Kalman Filter with Iteratively Sparse Regularization for Retrieving the Initial Field of the Convection-Diffusion Equation

Xuan Deng and Yuepeng Wang *

Posted Date: 19 March 2026

doi: 10.20944/preprints202603.1333.v1

Keywords: inverse problem; Kalman filter; l1 regularization; adaptive; data assimilation

Article

Kalman Filter with Iteratively Sparse Regularization for Retrieving the Initial Field of the Convection-Diffusion Equation

Xuan Deng and Yuepeng Wang *

School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China
* Correspondence: eduwyp@nuist.edu.cn

Abstract

Sparse regularization methods play an important role in inverse problems for extracting key features of underlying parameters and have attracted increasing attention in meteorological data assimilation. However, when the condition number of the background error covariance matrix is extremely large (e.g., 10^{12}), the instability of the inverse problem makes accurate reconstruction difficult. To address this issue, a gradient operator is incorporated into the sparse regularization term of the cost function, and a Kalman filter (KF) algorithm is developed within a majorization–minimization (MM) framework to solve the resulting optimization problem. The problem is reformulated as a weighted least-squares problem via the MM strategy and further decomposed into two subproblems in the null space and its oblique complementary space through oblique projection, which are then solved using the KF method. This approach avoids the use of an adjoint model typically required in four-dimensional variational data assimilation (4D-Var). In addition, a modified f-slope strategy with a constrained search interval is introduced to adaptively select the regularization parameter during computation. Numerical experiments on the initial-condition inversion of the advection–diffusion equation demonstrate that the proposed method achieves more accurate reconstruction of key features than the l_1-norm regularized 4D-Var method. The inversion errors remain low even when the condition number ranges from 10^8 to 10^{14},with relative MSE and MAE below 0.01 and relative bias below 0.005, indicating improved robustness and reconstruction accuracy.

Keywords: inverse problem; Kalman filter; l_1 regularization; adaptive; data assimilation

1. Introduction

The regularized least squares method is extensively applied in statistical modeling, signal processing, machine learning, and data assimilation[1–4]. It is especially important to explore the inversion of initial conditions using four-dimensional variational data assimilation (4D-Var) and the Kalman filter (KF) with sparse regularization techniques[5–7].

According to the literature[5], the application of the sparse regularized least squares method (generalized Lasso) in 4D-Var can be expressed as the following optimization problem.

$$\hat{\mathbf{m}} = \underset{\mathbf{m}\in\mathbb{R}^p}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{A}\mathbf{m}-\mathbf{b}\|^2_{\mathbf{R}^{-1}} + \frac{1}{2}\|\mathbf{m}-\mathbf{m}^b\|^2_{\mathbf{B}^{-1}} + \lambda\|\mathbf{L}\mathbf{m}\|_1, \tag{1}$$

where $\mathbf{m}$ denotes the initial conditions or parameters to be retrieved, $\mathbf{m}^b \in \mathbb{R}^{p\times 1}$ denotes the background information. $\mathbf{A} = [\mathbf{A}_1; \mathbf{A}_2; \ldots; \mathbf{A}_{N_h}]$, and $\mathbf{A}_k$ consists of the observation operator $\mathcal{H} \in \mathbb{R}^{r\times p}$ and the linear evolution operator $\mathcal{M}_{0,k} \in \mathbb{R}^{p\times p}$, i.e. $\mathbf{A}_k = \mathcal{H}\mathcal{M}_{0,k} \in \mathbb{R}^{r\times p}$. $\mathbf{b} = [\mathbf{b}_1; \mathbf{b}_2; \ldots; \mathbf{b}_{N_h}]$ contains N_h observation information $\mathbf{b}_k \in \mathbb{R}^{r\times 1}$ which is obtained by interval sampling, subscript k denotes the observation time t_k. $\mathbf{R} = \operatorname{diag}(\mathbf{R}_1, \ldots, \mathbf{R}_{N_h})$ is a block-diagonal matrix containing N_h observation error covariance matrices $\mathbf{R}_k \in \mathbb{R}^{r\times r}$. $\mathbf{B} \in \mathbb{R}^{p\times p}$ denotes the background error covariance

 doi:10.20944/preprints202603.1333.v1

matrix. If its condition number becomes excessively large, the stability of the resulting analysis may be adversely affected[5]. $\lambda\|\mathbf{Lm}\|_1$ is a sparsity-promoting regularization term that encourages most components of $\mathbf{Lm}$ to vanish, where $\mathbf{L} \in \mathbb{R}^{q\times p}$ is referred to as the regularization operator and $\lambda > 0$ as the regularization parameter. Note that the choice of $\mathbf{L}$ is critical to problem (1). When $\mathbf{L}$ is taken as the unit operator, the objective reduces to seeking a sparse solution for $\mathbf{m}$. In practice, many state variables are not inherently sparse in their original representations. Instead, they often exhibit sparsity or approximate sparsity under the action of certain transformation operators. Therefore, operators capable of promoting sparse representations will be considered in this work, such as wavelet basis[5] and the derivative operators[8,9]. In particular, the Laplace test reported in [10] suggests that selecting $\mathbf{L}$ as a first-order derivative operator yields improved reconstruction performance.

Algorithms for solving the generalized Lasso problem have been extensively developed, ranging from iterative thresholding algorithms (ISTA), coordinate descent methods, and path algorithms[11–13] to more advanced strategies such as variable-splitting techniques (e.g., the positive-negative split)[5,8], and the majorization-minimization (MM) schemes[14,15]. Furthermore, well-established software packages, including scikit-learn and glmnet, offer efficient implementations of these methods, which greatly facilitate the practical use and wider adoption of l_1 regularization. In [5], Ebtehaj et al. adopted a variable splitting strategy to handle the l_1 regularization term and introduced the Fourier and wavelet basis as the regularization operator. These operators simplify the computational procedure while improving the representation of discontinuous features in the solution. As a result, this approach provides important support for further studies of generalized Lasso in data assimilation. Nevertheless, the minimization of the objective function (1) in 4D-Var requires gradient information, which is typically computed using the tangent linear and adjoint model. For complex numerical systems, the construction of adjoint models is often challenging and costly to maintain, and may also suffer from theoretical limitations, thereby limiting computational efficiency and practical applicability.

The classical KF method is derived under the assumption of Gaussian statistics and is fundamentally rooted in least-squares estimation theory. Since its introduction, it has matured into a rigorous and widely adopted framework for addressing a broad spectrum of estimation problems. By systematically assimilating model predictions with observational data, KF produces updated state estimates accompanied by a quantitative assessment of their uncertainty via the analysis error covariance matrix. Unlike the 4D-Var method, it avoids adjoint construction and supports derivative-free optimization, which is well-suited to sequential data assimilation. Its core principles and associated algorithms can be extended to nonlinear systems through ensemble sampling strategies. Consequently, it has been widely applied in geosciences, engineering control, and biomedicine[16–19], emerging as an important framework for derivative-free optimization[20–26]. Ensemble Kalman inversion (EKI) and its variants incorporate prior information through the initial ensemble or via regularization[21,27,28], iteratively updating the ensemble based on the observation model to approximate the maximum a posteriori (MAP) estimate. However, when an l_1-norm term is included in the objective functional, the standard KF framework cannot be applied directly. A common approach is to replace the l_1 term with a suitable surrogate function[28,29]. For instance, Cao et al. employed the MM strategy to construct a surrogate representation of the Huber-norm based on Young's inequality[7,14]. The specific construction method is as follows:

$$\|m\|_1 = \sum_{i=1}^{p}|m_i| \le \sum_{i=1}^{p}\left(\frac{|m_i|^2}{2|\hat{m}_i|} + \frac{|\hat{m}_i|}{2}\right) = \frac{1}{2}\|C_m^{-\frac{1}{2}}m\|_2^2 + \frac{1}{2}\sum_{i=1}^{p}|\hat{m}_i|, \tag{2}$$

where $\hat{m}_i$ represents the i component of the optimal analysis value $\hat{m}$ of the last iteration, and the weight operator $C_m = \mathrm{diag}(|\hat{m}_1|, |\hat{m}_2|, \ldots, |\hat{m}_p|)$. This method provides a new idea for the application and solution of sparse regularization in KF data assimilation.

Since the constant term does not affect the location of the minimizer, Chartrand et al. demonstrated that it can be omitted in their algorithm[30]. Motivated by this observation, we incorporate the MM

 doi:10.20944/preprints202603.1333.v1

technique into problem (1) to facilitate sparse regularization, thereby reformulating the original problem.

$$\hat{\mathbf{m}}_k^{(j)} = \underset{\mathbf{m}\in\mathbb{R}^p}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{A}_k\mathbf{m}-\mathbf{b}_k\|^2_{\mathbf{R}_k^{-1}} + \frac{\lambda}{2}\|\mathbf{L}\mathbf{m}\|^2_{\mathbf{W}_j^{-1}} + \frac{1}{2}\|\mathbf{m}-\mathbf{m}_{k-1}\|^2_{\mathbf{P}_{k-1}^{-1}} \tag{3a}$$

$$= \underset{\mathbf{m}\in\mathbb{R}^p}{\operatorname{argmin}} \underbrace{\frac{1}{2}\|\mathbf{H}_k\mathbf{m}-\mathbf{y}_k\|_2^2 + \frac{\lambda}{2}\left\|\widetilde{\mathbf{L}}\mathbf{m}\right\|_2^2}_{\mathbf{J}_o(\mathbf{m})} + \underbrace{\frac{1}{2}\|\mathbf{m}-\mathbf{m}_{k-1}\|^2_{\mathbf{P}_{k-1}^{-1}}}_{\mathbf{J}_b(\mathbf{m})}. \tag{3b}$$

In the objective function (3), $\mathbf{J}_b(\mathbf{m})$ denotes the prior (background) regularization term, where $\mathbf{P}_{k-1}$ denotes the background error covariance matrix. When $k=1$, we set $\mathbf{m}_0=\mathbf{m}^b$ and $\mathbf{P}_0=\mathbf{B}$. $\mathbf{J}_o(\mathbf{m})$ represents the likelihood term associated with the observational information, where $\mathbf{H}_k=\mathbf{R}_k^{-\frac{1}{2}}\mathbf{A}_k$, $\mathbf{y}_k=\mathbf{R}_k^{-\frac{1}{2}}\mathbf{b}_k$, $\widetilde{\mathbf{L}}=\mathbf{W}_j^{-\frac{1}{2}}\mathbf{L}$ and j denotes the iteration index. In this study, the regularization operator in (3a) is chosen as the first-order difference operator $\mathbf{L}\in\mathbb{R}^{(p-1)\times p}$[8,9], which is introduced to mitigate the adverse effects caused by the ill-conditioning of the system matrix. The weight matrix $\mathbf{W}_j\in\mathbb{R}^{(p-1)\times(p-1)}$ is constructed adaptively based on the absolute value of the optimal analysis $\hat{\mathbf{m}}_k$ obtained at iteration $j-1$, which guarantees the positive definiteness of the weight matrix. Namely, $\mathbf{W}_j=\operatorname{diag}(|(\mathbf{L}\hat{\mathbf{m}}_k)_1|,\ldots,|(\mathbf{L}\hat{\mathbf{m}}_k)_{p-1}|)$. Furthermore, components of $\mathbf{L}\mathbf{m}$ with magnitude less than 10^{-3} are regarded as sparse in the present work.

At this stage, the generalized Lasso problem can be formulated and optimized in a derivative-free manner within the KF framework, thereby avoiding the explicit computation of adjoint operators. Nonetheless, its practical implementation requires addressing the following three components.

1. The gradient operator controls the smoothness of the regularized solution. However, since the regularization term is a seminorm, the standard least-squares approach cannot be applied directly[31]. To address the computational issues arising from the rank deficiency of the gradient matrix, the original problem is decomposed into two independent subproblems via oblique projection[10,32].
2. The choice of regularization parameters critically affects the quality of the reconstructed solution. In this work, the flattest slope (f-slope) method is employed for parameter selection. However, its computational efficiency is limited by the lack of predefined search intervals[33]. To overcome this limitation, we propose an adaptive regularization parameter selection scheme, wherein the search interval is dynamically adjusted to enhance the efficiency and robustness of the f-slope method.
3. To verify the effectiveness of the proposed method, we consider the advection diffusion equation as a test model, which is representative in physical simulations and suitable for evaluating the feasibility and adaptability of the inversion algorithm. Moreover, the present study conducts a comparative analysis between the proposed method and l_1-norm regularised four-dimensional variational assimilation (R4D-Var), focusing on key aspects such as the quality of feature reconstruction and the sparsity-promoting capability.

The specific structure of this paper is as follows. Section 2 introduces the construction process of the Adaptive iterative Kalman filter inversion algorithm, including a brief explanation of how to perform oblique projection decomposition of state variables in Section 2.1, the derivation of the iterative Kalman inversion algorithm in Section 2.2, and the adaptive selection of regularization parameters via the improved f-slope method in Section 2.3, along with the corresponding pseudocode. Section 3 aims to demonstrate the feasibility and adaptability of the proposed algorithm through the inversion of the initial conditions for the advection-diffusion equation. Sections 3.1 and 3.2 describe the numerical solution framework, parameter settings, and choice of evaluation metrics, while Section 3.3 presents the experimental results. Finally, Section 4 concludes the paper and discusses directions for future research.

 doi:10.20944/preprints202603.1333.v1

2. Adaptive Iterative Kalman Filter Inversion Framework

The key to solving the optimization problem (3) lies in the treatment of the regularization operator. When $\widetilde{\mathbf{L}}$ is square, the problem can be solved directly via matrix inversion. In the present work, the operator $\widetilde{\mathbf{L}}$ is rectangular, which gives rise to two challenges. First, conventional optimization approaches, such as variable splitting, cannot be applied directly [5]. Second, the regularization term is a semi-norm. The state variable $\mathbf{m}$ contains components lying both in the null space of $\widetilde{\mathbf{L}}$ and its complement, which prevents the regularization term from effectively constraining the null space components under the Tikhonov framework. Hence, it is necessary to decompose $\mathbf{m}$ appropriately. Such a separation not only facilitates the effective selection of regularization parameters but also provides a foundation for the subsequent construction of feasible solutions.

2.1. Oblique Projection Decomposition of State Variables

In the process of solving problem (3), $\widetilde{\mathbf{L}}$ is row full rank but not column full rank, and thus possesses a nontrivial null space $\mathcal{N}(\widetilde{\mathbf{L}})$. Under the standard Euclidean inner product, $\mathcal{N}(\widetilde{\mathbf{L}})$ and its orthogonal complement $\mathcal{N}(\widetilde{\mathbf{L}})^{\perp}$ are orthogonal. However, a standard orthogonal decomposition along $\mathcal{N}(\widetilde{\mathbf{L}})^{\perp}$ obstructs the decoupling of the observation term $\|\mathbf{H}_k\mathbf{m} - \mathbf{y}_k\|_2^2$ [32]. In this case, an oblique projection operator is introduced to enable decomposition along skew complement spaces.

First, the definition of null space $\mathcal{N}(\widetilde{\mathbf{L}})$ is

$$\mathbb{F} = \mathcal{N}(\widetilde{\mathbf{L}}) = \left\{ x \in \mathbb{R}^p \middle| \widetilde{\mathbf{L}}x = 0 \right\}.$$

Then, the weighted inner product induced by $\mathbf{H}_k$ is introduced, expressed as $\langle \mathbf{H}_k x, \mathbf{H}_k y \rangle = x^T {\mathbf{H}_k}^T \mathbf{H}_k y$, through which the weighted oblique complement of $\mathcal{N}(\widetilde{\mathbf{L}})$ is defined as

$$\mathbb{E} = \mathcal{N}(\widetilde{\mathbf{L}})^{\perp_{\mathbf{H}_k}} = \left\{ y \in \mathbb{R}^p \middle| x^T {\mathbf{H}_k}^T \mathbf{H}_k y = 0,\ \forall x \in \mathcal{N}\left(\widetilde{\mathbf{L}}\right) \right\}.$$

The above two subspaces satisfy $\mathbb{E} \oplus \mathbb{F} = \mathbb{R}^p, \mathbb{E} \cap \mathbb{F} = \{0\}$. If we take $\widetilde{\mathbf{L}}^{\dagger} \in \mathbb{R}^{p\times(p-1)}$ to represent the Moore Penrose pseudoinverse[34] of $\widetilde{\mathbf{L}}$, we can further construct the oblique pseudoinverse $\widetilde{\mathbf{L}}^{\dagger}_{\mathrm{H}_k} = \left(\mathbf{I} - \omega(\mathbf{H}_k\omega)^{\dagger}\mathbf{H}_k\right)\widetilde{\mathbf{L}}^{\dagger} \in \mathbb{R}^{p\times(p-1)}$[10], where $\omega \in \mathbb{R}^p$ is the right singular vector *citehan* corresponding to the zero singular values of $\tilde{\mathbf{L}}$ after SVD decomposition[10,32].

Based on the above construction, the oblique projection operator onto $\mathbb{E}$ along $\mathbb{F}$ is defined by $\mathcal{P}_{\mathbb{E},\mathbb{F}} = \widetilde{\mathbf{L}}^{\dagger}_{\mathbf{H}_k}\widetilde{\mathbf{L}} \in \mathbb{R}^{p\times p}$, then the complementary projection operator is given by $\mathcal{P}_{\mathbb{F},\mathbb{E}} = \mathbf{I} - \mathcal{P}_{\mathbb{E},\mathbb{F}}$. Given the above operators, we can write the solution $\mathbf{m} \in \mathbb{R}^p$ to probelm (3) as

$$\mathbf{m} = \mathbf{m}_{\mathbb{E}} + \mathbf{m}_{\mathbb{F}}, \quad \mathbf{m}_{\mathbb{E}} \in \mathbb{E},\ \mathbf{m}_{\mathbb{F}} \in \mathbb{F}, \tag{4}$$

where $\mathbf{m}_{\mathbb{E}} = \mathcal{P}_{\mathbb{E},\mathbb{F}}\mathbf{m}$, $\mathbf{m}_{\mathbb{F}} = \mathcal{P}_{\mathbb{F},\mathbb{E}}\mathbf{m}$. Compared to orthogonal projection, the advantage of this decomposition is that it ensures the orthogonality of the observed operator $\mathbf{H}_k$ mapping under weighted inner product, i.e. $\mathbf{H}_k\mathbf{m}_{\mathbb{E}} \perp \mathbf{H}_k\mathbf{m}_{\mathbb{F}}$. Consequently, $\mathbf{J}_o(\mathbf{m})$ can be decomposed as follows:

$$\begin{aligned} \mathbf{J}_o(\mathbf{m}) &= \mathbf{J}_o(\mathbf{m}_{\mathbb{E}}) + \mathbf{J}_o(\mathbf{m}_{\mathbb{F}}) \\ &= \frac{1}{2}\|\mathbf{H}_k\mathbf{m}_{\mathbb{E}} - \mathbf{y}_k\|_2^2 + \frac{\lambda}{2}\left\|\widetilde{\mathbf{L}}\mathbf{m}_{\mathbb{E}}\right\|_2^2 + \frac{1}{2}\|\mathbf{H}_k\mathbf{m}_{\mathbb{F}} - \mathbf{y}_k\|_2^2. \end{aligned} \tag{5}$$

Then, by introducing the pseudoinverse $\mathcal{P}^{\dagger}_{\mathbb{E},\mathbb{F}}$ of $\mathcal{P}_{\mathbb{E},\mathbb{F}}$, the relation $\mathcal{P}_{\mathbb{E},\mathbb{F}}\mathbf{m} = \mathbf{m}_{\mathbb{E}}$ can be equivalently written as $\mathbf{m} = \mathcal{P}^{\dagger}_{\mathbb{E},\mathbb{F}}\mathbf{m}_{\mathbb{E}}$. Substituting this expression into $\mathbf{J}_b(\mathbf{m})$ in Equation (3) yields $\mathbf{J}_b(\mathbf{m}_{\mathbb{E}}) = \frac{1}{2}\|\mathcal{P}^{\dagger}_{\mathbb{E},\mathbb{F}}\mathbf{m}_{\mathbb{E}} - \mathbf{m}_{k-1}\|^2_{\mathbf{P}^{-1}_{k-1}}$. This reformulates problem (3b) in $\mathbb{R}^p$ with respect to $\mathbf{m}$ as a

 doi:10.20944/preprints202603.1333.v1

least-squares problem in the subspace $\mathbb{E}$ with respect to $\mathbf{m}_{\mathbb{E}}$, whose solution can be obtained by solving the following recursive scheme.

$$\begin{cases} (\mathbf{m}_{\mathbb{E}})_k = (\mathbf{m}_{\mathbb{E}})_{k-1} + \varepsilon_{k-1}, \\ \mathbf{y}_k = (\mathbf{m}_{\mathbb{E}})_k + \mathbf{e}_k, \\ \mathbf{0} = \tilde{\mathbf{L}}(\mathbf{m}_{\mathbb{E}})_k + \mathbf{e}_j, \quad k = 1, 2, ..., N_h, \end{cases} \tag{6}$$

where $(\mathbf{m}_{\mathbb{E}})_{k-1} = \mathcal{P}_{\mathbb{E},\mathbb{F}}\mathbf{m}_{k-1}$ denotes the prior information of $\mathbf{m}_{\mathbb{E}}$, $\varepsilon_{k-1} \sim \mathcal{N}(0, \mathbf{C}_{k-1})$, $\mathbf{C}_{k-1} = \mathcal{P}_{\mathbb{E},\mathbb{F}}\mathbf{P}_{k-1}\mathcal{P}_{\mathbb{E},\mathbb{F}}^T \in \mathbb{R}^{p\times p}$ denotes the prior error covariance satisfied by $\mathbf{m}_{\mathbb{E}}$, $\mathbf{e}_k \sim \mathcal{N}(0, \mathbf{I}_{r\times r})$, and $\mathbf{e}_j \sim \mathcal{N}(0, \lambda^{-1}\mathbf{I}_{(p-1)\times(p-1)})$.

Proceeding analogously, we set $\mathbf{m} = \mathcal{P}_{\mathbb{F},\mathbb{E}}^{\dagger}\mathbf{m}_{\mathbb{F}}$ and substitute it into $\mathbf{J}_b(\mathbf{m})$, thus obtaining a null space optimization problem in terms of $\mathbf{m}_{\mathbb{F}}$.

$$\begin{cases} (\mathbf{m}_{\mathbb{F}})_k = (\mathbf{m}_{\mathbb{F}})_{k-1} + \epsilon_{k-1}, \\ \mathbf{y}_k = \mathbf{H}_k(\mathbf{m}_{\mathbb{F}})_k + \mathbf{e}_k, \quad k = 1, 2, ..., N_h, \end{cases} \tag{7}$$

where $(\mathbf{m}_{\mathbb{F}})_{k-1} = \mathcal{P}_{\mathbb{F},\mathbb{E}}\mathbf{m}_{k-1}$ denotes the prior information of $\mathbf{m}_{\mathbb{F}}$, $\epsilon_{k-1} \sim \mathcal{N}(0, \mathbf{D}_{k-1})$, and $\mathbf{D}_{k-1} = \mathcal{P}_{\mathbb{F},\mathbb{E}}\mathbf{P}_{k-1}\mathcal{P}_{\mathbb{F},\mathbb{E}}^T \in \mathbb{R}^{p\times p}$ denotes the prior error covariance satisfied by $\mathbf{m}_{\mathbb{F}}$.

The above decomposition reformulates the optimization problem in $\mathbb{R}^p$ into two subspace problems. This approach effectively solves the computational difficulties caused by non-square gradient operators and encourages us to continue with Kalman filtering for the solution.

2.2. Iterative Kalman Inversion Framework

Although the two subproblems have been formulated, further details must be considered to facilitate numerical implementation. In subproblem (6), the construction of the weight operator $\mathbf{W}_j$ depends entirely on the update of $\bar{\mathbf{m}} = \mathbf{L}\mathbf{m}$. In this section, we consider obtaining the solution expression of $\bar{\mathbf{m}}$ directly through problem (6).

According to the definition of the pseudoinverse, $\widetilde{\mathbf{L}}\mathbf{m} = \widetilde{\mathbf{L}}\mathbf{m}_{\mathbb{E}}$[10], the oblique complement subproblem (6) can be reformulated as the determination of $\bar{\mathbf{m}} = \mathbf{L}\mathbf{m}_{\mathbb{E}} \in \mathbb{R}^{(p-1)}$, which yields the following expression.

$$\begin{cases} \bar{\mathbf{m}}_k = \bar{\mathbf{m}}_{k-1} + \bar{\varepsilon}_{k-1}, & \text{(8a)} \\ \mathbf{y}_k = \bar{\mathbf{H}}_k\bar{\mathbf{m}}_k + \mathbf{e}_k, & \text{(8b)} \\ \mathbf{0} = \bar{\mathbf{m}}_k + \bar{\mathbf{e}}_j, \quad k = 1, 2, \ldots, N_h, & \text{(8c)} \end{cases}$$

where $\bar{\mathbf{m}}_{k-1} = \mathbf{L}\mathbf{m}_{k-1}$ denotes the prior information of $\bar{\mathbf{m}}$, $\bar{\varepsilon}_{k-1} \sim \mathcal{N}(0, \bar{\mathbf{C}}_{k-1})$, $\bar{\mathbf{C}}_{k-1} = \mathbf{L}\mathbf{P}_{k-1}\mathbf{L}^T \in \mathbb{R}^{(p-1)\times(p-1)}$ denotes the prior error covariance satisfied by $\bar{\mathbf{m}}$, $\bar{\mathbf{H}}_k = \mathbf{H}_k\widetilde{\mathbf{L}}_{\mathbf{H}_k}^{\dagger}\mathbf{W}_j^{-\frac{1}{2}} \in \mathbb{R}^{l\times(p-1)}$, and $\bar{\mathbf{e}}_j \sim \mathcal{N}(0, \lambda^{-1}\mathbf{W}_j)$.

Furthermore, the two subproblems are solved synchronously according to the definition in Section 2.1, with their prior information provided by the full space $\mathbb{R}^p$. In this context, the prior mean and the associated error covariance of $\mathbf{m}$ in $\mathbb{R}^p$ are denoted by $\mathbf{m}_{k-1}^a$ and $\mathbf{P}_{k-1}^a$, respectively. This formulation not only enables the iterative Kalman filtering schemes for subproblems (7) and (8) to be presented separately, but also ensures a strong coupling between them.

Specifically, for the subproblem defined on the oblique complement space $\mathbb{E}$, the prior information is provided by (8a). Based on the linear relationship between $\bar{\mathbf{m}}$ and the prior information in $\mathbb{R}^p$, the corresponding forecast step can be expressed as follows:

$$\bar{\mathbf{m}}_k^f = \mathbf{L}\mathbf{m}_{k-1}^a, \tag{9a}$$

$$\bar{\mathbf{C}}_k^f = \mathbf{L}\mathbf{P}_{k-1}^a\mathbf{L}^T, \tag{9b}$$

 doi:10.20944/preprints202603.1333.v1

where $\bar{\mathbf{m}}_k^f$ and $\bar{\mathbf{C}}_k^f$ denote the forecast mean and the forecast error covariance of $\bar{\mathbf{m}}$ in $\mathbb{E}$, respectively. When $k = N_h$, the weight operator is updated using the forecast mean as $\mathbf{W}_{j+1} = \text{diag}(|(\bar{\mathbf{m}}_{N_h}^f)_1|, |(\bar{\mathbf{m}}_{N_h}^f)_2|, \ldots, |(\bar{\mathbf{m}}_{N_h}^f)_{p-1}|)$, which constitutes the key step in applying the MM strategy within the KF framework to achieve iterative regularization.

Once the forecast information of the state variable $\bar{\mathbf{m}}$ is obtained as in Equation (9), the observational information in (8b) and (8c) can be incorporated by constructing the augmented observation equation $\mathbf{z}_k = \bar{\mathbf{F}}_k \bar{\mathbf{m}}_k^f + \mathbf{e}_{kj}$, where $\mathbf{z}_k = \left[\mathbf{y}_k^T, \mathbf{0}^T\right]^T$, $\bar{\mathbf{F}}_k = \left[\bar{\mathbf{H}}_k^T, \mathbf{I}_{(p-1)\times(p-1)}^T\right]^T$, $\mathbf{e}_{kj} = \left[\mathbf{e}_k^T, \bar{\mathbf{e}}_j^T\right]^T$, hereby enabling the update of the forecast state. The explicit formulation of the observation update step is given below.

$$\bar{\mathbf{m}}_k^a = \bar{\mathbf{m}}_k^f + \mathbf{K_1}\left(\mathbf{z}_k - \bar{\mathbf{F}}_k \bar{\mathbf{m}}_k^f\right), \tag{10a}$$

$$\bar{\mathbf{C}}_k^a = (\mathbf{I}_1 - \mathbf{K}_1 \bar{\mathbf{F}}_k)\bar{\mathbf{C}}_k^f, \tag{10b}$$

where $\bar{\mathbf{m}}_k^a$ and $\bar{\mathbf{C}}_k^a$ denote the analysis mean and the analysis error covariance of $\bar{\mathbf{m}}$, respectively. $\mathbf{I}_1 \in \mathbb{R}^{(p-1)\times(p-1)}$ is the identity matrix and $\mathbf{K}_1$ denotes the Kalman gain matrix associated with the iterative process in $\mathbb{E}$,

$$\mathbf{K}_1 = \bar{\mathbf{C}}_k^f \bar{\mathbf{F}}_k^T \left(\bar{\mathbf{F}}_k \bar{\mathbf{C}}_k^f \bar{\mathbf{F}}_k^T + \mathbf{\Sigma}\right)^{-1}. \tag{11}$$

In this formulation, $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{I}_{r\times r} & 0 \\ 0 & \lambda^{-1}\mathbf{W}_j \end{bmatrix}$ denotes the covariance matrix of the error term $\mathbf{e}_{kj}$ and serves as the pivotal mechanism through which sparse regularization is incorporated into the KF framework.

Subsequently, following the same treatment and incorporating the information in $\mathbb{R}^p$, the forecast step for the nullspace problem (7) in $\mathbb{F}$ can be formulated as follows:

$$(\mathbf{m}_{\mathbb{F}})_k^f = \mathcal{P}_{\mathbb{F},\mathbb{E}} \mathbf{m}_{k-1}^a, \tag{12a}$$

$$\mathbf{D}_k^f = \mathcal{P}_{\mathbb{F},\mathbb{E}} \mathbf{P}_{k-1}^a \mathcal{P}_{\mathbb{F},\mathbb{E}}^T, \tag{12b}$$

where $(\mathbf{m}_{\mathbb{F}})_k^f$ and $\mathbf{D}_k^f$ denote the forecast mean and the forecast error covariance of $\mathbf{m}_{\mathbb{F}}$, respectively.

In contrast to the observation update procedure in the oblique complement space $\mathbb{E}$, the state variable in $\mathbb{F}$ is updated directly from the forecast information according to the observation equation $\mathbf{y}_k = \mathbf{H}_k \mathbf{m}_{\mathbb{F}}^k + \mathbf{e}_k$ without invoking additional regularization. The corresponding update formulas are given by

$$(\mathbf{m}_{\mathbb{F}})_k^a = (\mathbf{m}_{\mathbb{F}})_k^f + \mathbf{K_2}\left(\mathbf{y}_k - \mathbf{H}_k (\mathbf{m}_{\mathbb{F}})_k^f\right), \tag{13a}$$

$$\mathbf{D}_k^a = (\mathbf{I}_2 - \mathbf{K}_2 \mathbf{H}_k)\mathbf{D}_k^f, \tag{13b}$$

where $(\mathbf{m}_{\mathbb{F}})_k^a$ and $\mathbf{D}_k^a$ represent the analysis mean and the analysis error covariance of $\mathbf{m}_{\mathbb{F}}$, respectively. $\mathbf{I}_2 \in \mathbb{R}^{p\times p}$ is the identity matrix and the Kalman gain matrix associated with the iterative process in $\mathbb{F}$ is defined as $\mathbf{K}_2 = \mathbf{D}_k^f \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{D}_k^f \mathbf{H}_k^T + \mathbf{I}_{r\times r}\right)^{-1}$.

Based on the above Kalman filtering procedures, the analysis means and analysis error covariances for the two subproblems can be obtained separately. According to the decomposition provided in Equation (4), $\bar{\mathbf{m}}_k^a$ and $(\mathbf{m}_{\mathbb{F}})_k^a$ are subsequently combined to construct the prior mean in $\mathbb{R}^p$ for the next computational step. Considering that the elements in the two subspaces are not statistically

 doi:10.20944/preprints202603.1333.v1

independent, the prior error covariance at the next step involves $\mathbf{D}_k^a$ and $\bar{\mathbf{C}}_k^a$, as well as their cross-covariance terms. The specific update formulation is given as follows:

$$\mathbf{m}_k^a = \tilde{\mathbf{L}}_{\mathrm{H}_k}^{\dagger} \mathbf{W}_j^{-\frac{1}{2}} \bar{\mathbf{m}}_k^a + (\mathbf{m}_{\mathbb{F}})_k^a, \tag{14a}$$

$$\mathbf{P}_k^a = (\tilde{\mathbf{L}}_{\mathrm{H}_k}^{\dagger} \mathbf{W}_j^{-\frac{1}{2}}) \bar{\mathbf{C}}_k^a (\tilde{\mathbf{L}}_{\mathrm{H}_k}^{\dagger} \mathbf{W}_j^{-\frac{1}{2}})^T + \mathbf{D}_k^a + \mathcal{P}_{\mathbb{E},\mathbb{F}} \mathbf{P}_{k-1}^a \mathcal{P}_{\mathbb{F},\mathbb{E}}^T + \mathcal{P}_{\mathbb{F},\mathbb{E}} \mathbf{P}_{k-1}^a \mathcal{P}_{\mathbb{E},\mathbb{F}}^T. \tag{14b}$$

Finally, by repeatedly evaluating Equations (9)-(14), $\mathbf{\Sigma}$ is updated iteratively, driving $\bar{\mathbf{m}}$ to progressively approximate the sparsity pattern associated with l_1 regularization. It should be emphasized that the regularization parameter λ embedded in $\mathbf{\Sigma}$ plays a crucial role in regulating the sparsity level, and its selection has a significant impact on the overall regularization performance.

2.3. Adaptive Regularization Parameter Selection

When applying the KF method for iterative sparse inversion, the selcetion of the regularization parameter λ determines the sparsity of the solution $\bar{\mathbf{m}}$ in the oblique complement space $\mathbb{E}$, and also influences the overall inversion quality of the solution $\mathbf{m}$ in $\mathbb{R}^p$. Since the initial conditions of the inversion problem are different and the solution tends to stabilize in the later stages of iteration, a fixed λ is often insufficient to accommodate the evolving regularization requirements. Therefore, an adaptive parameter selection strategy is necessary.

The traditional L-curve method determines the regularization parameter λ by balancing the residual norm and the solution norm, but the identification of the inflection point is often sensitive to the geometry of the curve, which limits the stability of curvature-based selection[35]. To address this limitation, the f-slope method provides a relatively stable strategy for selecting the regularization parameter. Specifically, it constructs the curve of the solution norm with respect to the regularization parameter and selects λ corresponding to the minimum slope of the curve. The reliability of this approach has been well validated[33].

Similar to the L-curve method, the f-slope method can also generate a candidate range for λ based on the singular value spectrum of the matrix $\bar{\mathbf{H}}_k$. Nevertheless, by setting the upper bound of the interval according to the maximum singular value, this approach tends to produce overly sparse regularization, which may degrade the inversion quality of the solution $\mathbf{m}$ in this paper (see Figure 7(f) for details). Inspired by the w-GCV method, the introduction of an interval constraint can effectively reduce the parameter search range and significantly improve selection efficiency. Accordingly, in this work, an interval constrained strategy is adopted to enhance the algorithm.

To effectively incorporate the iterative Kalman filter inversion framework with adaptive regularization parameter selection, the initial range of regularization parameters is first defined as $[\lambda_{\min}, \lambda_{\max}]$. Then, at iteration j, the optimal λ_k is determined based on the current observation $\mathbf{y}_k$ and the model operator $\bar{\mathbf{H}}_k$, following the slope function definition employed in the f-slope method:

$$H(\lambda) = \sum_{i=1}^{N} \gamma_i^2 \frac{\sigma_i^4 \lambda^2}{(\sigma_i^2 + \lambda^2)^3}, \tag{15}$$

where σ_i is the singular value of matrix $\hat{\mathbf{H}}_k$, which is obtained by SVD: $\hat{\mathbf{H}}_k = \sum_{i=1}^{N} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$, $\gamma_i = \mathbf{u}_i^T \mathbf{y}_k / \sigma_i$ is projection coefficient, and $\hat{\mathbf{H}}_k = \bar{\mathbf{H}}_k \mathbf{W}_j^{\frac{1}{2}} = \mathbf{H}_k \tilde{\mathbf{L}}_{\mathrm{H}_k}^{\dagger}$. Simultaneously, once λ_k is determined at each iteration, the upper bound of the search interval is updated to this value. The specific update is given by $[\lambda_{\min}, \lambda_{\max}^{(j)}] = [\lambda_{\min},\ \lambda_k]$. In addition, to prevent over sparsity and ensure effective sparsity, the upper limit of the search interval $\lambda_{\max}^{(j)} = \min(\lambda_k, \mu\sigma_{\max})$ is set in combination with σ_i when updating λ_k, where $0 < \mu < 1$ while lower limit $\lambda_{\min} = \max(\lambda_{\min}, \sigma_{\min})$. Finally, each iteration is updated with $\lambda_{\max}^{(j+1)} = \lambda_{\max}^{(j)}$, until the maximum number of iterations stops.

Based on the above improvements, the regularization parameters can be adaptively selected within the iterative Kalman filter framework. In addition, the iterative regularization process requires

 doi:10.20944/preprints202603.1333.v1

a well-defined convergence criterion[36]. According to Kalman filtering theory, the asymptotic decay of the gain matrix causes the analysis increment to approach zero. Therefore, in this work, the l_2 norm (root mean square error) of the difference between the analysis and the forecast of the state variable $\bar{\mathbf{m}}$ is adopted as the stopping criterion for the algorithm, that is

$$\left\| \bar{\mathbf{m}}_k^a - \bar{\mathbf{m}}_k^f \right\|_2 < \eta, \tag{16}$$

where η is the preset threshold.

The procedure for initial condition inversion using the Adaptive iterative Kalman filter inversion method is summarized in Algorithm 2.

Algorithm 1 Adaptive selection of regularization parameters $\lambda_* = \mathcal{G}(\mathbf{H}, \mathbf{y}, \lambda_{\min}, \lambda_{\max}, \mu)$

Input: $\mathbf{H}$, $\mathbf{y}$, $\lambda_{\min}$, $\lambda_{\max}$, μ

SVD: $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^T = \sum_{i=1}^{N} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$, $\gamma_i = \mathbf{u}_i^T \mathbf{y}/\sigma_i$
$\lambda_{\max} = \min(\lambda_{\max}, \mu\sigma_{\max})$; $\lambda_{\min} = \max(\lambda_{\min}, \sigma_{\min})$
$\mathbb{I} = [\lambda_{\min}, \lambda_{\max}]$
for $\lambda_* \in \mathbb{I}$ **do**
  $H(\lambda_*) = \sum_{i=1}^{N} \gamma_i^2 \frac{\sigma_i^4 \lambda_*^2}{(\sigma_i^2 + \lambda_*^2)^3}$
end for
Select the smallest λ_* in $H(\lambda)$: $\lambda_* = \underset{\lambda_* \in \mathbb{I}}{\operatorname{argmin}} H(\lambda_*)$

Output: regularization parameter λ_*

Algorithm 2 Adaptive iterative Kalman filter inversion (AIKFI)

Input: $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_n]^T$, $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n]^T$, $\mathbf{L}$, $\mathbf{m}_0 = \mathbf{m}^b$, $\mathbf{P}_0 = \mathbf{B}$, $\lambda_{\max}^{(0)} = \lambda_{\max}$, $\lambda_{\min}$, μ, η

for $j = 1 : iter$ **do**
  $\mathbf{m}_j = \mathbf{m}_{j-1}$
  $\mathbf{W}_j = \operatorname{diag}(|(\mathbf{L}\mathbf{m}_j)_1|, \ldots, |(\mathbf{L}\mathbf{m}_j)_{p-1}|)$
  $\mathbf{P}_j = \mathbf{P}_{j-1}$
  $\mathbf{m}_0^a = \mathbf{m}_j$, $\mathbf{P}_0^a = \mathbf{P}_j$; $\lambda_0 = \lambda_{\max}^{(j)}$
  for $k = 1 : N_h$ **do**
    Oblique projection: $\mathbf{m}_{k-1}^a$, $\mathbf{P}_{k-1}^a \begin{cases} \xrightarrow{\text{formula}(9)} & \bar{\mathbf{m}}_k^f,\ \bar{\mathbf{C}}_k^f \\ \xrightarrow{\text{formula}(12)} & (\mathbf{m}_{\mathbb{F}})_k^f,\ \mathbf{D}_k^f \end{cases}$
    $\lambda_k = \mathcal{G}(\mathbf{H}_k \widetilde{\mathbf{L}}_{\mathbf{H}_k}^{\dagger}, \mathbf{y}_k, \lambda_{\min}, \lambda_{k-1}, \mu)$, see algorithm 2
    $\Sigma = \begin{bmatrix} \mathbf{I}_{r\times r} & 0 \\ 0 & \lambda_k^{-1}\mathbf{W}_j \end{bmatrix} \xrightarrow{\text{formula}(11)} \text{formula}(10)$
    Observation update:
    $\bar{\mathbf{m}}_k^f$, $\bar{\mathbf{C}}_k^f \xrightarrow{\text{formula}(10)} \bar{\mathbf{m}}_k^a$, $\bar{\mathbf{C}}_k^a$
    $(\mathbf{m}_{\mathbb{F}})_k^f$, $\mathbf{D}_k^f \xrightarrow{\text{formula}(13)} (\mathbf{m}_{\mathbb{F}})_k^a$, $\mathbf{D}_k^a$
    Combined analysis value and error covariance:
    $\bar{\mathbf{m}}_k^a$, $(\mathbf{m}_{\mathbb{F}})_k^a \xrightarrow{\text{formula}(14a)} \mathbf{m}_k^a$
    $\bar{\mathbf{C}}_k^a$, $\mathbf{D}_k^a \xrightarrow{\text{formula}(14b)} \mathbf{P}_k^a$
    while $\left\| \bar{\mathbf{m}}_k^a - \bar{\mathbf{m}}_k^f \right\|_2 < \eta$ **do**
      Break
    end while
  end for
  $\mathbf{m}_j = \mathbf{m}_{N_h}^a$, $\mathbf{P}_j = \mathbf{P}_{N_h}^a$; $\lambda_{\max}^{(j)} = \lambda_{N_h}$
end for

Output: Optimal analysis value $\mathbf{m}_j$, regularization parameter $\lambda_{\max}^{(j)}$

3. Comparative Experiment Based on Linear Advection Diffusion Equation

To verify the reliability and effectiveness of the proposed algorithm and facilitate comparison with existing methods, the initial conditions reported in [5] are adopted, while the advection-diffusion equation serves as the numerical model.

 doi:10.20944/preprints202603.1333.v1

3.1. Linear Advection Diffusion Equation

The one-dimensional advection-diffusion equation can be regarded as a linear Burgers equation with a constant advection velocity. The initial value problem for the state variable $m(x,t)$ is given by:

$$\begin{cases} \dfrac{\partial m(x,t)}{\partial t} + v\dfrac{\partial m(x,t)}{\partial x} = \Upsilon\dfrac{\partial^2 m(x,t)}{\partial x^2}, & (x,t) \in [0,L]\times[0,T], \quad (17a) \\ m(0,t) = m(L,t), & t \in [0,T], \quad (17b) \\ m_x(0,t) = m_x(L,t), & t \in [0,T], \quad (17c) \\ m(x,0) = \phi(x), & x \in [0,L], \quad (17d) \end{cases}$$

where v denotes the advection velocity and Υ denotes the diffusion coefficient. In this paper, we set $L = 4\pi$, $T = 1$, $v = 2$ and $\Upsilon = 0.2$. $\phi(x)$ is the initial condition, two representative types are considered in this study. The first type is the Flat Top Hat (FTH) condition with

$$\phi(x) := \phi^{FTH}(x) = \begin{cases} 1, & x \in [0,\beta_1)\cup(\beta_2,L], \\ 2, & x \in [\beta_1,\beta_2], \end{cases} \tag{18}$$

and the second type is the Windowed Sine (WS) condition with

$$\phi(x) := \phi^{WS}(x) = \begin{cases} 1, & x \in [0,\beta_1)\cup(\beta_2,L], \\ 1 - \frac{1}{2}\sin\left(2\pi\dfrac{x-\beta_1}{\beta_2-\beta_1}\right), & x \in [\beta_1,\beta_2]. \end{cases} \tag{19}$$

To facilitate the subsequent numerical experiments, the problem is solved using a numerical discretization. The spatial domain is first discretized into p equidistant nodes, denoted by $x_s = s\Delta x$, $s = 0,1,\ldots,p-1$, where $p = 1024$ and the spatial step size is $\Delta x = \frac{L}{p}$. In Equations (18) and (19), the parameters are set as $\beta_1 = 180\Delta x$ and $\beta_2 = 300\Delta x$. The solution vector at the discrete points is defined as $\mathbf{m}(t) = [m(x_0,t), m(x_1,t), \ldots, m(x_{p-1},t)]^T$. The periodic boundary conditions are given by $m(x_{-1},t) \equiv m(x_{p-1},t)$ and $m(x_0,t) \equiv m(x_p,t)$. Using the central difference scheme, Equation (17) can then be transformed into the following initial-value problem of ordinary differential equations.

$$\begin{cases} \dfrac{d\mathbf{m}}{dt} = \mathbf{D}\mathbf{m}, & (20a) \\ \mathbf{m}(0) = [\phi(x_1), \phi(x_2), \ldots, \phi(x_{p-1})]. & (20b) \end{cases}$$

At this point, the solution to Equation (20) can be expressed as:

$$\mathbf{m}(t) = \mathcal{M}_{0,t}\mathbf{m}(0) = \mathbf{F}e^{\mathbf{\Lambda}t}\mathbf{F}^{-1}\mathbf{m}(0), \quad t \in [0,T], \tag{21}$$

where $\mathcal{M}_{0,t} \in \mathbb{R}^{p\times p}$ denotes the linear evolution operator, which can be expressed via the eigenvalue decomposition of the matrix $\mathbf{D}$ as $\mathbf{D} = \mathbf{F}\mathbf{\Lambda}\mathbf{F}^{-1}$. Here, $\mathbf{\Lambda} = \mathrm{diag}(\Lambda_0, \Lambda_1, \ldots, \Lambda_{p-1})$ is a diagonal matrix whose diagonal entries are the eigenvalues Λ_s of $\mathbf{D}$, given by $\Lambda_s = -\frac{4\Upsilon}{\Delta x^2}\sin^2\left(\frac{\pi s}{p}\right) + i\frac{v}{\Delta x}\sin\left(\frac{2\pi s}{p}\right)$ for $s = 0,1,\ldots,p-1$, where i is the imaginary unit. $\mathbf{F} \in \mathbb{C}^{p\times p}$ is the discrete Fourier transform matrix with entries $\mathbf{F}_{k,j} = e^{\frac{-2\pi ikj}{p}}$, satisfying $\mathbf{F}^{-1} = \frac{1}{p}\mathbf{F}^H$, the superscript H denoting the conjugate transpose.

3.2. Parameter Settings and Evaluation Metrics Selection

Once the accurate initial conditions are given as described in Equations (18) and (19), the state variable $\mathbf{m}(t)$ will provide correct forecast results in time under the forward prediction model (21). In practical applications, the initial conditions are usually difficult to obtain due to some factors such as estimation uncertainty or measurement inaccuracy. In such cases, it becomes necessary to correct the initial conditions using observational data. The initial guess to be corrected in this paper

 doi:10.20944/preprints202603.1333.v1

is $\phi_g(x) = \overline{\phi(x)} + \varepsilon$, where ε follows a Gaussian distribution $\mathcal{N}(0, \mathbf{B})$, and $\overline{\phi(x)}$ is set to either $\phi(x)$ or $\phi(x+\theta)+\vartheta$ (with $\theta = 10$, $\vartheta = -0.2$ in this study). The background error covariance is uniformly expressed as $\mathbf{B} = \sigma_b^2 \mathbf{C_B} \in \mathbb{R}^{p\times p}$, where $\sigma_b = 0.1$,

$$\mathbf{C_B} = \mathbf{C_{B_1}} = \mathbf{I}, \quad \mathbf{C_B} = \mathbf{C_{B_2}} = \begin{bmatrix} \rho(0) & \rho(1) & \cdots & \rho(p) \\ \rho(1) & \rho(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho(1) \\ \rho(p) & \cdots & \rho(1) & \rho(0) \end{bmatrix}. \tag{22}$$

Specifically, $\mathbf{C_{B_1}}$ corresponds to a white noise model with a small condition number, indicating that the state variables are uncorrelated across spatial grid points. In contrast, $\mathbf{C_{B_2}}$ represents a colored noise model with a larger condition number, reflecting that the state variables exhibit spatial correlations. Its structure function is given by $\rho(\tau) \propto e^{-\alpha|\tau|}(1+\alpha|\tau|)$, where τ denotes the spatiotemporal lag and α is a parameter controlling the decay rate of correlation. The characteristic correlation length of the model is therefore α^{-1}.

In this section, three types of initial guesses are configured as follows:

$$\phi_g^1 = \phi^{FTH}(x) + \varepsilon_1, \qquad \phi_g^2 = \phi^{WS}(x) + \varepsilon_1, \tag{23a}$$

$$\phi_g^3 = \phi^{FTH}(x+\theta) + \vartheta + \varepsilon_1, \qquad \phi_g^4 = \phi^{WS}(x+\theta) + \vartheta + \varepsilon_1, \tag{23b}$$

$$\phi_g^5 = \phi^{FTH}(x) + \varepsilon_2, \qquad \phi_g^6 = \phi^{WS}(x) + \varepsilon_2, \tag{23c}$$

where $\varepsilon_1 \sim \mathcal{N}(0, \sigma_b^2 \mathbf{C_{B_1}})$, $\varepsilon_2 \sim \mathcal{N}(0, \sigma_b^2 \mathbf{C_{B_2}})$. The initial guesses in Equation (23a) correspond to a background error covariance matrix with a small condition number. Those in Equation (23b) represent a large deviation between the background and true values, while those in Equation (23c) correspond to a background error covariance matrix with a large condition number.

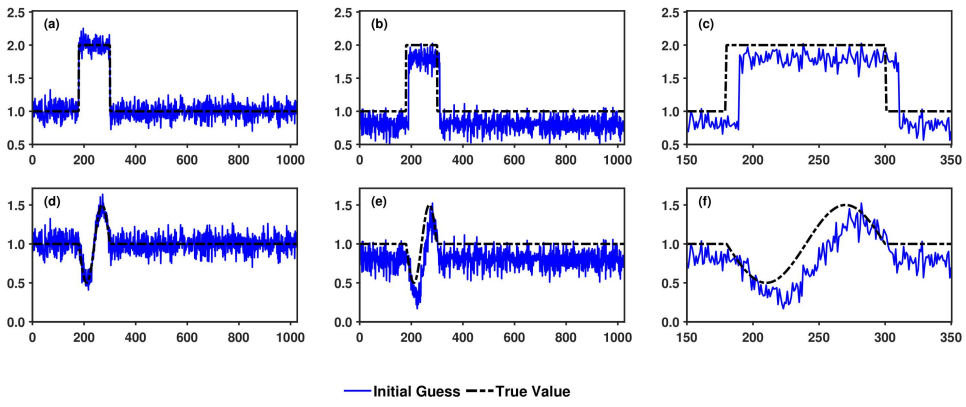


Figure 1. Different types of initial guesses. (a) and (d) correspond to ϕ_g^1 and ϕ_g^2. (b) and (e) correspond to ϕ_g^3 and ϕ_g^4. (c) and (f) show enlarged views of ϕ_g^3 and ϕ_g^4 over the interval $x \in [150\Delta x, 350\Delta x]$. The horizontal axis represents the spatial node index s in all subfigures, and the vertical axis denotes the corresponding solution $\phi(x_s)$, $s = 0, 1, \ldots, 1023$.

At this stage, the initial guess of the model must be corrected using actual observational data. Accordingly, the assimilation time window is uniformly divided into $N_t + 1$ time instants $t_n = n\Delta t$, $n = 0, 1, 2, \ldots, N_t$, with the number of time steps set to $N_t = 500$ and the time step size is $\Delta t = \frac{T}{N_t}$. Observations are taken at equally spaced intervals $t_k = t_0 + (k-1)\Delta h$, $k = 1, 2, \ldots, N_h$, where $\Delta h = 125\Delta t$ and the number of observations is $N_h = 5$. The observation operator $\mathcal{H} \in \mathbb{R}^{r\times p}$, with

 doi:10.20944/preprints202603.1333.v1

$r = \frac{p}{4} = 256$, is defined to map information from the high-dimensional state space to the low-dimensional observation space. Its explicit form is

$$\mathcal{H} = \frac{1}{4}\begin{bmatrix} 1,1,1,1, & 0,0,0,0, & \cdots & 0,0,0,0 \\ 0,0,0,0, & 1,1,1,1, & \cdots & 0,0,0,0 \\ \ddots & \vdots & \vdots & \vdots \\ 0,0,0,0, & 0,0,0,0, & \cdots & 1,1,1,1 \end{bmatrix}. \tag{24}$$

The observation data $\mathbf{b}_k$ are obtained from the observation equation $\mathbf{b}_k = \mathcal{H}\mathbf{m}_k + \mathbf{e}_k$, where $\mathbf{e}_k \sim \mathcal{N}(0, \mathbf{R})$ denotes the observation noise with standard deviation $\sigma_r = 0.08$. That is, the error covariance at all observation times is set to $\mathbf{R} = \sigma_r^2\mathbf{I} \in \mathbb{R}^{r\times r}$.

Three error metrics are employed in this study: the relative mean squared error (MSE_r), the relative mean absolute error (MAE_r), and the relative bias (BIAS_r).

$$\begin{aligned} \mathrm{MSE}_r &= \left\|m_0^t - m_0^a\right\|_2 / \left\|m_0^t\right\|_2, \\ \mathrm{MAE}_r &= \left\|m_0^t - m_0^a\right\|_1 / \left\|m_0^t\right\|_1, \\ \mathrm{BIAS}_r &= \left|E\left(m_0^t - m_0^a\right)\right| / \left|E\left(m_0^t\right)\right|, \end{aligned} \tag{25}$$

where m_0^t denotes the true initial value, m_0^a is the analysis, and $E(\cdot)$ represents the expected value. In fact, both MAE_r and MSE_r are average errors. MSE_r is more sensitive to outliers and thus generally larger than MAE_r, while BIAS_r measures the degree to which the analysis deviates from the truth.

3.3. Experimental Results

In this subsection, a comparison is first made between AIKFI and 4DVar, which illustrates the feasibility of the algorithm. Meanwhile, the process of adaptive selection of the regularization parameter is carried out to display its active role during two types of initial conditions. Then, the adaptability of the AIKFI algorithm was investigated using colored-noise models with different correlation lengths. This analysis highlights the advantages of the KF framework in simultaneously improving the analysis state estimate and providing the corresponding analysis error covariance. In addition, the influence of the regularization parameter on the algorithm's ability to correct the evolution from the initial state was further investigated.

The AIKFI algorithm aims to achieve l_1 sparsity of $\bar{\mathbf{m}}$ iteratively by adaptively adjusting the regularization parameter, thereby reconstructing the initial condition $\mathbf{m}(0)$ via Equation (14a). To validate the effectiveness of the proposed algorithm, several typical initial guesses are selected for correction, with the preset threshold uniformly set to $\eta = 10^{-3}$.

First, the initial guesses ϕ_g^1 and ϕ_g^2 are adopted to verify the feasibility of the algorithm under the setting of a background error covariance matrix with a small condition number. The details of these initial guesses are illustrated in Figure 1(a,d) and the corresponding correction results are compared with those obtained by 4D-Var. It is found that both the AIKFI and the l_1-norm R4D-Var[5] can effectively retrieve key features. The three error metrics (MSE_r, MAE_r, and BIAS_r) of the inversion results are presented in Table 1. For WS-type initial guesses, the AIKFI algorithm yields slightly better results than l_1-norm R4D-Var across all error metrics. For FTH-type initial guesses, the MAE_r of AIKFI is slightly higher than that of l_1-norm R4D-Var, but it still achieves an accuracy on the order of $O(10^{-2})$. This behavior arises because, unlike l_1 regularization, l_2 regularization compresses the state variable $\bar{\mathbf{m}}$ toward zero without forcing exact sparsity, producing a smoothing effect when inverting initial guesses with discontinuous features such as FTH.

The above procedure demonstrates the feasibility of the AIKFI algorithm in inverting initial conditions. To further investigate the adaptive capacity of the algorithm, the initial guesses ϕ_g^3 and ϕ_g^4 are selected for experimentation, as shown in Figure 1(b,e). In practical applications, insufficient or incomplete information of the initial conditions may readily lead to such types of initial guesses.

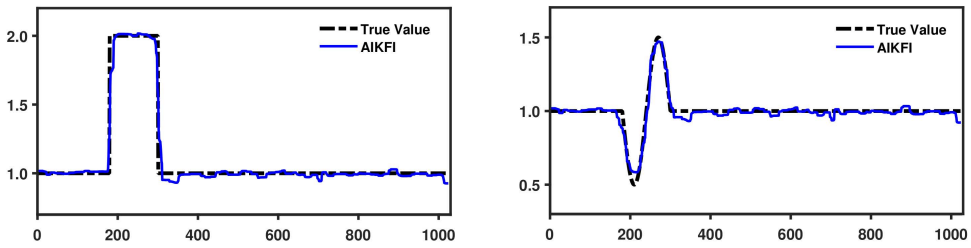
 doi:10.20944/preprints202603.1333.v1

Nevertheless, the developed algorithm is still capable of effectively retrieving the key features, with the inversion results presented in Figure 2(a,b).

Table 1. Comparison of the inversion results obtained by the AIKFI and l_1-norm R4D-Var algorithms for the initial guesses ϕ_g^1 and ϕ_g^2 in terms of MSE_r, MAE_r, and BIAS_r.

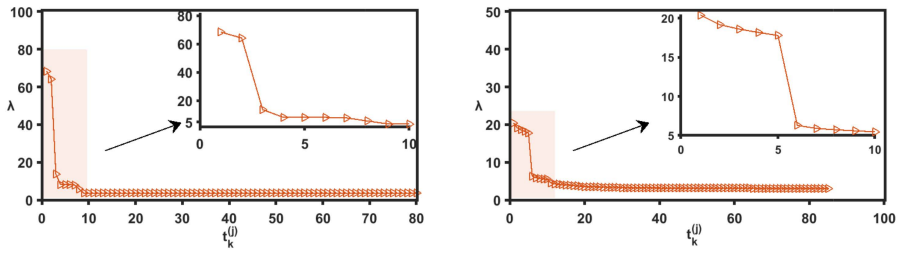
	$\mathbf{MSE}_r$		$\mathbf{MAE}_r$		$\mathbf{BIAS}_r$	
	AIKFI	R4D-Var	AIKFI	R4D-Var	AIKFI	R4D-Var
FTH	0.0174	0.0185	0.0104	0.0091	0.0014	0.0038
WS	0.0241	0.0288	0.0161	0.0215	0.0016	0.0037

During the correction process of initial guesses such as ϕ_g^3 and ϕ_g^4, the adaptive selection scheme for the regularization parameter employed by the AIKFI algorithm plays a crucial role. In this set of experiments, for the FTH-type initial guess, the initial selection interval of the regularization parameter is set to $[5, 100]$ with $\mu = 0.1$. As shown in Figure 3(a), the regularization parameter gradually decreases from 70 to 5 as the iteration proceeds, and remains close to 5 for computational times $t_k^{(j)} \geq 10$. In the experiment for the WS-type initial guess, the initial selection interval is set to $[1, 50]$ with $\mu = 0.03$ while the regularization parameter gradually decreases from 21 to 2 and stabilizes around 2 for $t_k^{(j)} \geq 20$, as depicted in Figure 3(b). These results indicate that the proposed adaptive algorithm automatically reduces the regularization parameter as the sparsity of the iterative solution increases, thereby demonstrating its adaptability under different initial conditions.



(a) The analysis value of ϕ_g^3 (b) The analysis value of ϕ_g^4

Figure 2. Correction results of the initial guesses ϕ_g^3 and ϕ_g^4 via the AIKFI algorithm.



(a) (b)

Figure 3. Adaptive regularization parameter selection during the correction process using the AIKFI algorithm. (a) Regularization parameter selection for the correction of ϕ_g^3. (b) Regularization parameter selection for the correction of ϕ_g^4. The horizontal axis denotes the computation time, whose relationship with the iteration index j and the observation time k is given by $t_k^{(j)} = k(j-1) + k$.

Next, under the background error covariance matrix with a large condition number, the initial guesses are taken as ϕ_g^5 and ϕ_g^6. For the correction experiment with ϕ_g^5, the correlation length in the structure function of the background error covariance matrix is set to $\alpha^{-1} = 25$, resulting in a condition number of 10^7 for the matrix. The initial interval for the regularization parameter is chosen to be $[20, 100]$ with $\mu = 0.5$. For the correction experiment with ϕ_g^6, the correlation length is set to $\alpha^{-1} = 50$, corresponding to a condition

 doi:10.20944/preprints202603.1333.v1

number of 10^8, and the initial interval for the regularization parameter is $[1, 50]$ with $\mu = 0.03$. The details of these two initial guesses are illustrated in Figure 4(a,d), respectively.

Figure 4 illustrates the iterative optimization process of this group of experiments. It can be observed that the initial guess deviates significantly from the true value and exhibits large uncertainty. After the iterative optimization via the AIKFI algorithm, not only are the key features effectively inverted, but the uncertainty of the state variables is also remarkably reduced, which preliminarily validates the numerical stability of the AIKFI algorithm under the background error covariance matrix with a large condition number.

Figure 5 further presents the sparsification effect in the gradient space corresponding to the initial guess correction results in Figure 4. It can be seen that the initial guesses ϕ_g^5 and ϕ_g^6 have no sparsity in the gradient space, and they gradually become sparse in the gradient space after the iterative optimization by the AIKFI algorithm, indicating that the corrected initial guess becomes gradually smooth, which is consistent with the results shown in Figure 4.

These results indicate that the proposed algorithm remains effective for large condition numbers of the background error covariance matrix, demonstrating its adaptability beyond the small-condition-number case. To further investigate the relationship between the inversion results and the condition number of the background error covariance matrix, the colored noise model $\mathbf{B} = \sigma_b^2 \mathbf{C}_{\mathbf{B}_2}$ is adopted as the background error covariance matrix, with a relatively wide range of applicable correlation lengths, $\alpha^{-1} \in \{1, 5, 25, 50, 100, 500, 1000, 3000\}$, which corresponding to the condition number of the background error covariance matrices ranging from 10^2 to 10^{14}.

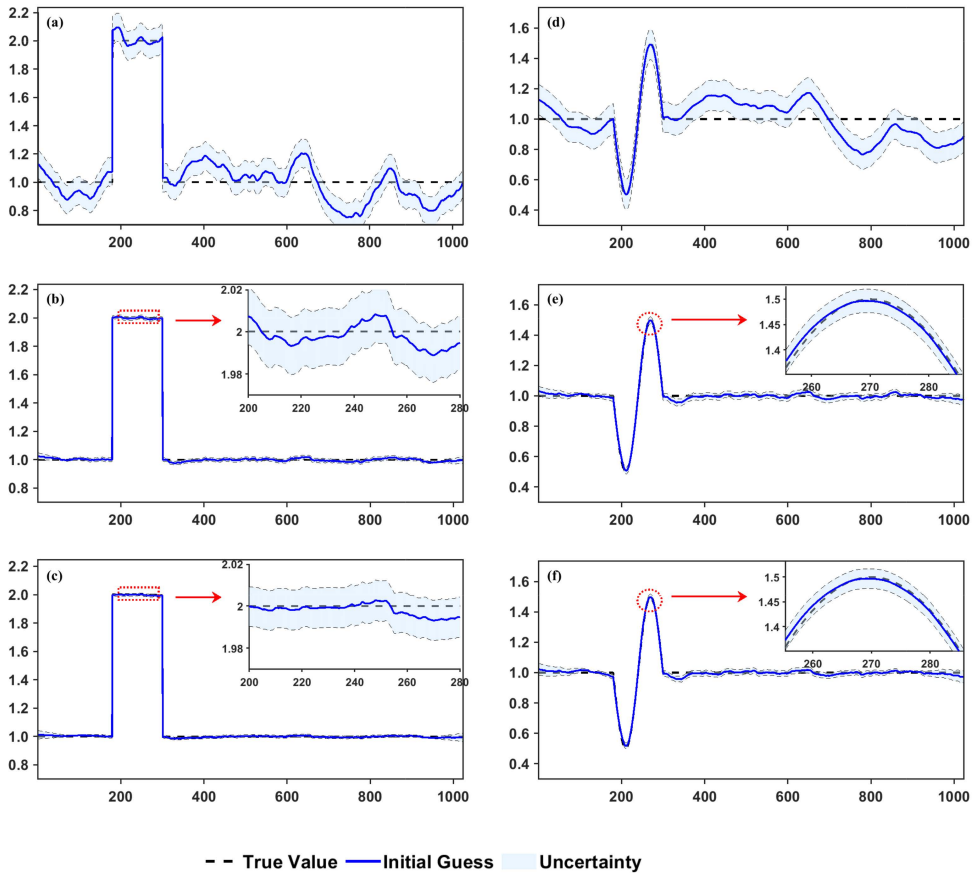
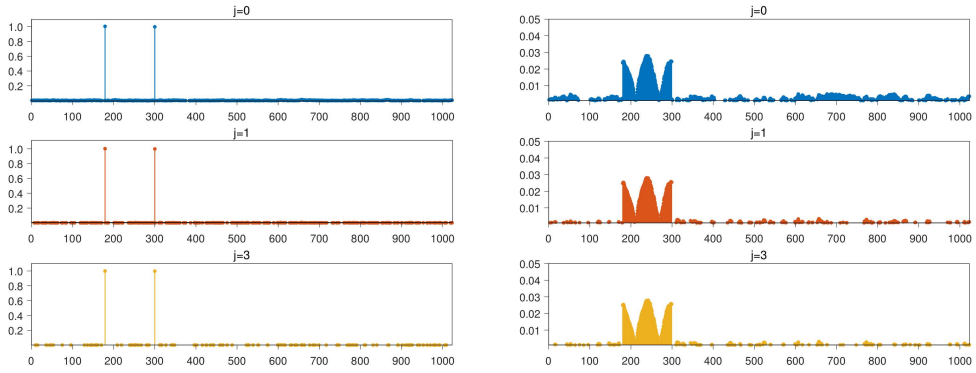


Figure 4. Iterative assimilation performance and uncertainty evolution during the correction process using the AIKFI algorithm. (a) and (d) show the initial guesses ϕ_g^5 and ϕ_g^6 together with their uncertainties, where the correlation lengths in the structure function of the background error covariance matrix are set to $\alpha^{-1} = 25$ and $\alpha^{-1} = 50$, respectively. (b) and (e) present the assimilation results and their uncertainties after 1st iteration of the AIKFI algorithm for ϕ_g^5 and ϕ_g^6. (c) and (f) display the final results, i.e., the assimilation results and their uncertainties after 3rd iteration.

 doi:10.20944/preprints202603.1333.v1



(a) (b)

Figure 5. Sparsity evolution in the gradient information during the iterative correction process using the AIKFI algorithm. (a) Sparsity evolution for the correction of the initial guess ϕ_g^5 in Figure 4(a). (b) Sparsity evolution for the correction of the initial guess ϕ_g^6 in Figure 4(d). The j denotes the iteration index. The horizontal axis represents point $s = 1, 2, \ldots, 1023$, while the vertical axis denotes the corresponding gradient information of the initial guess $\bar{\mathbf{m}} = \mathbf{Lm} \in \mathbb{R}^{1023}$.

Within the above range of condition numbers for the background error covariance matrix, as shown in Figure 6 for both FTH-type and WS-type initial guesses. Unlike l_1-norm R4D-Var, where all three types of errors gradually increase with the condition number, AIKFI algorithm shows a gradual decrease in error. Specifically, the MAE_r and MSE_r of the results corrected by the AIKFI algorithm are both below 0.035 and decrease further as the condition number increases. BIAS_r shows a stable trend as the condition number increases, and always remains below 0.005. In the colored noise model, this phenomenon arises because a higher condition number of the background error covariance matrix corresponds to a larger correlation length α^{-1}, which leads to stronger correlations among the elements in the matrix. This causes the initial guess to become smoother while gradually deviating from the true state, which is intuitively reflected in Figure 4(d). The algorithm in this paper is more robust to ill-conditioned phenomena and is therefore unaffected by the ill-conditioned matrix within the selected range.

It can thus be concluded that the proposed algorithm remains effective for background error covariance matrices with condition numbers ranging from 10^2 to 10^{14}. Moreover, according to the correction experiments for the initial guesses ϕ_g^3 and ϕ_g^4, the regularization parameters play a crucial role in the adaptability of an algorithm under different initial conditions.

Therefore, to further verify the crucial role of the adaptive regularization parameter selection scheme in improving inversion quality, two sets of experiments were conducted. In the first set, the initial guess ϕ_g^5, which is shown in Figure 4(a), was corrected using the AIKFI algorithm with a fixed regularization parameter $\lambda = 10$, and the results were compared with those shown in Figure 4(c), where the regularization parameter was selected adaptively. Regardless of whether adaptive regularization was employed, the AIKFI algorithm was able to correct the initial condition evolution up to $T = 3$, which significantly exceeds the assimilation window $[0, 1]$, as illustrated in Figures 7(b) and 7(c). This observation indicates that the adaptive algorithm proposed in this paper can gradually reduce the degree of regularization required by the problem during the iterative process, even when the upper bound of the regularization parameter is very large.

In the second set of experiments, the initial guess ϕ_g^6, which corresponds to the case shown in Figure 4(d), was corrected using the AIKFI algorithm in which the regularization parameter was selected adaptively by the f-slope algorithm without interval restrictions, and the results were compared with those presented in Figure 4(f). As shown in Figures 7(e,f), the f-slope algorithm without interval restrictions fails to achieve the performance of the interval-restricted f-slope algorithm, which is proposed in this study. Specifically, the key features of the state cannot be effectively recovered, which consequently prevents the algorithm from producing accurate temporal evolution. These results

 doi:10.20944/preprints202603.1333.v1

demonstrate that the adaptive scheme adopted in this study improves the adaptability of the AIKFI algorithm, thereby enhancing the overall inversion quality.

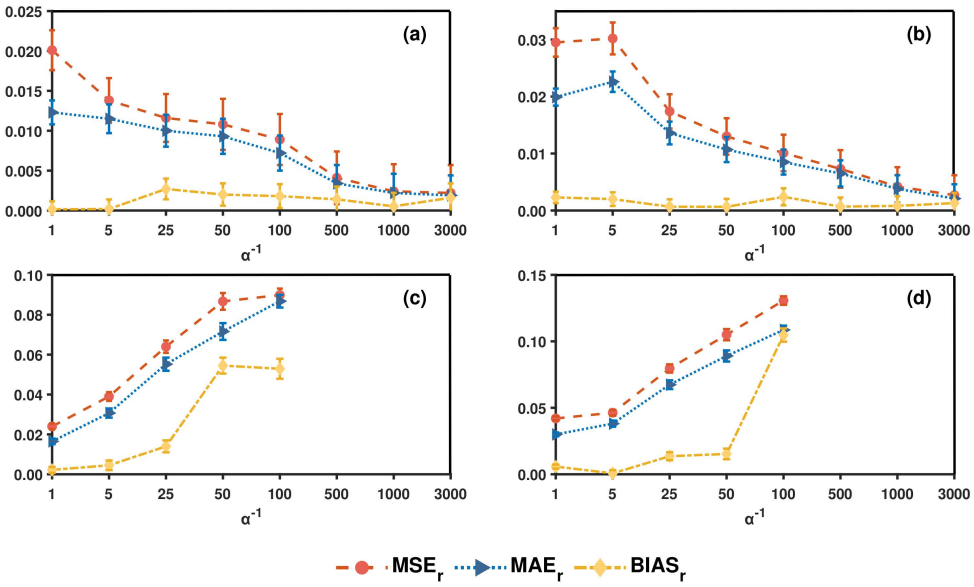


Figure 6. Variation of MSE_r, MAE_r, and BIAS_r with respect to the condition number of the background error covariance matrix. Error bars denote the error variance at each point. (a) and (b) show the errors of the corrected FTH- and WS-type initial guesses obtained using the AIKFI algorithm, respectively. (c) and (d) present the corresponding errors obtained using the l_1-norm R4D-Var algorithm, whcih uses cross-validation to select regularization parameters[5]. The initial guess is $\phi_g(x) = \phi(x) + \varepsilon$ same to (23c), the correlation length in the structure function of the background error covariance matrix varies within $\alpha^{-1} \in \{1, 5, 25, 50, 100, 500, 1000, 3000\}$, corresponding to condition numbers ranging from 10^2 to 10^{14}.

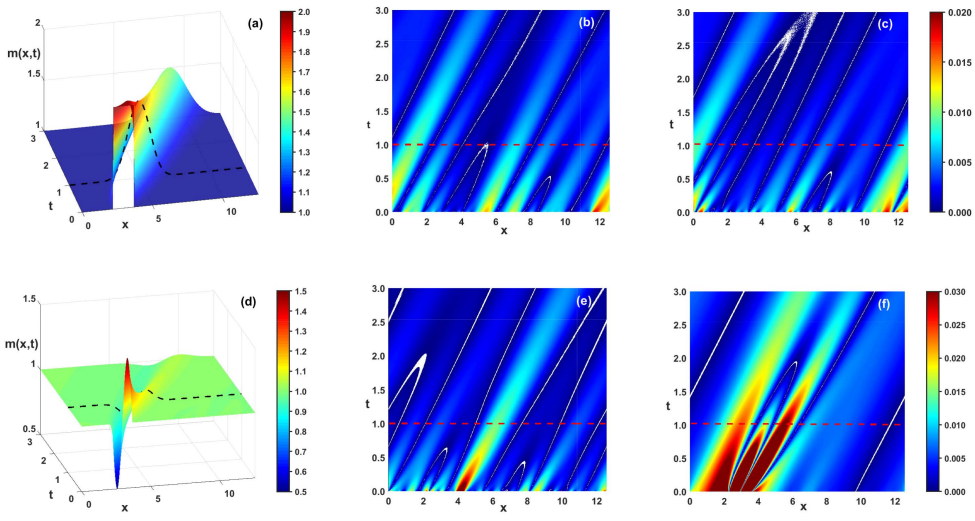


Figure 7. Error in the spatial evolution between the assimilation results and the true states. (a) and (d) show the evolution of the true states for FTH-type and WS-type, respectively. (b) and (e) present the absolute errors of the spatial evolution of the analyzed states corresponding to Figures 4(c) and 4(f) relative to the true evolution. (c) depicts the absolute error of the evolution result for the corrected WS-type initial guess in Figure 4(a) using a fixed regularization parameter in AIKFI, while (f) shows the corresponding error for the corrected WS-type initial guess in Figure 4(d) obtained with the unconstrained f-slope method.

 doi:10.20944/preprints202603.1333.v1

4. Discussion

In this work, we address the degradation of inversion quality caused by large condition numbers of the background error covariance matrix when applying generalized Lasso in 4D-Var. To this end, we develop an Adaptive iterative Kalman filter inversion (AIKFI) algorithm based on the majorization-minimization (MM) strategy. The method formulates the generalized Lasso objective using a gradient operator, transforms it into a weighted least-squares problem via the MM approach, and employs oblique projection to mitigate computational issues arising from the rank deficiency of the gradient matrix. Finally, an improved f-slope method is adopted to adaptively select the regularization parameter, enabling robust inversion within the Kalman filtering framework.

The feasibility and adaptability of the AIKFI algorithm are evaluated through numerical experiments on the initial-condition inversion of a one-dimensional advection–diffusion equation. For the first set of initial guesses ϕ_g^1 and ϕ_g^2, AIKFI is compared with the l_1-norm R4D-Var method, and both successfully recover the key features of the true state. For the second set ϕ_g^3 and ϕ_g^4, AIKFI still reconstructs the essential features despite limited initial information, demonstrating the role of adaptive regularization. The third set ϕ_g^5 and ϕ_g^6 further confirms the algorithm's adaptability, as AIKFI effectively recovers key features for background error covariance matrices with condition numbers ranging from 10^2 to 10^{14}. In particular, when the condition number lies between 10^8 and 10^{14}, the reconstruction errors remain small, with MSE_r and MAE_r below 0.01 and BIAS_r below 0.005. Additional experiments using a fixed regularization parameter and an unconstrained f-slope method further demonstrate the effectiveness of the proposed adaptive regularization scheme. It is worth noting that the lower bound of the search interval for the regularization parameter is critical to the inversion performance and should be determined based on the initial conditions.

The validity of the current algorithm was demonstrated through the retrieval of the initial field of the linear convection-diffusion equation. However, when a nonlinear forward prediction model is involved, the Ensemble Kalman filter can be considered to replace the linear Kalman filter. In this case, whether the method of adaptive selection of the regularization parameters is feasible still deserves further research, which will provide a potential application value for the sparsity of weight parameters in the neural network training process.

Author Contributions: X. D. and Y. W. conceived and designed the research; X. D. completed the code writing, implemented the numerical simulations, performed the data analysis, and conducted the experiments; X. D. and Y. W. designed the Adaptive iterative Kalman filter inversion framework and analysed the results; X. D. drafted the manuscript; X. D. and Y. W. revised the manuscript; Y. W. supervised the project. All authors participated in the discussion of the results.

Funding: This work reported here is supported by the National Natural Science Foundation of China (Grant No. 12241102 and 42275160).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Safi, S.K.; Alsheryani, M.; Alrashdi, M.; Suleiman, R.; Awwad, D.; Abdalla, Z.N. Optimizing linear regression models with lasso and ridge regression: A study on UAE financial behavior during COVID-19. *Migration Letters* **2023**, *20*, 139–153.
2. Wu, H.; Liang, L.; Mei, X.; Zhang, Y. A convex optimization approach for NLOS error mitigation in TOA-based localization. *IEEE Signal Processing Letters* **2022**, *29*, 677–681.
3. Li, S.; Peng, G.; Ji, M.; Cheng, F.; Chen, Z.; Li, Z. Impact identification of composite cylinder based on improved deep metric learning model and weighted fusion Tikhonov regularized total least squares. *Composite Structures* **2022**, *283*, 115144.

 doi:10.20944/preprints202603.1333.v1

4. Budd, C.; Freitag, M.A.; Nichols, N. Regularization techniques for ill-posed inverse problems in data assimilation. *Computers & fluids* **2011**, *46*, 168–173.
5. Ebtehaj, A.M.; Zupanski, M.; Lerman, G.; Foufoula-Georgiou, E. Variational data assimilation via sparse regularisation. *Tellus A: Dynamic Meteorology and Oceanography* **2014**, *66*, 21789.
6. Gen, W.; Shaoxue, S.; Huilan, L.; Rong, W.; Yin, Y. Discontinuous data 3D/4D variation fusion based on the constraint of L1 norm regularization term. *Advances in Earth Science* **2017**, *32*, 757–768.
7. Cao, L.; Qiao, D.; Chen, X. Laplace l_1 Huber based cubature Kalman filter for attitude estimation of small satellite. *Acta Astronautica* **2018**, *148*, 48–56.
8. Freitag, M.; Nichols, N.K.; Budd, C. Resolution of sharp fronts in the presence of model error in variational data assimilation. *Quarterly Journal of the Royal Meteorological Society* **2013**, *139*, 742–757.
9. Ebtehaj, A.M.; Foufoula-Georgiou, E. On variational downscaling, fusion, and assimilation of hydrometeorological states: A unified framework via regularization. *Water Resources Research* **2013**, *49*, 5944–5963.
10. Hansen, P.C. Oblique projections and standard-form transformations for discrete inverse problems. *Numerical linear algebra with applications* **2013**, *20*, 250–258.
11. Zhou, R.; Han, J.; Li, T.; Guo, Z. Fast independent component analysis denoising for magnetotelluric data based on a correlation coefficient and fast iterative shrinkage threshold algorithm. *IEEE transactions on geoscience and remote sensing* **2022**, *60*, 1–15.
12. Nie, F.; Chen, Q.; Yu, W.; Li, X. Row-sparse principal component analysis via coordinate descent method. *IEEE Transactions on Knowledge and Data Engineering* **2024**, *36*, 3460–3471.
13. Pastukhov, V. Fused lasso nearly-isotonic signal approximation in general dimensions. *Statistics and Computing* **2024**, *34*, 120.
14. Sun, Y.; Babu, P.; Palomar, D.P. Majorization-minimization algorithms in signal processing, communications, and machine learning. *IEEE Transactions on Signal Processing* **2016**, *65*, 794–816.
15. Tang, P.; Wang, C.; Jiang, B. A proximal-proximal majorization-minimization algorithm for nonconvex rank regression problems. *IEEE Transactions on Signal Processing* **2023**, *71*, 3502–3517.
16. Takano, T.; Nishida, K. Estimation of seismometer clock time offsets using Kalman Filter towards accurate seismic velocity change. *Geophysical Journal International* **2024**, *239*, 1087–1102.
17. Ge, Q.; Ma, Z.; Lu, Z.; Feng, X. Fault-Tolerant Cubature Kalman Filter for Engineering Estimation Control Systems. *IEEE Transactions on Cybernetics* **2024**, *54*, 3943–3953.
18. Yin, S.; Li, P.; Gu, X.; Yang, X.; Yu, L. Adaptive Kalman filter with LSTM network assistance for abnormal measurements. *Measurement Science and Technology* **2024**, *35*, 075113.
19. Zhao, Y.; Luong, F.; Teshuva, S.; Pelentritou, A.; Woods, W.; Liley, D.; Schmidt, D.F.; Boley, M.; Kuhlmann, L. Improved neurophysiological process imaging through optimization of Kalman filter initial conditions. *International journal of neural systems* **2023**, *33*, 2350024.
20. Kovachki, N.B.; Stuart, A.M. Ensemble Kalman inversion: a derivative-free technique for machine learning tasks. *Inverse Problems* **2019**, *35*, 095005.
21. Chada, N.K.; Stuart, A.M.; Tong, X.T. Tikhonov regularization within ensemble Kalman inversion. *SIAM Journal on Numerical Analysis* **2020**, *58*, 1263–1294.
22. Iglesias, M.; Yang, Y. Adaptive regularisation for ensemble Kalman inversion. *Inverse Problems* **2021**, *37*, 025008.
23. Weissmann, S.; Chada, N.K.; Schillings, C.; Tong, X.T. Adaptive Tikhonov strategies for stochastic ensemble Kalman inversion. *Inverse Problems* **2022**, *38*, 045009.
24. Zhang, X.L.; Xiao, H.; Luo, X.; He, G. Ensemble Kalman method for learning turbulence models from indirect observation data. *Journal of Fluid Mechanics* **2022**, *949*, A26.
25. Pensoneault, A.; Zhu, X. Uncertainty quantification for DeepONets with ensemble Kalman inversion. *Journal of computational physics* **2025**, *523*, 113670.
26. Liu, S.; Reich, S.; Tong, X.T. Dropout ensemble Kalman inversion for high dimensional inverse problems. *SIAM Journal on Numerical Analysis* **2025**, *63*, 685–715.
27. Iglesias, M.A.; Law, K.J.; Stuart, A.M. Ensemble Kalman methods for inverse problems. *Inverse Problems* **2013**, *29*, 045001.
28. Schneider, T.; Stuart, A.M.; Wu, J.L. Ensemble Kalman inversion for sparse learning of dynamical systems from time-averaged data. *Journal of Computational Physics* **2022**, *470*, 111559.
29. Lee, Y. l_p Regularization for Ensemble Kalman Inversion. *SIAM Journal on Scientific Computing* **2021**, *43*, A3417–A3437.

 doi:10.20944/preprints202603.1333.v1

30. Chartrand, R.; Yin, W. Iteratively reweighted algorithms for compressive sensing. In Proceedings of the 2008 IEEE international conference on acoustics, speech and signal processing. IEEE, 2008.
31. Hansen, P.C.; Jensen, T.K. Smoothing-norm preconditioning for regularizing minimum-residual methods. *SIAM journal on matrix analysis and applications* **2007**, *29*, 1–14.
32. Hansen, P.C. Oblique projections, pseudoinverses, and standard-form transformations. *Technical University of Denmark, Tech. Rep* **2004**.
33. Wu, L. A parameter choice method for Tikhonov regularization. *Electronic Transactions on Numerical Analysis* **2003**, *16*, 107–128.
34. Milne, R. An oblique matrix pseudoinverse. *SIAM Journal on Applied Mathematics* **1968**, *16*, 931–944.
35. Hansen, P.C. The L-curve and its use in the numerical treatment of inverse problems **1999**.
36. Hansen, P.C. Insight into semi-convergence of iterative regularization methods. *Linear Algebra and its Applications* **2025**.