Control and Cybernetics

vol. **48** (2019) No. 2

A Bayesian approach to parameter identification in gas networks*

by

Soheil Hajian[1], Michael Hintermüller[1,2], Claudia Schillings[3] and Nikolai Strogies[1]

[1]Institut für Mathematik, Humboldt–Universität zu Berlin,
Unter den Linden 6, 10099 Berlin, Germany
soheil.hajian@hu-berlin.de, strogies@mathematik.hu-berlin.de
[2]Weierstrass Institute, Mohrenstrasse 39,
10117 Berlin, Germany
michael.hintermueller@wias-berlin.de
[3]Institute for Mathematics, University of Mannheim,
B6 28-29, C 306; 68131 Mannheim, Germany
c.schillings@uni-mannheim.de

Dedicated to Günter Leugering on the occasion of His 65th birthday

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.

Keywords: Bayesian inversion, distributed friction coefficient, gas network/pipeline, hyperbolic PDE system

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. It is widely believed that during the transition from the current energy portfolio to one which has an emphasis on renewable energies

*Submitted: March 2019; Accepted: December 2019

and other energy carriers such as hydrogen, 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, originating 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 in a single pipe:

$$\begin{aligned} \partial_t p(x,t) + c^2 \partial_x q(x,t) &= 0 && \text{in } \Omega \times (0,T), \\ \partial_t q(x,t) + \partial_x p(x,t) &= \lambda(x) a(p(x,t), q(x,t)) && \text{in } \Omega \times (0,T), \end{aligned} \tag{1}$$

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 location $x \in \Omega$ and 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), $a(\cdot,\cdot) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the friction function and $\lambda \in \mathrm{L}^\infty(\Omega)$ denotes the *friction coefficient*. Loosely speaking, λ 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 Domschke et al., 2007), however, approximation formulas for the 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) is completed by the following initial and boundary conditions:

$$p(x,0) = p_0(x), q(x,0) = q_0(x) \quad \forall x \in \Omega, \tag{2a}$$

$$q(0,t) = g_L(t), q(1,t) = g_R(t) \quad \forall t > 0, \tag{2b}$$

where p_0, q_0, g_L, and g_R are given quantities. Combining (1) and (2) provides our overall model of gas flow. We refer to Domschke et al. (2007) and LeVeque (2002) for more on this and further details on the model reduction process.

Motivated by the 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 with the latter preventing the evaluation of the pressure precisely at the endpoints of the pipe. For a classical inverse problem relative to the conditional identification of a friction law for the isothermal Euler equations we refer to the work of Egger, Kugler and Strogies (2017), 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 (Assmann, Liers and Stingl, 2017) or the study of probabilistic constraints (Gonzalez Grandon, Heitsch and Henrion, 2017) both involving the friction coefficient which, as highlighted above, is uncertain during operation. The pertinent 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 (Kaipio and Somersalo, 2005; Stuart, 2010; Tarantola, 2005). 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, we introduce next the *parameter-to-observation map* $\mathcal{G}$, 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) applied to the friction function λ and a data formation operator. By measuring δp at finitely many time instances $t_j \in [0,T]$, $j = 1,\dots,K$, we obtain here

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

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 $\rho_\eta(\cdot)$.

Then, the conditional probability density of obtaining y for a given friction coefficient λ is $\rho_\eta(y - \mathcal{G}(\lambda))$, which is the *likelihood* of the data. Moreover, using Bayes' theorem, we can incorporate our "prior" knowledge on λ and provide its conditional probability density given the data y by

$$\rho_\lambda(\lambda|y) \propto \rho_\eta(y - \mathcal{G}(\lambda))\rho_\lambda(\lambda). \tag{3}$$

Our knowledge on λ, i.e., $\rho_\lambda(\lambda)$, is called *prior* (density) and the conditional probability density of λ given the data y, i.e., $\rho_\lambda(\lambda|y)$, is called *posterior* (density). 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 and 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.,

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(\lambda) \propto \rho_\eta(y - \mathcal{G}(\lambda)). \tag{4}$$

For an overview on the Bayesian approach to statistics in finite dimensions we refer the reader to Bernardo and Smith (1994). Bayesian inverse problems with an emphasis on modelling and computation are addressed in Kaipio and Somersalo (2005). We would also like to 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 Dashti and Stuart (2011) and viscous incompressible flow on a two-dimensional torus in Cotter et al. (2009).

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 characterization of the uncertainty via the posterior distribution; see, e.g., Dashti and Stuart (2016), Kaipio and Somersalo (2005), or Stuart (2010). 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, see Stuart (2010), 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 Dashti and Stuart (2016). However, computational challenges arise due to the complex structure of probability measures in high or infinite dimensional settings. 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., Cotter et al. (2013), Dick et al. (2016), El Moselhy and Marzouk (2012), Matthies et al. (2016), Scheichl, Stuart and Teckentrup (2017), or Schillings and Schwab (2013).

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 that the uncertainty quantification for hyperbolic differential equations with random data has been a very active field of research over the recent past; see Abgrall and Mishra (2017), Bijl et al. (2013) 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 (see Apte et al, 2007; Hayden, Olson and Titi, 2011; Majda and Harlim, 2012). However, the quantification of uncertainties in the inverse setting has been the subject of only a very small number of publications; compare Birolleau, Poëtte and Lucor (2014), Cotter et al. (2009). 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 Bayesian 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) in such a way 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) in the following abstract form of a dynamical system:

$$\begin{aligned} &\boldsymbol{u}'(t) + A\boldsymbol{u}(t) = \boldsymbol{f}(\lambda, \boldsymbol{u}(t), t) \quad (t > 0), \\ &\boldsymbol{u}(0) = \boldsymbol{u}_0, \end{aligned} \tag{5}$$

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

$$A := \begin{bmatrix} 0 & c^2\partial_x \\ \partial_x & 0 \end{bmatrix}, \tag{6}$$

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

$\mathrm{L_w}$ 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

$$\boldsymbol{f}(\lambda, \boldsymbol{u}, t) := \begin{pmatrix} -c^2(g_L(t) - g_R(t)) \\ \lambda\, a(p, \hat{q} + \tilde{q}) - x g_L'(t) - (1-x) g_R'(t) \end{pmatrix}. \tag{7}$$

Note that the 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 the continuity properties of $\boldsymbol{f}$.

Lemma 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

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

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

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

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 (5), 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 Pazy (1983, Definition 4.2.1). There are, however, two other notions of solution associated with (5): *strong solutions* (Pazy, 1983, Definition 4.2.8) and *mild solutions* (Pazy, 1983, Definition 4.2.3). We will work in this paper with strong solutions of (5):

Definition 1 (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 (5) 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.,

$$\begin{aligned} &\boldsymbol{u}'(t) + A\boldsymbol{u}(t) = \boldsymbol{f}(t) \quad (t > 0), \\ &\boldsymbol{u}(0) = \boldsymbol{u}_0, \end{aligned} \tag{8}$$

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 (8) 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 Ern and Guermond, 2004, Chapter 6.3):

Definition 2 (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*

$$\|Au\|_L \geq c_1 \|u\|_V - c_2 \|u\|_L \quad \forall u \in V, \tag{9}$$

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 (9) 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 Ern and Guermond (2004), 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 (5) for $\boldsymbol{u}_0 \in L$ and $\boldsymbol{f}(\lambda, \boldsymbol{u}(t), t) \in \mathrm{L}^1((0,T);L)$ if it satisfies

$$\boldsymbol{u}(t) = \mathbb{T}(t)\boldsymbol{u}_0 + \int_0^t \mathbb{T}(t-s)\boldsymbol{f}(\lambda, \boldsymbol{u}(s), s)\mathrm{d}s \quad \forall t \in [0,T]. \tag{10}$$

The following theorem states the existence and uniqueness of a mild solution which will be used later on to establish the analogous result for strong solutions. See Pazy (1983), Theorem 1.2, Chapter 6, for a proof.

Theorem 1 *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 (10).*

We now establish uniqueness and existence of a strong solution to (5) as well as boundedness and Lipschitz continuity of the solution with respect to the friction coefficient.

Theorem 2 *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 (5) a.e. in time and

$$\|\boldsymbol{u}(t)\|_L \leq C(T,\boldsymbol{u}_0,\lambda) \quad \forall t \in [0,T]. \tag{11}$$

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

$$\|p_1(t) - p_2(t)\|_{\mathrm{L}^2(\Omega)} \leq C\|\lambda_1 - \lambda_2\|_X \quad \forall t \in [0,T]. \tag{12}$$

Proof Let $\boldsymbol{u} \in \mathrm{C}([0,T];L)$ be a mild solution of (5), $\|\mathbb{T}(t)\|_L \leq M$ and $\|\boldsymbol{f}(\lambda,\boldsymbol{u}(t),t)\|_L \leq N$ for $t \in [0,T]$. For $h \in [0,t]$ we obtain

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

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

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

where we have used the fact 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 Pazy, 1983, 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 (11)). This, combined with Lip-schitz 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, Pazy (1983, 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 (5). Then, by Pazy (1983), 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, using the definition of mild solutions, we have for their difference

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

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

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

which completes the proof. □

From the definition of $\boldsymbol{f}$ in (7), Lemma 1 and Theorem 2, we obtain the following sufficient conditions on the boundary data g_L, g_R and the friction function such that the PDE in (1) admits a unique strong solution and the Lipschitz continuity result.

Corollary 1 *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). Further, 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 the Bayesian inverse problem

3.1. Prior modelling

Next, 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. For this purpose, let $\{\phi_j\}_{j=1}^{\infty}$ denote an infinite sequence in a Banach space X with the norm $\|\cdot\|_X$ associated with 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

$$\lambda = m_0 + \sum_{j=1}^{\infty} \lambda_j \phi_j, \tag{13}$$

where $m_0 \in X$ (not necessarily normalized) and $\{\lambda_j\}_{j=1}^{\infty}$ are random 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 also define the truncated series

$$\lambda^N = m_0 + \sum_{j=1}^{N} \lambda_j \phi_j. \tag{14}$$

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 Dashti and Stuart (2016), Theorem 2.1, states that $(\mathcal{X}, \|\cdot\|_X)$ is a Banach space.

Theorem 3 *The following holds $\mathbb{P}$-almost surely: The sequence of functions $\{\lambda^N\}_{N=1}^{\infty}$ given by (14) is Cauchy in $\mathcal{X}$ and the limiting function λ given by (13) satisfies*

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

3.2. Bayesian inverse problem

According to our assumptions, we have noisy observations of the pressure drop at our disposition, 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 $j = 1, \ldots, K$. 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

$$\mathcal{G} : \mathcal{X} \to \mathcal{Y}, \quad \mathcal{G}(\lambda) := (\delta p(t_1), \ldots, \delta p(t_K))^{\top}. \tag{15}$$

Note that Corollary 1 implies that

$$|\mathcal{G}(\lambda_1) - \mathcal{G}(\lambda_2)| \leq C \|\lambda_1 - \lambda_2\|_{\mathrm{L}^{\infty}(\Omega)}. \tag{16}$$

Moreover, Theorem 2 and Theorem 3 provide a constant $C' > 0$ such that

$$|\mathcal{G}(\lambda)| \le C' K^{1/2}. \tag{17}$$

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 (3).

For infinite-dimensional $\mathbb{X}$ the relation (4) between the posterior measure and the prior measure, based on the Radon-Nikodym derivative, yields

$$\frac{d\mu^y}{d\mu_0}(\lambda) = \frac{1}{Z}\exp\big(-\Phi(\lambda, y)\big), \tag{18a}$$

$$Z(y) = \int_X \exp\big(-\Phi(\lambda, y)\big)d\mu_0(\lambda), \tag{18b}$$

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 (18). The conditions furthermore guarantee well-posedness of the posterior distribution with respect to perturbations of the data. We first state the conditions and then check whether they are satisfied by our Φ.

Assumption 1 *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

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

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

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

(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 $\max\{\|\lambda_1\|_\Lambda, \|\lambda_2\|_\Lambda, \|y\|_Y\} < r$

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

(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 $\lambda \in \Lambda$*

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

The first and second condition in Assumption 1 ensures the boundedness of the normalization constant from above and below away from zero. The Lipschitz continuity of the potential Ψ in u, i.e. the third condition, ensures the measurability of Ψ w.r.t. the prior measure. The last condition allows to show stability results w.r.t. the data. We refer the interested reader to Dashti and Stuart (2011), and Sullivan (2015) for more details.

Using Corollary 1, we show in the following proposition that

$$\Phi : \mathbb{X} \times \mathcal{Y} \to \mathbb{R}, \quad \text{with} \quad \Phi(\lambda, y) = \frac{1}{2}|\Gamma^{-1/2}(y - \mathcal{G}(\lambda))|^2 \tag{19}$$

indeed satisfies Assumption 1.

Proposition 1 *Let $\boldsymbol{u} = (p, q)$ be the strong solution of (5), 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 (15). Then, the likelihood function Φ as defined in* (19) *satisfies Assumption 1.*

Proof In view of Assumption 1 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 (17), we have

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

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). □

We remark that in the uniform setting, boundedness of the potential follows directly from the model (13) and the properties of the forward problem. The existence and well-posedness results, presented in this section, are not limited to the uniform case and can be shown for rather general prior distributions provided that the conditions in Assumption 1 are satisfied. The following theorem asserts well-posedness of the posterior measure in our setting. It rests on the results from Dashti and Stuart (2016) and Sullivan (2015).

Theorem 4 *Let $\boldsymbol{u} = (p, q)$ be the strong solution of (5), given $\boldsymbol{u}_0 \in V$ and the prior measure μ_0 according to* (13)*. The uncertainty-to-observation map $\mathcal{G}$ is as in* (15)*. Then, the posterior measure μ^y satisfying* (18) *with* (19) *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 1, the likelihood function Φ satisfies Assumption 1. 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 Dashti and Stuart (2016), Theorem 3.4, the posterior μ^y is well-defined and fulfills (18). 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 1. Here, $L^1_{\mu_0}(\mathbb{X}, \mathbb{R})$ denotes the Bochner space of all measurable functions

$$f : \mathbb{X} \to \mathbb{R} \text{ with } \int_{\mathbb{X}} |f(\lambda)| \mathrm{d}\mu_0(\lambda) < \infty.$$

By Dashti and Stuart (2016), Theorem 4.5, the Lipschitz continuity of the posterior follows. □

4. Discretization of the forward problem and sampling

4.1. Discretization

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 (14) yielding

$$\lambda^N(x) = m_0(x) + \sum_{j=1}^{N} \lambda_j \phi_j(x), \tag{20}$$

where $m_0 \in \mathrm{L}^\infty(\Omega)$, $\lambda_j = \gamma_j \xi_j$, $j = 1, \ldots, N$ and $\phi_{2j} = \sin(2\pi(j-1)x)$, $j = 1, \ldots, \lfloor N/2 \rfloor$, $\phi_{2j-1} = \cos(2\pi(j-1)x)$, $j = 1, \ldots, \lfloor N/2 \rfloor + 1$. In order to ensure that $\lambda^N(x) \geq \nu > 0$ for all $x \in \Omega$, we choose γ_k with

$$\sum_{k\in\mathbb{N}} |\gamma_k| < \infty, \qquad \sum_{k\in\mathbb{N},|k|>j} |\gamma_k| < C\, j^{-\nu}, \tag{21}$$

for some $C > 0$ and $\nu > 0$; see Hoang, Schwab and Stuart (2013) and Schwab and Stuart (2012). The mean function $m_0(x)$ is chosen such that positivity of λ^N is ensured. The random variables ξ_k are assumed to be i.i.d. with $\xi_1 \sim U([-1, 1])$.

The truncation λ^N converges towards $\lambda \in \mathbb{X}$ at the rate ν, i.e. there exists $C > 0$ such that for all N and every $\lambda \in \mathbb{X}$, there holds

$$\|\lambda - \lambda^N\|_{\mathbb{X}} \leq CN^{-\nu}\,.$$

The Lipschitz continuity of the forward operator then leads to the estimate

$$|\mathcal{G}(\lambda) - \mathcal{G}(\lambda^N)| \leq CN^{-\nu}$$

with a possibly different constant $C > 0$. Instead of a fixed N-term truncation, random truncation leads to a variable dimension formulation. We refer to Cotter et al. (2013) for more details. By choosing a sufficiently large, but finite number N of terms, the truncated friction coefficient can be represented by using its degrees of freedoms (DOFs), i.e., we replace $\lambda^N(x)$ by

$$\boldsymbol{\lambda}^N = (\xi_1, \xi_2, \ldots, \xi_N)^\top \in \mathbb{U}.$$

with $\mathbb{U} = [-1, 1]^N$.

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)–(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):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}p_j(t) - \frac{1}{2\Delta t}\Big(p_{j+1}(t) - 2p_j(t) + p_{j-1}(t)\Big) + \frac{1}{c^2\,h_j}\Big(q_{j+1}(t) - q_{j-1}(t)\Big) \\ = \frac{1}{2c^2}\Big(\lambda_{j-1}a(p_{j-1}(t), q_{j-1}(t)) - \lambda_{j-1}a(p_{j+1}(t), q_{j+1}(t))\Big), \\ \frac{\mathrm{d}}{\mathrm{d}t}q_j(t) - \frac{1}{2\Delta t}\Big(q_{j+1}(t) - 2q_j(t) + q_{j-1}(t)\Big) + \frac{c^2}{h_j}\Big(p_{j+1}(t) - p_{j-1}(t)\Big) \\ = \frac{1}{2}\Big(\lambda_{j-1}a(p_{j-1}(t), q_{j-1}(t)) + \lambda_{j-1}a(p_{j+1}(t), q_{j+1}(t))\Big), \end{aligned} \tag{22}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}p_j(t_n) \approx \frac{1}{\Delta t}\Big[p_j^{n+1} - p_j^n\Big], \tag{23}$$

and analogously for q. In fact, substituting (23) into (22) provides the well known Lax-Friedrichs scheme for nonlinear conservation laws (see, e.g., Hajian, Huntermüller and Ulbrich, 2017; Hintermüller and Strogies, 2017a) combined with a special handling of the source term that, in particular, preserves steady states of (1) better than the usual techniques like splitting (see Toro, 2009) or direct incorporation (see Ulbrich, 2001). The exact derivation is based on the nature of broad solutions to semilinear systems of balance laws and can be found in Hintermüller and Strogies (2017b) where also a comparison of different strategies to incorporate source terms is provided. 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 Hintermüller and Strogies (2017b) 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

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

4.2. 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 (3) and obtain the posterior density function by

$$\pi(\lambda^N|y) = \frac{1}{Z} \exp\Big(- \frac{1}{2} |y - \mathcal{G}_h(\lambda^N)|_\Gamma^2 \Big) \pi_0(\lambda^N), \tag{24}$$

where $|\cdot|_\Gamma$ is the Euclidean norm induced by matrix Γ^{-1}, $Z > 0$ 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.

Algorithm 1 (Metropolis-Hastings)

Given $\lambda_{(0)} \in \mathbb{R}^{2N+1}$*, a proposal distribution* $\mathsf{q}(\lambda'|\lambda)$*,* $y \in \mathbb{R}^K$ *and* $M \in \mathbb{N}$.
Define $\pi_y(\lambda) := \exp\Big(- \frac{1}{2} |y - \mathcal{G}_h(\lambda)|_\Gamma^2 \Big) \pi_0(\lambda)$.
Define an empty list `output` $= [\,]$.
for $0 \le i \le M$ ***do***
    Draw a proposal λ' *from the proposal distribution function, i.e.,* $\mathsf{q}(\lambda'|\lambda_{(i)})$.
    Compute

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

    Set $\lambda_{(i+1)} := \lambda'$ *with probability* α *and* $\lambda_{(i+1)} := \lambda_{(i)}$ *with probability* $1 - \alpha$.
    Append $\lambda_{(i+1)}$ *into* `output`.
    $i \leftarrow i + 1$.
end for
return `output`

Algorithm 1 generates a Markov chain with stationary distribution $\pi(\lambda|y)$ (under additional assumptions on the proposal kernel). The first few samples are usually discarded for inference and are called *burn-in* samples. The

Metropolis–Hastings algorithm will be used in the following with two different proposal kernels, the Gaussian random walk proposal and the preconditioned Crank–Nicolson variant. For the Gaussian random walk proposal function and throughout this paper (when referring to MCMC) we use

$$\mathsf{q}(\lambda|\lambda') = \frac{1}{\sqrt{(2\pi)^{2N+1}\det(\Sigma)}}\exp(-|\lambda-\lambda'|_\Sigma^2/2) \tag{25}$$

with given symmetric positive definite matrix $\Sigma \in \mathbb{R}^{2N+1\times 2N+1}$.

4.3. Preconditioned Crank-Nicolson

Since the Gaussian random walk is known not to be robust with respect to the dimension of the underlying friction coefficient, we use a robust variant of the Metropolis-Hastings algorithm, called *preconditioned Crank-Nicolson* (pCN). Already in its basic form pCN is suitable for centered Gaussian priors, i.e., $\pi_0 \sim \mathcal{N}(0, C)$ where C is the covariance matrix. Algorithm 2 describes this method; see Chen et al. (2018).

Algorithm 2 (pCN with Gaussian prior)

Given $\lambda_{(0)} \in \mathbb{R}^{2N+1}$, $\beta \in (0,1]$, $y \in \mathbb{R}^K$ *and* $M \in \mathbb{N}$.

Define $\Phi(\lambda; y) := \frac{1}{2}|y - \mathcal{G}_h(\lambda)|_\Gamma^2$.

Define an empty list `output` $= [\,]$.

for $0 \le i \le M$ ***do***

  Propose $\lambda' = (1-\beta^2)^{\frac{1}{2}}\lambda_{(i)} + \beta\zeta_{(i)}$ *with* $\zeta_{(i)} \sim \mathcal{N}(0, C)$.

  Compute

$$\alpha = \min\left\{1, \exp\left(\Phi(\lambda_{(i)}; y) - \Phi(\lambda'; y)\right)\right\}$$

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

  Append $\lambda_{(i+1)}$ *into* `output`.

  $i \leftarrow i+1$.

end for

return `output`

In order to use pCN in our setting, i.e., with a uniform prior instead of a Gaussian prior, we need to rewrite the random representation of $\lambda(x)$ in a way suitable for Algorithm 2. Recall that the truncated random friction function is defined in an abstract setting by

$$\lambda^N(x) = m_0 + \sum_{j=1}^{N} \lambda_j\phi_j(x),$$

where $\lambda_j = \gamma_j\xi_j$. Here, $\{\gamma_j\}_{j=1}^N$ is a deterministic sequence and $\{\xi_j\}_{j=1}^N$ is an i.i.d. random sequence where $\xi_j \sim U([-1,1])$. We now represent ξ_j by

$\xi_j = G(\zeta_j)$ with

$$G(\zeta_j) := 2F(\zeta_j) - 1, \quad \zeta_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1),$$

and $F(z)$ is the cumulative distribution function of a standard normal distribution. Therefore λ^N can be represented using random numbers with normal distribution, i.e.,

$$\lambda^N = m_0 + \sum_{j=1}^{N} \gamma_j G(\zeta_j) \phi_j.$$

This enables us to define Algorithm 2 using uniform priors which we describe in Algorithm 3.

Algorithm 3 (pCN with uniform prior)

Given $\lambda_{(0)} \in \mathbb{R}^{2N+1}$, $\beta \in (0,1]$, $y \in \mathbb{R}^K$ *and* $M \in \mathbb{N}$.
Define $\Phi(\lambda; y) := \frac{1}{2}|y - \mathcal{G}_h(\lambda)|_\Gamma^2$.
Define an empty list `output` $= [\,]$.
for $0 \le i \le M$ ***do***
  Define $\hat{\zeta}_{(i)} := (G^{-1}(\lambda_{(i),0}), \ldots, G^{-1}(\lambda_{(i),2N+1}))^\top$.
  Propose $\zeta' = (1-\beta^2)^{\frac{1}{2}} \hat{\zeta}_{(i)} + \beta \zeta_{(i)}$ *with* $\zeta_{(i)} \sim \mathcal{N}(0, I)$.
  Define $\lambda' := (G(\zeta'_0), \ldots, G(\zeta'_{2N+1}))^\top$
  Compute

$$\alpha = \min\left\{1, \exp\left(\Phi(\lambda_{(i)}; y) - \Phi(\lambda'; y)\right)\right\}$$

  Set $\lambda_{(i+1)} := \lambda'$ *with probability* α *and* $\lambda_{(i+1)} := \lambda_{(i)}$ *with probability* $1 - \alpha$.
  Append $\lambda_{(i+1)}$ *into* `output`.
  $i \leftarrow i + 1$.
end for
return `output`

5. Numerical experiments

5.1. The setting

In this section we report on 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 consider a finite dimensional setting with only one unknown parameter. As the prior we choose a uniform one, i.e.,

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

where $\underline{\lambda} = 0.01$, $\overline{\lambda} = 0.5$ are two fixed constants and $Z = \overline{\lambda} - \underline{\lambda}$ is the normalization constant. This corresponds to $\lambda^1(x) = m_0(x) + \lambda_1$ with $m_0(x) = (\overline{\lambda} - \underline{\lambda})/2 + \underline{\lambda}$, $\lambda_1 = \gamma_1 \xi_1$, $\xi_1 \sim U([-1,1])$ and $\gamma_1 = (\overline{\lambda} - \underline{\lambda})/2$. We choose the boundary conditions for $q(x,t)$ on $\partial\Omega$ as

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

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

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

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 $t \in [0,5]$. We then add a mean zero Gaussian noise $\eta \in \mathbb{R}^K$ to y with covariance $\Gamma = I_K$, where I_K is the identity matrix of size $K = 20$; see Fig. 1 (left) for the comparison between true and noisy pressure drop. More precisely, we set $y = \mathcal{G}_h(\lambda_{\text{true}}) + \eta$ with $\lambda_{\text{true}} = 0.075$, and use MCMC to sample from the posterior distribution. 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. For the proposal function $\mathsf{q}(\lambda|\lambda')$, see (25), we choose $\Sigma = 0.25$ to achieve an acceptance rate of 20 % of the MCMC algorithm.

In Fig. 1 (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 $\lambda = 0.45$.

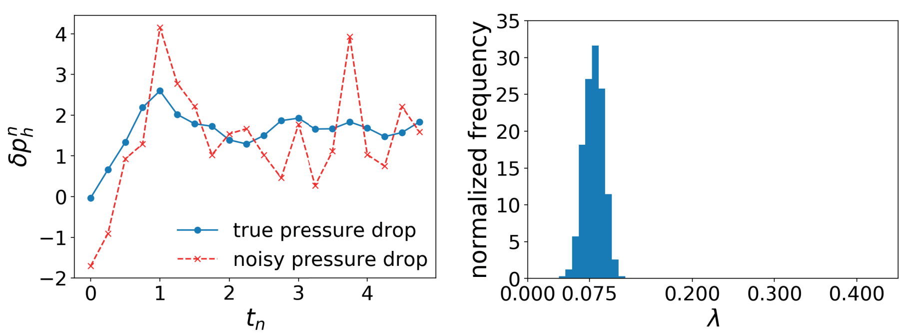

Figure 1. True and noisy pressure drop for a scalar friction coefficient with $\lambda_{\text{true}} = 0.075$ (left) and its corresponding histogram of the samples obtained from MCMC algorithm (right)

We finish this experiment by numerically checking for posterior consistency, i.e., the noise level in the observation will be reduced. Further, we increase the number of observation points to 100. In Fig. 2, the histogram of the identified friction constant is plotted for different noise levels, i.e., $\eta = \mathcal{N}(0, \gamma^2 \cdot I_K)$ for $\gamma = 0.1, \ldots, 0.5$. Observe that the posterior concentrates around the true value $\lambda_{\text{true}} = 0.075$ as the noise tends to zero.

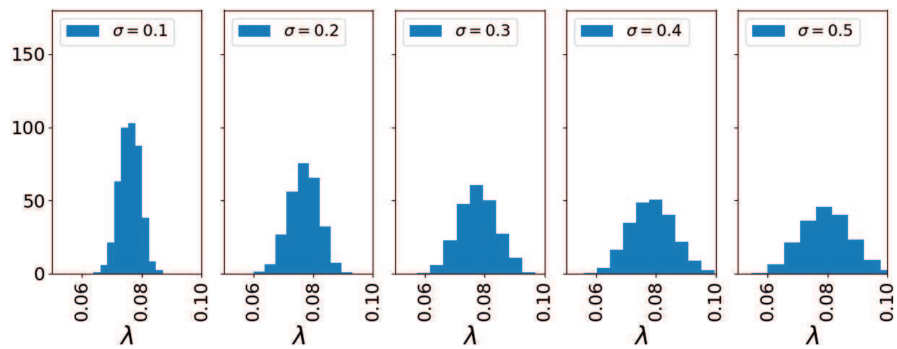

Figure 2. Effect of the noise in the observation on the posterior distribution.

In the next numerical example, we consider $\lambda(x)$ to be a function of the type (20) with $N = 5$. We set the true friction coefficient to be

$$\begin{aligned}\lambda_{\text{true}}(x) &= 2.0 + \frac{1}{4}\xi_1^{\text{true}} + \frac{1}{4}\xi_2^{\text{true}}\sin(2\pi x) + \frac{1}{9}\xi_3^{\text{true}}\cos(2\pi x) \\ &\quad + \frac{1}{16}\xi_4^{\text{true}}\sin(4\pi x) + \frac{1}{25}\xi_5^{\text{true}}\cos(4\pi x)\end{aligned}$$

where $\xi_1^{\text{true}} = \frac{1}{100}$, $\xi_2^{\text{true}} = \frac{1}{10}$, $\xi_3^{\text{true}} = \frac{2}{10}$, $\xi_4^{\text{true}} = \frac{4}{10}$, $\xi_5^{\text{true}} = \frac{1}{100}$. The prior distribution function is defined as

$$\pi_0(\boldsymbol{\lambda}) := \frac{1}{Z}\mathbf{1}_{[-1,1]^5}(\boldsymbol{\lambda}),$$

with normalization constant $Z > 0$. We infer the unknown from 100 observations with noise $\eta \sim N(0, 0.01^2 \times I_K)$. The computational setting is the same as in the one-dimensional example. In Fig. 3 (left) we observe the samples drawn from the posterior via MCMC. In Fig. 3 (right) we observe the pointwise mean friction coefficient obtained from the samples (compared to the initial one used as a starting point for the Markov chain).

5.2. 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 (20). Throughout this section we use the following double bump friction function:

$$\lambda_{\text{true}}(x) = \frac{1}{10} + \mathbf{1}_{[\frac{1}{8},\frac{3}{8}]}(x)\Big(\frac{1}{8} - \Big|x - \frac{1}{4}\Big|\Big) + 4 \times \mathbf{1}_{[\frac{5}{8},\frac{7}{8}]}(x)\Big(\frac{1}{8} - \Big|x - \frac{3}{4}\Big|\Big). \quad (26)$$

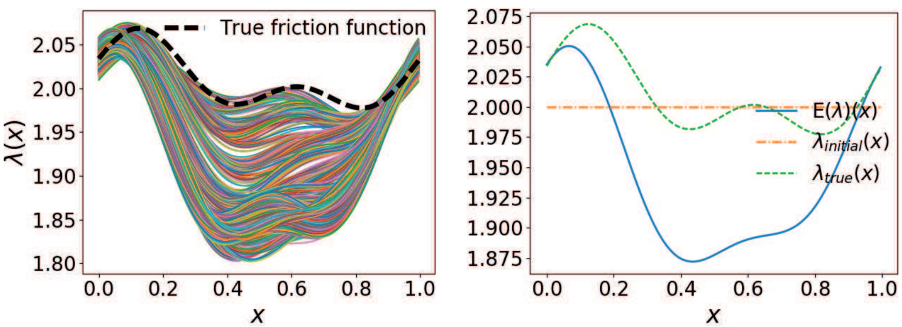

Figure 3. True and MCMC samples of the friction coefficient (left) and the pointwise mean friction function (right)

In Fig. 4 (left) we have plotted its approximation by the prior model (20) for $N = 0, \ldots, 20$.

The discretization settings are defined as in the previous section. The parametrization is assumed to be of the form

$$\lambda^N(x) = 0.24 + \sum_{j=1}^{N} \lambda_j \phi_j(x),$$

where $\lambda_j = \gamma_j \xi_j,\ j = 1, \ldots, N$ and $\phi_{2j} = \sin(2\pi(j-1)x),\ j = 1, \ldots, \lfloor N/2 \rfloor$, $\phi_{2j-1} = \cos(2\pi(j-1)x),\ j = 1, \ldots, \lfloor N/2 \rfloor + 1$. The coefficients γ_j are given as $\gamma_{1,2} = 0.06, \gamma_{3,4} = 0.001, \gamma_5 = 0.06, \gamma_6 = 0.03, \gamma_{7,8} = 0.001, \gamma_{9,10} = 0.01$ and $\gamma_j = 0.001 \cdot 1/j^2,\ j > 10$. We choose at first $N = 11$ and set the prior distribution function to

$$\pi_0(\boldsymbol{\lambda}) := \frac{1}{Z} \mathbf{1}_{[-1,1]^{11}}(\boldsymbol{\lambda}),$$

where $Z > 0$ is the normalization constant.

We consider 40 observations and add noise to the observation, i.e., we consider $y = \mathcal{G}(\lambda) + \eta$ where η is a Gaussian noise centered at zero with standard deviation 0.1. Therefore, $\Gamma = 0.1^2 \times I$. Algorithm 1 is used for sampling from the posterior distribution with 50,000 iterations. The mean is compared to the MAP, i.e. the minimizer of the negative log posterior, and to the underlying true parameter. The result is depicted in Fig. 5.

Both estimates lead to a satisfactory approximation of the observed data. Note that we considered in this example a specific prior distribution tailored to the problem. The prior distribution plays a crucial role in the accuracy of the estimates and the definition of the prior is in general a nontrivial task for real-world applications. We mention the hierarchical approach, which allows to learn additional parameters in the prior distribution from data and has the potential to significantly improve the accuracy of the estimates in practice.

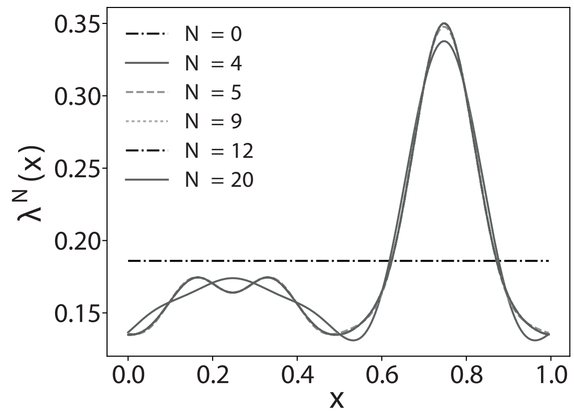

Figure 4. Best prior approximations of (26)

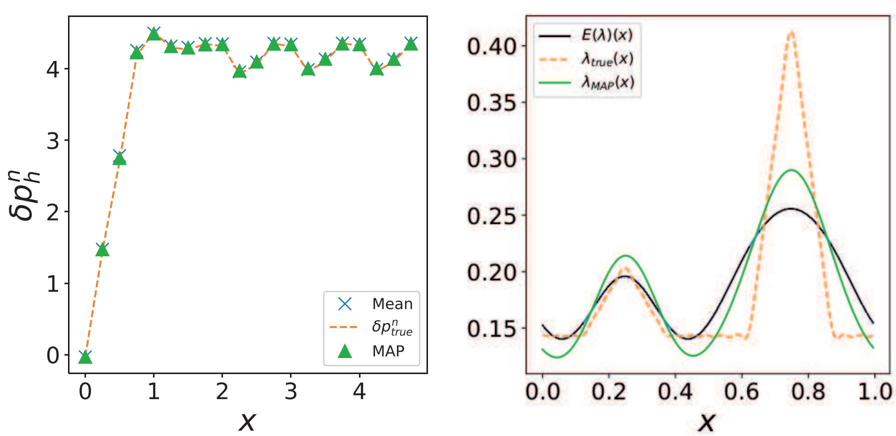

Figure 5. Observed pressure drop compared to the pressure drop computed by the conditional mean and the MAP (left) and true value of λ compared to the conditional mean and MAP (right)

Table 1. Number of accepted proposals for different N (the total number of samples is 10 000)

N	1	3	5	7	9	11	13	15
MCMC	2236	70	6	1	0	0	0	0
pCN	2011	1922	1860	1872	1930	1874	1846	1746

We refer, for example, to Kaipio and Somersalo (2005) for more details on the hierarchical setting. Furthermore, please note that we cannot expect to recover the true function from 40 noisy data points. Both estimates reflect the influence of the prior regularization, see Fig. 5. The concentration of the posterior distribution to the true parameter is a very interesting problem, both theoretically and numerically. The concentration is a highly desirable situation in practice, since it relates to informative or large data. However, it can pose a computational challenge for numerical methods based on the prior measure. See, e.g., Schillings, Sprungk and Wacker (2019) for more details.

5.3. Increasing the truncation parameter

We now increase the truncation parameter N in the prior model to investigate the dimension robust behavior of the pCN (in comparison to the Gaussian random walk). The setting is as follows: $\Gamma = 0.1^2 \times I$, 40 observations and 10 000 proposals, $N = 0, \dots, 7$, $C = I$, $\Sigma = \sigma^2 I$. The stepsizes of MCMC and pCN, i.e. σ, β are chosen such that the initial experiment shows approximately an acceptance rate of 20 %. We fix those parameters to observe the dimension robust behavior of the pCN. Observe that the acceptance rate for the Gaussian random walk decreases as N grows while pCN is robust with respect to N. We refer to Cotter et al. (2013) for a detailed discussion on the dimension robustness of pCN.

6. Conclusion and outlook

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 friction function. More precisely, provided that the friction function is of Lipschitz class, the underlying PDE has a strong solution (Definition 1) 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 results of this paper build a foundation for problems arising from gas networks. As a natural extension of this work, we are working towards Bayesian inverse problems related to leakage detection in gas pipes. Moreover, our results will be used to provide a robust friction function as an input parameter for gas pipe models within the collaborative research center SFB-TRR 154 on *Modelling, Simulation and Optimization at the Example of Gas Networks* (see `https://trr154.fau.de/index.php/en/`) for projects which relate to real applications. 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 and MCMC, 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

Abgrall R. and Mishra, S. (2017) Uncertainty quantification for hyperbolic systems of conservation Laws. In: *Handbook of Numerical Methods for Hyperbolic Problems*, R. Abgrall and Chi-Wang Shu, eds., *Handb. Numer. Anal.*, **18**. Elsevier/North-Holland, Amsterdam, 507–544.

Apte, A., Hairer, M., Stuart, A.M. and Voss, S. (2007) Sampling the posterior: an approach to non-Gaussian data assimilation. *Phys. D*, 230, 50–64.

Assmann, D., Liers, F. and Stingl, D. (2017) Decomposable robust two-stage optimization: An application to gas network operations under uncertainty. Preprint, TRR 154.

Bernardo, J.-M. and Smith, A. F. M. (1994) *Bayesian Theory. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons, Ltd., Chichester.

Bijl, H., Lucor, D., Mishra S. and Schwab, D. (2013) *Uncertainty Quantification in Computational Fluid Dynamics. Lecture Notes in Computational Science and Engineering*. Springer International Publishing.

Birolleau, A. , Poëtte, G. and Lucor, D. (2014) Adaptive Bayesian inference for discontinuous inverse problems, application to hyperbolic conservation laws. *Commun. Comput. Phys.*, 16, 1–34.

Chen, V., Dunlop, M. M., Papaspiliopoulos, O. and Stuart, A.M. (2018) Dimension Robust MCMC in Bayesian Inverse Problems. arXiv: 1803.03344.

Cotter, S.L., Dashti, M., Robinson, J.C. and Stuart, A.M. (2009) Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Problems*, 25, 115008, 43.

Cotter, S.L., Roberts, G. O., Stuart, A.M. and White, D. (2011) MCMC methods for functions: Modifying old algorithms to make them faster. *Statistical Science*, 28, 283–464.

Dashti, M. and Stuart, A.M. (2011) Uncertainty quantification and weak approximation of an elliptic inverse problem. *SIAM J. Numer. Anal.*, 49, 2524–2542.

Dashti, M. and Stuart, A.M (2016) *The Bayesian Approach to Inverse Problems.* Springer International Publishing, Cham, 1–118.

Dick, J., Gantner, R.N., Le Gia, Q. T. and Schwab, C. (2016) Higher order quasi-Monte Carlo integration for Bayesian estimation. *Tech. Report 2016-13, Seminar for Applied Mathematics*, ETH Zürich.

Domschke, P., Hiller, B., Lang, J. and Tischendorf, C. (2017) Modellierung von Gasnetzwerken: Eine Übersicht. Preprint, TRR 154.

Egger, H., Kugler,T. and Strogies, N. (2017) Parameter identiffcation in a semilinear hyperbolic System. *Inverse Problems*, 33, 055022, 25.

El Moselhy, T.A. and Marzouk, Y.M. (2012) Bayesian inference with optimal maps. *Journal of Computational Physics*, 231, 7815–7850.

Ern, A. and Guermond, J.-L. (2004) *Theory and Practice of Finite Elements. Applied Mathematical Sciences*, **159**. Springer-Verlag, New York.

Gonzalez Grandon, T., Heitsch, H. and Henrion, R. (2017) A joint model of probabilistic/robust constraints for gas transport management in stationary networks. *Computational Management Science*, 14, 443–460.

Hajian, S. Hintermüller, M. and Ulbrich, S. (2017) Total variation diminishing schemes in optimal control of scalar conservation laws. *IMA Journal of Numerical Analysis.*

Hayden, K., Olson, E. and Titi, E. S. (2011) Discrete data assimilation in the Lorenz and 2D Navier–Stokes equations. *Physica D: Nonlinear Phenomena*, 240, 1416 – 1425.

Hintermüller, M. and Strogies, N. (2017a) On the consistency of Runge–Kutta methods up to order three applied to the optimal control of scalar conservation laws. Preprint, TRR 154. Accepted in *Proceedings to NAOIV-2017.*

Hintermüller, M. and Strogies, N. (2017b) On the identiffcation of the friction coeffcient in a semilinear system for gas transport through a network. Preprint, TRR 154.

Hoang, V.H., Schwab, C. and Stuart, A.M. (2013) Complexity analysis of accelerated MCMC methods for Bayesian inversion. *Inverse Problems*, 29, 085010, 37.

Kaipio, J. and Somersalo, E. (2005) *Statistical and Computational Inverse Problems. Applied Mathematical Sciences*, **160**. Springer-Verlag, New York.

LeVeque, R. J. (2002) *Finite Volume Methods for Hyperbolic Problems. Cambridge Texts in Applied Mathematics.* Cambridge University Press, Cambridge.

Majda, A.J. and Harlim, J. (2012) *Filtering Complex Turbulent Systems.* Cambridge University Press, 1.

Matthies, H.G., Zander, E., Rosić, B.V. and Litvinenko, A. (2016) Parameter estimation via conditional expectation: a Bayesian inversion. *Advanced Modeling and Simulation in Engineering Sciences*, 3, 24.

Pazy, A. (1983) *Semigroups of Linear Operators and Applications to Partial Differential Equations. Applied Mathematical Sciences*, **44**. Springer-Verlag, New York.

Scheichl, R., Stuart, A. M. and Teckentrup, A.L. (2017) Quasi-Monte Carlo and multilevel Monte Carlo methods for computing posterior expectations in elliptic inverse problems. *SIAM/ASA J. Uncertain. Quantif.*, 5, 493–518.

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

Schillings, C., Sprungk, B. and Wacker, P. (2019) On the convergence of the Laplace approximation and noise-level-robustness of Laplace-based Monte Carlo methods for Bayesian inverse problems. arXiv: 1901.03958

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

Stuart, A. M. (2010) Inverse problems: a Bayesian perspective. *Acta Numerica*, 19.

Sullivan, T.J. (2015) *Introduction to Uncertainty Quantification. Texts in Applied Mathematics*, **63**. Springer, Cham.

Tarantola, A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

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

Ulbrich, S. (2001) Optimal Control of Nonlinear Hyperbolic Conservation Laws with Source Terms. Habilitation thesis, TUM – Technical University of Munich.