Gaussian functional regression for state prediction using linear PDE models and observations

N. C. Nguyen∗, H. Men†, R. M. Freund‡, and J. Peraire§

Abstract

Partial differential equations (PDEs) are commonly used to model a wide variety of physical phenomena. A PDE model of a physical problem is typically described by conservation laws, constitutive laws, material properties, boundary conditions, boundary data, and geometry. In most practical applications, however, the PDE model is only an approximation to the real physical problem due to both (i) the deliberate mathematical simplification of the model to keep it tractable, and (ii) the inherent uncertainty of the physical parameters. In such cases, the PDE model may not produce a good prediction of the true state of the underlying physical problem. In this paper, we introduce a Gaussian functional regression method that incorporates observations into a deterministic linear PDE model to improve its prediction of the true state. Our method is devised as follows. First, we augment the PDE model with a random Gaussian functional which serves to represent various sources of uncertainty in the model. We next derive a linear regression model for the Gaussian functional by utilizing observations and adjoint states. We then apply the principle of Gaussian process to determine the posterior distribution of the Gaussian functional. This allows us to compute the posterior distribution for our estimate of the true state and provide a statistical measure of the prediction error. Furthermore, we consider the problem of experimental design in this setting, wherein we develop an algorithm for designing experiments to efficiently reduce the variance of our state estimate. We provide several examples from heat conduction, the convection-diffusion equation, and the reduced wave equation, all of which that demonstrate the performance of the proposed methodology.

∗MIT Department of Aeronautics and Astronautics, 77 Massachusetts Ave., Cambridge, MA 02139, USA, email: cuongng@mit.edu
†MIT Department of Aeronautics and Astronautics, 77 Massachusetts Ave., Cambridge, MA 02139, USA, email: abbymen@mit.edu
‡MIT Sloan School of Management, 77 Massachusetts Ave., Cambridge, MA 02139, USA, email: rfreund@mit.edu
§MIT Department of Aeronautics and Astronautics, 77 Massachusetts Ave., Cambridge, MA 02139, USA, email: peraire@mit.edu
1 Introduction

Partial differential equations (PDEs) are commonly used to model a wide variety of physical processes such as heat transfer, fluid flows, electromagnetism, and structure deformations. A PDE model of a physical problem is typically described by conservation laws, constitutive laws, material properties, boundary conditions, boundary data, and geometry. In practical applications, however, the PDE model is only an approximation to the real physical problem due to (i) the deliberate mathematical simplification of the model to keep it tractable (by ignoring certain physics or certain boundary conditions that pose computational difficulties), and (ii) the uncertainty of the available data (by using geometry, material properties, and boundary data that are different from those of the physical problem due to noise, lack of measurements, etc.). These are referred to as model uncertainties to distinguish them from numerical error arising from the discretization of the governing PDE by a numerical method. In the presence of model uncertainties, the PDE model may not produce a good approximation of the true state of the underlying physical problem.

There are a number of different approaches for dealing with model uncertainties. One approach is to represent physical inputs as random parameters or random fields, thereby resulting in stochastic PDEs. There exist several numerical methods for solving stochastic PDEs such as Monte Carlo (MC) methods [11, 16], intrusive polynomial chaos [13, 39], non-intrusive polynomial chaos [38, 9, 18], stochastic collocation [38, 2, 28], response surface [4, 14], and Kriging [5, 23, 20, 40, 8].

Another approach to deal with model uncertainties is to determine the unknown inputs by matching the outputs to the observations. The observations may come from experimental data or higher-fidelity models. This approach is known as parameter estimation or model calibration. There exist several parameter estimation methods including least-square methods [6, 21, 37], maximum likelihood estimation [1], and Bayesian inference [12, 33, 35].

Both stochastic PDE modeling and parameter estimation require prior knowledge about the unknown inputs such as their origin (i.e., physical parameters, boundary conditions, source term, or geometry), their probability space as well as their probability density function. In practice, it can be difficult to obtain such prior information precisely. If such prior information are not accurate, they may lead to erroneous prediction of the true state.

Data assimilation is an alternative approach that combines different sources of data to yield an estimate of the state of a physical system. There are two different data assimilation methods, namely, stochastic data assimilation and variational data assimilation. In stochastic data assimilation [19, 24, 34], the state estimate is represented as a stochastic process and is determined by minimizing its variance. In variational data assimilation [21, 22, 41], the state estimate is defined as an optimal solution of a least-squares minimization principle.

Another approach is data interpolation [3, 10, 25, 26, 30, 36] which involves computing a collection of solutions (snapshots) of a parametrized or time-varying mathematical model and reconstructing the physical state by fitting the observation to the snapshots.

In this paper, we take a different approach recently introduced in [27], that requires less prior information and exploits the observations to improve the state prediction. First, we augment the PDE model with a random Gaussian functional which serves to represent various
sources of uncertainty in the model. This gives rise to a stochastic PDE model whose solution is characterized by the Gaussian functional. We next derive a linear regression model for the Gaussian functional by utilizing observations and adjoint states. We then apply the principle of Gaussian process (GP) regression [29, 31, 32] to determine the posterior distribution of the Gaussian functional. This allows us to compute the posterior distribution for our estimate of the unknown state and provide a statistical measure of the prediction error. A crucial ingredient in our method is the covariance operator representing the prior distribution of the Gaussian functional. The bilinear covariance operators considered incorporate a number of free parameters (the so-called hyperparameters) that are determined upon the observations by maximizing a marginal likelihood. Furthermore, we consider the problem of experimental design in this setting, wherein we develop an algorithm for designing experiments to efficiently reduce the variance of our state estimate.

We can relate our approach to existing data assimilation methods. Our approach seeks to characterize model uncertainties by introducing a Gaussian functional into the existing numerical model, while stochastic data assimilation directly represents the state estimate as a stochastic process and infer its posterior distribution from the observations, the model outputs, and the priors about the background state. We show in the Appendix B that our method and the Kalman method [19] yield exactly the same posterior distribution for a judicious choice of the priors. Our approach can also be shown to yield a posterior mean that satisfies the least-square minimization principle of 3D variational data assimilation [7, 22].

The proposed approach has the following features which can distinguish itself from the aforementioned approaches. First, it incorporates both the PDE model and the observations into the regression procedure. Second, it can handle the observations given in the form of linear functionals of the field variable. Third, the approach is non-parametric in the sense that it provides a systematic way to determine both the priors and the posteriors based on the observations. Fourth, it yields not only the mean estimate but also the posterior variance that quantifies the prediction error. And fifth, it provides a natural mechanism for designing experiments so as to reduce the posterior variance.

There are a number of new contributions relative to our previous work [27]. Herein we propose an efficient greedy algorithm to rationally select good observations from a large number of possible experiments. We show how our method can be related to Bayesian inference, 3D variational data assimilation, and the Kalman method. Finally, we apply our approach to a variety of PDEs including convection-diffusion equations and Helmholtz equations to demonstrate its performance.

The paper is organized as follows. In Section 2, we present a description of the problem of interest. In Section 3, we present the Gaussian functional regression method for state estimation using linear PDE models and observations. In Section 4, we describe a greedy algorithm for selecting the observations. In Section 5, we demonstrate our approach on several examples from heat conduction, convection-diffusion equation, and reduced wave equation. In Section 6, we conclude the paper with some remarks on future research. Finally, in the Appendix, we provide a detailed picture of our method in algebraic form and reveal its connection to 3D variational data assimilation and the Kalman method.


2 Problem Formulation

Let $\Omega$ denote a bounded open domain with Lipschitz boundary. Let $V$ be an appropriate finite element approximation space, that is defined on a triangulation of the domain $\Omega$. The Galerkin finite element formulation of a general linear PDE model can be stated as follows. We seek a solution $u^o \in V$ and an output vector $s^o \in \mathbb{C}^M$ such that

$$a(u^o, v) = \ell(v), \quad \forall v \in V,$$

$$s^o_i = c_i(u^o), \quad i = 1, \ldots, M. \quad (1a)$$

Here $a : V \times V \to \mathbb{C}$ is a continuous bilinear form, $\ell : V \to \mathbb{C}$ is a continuous and bounded linear functional, and $c_i : V \to \mathbb{C}, i = 1, \ldots, M$ are continuous and bounded linear functionals. The finite element approximation space $V$ is of $N$ dimensions. We shall take $N$ to be so large that the numerical solution $u$ is indistinguishable from the exact solution of the PDE model at any accuracy level of interest, i.e., we assume that the numerical error due to the finite element approximation is insignificant.

We assume that the underlying finite element (FE) model (1) is used to predict the true state of a deterministic and time-independent physical problem. We denote the true state by $u^{\text{true}}$ and the associated true output vector by $s^{\text{true}} \in \mathbb{C}^M$. Both $u^{\text{true}}$ and $s^{\text{true}}$ are not known. However, we assume that we are given a vector of $M$ observations $d \in \mathbb{C}^M$, which are the measurements of the true output vector $s^{\text{true}}$. We further assume that the observation vector differs from the true output vector $s^{\text{true}}$ by additive Gaussian noise, namely,

$$d = s^{\text{true}} + \varepsilon,$$  \hspace{1cm} (2)

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_M) \in \mathbb{C}^M$ are independent, identically distributed Gaussian random variables with zero mean and variance $\sigma^2$.

As mentioned in the Introduction, the solution $u^o$ of the FE model (1) may not be a good approximation to the true state $u^{\text{true}}$ because of various sources of uncertainty arising from the imprecise knowledge of boundary conditions, geometry, physical parameters, and governing equations. The uncertainty of our numerical model in predicting the true physical state prompts basic research questions such as:

- How do we combine the FE model (1) and the observations (2) to yield a more accurate prediction of the true state?
- How do we quantify the error in the prediction?, and
- How do we design experiments to improve the accuracy of the prediction?

In the remainder of this paper, we concentrate on addressing these questions. We address the first two questions by introducing a so-called Gaussian functional regression (GFR) method as described in the next section. We address with the third question by devising an experimental design procedure in Section 4.
3 Gaussian Functional Regression

3.1 Gaussian functional

We suppose that we are given a bounded linear functional \( g : V \to \mathbb{C} \), and we consider the following model: determine \( u \in V \) and \( s \in \mathbb{C}^M \) such that

\[
a(u, v) + g(v) = \ell(v), \quad \forall v \in V, \tag{3a}
\]

\[
s_i = c_i(u), \quad i = 1, \ldots, M. \tag{3b}
\]

Notice that the new model (3) differs from the original model (1) by the addition of the functional \( g \). Here we can exactly determine \( u \) and \( s \) only if \( g \) is known. Indeed, together with \( a(u, v) \) and \( \ell(v) \), the functional \( g \) determines the solution \( u \) and the output vector \( s \) of the model (3). Of course, if \( g(v) = \ell(v) - a(u^{\mathrm{true}}, v) \) then \( u = u^{\mathrm{true}} \) is the solution of (3). Unfortunately, this particular choice of \( g \) requires foreknowledge of the true state \( u^{\mathrm{true}} \), which we presumably do not know and indeed we seek to accurately predict.

In order to capture various sources of uncertainty in the original model (1), we hypothesize the linear functional \( g \) as a Gaussian process [31]. Specifically, the Gaussian functional \( g \) is assumed to have zero mean and covariance operator \( k \), namely:

\[
g(v) \sim \mathcal{GP}(0, k(w, v)), \quad \forall w, v \in V. \tag{4}
\]

We require that the covariance operator \( k : V \times V \to \mathbb{C} \) is symmetric positive-definite, namely:

\[
k(w, v) = k(v, w), \quad \text{and} \quad k(v, v) > 0 \quad \forall v \neq 0, \quad \text{and} \quad k(0, 0) = 0. \tag{5}
\]

As the covariance operator \( k \) characterizes the space of all possible functionals prior to taking into account the observations, it plays an important role in our method. The selection of a covariance operator will be discussed later.

Since the functional \( g \) is Gaussian, the model (3) becomes a stochastic system. In order to solve the stochastic system (3), we need to compute the posterior distribution of \( g \) on the basis of the observation vector \( d \).

3.2 Functional regression model

Here we develop a linear regression model for the Gaussian functional. To this end, we first obtain the adjoint solutions \( \phi_i \in V, i = 1, \ldots, M \) by solving the adjoint problems:

\[
a(v, \phi_i) = -c_i(v), \quad \forall v \in V. \tag{6}
\]

It follows from (1), (3), and (6) that

\[
g(\phi_i) = \ell(\phi_i) - a(u, \phi_i) = a(u^o, \phi_i) + c_i(u) = c_i(u) - c_i(u^o) = s_i - s_i^o, \tag{7}
\]

for \( i = 1, \ldots, M \). Moreover, we assume that the observation \( d_i \) differs from the output estimate \( s_i \) by Gaussian noise \( \varepsilon_i \sim \mathcal{N}(0, \sigma^2) \), namely,

\[
s_i = d_i - \varepsilon_i, \quad i = 1, \ldots, M. \tag{8}
\]
This equation is analogous to (2) which relates the observed data \( d \) to the true outputs \( s^\text{true} \). We substitute (8) into (7) to obtain

\[
d_i - s_i^\theta = g(\phi_i) + \varepsilon_i, \quad i = 1, \ldots, M.
\]

This equation can be viewed as a standard regression model, namely a linear model with additive Gaussian noise, and we shall now show that the model (9) allows us to determine the posterior distribution of \( g(v) \) for any given test function \( v \in V \) as follows.

Let \( \Phi = [\phi_1, \ldots, \phi_M] \) be a collection of \( M \) adjoint states as determined by (6). Let \( \Psi = [\psi_1 \in V, \ldots, \psi_N \in V] \) be a collection of \( N \) test functions, where \( \psi_j, 1 \leq j \leq N, \) are basis functions of the space \( V \). We would like to compute the posterior distribution of \( g \in \mathbb{C}^N \) with \( g_i = g(\psi_i), i = 1, \ldots, N \). According to the prior (4) and the regression model (9), the joint distribution of \( (d - s^\theta) \) and \( g \) is given by

\[
\begin{bmatrix}
d - s^\theta \\
g
\end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K(\Phi, \Phi) + \sigma^2 I & K(\Phi, \Psi) \\ K(\Psi, \Phi) & K(\Psi, \Psi) \end{bmatrix} \right),
\]

where \( K(\Phi, \Phi) \in \mathbb{C}^{M \times M}, K(\Phi, \Psi) \in \mathbb{C}^{M \times N}, K(\Psi, \Phi) \in \mathbb{C}^{N \times M}, \) and \( K(\Psi, \Psi) \in \mathbb{C}^{N \times N} \) have entries

\[
K_{ij}(\Phi, \Phi) = k(\phi_i, \phi_j), \quad i = 1, \ldots, M, j = 1, \ldots, M,
K_{ij}(\Phi, \Psi) = k(\phi_i, \psi_j), \quad i = 1, \ldots, M, j = 1, \ldots, N,
K_{ij}(\Psi, \Phi) = k(\psi_i, \phi_j), \quad i = 1, \ldots, N, j = 1, \ldots, M,
K_{ij}(\Psi, \Psi) = k(\psi_i, \psi_j), \quad i = 1, \ldots, N, j = 1, \ldots, N,
\]

respectively. It thus follows from (10) and the conditional normal distribution formula (see Appendix A.2 in Rasmussen and Williams [31]) that the posterior distribution of \( g \) is

\[
p(g | (d - s^\theta), \Phi, \Psi) \sim \mathcal{N}(\bar{g}, G),
\]

where the mean vector \( \bar{g} \in \mathbb{C}^N \) and the covariance matrix \( G \in \mathbb{C}^{N \times N} \) are given by

\[
\bar{g} = K(\Psi, \Phi) D^{-1} (d - s^\theta), \quad G = K(\Psi, \Psi) - K(\Psi, \Phi) D^{-1} K(\Phi, \Psi),
\]

with \( D = K(\Phi, \Phi) + \sigma^2 I \). The posterior mean and covariance (13) require the adjoint states (6), the inner products (11), and the inverse of \( D \). Since \( N \) is typically much larger than \( M \), the computational cost is dominated by the cost of computing the adjoint states.

We note that the posterior mean \( \bar{g} \) depends linearly on \( (d - s^\theta) \), which is the difference between the observation vector and the output vector of the original model (1). Another way to look at the posterior mean is to view it as a linear combination of \( M \) inner products, each one associated with one of the adjoint states, by writing

\[
\bar{g}_i = \sum_{j=1}^{M} \beta_j k(\psi_i, \phi_j), \quad i = 1, \ldots, N,
\]

where \( \beta = D^{-1} (d - s^\theta) \). We see that the discrepancy between the observed data and the original model (1) has an important effect on the posterior mean.
Also note that the posterior covariance $G$ in (13) does not explicitly depend on the values of the observed outputs, but only on the adjoint states $\Phi$ and the covariance operator $k(\cdot, \cdot)$. This is based on the premise that the covariance operator $k(\cdot, \cdot)$ is somehow known and given. In reality, of course, we create/invent $k(\cdot, \cdot)$ based on notions both of general principles as well as computational expedients. As will be seen in Section 3.4, we will construct/model a suitable covariance operator $k(\cdot, \cdot)$ based on data observations and maximum likelihood estimation. In this way the posterior covariance $G$ in (13) will implicitly depend on the observed outputs and our modeling assumptions.

Last of all, notice that the covariance $G$ in (13) is the difference between two terms: the first term $K(\Psi, \Psi)$ is simply the prior covariance, while the second term represents a reduction in the covariance due to the information that the observations furnish about the Gaussian functional.

### 3.3 Posterior distribution of the state estimate

It remains to compute the posterior distribution of the state estimate $u(x)$ for any spatial point $x \in \Omega$. To this end, we express the state estimate $u(x)$ as a linear combination of the basis functions $\psi_n(x), 1 \leq n \leq N$, as

$$u(x) = \sum_{n=1}^{N} u_n \psi_n(x),$$

where $u = (u_1, \ldots, u_N) \in \mathbb{C}^N$ denotes the state vector. Substituting (15) into (3) and choosing $v = \psi_i$ for $i = 1, \ldots, N$, we obtain the following linear system

$$Au = l - g,$$

where $A$, $l$, and $g$ have entries $A_{ij} = a(\psi_i, \psi_j), l_i = \ell(\psi_i)$, and $g_i = g(\psi_i)$ for $i, j = 1, \ldots, N$, respectively. It follows from (12) and (16) that the state vector $u$ obeys a normal distribution

$$u \sim \mathcal{N}(\bar{u}, U),$$

where the posterior mean vector $\bar{u} \in \mathbb{C}^N$ and the posterior covariance matrix $U \in \mathbb{C}^{N \times N}$ are given by

$$\bar{u} = A^{-1}(l - g), \quad U = A^{-1}GA^{-H},$$

where the superscript $H$ denotes the complex conjugate transpose. We can also write $\bar{u}$ as

$$\bar{u} = u^o - u^g,$$

where $u^o = A^{-1}l$ is the state vector of the original model (3) and $u^g = A^{-1}g$ is obtained by using the Gaussian functional regression model of Section 3.2. Similarly, we can write $U$ as

$$U = U^o - U^g,$$
where \( U^o = A^{-1}K(\Psi, \Psi)A^{-H} \) is the prior covariance matrix for the state estimate \( u \) and \( U^g = A^{-1}K(\Psi, \Phi)D^{-1}K(\Phi, \Psi)A^{-H} \) is the reduction in the covariance due to the observations. Note that the matrix \( U^o \) has full rank \( N \), whereas the rank of the matrix \( U^g \) is at most \( M \).

In the Appendix A, we present an alternative way to obtain the posterior distribution of the state vector \( u \) by applying Bayesian inference to the matrix system of the finite element weak formulation. Indeed, the Bayesian approach presented in the Appendix A and the Gaussian function regression described in this section yield the exact same result. The Appendix A specifically targeted to readers who are familiar with Bayesian analysis. The results in the Appendix A will be particularly useful in the development of the experimental design methodology in Section 4.

Let us now use the notation \( \psi(x) = (\psi_1(x), \ldots, \psi_N(x)) \in \mathbb{C}^N \) to denote a vector of values of the \( N \) basis functions at \( x \). We obtain from (15) and (17) that

\[
u(x) \sim N(\bar{\nu}(x), \eta(x)), \tag{21}\]

where the mean value function \( \bar{\nu}(x) \) and the variance function \( \eta(x) \) are given by

\[
\bar{\nu}(x) = (\psi(x))^H\bar{\nu}, \quad \eta(x) = (\psi(x))^H U \psi(x). \tag{22}\]

The mean value function represents the prediction of the true state \( u^{\text{true}}(x) \), while the variance function is a probabilistic measure of the prediction error. Note from (19) that the mean value function presented above in (22) is the difference between the original state function \( u^o(x) = (\psi(x))^H u^o \) and the function \( u^g(x) = (\psi(x))^H u^g \).

Not surprisingly, there is a connection between the mean value function (22) and a minimization principle. Let \( \bar{q}(x) = \sum_{m=1}^M \beta_m \phi_m(x) \), where \( \beta \) is the coefficient vector as in (14). Then the mean value function \( \bar{\nu}(x) \) is a solution of the following least-squares minimization problem:

\[
(\bar{\nu}(x), \bar{q}(x), \beta) = \arg \min_{(\nu, w, \gamma) \in V \times V \times \mathbb{C}^M} \frac{1}{2} k(w, w) + \frac{1}{2} \sigma^2 \gamma^H \gamma
\]

s.t. \( a(z, v) + k(w, v) = \ell(v), \quad \forall v \in V \),

\[
c_m(z) + \sigma^2 \gamma_m = d_m, \quad m = 1, \ldots, M . \tag{23}\]

A proof of the above optimization statement is given in the Appendix A. This result shows a connection between Gaussian functional regression and standard least-squares regression. In particular, the posterior mean state \( \bar{\nu}(x) \) is the optimal solution of a least-squares minimization problem. This is hardly a surprise as the posterior mean state is also the maximum a posteriori (MAP) estimate of the linear functional regression model in the Bayesian framework discussed in the Appendix. It is well known that the MAP estimate coincides with the least-squares solution. One advantage of the development herein over standard least-squares regression is that our analysis yields the posterior covariance state \( \eta(x) \) (22) in addition to the posterior mean state \( \bar{\nu}(x) \). Indeed, this will then allow us to use a maximum likelihood principle to choose/set the prior covariance operator \( k(\cdot, \cdot) \) based on data observations, whereas the standard least-squares regression does not provide any such mechanism. This is discussed in detail in Section 3.4 below.
3.4 Covariance operator

In order to determine the posterior distribution of the state estimate, the covariance operator \( k(\cdot, \cdot) \) needs to be specified. Indeed, this is a crucial aspect of our approach because it affects both the mean value function and the posterior variance function. We propose a class of bilinear covariance operators of the form:

\[
k(w, v; \theta) = \sum_{j=1}^{J} \theta_j k_j(w, v),
\]

where \( \theta = (\theta_1, \theta_2, \ldots, \theta_J) \) is a vector of free parameters (the so-called hyperparameters) and \( k_j(\cdot, \cdot), 1 \leq j \leq J \) are some given symmetric bilinear forms. Of special interest to us are those covariance operators of the form

\[
k(w, v; \theta) = \theta_1 \int_{\Omega} wvdx + \theta_2 \int_{\Omega} \nabla w \cdot \nabla vdx,
\]

which represents the \( H^1(\Omega) \) inner product for second-order linear PDEs.

Given the class of covariance operators (24), it remains to select the hyperparameter vector \( \theta \). We will do so guided by a maximum likelihood principle as follows. Note from (10) that the posterior distribution of \((d - s^0)\), conditional on \( \theta \), obeys the normal distribution:

\[
(d - s^0)|\theta \sim N(0, D(\theta)),
\]

where the matrix \( D(\theta) \) has entries

\[
D_{mq}(\theta) = \sum_{j=1}^{J} \theta_j k_j(\phi_m, \phi_q) + \sigma^2 \Delta_{mq}, \quad m, q = 1, \ldots, M,
\]

where \( \Delta_{mq} = 1 \) if \( m = q \) and \( \Delta_{mq} = 0 \) if \( m \neq q \). Let \( p((d - s^0)|\theta) \) denote the probability density function of (26), and let us choose the hyperparameter vector \( \theta \) guided by the maximum likelihood principle, which by monotonicity is equivalent to log-maximum likelihood. The log likelihood function is given by:

\[
\log p((d - s^0)|\theta) = -\frac{1}{2}(d - s^0)^T D(\theta)^{-1}(d - s^0) - \frac{1}{2} \log(\det(D(\theta))) - \frac{M}{2} \log(2\pi).
\]

We therefore propose to select \( \theta \) by seeking an approximate optimal solution to the log-likelihood optimization problem, yielding:

\[
\theta \approx \arg \max_{\theta' \in \mathbb{R}^J} \log p((d - s^0)|\theta').
\]

The optimization problem (29) does not have a closed form solution, and it is generically computationally challenging due to lack of convexity. We therefore seek approximate solutions of (29) in cases of interest. For example, when the covariance operator is given by (25), the hyperparameter space is the 2-dimensional plane, and simple grid generation/enumeration and function evaluation will be sufficient for our purposes.

After we determine the hyperparameter vector \( \theta \), we then compute the posterior distribution of the state estimate. Therefore our approach is non-parametric in the sense that the observations are used to determine both the prior distribution and the posterior distribution.
4 Experimental Design

4.1 Optimizing experimental design

In this section we consider the problem of designing experiments to obtain the observations. We assume that we are given a large set of \( L \) potential observation functionals \( b_i(\cdot), 1 \leq i \leq L \), each of which corresponds to a potential experiment to be carried out. Before actually performing any of these potential experiments, we face the following decision questions. Which experiment among the \( L \) potential experiments should be implemented first? Given the results of any previous experiments, which experiment should be chosen next? How many experiments are enough to provide an accurate prediction of the true state? These decision questions can be summed up as follows: how do we optimally choose a sequence of the actualized observation functionals \( c_m(\cdot), 1 \leq m \leq M \), among the potential observation functionals \( b_i(\cdot), 1 \leq i \leq L \)?

It will be more convenient to state our experimental design problem in the matrix form instead of in the weak form. The expressions involving \( K(\Psi, \Psi), K(\Phi, \Phi), K(\Phi, \Psi) \), etc., can be rather unwieldy, so we introduce a compact form of the notation by setting \( K = K(\Psi, \Psi) \). Moreover, if we let \( C \in \mathbb{C}^{N \times M} \) be a matrix with entries \( C_{nm} = c_m(\psi_n) \) for \( n = 1, \ldots, N \) and \( m = 1, \ldots, M \), then we have \( K(\Phi, \Phi) = -K A^{-H} C, K(\Phi, \Psi) = -C^H A^{-1} K \), and \( K(\Phi, \Psi) = C^H A^{-1} K A^{-H} C \). It follows that the posterior covariance matrix given in (18) can be written as

\[
U = A^{-1} (K - K A^{-H} C (\sigma^2 I + C^H A^{-1} K A^{-H} C)^{-1} C^H A^{-1} K) A^{-H}. \tag{30}
\]

(This expression coincides exactly with the expression (75) in the Appendix, which is obtained there by using the Bayesian approach.) We can also rewrite the posterior covariance matrix \( U \) as

\[
U(C) = U_0 - U_0 C (\sigma^2 I + C^H U_0 C)^{-1} C^H U_0,
\]

where we recall that \( U_0 = A^{-1} K A^{-H} \) is the prior covariance matrix. We see that the posterior covariance matrix depends only on the noise variance \( \sigma^2 \), the prior covariance matrix \( U_0 \), and the observation matrix \( C \).

We now let \( B \in \mathbb{C}^{N \times L} \) be a matrix with entries \( B_{nl} = b_l(\psi_n) \) for \( n = 1, \ldots, N \) and \( l = 1, \ldots, L \). Let \( M \) denote the number of experiments that we will implement, and consider the following set of matrices induced by the matrix \( B \):

\[
\mathcal{M}_M(B) := \{ E \in \mathbb{C}^{N \times M} : \text{each column of } E \text{ is chosen from among the columns of } B \}. \tag{32}
\]

In the experimental design problem, we seek an experiment matrix \( C \in \mathcal{M}_M(B) \) that minimizes the “size” of \( U(C) \), where size is measured in various ways in various design formulations such as the product of the eigenvalues of \( U(C) \), the largest eigenvalue of \( U(C) \), the sum of the eigenvalues of \( U(C) \), etc. In the interest of generality, let \( h(\cdot) : \mathbb{R}^N \to \mathbb{R} \) be a given monotone function of the ordered eigenvalues \( \lambda(U) := (\lambda_1(U), \ldots, \lambda_N(U)) \) of \( U \), and consider the following optimization problem:

\[
C^* = \arg \min_{E \in \mathcal{M}_M(B)} h(\lambda(U(E))), \tag{33}
\]

\[10\]
where \( U(E) \) is the posterior covariance matrix as given in (31) for \( E = C \). In other words, the optimal experiment design \( C^* \) is the matrix that minimizes a particular given monotone function of the eigenvalues of the posterior covariance matrix. Herein we develop a procedure for computing improved solutions of (33) that simultaneously applies to any (and hence every) monotone function \( h(\cdot) \) of the eigenvalues of \( U(E) \).

Note that the optimization problem (33) has a nonlinear (and typically non-convex) objective function, and that the set of feasible solutions is an exponentially large discrete set. Furthermore, the optimal observation matrix \( C^* \) is not hierarchical in the sense that increasing \( M \) might completely change the previously selected columns of \( C^* \). Because the optimal set of experiments might be very different for different values of \( M \), one cannot compute the exact optimum by inductively amending the previous optimal set of experiments. For all of the above reasons it therefore is likely to be too computationally expensive to compute an exact optimum of (33).

In Section 4.2 below we develop a greedy iterative procedure for choosing the experiments based on the eigenvector of the largest eigenvalue of \( U(C) \).

### 4.2 A greedy algorithm for choosing the experiments

Here we develop a greedy algorithm for choosing the experiments. We first consider the case where there is no noise in the observation vector \( d \), i.e., \( \sigma = 0 \) for the Gaussian noise. With \( \sigma = 0 \) it follows that \( u \) must satisfy \( C^H u = d \), so let us define:

\[
T := \{ u : C^H u = d \} \quad \text{and} \quad S := \{ u : C^H u = 0 \};
\]

here \( T \) is the affine set containing the possible values of \( u \) and \( S \) is the linear subspace parallel to \( T \). Let us assume without loss of generality that the columns of \( C \) are linearly independent, whereby \( \text{dim}(T) = \text{dim}(S) = N - M \). Recalling \( \bar{u} = A^{-1}(l - \bar{g}) \) from (18), it holds that \( \bar{u} \in T \) whereby \( S = T - \bar{u} \). The posterior covariance matrix (31) can be written as:

\[
U = U_0 - U_0 C (C^H U_0 C)^{-1} C^H U_0, \tag{34}
\]

where recall that \( U_0 = A^{-1} K A^{-H} \) is the prior covariance matrix. Let us also write the eigendecomposition of \( U \) as \( U = Q \Lambda Q^H \) for a suitable orthonormal matrix of eigenvectors \( Q \) and diagonal matrix \( \Lambda \) of corresponding eigenvalues. Since \( U \) is positive semi-definite, i.e., \( u^H U u \geq 0 \) for all \( u \), all eigenvalues of \( U \) are nonnegative, and we can partition the eigenvalues into those that are zero and those that are positive, and likewise partition the corresponding eigenvectors, and write:

\[
U = Q \Lambda Q^H = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \Lambda^+ \end{bmatrix} \begin{bmatrix} Q_1^H \\ Q_2^H \end{bmatrix} = Q_2 \Lambda^+ Q_2^H, \tag{35}
\]

where the columns of \( Q_1 \) are eigenvectors corresponding to eigenvalues with value 0 and the columns of \( Q_2 \) are eigenvectors corresponding to positive eigenvalues, and the diagonal matrix \( \Lambda^+ \) has positive diagonal entries. In a slight change of notation, let us denote these
entries as $\Lambda_i^+ > 0$ for $i = 1, \ldots, N - M$, and there is no loss of generality in assuming that $0 < \Lambda_1^+ \leq \Lambda_2^+ \leq \cdots \leq \Lambda_{N-M}^+$. It is straightforward to demonstrate that
\[
\begin{align*}
\mathbf{u}^H \mathbf{U} \mathbf{u} = 0 & \iff \mathbf{U} \mathbf{u} = 0 \iff \mathbf{u} \in \text{range}(\mathbf{C}) \iff \mathbf{u} \in \text{null}(\mathbf{Q}_2^H) . \quad (36)
\end{align*}
\]

For the given observation vector $\mathbf{d}$, $\mathbf{u}$ is restricted to lie in the affine set $\mathcal{T}$, and we write $\mathbf{u} | \mathbf{d} \sim N(\mathbf{u}, \mathbf{U})$ to designate that the posterior distribution of $\mathbf{u}$ is normally distributed on $\mathcal{T}$ with mean $\bar{\mathbf{u}}$ and covariance matrix $\mathbf{U}$. Even though $\mathbf{U}$ has no inverse, the probability density function of $\mathbf{u}$ on $\mathcal{T}$, which we denote as $p_{\mathcal{T}}(\mathbf{u}|\mathbf{d})$, is given by:
\[
p_{\mathcal{T}}(\mathbf{u}|\mathbf{d}) = \frac{1}{\pi^{(N-M)/2} \sqrt{\prod_{i=1}^{N-M} \Lambda_i^+}} e^{-\frac{1}{2} (\mathbf{u} - \bar{\mathbf{u}})^H \mathbf{Q}_2 (\Lambda^+)^{-1} \mathbf{Q}_2^H (\mathbf{u} - \bar{\mathbf{u}})} . \quad (37)
\]
Notice the precise way that the positive eigenvalues appear in (37).

Now let us suppose that we have already chosen $M$ experiments from among the $L$ experiments $b_i(\cdot)$, $i = 1, \ldots, L$, and let $\mathbf{C}$ denote the experiment matrix, and let us consider adding an additional experiment, which would be the $(M+1)^{\text{st}}$ experiment. Let this experiment be denoted as $\tilde{b}(\cdot)$, where $\tilde{b}(\cdot)$ is one of the $L$ experiments $b_i(\cdot)$, $i = 1, \ldots, L$. In the matrix form, let $\tilde{\mathbf{b}}$ be the vector with entries $\tilde{b}(\psi_n)$ for $n = 1, \ldots, N$. We will assume that $\tilde{\mathbf{b}} \notin \text{range}(\mathbf{C})$, so that the new/updated experiment matrix $\tilde{\mathbf{C}} := [\mathbf{C} \hspace{1em} \tilde{\mathbf{b}}]$ has linearly independent columns. Then letting $\tilde{\mathbf{U}} := \mathbf{U}(\tilde{\mathbf{C}})$, we have from (34) that
\[
\tilde{\mathbf{U}} = \mathbf{U}_0 - \mathbf{U}_0 \tilde{\mathbf{C}} (\tilde{\mathbf{C}}^H \mathbf{U}_0 \tilde{\mathbf{C}})^{-1} \tilde{\mathbf{C}}^H \mathbf{U}_0 . \quad (38)
\]
It turns out that $\tilde{\mathbf{U}}$ simplifies to:
\[
\tilde{\mathbf{U}} = \mathbf{U} - \left( \frac{1}{\tilde{\mathbf{b}}^H \mathbf{U} \tilde{\mathbf{b}}} \right) \mathbf{U} \tilde{\mathbf{b}} \tilde{\mathbf{b}}^H \mathbf{U} . \quad (39)
\]
(Equation (39) is derived by straightforward, but tedious, substitution using
\[
(\tilde{\mathbf{C}}^H \mathbf{U}_0 \tilde{\mathbf{C}})^{-1} = \begin{bmatrix} \mathbf{C}^H \mathbf{U}_0 \mathbf{C} & \mathbf{C}^H \mathbf{U}_0 \tilde{\mathbf{b}} \\ \tilde{\mathbf{b}}^H \mathbf{U}_0 \mathbf{C} & \tilde{\mathbf{b}}^H \mathbf{U}_0 \tilde{\mathbf{b}} \end{bmatrix}^{-1}
= \begin{bmatrix} (\mathbf{C}^H \mathbf{U}_0 \mathbf{C})^{-1} - \frac{\mathbf{C}^H \mathbf{U}_0 \mathbf{b} \tilde{\mathbf{b}}^H \mathbf{U}_0 \mathbf{C} (\mathbf{C}^H \mathbf{U}_0 \mathbf{C})^{-1}}{\alpha} & \frac{(\mathbf{C}^H \mathbf{U}_0 \mathbf{C})^{-1} \mathbf{C}^H \mathbf{U}_0 \tilde{\mathbf{b}}}{\alpha} \\ \frac{\tilde{\mathbf{b}}^H \mathbf{U}_0 \mathbf{C} (\mathbf{C}^H \mathbf{U}_0 \mathbf{C})^{-1}}{\alpha} & \frac{1}{\alpha} \end{bmatrix},
\]
where $\alpha := \tilde{\mathbf{b}}^H \mathbf{U} \tilde{\mathbf{b}}$, and using the block form of the inverse above in (38) and simplifying terms. Note that the assumption that $\tilde{\mathbf{b}} \notin \text{range}(\mathbf{C})$ implies via (36) that $\alpha = \tilde{\mathbf{b}}^H \mathbf{U} \tilde{\mathbf{b}} \neq 0$ so the above objects are all well-defined.)

Let us now compare the ordered eigenvalues of $\mathbf{U}$ and $\tilde{\mathbf{U}}$, which we denote, respectively, as the arrays
\[
\begin{align*}
\lambda & := (\lambda_1 \hspace{1em} \lambda_2 \hspace{1em} \cdots \hspace{1em} \lambda_N ) \\
\tilde{\lambda} & := (\tilde{\lambda}_1 \hspace{1em} \tilde{\lambda}_2 \hspace{1em} \cdots \hspace{1em} \tilde{\lambda}_N ).
\end{align*}
\]
From the linear independence assumption regarding the columns of \( \mathbf{C} \) and \( \tilde{\mathbf{C}} \) it follows using (36) that \( \lambda_1 = \cdots = \lambda_M = 0 \) and \( \lambda_{M+1} > 0 \), and also that \( \tilde{\lambda}_1 = \cdots = \tilde{\lambda}_{M+1} = 0 \) and \( \tilde{\lambda}_{M+2} > 0 \). Similar to (35), we will write

\[
\tilde{\mathbf{U}} = \begin{bmatrix} \tilde{\mathbf{Q}}_1 & \tilde{\mathbf{Q}}_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \Lambda^+ \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{Q}}_1^H \\ \tilde{\mathbf{Q}}_2^H \end{bmatrix} = \tilde{\mathbf{Q}}_2 \Lambda^+ \tilde{\mathbf{Q}}_2^H \, , \quad (40)
\]

where the columns of \( \tilde{\mathbf{Q}}_1 \) are eigenvectors corresponding to eigenvalues with value 0 and the columns of \( \tilde{\mathbf{Q}}_2 \) are eigenvectors corresponding to positive eigenvalues, and the diagonal matrix \( \tilde{\Lambda}^+ \) has positive diagonal entries. As with (40), we denote these entries as \( \tilde{\Lambda}^+_i > 0 \) for \( i = 1, \ldots, N - M - 1 \), and there is no loss of generality in assuming that \( 0 < \tilde{\Lambda}^+_1 \leq \tilde{\Lambda}^+_2 \leq \cdots \leq \tilde{\Lambda}^+_{N-M-1} \). Therefore we have:

\[
\tilde{\lambda}_1 \leq \lambda_1 \leq \tilde{\lambda}_2 \leq \lambda_2 \leq \cdots \leq \lambda_{N-1} \leq \lambda_N \leq \lambda_N \, . \quad (41)
\]

It therefore follows that:

\[
\Lambda^+_1 \leq \tilde{\Lambda}^+_1 \leq \Lambda^+_2 \leq \tilde{\Lambda}^+_2 \leq \cdots \leq \Lambda^+_{N-M-1} \leq \tilde{\Lambda}^+_{N-M-1} \leq \Lambda^+_{N-M} \, .
\]

This chain of inequalities is significant in that it provides lower bounds on the values of the updated eigenvalues \( \tilde{\Lambda}^+_i \) for \( i = 1, \ldots, N - M - 1 \):

\[
\tilde{\Lambda}^+_i \geq \Lambda^+_i \, , \text{for } i = 1, \ldots, N - M - 1 \, . \quad (42)
\]

Note that this result is true for any newly added experiment \( \tilde{\mathbf{b}} \) satisfying the condition that \( \tilde{\mathbf{C}} := [\mathbf{C} \; \tilde{\mathbf{b}}] \) has linearly independent columns.

Let us now consider if there might be an experiment \( \tilde{\mathbf{b}} \) for which (42) holds at equality for all \( i = 1, \ldots, N - M - 1 \), thus simultaneously providing the lowest value of the new eigenvalues \( \tilde{\Lambda}^+_i \) for \( i = 1, \ldots, N - M - 1 \). Let \( \mathbf{q}_N \) be an eigenvector of \( \mathbf{U} \) corresponding to the largest eigenvalue of \( \mathbf{U} \), namely \( \lambda_N = \Lambda^+_{N-M} \). Suppose that we are able to set the new experiment \( \tilde{\mathbf{b}} \) so that \( \tilde{\mathbf{b}} = \mathbf{q}_N \). In this case it follows from (39) and (35) that

\[
\tilde{\mathbf{U}} = \mathbf{U} - \left( \frac{1}{\mathbf{q}_N^H \mathbf{U} \mathbf{q}_N} \right) \mathbf{U} \mathbf{q}_N \mathbf{q}_N^H \mathbf{U} = \mathbf{Q} \Lambda \mathbf{Q}^H - \left( \frac{1}{\mathbf{q}_N^H \mathbf{Q} \Lambda \mathbf{Q}^H \mathbf{q}_N} \right) \mathbf{Q} \Lambda \mathbf{Q}^H \mathbf{q}_N \mathbf{q}_N \mathbf{Q} \Lambda \mathbf{Q}^H = \mathbf{Q} \Lambda \mathbf{Q}^H - \lambda_N \mathbf{q}_N \mathbf{q}_N^H \, ,
\]
and it follows that the columns of $Q$ comprise a set of orthonormal eigenvectors of $\hat{U}$ and that the nonzero eigenvalues of $\hat{U}$ are precisely $\Lambda_i^+$ for $i = 1, \ldots, N - M - 1$. Therefore the nonzero eigenvalues of $\hat{U}$ are in fact $\tilde{\Lambda}_i^+ = \Lambda_i^+$, for $i = 1, \ldots, N - M - 1$. We therefore see that when $\tilde{b} = q_N$ that (42) is satisfied at equality for all $i = 1, \ldots, N - M - 1$. In light of the lower bounds given in (42) it holds that $\tilde{b} = q_N$ is the ideal next experiment as it simultaneously minimizes any (and hence every) monotone function $h(\cdot)$ of the nonzero eigenvalues of $\hat{U}$.

From a practical point of view, we would not expect there to be an experiment $b_i(\cdot)$ whose corresponding $\tilde{b}_i$ vector satisfies $\tilde{b}_i = q_N$ (or $\tilde{b}_i = -q_N$). Instead, it makes good intuitive sense to seek an experiment $\tilde{b}_i$ that makes the smallest angle with $\pm q_N$. This leads to the criterion to choose the next experiment by finding the index $i^*_{\text{next}}$ such that

$$i^*_{\text{next}} = \arg\max_{i \in \{1, \ldots, L\}} \frac{|b_i^H q_N|}{\|b_i\|}.$$ 

This is formalized in the following algorithm.

0. (Initial values.) $M \leftarrow 0$. Given the matrix $K$, set the initial posterior covariance matrix $U$ to be the prior covariance matrix:

$$U = A^{-1}KA^{-H}. \quad (43)$$

1. (Compute eigenvector of largest eigenvalue.) Solve the eigenvalue problem $Uv = \lambda v$ to obtain the maximum eigenvalue $\lambda_{\text{max}}$ and the corresponding eigenvector $v_{\text{max}}$.

2. (Determine the next experiment to add.) Find an index $i^*_{\text{next}}$ such that

$$i^*_{\text{next}} = \arg\max_{i \in \{1, \ldots, L\}} \frac{|b_i^H v_{\text{max}}|}{\|b_i\|}, \quad (44)$$

and append $b_{i^*_{\text{next}}}$ to the current matrix $C$ to form the updated matrix $C \leftarrow [C \ b_{i^*_{\text{next}}}]$. Perform the new experiment to obtain the experiment outcome value $d_{M+1}$. Update the number of experiments: $M \leftarrow M + 1$. If $M \geq M_{\text{max}}$, Stop. (Optional: apply other stopping criterion to determine whether to Stop.)

3. (Update posterior covariance matrix.) Given the updated matrix $C$, calculate the updated posterior covariance matrix $U$ as

$$U = A^{-1} (K - KA^{-H}C(\sigma^2 I + C^H A^{-1}K A^{-H} C)^{-1} C^H A^{-1}K) A^{-H}. \quad (45)$$

4. (Optional: update hyperparameters.) Optionally, given the updated experiments and their outcomes, re-set the hyperparameters $\theta = (\theta_1, \ldots, \theta_J)$ of the covariance operator $k(\cdot, \cdot, \theta)$ in (24) by approximately solving the likelihood maximization problem (29) as described in Section 3.4.
5. Goto Step 1.

While the algorithm is best conceptualized for the case when $\sigma = 0$, we have found that in practice it also works well for $\sigma > 0$, using the intuitive notion that for small values of $\sigma$ the mathematics developed and used in this section remains approximately valid. Indeed, as we will see in Section 5, the algorithm works very well for $\sigma > 0$ in the applications of interest herein.

Notice that the computational complexity of this algorithm scales very favorably with the number of possible experiments $L$. Specifically, the algorithm needs to evaluate $L$ scalar products $b_i^H v_{max}$, $i = 1, \ldots, L$ only once at each of the $M$ steps. Therefore, the algorithm can accommodate a large list of possible experiments without too much computational effort so long as $M$ is relatively small.

Let us now discuss the optional Step 4 in the above algorithm. By our way of choosing the covariance operator discussed in Subsection 3.4, the matrix $K$ depends on the observations. Ultimately, the prior covariance matrix $K$ plays an important role in the posterior covariance matrix $U$ as it is given by (45). By updating the hyperparameters when a new experiment is chosen, we hope to further reduce the “size” of the posterior covariance matrix $U$ through a good choice of the prior covariance matrix $K$. In all numerical examples presented in the next section, we carry out the optional Step 4 in our greedy algorithm.

5 Numerical Results

We will demonstrate the proposed method on numerical examples from a variety of second-order linear PDEs such as heat equation, convection-diffusion equation, and Helmholtz equation. In order to verify the performance of our method, we will specify the true state and generate the observations by adding Gaussian noise to the true outputs.

5.1 Heat conduction example

We consider the following PDE model for heat conduction in a unit square domain:

$$-\Delta u^o + f^o, \text{ in } \Omega, \quad \text{and} \quad u^o = 0, \text{ on } \partial \Omega,$$  \hspace{1cm} (46)

where $\Omega \equiv (0, 1) \times (0, 1)$ and $f^o = 2\pi^2$. The Galerkin FE formulation of this model problem has the bilinear form and the linear functional

$$a(w, v) = \int_{\Omega} \nabla w \cdot \nabla v dx dy, \quad \ell(v) = \int_{\Omega} f^o v dx dy, \quad \forall w, v \in V,$$  \hspace{1cm} (47)

respectively. Here $V$ is a FE approximation space of piecewise linear polynomials defined on a finite element mesh of 5,000 elements. Figure 1 shows a plot of the finite element solution of our PDE model (46). The possible observation functionals are specified as

$$b_i(v) = \int_{\Omega} \exp \left( -\frac{(x - x_i)^2 + (y - y_i)^2}{0.25^2} \right) v dx dy, \quad i = 1, \ldots, L,$$  \hspace{1cm} (48)
where \((x_i, y_i), i = 1, \ldots, L = 961\) are the points of a \(31 \times 31\) uniform grid as shown in Figure 2. These spatial points represent possible measurement locations at which we obtain the observations. In this example, we assume that the observations are noise-free, i.e., \(\sigma = 0\).

Figure 1: Plot of \(u^o\) (left), \(u^{true}\) (middle), and their difference \(u^o - u^{true}\) (right).

Furthermore, we assume that the true state is \(u^{true} = \sin(\pi x) \sin(\pi y)\) as depicted in Figure 1. We note that the true state \(u^{true}\) is the solution of the PDE model (46) in which the source term \(f^o = 2\pi^2\) is replaced with \(f^{true} = 2\pi^2 \sin(\pi x) \sin(\pi y)\). Therefore, the source of uncertainty in our PDE model (46) comes from the source term. Due to this uncertainty, as shown in Figure 1, the solution \(u^o\) of the PDE model (46) is significantly different from the true state \(u^{true}\). In what follows, we combine the PDE model (46) with observations to improve the prediction of the true state by using the methodology proposed in this paper.

First, we consider a bilinear covariance operator of the form

\[
k(w, v; \theta) = \int_{\Omega} (\theta_1 w v + \theta_2 \nabla w \cdot \nabla v) \, dx dy,
\]

where the parameters \(\theta = (\theta_1, \theta_2)\) are chosen to maximize the log marginal likelihood on a uniform grid \(201 \times 201\) of the hyperparameter domain \([0.01, 100] \times [0, 1]\). We next pursue the greedy algorithm to determine the measurement locations, which are shown in Figure 2. We observe that the selected measurement locations are distributed over the whole domain. Figure 3 shows the hyperparameter \(\theta_1\) and the \(L^2(\Omega)\) norm of the standard function \(\|\sqrt{\eta}\|_{\Omega}\) as well as the error norm \(\|u^{true} - \tilde{u}\|_{\Omega}\) as a function of the number of measurements. It is interesting to note that the maximum marginal likelihood approach yields the second hyperparameter \(\theta_2 = 0\). It means that our bilinear covariance operator \(k\) is collinear with the \(L^2\) inner product. Figure 4 shows the posterior variance function and the selected measurement points at each iteration of our greedy algorithm. We observe that the variance reduces as we increase the number of measurements and that the measurement points are typically selected near the maximum peak of the posterior variance function.

We next consider a bilinear covariance operator of the form

\[
k(w, v; \theta) = \theta \int_{\Omega} \nabla w \cdot \nabla v dx dy,
\]
Figure 2: Potential measurement points (left) and selected measurement locations (right).

Figure 3: The first hyperparameter $\theta_1$ (left), and the error norm and the standard deviation norm (right) as a function of $M$. Note that $\theta_2 = 0$. 
Figure 4: The variance function $\eta(x)$ and the measurement points for $M = 1, \ldots, 12$, where the order goes from left to right and top to bottom.
We show in Figure 5 the selected measurement points. Figure 6 shows the hyperparameter \( \theta \) and the \( L^2(\Omega) \) norm of the standard function \( \| \sqrt{\eta} \|_\Omega \) as well as the error norm \( \| u^{\text{true}} - \bar{u} \|_\Omega \) as a function of the number of measurements. We see that both the error and the standard deviation function decreases as we increase \( M \). However, the error converges to zero much faster than the standard deviation. It is interesting to note that the choice of the covariance operator (50) yields smaller error yet larger standard deviation than the previous choice (49) for any \( M \). Hence, the former (49) yields better error estimate than the latter (50). This demonstrates that the structure of the covariance operator has a significant impact on the outcome of our estimate of the true state.

Figure 5: Measurement points selected by the greedy algorithm for the covariance operator of the form (50).

Figure 6: The parameter \( \theta \) (left), and the error norm and the standard deviation norm (right) as a function of \( M \) for the covariance operator of the form (50).
5.2 Convection-diffusion example

We next consider a convection-diffusion problem on a unit square \( \Omega \equiv (0,1) \times (0,1) \):

\[-\Delta u^o + c^o \cdot \nabla u^o = f^o, \quad \text{in } \Omega, \quad \text{and } u^o = 0, \quad \text{on } \partial \Omega, \quad (51)\]

where \( c^o = (10,10) \) and \( f^o = 10 \). The Galerkin FE formulation of the above PDE model is derived with the bilinear form and the linear functional

\[ a(w,v) = \int_{\Omega} \nabla w \cdot \nabla v \, dxdy + \int_{\Omega} (c^o \cdot \nabla w) v \, dxdy, \quad \ell(v) = \int_{\Omega} f^o v \, dxdy, \quad \forall w,v \in V, \quad (52)\]

where \( V \) is a FE approximation space of piecewise polynomials of degree \( p = 3 \) and is defined on a mesh of 800 elements. The possible measurement points and possible observation functionals are the same as in the previous example. We also assume that the observations are noise-free, namely, \( \sigma = 0 \).

To assess our method, we assume that the true state is the solution of the PDE model (51) in which the convective velocity \( c^o = (10,10) \) is replaced with \( c^\text{true} = (15,15) \). Hence, the source of uncertainty in our PDE model (51) comes from the convective velocity. Due to this uncertainty, as shown in Figure 1, the model state \( u^o \) is significantly different from the true state \( u^\text{true} \). In this example, the uncertainty is present in the bilinear form \( a \) (i.e., the matrix \( A \)), whereas, in the previous example, the uncertainty is present in the linear functional \( \ell \) (i.e., the vector \( l \)). This example serves to demonstrate the effectiveness of our method on cases where uncertainty affects PDE operators.

![Figure 7: Plot of \( u^o \) (left), \( u^\text{true} \) (middle), and their difference \( u^o - u^\text{true} \) (right).](image)

We consider a bilinear covariance operator of the form (49) and pursue the greedy algorithm to obtain the measurement locations as shown in Figure 8. We observe again that the selected measurement points are distributed over the whole domain. Figure 9 shows the hyperparameter \( \theta_1 \) and the norm of the standard function \( \| \sqrt{\eta} \|_\Omega \) as well as the error norm \( \| u^\text{true} - \bar{u} \|_\Omega \) as a function of the number of measurements. We note that \( \theta_2 = 0 \). As expected, both the error norm and the standard deviation norm decrease as \( M \) increases. Hence, increasing the number of measurements reduces the uncertainty in our prediction of the true state. This can be seen more clearly in Figure 10 which shows the absolute error.
function and the standard deviation function for three different values of $M = 20, 40, \text{and } 60$. We observe that the error near the top right corner of the domain is much larger than the one near the bottom left corner of the domain, while the standard deviation is spread over the physical domain much more evenly than the error. Note that both the mean prediction and the standard deviation vanish on the domain boundary because there is no uncertainty on the boundary condition in this example. In the next example, we consider a Helmholtz problem where the uncertainty source comes from the boundary condition.

![Figure 8: Measurement points selected by the greedy algorithm.](image)

![Figure 9: The first parameter $\theta_1$ (left), the error and the standard deviation (right) as a function of $M$. Note that $\theta_2 = 0$.](image)

### 5.3 A Helmholtz example

We consider the sound-hard scattering of an incident plane wave $u^{\text{inc}} = \exp(ikx)$ by a circle of radius $R = 1$, where $k$ is the wave number. The scattered field satisfies the exterior
Figure 10: The absolute error function $|\bar{u}(x) - u^{\text{true}}(x)|$ (top row) and the standard deviation function $\sqrt{\eta(x)}$ (bottom row) for $M = 20$ (left), $M = 40$ (middle), and $M = 60$ (right).

Helmholtz problem

$$\Delta u^{\text{true}} + k^2 u^{\text{true}} = 0, \quad \text{in } \mathbb{R}^2 \setminus \Omega_c,$$
$$\nabla u^{\text{true}} \cdot n + \nabla u^{\text{inc}} \cdot n = 0, \quad \text{on } \Gamma_c,$$
$$\lim_{r \to \infty} \sqrt{r} \left( \frac{\partial u^{\text{true}}}{\partial r} - iku^{\text{true}} \right) = 0,$$

where $\Omega_c$ is the domain of the unit circle and $\Gamma_c$ is the boundary of the unit circle. The last condition is known as Sommerfeld radiation condition. It is known that the solution of the above Helmholtz problem has the analytical form

$$u^{\text{true}}(x, y) = -\frac{J'_0(ka)}{H'_0(ka)} H^1_0(kr) - 2 \sum_{n=1}^{\infty} r^n \frac{J'_n(ka)}{H'_n(ka)} H^1_n(kr) \cos(n\theta),$$

where $(r, \theta)$ are polar coordinates of $(x, y)$, and $J_n$ and $H^1_n$ are the Bessel function and the Hankel function of the first kind, respectively [17]. Note that the prime denotes the derivative of a function with respect to its argument.

Since the scatter is simple the solution of the exterior Helmholtz problem (53) is known analytically. For more complicated scatterers, however, the problem is usually solved by using a numerical method. In the case of a finite element method, the unbounded domain
need to be truncated to a bounded domain and the Sommerfeld radiation condition should be replaced with a suitable absorbing boundary condition at the exterior boundary of the truncated domain. This gives rise to the following PDE model

\[
\begin{align*}
\Delta u^o + k^2 u^o &= 0, \quad \text{in } \Omega, \\
\nabla u^o \cdot n + \nabla u^{inc} \cdot n &= 0, \quad \text{on } \Gamma_c, \\
\nabla u^o \cdot n - iku^o &= 0, \quad \text{on } \Gamma_o,
\end{align*}
\]

(55)

where \( \Omega \) is the truncated domain as shown in Figure 11 and \( \Gamma_o \) is the exterior boundary of the truncated domain. Note here that we approximate the Sommerfeld radiation condition by the first-order absorbing boundary condition. The weak formulation of the PDE model (55) is given by

\[
a(u^o, v) = \ell(v), \quad \forall v \in X,
\]

(56)

where, for all \( w, v \in X \),

\[
a(w, v) = \int_\Omega (\nabla w \cdot \nabla v - k^2 w v) dxdy - \int_{\Gamma_o} ikwvdxdy,
\]

\[
\ell(v) = -\int_{\Gamma_c} (\nabla u^{inc} \cdot n) v dxdy,
\]

(57)

and \( X \equiv H^1(\Omega) \). In practice, we replace \( X \) with a finite dimensional space \( V \) which is a FE approximation space of piecewise polynomials of degree \( p = 3 \) and is defined on a mesh of 800 elements. This FE discretization is fine enough such that the numerical error is negligible.

This example serves to demonstrate our approach when the uncertainty source comes from the domain truncation and the boundary condition. In particular, we aim to improve the prediction of the scattered field \( u^{true} \) by combining the PDE model (55) with observations of the scattered field. Of particular interest is the value of the scattered field on the unit circle, as it ultimately determines the radar cross section. The possible observation functionals are specified as

\[
b_i(v) = \int_\Omega \exp \left( -\frac{(x-x_i)^2 + (y-y_i)^2}{0.25^2} \right) v dxdy, \quad i = 1, \ldots, L,
\]

(58)

where \((x_i, y_i), i = 1, \ldots, L = 3720\) coincide with the mesh points. In this example, we will consider three different noise levels \( \sigma = 0, 0.005, \) and 0.01.

We present in Figure 11 the selected measurement locations for three different noise levels. It is interesting to see that in all cases the measurement locations are quite close to the unit circle and distributed quite uniformly along the unit circle. We next show the true state, the mean prediction, and the 95% confidence region (shaded area) for \( \sigma = 0 \) in Figure 12, for \( \sigma = 0.005 \) in Figure 13, and for \( \sigma = 0.01 \) in Figure 14. Here the 95% confidence region is an area bounded by the mean prediction plus and minus two times the standard deviation function. We see that increasing \( M \) can improve the accuracy of the mean prediction for \( \sigma = 0 \) and \( \sigma = 0.005 \). However, for \( \sigma = 0.01 \), the mean prediction does not get better as we
increase $M$. This means that when $\sigma$ is equal or greater than 0.01, the observations do not add a positive contribution to the prediction of the true state. For $\sigma = 0$ the 95% confidence region shrinks quite rapidly as $M$ increases, whereas for $\sigma = 0.01$ the 95% confidence region does not shrink as $M$ increases. Therefore, when the observations become too noisy, adding more observations does not help improve the prediction. Clearly, the mean prediction and error estimate get worse as the noise level increases.

![Figure 11: The measurement locations obtained using the greedy algorithm for three different noise levels: $\sigma = 0$ (left), $\sigma = 0.005$ (middle), and $\sigma = 0.01$ (right).]

## 6 Conclusions

We have presented a statistical method that combines a linear PDE model with observations to predict the state of a physical problem. First, a random functional is introduced into the PDE model to account for various sources of uncertainty in the model. This random functional is posed as a Gaussian process with zero mean and prior covariance operator. Next, a linear regression model for the Gaussian functional is derived by utilizing the adjoint states and the observations. This regression model allows us to compute the posterior distribution of both the Gaussian functional and the state estimate. The method is thus named as Gaussian functional regression. A key ingredient of our method is the prior covariance operator of the Gaussian functional. We propose a class of affine bilinear forms for the prior covariance operator and determine the associated hyperparameters based on the observations. Furthermore, we devise a greedy algorithm to select observations among a large number of possible measurements. The proposed method has a number of attractive features as demonstrated by numerical examples from heat conduction, convection-diffusion, and wave propagation.

We would like to extend our approach to nonlinear PDEs. Two new challenges arise in the nonlinear case. First, although the functional $g$ is Gaussian, the state $u$ is no longer a Gaussian random field because of the nonlinearity. Second, the nonlinearity prevents a direct link between the adjoint solutions and the observations, thereby complicating the regression model. Our research will focus on addressing these issues.
Figure 12: Panels show the true state and the mean prediction as a function of $\theta$: the real part (left) and the imaginary part (right). These results are obtained using $M = 4$ (top row), $M = 8$ (middle row), and $M = 12$ (bottom row) observations with $\sigma = 0$. In these plots the shaded area represents the mean prediction plus and minus two times the standard deviation function (corresponding to the 95% confidence region).
Figure 13: Panels show the true state and the mean prediction as a function of $\theta$: the real part (left) and the imaginary part (right). These results are obtained using $M = 4$ (top row), $M = 8$ (middle row), and $M = 12$ (bottom row) observations with $\sigma = 0.005$. In these plots the shaded area represents the mean prediction plus and minus two times the standard deviation function (corresponding to the 95% confidence region).
Figure 14: Panels show the true state and the mean prediction as a function of $\theta$: the real part (left) and the imaginary part (right). These results are obtained using $M = 4$ (top row), $M = 8$ (middle row), and $M = 12$ (bottom row) observations with $\sigma = 0.01$. In these plots the shaded area represents the mean prediction plus and minus two times the standard deviation function (corresponding to the 95% confidence region).
References


A State Estimation with Bayesian Inference

A.1 Matrix representation

The matrix form of the Galerkin FE formulation (1) seeks \((u^o, s^o) \in \mathbb{C}^N \times \mathbb{C}^M\) that satisfies

\[
Au^o = l, \quad s^o = C^H u^o,
\]

(59)

where \(u^o \in \mathbb{C}^N\) is the vector of degrees of freedom of the FE solution \(u^o\). Recall that \(A, l,\) and \(C\) have entries \(A_{ij} = a(\psi_i, \psi_j), l_i = \ell(\psi_i), C_{im} = c_m(\psi_i)\) for \(i, j = 1, \ldots, N\) and \(m = 1, \ldots, M\).

Similarly, the matrix form of the stochastic weak formulation (3) seeks \((u, s) \in \mathbb{C}^N \times \mathbb{C}^M\) such that

\[
Au + g = l, \quad s = C^H u,
\]

(60)

where \(g\) is a Gaussian random vector with the prior distribution \(\mathcal{N}(0, K)\). Recall that the prior covariance matrix \(K\) has entries \(K_{ij} = k(\psi_i, \psi_j)\).

Below we describe a Bayesian approach to determine the posterior distribution of both \(g\) and \(u\) that incorporates the observation vector \(d\).

A.2 Bayesian analysis

We first need to define the likelihood function. To this end, we compute the adjoint state vectors stored in the matrix \(\Phi \in \mathbb{C}^{N \times M}\) by solving the adjoint problem:

\[
A^H \Phi = -C.
\]

(61)

It follows from (59), (60), and (61) that

\[
\Phi^H g = \Phi^H (l - Au) = \Phi^H (Au^o - Au) = -C^H (u^o - u) = s - s^o.
\]

(62)

We now substitute \(s = d - \varepsilon\) into (62) to arrive at the standard linear regression model:

\[
d - s^o = \Phi^H g + \varepsilon.
\]

(63)

Since \(\varepsilon \sim \mathcal{N}(0, \sigma^2 I)\), then \((d - s^o)|g \sim \mathcal{N}(\Phi^H g, \sigma^2 I)\). Therefore the likelihood function, i.e., the probability density function of \(d - s^o\) given the vector \(g\), is

\[
p((d - s^o)|g, \Phi) = \frac{1}{(2\pi \sigma^2)^{M/2}} \exp \left( -\frac{1}{2\sigma^2} \|d - s^o - \Phi^H g\|^2 \right),
\]

(64)
where \( \parallel \cdot \parallel \) denotes the Euclidean norm.

We now apply the Bayesian approach to the standard linear regression model (63). According to Bayes’s theorem (see Eq. (2.3) on page 17 in [33]), the posterior distribution of \( g \) is given by

\[
\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal likelihood}},
\]

where the marginal likelihood \( p((d - s^o) | \Phi) \) is given by

\[
p((d - s^o) | \Phi) = \int p((d - s^o) | \Phi, g) p(g) dg.
\]

Writing only the terms from the likelihood and the prior, we obtain

\[
p(g | (d - s^o), \Phi) \propto \exp \left( -\frac{1}{2\sigma^2} \|d - s^o - \Phi^H g\|^2 \right) \exp \left( -\frac{1}{2} g^H K^{-1} g \right),
\]

where \( \bar{g} \) and \( G \) are given by

\[
\bar{g} = \frac{1}{\sigma^2} G \Phi (d - s^o), \quad G = \left( \frac{1}{\sigma^2} \Phi \Phi^H + K^{-1} \right)^{-1}.
\]

However, since this formula requires the inverse of a matrix of size \( N \times N \), we will derive an equivalent expression which is more efficient to compute than (68).

We recall the Sherman-Morrison-Woodbury (SMW) formula for the matrix inversion

\[
(Z + UWV^H)^{-1} = Z^{-1} - Z^{-1}U(W^{-1} + V^H Z^{-1}U)^{-1}V^HZ^{-1}.
\]

Using the SMW formula with \( Z = K^{-1}, W = \sigma^{-2}I, U = V = \Phi \) we obtain

\[
G = K - K\Phi (\sigma^2 I + \Phi^H K\Phi)^{-1}\Phi^H K.
\]

It thus follows that

\[
G\Phi = K\Phi - K\Phi (\sigma^2 I + \Phi^H K\Phi)^{-1}\Phi^H K\Phi,
\]

\[
= K\Phi (\sigma^2 I + \Phi^H K\Phi)^{-1} (\sigma^2 I + \Phi^H K\Phi) - \Phi^H K\Phi,
\]

\[
= \sigma^2 K\Phi (\sigma^2 I + \Phi^H K\Phi)^{-1}.
\]

Hence, we obtain from (68)-(71) that

\[
\bar{g} = K\Phi D^{-1}(d - s^o), \quad G = K - K\Phi D^{-1}\Phi^H K,
\]

where \( D = \sigma^2 I + \Phi^H K\Phi \). In this way, we need to invert the matrix \( D \) of size \( M \times M \). Since \( M \) is typically much smaller than \( N \), the formula (72) is much more efficient to compute than the original one (68).
A.3 Posterior distribution of the state estimate

We note from the stochastic linear system (60) and the Gaussian property of $g$ that the posterior distribution of $u$ is a multivariate normal distribution

$$u|(d - s^o), \Phi \sim \mathcal{N}(\bar{u}, U),$$

where the posterior mean and the posterior covariance matrix are given by (18), namely:

$$\bar{u} = A^{-1}(l - \bar{g}), \quad U = A^{-1}GA^{-H}. \tag{74}$$

Substituting (72) into (74) yields

$$\bar{u} = A^{-1} \left( l - K\Phi(\sigma^2I + \Phi^H K\Phi)^{-1}(d - s^o) \right),$$

$$U = A^{-1} \left( K - K\Phi(\sigma^2I + \Phi^H K\Phi)^{-1}\Phi^H K \right) A^{-H}. \tag{75}$$

We observe that this result is exactly the same as the one obtained using the Gaussian functional regression model presented in Section 3.

Finally, we obtain from (59), (60) and (75) the following expression for the mean of the output vector

$$\bar{s} = C^H \bar{u} = s^o + \Phi^H K\Phi(\sigma^2I + \Phi^H K\Phi)^{-1}(d - s^o). \tag{76}$$

This equation shows among other things that $\bar{s} = d$ in the case that $\sigma = 0$. Therefore, in the case of noise-free observations, the mean output vector perfectly matches the observation vector. This is an important property of our method.

A.4 Relationship with a constrained least-squares regression

Let $\beta = D^{-1}(d - s^o)$ and $\bar{q} = \Phi\beta$. We now show that $(\bar{u}, \bar{q}, \beta)$ is the optimal solution of the following least squares minimization problem

$$\min_{(z, w, \gamma) \in \mathbb{C}^N \times \mathbb{C}^N \times \mathbb{C}^M} \frac{1}{2}w^H Kw + \frac{1}{2}\sigma^2 \gamma^H \gamma$$

s.t. $Az + Kw = l,$

$$C^H z + \sigma^2 \gamma = d. \tag{77}$$

We begin by introducing the Lagrangian

$$\mathcal{L}(z, w, \gamma, r, \lambda) = \frac{1}{2}w^H Kw + \frac{1}{2}\sigma^2 \gamma^H \gamma - (Az + Kw - l)^H r + (C^H z + \sigma^2 \gamma - d)^H \lambda \tag{78}$$

where $-r \in \mathbb{C}^N$ and $\lambda \in \mathbb{C}^M$ are the Lagrange multipliers of the constraints. The optimal solution $(z^*, w^*, \gamma^*)$ of the least squares problem (77), together with Lagrange multipliers $(r^*, \lambda^*)$ will satisfy the first-order condition that $\nabla \mathcal{L}(z^*, w^*, \gamma^*, r^*, \lambda^*) = 0$, which works
out to be the following equation system:

\[ A^H r^* - C \lambda^* = 0, \]  
\[ Kw^* - Kr^* = 0, \]  
\[ \gamma^* + \lambda^* = 0, \]  
\[ Az^* + Kw^* - l = 0, \]  
\[ C^H z^* + \sigma^2 \gamma^* - d = 0. \]  

The first three equations (79a), (79c) and (79c) yield that

\[ w^* = A^{-H} C \lambda^* = -\Phi \lambda^* = \Phi \gamma^*. \]  

Substituting (80) into (79d) and invoking (59) we obtain

\[ z^* = A^{-1} (l - Kw^*) = u^o - A^{-1} K \Phi \gamma^*. \]  

Inserting (81) into (79e) and invoking (59) we get

\[ (\Phi^H K \Phi + \sigma^2 I) \gamma^* = d - s^o. \]  

This equation shows that \( \beta = \gamma^* \). It follows from (80) that \( q = w^* \) and from (81) that \( \bar{u} = z^* \). Hence, we have shown that \((\bar{u}, \bar{q}, \beta) = (z^*, w^*, \gamma^*)\). This proves our claim.

### A.5 Relationship with 3D variational data assimilation

By eliminating the constraints in the least-squares problem (77) we obtain the following result

\[ \bar{u} := \arg \min_{z \in \mathbb{C}^N} \frac{1}{2} (Az - l)^H K^{-1} (Az - l) + \frac{1}{2} \sigma^{-2} (C^H z - d)^H (C^H z - d). \]  

We see that the posterior mean is optimal in the sense that it minimizes an error objective function, which is defined as the sum of the model error and the output error. Furthermore, we can write (83) as

\[ \bar{u} := \arg \min_{z \in \mathbb{C}^N} \frac{1}{2} (z - u^o)^H U_0^{-1} (z - u^o) + \frac{1}{2} \sigma^{-2} (C^H z - d)^H (C^H z - d). \]  

where \( U_0 = A^{-1} K A^{-H} \). The optimization formulation (84) is known as three-dimensional (3D) variational data assimilation [7, 22]. It shows the relation between our method and 3D variational data assimilation.
B State Estimation with the Kalman Method

Let us review the Kalman method [19] for state estimation. We begin by assuming that the state vector $u$ is Gaussian with the prior distribution

$$p(u) \propto \exp \left( -\frac{1}{2} (u - u^b)^H B^{-1} (u - u^b) \right),$$  
(85)

where $u^b$ is the background state vector and $B$ is the background covariance matrix. Furthermore, the data $d$ is assumed to have a Gaussian probability density function with covariance $R$ and mean $C^H u$. The likelihood function is thus given by

$$p(d|u) \propto \exp \left( -\frac{1}{2} (C^H u - d)^H R^{-1} (C^H u - d) \right).$$  
(86)

The posterior distribution of $u$ is given by Bayes’ theorem (see Eq. (2.3) on page 17 in [33]):

$$p(u|d) \propto p(d|u)p(u).$$  
(87)

It can be shown by algebraic manipulations that the posterior distribution is also Gaussian

$$p(u|d) \propto \exp \left( -\frac{1}{2} (u - \hat{u})^H \hat{U}^{-1} (u - \hat{u}) \right).$$  
(88)

Here the posterior mean $\hat{u}$ and covariance $\hat{U}$ are given by the Kalman formulas

$$\hat{u} = u^b + H (d - C^H u^b), \quad \hat{U} = B - HC^H B,$$  
(89)

where

$$H = BC \left( C^H BC + R \right)^{-1},$$  
(90)

is the so-called Kalman gain matrix.

We now show that our method is related to the Kalman method for a particular choice of the background state vector and the background covariance matrix. In particular, let us choose the prior information as

$$u^b = u^o, \quad B = A^{-1} K A^{-H}, \quad R = \sigma^2 I.$$  
(91)

It then follows from (89)-(91) that

$$\hat{u} = u^o + A^{-1} K A^{-H} C \left( C^H A^{-1} K A^{-H} C + \sigma^2 I \right)^{-1} (d - s^o),$$

$$\hat{U} = A^{-1} K A^{-H} - A^{-1} K A^{-H} C \left( C^H BC + \sigma^2 I \right)^{-1} C^H A^{-1} K A^{-H}.$$  
(92)

We see that the Kalman method yields exactly the same posterior distribution as our method when the priors are chosen by (91). Our method provides a systematic way to choose appropriate priors which are very important in the context of the Kalman method.