State estimation with nonlinear reduced models. Application to the reconstruction of blood flows with Doppler ultrasound images.

F. Galarce, J.F. Gerbeau, D. Lombardi, O. Mula

April 30, 2019

Abstract

Over the past years, several fast reconstruction algorithms based on reduced models have been proposed to address the state estimation problem of approximating an unknown function u of a Hilbert space V from measurement observations. Most strategies are however based on linear mappings where the reduced model is built a priori and independently of the observation space and the measurements. In this work we explore some nonlinear extensions that take these elements into account in the construction of the basis. The methodology is applied to the reconstruction of 3D blood flows from Doppler ultrasound images. The example not only shows the good performance of the nonlinear methods, but it also illustrates the potential of the methodology in medicine as tool to process and interpret data in a systematic manner which could help to build more robust and individualized diagnostics.

1 Introduction

In the field of medicine, modern diagnostics, therapy optimisation and decisions are based on relevant *quantities of interest* (QoI) which are very often estimated from data produced by measurement devices. Nowadays, one of the most common measuring technique is medical imaging, which has the advantage of being *non-invasive* for the patients. The rapid progress in image processing over the last decades has led to an increasing quantity and quality of the available data and the ability to process, interpret and use them in a systematic way has become a key objective in order to build more robust and individualized diagnostics. However, methods that are solely relying on data present certain shortcomings since in some cases it is difficult to relate the measurements with the desired QoI. This is due to several reasons. First, the nature of the device may not allow to measure directly the desired QoI and may require a complex post-processing to obtain an estimate of it. Note however that the available data could be used differently to estimate other QoI that have so far never been considered by medical doctors simply because the amount of information was not available in the past. Second, there could be morphological constraints that prevent from measuring at specific locations. One last obstruction comes when the desired QoI is a prediction, a forecast ahead in time, requiring a specific treatment to infer it.

In this context, this paper is a contribution to connect in a systematic manner data with classical QoI and also to build other types of outputs that could be useful for more individualized diagnostics. We propose to use *reduced models* of parametrized PDEs to make this link in a fast and reliable manner. To illustrate our methodology, we concentrate on the task of reconstructing in real-time 3D fields for which we are only given partial information coming in the form of images. The particular application that we address is the rapid reconstruction of the full 3D blood flow velocity field in an artery from Doppler ultrasound images taken on a restricted portion of the artery. This application could be useful, for instance, as an additional tool during medical examinations of patients presenting risks of stenosis in parts that are out of reach for the imaging device due to morphological constraints. We illustrate that we can reconstruct not only the whole velocity at every point, but also that the method can produce novel QoI that could be of interest for diagnostics.

Mathematically, the reconstruction of the blood velocity field can be seen as a state estimation problem where we have to recover an unknown function u from a certain Hilbert space V from the knowledge of measurement observations. There exist numerous methods to address this general problem in the fields of data assimilation and inverse problems (see [1, 2, 3] for extensive surveys).

In our case, contrary to classical inverse problem approaches where one seeks for the parameters of the PDE that best satisfy the measurements, we use the PDE models to learn fast reconstruction mappings which satisfy certain optimality properties. The high dimensionality of the problems that arise require the use of compression strategies which, in our case, are based on the model reduction of parametrized PDEs. The problem of state estimation using reduced modelling has attracted considerable attention in the last decade, see [4, 5, 6, 7, 8, 9, 10, 11, 12]. Our starting point is the Parametrized Background Data-Weak originally introduced in [13]. The method consists in a linear mapping which is built from a least squares fit between the measurements and a given reduced space. Like in most of the other approaches, the reduced space in PBDW is chosen *a priori* and independently of the measurements. In this paper, we introduce and compare several possible ways of constructing data-driven reduced models that yield to nonlinear reconstruction mappings. We show for the example of the artery that our proposed nonlinear reconstructions perform better than the classical linear method in terms of the reconstruction error.

The paper is organized as follows, in section 2 we describe mathematically what we understand by state estimation problems and reconstruction algorithms (section 2.1). We discuss what is known about optimal algorithms and justify the use of PBDW as a suboptimal yet good reconstruction method (section 2.2). We next introduce several nonlinear versions of PBDW that take into account the measurements in the reduced basis construction (section 2.3). We do a simple error analysis in order to explain the main mechanisms justifying the superiority of the nonlinear methods. We note that one of the nonlinear algorithms has first been proposed in [14]. We next apply the methodology to the reconstruction of 3D blood flows from Doppler images (section 3). The example is synthetic in the sense that the Doppler images have been synthetically generated. We build our reduced models upon a parametric Navier-Stokes PDE which is explained in section 3.1. Section 3.2 gives some details on Doppler ultrasound imaging and the way in which we have incorporated them to our methodology. Finally, sections 3.3 and 3.4 present numerical results in two examples. The results illustrate the superiority of the nonlinear reconstruction algorithms. In addition to this, our second example also shows that the method can be used to estimate quantities of interest which could be helpful in blockage detection.

2 Reconstruction methods

2.1 State estimation and recovery algorithms

Let Ω be a domain of $\mathbb{R}^d$ for a given dimension $d \geq 1$ and let V be a Hilbert space defined over Ω, with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Our goal is to recover an unknown function $u \in V$ from m measurement observations

$$\ell_i(u), \quad i = 1, \ldots, m, \tag{2.1}$$

where the ℓ_i are linearly independent linear forms over V. Each ℓ_i models a sensor device which is used to collect the measurement data $\ell_i(u)$. The Riesz representers of the ℓ_i are denoted by ω_i and span an m-dimensional space

$$W_m = \mathrm{span}\{\omega_1, \ldots, \omega_m\} \subset V.$$

The observations $\ell_1(u), \ldots, \ell_m(u)$ are thus equivalent to knowing the orthogonal projection

$$\omega = P_{W_m} u. \tag{2.2}$$

In this setting, the task of recovering u from the measurement observation ω can be viewed as building a recovery algorithm

$$A : W_m \mapsto V$$

such that $A(P_{W_m} u)$ is a good approximation of u in the sense that $\|u - A(P_{W_m} u)\|$ is small.

Recovering u from the measurements $P_{W_m} u$ is a very ill-posed problem since there are infinitely many $v \in V$ such that $P_{W_m} v = w$. It is thus necessary to add some a priori information on u in order to recover u up to a guaranteed accuracy. We are motivated by the setting where u is a solution to some parameter-dependent PDE of the general form

$$\mathcal{P}(u, y) = 0,$$

where $\mathcal{P}$ is a differential operator and y is a vector of parameters that describes some physical property and lives in a given set $\mathrm{Y} \subset \mathbb{R}^p$. Therefore, our prior on u is that it belongs to the set

$$\mathcal{M} := \{u(y) \in V \ : \ y \in \mathrm{Y}\}, \tag{2.3}$$

which is sometimes referred to as the *solution manifold*. The performance of a recovery mapping A is usually quantified in two ways:

- If the sole prior information is that u belongs to the manifold $\mathcal{M}$, the performance is usually measured by the worst case reconstruction error

$$E_{\mathrm{wc}}(A, \mathcal{M}) = \sup_{u \in \mathcal{M}} \|u - A(P_{W_m} u)\|.$$

- In some cases u is described by a probability distribution p on V supported on $\mathcal{M}$. This distribution is itself induced by a probability distribution on Y that is assumed to be known. In this Bayesian-type setting, the performance is usually measured in an average sense through the mean-square error

$$E_{\mathrm{ms}}^2(A, \mathcal{M}) = \mathbb{E}\left(\|u - A(P_{W_m} u)\|^2\right) = \int_V \|u - A(P_{W_m} u)\|^2 dp(u),$$

and it naturally follows that $E_{\mathrm{ms}}(A, \mathcal{M}) \leq E_{\mathrm{wc}}(A, \mathcal{M})$.

2.2 Optimal reconstruction algorithms

In general, one would like to use an algorithm A that is optimal in the sense of minimizing

$$\inf_{A:W_m \to V} E_{\mathrm{wc}}(A, \mathcal{M}), \quad \text{or} \quad \inf_{A:W_m \to V} E_{\mathrm{ms}}(A, \mathcal{M}).$$

However, as discussed in [15], optimal algorithms are difficult to compute and even to characterize for general sets $\mathcal{M}$. In this respect, the following is known (see [15] for further details):

- The problem of finding an algorithm A that minimizes $E_{\mathrm{wc}}(A, \mathcal{M})$ is called *optimal recovery*. It has been extensively studied for convex sets $\mathcal{M}$ that are balls of smoothness classes. This is however not the case in the current setting since the solution manifold $\mathcal{M}$ introduced in (2.3) usually has a complex geometry. We know that as soon as $\mathcal{M}$ is bounded there is a simple mathematical description of an optimal algorithm in terms of Chebyshev centers of certain sets. However, this algorithm cannot be easily computed due to the geometry and high dimensionality of the manifold.

- The problem of finding an algorithm A that minimizes $E_{\mathrm{ms}}(A, \mathcal{M})$ falls into the scope of *bayesian* or *learning problems*. As explained in [15], if the probability distribution on the manifold $\mathcal{M}$ is Gaussian, the optimal algorithm can easily be characterized and computed. However, the assumption on a Gaussian distribution is very strong and will not hold in general so finding a computable optimal algorithm in the mean-square sense is also an open problem.

These theoretical difficulties motivate the search for suboptimal yet fast and good recovery algorithms. One vehicle for this has been to build linear recovery algorithms $A \in \mathcal{L}(W_m, V)$ using *reduced modeling techniques*. However, since in general it is not clear that linear algorithms will be optimal, it is interesting to explore the advantage of reconstructing with computationally affordable nonlinear algorithms. In this work, we explore the performance of two different types of nonlinear algorithms which we introduce next. Unfortunately, they do not come with any theoretical guarantee of its superiority with respect to classical linear methods so we compare their performance with respect to classical linear algorithms in a non trivial example: the reconstruction of the blood velocity in an artery from a Doppler ultrasound image.

Remark 2.1. *Note that in the present setting the measurement space W_m is fixed and we will adhere to this assumption in the rest of the paper. This is reasonable since usually the nature and location of the sensors is fixed according to physical and financial constraints and also following the experience of the experts. A different, yet related problem, is to optimize the choice of the measurement space W_m. Two works on this topic involving greedy algorithms are [16, 17].*

2.3 Linear and nonlinear algorithms using reduced modeling

Reduced models are a family of methods that produce each a hierarchy of spaces $(V_n)_{n \geq 1}$ that approximate the solution manifold well in the sense that

$$\varepsilon_n := \sup_{u \in \mathcal{M}} \mathrm{dist}(u, V_n), \quad \text{or} \quad \delta_n^2 := \mathbb{E}\left(\mathrm{dist}(u, V_n)^2\right)$$

decays rapidly as n grows for certain classes of PDEs. Several methods exist to build these spaces among which stand the reduced basis method (see [18]), the (Generalized) Empirical Interpolation Method (see [19, 8, 16]), Principal Component Analysis (PCA) and low-rank methods (see [20, 21]).

3

Linear reconstruction algorithms that make use of reduced spaces V_n are the Generalized Empirical Interpolation Method (GEIM) introduced in [8] and further analyzed in [9, 16] and the Parametrized Background Data-Weak Approach (PBDW) introduced in [13] and further analyzed in [22]. Note that some modified versions have been proposed to address measurement noise (see, e.g., [11, 23]) and other recovery algorithms involving reduced modelling have also been recently proposed (see [12]).

Our nonlinear algorithms are built upon PBDW so we will first recall it before introducing the nonlinear extensions. We present a simple error analysis which helps to understand the main mechanisms driving the reconstruction quality.

2.3.1   PBDW, a linear recovery algorithm

Given a measurement space W_m and a reduced model V_n with $1 \leq n \leq m$, the PBDW algorithm

$$A_{m,n}^{(\mathrm{pbdw})} : W_m \to V$$

gives for any $\omega \in W_m$ a solution of

$$\min_{u=w+W^\perp} \mathrm{dist}(u, V_n).$$

Denoting

$$\beta(X, Y) := \inf_{x \in X} \sup_{y \in Y} \frac{\langle x, y \rangle}{\|x\| \, \|y\|} = \inf_{x \in X} \frac{\|P_Y x\|}{\|x\|} \in [0, 1] \tag{2.4}$$

for any pair of closed subspaces (X, Y) of V, the above optimization problem has a unique minimizer

$$A_{m,n}^{(\mathrm{pbdw})}(\omega) = u_{m,n}^*(w) := \arg\min_{u=w+W^\perp} \mathrm{dist}(u, V_n). \tag{2.5}$$

as soon as $n \leq m$ and $\beta(V_n, W_m) > 0$. We adhere to these two assumptions in the following.

As proven in appendix A, an explicit expression of $u_{m,n}^*(w)$ is

$$u_{m,n}^*(w) = v_{m,n}^*(w) + w - P_{W_m} v_{m,n}^*(w) \tag{2.6}$$

with

$$v_{m,n}^*(w) = \left(P_{V_n|W_m} P_{W_m|V_n} \right)^{-1} P_{V_n|W_m}(w), \tag{2.7}$$

where, for any pair of closed subspaces (X, Y) of V, $P_{X|Y} : Y \to X$ is the orthogonal projection into X restricted to Y. The invertibility of the operator $P_{V_n|W_m} P_{W_m|V_n}$ is guaranteed under the above conditions.

Formula (2.6) shows that $A_n^{(\mathrm{pbdw})}$ is a bounded linear map from W_m to $V_n \oplus (W_m \cap V_n^\perp)$. Depending on whether V_n is built to address the worst case or mean square error, the reconstruction performance is bounded by

$$e_{m,n}^{(\mathrm{wc, \, pbdw})} = E_{\mathrm{wc}}(A_{m,n}^{(\mathrm{pbdw})}, \mathcal{M}) \leq \beta^{-1}(V_n, W_m)\, \varepsilon_n, \tag{2.8}$$

or

$$e_{m,n}^{(\mathrm{ms, \, pbdw})} = E_{\mathrm{ms}}(A_{m,n}^{(\mathrm{pbdw})}, \mathcal{M}) := \mathbb{E}\left(\|u - A_n^{(PBDW)}(P_{W_m} u)\|^2 \right)^{1/2} \leq \beta^{-1}(V_n, W_m)\, \delta_n, \tag{2.9}$$

Note that $\beta(V_n, W_m)$ can be understood as a stability constant. It can also be interpreted as the cosine of the angle between V_n and W_m.

An important observation is that for a fixed measurement space W_m, the errors $e_{m,n}^{(\mathrm{wc, \, pbdw})}$ and $e_{m,n}^{(\mathrm{ms, \, pbdw})}$ reach a minimal value $e_{m,n_{\mathrm{wc}}^*}^{(\mathrm{wc, \, pbdw})}$ and $e_{m,n_{\mathrm{ms}}^*}^{(\mathrm{ms, \, pbdw})}$ as the dimension n varies from 1 to m. This behavior is due to the trade-off between the increase of the approximation properties of V_n as n grows and the degradation of the stability of the algorithm, given here by the decrease of $\beta(V_n, W_m)$ to 0 as $n \to m$. As a result, the best reconstruction performance with PBDW is given by

$$e_{m,n_{\mathrm{wc}}^*}^{(\mathrm{wc, \, pbdw})} = \min_{1 \leq n \leq m} e_{m,n}^{(\mathrm{wc, \, pbdw})}, \quad \mathrm{or} \quad e_{m,n_{\mathrm{ms}}^*}^{(\mathrm{ms, \, pbdw})} = \min_{1 \leq n \leq m} e_{m,n}^{(\mathrm{ms, \, pbdw})}.$$

2.3.2   Two nonlinear reconstruction algorithms

Taking PBDW as a starting point, we next describe two different strategies to build nonlinear algorithms. They all incorporate an affine extension of PBDW that we explain next.

Affine PBDW: A first simple step towards building a nonlinear algorithm is to formulate an affine version of PBDW where, given an average or nominal state $\bar{u} \in V$, we reconstruct with its equivalent affine version, which reads

$$A_{m,n}^{(\mathrm{aff})}(w) := \arg\min_{u \, \in \, w+W^\perp} \mathrm{dist}(u, \bar{u} + V_n). \tag{2.10}$$

4

Proceeding similarly as before, for a given $u \in \mathcal{M}$, the error is bounded by (see [15])

$$\|u - A_{m,n}^{(\text{aff})}(w)\| \le \beta^{-1}(V_n, W) \operatorname{dist}(u, \bar{u} + V_n) \tag{2.11}$$

and

$$e_{m,n}^{(\text{wc, aff})} := E_{\text{wc}}(A_{m,n}^{(\text{aff})}, \mathcal{M}) \le \beta^{-1}(V_n, W_m) \varepsilon_n^{(\text{aff})}, \tag{2.12}$$

or

$$e_{m,n}^{(\text{ms, pbdw})} := E_{\text{ms}}(A_{m,n}^{(\text{aff})}, \mathcal{M}) := \mathbb{E}\left(\|u - A_{m,n}^{(\text{aff})}(P_{W_m}u)\|^2\right)^{1/2} \le \beta^{-1}(V_n, W_m) \delta_n^{(\text{aff})}, \tag{2.13}$$

where

$$\varepsilon_n^{(\text{aff})} := \sup_{u \in \mathcal{M}} \operatorname{dist}(u, \bar{u} + V_n), \quad \text{or} \quad (\delta_n^{(\text{aff})})^2 := \mathbb{E}\left(\operatorname{dist}(u, \bar{u} + V_n)^2\right)$$

In the following, to simplify notation, we will use ε_n and δ_n to denote either the error in the linear PBDW or its affine version since the reasoning and the estimates that will be derived next have the same form.

Partition of $\mathcal{M}$: Our first strategy stems from the fact that the approximation errors $(\varepsilon_n)_n$ or $(\delta_n)_n$ provided by reduced basis may not always decrease rapidly to zero (the Kolmogorov n-width of $\mathcal{M}$ may decrease slowly). However, the physical structure of the problem could give a natural decomposition of the manifold $\mathcal{M}$ into different subdomains $\mathcal{M}^{(k)}$ that are better adapted for model reduction in the sense that the errors $(\varepsilon_n^{(k)})_n$ or $(\delta_n^{(k)})_n$ may decrease faster. This happens for example if we work in an application for which is possible to know exactly a subset of entries in y for any target $u(y)$, say $\bar{y} \in \mathbb{R}^{\bar{p}}$ (with $\bar{p} < p$). Given $\bar{y}$ and a data-base generated from the governing PDE $P(u, y)$ we may produce a disjoint union of K subsets $Y^{(k)}$ which yields a decomposition of $\mathcal{M}$ into subsets $\mathcal{M}^{(k)} = u(Y^{(k)})$. We can thus build reduced models $(V_n^{(k)})_{n=1}$ for each subset $\mathcal{M}^{(k)}$ and then reconstruct with the linear or affine PBDW. Proceeding similarly as in the previous section, the reconstruction performance on subset $\mathcal{M}^{(k)}$ is

$$e_{m,n}^{(\text{wc, pbdw, k})} = E_{\text{wc}}(A_{m,n}^{(\text{pbdw})}, \mathcal{M}^{(k)}) \le \beta^{-1}(V_n^{(k)}, W_m) \varepsilon_n^{(k)},$$

or

$$e_{m,n}^{(\text{ms, pbdw, k})} = E_{\text{ms}}(A_{m,n}^{(\text{pbdw})}, \mathcal{M}^{(k)}) := \mathbb{E}\left(\|u - A_{m,n}^{(PBDW)}(P_{W_m}u)\|^2\right)^{1/2} \le \beta^{-1}(V_n^{(k)}, W_m) \delta_n^{(k)}.$$

The best reconstruction performance for $\mathcal{M}^{(k)}$ is thus

$$e_{m,n_{\text{wc}}^*(k)}^{(\text{wc, pbdw, k})} = \min_{1 \le n \le m} e_{m,n}^{(\text{wc, pbdw, k})}, \quad \text{or} \quad e_{m,n_{\text{ms}}^*(k)}^{(\text{ms, pbdw, k})} = \min_{1 \le n \le m} e_{m,n}^{(\text{ms, pbdw, k})}.$$

It follows that the performance in $\mathcal{M} = \cup_{k=1}^K \mathcal{M}^{(k)}$ is

$$e_m^{(\text{wc, pbdw})} = \max_{1 \le n \le m} e_{m,n_{\text{ms}}^*(k)}^{(\text{ms, pbdw, k})}, \quad \text{or} \quad e_{m,n_{\text{ms}}^*(k)}^{(\text{ms, pbdw, k})} = \sum_{k=1}^K \omega_k e_{m,n_{\text{ms}}^*(k)}^{(\text{ms, pbdw, k})},$$

where $\omega_k = p(u \in \mathcal{M}^{(k)})$.

Data-based reduced models: The second strategy, originally introduced in [14], is motivated by the observation that the reduced models V_n or $V_n^{(k)}$ of the previous approaches are built independently of the given measurement space W_m. This may lead to stability issues that will manifest via very small values of $\beta(V_n, W_m) > 0$ for any dimension n. As a result, even if V_n might have very good approximation properties (ε_n or δ_n — resp. $\varepsilon_n^{(k)}$ or $\delta_n^{(k)}$— decrease rapidly with n), the reconstruction performance $e_{m,n_{\text{wc}}^*}^{(\text{wc, pbdw})}$ or $e_{m,n_{\text{ms}}^*}^{(\text{ms, pbdw})}$ might be poor because $\beta(V_n, W_m)$ may decrease very fast to 0 as $n \to m$. This motivates to adapt the construction of V_n to the given measurement space W_m or, even more, to adapt V_n to the given measurement observation $\omega \in W_m$. In our numerical tests, we follow the approach of [14] and, for a given $\omega \in W_m$, we build a data-driven $V_n(\omega)$ with a greedy algorithm that we explain next. Once this reduced model has been computed, we reconstruct with the data-driven affine version of PBDW,

$$A_{m,n}^{(\text{dd})}(w) := \operatorname*{arg\,min}_{u \in w + W^\perp} \operatorname{dist}(u, \bar{u} + V_n(w)), \tag{2.14}$$

where the difference with respect to (2.10) is that now V_n depends on ω. The reconstruction performance of this algorithm is bounded by

$$e_{m,n}^{(\text{wc, dd})} = E_{\text{wc}}(A_{m,n}^{(\text{dd})}, \mathcal{M}) \le \sup_{u \in \mathcal{M}} \beta^{-1}(V_n(P_{W_m}u), W) \operatorname{dist}(u, \bar{u} + V_n(P_{W_m}u)),$$

5

in the worst case setting. Similarly as before, $e_{m,n}^{(\mathrm{wc,\ dd})}$ reaches a minimum $e_{m,n_{(\mathrm{dd})}^*}^{(\mathrm{wc,\ dd})}$ when n varies from 1 to m. Since V_n is now adapted to the measurement observations, we expect that the current algorithm performs better than its classical linear counterpart.

We now recall the OMP greedy algorithm from [14] that we use. Let

$$\mathcal{D} := \{v = u/\|u\| \ : \ u \in \mathcal{M}\}$$

be the set of normalized functions from $\mathcal{M}$. If $\bar{u} = 0$, the first element φ_1 is chosen as

$$\varphi_1 \in \arg\max_{v \in \mathcal{D}} |\langle w, v \rangle| \tag{2.15}$$

For $n > 1$, given $V_n = \mathrm{span}\{\varphi_1, \ldots, \varphi_n\}$, we select

$$\varphi_{n+1} \in \arg\max_{v \in \mathcal{D}} \left| \left\langle w - P_{P_{W_m} V_n} w, \frac{P_{W_m} v}{\|P_{W_m} v\|} \right\rangle \right| \tag{2.16}$$

where $P_{W_m} V_n = \mathrm{span}\{P_{W_m}\varphi_1, \ldots, P_{W_m}\varphi_n\}$. We set $V_{n+1} = \mathrm{span}\{V_n, \varphi_{n+1}\}$.

Note that all operations in this algorithm are done in the space W_m. Hence we can do all calculations in $\mathbb{R}^m$, which makes this algorithm be very fast since it does not involve computations with functions from the whole space V.

In the case $\bar{u} \neq 0$, we introduce $\bar{w} = P_{W_m}\bar{u}$ and the shifted set

$$\delta_{\bar{u}}\mathcal{D} = \left\{v = \frac{u - \bar{u}}{\|u - \bar{u}\|} \ : \ u \in \mathcal{M}\right\}.$$

Now it suffices to apply the previous greedy algorithm to the target function $w - \bar{w}$ instead of w and do the search over $\delta_{\bar{u}}\mathcal{D}$ instead of $\mathcal{D}$.

3 Application: Reconstruction of 3D blood velocity fields from Doppler ultrasound images

We apply the above described methodology to reconstruct a 3D blood velocity field on a human carotid artery from Doppler ultrasound images. The images are synthetically generated and the use of data from real patients is differed to a future work. The main goal of the tests is twofold:

1. First and foremost, we aim to show that our methods give a good reconstruction accuracy in the 3D field and relevant quantities of interest and have thus potential to be used in diagnoses. Thanks to the use of reduced-models, the computing time for the reconstruction is very short and can be done in real-time.

2. Second, we aim to compare in a non trivial application the reconstruction performance between the classical PBDW, the nonlinear method involving a partition in $\mathcal{M}$ and the data-driven affine PBDW. As explained below, the problem is time-dependent and has a very specific temporal behavior dictated by the systole and the diastole of the cardiac cycle. This, together with other factors that we introduce next, yields a natural partition of $\mathcal{M} = \cup_{k=1}^K \mathcal{M}^{(k)}$ which we exploit to build reduced spaces $V_n^{(k)}$ for each $\mathcal{M}^{(k)}$.

The section is organized are follows. First, we present the parameter-dependent model that will define the manifold $\mathcal{M}$ on which we will rely to compute different reduced models. Second, we explain how to define a measurement space W_m from a Doppler velocity image. Finally, we present results on the comparison of the different methods based on their reconstruction performance.

3.1 The model: incompressible Navier-Stokes equations

Let Ω be a spatial bounded domain of $\mathbb{R}^3$ with the shape of a human carotid artery as given in Figure 1. The boundary $\Gamma := \partial\Omega$ is the union of the inlet part Γ_i where the blood in entering the domain, the outlets $\Gamma_{o,1}$ and $\Gamma_{o,2}$ where the blood is exiting the domain after a bifurcation, and the walls Γ_w.

We consider the following incompressible Navier-Stokes equations (NSE) on Ω and over the time interval $[0, T]$ for $T > 0$. For a fluid with density $\rho \in \mathbb{R}^+$ and dynamic viscosity $\mu \in \mathbb{R}^+$, we search for all $t \in [0, T]$ the couple $(u(t), p(t)) \in [H^1(\Omega)]^3 \times H^1(\Omega)$ of velocity and pressure such that

$$\begin{cases} \rho\dfrac{\partial u}{\partial t}(t) + \rho u(t)\nabla u(t) - \mu\Delta u(t) + \nabla p(t) = 0, & \text{in } \Omega \\ \\ \qquad\qquad\qquad\qquad\qquad\qquad \nabla \cdot u = 0, & \text{in } \Omega. \end{cases} \tag{3.1}$$

These equations are closed by adding a zero initial condition and the following boundary conditions:
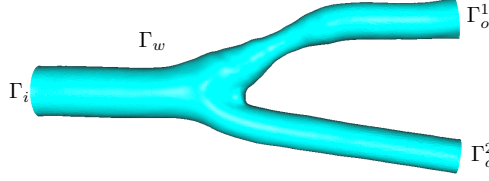
Figure 1: Domain Ω used in the simulations. Note the small stenosis in the upper part of the bifurcation.
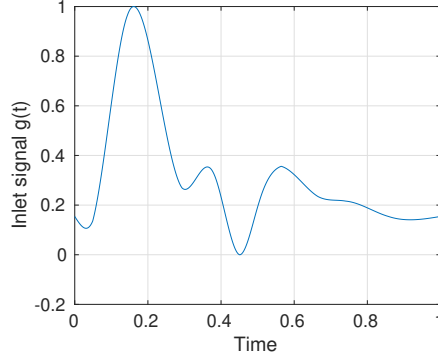

Figure 2: The function $g(t)$ for the inlet boundary condition.

- Boundary conditions:
    - No-slip condition for the vessel wall, that is, $u = (0,0,0)^T$ on Γ_w.
    - The inlet boundary Γ_i lies in the xz plane and we apply a Dirichlet condition for the velocity with a function $b(t, x, z) = u_0 \ g(t) f(x, z)$, where:
        * $u_0 \in \mathbb{R}^+$ is an scaling factor. The function $g(t)$ is built by interpolating experimental flow data in the common carotid area taken in [24]. Its behavior is given in Figure 2.
        * The function f is a 2D logit-normal distribution

        $$f(x) = \frac{1}{x(1-x)z(1-z)} \exp\left\{-0.5\left(\log\left(\frac{x}{1-x}\right) - s\right)^2 - 0.5\left(\log\left(\frac{z}{1-z}\right)\right)^2\right\},$$
        (3.2)

        where the parameter, $s \in \mathbb{R}^+$, controls the axial symmetry of the inlet flow.

    - For the outlet boundaries $\Gamma_{o,1}$ and $\Gamma_{o,2}$, we use the so-called Windkessel model (see [25]), which gives the average pressure over each $\Gamma_{o,k}$,

    $$\bar{p}_{o,k} = p_d^k + R_p^k \int_{\Gamma_{o,k}} u \cdot n, \quad k = 1, 2$$

    where $p_d^k \in \mathbb{R}$ is called *distal pressure* and is the solution to the ordinary differential equation:

    $$\begin{cases} C_d^k \dfrac{dp_d^k}{dt} + \dfrac{p_d^k}{R_d} &= \int_{\Gamma_{o,k}} u \cdot n \\ p_d^k(t=0) &= p_{d,k} \text{ given.} \end{cases}$$
    (3.3)

    This model aims to represent the cardiovascular system behavior beyond the boundaries of the working domain with a minimal increase in the computational cost. It is based on an analogy between flow and pressure with current and voltage in electricity. This is the reason why C_d^k is called distant capacitance and R_p^k and R_d^k are respectively called proximal and distant resistances. These three parameters are positive real numbers.

For the numerical solution of 3.1, we use the *FeLiSCe* library developed at Inria. We use a monolithic approach in which velocity and pressure are discretized in space with $\mathbb{P}_1 - \mathbb{P}_1$ Lagrange finite finite elements. Standard SUPG stabilization is used. Time is discretized with a semi-implicit

scheme with time-step $\delta t = 2 \cdot 10^{-3} s$. An explicit scheme is used to numerically solve the ODE on the distal pressure in the Windkessel model. In addition, a backflow stabilization is added in order to address potential instabilities in the outlet boundaries (see, e.g., [26]).

Remark 3.1. *Note that one could use more sophisticated models involving, for instance, fluid-structure interactions or more refined Winkessel models for the pressure. Our present model is a trade-off between its degree of realism and the difficulty and time to solve it. We refer to [25] for a detailed overview of cardiovascular modeling.*

Now that the model has been introduced, let us define the manifold of solutions that we consider in our numerical experiments. We set the following coefficients to a fixed value

$$
\begin{cases}
\rho & = 1 \text{ g/cm}^3 \\
\mu & = 0.03 \text{ Poise} \\
C_d^k & = 1.6 \times 10^{-5} \text{ for } k = 1, 2 \\
R_p^k & = 7501.5 \text{ for } k = 1, 2 \\
p_d^k & = 1.06 \times 10^5 \text{ for } k = 1, 2 \\
R_d^1 & = 60012
\end{cases}
\tag{3.4}
$$

We introduce the ratio of the distal resistances for the Windkessel model at the outlets of the geometry

$$
\eta := R_d^1 / R_d^2 = 6001.2 / R_d^2
$$

and we define the heart rate as

$$
\text{HR} := 60/T_c,
$$

where $T_c > 0$ is the cardiac cycle duration.

Our manifold $\mathcal{M}$ is generated by the variations of the six following parameters

$$
\begin{cases}
t & \in [0, T] \\
\text{HR} & \in [48, 120] \\
s & \in [0, 2] \\
T_{sis} & \in [0.2863, 0.3182] \text{ s.} \\
u_0 & \in [17, 20] \text{ cm/s} \\
\eta & \in [0.05, 0.2] \cup [0.5, 1.5] \cup [5, 20]
\end{cases}
\tag{3.5}
$$

Note that the time t is also seen as a parameter. The parameter set is thus

$$
\text{Y} = \{(t, HR, s, T_{sis}, u_0, \eta) \in \mathbb{R}^6 \,:\, t \in [0, T], \text{ HR} \in [48, 120], \, s \in [0, 0.2], \, \dots\} \subset \mathbb{R}^6
$$

and the manifold of solutions is

$$
\mathcal{M} := \{u(y) \in [H^1(\Omega)]^3 \,:\, y \in \text{Y}\}.
$$

At this point, several comments are in order:

- Note that we only consider velocities in the manifold because we are only interested in reconstructing the velocity field in this work. By proceeding analogously, we can consider the manifold of velocity and pressure pairs and reconstruct the pressure too. However, note that usually the pressure cannot be measured with non invasive methods in an artery so its reconstruction is a more involved task and it requires some specific developments which we will address in a forthcoming work.

- For each $y \in \text{Y}$, the velocity $u(y)$ is a function of $[H^1(\Omega)]^3$. In the following, we will view it as a function from

$$
V := [L^2(\Omega)]^3,
$$

which, endowed with the inner product,

$$
\langle (v_1, v_2, v_3), (w_1, w_2, w_3) \rangle := \sum_{i=1}^{3} \langle v_i, w_i \rangle_{L^2(\Omega)}, \quad \forall (v, w) \in [L^2(\Omega)]^3,
$$

defines a Hilbert space.

8

- Since the time variable has been included as a parameter, a simple way to build nonlinear reduced models is to set a window parameter $\tau > 0$ and consider the subset $\mathcal{M}^{(k)} = \mathcal{M}_{[t_k - \tau, t_k + \tau]} \subset \mathcal{M}$, where t is restricted to the interval $[t_k - \tau, t_k + \tau]$ of size 2τ centered around a given time t_k. We can then build reduced models to reconstruct this specific time interval. As we will see in the numerical experiments, this strategy is very effective in our problem because the velocity presents two regimes given by the systole and diastole periods.

- The computation of reduced models involves a discrete training subset $\widetilde{\mathcal{M}}_{\mathrm{train}} \subset \mathcal{M}$ which, in the experiments below, involves $\#\widetilde{\mathcal{M}}_{\mathrm{train}} = 78528$ snapshots $u(y)$. The parameters are chosen from a uniform random distribution and we only save the solutions during the second cardiac cycle of each simulation.

- For the purposes of illustrating the potential of the method for diagnoses, the notion of sickness will be understood in the following way.

  Definition 1 (Sick patient). We say that the output of the simulation corresponds to a healthy patient when $\eta \in [0.5, 1.5]$. Outside of this range, simulations correspond to sick patients.

  We thus have
  $$\mathcal{M} = \mathcal{M}_{\mathrm{healthy}} \cup \mathcal{M}_{\mathrm{sick}}$$
  with $\mathcal{M}_{\mathrm{healthy}} := \mathcal{M}_{\eta \in [0.5, 1.5]}$ and $\mathcal{M}_{\mathrm{sick}} = \mathcal{M}_{\eta \in [0.05, 0.2] \cup [5, 20]}$.

3.2 Measurements

At every time $t \in [0, T]$, we are given a Doppler ultrasound image that contains information on the blood velocity on a sub-domain of the carotid. From the image, we extract the measurement observations $\ell_i(u)$ that we will use to build a complete time-dependent 3D reconstruction of the blood velocity in the whole carotid Ω. As already brought up, the use of reduced-models will allow that the reconstructions are done in real-time.

Depending on the technology of the ultrasound device, there are two different types of velocity images. In most cases, ultrasound machines give a scalar mapping which is the projection of the velocity along the direction n of the ultasound probe. This mapping is called color flow image (CFI, see Figure 3(a)). In more modern devices, it is possible to get a 2D vector flow image corresponding to the projection of the velocity into the plane. This mapping is called vector flow image (VFI, see Figure 3(b)). For both imaging modes, the velocity is estimated by some time averaging techniques (we refer to [27] for further details).
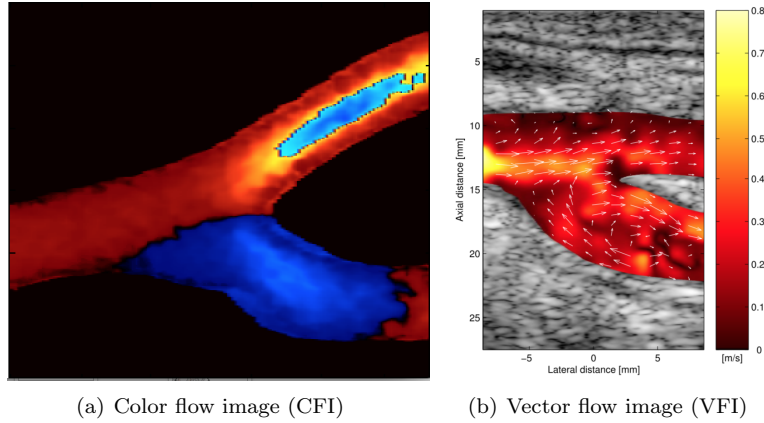

(a) Color flow image (CFI) (b) Vector flow image (VFI)

Figure 3: Velocity image of the common carotid bifurcation.

In the following, we work with an idealized version of CFI images. For each time t, a given image is a local average in space of the velocity projected into the direction in which the ultrasound probe is steered. More specifically, we consider a partition of $\Omega = \cup_{i=1}^m \Omega_i$ into m disjoint subdomains (voxels) Ω_i. Then, from each CFI image we collect

$$\ell_i(u) = \int_{\Omega_i} u \cdot n \, d\Omega_i, \quad 1 \leq i \leq m, \tag{3.6}$$

where n is a unitary vector giving the direction of the ultrasound beam. From (3.6), it follows that the Riesz representers of the ℓ_i in V are simply

$$\omega_i = \chi_{\Omega_i} n,$$

where χ_i denotes the characteristic function of the set Ω_i. Thus the measurement space is

$$W_m = W_m^{(CFI)} := \text{span}\{\omega_i\}_{i=1}^m.$$

Since the voxels Ω_i are disjoint from each other, the functions $\{\omega_i\}_{i=1}^m$ are orthogonal and therefore having a CFI image is equivalent to having

$$\omega = \mathbb{P}_{W_m} u = \sum_{i=1}^m \langle \omega_i, u \rangle \omega_i = \sum_{i=1}^m \ell_i(u) \omega_i. \tag{3.7}$$

Remark 3.2. *The case of VFI images can be treated similarly. This imaging mode gives $2m$ measurements*

$$\ell_i(u) = \int_{\Omega_i} u \cdot n \ d\Omega_i, \quad 1 \le i \le m,$$

and

$$\int_{\Omega_i} u \cdot n_\perp \ d\Omega_i, \quad 1 \le i \le m,$$

where n is again a unitary vector giving the direction of the ultrasound beam and $n_\perp$ is the unitary vector perpendicular to the bean and contained in the image plane. Therefore, $W_m^{(VFI)} = W_m^{(CFI)} \oplus \text{span}\{\chi_{\Omega_i} n_\perp\}_{i=1}^m$ which is a space of dimension $2m$. This clearly shows that the additional direction $n_\perp$ enriches the quality of the measurements in the sense that for any $u \in V$, the approximation error $\|u - P_{W_m} u\|$ will be smaller with the VFI mode than with the CFI one. As a result, the CFI mode which we consider in our examples is a more challenging case since the measurements contain less information.

3.3 Reconstruction on a first example with healthy patients

To validate our method, we first consider a simple example where we only work with healthy patients, so the manifold is $\mathcal{M}_{\text{healthy}}$. CFI mode allows us to define a measurement space W_m of dimension $m = 552$ (whereas for VFI, $m = 1104$) following the guidelines of section 3.2. As a result, each image can be seen as an observation $\omega \in W_m$ (see Figure 4). Our training set is composed of $\#\widetilde{\mathcal{M}}_{\text{train}} = 56383$ snapshots which have been generated from a set of 297 healthy patients. The performance of the algorithms is tested on a test set $\widetilde{\mathcal{M}}_{\text{test}}$ of 32 healthy patients.
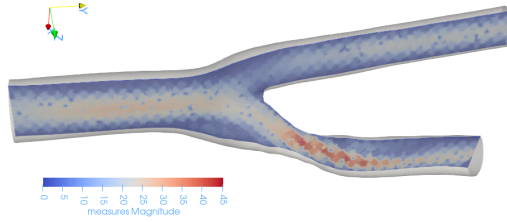

Figure 4: Example of synthetic CFI measures used in first example. The image leads to an space W_m of dimension $m = 552$

Using the training set of snapshots, we build the following reconstruction algorithms:

1. *Linear PBDW:* We build the linear spaces V_n from a classical singular value decomposition of the set $\widetilde{\mathcal{M}}_{\text{train}}$.

2. *Nonlinear algorithm with manifold partitioning:* In real medical examinations, the heart rate HR of the patient and the time t in which the ultrasound image is taken are known. We exploit this fact to decompose $\widetilde{\mathcal{M}}_{\text{train}}$ into $K = IJ$ subsets

$$\widetilde{\mathcal{M}}_{\text{train}} = \bigcup_{\substack{k=(i,j) \\ (i,j) \in \{1,...,I\} \times \{1,...,J\}}} \widetilde{\mathcal{M}}^{(k)}, \tag{3.8}$$

10

where, for each $k = (i,j)$ in $\{1, \ldots, I\} \times \{1, \ldots, J\}$,

$$\widetilde{\mathcal{M}}^{(k)} = \{u \in \widetilde{\mathcal{M}}_{\text{train}} : t \in [t_i - \tau, t_i + \tau], \text{ HR} \in [\text{HR}_j - \delta_{\text{HR}}, \text{HR}_j + \delta_{\text{HR}}]\}. \tag{3.9}$$

After an extensive error benchmark for the subset sizes τ and δ_{HR} we concluded that it is convenient to take in our tests, $\tau = T_c/10$ and $\delta_{\text{HR}} = 5$ beats per minute.

For each subset $\widetilde{\mathcal{M}}^{(k)}$ of the partition, we build in an offline phase a reduced model. Two constructions have been tested:

- A classical singular value decomposition of $\widetilde{\mathcal{M}}^{(k)}$.
- A classical greedy algorithm: for $n = 1$, we set $V_1^{(k)} = \text{span}\{u_1^{(k)}\}$ with

$$u_1^{(k)} = \frac{1}{\#\widetilde{\mathcal{M}}^{(k)}} \sum_{u \in \widetilde{\mathcal{M}}^{(k)}} \|u\|.$$

For $n > 1$, we select

$$u_n^{(k)} \in \arg\max_{u \in \widetilde{\mathcal{M}}^{(k)}} \|u - P_{V_{n-1}}u\|,$$

and set $V_n^{(k)} = \text{span}\{V_{n-1}^{(k)}, u_n^{(k)}\}$.

During the online reconstruction, given t and HR, we select the appropriate subset $\widetilde{\mathcal{M}}^{(k)}$ that includes t and HR and reconstruct with a linear PBDW with the reduced model corresponding to $\widetilde{\mathcal{M}}^{(k)}$.

3. *Data-driven nonlinear algorithm:* Since each CFI image can be seen as an observation $\omega \in W_m$, we run the Orthogonal Matching Pursuit algorithm of section 2.3.2 to build $V_n(\omega)$ and do the reconstruction. Note that the greedy search has to be done online since we need the knowledge of the measurement. To speed-up computations, instead searching in the whole training set

$$\delta_{\bar{u}}\widetilde{\mathcal{D}} = \left\{ v = \frac{u - \bar{u}}{\|u - \bar{u}\|} \; : \; u \in \widetilde{\mathcal{M}}_{\text{train}} \right\},$$

we restrict the search to the partition

$$\delta_{\bar{u}}\widetilde{\mathcal{D}}^{(k)} = \left\{ v = \frac{u - \bar{u}}{\|u - \bar{u}\|} \; : \; u \in \widetilde{\mathcal{M}}^{(k)} \right\}.$$

For each patient $u \in \widetilde{\mathcal{M}}_{\text{test}}$, we compute the average reconstruction error over the whole second cardiac cycle

$$e(u, A_{n,m}) = \left(\frac{\int_{T_c}^{2T_c} \|u(t) - A_{m,n}(P_{W_m}u(t))\|^2 \mathrm{d}t}{\int_{T_c}^{2T_c} \|u(t)\|^2 \mathrm{d}t} \right)^{1/2} \tag{3.10}$$

where $A_{m,n}(P_{W_m}u)$ denotes any of the above four reconstruction algorithms. We also consider the error in time

$$e(t, u, A_{m,n}) = \frac{\|u(t) - A_{m,n}(P_{W_m}u(t))\|}{\left(\int_{T_c}^{2T_c} \|u\|^2 \mathrm{d}t \right)^{1/2}} \tag{3.11}$$

Note that in both formulas we have divided by the total energy in the cardiac cycle $\left(\int_{T_c}^{2T_c} \|u(t)\|^2 \mathrm{d}t \right)^{1/2}$.

Figures 5(a) to 5(d) give the reconstruction errors $e(u, A_{n,m})$ as a function of the dimension n of V_n and for patients $u \in \widetilde{\mathcal{M}}_{\text{test}}$. The results are presented in the form of box plots to show the small deviations of the error around the mean value. Figures 6(a) and 6(b) give the average and worst case performance of the four methods,

$$e_{\text{av}}(A_{n,m}) = \sum_{u \in \widetilde{\mathcal{M}}_{\text{test}}} e(u, A_{n,m}), \quad e_{\text{wc}}(A_{n,m}) = \max_{u \in \widetilde{\mathcal{M}}_{\text{test}}} e(u, A_{n,m}).$$

Note that as n increases, the error decreases for all methods, except perhaps for the OMP approach where the error tends to stagnate for large values of n. This could be due to the fact that OMP heavily relies on the measurement information, which, in the present application, might not deliver enough information to learn reduced models $V_n(\omega)$ that improve the accuracy as n grows. We also see from the figure that the nonlinear method based on POD with windowing outperforms the rest in the sense that delivers a given target accuracy with a smaller dimension n of the reduced model. For instance, if we fix a target accuracy on the average performance to 10^{-2}, we see that the linear
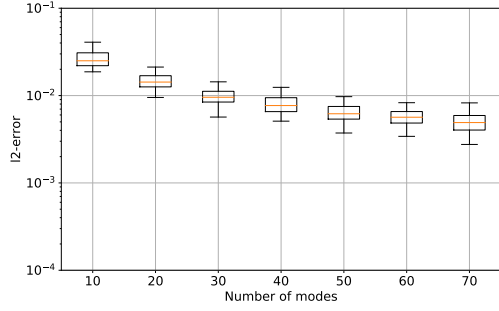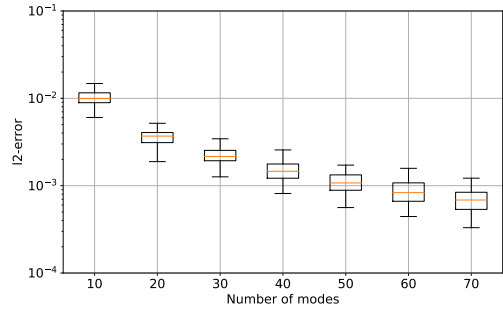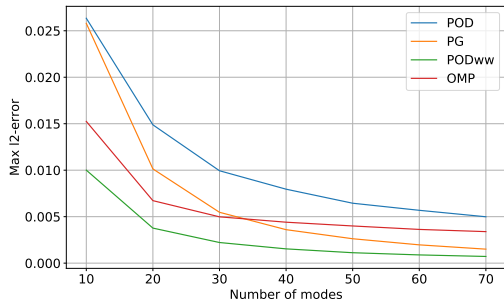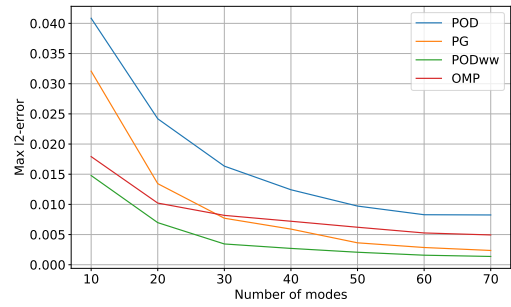
(a) POD
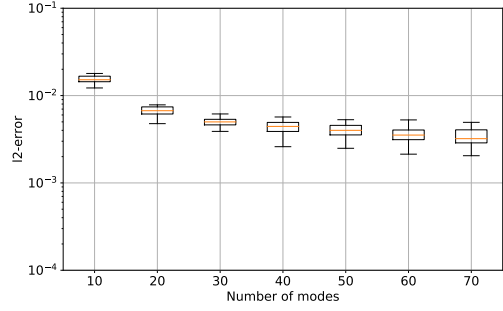
(b) OMP
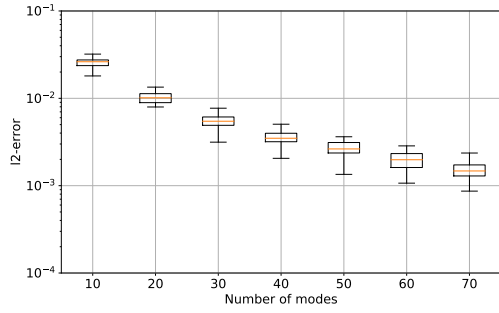
(c) PG

(d) POD with windowing

Figure 5: Reconstruction error $e(u, A_{n,m})$ for patients $u \in \widetilde{\mathcal{M}}_{\text{test}}$ in first example. The box-plots shows a regular distribution of the data for the 4 basis construction methods.

(a) Average error

(b) Worst error

Figure 6: Benchmark for first numerical example. The accuracy is evaluated using the error 3.10
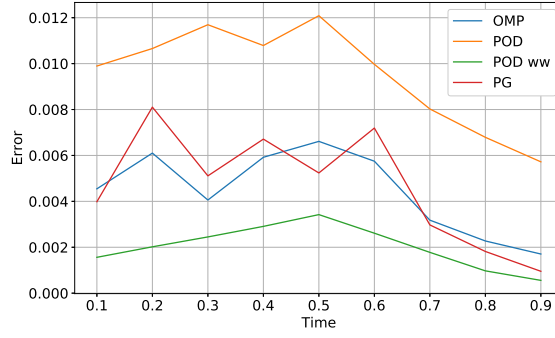
Figure 7: Average error comparison for test case described in section 3.3. The benchmark shows the temporal evolution of the quantity 3.11 during the cardiac cycle. The dimension of V_n is set to 30.

PBDW with POD requires 40 modes to achieve it, PG requires 20, OMP requires 17 and POD with windowing requires only 10 (see Figure 6(a)).

We next fix $n = 30$ and study the error in time $e(t, u, A_{m,30})$ on Figure 7. We observe that the reconstruction tends to be better during the late diastole phase of the cardiac cycle. This is probably due to the fact that the flow acceleration is much slower in this period, hence easier to compress in a reduced model.

As discussed in section 2.3, the inf-sup constant $\beta(V_n, W_m)$ might yield to stability issues when $n \to m$ since its value tends to zero (see equations (2.12) and (2.13)). Figures 8(a) and 8(b) show its behavior for the four methods during the systole and diastole period. We observe that the four methods performs similarly in terms of stability for the peak systole reconstruction. For the diastole phase, we observe that the inf-sup constant in OMP performs slightly worse than the rest. We think that this could be due to the fact that the measurement space W_m is not rich enough to allow OMP to properly learn reduced models when n becomes large.
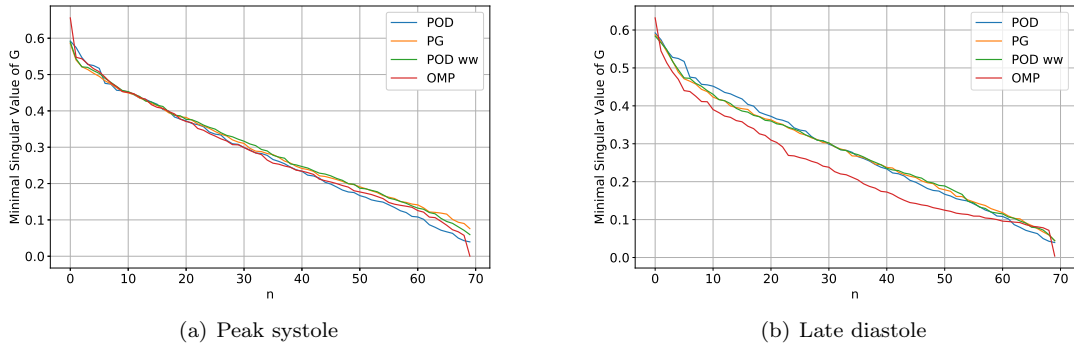

(a) Peak systole

(b) Late diastole

Figure 8: Behavior of $\beta(V_n, W_m)$ as a function of n.

3.4 Application to blockage detection

In this example we illustrate that even when the Doppler images do not give information on the whole carotid, we can nevertheless reconstruct the velocity field in the whole domain with our methodology. This is important for actual practice since doctors do not have images in the whole carotid due to morphological constraints. We also show in our example that the method has potential to efficiently estimate in real time relevant quantities of interest.

We illustrate these ideas in the following example: we consider the same setting as before but now the Doppler image does not provide information about the flow in the carotid bifurcations. Therefore, the image does not *see* the flow split in the common carotid downstream (see Figure 9). In this example we have tested the impact of working with CFI or VFI images.
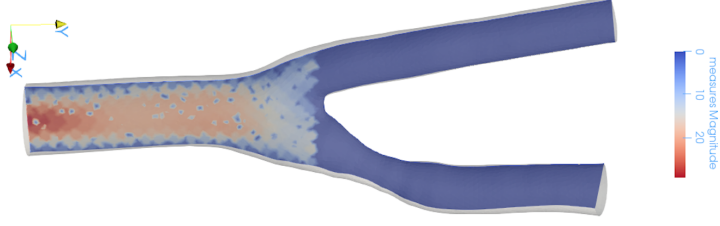
13

Figure 9: Synthetic CFI for the case of a blockage. Note that the image does not capture the flow after the bifurcation.

We train our reconstruction methods on a training set $\widetilde{\mathcal{M}}_{\text{train}}$ containing sick and healthy patients. Here, we only work with our two nonlinear algorithms:

1. the PBDW method with the same partition of the training set as in the previous example and with reduced models generated with SVD,

2. the data-driven OMP approach.

Figure 10 shows the average and worst case errors

$$e_{\text{av}}(A_{n,m}) = \sum_{u \in \widetilde{\mathcal{M}}_{\text{test}}} e(u, A_{n,m}), \quad e_{\text{wc}}(A_{n,m}) = \max_{u \in \widetilde{\mathcal{M}}_{\text{test}}} e(u, A_{n,m}),$$

as a function of the dimension n of V_n. Like in the previous example, both methods are delivering a very satisfactory accuracy: the average error is below 5.10^{-2} for both methods for all values of n. The method consisting in a partition of the manifold outperforms OMP.
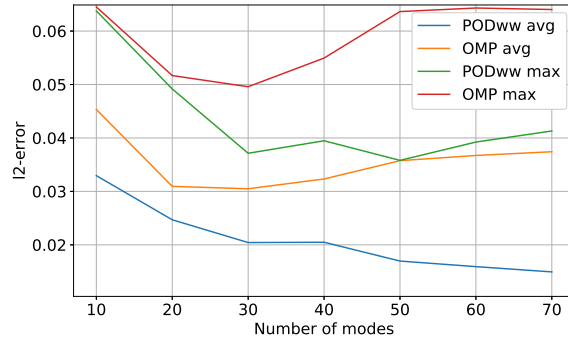

Figure 10: Average and worst reconstruction errors as a function of n.

We next show that the method is efficient to assist in the detection of blockages that may cause severe health problems like a stroke. Since a blockage alters the distribution of the velocity field after the bifurcation, a quantity of interest that could serve as a clinical index is the ratio

$$r = \frac{Q_2(t_{peak})}{Q_1(t_{peak})} \tag{3.12}$$

where

$$Q_i(t) := \int_{\Gamma_o^i} u(t) \cdot n$$

is the blood flow at the outlet Γ_o^i, $i = 1, 2$, and t_{peak} is the peak systole instant. Figure 11 shows the evolution of $Q_i(t)$ in time for a sick patient and its approximation with our two reconstruction methods. We observe that, regardless of the image format (CFI or VFI), both methods deliver very satisfactory predictions of the flow.

In Figure 12, we compare the value of the exact ratio r with the reconstructed one for sick and healthy patients $u \in \widetilde{\mathcal{M}}_{\text{test}}$.
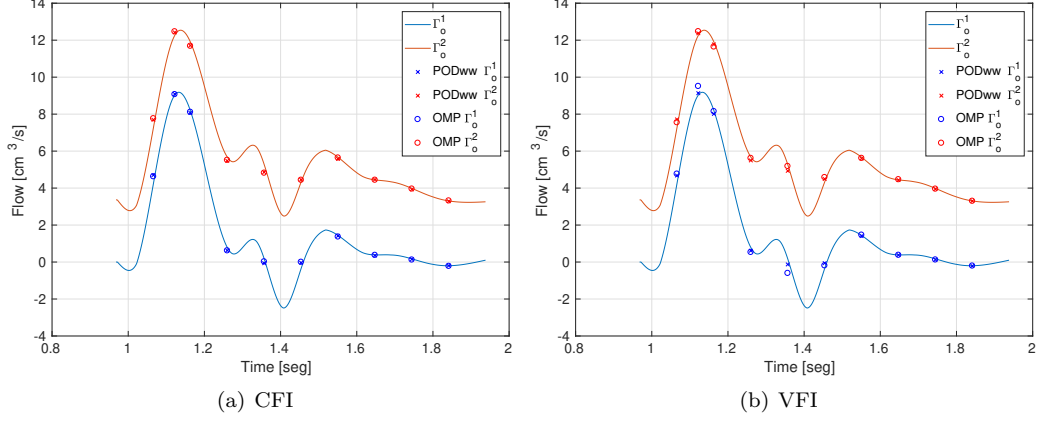
14

(a) CFI

(b) VFI

Figure 11: Comparison of simulated and reconstructed flow at the outlets of the domain.

To define a threshold ratio r^* to decide whether the patient has arterial blockage or not, we can take the average of the flow ratios between the healthier of the patients in the sick group and the sicker of the patients in the healthy group, namely,

$$r^* := \frac{\min_{v \in \widetilde{\mathcal{M}}_{\text{sick}}} r(v) + \max_{u \in \widetilde{\mathcal{M}}_{\text{healthy}}} r(u)}{2} \tag{3.13}$$

where $r(u)$ denotes the flow ratio associated to the velocity field u, as defined in 3.12. In our database, we obtain, $r^* = 1.25$, so any patient for which $r > 1.25$ will be considered as presenting high blockage risk. Note that the approximation is very close to the real value for moderate values of r regardless of the image modality. However, we tend to overestimate the value for $r > 1.7$. In presence of a blockage, r becomes significantly larger than one so the overestimation is by far more preferable than an underestimation. Indeed, the overestimation makes our method conservative and, in the worst case, we will conclude with a false positive. However, the method will not lead to a false negative diagnosis, which would leave a sick patient without treatment/surgery.
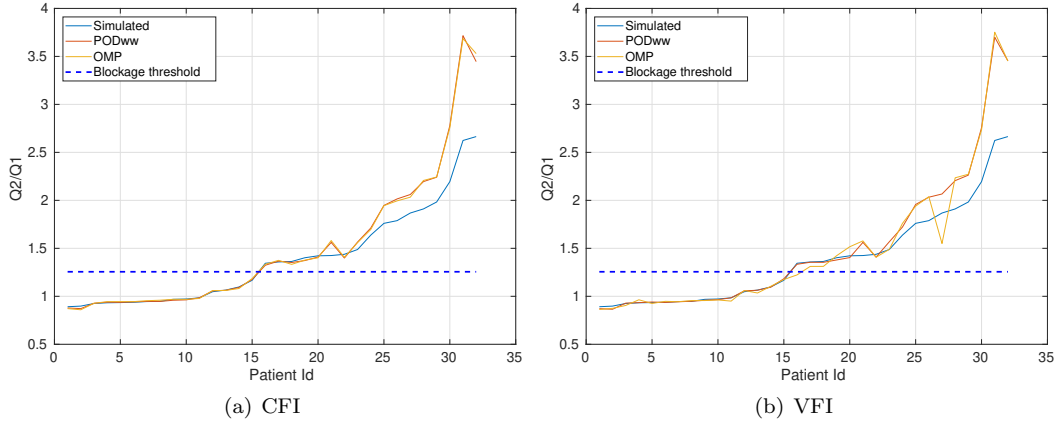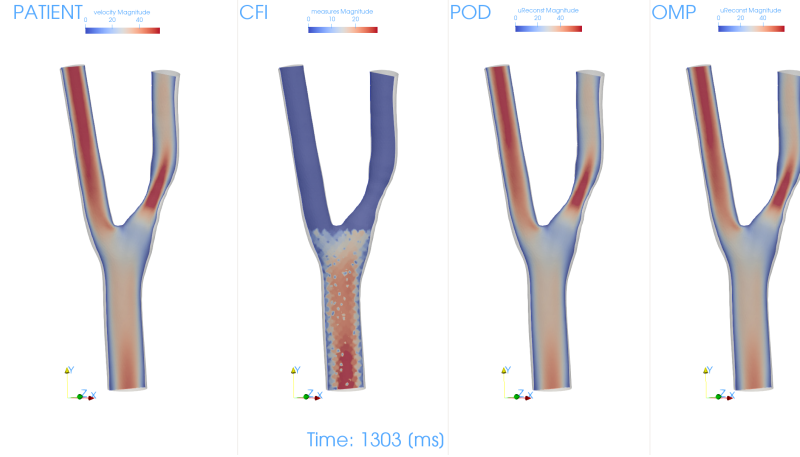

(a) CFI

(b) VFI

Figure 12: Flow ratio reconstructed for blockage detection. In this test the number of modes is fixed to 30.

Last but not least, Figure 13 shows an illustration of the reconstruction with our methods for a CFI and VFI Doppler image. From left to right, we see the exact flow, the Doppler image (CFI or VFI), the reconstruction with the POD with windowing and with OMP. Note that visually the reconstruction looks practically identical to the exact flow.

(a) CFI case
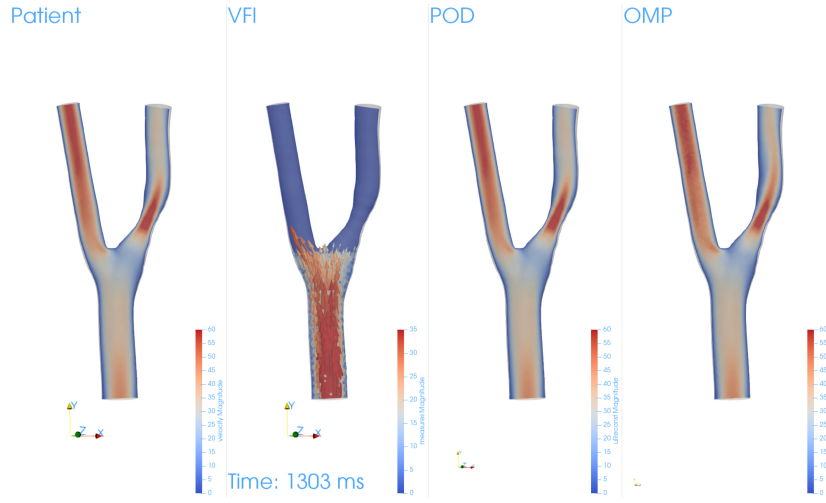
(b) VFI case

Figure 13: Reconstruction example for blockage case. From left to right, we see the target, the ideal image, the reconstruction with POD and the one with OMP.

4 Conclusions and perspectives

A comparison between different reconstruction methods to estimate flow velocity field starting from Doppler ultrasound images were proposed. A particular focus was put on the basis construction and we have explored the potential of nonlinear algorithms with respect to the classical linear PBDW method. The numerical tests were synthetically generated and mimic a real context in medical applications. From the numerical experiments it follows that a simple POD basis is not an optimal choice: this is primarily due to the fact that there are different flow regimes and patterns in a cardiac cycle and for different heart beat rates. The nonlinear reconstruction method built by POD with a partition in the parameters that are a priori known outperforms all the other bases choices, and it is a good trade off between simplicity and efficiency (the accuracy is in some cases even ten times better than a simple POD). The results give solid numerical evidence of the potential of the method as tool to assist in medical diagnoses. A current line of investigation is how to deal with realistic, noisy and time delayed measurements.

Acknowledgments

A   Explicit expression of $u^*_{m,n}(w)$, the function given by the linear PBDW algorithm

Let X and Y be two finite dimensional subspaces of V. In the following, the map

$$P_{X|Y} : Y \to X$$
$$y \mapsto P_X y$$

denotes the orthogonal projection into X restricted to Y.

Lemma A.1. *Let W_m and V_n be an observation space and a reduced basis of dimension $n \le m$ such that $\beta(V_n, W_m) > 0$. Then the linear PBDW algorithm is given by*

$$u^*_n(w) = w + v^*_n - P_W v^*_n, \tag{A.1}$$

with

$$v^*_n = \left(P_{V_n|W_m} P_{W_m|V_n}\right)^{-1} P_{V_n|W_m}(w). \tag{A.2}$$

Proof. By formula (2.5), $u^*_n(w)$ is a minimizer of

$$
\begin{aligned}
\min_{u \in w + W_m^\perp} \mathrm{dist}(u, V_n)^2 &= \min_{u \in w + W_m^\perp} \min_{v \in V_n} \|u - v\|^2 \\
&= \min_{v \in V_n} \min_{\eta \in W_m^\perp} \|w + \eta - v\|^2 \\
&= \min_{v \in V_n} \|w - v - P_{W_m^\perp}(w - v)\|^2 \\
&= \min_{v \in V_n} \|w - v + P_{W_m^\perp}(v)\|^2 \\
&= \min_{v \in V_n} \|w - P_{W_m}(v)\|^2.
\end{aligned}
$$

The last minimization problem is a classical least squares optimization. Any minimizer $v^*_n \in V_n$ satisfies the normal equations

$$P^*_{W_m|V_n} w = P^*_{W_m|V_n} P_{W_m|V_n} v^*_n,$$

where $P^*_{W_m|V_n} : V_n \to W_m$ is the adjoint operator of $P_{W_m|V_n}$. Note that $P^*_{W_m|V_n}$ is well defined since $\beta(V_n, W_m) = \min_{v \in V_n} \|P_{W_m|V_n} v\| / \|v\| > 0$, which implies that $P_{W_m|V_n}$ is injective and thus admits an adjoint. Furthermore, since for any $w \in W_m$ and $v \in V_n$, $\langle v, w \rangle = \langle P_{W_m|V_n} v, w \rangle = \langle v, P_{V_n|W_m} w \rangle$, it follows that $P^*_{W_m|V_n} = P_{V_n|W_m}$, which finally yields that the unique solution of the least squares problem is

$$v^*_n = \left(P_{V_n|W_m} P_{W_m|V_n}\right)^{-1} P_{V_n|W_m} w.$$

Therefore $u^*_n = w + \eta^*_n = w + v^*_n - P_{W_m} v^*_n$. $\square$

17

References

[1] B. Lahoz and R. Menard. *Data assimilation.* Springer, 2010.

[2] D. Simon. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches.* John Wiley & Sons, 2006.

[3] G. Evensen. *Data assimilation: the ensemble Kalman filter.* Springer Science & Business Media, 2009.

[4] M. Khalil, S. Adhikari, and A. Sarkar. Linear system identification using proper orthogonal decomposition. *Mechanical Systems and Signal Processing*, 21(8):3123–3145, 2007.

[5] P. Astrid, S. Weiland, K. Willcox, and T. Backx. Missing point estimation in models described by proper orthogonal decomposition. *IEEE Transactions on Automatic Control*, 53(10):2237–2251, 2008.

[6] M. Buffoni, S. Camarri, A. Iollo, E. Lombardi, and M. Salvetti. A non-linear observer for unsteady three-dimensional flows. *Journal of Computational Physics*, 227(4):2626–2643, 2008.

[7] R. Leroux, L. Chatellier, and L. David. Application of kalman filtering and partial least square regression to low order modeling of unsteady flows. In *TSFP Digital Library Online*. Begel House Inc., 2013.

[8] Y. Maday and O. Mula. A Generalized Empirical Interpolation Method: application of reduced basis techniques to data assimilation. In Franco Brezzi, Piero Colli Franzone, Ugo Gianazza, and Gianni Gilardi, editors, *Analysis and Numerics of Partial Differential Equations*, volume 4 of *Springer INdAM Series*, pages 221–235. Springer Milan, 2013.

[9] 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. *Computer Methods in Applied Mechanics and Engineering*, 287(0):310–334, 2015.

[10] M. Raiola, S. Discetti, and A. Ianiro. On PIV random error minimization with optimal POD-based low-order reconstruction. *Experiments in Fluids*, 56(4):75, 2015.

[11] 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 Marco L. Bittencourt, Ney A. Dumont, and Jan S. Hesthaven, editors, *Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2016: Selected Papers from the ICOSAHOM conference, June 27-July 1, 2016, Rio de Janeiro, Brazil*, pages 133–145, Cham, 2017. Springer International Publishing.

[12] M. Kärcher, S. Boyaval, M.A. Grepl, and K. Veroy. Reduced basis approximation and a posteriori error bounds for 4d-var data assimilation. *Optimization and Engineering*, 19(3):663–695, Sep 2018.

[13] 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. *International Journal for Numerical Methods in Engineering*, 102(5):933–965, 2015.

[14] A. Cohen, O. Mula, and J. Nichols. Data-driven reduced models. 2019.

[15] A. Cohen, R. Dahmen, W. an DeVore, J. Fadili, O. Mula, and J. Nichols. Optimal reduced model algorithms for data-based state estimation. *arXiv preprint arXiv:1903.07938*, 2019.

[16] Y. Maday, O. Mula, and G. Turinici. Convergence analysis of the Generalized Empirical Interpolation Method. *SIAM Journal on Numerical Analysis*, 54(3):1713–1731, 2016.

[17] P. Binev, A. Cohen, O. Mula, and J. Nichols. Greedy algorithms for optimal measurements selection in state estimation using reduced models. *SIAM/ASA Journal on Uncertainty Quantification*, 6(3):1101–1126, 2018.

[18] G. Rozza, D. B. P. Huynh, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3):1, Sep 2007.

[19] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An Empirical Interpolation Method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Acad. Sci. Paris, Série I.*, 339:667–672, 2004.

[20] A. Cohen, R. DeVore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's. *Analysis and Applications*, 09(01):11–47, 2011.

[21] A. Cohen and R. DeVore. Approximation of high-dimensional parametric pdes. *Acta Numerica*, 24:1–159, 2015.

[22] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Data assimilation in reduced modeling. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):1–29, 2017.

[23] T. Taddei. An adaptive parametrized-background data-weak approach to variational data assimilation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 51(5):1827–1858, 2017.

[24] P. Blanco, S. Watanabe, M. Passos, P. Lemos, and R. Feijoo. An anatomically detailed arterial network model for one-dimensional computational hemodynamics. *International Journal for Numerical Methods in Biomedical Engineering*, 2014.

[25] L. Formaggia, A. Quarteroni, and A. Veneziani. *Cardiovascular Mathematics. Modeling and simulation of the circulatory system.*, volume 1. 01 2009.

[26] C. Bertoglio, A. Caiazzo, Bazilevs Y., M. Braack, M. Esmaily-Moghadam, V. Gravemeier, A.L. Marsden, O. Pironneau, I.E. Vignon-Clementel, and W.A. Wall. Benchmark problems for numerical treatment of backflow at open boundaries. *preprint*, 2017.

[27] C. Kasai, K. Namekawa, A. Koyano, and R. Omoto. Real-time two-dimensional blood flow imaging using an autocorrelation technique. *IEEE transactions on sonics and ultrasonics*, 1985.