Application of Regularization Methods in the Sky Map Reconstruction of the Tianlai Cylinder Pathfinder Array (Postprint)

Authors: Kaifeng Yu, Shifan Zuo, Fengquan Wu, Yougang Wang and Xuelei Chen

Date: 2024-02-28T09:04:04+00:00

Abstract

The Tianlai cylinder pathfinder is a radio interferometer array to test 21 cm intensity mapping techniques in the post-reionization era. It works in passive drift scan mode to survey the sky visible in the northern hemisphere. To deal with the large instantaneous field of view and the spherical sky, we decompose the drift scan data into m-modes, which are linearly related to the sky intensity. The sky map is reconstructed by solving the linear interferometer equations. Due to incomplete uv coverage of the interferometer baselines, this inverse problem is usually ill-posed, and regularization method is needed for its solution. In this paper, we use simulation to investigate two frequently used regularization methods, the Truncated Singular Value Decomposition (TSVD), and the Tikhonov regularization techniques. Choosing the regularization parameter is very important for its application. We employ the generalized cross validation method and the L-curve method to determine the optimal value. We compare the resulting maps obtained with the different regularization methods, and for the different parameters derived using the different criteria. While both methods can yield good maps for a range of regularization parameters, in the Tikhonov method the suppression of noisy modes are more gradually applied, produce more smooth maps which avoids some visual artefacts in the maps generated with the TSVD method.

Full Text

Preamble

Research in Astronomy and Astrophysics, 24:025002 (12pp), 2024 February
 https://doi.org/10.1088/1674-4527/ad1223

Application of Regularization Methods in the Sky Map Reconstruction of the Tianlai Cylinder Pathfinder Array

Kaifeng Yu[1,2,3], Shifan Zuo[1,2], Fengquan Wu[1,2], Yougang Wang[1,2], and Xuelei Chen[1,2,3,4,5]

[1] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; xuelei@cosmology.bao.ac.cn [2] Key Laboratory of Radio Astronomy and Technology, Chinese Academy of Sciences, Beijing 100101, China [3] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China [4] Key Laboratory of Cosmology and Astrophysics (Liaoning) & College of Sciences, Northeastern University, Shenyang 110819, China [5] Center of High Energy Physics, Peking University, Beijing 100871, China

Received 2023 November 15; accepted 2023 November 27; published 2024 January 23

Abstract

The Tianlai cylinder pathfinder is a radio interferometer array designed to test 21 cm intensity mapping techniques in the post-reionization era. Operating in passive drift scan mode, it surveys the sky visible from the northern hemisphere. To handle the large instantaneous field of view and the spherical sky geometry, we decompose the drift scan data into m-modes, which are linearly related to the sky intensity. The sky map is reconstructed by solving the linear interferometer equations. Due to incomplete uv coverage of the interferometer baselines, this inverse problem is typically ill-posed, requiring regularization methods for its solution. In this paper, we use simulations to investigate two frequently used regularization methods: Truncated Singular Value Decomposition (TSVD) and Tikhonov regularization. Choosing the regularization parameter is critical for successful application.

We employ the generalized cross-validation method and the L-curve method to determine the optimal parameter values. We compare the resulting maps obtained with different regularization methods and with different parameters derived using the different criteria. While both methods can yield good maps for a range of regularization parameters, the Tikhonov method applies suppression of noisy modes more gradually, producing smoother maps that avoid some visual artifacts present in maps generated with the TSVD method.

Key words: techniques: interferometric –methods: numerical –cosmology: observations –radio continuum: general

1. Introduction

The Tianlai experiment is designed to develop and test the H I Intensity Mapping (IM) technique (Chen 2012; Li et al. 2020), in which the large-scale distribution of neutral hydrogen is observed without resolving individual galaxies. This allows for fast survey speed to cover the large volumes required for cosmological studies (Chang et al. 2008). This technique has been applied to observations with the Green Bank Telescope (GBT) and the Parkes telescope (Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013; Wolz et al. 2017; Anderson et al. 2018; Wolz et al. 2021). Other existing or ongoing H I IM experiments focusing on late-time cosmology include both single-dish telescopes and interferometers, such as FAST (Hu et al. 2020), BINGO (Battye et al. 2013), CHIME (CHIME Collaboration et al. 2022), and SKA in the near future (Square Kilometre Array Cosmology Science Working Group et al. 2020). However, H I IM observation is very challenging due to strong foreground radiation that is 4–5 orders of magnitude brighter than the cosmological signal. Additionally, instrumental systematics and other observational effects further complicate the separation of the signal from the foreground, requiring highly sophisticated analysis (Liu & Shaw 2020).

The Tianlai experiment includes a cylinder array and a dish array. The Tianlai cylinder pathfinder array is fixed on the ground and relies on the rotation of the Earth to survey the sky. It consists of three adjacent 15 m × 40 m cylindrical reflectors, equipped with 31, 32, and 33 feeds from east to west, respectively. The basic performance of the Tianlai arrays has been analyzed in Li et al. (2020) for the cylinder pathfinder and in Wu et al. (2021) for the dish pathfinder. Although making synthesis images of the sky from interferometric raw data is strictly speaking not needed for power spectrum estimation, in practice it is still an essential procedure to compress the data for further scientific analysis and to provide an intuitive means of checking data quality and algorithms applied.

The output of an interferometer, the interferometric visibility, is the cross-correlation between the voltage signals from two feed elements. For the unpolarized case, the visibility is given by

$$V_{ij} = \int d\hat{n}\, T(\hat{n}) A_i(\hat{n}) A_j^*(\hat{n}) e^{2\pi i \hat{n}\cdot \mathbf{u}_{ij}},$$

where $T(\hat{n})$ is the sky brightness temperature, $\hat{n}$ gives the direction on the sky, A_i is the beam of feed i, and $\mathbf{u}_{ij} = (\mathbf{r}_i - \mathbf{r}_j)/\lambda$ is the separation between feeds in units of the observed wavelength. In the second line, we introduce the beam transfer function B_{ij}. In discrete form, the integral is replaced by a sum over pixelated sky indexed by n_i as follows

$$V_{ij} = \sum_{n_i} T(n_i) B_{ij}(n_i) \Delta\Omega,$$

or in matrix-vector form,

$$\mathbf{V} = \mathbf{B} \cdot \mathbf{T}.$$

Given the measurements $\mathbf{V}$, synthesis imaging aims to estimate $T(\hat{n})$. In the flat-sky limit, the measured visibilities of an interferometer array correspond to Fourier modes of sky intensity variation, with orientation and angular scale determined by the direction and length of the baselines. However, in practice the baselines of an array often do not provide complete coverage of the spatial frequency domain, so reconstructing the sky map from the visibilities is not trivial, as it is mathematically an ill-posed inverse problem with no unique solution. To overcome gaps in measured visibilities on the uv plane, numerous techniques have been developed to reconstruct images in such cases, e.g., variants of the CLEAN algorithm (Högbom 1974) and burgeoning methods based on neural networks (Xu et al. 2020; Connor et al. 2022; Schmidt et al. 2022).

The general principle of sky image reconstruction for the Tianlai Array was investigated in Zhang et al. (2016a, 2016b). Zuo et al. (2021) developed a pipeline for the map-making process for the Tianlai array. In a recent paper (Yu et al. 2023, hereafter referred to as Paper I), we presented a simulation of the Tianlai cylinder pathfinder array through to the making of synthesis maps, taking into account both thermal noise and calibration error. This simulation revealed that although the Tianlai array is relatively compact and dense, there are still gaps in its baseline coverage, and as a result, the reconstructed image has artifacts arising from incomplete uv coverage or beam sidelobes. In the present paper, we investigate regularization methods for dealing with this ill-posed inverse problem.

The structure of this paper is as follows. In Section 2, we present our simulation setup and give a brief summary of the m-mode formalism for map-making. In Section 3, we apply the regularized m-mode formalism imaging to the simulation data and explore the choice of appropriate regularization parameters. In Section 4, we discuss the errors in these regularized maps, and finally in Section 5, we present our conclusions.

2.1. The m-mode Formalism

Although imaging the sky by solving for T in Equation (4) directly is intuitive, the computational cost is very large. If the interferometer has N_{bl} baselines and has run for N_{time} time samples, then for each single frequency and N_{pix} discrete sky pixels, the visibility vector $\mathbf{V}$ has size $N_{\mathrm{bl}} \times N_{\mathrm{time}}$, and the dimensions of the beam transfer matrix $\mathbf{B}$ will be $(N_{\mathrm{bl}} \times N_{\mathrm{time}}, N_{\mathrm{pix}})$. Zheng et al. (2017) demonstrated this method for a small array. For the Tianlai cylinder pathfinder, the number of non-redundant baselines is $N_{\mathrm{bl}} \approx 3300$, and the time samples are $N_{\mathrm{time}} \approx 21,600$ in a sidereal day for 4 s integration time. If we pixelate the whole sky within latitude $[-30^{\circ}, 90^{\circ}]$ using the HEALPix scheme with $N_{\mathrm{side}} = 256$ (Górski et al. 2005), the pixel number is $N_{\mathrm{pix}} \approx 6 \times 10^5$. Then the dimensions of the transfer matrix $\mathbf{B}$ would be $71\mathrm{M} \times 0.6\mathrm{M}$ per frequency. Apparently, solving Equation (4) directly would be intractable for a large array due to the enormous matrix involved.

The m-mode method (Shaw et al. 2014, 2015; Zhang et al. 2016b) provides a convenient and computationally efficient approach for data processing and analysis of drift-scan telescopes, especially wide-field telescopes. It has been applied in data analysis for LWA (Eastwood et al. 2018), EDA2 (Kriele et al. 2022), and CHIME (CHIME Collaboration et al. 2022). We have also implemented this method in the map-making procedure of the Tianlai data processing pipeline tlpipe (Zuo et al. 2021).

As the drift-scan telescope measures the sky periodically with the rotation of the Earth, we can decompose the sky brightness temperature $T(\hat{n})$ and the beam transfer function $B_{ij}(\hat{n}, f)$ in spherical harmonics. Taking the Fourier transform of the visibility $V_{ij}(f)$ can be written as

$$\tilde{V}_{ij}(m) = \sum_{\ell} B^{ij}_{\ell m} a_{\ell m},$$

and adding up the noise term gives the key equation in m-mode analysis. We can rewrite it in matrix form as

$$\mathbf{v}_m = \mathbf{B}_m \mathbf{a}_m + \mathbf{n}_m.$$

The label $(ij, \pm)$ indicates that positive and negative values of m are grouped together because the positive and negative m-modes measure the same real-valued sky, which gives $a_{\ell,-m} = (-1)^m a^*_{\ell,m}$. We can rewrite Equation (8) in a general form as

$$\mathbf{v} = \mathbf{B}\mathbf{a} + \mathbf{n},$$

where matrix $\mathbf{B}$ has a block diagonal structure, so we can treat each m-block independently. Then the Fourier transform of the visibility $\mathbf{v}$ is related to the spherical harmonics coefficients of the sky $\mathbf{a}$ by a simple linear equation. The imaging process is to solve for $\mathbf{a}$ given the observation $\mathbf{v}$ with prior information about $\mathbf{B}$ and noise $\mathbf{n}$. Compared with directly solving for $T(n_i)$ in Equation (3), the m-mode method handles matrix operations with much smaller matrices, which increases computational efficiency tremendously. The main computational cost is devoted to constructing the transfer matrix $\mathbf{B}$. However, once computed and stored, it can be reused in subsequent processing.

2.2. Simulation

We simulate the map-making process for the unpolarized case with a single frequency (750 MHz). We use the publicly available software cora[6] (Shaw et al. 2014, 2015) to generate our mock sky model, which contains foreground emission including diffuse emission from the Galaxy and extragalactic point sources, alongside the cosmological 21 cm signal. The diffuse emission is generated by extrapolating the Haslam map with a specified spectral index map and including random spectral and angular fluctuations. The extragalactic point sources consist of a catalog of real bright sources from NVSS and VLSS, a synthetic catalog of fainter sources, and a random background for even fainter sources. The cosmological 21 cm signal is generated by drawing Gaussian realizations of a power spectrum (Shaw et al. 2014, 2015).

We model the beam pattern, which is characterized as a long strip along the north-south direction, using the following analytical form (see Paper I for details), with parameters determined from a fit to electromagnetic simulations of the Tianlai cylinder array (Sun et al. 2022):

$$B(\hat{n}) = \exp\left[-\frac{1}{2}\left(\frac{\hat{n}\cdot\hat{x}}{\sigma_{EW}}\right)^2\right] \times \left[1 + F\cos\left(\frac{2\pi}{D_{NS}}\hat{n}\cdot\hat{y}\right)\right]^{\alpha},$$

where $\hat{x}$ and $\hat{y}$ are unit vectors pointing east and north, respectively, and

$$\sigma_{EW} = \frac{\theta_{EW}}{\sqrt{8\ln 2}},$$

where $D_{NS} = 0.3$ m is the size of the Tianlai cylinder feeds, and A_D is taken as the mean of the power beam of X and Y polarization at 750 MHz from Sun et al. (2022) to fit the models above. The fitting parameters are $\alpha = 1.04$, $F = 0.2$, $\theta_{EW} = 2.74^{\circ}$.

The noise in each m-mode (Shaw et al. 2015) is generated by sampling a Gaussian distribution with rms

$$\sigma_N = \frac{T_{\mathrm{sys}}}{\sqrt{2\Delta\nu t_{\mathrm{int}} N_{\mathrm{red}} N_{\mathrm{day}}}},$$

where T_{sys} is the system temperature, W is the geometric mean of individual beam solid angles, N_{day} is the number of observation days, N_{red} is the number of redundant baselines, t_{sid} is the sidereal day in seconds, $\Delta\nu$ is the bandwidth, and t_{int} is the single integration time. For the unpolarized case,

$$T_{\mathrm{sys}} = \sqrt{T_i^{\mathrm{sys}} T_j^{\mathrm{sys}}}.$$

Below we assume all feeds have the same system temperature and primary beam profile. We take the system temperature as 90 K for all feeds based on the estimation in Li et al. (2020), $t_{\mathrm{int}} = 4$ s and $\Delta\nu = 122$ kHz for the current configuration of the Tianlai cylinder pathfinder, and take $N_{\mathrm{day}} = 1$ for a single sidereal day observation, or $N_{\mathrm{day}} = 30$ to imitate stacking multiple sidereal days, which increases the signal-to-noise ratio.

2.3. Beam Coverage in Spherical Harmonics Space

The angular resolution and sensitivity of a telescope on the sky are limited by its size and configuration. We take the maximum spherical harmonics degree ℓ and order m that the Tianlai cylinder pathfinder is sensitive to at 750 MHz as

$$\ell_{\max} \approx \frac{2\pi D_{\max}}{\lambda}, \quad m_{\max} \approx \frac{2\pi D_{EW}}{\lambda \cos\delta},$$

where $D_{\max}$ is the maximum dimension of the entire array, D_{EW} is the physical size in the east-west direction, and δ is the latitude of the Tianlai site.

We can define the spherical harmonic beam coverage of the array by combining the beam transfer matrix from all baselines:

$$C_{\ell m} = \sum_{i=1}^{N_{\rm bl}} |B^i_{\ell m}|^2,$$

where $N_{\rm bl}$ is the total number of non-redundant baselines. This quantity can be used to describe the sensitivity of the array on the sky in spherical harmonic space, analogous to uv coverage, which defines the spatial frequencies measured by the array. In Figure 1 [Figure 1: see original paper], we show the beam coverage of the Tianlai cylinder pathfinder array.

The resolution in the east-west direction of a baseline is determined by its projected distance along this direction. The limit on m for a telescope is $m_{\max} \approx 2\pi D_{EW}/\lambda$ by taking into account that the m mode corresponds to the Fourier mode in the azimuthal direction, where $\delta \approx 44^\circ$ is the latitude of the Tianlai site, thus $m_{\max} \approx 510$ with $D_{EW} = 45$ m and $\lambda \approx 0.3997$ m at 750 MHz. The beam coverage is expected to center around $m \approx 2\pi D_w/\lambda$, where D_w equals 1 or 2 times the cylinder width. With $D_w = 15$ m or 30 m and $b = 0.4$ m, these are at $(\ell, m) \approx (235, 170)$ and $(470, 340)$, which is indeed the case as can be seen in Figure 1.

Roughly, the sensitivity of our model telescope is relatively high in the range of about $[m, m + 200]$ at $m \lesssim 400$, and decreases significantly at and outside the edge region (i.e., $m \sim 510$) that the telescope can reach. Consequently, when reconstructing the sky map from the spherical harmonic coefficients $a_{\ell m}$ in this paper, we filter out all modes with $\ell > 600$, as well as the $m = 0$ modes which measure the average over sidereal time.

2.4. Linear Least-squares Map-Maker

The imaging procedure can be seen as solving Equation (8) for the spherical harmonic coefficients $\mathbf{a}$ given the measurement $\mathbf{v}$ and the beam matrix $\mathbf{B}$. The least-squares solution, which minimizes $||\mathbf{v} - \mathbf{Ba}||^2$, is given by

$$\mathbf{a} = (\mathbf{B}^*\mathbf{B})^{-1}\mathbf{B}^*\mathbf{v},$$

where $*$ denotes the conjugate transpose.

In our current work, before making the general linear least-squares solution, we first subtract the contributions of four very strong radio sources from the visibilities: Cas A, Cyg A, Tau A, and Vir A, which can be treated as point sources at known locations for the current Tianlai cylinder pathfinder array. This reduces the dynamic range of the input map. If they are not subtracted, the reconstructed map will show apparent sidelobes around their positions.

The matrix $\mathbf{B}$ usually does not have full rank, as there are unmeasured modes due to incomplete uv-coverage and incomplete coverage of the full sky. In Figure 2 [Figure 2: see original paper], we plot the singular values of several $\mathbf{B}$ matrices; a cluster of small singular values approaching zero results in these matrices being rank-deficient. Hence $\mathbf{B}^*\mathbf{B}$ is not invertible, which makes the problem ill-posed. The solution to ill-posed inverse problems is usually unstable and may deviate greatly from the true solution in the presence of noise.

3. Regularization

A common approach to ill-posed problems is regularization (see, e.g., Engl et al. 1996; Hansen 1998), which provides a stable approximate solution by imposing additional constraints. Common regularization methods include Truncated Singular Value Decomposition (TSVD) and Tikhonov regularization.

3.1. Truncated Singular Value Decomposition

For the least-squares solution (Equation 13), the matrix $\mathbf{B}$ generally does not have full rank, so $\mathbf{B}^*\mathbf{B}$ is singular and not invertible. With the singular value decomposition of the $m \times n$ matrix $\mathbf{B}$,

$$\mathbf{B} = \mathbf{U}\ \ \mathbf{V}^* = \sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^*,$$

where $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ and $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ are unitary matrices, and the diagonal matrix $\ \ = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ has non-negative elements in non-increasing order. The Moore-Penrose pseudo-inverse of matrix $\mathbf{B}$ is defined as

$$\mathbf{B}^+ = \mathbf{V}\ \ ^+\mathbf{U}^*,$$

where $+$ denotes the pseudo-inverse, and the diagonal elements of matrix $\ ^+$ are the reciprocals of each non-zero element on the diagonal of $\ $, leaving the zeros in place. The minimum-norm least-squares solution to Equation (8) in terms of the Moore-Penrose pseudo-inverse is then given by

$$\mathbf{a} = \mathbf{B}^{+}\mathbf{v} = \sum_{i=1}^{r} \frac{\mathbf{u}_i^{*}\mathbf{v}}{\sigma_i}\mathbf{v}_i,$$

where $r = \mathrm{rank}(\mathbf{B})$. For noisy measurement $\mathbf{v} = \mathbf{v}_0 + \mathbf{n}$, where $\mathbf{v}_0 = \mathbf{B}\mathbf{a}_{\mathrm{true}}$, this solution becomes

$$\mathbf{a} = \sum_{i=1}^{r} \frac{\mathbf{u}_i^{*}\mathbf{v}_0}{\sigma_i}\mathbf{v}_i + \sum_{i=1}^{r} \frac{\mathbf{u}_i^{*}\mathbf{n}}{\sigma_i}\mathbf{v}_i.$$

The first term gives the true solution, while the second term gives the contribution from noise. The latter might be magnified by extremely small singular values σ_i of $\mathbf{B}$, making the solution unstable and heavily contaminated by noise.

Regularization solves this problem by dampening or filtering out unwanted components corresponding to small singular values, leading to an approximate but stable solution. In terms of the singular value decomposition of matrix $\mathbf{B}$, the regularized solution to Equation (8) can be written as

$$\mathbf{a}_{\mathrm{reg}} = \sum_{i=1}^{r} f_i \frac{\mathbf{u}_i^{*}\mathbf{v}}{\sigma_i}\mathbf{v}_i,$$

where f_i are filter factors. Different regularization methods can be defined by choosing suitable filter factors. For TSVD regularization, the pseudo-inverse $\mathbf{B}^{+}$ in Equation (14) is replaced by its rank-k approximation, with the remaining components filtered out. The filter factors are

$$f_i^{\mathrm{TSVD}} = \begin{cases} 1 & \text{if } \sigma_i \geq \epsilon \\ 0 & \text{if } \sigma_i < \epsilon \end{cases},$$

where ϵ is the truncation threshold.

In Paper I, we used TSVD regularization to compute the regularized solution of Equation (8), where the truncation satisfies $\sigma_i/\sigma_{\max} \geq \epsilon$ and we selected $\epsilon = 1 \times 10^{-3}$ as the default. As an illustration, we show in Figure 3 [Figure 3: see original paper] the reconstructed maps with three truncation thresholds ϵ, representing cases with too small, moderate, and too large ϵ, respectively. Noise in the reconstructed map is significantly amplified if there is no regularization or if the threshold value ϵ is too small, resulting in an image totally dominated by noise. For too large ϵ, although noise is not significantly amplified, the reconstructed signal may fail to capture the true signal completely in certain modes, so the resulting image is again not optimal. For a properly chosen regularization value, the reconstruction result is relatively good, with the input sky map well recovered. However, as we noted in Paper I, even in this case, near the bright part of the Galactic plane there are comb-like artifacts produced by incomplete reconstruction of certain truncated modes, especially those with small m values.

3.2. Tikhonov Regularization

Another widely used regularization method is Tikhonov regularization. In contrast to TSVD regularization, which imposes a hard threshold on components corresponding to small singular values, Tikhonov regularization aims to suppress undesirable modes by solving for $\mathbf{a}$ that minimizes

$$\arg\min_{\mathbf{a}} \left\{ ||\mathbf{Ba} - \mathbf{v}||_2^2 + \lambda^2 ||\mathbf{L}(\mathbf{a} - \mathbf{a}_0)||_2^2 \right\},$$

where $|| \cdot ||_2$ denotes the Frobenius norm, λ is the regularization parameter, $\mathbf{a}_0$ is a prior for $\mathbf{a}$ (which can be set to $\mathbf{a}_0 = \mathbf{0}$ if prior information is unavailable), and $\mathbf{L}$ is called the Tikhonov regularization matrix. There are several common choices for $\mathbf{L}$, including the identity matrix or a discrete approximation of the derivative operator. This minimization problem gives the solution

$$\mathbf{a}_\lambda = (\mathbf{B}^*\mathbf{B} + \lambda^2\mathbf{L}^*\mathbf{L})^{-1}\mathbf{B}^*\mathbf{v} + (\mathbf{B}^*\mathbf{B} + \lambda^2\mathbf{L}^*\mathbf{L})^{-1}\lambda^2\mathbf{L}^*\mathbf{L}\mathbf{a}_0.$$

In the simple case where $\mathbf{L} = \mathbf{I}$, the Tikhonov regularization is referred to as the standard form. The more general form, where $\mathbf{L} \neq \mathbf{I}$ and prior information about $\mathbf{a}$ and $\mathbf{n}$ is available, can be transformed into the standard form (Eldén 1977). For simplicity, we adopt $\mathbf{L} = \mathbf{I}$ and $\mathbf{a}_0 = \mathbf{0}$, so the objective function becomes

$$\arg\min_{\mathbf{a}} \left\{ ||\mathbf{Ba} - \mathbf{v}||_2^2 + \lambda^2 ||\mathbf{a}||_2^2 \right\},$$

and Equation (19) reduces to

$$\mathbf{a}_\lambda = (\mathbf{B}^*\mathbf{B} + \lambda^2\mathbf{I})^{-1}\mathbf{B}^*\mathbf{v}.$$

It can be expressed in the form of Equation (16) with filter factors

$$f_i^{\text{Tikhonov}} = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}.$$

Unlike TSVD regularization, which simply trims out modes corresponding to small eigenvalues, Tikhonov regularization applies gradual and smooth suppression to all modes. The degree of smoothness can be adjusted by the choice of the regularization parameter λ; a small regularization parameter yields a solution closer to that obtained from TSVD regularization.

In Figure 4 [Figure 4: see original paper], we illustrate Tikhonov-regularized sky maps with different regularization parameters λ. For a small λ, as shown in the left panel, noise in the regularized map is obviously amplified. The noise becomes less pronounced with a moderate λ value, as shown in the center panel. The comb-like artifact around the Galactic Center shown in Figure 3 is also less prominent, perhaps thanks to the gradual suppression of modes in Tikhonov regularization. However, for the case with an excessively large λ shown in the right panel, the underlying signal is also suppressed, generating a more bland map than the actual sky. Furthermore, accompanied by the diminished actual sky signal, the sidelobes from some bright sources become more pronounced; for example, an arc structure becomes noticeable near the north celestial pole and above Cyg A. An appropriate selection of λ that strikes a balance between noise amplification and information loss is expected to yield a satisfactory map containing moderate noise while faithfully preserving the true sky as much as possible.

The regularized solution error is given by

$$\mathbf{a}_\lambda - \mathbf{a}_{\text{true}} = \sum_{i=1}^{r} (f_i - 1) \frac{\mathbf{u}_i^* \mathbf{v}_0}{\sigma_i} \mathbf{v}_i + \sum_{i=1}^{r} f_i \frac{\mathbf{u}_i^* \mathbf{n}}{\sigma_i} \mathbf{v}_i,$$

where the first term corresponds to the regularization error, arising from regularization of the true signal $\mathbf{a}_{\text{true}}$, and the second term represents the perturbation error, attributed to the presence of noise. The filter factors f_i can be f_i^{TSVD} or f_i^{Tikhonov}. As $\lambda \to 0$, the filter factors $f_i^{\text{Tikhonov}} \to 1$, the regularization error approaches zero, but the perturbation error might be large, and the solution tends to be noisy. As λ increases, the filter factors f_i decrease, leading to smaller perturbation error but larger regularization error, and the regularized solution becomes oversmoothed and approaches zero. The key to obtaining a good regularized solution is to choose an optimal regularization parameter λ in Tikhonov regularization or k in TSVD regularization to balance these two errors.

3.3. The Choice of Regularization Parameter

Choosing the optimal regularization parameter that balances the trade-off between noise amplification and signal recovery is not always easy, as it depends on the specific problem, including factors such as the array configuration and the noise level in the data. Several approaches are available for selecting a regularization parameter near the optimal value. Below we adopt the notation used in Tikhonov regularization (i.e., take λ as the regularization parameter) for our description, though these methods are also applicable to TSVD regularization.

If the statistics of the noise are well known, λ can be chosen by applying the discrepancy principle (Engl 1987), such that the residual is at a comparable level to that of the noise:

$$||\mathbf{Ba}_\lambda - \mathbf{v}||_2 \leq \eta ||\mathbf{n}||_2,$$

where $\mathbf{a}_\lambda$ is the solution given by Equation (21) with regularization parameter value λ, and η is a user-specified constant constraining the bound. If the error is unknown, techniques such as generalized cross-validation (Golub et al. 1979) or the L-curve criterion (Hansen & O' Leary 1993) are usually applied to search for the appropriate regularization parameter.

Generalized Cross Validation (GCV)—The GCV method (Golub et al. 1979) is based on the following idea: for a linear system $\mathbf{Ax} = \mathbf{b}$, if we drop a data point b_i from vector $\mathbf{b}$ and obtain a regularized solution $\mathbf{x}_\lambda^{[i]}$ from the remaining vector $\mathbf{b}_{[i]}$, then the value $(\mathbf{Ax}_\lambda^{[i]})_i$ should be close to the excluded value b_i if a reasonable parameter λ is chosen (see Chapter 4 of Wahba 1990). This results in selecting a parameter λ that minimizes the GCV function

$$G(\lambda) = \frac{||(\mathbf{I} - \mathbf{BB}_\lambda^{\#})\mathbf{v}||_2^2}{\left[\mathrm{trace}(\mathbf{I} - \mathbf{BB}_\lambda^{\#})\right]^2},$$

where $\mathbf{B}_\lambda^{\#}$ is the regularized inverse and $\mathbf{a}_\lambda = \mathbf{B}_\lambda^{\#}\mathbf{v}$ is the corresponding regularized solution. For illustration, in Figure 5 [Figure 5: see original paper] we show the logarithmic plot of the calculated GCV function applied to data with 30 days integration noise for the case of $m = 200$. In our computation, λ values vary from 10^{-10} to 10^{-1} and are sampled evenly on a logarithmic scale. The GCV function decreases slowly as λ increases until at some point it rises rapidly, and the minimum is reached before this rapid rise, marked by a red star in the figure.

L-curve—In a log-log scale plot depicting the residual norm $||\mathbf{Ba}_\lambda - \mathbf{v}||_2$ versus the solution norm $||\mathbf{a}_\lambda||_2$, the resulting curve typically exhibits an L-shaped profile, comprising a flat part where the regularization error dominates at large λ, and a steep part where the perturbation error dominates at small λ. The optimal λ value should be chosen to balance these two errors, which corresponds to the corner of the L-curve and can be identified by locating the maximum curvature of the curve (Hansen 1999). Let $\hat{\mathbf{a}}_\lambda = \mathbf{a}_\lambda / ||\mathbf{a}_\lambda||_2$, then the curvature of the L-curve is given by

$$\kappa(\lambda) = \frac{\hat{\mathbf{a}}_\lambda'' \cdot \hat{\mathbf{a}}_\lambda' - ||\hat{\mathbf{a}}_\lambda'||^2}{\left(||\hat{\mathbf{a}}_\lambda'||^2\right)^{3/2}}.$$

In Figure 6 [Figure 6: see original paper] we plot the L-curve (left) and the corresponding curvature κ (right) applied to data with 30 days integration noise for the $m = 200$ case; the red star marks the optimal parameter λ determined using this method.

The optimal values for the regularization parameter λ as determined by the GCV criterion and the L-curve criterion for all m cases are shown in Figure 7 [Figure 7: see original paper]. Comparing the optimal values determined by these two criteria, we observe that the optimal λ obtained from the L-curve criterion is typically larger than that provided by the GCV criterion at the same m, especially for higher noise cases (e.g., 1 day integration noise level). The optimal λ values increase with m at $m \lesssim 100$ and then remain relatively stable for the remaining m values.

In Figure 8 [Figure 8: see original paper], we present a comparison between the input $a_{\ell m}$ and the Tikhonov-regularized solution for data with 30 days integration noise for the $m = 10$ (left) and $m = 200$ (right) cases, where the regularization parameter λ is given by the two criteria. The real and imaginary parts are plotted in the top and bottom sub-figures, and the residues are plotted in the bottom panel of each sub-figure. As expected, a smaller regularization parameter obtained from the GCV criterion results in a solution slightly closer to the true value, especially at smaller ℓ, but it may also amplify noise for certain modes, as illustrated here in the region $\ell \in [200, 400]$ for the $m = 10$ case. For the $m = 200$ case, with the smaller regularization parameter provided by the GCV criterion, the residue is smaller without the cost of amplifying the noise.

Moreover, we apply the L-curve criterion to choose the truncation threshold value ϵ or k in Equation (17) for TSVD regularization. We sample 100 values of ϵ evenly on a logarithmic scale ranging from 1×10^{-4} to 1×10^{-1} to compute points on the TSVD L-curve $(||\mathbf{B}\mathbf{a}_k - \mathbf{v}||_2, ||\mathbf{a}_k||_2)$. Unlike the case of Tikhonov regularization, the points on the TSVD regularization L-curve are discrete; a quadratic spline interpolation is applied to these discrete values to compute the curvature of the L-curve and select the sampled ϵ value closest to the maximum curvature. In Figure 9 [Figure 9: see original paper], we show results for data with 30 days integration noise. For comparison, the case for fixed $\epsilon = 1 \times 10^{-3}$ is also plotted. We can see that at large m the two curves almost coincide, but at some relatively smaller m values the L-curve suggests a smaller k, indicating that more low-sensitivity modes are trimmed off to avoid amplifying noise.

In the top panels of Figure 10 [Figure 10: see original paper] we illustrate the reconstructed maps with automatically chosen regularization parameters for data with 30 days noise level, and in the bottom panels we show their relative error (i.e., the fractional difference with the input map). We show cases of Tikhonov regularization using the GCV (left) and L-curve (middle) criteria, and TSVD regularization using the L-curve criterion (right).

It appears that the Tikhonov-regularized solutions generally yield visually better maps of the sky without the comb-like feature seen in the TSVD-regularized map. The errors are typically small except around locations where bright sources are situated. With the λ value obtained by the GCV criterion, which is smaller than that chosen with the L-curve criterion, the region near the Galactic Center is better reproduced, but the northern region tends to be noisier. With the larger λ value chosen using the L-curve criterion, the map is smoother. However, we note that the λ value chosen by both methods would vary with the noise level; therefore, in the present case the larger λ value obtained from the L-curve criterion produces a better overall visual impression, though this may change for different noise levels or setups. Owing to the flat tail of the GCV function curve (Figure 5), the λ value determined from the minimum of the GCV function might not always be accurate and satisfactory, while the L-curve method gives a more robust result. In the reconstructed maps, the smaller regularization parameter obtained from the GCV criterion leads to a noisier map compared to that obtained using the L-curve criterion. As for the TSVD-regularized map, it is generally similar to the map shown in Figure 3 and still shows the comb-like structure near the Galactic Center produced by truncation of certain modes.

3.4. An Overall Regularization Parameter

As matrix $\mathbf{B}$ in Equation (8) is block diagonal, the m-mode method takes advantage of this structure by processing each m-block independently. To produce regularized maps, we have optimized the choice of regularization parameter for each m separately using the GCV and L-curve criteria. For simplicity, it is also possible to adopt a single overall regularization parameter λ considering all m-blocks. We illustrate the L-curve criterion for Tikhonov regularization here. Then the point on the L-curve associated with the regularization parameter λ is given by

$$\left(\sum_m ||\mathbf{B}_m \mathbf{a}_{\lambda,m} - \mathbf{v}_m||_2^2, \sum_m ||\mathbf{a}_{\lambda,m}||_2^2 \right).$$

In Figure 11 [Figure 11: see original paper], we show the relevant L-curve for two cases with different noise levels. For the 1 day case, the noise level is relatively high, and the L-curve criterion yields a relatively large optimal regularization parameter value $\lambda = 2.31 \times 10^{-3}$, while for the 30 days case, the noise level is significantly lower, leading to a correspondingly smaller optimal regularization parameter value $\lambda = 5.34 \times 10^{-4}$.

In Figure 12 [Figure 12: see original paper] we illustrate the reconstructed map with these regularization parameters (the visibility data for imaging has the corresponding noise level). In the 1 day case, due to the relatively large noise and regularization parameter, the map shows more deviation than the 30 days map, and the sidelobe feature near the north celestial pole is quite obvious. However, we can see that for the 30 days noise level map, the deviation is also quite significant compared to the maps made by the m-by-m mode analysis shown in Figure 10. This is not surprising, since a single λ value may not be suitable for all different m-modes. For those less sensitive modes that are susceptible to noise, this value of λ may be too small to regularize the problem. On the other hand, for the modes to which the interferometer is sensitive, it may suppress too much of the true signal.

4. Discussions

Since the regularized solution is only an approximation of the true value, it inevitably introduces bias. In Figure 13 [Figure 13: see original paper], we show the fractional difference between the Tikhonov-regularized map from noise-free data and the input map, which reveals the bias from regularization. We can see that the major bias is around bright point sources and in regions near the horizon. The bias around bright point sources arises from their convolution with the point-spread function, which is expected to be mitigated through further deconvolution techniques such as the CLEAN algorithm.

We can also quantify the quality of the reconstructed map using the angular cross-power spectrum between the input map and the reconstructed maps. In the cross-power spectrum, defined as

$$C_\ell^{\mathrm{cross}} = \frac{1}{2\ell+1} \sum_m a_{\ell m}^{\mathrm{input}} (a_{\ell m}^{\mathrm{rec}})^*,$$

the noise from the maps being crossed is uncorrelated, so it does not contribute to the cross-power. In Figure 14 [Figure 14: see original paper], we show the cross-correlation coefficients, i.e., the ratio of the cross-power between the reconstructed and input maps to the square root of the product of the corresponding auto-power spectra. For Tikhonov regularization, we show cases where the regularization parameter λ is determined m-by-m using the GCV and L-curve criteria, and for TSVD regularization, the truncation threshold is determined based on the L-curve criterion.

In all cases, the correlation falls below 1 as ℓ increases, meaning that the correlation is not perfect due to reconstruction error. There is a general trend that all three curves follow. The result obtained from TSVD regularization has a correlation coefficient close to that of Tikhonov regularization with λ determined by the L-curve at $\ell \lesssim 470$, but beyond this scale there is significant deterioration. The result of Tikhonov regularization with λ determined by the GCV criterion exhibits higher correlation in the range $\ell \lesssim 150$, but shows lower correlation in the range $200 \lesssim \ell \lesssim 550$ than that of the L-curve criterion, where noise is prone to be amplified. However, despite comparable performance as measured by cross-correlation, the Tikhonov regularization produces a better visual impression.

5. Conclusion

In this paper we investigated regularization methods applied to sky map reconstruction from radio interferometer data taken by the Tianlai cylinder pathfinder array. Due to incomplete uv-coverage from the limited number of baselines, regularization is generally necessary in reconstructing the sky from interferometer data, though the strategy and method adopted differ case by case. In our previous paper (Yu et al. 2023), we investigated map reconstruction for the Tianlai cylinder array, with emphasis on assessing the impact of array calibration error and noise on the final map, where we used only the Moore-Penrose pseudo-inverse method in map reconstruction. However, there are various different regularization methods that also affect the map-making result, and exploration of these different regularization methods is the subject of the present paper.

In this work we studied TSVD regularization and Tikhonov regularization applied to our map reconstruction. TSVD regularization is in some sense similar or equivalent to the Moore-Penrose pseudo-inverse used in Paper I, which removes modes susceptible to noise. Tikhonov regularization, on the other hand, suppresses modes gradually—those susceptible to noise are more suppressed but not completely removed—consequently yielding a more smoothly regularized map. This smooth approach by Tikhonov regularization can avoid generating obvious artifacts from the sharp cutoff used in TSVD regularization, producing visually better maps, though it is not necessarily more accurate than TSVD regularization in a quantitative sense.

To obtain a high-fidelity sky map with regularization techniques, it is crucial not to over-regularize the data by using an excessively large regularization parameter. However, the map may be greatly affected by noise if too small a regularization parameter is chosen. We applied the GCV and L-curve methods to determine the optimal regularization parameter. We find that both methods generally produce good maps for reasonable noise levels. In our case the L-curve criterion provides a more stable regularization parameter, which also results in a map with better visual impression. However, we note that this result is specific to the case we investigated; for different data sets, noise levels, etc., the results can differ. Furthermore, although these methods can be used to optimize parameter selection, they are still based on simple reasoning, and additional tuning may be needed to achieve results that better meet specific requirements.

Acknowledgments

We acknowledge the support of the National SKA Program of China (Nos. 2022SKA0110100 and 2022SKA0110101), and the National Natural Science Foundation of China (Nos. 12203061, 12273070, and 12303004).

References

Anderson, C. J., Luciw, N. J., Li, Y. C., et al. 2018, MNRAS, 476, 3382 Battye, R. A., Browne, I. W. A., Dickinson, C., et al. 2013, MNRAS, 434, 1239 CHIME Collaboration, Amiri, M., Bandura, K., et al. 2023, ApJ, 947, 16 Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Natur, 466, 463 Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, PhRvL, 100, 091303 Chen, X. 2012, International Journal of Modern Physics Conference Series, 12, 256 Connor, L., Bouman, K. L., Ravi, V., & Hallinan, G. 2022, MNRAS, 514, 2614 Eastwood, M. W., Anderson, M. M., Monroe, R. M., et al. 2018, AJ, 156, 32 Eldén, L. 1977, BIT Numerical Mathematics, 17, 134 Engl, H. W. 1987, JOTA, 52, 209 Engl, H. W., Hanke, M., & Neubauer, A. 1996, Regularization of inverse problems (Dordrecht: Kluwer) Golub, G. H., Heath, M., & Wahba, G. 1979, Technometrics, 21, 215 Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759 Hansen, P. C. 1998, Rank-Deficient and Discrete Ill-Posed Problems (Philadelphia, PA: SIAM) Hansen, P. C. 1999, The L-Curve and Its Use in the Numerical Treatment of Inverse Problems Hansen, P. C., & O' Leary, D. P. 1993, SJSC, 14, 1487 Högbom, J. A. 1974, A&AS, 15, 417 Hu, W., Wang, X., Wu, F., et al. 2020, MNRAS, 493, 5854 Kriele, M. A., Wayth, R. B., Bentum, M. J., Juswardy, B., & Trott, C. M. 2022, PASA, 39, e017 Li, J., Zuo, S., Wu, F., et al. 2020, SCPMA, 63, 129862 Liu, A., & Shaw, J. R. 2020, PASP, 132, 062001 Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20 Schmidt, K., Geyer, F., Fröse, S., et al. 2022, A&A, 664, A134 Shaw, J. R., Sigurdson, K., Pen, U.-L., Stebbins, A., & Sitwell, M. 2014, ApJ, 781, 57 Shaw, J. R., Sigurdson, K., Sitwell, M., Stebbins, A., & Pen, U.-L. 2015, PhRvD, 91, 083514 Square Kilometre Array Cosmology Science Working Group, Bacon, D. J., Battye, R. A., et al. 2020, PASA, 37, e007 Sun, S., Li, J., Wu, F., et al. 2022, RAA, 22, 065020 Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, MNRAS, 434, L46 Wahba, G. 1990, Spline Models for Observational Data (Philadelphia, PA: SIAM) Wolz, L., Blake, C., Abdalla, F. B., et al. 2017, MNRAS, 464, 4938 Wolz, L., Pourtsidou, A., Masui, K. W., et al. 2022, MNRAS, 510, 3495 Wu F., Li, J., Zuo, S., et al. 2021, MNRAS, 506, 3455 Xu, L., Sun, W.-Q., Yan, Y.-H., & Zhang, W.-Q. 2020, RAA, 20, 170 Yu, K., Wu, F., Zuo, S., et al. 2023, RAA, 23, 105008 Zhang, J., Ansari, R., Chen, X., et al. 2016a, MNRAS, 461, 1950 Zhang, J., Zuo, S.-F., Ansari, R., et al. 2016b, RAA, 16, 158 Zheng, H., Tegmark, M., Dillon, J. S., et al. 2017, MNRAS, 464, 3486 Zuo, S., Li, J., Li, Y., et al. 2021, A&C, 34, 100439

Note: Figure translations are in progress. See original paper for figures.