A NOVEL METHOD AVOIDING INVERSE CRIME IN SOLVING INVERSE PROBLEMS OF PARABOLIC TYPE USING MODEL REDUCTION METHODS*

WENLONG ZHANG† AND ZHIWEN ZHANG‡

Abstract. In this paper, we propose novel proper orthogonal decomposition (POD)–based methods that effectively address the issue of inverse crime when solving parabolic inverse problems using model reduction methods. We study inverse source and initial value problems using our new methods. By exploiting the low-dimensional structures inherent in the solution space of parabolic equations and constructing POD basis functions, our methods significantly reduce computational costs while maintaining accuracy in solving parabolic inverse problems. In addition, we provide the convergence analysis of the proposed methods for solving these two types of parabolic inverse problems. Finally, we conduct numerical experiments to demonstrate the accuracy and efficiency of the proposed method. The results show that our method efficiently solves parabolic inverse problems and overcomes the inverse crime issues associated with traditional model reduction methods for such problems.

Key words. Parabolic inverse problem; Regularization method; Model reduction method; Inverse crime; Convergence analysis.

MSC codes. 35R30, 65J20, 65M12, 65N21, 78M34.

1. Introduction. Inverse problems associated with parabolic equations have garnered significant attention in mathematics and engineering research fields [25, 19]. These problems can be broadly categorized into several types: recovering source terms in the PDEs, determining the system's initial state, identifying physical parameters, and determining the boundary conditions. In this paper, we focus on the first two types of inverse problems and leave the other types as future work. We will refer to the first two types of inverse problems as inverse source problems and backward problems, respectively.

Inverse source problems, which involve reconstructing the source from final time observation, have attracted much attention from researchers in recent decades [23] and references therein. These problems have been extensively studied in the literature and applied to various physical and engineering source identification problems, such as groundwater migration, groundwater pollution detection, pollution source control, and environmental protection [8, 12, 16, 11, 23] and references therein. Accurately recovering pollutant sources is crucial for ensuring environmental safeguards in densely populated cities [9]. The estimation of the strength of acoustic sources from measurements can be found in e.g. [23, 13, 31]. Given the importance of inverse source problems in practical applications, numerous numerical methods have been extensively explored [8, 6, 29, 15, 16] and references therein.

For inverse source problems, iteration optimization methods are typically used to determine the true source term [6, 10, 24]. The forward parabolic equation must be solved one or two times in each iteration. However, as the size of discrete problems

*Submitted to the editors DATE.

†Corresponding author. Department of Mathematics & National Center for Applied Mathematics Shenzhen, Southern University of Science and Technology (SUSTech), 1088 Xueyuan Boulevard, University Town of Shenzhen, Xili, Nanshan, Shenzhen, Guangdong Province, P.R.China. (zhangwl@sustech.edu.cn).

‡Corresponding author. Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, P.R.China. Materials Innovation Institute for Life Sciences and Energy (MILES), HKU-SIRI, Shenzhen, P.R. China. (zhangzw@hku.hk).

This manuscript is for review purposes only.

increases (e.g. using the finite element method (FEM) or finite difference method (FDM)), the computational time increases rapidly, especially for time-evolution problems. As a result, the computation of the forward equation will consume the most time throughout the entire process.

To address the computational challenges associated with solving parabolic inverse source problems, we propose a data-driven model reduction method [35]. Our method consists of offline and online stages. In the offline stage, we exploit the low-dimensional structures in the solution space of parabolic equations in the forward problem, given a class of source functions, and construct a small number of POD basis functions to achieve significant dimension reduction. With the POD basis functions, we can rapidly solve the forward problem in the online stage. Consequently, we developed a fast algorithm called the POD algorithm to solve the optimization problem in the parabolic inverse source problems. Moreover, we derive an error estimate for the POD algorithm in parabolic inverse source problems. Numerical results demonstrate that the POD algorithm offers significant computational savings compared to the FEM while maintaining the same level of accuracy. However, we have to point out that our POD algorithm has a limitation: it requires assuming that the true source term belongs to a known function class, leading to the inverse crime [25, 22]. In this paper, we aim to develop novel model reduction methods to solve parabolic inverse problems and eliminate the inverse crime issue.

Backward problems are another important type of parabolic inverse problem that has been extensively studied in physics and engineering, particularly in the field of heat transfer. The main focus of these problems is to determine the initial condition from transient temperature measurements at the final time T. The main difficulty in solving the backward problems arises from the exponential decay of forward solutions of the parabolic equations with respect to the initial data. Therefore, backward problems are also ill-posed in the sense of Hadamard [23, 18], as the eigenvalues of elliptic operators decay exponentially fast, making them particularly unstable with respect to measurement data uncertainties. This lack of stability poses a significant challenge for numerical inversions, as even small changes in the data can lead to substantial differences in the reconstructed source strength.

In response to these challenges, many regularization techniques have been developed for solving backward problems. For instance, in [26], Sobolev error estimates and a prior parameter selection for semi-discrete Tikhonov regularization were derived. A backward problem for the one-dimensional heat conduction equation, with the measurements on a discrete set, was considered in [7], and the uniqueness of recovering the initial value was proved using the analytic continuation method. It is worth noting that in [30], a comparison of various inverse methods for estimating the initial condition of the heat equation was studied, demonstrating that explicit approaches to the backward problem yield disastrous results unless some form of regularization is utilized.

In our recent work [36], we study the stochastic convergence of regularized solutions for backward problems. We derive an error estimate for the least-squares regularized minimization problem within the framework of stochastic convergence. Our analysis reveals that the optimal error of the Tikhonov-type least-squares optimization problem depends on the noise level, the number of sensors, and the underlying ground truth. Additionally, we propose a self-adaptive algorithm to identify the optimal regularization parameter for the optimization problem without requiring knowledge of the noise level or any other prior information, which would be highly practical in applications. Numerical results demonstrate the effectiveness of our method in solving backward problems. By assuming the initial condition belongs to a known function class, it is straightforward to develop a POD method to solve backward problems using the stochastic convergence developed in [36]. However, the corresponding POD method also suffers from the issue of inverse crime [25, 22]. Once again, this motivates us to develop new model reduction methods to solve parabolic backward problems and eliminate the inverse crime issue.

As we have discussed, the issue of inverse crime arises when we develop the POD method for solving inverse problems [35], where the forward model used for generating the data is identical to the one employed for solving the inverse problem. This scenario can lead to overly optimistic results and underestimates the uncertainties associated with the solution. The inverse crime poses a significant challenge in inverse problem solving, as it fails to account for model errors and uncertainties that are inherent in real-world applications.

Motivated by an interesting observation in our previous work on numerical simulation of parabolic equations, we propose a novel POD-based model reduction method to address the issue of inverse crime encountered in solving inverse problems associated with parabolic equations. Specifically, we aim to develop a novel POD methodology that can be applied to both the inverse source problems and backward problems of the parabolic type. To start with, we consider a generic parabolic equation as follows:

$$
\begin{cases} u_t + \mathcal{L}u = f(x) & \text{in } \Omega \times (0,T), \\ u(x,t) = 0 & \text{on } \partial\Omega \times (0,T), \\ u(x,0) = g(x) & \text{in } \Omega\,, \end{cases} \tag{1.1}
$$

where $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ is a bounded domain with a C^2 boundary or a convex domain satisfying the uniform cone condition, $\mathcal{L}$ denotes a second-order elliptic operator given by $\mathcal{L}u = -\nabla \cdot (q(x)\nabla u) + c(x)u$, $f(x)$ is the source term, and $g(x)$ is the initial condition. We assume the elliptic operator $\mathcal{L}$ is uniform elliptic, i.e., there exist $q_{\min}, a_{\max} > 0$ such that $q_{\min} < q(x) < q_{\max}$ for all $x \in \Omega$. Additionally, we assume $q(x) \in C^1(\bar{\Omega})$, $c(x) \in C(\bar{\Omega})$ and $c(x) \geq 0$.

Let u represent the solution of the parabolic equation (1.1). We define the forward operator $\mathcal{S} : \mathcal{S}(f, g) = u(\cdot, T)$. The forward problem involves computing the solution $u(\cdot, t)$ for $t > 0$ given the source term $f(x)$ and initial condition $g(x)$. The inverse problem, on the other hand, aims to reconstruct $f(x)$ or $g(x)$ from the final time measurement $m = u(\cdot, T)$. We will solve two types of inverse problems as follows:

1. Inverse source problem: recover the source term $f(x)$ using the final time measurement $m = u(\cdot, T)$ and the known initial term $g(x)$.
2. Backward problem: recover the initial term $g(x)$ using the final time measurement $m = u(\cdot, T)$ and the known source term $f(x)$.

In this paper, we develop a novel POD-based model reduction method, called the adjoint-POD method, for solving inverse problems of parabolic types. We begin by developing the adjoint-POD method to solve the inverse source problem and construct basis functions for this problem. Specifically, we study the convergence of the POD basis functions obtained by our adjoint-POD method and prove their approximation property in Theorem 2.3. By leveraging this property, we derive the convergence analysis of the adjoint-POD method in solving inverse parabolic source problems in Theorem 2.4 and Theorem 2.5. We then extend our approach to the parabolic backward problem, where we construct the POD basis functions by solving the adjoint equation with the given final time measurement as the initial condition. We also prove the approximation property of the corresponding POD basis functions in Theorem 3.2 and the convergence analysis of the adjoint-POD method in solving the parabolic backward problem in Theorem 3.3 and Theorem 3.4. Finally, we present numerical experiments that demonstrate the effectiveness of our adjoint-POD method in solving parabolic inverse problems and its ability to overcome inverse crime.

The rest of the paper is organized as follows. In Section 2, we introduce the adjoint-POD method for solving parabolic inverse source problems and provide the error estimate for the proposed methods. Similarly, in Section 3, we propose the adjoint-POD method for solving parabolic backward problems and provide the corresponding error estimate. In Section 4, we present numerical results to demonstrate the accuracy of our methods. Finally, concluding remarks are made in Section 5.

2. Adjoint-POD method for parabolic inverse source problems.

Regularization methods are commonly used to solve inverse problems, where one applies iterative methods to solve the forward problem one or more times in each iterative step. As a result, most computational time is spent on solving the forward problem. To address this issue, it is natural to develop model reduction methods to decrease the computational cost of solving the forward problems.

One of the model reduction ideas in solving time-evolution problems is the POD method [33, 4]. The POD method uses data from an experiment or an accurate numerical simulation and extracts the most energetic modes in the system by using the singular value decomposition. This approach generates low-dimensional structures that can approximate the solutions to the time-evolution problem with high accuracy. The POD method has been applied successfully to solve many types of PDEs, including linear parabolic equations [34, 27], Navier-Stokes equations [27], viscous G-equations [17], Hamilton–Jacobi–Bellman (HJB) equations [28], and optimal control problems [2]. The interested reader is referred to [32, 3, 20] for a comprehensive introduction to the model reduction methods.

The traditional POD method has a significant drawback: to construct the POD basis functions, one must know the source term $f(x)$ or the initial condition $g(x)$. Therefore, directly using the POD method to solve inverse problems may result in the inverse crime issue. In our previous work [35], we mitigated this issue by assuming that the true source term belongs to a known function class. However, this approach does not fully address the issue of the inverse crime.

To tackle this challenge, we propose a novel method for model reduction in inverse problems: the adjoint-POD method. Unlike the traditional POD method, our method does not require any prior knowledge of the source or initial term. By integrating the adjoint method with the model reduction capabilities of the POD method, we will show that the adjoint-POD method can efficiently solve inverse problems while avoiding the inverse crime issue.

2.1. Adjoint POD method.

To demonstrate the idea of the adjoint-POD method, we will first apply it to solve the inverse source problem. This problem involves recovering the unknown source term $f(x)$ of the parabolic equation, given the final time measurement $m(x) = \mathcal{S}(f) = u(\cdot, T)$. Here, u satisfies the following equation:

$$
\begin{cases}
u_t + \mathcal{L}u = f(x) & \text{in } \Omega \times (0,T), \\
u(x,t) = 0 & \text{on } \partial\Omega \times (0,T), \\
u(x,0) = 0 & \text{in } \Omega\,.
\end{cases}
\tag{2.1}
$$

Here we assume that $u(x,t) = 0$ and $u(x,0) = 0$ for simplicity, otherwise, one just needs to subtract the background solution from the measurement $m(x)$. Since the source term $f(x)$ is unknown, we cannot use the traditional POD method to obtain snapshots. Instead, we will obtain the snapshots from the following equation:

$$\begin{cases} \tilde{u}_t + \mathcal{L}\tilde{u} = m(x) & \text{in } \Omega \times (0,T), \\ \tilde{u}(x,t) = 0 & \text{on } \partial\Omega \times (0,T), \\ \tilde{u}(x,0) = 0 & \text{in } \Omega\,. \end{cases} \tag{2.2}$$

We denote the snapshots as $\tilde{y}_k = \tilde{u}(\cdot, t_{k-1})$, $k = 1, \ldots, M+1$ with $M = \frac{T}{\Delta t}$, and $\tilde{y}_k = \overline{\partial}\tilde{u}(\cdot, t_{k-M-1})$, $k = M+2, \ldots, 2m+1$. Here, $\overline{\partial}\tilde{u}(\cdot, t_k) = \frac{\tilde{u}(\cdot,t_k) - \tilde{u}(\cdot,t_{k-1})}{\Delta t}$ for $k = 1, \ldots, M$. We then construct the new POD basis functions $\{\psi_1, ..., \psi_{N_{\text{pod}}}\}$ using the method described in Appendix A from the adjoint equation (2.2). Finally, we denote the linear space spanned by the POD basis functions as $V_{\text{POD}} = \text{span}\{\psi_1, ..., \psi_{N_{\text{pod}}}\}$.

We can use these new POD basis functions $\{\psi_1, ..., \psi_{N_{\text{pod}}}\}$ to approximate the forward problem and speed up the computation. We construct the fully discrete scheme on V_{pod}, and we denote the solution by U_k for $k = 1 \cdots M$. Specifically, we seek numerical solutions U_k such that:

$$(\bar{\partial} U_k, \psi) + a(U_k, \psi) = (f, \psi), \quad \forall \psi \in V_{\text{pod}}. \tag{2.3}$$

Here the bilinear form $a(u,v) = (q\nabla u, \nabla v) + (cu, v)$. We define the solution operator from the source term f to the final time solution U_M as $\mathcal{S}_{\text{pod}}$, such that $\mathcal{S}_{\text{pod}} f = U_M$. By using the new POD basis functions and the reduced-order model represented by $\mathcal{S}_{\text{pod}}$, we can efficiently solve the forward problem for each time step. This significantly reduces the computational cost compared to using the full-scale model. This approach is particularly useful when solving inverse problems, where multiple forward problem evaluations are required.

2.2. Convergence of the adjoint-POD method. We will first revisit an important property of the eigenvalue distribution for the classical elliptic operator $\mathcal{L}$ [1, 14].

PROPOSITION 2.1. *Suppose Ω is a bounded domain in $\mathbb{R}^d$ and $a(x), c(x) \in C^0(\bar{\Omega})$, $c(x) \geq 0$. Then, the eigenvalue problem*

$$\mathcal{L}\psi = \mu\,\psi \quad \textit{with} \quad \psi_{\partial\Omega} = 0 \tag{2.4}$$

has a countable set of positive eigenvalues $\mu_1 \leq \mu_2 \leq \cdots$, with corresponding eigenfunctions $\{\phi_k\}_{k=1}^{\infty}$ forming an orthogonal basis of $L^2(\Omega)$. Moreover, there exist positive constants $C_1, C_2 > 0$ such that $C_1 k^{2/d} \leq \mu_k \leq C_2 k^{2/d}$ for all $k = 1, 2, \cdots$.

According to the Proposition above, the eigenfunction set $\{\phi_k\}_{k=1}^{\infty}$ forms an orthogonal basis of $L^2(\Omega)$. Thus, for any $f \in L^2(\Omega)$, we can write $f = \sum_{k=1}^{\infty} f_k \phi_k$, where f_k's are coefficients. Similarly, let $u = \sum_{k=1}^{\infty} u_k(t)\phi_k$ be the solution of the problem (2.1). Substituting these expressions of f and u into the first equation of (2.1) and noting that $\mathcal{L}\phi_k = \mu_k \phi_k$, we can compare the coefficients of ϕ_k on both sides of the equation to obtain $u_k(0) = 0$ and

$$u_k'(t) + \mu_k u_k = f_k \qquad \text{in } (0,T)\,. \tag{2.5}$$

This equation expresses the time evolution of the coefficients $u_k(t)$ in terms of the coefficients f_k of the source term f. We can write the solution as $u_k(T) = \alpha_k f_k$, where $\alpha_k = e^{-\mu_k T}\int_0^T e^{\mu_k s} ds = \frac{1}{\mu_k}(1-e^{-\mu_k T})$. Noting that $Sf = u(\cdot,T) = \sum_{k=1}^{\infty} u_k(T)\phi_k$, we can formally write

$$S\Big(\sum_{k=1}^{\infty} f_k\phi_k\Big) = \sum_{k=1}^{\infty} \alpha_k f_k \phi_k.$$

This representation of the solution operator S provides a convenient way to compute the solution $u(\cdot,T)$ using the eigenfunctions ϕ_k and the coefficients α_k. To simplify the problem, we approximate the source term $f(x)$ by truncating it to a finite-dimensional space, i.e.

$$f_{\text{app}} = \sum_{k=1}^{L} f_k\phi_k. \tag{2.6}$$

Using this truncation, the solution $u(x,t)$ of the parabolic equation can be written as:

$$u(\cdot,T) = \sum_{k=1}^{L} \frac{1}{\mu_k}(1-e^{-\mu_k T}) f_k\phi_k. \tag{2.7}$$

After simple calculation, we will also derive that

$$\tilde{u}(x,t) = \sum_{k=1}^{L} \frac{1}{\mu_k}(1-e^{-\mu_k T})(1-e^{-\mu_k t}) f_k\phi_k. \tag{2.8}$$

It is worth noting that the POD basis (A.4) is simply the singular value decomposition of the matrix $\tilde{A} = (\tilde{y}_1, ..., \tilde{y}_M)$, where $\tilde{y}_j = (\tilde{u}(x_1,t_j), ..., \tilde{u}(x_N,t_j))^T$ and $x_1, ..., x_N$ are the finite element nodes in Ω. Specifically, if A has the singular value decomposition $A = U\Sigma V$, then the first M columns of U correspond exactly to the POD basis $\{\psi_k\}_{k=1}^{M}$.

Let us denote $A = (y_1, ..., y_M)$, $\tilde{A} = (\tilde{y}_1, ..., \tilde{y}_M)$, the matrix $\Phi = (\phi_1(\vec{x},t_1), ..., \phi_L(\vec{x},t_1))$, $F = \mathrm{diag}(f_1, ..., f_L)$, and $D = \mathrm{diag}(\frac{1}{\mu_k}(1-e^{-\mu_k T}), ..., \frac{1}{\mu_L}(1-e^{-\mu_L T}))$. In addition, let J be an $L \times M$ matrix with entries $J(i,j) = \frac{1}{\mu_i}(1-e^{-\mu_i t_j})$. Φ is a column orthogonal matrix due to the normal orthogonality of the eigenfunctions ϕ_k. Using the formulations of u and $\tilde{u}$, we can represent the matrices A and $\tilde{A}$ as follows:

$$A = \Phi F J, \quad \text{and} \quad \tilde{A} = \Phi D F J. \tag{2.9}$$

Proposition A.1 shows that the low-rank space V_{pod} provides the best N_{pod}-rank approximation of the column space of $\tilde{A}$. Our objective is to demonstrate that V_{pod} is also a good approximation of the column space of A, which will validate the effectiveness of the new POD method. To begin, we need to establish the relationship between the matrices A and $\tilde{A}$.

LEMMA 2.2. *If $L \le M$, then span$\{y_1, ..., y_M\}$ = span$\{\tilde{y}_1, ..., \tilde{y}_M\}$, i.e. $C(A) = C(\tilde{A})$.*

Proof. We will provide a concise proof for the case when the eigenvalues μ_j are distinct from each other. The proof for the case of repeated eigenvalues follows a similar approach. To demonstrate the desired results, we need to show that there exist matrices P and $\tilde{P}$ such that

$$\Phi DFJP = \Phi FJ, \quad \text{and} \quad \Phi DFJ = \Phi FJ\tilde{P}. \tag{2.10}$$

We will only present a brief proof for the first equality, as the second can be derived similarly. Since the columns of Φ are independent and the diagonal matrix F is invertible, we can show the existence of a matrix P by proving that the following equation holds

$$JP = DJ. \tag{2.11}$$

First, we will prove that the matrix $J'_{L\times L}$ with entries $J'(i,j) = 1 - e^{-\mu_i t_j}$ is invertible. To do this, we introduce the vector $\mathbf{e} = (1, ..., 1)^T$. It follows that we can write J' as the difference between the outer product of $\mathbf{e}$ and its transpose, $\mathbf{e}\mathbf{e}^T$, and the Vandermonde matrix V_L, where $V_L(i,j) = e^{-\mu_i t_j}$. Namely, we have $J' = \mathbf{e}\mathbf{e}^T - V_L$.

Next, we aim to show the invertibility of J' by contradiction. Suppose that J' is singular, which implies that there exists a nonzero vector $\mathbf{c} = (c_1, ..., c_L)^T$ satisfying $J'\mathbf{c} = 0$. Equivalently, we can express this as $V_L\mathbf{c} = \mathbf{e}\mathbf{e}^T\mathbf{c}$.

Let us consider the function $f(x) = \sum_{j=1}^{L} c_j e^{xt_j}$. With this assumption, we have that

$$f(0) = f(\mu_1) = f(\mu_2) = \cdots = f(\mu_L), \tag{2.12}$$

which implies that the function f has $L+1$ distinct zeros. Therefore, its derivative $f'(x) = \sum_{j=1}^{L} c_j t_j e^{xt_j}$ must has L distinct zeros. Since $\mathbf{c}$ is nonzero, and all μ_js are also nonzero, this implies that the Vandermonde matrix V_L is singular. However, this contradicts the fact that V_L is an invertible matrix. Consequently, J' must be a nonsingular matrix.

Since the invertibility of J', the first L columns of J are independent and thus form a basis for $\mathbb{R}^L$. Similarly, the matrix DJ also has independent columns that form a basis for $\mathbb{R}^L$. Therefore, there must exist a matrix P such that $JP = DJ$. This result establishes that the spaces spanned by the sets $\{y_1, ..., y_M\}$ and $\{\tilde{y}_1, ..., \tilde{y}_M\}$ are equivalent, i.e., $\text{span}\{y_1, \ldots, y_M\} = \text{span}\{\tilde{y}_1, \ldots, \tilde{y}_M\}$. This completes the proof. ∎

Proposition A.1 suggests that the new POD basis provides an effective approximation of the set $\{\tilde{y}_1, ..., \tilde{y}_M\}$. Given the previous results, we can now show that the new POD basis also serves as a good approximation for the original set $\{y_1, ..., y_M\}$.

THEOREM 2.3. *Using the same notation as in Proposition A.1, we can derive an approximation error bound if a sufficient number of snapshots are available, i.e. $L \le M$. In this case, the following error bound holds:*

$$\frac{\sum_{i=1}^{M} ||y_i - P_{pod} y_i||^2_{L^2(\Omega)}}{\sum_{i=1}^{M} ||y_i||^2_{L^2(\Omega)}} \le C L^{4/d} \rho, \tag{2.13}$$

where P_{pod} is the projection operator onto the adjoint-POD space $\text{span}\{\psi_1, \ldots, \psi_{N_{pod}}\}$ *and $\rho = \frac{\sum_{k=N_{pod}+1}^{2M+1} \lambda_k}{\sum_{k=1}^{2M+1} \lambda_k}$ is a parameter that depends on the decay speed of the eigenvalues of the correlation matrix.*

Proof. In the following proof, we assume $L = M$ for simplicity. For the case $L < M$, the proof is similar. Using the same notation of Lemma 2.2, Φ and J are both invertible square matrices. Then there exists a unique matrix P such that,

$$\Phi DFJP = \Phi FJ, \tag{2.14}$$

where $P = J^{-1}D^{-1}J$.

Hence, we have $y_j = \sum_{i=1}^{L} P_{ij}\tilde{y}_i$. Using the Cauchy-Schwarz inequality, we can show that for any $1 \leq j \leq L$,

$$\|y_j - P_{\text{pod}}y_j\|^2 \leq \sum_{i=1}^{L} P_{ij}^2 \sum_{i=1}^{L} \|\tilde{y}_i - P_{\text{pod}}\tilde{y}_i\|^2. \tag{2.15}$$

Moreover, we can get that

$$\sum_{j=1}^{L} \|y_j - P_{\text{pod}}y_j\|^2 \leq \sum_{i,j=1}^{L} P_{ij}^2 \sum_{i=1}^{L} \|\tilde{y}_i - P_{\text{pod}}\tilde{y}_i\|^2 = \|P\|_F^2 \sum_{i=1}^{L} \|\tilde{y}_i - P_{\text{pod}}\tilde{y}_i\|^2. \tag{2.16}$$

The rest is to estimate the Frobenius norm of P. Since $P = J^{-1}D^{-1}J$, we can define $\|P\|_d = \|D^{-1}\|_2$. It is easy to verify that $\|\cdot\|_d$ is a matrix norm. Then, we have

$$\|P\|_F \leq C\|P\|_d = C\|D^{-1}\|_2 \leq C\mu_L. \tag{2.17}$$

On the other hand, since Φ is an orthogonal matrix, we have,

$$\sum_{j=1}^{L} \|\tilde{y}_j\|^2 = \|\Phi DFJ\|_F^2 \tag{2.18}$$

$$= \|DFJ\|_F^2 \leq \|D\|_F^2 \|FJ\|_F^2 \tag{2.19}$$

$$\leq C\|FJ\|_F^2 \leq C\sum_{j=1}^{L} \|y_j\|^2. \tag{2.20}$$

Aligning with Proposition A.1, we can combine the previous inequalities to obtain

$$\frac{\sum_{i=1}^{M} ||y_i - P_{\text{pod}}y_i||_{L^2(\Omega)}^2}{\sum_{i=1}^{M} ||y_i||_{L^2(\Omega)}^2} \leq C\|P\|_F^2 \frac{\sum_{i=1}^{M} ||\tilde{y}_i - P_{\text{pod}}\tilde{y}_i||_{L^2(\Omega)}^2}{\sum_{i=1}^{M} ||\tilde{y}_i||_{L^2(\Omega)}^2} \tag{2.21}$$

$$\leq \mu_L^2 \rho. \tag{2.22}$$

The conclusion follows from the estimation $\mu_i \leq Ci^{2/d}$. ∎

2.3. Convergence of inverse parabolic source problem. To solve this inverse source problem, we use the well-established Tikhonov regularization method, which is expressed as

$$\min_{f\in X} \|\mathcal{S}(f) - m\|_{L^2(\Omega)}^2 + \lambda\|f\|_{L^2(\Omega)}^2. \tag{2.23}$$

However, in the conventional application of the POD method, the source term f must be determined initially to generate snapshots and obtain the POD basis functions. In the context of inverse problems, the only available information is the measurement $m(x)$. This predicament, referred to as the inverse crime, makes the conventional POD method impossible to implement in practice. Our new method overcomes this vital drawback by using the POD forward solver as the forward solver in the Tikhonov regularization method.

In the general discrete approximation of problem (2.23), we seek to solve the following least-squares regularized optimization problem:

$$\min_{f\in V_{\text{pod}}} \|\mathcal{S}_{\text{pod}}(f) - m\|^2_{L^2(\Omega)} + \lambda\|f\|^2_{L^2(\Omega)}. \tag{2.24}$$

Consider the functional $\mathcal{J}_{\text{pod}}[f] = \|\mathcal{S}_{\text{pod}}f - m\|^2_{L^2(\Omega)} + \lambda\|f\|^2_{L^2(\Omega)}$. By computing the Fréchet derivative of $\mathcal{J}_{\text{pod}}[f]$, we can derive the subsequent iterative scheme:

$$f_{k+1} = f_k - \beta d\mathcal{J}_{\text{pod}}[f_k], \quad \forall k\in\mathbb{N}, \tag{2.25}$$

where β is the step size, $d\mathcal{J}_{\text{pod}}[f] = \mathcal{S}^*_{\text{pod}}(\mathcal{S}_{\text{pod}}f - m) + \lambda f$ denotes the Fréchet derivative, and f_0 is an initial guess [35].

The above theory is based on the noise-free case, where the final time measurement $m = u(\cdot,T)$ is assumed to be precisely known. However, in practical applications, measurement data often contains uncertainties. We assume that the measurement data is blurred by noise and takes the discrete form

$$m_i^n = u(d_i,T) + e_i, \quad i = 1,\cdots,n, \tag{2.26}$$

where d_is represent the positions of detectors, and $\{e_i\}_{i=1}^n$ are independent and identically distributed (i.i.d.) random variables on an appropriate probability space $(\mathfrak{X},\mathcal{F},\mathbb{P})$.

Based on [6] and the analysis therein, we know that $\|u\|_{C([0,T];H^2(\Omega))} \le C\|f\|_{L^2(\Omega)}$. According to the embedding theorem of Sobolev spaces, we know that $H^2(\Omega)$ is continuously embedded into $C(\bar{\Omega})$ so that $u(\cdot,T)$ is well defined point-wisely for all $d_i\in\Omega$. Without loss of generality, we assume that the scattered locations $\{d_i\}_{i=1}^n$ are uniformly distributed in Ω. That is, there exists a constant $B>0$ such that $d_{\max}/d_{\min}\le B$, where $d_{\max}$ and $d_{\min}$ are defined by

$$d_{\max} = \sup_{x\in\Omega}\inf_{1\le i\le n}|x-d_i| \quad\text{and}\quad d_{\min} = \inf_{1\le i\ne j\le n}|d_i-d_j|. \tag{2.27}$$

We will first use the technique developed in [5] to recover the final time measurement $u(\cdot,T)$ from the noisy data m_i^n, for $i=1,...,n$. We approximate $u(\cdot,T)$ by solving the following minimization problem:

$$m = \operatorname*{argmin}_{u\in X}\frac{1}{n}\sum_{i=1}^n (u(x_i)-m_i^n)^2 + \alpha|u|^2_{H^2(\Omega)}. \tag{2.28}$$

The choice of the optimal parameter α typically depends on both the noise level and the unknown function u^*. In the case of measured data with uncertainty, an a posteriori method has been proposed and discussed in previous literature [5, 6, 36]. We list the algorithm below.

Assuming the pointwise noise e_i has a bounded variance σ, which is referred to as the noise level, [5] analyzed this problem and provided optimal convergence results. Moreover, they proposed an a posteriori algorithm to obtain the best approximation without knowing the true solution m and noise level σ. Here, we list their main results. If one chooses the optimal regularization parameter

$$\alpha^{1/2+d/8} = O(\sigma n^{-1/2}\|u(\cdot,T)\|^{-1}_{H^2(\Omega)}), \tag{2.29}$$

This manuscript is for review purposes only.

Algorithm 2.1 A self-consistent algorithm for finding the optimal α

Given an initial guess of α_0; for $j = 0, 1, \cdots$, do the following
Solve (2.28) for u_h with α replaced by the current value of α_j on the mesh;
Update α_{j+1}: $\alpha_{j+1}^{1/2+d/8} = n^{-1/2}\|u_h - m^n\|_n \, |u_h|_{H^2(\Omega)}^{-1}$.
The algorithm stops if $|\alpha_j - \alpha_{j+1}| < 10^{-10}$.

then the solution m of (2.28) achieves the optimal convergence

$$\mathbb{E}\big[\|u(\cdot,T) - m\|_{L^2(\Omega)}\big] \le C\alpha^{1/2}\|u(\cdot,T)\|_{H^2(\Omega)}. \tag{2.30}$$

If the noise $\{e_i\}_{i=1}^n$ are independent Gaussian random variables with variance σ, we further have,

$$\mathbb{P}(\|u(\cdot,T) - m\|_{L^2(\Omega)} \ge \alpha^{1/2}\|u(\cdot,T)\|_{H^2(\Omega)} z) \le 2e^{-Cz^2}. \tag{2.31}$$

Using this recovered function $m(x)$, we generate the adjoint POD basis functions in Section 2.1. It can be easily shown that, with uncertainty, the POD basis functions are still a good low-rank approximation of the snapshots $\{y_1, ..., y_M\}$. Combining Theorem 2.3 and (2.30), we have that for any $1 \le i \le M$,

$$||y_i - P_{\text{pod}} y_i||_{L^2(\Omega)}^2 \le C(ML^{4/d}\rho + \alpha)\|f\|_{L^2(\Omega)}^2. \tag{2.32}$$

Since we replace the source term by a finite truncation (2.6) and if $f \in H^1(\Omega)$, we have

$$\|f - f_{\text{app}}\|_{L^2} \le C\frac{\|\nabla f\|_{L^2}}{\sqrt{\mu_L}} \le C\frac{\|\nabla f\|_{L^2}}{L^{1/d}}. \tag{2.33}$$

If $f \in L^2(\Omega)$, then $f_{\text{app}} \to f$ as $L \to +\infty$. We assume

$$\|f - f_{app}\|_{L^2}^2 \le \varepsilon, \tag{2.34}$$

where ε depends on L. With these results, using a similar technique to prove Theorem 4.1 in [35], we can obtain the following convergence results.

THEOREM 2.4. *Let $\{e_i\}_{i=1}^n$ be independent random variables satisfying $\mathbb{E}[e_i] = 0$ and $\mathbb{E}[e_i^2] \le \sigma^2$ for $i = 1, \cdots, n$. Set $\alpha^{1/2+d/8} = O(\sigma n^{-1/2}\|u^*(\cdot,T)\|_{H^2(\Omega)}^{-1})$ in (2.28), and $\lambda = O(ML^{4/d}\rho + \alpha)$. Then, we have the following error estimates*

$$\mathbb{E}\big[\|\mathcal{S}f^* - \mathcal{S}_{pod}f_{pod}\|_{L^2(\Omega)}^2\big] \le C\lambda\|f^*\|_{L^2(\Omega)}^2 + C\varepsilon, \tag{2.35}$$

$$\mathbb{E}\big[\|f^* - f_{pod}\|_{L^2(\Omega)}^2\big] \le C\|f^*\|_{L^2(\Omega)}^2 + C\varepsilon, \tag{2.36}$$

and

$$\mathbb{E}\big[\|f^* - f_{pod}\|_{H^{-1}(\Omega)}^2\big] \le C\lambda^{1/2}\|f^*\|_{L^2(\Omega)}^2 + C\varepsilon. \tag{2.37}$$

Furthermore, if we assume the noise $\{e_i\}_{i=1}^n$ are independent Gaussian random variables with variance σ, we can obtain a stronger type of convergence. A similar proof can be found in [6]. We list the results here for completeness.

This manuscript is for review purposes only.

THEOREM 2.5. *Let* $\{e_i\}_{i=1}^n$ *be independent Gaussian random variables with variance* σ. *Set* $\alpha^{1/2+d/8} = O(\sigma n^{-1/2}\|u^*(\cdot,T)\|_{H^2(\Omega)}^{-1})$ *in* (2.28)*, and* $\lambda = O(ML^{4/d}\rho+\alpha)$. *Then, there exists a constant* C*, for any* $z > 0$*, such that*

$$\mathbb{P}(\|S_{pod}f_{pod} - Sf^*\|_{L^2(\Omega)} \geq (\lambda^{1/2}\|f^*\|_{L^2} + \varepsilon)z) \leq 2e^{-Cz^2}, \tag{2.38}$$

$$\mathbb{P}(\|f_{pod} - f^*\|_{L^2(\Omega)} \geq (\|f^*\|_{L^2} + \varepsilon)z) \leq 2e^{-Cz^2}, \tag{2.39}$$

and

$$\mathbb{P}(\|f_{pod} - f^*\|_{H^{-1}(\Omega)} \geq (\lambda^{1/4}\|f^*\|_{L^2} + \varepsilon)z) \leq 2e^{-Cz^2}. \tag{2.40}$$

Theorems 2.4 and 2.5 provide the stochastic convergence rate for the error $\|\mathcal{S}f^* - \mathcal{S}_{\text{pod}}f_{\text{pod}}\|_{L^2(\Omega)}^2 = O_p(\lambda\|f^*\|_{L^2(\Omega)}^2 + \varepsilon)$, and $\|f^* - f_{\text{pod}}\|_{H^{-1}(\Omega)}^2 = O_p(\lambda^{1/2}\|f^*\|_{L^2(\Omega)}^2 + \varepsilon)$ in terms of the values λ and ε. These results provide useful guidance for practical numerical simulations.

3. Parabolic backward problem.

For the backward problem of the parabolic equation, our goal is to recover the initial condition $g(x)$, given the final time measurement $m = \mathcal{S}(g) = u(\cdot,T)$. In this case, u satisfies the following equation:

$$\begin{cases} u_t + \mathcal{L}u = 0 & \text{in } \Omega\times(0,T), \\ u(x,t) = 0 & \text{on } \partial\Omega\times(0,T), \\ u(x,0) = g(x) & \text{in } \Omega\,, \end{cases} \tag{3.1}$$

Unlike the traditional POD method, we obtain the snapshots from the adjoint equation as follows:

$$\begin{cases} \tilde{u}_t + \mathcal{L}\tilde{u} = 0 & \text{in } \Omega\times(0,T), \\ \tilde{u}(x,t) = 0 & \text{on } \partial\Omega\times(0,T), \\ \tilde{u}(x,0) = m(x) & \text{in } \Omega\,, \end{cases} \tag{3.2}$$

In this case, the snapshots are generated by solving the adjoint equation (3.2) with the given final time measurement $m(x)$ as the initial condition.

Following the standard procedure in section 2.2, we generate the new POD basis functions ψ_k's from the snapshots $\{\tilde{u}(\cdot,t_0), \tilde{u}(\cdot,t_1), \ldots, \tilde{u}(\cdot,t_M)\}$, where $t_k = k\Delta t$ with $\Delta t = \frac{T}{M}$ and $k = 0,\ldots,M$. Then, we obtain the following error formula, similar to Proposition A.1:

$$\frac{\sum_{i=1}^{2M+1}\left\|\tilde{y}_i - \sum_{k=1}^{N_{\text{pod}}}\left(\tilde{y}_i, \psi_k(\cdot)\right)_{L^2(\Omega}\psi_k(\cdot)\right\|_{L^2(\Omega)}^2}{\sum_{i=1}^{2M+1}\|\tilde{y}_i\|_{L^2(\Omega)}^2} = \rho, \tag{3.3}$$

where the number N_{pod} is determined according to the decay of the ratio $\rho = \frac{\sum_{k=N_{\text{pod}}+1}^{2M+1}\lambda_k}{\sum_{k=1}^{2M+1}\lambda_k}$.

We consider using these new POD basis functions to approximate the forward problem to accelerate the computation. The fully discrete scheme is constructed on V_{pod}, and the solution is denoted by U_k for $k = 1\cdots M$ with $M = \frac{T}{\Delta t}$. Specifically, we seek numerical solutions U_k such that

$$(\bar{\partial}U_k, \psi) + a(U_k, \psi) = 0, \quad \forall \psi \in V_{pod}, \tag{3.4}$$

with $U_0 = g(x)$. We define the solution operator from the initial condition $g(x)$ to the final time solution U_M as $\mathcal{S}_{\text{pod}}$, i.e., $\mathcal{S}_{\text{pod}}g = U_M$.

3.1. Convergence of the adjoint-POD method. For any $g \in L^2(\Omega)$, we can write $g = \sum_{k=1}^{\infty} g_k \phi_k$ for a set of coefficients g_k. Let $u = \sum_{k=1}^{\infty} u_k(t)\phi_k$ be the solution of the problem (2.1). Substituting these two expressions of g and u into the first equation of (2.1), we get, by noting the fact that $L\phi_k = \mu_k \phi_k$ and comparing the coefficients of ϕ_k on both sides of the equation, that $u_k(0) = g_k$ and

$$u_k'(t) + \mu_k u_k = 0 \quad \text{in} \ \ (0, T).$$

We can write the solution as $u_k(T) = \alpha_k \, g_k$, with $\alpha_k = e^{-\mu_k T}$. Noting that $Sg = u(\cdot, T) = \sum_{k=1}^{\infty} u_k(T)\phi_k$, we can formally write

$$S\Big(\sum_{k=1}^{\infty} g_k \phi_k\Big) = \sum_{k=1}^{\infty} \alpha_k g_k \phi_k.$$

This representation shows the relationship between the initial condition g and the solution $u(\cdot, T)$ at the final time T. The operator S maps the initial condition to the solution at time T through the coefficients α_k, which depend on the eigenvalues μ_k of the operator L and the final time T. This relationship can be used to analyze the properties of the solution and the backward problem.

For simplicity, we approximate the source term $g(x)$ by a finite-dimensional truncation, i.e.

$$g_{\text{app}} = \sum_{k=1}^{L} g_k \phi_k. \tag{3.5}$$

Then, the solution $u(x,t)$ of the parabolic equation has the form:$u(\cdot, T) = \sum_{k=1}^{L} e^{-\mu_k T} f_k \phi_k$. After simple calculation, we can also have that $\tilde{u}(x,t) = \sum_{k=1}^{L} e^{-\mu_k T} e^{-\mu_k t} f_k \phi_k$. Constructing the POD basis functions is equivalent to computing the singular value decomposition of the matrix $\tilde{A} = (\tilde{y}_1, ..., \tilde{y}_M)$, where $\tilde{y}_j = (\tilde{u}(x_1, t_j), ..., \tilde{u}(x_N, t_j))^T$. Here $x_1, ..., x_N$ are the finite element nodes in Ω. Suppose A has the singular value decomposition: $A = U\Sigma V$, then the $\psi_k s$ are exactly the first M columns of U.

Let us denote $A = (y_1, ..., y_M)$ and $\tilde{A} = (\tilde{y}_1, ..., \tilde{y}_M)$, the matrix $\Phi = (\phi_1(\vec{x}, t_1), ..., \phi_L(\vec{x}, t_1))$, $F = \text{diag}(f_1, ..., f_L)$ and $D = \text{diag}(e^{-\mu_1 T}, ..., e^{-\mu_L T})$. We also define the $L \times M$ matrix J with entries $J(i,j) = e^{-\mu_i t_j}$. Since the normal orthogonality of the eigenfunctions ϕ_k's, Φ is a column orthogonal matrix. Using the formulations of u and $\tilde{u}$, we can represent the matrix A and $\tilde{A}$ as follows:

$$A = \Phi F J, \quad \text{and} \quad \tilde{A} = \Phi D F J. \tag{3.6}$$

The matrix representations of A and $\tilde{A}$ provide a concise way to express the relationship between the coefficients of the eigenfunctions ϕ_k and the solutions $u(x,t)$ and $\tilde{u}(x,t)$.

Proposition A.1 shows that the low-rank space V_{pod} provides the best N_{pod}-rank approximation of the column space of $\tilde{A}$. We aim to prove that V_{pod} is also a good approximation of the column space of A, which will confirm the effectiveness of the new POD method. To begin, we establish the connection between the matrices A and $\tilde{A}$.

LEMMA 3.1. *If $L \le M$, it follows that span$\{y_1, ..., y_M\}$ = span$\{\tilde{y}_1, ..., \tilde{y}_M\}$. Consequently, the column spaces of A and $\tilde{A}$ are identical, i.e., $C(A) = C(\tilde{A})$.*

Proof. The proof of this statement is analogous to that of Lemma 2.2, using the fact that the matrix J is a Vandermonde matrix. ∎

Using (3.3), we observe that the new POD basis provides a good approximation of $\{\tilde{y}_1, ..., \tilde{y}_M\}$. With the preceding analysis, we aim to demonstrate that the new POD basis is also an effective approximation of $\{y_1, ..., y_M\}$.

THEOREM 3.2. *Using the same notation in this section, we note that if a sufficient number of snapshots are available, i.e., $L \leq M$, then we have*

$$\frac{\sum_{i=1}^{M} ||y_i - P_{pod} y_i||^2_{L^2(\Omega)}}{\sum_{i=1}^{M} ||y_i||^2_{L^2(\Omega)}} \leq C e^{2\mu_L T} \rho, \tag{3.7}$$

where P_{pod} denotes the projection operator on the adjoint-POD space span$\{\psi_1, ...,$ $\psi_{N_{pod}}\}$, and $\rho = \frac{\sum_{k=N_{pod}+1}^{M} \lambda_k}{\sum_{k=1}^{M} \lambda_k}$.

Proof. In the following proof, we assume $L = M$ for simplicity. For the case $L < M$, the proof is similar. Using the same notation of Lemma 3.1, Φ and J are both invertible square matrices. Then, there exists a unique matrix P such that,

$$\Phi D F J P = \Phi F J, \tag{3.8}$$

where $P = J^{-1} D^{-1} J$. Therefore, we have $y_j = \sum_{i=1}^{L} P_{ij} \tilde{y}_i$.

Using the Cauchy-Schwarz inequality, we obtain the following estimate for any $1 \leq j \leq L$,

$$\|y_j - P_{\text{pod}} y_j\|^2 \leq \sum_{i=1}^{L} P_{ij}^2 \sum_{i=1}^{L} \|\tilde{y}_i - P_{\text{pod}} \tilde{y}_i\|^2. \tag{3.9}$$

Hence, we can obtain that

$$\sum_{j=1}^{L} \|y_j - P_{\text{pod}} y_j\|^2 \leq \sum_{i,j=1}^{L} P_{ij}^2 \sum_{i=1}^{L} \|\tilde{y}_i - P_{\text{pod}} \tilde{y}_i\|^2 = \|P\|_F^2 \sum_{i=1}^{L} \|\tilde{y}_i - P_{\text{pod}} \tilde{y}_i\|^2. \tag{3.10}$$

The remaining part is to estimate the Frobenius norm of P. Since $P = J^{-1} D^{-1} J$, we define that $\|P\|_d = \|D^{-1}\|_2$. It is easy to verify that $\|\cdot\|_d$ is a matrix norm. Then, we obtain

$$\|P\|_F \leq C \|P\|_d = C \|D^{-1}\|_2 \leq C e^{\mu_L T}. \tag{3.11}$$

On the other hand, since Φ is an orthogonal matrix, we have

$$\sum_{j=1}^{L} \|\tilde{y}_j\|^2 = \|\Phi D F J\|_F^2 = \|D F J\|_F^2 \leq \|D\|_F^2 \|F J\|_F^2 \tag{3.12}$$

$$\leq C \|F J\|_F^2 \leq C \sum_{j=1}^{L} \|y_j\|^2. \tag{3.13}$$

Aligned with the equation (3.3), we can derive the following estimate:

$$\frac{\sum_{i=1}^{M} ||y_i - P_{pod} y_i||^2_{L^2(\Omega)}}{\sum_{i=1}^{M} ||y_i||^2_{L^2(\Omega)}} \leq C \|P\|_F^2 \frac{\sum_{i=1}^{M} ||\tilde{y}_i - P_{pod} \tilde{y}_i||^2_{L^2(\Omega)}}{\sum_{i=1}^{M} ||\tilde{y}_i||^2_{L^2(\Omega)}} \tag{3.14}$$

$$\leq C e^{2\mu_L T} \rho. \tag{3.15}$$

Therefore, we can conclude the inequality with the estimated value of $\mu_i \leq C i^{2/d}$. ∎

3.2. Stochastic convergence of backward problem. To solve this backward problem, we use the traditional Tikhonov regularization method,

$$\min_{g\in X} \|\mathcal{S}(g) - m\|^2_{L^2(\Omega)} + \lambda\|g\|^2_{L^2(\Omega)}, \tag{3.16}$$

where λ is the regularization parameter.

In the discrete approximation of the regularization problem (3.16), we solve the following least-squares regularized optimization problem:

$$\min_{g\in V_{\rm pod}} \|\mathcal{S}_{\rm pod}(g) - m\|^2_{L^2(\Omega)} + \lambda\|g\|^2_{L^2(\Omega)}. \tag{3.17}$$

We define the functional $\mathcal{J}_{\rm pod}[g] = \|\mathcal{S}_{\rm pod}g - m\|^2_{L^2(\Omega)} + \lambda\|g\|^2_{L^2(\Omega)}$. We can compute the Fréchet derivative of $\mathcal{J}_{\rm pod}[g]$ and obtain the following iterative scheme:

$$g_{k+1} = g_k - \beta d\mathcal{J}_{pod}[g_k], \quad \forall k \in \mathbb{N}, \tag{3.18}$$

where β is the step size, $d\mathcal{J}_{\rm pod}[g] = \mathcal{S}^*_{\rm pod}(\mathcal{S}_{\rm pod}g - m) + \lambda f$, and g_0 is an initial guess.

The above theory is based on the assumption of a noise-free case, where the final time measurement $m = u(\cdot, T)$ is known exactly. However, in practical applications, the measurement data often contains uncertainty. To account for this uncertainty, we assume that the measurement data takes the discrete form $m^n_i = u(d_i, T) + e_i$ for $i = 1, \cdots, n$, where e_i denotes the measurement error at the i-th point. Based on the properties of parabolic equations [36], we know that $\|u\|_{C([0,T];H^2(\Omega))} \le C\|g\|_{L^2(\Omega)}$. Furthermore, according to the embedding theorem of Sobolev spaces, we know that $H^2(\Omega)$ is continuously embedded into $C(\bar{\Omega})$, ensuring that $u(\cdot, T)$ is well-defined pointwise for all $d_i \in \Omega$.

We repeat the procedure of (2.28) in section 2.3. If we choose the optimal regularization parameter $\alpha^{1/2+d/8} = O(\sigma n^{-1/2}\|u(\cdot, T)\|^{-1}_{H^2(\Omega)})$ (in practice, we chose the optimal parameter using Algorithm 2.3), then the denoised solution m achieves the optimal convergence with noise of bounded variance σ^2 given by

$$\mathbb{E}\big[\|u(\cdot, T) - m\|_{L^2(\Omega)}\big] \le C\alpha^{1/2}\|u(\cdot, T)\|_{H^2(\Omega)}. \tag{3.19}$$

Assuming that the noise $\{e_i\}^n_{i=1}$ are independent Gaussian random variables with a variance of σ, we can obtain a further result. Specifically, we have:

$$\mathbb{P}(\|u(\cdot, T) - m\|_{L^2(\Omega)} \ge \alpha^{1/2}\|u(\cdot, T)\|_{H^2(\Omega)}z) \le 2e^{-Cz^2}. \tag{3.20}$$

In practice, the noise level and the true initial term may be unknown. To determine the optimal regularization parameter α for denoising, we can apply Algorithm 2.3. This algorithm will select a value of α that balances the trade-off between reducing noise and preserving important features of the solution. Using the recovered function $m(x)$, we generate the adjoint POD basis in Section 2.1. It can be shown that, even with uncertainty, the POD basis still provides a good low-rank approximation of the snapshots $\{y_1, ..., y_M\}$. Combining Theorem 3.2 and (3.19), we can deduce that for any $1 \le i \le M$, the following inequality holds:

$$||y_i - P_{\rm pod}y_i||^2_{L^2(\Omega)} \le C(Me^{2\mu_L T}\rho + \alpha)\|g\|^2_{L^2(\Omega)}. \tag{3.21}$$

Since we replace the source term with a finite truncation (3.5), we make the assumption that:

$$\|g - g_{\rm app}\|^2_{L^2} \le \varepsilon(L). \tag{3.22}$$

Using similar techniques to those employed in the proof of Theorem 3.7 in [36], we can derive the following convergence results based on the aforementioned assumptions.

THEOREM 3.3. *Let $\{e_i\}_{i=1}^n$ be independent random variables satisfying $\mathbb{E}[e_i] = 0$ and $\mathbb{E}[e_i^2] \le \sigma^2$ for $i = 1, \cdots, n$. If we set $\alpha^{1/2+d/8} = O(\sigma n^{-1/2} \|u(\cdot, T)\|_{H^2(\Omega)}^{-1})$ in (2.28) and $\lambda = O(Me^{2\mu_L T}\rho + \alpha)$, then we obtain:*

$$\mathbb{E}\big[\|\mathcal{S}g^* - \mathcal{S}_{pod}g_{pod}\|_{L^2(\Omega)}^2\big] \le C\lambda \|g^*\|_{L^2(\Omega)}^2 + C\varepsilon, \tag{3.23}$$

and

$$\mathbb{E}\big[\|g^* - g_{pod}\|_{L^2(\Omega)}^2\big] \le C\|g^*\|_{L^2(\Omega)}^2 + C\varepsilon. \tag{3.24}$$

Furthermore, if we assume that the noise $\{e_i\}_{i=1}^n$ are independent Gaussian random variables with a variance of σ, we can obtain a stronger type of convergence. A similar proof can be found in [36]. Here, we only present the resulting convergence results.

THEOREM 3.4. *Let $\{e_i\}_{i=1}^n$ be independent Gaussian random variables with a variance of σ. If we set $\alpha^{1/2+d/8} = O(\sigma n^{-1/2} \|u(\cdot, T)\|_{H^2(\Omega)}^{-1})$ in (2.28) and $\lambda = O(Me^{2\mu_L T}\rho + \alpha)$, then for any $z > 0$, there exists a constant C such that:*

$$\mathbb{P}(\|S_{pod}g_{pod} - Sg^*\|_{L^2(\Omega)} \ge (\lambda^{1/2}\|g^*\|_{L^2(\Omega)} + \varepsilon)z) \le 2e^{-Cz^2}, \tag{3.25}$$

and

$$\mathbb{P}(\|g_{pod} - g^*\|_{L^2(\Omega)} \ge (\|g^*\|_{L^2(\Omega)} + \varepsilon)z) \le 2e^{-Cz^2}. \tag{3.26}$$

Theorem 3.3 and Theorem 3.4 provide the stochastic convergence rate as $\|\mathcal{S}g^* - \mathcal{S}_{\text{pod}}g_{\text{pod}}\|_{L^2(\Omega)}^2 = O_p(\lambda\|g^*\|_{L^2(\Omega)}^2 + \varepsilon)$ in terms of the values λ and ε. These results also provide useful guidance for practical numerical simulations.

4. Numerical examples.

This section presents several numerical examples to demonstrate the reconstruction results for the inverse source problem and the backward problem discussed in this paper. Specifically, we consider a parabolic equation as follows:

$$\begin{cases} u_t - \Delta u = f(x) & \text{in } \Omega \times (0, T), \\ u(x,t) = 0 & \text{on } \partial\Omega \times (0, T), \\ u(x,0) = g(x) & \text{in } \Omega\,, \end{cases} \tag{4.1}$$

where the domain $\Omega = [0, \pi]^2$, $f(x)$ is the source term, and $g(x)$ is the initial condition. For each observation data set, we first use the backward Euler scheme in time and the linear finite element method (specifically, the P1 element) in space with a mesh size of $h = 1/50$ and a time step of $\Delta t = T/400$. We use 9 POD basis functions (usually less basis is enough) to compute the inverse problems unless otherwise specified. For the FEM, there are approximately 2500 basis functions.

In [35], the authors have already compared the efficiency of the POD method and the finite element method for solving this inverse problem. They have demonstrated that the POD method achieves a speed-up of at least 6 times, even with 400 finite element basis functions. Therefore, we do not include a comparison with the finite element method in this paper. However, it is worth noting that as the number of finite element basis functions increases, the potential for the POD method to achieve greater speed-up also grows correspondingly.

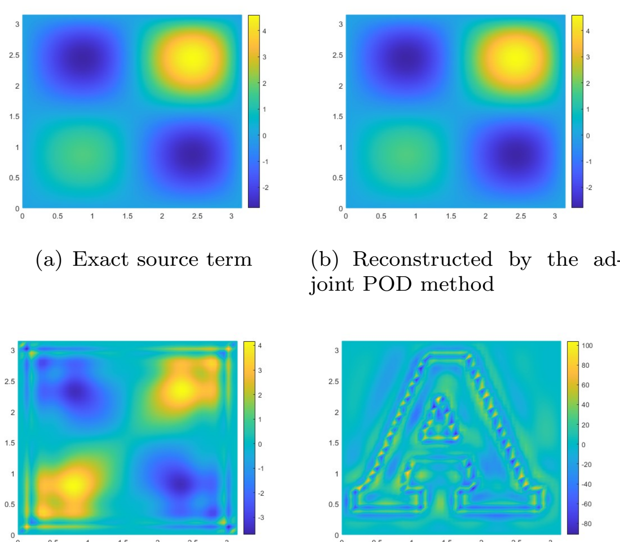
4.1. Numerical results for inverse source problems. In the following examples, we use the adjoint POD method to recover the source term f as described in Section 2. We obtain the data for the forward problem with the exact source term f at the final time $T = 1$.

Example 4.1 We first demonstrate the importance of choosing the appropriate POD basis functions. For the same source term, we apply different right-hand sides of equation (2.1) to obtain the POD basis functions. Subsequently, we solve the inverse source problem using these different POD basis functions.

Figure 1 illustrates this process. The true source term is $\sin(2x)\sin(2y)e^{\frac{x+y}{\pi}}$, and its surface plot is shown in Figure 1 (a). Figure 1 (b) presents the reconstruction result obtained using the adjoint POD method proposed in this paper, indicating that our new method effectively and efficiently recovers the source term. Figure 1 (c) displays the result when an incorrect right-hand side is used to generate the POD basis functions. In this case, we use $\sin(x)\sin(y)$ as the right-hand side in (2.1) to generate the POD basis functions. Figure 1 (d) shows the result when we use an A-shaped function as the right-hand side to derive the POD basis.

As can be seen, Figure 1 (b) provides a good reconstruction, whereas Figures 1 (c) and (d) yield inaccurate results. The result in Figure 1 (d) is particularly striking, as the recovered image deviates significantly from the exact source term.

In this example, we demonstrate the importance of selecting appropriate basis functions for solving inverse problems. Using an unsuitable set of basis functions can lead to incorrect results. Our proposed adjoint POD method offers a set of suitable basis functions for such problems. In the following examples, we will compare our adjoint POD basis functions with the original true POD basis to validate Theorem 2.3.



(a) Exact source term

(b) Reconstructed by the adjoint POD method

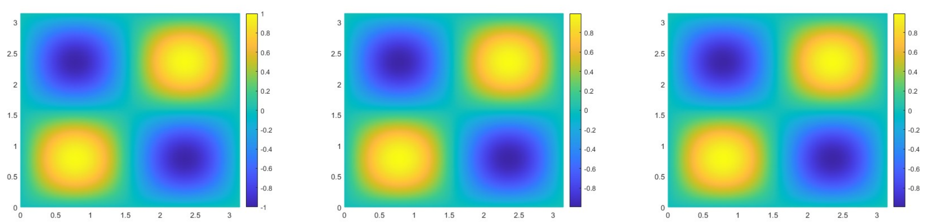
(c) Using the POD basis generated with the right hand side $\sin(x)\sin(y)$

(d) Using the POD basis generated with the right hand side of A shaped function

FIG. 1. *The importance of choice of POD basis*

Example 4.2 In this example, we first use the true source term as the right-hand side in (2.1) to generate the POD basis. Then, we generate the POD basis using our proposed adjoint POD method. To validate Theorem 2.3, we compare both sets of basis functions and assess their similarity. Figure 2 shows the results when using the exact source term $f^* = \sin(2x)\sin(2y)$. Figure 2 (b) and Figure 2 (c) show that both the traditional POD method and our adjoint POD method recover the true source term accurately. However, our adjoint POD method does not require prior knowledge of the exact source term, while the traditional POD method does, leading to the so-called inverse crime.

Figure 3 presents the results for an exact source term f^* in the form of a Z-shaped function. This example also illustrates the efficiency of the POD method in solving inverse problems compared to the finite element method. Figure 3 (c) and Figure 3 (e) show that both traditional POD and our adjoint POD method yield basis functions that contain critical information about the exact source term we aim to recover. In contrast, the basis functions of the finite element method do not contain any prior information about the true function we need to recover.



(a) Exact source term (b) Recovered result by traditional POD with exact right hand side (c) Result by the adjoint POD

FIG. 2. *Comparison of basis between traditional POD and the adjoint POD for* $f = \sin(2x)\sin(2y)$.

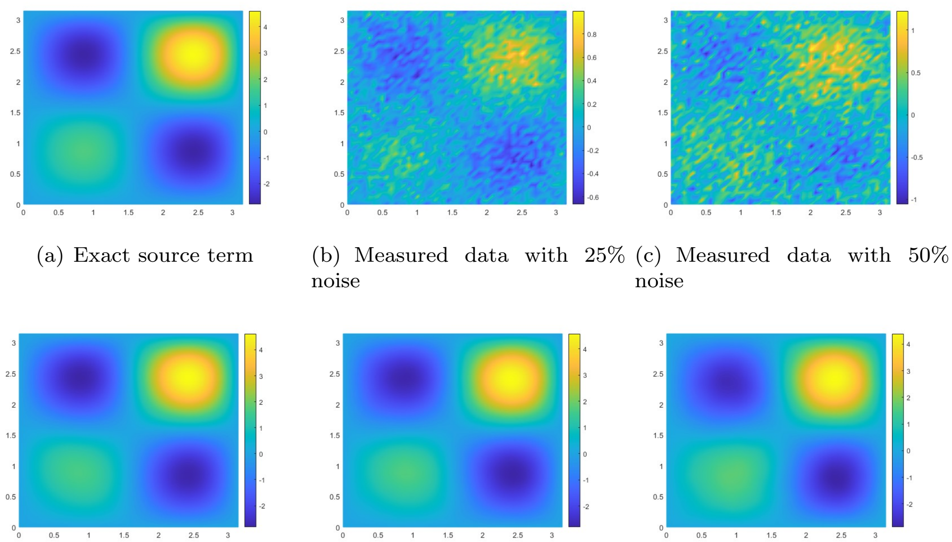
(a) Exact source term (b) Recovered result by traditional POD with exact right hand side (c) Result by the adjoint POD

FIG. 3. *Comparison of basis between traditional POD and the adjoint POD for* f^* *of Z-shaped function.*

In the previously mentioned cases, all measured data were noise-free. We will now examine highly challenging cases with noise levels ranging from 10% to 50% to test the denoising method described in Section 2.3.

Example 4.3 In this example, we will evaluate the robustness of the adjoint POD method in the presence of noise. We consider the measurement data to be $m_i^n = u(d_i, T) + \sigma e_i$, $i = 1, \cdots, n$, where d_i represents positions within the domain Ω, and $\{e_i\}_{i=1}^n$ are independent standard normal random variables. We will take 2500 positions d_i uniformly distributed over the domain Ω.

Figure 4 demonstrates the robustness of our method in the presence of significant noise. Even with a 50% noise level, where the measured data is entirely obscured by noise as shown in Figure 4 (c), our method is still able to recover the source term as depicted in Figure 4 (f).



(a) Exact source term (b) Measured data with 25% noise (c) Measured data with 50% noise

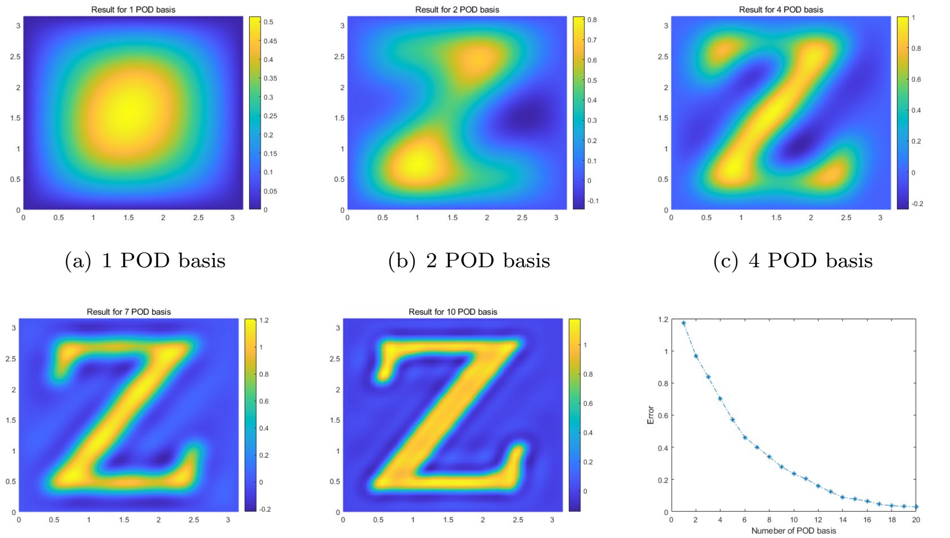
(d) Recovered result for 10% noise (e) Recovered result for 25% noise (f) Recovered result for 50% noise

FIG. 4. *Robustness of the adjoint POD against the noise for* $f = sin(2x)sin(2y)e^{\frac{x+y}{\pi}}$.

Example 4.4 In this example, we demonstrate the effect of choosing different numbers of our new POD basis and verify the convergence of our POD method. Figure 5 illustrates this test. We see that as the number of our POD basis increases, the reconstruction error decays rapidly; see Figure 5 (f). We observe exponential decay from our numerical experiments. Even for a complicated function with a 'Z' shape, we can obtain satisfactory results with just 10 POD basis functions. As the number of our POD basis increases, the reconstruction becomes increasingly accurate, as depicted in Figure 5 (a) - (e).

4.2. Numerical results for backward problems. In this subsection, we apply the new POD method to recover the initial condition g, as discussed in Section 3. We collect the data at the time $T = 0.05$. Since most of the examples are similar to what we did in section 4.1, we may not explain the setting of numerical examples in detail.

Example 4.5 In this example, we demonstrate the importance of POD basis, similar to what we did in Example 4.1. We solve the backward problem using different POD basis functions.



(a) 1 POD basis (b) 2 POD basis (c) 4 POD basis

(d) 7 POD basis (e) 10 POD basis (f) Error decay as the number of POD basis grows

FIG. 5. *Effects of the numbers of POD basis.*

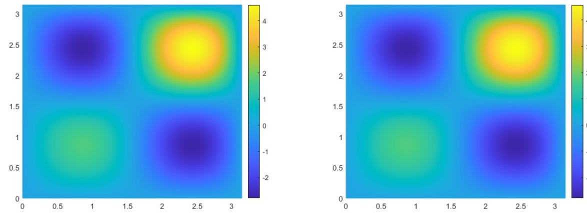
Figure 6 illustrates the corresponding results. The true initial condition is $\sin(2x)\sin(2y)e^{\frac{x+y}{\pi}}$, and its surface plot is shown in Figure 6 (a). Figure 6 (b) presents the reconstruction result using our adjoint POD method. Figure 6 (c) shows the result when we use an incorrect initial condition to generate the POD basis functions, and Figure 6 (d) displays the result when we use an A-shaped function as the initial condition to generate the POD basis functions.

Example 4.6 In the following two examples, we adopt the approach of example 4.2. We first use the true initial condition as the initial condition in Eq. (3.1) to generate the POD basis functions. Then, we generate the POD basis functions using our proposed adjoint POD method.

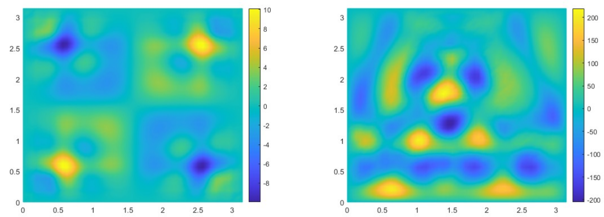
Figure 7 shows the results obtained by using the exact source term $g^* = \sin(2x)\sin(2y)e^{\frac{x+y}{\pi}}$. Figures 7 (b) and 7 (c) demonstrate that both the traditional POD and our adjoint POD methods work well in recovering the true initial condition. However, our adjoint POD method does not require prior knowledge of the exact initial condition. Figure 8 shows the results using the exact initial condition g^* of the A-shaped function.

In the above numerical examples in this subsection, all the measured data are noise-free. Now, we test the denoising method discussed in Section 3.2 by examining very challenging cases with noise levels ranging from 10% to 50%.

Example 4.7 Similar to Example 4.3, we test the robustness of the adjoint POD method against noise. We take the measurement data as $m_i^n = u(d_i, T) + \sigma e_i$, for $i = 1, \cdots, n$, where d_is represent the positions inside Ω, $\{e_i\}_{i=1}^n$ are independent standard normal random variables, and σ is the noise level. We use 2500 positions d_i uniformly distributed over the domain Ω. Figure 9 demonstrates that our method is robust even in the presence of significant noise. Remarkably, even with 50% noise, when the measured data is completely obscured by noise as shown in Figure 9 (c), we

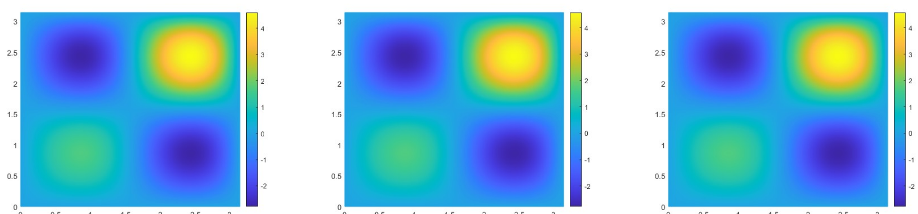


(a) Exact initial condition (b) POD method proposed in this paper



(c) Using the POD basis generated with the initial term $sin(x)sin(y)$ (d) Using the POD basis generated with the initial term of A-shaped function

FIG. 6. *The importance of choice of POD basis*

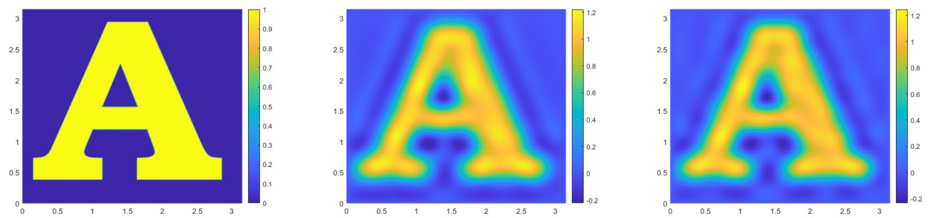


(a) Exact initial condition (b) Recovered result by traditional POD with exact initial term (c) Result by the adjoint POD

FIG. 7. *Comparison of basis between traditional POD and the adjoint POD for* $f = sin(2x)sin(2y)e^{\frac{x+y}{\pi}}$.

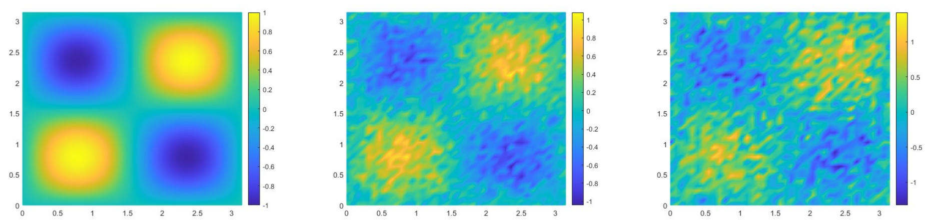
can still recover the initial condition, as seen in Figure 9 (f).

Example 4.8 Finally, we study a more interesting case. Although the inverse source problem and the backward problem are two distinct problems, we have observed from the previous numerical examples that they share some commonalities. Specifically, the POD basis functions for both problems contain critical information about the functions one wants to recover. As a result, we employ the POD basis functions derived from the inverse source problem to solve the backward problem. Please refer to Figure 10 for the reconstruction results. This example also shows that the inverse source problem and the backward problem share similar basis functions from the point of view of the POD method. This similarity can also be explained by the theory of semi-groups and spectral analysis.

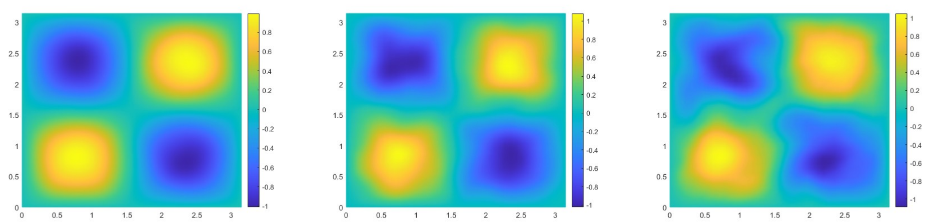


(a) Exact initial condition (b) Recovered result by traditional POD with exact initial term (c) Result by the adjoint POD

FIG. 8. *Comparison of basis between traditional POD and the adjoint POD for f^* of A shaped function.*



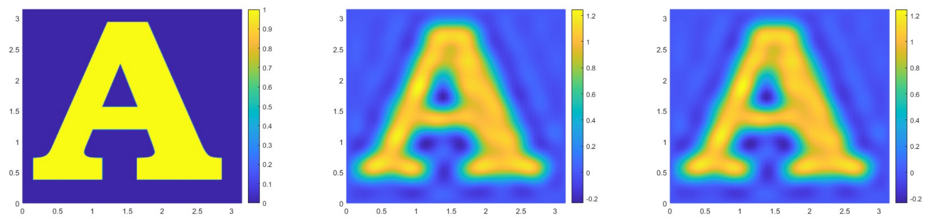
(a) Exact initial condition (b) Measured data with 25% noise (c) Measured data with 50% noise



(d) Recovered result for 10% noise (e) Recovered result for 25% noise (f) Recovered result for 50% noise

FIG. 9. *Robustness of the adjoint POD against the noise for $g = sin(2x)sin(2y)$.*

5. Conclusion. In this paper, we have developed a novel POD-based model reduction method for solving parabolic inverse problems, specifically focusing on inverse source and inverse initial value problems. By leveraging the intrinsic low-dimensional structures in the parabolic equations, we have successfully developed POD basis functions that reduce computational costs while maintaining accuracy in solving parabolic inverse problems. Our primary contribution is the effective addressing of the inverse crime issue associated with traditional model reduction methods for solving PDEs, which is a common challenge encountered by traditional POD methods for solving inverse problems. Furthermore, we have provided convergence analysis for our proposed method and demonstrated its accuracy and efficiency through numerical examples. The results show that our proposed method overcomes the inverse crime issue and produces superior results compared to traditional methods. Therefore, our proposed



(a) Exact initial condition

(b) Recovered result using the POD basis from backward problem

(c) Recovered result using the POD basis from inverse source problem

FIG. 10. *Solve the backward problem using the basis from the inverse source problem.*

method provides a promising avenue for solving parabolic inverse problems with high accuracy and computational efficiency.

Acknowledgement. The research of W. Zhang is supported by the National Natural Science Foundation of China under grant numbers 12371423 and 12241104. The research of Z. Zhang is supported by the National Natural Science Foundation of China under grant number 12171406, Hong Kong RGC grant projects 17307921 and 17304324, Seed Funding for Strategic Interdisciplinary Research Scheme 2021/22 (HKU), and Seed Funding from the HKU-TCL Joint Research Centre for Artificial Intelligence.

Appendix A. Proper orthogonal decomposition (POD) method. Assuming that $u \in H_0^1(\Omega)$ is the solution to the weak formulation of the parabolic equation (1.1), the construction of POD basis functions requires solution snapshots. These solution snapshots can be obtained through various technological means related to a specific application, such as experimental data or numerical methods. For instance, one can measure the solution at different times using sensors in experimental settings to obtain the solution snapshots. Alternatively, one can use FEM to solve the parabolic equation numerically and obtain the solution snapshots.

Given a set of solutions at different time instances $\big\{u(\cdot,t_0), u(\cdot,t_1), \ldots, u(\cdot,t_M)\big\}$, where $t_k = k\Delta t$ with $\Delta t = \frac{T}{M}$ and $k = 0, \ldots, M$, we first obtain the solution snapshots $\{y_1, \ldots, y_{M+1}, y_{M+2}, \ldots, y_{2M+1}\}$, where $y_k = u(\cdot, t_{k-1})$, $k = 1, \ldots, M+1$, and $y_k = \overline{\partial} u(\cdot, t_{k-M-1})$, $k = M+2, \ldots, 2m+1$ with $\overline{\partial} u(\cdot, t_k) = \frac{u(\cdot,t_k) - u(\cdot,t_{k-1})}{\Delta t}$, $k = 1, \ldots, M$.

The POD basis functions $\{\psi_k\}_{k=1}^{N_{\text{pod}}}$ are constructed by minimizing the following projection error:

$$\text{(A.1)} \qquad \frac{1}{2m+1}\Big(\sum_{j=0}^{M} \big\| u(t_j) - \sum_{k=1}^{N_{\text{pod}}} (u(t_j), \psi_k)_{L^2(\Omega)} \psi_k \big\|_{L^2(\Omega)}^2$$

$$\text{(A.2)} \qquad + \sum_{j=1}^{M} \big\| \overline{\partial} u(t_j) - \sum_{k=1}^{N_{\text{pod}}} (\overline{\partial} u(t_j), \psi_k)_{L^2(\Omega)} \psi_k \big\|_{L^2(\Omega)}^2 \Big)$$

subject to the constraints that $\big(\psi_{k_1}(\cdot), \psi_{k_2}(\cdot)\big)_{L^2(\Omega)} = \delta_{k_1 k_2}$, $1 \le k_1, k_2 \le N_{\text{pod}}$, where $\delta_{k_1 k_2} = 1$ if $k_1 = k_2$, otherwise $\delta_{k_1 k_2} = 0$. Here, N_{pod} denotes the number of POD basis functions that will be extracted from solution snapshots.

This manuscript is for review purposes only.

Let $V_{\text{pod}} = \text{span}\{\psi_1, \dots, \psi_{N_{\text{pod}}}\}$ denote the finite-dimensional space spanned by the POD basis functions. Using the method of snapshot proposed by Sirovich [33], we know that the minimizing problem can be reduced to the following eigenvalue problem:

$$Kv = \mu v, \tag{A.3}$$

where the correlation matrix K is computed from the solution snapshots $\{y_1, y_2, \dots, y_{2M+1}\}$ with entries $K_{ij} = (y_i, y_j)_{L^2(\Omega)}$, $i, j = 1, \dots, 2M+1$, and K is symmetric and semi-positive definite. We sort the eigenvalues in a decreasing order as $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{2m+1}$ and the corresponding eigenvectors are denoted by v_k, $k = 1, ..., 2M+1$. It can be shown that if the POD basis functions are constructed by

$$\varphi_k(\cdot) = \frac{1}{\sqrt{\lambda_k}} \sum_{j=1}^{2M+1} (v_k)_j u(\cdot, t_j), \quad 1 \leq k \leq N_{\text{pod}}, \tag{A.4}$$

where $(v_k)_j$ is the j-th component of the eigenvector v_k, they minimize the projection error.

The approximation error for the POD method has been studied extensively in the literature, particularly in the works [21] and [3].

Proposition A.1 (Section 3.3.2, [21] or p. 502, [3]). *Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{2M+1} \geq 0$ denote the non-negative eigenvalues of the correlation matrix K in the eigenvalue problem* (A.3)*. Then, $\{\psi_k\}_{k=1}^{N_{pod}}$ constructed according to the method of snapshots* (A.4) *is the set of POD basis functions, and we have the following error formula:*

$$\frac{\sum_{i=1}^{2M+1} \left|\left|\tilde{y}_i - \sum_{k=1}^{N_{pod}} \left(\tilde{y}_i, \psi_k(\cdot)\right)_{L^2(\Omega}\psi_k(\cdot)\right|\right|^2_{L^2(\Omega)}}{\sum_{i=1}^{2M+1} ||\tilde{y}_i||^2_{L^2(\Omega)}} = \frac{\sum_{k=N_{pod}+1}^{2M+1} \lambda_k}{\sum_{k=1}^{2M+1} \lambda_k}, \tag{A.5}$$

where the number N_{pod} is determined by the decay of the ratio $\rho = \frac{\sum_{k=N_{pod}+1}^{2M+1} \lambda_k}{\sum_{k=1}^{2M+1} \lambda_k}$.

REFERENCES

[1] S. Agmon, *Lectures on Elliptic Boundary Value Problems*, Van Norstrand, Princeton, NJ, 1965.

[2] A. Alla and M. Falcone, *A time-adaptive POD method for optimal control problems*, IFAC Proceedings Volumes, 46 (2013), pp. 245–250.

[3] P. Benner, S. Gugercin, and K. Willcox, *A survey of projection-based model reduction methods for parametric dynamical systems*, SIAM Review, 57 (2015), pp. 483–531.

[4] G. Berkooz, P. Holmes, and J. Lumley, *The proper orthogonal decomposition in the analysis of turbulent flows*, Annual review of fluid mechanics, 25 (1993), pp. 539–575.

[5] Z. Chen, R. Tuo, and W. Zhang, *Stochastic convergence of a nonconforming finite element method for the thin plate spline smoother for observational data*, SIAM Journal on Numerical Analysis, 56 (2018), pp. 635–659.

[6] Z. Chen, W. Zhang, and J. Zou, *Stochastic convergence of regularized solutions and their finite element approximations to inverse source problems*, SIAM Journal on Numerical Analysis, 60 (2022), pp. 751–780.

[7] J. Cheng, Y. Ke, and T. Wei, *The backward problem of parabolic equations with the measurements on a discrete set*, Journal of Inverse and Ill-posed Problems, 28 (2020), pp. 137–144.

[8] J. Cheng and J. Liu, *An inverse source problem for parabolic equations with local measurements*, Applied Mathematics Letters, 103 (2020), p. 106213.

[9] W. Cheng, L. Zhao, and C. Fu, *Source term identification for an axisymmetric inverse heat conduction problem*, Computers & mathematics with applications, 59 (2010), pp. 142–148.
[10] S. D'haeyer, B. Johansson, and M. Slodička, *Reconstruction of a spacewise-dependent heat source in a time-dependent heat diffusion process*, IMA Journal of Applied Mathematics, 79 (2014), pp. 33–53.
[11] A. El Badia, A. El Hajj, M. Jazar, and H. Moustafa, *Lipschitz stability estimates for an inverse source problem in an elliptic equation from interior measurements*, Applicable Analysis, 95 (2016), pp. 1873–1890.
[12] A. El Badia, T. Ha Duong, and F. Moutazaim, *Numerical solution for the identification of source terms from boundary measurements*, Inverse Problems in Engineering, 8 (2000), pp. 345–364.
[13] A. El Badia and T. Nara, *An inverse source problem for Helmholtz's equation from the Cauchy data with a single wave number*, Inverse Problems, 27 (2011), p. 105001.
[14] J. Fleckinger and M. Lapidus, *Eigenvalues of elliptic boundary value problems with an indefinite weight function*, Transactions of the American Mathematical Society, 295 (1986), pp. 305–324, http://www.jstor.org/stable/2000158.
[15] G. Garcia, A. Osses, and M. Tapia, *A heat source reconstruction formula from single internal measurements using a family of null controls*, Journal of Inverse and Ill-posed Problems, 21 (2013), pp. 755–779.
[16] S. Gorelick, B. Evans, and I. Remson, *Identifying sources of groundwater pollution: an optimization approach*, Water Resources Research, 19 (1983), pp. 779–790.
[17] H. Gu, J. Xin, and Z. Zhang, *Error estimates for a POD method for solving viscous G-equations in incompressible cellular flows*, SIAM Journal on Scientific Computing, 43 (2021), pp. A636–A662.
[18] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*, Courier Corporation, 2003.
[19] A. Hasanoğlu and V. Romanov, *Introduction to inverse problems for differential equations*, Springer, 2021.
[20] J. Hesthaven, G. Rozza, and B. Stamm, *Certified reduced basis methods for parametrized partial differential equations*, Springer, 2016.
[21] P. Holmes, J. Lumley, and G. Berkooz, *Turbulence, coherent structures, dynamical systems and symmetry*, Cambridge University Press, 1998.
[22] J. Idier, *Bayesian approach to inverse problems*, John Wiley & Sons, 2013.
[23] V. Isakov, *Inverse problems for partial differential equations*, vol. 127, Springer, 2006.
[24] T. Johansson and D. Lesnic, *Determination of a spacewise dependent heat source*, Journal of computational and Applied Mathematics, 209 (2007), pp. 66–80.
[25] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, vol. 160, Springer, New York, 2005.
[26] J. Krebs, A. Louis, and H. Wendland, *Sobolev error estimates and a priori parameter selection for semi-discrete tikhonov regularization*, Journal of Inverse and Ill-posed Problems, 17 (2009), pp. 845–869.
[27] K. Kunisch and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parabolic problems*, Numerische Mathematik, 90 (2001), pp. 117–148.
[28] K. Kunisch, S. Volkwein, and L. Xie, *HJB-POD-based feedback design for the optimal control of evolution problems*, SIAM Journal on Applied Dynamical Systems, 3 (2004), pp. 701–722.
[29] X. Liu and Z. Zhai, *Inverse modeling methods for indoor airborne pollutant tracking: literature review and fundamentals.*, Indoor air, 17 (2007), pp. 419–438.
[30] W. Muniz, H. Velho, and F. Ramos, *A comparison of some inverse methods for estimating the initial condition of the heat equation*, Journal of Computational and Applied Mathematics, 103 (1999), pp. 145–163.
[31] P. Nelson and S. Yoon, *Estimation of acoustic source strength by inverse methods: Part i, conditioning of the inverse problem*, Journal of sound and vibration, 233 (2000), pp. 639–664.
[32] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced basis methods for partial differential equations: an introduction*, vol. 92, Springer, 2015.
[33] L. Sirovich, *Turbulence and the dynamics of coherent structures. I. Coherent structures*, Quarterly of applied mathematics, 45 (1987), pp. 561–571.
[34] S. Volkwein, *Proper orthogonal decomposition: Theory and reduced-order modelling*, Lecture Notes, University of Konstanz, 4 (2013).
[35] Z. Wang, W. Zhang, and Z. Zhang, *A data-driven model reduction method for parabolic inverse source problems and its convergence analysis*, Journal of Computational Physics,

487 (2023), p. 112156.

[36] Z. Wang, W. Zhang, and Z. Zhang, *Stochastic convergence of regularized solutions for backward heat conduction problems*, arXiv:2311.03623, (2023).