Goal-Oriented Inference: Approach, Linear Theory, and Application to Advection Diffusion

SIAM J. SCI. COMPUT.
Vol. 34, No. 4, pp. A1880–A1904

GOAL-ORIENTED INFERENCE: APPROACH, LINEAR THEORY, AND APPLICATION TO ADVECTION DIFFUSION*

CHAD LIEBERMAN† AND KAREN WILLCOX†

Abstract. Inference of model parameters is one step in an engineering process often ending in predictions that support decision in the form of design or control. Incorporation of end goals into the inference process leads to more efficient goal-oriented algorithms that automatically target the most relevant parameters for prediction. In the linear setting the control-theoretic concepts underlying balanced truncation model reduction can be exploited in inference through a dimensionally optimal subspace regularizer. The inference-for-prediction method exactly replicates the prediction results of either truncated singular value decomposition, Tikhonov-regularized, or Gaussian statistical inverse problem formulations independent of data; it sacrifices accuracy in parameter estimate for online efficiency. The new method leads to low-dimensional parameterization of the inverse problem enabling solution on smartphones or laptops in the field.

Key words. inference, observability, model reduction, balanced truncation, PDEs

AMS subject classifications. 35R15, 65M60, 62F30, 93B07, 15A18

DOI. 10.1137/110857763

1. Introduction.

The process of utilizing experimental data to estimate unknown parameters is central to many important problems in science and engineering. Inference problems arise in medical imaging [3], geophysics [7], meteorology and oceanography [13], heat transfer [1], electromagnetic scattering [11], and electrical impedance tomography [2], among many other disciplines.

Many inverse problems are ill-posed; the data do not determine a unique solution. Inference approaches, therefore, rely on the injection of prior information. In deterministic formulations [9], this prior information is often manifested as a form of regularization. In Bayesian statistical formulations [17], the prior information is used to formulate a prior distribution reflecting the belief in probable parameter values. As a result, the distinction becomes blurred between inferred parameter modes informed by data and modes influenced largely or wholly by prior information. Without careful design of prior information, this injection of outside information can overshadow the information contained in the limited data that are obtained. Although ill-posedness will always be an issue to some extent in limited data settings, in this paper we show that it is possible to partially circumvent the deleterious effects of the use of regularizers or prior information by incorporating end goals.

While in some cases estimation of unknown parameters is the end goal, there are many engineering processes where parameter estimation is one step in a multi-step process ending with design. In such scenarios, engineers often define output quantities of interest to be optimized by the design. In consideration of this fact, we propose a goal-oriented approach to inference that accounts for the output quantities of interest. Generally, in an abstract sense, our experimental data are informative

*Submitted to the journal's Methods and Algorithms for Scientific Computing section December 5, 2011; accepted for publication (in revised form) May 15, 2012; published electronically July 3, 2012. This work was supported by the Air Force Office of Scientific Research MURI Program, program manager Dr. F. Fahroo, and Department of Energy Applied Mathematics Program grant DE-FG02-08ER25858, program manager A. Landsberg.
http://www.siam.org/journals/sisc/34-4/85776.html

†Department of Aeronautics & Astronautics, MIT, Cambridge, MA 02139 (celieber@mit.edu, kwillcox@mit.edu).

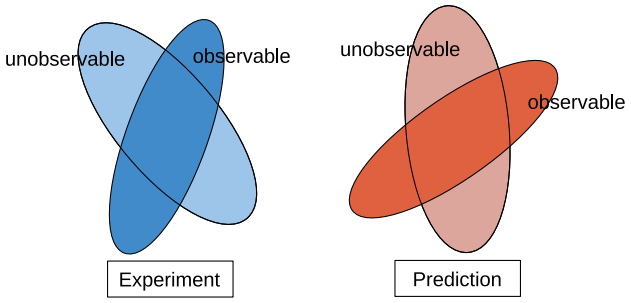

FIG. 1.1. *The two separate decompositions of the parameter space based on the experiment (left) and prediction (right).*

about certain modes in the parameter space and other modes in parameter space are required to accurately estimate the output quantities of interest. Our philosophy is to understand the relationship between these two sets of modes and to modify our approach to inference based on that information. The goal-oriented inference method involves identifying parameter modes that are both informed by experiment and also required for estimating output quantities of interest. In what follows, we refer to the output quantities of interest as predictions, although predictions need not be outputs of a system but instead could be, for example, the evaluation of the objective function in a design optimization problem. We call it the *inference-for-prediction* (IFP) method.

The two decompositions of parameter space based on the experimental process and the prediction process are shown notionally in Figure 1.1. We consider abstractly the decomposition into experimentally observable and experimentally unobservable modes on the left side of Figure 1.1. Experimentally observable modes are informed by the experimental data while experimentally unobservable modes are not. The analogous decomposition for prediction is shown on the right side of Figure 1.1. It is informative to explore the combinations of observable modes from the two processes. We explicitly identify here three types of parameter modes. Modes informed by experiment and required for prediction are targeted by the IFP method. Modes that are informed by experiment but not required for prediction represent inefficiencies in the experimental data acquisition. Finally, modes that are required for prediction but are uninformed by experiment lead to uncertainty in the prediction and may guide future experimentation.

There are many advantages to this way of thinking, including computational efficiency in the inference step; enabling deployment on lightweight, portable devices (e.g., smartphones and laptops) in the field; understanding of the effects of regularization and prior information on predictions; identification of inefficiencies in experimental data acquisition to focus efforts on data informative about modes required for predictions; and understanding of vulnerabilities in predictions.

Our approach exploits a compelling connection to balanced truncation model reduction [16]. In balanced truncation one obtains a projection-based reduced model that balances the controllability and observability of the underlying transformed system. The control-theoretic concepts underlying balanced truncation suggest a new form of subspace regularization exploiting the similarity in spatial patterns between permeability and pressure in subsurface identification and control problems [18]. In the goal-oriented inference context, we present an approach that leads to a balance

between the experiment and prediction observability associated with parameter modes. In the linear case, this approach leads to a form of subspace regularization that is dimensionally optimal in the sense that we infer the fewest number of parameters necessary to accurately estimate the predictions required to meet end goals.

This paper is organized as follows. In section 2 we provide background material on the control-theoretic concepts. We review the principles of and give an algorithm for balanced truncation model reduction. In section 3 we define experiment and prediction observability, we delineate the IFP algorithm for the truncated singular value decomposition inverse problem formulation, and we provide discussion on the numerical implementation and computational cost associated with the new method. The theoretical underpinnings of the method are provided in section 4, including a geometric interpretation and proof of the dimensional optimality of the critical subspace. In section 5 we extend the approach to Tikhonov-regularized and Gaussian statistical inverse problems. Section 6 discusses the application of the new method to a model problem in contaminant identification and presents numerical results that corroborate the theory. Conclusions are provided in section 7.

2. Background. In this section we provide background material in control-theoretic concepts and balanced truncation model reduction. Controllability and observability gramians are defined in section 2.1 and their properties discussed. In section 2.2 balanced truncation model reduction is described. This is background material in control theory and model reduction that we will lean on for the development of our goal-oriented inference approach.

Consider the time-invariant discrete-time linear system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad k = 0, 1, \ldots, \tag{2.1}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k, \quad k = 0, 1, \ldots, \tag{2.2}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state at time step k, $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times n_i}$, $\mathbf{u}_k \in \mathbb{R}^{n_i}$ is the input at time step k, $\mathbf{C} \in \mathbb{R}^{n_o\times n}$, and $\mathbf{y}_k \in \mathbb{R}^{n_o}$ is the output at time step k. The system has given initial condition $\mathbf{x}_0$. We assume that (2.1) is stable; i.e., the spectral radius $\rho(\mathbf{A}) < 1$.

2.1. Controllability and observability. Controllability and observability are two important properties of the system (2.1)–(2.2) [12]. The information contained within them is exploited in balanced truncation model reduction, as we describe in section 2.2.

A measure of the *controllability* $L_c(\mathbf{x})$ of a state $\mathbf{x}$ is the minimum input energy required to drive the system to zero when initialized at $\mathbf{x}_0 = \mathbf{x}$; i.e.,

$$L_c(\mathbf{x}) = \min_{\mathbf{u}_k\ \forall k} \sum_{k=0}^{\infty} \|\mathbf{u}_k\|^2 \quad \text{s.t. } \mathbf{x}_0 = \mathbf{x},\ \lim_{k\to\infty} \mathbf{x}_k = \mathbf{0}.$$

Let $\mathbf{P} = \sum_{k=0}^{\infty} \mathbf{A}^k\mathbf{B}\mathbf{B}^\top(\mathbf{A}^\top)^k \in \mathbb{R}^{n\times n}$ be the controllability gramian. The system (2.1)–(2.2) is *controllable* if $\mathbf{P}$ is full rank. Then we may write $L_c(\mathbf{x}) = \mathbf{x}^\top\mathbf{P}^{-1}\mathbf{x}$.

A measure of the *observability* $L_o(\mathbf{x})$ of a state $\mathbf{x}$ is the total output energy generated by the unforced ($\mathbf{u}_k = \mathbf{0}\ \forall k$) system initialized at $\mathbf{x}_0 = \mathbf{x}$; i.e.,

$$L_o(\mathbf{x}) = \sum_{k=0}^{\infty} \|\mathbf{y}_k\|^2 = \sum_{k=0}^{\infty} \|\mathbf{C}\mathbf{A}^k\mathbf{x}\|^2.$$

Let $\mathbf{Q} = \sum_{k=0}^{\infty} (\mathbf{A}^\top)^k \mathbf{C}^\top \mathbf{C} \mathbf{A}^k \in \mathbb{R}^{n \times n}$ be the observability gramian. Then the observability associated with a state $\mathbf{x}$ is $L_o(\mathbf{x}) = \mathbf{x}^\top \mathbf{Q} \mathbf{x}$. The system (2.1)–(2.2) is *observable* if $\mathbf{Q}$ is full rank.

The controllability and observability gramians are usually computed as solutions to the Stein equations

$$-\mathbf{P} + \mathbf{A}\mathbf{P}\mathbf{A}^\top = -\mathbf{B}\mathbf{B}^\top, \qquad -\mathbf{Q} + \mathbf{A}^\top \mathbf{Q} \mathbf{A} = -\mathbf{C}^\top \mathbf{C},$$

respectively.

2.2. Balanced truncation model reduction. A projection-based reduced model of the system (2.1)–(2.2) is given by

$$\begin{aligned} \hat{\mathbf{x}}_{k+1} &= \hat{\mathbf{A}} \hat{\mathbf{x}}_k + \hat{\mathbf{B}} \mathbf{u}_k, \quad && k = 0, 1, \ldots, \\ \hat{\mathbf{y}}_k &= \hat{\mathbf{C}} \hat{\mathbf{x}}_k, && k = 0, 1, \ldots, \end{aligned}$$

where $\hat{\mathbf{A}} = \mathbf{U}^\top \mathbf{A} \mathbf{V} \in \mathbb{R}^{m \times m}$, $\hat{\mathbf{B}} = \mathbf{U}^\top \mathbf{B} \in \mathbb{R}^{m \times n_i}$, $\hat{\mathbf{C}} = \mathbf{C}\mathbf{V} \in \mathbb{R}^{n_o \times m}$, $\hat{\mathbf{x}}_k \in \mathbb{R}^m$ is the reduced state at time step k, and $\hat{\mathbf{y}}_k \in \mathbb{R}^{n_o}$ is the output of the reduced model at time step k. The left basis $\mathbf{U} \in \mathbb{R}^{n \times m}$ and right basis $\mathbf{V} \in \mathbb{R}^{n \times m}$ span subspaces of dimension $m \ll n$ and satisfy $\mathbf{U}^\top \mathbf{V} = \mathbf{I}$.

Balanced truncation model reduction is one method for selecting the left and right bases [16]. Conceptually, balanced truncation can be understood in two distinct steps. The first step is a similarity transformation to describe the state space of the system (2.1)–(2.2) in a way that balances each coordinate direction's combined measure of controllability and observability. In particular, the controllability and observability gramians of the transformed system are diagonal and equal. The second step is truncation, retaining only some of the states in the transformed model and discarding the rest. For example, any coordinate directions having zero combined measure of controllability and observability can be truncated without affecting the system's input-output behavior.

We first identify the similarity transformation. The balanced truncation left and right bases can be obtained using general matrix factors of the controllability and observability gramians [4, 15]. For purposes of exposition, we will assume the factors are square. Let $\mathbf{P} = \mathbf{S}\mathbf{S}^\top$ and $\mathbf{Q} = \mathbf{R}\mathbf{R}^\top$. Consider the similarity transformation defined by $\mathbf{T}^{-1} = \mathbf{\Sigma}^{-1/2} \mathbf{M}^\top \mathbf{R}^\top$ and $\mathbf{T} = \mathbf{S}\mathbf{N}\mathbf{\Sigma}^{-1/2}$, where $\mathbf{M}\mathbf{\Sigma}\mathbf{N}^\top$ is the singular value decomposition (SVD) of $\mathbf{R}^\top \mathbf{S}$. The transformed system

$$\begin{aligned} \tilde{\mathbf{x}}_{k+1} &= \mathbf{T}\mathbf{A}\mathbf{T}^{-1} \tilde{\mathbf{x}}_k + \mathbf{T}\mathbf{B}\mathbf{u}_k, \quad && k = 0, 1, \ldots, \\ \mathbf{y}_k &= \mathbf{C}\mathbf{T}^{-1} \tilde{\mathbf{x}}_k, && k = 0, 1, \ldots, \end{aligned}$$

has diagonal and equal controllability and observability gramians,

$$\mathbf{T}\mathbf{P}\mathbf{T}^\top = \mathbf{T}^{-T} \mathbf{Q} \mathbf{T}^{-1} = \mathbf{\Sigma}.$$

The coefficients $\sigma_1, \sigma_2, \ldots, \sigma_n$ on the diagonal of $\mathbf{\Sigma}$ are known as the Hankel singular values, which represent a joint measure of the controllability and observability of the modes in the transformed system. The second step is the truncation of the transformed state $\tilde{\mathbf{x}} \in \mathbb{R}^n$ to $\hat{\mathbf{x}} \in \mathbb{R}^m$. The truncation eliminates the least controllable and observable modes of the system based on the Hankel singular values.

In practice the balanced truncation reduced model can be obtained by directly identifying left and right bases $\mathbf{U}$ and $\mathbf{V}$ by Algorithm 1. Although balanced truncation is not optimal, there exist bounds on the $\mathcal{H}_\infty$-norm of the error system related to

Algorithm 1 Balanced truncation model reduction left and right bases

1: Compute the first m normalized eigenvectors $\boldsymbol{\psi}_i$ of $\mathbf{S}^\top\mathbf{QS}$ with corresponding eigenvalues σ_i^2; i.e.,

$$\mathbf{S}^\top\mathbf{QS}\boldsymbol{\psi}_i = \sigma_i^2\boldsymbol{\psi}_i, \quad \|\boldsymbol{\psi}\|^2 = 1,\ \boldsymbol{\psi}_i^\top\boldsymbol{\psi}_j = \delta_{ij},\ i = 1, 2, \ldots, m.$$

2: Compute the first m left eigenvectors $\boldsymbol{\phi}_i$ of $\mathbf{R}^\top\mathbf{PR}$ also having eigenvalues σ_i^2; i.e.,

$$\boldsymbol{\phi}_i = \sigma_i^{-1}\boldsymbol{\psi}_i^\top\mathbf{S}^\top\mathbf{R}, \quad i = 1, 2, \ldots, m.$$

3: Then define the left and right bases

$$\mathbf{U} = \mathbf{R}\left[\begin{array}{ccc} \sigma_1^{-1/2}\boldsymbol{\phi}_1^\top & \cdots & \sigma_m^{-1/2}\boldsymbol{\phi}_m^\top \end{array}\right], \qquad \mathbf{V} = \mathbf{S}\left[\begin{array}{ccc} \sigma_1^{-1/2}\boldsymbol{\psi}_1 & \cdots & \sigma_m^{-1/2}\boldsymbol{\psi}_m \end{array}\right].$$

the truncated Hankel singular values [16]. There exist algorithms to obtain a balanced reduced model via approximate computation of the gramians for large-scale systems [15, 10].

3. IFP method in the linear setting. Let $\boldsymbol{\mu} \in \mathbb{R}^q$ be an unknown parameter defining a system of interest. We assume that q is large; i.e., there are many more parameters to infer than experiments we can afford to perform or predictions we wish to make. Let $\mathbf{O}_e \in \mathbb{R}^{r\times q}$ be the linear observation operator representing the (usually indirect) measurement process mapping the parameter space to the space of experimental observables of dimension $r < q$. We write experimental outputs $\mathbf{y}_e = \mathbf{O}_e\boldsymbol{\mu}$. In many instances it will be appropriate to model sensor error, in which case we obtain $\mathbf{y}_d = e(\mathbf{y}_e, \epsilon)$ for some error model e and a measure of error ϵ. Our formulation will utilize the experimental output matrix $\mathbf{O}_e$, but our algorithms will be data-independent and therefore admit any form of the error model e. For many applications of interest, $\mathbf{O}_e$ will be the composition of a PDE operator and an observation operator. Likewise, the prediction operator $\mathbf{O}_p \in \mathbb{R}^{s\times q}$ is analogous to $\mathbf{O}_e$ but instead measures prediction output quantities of interest in the space of dimension $s < r$.[1] We write prediction $\mathbf{y}_p = \mathbf{O}_p\boldsymbol{\mu}$.[2]

In section 3.1 we define experiment and prediction observability and the associated gramians. Section 3.2 states our assumptions, gives important definitions, and establishes the IFP property. We conclude with an algorithm for obtaining a basis for efficient inversion. Finally, in section 3.3 we discuss the numerical implementation of the algorithm and analyze the computational complexity.

3.1. Experiment and prediction observability. Experiment and prediction observability extend the concept of observability of linear systems described in section 2.1 to the goal-oriented inference setting.

[1]The assumption $s < r$ is not necessary, but it does simplify the exposition. For almost all of the following, one could generalize by replacing s with $\min(r, s)$. It is essential only that $r \ll q$ and $s \ll q$.

[2]Note that the physics underlying the PDE operators (if present) in $\mathbf{O}_e$ and $\mathbf{O}_p$ need not be the same. Typically, experimental conditions will differ from operational conditions, and our method admits that naturally.

A measure of the *experiment observability* of a parameter $\boldsymbol{\mu}$ is given by the experimental output energy associated with it. We define $L_e(\boldsymbol{\mu}) = \|\mathbf{y}_e\|^2 = \|\mathbf{O}_e\boldsymbol{\mu}\|^2$. Consequently, the experiment observability gramian $\mathbf{H}_e = \mathbf{O}_e^\top \mathbf{O}_e$ can be defined since $L_e(\boldsymbol{\mu}) = \boldsymbol{\mu}^\top \mathbf{H}_e \boldsymbol{\mu}$. Since the experiment observability gramian is symmetric and positive semidefinite, it admits the decomposition $\mathbf{H}_e = \mathbf{V}_e \mathbf{L}_e \mathbf{V}_e^\top$, where $\mathbf{V}_e \in \mathbb{R}^{q\times r}$ is orthogonal and $\mathbf{L}_e \in \mathbb{R}^{r\times r}$ is diagonal with positive entries. The columns of $\mathbf{V}_e$ are eigenvectors of $\mathbf{H}_e$ with corresponding eigenvalues on the diagonal of $\mathbf{L}_e$. When we solve the inverse problem, the pseudoinverse $\mathbf{H}_e^\dagger = \mathbf{V}_e \mathbf{L}_e^{-1} \mathbf{V}_e^\top$ and its matrix factor $\mathbf{G}_e = \mathbf{V}_e \mathbf{L}_e^{-1/2}$ will play an important role.

Let $\mathbf{V}_{e^\perp}$ be an orthogonal basis whose range is the orthogonal complement to the range of $\mathbf{V}_e$. Then any parameter $\boldsymbol{\mu}$ can be decomposed as $\boldsymbol{\mu} = \mathbf{V}_e \mathbf{V}_e^\top \boldsymbol{\mu} + \mathbf{V}_{e^\perp} \mathbf{V}_{e^\perp}^\top \boldsymbol{\mu}$. The first component influences the data $\mathbf{y}_e$ observed for parameter $\boldsymbol{\mu}$ while the second component produces exactly zero experimental output. When we utilize data $\mathbf{y}_e$ to infer the parameter $\boldsymbol{\mu}$, the second component is determined by injecting outside information through regularization and prior distribution in the deterministic and statistical approaches, respectively.

A measure of the *prediction observability* of a parameter $\boldsymbol{\mu}$ is given by the prediction output energy associated with it. Define $L_p(\boldsymbol{\mu}) = \|\mathbf{y}_p\|^2 = \|\mathbf{O}_p\boldsymbol{\mu}\|^2$. The prediction observability gramian $\mathbf{H}_p = \mathbf{O}_p^\top \mathbf{O}_p$ then follows since $L_p(\boldsymbol{\mu}) = \boldsymbol{\mu}^\top \mathbf{H}_p \boldsymbol{\mu}$. It is also symmetric and positive semidefinite and therefore has a decomposition $\mathbf{H}_p = \mathbf{V}_p \mathbf{L}_p \mathbf{V}_p^\top$ analogous to $\mathbf{H}_e$ above. Similarly, any parameter $\boldsymbol{\mu}$ can be decomposed as $\boldsymbol{\mu} = \mathbf{V}_p \mathbf{V}_p^\top \boldsymbol{\mu} + \mathbf{V}_{p^\perp} \mathbf{V}_{p^\perp}^\top \boldsymbol{\mu}$. The first component will pass through to predictions $\mathbf{y}_p$ and is therefore necessary to accurately estimate; on the other hand, the second component is in the kernel of $\mathbf{O}_p$ and will therefore not contribute to $\mathbf{y}_p$. Thus, the second component need not be accurately estimated, or even estimated at all, to achieve accurate estimates of $\mathbf{y}_p$.

3.2. IFP algorithm. The IFP method will lead to a well-chosen basis for inference spanning the low-dimensional subspace of the parameter space that will result in replication of the predictions obtained by a traditional approach to the linear inverse problem. In this section, we will treat the truncated SVD approach to the linear inverse problem, and we will extend the method to the Tikhonov-regularized inverse problem and Gaussian statistical inverse problem in section 5.

We begin with a truncated singular value decomposition (TSVD) approach to the linear inverse problem. The inverse problem uses data $\mathbf{y}_d$ and knowledge of $\mathbf{O}_e$ to estimate the unknown parameter $\boldsymbol{\mu}$. In many applications, however, the inverse problem is ill-posed due to the vast null space of $\mathbf{O}_e$. This difficulty is usually overcome by regularization. In this section, we consider a form of subspace regularization by seeking a solution only in the row space of $\mathbf{O}_e$. In section 5.1, we will consider regularization in the form of a penalty in the objective function.

Let $\mathbf{PSV}^\top = \mathbf{O}_e$ be the SVD with $\mathbf{P} \in \mathbb{R}^{r\times r}$, $\mathbf{S} \in \mathbb{R}^{r\times q}$, and $\mathbf{V} \in \mathbb{R}^{q\times q}$. Let $\mathbf{V}_e \in \mathbb{R}^{q\times r}$ and $\mathbf{V}_e^\perp \in \mathbb{R}^{q\times(q-r)}$ span the row space and null space of $\mathbf{O}_e$, respectively, such that $\mathbf{V} = [\mathbf{V}_e, \mathbf{V}_e^\perp]$. Let $\mathcal{V}_e \subset \mathbb{R}^q$ be the r-dimensional subspace spanned by the columns of $\mathbf{V}_e$. The TSVD approach searches for $\boldsymbol{\mu} \in \mathcal{V}_e$ that reproduces the observed data with minimal error in ℓ_2-norm. That is,

$$\boldsymbol{\mu}^{\mathrm{TSVD}} = \arg\min_{\boldsymbol{\mu}\in\mathcal{V}_e} \frac{1}{2}\|\mathbf{y}_d - \mathbf{O}_e\boldsymbol{\mu}\|_2^2. \tag{3.1}$$

The first-order optimality condition for (3.1) is obtained by imposing the constraint and setting the first derivative of the objective function to zero; i.e.,

$$\mathbf{V}_e^\top \mathbf{O}_e^\top \mathbf{O}_e \mathbf{V}_e \mathbf{a} = \mathbf{V}_e^\top \mathbf{O}_e^\top \mathbf{y}_d, \tag{3.2}$$

where $\mathbf{a} \in \mathbb{R}^r$ is the vector of modal coefficients in the expansion $\mathbf{x}^{\mathrm{TSVD}} = \mathbf{V}_e \mathbf{a}$.

Substituting the reduced eigendecomposition of $\mathbf{O}_e^T \mathbf{O}_e = \mathbf{H}_e$ and noting that $\mathbf{V}_e^T \mathbf{V}_e = \mathbf{I}$, (3.2) reduces to $\mathbf{a} = \mathbf{L}_e^{-1} \mathbf{V}_e^\top \mathbf{O}_e^\top \mathbf{y}_d$. Therefore, the TSVD parameter estimate is given by

$$\boldsymbol{\mu}^{\mathrm{TSVD}} = \mathbf{V}_e \mathbf{L}_e^{-1} \mathbf{V}_e^\top \mathbf{O}_e^\top \mathbf{y}_d.$$

In the traditional two-step approach, this estimate of $\boldsymbol{\mu}$ would then be utilized in simulation to predict output quantities of interest

$$\mathbf{y}_p^{\mathrm{TSVD}} = \mathbf{O}_p \boldsymbol{\mu}^{\mathrm{TSVD}}.$$

It is precisely these prediction outputs that the IFP method will reproduce.

The following derivation of the IFP basis resembles, and in fact was inspired by, balanced truncation model reduction [16]. It is not necessary, however, for there to exist an underlying state space system to use the IFP method; it is sufficient to have models only for the experiment and prediction operators $\mathbf{O}_e$ and $\mathbf{O}_p$.

Before stating the basis generation algorithm, we first define the key property of the IFP method.

PROPERTY 1. *A parameter estimate $\boldsymbol{\mu}^*$ has the IFP property if it results in prediction equal to that of the prediction resulting from the TSVD parameter estimate; i.e., $\mathbf{y}_p(\boldsymbol{\mu}^*) = \mathbf{O}_p \boldsymbol{\mu}^* = \mathbf{y}_p^{\mathrm{TSVD}}$.*

Our goal is to find an s-dimensional subspace $\mathcal{W} \subset \mathbb{R}^q$ such that the IFP solution

$$\boldsymbol{\mu}^{\mathrm{IFP}} = \arg\min_{\boldsymbol{\mu} \in \mathcal{W}} \frac{1}{2} \|\mathbf{y}_d - \mathbf{O}_e \boldsymbol{\mu}\|_2^2 \tag{3.3}$$

has Property 1. For now, we assume that such a subspace exists and is spanned by the columns of the matrix $\mathbf{W} \in \mathbb{R}^{q \times s}$. Then, we may write

$$\boldsymbol{\mu}^{\mathrm{IFP}} = \mathbf{W} (\mathbf{W}^\top \mathbf{H}_e \mathbf{W})^{-1} \mathbf{W}^\top \mathbf{O}_e^\top \mathbf{y}_d. \tag{3.4}$$

We will also need an important assumption regarding the geometry of the experiment and prediction observable subspaces.

Assumption 1. We will assume throughout that $\mathrm{rank}(\mathbf{V}_p^\top \mathbf{V}_e) = s$.

Assumption 1 is not a limiting assumption; however, if $\mathrm{rank}(\mathbf{V}_p^\top \mathbf{V}_e) < s$, this implies that there are some linear combinations of experiments orthogonal to predictions; i.e., there should be a suitable reformulation such that $\mathbf{V}_p^\top \mathbf{V}_e$ is full rank.

We now define the IFP subspace.

DEFINITION 3.1. *An IFP subspace is an s-dimensional subspace $\mathcal{W}$ such that the solution $\boldsymbol{\mu}^{\mathrm{IFP}}$ to (3.3) has Property 1 independent of the data $\mathbf{y}_d$.*

The definition of an IFP basis follows naturally.

DEFINITION 3.2. *Any basis $\mathbf{W} \in \mathbb{R}^{q \times s}$ is an IFP basis if its columns span an IFP subspace $\mathcal{W}$.*

We now present an algorithm for obtaining an IFP basis $\mathbf{W}$ (we prove it in section 4.1) that simultaneously diagonalizes the projected experiment and prediction observability gramians. Although the simultaneous diagonalization is not necessary

Algorithm 2 IFP Basis Generation for TSVD approach

1: Define $\mathbf{G}_e = \mathbf{V}_e \mathbf{L}_e^{-1/2}$.
2: Compute the reduced eigendecomposition $\boldsymbol{\Psi}\boldsymbol{\Sigma}^2\boldsymbol{\Psi}^\top$ of $\mathbf{G}_e^\top \mathbf{O}_p^\top \mathbf{O}_p \mathbf{G}_e$.
3: Define $\mathbf{W} = \mathbf{G}_e \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2}$.

to replicate the TSVD predictions (any basis for $\mathcal{W}$ will do), it does provide a measure by which further reduction can be performed if desired.

The singular values on the diagonal of $\boldsymbol{\Sigma}$ are analogous to the Hankel singular values of balanced truncation. They represent a joint measure of the experiment and prediction observability. While Algorithm 2 identifies a low-dimensional basis for the IFP subspace, it is possible to truncate further. In this basis, eliminating columns of $\mathbf{W}$ from the right is analogous to removing the least experiment and prediction observable modes according to the joint measure reflected by the singular values.

3.3. IFP implementation and computational complexity. Algorithm 2 has two major computational steps. In step 1 we require the eigendecomposition of the experiment observability gramian $\mathbf{H}_e = \mathbf{O}_e^\top \mathbf{O}_e$, which has rank r. Step 2 involves an eigendecomposition of a matrix of rank s.

For step 1, efficient implementation should include a code to perform the matrix-vector product $\mathbf{H}_e\mathbf{v}$ as efficiently as possible, both in terms of storage and operation cost. Oftentimes, particularly when the governing equations are given by PDEs, this implies a matrix-free implementation. For problems of interest, the action $\mathbf{H}_e\mathbf{v}$ is one forward and one adjoint solution starting from initial condition $\mathbf{v}$ [5]. Let $\alpha(n)$ be the cost of one time step of the PDE (depending on the mesh DOFs n) and K_e be the number of time steps until the last data are collected. Then the forward and adjoint solutions cost $2\alpha(n)K_e$. Since the experiment observability gramian has rank r, an iterative eigenvalue solver like Lanczos iteration will require at least $2r\alpha(n)K_e$ to obtain the eigendecomposition, but it may not exceed this cost by much if the eigenvalues are well separated [8]. Note here that r is independent of n and that $\alpha(n) \sim n$ if appropriate preconditioners are used. If all operations are performed iteratively, the storage requirements should not exceed a small constant number of parameter vectors and therefore scale linearly with q. However, we do assume here that we store the r eigenvalues and eigenvectors for a total cost of $(q+1)r = qr + r$.

The computation in step 2 contains two parts. First, the implementation should include a matrix-free code for computing $\mathbf{H}_p\mathbf{v} = \mathbf{O}_p^\top \mathbf{O}_p\mathbf{v}$. Second, a rank s eigendecomposition must be computed. The code for the action of the prediction observability gramian on a vector $\mathbf{H}_p\mathbf{v}$ will also manifest in forward and adjoint solves, although the final time of the simulation $K_p > K_e$. Thus, the cost is approximately $2\alpha(n)K_p$ for each matrix-vector product. This computation is a part of the code that computes the matrix-vector product $\mathbf{G}_e^\top \mathbf{O}_p^\top \mathbf{O}_p \mathbf{G}_e \mathbf{v} = \mathbf{L}_e^{-1/2} \mathbf{V}_e^\top \mathbf{H}_p \mathbf{V}_e \mathbf{L}_e \mathbf{v}$ utilized by the eigenvalue solver. Each such product requires in order (from right to left) r scalar products, nr scalar products, r q-vector sums, $2\alpha(n)K_p$ for the action of $\mathbf{H}_p$, r q-vector inner products, and finally another r scalar products. That is a total cost of $2(4q+2)r\alpha(n)K_p$ for each matrix-vector product $\mathbf{G}_e^\top \mathbf{O}_p^\top \mathbf{O}_p \mathbf{G}_e \mathbf{v}$. Since this matrix has rank s, we can expect approximately s such iterations giving a total cost for step 2 of approximately $2(4q+2)rs\alpha(n)K_p$. There is negligible additional storage required at this step since the storage of the eigenvectors $\mathbf{V}_e$ will dominate. We

store the resulting s eigenvalues and r-dimensional eigenvectors for a storage cost of $(r+1)s$.

If we combine the cost of the first two steps and then account for the final matrix multiplication to obtain $\mathbf{W}$, we have a total operation cost of approximately $2r\alpha(n)K_e + 2(4q+2)rs\alpha(n)K_p + qrs^2$ and total storage cost of approximately $(q+1)r+(r+1)s+qs$. While the IFP method may be more computationally expensive than traditional inference procedures for the solution of one-off inverse problems, the benefits of the IFP method are three-fold. First, if data are collected repeatedly under the same experimental observation operator, then the cost of determining the IFP basis can be amortized over the experiments. Second, the IFP basis encodes important information about the process relating inference and prediction, in particular through an analysis of the range of the IFP basis as it compares to the ranges of the gramians of the experiment and prediction processes. The study of this relationship could play a role in determining the effects of regularization and in designing future experiments. Last, and perhaps most importantly, since the IFP method has data-independent theory, it is feasible to move all of this computation offline and utilize only the resulting basis in an online deployment. This offline-online decomposition of cost can make the IFP approach efficient for inverse problem solutions on lightweight, portable devices in the field.

4. IFP linear theory. In this section we develop the relevant theory for the IFP method in linear inverse problems. We first prove that the basis generated by Algorithm 2 leads to Property 1. We give a geometric interpretation of the IFP method in this setting. Finally, the section is concluded with a proof of dimensional optimality, showing that there is no subspace of dimension less than s that makes the solution of (3.3) have Property 1.

4.1. Prediction exactness. We first show that the basis generated by Algorithm 2 defines an IFP subspace, i.e., that the solution to (3.3) has Property 1.

THEOREM 4.1. *Algorithm* 2 *leads to a basis* $\mathbf{W}$ *whose columns span an IFP subspace. Therefore, the solution* $\boldsymbol{\mu}^{\mathrm{IFP}}$ *to* (3.3) *with* $\mathcal{W} = \mathrm{range}(\mathbf{W})$ *has Property* 1.

Proof. Let $\mathbf{U} = \mathbf{O}_p^\top \mathbf{O}_p \mathbf{G}_e \boldsymbol{\Psi} \boldsymbol{\Sigma}^{-3/2}$ and note that the columns of $\mathbf{U}$ are a basis for the row space of $\mathbf{O}_p$. Thus, any two parameter estimates $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ satisfying $\mathbf{U}^\top(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$ will have the same prediction $\mathbf{O}_p\boldsymbol{\mu}_1 = \mathbf{O}_p\boldsymbol{\mu}_2$. We will now show $\mathbf{U}^\top(\boldsymbol{\mu}^{\mathrm{IFP}} - \boldsymbol{\mu}^{\mathrm{TSVD}}) = \mathbf{0}$. Substituting the optimality conditions for the TSVD and IFP optimization problems, we find

$$\mathbf{U}^\top(\boldsymbol{\mu}^{\mathrm{IFP}} - \boldsymbol{\mu}^{\mathrm{TSVD}}) = \mathbf{U}^\top(\mathbf{W}(\mathbf{W}^\top\mathbf{H}_e\mathbf{W})^{-1}\mathbf{W}^\top\mathbf{O}_e^\top - \mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top\mathbf{O}_e^\top)\mathbf{y}_d. \tag{4.1}$$

By construction, we have

$$\mathbf{U}^\top\mathbf{W} = \boldsymbol{\Sigma}^{-3/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_p^\top\mathbf{O}_p\mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-3/2}\boldsymbol{\Sigma}^2\boldsymbol{\Sigma}^{-1/2} = \mathbf{I},$$

where we have used the orthonormality of $\boldsymbol{\Psi}$ and the eigendecomposition from Algorithm 2. Furthermore, the basis $\mathbf{W}$ satisfies the relation

$$\begin{aligned}\mathbf{W}^\top\mathbf{H}_e\mathbf{W} &= \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{V}_e\mathbf{L}_e\mathbf{V}_e^\top\mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2},\\ &= \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Psi}^\top\mathbf{L}_e^{-1/2}\mathbf{V}_e^\top\mathbf{V}_e\mathbf{L}_e\mathbf{V}_e^\top\mathbf{V}_e\mathbf{L}_e^{-1/2}\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2}\\ &= \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Psi}^\top\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1}.\end{aligned}$$

Using these facts, (4.1) reduces to

$$\mathbf{U}^\top(\boldsymbol{\mu}^{\mathrm{IFP}} - \boldsymbol{\mu}^{\mathrm{TSVD}}) = (\boldsymbol{\Sigma}\mathbf{W}^\top\mathbf{O}_e^\top - \mathbf{U}^\top\mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top\mathbf{O}_e^\top)\mathbf{y}_d.$$

We now have $\boldsymbol{\Sigma}\mathbf{W}^\top\mathbf{O}_e^\top = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_e^\top$ and

$$\mathbf{U}^\top\mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top\mathbf{O}_e^\top = \boldsymbol{\Sigma}^{-3/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_p^\top\mathbf{O}_p\mathbf{G}_e\mathbf{G}_e^\top\mathbf{O}_e^\top = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_e^\top.$$

This demonstrates that $\mathbf{U}^\top(\boldsymbol{\mu}^{\mathrm{IFP}}-\boldsymbol{\mu}^{\mathrm{TSVD}}) = (\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_e^\top-\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_e^\top)\mathbf{y}_d = \mathbf{0}$ and therefore proves that $\mathbf{y}_p^{\mathrm{IFP}} = \mathbf{y}_p^{\mathrm{TSVD}}$ for all data $\mathbf{y}_d$. ∎

Note that Theorem 4.1 holds irrespective of data $\mathbf{y}_d$. The IFP method inherits the sensitivity to noise of the TSVD approach.

4.2. Geometric interpretation. Of particular interest is the geometry of the approach. The solution $\boldsymbol{\mu}^{\mathrm{IFP}}$ is obtained as the oblique projection of $\boldsymbol{\mu}^{\mathrm{TSVD}}$ based on the projector $\boldsymbol{\Pi} = \mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{H}_e$. That is to say, $\boldsymbol{\mu}^{\mathrm{IFP}} = \boldsymbol{\Pi}\boldsymbol{\mu}^{\mathrm{TSVD}}$ independent of the data. We show first that $\boldsymbol{\Pi}$ is an oblique projector.

THEOREM 4.2. *The matrix* $\boldsymbol{\Pi} = \mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{H}_e$ *is an oblique projector.*

Proof. We first show that $\boldsymbol{\Pi}$ is a projector, and then we establish that its range and null space are not orthogonal complements. We have

$$\boldsymbol{\Pi}^2 = \mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{H}_e\mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{H}_e = \mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{H}_e = \mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{H}_e = \boldsymbol{\Pi}.$$

Since $\boldsymbol{\Pi}^2 = \boldsymbol{\Pi}$, $\boldsymbol{\Pi}$ is a projector. An orthogonal projector has orthogonal range and null spaces. Any projector that is not an orthogonal projector is an oblique projector. Therefore, it suffices for us to obtain a vector $\mathbf{v} \in \mathbb{R}^q$ such that $(\boldsymbol{\Pi}\mathbf{v})^\top(\mathbf{v} - \boldsymbol{\Pi}\mathbf{v}) \neq 0$ to show that $\boldsymbol{\Pi}$ is an oblique projector since $\boldsymbol{\Pi}\mathbf{v} \in \mathrm{range}(\boldsymbol{\Pi})$ and $\mathbf{v} - \boldsymbol{\Pi}\mathbf{v} \in \mathrm{null}(\boldsymbol{\Pi})$.

We assume that $\mathbf{L}_e \neq \mathbf{I}$ in general; if it is, then $\boldsymbol{\Pi}$ is an orthogonal projector. Let $\mathbf{z} \in \mathbb{R}^r$ be chosen such that $\boldsymbol{\Psi}^\top\mathbf{z} = \mathbf{0}$ but that $\boldsymbol{\Psi}^\top\mathbf{L}_e\mathbf{z} \neq \mathbf{0}$. Then define $\mathbf{v} = \mathbf{V}_e\mathbf{L}_e^{1/2}\mathbf{z}$. Then if we write out the expression above, we find

$$(\boldsymbol{\Pi}\mathbf{v})^\top(\mathbf{v} - \boldsymbol{\Pi}\mathbf{v}) = \mathbf{v}^\top\mathbf{V}_e\boldsymbol{\Lambda}\mathbf{V}_e^\top\mathbf{v} - \mathbf{v}^\top\mathbf{V}_e\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\mathbf{V}_e^\top\mathbf{v}, \tag{4.2}$$

where $\boldsymbol{\Lambda} = \mathbf{L}_e^{1/2}\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{L}_e^{-1/2}$. Based on our choice of $\mathbf{v}$ above, we find that the first term on the right-hand side vanishes; i.e.,

$$\mathbf{v}^\top\mathbf{V}_e\boldsymbol{\Lambda}\mathbf{V}_e^\top\mathbf{v} = \mathbf{z}^\top\mathbf{L}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{z} = \mathbf{0}.$$

The second term on the right-hand side of (4.2) can be rewritten as $\|\boldsymbol{\Lambda}^\top\mathbf{V}_e^\top\mathbf{v}\|^2 = \|\mathbf{L}_e^{-1/2}\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{L}_e\mathbf{z}\|^2 \geq 0$. Since we chose $\mathbf{z}$ such that $\boldsymbol{\Psi}^\top\mathbf{L}_e\mathbf{z} \neq \mathbf{0}$, the second term is positive. This implies that we have found a $\mathbf{v}$ such that $(\boldsymbol{\Pi}\mathbf{v})^\top(\mathbf{v} - \boldsymbol{\Pi}\mathbf{v}) \neq 0$, and therefore $\boldsymbol{\Pi}$ is an oblique projector. ∎

THEOREM 4.3. *The parameter estimate* $\boldsymbol{\mu}^{\mathrm{IFP}}$ *obtained using the IFP method is the oblique projection under* $\boldsymbol{\Pi}$ *of the TSVD solution* $\boldsymbol{\mu}^{\mathrm{TSVD}}$.

Proof. Using the basis $\mathbf{W}$ obtained by Algorithm 2 the IFP solution is computed as

$$\boldsymbol{\mu}^{\mathrm{IFP}} = \mathbf{W}(\mathbf{W}^\top\mathbf{H}_e\mathbf{W})^{-1}\mathbf{W}^\top\mathbf{O}_e^\top\mathbf{y}_d.$$

Since the TSVD solution obtained always reproduces the data exactly (even under noisy data conditions), we can write $\mathbf{y}_d = \mathbf{O}_e\boldsymbol{\mu}^{\mathrm{TSVD}}$. Therefore,

$$\boldsymbol{\mu}^{\mathrm{IFP}} = \mathbf{W}(\mathbf{W}^\top\mathbf{H}_e\mathbf{W})^{-1}\mathbf{W}^\top\mathbf{H}_e\boldsymbol{\mu}^{\mathrm{TSVD}}.$$

Recalling that $\mathbf{W}^\top\mathbf{H}_e\mathbf{W} = \boldsymbol{\Sigma}^{-1}$ and that $\mathbf{W} = \mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2}$, we have

$$\boldsymbol{\mu}^{\mathrm{IFP}} = \mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{H}_e\boldsymbol{\mu}^{\mathrm{TSVD}} = \boldsymbol{\Pi}\boldsymbol{\mu}^{\mathrm{TSVD}}.$$ ∎

4.3. Dimensional optimality of the IFP subspace. It is natural to ask whether the IFP subspace of dimension s is the subspace of minimum dimension such that the solution of (3.3) has Property 1. We now show that there does not exist an $\tilde{s}$-dimensional subspace $\tilde{\mathcal{W}}$ for $\tilde{s} < s$ such that the solution of (3.3) has Property 1.

THEOREM 4.4. *The IFP subspace $\mathcal{W}$ is the subspace of minimum dimension such that the solution to* (3.3) *has Property* 1.

Proof. In view of Assumption 1, the predictable component of the TSVD solution is obtained from the data by the operation

$$\boldsymbol{\mu}_p^{\mathrm{TSVD}} = \mathbf{V}_p\mathbf{V}_p^\top\boldsymbol{\mu}^{\mathrm{TSVD}} = \mathbf{V}_p\mathbf{V}_p^\top\mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top\mathbf{O}_e^\top\mathbf{y}_d,$$

where the matrix $\mathbf{V}_p\mathbf{V}_p^\top\mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top\mathbf{O}_e^\top$ transforming data $\mathbf{y}_d$ to $\boldsymbol{\mu}_p^{\mathrm{TSVD}}$ has rank s. Let $\tilde{s} < s$ and $\tilde{\mathbf{W}} \in \mathbb{R}^{q\times\tilde{s}}$ be a basis for any $\tilde{s}$-dimensional subspace $\tilde{\mathcal{W}}$. Based on the IFP formulation the matrix from data $\mathbf{y}_d$ to predictable component of the IFP estimate $\boldsymbol{\mu}_p^{\mathrm{IFP}} = \mathbf{V}_p\mathbf{V}_p^\top\boldsymbol{\mu}^{\mathrm{IFP}}$ is $\mathbf{V}_p\mathbf{V}_p^\top\tilde{\mathbf{W}}(\tilde{\mathbf{W}}^\top\mathbf{H}_e\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{W}}^\top\mathbf{O}_e^\top$. In order for $\mathbf{x}_p^{\mathrm{IFP}} = \mathbf{x}_p^{\mathrm{TSVD}}$ for arbitrary $\mathbf{y}_d$ it must be the case that $\mathbf{V}_p\mathbf{V}_p^\top\mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top\mathbf{O}_e^\top = \mathbf{V}_p\mathbf{V}_p^\top\tilde{\mathbf{W}}(\tilde{\mathbf{W}}^\top\mathbf{H}_e\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{W}}^\top\mathbf{O}_e^\top$. However, we know that the matrix on the left has rank s and the matrix on the right has rank $\tilde{s} \neq s$, establishing a contradiction. Therefore, a basis permitting Property 1 must have dimension at least s. Thus, an IFP subspace, which has dimension s, is dimensionally optimal. ∎

5. Extensions of the IFP method. We extend the IFP method to Tikhonov-regularized inverse problems and Gaussian statistical inverse problems in this section. It is shown that only a small modification to the IFP method above is necessary to apply the goal-oriented approach to these cases.

5.1. Tikhonov-regularized inverse problem. Another method for regularizing ill-posed inverse problems is adding a penalty term to the objective function [9]. The idea is to select the parameter that most closely matches the experimental data and is minimum under some prescribed norm. The main effect is a modification of the experiment observability gramian in the algorithm.

A Tikhonov-regularized inverse problem [9] has the form

$$\boldsymbol{\mu}^{\mathrm{TR}} = \arg\min_{\boldsymbol{\mu}\in\mathbb{R}^q} \frac{1}{2}\|\mathbf{y}_d - \mathbf{O}_e\boldsymbol{\mu}\|_2^2 + \frac{1}{2}\|\mathbf{R}\boldsymbol{\mu}\|_2^2, \tag{5.1}$$

where we assume the regularization parameter weighting the two terms has been incorporated into the regularization matrix $\mathbf{R}$. For these formulations, the experiment observability gramian becomes $\mathbf{O}_e^\top\mathbf{O}_e + \mathbf{R}^\top\mathbf{R}$, where $\mathbf{R}^\top\mathbf{R}$ is assumed to fill at least the null space of $\mathbf{O}_e^\top\mathbf{O}_e$, making the problem (5.1) well-posed.

The optimality condition for the Tikhonov-regularized inverse problem (5.1) is given by

$$(\mathbf{O}_e^\top\mathbf{O}_e + \mathbf{R}^\top\mathbf{R})\boldsymbol{\mu}^{\mathrm{TR}} = \mathbf{O}_e^\top\mathbf{y}_d,$$

where we assume that $\mathbf{R}$ is chosen such that $\mathbf{O}_e^\top\mathbf{O}_e + \mathbf{R}^\top\mathbf{R}$ has rank q and is positive definite. The solution of (5.1) is then given by

$$\boldsymbol{\mu}^{\mathrm{TR}} = (\mathbf{O}_e^\top\mathbf{O}_e + \mathbf{R}^\top\mathbf{R})^{-1}\mathbf{O}_e^\top\mathbf{y}_d,$$

and the associated prediction is

$$\mathbf{y}_p^{\mathrm{TR}} = \mathbf{O}_p\boldsymbol{\mu}^{\mathrm{TR}}.$$

We will now show that we can modify the IFP method of section 3.2 for the inverse problem (5.1) to once again replicate the predictions without inverting for all of the parametric modes. The key here is a modification to the experiment observability gramian. In particular, we have $\mathbf{H}_e = \mathbf{O}_e^\top \mathbf{O}_e + \mathbf{R}^\top \mathbf{R}$. Given an IFP subspace $\mathcal{W}$, we obtain the IFP solution

$$\boldsymbol{\mu}^{\text{IFP}} = \arg\min_{\boldsymbol{\mu}\in\mathcal{W}} \frac{1}{2}\|\mathbf{y}_d - \mathbf{O}_e\boldsymbol{\mu}\|_2^2 + \frac{1}{2}\|\mathbf{R}\boldsymbol{\mu}\|_2^2. \tag{5.2}$$

However, the IFP basis $\mathbf{W}$ is now obtained by Algorithm 3.

Algorithm 3 IFP Basis Generation for Tikhonov-regularized approach

1: Define $\mathbf{G}_e = \mathbf{V}_e \mathbf{L}_e^{-1/2}$, where $\mathbf{V}_e \mathbf{L}_e \mathbf{V}_e^\top$ is the eigendecomposition of $\mathbf{O}_e^\top \mathbf{O}_e + \mathbf{R}^\top \mathbf{R}$.
2: Compute the reduced eigendecomposition $\boldsymbol{\Psi}\boldsymbol{\Sigma}^2\boldsymbol{\Psi}^\top$ of $\mathbf{G}_e^\top \mathbf{O}_p^\top \mathbf{O}_p \mathbf{G}_e$.
3: Define $\mathbf{W} = \mathbf{G}_e \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2}$.

The eigendecomposition in step 1 of Algorithm 3 will lead to square eigenvector matrices $\mathbf{V}_e \in \mathbb{R}^{n\times n}$ since $\mathbf{H}_e$ is full rank by design of $\mathbf{R}$, which increases both the operation and storage cost of the algorithm. It will be necessary to obtain n eigenvectors to maintain prediction exactness, and this cost could be prohibitive. Otherwise, one may choose to truncate the eigendecomposition based on the decay of the eigenvalues and obtain an approximation to our method. It should be mentioned, however, that since $\mathbf{R}$ is specified, the cost of each matrix-vector product $\mathbf{H}_e\mathbf{v}$ should not be much greater than the cost for the unregularized experiment observability gramian in the TSVD approach in section 3.2. So although there are a total of n (instead of r) eigenvectors to compute, the cost of obtaining each one is still about the same. It may also be possible to obtain the eigendecomposition more efficiently by leveraging the symmetry of the two matrices $\mathbf{O}_e^\top \mathbf{O}_e$ and $\mathbf{R}^\top \mathbf{R}$ by providing an estimate of the eigenvalues based on the eigenvalues of $\mathbf{O}_e^\top \mathbf{O}_e$ and $\mathbf{R}^\top \mathbf{R}$ separately [14].

THEOREM 5.1. *The predictions* $\mathbf{y}_p^{\text{IFP}}$ *arising from the IFP solution (see* (3.4)*)* $\boldsymbol{\mu}^{\text{IFP}}$ *of* (5.2) *with basis* $\mathbf{W}$ *defined by Algorithm* 3 *are identical to the Tikhonov-regularized predictions* $\mathbf{y}_p^{\text{TR}}$.

Proof. The proof is exactly the same as the proof of Theorem 4.1. Algorithms 2 and 3 work with the eigendecomposition of $\mathbf{H}_e$, which has been suitably redefined for the Tikhonov-regularized inverse problem here. ∎

5.2. Linear Gaussian statistical inverse problem. One way to account for uncertainty in prior knowledge and uncertainty in sensor measurements is through a statistical formulation of the inverse problem. In this section, we demonstrate how the IFP methodology can be extended to the statistical setting using a Bayesian approach with a Gaussian prior and Gaussian likelihood. The solution to the statistical inverse problem is a random variable and therefore has a distribution, which in this case is also Gaussian due to the linearity. That distribution over the parameter is then propagated through to the prediction, resulting in a distribution over predictions that we refer to as the *posterior predictive*. Instead of finding a single estimate of the predictions, we will determine a mean and covariance estimate. The mean estimate is obtained by the IFP method for a specific Tikhonov-regularized inverse problem; i.e., the procedure discussed in section 5.1 is all that is required. We show that the covariance estimate can be obtained at minimal additional cost through matrix multiplications involving the IFP basis $\mathbf{W}$ and singular values $\boldsymbol{\Sigma}$.

Let $\boldsymbol{\mu} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}_0)$ be the multivariate Gaussian random variable with mean $\mathbf{0}$ and covariance $\boldsymbol{\Gamma}_0$ representing our prior knowledge of the unknown parameter.[3] We assume that the measurements we make are corrupted by independent additive Gaussian errors $\boldsymbol{\epsilon} = \mathbf{y}_d - \mathbf{O}_e\boldsymbol{\mu} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ with zero mean and variance σ^2.

Given that the map from parameters to experimental outputs is linear, by Bayes's rule, we write the posterior estimate of the parameter

$$\boldsymbol{\mu}|\mathbf{y}_d \sim \mathcal{N}(\boldsymbol{\mu}_\pi, \boldsymbol{\Gamma}_\pi),$$

where

$$\begin{aligned} \boldsymbol{\mu}_\pi &= \sigma^{-2}(\boldsymbol{\Gamma}_0^{-1} + \sigma^{-2}\mathbf{O}_e^\top\mathbf{O}_e)^{-1}\mathbf{O}_e^\top\mathbf{y}_d, \\ \boldsymbol{\Gamma}_\pi &= (\boldsymbol{\Gamma}_0^{-1} + \sigma^{-2}\mathbf{O}_e^\top\mathbf{O}_e)^{-1}. \end{aligned} \tag{5.3}$$

Recall, however, that we are interested only in the statistics of the prediction arising from simulations utilizing this parameter. That is, the posterior predictive

$$\mathbf{y}_p|\mathbf{y}_d \sim \mathcal{N}(\mathbf{O}_p\boldsymbol{\mu}_\pi, \mathbf{O}_p\boldsymbol{\Gamma}_\pi\mathbf{O}_p^\top).$$

It is this posterior predictive distribution $\mathbf{y}_p|\mathbf{y}_d$ that will be replicated by the IFP method.

We will now show that the IFP approach can obtain the posterior predictive. First note that the posterior predictive mean is obtained as the solution to a Tikhonov-regularized inverse problem.

THEOREM 5.2. *The posterior predictive mean $\mathbf{O}_p\boldsymbol{\mu}_\pi$ is obtained by solving* (5.2) *with $\mathbf{W}$ generated by Algorithm* 3*, where $\mathbf{R}$ is chosen such that $\mathbf{R}^\top\mathbf{R} = \sigma^2\boldsymbol{\Gamma}_0^{-1}$.*

Proof. We first rewrite the Tikhonov-regularized inverse problem (5.1) to account for the sensor error and prior knowledge; i.e., we search for the parameter

$$\begin{aligned} \boldsymbol{\mu}^* &= \arg\min_{\boldsymbol{\mu}\in\mathbb{R}^q} \frac{1}{2}\|\mathbf{y}_d - \mathbf{O}_e\boldsymbol{\mu}\|_2^2 + \frac{1}{2}\sigma^2\boldsymbol{\mu}^\top\boldsymbol{\Gamma}_0^{-1}\boldsymbol{\mu} \\ &= \arg\min_{\boldsymbol{\mu}\in\mathbb{R}^q} \frac{1}{2\sigma^2}\|\mathbf{y}_d - \mathbf{O}_e\boldsymbol{\mu}\|_2^2 + \frac{1}{2}\boldsymbol{\mu}^\top\boldsymbol{\Gamma}_0^{-1}\boldsymbol{\mu}. \end{aligned} \tag{5.4}$$

We now show that this is precisely the posterior mean. The first-order optimality condition of (5.4) is given by

$$(\boldsymbol{\Gamma}_0^{-1} + \sigma^{-2}\mathbf{O}_e^\top\mathbf{O}_e)\boldsymbol{\mu}^* = \sigma^{-2}\mathbf{O}_e^\top\mathbf{y}_d,$$

whose solution is $\boldsymbol{\mu}^* = \sigma^{-2}(\boldsymbol{\Gamma}_0^{-1} + \sigma^{-2}\mathbf{O}_e^\top\mathbf{O}_e)^{-1}\mathbf{O}_e^\top\mathbf{y}_d$. This is equal to $\boldsymbol{\mu}_\pi$ given in (5.3). The remainder of the proof is completely analogous to that of Theorem 5.1. ∎

The following theorem states that the posterior predictive covariance can be recovered by a matrix multiplication involving the IFP basis $\mathbf{W}$ and the diagonal matrix of singular values $\boldsymbol{\Sigma}$ already computed in the posterior predictive mean obtained above.

THEOREM 5.3. *The posterior predictive covariance can be obtained as a matrix multiplication involving the IFP basis $\mathbf{W}$ and singular values $\boldsymbol{\Sigma}$ from Algorithm* 3 *since*

$$\mathbf{O}_p\boldsymbol{\Gamma}_\pi\mathbf{O}_p^\top = \mathbf{O}_p\mathbf{W}\boldsymbol{\Sigma}\mathbf{W}^\top\mathbf{O}_p^\top.$$

[3]The method readily admits priors with nonzero mean. Both the traditional approach and IFP method would then target the deviation from the mean; the covariance remains unchanged.

Proof. The posterior predictive covariance is given by

$$\mathbf{O}_p\boldsymbol{\Gamma}_\pi\mathbf{O}_p^\top = \mathbf{O}_p(\boldsymbol{\Gamma}_0^{-1} + \sigma^{-2}\mathbf{O}_e^\top\mathbf{O}_e)^{-1}\mathbf{O}_p^\top. \tag{5.5}$$

Recall that $\mathbf{H}_e = \boldsymbol{\Gamma}_0^{-1} + \sigma^{-2}\mathbf{O}_e^\top\mathbf{O}_e$ is full rank; therefore, $\mathbf{H}_e^{-1} = \mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top$ and $\mathbf{V}_e\mathbf{L}_e^{-1}\mathbf{V}_e^\top = \mathbf{G}_e\mathbf{G}_e^\top$. Substituting into (5.5), we find

$$\mathbf{O}_p\boldsymbol{\Gamma}_\pi\mathbf{O}_p^\top = \mathbf{O}_p\mathbf{G}_e\mathbf{G}_e^\top\mathbf{O}_p^\top. \tag{5.6}$$

Since $\text{range}(\boldsymbol{\Psi})^\perp \subset \text{null}(\mathbf{O}_p\mathbf{G}_e)$ and $(\mathbf{I} - \boldsymbol{\Psi}\boldsymbol{\Psi}^\top)$ is the orthogonal projector onto $\text{range}(\boldsymbol{\Psi})^\perp$, we have $\mathbf{O}_p\mathbf{G}_e(\mathbf{I} - \boldsymbol{\Psi}\boldsymbol{\Psi}^\top) = \mathbf{0}$ and therefore $\mathbf{O}_p\mathbf{G}_e = \mathbf{O}_p\mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top$. Substituting into (5.6), we find

$$\mathbf{O}_p\boldsymbol{\Gamma}_\pi\mathbf{O}_p^\top = \mathbf{O}_p\mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_p^\top.$$

Inserting the identity $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2} = \mathbf{I}$, we obtain

$$\mathbf{O}_p\boldsymbol{\Gamma}_\pi\mathbf{O}_p^\top = \mathbf{O}_p\mathbf{G}_e\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Psi}^\top\mathbf{G}_e^\top\mathbf{O}_p^\top = \mathbf{O}_p\mathbf{W}\boldsymbol{\Sigma}\mathbf{W}^\top\mathbf{O}_p^\top. \qquad \square$$

In the next section, we will apply the IFP method to a two-dimensional (2-D) advection-diffusion model problem simulating contaminant transport.

6. Application to 2-D advection diffusion. Consider a contaminant inversion and prediction problem modeled by advection diffusion in two dimensions. A one-time contaminant release advects and diffuses throughout the domain. A handful of sensors measuring concentration of the contaminant give localized readings at discrete times. We wish to predict output quantities of interest depending on the contaminant concentration in the domain at later times. In section 6.1 we provide details about the problem description. In section 6.2 we present the numerical results corroborating the theory developed above.

6.1. Problem description. Let $\mathbf{z} = (z_1, z_2)$ be the spatial coordinates of a 2-D rectangular domain $\Omega = \{(z_1, z_2) \mid 0 \le z_1 \le 1,\ 0 \le z_2 \le 0.4\}$. Denote by $\partial\Omega$ the boundary of that domain. Let $c(\mathbf{z}, t) : \Omega \times \mathbb{R}_+ \to \mathbb{R}_+$ be the contaminant concentration at $\mathbf{z}$ and time t, where $\mathbb{R}_+ = [0, \infty)$. We prescribe ambient wind velocity $\mathbf{u} = (1.5, 0.4)$ constant throughout the domain. Let the diffusivity $\kappa = 0.02$ also be constant. Given initial condition $c_0(\mathbf{z}) = c(\mathbf{z}, 0)$, the contaminant evolves in time according to the advection-diffusion equation

$$\frac{\partial c}{\partial t} = -\kappa\nabla^2 c + \mathbf{u}\cdot\nabla c, \qquad \mathbf{z} \in \Omega, t > 0, \tag{6.1}$$

$$\nabla c \cdot \mathbf{n} = 0, \qquad \mathbf{z} \in \partial\Omega, t > 0, \tag{6.2}$$

where $\nabla = (\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2})$ and $\mathbf{n}$ denotes the outward-pointing unit normal on each of the four segments of $\partial\Omega$.

The experimental outputs $\mathbf{y}_e(t) = (y_{e_1}(t), y_{e_2}(t), \ldots, y_{e_{n_s}}(t))$ at time t are given by localized integrals of the contaminant concentration, i.e.,

$$y_{e_i}(t) = \int_\Omega c(\mathbf{z}, t) \exp\left\{-\frac{1}{2\sigma_e^2}\|\mathbf{z} - \mathbf{z}_i\|^2\right\} d\mathbf{z}, \quad i = 1, 2, \ldots, n_s, \tag{6.3}$$

where $\mathbf{z}_i$ is the location of the ith sensor, $\sigma_e = 0.01$ is a measure of the sensing radius for all sensors, $\|\cdot\|$ represents the Euclidean norm in $\mathbb{R}^2$, and n_s is the number of

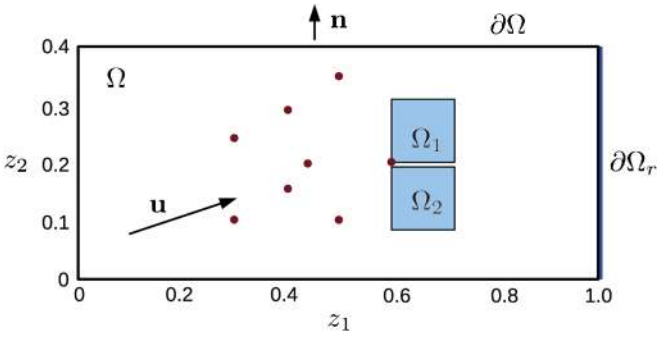

FIG. 6.1. *The domain and the eight sensor center locations.*

sensors distributed in the domain. Contaminant concentration readings are available only at discrete times $t = t_0, t_1, \ldots, t_{n_r}$, where n_r is the number of readings. In what follows, we will denote the concatenation of experimental outputs as

$$\mathbf{y}_e = \left[\begin{array}{cccc} \mathbf{y}_e^\top(t_0) & \mathbf{y}_e^\top(t_1) & \cdots & \mathbf{y}_e^\top(t_{n_r}) \end{array} \right]^\top. \tag{6.4}$$

Then, $\mathbf{y}_e \in \mathbb{R}^r$, where $r = n_s n_r$. In our numerical experiments, we use eight sensors. The domain and sensor locations are shown in Figure 6.1. The sensors are placed in the domain with knowledge of the synthetic initial contaminant concentration but are not chosen with consideration for any of the outputs of interest. Both the IFP and traditional approaches utilize the same sensor configuration. We make measurements at time instants $t = \Delta t, 2\Delta t, \ldots, 30\Delta t$, where $\Delta t = 5 \times 10^{-3}$.

For the numerical experiments we compare prediction outputs from the three traditional methods to their respective IFP implementations. We define three time-dependent prediction outputs of interest and two scalar prediction outputs.

Let $\partial\Omega_r = \{(z_1, z_2) \mid z_1 = 1, 0 < z_2 < 0.4\}$ denote the right boundary of the domain. One prediction output quantity of interest is the total contaminant propagating outward through this boundary as a function of time in the interval $60\Delta t \le t \le 70\Delta t$, i.e.,

$$y_{p_0}(t) = \int_{\partial\Omega_r} c(\mathbf{z}, t)\mathbf{u} \cdot \mathbf{n}_r \, dz_2, \quad 60\Delta t \le t \le 70\Delta t,$$

where $\mathbf{n}_r = (1, 0)$ is the outward-pointing unit normal for the right boundary.

We define a second prediction output of interest that is the total contaminant contained within a box on the interior of the domain. Let $\Omega_1 = \{(z_1, z_2) \mid 0.6023 \le z_1 \le 0.6932, 0.2000 \le z_2 \le 0.2909\}$ and define

$$y_{p_1}(t) = \int_{\Omega_1} c(\mathbf{z}, t) dz_1 \, dz_2, \quad 25\Delta t \le t \le 50\Delta t.$$

For the demonstration of the IFP methodology in the statistical setting, we will use two scalar prediction output quantities of interest. Let $\Omega_2 = \{(z_1, z_2) \mid 0.6023 \le z_1 \le 0.6932, 0.1000 \le z_2 \le 0.1909\}$; see Figure 6.1. Define

$$y_{p_2}(t) = \int_{\Omega_2} c(\mathbf{z}, t) dz_1 \, dz_2, \quad 25\Delta t \le t \le 50\Delta t.$$

TABLE 6.1
Standard deviations and center locations for the five Gaussian plumes summed to form the initial condition (6.5) *used to generate the synthetic data for the numerical experiments. The initial condition is pictured in Figure* 6.2.

i	1	2	3	4	5
α_i	0.07	0.05	0.07	0.05	0.05
$z1_i$	0.20	0.25	0.35	0.45	0.55
$z2_i$	0.15	0.15	0.20	0.20	0.12

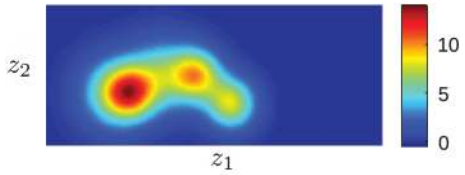

FIG. 6.2. *The initial contaminant concentration* $c_0(\mathbf{z})$ *used to generate synthetic data for the numerical experiments.*

Our third and fourth prediction outputs of interest are the time-integrated quantities

$$y_{p_3} = \int_{t=25\Delta t}^{50\Delta t} y_{p_1}(t)\, dt \qquad \text{and} \qquad y_{p_4} = \int_{t=25\Delta t}^{50\Delta t} y_{p_2}(t)\, dt.$$

The IFP method utilizes the experimental data to infer those components of the parameter that are relevant for predicting the output quantities of interest. Our numerical experiments generate synthetic data by prescribing an initial condition that is a sum of Gaussian plumes, i.e.,

$$c_0(\mathbf{z}) = \sum_{i=1}^{5} \frac{1}{\alpha_i \sqrt{2\pi}} \exp\left\{ -\frac{1}{2\alpha_i^2} \|\mathbf{z} - \mathbf{z}_i\|^2 \right\}, \tag{6.5}$$

where the standard deviations α_i and centers $\mathbf{z}_i$, $i = 1, 2, \ldots, 5$, are given in Table 6.1.

The initial condition is pictured in Figure 6.2. For reference we present four snapshots of the contaminant concentration in the domain at times $t = 10\Delta t, 30\Delta t, 50\Delta t, 70\Delta t$ in Figure 6.3. The synthetic data is corrupted by noise for our experi-

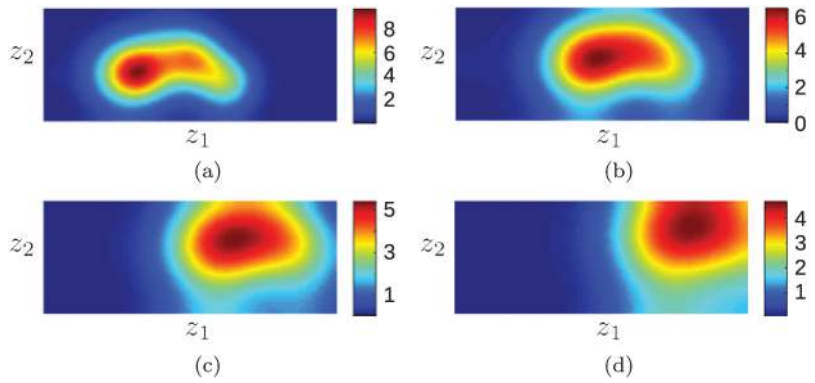

FIG. 6.3. *The evolution of the contaminant whose initial concentration is shown in Figure* 6.2 *at time steps* (a) $t = 10\Delta t$, (b) $t = 30\Delta t$, (c) $t = 50\Delta t$, *and* (d) $t = 70\Delta t$.

ments by adding random errors distributed normally with zero mean and variance $\sigma^2 = 0.01$.

The goal of our IFP method is not to obtain the true output of interest based on the true parameter but rather to match the prediction obtained by employing any of the traditional inference formulations discussed above. In other words, we are not proposing to improve accuracy of inference but rather to exploit final goals to make existing inference methods more efficient online and more transparent with respect to injected information.

For numerical simulation we discretize the continuous formulation (6.1)–(6.2) in space and time. We discretize in space by the finite element method (FEM) using a regular simplicial mesh with 44 and 88 elements each on the short and long boundary edges, respectively. The mesh has 7744 elements and 4005 nodes. We use a linear nodal basis to approximate the numerical solution. The numerical instability due to the advection term is treated by a streamline upwind Petrov–Galerkin (SUPG) correction [6]. The semidiscrete equation is time-stepped by Crank–Nicolson, leading to a system of the form (2.1) with $\mathbf{u}_k = \mathbf{0}\ \forall k$.

The integral computations for calculating the experimental outputs and the prediction output are also approximated by the discrete solution. For the experimental outputs, the integral is computed using a mass-matrix-weighted inner product between the rapidly decaying Gaussian sensor in the integrand of (6.3) and the solution vector $\mathbf{x}_k$ at time step k. The prediction output quantity of interest is estimated by using a midpoint rule in time, and the linear nodal basis leads to a midpoint integration rule in space as well. In both experiment and prediction, the outputs are linear functions of the initial condition. For example, define $\mathbf{C}_e$ such that the experimental outputs (6.4) are given by

$$\mathbf{y}_e(k\Delta t) = \mathbf{C}_e\mathbf{x}_k, \qquad k = 1, 2, \ldots, 30.$$

Let $\boldsymbol{\mu} = \mathbf{x}_0$; then $\mathbf{y}_e = \mathbf{O}_e\boldsymbol{\mu}$, where $\mathbf{O}_e = \begin{bmatrix} (\mathbf{C}_e\mathbf{A})^\top & (\mathbf{C}_e\mathbf{A}^2)^\top & \cdots & (\mathbf{C}_e\mathbf{A}^{30})^\top \end{bmatrix}^\top$. Similarly, define $\mathbf{C}_{p_0}$ such that $\mathbf{y}_{p_0}(k\Delta t) = \mathbf{C}_{p_0}\mathbf{x}_k$ for $k = 60, \ldots, 70$; then $\mathbf{y}_{p_0} = \mathbf{O}_p\boldsymbol{\mu}$, where $\mathbf{O}_p = \begin{bmatrix} (\mathbf{C}_{p_0}\mathbf{A}^{60})^\top & (\mathbf{C}_{p_0}\mathbf{A}^{61})^\top & \cdots & (\mathbf{C}_{p_0}\mathbf{A}^{70})^\top \end{bmatrix}^\top$. For the other outputs of interest, we need only redefine $\mathbf{O}_p$ appropriately.

6.2. Results for numerical experiments. In this section we present results for the 2-D advection-diffusion application described in the preceding section. We will demonstrate the IFP methodology in each of the three inverse problem formulations described above: TSVD, Tikhonov-regularized, and Gaussian statistical. All of the problems are implemented in MATLAB and utilize the built-in LAPACK eigenvalue solver.

6.2.1. TSVD approach. The IFP method was applied in the context of the TSVD approach to the initial condition problem described above. In this case, we focus on the time-dependent output $y_{p_0}(t)$. Similar results are obtained for all of the outputs.

Algorithm 2 is implemented to obtain the IFP basis $\mathbf{W} \in \mathbb{R}^{q\times s}$ whose first four modes are plotted in Figure 6.4. The high frequency characteristics are inherited from the eigenmodes $\mathbf{V}_e$. For this problem, there are 240 experimental outputs (eight concentration sensors over 30 time steps) and there are 11 prediction outputs (right-side flux over 11 time steps). Although $\mathbf{H}_e$ mathematically has rank 240, the reduced eigendecomposition reveals that it can be approximated almost exactly by a rank-54 matrix; this is due to the numerical implementation and tolerance in the eigensolver.

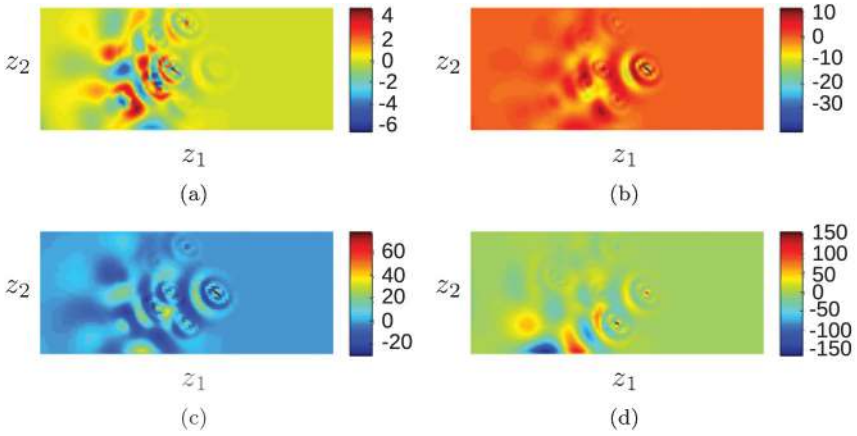

FIG. 6.4. *The first four modes* (a)–(d) *of the IFP basis* $\mathbf{W}$ *for the TSVD approach.*

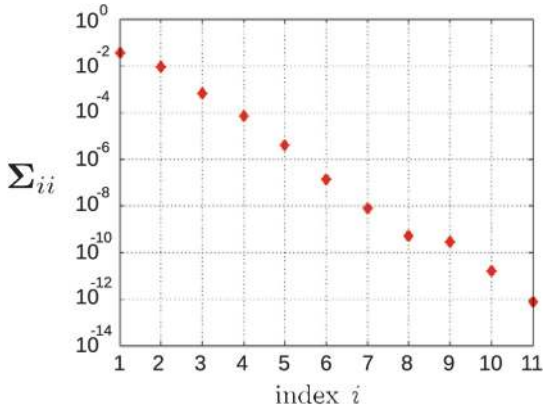

FIG. 6.5. *The singular values on the diagonal of* $\mathbf{\Sigma}$ *reflecting the joint measure of experiment and prediction observability for the TSVD approach.*

The singular values $\mathbf{\Sigma}_{ii}$ indicate that there is a subspace of dimension $s = 11$ for which there will be no information loss in the inference-to-prediction process; therefore, the IFP method yields a basis $\mathbf{W} \in \mathbb{R}^{q \times s}$. Decay of the singular values (see Figure 6.5) indicates that further truncation to fewer than 11 modes is possible; the IFP solution would then not result in exact predictions, but the error incurred by truncating the last three or four modes would be very small.

We now turn to the results of the inversion. In Figure 6.6 we plot (a) the real initial condition, (b) the TSVD estimate, (c) the IFP estimate, and (d) the difference or error $\boldsymbol{\mu}_e = \boldsymbol{\mu}^{\mathrm{TSVD}} - \boldsymbol{\mu}^{\mathrm{IFP}}$. It is important to recall here that the IFP approach targets prediction outputs and is not designed to accurately infer the unknown parameter. Clearly the traditional inference method is more proficient at that.

What is relevant, though, is the propagation of the error $\boldsymbol{\mu}_e$ to the prediction output $y_{p_0}(t)$. If the IFP estimate μ^{IFP} results in the same predictions as the TSVD estimate μ^{TSVD} as the theory claims, then we expect that the error initial condition μ_e will lead to zero prediction. In Figure 6.7 we plot snapshots of the evolving error field beginning with initial condition $\boldsymbol{\mu}_e$ at four time steps within the prediction time

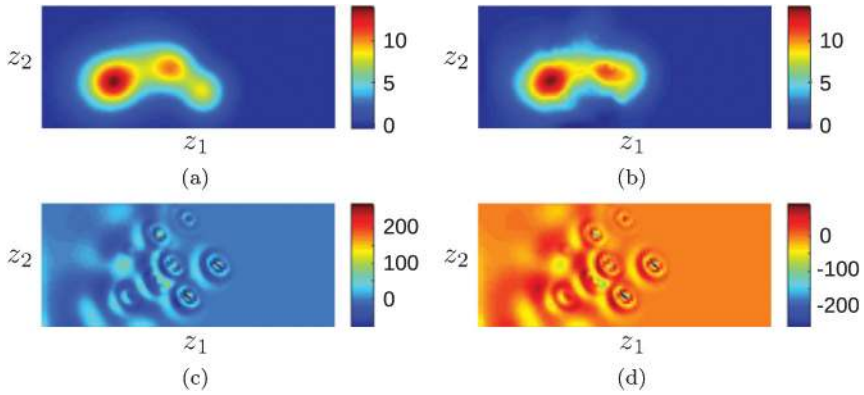

FIG. 6.6. *The* (a) *real initial condition,* (b) *TSVD estimate,* (c) *IFP estimate, and* (d) *error* $\boldsymbol{\mu}_e = \boldsymbol{\mu}^{TSVD} - \boldsymbol{\mu}^{IFP}$. *In Figure* 6.7 *we show the propagation of* $\boldsymbol{\mu}_e$ *to the output time steps.*

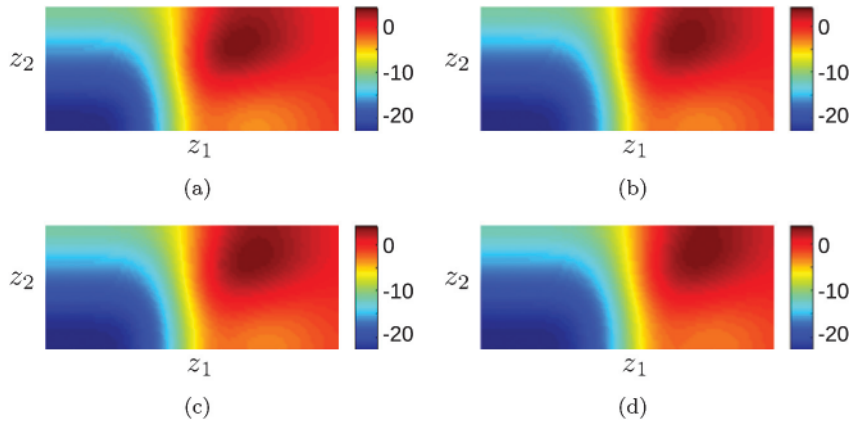

FIG. 6.7. *The propagation of* $\boldsymbol{\mu}_e$ *according to the advection-diffusion equation to the prediction output* y_{p_0} *at time steps* (a) $t = 60\Delta t$, (b) $t = 64\Delta t$, (c) $t = 65\Delta t$, *and* (d) $t = 70\Delta t$. *The integrated flux through the right boundary is negligible.*

region $t \in [60\Delta t, 70\Delta t]$. It can be seen that the error propagation leads to negligible flux through the right boundary, as the theory predicts.

In Figure 6.8 we plot the prediction outputs for both the TSVD and IFP approaches, as well as the error in the outputs. The prediction output curves are directly on top of each other, and the error is seven orders of magnitude less than the output predictions themselves. The error is not identically zero due to the numerical approximations, e.g., in the eigenvector solver, where tolerances are used.

Although our results above do not involve further truncation from the original IFP basis in $s = 11$ dimensions, we show in Figure 6.9 the error in prediction outputs as it varies with the number of basis vectors included in the IFP estimate. The error is significant if one includes just a few basis vectors, but as soon as six vectors are included the error drops to 10^{-6}.

In the next section, we demonstrate the approach for a Tikhonov-regularized inverse problem.

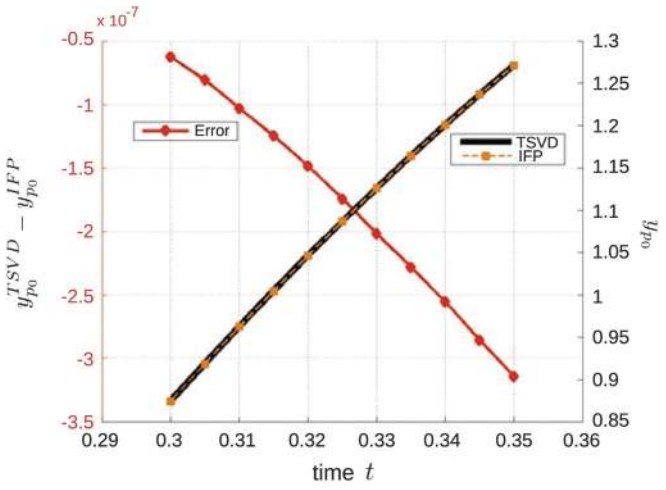

FIG. 6.8. *The (left ordinate axis) error between the prediction outputs from the TSVD and IFP approaches (red, diamonds) and the (right ordinate axis) predictions themselves based on TSVD (black, solid) and IFP (orange, dashed, squares) approaches.*

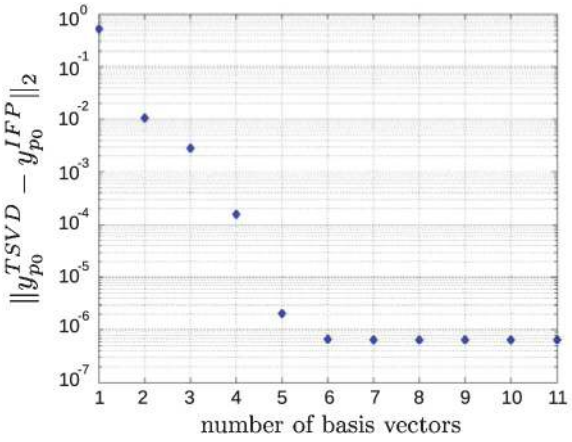

FIG. 6.9. *The error in prediction outputs* $\|y_{p_0}^{\mathrm{TSVD}} - y_{p_0}^{\mathrm{IFP}}\|_2$ *between the TSVD and IFP predictions vs. the number of IFP basis vectors included in* $\mathbf{W}$.

6.2.2. Tikhonov-regularized approach. In the Tikhonov-regularized approach, we define the matrix $\mathbf{R}$ implicitly by setting the diagonal matrix $\left(\mathbf{R}^\top \mathbf{R}\right)_{jj} = 0.1\lambda_{\min}(1+(99j/4004))$ where $\lambda_{\min}$ is the smallest nonzero eigenvalue of the experiment observability gramian. This spreads the eigenvalues of $\mathbf{R}^\top \mathbf{R}$ evenly between approximately 0.0660 and 6.6028. We focus in this section on the output $y_{p_1}(t)$ defined above. Results are similar for the other outputs of interest.

For this experiment, we find that $s = 26$ is the dimension of the IFP basis and here $r = q = 4005$ since the regularization fills the null space of the experiment observability gramian. The first four basis modes are plotted in Figure 6.10, and the singular values are shown in Figure 6.11.

In Figure 6.12 we show the (a) real initial condition, (b) Tikhonov-regularized (TR) estimate, (c) IFP estimate, and the (d) error $\boldsymbol{\mu}_e = \boldsymbol{\mu}^{\mathrm{TR}} - \boldsymbol{\mu}^{\mathrm{IFP}}$. The evolution

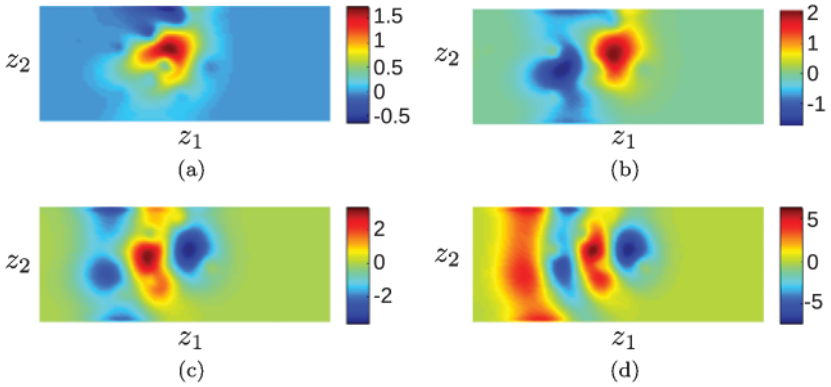

FIG. 6.10. *The first four modes* (a)–(d) *of the IFP basis* $\mathbf{W}$ *for the Tikhonov-regularized approach.*

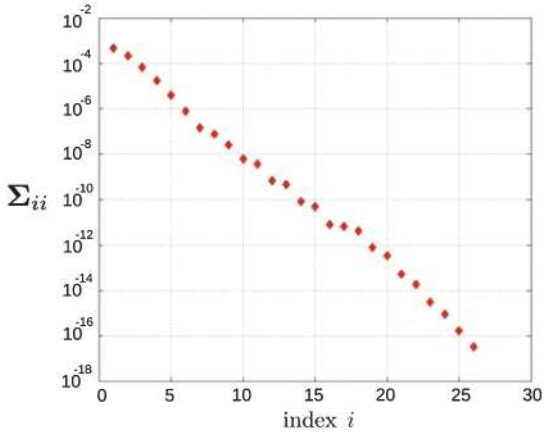

FIG. 6.11. *The singular values on the diagonal of* $\mathbf{\Sigma}$ *reflecting the joint measure of experiment and prediction observability for the Tikhonov-regularized inverse problem approach.*

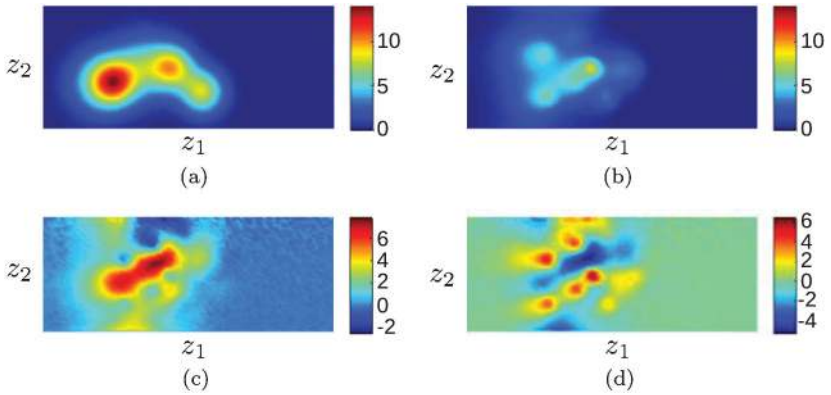

FIG. 6.12. *The* (a) *real initial condition,* (b) *Tikhonov-regularized (TR) estimate,* (c) *IFP estimate, and* (d) *error* $\boldsymbol{\mu}_e = \boldsymbol{\mu}^{TR} - \boldsymbol{\mu}^{IFP}$. *In Figure* 6.13 *we show the propagation of* $\boldsymbol{\mu}_e$ *to the output time steps.*

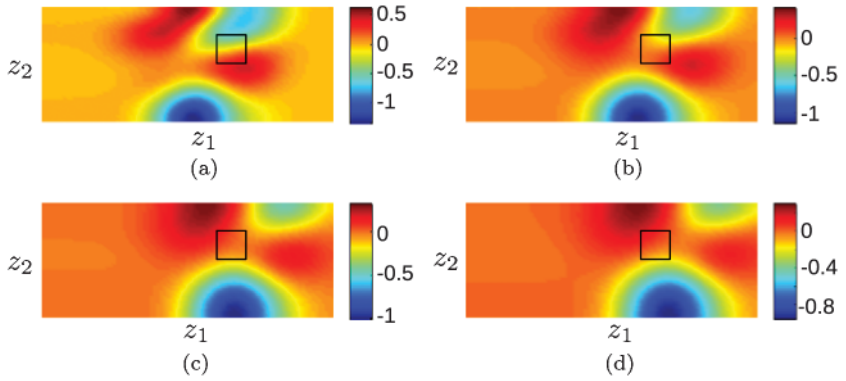

FIG. 6.13. *The propagation of* $\boldsymbol{\mu}_e$ *according to the advection-diffusion equation to the prediction output* y_{p_1} *at time steps* (a) $t = 25\Delta t$, (b) $t = 35\Delta t$, (c) $t = 45\Delta t$, *and* (d) $t = 50\Delta t$. *The average concentration inside of* Ω_1 *(box) is negligible for all time steps* $t \in [25\Delta t, 50\Delta t]$.

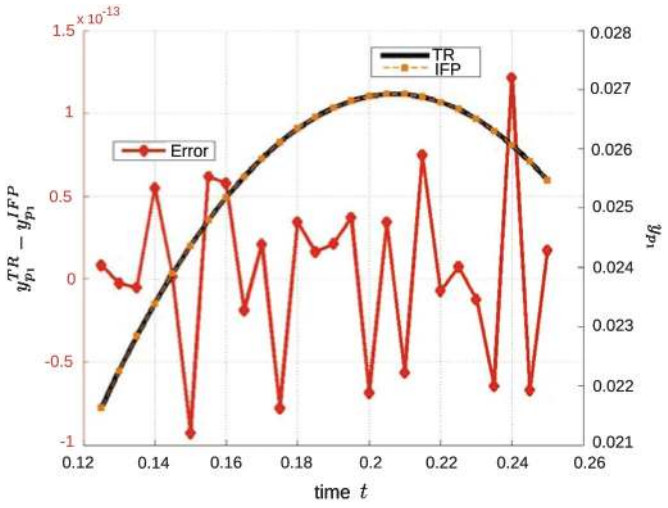

FIG. 6.14. *The (left ordinate axis) error between the prediction outputs from the TR and IFP approaches (red, diamonds) and the (right ordinate axis) predictions themselves based on TR (black, solid) and IFP (orange, dashed, squares) approaches.*

of the error $\boldsymbol{\mu}_e$ through the advection-diffusion equation is shown in Figure 6.13 for four time steps in the temporal range of the predictions $t \in [25\Delta t, 50\Delta t]$. Consistently with the equal predictions based on TR and IFP parameter estimates, we see that the error initial condition leads to negligible average contaminant concentration within Ω_1 over this time period.

We show the error in predictions and the predictions themselves in Figure 6.14. The errors are again many orders of magnitude smaller than the predictions, and the predictions themselves lie directly on top of each other. This result is consistent with the theory presented in the preceding sections.

6.2.3. Gaussian statistical approach. For the statistical approach, we specify a prior distribution on the parameter $\boldsymbol{\mu}$. We use a multivariate normal prior with mean zero and covariance matrix $\boldsymbol{\Gamma}_0$ with the (i, j)th element given by

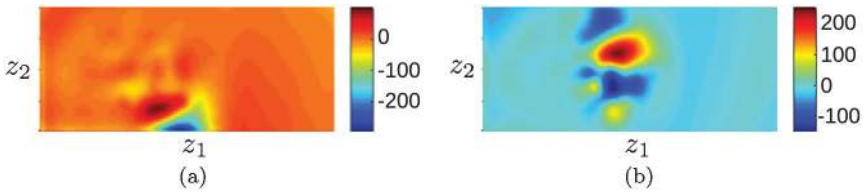

FIG. 6.15. *The only two modes* (a) *and* (b) *of the IFP basis* $\mathbf{W}$ *for the Gaussian statistical approach.*

TABLE 6.2

Means and covariances for the prediction outputs. In each cell, we list the result from the traditional approach (TA), the result from the IFP approach (IFP), and the absolute value of the error (E). Equiprobable contours for the associated probability density functions are pictured in Figure 6.16.

		Mean	Covariance y_{p_3}	Covariance y_{p_4}
y_{p_3}	TA	5.1983E-1	1.9376E-8	2.0276E-9
	IFP	5.1983E-1	1.9376E-8	2.0276E-9
	E	2.3309E-11	1.3235E-23	4.6736E-23
y_{p_4}	TA	3.0017E-1	2.0276E-9	8.1433E-8
	IFP	3.0017E-1	2.0276E-9	8.1433E-8
	E	1.0047E-10	4.6736E-23	7.9409E-23

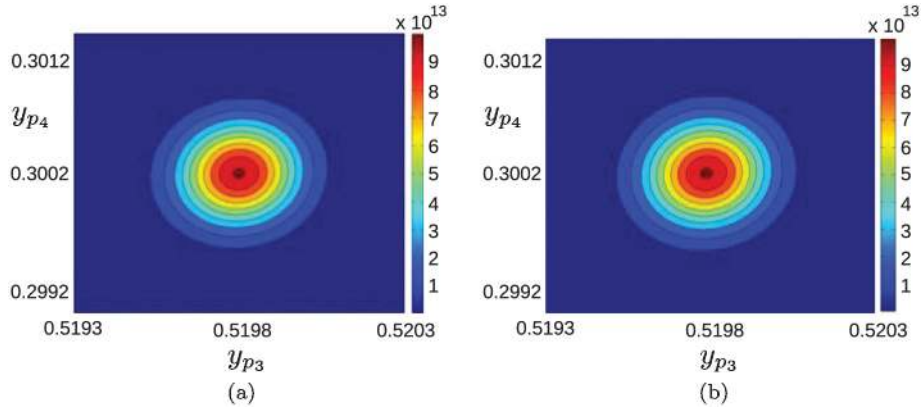

FIG. 6.16. *Contour plots of the joint probability density function over the outputs* (y_{p_3}, y_{p_4}) *for the* (a) *traditional approach and the* (b) *IFP approach. The means and covariance matrices corresponding to both approaches are also given in Table* 6.2.

$$\mathbf{\Gamma}_{0_{ij}} = a \exp\left\{\frac{-\|\mathbf{z}_i - \mathbf{z}_j\|_2^2}{2b^2}\right\} + c\mathbf{I}, \quad 1 \leq i, j, \leq n,$$

with constants $a = 0.001$, $b = 0.5$, and $c = 0.1$. We assume the sensors are corrupted by additive Gaussian noise that is independent and identically distributed each with zero mean and variance $\sigma^2 = 0.01$. For this numerical experiment we focus on the scalar outputs y_{p_3} and y_{p_4}. We present results here for the posterior predictive mean and posterior predictive covariance; however, more attention is given to the covariance since the mean computation is analogous to the Tikhonov-regularized problem in the preceding section.

In this case again the prior distribution affects every mode of the parameter so that $r = q = 4005$. On the other hand, there are only two scalar outputs of interest, so we find that the IFP basis has dimension $s = 2$. In Figure 6.15 we plot these two basis vectors. The singular values are $\mathbf{\Sigma}_{11} = 2.855 \times 10^{-4}$ and $\mathbf{\Sigma}_{22} = 1.390 \times 10^{-4}$.

The results are presented in Table 6.2 and Figure 6.16. The estimated posterior predictive means and covariances are nearly identical, having componentwise errors many orders of magnitude smaller than the values themselves. Once again, the numerical results reflect the theory. Inverting for just two modes of the parameter is sufficient to exactly obtain the posterior predictive distribution. In Figure 6.16 we plot equiprobability contours of the posterior predictive distribution from the (a) traditional and (b) IFP approaches.

7. Conclusion. A goal-oriented inference approach can exploit low-dimensional prediction output quantities of interest by inferring many fewer parameter modes. Errors are tolerated in the parameter estimate without sacrificing any accuracy in the predictions. In the linear setting, the theoretical underpinnings of the IFP method are established for its use in three traditional inverse problem formulations. Prediction exactness holds for all cases, regardless of whether or not prediction modes are well informed by experimental data. The approach has a compelling connection to the balanced truncation model reduction. As a result the number of parameter modes to be inferred in the method can be further truncated based on a joint measure of experiment and prediction observability given by the associated Hankel singular values, which rank the importance of parameter modes to the data-to-prediction map. The new approach admits an offline/online decomposition that may make inference feasible on lightweight, portable devices in the field.

Acknowledgment. The authors wish to express their gratitude to Tarek El Moselhy for his insightful comments and careful review of the proofs.

REFERENCES

[1] O. Alifanov, *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin, 1994.
[2] H. Ammari, E. Bonnetier, Y. Capdeboscq, M. Tanter, and M. Fink, *Electrical impedance tomography by elastic deformation*, SIAM J. Appl. Math., 68 (2008), pp. 1557–1573.
[3] S. R. Arridge, *Optical tomography in medical imaging*, Inverse Problems, 15 (1999), pp. R41–R93.
[4] P. Benner, E. S. Quintana-Orti, and G. Quintana-Orti, *Balanced truncation model reduction of large-scale dense systems on parallel computers*, Math. Comput. Model. Dyn. Syst., 6 (2000), pp. 383–405.
[5] L. T. Biegler, O. Ghattas, M. Heinkenschloss, D. Keyes, and B. van Bloemen Waanders, eds., *Real-Time PDE- Constrained Optimization*, SIAM, Philadelphia, 2007.
[6] A. Brooks and T. J. R. Hughes, *Streamline upwind/Petrov-Galerkin methods for advection-dominated flows*, in Proceedings of the Third International Conference on Finite Element Methods in Fluid Flows, 1980, pp. 283–292.
[7] C. Burstedde and O. Ghattas, *Algorithmic strategies for full waveform inversion: 1d experiments*, Geophysics, 74 (2009), pp. WCC37–WCC46.
[8] J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations: Vol.* I: *Theory*, SIAM, Philadelphia, 2002.
[9] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
[10] S. Gugercin and A. C. Antoulas, *A survey of model reduction by balanced truncation and some new results*, Internat. J. Control, 77 (2004), pp. 748–766.
[11] T. Hohage and S. Langer, *Acceleration techniques for regularized Newton methods applied to electromagnetic inverse medium scattering problems*, Inverse Problems, 26 (2010), 074011.
[12] R. E. Kalman, *On the general theory of control systems*, in Proceedings of the First International Congress on Automatic Control, 1960, pp. 481–491.

[13] E. Kalnay, *Atmospheric Modeling, Data Assimilation, and Predictability*, Cambridge University Press, Cambridge, UK, 2003.
[14] A. Knutson and T. Tao, *Honeycombs and sums of Hermitian matrices*, Notices Amer. Math. Soc., 48 (2001), pp. 175–186.
[15] J. Li and J. White, *Reduction of large circuit models via low rank approximate gramians*, Int. J. Appl. Math. Comput. Sci., 11 (2001), pp. 101–121.
[16] B. C. Moore, *Principal component analysis in linear systems: Controllability, observability, and model reduction*, IEEE Trans. Automat. Control, 26 (1981), pp. 17–31.
[17] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, 2005.
[18] M. J. Zandvliet, J. F. M. van Doren, O. H. Bosgra, J. D. Jansen, and P. M. J. van den Hof, *Controllability, observability and identifiability in single-phase porous media flow*, Comput. Geosci., 12 (2008), pp. 605–622.