Surrogate modelling and uncertainty quantification for multiscale simulation

Ye, D.

Publication date
2022

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
4.1 Introduction

Computational modelling has demonstrated significant efficiency and flexibility in cardiovascular science for the study of mechanisms of disease, biological/pathological processes and designs of medical devices [89, 133–136]. Computational models describe the behaviour of systems using governing equations, usually via partial differential equations (PDEs) with particular initial and boundary conditions, but also using cell-based or particle based approaches. To capture the full complexity of the behavior of the system, the models are generally sophisticated and involve domains with a complex shape. Numerical methods, such as finite element methods (FEM) [137] and finite volume methods [138] are applied to solve the model equations. The constructed models subsequently contribute to studying the underlying dynamics of the system and
making predictions to support decision-making in future clinical applications.

The evaluation of fluid dynamics, especially hemodynamics, is one of the important topics in cardiovascular science. The mechanical and biological interaction between blood flow and the vessel wall has a significant impact on the initiation and progression of vascular diseases, e.g. the development of atherosclerotic plaque [139–141], or restenosis of arteries after percutaneous intervention [142, 143]. Therefore computational fluid dynamics (CFD) techniques are widely applied to perform high-fidelity simulations to mimic and predict the behavior of blood flow, to further support the clinical practice or in-depth study of physiology and pathology [88, 144–146].

However, owing to the complexity and scale of cardiovascular flow problems, the evaluation of the model could be computationally expensive, especially in the cases where a large number of evaluations are required, such as design optimisation, reliability analysis and uncertainty quantification [6, 28]. Reduced order modelling (ROM) provides a high-fidelity framework to reduce the computational complexity for solving parametric PDEs [70, 147]. It exploits the intrinsic correlation of the full order model (FOM) solution over a physical domain, time evolution, or parameter space, and projects the system operation to a low-dimensional reduced space. One of the widely applied reduced-order methods is Proper Orthogonal Decomposition (POD) [67]. The POD can be realized using principal component analysis (PCA), or the singular value decomposition (SVD). The projection coefficients can be computed either in an intrusive manner by manipulating the governing equations with Galerkin methods [68, 69] or in a non-intrusive manner by formulating it as an interpolation problem [60, 70]. Dynamic mode decomposition is another projection-based method for the temporal decomposition of CFD [71, 72]. The empirical interpolation method [73] and the discrete empirical interpolation method [74] were proposed to recover an affine expansion for the nonlinear problem. Operator inference methods can directly approximate the reduced-order differential operator from data without knowing the full-order operators [75, 76].

In cardiovascular engineering, the shapes of domains for blood flow simulations are either segmented from clinical image data or generated indirectly
by computational models [148]. These anatomical geometries of a kind, such as the segments of arteries and organs are similar in terms of general shape, but different in details. These shapes are typically irregular and unparametrised. Apart from the parametrisation issue, the difference in the shapes of the domain also leads to distinct spatial discretisations. To ensure the spatial compatibility among snapshots for ROM, a parameter-independent reference domain is needed hence all the discrete solutions generated from FEM share the same discretisation. Such a method requires a mapping between the reference and target domains that can be computed via an interpolation based on point-to-point (or landmark) correspondence between the two domains, which is not a general known condition in cardiovascular engineering.

In this chapter, a data-driven surrogate model based on the non-intrusive reduced order modelling method is proposed. The surrogate model aims to approximate the fluid dynamics of blood flow within vessels of distinct but similar shapes efficiently. Different from existing approaches, the proposed surrogate model performs a surface registration based on currents to approximate the shape of the target domain from the reference shape. The diffeomorphism constructed during surface registration provides the mapping between the reference domain and the target domain, as well as the parametrisation of the shape. The evaluations are subsequently performed on the reference domain to ensure the spatial compatibility of the snapshots. POD is applied to construct reduced bases and the projection coefficients are computed via a radial basis function (RBF) interpolation method.

The chapter is arranged as follows. The details of the surrogate model are described in Section 4.2. Two numerical examples are presented and discussed in Section 4.3, followed by a conclusion in Section 4.4.

4.2 Method

4.2.1 Surfaces registration with current

Here we introduce an approach based non-parametric representation of surfaces using currents and its deformation framework [149–151]. The method
characterises shapes in the form of currents which enables one to quantify the dissimilarity between two shapes without point correspondences, and subsequently can be applied to perform surface registration. We assume that the shapes of domains in the blood flow simulations are similar and can be achieved by a deformation of the reference domain. A diffeomorphic registration, which transforms shapes from a source coordinate system to a target coordinate system, is generated via minimising the difference between two surfaces and geodesic distance.

Represent surfaces using currents

Given a continues vector fields $\omega : \mathbb{R}^d \rightarrow \mathbb{R}^d$ in the ambient space $\mathbb{R}^d$, where $d = 2$ or $3$, an oriented surface $S$ in $\mathbb{R}^3$ can be characterised by the flux going through the surface,

$$[S](\omega) = \int_S \langle \omega(x), n(x) \rangle dS(x), \quad (4.1)$$

where $n(x)$ denotes the unit normal vector of the surface $S$ at point $x$. For a curve in two dimensional case ($d = 2$), $n(x)$ is replaced by a tangent vector. The computation of the flux defines a linear mapping from a space of vector field $W$ to $\mathbb{R}$. This mapping is called the current associated with surface $S$. We denote the space of current as $W'$. By defining the space $W$ as a Reproducing Kernel Hilbert Space (RKHS), the vector field can be represented in the form of a kernel function $K : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$, therefore a vector field $\omega$ in $W$ can be reformulated with fixed points $\{x_i^*\}_{i=1,2,\ldots,n}$ and their corresponding vectors $\beta$,

$$\omega(x) = K^W(x,x^*)\beta. \quad (4.2)$$

The space $W$ equips with inner product $\langle K(\cdot,x)\alpha, K(\cdot,x')\beta \rangle_W = \alpha^T K(x,x') \beta$. Generally a Gaussian kernel $K^W(x_i,x_j) = \exp(-\frac{|x_i-x_j|^2}{\lambda_W})$ is applied, where $\lambda_W$ denotes the scale at which $W$ may spatially vary. If the normal vector field associated to a surface is presented in the kernel form, equation 4.1 can be written into,
4.2. Method

\[ [S](\omega) = \int_S \langle \omega, K(\cdot, x)n(x) \rangle_W dS(x) \]  
\[ = \langle \int_S K(\cdot, x)n(x)dS(x), \omega \rangle_W . \]  

(4.3)

It shows that \( \int_S K(\cdot, x)n(x)dS(x) \) is a unique Riesz representation in \( W \) of current \( [S] \) in \( W' \) \[152\]. Therefore, the inner produce of two currents \( [S_1] \) and \( [S_2] \) can be obtained by,

\[ \langle [S_1], [S_2] \rangle_{W'} := \int_{S_1} \int_{S_2} K(x, x')\langle n_1(x), n_2(x') \rangle dS_1(x)dS_2(x'). \]  

(4.4)

The associated norm can be applied to measure the dissimilarity between two shapes,

\[ \| [S_1] - [S_2] \|^2_{W'} = \langle [S_1], [S_1] \rangle_{W'} - 2\langle [S_1], [S_2] \rangle_{W'} + \langle [S_2], [S_2] \rangle_{W'}. \]  

(4.5)

**Diffeomorphism construction**

The shape of a reference domain can be represented by a series of control points (CPs) \( \{x_i^{cp}\}_{i=1,...,N_{cp}} \) on a boundary (the boundary vertices in a finite element mesh). The surface registration of a reference shape \( S_{ref} \) to a target shape \( S_{tar} \) can be performed by finding a flow of diffeomorphism \( \varphi(\cdot, t) : \mathbb{R}^d \to \mathbb{R}^d \) such that the dissimilarity of the shapes, as well as geodesic distance of the process, are minimised. The trajectory of these points at a given time \( t \), can be described using a flow differential equation:

\[ \frac{\partial \varphi(x, t)}{\partial t} = \nu(\varphi(x, t)), \]  

(4.6)

where \( \nu(\varphi(x, t)) \) is a vector field representing the velocity of the deformation. Similar to equation 4.2, the velocity vector field can be approximated by a summation of kernel function at discrete CPs,

\[ \nu(x) = \sum_{i=1}^{N_{cp}} K(x, x_i^{cp})\alpha_i, \]  

(4.7)
where $\alpha_i$ denotes the basis vector of the field and $K(\cdot, \cdot)$ is a kernel function that decides the weights of each basis based on the distance between two points. Together with equation 4.6, the flow of diffeomorphism can be reformulated as the motion of CPs,

$$\frac{\partial \varphi_t(x_j)}{\partial t} = \sum_{i=1}^{N_{cp}} K(x_j, x_i^{cp})\alpha_i(t), \quad j = 1, \cdots, N_{cp}. \quad (4.8)$$

To perform a surface registration from a reference shape $S_{\text{ref}}$ to a target shape $S_{\text{tar}}$, a proper velocity vector field can be computed via an optimisation problem formulated as,

$$\arg\min_v J(v) = \|\varphi(S_{\text{ref}}) - [S_{\text{tar}}]\|^2_{W'} + \int_0^1 \|v\|^2_W dt, \quad (4.9)$$

where $\|\varphi(S_{\text{ref}}) - [S_{\text{tar}}]\|^2_{W'}$ measures the difference between the deformed shape and target shape using currents; $\int_0^1 \|v\|^2_W dt$ denotes the total kinetic energy that required to deform the shape from its initial state.

**Radial basis function interpolation**

The radial basis function method is one of the state-of-the-art methods for multivariate interpolation [153]. The method approximates the desired function by a linear combination of basis functions based on the known collocation points. It has been widely applied in computational science and engineering [154, 155]. In this work, RBF interpolation is applied for both computing the mapping between reference and target domain, and predicting the reduced coefficients for non-intrusive ROM. Here we present the description based on a general form of RBF interpolation.

Assume that $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is the latent function between inputs and outputs, and a collection of collocation points $\{f(x_i)\}_{i=1}^{\ell}$ is known. A general form of RBF interpolator $\hat{f}$ can be written as:

$$\hat{f}(x) = \sum_{i=1}^{\ell} \lambda_i \psi(\|x - x_i\|) + P(x), \quad (4.10)$$
such that
\[
\hat{f}(x_j) = \sum_{i=1}^{\ell} \lambda_i \psi(||x_j - x_i||) + P(x_j) = f(x_j), \text{ for } j = 1, \cdots, \ell, \tag{4.11}
\]
where \(\psi : [0, +\infty] \rightarrow \mathbb{R}\) is the radial basis function which is also known as kernel function. It computes similarity between a prediction point and known collocation points. Multiple choices of kernels are available such as cubic kernel, thin-plate spline kernel or Gaussian kernel. Each kernel leads to its unique asymptotic behaviour, see [156] for more details; \(\lambda_i\) denotes the expansion coefficients (weights) for each collocation points; \(P(x) = \sum_{i=1}^{q} \eta_i p_i(x)\) denotes a low-order polynomial function for conditionally positive definite radial functions (e.g. thin-plate spline kernel), where \(p_1, p_2, \cdots, p_q\) are the polynomial bases of order not exceeding \(q\). The additional degrees of freedom introduced by the polynomial function are compensated by,

\[
\sum_{i=1}^{\ell} \lambda_i p_j(x_i) = 0, \quad j = 1, \cdots, q. \tag{4.12}
\]

Together with the constraint from equation 4.11, a linear system can be formed,

\[
\begin{pmatrix}
\Psi \\
\mathbf{P}^T
\end{pmatrix}
\begin{pmatrix}
\lambda \\
\eta
\end{pmatrix} =
\begin{pmatrix}
f \\
0
\end{pmatrix}, \tag{4.13}
\]

where
\[
\Psi = \begin{pmatrix}
\psi(||x_1 - x_1||) & \cdots & \psi(||x_1 - x_\ell||) \\
\vdots & \ddots & \vdots \\
\psi(||x_\ell - x_1||) & \cdots & \psi(||x_\ell - x_\ell||)
\end{pmatrix}, \quad
\mathbf{P} = \begin{pmatrix}
p_1(x_1) & \cdots & p_q(x_1) \\
\vdots & \ddots & \vdots \\
p_1(x_\ell) & \cdots & p_q(x_\ell)
\end{pmatrix},
\]

and \(\lambda = [\lambda_1, \cdots, \lambda_\ell]^T\), \(\eta = [\eta_1, \cdots, \eta_q]^T\) and \(f = [f(x_1), \cdots, f(x_\ell)]^T\) being the vector form of expansion coefficients for the kernel function and the polynomial function, and output vector, respectively. By solving the linear system above, the expansion and polynomial coefficients can be obtained and hence an interpolation for a new point \(\hat{x}\) can be estimated based on equation 4.10.
4.2.2 FOM of a blood flow simulation on reference domain

Consider the blood flow with moderate Reynolds number as an incompressible Newtonian fluid to be modelled with the steady Navier-Stokes equations. Within a parametric vessel lumen as the spatial domain $\Omega(\gamma) \in \mathbb{R}^d$, where $d = 2$ or $3$, such problem can be formulated as: find vectorial velocity field $u(x; \gamma) : \Omega(\gamma) \to \mathbb{R}^d$, and scalar pressure field $p(x; \gamma) : \Omega(\gamma) \to \mathbb{R}$, such that:

\[
(u \cdot \nabla)u - \nu \nabla^2 u + \frac{1}{\rho} \nabla p = f, \\
\nabla \cdot u = 0,
\]

and subject to boundary conditions:

\[
\begin{align*}
    u &= g_D, & x &\in \partial \Omega_D(\gamma), \\
    -pn + \nu (n \cdot \nabla) u &= g_N, & x &\in \partial \Omega_N(\gamma),
\end{align*}
\]

where $\gamma \in \mathcal{P} \subset \mathbb{R}^{N_p}$ stands for the geometric parameters; $\nu$ and $\rho$ denotes the dynamics viscosity and density of the fluid respectively; $f$ denotes the body force. Equation 4.15a and 4.15b represent Dirichlet and Neumann boundary condition on $\partial \Omega_D$ and $\partial \Omega_N$, respectively. Dirichlet boundary condition is generally defined for the inlet and wall boundary of a vessel, while the Neumann boundary condition is applied to the outlet. For the sake of simplicity, the body force term is considered to be zero in this work.

As mentioned before, the governing equations should be solved on a parameter-independent reference domain $\Omega^*$ to ensure the spatial compatibility of the snapshot through a parameterised mapping $\mathcal{X}(\xi; \gamma)$, i.e. $x = \mathcal{X}(\xi; \gamma)$. The variational form of the system is therefore obtained by introducing two functional spaces $V := \{w \in H^1(\Omega^*) \mid w = 0 \text{ on } \partial \Omega_D^*\}$ of weighting functions for momentum conservation and $Q := \{q \in L^2(\Omega^*)\}$ for incompressible constraint over the reference domain. $Q$ can be also applied as the solution space for pressure. Besides, solution spaces $S := \{u \in H^1(\Omega^*) \mid u = g_D \text{ on } \partial \Omega_D^*\}$ is defined over the domain for velocity. The governing equations are then projected to the weighting spaces by multiplying weighting functions and integrating over the domain, which reformulates the problem 4.14 to: find $u(\xi) \in S$ and
\( p(\xi) \in \mathcal{Q} \) such that for all \( w \) and \( q \),

\[
\begin{align*}
    a(w, u; \gamma) + c(u, w, u; \gamma) + b(w, p; \gamma) &= (w, g_N), \\
    b(u, q; \gamma) &= 0
\end{align*}
\]  
(4.16)

where

\[
\begin{align*}
    a(w, u; \gamma) &= \int_{\Omega^*} \sum_{i=1}^{d} \sum_{j=1}^{d} \frac{\partial u}{\partial \xi_i} \kappa_{ij}(\xi; \gamma) \frac{\partial w}{\partial \xi_j} d\Omega^*, \\
    b(w, p; \gamma) &= -\int_{\Omega^*} q \sum_{i=1}^{d} \sum_{j=1}^{d} \zeta_{ij}(\xi; \gamma) \frac{\partial u_j}{\partial \xi_i} d\Omega^*, \\
    c(u, w, u; \gamma) &= \int_{\Omega^*} \sum_{i=1}^{d} \sum_{j=1}^{d} \sum_{k=1}^{d} u_i \zeta_{ij}(\xi; \gamma) \frac{\partial w_k}{\partial \xi_j} u_k d\Omega^*.
\end{align*}
\]

These are the bilinear and trilinear forms for diffusion term, pressure-divergence term and convection term of the system, respectively. The elements \( \kappa_{ij}(\xi; \mu) \) and \( \zeta_{ij}(x; \gamma) \) in the compact forms come from the tensor representation of functions associated with coordinate transformation,

\[
\begin{align*}
    \kappa(\xi, \gamma) &= \nu(J_X(\xi; \gamma))^{-1}(J_X(\xi; \gamma))^{-T} |J_X|, \\
    \zeta(\xi, \gamma) &= (J_X(\xi; \gamma))^{-1} |J_X|,
\end{align*}
\]  
(4.17a, 4.17b)

where \( J_X(\xi; \gamma) \) denotes Jacobian matrix of mapping \( X(\cdot; \gamma) \) and \( |J_X| \) is the determinant of the matrix.

Galerkin spatial discretisation approximates the solution by seeking solutions within corresponding finite dimensional subspace \( \mathcal{V}^h \subset \mathcal{V} \) and \( \mathcal{Q}^h \subset \mathcal{Q} \), and the problem is subsequently reformulated to: seek velocity field \( u^h \in \mathcal{S}^h \) and pressure \( p^h \in \mathcal{Q}^h \), such that for all \( (w^h, q^h) \in \mathcal{V}^h \times \mathcal{Q}^h \),

\[
\begin{align*}
    a(w^h, u^h; \gamma) + c(u^h, w^h, u^h; \gamma) + b(w^h, p^h; \gamma) &= (w^h, g_N^h), \\
    b(u^h, q^h; \gamma) &= 0.
\end{align*}
\]  
(4.18)

The discrete solution of finite element approximation to the problem can be obtained by solving the system above with the given boundary conditions.
Chapter 4

The computational accuracy of the discretised system derived is related to the
degrees of polynomial chosen for shape functions and the size of the elements.
In the remaining part of paper, we denotes the discrete solutions in general
of a FOM blood flow simulation with geometric parameters $\gamma$ on a reference
domain $\Omega^*$ as $u^h(\xi; \gamma)$, which are also known as snapshots in context of ROM.

4.2.3 Reduced order model

The application of ROM is motivated by reducing the high computational
cost of a FOM owing to a large number of degrees of freedom (DOF) involved
in a large-scale system. In other scenarios, e.g. UQ, design optimisation, or
reliability analysis where a considerable amount of evaluations are required,
ROM frequently serves as a surrogate model technique to make the evaluations
computationally tractable.

ROM seeks hidden low-dimensional patterns behind the FOM and ap-
proximates the FOM with high fidelity. Here ROM is applied to extract the
low-dimensional representation over time and geometric parameters and the
FOM solution therefore can be approximated by,

$$u^h(\xi; \gamma) \approx \hat{u}^h(\xi, \gamma) = \sum_{i=1}^{N_l} u^r_{i \gamma}(\gamma) \phi_i(\xi),$$  (4.19)

where $\{u^r_{i \gamma}(\gamma)\}_{i=1, \ldots, N_l}$ denotes the reduced coefficients which provide the
weighting of each reduced basis and $\phi_i(\xi)$ is the corresponding reduced basis.
The decomposition can be applied to any flow fields such as velocities in
different directions, pressure, and shear stress fields. We denote $N$ and $N_\gamma$ as
the number of nodes on a reference mesh and the number of snapshots available
(corresponding to geometric parameters), respectively in the ROM.

A non-intrusive ROM framework consists of two stages, offline and online.
The offline stage prepares the necessary ingredients for prediction, including
snapshot generation, construction of reduced bases with the dimensionality re-
duction method, and training an interpolator for predicting reduced coefficients
for each basis. During the online stage, the interpolator is applied to predict the
reduced coefficients according to new parameters, and together with the reduced
4.2. Method

bases, an approximation of the system can be subsequently reconstructed via equation 4.19. One of the widely applied reduced-order methods is POD [67]. The POD can be realized using PCA, or SVD. In this work, SVD is applied to construct reduced bases.

Proper orthogonal decomposition

Consider a snapshot matrix \( M \in \mathbb{R}^{N \times N_\gamma} \) consisting of snapshots with respect to corresponding geometric parameters,

\[
M = [u^h(\gamma_1)|u^h(\gamma_2)| \cdots |u^h(\gamma_{N_\gamma})].
\]  

(4.20)

By performing SVD, such a snapshot matrix can be written in a factorization form,

\[
M = Z \Sigma V^T,
\]  

(4.21)

where \( Z \) and \( V \) denote left and right orthonormal matrices. The diagonal matrix \( \Sigma \) consists of singular values \( \sigma_i \), where \( i = 1, ..., N_\gamma \) and \( \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_{N_\gamma} \geq 0 \). The objective of POD is to find out a set of orthogonal basis \( \Phi^* = \{\phi_1, \phi_2, \cdots, \phi_{N_l}\} \) from the space \( Z = \{\Phi \in \mathbb{R}^{N \times N_l} : \Phi^T \Phi = I\} \) containing all possible orthogonal bases, such that the projection residual is minimized:

\[
\min_{\Phi \in Z} \|M - \Phi \Phi^T M\|_F^2,
\]

(4.22)

where \( \|\cdot\|_F \) denotes the Frobenius norm. The Eckart-Young theorem [119] shows that the orthogonal bases constructed by the basis vectors \( \{z_i\}_{i=1}^{N_l} \) taken from the \( i \)th column of \( Z \) is the unique solution to the optimization problem. The cumulative energy captured by \( N_l \) number of reduced bases can be evaluated by the ratio between truncated and complete singular values,

\[
R_{en} = \frac{\|M - \tilde{\Phi} \tilde{\Phi}^T M\|_F^2}{\|M\|_F^2} = \frac{\sum_{i=1}^{N_l} \sigma_i^2}{\sum_{i=1}^{N_\gamma} \sigma_i^2}.
\]

(4.23)

The values of \( \sigma_i \) decay rapidly which means a small number of bases are sufficient to approximate the data well.
Algorithm 4.1: Offline stage of the surrogate model

**Surface registration**
1. Collect data of vessel shapes represented by a collection of boundary points \(\{\{x_{1,i}\}_{i=1}^{b_1}, \{x_{2,i}\}_{i=1}^{b_2}, \cdots, \{x_{N_s,i}\}_{i=1}^{b_{N_s}}\}\), where \(b_i\) denotes the number of the boundary points for each shape and \(N_s\) denotes the total number of shapes. For convenience, we denote the boundary points for each shape as \(X_j = \{x_{j,i}\}_{i=1}^{b_j}\).
2. Define a reference shape \(\{x_{*,i}\}_{i=1}^{b_*}\) or perform atlas regression based on the shape dataset.
3. Perform surface registration based on currents representation, and compute the corresponding diffeomorphism for each shape in the dataset, \(\{\varphi_1, \varphi_2, \cdots, \varphi_{N_s}\}\).

**Shape parameterisation**
1. Parameterise the shapes using displacement of boundary points \(\{\gamma_i\}_{i=1}^{N_s} = \{\varphi_i(X_i) - X_i\}_{i=1}^{N_s}\).

**Coordinate transformation**
1. Train RBF interpolators which approximate the mapping \(\mathcal{X}_i(\cdot, \gamma)\) between the reference domain and a target domain based on the input data \(X_i\) and output \(\varphi_i(X_i)\), for \(i = 1, 2, \cdots, N_s\).

**Model evaluation**
1. Compute the Jacobian matrix \(J_{\mathcal{X}_i}\) of \(\mathcal{X}_i(\cdot, \gamma)\) and its determinant \(|J_{\mathcal{X}_i}|\).
2. Solve Navier-Stokes equations for each shape on the reference domain \(\Omega^*\) and corresponding discrete solution \(u^h(\xi; \gamma)\) is obtained.

**Reduced-order modelling**
1. Construct the snapshot matrix \([u^h(\gamma_1) | \cdots | u^h(\gamma_{N_s})]\) and perform POD to obtain orthogonal reduced bases \(\Phi\). \(N_l\) is the number of bases that the cumulative energy \(\epsilon\) of which has reached 99.9%.
2. Project the output of training data to reduced bases and compute the reduced coefficients \(\{u_i^{rb}(\gamma)\}_{i=1}^{N_l}\).
3. Apply POD to the geometric parameters \(\gamma\) and obtain the reduced coefficients of geometric parameters \(\gamma^{rb}\) which can be applied to fully represent the geometric parameters.
4. Train a RBF interpolator with reduced coefficients of geometric parameters \(\gamma^{rb}\) and reduced coefficients of flow field \(u_i^{rb}(\gamma^{rb})\), where \(i = 1, \cdots, N_l\).
4.2. Method

Algorithm 4.2: Online stage of the surrogate model

**Surface registration**
1. Collect boundary data of a vessel for prediction, \( \hat{X} = \{\hat{x}_i\}_{i=1}^b \)
2. Perform surface registration based on currents representation, and compute the corresponding diffeomorphism for the shape, \( \hat{\phi} \).

**Shape parameterisation**
1. Parameterise the shapes using displacement of boundary points
   \[
   \hat{\gamma} = \hat{\phi}(\hat{X}) - \hat{X}. 
   \]
2. Project \( \hat{\gamma} \) to the reduced space constructed during the online stage and compute the reduced coefficients for the geometric parameters \( \hat{\gamma}^{rb} \).

**Coordinate transformation**
1. Train a RBF interpolator which approximate the mapping \( \hat{X}(\cdot, \hat{\gamma}) \) between the reference domain and a target domain based on the input data \( \hat{X} \) and output \( \hat{\phi}(\hat{X}) \).

**Prediction via ROM**
1. Apply trained RBF interpolator to predict the reduced coefficients \( \hat{u}^{rb}(\hat{\gamma}^{rb}) \).
2. Reconstruct the complete flow fields on the reference domain,
   \[
   \hat{u}^h(\xi; \gamma) = \sum_{i}^{N_i} \hat{u}_i^{rb}(\hat{\gamma}) \phi_i(\xi) \]
   and map it back to its original domain \( \hat{u}^h(\hat{X}^{-1}(\hat{x}); \hat{\gamma}) \).

Online-offline framework of ROM

In the ROM, the framework can be separated into two stages, online and offline. The offline stage prepares essential ingredients while the online stage makes the prediction. The offline stage mainly consists of surface registration, model evaluations on the reference domain and the construction of ROM, while the online stages extract the geometric parameters from a new shape and predict the flow profile based on the ROM. The details of the online-offline framework are outlined by algorithm 4.1 and algorithm 4.2.

The offline stage collects and prepares the essential ingredients for online prediction. Assume that we are interested in predicting the flow fields in different but similar shapes of the domains. These domains are usually segments of arteries generated either from the segmentation of medical image data or from a computational model. The shapes of the domains are assumed to be similar to
each other such that all the shapes can be considered as small deformations from a reference domain. For a computational model simulating the time-dependent tissue growth in the artery, the initial shape of the domain can be chosen as the reference shape. For more general situations, the reference shape can be chosen as the average of the shapes, namely the atlas construction [157, 158].

By performing the surface registration between the reference shape and target shapes, the diffeomorphisms between the shapes are constructed and subsequently applied in both shape parametrisation and mapping construction. The displacements of the CPs of each shape are viewed as the geometric parameters of the parametric PDE problem described in Section 4.2.2. However, such an approach to parametrise the shape may lead to the high dimensionality of parameters owing to the number of the CPs of a two- or three-dimensional shape. Besides, the displacements of CPs only represent local information and correlations can be found between the displacements. Therefore, the dimensionality reduction technique can be employed to further extract the latent low-dimensional representation of these parameters. In this work, POD is also applied to extract the reduced coefficients of the geometrical parameters.

On the other hand, the constructed diffeomorphism also provides a mapping between the reference domain and a target domain. Such a mapping can be approximated by RBF interpolation based on the position of the control points before and after the deformation. These mappings are further used to project the Navier-Stokes equations on the reference domain using the Jacobian matrix and its determinant, through which all the evaluations of the blood flow simulations in different shapes of the domain can be evaluated on the same mesh. This ensures the spatial compatibility over the snapshots corresponding to the geometric parameters for ROM. Therefore the reduced bases can be constructed by applying POD to the snapshot matrix, and the reduced coefficients of the flow field are considered as a function of the reduced coefficients of the geometrical parameters. In this work, RBF interpolation is also applied to approximate the function between the reduced coefficients of geometric parameters and flow fields.

During the online stage, reduced coefficients of geometric parameters of a new
shape can be again computed through surface registration and subsequently, used to estimate the reduced coefficients of corresponding flow fields. The estimated reduced coefficients together with the reduced bases derived in the offline stage reconstruct the entire flow field on the reference domain, which is mapped back to its original domain at the final step.

4.2.4 Error estimation

In order to present the accuracy and performance of the proposed ROM, several measurements were applied. The relative error was employed to visualise the relative difference between two flow fields at a specific point \( x \),

\[
e(\mathbf{x}) = \frac{|\mathbf{u}^h(x) - \hat{\mathbf{u}}^h(x)|}{\max\{|\mathbf{u}^h|\}}.
\]

(4.24)

where \( \max\{|\mathbf{u}^h|\} \) denotes the maximum absolute value of the vector \( \mathbf{u}^h \). Relative \( L^2 \) error corresponding to geometric parameters \( \gamma \) is employed to measure the overall difference of the estimations between a FOM and the proposed ROM over the domain,

\[
\epsilon_{rb}(\gamma) = \frac{\|\mathbf{u}^h(\gamma_i) - \hat{\mathbf{u}}^h(\gamma_i)\|_{L^2}}{\|\mathbf{u}^h(\gamma_i)\|_{L^2}} = \frac{\|\mathbf{u}^h(\gamma_i) - \tilde{\Phi}\mathbf{u}^{rb}(\gamma_i)\|_{L^2}}{\|\mathbf{u}^h(\gamma_i)\|_{L^2}}.
\]

(4.25)

where \( \| \cdot \|_{L^2} \) denotes \( L^2 \) norm and \( \mathbf{u}^{rb}(\gamma) = [u_1^{rb}(\gamma), \cdots, u_{N_l}^{rb}(\gamma)] \) denotes the vector consisting of the reduced coefficients. Besides, the projection error introduced by POD is also measured by the relative \( L^2 \) error,

\[
\epsilon_{POD}(\gamma) = \frac{\|\mathbf{u}^h(\gamma) - \tilde{\Phi}\tilde{\Phi}^T\mathbf{u}^h(\gamma)\|_{L^2}}{\|\mathbf{u}^h(\gamma)\|_{L^2}}.
\]

(4.26)

4.3 Numerical examples

4.3.1 Case one: blood flow in a stenosis

The simulation of blood flowing through a stenotic vessel is a common scenario in cardiovascular science to study the development of arteriosclerosis.
or in-stent restenosis. We constructed an idealised 2D example of blood flow passing through a stenotic coronary artery. The geometric configuration of the domain is shown in Figure 4.1(a). The position and severity of the stenosis of the upper and lower boundaries are described by the means and standard
deviations of two Gaussian functions,

\[ L_{\text{low}}(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right), \]

\[ L_{\text{up}}(x) = L_0 - \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right), \]

where \( L_0 = 2 \text{ mm} \) is the width of the inlet. 500 different shapes of domains were generated by sampling \( \sigma_1 \) and \( \sigma_2 \) within the ranges \([0.8, 2]\), \( \mu_1 \) and \( \mu_2 \) within the range \([0.3L_1, 0.7L_1]\) using the Latin hypercube method [159]. 400 of the samples were applied for the construction of the ROM in offline stage, while the left 100 samples were used for testing during the online stage. The dynamic viscosity and density of the flow were set to be \( 3.5 \times 10^{-3} \text{ g/(mm \cdot s)} \) and \( 1.06 \times 10^{-3} \text{ g/mm}^3 \). A constant parabolic velocity profile \( (u_{\text{inlet}} = 200 \text{ mm/s}) \) was prescribed at the inlet. A non-slip boundary condition was prescribed for the upper and lower boundaries, and a homogeneous Neumann condition was forced on the outlet. The simulation including mesh generation and finite element method was implemented using the open-source PDE solver FreeFEM [160]. A rectangular shape was chosen for the shape of the reference domain, as shown in Figure 4.1(b) and a Taylor-Hood P2-P1 finite element spatial discretisation was applied.

The surface registration was performed via statistical analysis software Deformetrica [161]. One of the surface registration results is demonstrated in Figure 4.2. The result shows that the reference shape deformed gradually toward the target shape and aligned well at the final time step. The average residual of the surface registration over 500 samples is \( 1.31 \times 10^{-7} \) (measured on currents). The results of surface registration were subsequently applied to compute the mappings between the domains of different shapes using RBF interpolation. A cubic kernel is employed with a polynomial function of the first order. Owing to the limited amount of mapping information we have, it is difficult to validate the interpolation itself. Instead, the validation was performed based on the results of the FEM simulation. The result computed on the reference domain was mapped back to its original domain and compared to the result directly
simulated on the target domain. The accuracy of the simulation result based on the reference domain reflects the accuracy of the mapping. Figure 4.3 shows the comparison of one sample. The regions with relatively higher errors are located close to the deformed boundaries, which means that RBF interpolation introduced minor errors into the snapshots. The average relative $L^2$ errors of velocities in x and y directions, and pressure of 400 training samples are 0.21%, 0.42% and 0.042%, respectively. It implies that the overall errors introduced by mapping are small, therefore the mapping approximated by the RBF interpolation is accurate enough for the FE simulation. We therefore considered these evaluations based on the reference domain as the snapshots for the reduced model.

Figure 4.3: A comparison of the flow fields solved on the reference domain and its original domain. a) Velocity fields and corresponding relative error in the x direction. b) Velocity field and corresponding relative error in the y direction.
4.3. **Numerical examples**

Figure 4.4: (a) Percentage of energy captured by the number of the reduced bases for the velocity field, pressure field and geometric parameters.

The surface registration results have also been applied to parameterise the configuration of the domains. In this case, there are 360 points on the boundary and therefore leads to 720 geometric parameters in a two-dimensional problem. POD was applied to further reduce the dimensionality of the parameters to 17, which has already captured 99.9% of the total variance. The percentage of the total variance captured by the number of reduced bases is demonstrated in Figure 4.4.

The discrete results of simulations on the reference domain are subsequently taken as the snapshots of ROM, and POD was performed to build reduced bases. The percentage of the total variance of the velocity and pressure data captured by the number of reduced bases is demonstrated in Figure 4.4. Similar to the dimensionality reduction to the geometric parameters, we assume that if more than 99.9% of the variance has been recovered by POD, the approximation reconstructed by the first $N_l$ bases performs well enough. More rigorous criteria could be set but the benefits are marginal as long as the main patterns have been already covered.

RBF interpolation again was utilised to predict reduced coefficients of flow fields based on reduced coefficients of the geometric parameters. A comparison of the expected and predicted reduced coefficients of one test case is shown in Figure 4.5. The points cluster around the diagonal line, indicating that the RBF interpolation predicted the reduced coefficients well. Figure 4.6 visualises the
velocity fields reconstructed from predicted reduced coefficients and compares the result to the evaluation of the corresponding FOM. Note that in this comparison, the evaluations of the FOM were based on the reference domain. The average relative $L^2$ errors over 100 test samples are 0.22%, 2.95% and 0.39% for velocity in x and y directions and pressure, respectively. The result shows that the prediction of the surrogate model matches the finite element simulation result truth well. The relative $L^2$ error is comparatively large in the y direction owing to its small magnitude compared to the velocity in the x direction.
Figure 4.6: A comparison between FE simulation results and ROM prediction for velocities and pressure and their corresponding relative errors.
4.3.2 Case two: blood flow in a bifurcation

We present the second example of blood flow simulations through an artery bifurcation, another common scenario in cardiovascular sciences. The basic configuration of the flow domain is shown in Figure 4.7, which is also considered as the reference shape. A series of new shapes are generated via deformations based on the basic configuration. The four red dots in Figure 4.7 are the chosen points to control the deformation. We consider the displacement of these points in the x and y directions as random variables and sample within the range $[-0.5, 0.5](\text{mm})$. The new shapes are therefore imposed by an interpolation. Similar to case one, 500 samples were generated by the Latin hypercube method and the same fluid properties and boundary conditions were applied. The Reynolds number for the case is 60.

The surface registration was performed to compute the geometric parameters and the corresponding mapping between the reference shape and other shapes. The average residual of the registration is $6.4 \times 10^{-6}$. There are 580 nodes on the boundary of the reference shapes which leads to 1160 geometric parameters in the two-dimensional case. Owing to the POD, the dimension of the geometric
4.3. Numerical examples

parameters was reduced to 8. The accuracy of the mapping is measured by the accuracy of the result solved on the reference domain. The average relative $L^2$ errors of velocities in x and y directions, and pressure of 400 training samples are 0.033%, 0.069% and 0.026%, respectively.

The bases of the ROM were subsequently constructed. The velocities in the x and y directions require 17 and 21 reduced bases to capture 99.9% of the total variance of the data, while pressure needs 8 reduced bases. The predictions of the ROM for one of the test cases are shown in Figure 4.8. The average $L^2$ errors of velocities in the x and y directions and pressure over 100 test cases are 0.70%, 1.53% and 2.65%, respectively.

![Figure 4.8](image-url)

Figure 4.8: A comparison between FE simulation results and ROM prediction for velocities and pressure and their corresponding relative errors.
4.4 Conclusion

In this chapter, a surrogate model for blood flow simulations in an unparameterised vessel lumen is presented. The surrogate model is constructed based on surface registration and non-intrusive reduced-order modelling. The surface registration with currents provides a mapping between reference and target coordinate systems and parameterises the shape without point-to-point correspondence. With the coordinate mapping, all the evaluations of FOM are performed on a reference domain which ensures the spatial compatibility of snapshots. The non-intrusive reduced order model is subsequently constructed using POD and the RBF interpolator is trained for predicting the reduced coefficients of ROM based on reduced coefficients of geometric parameters of the shape. Two examples of blood flow simulations based on stenosis vessels and bifurcation vessels are presented and discussed.

The result shows that the proposed surrogate model can predict the desired flow profile accurately with different geometries. However, the two presented examples are two-dimensional cases, which are not capable of demonstrating a significant speedup of the method compared to a traditional full-order model. The final objective of the research is to apply this surrogate model to a production model, for example, ISR3D with patient-specific cases and reduce the computational cost for evaluations. The method can be further leveraged for more demanding scenarios such as UQ and design optimisation.