Stochastic Model-Based Source Identification

Luke Calkins, Reza Khodayi-mehr, Wilkins Aquino, and Michael Zavlanos

Abstract—**In this paper we investigate the use of Stochastic Reduced Order Models (SROMs) for solving Stochastic Source Identification (SSI) problems in steady-state transport phenomena given statistics of the system state at a small number of locations. We capture the physics of the transport phenomenon by a Partial Differential Equation (PDE) which we discretize using the finite element method. The SSI problem is then formulated as a stochastic optimization problem constrained by the PDE, and then transformed into a deterministic one after representing the random quantities with a low-dimensional discrete SROM. The small number of samples given by SROMs requires only a small number of PDE solves at each optimization iteration in order to obtain a solution to the SSI problem, defined as a distribution of possible source locations and intensities. We provide simulations to demonstrate the effectiveness of SROMs in capturing uncertainty. We also demonstrate the ability of SROMs to capture multiple independent sources of uncertainty, in particular, we consider uncertainty in the location of the measurements which has practical implications in robotics applications.**

I. Introduction

Source Identification (SI) refers to the estimation of the location, intensity, and shape of a source given a set of measurements of the quantity generated by that source. The SI problem has a wide variety of applications spanning from environmental protection to human safety [1], [2]. SI falls under a broader class of problems known as Inverse Problems (IPs) [3], [4]. In this setting, one is provided with observations of a system state and uses these observations to estimate some unknown parameters of the source. A common approach to solving IPs is to formulate them as deterministic optimization problems that minimize an appropriate objective function involving the parameters of interest and the observed data subject to the constraint that some model of the underlying phenomenon is satisfied. However, when solving IPs in practice, uncertainty is unavoidable (e.g., measurement noise, boundary conditions) and incorporating these uncertainties is crucial to understanding the degree of uncertainty in the solution. However, quantifying this uncertainty comes at the cost of increased computation. The method presented in this paper focuses on an efficient method to solve the SI problem under uncertainty.

Existing methods to handle uncertainty in IPs fall into two categories, namely, Bayesian inference methods and the stochastic optimization approaches. In Bayesian methods, a prior distribution of the unknown parameters and a likelihood model are used to form a posterior distribution. Often, this distribution can only be estimated via Markov Chain Monte Carlo (MCMC) methods. In [5] and [6], standard MCMC algorithms are used to solve IPs in elastography and heat conduction respectively. In [7] these methods are adapted using a sparse grid collocation approach. While these methods have been shown to work well for various applications, they require the costly solution of a deterministic forward simulation for every parameter sample. Furthermore, the solution can be significantly different with small fluctuations in the observed data, or by choosing a different prior distribution, which can often be arbitrary.

Stochastic optimization approaches differ from Bayesian approaches in that the inputs to such methods are the statistics of the state of the system rather than a single observation of the random state. An approach using stochastic optimization for IPs was first developed in [8] for the inverse heat conduction problem. In this work, the unknown parameters were represented using generalized polynomial chaos expansions. The work was later extended by representing uncertainty with a sparse grid collocation approach [9], [10].

In this paper, we utilize Stochastic Reduced Order Models (SROMs) to approximate the random quantities. A SROM is a low-dimensional discrete approximation to a continuous random quantity consisting of a finite and generally small set of samples and corresponding probabilities. The sample-probability pairs are chosen such that the resulting discrete distribution is close in a statistical sense to the underlying continuous distribution that it is approximating. SROMs were first proposed in [11] and further refined in [12]. SROMs have seen success in forward uncertainty propagation in applications such as linear dynamical systems [13] and interangular corrosion rates [14]. The SROM methodology for general IPs under uncertainty was developed in [15]. The contribution of this paper is the use of SROMs for stochastic SI in advection-diffusion problems under uncertainty in the measurement locations, a problem of practical importance especially in controls and robotics applications that has not been explored in relevant literature on SROMs. Our results suggest that SROMs, which have only been recently investigated in uncertainty quantification and IP literature, can present a potentially powerful alternative to Bayesian methods that are predominantly used to solve estimation problems in the controls and robotics.

The remaining sections of this paper are organized as follows. In Section II we define the stochastic SI problem in an advection-diffusion transport system. Section III introduces SROMs and describes how they can be used to propagate uncertainty in forward problems. Section IV describes how

This work is supported in part by NSF awards CNS #1261828 and DGE #1644868

Luke Calkins, Reza Khodayi-mehr and Michael M. Zavlanos are with the Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA, {william.calkins, reza.khodayi.mehr, michael.zavlanos}@duke.edu. Wilkins Aquino is with the Department of Civil and Environmental Engineering, Duke University, Durham, NC 27708, USA, wilkins.aquino@duke.edu.

SROMs can be used to solve stochastic IPs and discusses our solution strategy for the stochastic SI problem. Section V provides numerical simulations demonstrating the effectiveness of the method. We conclude in Section VI with a summary of this work as well as future directions.

II. Problem Definition

A. Advection-Diffusion Partial Differential Equation

Consider a physical domain of interest $\Omega \in \mathbb{R}^\ell$ where $1 \le \ell \le 3$ and some nonnegative source function $f : \Omega \to \mathbb{R}_+$. Denote the measurable concentration quantity by $c : \Omega \to \mathbb{R}_+$. Let the velocity field within the domain be denoted by $\mathbf{q} \in \mathbb{R}^\ell$ and the diffusivity of the medium by $D \in \mathbb{R}_+$. Assuming steady-state conditions and a zero-valued (Dirichlet) condition on the boundary Γ of the domain, transport of the concentration in the domain is modeled with the following Boundary Value Problem (BVP):

$$\begin{aligned} -D\nabla^2 c + \mathbf{q}\cdot\nabla c - f &= 0 \qquad \text{in } \Omega \\ c &= 0 \qquad \text{on } \Gamma. \end{aligned} \tag{1}$$

After discretization of the BVP (1) using the Finite Element (FE) Method with a mesh with $\hat{N}$ total nodes, we get a finite dimensional approximation of the solution given by the following algebraic system of equations:

$$[\mathbf{K}]\hat{\mathbf{c}} = [\mathbf{R}]\hat{\mathbf{f}}, \tag{2}$$

where the matrices $[\mathbf{K}]$ and $[\mathbf{R}]$ are $\hat{N}\times\hat{N}$ sparse matrices. $\hat{\mathbf{c}}$ and $\hat{\mathbf{f}}$ are the $\hat{N}$ dimensional concentration and source vectors respectively, with entries being the value of the concentration and source functions evaluated at the nodes of the FE mesh. Therefore, given a source vector $\hat{\mathbf{f}}$, a unique $\hat{\mathbf{c}} = [\mathbf{K}]^{-1}[\mathbf{R}]\hat{\mathbf{f}}$ can be determined. See [16] for details.

Let the matrix $[\mathbf{Q}] \in \mathbb{R}^{n\times\hat{N}}$ be an indicator matrix such that after multiplying $[\mathbf{K}]^{-1}[\mathbf{R}]\hat{\mathbf{f}}$, returns the $n \ll \hat{N}$ entries of $[\mathbf{K}]^{-1}[\mathbf{R}]\hat{\mathbf{f}}$ corresponding to specific locations in the FE mesh, i.e., a subset of the full $\hat{N}$-dimensional vector $\hat{\mathbf{c}}$. Furthermore, assume that the source vector $\hat{\mathbf{f}}$ can be parametrized by $d \ll \hat{N}$ parameters denoted by $\mathbf{s} \in \mathbb{R}^d$. Given $\mathbf{s}$, the $\hat{N}$-dimensional vector $\hat{\mathbf{f}}$ can be constructed, and then the n-dimensional concentration vector can be solved for $\mathbf{c} = [\mathbf{Q}][\mathbf{K}]^{-1}[\mathbf{R}]\,\hat{\mathbf{f}}(\mathbf{s}) \in \mathbb{R}^n$. Lastly, let $\boldsymbol{\theta} \in \mathbb{R}^q$ denote some other parameters in the BVP (1) different from the source. Now, we can define a finite dimensional abstract model for (1) in the reduced space of source and concentration vectors,

$$\mathbf{M}(\mathbf{c}, \mathbf{s}; \boldsymbol{\theta}) = \mathbf{0} \quad \mathbf{c} \in \mathbb{R}^n, \;\; \mathbf{s} \in \mathbb{R}^d \quad \boldsymbol{\theta} \in \mathbb{R}^q. \tag{3}$$

Consider a set of n stationary sensors within the domain Ω that can take measurements of the concentration field. Specifically, these sensors are located at the locations corresponding to the entries of $\mathbf{c}$. These sensors observe the concentration under the influence of noise. Let $\mathbf{C} \in \mathbb{R}^n$ denote the random concentration vector. Assume that the noise follows a multiplicative model such that $\mathbf{C}_i = \mathbf{c}_i(1+\delta\epsilon)$ where $\mathbf{C}_i$ and $\mathbf{c}_i$ denote the i'th elements of the noise-corrupted and true concentration vectors respectively. δ is the prescribed noise level in the system, and $\epsilon \sim \mathcal{N}(0,1)$. Therefore, the distribution of concentration at each measurement location is a scaled and shifted normal, i.e.,

$$\mathbf{C}_i \sim \mathcal{N}(\mathbf{c}_i, \mathbf{c}_i^2\delta^2). \tag{4}$$

B. Stochastic Source Identification Problem

Given statistics of the observable system state, denoted by $\mathbf{C}_m$, we wish to determine the random source and concentration elements $\mathbf{S}$ and $\mathbf{C}$ that best approximate the observed measurements in a statistical sense. We require that these random elements satisfy the governing physical model (3). Furthermore, let $\boldsymbol{\Theta} \subset \mathbb{R}^q$ denote some additional known source of uncertainty such as uncertainty in the diffusivity D, velocity field $\mathbf{q}$, boundary conditions, or location of the measurements. The abstract model then becomes

$$\mathbf{M}(\mathbf{C}, \mathbf{S}; \boldsymbol{\Theta}) = \mathbf{0}. \tag{5}$$

Problem II.1. *Given the distribution of concentration at n locations in the domain Ω and the probability law of random parameters $\boldsymbol{\Theta}$, find the probability law of the source that generated the observed measurement statistics.*

In order to solve Problem II.1, we will define an objective function to be minimized $\mathcal{J}(\mathbf{C}, \mathbf{S}; \boldsymbol{\Theta})$ where $\mathcal{J} : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}$. In the context of inverse problems, this objective function can be defined as:

$$\mathcal{J}(\mathbf{C}, \mathbf{S}; \boldsymbol{\Theta}) := \mathcal{D}(\mathbf{C}, \mathbf{C}_m; \boldsymbol{\Theta}) + \mathcal{R}(\mathbf{S}, \mathbf{S}_0), \tag{6}$$

where the operators $\mathcal{D}(\cdot,\cdot)$ and $\mathcal{R}(\cdot,\cdot)$ represent functions measuring distance between random elements. These can be thought of as distance metrics in the sense of statistics. $\mathcal{R}(\cdot,\cdot)$ along with $\mathbf{S}_0$ represent *a prior* knowledge of the statistics of the source variable of interest $\mathbf{S}$. With the objective function defined, we can now define a stochastic optimization problem:

$$\begin{aligned} \min_{\mathbf{C}\in\mathcal{C}, \mathbf{s}\in\mathcal{S}} \quad & \mathcal{J}(\mathbf{C}, \mathbf{S}; \boldsymbol{\Theta}) \\ \text{s.t.} \quad & \mathbf{M}(\mathbf{C}, \mathbf{S}; \boldsymbol{\Theta}) = \mathbf{0}. \end{aligned} \tag{7}$$

Assuming that a number of measurements is available and so is the distribution at each measurement location, along with the known probability law of $\boldsymbol{\Theta}$, we wish to estimate the probability law of the source vector $\mathbf{S}$.

III. Stochastic Reduced Order Models

A Stochastic Reduced Order Model (SROM) $\tilde{\mathbf{X}}$ for a random quantity $\mathbf{X}$ is comprised of a finite set of samples $\{\tilde{\mathbf{x}}^{(1)}, \ldots, \tilde{\mathbf{x}}^{(m)}\}$ and corresponding probabilities $\mathbf{p} = \{p^{(1)}, \ldots p^{(m)}\}$ that satisfy $p^{(j)} \ge 0$ and $\sum_{j=1}^m p^{(j)} = 1$, such that the discrete distribution $\{\tilde{\mathbf{x}}^{(j)}, p^{(j)}\}_{j=1}^m$ is close to $\mathbf{X}$ in a statistical sense. The SROM is completely defined by the size m of samples and the corresponding probabilities. SROMs were first developed for propagation of uncertainty through forward models [11]. In this setting, it is assumed that a statistical description of $\mathbf{X}$ is known *a priori*. This statistical description could include its marginal distributions, moments up to order q, and correlation matrix, i.e.,

$$F_i(x_i) := P(X_i \leq x_i) \tag{8a}$$
$$\mu_i(q) := \mathbb{E}[X_i^q] \tag{8b}$$
$$\mathbf{r} := \mathbb{E}[\mathbf{X}\mathbf{X}^T]. \tag{8c}$$

The statistical quantities (8) can be estimated by

$$\tilde{F}_i(x_i) := \sum_{j=1}^{m} p^{(j)} \mathbf{1}\left(\tilde{x}_i^{(j)} \leq x_i\right) \tag{9a}$$
$$\tilde{\mu}_i(q) := \sum_{j=1}^{m} p^{(j)} (\tilde{x}_i^{(j)})^q \tag{9b}$$
$$\tilde{r}(i,k) := \sum_{j=1}^{m} p^{(j)} \tilde{x}_i^{(j)} \tilde{x}_k^{(j)}, \tag{9c}$$

where $\mathbf{1}(\cdot)$ is the indicator function.

A. SROM Construction

We wish to construct a SROM $\tilde{\mathbf{X}}$ for random vector $\mathbf{X}$ such that the statistics of $\tilde{\mathbf{X}}$ given by (9) best approximate those of $\mathbf{X}$ (8). This can be achieved by solving an optimization problem that measures the discrepancy between the SROM statistics and the true statistics. Specifically, upon defining the following error terms:

$$e_1(\tilde{\mathbf{x}}, \mathbf{p}) = \frac{1}{2} \sum_{i=1}^{d} \int_{I_i} \left(\tilde{F}_i(x_i) - F_i(x_i)\right)^2 dx_i \tag{10a}$$
$$e_2(\tilde{\mathbf{x}}, \mathbf{p}) := \frac{1}{2} \sum_{i=1}^{d} \sum_{q=1}^{\bar{q}} \left(\frac{\tilde{\mu}_i(q) - \mu_i(q)}{\mu_i(q)}\right)^2 \tag{10b}$$
$$e_3(\tilde{\mathbf{x}}, \mathbf{p}) := \frac{1}{2} \sum_{i,=1;j>i}^{d} \left(\frac{\tilde{r}(i,j) - r(i,j)}{r(i,j)}\right)^2, \tag{10c}$$

where I_i is the support of distribution function F_i and d is the dimension of the random vector $\mathbf{X}$, the SROM parameters can be selected by solving the following optimization problem:

$$\tilde{\mathbf{X}} := \underset{\{\tilde{\mathbf{x}}\},\, \mathbf{p}}{\text{argmin}} \quad (\sum_{i=1}^{3} \alpha_i e_i(\{\tilde{\mathbf{x}}\}, \mathbf{p})) \tag{11}$$
$$\text{s.t.} \quad \sum_{j=1}^{m} p^{(j)} = 1$$
$$p^{(j)} \geq 0,$$

where $\{\alpha_i \geq 0\}_{i=1}^3$ are weight coefficients controlling the influence of each error term in the objective function. The SROM model size m can be selected based on computational resources available.

B. Forward Uncertainty Propagation Using SROMs

In the setting of advection-diffusion transport, given a system of algebraic equations (3) relating the source and concentration vectors $\mathbf{s}$ and $\mathbf{c}$, consider the case where the statistics of the source parameters $\mathbf{s}$ are known such that a SROM $\tilde{\mathbf{S}} = \{\tilde{\mathbf{s}}^{(i)}, p_s^{(i)}\}_{i=1}^{m_s}$ of size m_s has been constructed. Also, the statistics of some other source of uncertainty are known such that a SROM $\tilde{\mathbf{\Theta}} = \{\tilde{\boldsymbol{\theta}}^{(j)}, p_\theta^{(j)}\}_{j=1}^{m_\theta}$ of size m_θ has also been constructed. Now, in the "forward problem", the goal is to estimate the statistics of the state variable $\mathbf{c}$. We can do so by solving the model (3) $m_s \times m_\theta$ times for the state samples $\{\tilde{\mathbf{c}}\}$ corresponding to each source sample $\{\tilde{\mathbf{s}}\}$ and parameter sample $\{\tilde{\boldsymbol{\theta}}\}$, i.e. by solving

$$\mathbf{M}(\tilde{\mathbf{c}}^{(k)}, \tilde{\mathbf{s}}^{(i)}; \tilde{\boldsymbol{\theta}}^{(j)}) = \mathbf{0} \text{ for } i = 1, \ldots, m_s \quad j = 1, \ldots, m_\theta, \tag{12}$$

where $k = (j-1) \times m_s + i$ and the corresponding probabilities of each concentration sample are

$$p_c^{(k)} = p_s^{(i)} \times p_\theta^{(j)} \text{ for } k = 1, \ldots, m_{\mathbf{s}} \times m_\theta. \tag{13}$$

The resulting concentration state samples $\{\tilde{\mathbf{c}}^{(i)}\}_{i=1}^{m_s \times m_\theta}$ along with the probabilities $\mathbf{p}_c$ given by equation (13), define a SROM $\tilde{\mathbf{C}}$ for the random concentration vector $\mathbf{C}$. Then, we can use this new SROM to estimate the statistics of the state variable using (9).

Remark III.1. *In the absence of uncertainty in other parameters ($\boldsymbol{\theta}$), the concentration SROM $\tilde{\mathbf{C}} = \{\tilde{\mathbf{c}}^{(i)}, p_c^{(i)}\}_{i=1}^{m_s}$ is formed by solving the model m_s times and keeping the original source probabilities i.e. $\mathbf{p}_c = \mathbf{p}_s$.*

The proposed SROM framework can be easily extended to handle multiple (independent) sources of uncertainty. It can be seen that propagating uncertainty with SROMs is non-intrusive in a way analogous to Monte Carlo simulation. However, SROMs are constructed by assigning non-uniform probabilities to certain samples in order to better represent the random element. Using optimization to construct the distribution with only a few samples results in much fewer model evaluations compared with sampling methods [11] - [14].

IV. SROMs for Inverse Problems

In this paper we use SROMs to solve the stochastic source identification problem in advection-diffusion transport systems under uncertainty in the concentration measurements and measurement locations. We focus on uncertainty in the measurement locations due to its practical importance for robotics applications and because it has not been considered using SROMs before. In this setting, the role of the concentration and source is reversed from the previous Section III-B. Specifically, given statistics of the observable system state $\mathbf{C}_m$ at n locations that themselves are uncertain, we wish to determine the probability law of the source $\mathbf{S}$. The general idea is to reformulate the stochastic constrained inverse problem (7) into a deterministic one using a SROM approximation to the random quantities. Specifically, we will use a "Discretize-Optimize" approach that employs the SROM approximations $\tilde{\mathbf{C}}$, $\tilde{\mathbf{S}}$, and $\tilde{\mathbf{\Theta}}$ for the random elements $\mathbf{C}$, $\mathbf{S}$, and $\mathbf{\Theta}$. We will adopt the following notation for the SROM samples using block vectors $\{\tilde{\mathbf{c}}\} := \{\tilde{\mathbf{c}}^{(1)}, \dots, \tilde{\mathbf{c}}^{(m_s \times m_\theta)}\}^T \in \mathbb{R}^N$, $\{\tilde{\mathbf{s}}\} := \{\tilde{\mathbf{s}}^{(1)}, \dots, \tilde{\mathbf{s}}^{(m_s)}\}^T \in \mathbb{R}^D$, and $\{\tilde{\boldsymbol{\theta}}\} := \{\tilde{\boldsymbol{\theta}}^{(1)}, \dots, \tilde{\boldsymbol{\theta}}^{(m_\theta)}\}^T \in \mathbb{R}^Q$, where $N = n \times m_s \times m_\theta$, $D = d \times m_s$, and $Q = q \times m_\theta$. Similarly, the model constraint (7) can be replaced with the SROM approximations giving:

$$\tilde{\mathbf{M}} := \left\{ \begin{array}{c} \mathbf{M}(\tilde{\mathbf{c}}^{(1)}, \tilde{\mathbf{s}}^{(1)}; \tilde{\boldsymbol{\theta}}^{(1)}) \\ \vdots \\ \mathbf{M}(\tilde{\mathbf{c}}^{(k)}, \tilde{\mathbf{s}}^{(i)}; \tilde{\boldsymbol{\theta}}^{(j)}) \\ \vdots \\ \mathbf{M}(\tilde{\mathbf{c}}^{(m_\mathbf{s} \times m_{\boldsymbol{\theta}})}, \tilde{\mathbf{s}}^{(m_\mathbf{s})}; \tilde{\boldsymbol{\theta}}^{(m_{\boldsymbol{\theta}})}) \end{array} \right\} = \mathbf{0} \in \mathbb{R}^N, \quad (14)$$

so that the stochastic constraint equation (7) has been decoupled into $m_s \times m_\theta$ sets of deterministic equality constraints. Using these notations, the objective function (7) can be expressed in terms of the SROM approximations as

$$\mathcal{J}(\mathbf{C}, \mathbf{S}; \mathbf{\Theta}) \approx \mathcal{J}(\tilde{\mathbf{C}}, \tilde{\mathbf{S}}; \tilde{\mathbf{\Theta}}) := \mathcal{J}(\{\tilde{\mathbf{c}}\}, \{\tilde{\mathbf{s}}\}, \{\tilde{\boldsymbol{\theta}}\}, \mathbf{p}_s, \mathbf{p}_\theta). \quad (15)$$

Since the state sample $\tilde{\mathbf{c}}^{(k)}$ is an implicit function of the parameter samples $\tilde{\mathbf{s}}^{(i)}$ and $\tilde{\boldsymbol{\theta}}^{(j)}$ by the properties of the model (3), we can bring the model constraint (14) into the objective function, i.e.,

$$\hat{\mathcal{J}}(\{\tilde{\mathbf{s}}\}, \mathbf{p}_s) := \mathcal{J}(\{\ \tilde{\mathbf{c}}(\{\tilde{\mathbf{s}}\}, \{\tilde{\boldsymbol{\theta}}\}, \mathbf{p}_\theta)\ \}, \{\tilde{\mathbf{s}}\}, \mathbf{p}_s). \quad (16)$$

Finally, we can define the SROM solution to the stochastic optimization problem (7) as:

$$\tilde{\mathbf{S}}^* := \underset{\{\tilde{\mathbf{s}}\} \in \mathcal{Q}, \mathbf{p}_s \in \mathcal{P}}{\text{argmin}} \quad \hat{\mathcal{J}}(\{\tilde{\mathbf{s}}\}, \mathbf{p}_s), \quad (17)$$

where $\mathcal{P}$ is the feasible set of probabilities

$$\mathcal{P} := \left\{ \mathbf{p}_s |\ \ \mathbf{p}_s \in \mathbb{R}^{m_\mathbf{s}}, \ \ \sum_{i=1}^{m_\mathbf{s}} p_s^{(i)} = 1 \ \text{ and } \ p_s^{(i)} \geq 0, \ \ \forall i \right\}, \quad (18)$$

and $\mathcal{Q}$ is the feasible set of source samples which can include *a priori* bounds on the location of the source within the domain of interest Ω or on the intensity of the source. In essence, we are constructing a SROM for the source $\{\tilde{\mathbf{s}}^{(i)}, p_s^{(i)}\}_{i=1}^{m_\mathbf{s}}$ such that the corresponding concentration SROM, which also depends on the uncertainty in the measurement locations captured by $\{\tilde{\boldsymbol{\theta}}^{(i)}, p_\theta^{(i)}\}_{i=1}^{m_\theta}$, is close to the observed concentration distributions in a statistical sense. The optimization variables are the source samples and probabilities $\{\tilde{\mathbf{s}}^{(i)}, \mathbf{p}_s^{(i)}\}_{i=1}^{m_\mathbf{s}}$, and the objective function will have the form of the usual SROM objective (11), however, the objective function will be minimizing discrepancy between the concentration SROM and the observed concentration statistics.

A. Solution Strategy Using SROMs

To solve (17), we need to derive the gradient of the objective function to supply to a solver. Let, $\mathbf{C}_m^q \in \mathbb{R}_+^n$ be the vector of the q'th moments of the concentration distribution at each measurement location. The optimization problem (17) finds a set of source samples and probabilities such that along with the known uncertainty in the location of the measurements according to $\tilde{\mathbf{\Theta}}$, the corresponding concentration samples according to (3) and set of probabilities (13), provide a good estimate to the observed statistics $\mathbf{C}_m$ as determined by (10).

To illustrate how the finite element model is utilized in the objective function, we present an example of the objective function (17) that utilizes a single error term from (11). Specifically, the case where $\alpha_2 = 1$ and $\alpha_1 = \alpha_3 = 0$ in (11). Before solving the inverse problem for the source, we first construct a SROM for each measurement location that adequately captures the degree of uncertainty in the location of the measurement points. The SROMs for the measurement locations enter the optimization problem by mapping the source samples to the respective locations in the set of location SROM samples and multiplying by the corresponding probability that the measurements were observed at that location. The objective function can now be defined:

$$\tilde{\mathbf{S}} := \underset{\{\tilde{\mathbf{s}}, \mathbf{p}_s\}}{\text{argmin}} \frac{1}{2} \sum_{i=1}^{n} \sum_{q=1}^{\bar{q}} \left(\frac{\mathbf{C}_m^q(i) - \sum_{j=1}^{m_s} \sum_{k=1}^{m_\theta} p_s^{(j)} p_\theta^{(k)} (\mathbf{c}_i(\mathbf{s}^{(j)}, \boldsymbol{\theta}^{(k)}))^q}{\mathbf{C}_m^q(i)} \right)^2, \quad (19)$$

where $\mathbf{c}_i(\mathbf{s}^{(j)}, \boldsymbol{\theta}^{(k)})$ is the i'th entry of measurement vector $\mathbf{c}$ corresponding to measurement location i when solving the model (3) for source $\mathbf{s}^{(j)}$ and measurement location $\boldsymbol{\theta}^{(k)}$.

We are now in position to derive the gradient of the optimization problem (19) with respect to the source parameters $\mathbf{s}^{(j)}$ and probabilities $p_s^{(j)}$:

$$\frac{\partial \hat{\mathcal{J}}}{\partial \mathbf{s}^{(j)}} = \sum_{i=1}^{n} \sum_{q=1}^{\bar{q}} \left(\frac{\mathbf{C}_m^q(i) - \sum_{j=1}^{m_s} \sum_{k=1}^{m_\theta} p_s^{(j)} p_\theta^{(k)} (\mathbf{c}_i(\mathbf{s}^{(j)}, \boldsymbol{\theta}^{(k)}))^q}{\mathbf{C}_m^q(i)} \right) \times \left(\frac{-q\, p_s^{(j)} \sum_{k=1}^{m_\theta} p_\theta^{(k)} (\mathbf{c}_i(\mathbf{s}^{(j)}, \boldsymbol{\theta}^{(k)}))^{(q-1)} (\partial \mathbf{c}_i / \partial \mathbf{s}_j)}{\mathbf{C}_m^q(i)} \right), \quad (20)$$

where $\partial \mathbf{c} / \partial \mathbf{s}^{(j)} = [\mathbf{Q}][\mathbf{K}]^{-1}[\mathbf{R}]\, \partial \hat{\mathbf{f}} / \partial \mathbf{s}^{(j)}$. The gradient $\partial \hat{\mathcal{J}} / \partial p_s^{(j)}$ is simpler and omitted for brevity. These gradients can be supplied to a deterministic solver to solve for the source samples and probabilities $\{\mathbf{s}^{(j)}, p_s^{(j)}\}_{j=1}^{m_s}$.

V. Numerical Experiments

In this section we provide numerical simulations to demonstrate the effectiveness of the proposed SROM optimization framework to identify sources in advection-diffusion transport

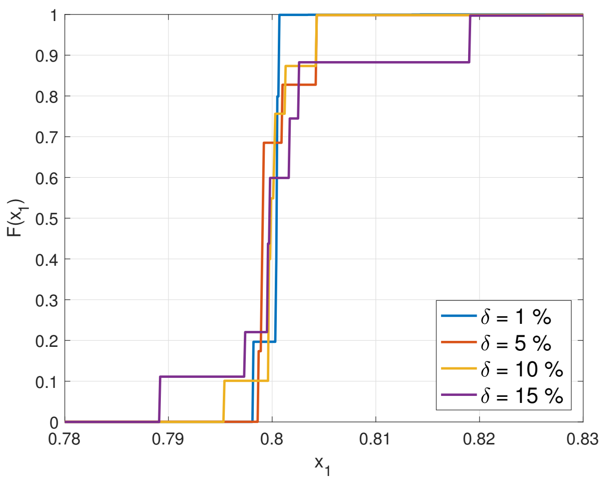

Fig. 1: CDF plot of x_1 for $\delta = 0.01, 0.05, 0.1$ and 0.15. True value of $x_1 = 0.8$.

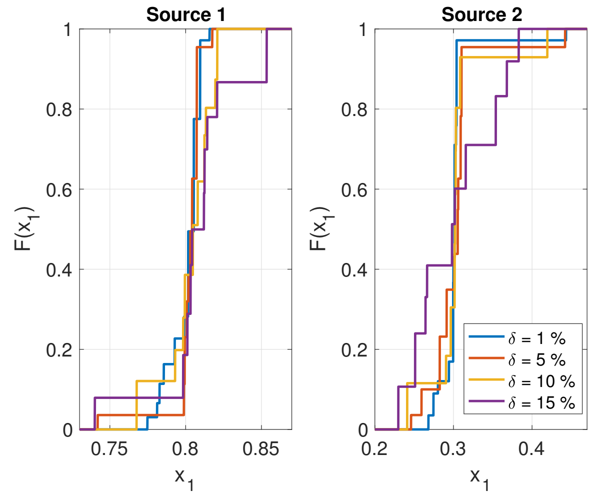

Fig. 2: CDF plots of x_{c_1} parameter for both sources. True values were $x_{c_1}^{(1)} = 0.8$ and $x_{c_1}^{(2)} = 0.3$

systems under uncertainty in the concentration measurements as well as the measurement locations. We solved the nonlinear optimization problem (17) using the MATLAB optimization toolbox with an interior-point algorithm for which we provide the analytical gradient. In all of the simulations, we consider 2D square domains with dimension 1×1. A Finite Element (FE) mesh size of $\hat{N} = 41 \times 41 = 1681$ nodes was utilized. An in-house FE code was used to build a mesh and construct matrices $[\mathbf{K}]$ and $[\mathbf{R}]$ in (2). Measurements were taken on a lattice of $n = 8$ equally spaced locations throughout the domain. We represented the source with 3 parameters, namely its center $\mathbf{x}_c \in [0,1] \times [0,1]$ and its peak intensity $\beta \in \mathbb{R}_+$ such that $\mathbf{s} = [\mathbf{x}_c, \beta]^T \in \mathbb{R}^3$. We used the following Gaussian-like function to represent the source, $f(x;\mathbf{s}) = \beta \exp\{-1/(2\gamma^2) \left\|x - x_c\right\|^2\}$, where it is assumed that the parameter γ is given. We assume that the distribution on the location of the measurements is a 2-dimensional symmetric uncorrelated Gaussian centered at the true location where the measurement statistics were generated. Without loss of generality, all of the simulations presented are diffusion dominated, i.e. $\mathbf{q} = 0$ in (1). The diffusivity of the medium is set to $D = 1$.

A. Source Identification with Known Measurement Locations

We first present a simulation with a single source in a domain with known measurement locations. The parameters of the source were set to $[x_{c_1}, x_{c_2}, \beta] = [0.8, 0.4, 1000]$. Given this source, we then generate concentration samples using Monte Carlo sampling according to (4). These samples are the data $\mathbf{C}_m$ used to solve the inverse problem. Source samples were generated randomly to initialize the problem. All of the probabilities were initialized as $p_{\mathbf{s}}^{(i)} = 1/m_{\mathbf{s}}$, i.e., a uniform distribution. For the objective function in the optimization problem, α_2 was set to 1 (with $\bar{q} = 8$) while α_1 and α_3 were set to 0 in (11). Therefore, we solved the SROM optimization problem (17) for the source samples and probabilities such that the corresponding concentration samples minimized the discrepancy between the first eight moments of the observed concentration distributions. In this simulation, $m_{\mathbf{s}} = 10$ SROM samples were used. The optimization problem was solved for $\delta = 1\%$, 5%, 10% and 15% noise levels according to (4). Plots of the SROM CDF given by (9a) for x_{c_1} of the random vector $\mathbf{s}$ is shown in Figure 1. The SROM CDF can be seen to be centered around the true value, with a higher variance for the cases where the observed concentration variance was higher. CDF plots for x_{c_2} and β are similar but omitted for brevity.

Next, we present a more complex scenario with two sources within the domain. This simulation demonstrates the ability of SROMs to handle higher dimensional parameter spaces. In this setting, the number of unknowns is six, i.e., $\mathbf{s} = [x_{c_1}^{(1)}, x_{c_2}^{(1)}, \beta^{(1)}, x_{c_1}^{(2)}, x_{c_2}^{(2)}, \beta^{(2)}]$. When the forward problem is solved in each iteration of the optimization, the right hand side vector $\hat{\mathbf{f}}$ in (2) is formed by first forming the individual source vectors and then adding them together, i.e. $\hat{\mathbf{f}} = \hat{\mathbf{f}}_1(x_{c_1}^{(1)}, x_{c_2}^{(1)}, \beta^{(1)}) + \hat{\mathbf{f}}_2(x_{c_1}^{(2)}, x_{c_2}^{(2)}\beta^{(2)})$. Again, $m_{\mathbf{s}} = 10$ SROM samples were used. Plots of the CDF for the x_{c_1} parameter of both sources is shown in Figure 2 for the same noise levels δ as in the single source case. CDF plots of x_{c_2} and β of each source are similar but omitted for brevity.

B. Uncertain Measurement Locations

In this simulation, we incorporate the uncertainty in the location of the measurements in addition to measurement noise when solving the source identification problem. This is of practical importance in robotics applications, where localization uncertainty is typically present. It is assumed that this uncertainty in the measurement locations is known ahead of time such that a SROM of size $m_{\boldsymbol{\theta}}$ for each measurement location can be constructed to reflect this uncertainty as discussed in Section IV-A. By selecting nearby points within the FE mesh, we solved a SROM optimization problem for the

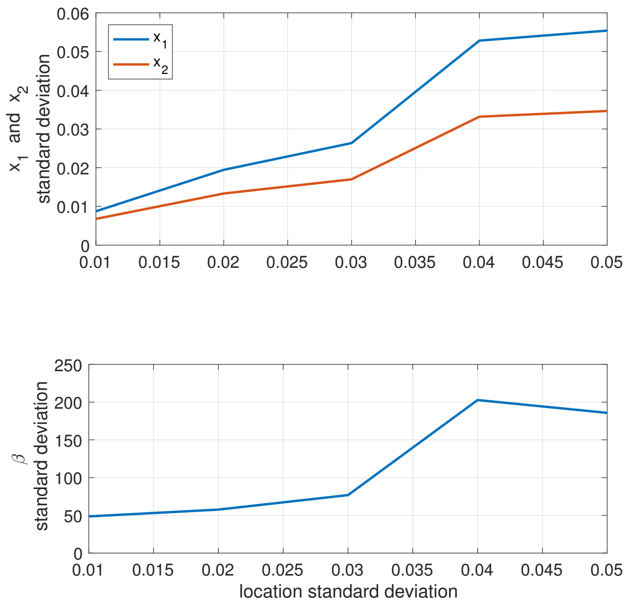

Fig. 3: x_1, x_2, and β standard deviation versus measurement location standard deviation for $n = 8$ measurements and $\delta = 1\%$ measurement noise.

probabilities of these nearby points such that the corresponding SROM statistics satisfied the prescribed distribution on the location of the measurements. Therefore, we are able to examine how the SROM solution for the source parameters changes as we increase the uncertainty in the location of the measurements.

Results are shown in Figure 3 for $m_{\mathbf{s}} = 5$ source samples and $m_{\boldsymbol{\theta}} = 5$ location samples corresponding to each measurement location with a fixed measurement noise level of $\delta = 1\%$. Again, the true source parameters were set to $[x_{c_1}, x_{c_2}, \beta] = [0.8, 0.4, 1000]$. The uncertainty in the source parameters is shown for the different prescribed standard deviations on the location (symmetric in x_1 and x_2 directions) of the measurements $\sigma_{\boldsymbol{\theta}} = [0.01,\ 0.02,\ 0.03,\ 0.04,\ 0.05]$, i.e., up to 5% of the total domain length. It can be seen that the SROM solution correctly captures the growing uncertainty in the location and intensity of the source as the uncertainty in the measurement locations grows. Although the uncertainty in the source intensity β shown in Figure 3 is not a monotonically increasing function of the the measurement location uncertainty $\sigma_{\boldsymbol{\theta}}$, the uncertainty in the source location parameter $\mathbf{x}_c$ is monotonically increasing with $\sigma_{\boldsymbol{\theta}}$ and thus reflects an overall increase in the uncertainty of the source as we become more uncertain about the location of the measurements.

VI. Conclusion

In this paper we investigated the use of Stochastic Reduced Order Models (SROMs) for solving stochastic source identification problems, specifically in advection-diffusion transport systems in steady state. Given statistics of concentration measurements within a domain, we formulated a stochastic optimization problem to solve for the probability law of the unknown source. Our approach requires significantly fewer samples and thus much less model evaluations than Monte Carlo sampling methods. We considered different sources of uncertainty, including uncertainty in the measurement locations which is of practical importance in robotics applications where localization errors are common. We also provided simulation results that demonstrated the ability of our method to effectively quantify uncertainty in the source parameters. Our results suggest that SROMs can present a potentially powerful alternative to Bayesian methods that are predominantly used to solve estimation problems in the controls and robotics literature. Future research includes introducing mobility and constructing optimal sequences of measurements to decrease estimation uncertainty.

References

[1] G. Kowadlo and R. A. Russell, "Robot odor localization: a taxonomy and survey," *The International Journal of Robotics Research*, vol. 27, no. 8, pp. 869–894, 2008.

[2] A. El Badia, T. Ha-Duong, and A. Hamdi, "Identification of a point source in a linear advection–dispersion–reaction equation: application to a pollution source problem," *Inverse Problems*, vol. 21, no. 3, p. 1121, 2005.

[3] R. Khodayi-mehr, W. Aquino, and M. M. Zavlanos, "Model-based sparse source identification," in *American Control Conference (ACC), 2015*, pp. 1818–1823, IEEE, 2015.

[4] R. Khodayi-mehr, W. Aquino, and M. M. Zavlanos, "Nonlinear reduced order source identification," in *American Control Conference (ACC), 2016*, pp. 6302–6307, IEEE, 2016.

[5] P.-S. Koutsourelakis, "A novel bayesian strategy for the identification of spatially varying material properties and model validation: an application to static elastography," *International Journal for Numerical Methods in Engineering*, vol. 91, no. 3, pp. 249–268, 2012.

[6] J. Wang and N. Zabaras, "A bayesian inference approach to the inverse heat conduction problem," *International Journal of Heat and Mass Transfer*, vol. 47, no. 17, pp. 3927–3941, 2004.

[7] X. Ma and N. Zabaras, "An efficient bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method," *Inverse Problems*, vol. 25, no. 3, p. 035013, 2009.

[8] V. A. B. Narayanan and N. Zabaras, "Stochastic inverse heat conduction using a spectral approach," *International Journal for Numerical Methods in Engineering*, vol. 60, no. 9, pp. 1569–1593, 2004.

[9] N. Zabaras and B. Ganapathysubramanian, "A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach," *Journal of Computational Physics*, vol. 227, no. 9, pp. 4697–4735, 2008.

[10] D. P. Kouri, M. Heinkenschloss, D. Ridzal, and B. G. van Bloemen Waanders, "A trust-region algorithm with adaptive stochastic collocation for pde optimization under uncertainty," *SIAM Journal on Scientific Computing*, vol. 35, no. 4, pp. A1847–A1879, 2013.

[11] M. Grigoriu, "Reduced order models for random functions. application to stochastic problems," *Applied Mathematical Modelling*, vol. 33, no. 1, pp. 161–175, 2009.

[12] J. E. Warner, M. Grigoriu, and W. Aquino, "Stochastic reduced order models for random vectors: application to random eigenvalue problems," *Probabilistic Engineering Mechanics*, vol. 31, pp. 1–11, 2013.

[13] M. Grigoriu, "Solution of linear dynamic systems with uncertain properties by stochastic reduced order models," *Probabilistic Engineering Mechanics*, vol. 34, pp. 168–176, 2013.

[14] S. Sarkar, J. E. Warner, W. Aquino, and M. D. Grigoriu, "Stochastic reduced order models for uncertainty quantification of intergranular corrosion rates," *Corrosion Science*, vol. 80, pp. 257–268, 2014.

[15] J. E. Warner, W. Aquino, and M. D. Grigoriu, "Stochastic reduced order models for inverse problems under uncertainty," *Computer methods in applied mechanics and engineering*, vol. 285, pp. 488–514, 2015.

[16] D. L. Logan, *A first course in the finite element method*. Cengage Learning, 2011.