HAL open science

Reducing sensors for transient heat transfer problems by means of variational data assimilation

Amina Benaceur

To cite this version:

Amina Benaceur. Reducing sensors for transient heat transfer problems by means of variational data assimilation. SMAI Journal of Computational Mathematics, 2021, 7, pp.1-25. 10.5802/smai-jcm.68 . hal-02265533v2

HAL Id: hal-02265533

https://hal.science/hal-02265533v2

Submitted on 25 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SMAI-JCM
SMAI Journal of Computational Mathematics

Reducing sensors for transient heat transfer problems by means of variational data assimilation

Amina Benaceur

Volume 7 (2021), p. 1-25.
<http://smai-jcm.centre-mersenne.org/item?id=SMAI-JCM_2021__7__1_0>

© Société de Mathématiques Appliquées et Industrielles, 2021
Certains droits réservés.

Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/

Soumission sur https://smai-jcm.centre-mersenne.org/ojs/submission

SMAI Journal of Computational Mathematics
Vol. 7, 1-25 (2021)

Reducing sensors for transient heat transfer problems by means of variational data assimilation

Amina Benaceur [1]

[1] Massachusetts Institute of Technology, Cambridge, USA.
University Paris-Est, CERMICS (ENPC), 77455 Marne la Vallée Cedex 2 and INRIA Paris, 75589 Paris, France.
EDF Lab Les Renardières, 77250 Ecuelles Cedex, France.
E-mail address: benaceur@mit.edu.

Abstract. We propose a contribution that combines model reduction with data assimilation. A dedicated Parametrized Background Data-Weak (PBDW) approach has been introduced in the literature so as to combine numerical models with experimental measurements. We extend the approach to a time-dependent framework by means of a `POD-greedy` reduced basis construction. Since the construction of the basis is performed offline, the algorithm addresses the time dependence of the problem while the time stepping scheme remains unchanged. Moreover, we devise a new algorithm that exploits offline state estimates in order to diminish both the dimension of the online PBDW statement and the number of required sensors collecting data. The idea is to exploit *in situ* observations in order to update the best-knowledge model, thereby improving the approximation capacity of the background space.

2020 Mathematics Subject Classification. 65K10.

Keywords. PBDW, model reduction, data assimilation, sensor reduction, heat transfer.

1. Introduction

State estimation is a task in which the quantity of interest is the "true" state u^{true} of a physical system over a space or space-time domain of interest. However, numerical prediction based on a given mathematical model may be deficient due to limitations imposed by available knowledge. In other words, the mathematical model can only take anticipated or parametric uncertainty into account. A more accurate prediction requires the incorporation of experimental observations in order to accommodate unanticipated or non-parametric uncertainty.

For this reason, data assimilation methods have been widely explored in the literature. The goal of these methods is to use *a priori* information to deduce the best mathematical model, while using available experimental data to produce the most accurate approximation of a physical system. Many data assimilation methods involve the minimization of a cost function, such as least-squares type methods, designed to compute the mismatch between the model approximation and given observations. For instance, we mention the Stable Least Squares approximation [7, 8] and the linearized Structured Total Least Squares [9]. As opposed to statistical techniques [40, 22], variational data assimilation relies more heavily on the mathematical model. Considering the adjoint method [34, 39], the goal is to reconstruct a physical state through the minimization of a cost function in order to optimize the parameters of the model with respect to experimental data. However, this method is only based on the precision of the mathematical model and mainly uses experimental data to correct the input, but not the final state. The Proper Orthogonal Decomposition (POD) [24] has also been exploited in a data assimilation perspective, such as for the Gappy-POD [41, 12], which was introduced to handle incomplete data-sets, or as in [33] where data assimilation is enhanced using the POD method. We also mention other existing methods such as the 3D-VAR [26] and the 4D-VAR [23] for variational data assimilation. One of the drawbacks of many variational data assimilation methods is their computational intrusivity, which means that at any stage, computational procedures need to access the

model in order to perform their calculations. Intrusivity is very inconvenient in many contexts, for instance when using industrial high-fidelity black-box solvers. Hence, non-intrusive or partially-intrusive options can be valuable. The methods used in [41, 33] are less intrusive. Nevertheless, they have not been applied with reduced order models.

The Parameterized-Background Data-Weak (PBDW) formulation for variational data assimilation is a data-driven reduced order modeling approach that was initially devised in [30] so as to merge prediction by model with prediction by data, without being intrusive. It can represent the physics of the state using a sophisticated model, and applies non-intrusive and non-iterative real-time variational data assimilation employing Reduced Basis (RB) methods [21, 35], with correction of model error. The PBDW relies on the knowledge of some particular solutions to the parameterized model, and some measurements over the physical state to be approximated. Unlike classical design-of-experiment approaches [13], the PBDW emphasised state estimation rather than parameter estimation. The PBDW approach has been developed in order to estimate the true state u^{true} for several configurations of a physical system. Supposing that the true state u^{true} depends on some unknown parameter ω in an unknown parameter set Θ that represents the unanticipated uncertainty, the goal is to account for the dependency of the true state $u^{\text{true}}(\omega)$ on uncertain parameters by means of the sole knowledge of data. In this paper, whenever the context is unambiguous, the parameter ω is dropped.

The formulation combines a so-called "best-knowledge" (`bk`) model represented by a parametrized partial differential equation (PDE) and experimentally observable measurements. The use of data in the PBDW approach is fundamental not only to reconstruct the quantities of interest, but also to correct the possible bias in the mathematical `bk` model. The PBDW approach provides the following attractive features:

- The PBDW variational formulation simplifies the construction of *a priori* error estimates which can guide the optimal choice of the experimental observations.
- The PBDW formulation uses a `bk` model that accommodates anticipated uncertainty associated with the parameters of the model in a computationally convenient way. This `bk` model is typically built using model-order reduction techniques.

Note that the PBDW formulation does not explicitly include the equations of the `bk` model, but only a finite collection of solutions to the `bk` model. Thus, another important feature of the PBDW approach is its non-intrusivity. In fact, once a `bk` space has been generated, we no longer need the `bk` model.

The PBDW approach has been subject to active research in recent years. It has been used for several applications. For instance, [38] applies the PBDW for structural health monitoring; [19] proposes a non-intrusive PBDW with application to urban dispersion modeling frameworks; and [1, 2] exploit the generalized empirical interpolation method [27, 28, 29] in a data interpolation perspective. As a further step towards efficient industrial implementation, [31] develops a PBDW approach based on noisy observations and [32] introduces an adaptive PBDW approach with a user-defined update space. Moreover, [14] tackles the case of nonlinear problems with an unsteady application. Besides, the issue of noisy observations is also addressed in [15]. Further work regarding the selection of experimental observation functionals has been conducted in [37], whereas [36] proposes a localization procedure that treats the uncertainty related to boundary conditions. Finally, [6] proposes optimal choices for the reduced models and [5] addresses optimal measurements in state estimation using reduced models.

The PBDW approach was devised in [30] for steady problems. In this paper, we propose, as initiated in [3], an extension of the PBDW approach to time-dependent state estimation. Two main contributions to the standard PBDW approach are presented:

- We build appropriate background spaces for the time-dependent setting using the `POD-greedy` algorithm [16]. Here, the `bk`-model is approximated in space using the Finite Element Method (FEM). The `POD-greedy` procedure has been combined with FEM in previous work [25, 18, 4].

- We propose a modified offline stage so as to alleviate its computational cost which can be sizeable in a time-dependent setting. The new offline stage allows for better computational efficiency owing to a smaller online system. Moreover, it achieves substantial cost savings associated with data collection since it diminishes the number of observation sensors needed online. Note that this modified offline stage can also be applied to a steady framework.

This paper is organized as follows. Section 2 sets the notation and introduces the PBDW setting in the stationary framework as introduced in the literature. The reader familiar with this material can directly skip to Section 3 which extends the PBDW approach to the time-dependent framework. In Section 4, we discuss the offline stage and present the suggested algorithm in detail. Finally, in Section 5, we illustrate our method by some numerical results. We focus on standard parabolic problems exemplified by the linear and nonlinear heat problem. We expect similar behaviour if a moderate advection reaction term is added to the model. The more delicate question of dominant low-order perturbations is left for future work. The main idea and results presented in this work are inspired from the PhD thesis of the author [3].

2. Basic notation and stationary PBDW

We consider a spatial domain (open, bounded, connected subset) $\Omega \subset \mathbb{R}^d$, $d \geq 1$, with a Lipschitz boundary. We introduce a Hilbert space $\mathcal{U}$ composed of functions defined over Ω. The space $\mathcal{U}$ is endowed with an inner product $(\cdot,\cdot)$ and we denote by $\|\cdot\|$ the induced norm; $\mathcal{U}$ consists of functions $\{w : \Omega \to \mathbb{R} \mid \|w\| < \infty\}$. To fix the ideas, we assume that $H_0^1(\Omega) \subset \mathcal{U} \subset H^1(\Omega)$, and we denote the dual space of $\mathcal{U}$ by $\mathcal{U}'$. The Riesz operator $R_{\mathcal{U}} : \mathcal{U}' \to \mathcal{U}$ satisfies, for each $\ell \in \mathcal{U}'$, and for all $v \in \mathcal{U}$, the equality $(R_{\mathcal{U}}(\ell), v) = \ell(v)$. For any closed subspace $\mathcal{Q} \subset \mathcal{U}$, the orthogonal complement of $\mathcal{Q}$ is defined as $\mathcal{Q}^\perp := \{w \in \mathcal{U} \mid (w, v) = 0,\ \forall v \in \mathcal{Q}\}$. Finally, we introduce a parameter set $\mathcal{P} \subset \mathbb{R}^p,\ p \geq 1$, whose elements are generically denoted by $\mu \in \mathcal{P}$.

The first source of information we shall afford ourselves in the PBDW approach is a so-called "best-knowledge" (`bk`) mathematical model in the form of a parametrized PDE posed over the domain Ω (or more generally, over a domain Ω^{bk} such that $\Omega \subset \Omega^{\mathrm{bk}}$). Given a parameter value μ in the parameter set $\mathcal{P}$, we denote the solution to the `bk` parametrized PDE as $u^{\mathrm{bk}}(\mu) \in \mathcal{U}$. Then, we introduce the manifold associated with the solutions of the `bk` model $\mathcal{M}^{\mathrm{bk}} := \{u^{\mathrm{bk}}(\mu) \mid \mu \in \mathcal{P}\} \subset \mathcal{U}$. In ideal situations, the true solution u^{true} is well approximated by the `bk` manifold, i.e., the model error

$$\epsilon^{\mathrm{bk}}_{\mathrm{mod}}(u^{\mathrm{true}}) := \inf_{z \in \mathcal{M}^{\mathrm{bk}}} \|u^{\mathrm{true}} - z\|, \tag{2.1}$$

is very small.

We introduce nested background subspaces $\mathcal{Z}_1 \subset \ldots \subset \mathcal{Z}_N \subset \ldots \subset \mathcal{U}$ that are generated to approximate the `bk` manifold $\mathcal{M}^{\mathrm{bk}}$ to a certain accuracy. These subspaces can be built using various model-order reduction techniques, for instance, the Reduced Basis method. Note that the indices of the subspaces conventionally indicate their dimensions. To measure how well the true solution is approximated by the background space $\mathcal{Z}_N$, we define the quantity $\epsilon^{\mathrm{bk}}_N(u^{\mathrm{true}}) := \inf_{z \in \mathcal{Z}_N} \|u^{\mathrm{true}} - z\|$. The background space is built so that $\epsilon^{\mathrm{bk}}_N(u^{\mathrm{true}}) \underset{N \to +\infty}{\longrightarrow} \epsilon^{\mathrm{bk}}_{\mathrm{mod}}(u^{\mathrm{true}})$. Moreover, we introduce the reduction error $\epsilon^{\mathrm{bk}}_{\mathrm{red},N} := \sup_{u \in \mathcal{M}^{\mathrm{bk}}} \inf_{z \in \mathcal{Z}_N} \inf \|u - z\|$, which encodes the loss of accuracy caused by solving the `bk` model in the N-dimensional background space $\mathcal{Z}_N$. For later purposes, we introduce $\Pi_{\mathcal{Z}_N}(u^{\mathrm{true}})$ as the closest point to u^{true} in $\mathcal{Z}_N$. Note that $\Pi_{\mathcal{Z}_N}$ is the $\mathcal{U}$-orthogonal projection onto $\mathcal{Z}_N$. The background space $\mathcal{Z}_N$ can be interpreted as a prior space that approximates the `bk` manifold which we hope approximates well the true state u^{true}. As previously alluded to, u^{true} rarely lies in $\mathcal{M}^{\mathrm{bk}}$ in realistic engineering study cases.

The remainder of this Section is a brief review of the PBDW method in the stationary case.

2.1. **Unlimited-observations statement**

Let us first describe an ideal situation. The unlimited-observations PBDW statement reads: find $(u_N^*, z_N^*, \eta_N^*) \in \mathcal{U} \times \mathcal{Z}_N \times \mathcal{U}$ such that

$$(u_N^*, z_N^*, \eta_N^*) = \underset{\substack{u_N \in \mathcal{U} \\ z_N \in \mathcal{Z}_N \\ \eta_N \in \mathcal{U}}}{\operatorname{arginf}} \|\eta_N\|^2, \tag{2.2a}$$

$$\text{subject to } (u_N, v) = (\eta_N, v) + (z_N, v), \quad \forall v \in \mathcal{U}, \tag{2.2b}$$

$$(u_N, \phi) = (u^{\text{true}}, \phi), \qquad \forall \phi \in \mathcal{U}. \tag{2.2c}$$

A direct consequence of (2.2b) is that

$$u_N^* = z_N^* + \eta_N^*, \tag{2.3}$$

whereas (2.2c) implies that $u_N^* = u^{\text{true}}$.

The so-called "deduced background estimate" z_N^* can only represent anticipated uncertainty. Since the `bk` mathematical model of a physical system is often deficient, one cannot realistically assume that the state estimate u_N^* of u^{true} lies completely in the `bk` manifold (or in the background space $\mathcal{Z}_N$). Therefore, the so-called "update estimate" η_N^* is meant to cure the deficiency of the `bk` model by capturing unanticipated uncertainty. In other words, the key idea of the PBDW statement (2.2) is to search for the smallest correction to the `bk` manifold. The following result is proved in [30].

The Euler–Lagrange saddle-point problem associated with the PBDW statement (2.2) reads: find $(z_N^*, \eta_N^*) \in \mathcal{Z}_N \times \mathcal{U}$ such that

$$\begin{cases} (\eta_N^*, q) + (z_N^*, q) = (u^{\text{true}}, q), & \forall q \in \mathcal{U}, \\ (\eta_N^*, p) = 0, & \forall p \in \mathcal{Z}_N, \end{cases} \tag{2.4}$$

and set $u_N^* = z_N^* + \eta_N^*$.

2.2. **Observable space**

The evaluation of the right-hand side (u^{true}, q) in (2.4) requires the full knowledge of the true state u^{true} which is unrealistic. In practice, one can only afford a limited number of experimental observations of the true state u^{true}. In the present setting, the experimental observations are interpreted as the application of prescribed observation functionals $\ell_m^{\text{obs}} \in \mathcal{U}'$ for all $m \in \{1, \dots, M\}$ such that the m-th experimental observation is given by $\ell_m^{\text{obs}}(u^{\text{true}}) \in \mathbb{R}, \quad \forall m \in \{1, \dots, M\}$. One can consider any observation functional that renders the behavior of some physical sensor.

Generally, we introduce an experimentally observable space $\mathcal{U}_M := \operatorname{Span}\{q_m\}_{1 \le m \le M} \subset \mathcal{U}$, where $q_m := R_{\mathcal{U}}(\ell_m^{\text{obs}})$ is the Riesz representation of $\ell_m^{\text{obs}} \in \mathcal{U}'$, for all $m \in \{1, \dots, M\}$. The experimental observations of the true state satisfy $(u^{\text{true}}, q_m) = \ell_m^{\text{obs}}(u^{\text{true}})$, for all $m \in \{1, \dots, M\}$. Hence, for all $q \in \mathcal{U}_M$ such that $q = \sum_{m=1}^M \alpha_m q_m$ and $(\alpha_m)_{1 \le m \le M} \in \mathbb{R}^M$, the inner product (u^{true}, q) can be deduced from the experimental observations as a linear combination of the M available observations:

$$\left(u^{\text{true}}, q\right) = \sum_{m=1}^M \alpha_m \left(u^{\text{true}}, q_m\right) = \sum_{m=1}^M \alpha_m \ell_m^{\text{obs}}(u^{\text{true}}). \tag{2.5}$$

2.3. **Limited-observations statement**

Let us now describe the PBDW statement in the case of limited observations. Henceforth, we make the crucial assumption that

$$\mathcal{Z}_N \cap \mathcal{U}_M^\perp = \{0\}, \tag{2.6}$$

which is meant to ensure the well-posedness of the PBDW statement with limited observations [3]. The limited-observations PBDW statement reads: find $(u^*_{N,M}, z^*_{N,M}, \eta^*_{N,M}) \in \mathcal{U} \times \mathcal{Z}_N \times \mathcal{U}$ such that

$$(u^*_{N,M}, z^*_{N,M}, \eta^*_{N,M}) = \underset{\substack{u_{N,M} \in \mathcal{U} \\ z_{N,M} \in \mathcal{Z}_N \\ \eta_{N,M} \in \mathcal{U}}}{\operatorname{arginf}} \|\eta_{N,M}\|^2, \tag{2.7a}$$

$$\text{subject to} \quad (u_{N,M}, v) = (\eta_{N,M}, v) + (z_{N,M}, v), \quad \forall v \in \mathcal{U}, \tag{2.7b}$$

$$(u_{N,M}, \phi) = (u^{\text{true}}, \phi), \quad \forall \phi \in \mathcal{U}_M. \tag{2.7c}$$

As above, (2.7b) implies that the limited-observations state estimate $u^*_{N,M}$ satisfies

$$u^*_{N,M} = z^*_{N,M} + \eta^*_{N,M}. \tag{2.8}$$

One can show (e.g., by introducing the Lagrangian) that the limited-observations problem (2.7) is equivalent to the limited-observations saddle-point problem: find $(z^*_{N,M}, \eta^*_{N,M}) \in \mathcal{Z}_N \times \mathcal{U}_M$ such that

$$(\eta^*_{N,M}, q) + (z^*_{N,M}, q) = (u^{\text{true}}, q), \quad \forall q \in \mathcal{U}_M, \tag{2.9a}$$

$$(\eta^*_{N,M}, p) = 0, \quad \forall p \in \mathcal{Z}_N, \tag{2.9b}$$

and define $u^*_{N,M}$ according to (2.8). The linear system (2.9) is well posed under assumption (2.6).

Remark 2.1 (Perfect background space)**.** The choice of the background space $\mathcal{Z}_N$ and of the observable space $\mathcal{U}_M$ may lead to several specific configurations. In particular, the background space $\mathcal{Z}_N$ is said to be perfect if the reduction error vanishes, i.e., $u^{\text{true}} \in \mathcal{Z}_N$. In this case, the pair $(u^{\text{true}}, 0) \in \mathcal{Z}_N \times \mathcal{U}_M$ is the unique solution to (2.9). Hence, the state estimate $u^*_{N,M}$ also belongs to $\mathcal{Z}_N$ and the update estimate satisfies $\eta^*_{N,M} = 0$.

2.4. **Algebraic formulation**

We now present the algebraic formulation of the limited-observations PBDW statement. We first introduce an $\mathcal{N}$-dimensional approximation space $\mathcal{U}^{\mathcal{N}}$ of the infinite-dimensional space $\mathcal{U}$ as well as discrete approximation spaces $\mathcal{Z}_N \subset \mathcal{U}^{\mathcal{N}}$ and $\mathcal{U}_M \subset \mathcal{U}^{\mathcal{N}}$ of the subspaces $\mathcal{Z}_N$ and $\mathcal{U}_M$, respectively. These spaces are built using finite elements [10]. We assume that the size of the mesh is small enough so that the $\mathcal{N}$-dimensional space discretization delivers High-Fidelity (HF) approximations within the requested level of accuracy. To alleviate the notation, we have dropped the superscript $\mathcal{N}$; hence, the discrete FEM spaces are denoted $\mathcal{Z}_N$ and $\mathcal{U}_M$ instead of $\mathcal{Z}_N^{\mathcal{N}}$ and $\mathcal{U}_M^{\mathcal{N}}$, but we still keep the notation $\mathcal{U}^{\mathcal{N}}$ for the FEM-discretization space. Then, we introduce a basis for the background space $\mathcal{Z}_N := \text{Span}\{\zeta_n\}_{1 \le n \le N}$. The update space is spanned by the Riesz representations of the observation functionals in $\mathcal{U}^{\mathcal{N}}$, i.e., $\mathcal{U}_M := \text{Span}\{q_m\}_{1 \le m \le M}$, where $q_m \in \mathcal{U}^{\mathcal{N}}$, for all $m \in \{1, \dots, M\}$. The high-fidelity (HF) discretization of the saddle-point problem (2.9) is: Find $(z^*_{N,M}, \eta^*_{N,M}) \in \mathcal{Z}_N \times \mathcal{U}_M$ such that

$$(\eta^*_{N,M}, q) + (z^*_{N,M}, q) = (u^{\text{true}}, q), \quad \forall q \in \mathcal{U}_M, \tag{2.10a}$$

$$(\eta^*_{N,M}, p) = 0, \quad \forall p \in \mathcal{Z}_N. \tag{2.10b}$$

The solution of (2.10) is then searched under the form

$$z^*_{N,M} = \sum_{n=1}^{N} z_n \zeta_n, \quad \text{and} \quad \eta^*_{N,M} = \sum_{m=1}^{M} \eta_m q_m, \tag{2.11}$$

and we introduce the component vectors $\boldsymbol{z}^*_{N,M} := (z_n)_{1 \le n \le N} \in \mathbb{R}^N$ and $\boldsymbol{\eta}^*_{N,M} := (\eta_m)_{1 \le m \le M} \in \mathbb{R}^M$. We also introduce the basis matrices $\mathbf{Z}_N \in \mathbb{R}^{\mathcal{N} \times N}$ and $\mathbf{U}_M \in \mathbb{R}^{\mathcal{N} \times M}$ whose column vectors are the

components of the functions $\{\zeta_n\}_{1\le n\le N}$ and $\{q_m\}_{1\le m\le M}$ respectively in the basis of $\mathcal{U}^{\mathcal{N}}$. In algebraic form, the FEM-discretized saddle-point problem (2.10) reads: find $(\boldsymbol{z}^*_{N,M}, \boldsymbol{\eta}^*_{N,M}) \in \mathbb{R}^N \times \mathbb{R}^M$ such that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\eta}^*_{N,M} \\ \boldsymbol{z}^*_{N,M} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\ell}^{\text{obs}}_M \\ \mathbf{0} \end{pmatrix}, \tag{2.12}$$

with the matrices

$$\mathbf{A} = \Big((q_{m'}, q_m)\Big)_{1\le m,m'\le M} \in \mathbb{R}^{M\times M},\ \mathbf{B} = \Big((\zeta_n, q_m)\Big)_{1\le m\le M, 1\le n\le N} \in \mathbb{R}^{M\times N}, \tag{2.13}$$

and the vector of observations $\boldsymbol{\ell}^{\text{obs}}_M = \big(\ell^{\text{obs}}_m(u^{\text{true}})\big)_{1\le m\le M} \in \mathbb{R}^M$.

In practice, the matrices $\mathbf{A}$ and $\mathbf{B}$ are computed using the algebraic formulas $\mathbf{A} = \mathbf{U}_M\mathbf{M}\mathbf{U}_M$, and $\mathbf{B} = \mathbf{Z}_N\mathbf{M}\mathbf{U}_M$, where $\mathbf{M}$ is the Gram matrix of the inner product in $\mathcal{U}$. Thus, solving (2.12) allows for a straightforward reconstruction of the components of the state estimate in the basis of $\mathcal{U}^{\mathcal{N}}$ as follows:

$$\boldsymbol{u}^*_{N,M} = \mathbf{Z}_N\boldsymbol{z}^*_{N,M} + \mathbf{U}_M\boldsymbol{\eta}^*_{N,M}. \tag{2.14}$$

3. Time-dependent PBDW

Consider a finite time interval $I = [0, T]$, with $T > 0$. To discretize in time, we consider an integer $K \ge 1$, we define $0 = t^0 < \cdots < t^K = T$ as $(K + 1)$ distinct time nodes over I, and we set $\mathbb{K}^{\text{tr}} = \{1, \ldots, K\}$, $\overline{\mathbb{K}}^{\text{tr}} = \{0\} \cup \mathbb{K}^{\text{tr}}$ and $I^{\text{tr}} = \{t^k\}_{k\in\overline{\mathbb{K}}^{\text{tr}}}$. This section aims at deriving a state estimate for a time-dependent solution.

Remark 3.1 (Initial condition)**.** In the present setting, we choose not to solve the PBDW statement for the initial time node $k = 0$. It is straightforward to consider a setting where the initial time node is also included.

3.1. Unlimited-observations statement

In this ideal setting, we assume that $u^{\text{true}} \in \mathcal{C}^0(I;\mathcal{U})$. Here, the observations are considered unlimited in both space and time. The time-dependent unlimited-observations PBDW statement reads: for each $k \in \mathbb{K}^{\text{tr}}$, find $(u^{k,*}_N, z^{k,*}_N, \eta^{k,*}_N) \in \mathcal{U} \times \mathcal{Z}_N \times \mathcal{U}$ such that

$$(u^{k,*}_N, z^{k,*}_N, \eta^{k,*}_N) = \underset{\substack{u_N\in\mathcal{U}\\ z_N\in\mathcal{Z}_N\\ \eta_N\in\mathcal{U}}}{\text{arginf}}\ \|\eta_N\|^2, \tag{3.1a}$$

$$\text{subject to } (u_N, v) = (\eta_N, v) + (z_N, v), \qquad \forall v \in \mathcal{U}, \tag{3.1b}$$

$$(u_N, \phi) = (u^{k,\text{true}}, \phi), \qquad \forall \phi \in \mathcal{U}. \tag{3.1c}$$

where $u^{k,\text{true}} := u^{\text{true}}(t^k, \cdot)$.

Lemma 3.2 (Solution)**.** *For each $k \in \mathbb{K}^{\text{tr}}$, the solution to* (3.1) *is given by*

$$u^{k,*}_N = u^{k,\text{true}}, \quad z^{k,*}_N = \Pi_{\mathcal{Z}_N}(u^{k,\text{true}}), \quad \eta^{k,*}_N = \Pi_{\mathcal{Z}_N^\perp}(u^{k,\text{true}}). \tag{3.2}$$

Proof. From (3.1c), we infer that $u^{k,*}_N = u^{k,\text{true}}$. Next, we deduce from (3.1b) that $u^{k,\text{true}} = z^{k,*}_N + \eta^{k,*}_N$. Since (3.1a) is a minimization of $\|\eta^{k,*}_N\|$, it follows that $z^{k,*}_N = \Pi_{\mathcal{Z}_N}(u^{k,\text{true}})$. Thus, $\eta^{k,*}_N = \Pi_{\mathcal{Z}_N^\perp}(u^{k,\text{true}})$. ∎

The Euler–Lagrange saddle-point problem associated with the time-dependent PBDW statement (3.1) reads: for each $k \in \mathbb{K}^{\rm tr}$, find $(z_N^{k,*}, \eta_N^{k,*}) \in \mathcal{Z}_N \times \mathcal{U}$ such that

$$(\eta_N^{k,*}, q) + (z_N^{k,*}, q) = (u^{k,\rm true}, q), \quad \forall q \in \mathcal{U}, \tag{3.3a}$$

$$(\eta_N^{k,*}, p) = 0, \qquad \forall p \in \mathcal{Z}_N. \tag{3.3b}$$

The unlimited-observations state estimate is then

$$u_N^{k,*} = z_N^{k,*} + \eta_N^{k,*}, \quad \forall k \in \mathbb{K}^{\rm tr}. \tag{3.4}$$

The saddle-point problem (3.3) is purely geometric and does not include any explicit reference to the `bk` model. The unique link to the `bk` model is through the background space $\mathcal{Z}_N$. Therefore, the PBDW approach is applicable to a wide class of engineering problems. Moreover, the non-intrusivity of (3.3) simplifies its implementation. In practice, non-intrusivity means that computational procedures do not need to access the bk-model in order to perform their calculations.

3.2. **Limited-observations statement**

We now weaken the regularity assumption on the true state and only assume that $u^{\rm true} \in L^1(I;\mathcal{U})$. We introduce the time-integration intervals

$$\mathcal{I}^k = [t^k - \delta t_k, t^k + \delta t_k], \quad \forall k \in \mathbb{K}^{\rm tr}, \tag{3.5}$$

where $\delta t^k > 0$ is a parameter related to the precision of the sensor (ideally, $\delta t^k < \min(t^{k+1} - t^k, t^k - t^{k-1})$ with obvious adaptation if k=K). Then, for any function $v \in L^1(I;\mathcal{U})$, we define the time-averaged snapshots

$$v^k(x) := \frac{1}{|\mathcal{I}^k|} \int_{\mathcal{I}^k} v(t,x)\ dt \in \mathcal{U}, \quad \forall k \in \mathbb{K}^{\rm tr}. \tag{3.6}$$

The evaluation of the right-hand side $(u^{k,\rm true}, q)$ in (3.3) requires the full knowledge of the true state $u^{k,\rm true}$, which is unrealistic. In practice, one can only afford a limited number of experimental observations of the true state $u^{k,\rm true}$. Therefore, we consider observation functionals that render the behavior of given sensors. We use the same observation functionals as in the time-independent setting [30], but we let them act on the time-averaged snapshots of the true solution, i.e., we consider

$$\ell_m^{k,\rm obs}(u^{\rm true}) := \ell_m^{\rm obs}(u^{k,\rm true}), \quad \forall m \in \{1, \dots, M\},\ \forall k \in \mathbb{K}^{\rm tr}. \tag{3.7}$$

For instance, if the sensors act through local uniform time integration, we have

$$\ell_m^{k,\rm obs}(u^{\rm true}) = \frac{1}{|\mathcal{R}_m|} \int_{\mathcal{R}_m} u^{k,\rm true}(x)\ dx = \frac{1}{|\mathcal{R}_m|} \frac{1}{|\mathcal{I}_k|} \int_{\mathcal{R}_m} \int_{\mathcal{I}_k} u^{\rm true}(t,x)\ dxdt, \tag{3.8}$$

whereas if the sensors act through integration against a Gaussian, we have

$$\begin{aligned} \ell_m^{k,\rm obs}(u^{\rm true}) &= \frac{1}{\sqrt{2\pi r_m^2}} \int_{\mathcal{R}_m} u^{k,\rm true}(x) \exp\left(\frac{-(x - x_m^c)^2}{2r_m^2}\right)\ dxdt, \\ &= \frac{1}{|\mathcal{I}_k|} \frac{1}{\sqrt{2\pi r_m^2}} \int_{\mathcal{I}_k} \int_{\mathcal{R}_m} u^{\rm true}(x) \exp\left(\frac{-(x - x_m^c)^2}{2r_m^2}\right)\ dxdt. \end{aligned} \tag{3.9}$$

Generally, we introduce the time-independent observable space $\mathcal{U}_M \subset \mathcal{U}$ such that

$$\mathcal{U}_M = \mathrm{Span}\{q_1, \dots, q_M\}. \tag{3.10}$$

Remark 3.3 (Sensors)**.** The most convenient configuration to collect observations in industrial contexts is to measure the quantities at user-defined space-time locations. In actual practice, sensors do not take pointwise measures but localized ones. A sensor collects the data that is enclosed in a small

area centered at the sensor location. Hence, equation (3.8) means that the sensor returns a measurement that is equal to the space-time averaged quantity we are collecting. The setup defined by equation (3.9) means that the sensor returns a Gaussian-like measurement over a centered space-time domain. Moreover, the observation functionals are Riesz representors of any type of physical data available for the user, and not necessarily pointwise or pointwise-like (integrations over small patches) observations.

Remark 3.4 (Functionals)**.** The functionals ℓ_m^{obs} in (3.7) are fixed over time. If they change over time, the space $\mathcal{U}_M$ in (3.10) will change at every time step. Another possibility is to have a larger space $\mathcal{U}_M$ with optional information in specific directions (corresponding to physical locations of the sensors) at each time step as soon as stability is granted. Yet, such an approach can suffer from continuity issues due to the change of directions with available measurements between time steps.

The observation functionals in $\mathcal{U}'$ are then defined as

$$\ell_m^{k,\text{obs}}(u^{\text{true}}) = (u^{k,\text{true}}, q_m), \quad \forall m \in \{1, \dots, M\}, \quad \forall k \in \mathbb{K}^{\text{tr}}. \tag{3.11}$$

Note that, for fixed sensor locations, the computational effort to compute the Riesz representations of the observation functionals is time-independent and is incurred only once so that the experimental observations of the true state satisfy

$$\ell_m^{k,\text{obs}}(u^{\text{true}}) = \left(u^{k,\text{true}}, q_m\right) = \frac{1}{|\mathcal{I}_k|} \int_{\mathcal{I}^k} \ell_m^{\text{obs}}(u^{\text{true}}(t,\cdot))dt, \ \forall m \in \{1, \dots, M\}, \ \forall k \in \mathbb{K}^{\text{tr}}. \tag{3.12}$$

Hence, for all $q \in \mathcal{U}_M$ such that,

$$q = \sum_{m=1}^{M} \alpha_m q_m, \tag{3.13}$$

the inner product $\left(u^{k,\text{true}}, q\right)$ is deduced from the experimental observations as follows:

$$\left(u^{k,\text{true}}, q\right) = \frac{1}{|\mathcal{I}_k|} \int_{\mathcal{I}^k} \sum_{m=1}^{M} \alpha_m \left(u^{\text{true}}(t,\cdot), q_m\right) dt = \frac{1}{|\mathcal{I}_k|} \sum_{m=1}^{M} \alpha_m \int_{\mathcal{I}^k} \ell_m^{\text{obs}}(u^{\text{true}}(t,\cdot))dt. \tag{3.14}$$

Henceforth, we make the crucial assumption that

$$\mathcal{Z}_N \cap \mathcal{U}_M^{\perp} = \{0\}, \tag{3.15}$$

which is also equivalent to

$$\beta_{N,M} := \inf_{w \in \mathcal{Z}_N} \sup_{v \in \mathcal{U}_M} \frac{(w,v)}{\|w\| \ \|v\|} \in (0,1], \tag{3.16}$$

where $\beta_{N,M}$ is the so-called stability constant (the reader can refer to [3] for a proof). Assumption (3.15) can be viewed as a requirement to have enough sensors (note that $\mathcal{Z}_N \cap \mathcal{U}^{\perp} = \{0\}$). Under this assumption, the limited-observations PBDW statement reads: for each $k \in \mathbb{K}^{\text{tr}}$, find $(u_{N,M}^{k,*}, z_{N,M}^{k,*}, \eta_{N,M}^{k,*}) \in \mathcal{U} \times \mathcal{Z}_N \times \mathcal{U}$ such that

$$(u_{N,M}^{k,*}, z_{N,M}^{k,*}, \eta_{N,M}^{k,*}) = \underset{\substack{u_{N,M} \in \mathcal{U} \\ z_{N,M} \in \mathcal{Z}_N \\ \eta_{N,M} \in \mathcal{U}}}{\text{arginf}} \ \|\eta_{N,M}\|^2, \tag{3.17a}$$

$$\text{subject to } (u_{N,M}, v) = (\eta_{N,M}, v) + (z_{N,M}, v), \qquad \forall v \in \mathcal{U}, \tag{3.17b}$$

$$(u_{N,M}, \phi) = (u^{k,\text{true}}, \phi), \qquad \forall \phi \in \mathcal{U}_M. \tag{3.17c}$$

The limited-observations saddle-point problem associated with (3.17) reads: for each $k \in \mathbb{K}^{\mathrm{tr}}$, find $(z_{N,M}^{k,*}, \eta_{N,M}^{k,*}) \in \mathcal{Z}_N \times \mathcal{U}_M$ such that

$$(\eta_{N,M}^{k,*}, q) + (z_{N,M}^{k,*}, q) = (u^{k,\mathrm{true}}, q), \quad \forall q \in \mathcal{U}_M, \tag{3.18a}$$

$$(\eta_{N,M}^{k,*}, p) = 0, \qquad \forall p \in \mathcal{Z}_N, \tag{3.18b}$$

and the limited-observations state estimate is

$$u_{N,M}^{k,*} = z_{N,M}^{k,*} + \eta_{N,M}^{k,*}, \quad \forall k \in \mathbb{K}^{\mathrm{tr}}. \tag{3.19}$$

Remark 3.5 (Pointwise measurements)**.** For simplicity of implementation, assuming that $u^{\mathrm{true}} \in \mathcal{C}^0(I;\mathcal{U})$, one may consider pointwise measurements in time, i.e.,

$$\left(u^{k,\mathrm{true}}, q_m\right) = \ell_m^{\mathrm{obs}}(u^{\mathrm{true}}(t^k, \cdot)), \quad \forall m \in \{1, \dots, M\}, \forall k \in \mathbb{K}^{\mathrm{tr}}. \tag{3.20}$$

The assumption (3.20) is typically reasonable for a sensor of small precision δt^k.

In algebraic form, the limited-observations PBDW statement reads: for each $k \in \mathbb{K}^{\mathrm{tr}}$, find $(\boldsymbol{z}_{N,M}^{k,*}, \boldsymbol{\eta}_{N,M}^{k,*}) \in \mathbb{R}^N \times \mathbb{R}^M$ such that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\eta}_{N,M}^{k,*} \\ \boldsymbol{z}_{N,M}^{k,*} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\ell}_M^{k,\mathrm{obs}} \\ \mathbf{0} \end{pmatrix}, \tag{3.21}$$

with the matrices

$$\mathbf{A} = \Big((q_{m'}, q_m)\Big)_{1 \le m, m' \le M} \in \mathbb{R}^{M \times M}, \ \mathbf{B} = \Big((\zeta_n, q_m)\Big)_{1 \le m \le M, 1 \le n \le N} \in \mathbb{R}^{M \times N}, \tag{3.22}$$

and the vector of observations

$$\boldsymbol{\ell}_M^{k,\mathrm{obs}} = (\ell_m^{\mathrm{obs}}(u^{k,\mathrm{true}}))_{1 \le m \le M} \in \mathbb{R}^M. \tag{3.23}$$

Offline/online procedure: Since several realizations $u^{k,\mathrm{true}}(\omega)$ of the true state are considered, an offline/online procedure can be employed to solve (3.21). During the offline stage, one precomputes the RB functions $(\zeta_n)_{1 \le n \le N}$ and the Riesz representers $(q_m)_{1 \le m \le M}$ leading to the matrices $\mathbf{A} \in \mathbb{R}^{M \times M}$ and $\mathbf{B} \in \mathbb{R}^{N \times M}$ once and for all. Then, during the online stage, for each new set of observations corresponding to a new realization of the true state $u^{k,\mathrm{true}}(\omega)$, all that remains to be performed is to form the vector of observations $\boldsymbol{\ell}_M^{k,\mathrm{obs}}(\omega)$ and to retrieve the deduced background estimate $\boldsymbol{z}_{N,M}^{k,*}(\omega)$ and the update estimate $\boldsymbol{\eta}_{N,M}^{k,*}(\omega)$ by solving the $(N+M)$-dimensional linear problem (3.21). The PBDW state estimate $\boldsymbol{u}_{N,M}^{k,*}(\omega)$ is then computed using (3.19).

Remark 3.6 (PBDW matrices)**.** Notice that the PBDW matrices $\mathbf{A}$ and $\mathbf{B}$ are time-independent; only the right-hand side in (3.21) depends on k.

Remark 3.7 (Choice of spaces)**.** The subspaces $\mathcal{Z}_N$ and $\mathcal{U}_M$ must be chosen carefully. In fact, we want a small angle between the spaces $\mathcal{Z}_N$ and $\mathcal{U}_M$ in order to increase the stability constant $\beta_{N,M}$, but we need some overlap between the spaces $\mathcal{Z}_N^\perp$ and $\mathcal{U}_M$ to improve the approximation capacity of $\mathcal{Z}_N^\perp \cap \mathcal{U}_M$. We refer the reader to [30] for a detailed *a priori* error analysis and to [3] for an overview on the influence of the spaces $\mathcal{Z}_N$ and $\mathcal{U}_M$ on the accuracy of the results. We highlight that using multiple time-dependent bases did not appear to be beneficial in our numerical experiments. In particular, this choice leads to higer storage requirements and the need for higher time sampling.

4. Offline stage

In this section, we discuss the offline stage. Our main goal is to address the construction of the background space $\mathcal{Z}_N$.

4.1. Background space construction via `POD-greedy`

Suppose that we have computed a set of high-fidelity (HF) trajectories

$$\mathcal{S} = (\mathcal{S}_k)_{k\in\mathbb{K}^{\mathrm{tr}}} = \big((u^k(\mu))_{\mu\in\mathcal{P}^{\mathrm{tr}}}\big)_{k\in\mathbb{K}^{\mathrm{tr}}}, \tag{4.1}$$

where $u^k(\mu) := u(\mu)(t^k, \cdot)$, for all $k \in \mathbb{K}^{\mathrm{tr}}$. If we were to consider the PBDW statement (3.17) for each $k \in \mathbb{K}^{\mathrm{tr}}$ as an independent steady PBDW statement, we would be working with the time-dependent background spaces

$$\mathcal{Z}^k_{N^k} = \mathtt{POD}(\mathcal{S}_k, \epsilon_{\mathrm{POD}}), \quad \forall k \in \mathbb{K}^{\mathrm{tr}}, \tag{4.2}$$

where the procedure `POD` is introduced in [24]. However, this strategy is not convenient since the sizes N^k of the background spaces $\mathcal{Z}^k_{N^k}$ would depend on k. Since the observable space $\mathcal{U}_M$ is fixed, the same non-homogeneity between time nodes would also arise in the stability constant $\beta_{N^k,M}$. Thus, we propose to apply a `POD-greedy` algorithm [16, 17] in order to build a time-independent background space $\mathcal{Z}_N$ that will be used for all $k \in \mathbb{K}^{\mathrm{tr}}$. The advantage is that the PBDW matrices $\mathbf{A}$ and $\mathbf{B}$ and the stability constant $\beta_{N,M}$ remain unchanged regardless of the discrete time node. The offline stage using the `POD-greedy` algorithm is summarized in Algorithm 1.

Algorithm 1 Offline stage via `POD-greedy`

Input : $\mathcal{S}$ and ϵ_{POD}.
$\mathcal{Q}^{\mathrm{init}}$: an initial set of Riesz representations for the observations.

1: Compute $\mathcal{Z}_N := \mathtt{POD\text{-}greedy}(\mathcal{S}, \epsilon_{\mathrm{POD}})$.
2: Set $\mathcal{U}_M := \mathrm{Span}\{\mathcal{Q}^{\mathrm{init}}\}$.
3: Compute the matrices $\mathbf{A}$ and $\mathbf{B}$ using $\mathcal{Z}_N$ and $\mathcal{U}_M$.

Output : $\mathcal{Z}_N$, $\mathcal{U}_M$, $\mathbf{A}$ and $\mathbf{B}$.

4.2. Background space construction via state estimation

We now devise a new algorithm in the context of time-dependent PBDW to perform the offline stage.

4.2.1. *Main idea*

Here, the construction of the background space $\mathcal{Z}_N$, the choice of the observation space $\mathcal{U}_M$ and the PBDW matrices are modified. The key idea of the new procedure is to precompute the PBDW state estimates of the parameters in the training set $\mathcal{P}^{\mathrm{tr}}$ during the offline stage. The background space is then deduced from these PBDW state estimates. The benefit is that the newly created background space incorporates data-based knowledge. The modified offline stage of the PBDW for time-dependent problems is described in Algorithm 2. Within the modified offline stage, we use the so-called "Greedy stability maximization" (`S-Greedy`) algorithm (considered in [30]) in line 6 in order to identify the least stable mode and then take the best measurement. The algorithm uses an input space $\mathcal{Z}_N$ that results from a `POD-greedy` procedure so that $\mathcal{Z}_1$ contains the dominant mode, and so forth. The `S-greedy` algorithm selects the observations progressively. Thus, the enrichment of the observable

space $\mathcal{U}_M$ stops once the minimum stability β_{MIN} is reached. The procedure `S-greedy` is described in Algorithm 3 below. Altogether, the modified offline stage in the proposed algorithm offers four major

Algorithm 2 Modified offline stage of the time-dependent PBDW

Input : $\mathcal{P}^{\text{tr}}$, $\mathbb{K}^{\text{tr}}$, $\mathcal{S}$, ϵ_{POD}, $\epsilon_{\text{POD}}^{\text{init}}$ and β_{MIN}.
$\mathcal{Q}^{\text{init}}$: an initial set of Riesz representations of observations.

1: Compute $\mathcal{Z}_{N^{\text{init}}}^{\text{init}} := \texttt{POD-greedy}(\mathcal{S}, \epsilon_{\text{POD}}^{\text{init}})$.
2: Set $\mathcal{U}_{M^{\text{init}}}^{\text{init}} := \text{Span}\{\mathcal{Q}^{\text{init}}\}$.
3: Compute the matrices $\mathbf{A}^{\text{init}}$ and $\mathbf{B}^{\text{init}}$ using $\mathcal{Z}_{N^{\text{init}}}^{\text{init}}$ and $\mathcal{U}_{M^{\text{init}}}^{\text{init}}$.
4: Estimate the state $u^{k,*}(\mu)$ for all $(\mu, k) \in \mathcal{P}^{\text{tr}} \times \mathbb{K}^{\text{tr}}$.
5: Compute $\mathcal{Z}_N := \texttt{POD-greedy}(\{u^{k,*}(\mu)\}_{\mu \in \mathcal{P}^{\text{tr}}, k \in \mathbb{K}^{\text{tr}}}, \epsilon_{\text{POD}})$.
6: Compute $\mathcal{U}_M := \texttt{S-Greedy}(\mathcal{P}^{\text{tr}}, \mathbb{K}^{\text{tr}}, N, \{\mathcal{Z}_n\}_{n=1}^N, \beta_{\text{MIN}}, \mathcal{Q}^{\text{init}})$.
7: Compute the matrices $\mathbf{A}$ and $\mathbf{B}$ using $\mathcal{Z}_N$ and $\mathcal{U}_M$.

Output : $\mathcal{Z}_N$, $\mathcal{U}_M$, $\mathbf{A}$ and $\mathbf{B}$.

advantages:

(1) **Improved background space:** Since the background space $\mathcal{Z}_N$ is built using both the `bk` model and the observations, it is expected to have better approximation capacities of the true state.

(2) **Reduced number of online observations:** In line 6 of Algorithm 2, we select each new data point so as to maximize the stability constant $\beta_{N,M}$. Thus, the observations that will be used during the online stage are mainly needed only for stability and not for accuracy.

(3) **Reduced dimension of the online PBDW statement:** Since the number of observations is significantly reduced, the modified PBDW matrices are of smaller size compared to the matrices of the standard PBDW. Thus, using the modified offline algorithm, the online PBDW formulation is solved faster.

(4) **Reduced storage cost:** Owing to the reduced number of online measurements, the dimensions of the observable space $\mathcal{U}_M$ and of the matrices $\mathbf{A}$ and $\mathbf{B}$ are smaller, whence the storage gain.

Regarding computational efficiency, the modified procedure consists of more steps than in the standard PBDW. However, all the additional steps of the algorithm are performed offline. As for all reduced-order modeling techniques, the goal of the algorithm is to further improve the online efficiency. Hence, the computational savings brought by the new PBDW formulation come, in our opinion, at a reasonable offline price. Indeed, the resolution of the (online) standard PBDW statement for each parameter $\mu \in \mathcal{P}^{\text{tr}}$ has a reduced computational cost. The only relevant additional computational cost incurred offline is related to the second `POD-greedy` algorithm (cf. line 5 of Algorithm 2). We believe this computational effort remains acceptable.

Remark 4.1 (Least stable mode)**.** Line 6 of Algorithm 3 may return several infima. Among these infima, we select a function whose norm in $\mathcal{U}$ is maximal.

Remark 4.2 (Steady setting)**.** In a time-dependent framework, the computational savings induced by the modified offline stage are substantial, in particular because of the influence of the time steps. However, Algorithm 3 can be applied in the steady setting as well.

Algorithm 3 `S-Greedy`: Stability-maximization algorithm

Input : N, $\mathcal{Z}_N$ and $\beta_{\text{MIN}} \in (0, 1]$.
$\mathcal{Q}^{\text{init}}$: an initial set of Riesz representations of the observations.

1: Choose a random $q_1 \in \mathcal{Q}^{\text{init}}$.
2: Set $\mathcal{U}_1 := \text{Span}\{q_1\}$.
3: Compute the stability constant $\beta_{1,1}$ using $\mathcal{Z}_1$ and $\mathcal{U}_1$.
4: Set $m := 2$.
5: **while** $\beta_{N,m-1} < \beta_{\text{MIN}}$ or $m < M$ **do**
6: Compute the least stable mode and the associated supremizer

$$w_{\text{inf}} \in \underset{w \in \mathcal{Z}_N}{\text{arginf}} \sup_{v \in \mathcal{U}_{m-1}} \frac{(w, v)}{\|w\| \, \|v\|}, \quad \text{and} \quad v_{\text{sup}} = \Pi_{\mathcal{U}_{m-1}}(w_{\text{inf}}).$$

7: Identify the least well-approximated vector $q_m = \underset{q \in \mathcal{Q}^{\text{init}}}{\text{argsup}} \, |(q, w_{\text{inf}} - v_{\text{sup}})|$.
8: Set $\mathcal{U}_m := \text{Span}\{\mathcal{U}_{m-1}, q_m\}$.
9: Compute the stability constant $\beta_{N,m}$.
10: $m = m + 1$.
11: **end while**
12: $M := m$.

Output : $\mathcal{U}_M$.

4.2.2. *Practical aspects*

As the locations of the sensors are fixed in many industrial applications, the offline optimized strategy does not always assume the freedom to place the sensors. In fact, the sensor placement is an option in the procedure to maximize accuracy, but the overall strategy proves efficient even for fixed sensor locations as illustrated in the applications of Section 5. To be more precise, the initial choice of the Riesz representations of observations $\mathcal{Q}^{\text{init}}$ considered as an input in Algorithm 2 can be left to the user depending on their industrial constraints, and must not necessarily be performed using the `S-Greedy` algorithm. The same reasoning can be applied when updating the observable space in line 6 of Algorithm 2. However, we emphasize that the final set of observations must be a subset of the initial set $\mathcal{Q}^{\text{init}}$.

Most practical applications in the literature are cases for which measurement values must come from experiments that were performed offline during a one-shot data collection step, meaning implicitly that the user is allowed to work with the same number of sensors throughout the entire experiment. On the contrary, the optimized offline procedure also handles practical cases for which the sensor placement may be reduced within the offline stage. This context is related to real case scenarii in which a first data collection campaign has been performed at a first occasion with insufficient results. When the experimental data collection needs to be redeployed for further investigation of a physical system, the optimized offline stage diminishes substantially its cost by taking into account the previous data, complemented by a less expensive renewed sensor placement. This can typically be the case in industrial research and development engineering units, such as that which hosted the author of the paper during the elaboration of this work. Yet, the optimized offline procedure can still be used for the usual one-shot data collection framework. In this case, the benefit of the strategy is to further diminish the size of the online system, thereby improving the storage and resolution performances of the online stage.

5. Numerical results

In this section, we illustrate the above developments on test cases related to the heat equation. The goal

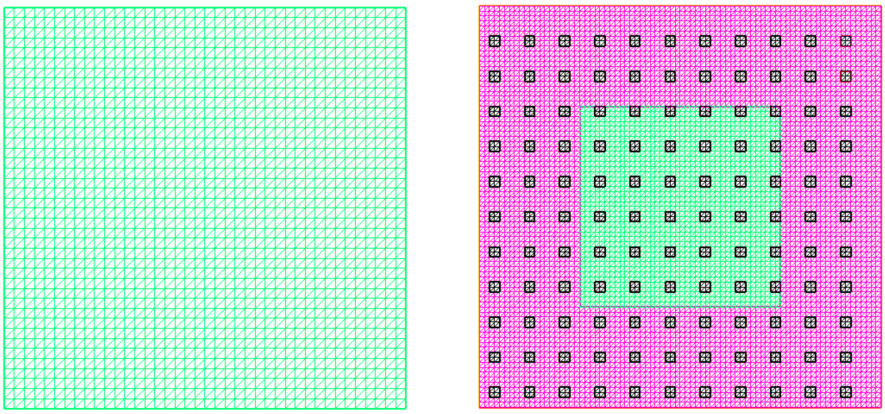

Figure 5.1. Computational domain and mesh with $\mathcal{N} = 6561$. The little black squares are observation subsets $\{\mathcal{R}_m\}_{m=1}^{121}$. Left: Mono-material plate. Right: Bi-material plate.

is to illustrate the computational performance of our algorithms. In all our test cases, we consider a two-dimensional setting based on the plate illustrated in the left panel of Figure 5.1 with $\Omega = (-2, 2)^2 \subset \mathbb{R}^2$. We use a finite element (FE) [11] subspace $\mathcal{U}^{\mathcal{N}} \subset \mathcal{U} = H^1(\Omega)$ consisting of continuous, piecewise affine functions in order to generate HF trajectories. The FE subspace $\mathcal{U}^{\mathcal{N}}$ is based on a mesh that contains $\mathcal{N} = 6561$ nodes. The experimental data is generated synthetically and the observation subsets $\{\mathcal{R}_m\}_{1\leq m\leq M}$ are uniformly selected over the plate as illustrated in the left panel of Figure 5.1. Regarding the implementation, the HF computations use the software `FreeFem++` [20], whereas the reduced-order modeling and the PBDW-related algorithms have been developed in `Python`.

5.1. Physical model problem

We apply the above methodology to the following parabolic PDE: For many values of the parameter $\mu \in \mathcal{P}$, find $u(\mu) : I \times \Omega \to \mathbb{R}$ such that

$$\begin{cases} \dfrac{\partial u(\mu)}{\partial t} - \nabla \cdot (D(\mu)\nabla u(\mu)) = 0, & \text{in } I \times \Omega, \\ u(\mu)(t = 0, \cdot) = u_0, & \text{in } \Omega, \\ \text{Boundary conditions}, & \text{on } I \times \partial\Omega, \end{cases} \tag{5.1}$$

where $u_0 = 293.15\text{K}$ (20^oC). We will supplement (5.1) with two types of boundary conditions:

(1) **Linear heat equation:** We apply a homogeneous Neumann boundary condition on $\partial\Omega_0$ and a non-homogeneous Neumann boundary condition on $\partial\Omega_n$, i.e.,

$$\begin{cases} -D(\mu)\dfrac{\partial u(\mu)}{\partial n} = 0, & \text{on } I \times \partial\Omega_0, \\ -D(\mu)\dfrac{\partial u(\mu)}{\partial n} = \phi_e, & \text{on } I \times \partial\Omega_n, \end{cases} \tag{5.2}$$

with $\phi_e = 3\text{K}\cdot\text{m}\cdot\text{s}^{-1}$, $\partial\Omega_0 = (-2, 2) \times \{2\} \cup \{2\} \times (-2, 2)$, and $\partial\Omega_n = (-2, 2) \times \{-2\} \cup \{-2\} \times (-2, 2)$. Thus, the resulting problem (5.1)–(5.2) is linear. Note that $\partial\Omega_0$ consists of the upper and right sides of the plate and $\partial\Omega_n$ consists of its lower and left sides, so that $\partial\Omega = \overline{\partial}\Omega_0 \cup \overline{\partial}\Omega_n$.

(2) **Nonlinear heat equation:** We apply Stefan–Boltzmann boundary conditions on $\partial\Omega$, i.e.,

$$-D(\mu)\frac{\partial u}{\partial n} = \sigma\varepsilon(u^4 - u_r^4), \qquad \text{on } I \times \partial\Omega, \tag{5.3}$$

where $u_r = 303.15\text{K}$ (30^oC) is an enclosure temperature, $\sigma = 5.67 \times 10^{-8}\text{W.m}^{-2}.\text{K}^{-4}$ is the Stefan–Boltzmann constant and $\varepsilon = 3.10^{-3}$ is the emissivity. The Stefan–Boltzmann boundary condition is nonlinear and so is the resulting problem (5.1)–(5.3).

In what follows, the background spaces $\mathcal{Z}_N$ will be generated by solving either the linear PDE (5.1)–(5.2) or the nonlinear PDE (5.1)-(5.3) with a uniform diffusivity function $D(\mu)$ such that for all $x \in \Omega$, $D(\mu)(x) = D_{\text{uni}}(\mu)(x) := \mu \mathbf{1}_\Omega(x)$.

5.2. **Synthetic data generation**

We synthesize the data by first synthesizing a true solution and then applying to it the linear functionals by means of their Riesz representations in the observable space $\mathcal{U}_M$. In order to synthesize the true solution, we consider a "true model" based on the bi-material plate (cf. right panel of Figure 5.1) where we choose a fixed internal diffusivity $D_{\text{int}} = 1$ and define, for each $\mu \in \mathcal{P}$, the diffusivity function $D(\mu)$ as $D(\mu)(x) = D_{\text{syn}}(\mu)(x) := \mu D_{\text{int}} \mathbf{1}_{\Omega_{\text{ext}}}(x) + D_{\text{int}} \mathbf{1}_{\Omega_{int}}(x)$, for all $x \in \Omega$, where $\Omega_{\text{int}} = (-1,1)^2$ and $\Omega_{\text{ext}} = (-2,2)^2 \setminus (-1,1)^2$, so that $\overline{\Omega} = \overline{\Omega}_{\text{int}} \cup \overline{\Omega}_{\text{ext}}$ and $\Omega_{\text{int}} \cap \Omega_{\text{ext}} = \emptyset$. The synthetic true solutions are then defined as the solutions of (5.1) for all $\mu \in \mathcal{P}$, with either the linear boundary condition (5.2) or the nonlinear boundary condition (5.3).

5.2.1. *Test configurations*

In order to investigate the PBDW formulation, we perform test cases on two distinct configurations:

(1) **Perfect model:** The `bk` model is said to be perfect when $\epsilon_{\text{mod}}^{\text{bk}}(u^{\text{true}}(\omega)) = 0$, for every $\omega \in \Theta$ (see (2.1)) (we recall that ω represents the unanticipated uncertainty). In this situation, $u^{\text{true}}(\omega) \in \mathcal{M}^{\text{bk}}$ for all $\omega \in \Theta$. Although the model is perfect, some discrepancies between the HF solutions and the measurements might arise from model-order reduction since $\mathcal{M}^{\text{bk}} \neq \mathcal{Z}_N$ (cf. Figure **??**). Note that this scenario seldom occurs in engineering situations. This test configuration is meant to assess the accuracy of the PBDW formulation when the observable space $\mathcal{U}_M$ scarcely has additional information compared to $\mathcal{Z}_N$.

(2) **Imperfect model:** The `bk` model is said to be imperfect when the modeling error does not vanish. In this situation, there exists at least one (and in general many) $\omega \in \Theta$ such that $\epsilon_{\text{mod}}^{\text{bk}}(u^{\text{true}}(\omega)) \neq 0$, i.e., $u^{\text{true}}(\omega) \notin \mathcal{M}^{\text{bk}}$. Consider for instance the plates in Figure 5.1. If the true solution is generated synthetically using the bi-material plate, an example of an imperfect `bk` model can be the one for which we solve the same PDE that has generated the true states without accounting for the difference in diffusivity between the subdomains of the plate.

5.3. **Background space construction via `POD-greedy`**

In this section, four test cases are considered to study the PBDW approach.

- Test case (a): Linear perfect.
- Test case (b): Linear imperfect.
- Test case (c): Nonlinear perfect.
- Test case (d): Nonlinear imperfect.

5.3.1. *Linear case*

Regarding time discretization, we consider the time interval $I = [0, 10]$s, the set of discrete time nodes $\mathbb{K}^{\text{tr}} = \{1, \ldots, 200\}$, and a constant time step $\Delta t^k = 0.05$s for all $k \in \mathbb{K}^{\text{tr}}$. Finally, we introduce the parameter interval $\mathcal{P} = [0.05, 1]$ and the training set $\mathcal{P}^{\text{tr}} = 0.05 \times \{1, \ldots, 20\}$. In Figure 5.2,

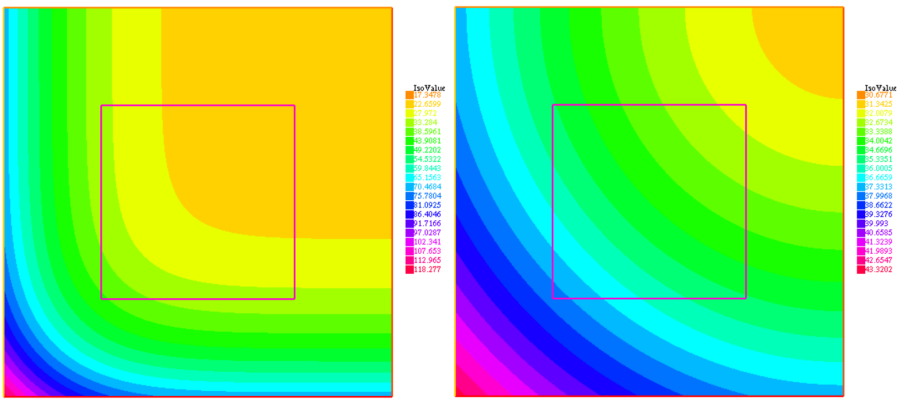

Figure 5.2. Test cases (a) and (b) : HF solutions for the `bk` model with Neumann boundary conditions. Left: $\mu = 1$ in D_{uni} (values from 17.3^{o}C to 118.3^{o}C). Right: $\mu = 20$ in D_{uni} (values from 30.7^{o}C to 43.3^{o}C).

we show the HF temperature profiles for the model problem (5.1)–(5.2) over the homogeneous plate at the end of the simulation, i.e., for $t^K = 10$s and for two parameter values. We recall that these solutions will be used as true solutions for the perfect linear case. As the simulation time evolves, the energy related to the flux ϕ_e propagates through the plate which is progressively heated. Moreover, the overall temperature is higher for smaller values of the parameter μ than for larger values. As physically expected, the thermal diffusion over the plate is stronger for larger values of μ than for smaller values.

Test case (a): Linear perfect model. We consider the case of a perfect `bk` model for which the diffusivity is uniform over the entire domain Ω. Thus, the true solutions correspond to the HF computations of the `bk` model. The resulting trajectories are reduced using the `POD-greedy` algorithm. For instance, for a tolerance value $\epsilon_{\text{POD}} = 10^{-2}$, the background space $\mathcal{Z}_N$ is composed of $N = 5$ modes. Regarding observations, the initial set $\mathcal{Q}^{\text{init}}$ is obtained using $M = \text{Card}(\mathcal{Q}^{\text{init}}) = 121$ sensors that are uniformly placed over the plate (see Figure 5.1). Using both the background space $\mathcal{Z}_N$ and the observable space $\mathcal{U}_M(\mathcal{Q}^{\text{init}})$, we build the offline matrices $\mathbf{A}$ and $\mathbf{B}$. During the online stage, we estimate the state $u^*_{N,M}$ for every parameter μ in the training set $\mathcal{P}^{\text{tr}}$. Using the weighted H^1-norm, the state estimation relative H^1-error $e^k(\mu)$ defined as

$$e^k(\mu) := \frac{\|u^{k,\text{true}}(\mu) - u^{k,*}_{N,M}(\mu)\|_{H^1(\Omega)}}{\|u^{k,\text{true}}(\mu)\|_{H^1(\Omega)}}, \quad \forall \mu \in \mathcal{P}, \tag{5.4}$$

is displayed in Figure 5.3 as a function of the value of the parameter μ for several values of ϵ_{POD}. In this first configuration, one can notice that the error decreases for smaller tolerances ϵ_{POD}, i.e., with the dimension N of the background space $\mathcal{Z}_N$. However, the right panel of Figure 5.3 shows a starting increase in the relative H^1-error $e^k(\mu)$ for $\epsilon_{\text{POD}} = 10^{-6}$ and an oscillatory behavior of the relative H^1-error $e^k(\mu)$. Although counter-intuitive in the reduced-basis context, this phenomenon is due to the deterioration of the stability constant $\beta_{N,M}$. This observation confirms the claims made in Remark 3.7.

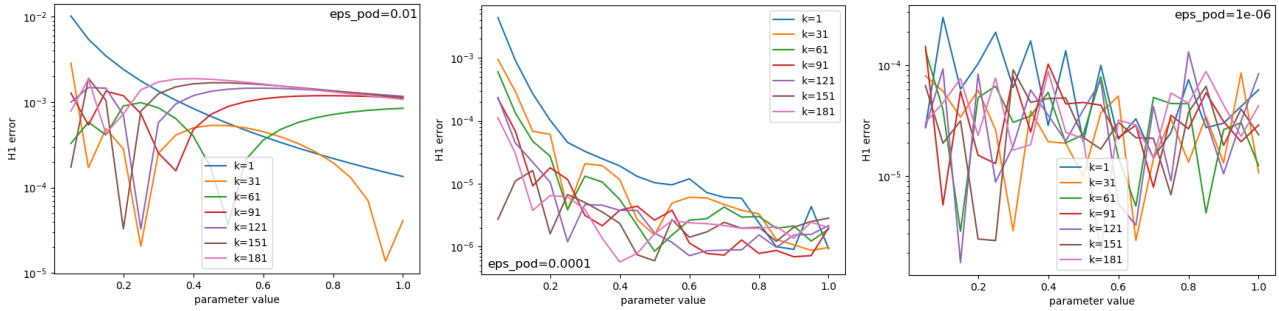

Figure 5.3. Test case (a): Relative H^1-error $e^k(\mu)$ for some time nodes $k \in \mathbb{K}^{\mathrm{tr}}$. Left: $\epsilon_{\mathrm{POD}} = 10^{-2}$ ($N = 5$). Middle : $\epsilon_{\mathrm{POD}} = 10^{-4}$ ($N = 10$). Right: $\epsilon_{\mathrm{POD}} = 10^{-6}$ ($N = 15$).

Test case (b): Linear imperfect model. This second test investigates the case of a linear imperfect `bk` model. In Figure 5.4, we show the HF temperature profiles for the true solutions over the

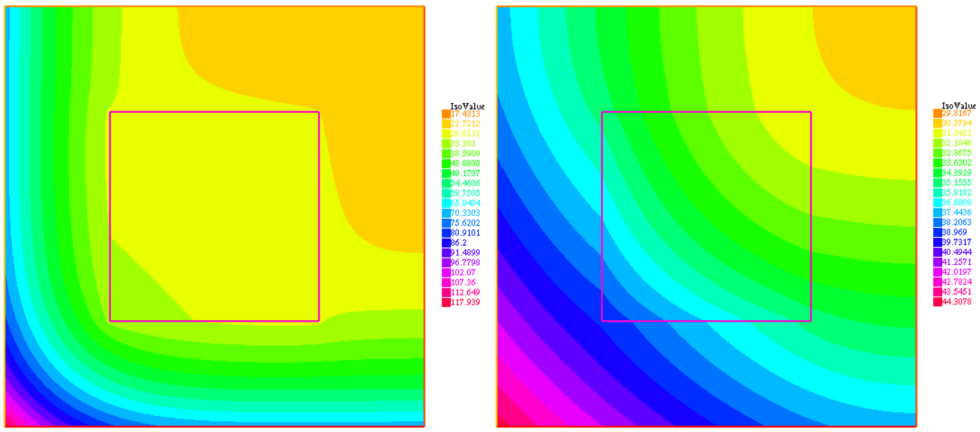

Figure 5.4. Test case (b) : Synthetic true solutions with Neumann boundary conditions. Left: $\mu = 1$ in $D_{\mathrm{syn}}(\mu)$ (values from 17.4^{o}C to 117.9^{o}C). Right: $\mu = 20$ in $D_{\mathrm{syn}}(\mu)$ (values from 29.8^{o}C to 44.3^{o}C).

bi-material plate at the end of the simulation, i.e., at $t^K = 10$s and for two different parameter values. The temperature fields exhibit the same overall behavior as in Figure 5.2. Additionally, we notice that the difference in diffusivity between Ω_{int} and Ω_{ext} leads as expected to a kink in the temperature isolines. When $\mu < 1$, the thermal diffusion is stronger in the inner plate corresponding to Ω_{int}, whereas for $\mu > 1$, the thermal diffusion is weaker in the inner plate. Using the HF trajectories produced by the `bk` model, we generate a background space $\mathcal{Z}_N$ by means of a `POD-greedy` algorithm. We use $M = 121$ observations to build the observable space $\mathcal{U}_M$. The relative H^1-errors $e^k(\mu)$ defined in (5.4) are shown in Figure 5.5 as a function of the value of the parameter μ. For instance, for a tolerance value $\epsilon_{\mathrm{POD}} = 10^{-3}$, $\mathcal{Z}_N$ is spanned by $N = 7$ vectors. One can notice that the error for the first time step $k = 1$ is lower than for the other time steps. This behavior is due to the time-marching which induces that the solution at the very first time step is close to the initial condition. Hence, the constant mode is the most prominent in the representation of the solution at $k = 1$. Besides, the error vanishes for $\mu = 0.5$ since this configuration is equivalent to a perfect `bk` model. However, the bottom panels of Figure 5.5 show a gradual error increase with the dimension N of the `bk` space. This tendency was already observed for the linear perfect test case, although in smaller proportions. As before, the

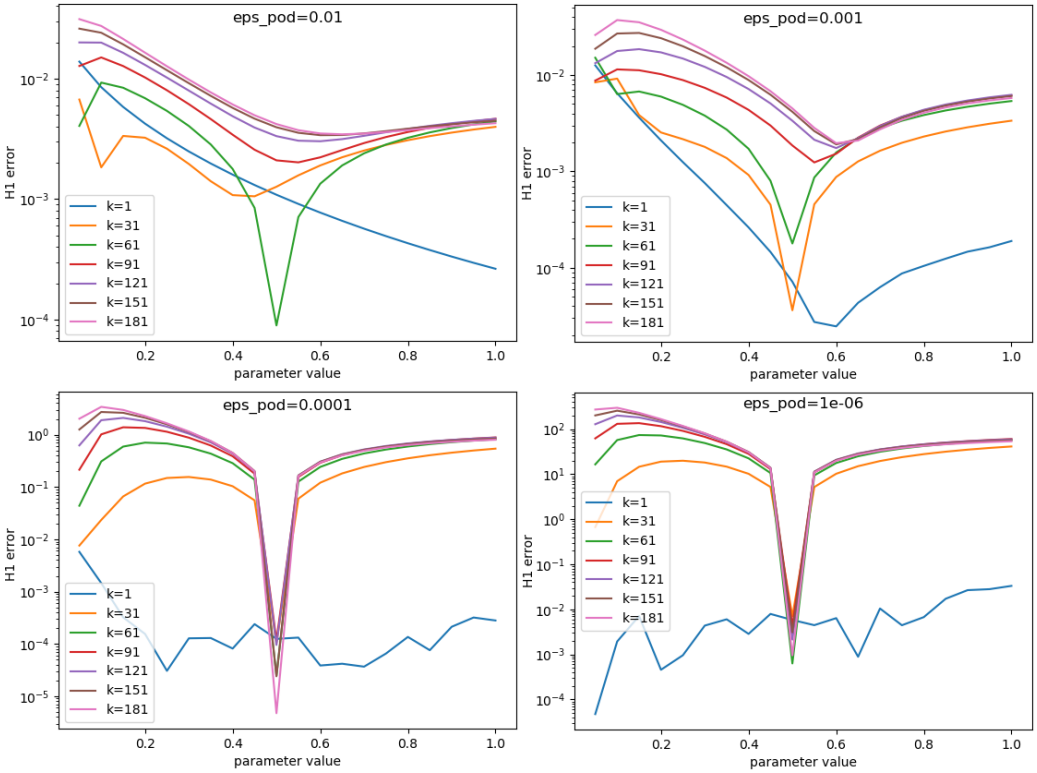

Figure 5.5. Test case (b) : Relative H^1-error $e^k(\mu)$ for some time nodes $k \in \mathbb{K}^{\text{tr}}$ and $M = 121$. Top left: $\epsilon_{\text{POD}} = 10^{-2}$ ($N = 5$). Top right: $\epsilon_{\text{POD}} = 10^{-3}$ ($N = 7$). Bottom left: $\epsilon_{\text{POD}} = 10^{-4}$ ($N = 10$). Bottom right: $\epsilon_{\text{POD}} = 10^{-6}$ ($N = 15$).

stability constant $\beta_{N,M}$ is degraded when increasing the dimension N of the background space $\mathcal{Z}_N$. Moreover, the enrichment of $\mathcal{Z}_N$ does not add relevant modes anymore (in terms of associated singular values). For the sake of comparison, we enrich the observable space $\mathcal{U}_M$ such that $M = 676$ and plot the relative H^1-errors $e^k(\mu)$ for the same values of ϵ_{POD} in Figure 5.6. Our interpretation is confirmed since the stability issues do not arise anymore for $\epsilon_{\text{POD}} = 10^{-4}$. Owing to the increase of M, the stability decrease with respect to N is somewhat compensated. Finally, the bottom-right panel of Figure 5.6 shows the beginning of an error increase. Using the same reasoning as above, we conclude that more observations are needed for $\epsilon_{\text{POD}} = 10^{-6}$.

Finally, we visualize the stability constant $\beta_{N,M}$ as a function of the number of observations M in Figure 5.7. The left panel of the figure shows a single curve for clarity, whereas the right panel includes curves for several values of the tolerance ϵ_{POD} (note that the two panels do not use the same rule). As expected, for a constant value of N, the more the observations, the better the stability. For a number of observations $M = 3000$, the PBDW formulation is perfectly stable (or close to) for all the considered values of ϵ_{POD}.

5.3.2. *Nonlinear case*

Here, we consider the PDE (5.1)–(5.3) with $u_r = 303.15$K, $\sigma = 5.67 \times 10^{-8}$W.m^{-2}.K^{-4} and $\varepsilon = 3.10^{-3}$. Except for the parameter interval $\mathcal{P} = [0.1, 2]$, the set $\mathcal{P}^{\text{tr}} = \{0.1i, 1 \leq i \leq 20\}$ and the time step $\Delta t^k = 0.1$, all the other numerical data remain the same as for the linear test case from the previous section.

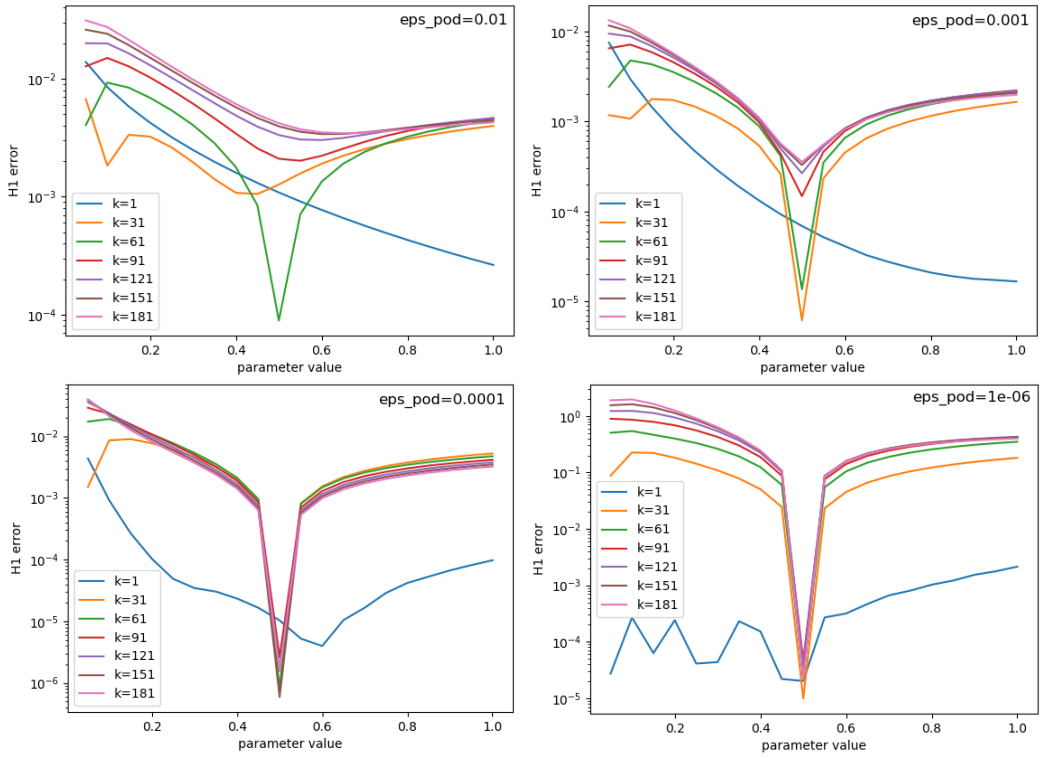

Figure 5.6. Test case (b) : Relative H^1-error $e^k(\mu)$ for some time nodes $k \in \mathbb{K}^{\mathrm{tr}}$ and $M = 676$. Top left: $\epsilon_{\mathrm{POD}} = 10^{-2}$ ($N = 5$). Top right: $\epsilon_{\mathrm{POD}} = 10^{-3}$ ($N = 7$). Bottom left: $\epsilon_{\mathrm{POD}} = 10^{-4}$ ($N = 10$). Bottom right: $\epsilon_{\mathrm{POD}} = 10^{-6}$ ($N = 15$).

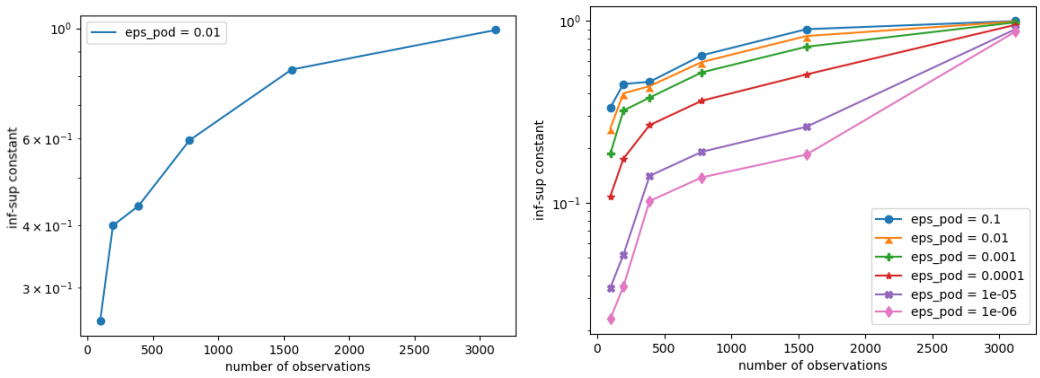

Figure 5.7. Test case (b) : Stability constant $\beta_{N,M}$ as a function of M. On the right panel, the values of N are respectively 3, 5, 7, 10, 13, 15 for the values of ϵ_{POD} in decreasing order.

Test case (c): Nonlinear perfect model. We consider the case with a perfect `bk` model. Thus, the true solutions correspond to the HF computations of the `bk` model (cf. left panel of Figure 5.8). The resulting trajectories are reduced using the `POD-greedy` algorithm. For instance, for a tolerance value $\epsilon_{\mathrm{POD}} = 10^{-2}$, the background space $\mathcal{Z}_N$ consists of $N = 3$ modes. Regarding observations, the initial set $\mathcal{Q}^{\mathrm{init}}$ is obtained using $M = \mathrm{Card}(\mathcal{Q}^{\mathrm{init}}) = 121$ sensors that are uniformly placed over the plate (see Figure 5.1). During the online stage, we estimate the state $u^*_{N,M}$ for every parameter μ in

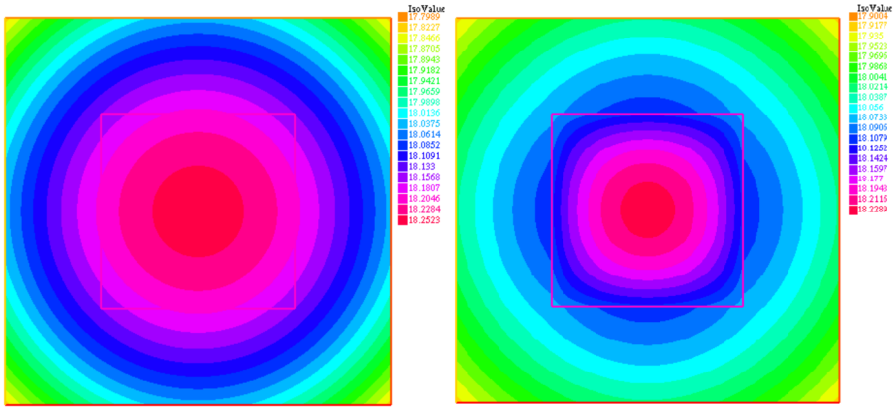

Figure 5.8. Test cases (c) and (d) : Left: HF solution for the `bk` model (values from 17.80^{o}C to 18.25^{o}C). Right: Synthetic true solution using a bi-material plate (values from 17.90^{o}C to 18.23^{o}C).

the training set $\mathcal{P}^{\text{tr}}$. In Figure 5.9, we display the state estimation relative H^1-error $e^k(\mu)$ defined in (5.4) as a function of the value of the parameter μ for several values of ϵ_{POD}. In contrast to the

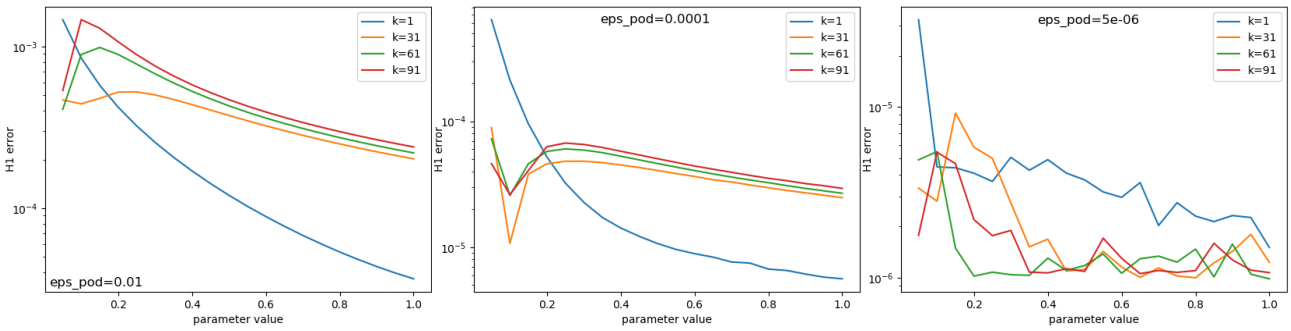

Figure 5.9. Test case (c) : Relative H^1-error $e^k(\mu)$ for some time nodes $k \in \mathbb{K}^{\text{tr}}$ and $M = 121$. Left: $\epsilon_{\text{POD}} = 10^{-2}$ ($N = 3$). Middle: $\epsilon_{\text{POD}} = 10^{-4}$ ($N = 7$). Right: $\epsilon_{\text{POD}} = 5.10^{-6}$ ($N = 11$).

linear case, the error always decreases for smaller tolerances ϵ_{POD}, i.e., with the dimension N of the background space $\mathcal{Z}_N$. However, we expect that, for some very small tolerance value (e.g. ϵ_{POD} such that $N > M$), the stability issues mentioned above would arise again.

Test case (d): Nonlinear imperfect model. This test case investigates a nonlinear imperfect `bk` model for which the HF `bk` solutions and the true solutions are respectively displayed in the left and right panels of Figure 5.8. The temperature profile for the true solution over the bi-material plate at the end of the simulation, i.e., at $t^K = 10$s clearly shows a different behavior at the boundaries of the inner material. Regarding the PBDW state estimation, Figure 5.10 shows the relative H^1-error $e^k(\mu)$ defined in (5.4) using $M = 121$ observations to build the observable space $\mathcal{U}_M$. For $\epsilon_{\text{POD}} = 10^{-4}$, $\mathcal{Z}_N$ is spanned by $N = 7$ vectors. Notice that the error vanishes for $\mu = 0.25$ since this configuration is equivalent to a perfect `bk` model. We notice that the relative H^1-error $e^k(\mu)$ increases because the stability constant decreases. Figure 5.11 visualizes the relative H^1-error $e^k(\mu)$ for a higher number of observations $M = 676$. We observe that augmenting the dimension of the observable space $\mathcal{U}_M$ cures the stability issues. Also, the errors are lower owing to the higher number of observations. Finally,

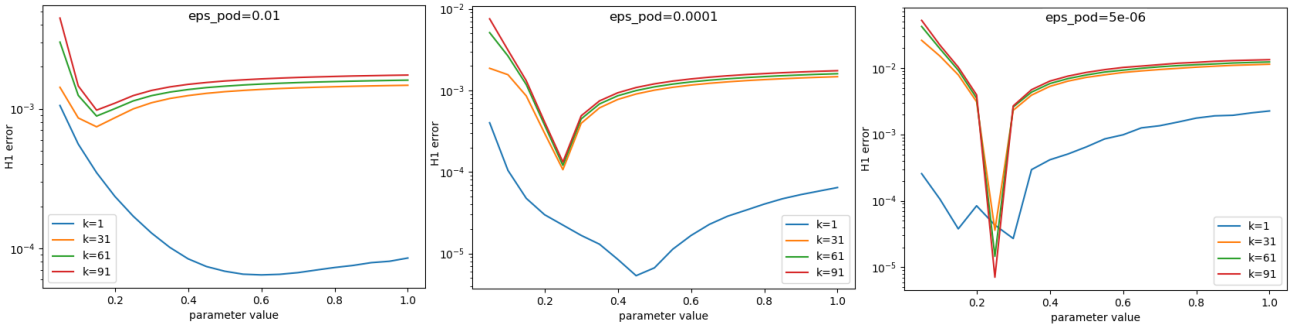

Figure 5.10. Test case (d) : Relative H^1-error $e^k(\mu)$ for some time nodes $k \in \mathbb{K}^{\mathrm{tr}}$ and $M = 121$. Left: $\epsilon_{\mathrm{POD}} = 10^{-2}$ $(N = 3)$. Middle: $\epsilon_{\mathrm{POD}} = 10^{-4}$ $(N = 7)$. Right: $\epsilon_{\mathrm{POD}} = 10^{-6}$ $(N = 11)$.

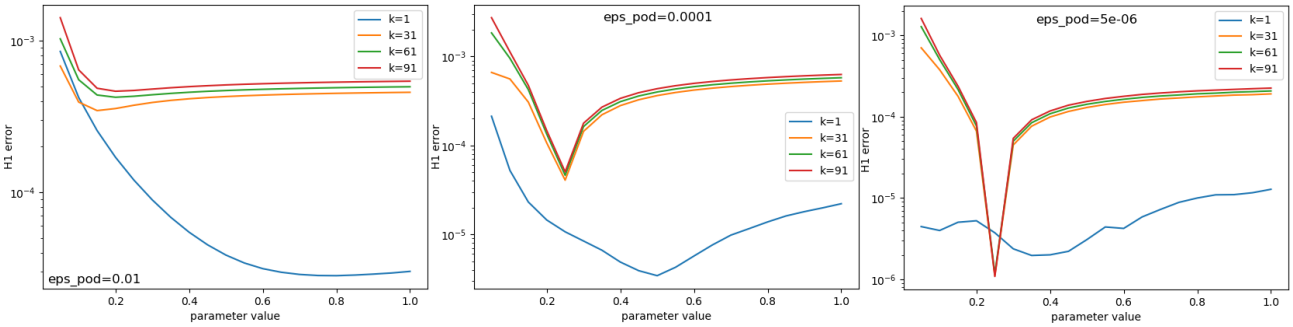

Figure 5.11. Test case (d) : Relative H^1-error $e^k(\mu)$ for some time nodes $k \in \mathbb{K}^{\mathrm{tr}}$ and $M = 676$. Left: $\epsilon_{\mathrm{POD}} = 10^{-2}$ $(N = 3)$. Middle: $\epsilon_{\mathrm{POD}} = 10^{-4}$ $(N = 7)$. Right: $\epsilon_{\mathrm{POD}} = 5.10^{-6}$ $(N = 11)$.

Figure 5.12 shows the stability constant $\beta_{N,M}$ as a function of the number of observations M. The behavior is quite similar to the linear case. Hence, the nonlinear character of the problem does not influence the overall features of the PBDW statement. This observation corroborates the independence with regard to the `bk` model.

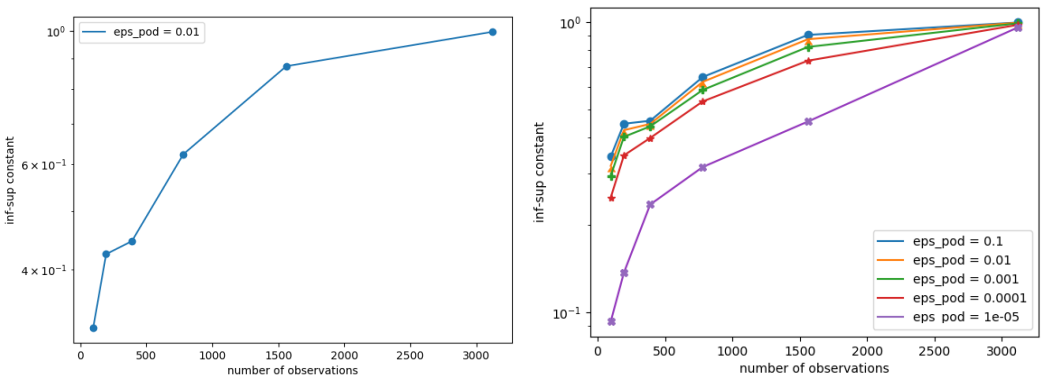

Figure 5.12. Test case (d) : Stability constant $\beta_{N,M}$. On the right panel, the values of N are respectively 2, 3, 5, 7, 11 for the values of ϵ_{POD} in decreasing order.

5.4. **Background space construction via state estimation**

We now illustrate the performances of Algorithm 2 for the following linear imperfect case:

- Test case (e): We consider a simulation duration $T = 4$s and a time step $\Delta t = 0.1$s. Test truths are synthesized with an internal diffusivity $D_{\text{int}} = 0.2$.

As opposed to the previous section, we choose a non-parametric `bk` model based on an HF computation for $\mu = 0.5$. The resulting unique trajectory is then reduced using a `POD` algorithm, which is equivalent to a `POD-greedy` for a single trajectory (cf. line 1 of Algorithm 2). For a tolerance value $\epsilon_{\text{POD}}^{\text{init}} = 10^{-2}$, we obtain a background space $\mathcal{Z}_{N^{\text{init}}}^{\text{init}}$ composed of $N^{\text{init}} = 34$ modes. As regards observations, the initial set $\mathcal{Q}^{\text{init}}$ consists of $M^{\text{init}} = \text{Card}(\mathcal{Q}^{\text{init}}) = 1521$ sensors that are uniformly placed over the plate (cf. line 2 of Algorithm 2). Using both the background space $\mathcal{Z}_{N^{\text{init}}}^{\text{init}}$ and the observable space $\mathcal{U}_{M^{\text{init}}}^{\text{init}}(\mathcal{Q}^{\text{init}})$, we estimate the state $u^*_{N^{\text{init}},M^{\text{init}}}$ for every parameter μ in the training set $\mathcal{P}^{\text{tr}} = \{0, 4, 8, 12, 16\}$ (cf. line 4 of Algorithm 2). The state estimation leads to the relative H^1-error $e^k(\mu)$ shown in Figure 5.13. We also plot in Figure 5.14 the absolute H^1-norms of the deduced background estimate $z^*_{N^{\text{init}},M^{\text{init}}}$

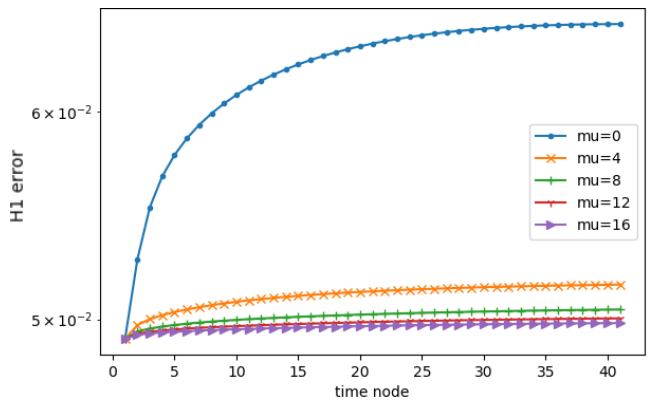

Figure 5.13. Test case (e) : Relative H^1-error $e^k(\mu)$ for the state estimate as a function of the time nodes. The curves correspond to different values of μ.

and the update estimate $\eta^*_{N^{\text{init}},M^{\text{init}}}$. One can notice that the latter is non-negligible compared to the former. Once the first part of the modified offline stage has been performed, we use the resulting state

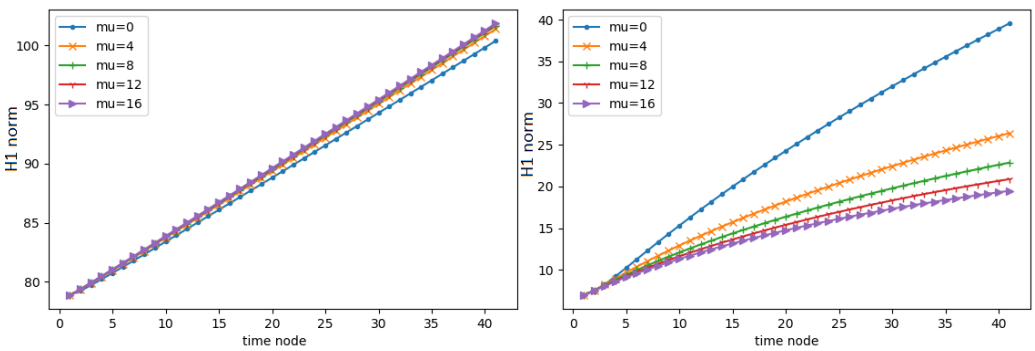

Figure 5.14. Test case (e) : Absolute H^1-norms of the contributions $z^*_{N^{\text{init}},M^{\text{init}}}$ and $\eta^*_{N^{\text{init}},M^{\text{init}}}$ as a function of the time nodes. The various curves correspond to the different values of μ.

estimates in order to build the modified background space (cf. line 5 of Algorithm 2). For a tolerance $\epsilon_{\text{POD}} = 5.10^{-2}$, the `POD-greedy` algorithm selects four modes. Then, we build the observable space $\mathcal{U}_M$ using $M = 121$ uniformly distributed sensors (the optimal choice can be made using the `S-Greedy` algorithm, see Algorithm 3). Figure 5.15 displays the errors for the verification set $\mathcal{P}^{\text{verif}} = \{0, \ldots, 19\}$. The state estimation relative H^1-error $e^k(\mu)$ remains comparable to that of the five parameters used for the offline construction. Regarding the online observations, we highlight that the online results are achieved using only $M \approx 8\% M^{\text{init}}$. Finally, Figure 5.16 shows the absolute H^1-norms of the deduced

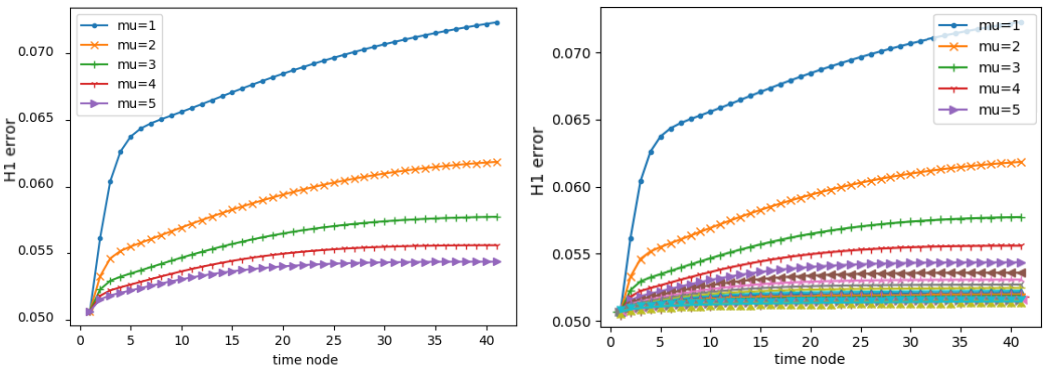

Figure 5.15. Test case (e) : Relative H^1-error $e^k(\mu)$ for the state estimate as a function of the time nodes during the online stage. The various curves correspond to the different values of μ. Left: for all $\mu \in \mathcal{P}^{\text{tr}}$. Right: for all $\mu \in \mathcal{P}^{\text{verif}}$.

background estimate $z^*_{N,M}$ and the update estimate $\eta^*_{N,M}$. We observe that the update estimate $\eta^*_{N,M}$ has a lower norm compared to Figure 5.14, whereas the deduced background estimate $z^*_{N,M}$ has a larger norm. This is due to the offline inclusion of observations in the new background space $\mathcal{Z}_N$ through offline state estimation. Therefore, we deduce that the modified offline algorithm achieves the expected objective.

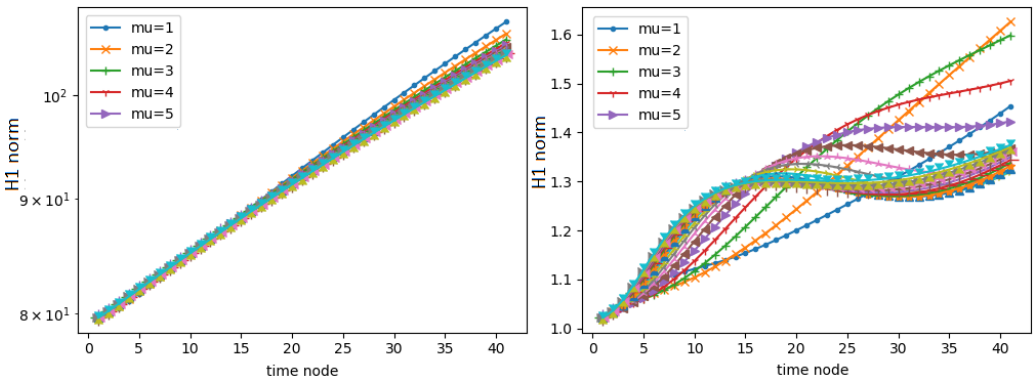

Figure 5.16. Test case (e) : Absolute H^1-norms of the contributions $z^*_{N,M}$ and $\eta^*_{N,M}$ as a function of the time nodes. The various curves correspond to the different values of μ.

6. Conclusion and perspectives

We have presented a time-dependent extension of the PBDW approach with a modified offline stage that suits both the steady and time-dependent cases. As regards the time-dependent setting, numerical tests on both linear and nonlinear cases assess the efficiency of the method for well chosen dimensions of the `bk` space $\mathcal{Z}_N$ and of the observable space $\mathcal{U}_M$. The test cases show that augmenting the dimension N of $\mathcal{Z}_N$ is counter-productive starting from a certain rank due of the deterioration of the stability constant $\beta_{N,M}$. In such cases, increasing the dimension of the observable space $\mathcal{U}_M$ is an alternative that restores a good stability of the problem. However, the measurements are expensive to obtain in engineering scenarios. In this context, a modified offline stage is introduced so as to reduce the number of final observations that are required within the online stage. The numerical performances obtained produce accuracy levels that are comparable to the standard method. A promising application is the assessment of the proposed methodology for other types of time-dependent problems as well as for three-dimensional industrial cases. Another interesting research direction is the inclusion of noise in the time-dependent framework.

Acknowledgements

This work is partially supported by Electricité De France (EDF) and a CIFRE PhD fellowship from ANRT. The author is thankful to A. T. Patera, A. Ern and V. Ehrlacher for proof-reading and stimulating discussions.

References

[1] J.-P. Argaud, B. Bouriquet, F. de Caso, H. Gong, Y. Maday, and O. Mula. Sensor placement in nuclear reactors based on the generalized empirical interpolation method. *J. Comput. Phys.*, 363:354–370, 2018.

[2] J.-P. Argaud, B. Bouriquet, H. Gong, Y. Maday, and O. Mula. Stabilization of (G)EIM in presence of measurement noise: application to nuclear reactor physics. In *Spectral and high order methods for partial differential equations—ICOSAHOM 2016*, volume 119 of *Lect. Notes Comput. Sci. Eng.*, pages 133–145. Springer, 2017.

[3] A. Benaceur. *Model reduction for nonlinear problems in thermics and mechanics.* Thesis, University Paris-Est Marne la Vallée, 2018.

[4] A. Benaceur, V. Ehrlacher, A. Ern, and S. Meunier. A Progressive Reduced Basis/Empirical Interpolation Method for Nonlinear Parabolic Problems. *SIAM J. Sci. Comput.*, 40(5):A2930–A2955, 2018.

[5] P. Binev, A. Cohen, O. Mula, and J. Nichols. Greedy Algorithms for Optimal Measurements Selection in State Estimation Using Reduced Models. *SIAM/ASA J. Uncertain. Quantif.*, 6(3):1101–1126, 2018.

[6] A. Cohen, W. Dahmen, R. DeVore, J. Fadili, O. Mula, and J. Nichols. Optimal reduced model algorithms for data-based state estimation. `https://arxiv.org/abs/1903.07938`, 2019.

[7] A. Cohen, M. A. Davenport, and D. Leviatan. On the stability and accuracy of least squares approximations. *Found. Comput. Math.*, 13(5):819–834, 2013.

[8] A. Cohen, M. A. Davenport, and D. Leviatan. Correction to: On the stability and accuracy of least squares approximations [MR3105946]. *Found. Comput. Math.*, 19(1):239, 2019.

[9] B. De Moor. Structured total least squares and L_2 approximation problems. *Linear Algebra Appl.*, 188/189:163–205, 1993.

[10] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer, 2004.

[11] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer, 2004.

[12] R. Everson and L. Sirovich. Karhunen–Loeve procedure for gappy data. *JOSA A*, 12(8):1657–1664, 1995.

[13] G. Franceschini and S. Macchietto. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19):4846–4872, 2008. Model-Based Experimental Analysis.

[14] F. Galarce, J.-F. Gerbeau, D. Lombardi, and O. Mula. State estimation with nonlinear reduced models. Application to the reconstruction of blood flows with Doppler ultrasound images. `https://arxiv.org/abs/1904.13367`, 2019.

[15] H. Gong, Y. Maday, O. Mula, and T. Taddei. Pbdw method for state estimation: error analysis for noisy data and nonlinear formulation. `https://arxiv.org/abs/1906.00810`, 2019.

[16] B. Haasdonk. Convergence rates of the POD-greedy method. *ESAIM, Math. Model. Numer. Anal.*, 47(3):859–873, 2013.

[17] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *M2AN, Math. Model. Numer. Anal.*, 42(2):277–302, 2008.

[18] B. Haasdonk, J. Salomon, and B. Wohlmuth. A Reduced Basis Method for the Simulation of American Options. *Numerical Mathematics and Advanced Applications*, 2012.

[19] J. K. Hammond, R. Chakir, F. Bourquin, and Y. Maday. PBDW: A non-intrusive Reduced Basis Data Assimilation method and its application to an urban dispersion modeling framework. *Appl. Math. Model.*, 76:1–25, 2019.

[20] F. Hecht. New developments in FreeFem++. Open source on `http://www.freefem.org`, 2012.

[21] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations*. SpringerBriefs in Mathematics. Springer; BCAM Basque Center for Applied Mathematics, Bilbao, 2016.

[22] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME Ser. D. J. Basic Engrg.*, 82(1):35–45, 1960.

[23] Mark Kärcher, Sébastien Boyaval, Martin A. Grepl, and Karen Veroy. Reduced basis approximation and a posteriori error bounds for 4D-Var data assimilation. *Optim. Eng.*, 19(3):663–695, 2018.

[24] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.*, 90(1):117–148, 2001.

[25] Q. Li, Y. Wang, and M. Vasilyeva. Multiscale model reduction for fluid infiltration simulation through dual-continuum porous media with localized uncertainties. *J. Comput. Appl. Math.*, 336:127–146, 2018.

[26] A. C. Lorenc. A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review*, 109(4):701–721, 1981.

[27] Y. Maday and O. Mula. A generalized empirical interpolation method: application of reduced basis techniques to data assimilation. In *Analysis and numerics of partial differential equations*, volume 4 of *Springer INdAM Series*, pages 221–235. Springer, 2013.

[28] Y. Maday, O. Mula, A. T. Patera, and M. Yano. The Generalized Empirical Interpolation Method: Stability theory on Hilbert spaces with an application to the Stokes equation. *Comput. Methods Appl. Mech. Eng.*, 287:310–334, 2015.

[29] Y. Maday, O. Mula, and G. Turinici. Convergence analysis of the generalized empirical interpolation method. *SIAM J. Numer. Anal.*, 54(3):1713–1731, 2016.

[30] Y. Maday, A. T. Patera, J. D. Penn, and M. Yano. A parameterized-background data-weak approach to variational data assimilation: formulation, analysis, and application to acoustics. *Int. J. Numer. Meth. Engng.*, 102(5):933–965, 2015.

[31] Y. Maday, A. T. Patera, J. D. Penn, and M. Yano. PBDW state estimation: noisy observations; configuration-adaptive background spaces; physical interpretations. In *CANUM 2014—42e Congrès National d'Analyse Numérique*, volume 50 of *ESAIM Proc. Surveys*, pages 144–168. EDP Sci., Les Ulis, 2015.

[32] Y. Maday and T. Taddei. Adaptive PBDW Approach to State Estimation: Noisy Observations; User-Defined Update Spaces. *SIAM J. Sci. Comput.*, 41(4):B669–B693, 2019.

[33] V. Mons, L. Margheri, J.-C. Chassaing, and P. Sagaut. Data assimilation-based reconstruction of urban pollutant release characteristics. *Journal of Wind Engineering and Industrial Aerodynamics*, 169:232–250, 2017.

[34] A. Nassiopoulos and F. Bourquin. Fast three-dimensional temperature reconstruction. *Comput. Methods Appl. Mech. Eng.*, 199(49-52):3169–3178, 2010.

[35] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations.* La Matematica per il 3+2. Springer, 2016.

[36] T. Taddei and A. T. Patera. A localization strategy for data assimilation; application to state estimation and parameter estimation. *SIAM J. Sci. Comput.*, 40(2):B611–B636, 2018.

[37] T. Taddei, J. D. Penn, and A. T. Patera. Validation by Monte Carlo sampling of experimental observation functionals. *Int. J. Numer. Meth. Engng.*, 112(13):2135–2150, 2017.

[38] T. Taddei, J. D. Penn, M. Yano, and A. T. Patera. Simulation-based classification; a model-order-reduction approach for structural health monitoring. *Arch. Comput. Methods Eng.*, 25(1):23–45, 2018.

[39] J. Waeytens, P. Chatellier, and F. Bourquin. Inverse computational fluid dynamics: Influence of discretization and model errors on flows in water network including junctions. *Journal of Fluids Engineering*, 137(9):091401, 2015.

[40] C. K. Wikle and L. M. Berliner. A Bayesian tutorial for data assimilation. *Physica D*, 230(1-2):1–16, 2007.

[41] K. Willcox. Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition. *Computers & Fluids*, 35(2):208–226, 2006.