A BAYESIAN APPROACH TO PARAMETER IDENTIFICATION IN GAS NETWORKS

SOHEIL HAJIAN, MICHAEL HINTERMÜLLER, CLAUDIA SCHILLINGS AND NIKOLAI STROGIES

Abstract. The inverse problem of identifying the friction coefficient in an isothermal semilinear Euler system is considered. Adopting a Bayesian approach, the goal is to identify the distribution of the quantity of interest based on a finite number of noisy measurements of the pressure at the boundaries of the domain. First well-posedness of the underlying non-linear PDE system is shown using semigroup theory, and then Lipschitz continuity of the solution operator with respect to the friction coefficient is established. Based on the Lipschitz property, well-posedness of the resulting Bayesian inverse problem for the identification of the friction coefficient is inferred. Numerical tests for scalar and distributed parameters are performed to validate the theoretical results.

Key words. Bayesian inversion, distributed friction coefficient, gas network/pipeline, hyperbolic PDE system.

AMS subject classifications.

1. Introduction. In many countries, the turnaround in energy policy is one of the main focus areas of political decision making and public opinion in the energy sector. In particular, the shift away from nuclear energy supply will only be possible by exploring new sustainable resources. During the transition from the current energy portfolio to one which has an emphasis on renewable energies and other energy carriers such as hydrogen, it is widely believed that an optimized use of natural gas will play a key role. This is in particular plausible when considering the currently known available gas resources, the transportability of gas over long distances, its storage capacity, and the fact that it can be traded on markets which helps an efficient distribution.

The transport of natural gas is typically achieved through a complex system of pipelines emanating at production sites or storage facilities and ending at customer locations. Mathematically and generally speaking, the gas transport in pipes is described by the compressible Euler equations, a system of hyperbolic partial differential equations (PDEs). It provides the dynamics of the density, momentum and energy of the underlying gas. As in our target application the diameter of a pipe is much smaller than the length, the study of a one-dimensional version of the PDE model is sufficient. Moreover, since the flow in pipes is usually assumed to start at a stationary state and to evolve smoothly due to industry regulations (thus preventing shock formation), the one-dimensional Euler system simplifies to isothermal equations where the unknowns are density and momentum. Now, under the assumption that the speed of the gas is significantly smaller than the speed of sound, we arrive at the following semi-linear PDE system describing the gas dynamics on a single pipe:

$$\begin{aligned} \partial_t p(x,t) + c^2 \partial_x q(x,t) &= 0 && \text{in } \Omega \times (0,T), \\ \partial_t q(x,t) + \partial_x p(x,t) &= \lambda a(p(x,t), q(x,t)) && \text{in } \Omega \times (0,T), \end{aligned} \tag{1.1}$$

where Ω is the physical domain, which is–without loss of generality–assumed to be $\Omega := (0,1)$, $T > 0$ is some finite time horizon, $p(x,t)$ is the pressure of the gas in the pipe at the location $x \in \Omega$ and at time $t \in [0,T]$, and $q(x,t)$ is the momentum density. The parameter $c > 0$ relates to the speed of sound. In (1.1), $a(\cdot,\cdot) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the friction function and $\lambda \in \mathrm{L}^\infty(\Omega)$ denotes the *friction coefficient*. Loosely speaking, λa describes the roughness of the interior walls of the pipes and influences

the transport in a decisive way. While a scalar value for λ is typically provided for newly produced pipes by manufacturers, its value and in particular its spatial distribution during operation (over time) is neither accessible to direct measurements nor known in general. For specific settings (see [12]), however, approximation formulas for scalar λ are known from the engineering literature. Under a regular operating mode and assuming usual manufacturing and quality conditions in the production of pipeline tubes, $a(\cdot,\cdot)$ can be assumed Lipschitz continuous with respect to each argument. According to the application in mind, system (1.1) is completed by the following initial and boundary conditions:

$$p(x,0) = p_0(x), q(x,0) = q_0(x) \quad \forall x \in \Omega, \tag{1.2a}$$

$$q(0,t) = g_L(t), q(1,t) = g_R(t) \quad \forall t > 0, \tag{1.2b}$$

where p_0, q_0, g_L, and g_R are given quantities. Combining (1.1) and (1.2) provides our overall model of gas flow. We refer to [12, 23] for more on this and further details on the model reduction process.

Motivated by problems of simulation or optimization of the aforementioned gas pipeline network, we are interested in the inverse problem of identifying the friction coefficient of a gas pipe from a finite set of noisy observations of the pressure drop at both ends of the pipe. Here, the pressure drop is given by

$$\delta p(t) := \left| \|p(\cdot,t)\|_{\mathrm{L}^2(0,\epsilon)} - \|p(\cdot,t)\|_{\mathrm{L}^2(1-\epsilon,1)} \right| \quad \forall t \in [0,T],$$

for $0 < \epsilon \ll 1$, which accommodates the L^1-regularity of p. For a classical inverse problem for the conditional identification of a friction law for the isothermal Euler equations we refer to the work [13] which relies on regularity of the solution and differentiability of the underlying solution operator.

Important questions in the optimization of gas networks are related to robust control of the system [3] or the study of probabilistic constraints [16], both involving the friction coefficient which is uncertain during operation as highlighted above. The associated mathematical formulations indeed depend on statistical properties of the associated uncertain quantities such as the mean value, standard deviation or even the entire distribution. As a consequence, this paper addresses the aforementioned inverse problem by employing a Bayesian approach [22, 30, 32]. Here we consider both, a finite dimensional and an infinite dimensional friction coefficient, thus necessitating the application of the respectively associated version of Bayes' rule. Adapting the infinite dimensional approach (rather than considering discretized versions of the friction coefficient only) is beneficial as it provides us with algorithms which are robust with respect to discretization refinements.

For setting up the Bayesian framework in our context, next we introduce the *uncertainty-to-observation operator* $\mathcal{G}$, which maps the underlying unknown (i.e., the friction coefficient) onto the data y. It is the composition of the solution operator of the forward problem (1.1) applied to the friction function λ and a data formation operator. Measuring δp at finitely many time instances $t_j \in [0,T]$, $j = 1, \ldots, K$, we have here

$$\mathcal{G}(\lambda) = (\delta p(t_1), \ldots, \delta p(t_K))^\top.$$

As observations are inevitably noisy we arrive at

$$y = \mathcal{G}(\lambda) + \eta,$$

where $\eta \in \mathbb{R}^K$ represents Gaussian noise with mean zero and associated covariance matrix $\Gamma \in \mathbb{R}^{K\times K}$. The pertinent probability density function (PDF) is denoted by $\rho(\cdot)$.

When λ is finite dimensional, then the probability of obtaining y for a given friction coefficient λ is $\mathbb{P}(y|\lambda) = \rho(y - \mathcal{G}(\lambda))$, which is the *likelihood* of the data. Moreover, using Bayes' theorem, we can incorporate our "prior" knowledge on λ and provide its probability given the data y by

$$\mathbb{P}(\lambda|y) \propto \mathbb{P}(y|\lambda)\mathbb{P}(\lambda). \tag{1.3}$$

Our knowledge on λ, i.e., $\mathbb{P}(\lambda)$, is called *prior* and the probability of the function λ given the data y, i.e., $\mathbb{P}(\lambda|y)$, is called *posterior*. In an infinite dimensional setting, e.g., when $\lambda \in \mathrm{L}^\infty(\Omega)$, then there is no density with respect to the Lebesgue measure the Bayes rule is understood as the Radon-Nikodym derivative of the posterior measure $\mu^y(\mathrm{d}\lambda)$ with respect to the prior measure $\mu_0(\mathrm{d}\lambda)$, i.e.,

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(\lambda) \propto \rho(y - \mathcal{G}(\lambda)). \tag{1.4}$$

For an overview on the Bayesian approach to statistics in finite dimensions we refer the reader to [4]. Bayesian inverse problems with an emphasis on modelling and computation are addressed in [22]. We also mention that the Bayesian inversion methods that will be used in this paper have been successfully applied to linear elliptic problems such as Darcy's flow in [9] and viscous incompressible flow on a two-dimensional torus in [7].

Uncertainty quantification for inverse problems has become a very active field of research over the recent years. Adopting, in this context, the Bayesian viewpoint leads to a complete characterisation of the uncertainty via the posterior distribution; see, e.g., [10, 22, 30]. In a general inverse problem setting, the goal is to recover unknown parameters from noisy measurements of system quantities. Bayesian inversion, however, interprets the unknown as a random variable and computes the conditional distribution of the unknown parameters given noisy measurements and prior distribution. It is known [30] that the latter approach is well defined in the infinite-dimensional setting. Thus, it is suitable for the identification of parameter functions belonging to some infinite dimensional Banach space.

Under certain regularity assumptions on the forward problem describing the underlying physics, well-posedness and stability results can be established for the Bayesian problem. In particular, robustness with respect to numerical approximations of the forward problem is of interest to ensure a stable inversion of the problem; see [10]. However, computational challenges arise due to the complex structure of probability measures in the high or infinite dimensional setting. In order to circumvent the curse of dimensionality, there is substantial interest in the development of dimension independent methods, i.e., algorithms which are robust with respect to the dimension of the parameter space and thus, applicable to high-dimensional real-world problems; see, e.g., [8, 11, 14, 25, 27, 28].

In this paper, we will focus on Markov-Chain-Monte-Carlo (MCMC) methods formulated in function spaces. For our focus application, the identification of the friction coefficient in an isothermal Euler system, it represents a suitable choice due to the low requirements on the regularity of the forward problem. It is well known that, in general, solutions of hyperbolic systems develop discontinuities in finite time

and therefore pose additional difficulties in the efficient treatment of uncertainties due to the lack of smoothness with respect to uncertain inputs.

As our forward (or state) system is hyperbolic, we mention the uncertainty quantification for hyperbolic differential equations with random data has been a very active field of research over the recent past; see [1, 5] and the references therein. In the data assimilation context, especially for weather forecasting applications, efficient methods for state estimation have received particular attention in recent years [2, 18, 24]. However, the quantification of uncertainties in the inverse setting has been the subject of only a very small number of publications; compare [6, 7]. None of these is related to our work.

The rest of this paper is organized as follows. In section 2 the PDE system is studied with respect to existence and Lipschitz stability. Prior modelling and the Bayesian framework are the subjects of section 3. The numerical realisation of the forward problem is considered in section 4, and numerical results are provided in section 5.

2. Properties of the underlying PDE system. In order to establish well-posedness of the Baysian inverse problem, it is sufficient to show that the underlying PDE is well-posed and that the solution operator is Lipschitz continuous with respect to the friction coefficient.

It is convenient to reformulate the PDE (1.1) such that the flux function $q(x,t)$ has vanishing traces. For this purpose, let us define $q = \tilde{q} + \hat{q}$ where $\tilde{q}(x,t) := xg_L(t) + (1-x)g_R(t)$. Then, since $\hat{q}(0,t) = \hat{q}(1,t) = 0$, we can write the underlying PDE (1.1) in the following abstract form of a dynamical system:

$$\begin{aligned} &\boldsymbol{u}'(t) + A\boldsymbol{u}(t) = \boldsymbol{f}(\lambda, \boldsymbol{u}(t), t) \quad (t > 0), \\ &\boldsymbol{u}(0) = \boldsymbol{u}_0, \end{aligned} \tag{2.1}$$

where $\boldsymbol{u}(t) := (p(\cdot,t), \hat{q}(\cdot,t))^\top$ and $\boldsymbol{u}_0 := (p_0(\cdot), q_0(\cdot) - \tilde{q}(\cdot,0))^\top$. Here, $A : V \to L$ is given by

$$A := \begin{bmatrix} 0 & c^2\partial_x \\ \partial_x & 0 \end{bmatrix}, \tag{2.2}$$

where $V := \mathrm{H}^1(\Omega) \times \mathrm{H}^1_0(\Omega)$, $L := \mathrm{L}^2_{\mathrm{w}}(\Omega) \times \mathrm{L}^2(\Omega)$ with the weighted L^2-space L_{w} defined by

$$\mathrm{L}^2_{\mathrm{w}}(\Omega) := \Big\{ v : \Omega \to \mathbb{R} : \int_\Omega \frac{v^2(x)}{c^2} \mathrm{d}x < \infty \Big\}.$$

The right-hand side $\boldsymbol{f} : X \times L \times [0,T] \to L$, where $X := \mathrm{L}^\infty(\Omega)$, is given by

$$\boldsymbol{f}(\lambda, \boldsymbol{u}, t) := \begin{pmatrix} -c^2(g_L(t) - g_R(t)) \\ \lambda\, a(p, \hat{q} + \tilde{q}) - xg_L'(t) - (1-x)g_R'(t) \end{pmatrix}. \tag{2.3}$$

Note that properties of boundary conditions as well as the friction function $a(\cdot,\cdot)$ have an impact on $\boldsymbol{f}(\cdot,\cdot,\cdot)$. In the following lemma we state continuity properties of $\boldsymbol{f}$.

LEMMA 2.1. *Let $g_L(t), g_R(t) \in \mathrm{C}^1([0,T])$, and suppose that $a(\cdot,\cdot) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous with respect to its arguments. Then $\boldsymbol{f} : X \times L \times [0,T] \to L$ is continuous in time and Lipschitz continuous in L for a fixed $\lambda \in X$. Moreover, if g_L' and g_R' are Lipschitz continuous on $[0,T]$, then $\boldsymbol{f}$ is Lipschitz continuous in both variables.*

Proof. We first prove Lipschitz continuity with respect to $\boldsymbol{u}$. Note that for a fixed time t and λ we have

$$\boldsymbol{f}(\lambda, \boldsymbol{u}_1, t) - \boldsymbol{f}(\lambda, \boldsymbol{u}_2, t) = \begin{pmatrix} 0 \\ \lambda[a(p_1, \hat{q}_1 + \tilde{q}) - a(p_2, \hat{q}_2 + \tilde{q})] \end{pmatrix},$$

for all $\boldsymbol{u}_1 = (p_1, \hat{q}_1) \in L$ and $\boldsymbol{u}_2 = (p_2, \hat{q}_2) \in L$. Therefore we get

$$\begin{aligned} \|\boldsymbol{f}(\lambda, \boldsymbol{u}_1, t) - \boldsymbol{f}(\lambda, \boldsymbol{u}_2, t)\|_L &= \|\lambda[a(p_1, \hat{q}_1 + \tilde{q}) - a(p_2, \hat{q}_2 + \tilde{q})]\|_{\mathrm{L}^2(\Omega)} \\ &\leq C_L \|\lambda\|_{\mathrm{L}^2(\Omega)} \|u_1 - u_2\|_L, \end{aligned}$$

where C_L is the Lipschitz constant of $a(\cdot, \cdot)$. Similarly, continuity in time follows from

$$\begin{aligned} \|\boldsymbol{f}(\lambda, \boldsymbol{u}, t_1) - \boldsymbol{f}(\lambda, \boldsymbol{u}, t_2)\|_L \leq \; & \max(1, C_L \|\lambda\|_{\mathrm{L}^2(\Omega)}) \Big(|g_L(t_1) - g_L(t_2)| + |g_L'(t_1) - g_L'(t_2)| \\ & + |g_R'(t_1) - g_R'(t_2)| + |g_R(t_1) - g_R(t_2)| \Big), \end{aligned}$$

with $t_1, t_2 \in [0, T]$. Since g_L and g_R are continuously differentiable we conclude that $\boldsymbol{f}$ is continuous in time. Finally, if both $g_L'(t)$ and $g_R'(t)$ are Lipschitz continuous, then we conclude that $\boldsymbol{f}$ is Lipschitz continuous in both variables. ∎

From the definition of the dynamical system in (2.1), it is natural to look for solutions $\boldsymbol{u} \in \mathrm{C}^1([0, T]; L) \cap \mathrm{C}^0([0, T]; V)$ which are called classical solutions in the context of semigroup theory; see [26, Definition 4.2.1]. There are, however, two other notions of solution associated with (2.1): *strong solutions* ([26, Definition 4.2.8]) and *mild solutions* ([26, Definition 4.2.3]). We will work in this paper with strong solutions of (2.1):

Definition 2.2 (strong solutions). *A function $\boldsymbol{u}$ which is differentiable almost everywhere on $[0, T]$ such that $\boldsymbol{u}' \in \mathrm{L}^1(0, T; L)$ is called a strong solution of the initial value problem (2.1) if $\boldsymbol{u}(0) = \boldsymbol{u}_0$ and $\boldsymbol{u}' + A\boldsymbol{u}(t) = \boldsymbol{f}(\lambda, \boldsymbol{u}(t), t)$ almost everywhere (a.e). on $(0, T]$.*

For the moment consider that the right-hand side is fixed, i.e.,

$$\begin{aligned} &\boldsymbol{u}'(t) + A\boldsymbol{u}(t) = \boldsymbol{f}(t) \quad (t > 0), \\ &\boldsymbol{u}(0) = \boldsymbol{u}_0, \end{aligned} \tag{2.4}$$

for a given $\boldsymbol{f}(t)$. We then seek $\boldsymbol{u}(t) \in \mathrm{C}([0, T]; L) \cap \mathrm{L}^1(0, T; V)$ with $\boldsymbol{u}' \in \mathrm{L}^1(0, T; L)$ such that (2.4) is satisfied a.e. in $[0, T]$. Semigroup theory is used to establish existence and uniqueness of the solution as A is non-coercive. For this purpose, recall the following definition (see [15, Chapter 6.3]):

Definition 2.3 (Monotone and maximal operator). *The operator $A : V \to L$ is said to be monotone* if and only if *for all $u \in V$, $(Au, u)_L \geq 0$. Moreover, A is said to be maximal* if and only if *for all $\hat{f} \in L$, there exists $u \in V$ such that $u + Au = \hat{f}$. If A is monotone, L is reflexive, and*

$$\|Au\|_L \geq c_1 \|u\|_V - c_2 \|u\|_L \quad \forall u \in V, \tag{2.5}$$

for some $c_1, c_2 > 0$ then we can conclude that A is maximal.

Note that from the definition of A, as well as the spaces L and V, we have

$$(Av, v)_L = \int_\Omega p\, \partial_x q + q\, \partial_x p = (p\, q)|_{\partial\Omega} = 0.$$

Since A is monotone, L is reflexive and (2.5) holds trivially due to the definition of L and V, we conclude that A is also maximal. For what follows, let $D(A)$ denote the domain of A. Note that $D(A)$ is a linear subspace of L. Now, if the operator $A : D(A) \subset L \to L$ is maximal and monotone then it holds that

1. $D(A)$ is dense in L;
2. the graph of A is closed;
3. $\forall \eta > 0$, $I + \eta A \in \mathcal{L}(D(A); L)$ is bijective and $\|(I + \eta A)^{-1}\|_{\mathcal{L}(L;L)} \leq 1$;

see [15, Lemma 6.51] and the references therein. The properties 1.-3. enable us to apply the Hille-Yosida theorem to show that A generates a C_0-semigroup of contractions $\mathbb{T}(t)$, for $t > 0$. Then the function $\boldsymbol{u} \in \mathrm{C}([0,T]; L)$ is a *mild solution* of (2.1) for $\boldsymbol{u}_0 \in L$ and $\boldsymbol{f}(\lambda, \boldsymbol{u}(t), t) \in \mathrm{L}^1((0,T); L)$ if it satisfies

$$\boldsymbol{u}(t) = \mathbb{T}(t)\boldsymbol{u}_0 + \int_0^t \mathbb{T}(t-s)\boldsymbol{f}(\lambda, \boldsymbol{u}(s), s)\mathrm{d}s \quad \forall t \in [0,T]. \tag{2.6}$$

The following theorem states existence and uniqueness of a mild solution which will be used later on to establish the analogous result for of strong solutions. See [26, Theorem 1.2, Chapter 6] for a proof.

THEOREM 2.4. *Let $\boldsymbol{f} : X \times L \times [0,T] \to L$ be continuous in t on $[0,T]$ and uniformly Lipschitz continuous on L. Then for every $\boldsymbol{u}_0 \in L$, there exists a unique mild solution $\boldsymbol{u} \in \mathrm{C}([0,T]; L)$ satisfying (2.6).*

We now establish a uniqueness and existence of a strong solution to (2.1) as well as boundedness and Lipschitz continuity of the solution with respect the friction coefficient.

THEOREM 2.5. *Let $\boldsymbol{f} : X \times L \times [0,T] \to L$ be uniformly Lipschitz continuous in time and on L for a given $\lambda \in X$, i.e., for some $C > 0$ depending on λ we have*

$$\|\boldsymbol{f}(\lambda, \boldsymbol{v}_1, t_1) - \boldsymbol{f}(\lambda, \boldsymbol{v}_2, t_2)\|_L \leq C(\|\lambda\|_X)(|t_1 - t_2| + \|\boldsymbol{v}_1 - \boldsymbol{v}_2\|_L) \quad \forall t_1, t_2 \in [0,T],$$

for all $\boldsymbol{v}_1, \boldsymbol{v}_2 \in L$. Then, for $\boldsymbol{u}_0 \in D(A)$ there exists a unique strong solution $\boldsymbol{u} \in \mathrm{L}^1(0,T;V) \cap \mathrm{C}([0,T]; L)$ and $\boldsymbol{u}' \in \mathrm{L}^1(0,T;L)$ satisfying (2.1) a.e. in time and

$$\|\boldsymbol{u}(t)\|_L \leq C(T, \boldsymbol{u}_0, \lambda) \quad \forall t \in [0,T]. \tag{2.7}$$

Moreover, if $\|\lambda\|_X \leq c$ for $\lambda \in X$ and $\boldsymbol{f}$ is (locally) Lipschitz continuous with respect to λ, then we have

$$\|p_1(t) - p_2(t)\|_{\mathrm{L}^2(\Omega)} \leq C\|\lambda_1 - \lambda_2\|_X \quad \forall t \in [0,T]. \tag{2.8}$$

Proof. Let $\boldsymbol{u} \in \mathrm{C}([0,T]; L)$ be a mild solution of (2.1), $\|\mathbb{T}(t)\|_L \leq M$ and $\|\boldsymbol{f}(\lambda, \boldsymbol{u}(t), t)\|_L \leq N$ for $t \in [0,T]$. For $h \in [0,t]$ we obtain

$$\begin{aligned} \boldsymbol{u}(t+h) - \boldsymbol{u}(t) &= \mathbb{T}(t+h)\boldsymbol{u}_0 - \mathbb{T}(t)\boldsymbol{u}_0 \\ &\quad + \textstyle\int_0^t \mathbb{T}(t-s)[\boldsymbol{f}(\lambda, \boldsymbol{u}(s+h), s+h) - \boldsymbol{f}(\lambda, \boldsymbol{u}(s), s)]\mathrm{d}s \\ &\quad + \textstyle\int_0^h \mathbb{T}(t+h-s)\boldsymbol{f}(\lambda, \boldsymbol{u}(s), s)\mathrm{d}s. \end{aligned}$$

Taking norms on both sides, applying the triangle inequality and estimating yield

$$\begin{aligned} \|\boldsymbol{u}(t+h) - \boldsymbol{u}(t)\|_L &\leq hM\|A\boldsymbol{u}_0\|_L + MC\textstyle\int_0^t \|\boldsymbol{u}(s+h) - \boldsymbol{u}(s)\|_L \mathrm{d}s + hMN \\ &\leq C'h + MC\textstyle\int_0^t \|\boldsymbol{u}(s+h) - \boldsymbol{u}(s)\|_L \mathrm{d}s, \end{aligned}$$

where we have used that $\mathbb{T}(t+h)\boldsymbol{u}_0 - \mathbb{T}(t)\boldsymbol{u}_0 = \int_t^{t+h} \mathbb{T}(s)A\boldsymbol{u}_0 \mathrm{d}s$ (see [26, Theorem 1.2.4]). Then Grönwall's inequality implies

$$\|\boldsymbol{u}(t+h) - \boldsymbol{u}(t)\|_L \leq C'' \exp(T\,M\,C)\,h,$$

which shows that $\boldsymbol{u}(t)$ is Lipschitz continuous (and it also proves (2.7)). This, combined with Lipschitz continuity of $\boldsymbol{f}$, implies that $t \mapsto \boldsymbol{f}(\lambda, \boldsymbol{u}(t), t)$ is Lipschitz continuous on $[0, T]$. Since L is reflexive, $\boldsymbol{u}_0 \in D(A)$, and $\boldsymbol{f}(\lambda, \boldsymbol{u}(t), t)$ is Lipschitz continuous, [26, Corollary 4.2.11] implies the existence of a unique strong solution on $[0, T]$ for the following problem

$$\begin{aligned} &\boldsymbol{v}'(t) + A\boldsymbol{v}(t) = \boldsymbol{f}(\lambda, \boldsymbol{u}(t), t) \quad (t \geq 0), \\ &\boldsymbol{v}(0) = \boldsymbol{u}_0. \end{aligned}$$

Since a strong solution is a mild solution, we have

$$\boldsymbol{v}(t) = \mathbb{T}(t)\boldsymbol{u}_0 + \int_0^t \mathbb{T}(t-s)\boldsymbol{f}(\lambda, \boldsymbol{u}(s), s)\mathrm{d}s = \boldsymbol{u}(t),$$

and, thus, $\boldsymbol{u}(t)$ is a strong solution to (2.1). Then by [26, Theorem 4.2.9], we have $\boldsymbol{u} \in \mathrm{L}^1(0, T; V)$ and $\boldsymbol{u}' \in \mathrm{L}^1(0, T; L)$.

Suppose $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are solutions obtained from two friction coefficients, λ_1 and λ_2, respectively. Then for their difference we have, using the definition of mild solutions, that

$$\begin{aligned} \|\boldsymbol{u}_1(t) - \boldsymbol{u}_2(t)\|_L &\leq M \textstyle\int_0^t \|\boldsymbol{f}(\lambda_1, \boldsymbol{u}_1(s), s) - \boldsymbol{f}(\lambda_2, \boldsymbol{u}_2(s), s)\|_L \,\mathrm{d}s \\ &\leq M\|\lambda_1 - \lambda_2\|_X C(\boldsymbol{u}_1(t)) + \|\lambda_2\|_X \textstyle\int_0^T \|\boldsymbol{u}_1(s) - \boldsymbol{u}_2(s)\|_L \,\mathrm{d}s \\ &\leq c_1\|\lambda_1 - \lambda_2\|_X + c_2 \textstyle\int_0^t \|\boldsymbol{u}_1(s) - \boldsymbol{u}_2(s)\|_L \mathrm{d}s. \end{aligned}$$

Then an application of Grönwall inequality yields Lipschitz continuity of the pressure difference with respect to friction coefficient, i.e.,

$$\|p_1(t) - p_2(t)\|_{\mathrm{L}^2(\Omega)} \leq \|\boldsymbol{u}_1(t) - \boldsymbol{u}_2(t)\|_L \leq c_3\|\lambda_1 - \lambda_2\|_X.$$

This completes the proof. □

From the definition of $\boldsymbol{f}$ in (2.3), Lemma 2.1 and Theorem 2.5, we obtain the following sufficient conditions on the boundary data g_L, g_R and the friction function such that the PDE in (1.1) admits a unique strong solution and the Lipschitz continuity result.

COROLLARY 2.6. *Let $g_L(t), g_R(t) \in \mathrm{C}^1([0, T])$ and g_L', g_R' be Lipschitz continuous on $[0, T]$. Moreover, suppose that $a(\cdot, \cdot)$ is uniformly Lipschitz continuous with respect to its arguments. Then, for every $(p_0, q_0) \in V$, there exists a unique strong solution $(p, q) \in \mathrm{C}([0, T]; L) \cap \mathrm{L}^1(0, T; V)$ with $(p', q') \in \mathrm{L}^1(0, T; L)$ to (1.1). Furthermore we have*

$$\|(p(t), q(t))\|_L \leq C(T) \quad \forall t \in [0, T],$$

and if $\|\lambda\|_X \leq C'$ uniformly, then it holds that

$$\|p_1(t) - p_2(t)\|_{\mathrm{L}^2(\Omega)} \leq C\|\lambda_1 - \lambda_2\|_X \quad \forall t \in [0, T].$$

3. Prior modelling and Bayesian inverse problem. In this section we construct probability measures on a function space. It is natural to use separable Banach spaces to define random functions using a countable infinite sequence in a Banach space. Let $\{\phi_j\}_{j=1}^{\infty}$ denote an infinite sequence in a Banach space X with the norm $\|\cdot\|_X$ on a bounded domain $\Omega \subset \mathbb{R}$. The functions are assumed to be normalized, i.e., $\|\phi_j\|_X = 1$. We then define the randomized function

$$\lambda = m_0 + \sum_{j=1}^{\infty} \lambda_j \phi_j, \tag{3.1}$$

where $m_0 \in X$ (not necessarily normalized) and $\{\lambda_j\}_{j=1}^{\infty}$ are randomized numbers defined by $\lambda_j = \gamma_j \xi_j$. Here $\{\gamma_j\}_{j=1}^{\infty}$ is a deterministic sequence, $\{\xi_j\}_{j=1}^{\infty}$ is an independent and identically distributed (i.i.d.) random sequence and we assume that ξ_j has mean zero for all $j \in \mathbb{N}$. For $N \in \mathbb{N}$, we furthermore define the truncated series by

$$\lambda^N = m_0 + \sum_{j=1}^{N} \lambda_j \phi_j. \tag{3.2}$$

As before, we choose $X = \mathrm{L}^\infty(\Omega)$, $\gamma = \{\gamma_j\}_{j=1}^{\infty} \in \ell^1$, and $\xi_j \sim U([-1,1])$ for all $j \in \mathbb{N}$. Moreover, we assume that there exist positive constants $m_{\min} \leq m_{\max}$ and $\delta > 0$ such that

$$\operatorname{ess\,inf}_{x\in\Omega} m_0(x) \geq m_{\min}, \quad \operatorname{ess\,sup}_{x\in\Omega} m_0(x) \leq m_{\max}, \quad \|\gamma\|_{\ell^1} = \tfrac{\delta}{1+\delta} m_{\min}.$$

Since X is not separable, we work with the closure of the linear span of functions $(m_0, \{\phi_j\}_{j=1}^{\infty})$ with respect to $\|\cdot\|_X$. The resulting space is denoted by $\mathcal{X}$ in what follows. The next result, taken from [10, Theorem 2.1], states that $(\mathcal{X}, \|\cdot\|_X)$ is a Banach space.

THEOREM 3.1. *The following holds $\mathbb{P}$-almost surely: The sequence of functions $\{\lambda^N\}_{N=1}^{\infty}$ given by (3.2) is Cauchy in $\mathcal{X}$ and the limiting function λ given by (3.1) satisfies*

$$\frac{1}{1+\delta} m_{\min} \leq \lambda(x) \leq m_{\max} + \frac{\delta}{1+\delta} m_{\min} \quad \textit{for almost every } x \in \Omega.$$

3.1. Bayesian inverse problem. According to our assumptions we have noisy observations of the pressure drop, here denoted by $y = \{y_j\}_{j=1}^{K} \in \mathbb{R}^K$, for $K \in \mathbb{N}$, at times $t_j \in (0,T]$ for $j = 1, \ldots, K$, at our disposition. Now, let $\mathcal{Y} := \mathbb{R}^K$, with norm $|\cdot|$, and $\mathcal{X}$ from above. Then the uncertainty-to-observation operator $\mathcal{G}$ is defined by

$$\mathcal{G} : \mathcal{X} \to \mathcal{Y}, \quad \mathcal{G}(\lambda) := (\delta p(t_1), \ldots, \delta p(t_K))^\top. \tag{3.3}$$

Note that Corollary 2.6 implies that

$$|\mathcal{G}(\lambda_1) - \mathcal{G}(\lambda_2)| \leq C \|\lambda_1 - \lambda_2\|_{\mathrm{L}^\infty(\Omega)}. \tag{3.4}$$

Moreover, Theorems 2.5 and 3.1 provide a constant $C' > 0$ such that

$$|\mathcal{G}(\lambda)| \leq C'\, K^{1/2}. \tag{3.5}$$

We use an additive linear noise model in our observations, i.e.,

$$y = \mathcal{G}(\lambda) + \eta,$$

where $\eta = \{\eta_j\}_{j=1}^{K}$ is Gaussian observation noise with mean zero and positive definite covariance matrix $\Gamma \in \mathbb{R}^{K\times K}$. In general, the observation operator is a map $\mathcal{G} : \mathbb{X} \to \mathcal{Y}$, where we consider $\mathbb{X}$ either finite-dimensional ($\mathbb{X} = \mathbb{R}^M, M \in \mathbb{N}$) or infinite dimensional ($\mathbb{X} = \mathcal{X}$) and $\mathcal{Y} = \mathbb{R}^K$. In the case where $\mathbb{X}$ is finite dimensional, the posterior distribution is obtained from (1.3). For infinite-dimensional $\mathbb{X}$ the relation

(1.4) between the posterior measure and the prior measure based on the Radon-Nikodym derivative yields

$$\frac{d\mu^y}{d\mu_0}(\lambda) = \frac{1}{Z}\exp\big(-\Phi(\lambda, y)\big), \tag{3.6a}$$

$$Z(y) = \int_X \exp\big(-\Phi(\lambda, y)\big) d\mu_0(\lambda), \tag{3.6b}$$

where $Z = Z(y)$ is a normalization constant, $\Phi(\lambda, y) := \frac{1}{2}\big|\Gamma^{-1/2}\big(y - \mathcal{G}(\lambda)\big)\big|^2$, and $|\cdot|$ is the usual Euclidean norm.

The following, rather general conditions on a function Ψ and the prior μ_0 are sufficient to guarantee that μ^y is a well-defined probability measure on $\mathbb{X}$ if Ψ replaces Φ in (3.6). We first state the conditions and then check whether they are satisfied by our Φ.

ASSUMPTION 3.2. *Let Λ and Y be Banach spaces. The function $\Psi : \Lambda \times Y \to \mathbb{R}$ satisfies the following conditions.*

(i) For every $\epsilon > 0$ and $r > 0$ there is $\mathfrak{M} = \mathfrak{M}(\epsilon, r) \in \mathbb{R}$, such that for all $\lambda \in \Lambda$, and for all $y \in Y$ with $\|y\|_Y < r$, it holds that

$$\Psi(\lambda, y) \geq \mathfrak{M} - \epsilon\|\lambda\|_\Lambda^2.$$

(ii) For every $r > 0$ there exists $\mathfrak{K} = \mathfrak{K}(r) > 0$ such that for all $\lambda \in \Lambda$, $y \in Y$ with $\max\{\|\lambda\|_\Lambda, \|y\|_Y\} < r$

$$\Psi(\lambda, y) \leq \mathfrak{K}.$$

(iii) For every $r > 0$ there exists $\mathfrak{R} = \mathfrak{R}(r) > 0$ such that for all $\lambda_1, \lambda_2 \in \Lambda$ and $y \in Y$ with $\max\{\|\lambda_1\|_\Lambda, \|\lambda_2\|_\Lambda, \|y\|_Y\} < r$

$$|\Psi(\lambda_1, y) - \Psi(\lambda_2, y)| \leq \mathfrak{R}\|\lambda_1 - \lambda_2\|_\Lambda.$$

(iv) For every $\epsilon > 0$ and $r > 0$, there is $\mathfrak{C} = \mathfrak{C}(\epsilon, r) \in \mathbb{R}$ such that for all $y_1, y_2 \in Y$ with $\max\{\|y_1\|_Y, \|y_2\|_Y\} < r$ and for every $\lambda \in \Lambda$

$$|\Psi(\lambda, y_1) - \Psi(\lambda, y_2)| \leq \exp(\epsilon\|\lambda\|_\Lambda^2 + \mathfrak{C})\|y_1 - y_2\|_Y.$$

Using Corollary 2.6, we show in the following proposition that

$$\Phi : \mathbb{X} \times \mathcal{Y} \to \mathbb{R}, \quad \text{with} \quad \Phi(\lambda, y) = \frac{1}{2}|\Gamma^{-1/2}(y - \mathcal{G}(\lambda))|^2 \tag{3.7}$$

indeed satisfies Assumption 3.2.

PROPOSITION 3.3. *Let $\boldsymbol{u} = (p, q)$ be the strong solution of (2.1), given $\boldsymbol{u}_0 \in V$ and $\lambda \in \mathbb{X}$ such that $\|\lambda\|_X \leq C$ for some $C > 0$. Moreover, let the observation operator $\mathcal{G}$ be defined as in (3.3). Then the likelihood function Φ as defined in* (3.7) *satisfies Assumption 3.2.*

Proof. In view of Assumption 3.2 we set $\Lambda := \mathbb{X}$, $Y := \mathcal{Y}$, and $\Psi \equiv \Phi$. Assumption (i) is satisfied trivially since $\Phi(\lambda, y) \geq 0$. For assumption (ii), using (3.5), we have

$$\Phi(\lambda, y) \leq \frac{1}{2}\|\Gamma^{-1}\|_{\mathcal{L}(\mathcal{Y};\mathcal{Y})}(\|y\|_\mathcal{Y} + \|\mathcal{G}(\lambda)\|_\mathcal{Y}) \leq C.$$

for all $y \in \mathcal{Y}$ and $\lambda \in \mathbb{X}$ satisfying $\max\{\|y\|_{\mathcal{Y}}, \|\lambda\|_{\mathbb{X}}\} \leq \mathfrak{R}$. Assumption (iii) is fulfilled by the Lipschitz continuity of the pressure drop, i.e.,

$$\begin{aligned} |\Phi(\lambda_1, y) - \Phi(\lambda_2, y)| &= \left| \langle \mathcal{G}(\lambda_1) - \mathcal{G}(\lambda_2), 2y + \mathcal{G}(\lambda_1) + \mathcal{G}(\lambda_2) \rangle_{\Gamma^{-1}} \right| \\ &\leq \|\Gamma^{-1}\|_{\mathcal{L}(\mathcal{Y};\mathcal{Y})} \|G(\lambda_1) - G(\lambda_2)\|_{\mathcal{Y}} \|2y + \mathcal{G}(\lambda_1) + \mathcal{G}(\lambda_2)\|_{\mathcal{Y}} \\ &\leq C \|\Gamma^{-1}\|_{\mathcal{L}(\mathcal{Y};\mathcal{Y})} \|G(\lambda_1) - G(\lambda_2)\|_{\mathcal{Y}} \\ &\leq C' \|\Gamma^{-1}\|_{\mathcal{L}(\mathcal{Y};\mathcal{Y})} \|\lambda_1 - \lambda_2\|_{\mathbb{X}}. \end{aligned}$$

where $\langle \cdot, \cdot \rangle_{\Gamma^{-1}}$ is the Euclidean inner product induced by matrix Γ^{-1}. The proof of assumption (iv) is similar to the one of assumption (iii). ∎

The following theorem asserts well-posedness of the posterior measure. It rests on the results in [10, 31].

THEOREM 3.4. *Let $\boldsymbol{u} = (p, q)$ be the strong solution of (2.1), given $\boldsymbol{u}_0 \in V$ and the prior measure μ_0 according to* (3.1)*. The uncertainty-to-observation map $\mathcal{G}$ is as in* (3.3)*. Then, the posterior measure μ^y satisfying* (3.6) *with* (3.7) *is a well-defined probability measure on $\mathbb{X}$ for each observation $y \in \mathcal{Y} = \mathbb{R}^K$. Furthermore, the inference is well-posed with respect to perturbations in the data, i.e., there exists a constant $C \geq 0$ such that*

$$d_H(\mu^y, \mu^{\tilde{y}}) \leq C|y - \tilde{y}|, \quad \forall y, \tilde{y} \in \mathcal{Y},$$

where

$$d_H(\mu^y, \mu^{\tilde{y}}) = \Big[\int_{\mathbb{X}} \Big(\sqrt{\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(\lambda)} - \sqrt{\frac{\mathrm{d}\mu^{\tilde{y}}}{\mathrm{d}\mu_0}(\lambda)} \Big)^2 \mathrm{d}\mu_0(\lambda) \Big]^{1/2}$$

denotes the Hellinger distance of the measures $\mu^y, \mu^{\tilde{y}}$.

Proof. By Proposition 3.3, the likelihood function Φ satisfies Assumption 3.2. The local Lipschitz continuity of the potential implies the measurability of Φ with respect to the product measure $\nu_0(\mathrm{d}y, \mathrm{d}\lambda) = \rho(\mathrm{d}y)\mu_0(\mathrm{d}\lambda)$, where $\rho = \mathcal{N}(0, \Gamma)$, i.e., $\eta \sim \rho$. The boundedness of Φ further implies that $Z(y) > 0$. By [10, Theorem 3.4], the posterior μ^y is well-defined and fulfills (3.6). The boundedness of the potential implies $\exp(-M + \epsilon\|\lambda\|_{\mathbb{X}}^2)(1 + \exp(\epsilon\|\lambda\|_{\mathbb{X}}^2 + C)^2) \in L^1_{\mu_0}(\mathbb{X}, \mathbb{R})$ for every $r, \epsilon > 0$, with constants M, C given in Assumption 3.2. Here, $L^1_{\mu_0}(\mathbb{X}, \mathbb{R})$ denotes the Bochner space of all measurable functions $f : \mathbb{X} \to \mathbb{R}$ with $\int_{\mathbb{X}} |f(\lambda)| \mathrm{d}\mu_0(\lambda) < \infty$. By [10, Theorem 4.5], the Lipschitz continuity of the posterior follows. ∎

4. Discretization of the forward problem and sampling. The statistical inverse problem posed in Section 1 is now solved by using a discretization of the forward problem and the MCMC algorithm. More specifically, our goal is to reconstruct the friction coefficient $\lambda(x)$ from noisy observations of the pressure drop.

As motivated above, we model the truncated friction coefficient according to (3.2) yielding

$$\lambda^N(x) = m_0(x) + \sum_{k=-N}^{N} \lambda_k \phi_k(x),$$

where $m_0 \in \mathrm{L}^\infty(\Omega)$ and $\phi_k(x)$ are scaled Fourier coefficients, i.e.,

$$\phi_k(x) = a_k \exp(i\, kx) \quad \text{for all } k \in \mathbb{Z},$$

$\{\lambda_k\}$ are complex coefficients that satisfy $\lambda_k^* = \lambda_{-k}$ and $\{a_k\}$ are real coefficients. In order to ensure that $\lambda^N(x) \geq \nu > 0$ for all $x \in \Omega$, we choose a_k with

$$\sum_{k\in\mathbb{Z}} |a_k| < \infty, \qquad \sum_{k\in\mathbb{Z},|k|>j} |a_k| < C\, j^{-\nu}, \tag{4.1}$$

for some $C > 0$ and $\nu > 0$; see [21, 29]. In particular, we use $a_k := \bar{a}|k|^2$ in order to ensure that (4.1) is satisfied. Then the mean function $m_0(x)$ can be chosen to ensure positivity of λ^N.

For the convenience of implementation, we reformulate the prior model as

$$\lambda^N(x) = m_0(x) + \sum_{k=0}^{N} a_k \Big(\alpha_k \sin(2\pi k x) + \beta_k \cos(2\pi k x)\Big), \tag{4.2}$$

where we have the following relation between $\{\lambda_k\}$ and $\{\alpha_k, \beta_k\}$

$$\lambda_k = \begin{cases} \frac{1}{2}(\beta_k - i\alpha_k) & \text{for } k > 0, \\ \beta_0 & \text{for } k = 0, \\ \lambda_{-k}^* & \text{for } k < 0. \end{cases}$$

Here λ_{-k}^* denotes the complex conjugate of λ_{-k}. Then, the truncated friction coefficient can be represented by using its degrees of freedoms (DOFs), i.e., we replace $\lambda^N(x)$ by

$$\boldsymbol{\lambda}^N = (\beta_0, \alpha_1, \beta_1, \ldots, \alpha_N, \beta_N)^\top \in \mathbb{R}^{2N+1}.$$

In our finite dimensional version of $\mathbb{X}$ this yields $\mathbb{X} = \mathbb{R}^M$ with $M = 2N + 1$.

We seek an algorithm that finds the statistical properties of the components of $\boldsymbol{\lambda}^N$. In other words, providing a prior distribution for $\boldsymbol{\lambda}^N$ (for a fixed N) and a finite number of observations on the pressure drop, the MCMC algorithm aims to find a posterior probability distribution for the components of $\boldsymbol{\lambda}^N$. The method requires to solve the forward problem (1.1)–(1.2) frequently.

In order to discretize the forward problem, we first partition the domain Ω into N_h elements of non-overlapping intervals $I_j := (x_{j-1/2}, x_{j+1/2}]$, where $x_{j-1/2} < x_{j+1/2}$ for $j = 1, \ldots, N_h$. Denoting the semi-discrete approximation of $p(x,t)$ and $q(x,t)$ at time t by

$$\boldsymbol{p}(t) := \big(p_1(t), p_2(t), \ldots, p_{N_h}(t)\big), \quad \boldsymbol{q}(t) := \big(q_1(t), q_2(t), \ldots, q_{N_h}(t)\big),$$

respectively, where

$$p_j(t) \approx \frac{1}{h_j} \int_{x_{j-1/2}}^{x_{j+1/2}} p(x,t)\,\mathrm{d}x, \quad q_j(t) \approx \frac{1}{h_j} \int_{x_{j-1/2}}^{x_{j+1/2}} q(x,t)\,\mathrm{d}x \quad \forall j = 1, \ldots, N_h,$$

with $h_j = x_{j+1/2} - x_{j-1/2}$, we utilize the following semi-discrete scheme for the forward

problem (1.1):

(4.3)
$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}p_j(t) - \frac{1}{2\Delta t}\Big(p_{j+1}(t) - 2p_j(t) + p_{j-1}(t)\Big) + \frac{1}{c^2\, h_j}\Big(q_{j+1}(t) - q_{j-1}(t)\Big) \\
&\qquad\qquad = \frac{1}{2c^2}\Big(\lambda_{j-1}a(p_{j-1}(t), q_{j-1}(t)) - \lambda_{j-1}a(p_{j+1}(t), q_{j+1}(t))\Big), \\
&\frac{\mathrm{d}}{\mathrm{d}t}q_j(t) - \frac{1}{2\Delta t}\Big(q_{j+1}(t) - 2q_j(t) + q_{j-1}(t)\Big) + \frac{c^2}{h_j}\Big(p_{j+1}(t) - p_{j-1}(t)\Big) \\
&\qquad\qquad = \frac{1}{2}\Big(\lambda_{j-1}a(p_{j-1}(t), q_{j-1}(t)) + \lambda_{j-1}a(p_{j+1}(t), q_{j+1}(t))\Big),
\end{aligned}$$

for all $j = 1, \ldots, N_h$. In order to discretize in time and obtain the full discrete scheme, we partition the time direction into time slabs t_n for $n \in \mathbb{Z}_+$, where $t_n < t_{n+1}$. For simplicity we assume a uniform time-step, i.e., $t_{n+1} - t_n = \Delta t$ for all $n \in \mathbb{Z}_+$ and similarly a uniform mesh-size, i.e., $h_j = h$ for all $j \in \mathbb{Z}$. Then the time derivatives are approximated by

$$\frac{\mathrm{d}}{\mathrm{d}t}p_j(t_n) \approx \frac{1}{\Delta t}\Big[p_j^{n+1} - p_j^n\Big], \tag{4.4}$$

and analogously for q. In fact, substituting (4.4) into (4.3) provides the well known Lax-Friedrichs scheme for nonlinear conservation laws (see, e.g., [17, 19]) combined with a special handling of the source term that in particular preserves steady states of (1.1) better than the usual techniques like splitting (see [33]) or direct incorporation (see [34]). The exact derivation is based on the nature of broad solutions to semilinear systems of balance laws and can be found in [20] where also a comparison of different strategies to incorporate source terms can be found. The initial conditions are imposed weakly through

$$p_j^0 = \frac{1}{h_j}\int_{x_{j-1/2}}^{x_{j+1/2}} p_0(x)\,\mathrm{d}x, \quad q_j^0 = \frac{1}{h_j}\int_{x_{j-1/2}}^{x_{j+1/2}} q_0(x)\,\mathrm{d}x \quad \forall j = 1, \ldots, N_h,$$

and the boundary conditions are realized through ghost cells yielding

$$q_0^n = g_L(t_n), \quad q_{N_h+1}^n = g_R(t_n) \quad \text{for all } t_n.$$

Since the boundary conditions merely prescribe q, the boundary values of p have to be computed. Here the flow into the ghost cells based on characteristic lines is utilized again, providing

$$p_{0/N_h+1}^{n+1} = p_{1/N_h}^n \mp \tfrac{1}{c}q_{1/N_h}^n \pm \tfrac{1}{c}q_{0/N_h+1}^{n+1} \mp \tfrac{\Delta t}{c}\lambda_{1/N_h}a(q_{1/N_h}^n, p_{1/N_h}^n).$$

We again refer to [20] for details. The discrete pressure drop at time t_n is defined by

$$\delta p_h^n := \Big|\|p_h^n\|_{\mathrm{L}^2(0,\epsilon)} - \|p_h^n\|_{\mathrm{L}^2(1-\epsilon,1)}\Big|,$$

where $p_h^n \in \mathrm{L}^2(\Omega)$ and $q_h^n \in \mathrm{L}^2(\Omega)$ are piecewise approximate solutions of the form

$$p_h^n(x) = \sum_{j=1}^{N_h} p_j^n\, \mathbf{1}_{I_j}(x), \quad q_h^n(x) = \sum_{j=1}^{N_h} q_j^n\, \mathbf{1}_{I_j}(x) \quad \forall x \in \Omega.$$

Here $\mathbf{1}_{I_j}(x)$ is the characteristic function of the interval $I_j = (x_{j-1/2}, x_{j+1/2}]$. When discretizing the infinite-dimensional setting, the discrete uncertainty-to-observation operator $\mathcal{G}_h : \mathbb{R}^{2N+1} \to \mathbb{R}^K$ is defined by

$$\mathcal{G}_h(\boldsymbol{\lambda}^N) := \big(\delta p_h^{n_1}, \delta p_h^{n_2}, \dots, \delta p_h^{n_K}\big).$$

4.1. Markov Chain Monte Carlo. Following the discretization of the forward problem and representing the friction function by a truncated series we can use the Bayesian framework (1.3) and obtain the posterior density function by

$$\pi(\lambda^N|y) = \frac{1}{Z}\exp\Big(-\frac{1}{2}|y - \mathcal{G}_h(\lambda^N)|_\Gamma^2\Big)\pi_0(\lambda^N), \tag{4.5}$$

where $Z > 0$ is the normalization constant and $\pi_0(\cdot)$ is the prior density function. MCMC is a method to sample from a given distribution whose normalization constant is unknown or intractable to compute. In our context, we use the MCMC algorithm as stated in Algorithm 1. In its formulation and throughout this paper we use $\mathsf{q}(\lambda|\lambda') = \exp(-|\lambda - \lambda'|^2/2\sigma^2)/\sqrt{2\pi\sigma^2}$.

Algorithm 1 (MCMC).

1: *Given* $\lambda_{(0)} \in \mathbb{R}^{2N+1}$, *a proposal distribution* $\mathsf{q}(\lambda'|\lambda)$, $y \in \mathbb{R}^K$.
2: *Define* $\pi_y(\lambda) := \exp\Big(-\frac{1}{2}|y - \mathcal{G}_h(\lambda)|_\Gamma^2\Big)\pi_0(\lambda)$.
3: ***for*** $i \geq 0$ ***do***
4: *Draw a proposal* λ' *from the proposal distribution function, i.e.,* $\mathsf{q}(\lambda'|\lambda_{(i)})$.
5: *Compute*

$$\alpha = \min\Big\{1, \frac{\pi_y(\lambda')\,\mathsf{q}(\lambda_{(i)}|\lambda')}{\pi_y(\lambda_{(i)})\,\mathsf{q}(\lambda'|\lambda_{(i)})}\Big\}.$$

6: *Set* $\lambda_{(i+1)} := \lambda'$ *with probability* α *and* $\lambda_{(i+1)} := \lambda_{(i)}$ *with probability* $1 - \alpha$.
7: ***end for***

Algorithm 1 generates a Markov chain with stationary distribution $\pi(\lambda|y)$ and $\{\lambda_{(i)}\}_{i\in\mathbb{N}}$ with states $\lambda_{(i)}$ is used for inference. However, the first few samples are usually discarded for inference and are called *burn-in* samples.

5. Numerical experiments. In this section we perform numerical experiments using the forward scheme and the MCMC algorithm as described in Section 4. The forward solver is implemented in C++ in order to perform fast numerical computation while the MCMC algorithm is implemented in Python scripting language with calls to the C++ code. The software package can be downloaded from `http://github.com/fg8/UQ` for academic purposes.

In the first numerical experiment we choose the truncation parameter in the friction coefficient to be $N = 0$ and choose $m_0 = 0$. This yields a scalar friction coefficient, i.e., $\lambda^0(x) = \beta_0$. As the prior we choose a uniform one, i.e.,

$$\pi_0(\lambda) := \frac{1}{Z}\mathbf{1}_{[\underline{\lambda},\overline{\lambda}]}(\lambda),$$

where $\underline{\lambda} < \overline{\lambda}$ are two fixed constants and $Z = \overline{\lambda} - \underline{\lambda}$ is the normalization constant. We choose the boundary conditions for $q(x,t)$ on $\partial\Omega$ as

$$q(0,t) = q(1,t) = 10 - \sin(2\pi t),$$

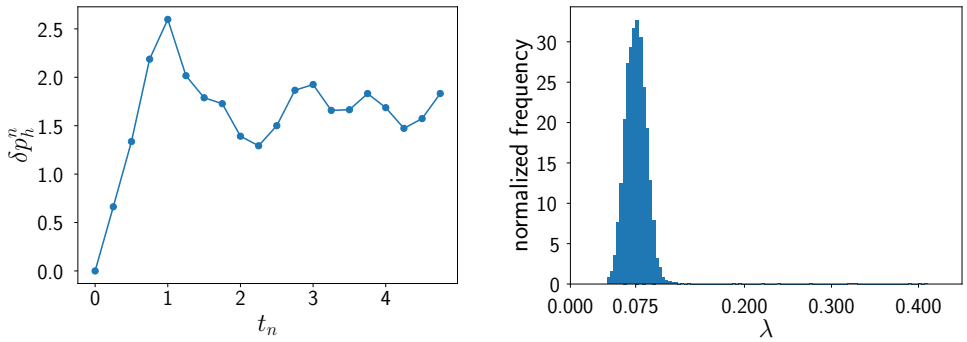

FIG. 5.1. *Pressure drop for a scalar friction coefficient with* $\lambda_{true} = 0.075$ *(left) and its corresponding histogram of the samples obtained from MCMC algorithm (right).*

and the initial conditions as

$$p(x,0) = q(x,0) = 10 + \sin(2\pi x).$$

The model problem is completed by choosing the friction function $a(\cdot,\cdot)$ as

$$a(p,q) = -q|q|.$$

In order to validate the forward and the MCMC solvers, we set $\Gamma = I$ and $y_{\text{true}} = \mathcal{G}(\lambda_{\text{true}})$, with $\lambda_{\text{true}} = 0.075$. Then we choose $\underline{\lambda} = 0$ and $\overline{\lambda} = 0.5$ to fix the prior distribution. We fix the number of cells for the forward solver to be $N_h = 200$ and measure the pressure drop at t_{n_ℓ} for $\ell = 1, ..., 20$, which are distributed uniformly in the interval $t \in [0, 5]$ (see Figure 5.1). We then run the MCMC of Algorithm 1 with 10,000 iterations to generate the Markov chain and we let 1,000 burn-in iterations. The histogram of the Markov Chain is plotted in Figure 5.1 (right). Observe that the samples cluster around the true value, i.e., $\lambda_{\text{true}} = 0.075$ while the starting point of the Markov Chain was set to be $\lambda_{(0)} = 0.45$, i.e., far from the true value.

We then add a mean zero Gaussian noise $\eta \in \mathbb{R}^K$ to y with covariance $\Gamma = 0.25 \times I_K$, where I_K is the identity matrix of size $K = 20$; see Figure 5.2 (left) for the comparison between true and noisy pressure drop. More precisely we set $y = \mathcal{G}_h(\lambda_{\text{true}}) + \eta$, and use MCMC to sample from the posterior distribution. The settings are the same as in the previous example. In Figure 5.2 (right) we see that the posterior distribution clusters around the true friction constant, i.e., $\lambda_{\text{true}} = 0.075$, despite the initial sample of MCMC to be $\lambda = 0.45$.

In the next numerical example, we consider $\lambda(x)$ to be a function of the type (4.2) with $N = 3$. We set the true friction coefficient to be

$$\lambda_{\text{true}}(x) = 0.075 + \sum_{k=1}^{N} a_k \Big(\alpha_k \sin(2\pi k x) + \beta_k \cos(2\pi k x) \Big),$$

where

$$\alpha_1 = \frac{15}{1000}, \quad \beta_1 = \frac{35}{1000}, \quad \alpha_2 = \frac{25}{1000}, \quad \beta_2 = \frac{35}{1000}, \quad \alpha_3 = \frac{1}{1000}, \quad \beta_3 = \frac{30}{1000}.$$

The prior distribution function is then defined as

$$\pi_0(\boldsymbol{\lambda}) := \mathbf{1}_{[0.0,0.5]\times[0.0,0.05]^4}(\boldsymbol{\lambda}),$$

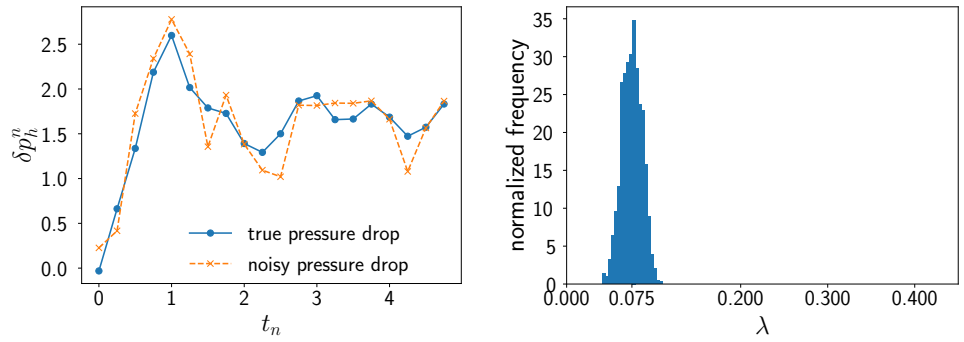

FIG. 5.2. *True and noisy pressure drop for a scalar friction coefficient with* $\lambda_{true} = 0.075$ *(left) and its corresponding histogram of the samples obtained from MCMC algorithm (right).*

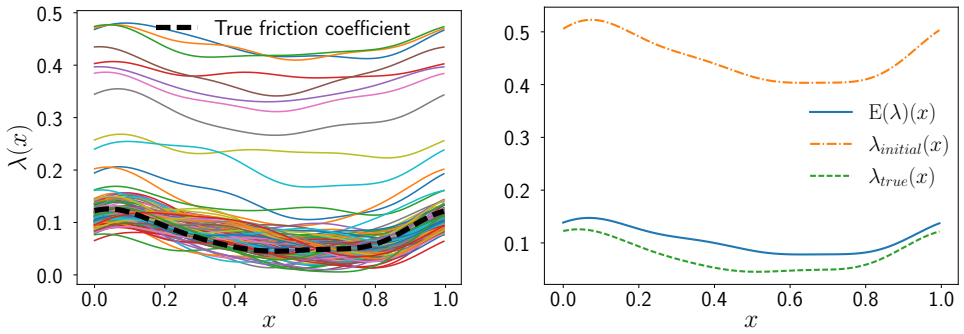

FIG. 5.3. *True and MCMC samples of the friction coefficient (left) and the pointwise mean friction function (right).*

and the starting point of MCMC $\boldsymbol{\lambda}_{(0)} = (0.45, 0.04, 0.04, 0.04, 0.04, 0.04, 0.04)^\top$. In Figure 5.3 (left) we observe the samples drawn from the MCMC. Note that the initial starting point of the MCMC is far from the true friction coefficient and as the Markov Chain grows we converge to the true friction coefficient. In Figure 5.3 (right) we observe the pointwise mean friction coefficient obtained from the samples.

5.1. A double bump friction function. In this section we perform experiments with a true friction coefficient that can be approximated only in the limit by the truncated prior model (4.2). Throughout this section we use the following double bump friction function:

$$\lambda_{\text{true}}(x) = \frac{1}{10} + \mathbf{1}_{[\frac{1}{8},\frac{3}{8}]}(x)\Big(\frac{1}{8} - \Big|x - \frac{1}{4}\Big|\Big) + 4 \times \mathbf{1}_{[\frac{5}{8},\frac{7}{8}]}(x)\Big(\frac{1}{8} - \Big|x - \frac{3}{4}\Big|\Big). \tag{5.1}$$

In Figure 5.4 (left) we have plotted its approximation by the prior model (4.2) for $N = 0, \ldots, 20$.

The discretization settings are defined as in the previous section. We choose at first $N = 4$ and set the prior distribution function to

$$\pi_0(\boldsymbol{\lambda}) := \frac{1}{Z}\mathbf{1}_{[0.0,0.45]\times[-0.025,0.025]^4\times[-0.01,0.01]^4}(\boldsymbol{\lambda}),$$

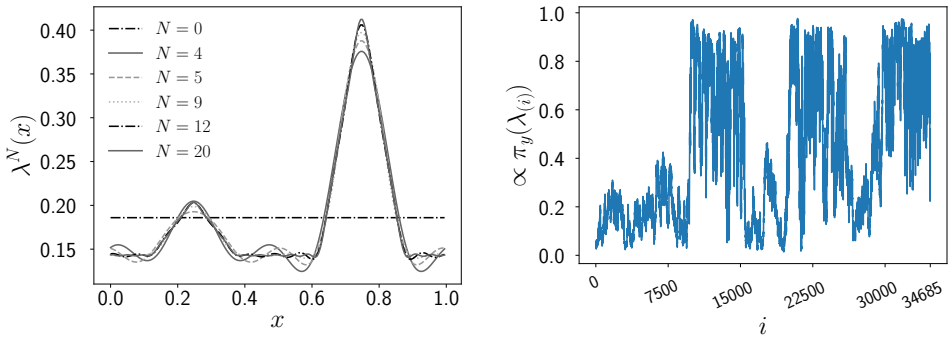

FIG. 5.4. *Fourier approximations of (5.1) (left) and posterior density function (upto a constant) for the Markov chain obtained from MCMC (right).*

where $Z > 0$ is the normalization constant and the starting point of the Markov chain is chosen as

$$\boldsymbol{\lambda}_{(0)} = (0.25, 0, 0, 0, 0, 0, 0, 0, 0)^\top.$$

We fix $\Gamma = I$, $y_{\text{true}} := \mathcal{G}(\lambda_{\text{true}})$ and run the MCMC to sample from the posterior distribution function.

In Figure 5.4, we plot the value of the posterior distribution function (upto a constant) evaluated at samples of the Markov chain obtained from Algorithm 1. Observe that there are three distinctive regions where the posterior density function is maximized, i.e., for i in $[7500, 15000]$, $[22000, 27000]$ and $[30000, 34685]$. This observation may hint that there are multiple friction functions that maximize the posterior distribution. In other words, for a given pressure drop, there might exists multiple friction functions, i.e., the uncertainty-to-observation operator is not *surjective*.

In order to identify these friction functions we use a clustering algorithm to categorize friction functions that maximize the posterior distribution. Recall that given the truncation constant N we have for each sample in the Markov chain that $\boldsymbol{\lambda}_{(i)} \in \mathbb{R}^{2N+1}$. We first select the samples in the Markov chain that maximize the posterior distribution upto a threshold. Let us denote the number of such samples by $N_{\max}$. Then, we concatenate the maximizing samples to form a feature matrix $F \in \mathbb{R}^{N_{\max} \times (2N+1)}$, i.e.,

$$F := (\boldsymbol{\lambda}_{(i_0)}, \boldsymbol{\lambda}_{(i_1)}, \dots, \boldsymbol{\lambda}_{(i_{N_{\max}})})^\top. \tag{5.2}$$

Using the feature matrix F and a clustering algorithm, e.g., k-means, we can classify samples that maximize the posterior distribution. We summarize the above post-processing step in the following algorithm:

ALGORITHM 2 (Post-processing).

1: *Given the Markov chain* $\{\boldsymbol{\lambda}_{(i)}\}_{i=0}^{N_{\text{samples}}}$*, number of clusters, i.e.,* n_{cluster}*, and a threshold* `threshold` < 1.

2: *Compute maximizing samples:* $\{\boldsymbol{\lambda}_{(i_j)}\}_{j=0}^{N_{\max}}$ *such that*

$$\boldsymbol{\lambda}_{(i_j)} > \texttt{threshold} \times \Big(\max_{i=1,\dots,N_{\text{samples}}} \pi_y(\boldsymbol{\lambda}_{(i)}) \Big).$$

3: *Set the feature matrix* F *using (5.2).*

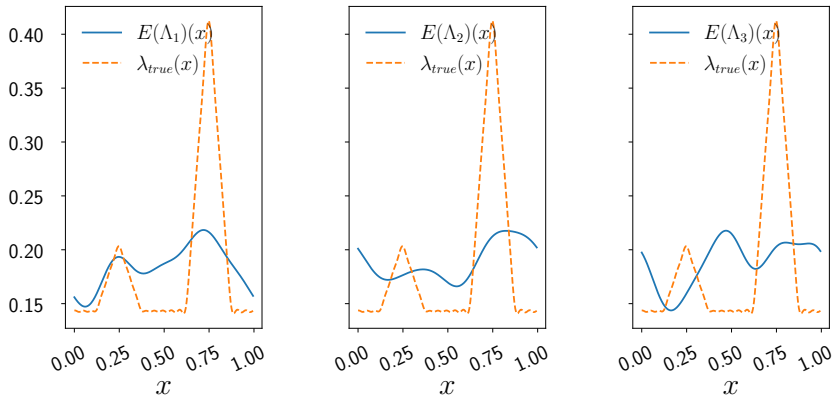

FIG. 5.5. *Mean functions of the three clusters obtained from Algorithm 2 ($N = 4$).*

4: Compute the clusters using k-means:

$$\{\Lambda_1, \Lambda_2, \cdots, \Lambda_{n_{\text{cluster}}}\} = \texttt{KMeans}(F, n_{\text{cluster}}),$$

where Λ_ℓ contains samples that form the ℓth cluster.

5: ***return*** $\{\Lambda_\ell\}_{\ell=1}^{n_{\text{cluster}}}$

We apply Algorithm 2 to our Markov chain with `threshold` $= 0.95$ and $n_{\text{cluster}} = 3$. In Figure 5.5, we plot the mean function obtained from samples of each of the three clusters. Note that all three mean functions maximize the posterior density function. However the first mean function, i.e., $E(\Lambda_1)(x)$, captures the true friction function in the sense that it captures the right position of the bumps. In Figure 5.6, the pressure drops from the above mean functions are plotted against the true pressure drop. Note that all pressure drops are a good approximation of the true one which shows that the uncertainty-to-observation is not surjective.

Adding noise. In this section we add noise into the observation, i.e., we consider $y = \mathcal{G}(\lambda) + \eta$ where η is a Gaussian noise centered at zero with standard deviation 10^{-4}. Therefore $\Gamma = 10^{-4} \times I$. The setting is same as in the noise-free experiment and Algorithm 1 is used for sampling along with the post-processing step of Algorithm 2 with $n_{\text{cluster}} = 3$. The result is depicted in Figure 5.7.

Increasing the truncation parameter. We now increase the truncation parameter N in the prior model (4.2). Observe that in Figure 5.8, the approximation has improved in the sense that we can capture the second bump located at $x = 0.75$ more accurately, and the gap between the first and second gap in the interval $[0.25, 0.75]$ is better approximated compare to Figure 5.5 (where $N = 4$). Similar results can be obtained for $N = 6$, which we plot in Figure 5.9. For the last experiment, we have chosen the number of clusters to be $n_{\text{cluster}} = 6$ and plotted the first four clusters.

6. Conclusion. In this paper we considered a semi-linear isothermal Euler equation for the modelling of the gas flow in pipes and showed well-posedness of the underlying PDE as well as continuous dependence of the unknowns with respect to the

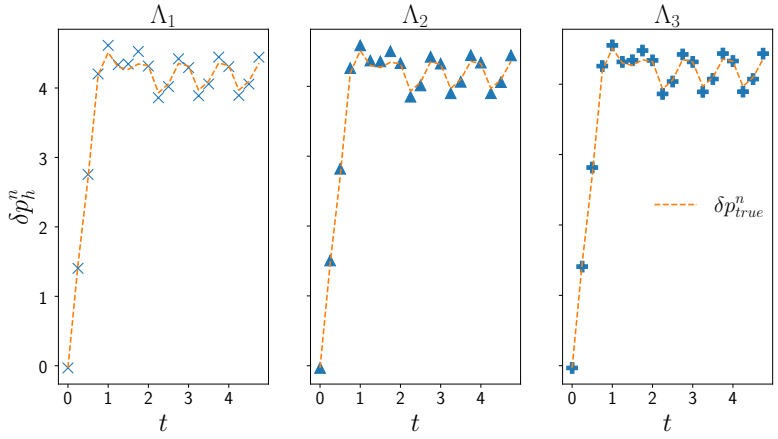

FIG. 5.6. *True pressure drop and the pressure drop associated to mean functions of the three clusters obtained from Algorithm 2 ($N = 4$).*

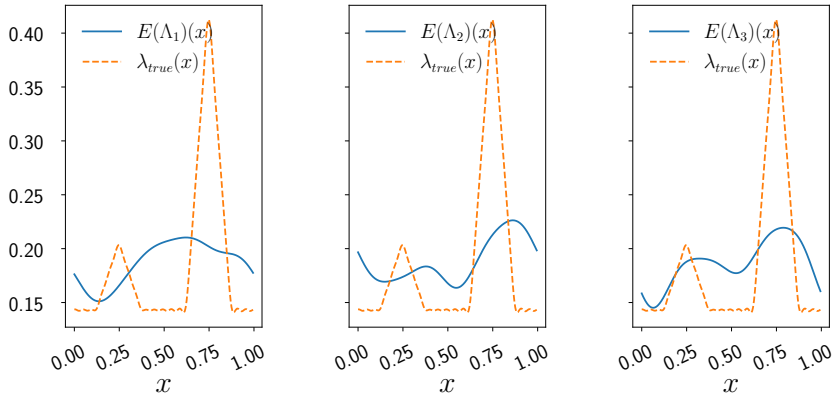

FIG. 5.7. *Mean functions of the three clusters obtained from Algorithm 2 when noise is present ($N = 4$).*

friction function. More precisely, provided that the friction function is of Lipschitz class, the underlying PDE has a strong solution (Definition 2.2) and is Lipschitz continuous with respect to the friction function. We then focused on the Bayesian inverse problem of the identification of the friction coefficient using finite (noisy) observations of the pressure drop along the pipe. We showed that the underlying Bayesian inverse problem is well-posed. This enabled us to have a well-posed formulation of the MCMC algorithm both at the continuous and the discrete level. We then discretized the underlying friction coefficient and the PDE in order to identify numerically the friction function using MCMC. The numerical results revealed that the uncertainty-to-observation operator lacks surjectivity which leads to multiple friction functions that maximize the likelihood function. Finally, in order to characterize these friction

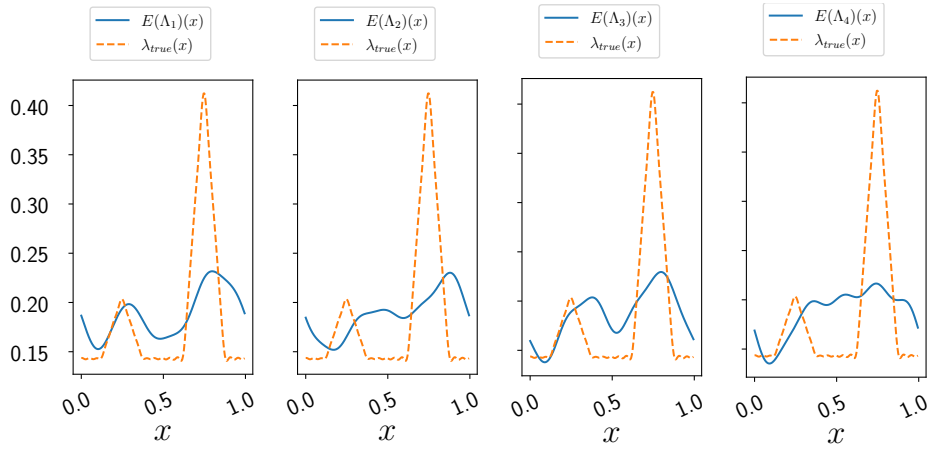

FIG. 5.8. *Mean functions of the four clusters obtained from Algorithm 2 ($N = 5$).*

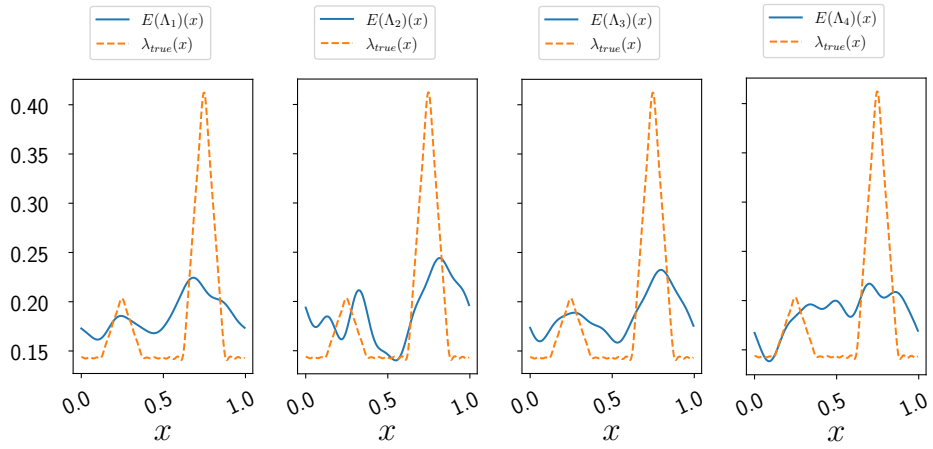

FIG. 5.9. *Mean functions of the four clusters obtained from Algorithm 2 ($N = 6$).*

coefficients, we have used a clustering algorithm.

We would like to finish this conclusion by mentioning that the convergence of the overall numerical method, i.e., discretization of the forward problem, friction coefficient, MCMC and the post-processing step, is an interesting question on its own.

Acknowledgment. This research was supported by the German Research Foundation DFG through the SFB-TRR 154 and by the Research Center MATHEON through project OT6 funded by the Einstein Center for Mathematics Berlin.

REFERENCES

[1] R. ABGRALL AND S. MISHRA, *Uncertainty quantification for hyperbolic systems of conservation laws*, in Handbook of numerical methods for hyperbolic problems, Rmi Abgrall and Chi-

Wang Shu, eds., vol. 18 of Handb. Numer. Anal., Elsevier/North-Holland, Amsterdam, 2017, pp. 507–544.
[2] A. Apte, M. Hairer, A. M. Stuart, and J. Voss, *Sampling the posterior: an approach to non-Gaussian data assimilation*, Phys. D, 230 (2007), pp. 50–64.
[3] D. Assmann, F. Liers, and M. Stingl, *Decomposable robust two-stage optimization: An application to gas network operations under uncertainty*, preprint, TRR 154, 2017.
[4] J.-M. Bernardo and A. F. M. Smith, *Bayesian theory*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Ltd., Chichester, 1994.
[5] H. Bijl, D. Lucor, S. Mishra, and C. Schwab, *Uncertainty Quantification in Computational Fluid Dynamics*, Lecture Notes in Computational Science and Engineering, Springer International Publishing, 2013.
[6] A. Birolleau, G. Poëtte, and D. Lucor, *Adaptive Bayesian inference for discontinuous inverse problems, application to hyperbolic conservation laws*, Commun. Comput. Phys., 16 (2014), pp. 1–34.
[7] S. L. Cotter, M. Dashti, J. C. Robinson, and A. M. Stuart, *Bayesian inverse problems for functions and applications to fluid mechanics*, Inverse Problems, 25 (2009), pp. 115008, 43.
[8] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White, *MCMC methods for functions: Modifying old algorithms to make them faster*, Statistical Science, 28 (2013), pp. 283 – 464.
[9] M. Dashti and A. M. Stuart, *Uncertainty quantification and weak approximation of an elliptic inverse problem*, SIAM J. Numer. Anal., 49 (2011), pp. 2524–2542.
[10] ———, *The Bayesian Approach to Inverse Problems*, Springer International Publishing, Cham, 2016, pp. 1–118.
[11] J. Dick, R. N. Gantner, Q. T. Le Gia, and C. Schwab, *Higher order quasi-Monte Carlo integration for Bayesian estimation*, Tech. Report 2016-13, Seminar for Applied Mathematics, ETH Zürich, 2016.
[12] P. Domschke, B. Hiller, J. Lang, and C. Tischendorf, *Modellierung von Gasnetzwerken: Eine Übersicht*, preprint, TRR 154, 2017.
[13] H. Egger, T. Kugler, and N. Strogies, *Parameter identification in a semilinear hyperbolic system*, Inverse Problems, 33 (2017), pp. 055022, 25.
[14] T. A. El Moselhy and Y. M. Marzouk, *Bayesian inference with optimal maps*, Journal of Computational Physics, 231 (2012), pp. 7815–7850.
[15] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.
[16] T. Gonzalez Grandon, H. Heitsch, and R. Henrion, *A joint model of probabilistic/robust constraints for gas transport management in stationary networks*, Computational Management Science, 14 (2017), pp. 443–460.
[17] S. Hajian, M. Hintermüller, and S. Ulbrich, *Total variation diminishing schemes in optimal control of scalar conservation laws*, IMA Journal of Numerical Analysis, (2017).
[18] K. Hayden, E. Olson, and E. S. Titi, *Discrete data assimilation in the Lorenz and 2D Navier–Stokes equations*, Physica D: Nonlinear Phenomena, 240 (2011), pp. 1416 – 1425.
[19] M. Hintermüller and N. Strogies, *On the consistency of Runge–Kutta methods up to order three applied to the optimal control of scalar conservation laws*, preprint, TRR 154, 2017. Accepted in Proceedings to NAOIV-2017.
[20] ———, *On the identification of the friction coefficient in a semilinear system for gas transport through a network*, preprint, TRR 154, 2017.
[21] V. H. Hoang, C. Schwab, and A. M. Stuart, *Complexity analysis of accelerated MCMC methods for Bayesian inversion*, Inverse Problems, 29 (2013), pp. 085010, 37.
[22] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*, vol. 160 of Applied Mathematical Sciences, Springer-Verlag, New York, 2005.
[23] R. J. LeVeque, *Finite volume methods for hyperbolic problems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2002.
[24] A. J. Majda and J. Harlim, *Filtering complex turbulent systems*, Cambridge University Press, 1 2012.
[25] H. G. Matthies, E. Zander, B. V. Rosić, and Alexander Litvinenko, *Parameter estimation via conditional expectation: a Bayesian inversion*, Advanced Modeling and Simulation in Engineering Sciences, 3 (2016), p. 24.
[26] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, vol. 44 of Applied Mathematical Sciences, Springer-Verlag, New York, 1983.
[27] R. Scheichl, A. M. Stuart, and A. L. Teckentrup, *Quasi-Monte Carlo and multilevel*

Monte Carlo methods for computing posterior expectations in elliptic inverse problems, SIAM/ASA J. Uncertain. Quantif., 5 (2017), pp. 493–518.

[28] C. Schillings and C. Schwab, *Sparse, adaptive Smolyak quadratures for Bayesian inverse problems*, Inverse Problems, 29 (2013), pp. 065011:1–28.

[29] C. Schwab and A. M. Stuart, *Sparse deterministic approximation of Bayesian inverse problems*, Inverse Problems, 28 (2012), pp. 045003, 32.

[30] A. M. Stuart, *Inverse problems: a Bayesian perspective*, in Acta Numerica 2010, vol. 19.

[31] T. J. Sullivan, *Introduction to uncertainty quantification*, vol. 63 of Texts in Applied Mathematics, Springer, Cham, 2015.

[32] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.

[33] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer-Verlag, Berlin, third ed., 2009. A Practical Introduction.

[34] S. Ulbrich, *Optimal Control of Nonlinear Hyperbolic Conservation Laws with Source Terms*, habilitation, TUM, 2001.