}}|\mathfrak{t}_\chi(f)|^2$.

A fixed-aperture layout was obtained using simulated annealing [25] to minimize the maximum absolute Pearson correlation between the spatial power patterns of the receiving responses. The same layout was used for all frequency-selection methods. This power-pattern metric differs from the complex-field correlation used in the frequency-set constraint.

Appendix A.2. Equivalent Receiving Response

With the imaging plane at $z = 0$, write $\bar{\mathbf{r}} = [\mathbf{r}^{\mathsf{T}}, 0]^{\mathsf{T}}$ and $\mathbf{o}_{\text{MS}}(t) = [\mathbf{u}_{\parallel}^{\mathsf{T}}(t), z_0]^{\mathsf{T}}$. The known aperture orientation is specified by orthonormal tangent vectors $\mathbf{e}_{\text{MS},1}(t)$ and $\mathbf{e}_{\text{MS},2}(t)$. Unit positions are

$$\mathbf{p}_n^{\text{MS}}(t) = \mathbf{o}_{\text{MS}}(t) + d_{\text{MS}}\big[(n_1 - 13)\mathbf{e}_{\text{MS},1}(t) + (n_2 - 13)\mathbf{e}_{\text{MS},2}(t)\big], \tag{A1}$$

where $n_1, n_2 \in \{1, \ldots, 25\}$, $n = 25(n_1 - 1) + n_2$, and $d_{\text{MS}} = 5$ mm is the element pitch.

Let $d_n^{\text{rx}} = \|\bar{\mathbf{r}} - \mathbf{p}_n^{\text{MS}}(t)\|_2$, $k_{\text{em}}(f) = 2\pi f/c$, and $\mathfrak{g}_f(d) = e^{-ik_{\text{em}}(f)d}/(4\pi d)$. Under line-of-sight propagation and single scattering, scalar superposition gives [26]

$$\Psi_j(\mathbf{r}, f_p; t) = \sum_{n=1}^{N_{\text{MS}}} v_n(f_p)\mathfrak{g}_{f_p}(d_n^{\text{rx}}), \qquad v_n(f_p) = \zeta_n(f_p)\mathfrak{t}_{\chi_n}(f_p), \tag{A2}$$

where $N_{\text{MS}} = 625$ is the number of metasurface elements, χ_n identifies the geometry assigned to element n, and ζ_n represents the transfer from element n to the receiver. At fixed polarization, the scalar aperture model used $\zeta_n = 1$ and neglected incidence-angle dependence and coupling between unlike units. The same free-space factor was used on the transmit path.

Unit-dependent carrier phases were retained, while the common envelope delay in (1) was referenced to the aperture center. The inter-unit delay spread was assumed small relative to the envelope variation time. The unit transmission and receiver transfer responses were approximated as constant at their carrier-frequency values within the full-width-at-half-maximum bandwidth of each Gaussian pulse's power spectrum, approximately 0.22 GHz.

Appendix A.3. Position Derivatives

In the following derivations, time, UAV position, and aperture orientation are held fixed, and ∇ denotes differentiation with respect to $\mathbf{r}$. For any distance $d = \sqrt{\|\mathbf{r} - \mathbf{p}_{\parallel}\|_2^2 + p_z^2}$, define $\mathbf{q} = (\mathbf{r} - \mathbf{p}_{\parallel})/d$. Then $\nabla d = \mathbf{q}$ and $\nabla^2 d = (\mathbf{I}_2 - \mathbf{q}\mathbf{q}^{\mathsf{T}})/d$.

The first radial derivative is $\mathfrak{g}'_f(d) = (-ik_{\mathrm{em}} - d^{-1})\mathfrak{g}_f(d)$, and the second is $\mathfrak{g}''_f(d) = (-k_{\mathrm{em}}^2 + 2ik_{\mathrm{em}}d^{-1} + 2d^{-2})\mathfrak{g}_f(d)$. Since the unit coefficients are position independent,

$$\begin{aligned}\nabla\Psi_j &= \sum_n v_n \mathfrak{g}'_{f_p}(d_n^{\mathrm{rx}})\nabla d_n^{\mathrm{rx}},\\ \nabla^2\Psi_j &= \sum_n v_n \left[\mathfrak{g}''_{f_p}(d_n^{\mathrm{rx}})\nabla d_n^{\mathrm{rx}}(\nabla d_n^{\mathrm{rx}})^{\mathsf{T}} + \mathfrak{g}'_{f_p}(d_n^{\mathrm{rx}})\nabla^2 d_n^{\mathrm{rx}}\right].\end{aligned} \tag{A3}$$

Applying the same differentiation rules yields the derivatives of $g_T = \mathfrak{g}_{f_p}(R)$ and $\tau = 2R/c$, where $R = \sqrt{\|\mathbf{r} - \mathbf{u}_\parallel(t)\|_2^2 + z_0^2}$.

For an entry of the observation matrix, set $b = g_T\Psi_j$ and $A = A_{j,p}(t-\tau)$, so that $h_{j,p,l} = bA$. Primes on A denote derivatives with respect to its time argument. The product and chain rules give $\nabla h = A\nabla b - bA'\nabla\tau$ and

$$\nabla^2 h = A\nabla^2 b - A'\left[\nabla b(\nabla\tau)^{\mathsf{T}} + \nabla\tau(\nabla b)^{\mathsf{T}}\right] + bA''\nabla\tau(\nabla\tau)^{\mathsf{T}} - bA'\nabla^2\tau, \tag{A4}$$

where $\nabla b = \Psi_j\nabla g_T + g_T\nabla\Psi_j$ and

$$\nabla^2 b = \Psi_j\nabla^2 g_T + g_T\nabla^2\Psi_j + \nabla g_T(\nabla\Psi_j)^{\mathsf{T}} + \nabla\Psi_j(\nabla g_T)^{\mathsf{T}}.$$

Since $\partial\mathbf{r}_m/\partial\mathbf{s}^{\mathsf{T}} = \mathbf{T}(t - t_j)$, (8) follows directly.

Appendix B. Mismatch and Recovery-Error Bounds

Appendix B.1. Observation-Matrix Mismatch

For fixed j and $\mathcal{F}$, define $\mathbf{A}_{\mathcal{F}}(\mathbf{u}) = \mathbf{\Phi}_{j,\mathcal{F}}(\widehat{\mathbf{s}}_j^{-} + \mathbf{L}_j\mathbf{u})$ on $\|\mathbf{u}\|_2 \le r_0$. Its derivative at zero is $\mathcal{D}_{\mathcal{F}}[\mathbf{v}] = \sum_{a=1}^{6}[\mathbf{L}_j\mathbf{v}]_a\mathbf{D}_{j,\mathcal{F}}^{(a)}$. The first-order coefficient and a second-order bound satisfy

$$L_{\mathcal{F}} = \sup_{\|\mathbf{v}\|_2=1}\|\mathcal{D}_{\mathcal{F}}[\mathbf{v}]\|_2, \qquad H_{\mathcal{F}} \ge \sup_{\substack{\|\mathbf{u}\|_2\le r_0\\ \|\mathbf{v}\|_2=1}}\|D^2\mathbf{A}_{\mathcal{F}}(\mathbf{u})[\mathbf{v},\mathbf{v}]\|_2, \tag{B1}$$

where the directions are real. For twice continuously differentiable entries, Taylor's theorem with integral remainder gives, for $\|\mathbf{u}\|_2 \le r \le r_0$,

$$\|\Delta\mathbf{\Phi}_{j,\mathcal{F}}\|_2 \le L_{\mathcal{F}}\|\mathbf{u}\|_2 + \int_0^1 (1-\lambda)H_{\mathcal{F}}\|\mathbf{u}\|_2^2\,\mathrm{d}\lambda \le L_{\mathcal{F}}r + \tfrac{1}{2}H_{\mathcal{F}}r^2, \tag{B2}$$

which establishes (10). The affine motion model gives the entrywise Hessian $\nabla_{\mathbf{u}}^2 h = \mathbf{L}_j^{\mathsf{T}}\mathbf{T}^{\mathsf{T}}(\delta)\nabla_{\mathbf{r}}^2 h\,\mathbf{T}(\delta)\mathbf{L}_j$, where $\delta = t_{j,p,l} - t_j$. Placing each value of $\mathbf{v}^{\mathsf{T}}\nabla_{\mathbf{u}}^2 h\mathbf{v}$ at the corresponding position in the observation matrix gives $D^2\mathbf{A}_{\mathcal{F}}[\mathbf{v},\mathbf{v}]$. For conservative evaluation, let $\mathbf{B}_b = \sum_a[\mathbf{L}_j]_{ab}\mathbf{D}_{j,\mathcal{F}}^{(a)}$, and let $\kappa_{p,l,m}$ bound the corresponding $\|\nabla_{\mathbf{u}}^2 h\|_2$ throughout the error set $\|\mathbf{u}\|_2 \le r_0$. A conservative upper bound for $L_{\mathcal{F}}$ and a valid choice of $H_{\mathcal{F}}$ are then given by

$$L_{\mathcal{F}} \le \overline{L}_{\mathcal{F}} := \left(\sum_{b=1}^{6}\|\mathbf{B}_b\|_{\mathrm{F}}^2\right)^{1/2}, \qquad H_{\mathcal{F}} = \left(\sum_{p:f_p\in\mathcal{F}}\sum_{l=1}^{L}\sum_{m=1}^{M}\kappa_{p,l,m}^2\right)^{1/2} \tag{B3}$$

where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm. The conservative coefficient $\overline{L}_{\mathcal{F}}$ was used in both the mismatch-bound and tolerance calculations. At each sample and grid point, the position variation over $\|\mathbf{u}\|_2 \le r_0$ is at most $r_0\|\mathbf{T}(\delta)\mathbf{L}_j\|_2$. Substitution of the propagation-distance and pulse-envelope derivative bounds over this region into the expressions in Appendix A.3 gives $\kappa_{p,l,m}$.

Appendix B.2. Recovery-Error and State-Error Tolerance

Under the conditions in Section 2.2.2, define $\mathbf{e} = \widehat{\mathbf{x}}_j - \mathbf{x}_j$. The echo model (7) gives $\widehat{\mathbf{\Phi}}_{j,\mathcal{F}}\mathbf{e} = \left(\widehat{\mathbf{\Phi}}_{j,\mathcal{F}}\widehat{\mathbf{x}}_j - \mathbf{y}_{j,\mathcal{F}}\right) + \Delta\mathbf{\Phi}_{j,\mathcal{F}}\mathbf{x}_j + \mathbf{n}_{j,\mathcal{F}}$. Since $\mathbf{e}$ is supported on $\mathcal{I}_\omega$ and has at most $2K$ nonzero entries, the definition of $\alpha_\mathcal{F}$, the triangle inequality, and (10) give

$$\alpha_\mathcal{F}\|\mathbf{e}\|_2 \le \|\widehat{\mathbf{\Phi}}_{j,\mathcal{F}}\mathbf{e}\|_2 \le \varepsilon_{\mathrm{rec},\mathcal{F}} + \varepsilon_{n,\mathcal{F}} + (L_\mathcal{F} r + \tfrac{1}{2}H_\mathcal{F} r^2)\|\mathbf{x}_j\|_2. \tag{B4}$$

Dividing by $\alpha_\mathcal{F}\|\mathbf{x}_j\|_2$ and using $\|\mathbf{x}_j\|_2 \ge x_{\min} > 0$ establishes (12).

For $m_\mathcal{F} > 0$ and $H_\mathcal{F} > 0$, the left-hand side of $L_\mathcal{F} r + \frac{1}{2}H_\mathcal{F} r^2 \le m_\mathcal{F}$ is nondecreasing for $r \ge 0$. Solving the corresponding quadratic equation for its positive root and taking the minimum of that root and r_0 gives (13). For $m_\mathcal{F} > 0$ and $H_\mathcal{F} = 0$, the tolerance is $\min\{r_0, m_\mathcal{F}/L_\mathcal{F}\}$ when $L_\mathcal{F} > 0$, and r_0 when $L_\mathcal{F} = 0$.

Finally, $\delta\mathbf{r}_j(t) = \mathbf{T}(t - t_j)\mathbf{L}_j\mathbf{u}$ implies $\sup_{\|\mathbf{u}\|_2 \le r_{\mathrm{safe}}} |\mathbf{d}^\mathsf{T}\delta\mathbf{r}_j(t)| = r_{\mathrm{safe}}\|\mathbf{L}_j^\mathsf{T}\mathbf{T}^\mathsf{T}(t - t_j)\mathbf{d}\|_2$. Applying Cauchy–Schwarz and substituting $\mathbf{L}_j\mathbf{L}_j^\mathsf{T} = \mathbf{P}_j^-$ yields (14).

Appendix C. Simulation and Calibration Settings

Appendix C.1. Offline Frequency Selection

The selected frequency set was $\mathcal{F}_{\mathrm{sel}} = \{(32 + 0.05q)\ \mathrm{GHz} : q \in \mathcal{Q}_{\mathrm{sel}}\}$, where $\mathcal{Q}_{\mathrm{sel}} = \{0, 5, 9, 12, 17, 20, 24, 29, 32, 36, 40, 43, 47, 51, 55, 59, 63, 68, 72, 76, 80\}$. UAV offsets relative to the predicted centroid were sampled separately along the x and y axes, from -1.5 to 1.5 m in 0.1 m increments. For each geometry, the admissible support set comprised the 3×3 grid neighborhoods centered at the two scatterers.

The conditional tolerance was evaluated with $K = 2$, using the same first-frame prediction covariance $\mathbf{P}_1^-$ for all geometries and the initialization and process-noise settings specified below. The first-frame covariance was adopted because it is determined by the prescribed initialization and motion model before image-derived feedback becomes available, thereby providing a consistent uncertainty model for comparing candidate frequency sets. During online tracking, the prediction covariance $\mathbf{P}_j^-$ evolves through successive prediction and observation updates, while the selected frequency set $\mathcal{F}_{\mathrm{sel}}$ is kept fixed. Therefore, changes in the online covariance do not initiate a new frequency-set search. The offline tolerance values characterize the prescribed design conditions, and the performance of the fixed frequency set under evolving closed-loop conditions is evaluated through the tracking and imaging simulations. The offline calculation used $K = 2$, $r_0 = 2$, $\rho_{\max} = 0.95$, and $x_{\min} = \sqrt{2}$. At an SNR of 10 dB, the noise vector for each frequency set contained $LN_F = 1344$ independent circularly symmetric complex Gaussian samples with per-sample variance σ_n^2. Let $\mathbf{y}_{\mathrm{ref}}$ denote the noiseless reference echo obtained using $\mathcal{F}_0$ for the two-scatterer target at the first frame with zero position offset. The 99th-percentile noise-norm bound was $\varepsilon_{n,\mathcal{F}} = 37.825\sigma_n$. For the conditional tolerance calculation, the noise and reconstruction-residual budgets were prescribed to satisfy

$$\sigma_n = \frac{\|\mathbf{y}_{\mathrm{ref}}\|_2}{\sqrt{1344 \times 10}}, \qquad \varepsilon_{n,\mathcal{F}} = 37.825\sigma_n,$$

$$\varepsilon_{\mathrm{rec},\mathcal{F}} = 0.001800\sqrt{2} - 37.825\sigma_n > 0.$$

The recovery-error bound applies when the actual noise norm and reconstruction residual satisfy their respective budgets.

For each frequency set and geometry, $\alpha_{\mathcal{F}}$ was evaluated over all $\binom{18}{4}$ four-column submatrices using (C1). In general, with $s = \min(2K, |\mathcal{I}_\omega|)$,

$$\alpha_{\mathcal{F}}^2 = \min_{\substack{\mathcal{S}\subseteq\mathcal{I}_\omega \\ |\mathcal{S}|=s}} \lambda_{\min}\Big(\widehat{\mathbf{\Phi}}_{j,\mathcal{F},\mathcal{S}}^{\mathsf{H}}\widehat{\mathbf{\Phi}}_{j,\mathcal{F},\mathcal{S}}\Big). \tag{C1}$$

Here, $\mathcal{S}$ indexes a column subset and $\lambda_{\min}$ denotes the smallest eigenvalue. The conservative coefficient $\overline{L}_{\mathcal{F}}$ in (B3) was substituted for $L_{\mathcal{F}}$ in frequency selection and tolerance evaluation. The Hessian bounds used to form $H_{\mathcal{F}}$ were obtained over $\|\mathbf{u}\|_2 \le r_0$ using the distance and pulse-envelope derivative bounds described in Appendix B. The proposed criterion and the three alternative criteria (CL-Corr, CL-Dist, and CL-Sens) were optimized using the same single-frequency exchange procedure, initialized with $\mathcal{F}_0$. At each iteration, all exchanges of one selected interior frequency with one unselected candidate were evaluated, with the band endpoints fixed at 32 and 36 GHz. The feasible exchange providing the largest improvement in the corresponding objective was accepted. The search was terminated when a complete scan yielded no feasible improving exchange. The same scenario set and feasibility constraints were used for all four criteria.

For Figure 4, a fixed-direction simulation was performed at the zero-offset geometry ω_{rep} using the first-frame prediction covariance. The state perturbation was specified as

$$\delta\mathbf{s} = \mathbf{L}_1 r\mathbf{v}, \qquad \mathbf{v} = [1,0,0,0,0,0]^{\mathsf{T}}, \qquad \mathbf{L}_1 = (\mathbf{P}_1^-)^{1/2}.$$

The observation-matrix perturbation and normalized recovery error were calculated as

$$\big\|\mathbf{\Phi}_{1,\mathcal{F}}(\widehat{\mathbf{s}}_1^- + \delta\mathbf{s}) - \mathbf{\Phi}_{1,\mathcal{F}}(\widehat{\mathbf{s}}_1^-)\big\|_2 \quad \text{and} \quad \frac{\|\widehat{\mathbf{x}}_1 - \mathbf{x}_1\|_2}{\|\mathbf{x}_1\|_2}.$$

A single trial was performed at each r, with the same state perturbation and complex noise vector used for both frequency sets. For the recovery-error evaluation in Figure 4b, only the two largest-magnitude SBL coefficients within the prescribed 18-point support set were retained, and all other coefficients were set to zero. Both the reconstruction residual $\|\widehat{\mathbf{\Phi}}_{1,\mathcal{F}}\widehat{\mathbf{x}}_1 - \mathbf{y}_{1,\mathcal{F}}\|_2$ and the normalized recovery error were calculated using this truncated estimate. The true and truncated coefficient vectors therefore each had at most two nonzero entries within the prescribed support set. The actual noise norm and reconstruction residual satisfied their respective budgets. This truncation was used only for the conditional recovery-bound evaluation in Figure 4b; the subsequent imaging and tracking evaluations used full-grid SBL reconstruction.

Appendix C.2. Tracking and Imaging

The Gaussian pulse envelope was evaluated from 0 to 8 ns relative to the start of each carrier transmission, with its peak at 4 ns. For OL, CL, and CL-FS, the diagonal prior-covariance matrix $\mathbf{\Gamma}_j$ in SBL was initialized as $\mathbf{\Gamma}_j^{(0)} = \mathbf{I}_M$. Iterations terminated when the relative change in its diagonal entries fell below 10^{-4} or the iteration count reached 200.

The initial state estimate was generated as $\widehat{\mathbf{s}}_0^+ = \mathbf{s}_0 + \mathbf{e}_0$, where $\mathbf{e}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_0^+)$ and $\mathbf{P}_0^+ = \text{diag}(0.04^2, 0.04^2, 0.20^2, 0.20^2, 0.10^2, 0.10^2)$. In the baseline tracking and imaging simulations, the true target followed a constant-acceleration trajectory, and $\mathbf{Q}_j = \mathbf{0}$ was used by the filters. For OL, the simulated position observation was $\mathbf{z}_j^{\text{OL}} = \mathbf{H}\mathbf{s}_j + \mathbf{v}_j^{\text{OL}}$, where $\mathbf{v}_j^{\text{OL}} \sim \mathcal{N}(\mathbf{0}, \sigma_{\text{trk}}^2\mathbf{I}_2)$ was independent of the state-prediction error and echo noise. The variance in each coordinate was $\sigma_{\text{trk}}^2 = (0.01\text{ m})^2 + (0.1\text{ m})^2\, 10^{-\text{SNR}/10}$. For OL-FS-M, independent position observations had zero-mean errors with covariance matched to

the calibrated pre-update image-centroid observation-error covariance of CL-FS at the corresponding SNR.

The intensity threshold used for centroid extraction in CL and CL-FS and for scatterer detection in the localization evaluation was set to $\eta_j = \max\{0.1 \max_m I_{j,m}, \eta_{\text{det}}\}$. At each SNR, η_{det} was the 99th percentile of the maximum intensities in 1000 noise-only calibration images and was held fixed during evaluation. For each of CL and CL-FS, 1000 independent calibration trajectories were generated at each SNR. At each frame, only trajectories with $\mathcal{M}_j \neq \varnothing$ contributed to the estimates of $\bar{\boldsymbol{\epsilon}}_j$, $\mathbf{R}_j$, and $\mathbf{C}_j$. These estimates were used to update the state estimates in the calibration trajectories. The resulting frame-dependent calibration statistics were then held fixed during the independent evaluation simulations. Trajectories with $\mathcal{M}_j = \varnothing$ followed the skipped-update rule in Section 2.1.3.

For the extended-target simulations, 18 scatterers were fixed relative to the target centroid. At time t, their positions were $\mathbf{c}(t) + 0.02 \text{ m } [a, b]^{\mathsf{T}}$, where $a \in \{-1, 0, 1\}$ and $b \in \{-4, -3, -2, 2, 3, 4\}$. Their fixed coefficients were $x_{a,b} = \exp[i\pi(a + 2b)/4]$. Their equal magnitudes and symmetric positions place the energy centroid at $\mathbf{c}(t)$.

Appendix D. Reference-Method Adaptations and Performance Evaluation

For CADMM, the complex echoes and predicted observation matrices were partitioned by carrier frequency, and image reconstruction followed the CADMM procedure in [9]. Its ISAR-specific preprocessing and orthogonal matching pursuit (OMP)-based motion-phase estimation were omitted. For ESBL, the sparse Bayesian coefficient and hyperparameter updates of [24] were applied to the complex-observation model; its high-resolution range profile (HRRP)-specific target separation and velocity estimation were omitted.

For the calculation of PSNR and SSIM, the stationary-target image reconstructed using $\mathcal{F}_0$ at an SNR of 20 dB, as shown in Figure 8a, was adopted as the common reference. Each reconstructed image was bilinearly interpolated onto a grid covering a 0.8 m $\times$ 0.8 m region centered at the true target position, with values outside the reconstruction grid set to zero. Both metrics were calculated over this full evaluation region rather than the central region displayed in Figure 8. A common intensity scale was adopted, with the peak value of the reference image set to 255. SSIM was calculated using an 11 $\times$ 11 Gaussian window with a standard deviation of 1.5 pixels and stability constants $C_1 = (0.01 \times 255)^2$ and $C_2 = (0.03 \times 255)^2$. For all methods, local intensity maxima above η_j were retained as scatterer detections. The noise threshold η_{det} was calibrated separately for each method and SNR as in Appendix C, then fixed during evaluation. An estimated scatterer associated with grid point m was assigned the position $\hat{\mathbf{c}}_j^- + \boldsymbol{\xi}_m$. A true scatterer at grid index $m^\star$ had position $\mathbf{c}_j + \boldsymbol{\xi}_{m^\star}$. Estimated and true positions were matched one-to-one by minimizing the total squared distance, with the number of pairs equal to the smaller of the two set sizes and no distance gate applied. For N matched pairs, the localization RMSE is $\text{RMSE}_{\text{loc}} = \sqrt{N^{-1} \sum_{i=1}^{N} \|\hat{\mathbf{r}}_i - \mathbf{r}_i\|_2^2}$, where $\hat{\mathbf{r}}_i$ and $\mathbf{r}_i$ are the estimated and true positions. The plotted value is $\text{RMSE}_{\text{loc,dB}} = 20 \log_{10}(\text{RMSE}_{\text{loc}}/(0.1 \text{ m}))$. The RMSE was calculated over all matched pairs; unmatched positions and frames without detections made no contribution.

Detection coverage was evaluated using one-to-one matching that maximized the number of pairs with a separation no greater than $\tau_{\text{det}} = 0.04$ m, using the same threshold for all methods. For N_{run} trajectories of J frames with N_{sc} true scatterers per frame, Coverage $= 100\% \, (N_{\text{run}} J N_{\text{sc}})^{-1} \sum_{n=1}^{N_{\text{run}}} \sum_{j=1}^{J} D_{n,j}$, where $D_{n,j}$ is the number of admissible matched pairs in frame j of trajectory n and equals zero when no scatterers are detected. The denominator includes all true scatterers in all frames; $N_{\text{sc}} = 18$ for the extended target.

Appendix E. Computational Complexity

The computational complexity of dense-matrix operations is analyzed for a fixed six-dimensional state and a fixed metasurface with $N_{\mathrm{MS}} = 625$ elements. Because the element count is fixed, the cost of summing over the elements when evaluating responses and derivatives is included in the constant factors of the $\mathcal{O}(\cdot)$ notation. The resulting complexity estimates describe how computational cost scales with the number of frequencies, the number of samples per frequency, and the number of reconstruction grid points for this fixed hardware configuration. Let $Q = LN_F$ denote the number of observations per frame.

Appendix E.1. Offline Frequency Selection

For each scenario, with the number of metasurface elements fixed, the candidate responses and their derivatives at the specified positions are evaluated with a computational cost of $\mathcal{O}(N_c LM)$, while the pairwise field correlations are computed with a cost of $\mathcal{O}(N_c^2 M)$. Each exchange scan evaluates at most $N_F(N_c - N_F)$ candidates. With T_{ex} scans, including the final scan, and cost C_{eval} per combination and scenario, the total cost is

$$C_{\mathrm{off}} = \mathcal{O}\Big(|\Omega|\big[N_c LM + N_c^2 M + (1 + T_{\mathrm{ex}} N_F(N_c - N_F))C_{\mathrm{eval}}\big]\Big). \tag{E1}$$

The cost refers to frequency selection with a fixed metasurface and aperture layout. For $M_{\mathcal{I},\omega} = |\mathcal{I}_\omega|$ and $s = \min(2K, M_{\mathcal{I},\omega})$, exact evaluation of (C1) costs $\mathcal{O}\Big(\binom{M_{\mathcal{I},\omega}}{s}[Qs^2 + s^3]\Big)$. Once the entrywise Hessian bounds are available, assembling (B3) costs $\mathcal{O}(QM)$. The evaluation cost C_{eval} includes these operations, constraint checks, and tolerance calculation. The cost of establishing derivative bounds over the entire neighborhood is included in C_{eval}, or accounted for separately as preprocessing if these bounds are computed in advance.

Appendix E.2. Online Tracking and Imaging

Let T_{SBL} denote the number of SBL iterations.

With N_{MS} fixed, rebuilding the observation matrix costs $\mathcal{O}(QM)$ per frame. With observation-space SBL updates [21] and without forming the full posterior covariance, each iteration costs $\mathcal{O}(Q^2 M + Q^3)$. Centroid extraction costs $\mathcal{O}(M)$, and the fixed-dimensional state update costs $\mathcal{O}(1)$ when the required covariances are supplied. Hence, for $T_{\mathrm{SBL}} \geq 1$, $C_{\mathrm{on}} = \mathcal{O}\big(T_{\mathrm{SBL}}[Q^2 M + Q^3]\big)$. With calibration statistics obtained offline, CL and CL-FS have the same asymptotic online computational complexity for identical dimensions, implementations, and iteration counts; frequency selection incurs additional offline computational cost.

Appendix F. Additional Frequency-Selection Results

Appendix F.1. Resolution Under y-Direction Offsets

Figure A2 extends the comparison to y-direction offsets, with CL-FS achieving an 84% success rate at 0.08 m, compared with 30% for CL.

Appendix F.2. Resolution Performance Across Frequency Sets

The uniform set, the proposed set, and four additional frequency sets are compared in Table A1. Four additional sets were randomly sampled from the same candidate grid, each containing 21 distinct frequencies and both band endpoints. They are written as $\mathcal{F}_{\mathrm{R}i} = \{(32 + 0.05q)\ \mathrm{GHz} : q \in \mathcal{Q}_{\mathrm{R}i}\}$, $i = 1, \ldots, 4$, with the indices listed below. The peak field-correlation coefficients $\rho^{\Psi}_{\mathrm{peak}}$ for $\mathcal{F}_{\mathrm{R1}}$, $\mathcal{F}_{\mathrm{R2}}$, $\mathcal{F}_{\mathrm{R3}}$, and $\mathcal{F}_{\mathrm{R4}}$ were 0.0361, 0.0420, 0.0485, and 0.0556, respectively.

Table A1. Two-target resolution success rates for six frequency sets. Brackets indicate Wilson 95% confidence intervals calculated from 100 independent trials per entry.

Metric	$\mathcal{F}_0$	$\mathcal{F}_{R1}$	$\mathcal{F}_{R2}$	$\mathcal{F}_{R3}$	$\mathcal{F}_{R4}$	$\mathcal{F}_{sel}$
x-direction success rate (%)	20 [13.3, 28.9]	35 [26.4, 44.7]	45 [35.6, 54.8]	60 [50.2, 69.1]	50 [40.4, 59.6]	80 [71.1, 86.7]
y-direction success rate (%)	30 [21.9, 39.6]	42 [32.8, 51.8]	56 [46.2, 65.3]	52 [42.3, 61.5]	70 [60.4, 78.1]	84 [75.6, 89.9]

Index set	Frequency indices
$\mathcal{Q}_{R1}$	$\{0, 1, 2, 3, 7, 8, 27, 29, 30, 41, 49, 50, 52, 62, 64, 70, 72, 74, 75, 79, 80\}$
$\mathcal{Q}_{R2}$	$\{0, 2, 7, 8, 12, 14, 15, 16, 18, 19, 20, 22, 23, 33, 34, 40, 46, 47, 69, 76, 80\}$
$\mathcal{Q}_{R3}$	$\{0, 2, 11, 13, 14, 20, 21, 27, 28, 36, 39, 43, 44, 50, 51, 55, 60, 64, 73, 74, 80\}$
$\mathcal{Q}_{R4}$	$\{0, 4, 5, 6, 10, 11, 23, 31, 33, 34, 35, 39, 41, 42, 47, 48, 68, 71, 77, 79, 80\}$

Resolution was evaluated at an SNR of 10 dB with a 0.08 m offset applied separately along the x and y axes. Each entry was obtained from 100 independent simulation trials, with Wilson 95% confidence intervals calculated from the success counts.

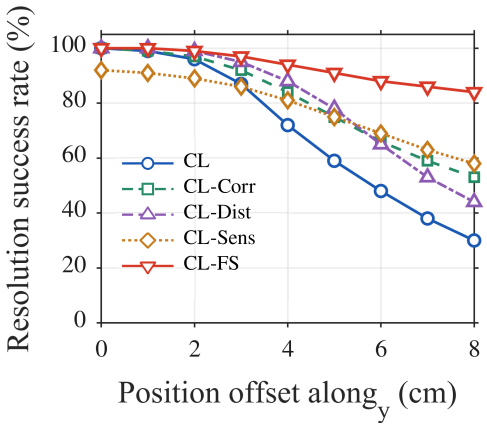


Figure A2. Two-target resolution success rates under y-direction offsets. Success rates were evaluated at an SNR of 10 dB using 100 independent noise realizations for each offset. The frequency-selection methods are defined in the caption of Figure 6.

Appendix G. Tracking Under Random Acceleration Disturbances

Tracking performance for the extended target was evaluated at an SNR of 0 dB by adding random acceleration disturbances $\mathbf{w}_j = \mathbf{G}_j\boldsymbol{\xi}_j$ to the true state transition in Section 2.1.1. The increments $\boldsymbol{\xi}_j \sim \mathcal{N}(\mathbf{0}, \sigma_{\Delta a}^2\mathbf{I}_2)$ were independent across frames and trajectories, with $\sigma_{\Delta a} = 0.1$ m/s^2 per component. The filters for CL, CL-FS, and CL-FS-C0 all used the process-noise covariance $\mathbf{Q}_j = \sigma_{\Delta a}^2\mathbf{G}_j\mathbf{G}_j^{\mathsf{T}}$, where $\mathbf{G}_j = [\frac{1}{2}\Delta t_j^2\mathbf{I}_2, \Delta t_j\mathbf{I}_2, \mathbf{I}_2]^{\mathsf{T}}$. The target configuration, frequency sets, and other simulation settings were the same as those specified in Appendix C. All methods used a common prescribed UAV trajectory over 150 frames with an inter-frame interval of 0.01 s. The bias, observation-error covariance, and cross-covariance were calibrated on 1000 independent constant-acceleration trajectories without added motion disturbances, using filters with the same $\mathbf{Q}_j$, and then fixed for evaluation. CL-FS-C0 retained the CL-FS calibration but set $\mathbf{C}_j = \mathbf{0}$ in the update. For each frame, including frames in which the observation update was skipped, the two-dimensional centroid RMSE was calculated from the updated state estimates over 1000 independent evaluation trajectories. The same true trajectories and initial estimation errors were used for all methods. In Figure A3, lower centroid RMSE was obtained with CL-FS than with CL and CL-FS-C0 under the evaluated motion disturbances.

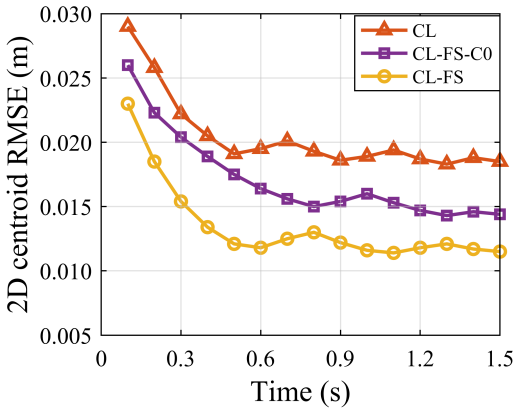


Figure A3. Two-dimensional centroid RMSE under random acceleration disturbances.

References

1. Lyu, M.; Zhao, Y.; Huang, C.; Huang, H. Unmanned Aerial Vehicles for Search and Rescue: A Survey. *Remote Sens.* **2023**, *15*, 3266. [CrossRef]
2. Song, F.; Zhang, W.; Yuan, T.; Ji, Z.; Cao, Z.; Xu, B.; Lu, L.; Zou, S. UAV Quantitative Remote Sensing of Riparian Zone Vegetation for River and Lake Health Assessment: A Review. *Remote Sens.* **2024**, *16*, 3560. [CrossRef]
3. Harlow, K.; Jang, H.; Barfoot, T.D.; Kim, A.; Heckman, C. A New Wave in Robotics: Survey on Recent MmWave Radar Applications in Robotics. *IEEE Trans. Robot.* **2024**, *40*, 4544–4560. [CrossRef]
4. Lu, C.X.; Rosa, S.; Zhao, P.; Wang, B.; Chen, C.; Stankovic, J.A.; Trigoni, N.; Markham, A. See Through Smoke: Robust Indoor Mapping with Low-cost mmWave Radar. In *Proceedings of the 18th Annual International Conference on Mobile Systems, Applications, and Services*; MobiSys'20; Association for Computing Machinery: New York, NY, USA, 2020; pp. 14–27. [CrossRef]
5. Zhao, S.; Li, K.; Luo, Y.; Wang, Q.; Zhao, X.; Zhang, Z. Structure-Preserving Feature Disentanglement with an Adaptive Transformer for Unsupervised Cross-Domain Object Detection in Heterogeneous SAR Imagery. *ISPRS J. Photogramm. Remote Sens.* **2026**, *242*, 490–504. [CrossRef]
6. Zhao, S.; Wang, Q.; Zhao, X.; Kang, Y.; Zhang, C.; Luo, Y. Frequency-Adaptive Disentanglement and Topological Graph Alignment for Cross-Modal Optical–SAR Ship Reidentification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2026**, *19*, 23017–23030. [CrossRef]
7. Li, D.; Li, X.; Qin, Y.; Cheng, Y.; Wang, H. Radar Coincidence Imaging: An Instantaneous Imaging Technique With Stochastic Signals. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2261–2277. [CrossRef]
8. Xu, H.; Xu, F.; Jin, Y.Q. Variable-Resolution SAR Imaging Mode With the Principle of Maximum Mutual Information. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5201914. [CrossRef]
9. Gui, S.; Yang, Y.; Hu, R.; Yan, F.; Tian, Z.; Pi, Y. Dynamic ISAR Imaging Method for Multiple Moving Vehicles Based on OMP-CADMM. *IEEE Trans. Veh. Technol.* **2022**, *71*, 10948–10959. [CrossRef]
10. Zhang, F.; Liu, X.; Zhou, X.; Wang, X.; Liu, W. Autofocus technique for radar coincidence imaging with model error via iterative maximum a posteriori. *J. Eng.* **2019**, *2019*, 5837–5840. [CrossRef]
11. Li, H.; Li, S.; Li, Z.; Dai, Y.; Jin, T. Compressed Sensing Imaging with Compensation of Motion Errors for MIMO Radar. *Remote Sens.* **2021**, *13*, 4909. [CrossRef]
12. Chen, M.; Qiu, X.; Cheng, Y.; Shang, M.; Li, R.; Li, W. Two-Dimensional Autofocus for Ultra-High-Resolution Squint Spotlight Airborne SAR Based on Improved Spectrum Modification. *Remote Sens.* **2024**, *16*, 2158. [CrossRef]
13. Zhou, Y.; Nan, Y.; Liang, D.; Xue, Z.; Chen, Y.; Zhou, H.; Zhao, Y. Auto-Focusing Imaging and Performance Analysis of Ka-Band Carrier-Frequency-Agility SAR. *Remote Sens.* **2026**, *18*, 1197. [CrossRef]
14. Ma, Y.; Ma, R.; Lin, Z.; Miao, C.; Zhang, R.; Long, W.; Wu, W.; Wang, J. Distributed Split Single-Sideband Time-Modulated Arrays for Secure Communications. *IEEE Internet Things J.* **2026**, *13*, 23862–23875. [CrossRef]
15. Ma, Y.; Ma, R.; Lin, Z.; Zhang, R.; Cai, Y.; Wu, W.; Wang, J. Improving Age of Information for Covert Communication With Time-Modulated Arrays. *IEEE Internet Things J.* **2025**, *12*, 1718–1731. [CrossRef]
16. Zhao, M.; Zhu, S.; Fromenteze, T.; Abbasi, M.A.B.; Chen, X.; Huang, Y.; Yurduseven, O. Dual-Polarized Frequency-Diverse Metaimager for Computational Polarimetric Imaging. *IEEE Trans. Antennas Propag.* **2024**, *72*, 5479–5488. [CrossRef]
17. Li, A.; Zhao, M.; Abbasi, M.A.B.; Yurduseven, O. A Measurement Mode Selection Method for Computational Microwave Imaging. *IEEE Antennas Wirel. Propag. Lett.* **2024**, *23*, 3382–3386. [CrossRef]
18. Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.* **1960**, *82*, 35–45. [CrossRef]

19. Lipworth, G.; Mrozack, A.; Hunt, J.; Marks, D.L.; Driscoll, T.; Brady, D.; Smith, D.R. Metamaterial Apertures for Coherent Computational Imaging on the Physical Layer. *J. Opt. Soc. Am. A* **2013**, *30*, 1603–1612. [CrossRef]
20. Zhang, J.; Yuan, B.; Jiang, Z.; Guo, Y.; Wang, D. Microwave Staring Correlated Imaging Method Based on Steady Radiation Fields Sequence. *Sensors* **2020**, *20*, 6859. [CrossRef]
21. Wipf, D.P.; Rao, B.D. Sparse Bayesian Learning for Basis Selection. *IEEE Trans. Signal Process.* **2004**, *52*, 2153–2164. [CrossRef]
22. Nocedal, J.; Wright, S.J. *Numerical Optimization*, 2nd ed.; Springer: New York, NY, USA, 2006. [CrossRef]
23. Miller, A.J.; Nguyen, N.K. A Fedorov Exchange Algorithm for D-Optimal Design. *J. R. Stat. Soc. Ser. C (Applied Stat.)* **1994**, *43*, 669–677. [CrossRef]
24. Dong, H.; Dai, F.; Zhang, J.; Lu, X. HRRPs Reconstruction of Multiple High-Speed Targets With Large Time-Bandwidth Product LFM Waveform Based on Efficient SBL. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 5322–5339. [CrossRef]
25. Kirkpatrick, S.; Gelatt, Jr., C.D.; Vecchi, M.P. Optimization by Simulated Annealing. *Science* **1983**, *220*, 671–680. [CrossRef]
26. Chew, W.C. *Waves and Fields in Inhomogeneous Media*; IEEE Press: Piscataway, NJ, USA, 1999.

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.