



**HAL**  
open science

# An optimization-based registration approach to geometry reduction

Tommaso Taddei

► **To cite this version:**

Tommaso Taddei. An optimization-based registration approach to geometry reduction. 2023. hal-04173723

**HAL Id: hal-04173723**

**<https://hal.science/hal-04173723>**

Preprint submitted on 31 Jul 2023

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# An optimization-based registration approach to geometry reduction

Tommaso Taddei<sup>1</sup>

<sup>1</sup> *IMB, UMR 5251, Univ. Bordeaux; 33400, Talence, France. Inria Bordeaux Sud-Ouest, Team MEMPHIS; 33400, Talence, France,* tommaso.taddei@inria.fr

## Abstract

We develop and assess an optimization-based approach to parametric geometry reduction. Given a family of parametric domains, we aim to determine a parametric diffeomorphism  $\Phi$  that maps a fixed reference domain  $\Omega$  into each element of the family, for different values of the parameter; the ultimate goal of our study is to determine an effective tool for parametric projection-based model order reduction of partial differential equations in parametric geometries. For practical problems in engineering, explicit parameterizations of the geometry are likely unavailable: for this reason, our approach takes as inputs a reference mesh of  $\Omega$  and a point cloud  $\{y_i^{\text{raw}}\}_{i=1}^Q$  that belongs to the boundary of the target domain  $V$  and returns a bijection  $\Phi$  that approximately maps  $\Omega$  in  $V$ . We propose a two-step procedure: given the point clouds  $\{x_j\}_{j=1}^N \subset \partial\Omega$  and  $\{y_i^{\text{raw}}\}_{i=1}^Q \subset \partial V$ , we first resort to a point-set registration algorithm to determine the displacements  $\{v_j\}_{j=1}^N$  such that the deformed point cloud  $\{y_j := x_j + v_j\}_{j=1}^N$  approximates  $\partial V$ ; then, we solve a nonlinear non-convex optimization problem to build a mapping  $\Phi$  that is bijective from  $\Omega$  in  $\mathbb{R}^d$  and (approximately) satisfies  $\Phi(x_j) = y_j$  for  $j = 1, \dots, N$ . We present a rigorous mathematical analysis to justify our approach; we further present thorough numerical experiments to show the effectiveness of the proposed method.

## 1 Introduction

A broad range of problems in science and engineering requires parametric studies to assess the influence of geometry on the solution to a given partial differential equation (PDE). Given the family of parametric domains  $\{\Omega_\mu : \mu \in \mathcal{P}\}$  in  $\mathbb{R}^d$ , where  $\mu$  denotes a vector of parameters in the compact set  $\mathcal{P}$  in  $\mathbb{R}^p$ , we seek a reference domain  $\Omega$  and a parameterized mapping  $\Phi : \Omega \times \mathcal{P} \rightarrow \mathbb{R}^d$  such that  $\Phi_\mu(\Omega)$  approximates  $\Omega_\mu$  — in a sense to be defined — and is bijective in  $\Omega$  for all  $\mu \in \mathcal{P}$ . The aim of this work is to develop and analyze an optimization-based geometry registration and reduction (GRR) framework for the construction of the mapping  $\Phi$  and for its rapid evaluation for new configurations. In this work, we refer to the problem of determining a mapping from the reference domain  $\Omega$  to the target domain  $\Omega_\mu$  as *geometry registration*; on the other hand, we refer to the process of determining a low-rank representation of the parametric mapping  $\Phi$  as *geometry reduction*.

The ultimate goal of our study is parametric (projection-based) model order reduction (pMOR, [9, 27]) for PDEs in parametric geometries. Given the parametric PDE model  $\mathcal{L}_\mu(\mathfrak{u}_\mu) = 0$  defined over the domain  $\Omega_\mu$ , pMOR methods rely on a parametric mapping  $\Phi$  to recast the problem over a parameter-independent configuration  $\Omega$ ; then, pMOR methods resort to the projection of the PDE over suitable parameter-independent low-dimensional trial and test spaces to approximate the mapped field  $\mathfrak{u}_\mu \circ \Phi_\mu$ . As discussed in Remark 2.1, the formulation in the reference domain might be performed implicitly (discretize-then-map) or explicitly (map-then-discretize): for finite element (FE) discretizations, both approaches rely on the introduction of a mesh  $\mathcal{T}^{\text{hf}} = \left(\{x_i^{\text{hf}}\}_{i=1}^{N_{\text{hf}}}, \mathbf{T}\right)$  with nodes  $\{x_i^{\text{hf}}\}_{i=1}^{N_{\text{hf}}}$  and connectivity matrix  $\mathbf{T}$ . Despite the many contributions to the field [16, 17, 21, 22, 29], development of rapid and reliable geometry reduction techniques for large deformations is still a challenging task.

Effective GRR techniques should fulfill the following requirements.

- *Hidden parameterization.* In many applications (e.g., modeling of biological processes in patient-specific geometries), the parameterization is unknown and the only available information is a set of points  $\{y_i^{\text{raw}}\}_{i=1}^Q$  that belong to the boundary of  $\Omega_\mu$ . GRR algorithms should thus take as input the mesh  $\mathcal{T}^{\text{hf}}$  and the point cloud  $\{y_i^{\text{raw}}\}_{i=1}^Q \subset \partial\Omega_\mu$  and return the deformed mesh  $\Phi_\mu(\mathcal{T}^{\text{hf}})$  with deformed nodes  $\{\Phi_\mu(x_i^{\text{hf}})\}_{i=1}^{N_{\text{hf}}}$  and fixed connectivity matrix  $\mathbf{T}$ .
- *Real-time computations.* In many scenarios, the computation of the deformed mesh should be performed extremely rapidly, possibly with little computing resources. To achieve this goal, effective reduction strategies based on the low-rank approximation of the sought mapping should be developed.
- *Control of mesh quality and geometric error.* GRR methods should explicitly ensure that the geometric error — measured by the Hausdorff distance between  $\Phi_\mu(\Omega)$  and  $\Omega_\mu$  (see section 3) — and the quality of the deformed mesh meet user-defined tolerances.

In this work, we propose a two-step procedure: given the point clouds  $\{x_j\}_{j=1}^N \subset \partial\Omega$  and  $\{y_i^{\text{raw}}\}_{i=1}^Q \subset \partial\Omega_\mu$ , we first resort to a point-set registration (PSR, [10, 18, 19, 20, 40]) algorithm to determine the displacements  $\{v_j\}_{j=1}^N$  such that the deformed point cloud  $\{y_j := x_j + v_j\}_{j=1}^N$  approximates  $\partial\Omega_\mu$ ; then, we seek a mapping  $\Phi_\mu$  that is bijective in  $\Omega$  and (approximately) satisfies  $\Phi_\mu(x_j) = y_j$  for  $j = 1, \dots, N$ . To accomplish the first task, we rely on the coherent point drift (CPD, [25, 24]) method: CPD is a well-established non-rigid PSR technique that is broadly used for image processing and pattern recognition applications and is also the point of departure of several more recent methods. On the other hand, following [31], we recast the problem of finding the mapping  $\Phi$  as an optimization problem: the optimization framework allows us to directly control the geometry error and also the quality of the deformed mesh; furthermore, the spectacular advances in non-convex optimization enable the solution to highly-nonlinear and non-convex problems in a reasonable time frame. Finally, we resort to proper orthogonal decomposition (POD, [36]) to identify a low-rank approximation space for the mapping and ultimately speed up the registration process.

We present a rigorous mathematical analysis of the geometry error that provides a rigorous foundation for the optimization approach. In more detail, we show that for smooth domains, under mild assumptions on  $\Phi$  the Hausdorff distance between  $\Omega_\mu$  and  $\Phi_\mu(\Omega)$  is controlled by  $\max_{j=1, \dots, N} \|\Phi_\mu(x_j) - y_j\|_2$  (see Propositions 3.1 and 3.2); we also show that the same result does not hold for Lipschitz domains with corners. Moreover, we establish a connection between elasticity-based mesh morphing and registration methods (cf. Remark 2.2). Finally, we investigate the performance of several variants of the optimization statement considered in [31] for a number of two-dimensional model problems.

This work is also linked to optimization-based mesh morphing, which rely on the solution to a (convex) optimization problem to determine the deformed mesh (see the review [30]); here, we emphasize the application to parametric problems and the use of a global — as opposed to local, compactly-supported — basis for the mapping. Furthermore, our approach exploits the connection between GRR and the problem of point-set registration in bounded domains, which was previously considered in [12] for MOR applications.

We further remark that several authors have proposed mesh moving strategies for unsteady PDEs that rely on multiple incremental solutions to a suitable elasticity problem with mesh-dependent properties (see [34, 35] and references therein). In the parametric setting, we can interpret this strategy as feed-forward maps of the form  $\Phi = \Phi_N \circ \dots \circ \Phi_1$  where the  $i$ -th map  $\Phi_i$  maps  $\Omega_i$  into  $\Omega_{i+1}$  with  $\Omega_1 = \Omega$  and  $\Omega_{N+1} = \Omega_\mu$ . In this work, we consider instead linear approximations (cf. (13)) that are determined by solving a single highly-nonlinear optimization problem: this choice provides a natural framework for linear-subspace data compression methods (such as POD) and is thus well-suited for MOR applications.

The outline of the paper is as follows. In section 2, we present the methodology: first, we consider the problem of point-set registration in bounded domains; second, we present the extension to geometry registration; third, we comment on dimensionality reduction. In section 3, we present the analysis of the geometry error; in section 4, we present the results of the numerical experiments; finally, in section 5, we provide a summary of the contributions and of the numerical results. Several appendices conclude the paper.

## 2 Optimization-based registration

In section 2.1, we introduce relevant notation and preliminary definitions that are useful for the subsequent discussion. Then, in section 2.2, we discuss the problem of point-set registration in bounded domains: we provide a complete discussion for tensorized domains and we comment on the extension to more general domains. In section 2.3, we discuss the problem of geometry registration and reduction. In the remainder, the spatial dimension  $d$  is either  $d = 2$  or  $d = 3$ .

### 2.1 Notation and preliminary definitions

We denote by  $x$  a generic point of the domain  $\Omega$ , and by  $\mathbf{n}(x)$  the outward normal to  $\Omega$  at  $x \in \partial\Omega$ . We define the identity map  $\text{id} : \Omega \rightarrow \Omega$  such that  $\text{id}(x) = x$ ; then, given the mapping  $\Phi : \Omega \rightarrow \mathbb{R}^d$ , we denote by  $u = \Phi - \text{id}$  the displacement field and by  $J(\Phi) = \det(\nabla\Phi)$  the Jacobian determinant. We denote by  $\mathcal{T}_{\text{hf}}$  a finite element (FE) mesh of  $\Omega$  with  $N_e$  elements  $\{\mathbb{D}_k\}_{k=1}^{N_e}$ ; we define the reference element  $\widehat{\mathbb{D}} = \{\tilde{x} \in (0, 1)^d : \sum_{i=1}^d (\tilde{x})_i < 1\}$  and the shape functions  $\{\ell_i\}_{i=1}^{n_{\text{lp}}}$  of the polynomial space  $\mathbb{P}_p(\widehat{\mathbb{D}})$  associated with the nodes  $\{\tilde{x}_i\}_{i=1}^{n_{\text{lp}}}$ . Then, we represent the elemental mappings  $\Psi_k^{\text{hf}}$  from  $\widehat{\mathbb{D}}$  to  $\mathbb{D}_k$  for  $k = 1, \dots, N_e$  as

$$\Psi_k^{\text{hf}}(\tilde{x}) = \sum_{i=1}^{n_{\text{lp}}} x_{i,k}^{\text{hf}} \ell_i(\tilde{x}), \quad (1)$$

where  $\{x_{i,k}^{\text{hf}} := \Psi_k^{\text{hf}}(\tilde{x}_i) : i = 1, \dots, n_{\text{lp}}, k = 1, \dots, N_e\}$  are the nodes of the mesh. To simplify the presentation, we here consider the same basis for both FE fields and FE elemental mappings (isoparametric FE); the discussion below can be trivially extended to sub- and sup-parametric discretizations. We say that the mapping  $\Phi$  is

bijective if  $J(\Phi)$  is strictly positive in  $\Omega$ ; on the other hand, we say that  $\Phi$  is discretely bijective with respect to the mesh  $\mathcal{T}_{\text{hf}}$  if the elemental mappings

$$\Psi_{\Phi,k}^{\text{hf}}(\tilde{x}) = \sum_{i=1}^{n_{\text{lp}}} \Phi(x_{i,k}^{\text{hf}}) \ell_i(\tilde{x}) \quad (2)$$

are all bijective. Note that the function  $\Phi : \Omega \rightarrow \mathbb{R}^d$  is bijective with respect to  $\mathcal{T}_{\text{hf}}$  if  $\Phi_{\text{hf}} : \Omega \rightarrow \mathbb{R}^d$  given by

$$\Phi_{\text{hf}}(x) = \Psi_{\Phi,k}^{\text{hf}}\left(\left(\Psi_k^{\text{hf}}\right)^{-1}(x)\right), \quad \forall x \in \mathcal{D}_k, \quad k = 1, \dots, N_e, \quad (3)$$

is a bijection. Bijectivity and discrete bijectivity are two independent conditions, which are both important for model reduction of parametric systems in parameterized geometries as discussed in the following remark.

**Remark 2.1.** *To clarify the distinction between bijectivity and discrete bijectivity, consider the problem of approximating the solution to the following Laplace equation in the family parameterized domains  $\{\Omega_\mu : \mu \in \mathcal{P}\}$ :*

$$-\Delta U_\mu = f_\mu \text{ in } \Omega_\mu, \quad U_\mu|_{\partial\Omega_\mu} = 0, \quad (4)$$

for some  $f_\mu \in H^{-1}(\Omega_\mu)$ . Given the reference domain  $\Omega$ , we denote by  $\Phi : \Omega \times \mathcal{P} \rightarrow \mathbb{R}^d$  a mapping such that  $\Phi_\mu(\Omega) = \Omega_\mu$  for all  $\mu \in \mathcal{P}$ , by  $\mathcal{T}_{\text{hf}}$  a FE mesh of  $\Omega$ , and by  $\mathcal{X}_{\text{hf}} \subset H_0^1(\Omega)$  the corresponding FE space. To devise a FE discretization of (4), we first observe that the mapped field  $\tilde{U}_\mu := U_\mu \circ \Phi_\mu$  satisfies

$$\int_{\Omega} J(\Phi_\mu) \left( \nabla \Phi_\mu^{-1} (\nabla \Phi_\mu)^{-*} \nabla \tilde{U}_\mu \cdot \nabla v - f_\mu \circ \Phi_\mu v \right) dx = 0 \quad \forall v \in H_0^1(\Omega), \quad (5)$$

where  $\nabla \Phi_\mu^*$  denotes the transpose of  $\nabla \Phi_\mu$ ; then, we proceed to discretize (5) by projecting the equations over the FE space  $\mathcal{X}_{\text{hf}}$ . Alternatively, we might consider the formulation: find  $U_\mu^{\text{hf}} \in \mathcal{X}_{\text{hf},\mu}$  such that

$$\sum_{k=1}^{N_e} \int_{\mathcal{D}_{\mu,k}} (\nabla U_\mu^{\text{hf}} \cdot \nabla v - f_\mu v) dx = 0 \quad \forall v \in \mathcal{X}_{\text{hf},\mu}, \quad (6)$$

where  $\mathcal{X}_{\text{hf},\mu}$  is the FE space associated with the deformed FE mesh  $\Phi_\mu(\mathcal{T}_{\text{hf}})$ , while  $\mathcal{D}_{\mu,k} = \{\Psi_{\mu,k}^{\text{hf}}(\tilde{x}) : \tilde{x} \in \widehat{\mathcal{D}}\}$  denotes the  $k$ -th deformed element of the mesh. Clearly, (5) and (6) are equivalent if and only if  $\Phi = \Phi_{\text{hf}}$  (see (3)). Both approaches (5) and (6) are broadly used for the treatment of parameterized geometries in *pMOR*: following [32], we refer to (5) as to *map-then-discretize* approach and to (6) as to *discretize-then-map* approach. Note that (5) requires the bijectivity of the mapping  $\Phi$ , while (6) requires the discrete bijectivity with respect to the mesh  $\mathcal{T}_{\text{hf}}$ .

## 2.2 Point-set registration in bounded domains

### 2.2.1 Optimization statement

Given the domain  $\Omega \subset \mathbb{R}^d$ , and the point clouds  $\{x_i\}_{i=1}^N$  and  $\{y_i\}_{i=1}^N$ , we consider the problem of finding a mapping  $\Phi = \text{id} + u$  such that  $\Phi(x_i) = y_i$  for  $i = 1, \dots, N$  and  $\Phi$  is bijective from  $\Omega$  in itself. Towards this end, we define the operator  $B : [C(\Omega)]^d \rightarrow \mathbb{R}^{d \cdot N}$  and the vector  $\mathbf{z}$  such that

$$B\Phi = \begin{bmatrix} (\Phi(x_1))_1 \\ \vdots \\ (\Phi(x_N))_1 \\ \vdots \\ (\Phi(x_N))_d \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} (y_1)_1 \\ \vdots \\ (y_1)_d \\ \vdots \\ (y_N)_d \end{bmatrix}. \quad (7)$$

Then, we can state the constrained optimization statement as follows: given (i) the set of vector-valued functions  $\mathcal{X}$ , (ii) the objective function  $f^{\text{obj}} : \mathcal{X} \rightarrow \mathbb{R}$ , we seek  $\Phi \in \mathcal{X}$  to minimize

$$\min_{\Phi \in \mathcal{W}} f^{\text{obj}}(\Phi) \quad \text{s.t.} \quad B\Phi = \mathbf{z}. \quad (8)$$

Clearly, the performance of registration relies on the choice of  $\mathcal{W}$ ,  $f^{\text{obj}}$  and also on the optimization algorithm that is employed to determine local minima of (8). We address these points in the next sections.

In many contexts, the hard constraint  $B\Phi = \mathbf{z}$  is either inappropriate (e.g., if the point clouds are subject to noise) or might lead to an ill-posed statement (e.g., if two points are coincident). Exploiting classical theory in inverse problems, we consider two regularization strategies for (8):

$$\min_{\Phi \in \mathcal{W}} \xi \mathfrak{f}^{\text{obj}}(\Phi) + \frac{1}{2} \|B\Phi - \mathbf{z}\|_2^2, \quad (9)$$

and

$$\min_{\Phi \in \mathcal{W}} \mathfrak{f}^{\text{obj}}(\Phi) \quad \text{s.t.} \quad \|B\Phi - \mathbf{z}\|_\infty \leq \delta. \quad (10)$$

Following taxonomy from the inverse problem literature [13], we refer to (9) as to Tykhonov-regularized registration statement and to (10) as to Morozov-regularized registration statement. We observe that both approaches rely on the introduction of an hyper-parameter —  $\xi$  and  $\delta$ , respectively.

Morozov regularization allows to directly control the geometry error through the choice of the parameter  $\delta$ ; if  $\mathcal{W}$  is a linear space, the choice of the norm  $\|\cdot\|_\infty$  in (10) leads to an optimization statement with linear inequality constraints, which can be effectively tackled using interior point methods. Tykhonov regularization does not enable direct control of the geometry error and requires a careful selection of the parameter  $\xi$ ; however, the resulting optimization problem is unconstrained and, in our experience, can be solved more efficiently than (10) using quasi-Newton methods, particularly for the choice of the squared 2-norm in (9).

### 2.2.2 Choice of the objective function and of the search space $\mathcal{W}$ for tensorized domains

The choice of the space  $\mathcal{W}$  and of the objective function  $\mathfrak{f}^{\text{obj}}$  should ensure that (i)  $B$  is a bounded functional over  $\mathcal{W}$ , (ii) any minimizer  $\Phi$  of the problem (9) or (10) should be a bijection in  $\Omega$  and/or should be discretely bijective with respect to the target mesh  $\mathcal{T}_{\text{hf}}$ , and (iii) the space  $\mathcal{W}$  is well-suited to fit a broad range of bijections in  $\Omega$  (*high expressive power*). In this section, we discuss the choice of  $\mathcal{W}$  and  $\mathfrak{f}^{\text{obj}}$  for tensorized domains. Without loss of generality, we consider  $\Omega = (0, 1)^d$ .

If  $\Omega$  is a tensorized domain, we can prove that  $\Phi = \text{id} + u$  is bijective in  $\Omega$  if  $u \cdot \mathbf{n}|_{\partial\Omega} = 0$  and  $J(\Phi)$  is strictly positive in  $\bar{\Omega}$  (cf. Proposition 2.3 in [31]). We can thus choose the affine space  $\mathcal{W}$  as

$$\mathcal{W} = \{\text{id} + \varphi : \varphi \in [H^2(\Omega)]^d, \varphi \cdot \mathbf{n}|_{\partial\Omega} = 0\}. \quad (11)$$

Note that  $\mathcal{W}$  is a linear affine space of continuous functions for  $d \leq 3$ ; more precisely, in two and three dimensions  $H^2$  is the Sobolev space of minimal regularity that is contained in the space of continuous functions [1]. For practical computations, we replace the infinite-dimensional space  $H^2(\Omega)$  in (11) with a finite-dimensional tensorized polynomial space of dimension  $(n_{\text{lp}} + 1)^d$ , that is  $\mathcal{W}_{\text{hf}} = \text{id} + \mathcal{U}_{\text{hf}}$  with

$$\mathcal{U}_{\text{hf}} = \{\varphi \in [\mathbb{Q}_{n_{\text{lp}}}]^d : \varphi \cdot \mathbf{n}|_{\partial\Omega} = 0\}, \quad \mathbb{Q}_{n_{\text{lp}}} = \text{span} \left\{ \varphi(x) = \prod_{i=1}^d \ell_i((x)_i) : \ell_1, \dots, \ell_d \in \mathbb{P}_{n_{\text{lp}}}(\mathbb{R}) \right\}, \quad (12)$$

where  $\mathbb{P}_{n_{\text{lp}}}(\mathbb{R})$  denotes the space of one-dimensional polynomials of degree up to  $n_{\text{lp}}$ . We observe that  $\mathcal{W}_{\text{hf}}$  is dense in  $\mathcal{W}$  in the limit  $n_{\text{lp}} \rightarrow \infty$ ; we further observe that the space  $\mathcal{U}_{\text{hf}}$  is of dimension  $M = d(n_{\text{lp}} + 1)^d - 2d(n_{\text{lp}} + 1)^{d-1}$ . Given the orthonormal basis of  $\mathcal{U}_{\text{hf}}$   $\{\psi_m\}_{m=1}^M$ , we can introduce the operator  $\mathbb{N} : \mathbb{R}^M \rightarrow \mathcal{W}_{\text{hf}}$  such that

$$\mathbb{N}(x; \mathbf{a}) := x + \sum_{m=1}^M (\mathbf{a})_m \psi_m(x); \quad (13)$$

note that  $\mathbb{N}$  is a linear affine operator and  $\mathbb{N}(\cdot; \mathbf{0}) = \text{id}$ .

We propose two different objective functions: the former is designed to enforce bijectivity of  $\Phi$ , while the latter is designed to enforce discrete bijectivity. In both cases, we add the regularization term

$$\mathfrak{R}(\Phi) = \frac{1}{2|\Omega|} |\Phi|_{H^2(\Omega)}^2 = \frac{1}{2|\Omega|} \int_{\Omega} \sum_{i,j,k=1}^d (\partial_{i,j}(\Phi)_k)^2 dx, \quad (14)$$

which is designed to promote smoothness: the regularization  $\mathfrak{R}$  (14) ensures that the  $H^2$  norm of any global minimum of (9) or (10) is uniformly bounded in  $H^2$  for  $M \rightarrow \infty$ . Given  $\epsilon, C_{\text{exp}} > 0$ , we define the objective function  $\mathfrak{f}^{\text{obj}} = \mathfrak{f}_{\text{jac}} + \mathfrak{R}$  such that

$$\mathfrak{f}_{\text{jac}}(\Phi) = \frac{1}{|\Omega|} \int_{\Omega} \exp\left(\frac{\epsilon - J(\Phi)}{C_{\text{exp}}}\right) dx. \quad (15)$$

The parameter  $\epsilon$  provides a weak lower bound for the value of the Jacobian determinant of the mapping  $\Phi$ : provided that  $e^{\epsilon/C_{\text{exp}}} \gg 1$ , the objective function (15) forces the optimizer to ensure the condition  $J(\Phi) \gtrsim \epsilon$

everywhere in  $\Omega$ ; the  $H^2$  penalization enforces smoothness of  $\Phi$  and is also instrumental to ensure bijectivity (see [31, section 2.2]). Given the constant  $\kappa_{\text{msh}}$ , following [15] (see also [39]), we define the objective function  $\mathfrak{f}^{\text{obj}} = \mathfrak{f}_{\text{msh}} + \mathfrak{P}$  such that

$$\mathfrak{f}_{\text{msh}}(\Phi) = \frac{1}{|\Omega|} \sum_{k=1}^{N_e} |\mathbb{D}_k| \int_{\mathbb{D}} \exp(\kappa_{\text{msh}} - q_{\Phi,k}) d\tilde{x}, \quad q_{\Phi,k} := \frac{1}{d^2} \left( \frac{\|\nabla \Psi_{\Phi,k}^{\text{hf}}\|_{\mathbb{F}}^2}{(\det(\nabla \Psi_{\Phi,k}^{\text{hf}}))_+^{2/d}} \right)^2, \quad (16)$$

where  $\Psi_{\Phi,k}^{\text{hf}}$  is defined in (2),  $\|\cdot\|_{\mathbb{F}}$  is the Frobenius norm and  $(\cdot)_+ = \max(\cdot, 0)$ . The ratio  $q_{\Phi,k}$  measures the degree of anisotropy of the mesh: it can thus be interpreted as a measure of the quality of the deformed mesh. Note that if  $\Psi_{\Phi,k}$  is perfectly isotropic,  $q_{\Phi,k} = 1$ . Similarly to (15),  $\mathfrak{f}^{\text{obj}} = \mathfrak{f}_{\text{msh}} + \mathfrak{P}$  strongly penalizes deformations for which  $q_{\Phi,k}$  exceeds  $\kappa_{\text{msh}}$  for some element  $\mathbb{D}_k$  of the mesh, and promotes smoothness of the mapping in regions where  $q_{\Phi,k} < \kappa_{\text{msh}}$ .

We observe that (15) and to (16) depend on several hyper-parameters —  $\epsilon$ ,  $C_{\text{exp}}$  and  $\kappa_{\text{msh}}$ . In our numerical results, we investigate the sensitivity with respect to  $\epsilon$  and  $\kappa_{\text{msh}}$ ; on the other hand, we set  $C_{\text{exp}} = 0.025\epsilon$ . We further observe that neither (15) nor (16) are convex: therefore, we do not expect that problems (9) and (10) admit a unique solution; we provide extensive numerical investigations of the registration approach based on the two choices of the objective function in section 4. In the remainder, we refer to (15) and to (16) as to “exp-jac” objective and “exp-mesh” objective, respectively. Finally, we observe that (14), (15) and (16) — and thus also the objective function  $\mathfrak{f}^{\text{obj}}$  — are invariant under translations and rotations, that is

$$\mathfrak{P}(\mathbf{b} + \mathbf{R}\Phi) = \mathfrak{P}(\Phi), \quad \mathfrak{f}_{\text{jac}}(\mathbf{b} + \mathbf{R}\Phi) = \mathfrak{f}_{\text{jac}}(\Phi), \quad \mathfrak{f}_{\text{msh}}(\mathbf{b} + \mathbf{R}\Phi) = \mathfrak{f}_{\text{msh}}(\Phi),$$

for any  $\mathbf{b} \in \mathbb{R}^d$  and any rotation  $\mathbf{R} \in \mathbb{R}^{d \times d}$ .

**Remark 2.2.** *We might attempt to establish a direct link between optimization-based registration and elasticity-based mesh morphing. We might indeed formally consider the statement*

$$\min_{\Phi \in \mathcal{W}} \mathfrak{f}^{\text{el}}(\Phi) \quad \text{s.t.} \quad B\Phi = \mathbf{z}, \quad (17)$$

where  $\mathfrak{f}^{\text{el}}(\Phi) = \int_{\Omega} \psi^{\text{se}}(x; \Phi) dx$  is the integral over  $\Omega$  of the strain energy  $\psi^{\text{se}}$ . To provide concrete references that are considered in the numerical experiments, we might consider the linear isotropic model:

$$\psi^{\text{se}}(\cdot; \Phi = \mathbf{id} + u) = \lambda_1 (\nabla \cdot u)^2 + 2\lambda_2 \left\| \frac{1}{2} (\nabla u + (\nabla u)^*) \right\|_{\mathbb{F}}^2, \quad (18)$$

or the nonlinear Neo-Hookean model (for  $d = 2$ ):

$$\psi^{\text{se}}(\cdot; \Phi) = \frac{1}{2} \lambda_2 \|\nabla \Phi\|_{\mathbb{F}}^2 - \lambda_2 \log(J(\Phi)) + \lambda_1 (\log(J(\Phi)))^2, \quad (19)$$

where  $\lambda_1, \lambda_2$  are the Lamé constants.

We observe that the models (18) and (19) are naturally defined in  $H^1(\Omega)$ , which is not contained in the space of continuum functions in  $\Omega$  for  $d \geq 1$ . On the other hand, the objective functions (15) and (16) involve derivatives of the deformation gradient  $\nabla \Phi$ : therefore, they cannot be interpreted as strain energies of a suitable hyper-elastic material.

### 2.2.3 Generalization to arbitrary domains

The extension of the registration approach to non-tensorized domains is particularly challenging due to the difficulty to ensure the condition  $\Phi(\Omega) = \Omega$ . From an algorithmic standpoint, we should identify a new operator  $\mathbb{N}$  (cf. (13)) and a new penalization term  $\mathfrak{P}$  (cf. (14)). A potential strategy based on a spectral element approximation to deal with this case is proposed in [33]. In this paper, we do not address this issue.

### 2.2.4 Generalization to unsorted point clouds

In many applications, the reference and target point clouds  $\{x_i\}_{i=1}^N$  and  $\{y_i^{\text{raw}}\}_{i=1}^Q$  might be of different cardinality (i.e.,  $N \neq Q$ ) and/or might not be properly sorted. We are hence in need of a point-set registration (PSR) algorithm that aligns the two point clouds, that is, it finds  $\{v_i\}_{i=1}^N$  such that  $\{y_i := x_i + v_i\}_{i=1}^N$  approximates — in a sense to be defined — the target point cloud  $\{y_i^{\text{raw}}\}_{i=1}^Q$ . If we define the matrices  $\mathbf{X} = [x_1, \dots, x_N]^* \in \mathbb{R}^{N \times d}$ ,  $\mathbf{Y}^{\text{raw}} = [y_1^{\text{raw}}, \dots, y_Q^{\text{raw}}]^* \in \mathbb{R}^{Q \times d}$  and  $\mathbf{Y} = [x_1 + v_1, \dots, x_N + v_N]^* \in \mathbb{R}^{N \times d}$ , we can formalize an abstract PSR algorithm as

$$\mathbf{Y} = \text{PSR}(\mathbf{X}, \mathbf{Y}^{\text{raw}}).$$

The literature on PSR is extremely vast and a thorough review of the subject is beyond the scope of the present work. In this paper, we rely on the coherent point drift (CPD, [25, 24]) algorithm: CPD is a well-established non-rigid PSR technique that is broadly used and is also the point of departure of several more recent methods.

The CPD method seeks mappings of the form

$$\Phi^{\text{cpd}}(x) = x + \sum_{i=1}^N w_i \phi(\|x - x_i\|_2), \quad \text{with } \phi(r) = \exp\left(-\frac{1}{2\beta^2}r^2\right). \quad (20)$$

where  $\mathbf{W} = [w_1, \dots, w_N]^* \in \mathbb{R}^{N \times d}$  is a matrix of coefficients that is chosen through an iterative procedure; therefore, the CPD method seeks mappings in the affine space

$$\mathcal{W}_{\text{hf}}^{\text{cpd}} = \text{id} + \mathcal{U}_{\text{hf}}^{\text{cpd}}, \quad \text{with } \mathcal{U}_{\text{hf}}^{\text{cpd}} = \text{span}\{\phi(\|\cdot - x_i\|_2)\mathbf{e}_j : i = 1, \dots, N, j = 1, \dots, d\}, \quad (21)$$

where  $\mathbf{e}_1, \dots, \mathbf{e}_d$  are the elements of the canonical basis of  $\mathbb{R}^d$ . Since the CPD methodology is at this stage well-established, we postpone the detailed description to Appendix A. Here, we anticipate that the method can naturally cope with low-dimensional search spaces  $\mathcal{U}_M^{\text{cpd}} \subset H^2(\mathbb{R}^d)$ : we discuss in section 2.3.2 the importance of this observation.

## 2.3 Geometry registration and reduction

### 2.3.1 Optimization statement

We adapt the framework of section 2.2 to the problem of geometry registration: given the reference domain  $\Omega$  and the point cloud  $\{y_i^{\text{raw}}\}_{i=1}^M$  that belongs to the boundary of the domain  $V$ , we seek a bijective mapping  $\Phi$  from  $\Omega$  to  $V$ . This task is propedeutic to the problem of geometry reduction, which is addressed in the next section. In the remainder, we denote by  $\{x_i\}_{i=1}^N$  the discretization of the boundary  $\partial\Omega$  that is used for registration; we also denote by  $\mathcal{T}_{\text{hf}}$  a FE mesh in  $\Omega$ .

We introduce the hyper-rectangle  $\Omega_{\text{box}}$  that strictly contains  $\Omega$  and the point cloud  $\{y_i^{\text{raw}}\}_i$ ; we define the affine space  $\mathcal{W}_{\text{hf}} = \text{id} + \mathcal{U}_{\text{hf}}$  with  $\mathcal{U}_{\text{hf}} = [\mathbb{Q}_{n_{\text{ip}}}]^d$ ; given  $\Phi \in \mathcal{W}_{\text{hf}}$ , we further define the counterparts of the quantities in section 2.2.2  $\mathfrak{F}(\Phi) = \frac{1}{2|\Omega_{\text{box}}|} |\Phi|_{H^2(\Omega_{\text{box}})}^2$  (cf. (14)),  $\mathfrak{f}_{\text{jac}}(\Phi) = \frac{1}{|\Omega_{\text{box}}|} \int_{\Omega_{\text{box}}} \exp\left(\frac{\epsilon - J(\Phi)}{C_{\text{exp}}}\right) dx$  (cf. (15)), and  $\mathfrak{f}_{\text{msh}}(\Phi)$  as in (16). Then, we consider the following two-step registration procedure: first, we resort to PSR to determine the deformed points  $\{y_i = x_i + v_i\}_{i=1}^N$  (cf. section 2.2.4); second, we solve a registration problem in  $\Omega_{\text{box}}$  with inputs  $\{x_i\}_{i=1}^N$  and  $\{y_i\}_{i=1}^N$ . Towards this end, we consider the Tykhonov-regularized registration statement

$$\min_{\Phi \in \mathcal{W}_{\text{hf}}} \xi \mathfrak{f}^{\text{obj}}(\Phi) + \frac{1}{2} \|B\Phi - \mathbf{z}\|_2^2; \quad (22)$$

and the Morozov-regularized registration statement

$$\min_{\Phi \in \mathcal{W}_{\text{hf}}} \mathfrak{f}^{\text{obj}}(\Phi) \quad \text{s.t.} \quad \|B\Phi - \mathbf{z}\|_{\infty} \leq \delta. \quad (23)$$

The operator  $B$  and the vector  $\mathbf{z}$  are defined as in (7); as in section 2.2.2, the objective  $\mathfrak{f}^{\text{obj}}$  is given by  $\mathfrak{f}^{\text{obj}} = \mathfrak{f}_{\text{jac}} + \mathfrak{F}$  (“exp-jac”) or by  $\mathfrak{f}^{\text{obj}} = \mathfrak{f}_{\text{msh}} + \mathfrak{F}$  (“exp-mesh”).

In view of the numerical assessment, we further introduce the formulation

$$\min_{\Phi \in \mathcal{W}_{\text{hf}}} \frac{1}{2} \|B\Phi - \mathbf{z}\|_2^2 + \xi \mathfrak{F}(\Phi) \quad \text{s.t.} \quad \mathfrak{f}_{\star}(\Phi) - \delta_{\text{con}} \leq 0, \quad (24)$$

which was considered in [31]. Statement (24) reads as a nonlinear constrained optimization problem with quadratic objective function and nonlinear non-convex inequality constraint. We envision that the advantage of (24) is that it allows to control more explicitly the minimum of the Jacobian determinant (for  $\star = \text{jac}$ ) or the mesh quality (for  $\star = \text{msh}$ ); on the other hand, it does not explicitly control the geometry error (unlike (23)) and it involves a nonlinear constraint. We compare performance of (24) with (22) and (23) in the numerical experiments.

We here consider a polynomial space  $\mathcal{W}_{\text{hf}}$  in (22), (23), (24). This choice is motivated by the need to compute second-order derivatives for the regularization term and by the strong approximation properties of polynomials in moderate dimensions. A thorough assessment of other approximation classes is beyond the scope of the present work.

We remark that the goal of geometry registration is to find a mapping  $\Phi$  such that  $\Phi(\Omega)$  is close (in the sense of Hausdorff) to the domain  $V$ . Statements (22) and (23) (24) control the error  $B\Phi - \mathbf{z}$  in a convenient norm; in section 3, we discuss under what conditions the control of the difference  $B\Phi - \mathbf{z}$  ensures appropriate reconstruction of  $V$ .

### 2.3.2 Dimensionality reduction

Given the results of  $n_{\text{train}}$  geometry registration problems  $\{\Phi_i = \text{id} + u_i\}_{i=1}^{n_{\text{train}}}$  for different configurations, we might apply dimensionality reduction techniques to identify a low-dimensional representation of the mapping. The objective of dimensionality reduction is twofold: first, by reducing the number of unknowns, we might significantly speed up computations; second, dimensionality reduction techniques might also act as low-pass filters to avoid overfitting, especially for noisy datasets.

In this work, we rely on proper orthogonal decomposition (POD, [36]) based on the  $H^2(\Omega_{\text{box}})$  inner product to determine a  $M$ -dimensional reduced space  $\mathcal{U}_M = \text{span}\{\psi_m\}_{m=1}^M \subset \mathcal{U}_{\text{hf}}$ . We first assemble the Gramian matrix  $\mathbf{C} \in \mathbb{R}^{n_{\text{train}} \times n_{\text{train}}}$  such that  $(\mathbf{C})_{i,j} = (u_i, u_j)_{H^2(\Omega_{\text{box}})}$  for  $i, j = 1, \dots, n_{\text{train}}$ ; then, we compute the eigenpairs of  $\mathbf{C}$   $\{(\lambda_i, \psi_i)\}_i$  such that  $\lambda_1 \geq \dots \geq \lambda_{n_{\text{train}}} \geq 0$ . Finally, we define the POD space

$$\psi_m = \frac{1}{\sqrt{\lambda_m}} \sum_{i=1}^{n_{\text{train}}} u_i (\psi_m)_i, \quad m = 1, \dots, M.$$

The size  $M$  of the reduced space is chosen according to the criterion

$$M := \min \{m : \text{EC}_m(\boldsymbol{\lambda}) \geq 1 - \text{tol}_{\text{pod}}\}, \quad \text{where } \text{EC}_m(\boldsymbol{\lambda}) = \left( \sum_{j=1}^{n_{\text{train}}} \lambda_j \right)^{-1} \sum_{i=1}^m \lambda_i. \quad (25)$$

In the remainder,  $\text{EC}_m(\boldsymbol{\lambda})$  is referred to as the relative energy content of the POD space  $\mathcal{U}_M$ .

We observe that the cost of solving the optimization problems (22), (23), (24) with reduced search space  $\text{id} + \mathcal{U}_M$  might still be significant due to the need to compute the function (16) over all elements of the mesh. To address this issue, we might extend reduced quadrature techniques developed in the pMOR framework[5, 38] to the registration framework. This extension is beyond the scope of the present paper.

As discussed above, we first resort to PSR to determine the sorted deformed points  $\{y_i\}_{i=1}^N$  and then to one of the statements (22), (23), (24) to obtain the bijective mapping. So far, we proposed to apply POD to reduce the cost of the second step of the procedure. Even if the latter dominates by far computational costs, since the problem of interest is inherently non-convex, the difference in the ansatz spaces employed by CPD (cf. (21)) and by registration might lead to inaccurate performance. To investigate this issue, in the numerical results we consider two distinct strategies: **(i)** apply CPD with  $\mathcal{W}_{\text{hf}}^{\text{cpd}}$ ; **(ii)** apply CPD with

$$\mathcal{W}_M^{\text{cpd}} = \text{id} + \text{span}\{\psi_m\}_{m=1}^M. \quad (26)$$

The second strategy requires to slightly modify the original CPD formulation that are described in Appendix A.

## 3 Analysis

We denote by  $U$  and  $V$  two Lipschitz domains that are isomorphic to the unit ball and are compactly embedded in the hyper-cube  $\Omega_{\text{box}}$ . We denote by  $\text{dist}_{\text{H}}(U, V)$  the Hausdorff distance between  $U$  and  $V$  such that

$$\text{dist}_{\text{H}}(U, V) = \max \left\{ \sup_{x \in U} \text{dist}(x, V), \sup_{x \in V} \text{dist}(x, U) \right\}, \quad (27)$$

where  $\text{dist}(x, \omega) = \inf_{y \in \omega} \|x - y\|_2$  for any  $x \in \mathbb{R}^d$  and any measurable set  $\omega \subset \mathbb{R}^d$ . Given the target domain  $V$  and the reference domain  $\Omega$ , the goal of geometry registration algorithms is to determine a bijective mapping  $\Phi$  from  $\Omega$  to  $\mathbb{R}^d$  such that the Hausdorff distance  $\text{dist}_{\text{H}}(\Phi(\Omega), V)$  is below a given tolerance.

In order to bridge the gap between the target (27) and the computational procedure introduced in this paper, we introduce the non-symmetric boundary distance:

$$\text{dist}_{\text{bnd}}(U; V) = \sup_{x \in \partial U} \text{dist}(x, \partial V). \quad (28)$$

Next Proposition clarifies the link between (28) and the geometry constraint in (22) and (23); the proof is contained in Appendix B.

**Proposition 3.1.** *Let  $\{x_i\}_{i=1}^N \subset \partial U$  be an  $\epsilon$ -cover of  $\partial \Omega$ , that is  $\sup_{x \in \partial \Omega} \text{dist}(x, \{x_i\}_i) = \epsilon$ . Assume that  $\Phi : \Omega \rightarrow \mathbb{R}^d$  is a Lipschitz bijective map with Lipschitz constant  $K$  and let  $\{y_i\}_{i=1}^N \subset \partial V$ . Then, we have*

$$\text{dist}_{\text{bnd}}(\Phi(U); V) \leq \max_{i=1, \dots, N} \|\Phi(x_i) - y_i\|_2 + K\epsilon. \quad (29)$$

Note that  $\max_{i=1, \dots, N} \|\Phi(x_i) - y_i\|_2 \leq \sqrt{d} \max_{i=1, \dots, N} \|\Phi(x_i) - y_i\|_{\infty} = \sqrt{d} \|B\Phi - \mathbf{z}\|_{\infty}$ , which is the constraint of the optimization statement (23).

The objective of this section is to establish conditions for which (28) is equal to (27). Towards this end, we first present a number of definitions and preliminary results (cf. section 3.1), then we present the main proposition for smooth domains (cf. section 3.2) and we show through the vehicle of a two-dimensional example that (27) and (28) cannot be equal for Lipschitz domains with corners (cf. section 3.3). Proofs of the Lemmas below are contained in Appendix B.

### 3.1 Preliminary results and definitions

Given  $\delta > 0$ , we define the  $\delta$ -neighborhood of  $\partial U$  as

$$\text{Neigh}_\delta(\partial U) := \{x \in \Omega_{\text{box}} : \text{dist}(x, \partial U) < \delta\}, \quad (30)$$

and the tubular neighborhood of  $\partial U$  as

$$\text{Neigh}_\delta^t(\partial U) := \{y \in \mathbb{R}^d : y = x + t\mathbf{n}(x), |t| < \delta, x \in \partial U\}. \quad (31)$$

Next Lemma summarizes important properties of  $\text{Neigh}_\delta(\partial U)$ ,  $\text{Neigh}_\delta^t(\partial U)$  that will be used below — we remark that the second statement of Lemma 3.1 is an immediate consequence of the Weyl’s tube formula[7]. Given the set  $A$ , we denote by  $|A|_{(d-1)}$  (resp.,  $|A|_{(d)}$ ) the  $(d-1)$ -dimensional (resp.,  $d$ -dimensional) measure of  $A$  — to fix ideas, if  $d = 2$ ,  $|A|_{(1)}$  is the length of the curve  $A$  embedded in  $\mathbb{R}^2$ , while  $|A|_{(2)}$  is the area of the region  $A$ . If not specified otherwise, we shall assume that  $\delta$  is small enough so that  $\text{Neigh}_\delta(\partial U)$  is compactly embedded in  $\Omega_{\text{box}}$ .

**Lemma 3.1.** *Let  $U \subset \Omega_{\text{box}}$  be a  $C^1$  domain isomorphic to the unit ball. Then, (i)  $\text{Neigh}_\delta(\partial U) = \text{Neigh}_\delta^t(\partial U)$ , and (ii)*

$$|\text{Neigh}_\delta(\partial U)|_{(d)} \leq \begin{cases} 2\delta|\partial U|_{(d-1)} & \text{if } d = 2, \\ 2\delta|\partial U|_{(d-1)} + \frac{8\pi^2}{3}\delta^3 & \text{if } d = 3. \end{cases} \quad (32)$$

We introduce a special class of domains.

**Definition 3.1.** *We say that  $U$  is  $\delta$ -regular if for any given  $x \in \partial U$  any set  $V$  diffeomorphic to the unit ball such that  $\partial V \subset \text{Neigh}_\delta(\partial U) \setminus \mathcal{B}_\delta(x)$  is contained in  $\text{Neigh}_\delta(\partial U)$ .*

Figure 1 provides an interpretation of Definition 3.1: for a circle  $U = \mathcal{B}_1(0)$  (cf. Figure 1(a)), we find by inspection that any  $V$  such that  $\partial V \subset \text{Neigh}_\delta(\partial U) \setminus \mathcal{B}_\delta(x)$  for some  $x \in \partial U$  must be contained in the annulus  $\mathcal{B}_{1+\delta}(0) \setminus \mathcal{B}_{1-\delta}(0)$ , provided that  $0 < \delta < 1$ . For any fixed  $\delta > 0$ , we might construct smooth domains with bounded curvature that are not  $\delta$ -regular (cf. example in Figure 1(b)). The example in Figure 1(b) also shows that  $\delta$ -regularity is a global property of  $\partial U$  in the sense that does not uniquely depend on local properties (e.g., curvature, Lipschitz constant) of the boundary.

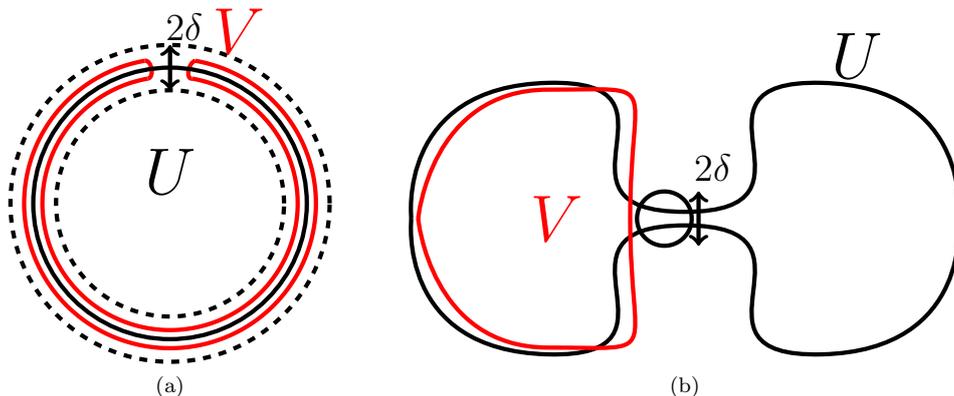


Figure 1: geometric interpretation of Definition 3.1. Dashed lines in (a) denote the boundary of  $\text{Neigh}_\delta(\partial U)$ .

Next three Lemmas offer further interpretations and properties of  $\delta$ -regular domains: Lemma 3.2 provides a sufficient condition for  $\delta$ -regularity; Lemma 3.3 shows that smooth domains are  $\delta$ -regular for sufficiently small values of  $\delta$ ; Lemma 3.4 shows an interesting property of  $\delta$ -regular domains that is exploited below. As currently stated, Lemma 3.3 does not provide an explicit estimate of  $\delta_0$ ; in Appendix B we provide a more detailed expression for  $\delta_0$ . We anticipate that  $\delta_0$  depends on the maximum principal curvature over  $\partial U$  and also on a global property of the boundary.

**Lemma 3.2.** *Let  $U$  be diffeomorphic to the unit ball. Assume that the hyper-surfaces  $\{x \pm \delta \mathbf{n}(x) : x \in \partial U\}$  do not have self intersections. Then,  $U$  is  $\delta$ -regular.*

**Lemma 3.3.** *Let  $U$  be a  $C^2$  domain isomorphic to the unit ball. Then, there exists  $\delta_0 > 0$  such that  $U$  is  $\delta$ -regular for any  $\delta < \delta_0$ .*

**Lemma 3.4.** *Let  $U, V$  be Lipschitz domains that are isomorphic to the unit ball. Assume that (i)  $V$  is  $\delta$ -regular for some  $\delta > 0$ ; (ii)  $|U|_{(d)} > |\text{Neigh}_\delta(\partial V)|_{(d)}$ ; (iii)  $\text{dist}_{\text{bnd}}(U, V) = \max_{x \in \partial U} \text{dist}(x, \partial V) \leq \delta$ . Then,  $\text{dist}_{\text{H}}(\partial U, \partial V) \leq \delta$  and  $\text{dist}_{\text{H}}(U, V) \leq \delta$ .*

### 3.2 Equivalence of (27) and (28) for smooth domains

Proposition 3.2 contains the result that is relevant for our discussion.

**Proposition 3.2.** *Let  $\Phi : \Omega_{\text{box}} \rightarrow \mathbb{R}^d$  be a bijection in  $\Omega_{\text{box}}$  and let  $\Omega, V \subset \Omega_{\text{box}}$  be diffeomorphic to the unit ball. Define  $\epsilon := \min_{x \in \bar{\Omega}} \det(\nabla \Phi)$ ,  $\delta := \text{dist}_{\text{bnd}}(\Phi(\Omega); V)$ . Assume that (i)  $\epsilon |\Omega|_{(d)} > |\text{Neigh}_\delta(\partial V)|_{(d)}$  and (ii)  $V$  is  $\delta$ -regular. Then, we have  $\text{dist}_{\text{H}}(\Phi(\Omega), V) \leq \delta$ ,  $\text{dist}_{\text{H}}(\partial \Phi(\Omega), \partial V) \leq \delta$ .*

*Proof.* We define  $U = \Phi(\Omega)$ . Clearly,  $U$  is isomorphic to the unit ball; furthermore, using the change-of-variable formula and the definition of  $\epsilon$ , we obtain

$$|U|_{(d)} = \int_{\Omega} \det(\nabla \Phi) dx \geq \epsilon |\Omega|_{(d)} > |\text{Neigh}_\delta(\partial V)|_{(d)}.$$

Using the definition (28) and the identity  $\partial \Phi(\Omega) = \Phi(\partial \Omega)$ , which is valid for any bijection  $\Phi$ , we also find

$$\max_{x \in \partial U} \text{dist}(x, \partial V) = \max_{x \in \partial \Omega} \text{dist}(\Phi(x), \partial V) \leq \delta.$$

Therefore,  $U, V$  satisfy the hypotheses of Lemma 3.4: we conclude that  $\text{dist}_{\text{H}}(U, V), \text{dist}_{\text{H}}(\partial U, \partial V) \leq \delta$ .  $\square$

Proposition 3.2 — together with Proposition 3.1 — provides a rigorous justification of the geometry registration strategy proposed in section 2. We observe that the value  $\epsilon$  in Proposition 3.2 is weakly controlled by the exp-jac objective function (15). In practice, we expect  $\epsilon$  to be significantly larger than the geometric error  $\delta$  at the boundary: therefore, the condition  $\epsilon |\Omega|_{(d)} \geq |\text{Neigh}_\delta(\partial V)|_{(d)}$  should in practice be easy to enforce, even for slender domains with small  $\frac{|V|_{(d)}}{|\partial V|_{(d-1)}}$ .

On the other hand, we cannot establish a rigorous connection between the exp-mesh objective function (16) and the quantity  $\epsilon$  in Proposition 3.2. In this respect, we recall that in the discretize-then-map framework (cf. Remark 2.1) the mapping to be considered in Proposition 3.2 for the mesh  $\mathcal{T}_{\text{hf}}$  is  $\Phi_{\text{hf}}$  (cf. (3)): for purely isotropic P1 FE discretizations, it is easy to verify that  $\nabla \Phi_{\text{hf}}|_{D_k} = \frac{|D_{\Phi, k}|}{|D_k|} \mathbb{1}$ , while the ratio  $q_{\Phi, k}$  in (16) is equal to one. We conclude that the solution to the registration problem with objective given by (16) might lead to extremely small elements (see Figure 3 in the numerical example of section 4.1), which might prevent the application of Proposition 3.2. If large deformations are expected in the proximity of the boundary, it might thus be necessary to combine (15) and (16).

### 3.3 Strict inequality for Lipschitz domains

The analysis of the previous section cannot be readily extended to Lipschitz domains with corners. To investigate the problem, we shall consider the domain depicted in Figure 2: we are here interested in the neighborhood of the vertex  $x^*$ ; we assume that  $V$  is of class  $C^2$  elsewhere. We denote by  $2\alpha \in [0, \pi]$  the angle associated with the corner at  $x^*$ ; the case  $2\alpha \in [\pi, 2\pi]$  is analogous.

We first consider the domains  $U$  and  $V$  in Figure 2(a): the dashed lines denote the boundary of  $\text{Neigh}_\delta(\partial V)$ . Clearly,  $U$  belongs to  $\text{Neigh}_\delta(\partial U)$ ; the point  $x^*$  satisfies  $\text{dist}(x^*, \partial U) = \frac{\delta}{\sin(\alpha)} =: r_\delta > \delta$ . The example shows that, for a domain  $V$  with a corner of angle  $2\alpha$  we might construct a domain  $U$  such that  $U$  is isomorphic to  $\mathcal{B}_1(0)$ ,  $|U|_{(d)} > 2\delta |\partial V|_{(d-1)}$ ,  $\partial U \subset \text{Neigh}_\delta(\partial V)$ , and  $\max_{x \in \partial V} \text{dist}(x, U) = r_\delta > \delta$ .

We seek an upper bound for  $\max_{x \in \partial V} \text{dist}(x, U)$  among all domains  $U$  satisfying (i)  $U$  is isomorphic to the unit ball  $\mathcal{B}_1(0)$ , (ii)  $|U|_{(d)} > 2\delta |\partial V|_{(d-1)}$ , (iii)  $\partial U \subset \text{Neigh}_\delta(\partial V)$ , and (iv)  $x^* \in \partial U$ . Exploiting the same argument as in Lemma 3.3 (cf. Appendix B), we find that there exists  $\delta_0 > 0$  such that the curves  $\gamma_\delta^\pm = \gamma \pm \delta \mathbf{n}$  do not have self-intersections in  $\partial V \setminus \mathcal{B}_{r_\delta}(x^*)$  for all  $\delta \leq \delta_0$ . Therefore, we have

$$\max_{x \in \partial V \setminus \mathcal{B}_{r_\delta}(x^*)} \text{dist}(x, \partial U) \leq \delta,$$

for any  $U$  that satisfies (i)-(iv). On the other hand, in the neighborhood of  $x^*$ , we can show that the domain  $U$  depicted in Figure 2(b) maximizes  $\max_{x \in \partial V \cap \mathcal{B}_{r_\delta}(x^*)} \text{dist}(x, \partial U)$ . Note that  $U$  is the limit of domains isomorphic to the unit ball; therefore, computation of the distance of  $\partial V$  from  $U$  provides an upper bound for

$\max_{x \in \partial V \setminus \mathcal{B}_{r_\delta}(x^*)} \text{dist}(x, \partial U)$  among Lipschitz domains  $U$  that satisfy (i)-(iv). Exploiting well-known trigonometric identities, we find that the point  $x(t)$  in Figure 2(b) satisfies

$$\|x(t) - A^*\|_2 = \sqrt{(r_\delta - t)^2 + (\tan(\alpha)t)^2} = \delta \sqrt{(1/\sin(\alpha) - t')^2 + (\tan(\alpha)t')^2}, \quad t' = \frac{t}{\delta},$$

and

$$\|x(t) - x^*\|_2 = \delta \frac{t'}{\cos(\alpha)}, \quad \|x(t) - B^*\|_2 = (1 + 2 \sin(\alpha)t') \delta.$$

In conclusion, we obtain

$$\max_{x \in \partial U} \text{dist}(x, \partial V) = \max_{t \in [0, \frac{\delta}{\sin(\alpha)} - \delta \sin(\alpha)]} \text{dist}(x(t), \partial U) = \mathcal{C}(\alpha) \delta, \quad (33a)$$

where

$$\mathcal{C}(\alpha) = \max_{t' \in [0, \frac{1}{\sin(\alpha)} - \sin(\alpha)]} \min \left\{ \sqrt{(1/\sin(\alpha) - t')^2 + (\tan(\alpha)t')^2}, \frac{t'}{\cos(\alpha)}, 1 + 2 \sin(\alpha)t' \right\}. \quad (33b)$$

In Figure 2(c), we show the behavior of  $\mathcal{C}(\alpha)$  with respect to  $\alpha$  and we compare it with  $\frac{1}{\sin(\alpha)}$ . Note that  $\mathcal{C}(\alpha) \leq \min\{3, \frac{1}{\sin(\alpha)}\}$  for all  $\alpha \in [0, \pi/2]$ ;  $\mathcal{C}$  is monotonic decreasing with  $\alpha$ ; and  $\mathcal{C}(\alpha = \pi/4) = 1$ . The assumption  $x^* \in \partial U$  leads to a reduction of the worst-case scenario for  $\max_{x \in \partial V} \text{dist}(x, \partial U)$  (and thus for  $\text{dist}_H(\partial V, \partial U)$ ) but does not guarantee that  $\max_{x \in \partial V} \text{dist}(x, \partial U) = \delta$  asymptotically. We note, however, that, while the domain  $U$  depicted in Figure 2(a) is regular, the domain in Figure 2(b) is not Lipschitz and thus the upper bound in (33) cannot be achieved.

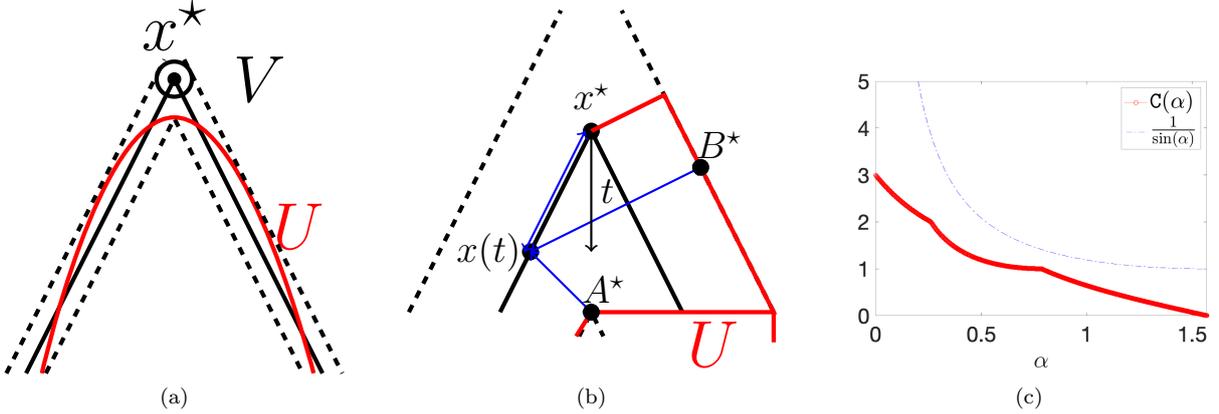


Figure 2: analysis for Lipschitz domains. (a) example of domains  $U, V$  isomorphic to  $\mathcal{B}_1(0)$  with  $|U|_{(2)} > 2\delta|\partial V|_{(1)}$  and  $\partial U \subset \text{Neigh}_\delta(\partial V)$  such that  $\text{dist}_H(\partial U, \partial V) = \frac{\delta}{\sin(\alpha)}$ . (b) domain  $U$  satisfying (i)-(iv) such that  $\text{dist}_H(\partial U, \partial V) = \mathcal{C}(\alpha)\delta$ . (c) behavior of  $\frac{1}{\sin(\alpha)}$  and  $\mathcal{C}(\alpha)$  with respect to  $\alpha \in [0, \pi/2]$ .

The analysis of this section shows that the approximation of slender bodies and/or of bodies with cusps and corners is challenging for the optimization-based registration procedure discussed in this paper. However, if we are able to ensure appropriate approximation — via interpolation — of the target shape at corners and cusps, we can recover near-optimal bounds for the geometry error.

## 4 Numerical results

We present several numerical experiments for a number of two-dimensional model problems. In order to assess the performance of a given mapping  $\Phi$ , we report the behavior of the minimum value  $q_{\min}$  of the radius ratio<sup>1</sup> over the elements of the deformed target mesh  $\mathcal{T}_{\text{hf}}$ . We also report the minimum value of the Jacobian determinant over the domain  $\bar{\Omega}$ ,  $J_{\min} = \min_{x \in \bar{\Omega}} J(\Phi)(x)$ . All numerical simulations are performed in Matlab 2020b on a commodity laptop.

<sup>1</sup>The radius ratio is the ratio between the radius of the circle inscribed in the triangle and the radius of the circle circumscribed around the triangle.

## 4.1 Three-point registration

Given  $\Omega = (0, 1)^2$ , we consider the problem of deforming the points  $x_1 = (1/2, 1/2)$ ,  $x_2 = (1/4, 1/4)$ ,  $x_3 = (3/4, 1/4)$  into the points  $y_1 = (1/4, 3/4)$ ,  $y_2 = (1/16, 1/16)$ ,  $y_3 = (1/2, 1/4)$ . The aim of this test is to illustrate the impact of the objective function on performance. Towards this end, we consider the Morozov-regularized formulation (10) with objective function equal to  $f^{\text{obj}} = \mathfrak{F}$  (dubbed “H2” below),  $f^{\text{obj}} = f_{\text{jac}} + \mathfrak{F}$  (dubbed “exp-jac”),  $f^{\text{obj}} = f_{\text{msh}} + \mathfrak{F}$  (dubbed “exp-mesh”) — we refer to section 2.2.2 for the definitions. We also investigated other objective functions including the ones associated with the linear elasticity strain energy (18) and with the neohookean strain energy (19): we provide a representative test for these objective functions in Appendix C.

Figure 3(a) shows the reference and deformed points and the mesh  $\mathcal{T}_{\text{hf}}$  that is used to assess performance. We resort to the Matlab function `fmincon` based on the interior-point method; in all our tests, we consider the initial condition  $\Phi = \text{id}$ . Figures 3(b) and (c) show the deformed mesh obtained using the exp-jac objective and the exp-mesh objective, respectively. We consider  $\delta = 10^{-6}$  in (10), and we set  $\epsilon = 0.1$  for the exp-jac objective, and  $\kappa_{\text{msh}} = 10$  for the exp-mesh objective. Note that the two deformed meshes significantly differ from each other in the proximity of the origin.

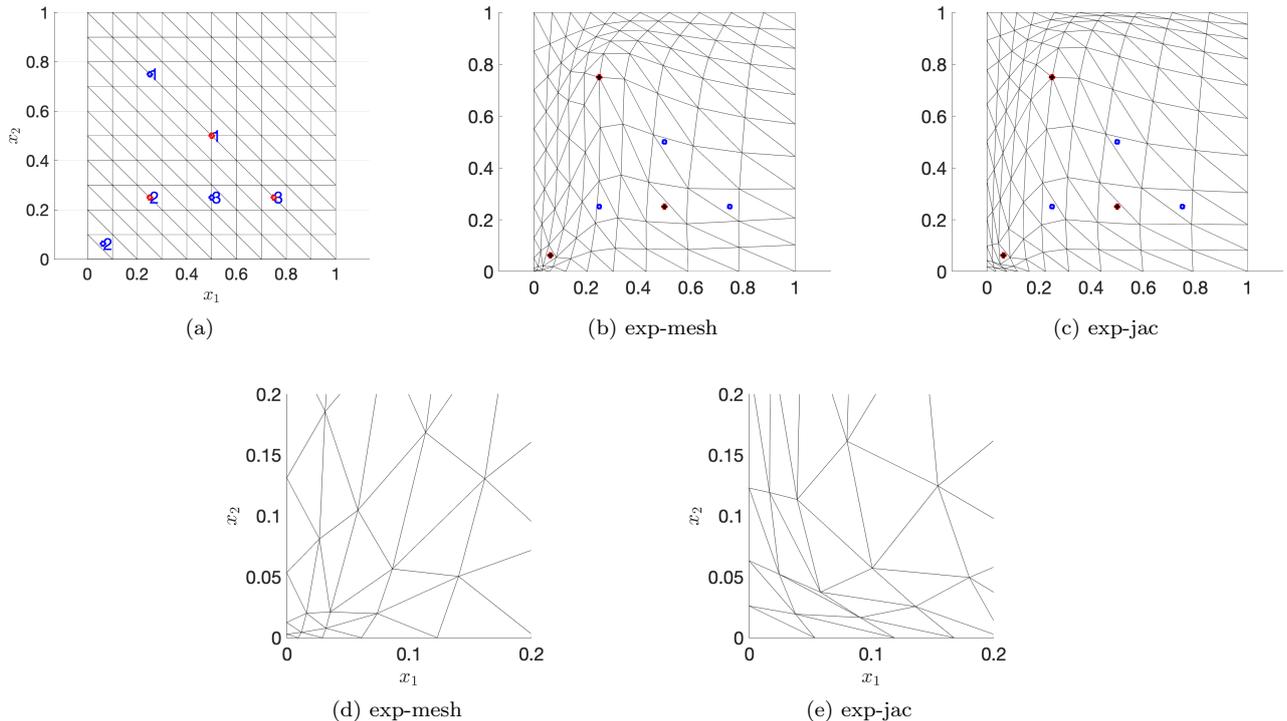


Figure 3: three-point registration problem; visualization of the reference mesh and of the deformed mesh for two different registration algorithms ( $\delta = 10^{-6}$ ,  $\epsilon = 0.1$ ,  $\kappa_{\text{msh}} = 10$ ). Figures (d)-(e) show a zoom of the deformed meshes in the proximity of the origin.

Figure 4 shows the performance of the Morozov-regularized statement (10) for the objective functions H2, exp-jac and exp-mesh, for three choices of  $\delta$  and for the deformation points  $\{y_i(t) = (1 - t)x_i + ty_i\}_{i=1}^3$  for  $t \in [0, 1]$ . The quadratic objective function  $H^2$  fails to deliver a proper mesh for  $t > 0.6$  — i.e., the deformed mesh contains inverted elements. On the other hand, the two nonlinear approaches are able to deliver bijective maps that are also bijective with respect to the target mesh  $\mathcal{T}_{\text{hf}}$ .

Figure 5 compares the performance of registration based on the “full” polynomial space (cf. (12)) with the performance of registration based on the “potential space”  $\mathcal{W}_{\text{hf}}^{\text{pot}} = \{\text{id} + \nabla\phi : \phi \in \mathcal{Q}_{n_{\text{lp}}}, \partial_n\phi|_{\partial\Omega} = 0\}$ . We recall that the condition  $\det(\nabla\Phi) > 0$  for  $\Phi = \text{id} + \nabla\phi \in \mathcal{W}_{\text{hf}}^{\text{pot}}$  is equivalent to the convexity of  $\phi$ . The interest for this particular choice of the search space is that there exist several effective convexification procedures, which have been recently developed for optimal transport problems (see, e.g., [14]), that might pave the way for effective registration strategies. To ensure the fairness of the test, we choose two different polynomial degrees in order to have trial spaces of comparable dimension —  $M = 1150$  for the “full” space and  $M = 1155$  for the “potential” space. We observe that the potential approximation significantly deteriorates the performance of registration for large deformations.

Figure 6 shows the sensitivity of the algorithm with respect to the choice of the polynomial order  $n_{\text{lp}}$  in (12). We here consider the exp-mesh objective function with  $\kappa_{\text{msh}} = 10$ . We observe that for this particular test case results are nearly independent of  $n_{\text{lp}}$  for  $n_{\text{lp}} \gtrsim 10$ .

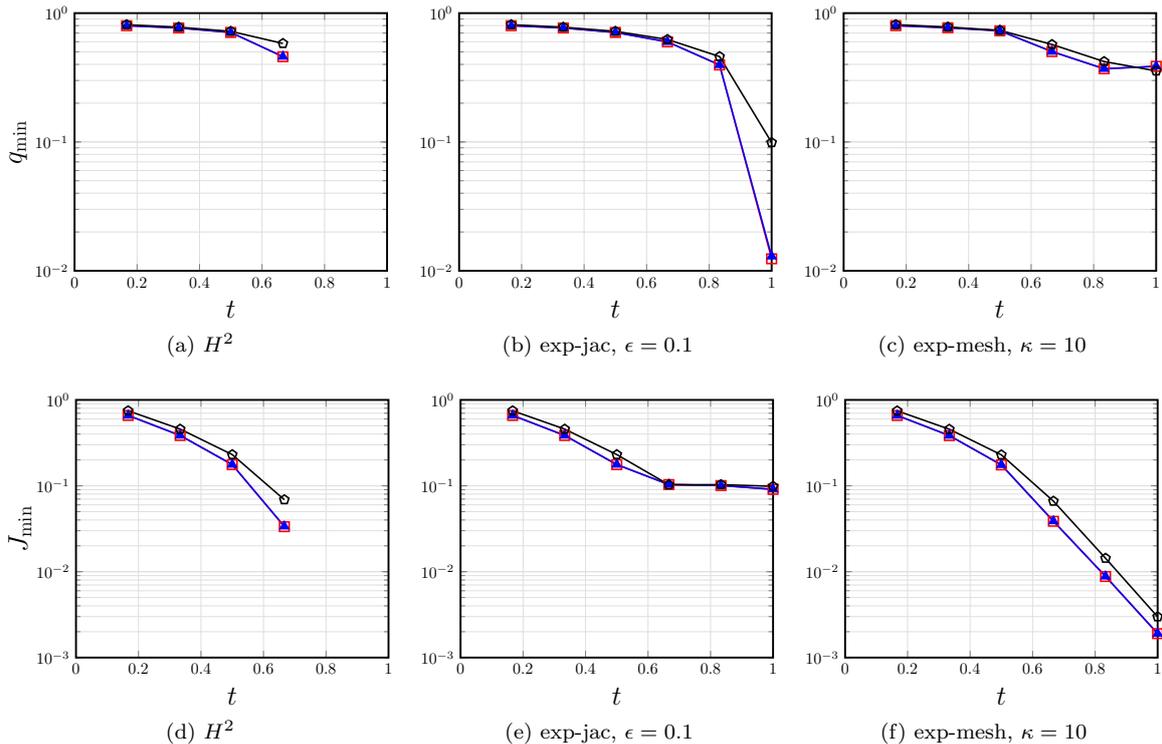


Figure 4: three-point registration problem. Performance of the Morozov-regularized statement (10) for the objective functions  $H^2$ , exp-jac and exp-mesh, for three choices of  $\delta$  and for the deformation points  $\{y_i(t) = (1-t)x_i + ty_i\}_{i=1}^3$  for  $t \in [0, 1]$ .  $\delta = 10^{-6}$   $\square$ ;  $\delta = 10^{-4}$   $\blacktriangle$ ,  $\delta = 10^{-2}$   $\blacklozenge$ .

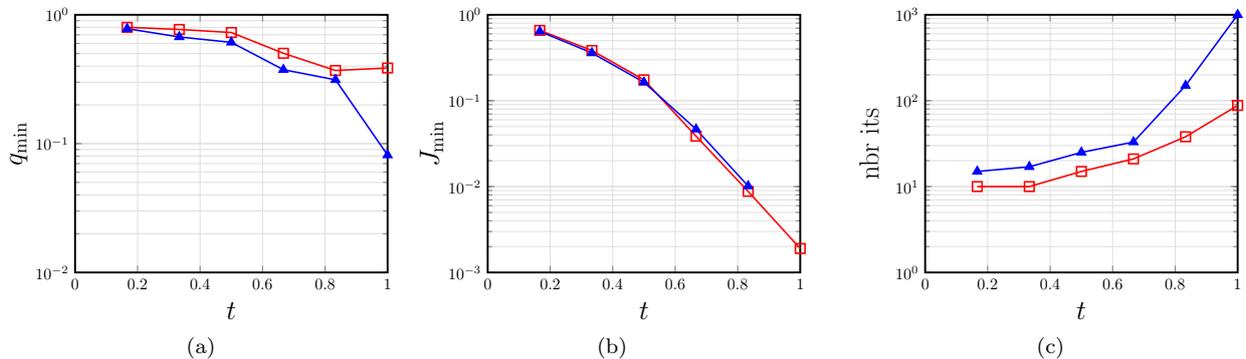


Figure 5: three-point registration problem. Comparison of performance for “full” space and “potential” space. Objective function exp-mesh,  $\kappa_{\text{msh}} = 10$ ,  $\delta = 10^{-6}$  ( full basis  $\square$ ; potential  $\blacktriangle$ ).

## 4.2 Large deformation mesh morphing

We consider the domain  $\Omega = (0, 1)^2 \setminus (0.4, 0.6)^2$  in Figure 7(a) and we consider a P2 FE triangulation  $\mathcal{T}_{\text{hf}}$  with linear elements of  $\Omega$  that is also depicted in Figure 7(a). We fix the outer boundary and we rotate the inner boundary about the center  $\bar{x} = [0.5, 0.5]$  by an angle  $\theta \in (0, 120^\circ]$ . The same test case was previously considered in [6] to demonstrate the quality of an elasticity-based mesh deformation method.

We consider the registration statement (10) with  $\delta = 10^{-6}$ ,  $\{x_i\}_{i=1}^N$  equal to the points of the mesh on the interior boundary  $\partial\Omega_{\text{in}}$  and  $\{y_i = \mathbf{Rot}(\theta)(x_i - \bar{x})\}_{i=1}^N$ ; we consider the exp-jac objective with  $\epsilon = 0.05$  and the exp-mesh objective with  $\kappa_{\text{msh}} = 10$ . We consider polynomials of degree  $n_{\text{lp}} = 25$ . In order to solve the optimization problem, we split the interval  $(0, 120^\circ]$  into  $N_\theta = 15$  equispace sub-intervals  $(\theta_{k-1}, \theta_k)$  with  $k = 1, \dots, N_\theta$ ; then, we consider a continuation strategy in which we initialize the optimizer for  $\theta = \theta_k$  using the previously-computed solution for  $\theta = \theta_{k-1}$ .

Figures 7(b) and (c) show the deformed meshes for  $\theta = 60^\circ$  and  $\theta = 120^\circ$  for the exp-mesh objective. Figures 8(a)-(b)-(c) compare the behavior of the minimum radius  $q_{\min}$ , the minimum Jacobian determinant  $J_{\min}$ , and the number of iterations of the interior point method for the two choices of the objective function and for several

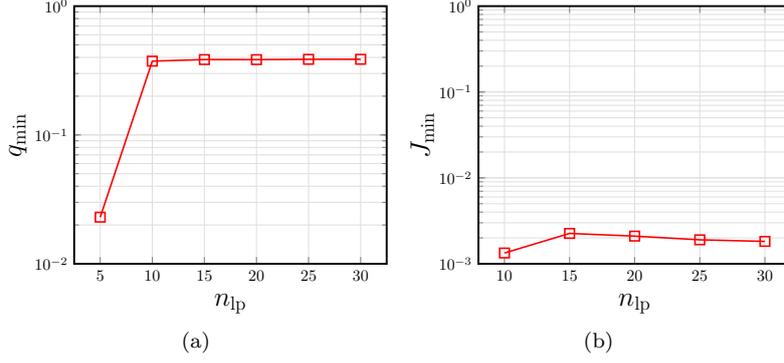


Figure 6: three-point registration problem; sensitivity with respect to the polynomial order  $n_{lp}$  for  $t = 1$ . Objective function: exp-mesh;  $\delta = 10^{-6}$ ,  $\kappa_{\text{msh}} = 10$ .

values of  $\theta$ .

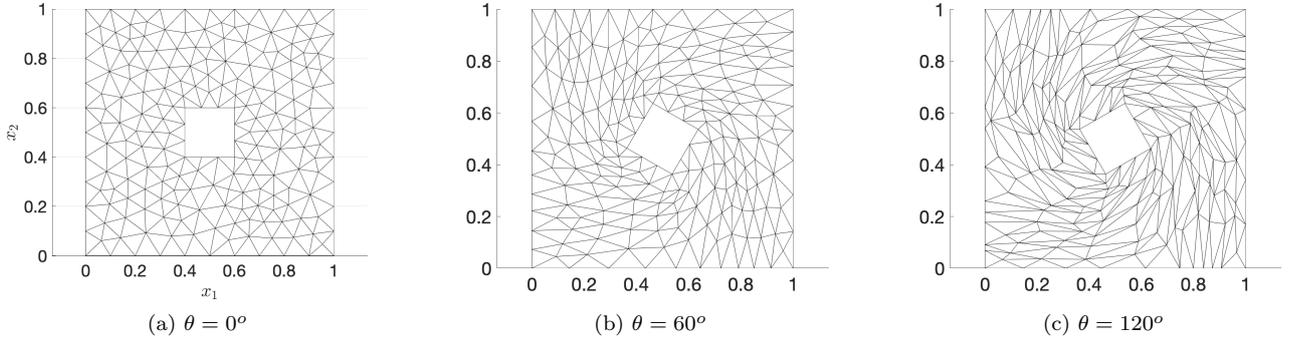


Figure 7: large deformation mesh morphing. (a) reference configuration. (b)-(c) deformed meshes obtained using (10) with exp-mesh objective,  $\delta = 10^{-6}$ ,  $\kappa_{\text{msh}} = 10$ ,  $n_{lp} = 25$ .

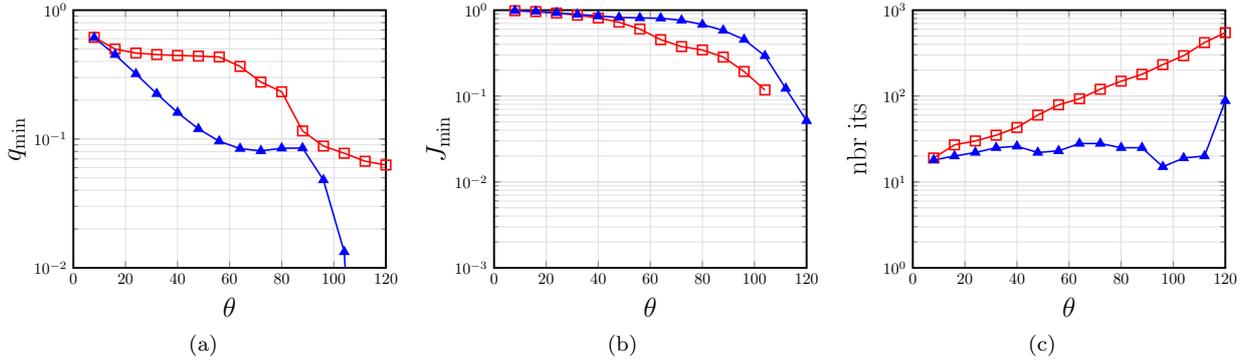


Figure 8: Large deformation mesh morphing. Behavior of the minimum radius ratio  $q_{\min}$ , the minimum Jacobian  $J_{\min}$ , and the number of iterations for exp-jac and exp-mesh objective function and for several values of  $\theta$ . ( $\delta = 10^{-6}$ ,  $\epsilon = 0.05$ ,  $\kappa_{\text{msh}} = 10$ ,  $n_{lp} = 25$ ). exp-mesh  $\square$ ; exp-jac  $\blacktriangle$ .

Some comments are in order. Both approaches succeed to deliver valid deformations with respect to their target — that is, the bijectivity of  $\Phi$  for the exp-jac objective and discrete bijectivity for the exp-mesh objective. Note, however, that for  $\theta \gtrsim 100^\circ$  the solution to (10) with exp-jac objective fails to deliver a valid deformed mesh; similarly, the solution to (10) with exp-mesh objective is not bijective over  $\Omega$ . Recalling Remark 2.1, we conclude that the choice of the objective function should be driven by the choice of the strategy to deal with geometry variations — namely, discretize-then-map or map-then-discretize.

### 4.3 Geometry reduction of a parametric two-dimensional domain

We consider a variant of the parametric geometry reduction problem considered in [31]. We introduce  $\Omega_{\text{box}} = (-2, 2)^2$  and we define the reference domain  $\Omega = \Omega_{\text{box}} \setminus \bigcup_{i=1}^2 \Omega_{\text{in}}^{(i)}$  where  $\Omega_{\text{in}}^{(1)}, \Omega_{\text{in}}^{(2)}$  are two circles of radius  $r = 1/2$  centered in  $\bar{x}_1 = (-1, 0)$  and  $\bar{x}_2 = (1, 0)$ , respectively. We parameterize the inner circles using the functions  $\gamma_{\text{ref}, \text{in}}^{(i)} : [0, 2\pi] \rightarrow \mathbb{R}^2$  with  $\gamma_{\text{ref}, \text{in}}^{(i)}(t) = \bar{x}_i + \frac{1}{2} [\cos(t), \sin(t)]$ . We further introduce the parameter domain  $\mathcal{P}_{\text{in}} = [0.1, 0.4]^2 \times [0, \pi/4]$  and the curves

$$\gamma_{\text{in}, \nu^{(i)}}^{(i)}(t) = \bar{x}_i + \frac{1}{2} \begin{bmatrix} \cos(t) \left( 1 + \nu_1^{(i)} \left( \cos(t + \nu_3^{(i)}) \right)^2 + 2 \cdot 10^{-3} ((2\pi - t)t)^2 \right) \\ \sin(t) \left( 1 + \nu_2^{(i)} \left( \sin(t + \nu_3^{(i)}) \right)^2 + 2 \cdot 10^{-3} ((2\pi - t)t)^2 \right) \end{bmatrix}, \quad (34)$$

with  $\bar{x}_1 = [-1, 0]$ ,  $\bar{x}_2 = [1, 0]$  and  $r = 1/2$ . Then, we define the vector of parameters  $\mu = [\nu^{(1)}, \nu^{(2)}]$  in the parameter region  $\mathcal{P} = \mathcal{P}_{\text{in}} \times \mathcal{P}_{\text{in}}$  and we introduce the family of parameterized domains

$$\Omega_\mu = \Omega_{\text{box}} \setminus \left( \Omega_{\text{in}, \nu^{(1)}}^{(1)} \cup \Omega_{\text{in}, \nu^{(2)}}^{(2)} \right), \quad \text{with } \partial\Omega_{\text{in}, \nu^{(i)}}^{(i)} = \gamma_{\text{in}, \nu^{(i)}}^{(i)}([0, 2\pi]), \quad i = 1, 2. \quad (35)$$

Figure 9(a) shows the reference configuration and the mesh  $\mathcal{T}_{\text{hf}}$  employed for the numerical investigations; Figures 9(b) and (c) show two elements of the family  $\{\Omega_\mu : \mu \in \mathcal{P}\}$ .

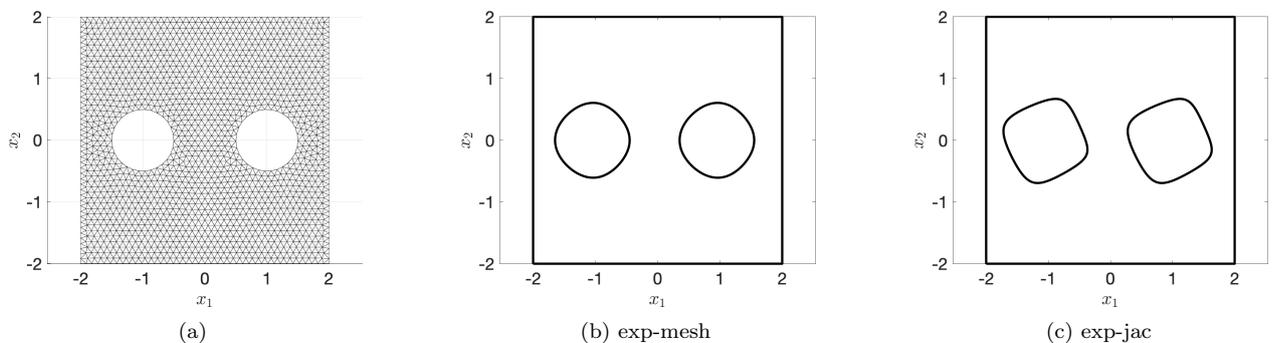


Figure 9: geometry reduction of a parametric two-dimensional domain. (a) reference configuration. (b) domain  $\Omega_\mu$  for  $\mu = [\nu_{\text{min}}, \nu_{\text{min}}]$ , with  $(\nu_{\text{min}})_j = \min_{\nu \in \mathcal{P}_{\text{in}}}(\nu)_j$ ,  $j = 1, 2, 3$ . (c) domain  $\Omega_\mu$  for  $\mu = [\nu_{\text{max}}^{(1)}, \nu_{\text{max}}^{(2)}]$ , with  $(\nu_{\text{max}})_j = \max_{\nu \in \mathcal{P}_{\text{in}}}(\nu)_j$ ,  $j = 1, 2, 3$ .

We consider  $n_{\text{lp}} = 20$  and we enforce  $(\Phi - \text{id}) \cdot \mathbf{n}|_{\partial\Omega_{\text{box}}} = 0$  to ensure bijectivity over  $\Omega_{\text{box}}$  — note that the resulting affine space  $\mathcal{W}_{\text{hf}}$  is of dimension  $M = 720$ . We generate training and test sets  $\mathcal{P}_{\text{train}}, \mathcal{P}_{\text{test}} \subset \mathcal{P}$  of cardinality  $n_{\text{train}} = 100$  and  $n_{\text{test}} = 20$  using independent realizations of an uniform random variable over  $\mathcal{P}$ . We apply registration with exp-mesh objective function ( $\kappa_{\text{msh}} = 10$ ) and we describe the internal reference and parameterized domains using  $N_v = 100$  points,

$$\begin{cases} x_{i:=k+(j-1)N_v} = \gamma_{\text{in}, \text{ref}}^{(j)}(t_k) : k = 1, \dots, N_v, j = 1, 2 \end{cases}, \quad \text{with } t_k = \frac{k-1}{N_v} \quad (36)$$

$$\begin{cases} y_{i:=k+(j-1)N_v}(\mu) = \gamma_{\text{in}, \nu^{(j)}}^{(j)}(t_k) : k = 1, \dots, N_v, j = 1, 2 \end{cases}.$$

Note that the reference and target points are sorted *a priori* and that  $N = 2N_v = 200$ . In all our tests, we initialize the registration algorithm with  $\Phi = \text{id}$ .

We here design a GRR algorithm that takes as input the points  $\{x_i\}_{i=1}^N$  and  $\{y_i^{\text{raw}}\}_{i=1}^Q$  and the mesh  $\mathcal{T}_{\text{hf}}$  of  $\Omega$  and returns a deformed mesh  $\Phi(\mathcal{T}_{\text{hf}})$ ,

$$\Phi(\mathcal{T}_{\text{hf}}) = \text{registration} \left( \mathcal{T}_{\text{hf}}, \{x_i\}_{i=1}^N, \{y_i^{\text{raw}}\}_{i=1}^Q \right). \quad (37)$$

We do not exploit the knowledge of the parameterization. We first consider the simplified case of “sorted” points (cf. (36)) in which we feed the algorithm (37) with  $\{y_i^{\text{raw}} = y_i\}_{i=1}^{Q=N}$ ; then, we consider the more challenging problem of unsorted and sparse points in which we feed (37) with

$$\{y_i^{\text{raw}} = y_{\mathbf{I}_i}\}_{i=1}^Q, \quad Q = 0.8N, \quad \mathbf{I} \subset \{1, \dots, N\}$$

where  $\mathbf{I}$  is obtained by selecting the first  $Q$  points of a random permutation of  $\{1, \dots, N\}$ . We envision that the unsorted problem is an adequate proxy of the geometry registration problems encountered in practical

applications. Table 1 summarizes the choice of the hyper-parameters considered in the numerical experiments. Here, the geometric error is practically estimated as

$$E_{\text{geo}}(\Phi) = \max_{i=1,\dots,N} \min_{j=1,\dots,5N} \|\Phi(x_i) - y_j^{\text{test}}\|_2, \quad (38)$$

where  $\{y_j^{\text{test}}\}_{j=1}^{5N}$  are defined as in (36) with  $N_v = 500$ .

Method	Hyper-parameters	Optimization method
Tykhonov (22)	exp-mesh objective, $\kappa_{\text{msh}} = 10$ , $\xi \in \{10^{-4}, 10^{-5}\}$	quasi-Newton
Morozov (23)	exp-mesh objective, $\delta \in \{10^{-3}, 10^{-4}\}$	interior point (linear constraints)
Inverted (24)	exp-mesh constraint, $\xi \in \{10^{-4}, 10^{-5}\}$ , $\delta_{\text{con}} = 1$	interior point (nonlinear constraints)
Coherent point drift	$\beta = 1$ , $\lambda = 1$ , $w = 0$	Algorithm 2, Appendix A

Table 1: geometry reduction of a parametric two-dimensional domain. Details of the optimization method.

Table 2 and Table 3 summarize the average and worst performance of the registration methods of section 2.3 on the training and the test set, respectively. In more detail, in Table 2, we consider the statements (22), (23), and (24): Tykhonov regularization enables faster predictions, particularly in the worst-case cost; Morozov regularization enables a sharp control of the geometric error — which is in the order of  $\sqrt{\delta}$  for all numerical experiments — at the price of significantly larger worst-case costs; the inverted formulation leads to results that are close to Tykhonov regularization. In Table 3, we compare performance of the same registration algorithms over the test set based on the full space ( $M = 720$ ) and on the reduced space obtained using POD with tolerance  $tol_{\text{pod}} = 10^{-5}$  (cf. (25)): we set  $\xi = 10^{-5}$  in (22) and (24), and  $\delta = 10^{-4}$  in (23). we observe that POD enables speedups in the order of  $\mathcal{O}(100)$  without any significant deterioration of performance. Even more, we observe that dimensionality reduction slightly improves reconstruction performance for certain configurations: this empirical finding suggests that the reduction of the number of unknowns simplifies the optimization task and might prevent convergence to suboptimal local minima.

Training (sorted)	$q_{\text{min}}$		geo error		nbr its		cost [s]	
	avg	min	avg	max	avg	max	avg	max
Tykhonov ( $\xi = 10^{-4}$ )	0.44	0.32	$9.53 \cdot 10^{-5}$	$1.98 \cdot 10^{-4}$	287.5	401	31.85	44.04
Tykhonov ( $\xi = 10^{-5}$ )	0.42	0.27	$8.18 \cdot 10^{-5}$	$1.17 \cdot 10^{-4}$	274.9	383	30.62	42.39
Morozov ( $\delta = 10^{-3}$ )	0.45	0.34	$1.40 \cdot 10^{-3}$	$1.42 \cdot 10^{-3}$	78.6	176	26.58	68.18
Morozov ( $\delta = 10^{-4}$ )	0.45	0.34	$1.36 \cdot 10^{-4}$	$1.42 \cdot 10^{-4}$	75.7	211	30.06	108.04
Inverted ( $\xi = 10^{-4}$ )	0.41	0.24	$3.37 \cdot 10^{-4}$	$6.20 \cdot 10^{-4}$	199.6	285	33.03	49.38
Inverted ( $\xi = 10^{-5}$ )	0.41	0.24	$8.43 \cdot 10^{-5}$	$1.29 \cdot 10^{-4}$	204.14	349	32.92	55.89

Table 2: geometry reduction of parametric two-dimensional domain. Registration performance on the training set  $\mathcal{P}_{\text{train}}$  for sorted data,  $M = 720$ .

Test (sorted)	$q_{\text{min}}$		geo error		nbr its		cost [s]	
	avg	min	avg	max	avg	max	avg	max
Tykhonov, full ( $M = 720$ )	0.41	0.30	$0.81 \cdot 10^{-4}$	$1.01 \cdot 10^{-4}$	270.0	378	30.10	41.91
Tykhonov, reduced ( $M = 20$ )	0.42	0.31	$0.73 \cdot 10^{-4}$	$1.01 \cdot 10^{-4}$	59.3	91	0.23	0.31
Morozov, full ( $M = 720$ )	0.44	0.34	$1.37 \cdot 10^{-4}$	$1.40 \cdot 10^{-4}$	68.3	86	26.31	73.12
Morozov, reduced ( $M = 21$ )	0.44	0.35	$1.35 \cdot 10^{-4}$	$1.42 \cdot 10^{-4}$	47.1	61	0.51	0.63
Inverted, full ( $M = 720$ )	0.41	0.27	$0.79 \cdot 10^{-4}$	$1.13 \cdot 10^{-4}$	217.0	320	34.68	51.71
Inverted, reduced ( $M = 21$ )	0.40	0.26	$0.69 \cdot 10^{-4}$	$1.20 \cdot 10^{-4}$	77.5	117	0.35	0.51

Table 3: geometry reduction of a parametric two-dimensional domain. Registration performance on the test set  $\mathcal{P}_{\text{test}}$  for sorted data.

Table 4 shows performance for the unsorted problem; in this test, we consider unsorted data both for training and assessment, and we resort to the full CPD space. We observe that both Tykhonov and the inverted formulation lead to performance that are comparable to the performance obtained using CPD in terms of geometric error. Note also that the number of POD modes required to achieve the tolerance  $tol_{\text{pod}} = 10^{-5}$  in (25) is significantly larger (see also Figure 11) for the unsorted dataset: this empirical finding suggests that the application of CPD to the unsorted dataset with missing points might introduce numerical noise that ultimately hinders the compressibility of the manifold associated with the mapping  $\Phi$ .

Test (unsorted)	$q_{\min}$		geo error		nbr its		cost [s]	
	avg	min	avg	max	avg	max	avg	max
CPD	—	—	$6.68 \cdot 10^{-3}$	$8.61 \cdot 10^{-3}$	53.0	97	0.03	0.05
Tykhonov, full ( $M = 720$ )	0.43	0.29	$6.93 \cdot 10^{-3}$	$8.61 \cdot 10^{-3}$	334.5	401	37.39	44.75
Tykhonov, reduced ( $M = 48$ )	0.43	0.29	$6.93 \cdot 10^{-3}$	$8.67 \cdot 10^{-3}$	181.1	237	1.19	1.52
Inverted, full ( $M = 720$ )	0.40	0.26	$7.03 \cdot 10^{-3}$	$8.86 \cdot 10^{-3}$	208.0	245	33.91	41.41
Inverted, reduced ( $M = 40$ )	0.40	0.26	$7.32 \cdot 10^{-3}$	$9.03 \cdot 10^{-3}$	130.3	143	0.85	1.02

Table 4: geometry reduction of a parametric two-dimensional domain. Registration performance on the test set  $\mathcal{P}_{\text{test}}$  for unsorted data, with 80% training data ( $\xi = 10^{-4}$ ). CPD is based on the full space (21).

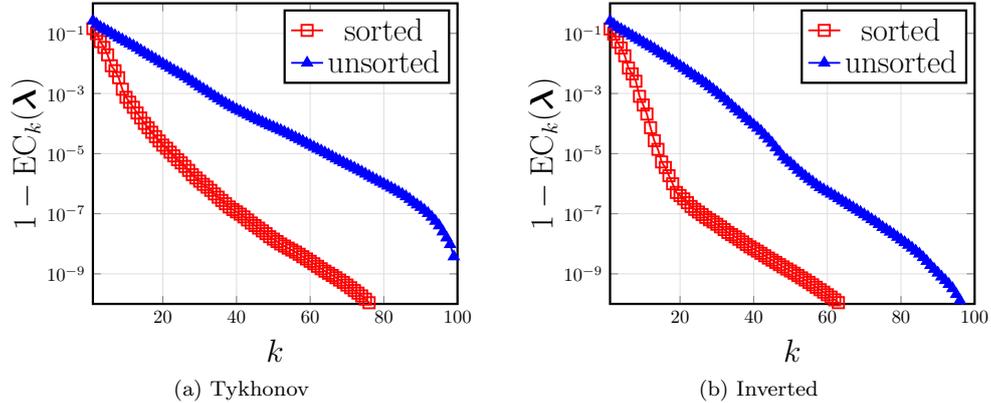


Figure 10: geometry reduction of a parametric two-dimensional domain. POD eigenvalues for sorted and unsorted data for Tykhonov and inverted formulations with  $\xi = 10^{-4}$ ; CPD is based on the full space (21).

Figure 11 shows the deformed meshes obtained for an out-of-sample parameter using Tykhonov regularization with full space and sorted data ( $\xi = 10^{-5}$ ), Morozov regularization with full space and sorted data ( $\delta = 10^{-4}$ ), Tykhonov regularization (in combination with CPD) with reduced space and unsorted data, respectively ( $\xi = 10^{-5}$ ,  $M = 48$ ). For this test case, we obtain that the geometric error (38) is given by  $7.8 \cdot 10^{-5}$ ,  $1.4 \cdot 10^{-4}$ ,  $5.9 \cdot 10^{-3}$ , respectively; the geometric error of CPD is  $5.9 \cdot 10^{-3}$ .

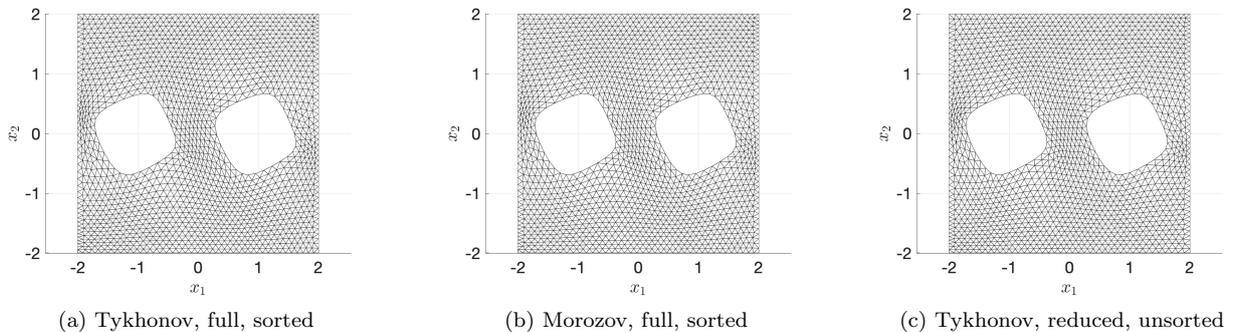


Figure 11: geometry reduction of a parametric two-dimensional domain. Deformed mesh obtained using three different methods for  $\mu = [\nu_{\max}, \nu_{\max}]$ .

Table 5 investigates performance of the registration method based on sorted training on a unsorted test dataset: we envision that this scenario is of interest for real-time applications for which online data are considerably more noisy than data used for training. We consider the Tykhonov formulation with  $\xi = 10^{-4}$ ; we perform POD with tolerance  $tol_{\text{pod}} = 10^{-5}$  to obtain a reduced space of cardinality  $M = 22$ ; then, we compare performance obtained using CPD based on the full space (21) with the performance obtained using CPD with the reduced space (26). We observe that dimensionality reduction in the CPD algorithm has a beneficial effect on the accuracy of the registration procedure and also reduces the cost of CPD.

## 5 Summary and discussion

We presented an optimization-based approach to the problem of geometry registration and reduction (GRR). GRR is of paramount importance in pMOR to deal with PDEs in parametric geometries. We presented a

Full-space CPD	$q_{\min}$		geo error		nbr its		cost [s]	
	avg	min	avg	max	avg	max	avg	max
CPD			$6.68 \cdot 10^{-3}$	$8.61 \cdot 10^{-3}$	53.0	97	0.04	0.08
Tykhonov, reduced ( $M = 22$ )	0.46	0.30	$2.95 \cdot 10^{-2}$	$5.91 \cdot 10^{-2}$	95.3	119	0.33	0.39

Reduced-space CPD	$q_{\min}$		geo error		nbr its		cost [s]	
	avg	min	avg	max	avg	max	avg	max
CPD			$6.09 \cdot 10^{-3}$	$1.13 \cdot 10^{-2}$	26.4	39	0.01	0.02
Tykhonov, reduced ( $M = 22$ )	0.41	0.34	$3.29 \cdot 10^{-3}$	$1.14 \cdot 10^{-2}$	68.4	116	0.24	0.38

Table 5: geometry reduction of parametric two-dimensional domain. Registration performance on the test set  $\mathcal{P}_{\text{test}}$  for unsorted data; training performed on sorted data,  $\xi = 10^{-4}$ ,  $\text{tol}_{\text{pod}} = 10^{-5}$ .

thorough mathematical analysis that offers the theoretical foundations for the methodology; we further presented thorough numerical investigations for three two-dimensional model problems.

The analysis of section 3 rigorously justifies the proposed approach for smooth  $C^2$  domains; it also illustrates the issues that the approach faces when dealing with slender bodies and/or domains with corners. In particular, ensuring appropriate approximation — via interpolation — of the target shape at corners and cusps is key to recover near-optimal bounds for the geometry error.

The numerical results of section 4 provide a thorough overview of the performance of optimization-based registration methods, for several choices of the objective function and of the parameters. First, Tykhonov regularization appears to be superior in terms of computational performance at the price of a less straightforward choice of the free parameter  $\xi$  (compared to Morozov regularization). In addition, the combination with a representative PSR procedure did not lead to any unstable behavior and/or lack of convergence. Second, we observe that potential approximations of the form  $\Phi = \text{id} + \nabla\phi$  are highly suboptimal (cf. Figure 5) in our setting: this is in striking contrast with the related problem of optimal transportation for which the solution is guaranteed to be the gradient of a convex function [3]. Third, the results of section 4.3 show that geometry reduction based on POD might greatly reduce the cost of registration. In this respect, the development of specialized hyper-reduction (see, e.g., [5, 38]) techniques might lead to much more significant computational gains for large-scale problems. Fourth, we observed that the deformation is highly sensitive to the choice of the objective function (see, e.g., Figure 3 and Figure 8): for projection-based pMOR applications, this choice should thus be driven by the strategy to deal with geometry variations — namely, discretize-then-map or map-then-discretize. Fifth, we observe that the approach can also cope with unsorted data when trained on sorted data (cf. Table 5): towards this end, the application of the reduced-space CPD seems crucial to achieve accurate reconstructions.

## Acknowledgement

The author thanks Prof. Angelo Iollo (University of Bordeaux, Inria) for fruitful discussions; he also thanks Dr. Pierre Mounoud (University of Bordeaux) for extensive discussions on the proofs of section 3.

## A Coherent point drift

In this Appendix, we review the coherent point drift (CPD) algorithm first introduced in [24]. CPD is a probabilistic method based on a Gaussian mixture model (GMM): the PSR problem is formulated as a maximum likelihood estimation problem with a motion coherence constraint and is solved using an expectation-maximization (EM) procedure. First, we briefly present EM algorithms for maximum likelihood estimation; then, we present the probabilistic model of CPD and we discuss the application of EM; finally, we comment on dimensionality reduction. In Appendix A.1 and Appendix A.2, we denote by  $\mathbf{Y} = [y_1, \dots, y_Q]^*$  the observed data to shorten notation.

### A.1 Expectation maximization procedures

The EM algorithm is a prominent method in machine learning to find maximum likelihood solutions for models with latent variables; EM is used when the maximization of the likelihood of the observed variables is difficult but can be made easier by enlarging the sample with latent (unobserved) data. We refer to [2] Chapter 9, and to [8, Chapter 8] for a thorough review of the methodology and for a rigorous theoretical justification.

We denote by  $\Theta$  the parameter values of the model, by  $\mathbf{Y} = [y_1, \dots, y_Q]^* \in \mathbb{R}^{Q \times d}$  the observed data and by  $\mathbf{Z} = [z_1, \dots, z_Q]^* \in \{1, \dots, N\}^Q$  the latent variables for some  $N \in \mathbb{N}$  — we here consider the case of a single

discrete latent variable; however, the method can also deal with continuous latent variables. We denote by  $p(\mathbf{Y}, \mathbf{Z}|\Theta)$  the joint likelihood; recalling the Bayes' theorem, we find the expression for the posterior of  $\mathbf{Z}$ :

$$\Pr(\mathbf{Z}|\mathbf{Y}, \Theta) = \frac{\Pr(\mathbf{Y}, \mathbf{Z}|\Theta)}{\Pr(\mathbf{Z}|\Theta)}. \quad (39)$$

Assuming that the samples are independent identically distributed, the EM procedure can be stated as in Algorithm 1. Note that, since the logarithm acts directly on the joint distribution, the M-step maximization might be tractable. The choice of the expectation in the M-step is rigorously justified by interpreting EM as an alternative minimization algorithm (see, [2, Chapter 9.4] [8, Chapter 8.5.3]). Several authors have also considered to partition the parameters  $\Theta$  into groups and then break down the M-step into multiple steps each of which involves optimizing one of the subsets with the remainders held fixed [23].

Algorithm 1: Expectation-Maximization algorithm.

- 1: Set  $\Theta^{(k)} = \Theta_0$ .
- 2: **for**  $k = 1, 2, \dots$ , **until convergence do**
- 3:   **E-step:** Evaluate  $(\mathbf{P})_{i,j} = \Pr(Z = z_j | \mathbf{Y} = y_i, \Theta^{(k)})$  using (39).
- 4:   **M-step:** Maximize the expectation of the complete data log-likelihood  

$$\Theta \mapsto \Omega(\Theta; \mathbf{P}) := \sum_{i=1}^Q \sum_{j=1}^K (\mathbf{P})_{i,j} \log(\Pr(Z = z_i, Y = y_j | \Theta))$$
 to find  $\Theta^{(k+1)}$ .
- 5: **end for**

## A.2 Probabilistic model of CPD

We introduce the mixture model

$$Y = \sum_{j=1}^N \mathbb{1}(Z = j)Y^{(j)} + \mathbb{1}(Z = N + 1)Y^{(N+1)}, \quad (40)$$

where  $Y^{(j)} \sim \mathcal{N}(\mu_j, \sigma^2 \mathbb{1})$  for  $j = 1, \dots, N$ ,  $Y^{(N+1)} \sim \text{Uniform}(\mathcal{D})$ , and  $Z \sim \text{Multinomial}(\{1, \dots, N + 1\})$  are independent random variables. Given  $w \in [0, 1)$ , we set

$$\Pr(Z = j) = \begin{cases} \frac{1-w}{N} & j = 1, \dots, N \\ w & j = N + 1 \end{cases} \quad |\mathcal{D}| = Q.$$

Exploiting the previous hypotheses, we find

$$\Pr(Y = y, Z = j) = \Pr(Y^{(j)} = y) \Pr(Z = j) = \begin{cases} \frac{1-w}{N} \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{\|y - \mu_j\|_2^2}{2\sigma^2}\right) & \text{if } j \in \{1, \dots, N\} \\ \frac{w}{Q} & \text{if } j = N + 1. \end{cases}$$

We consider the set of tunable parameters  $\Theta = \{\mu_1, \dots, \mu_N, \sigma^2\}$ , while we fix  $w, \mathcal{D}$  *a priori*. We observe that (40) reads as a GMM with a perturbation given by the uniform random variable  $Y^{(N+1)}$ : the latter is intended to account for noise and outliers in the dataset. We do not explicitly construct the domain  $\mathcal{D}$ : we simply assume that all the observed datapoints belong to  $\mathcal{D}$ .

Given  $j \in \{1, \dots, N + 1\}$  and  $y \in \mathcal{D}$ , using (39), we find

$$\Pr(Z = j | Y = y) = \frac{\Pr(Y^{(j)} = y, Z = j)}{\sum_{k=1}^{N+1} \Pr(Z = k) \Pr(Y^{(k)} = y)} = \frac{\Pr(Y^{(j)} = y) \Pr(Z = j)}{\sum_{k=1}^{N+1} \Pr(Z = k) \Pr(Y^{(k)} = y)} = \frac{\exp\left(-\frac{\|y - \mu_j\|_2^2}{2\sigma^2}\right)}{\sum_{k=1}^N \exp\left(-\frac{\|y - \mu_k\|_2^2}{2\sigma^2}\right) + c}, \quad (41)$$

with  $c = \frac{w}{1-w} \frac{N}{Q} (2\pi\sigma^2)^{d/2}$ , for  $j = 1, \dots, N$ . We further introduce the matrix  $\mathbf{P} \in \mathbb{R}^{Q \times N}$  such that  $(\mathbf{P})_{i,j} = \Pr(Z = j | Y = y_i)$  for  $i = 1, \dots, Q$  and  $j = 1, \dots, N$  and we define  $Q_p = \sum_{i,j} (\mathbf{P})_{i,j}$ ; then, assuming that

$y_1, \dots, y_Q \in \mathcal{D}$ , we obtain

$$\begin{aligned}
-\Omega(\Theta; \mathbf{P}) &= -\sum_{i=1}^Q \left( \left( \sum_{j=1}^N (\mathbf{P})_{i,j} \log \Pr(Y = y_i, Z = j | \Theta) \right) + \left( 1 - \sum_{k=1}^N (\mathbf{P})_{i,k} \right) \log(1/|\mathcal{D}|) \right) \\
&= \sum_{i=1}^Q \sum_{j=1}^N (\mathbf{P})_{i,j} \left( \frac{1}{2\sigma^2} \|y_i - \mu_j\|_2^2 + \frac{d}{2} \log(\sigma^2) + \log \left( \frac{1-w}{N(2\pi)^{d/2}} \right) \right) + C_1; \\
&= \left( \frac{1}{2\sigma^2} \sum_{i=1}^Q \sum_{j=1}^N (\mathbf{P})_{i,j} \|y_i - \mu_j\|_2^2 \right) + \frac{dQ_P}{2} \log(\sigma^2) + C_2,
\end{aligned} \tag{42}$$

where  $C_1, C_2$  are constants that are independent of  $\Theta$ .

So far, we have not forced the GMM centroids to move coherently; towards this end, we propose the model  $\mu_j = x_j + v(x_j)$  where  $\{x_j\}_{j=1}^N$  is the reference point cloud that we wish to deform and  $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a displacement field that is assumed to belong to the native Reproducing Kernel Hilbert space (RKHS)  $\mathcal{H}_\phi$  associated with the radial basis function (RBF)  $\phi : r \mapsto \exp\left(-\frac{r^2}{2\beta^2}\right)$ . Then, we introduce the objective function  $\mathfrak{E}$  for the M-step by adding the regularization  $\frac{\lambda}{2} \|v\|_{\mathcal{H}_\phi}^2$  to (42):

$$\mathfrak{E}(v, \sigma^2 | \mathbf{P}) = \frac{1}{2\sigma^2} \sum_{i=1}^Q \sum_{j=1}^N (\mathbf{P})_{i,j} \|y_i - x_j - v(x_j)\|_2^2 + \frac{dQ_P}{2} \log(\sigma^2) + \frac{\lambda}{2} \|v\|_{\mathcal{H}_\phi}^2. \tag{43}$$

Exploiting the representation theorem for RKHS (e.g., [37]), we find that minimizers of (43) belong to the space  $\mathcal{U}_{\text{hf}}^{\text{cpd}}$  introduced in (21); we also observe that in the Bayesian setting the regularization  $\frac{\lambda}{2} \|v\|_{\mathcal{H}_\phi}^2$  might be associated to the logarithm of a prior on the displacement field.

We observe that the probabilistic model introduced in this section depends on three free parameters:  $w, \beta, \lambda$ . The parameter  $w \in [0, 1]$  is designed to account for noise and outliers in the datasets: it should be set to zero for noiseless data. The parameter  $\beta$  reflects the strength of interaction between points: the value of  $\beta$  should thus depend on the characteristic length-scale of the displacement field we wish to approximate. The value of  $\lambda$  is associated with the Tikhonov regularization of the  $E$ -step and thus reflects the trade-off between data fitting and smoothness regularization. Finally, the algorithm also depends on the choice of the RBF  $\phi$ : following [24], we here consider the Gaussian kernel; however, we observe that several other choices are possible and there is an extensive literature on the comparison between different RBFs [28].

### A.3 Coherent point drift procedure

We adapt the EM algorithm 1 to the mixture model (40): we present the detailed procedure in Algorithm 2. To facilitate the M-step, we first solve for the displacement  $v$  and then for  $\sigma^2$ .

Algorithm 2: Coherent point drift [24].

*Inputs:*  $\mathbf{X} = [x_1, \dots, x_N]^* \in \mathbb{R}^{N \times d}$ ,  $\mathbf{Y}^{\text{raw}} = [y_1^{\text{raw}}, \dots, y_Q^{\text{raw}}]^* \in \mathbb{R}^{Q \times d}$ .

*Hyper-parameters:*  $w \in [0, 1], \beta, \lambda > 0$ .

*Outputs:*  $\mathbf{Y} = [x_1 + v(x_1), \dots, x_N + v(x_N)]^* \in \mathbb{R}^{N \times d}$ .

- 1: Set  $v^{(0)} = 0$  and  $\sigma^{2,(0)} = \frac{1}{dQN} \sum_{i=1}^Q \sum_{j=1}^N \|x_j - y_i^{\text{raw}}\|_2^2$ .
- 2: **for**  $k = 1, 2, \dots$ , **until convergence do**
- 3:   **E-step:** Find  $\mathbf{P}^{(k)} \in \mathbb{R}^{Q \times N}$  such that  $(\mathbf{P}^{(k)})_{i,j} = \Pr(Z = j | Y = y_i^{\text{raw}})$  using (41) with  $\mu_j = x_j + v^{(k-1)}(x_j)$  and  $\sigma^2 = \sigma^{2,(k-1)}$ .
- 4:   **M-step (I):** find  $v^{(k)} = \arg \min_{v \in \mathcal{H}_\phi} \mathfrak{E}(v, \sigma^{2,(k-1)} | \mathbf{P}^{(k)})$  (cf. (43)).
- 5:   **M-step (II):** find  $\sigma^{2,(k)} = \arg \min_{\sigma^2 \in \mathbb{R}_+} \mathfrak{E}(v^{(k)}, \sigma^2 | \mathbf{P}^{(k)})$ .
- 6: **end for**

Exploiting the representation theorem for  $v$ , we find that  $v^{(k)}(\cdot) = \sum_{j=1}^N w_j^{(k)} \phi(\|\cdot - x_j\|_2)$  and  $\mathbf{W}^{(k)} = [w_1^{(k)}, \dots, w_N^{(k)}]^* \in \mathbb{R}^{N \times d}$  solves the linear system (we omit dependence on the iteration count  $k$  at the right-

hand-side to shorten notation)

$$\mathbf{W}^{(k)} = \left( \mathbf{G} + \lambda \sigma^2 (\mathbf{d}(\mathbf{P}\mathbf{1}_Q))^{-1} \right)^{-1} \left( (\mathbf{d}(\mathbf{P}\mathbf{1}_Q))^{-1} \mathbf{P}\mathbf{Y}^{\text{raw}} - \mathbf{X} \right), \quad (44)$$

where  $\mathbf{G} \in \mathbb{R}^{N \times N}$  satisfies  $(\mathbf{G})_{i,j} = \exp\left(-\frac{1}{2\beta^2} \|x_i - x_j\|_2^2\right)$ ,  $i, j = 1, \dots, N$ ,  $\mathbf{1}_Q$  is the  $Q$ -dimensional vector with entries all equal to one, and  $\mathbf{d}(\mathbf{P}\mathbf{1}_Q) = \text{diag}((\mathbf{P}\mathbf{1}_Q)_1, \dots, (\mathbf{P}\mathbf{1}_Q)_N) \in \mathbb{R}^{N \times N}$ . By tedious but straightforward calculations, we also find the explicit expression for  $\sigma^{2,(k)}$

$$\sigma^{2,(k)} = \frac{1}{Q_p d} \left( \text{Tr}((\mathbf{Y}^{\text{raw}})^* \mathbf{d}(\mathbf{P}^* \mathbf{1}_N) \mathbf{Y}^{\text{raw}}) - 2 \text{Tr}((\mathbf{P}\mathbf{Y}^{\text{raw}})^* \mathbf{Y}) + \text{Tr}(\mathbf{Y}^* \mathbf{d}(\mathbf{P}^* \mathbf{1}_N) \mathbf{Y}) \right). \quad (45)$$

with  $Q_p = \sum_{i,j} (\mathbf{P}^{(k)})_{i,j}$  and  $\mathbf{Y} = \mathbf{X} + \mathbf{G}\mathbf{W}^{(k)}$ .

Since the seminal work [24], several authors have discussed how to effectively implement Algorithm 2 to enable its application to large datasets and also to prevent unstable behaviors. In this work, we resort to an eigenvalue decomposition of the matrix  $\mathbf{G}$  and to the Woodbury identity to invert the system in (44) (cf. [24, section 6]). Furthermore, before solving the system, we remove all lines  $i$  for which  $\sum_{j=1}^Q (\mathbf{P})_{i,j} < 10^{-12}$  and we set  $\mathbf{W}(i, \cdot) = 0$ . In all our numerical experiments, we consider the termination condition

$$\|\mathbf{W}^{(k)} - \mathbf{W}^{(k-1)}\|_2 < 10^{-4} \quad \text{OR} \quad \max_{j=1, \dots, Q} \min_{i=1, \dots, N} \|\mathbf{Y}^{\text{raw}}(j, \cdot) - \mathbf{Y}(i, \cdot)\|_2 < 10^{-5}. \quad (46)$$

#### A.4 Low-dimensional representation of the displacement field $v$

As discussed in section 2.3.2, and also shown in the numerical experiment of section 4.3, we might wish to restrict computations over a low-dimensional subspace  $\mathcal{U}_M^{\text{cpd}}$ . Since we are ultimately interested in bounded-domain registration, the natural choice is  $\mathcal{U}_M^{\text{cpd}} = \mathcal{U}_M$ . Since in general  $\mathcal{U}_M$  is not contained in  $\mathcal{H}_\phi$ , we propose to replace the regularization  $\|\cdot\|_{\mathcal{H}_\phi}^2$  in (43) with  $\|\cdot\|_{H^2(\Omega_{\text{box}})}^2$ . In conclusion, we replace Line 4 of Algorithm 2 with

$$v^{(k+1)} = \arg \min_{v \in \mathcal{U}_M} \frac{1}{2\sigma^2} \sum_{i=1}^Q \sum_{j=1}^N (\mathbf{P})_{i,j} \|y_i - x_j - v(x_j)\|_2^2 + \frac{\lambda}{2} \|v\|_{H^2(\Omega_{\text{box}})}^2. \quad (47)$$

Provided that  $\{\psi_m\}_{m=1}^M$  is an orthonormal basis of  $\mathcal{U}_M$ , by tedious but straightforward calculations, we find that  $v^{(k+1)}(\cdot) = \sum_{m=1}^M (\mathbf{A}^{-1} \mathbf{b})_m \psi_m(\cdot)$  where

$$\begin{aligned} (\mathbf{A})_{m,m'} &= \lambda \sigma^2 \delta_{m,m'} + \sum_{j=1}^N \sum_{\ell=1}^d (\psi_m(x_j))_\ell (\mathbf{P}\mathbf{1}_Q)_j (\psi_{m'}(x_j))_\ell \\ (\mathbf{b})_m &= \sum_{j=1}^N \sum_{\ell=1}^d (\psi_m(x_j))_\ell (\mathbf{P}\mathbf{Y}^{\text{raw}} - \mathbf{d}(\mathbf{P}\mathbf{1}_Q)\mathbf{X})_{j,\ell}, \end{aligned}$$

for  $m, m' = 1, \dots, M$ .

## B Proofs

### B.1 Proposition 3.1

We first observe that if  $\{x_i\}_{i=1}^N$  is an  $\epsilon$ -cover of  $\partial\Omega$  then  $\{\Phi(x_i)\}_{i=1}^N$  is  $K\epsilon$ -cover of  $\partial\Phi(\Omega)$ . To prove this statement, consider  $\tilde{x} \in \partial\Phi(\Omega)$ ; since  $\Phi$  is bijective,  $\partial\Phi(\Omega) = \Phi(\partial\Omega)$  and thus there exists  $x \in \partial\Omega$  such that  $\tilde{x} = \Phi(x)$ . We denote by  $i^* \in \{1, \dots, N\}$  the index that satisfies  $\|x - x_{i^*}\|_2 = \text{dist}(x, \{x_i\}_{i=1}^N)$ : since  $\{x_i\}_{i=1}^N$  is an  $\epsilon$ -cover of  $\partial\Omega$ , we have  $\|x - x_{i^*}\|_2 \leq \epsilon$ . Then, we find

$$\text{dist}(\tilde{x}, \{\Phi(x_i)\}_{i=1}^N) = \text{dist}(\Phi(x), \{\Phi(x_i)\}_{i=1}^N) \leq \|\Phi(x) - \Phi(x_{i^*})\|_2 \leq K\epsilon.$$

Exploiting the previous estimate, we obtain that for all  $\tilde{x} = \Phi(x) \in \partial\Phi(\Omega)$  we have

$$\begin{aligned} \text{dist}(\tilde{x}, \partial V) &= \inf_{y \in \partial V} \|\Phi(x) - y\|_2 \leq \min_{i=1, \dots, N} \|\Phi(x) - y_i\|_2 \\ &\leq \min_{i=1, \dots, N} (\|\Phi(x_i) - y_i\|_2 + \|\Phi(x_i) - \Phi(x)\|_2) \\ &\leq \left( \max_{i=1, \dots, N} \|\Phi(x_i) - y_i\|_2 \right) + \left( \min_{i=1, \dots, N} \|\Phi(x_i) - \Phi(x)\|_2 \right) \leq \max_{i=1, \dots, N} \|\Phi(x_i) - y_i\|_2 + K\epsilon. \quad \square \end{aligned}$$

## B.2 Lemma 3.1

Proof of the second statement is an immediate consequence of the Weyl's tube formula [7], which states that if the curves (or surfaces for  $d = 3$ )  $\{x \pm \delta \mathbf{n}(x) : x \in \partial U\}$  do not have self-intersections, we have

$$\text{Neigh}_\delta(\partial U) = 2\delta|\partial U|_{(d-1)} + \frac{4\pi}{3}\chi(\partial U)\delta^3,$$

where  $\chi(\partial U)$  is the so-called Euler characteristics. The quantity  $\chi(\partial U)$  is invariant under isomorphisms and is equal to 0 for the unit circle and to 2 for the unit sphere.

We now prove the first statement. In more detail, we show that any  $y \in \text{Neigh}_\delta^t(\partial U)$  belongs to  $\text{Neigh}_\delta(\partial U)$  and then that any  $y \in \text{Neigh}_\delta(\partial U)$  belongs to  $\text{Neigh}_\delta^t(\partial U)$ .

Let  $y = x + t\mathbf{n}(x) \in \text{Neigh}_\delta^t(\partial U)$ . Clearly, we have  $\text{dist}(y, \partial U) \leq \|y - x\|_2 = |t| < \delta$ , which implies that  $y \in \text{Neigh}_\delta(\partial U)$ .

Let  $y \in \text{Neigh}_\delta(\partial U)$ . We define  $f : \partial U \rightarrow \mathbb{R}$  such that  $f(x) = \|x - y\|_2^2$ ; we denote by  $\nabla_{\partial U} f := \nabla f - \mathbf{n}(\mathbf{n} \cdot \nabla f)$  the surface gradient of  $f$ . Since  $\partial U$  is a closed set in  $\mathbb{R}^d$ , there exists a minimizer  $\tilde{x} \in \partial U$  that  $f(\tilde{x}) = \inf_{x \in \partial U} f(x)$ ; furthermore,

$$\|\tilde{x} - y\|_2^2 = \inf_{x \in \partial U} f(x) = (\text{dist}(y, \partial U))^2 < \delta^2 \quad \Rightarrow \|\tilde{x} - y\|_2 < \delta.$$

Since  $\partial U$  is a closed hyper-surface without boundary,  $\tilde{x}$  must be a stationary point of  $f$  over  $\partial U$ , that is  $\nabla_{\partial U} f(\tilde{x}) = 0$ . Exploiting the definition of the surface gradient, we find that  $\tilde{x} - y = \tilde{\mathbf{n}}(\tilde{\mathbf{n}} \cdot (x^* - y))$  with  $\tilde{\mathbf{n}} = \mathbf{n}(\tilde{x})$ , which implies that  $\tilde{x} - y$  is parallel to  $\tilde{\mathbf{n}}$  and then that

$$y = \tilde{x} - (\theta \|\tilde{x} - y\|_2) \tilde{\mathbf{n}},$$

for either  $\theta = 1$  or  $\theta = -1$ . Recalling (31), we find that  $y \in \text{Neigh}_\delta^t(\partial U)$ .  $\square$

## B.3 Lemma 3.2

The no-intersection hypothesis implies that  $\text{Neigh}_\delta(\partial U)$  is diffeomorphic to a spherical shell. Therefore, there exists a mapping  $\Phi$  that transforms  $U$  into the unit ball and  $\text{Neigh}_\delta(\partial U)$  into the shell  $\mathcal{B}_{1+\delta}(0) \setminus \mathcal{B}_{1-\delta}(0)$ . We conclude that it suffices to prove the result for  $U$  equal to the unit ball.

By contradiction, let  $y \in (V \setminus \text{Neigh}_\delta(\partial U)) \cap \mathcal{B}_{1-\delta}(0)$  and consider  $p = x - \delta \mathbf{n}(x) \in U \setminus \bar{V}$ ; note that  $p \in \partial \mathcal{B}_{1-\delta}(0)$ . Clearly, the segment  $\Gamma = \{tp + (1-t)y : t \in (0, 1)\}$  is contained in  $\mathcal{B}_{1-\delta}(0)$  and — since  $y \in V$  and  $p \notin V$  — intersects  $\partial V$  at some point  $z \in \partial V \cap \Gamma$ . We thus proved the existence of a point  $z \in \partial V$  that does not belong to  $\text{Neigh}_\delta(\partial U)$ , which is a contradiction. The case  $y \in (V \setminus \text{Neigh}_\delta(\partial U)) \setminus \mathcal{B}_{1+\delta}(0)$  can be treated similarly.  $\square$

## B.4 Lemma 3.3.

The aim of this section is to further discuss Definition 3.1 and to prove a refined version of Lemma 3.3. We focus on the three-dimensional case: the same argument applies to two-dimensional domains with minor modifications.

We first recall the following elementary result.

**Lemma B.1.** *Let  $f : [a, b] \rightarrow \mathbb{R}^k$  be differentiable. Then,  $\|f(b) - f(a)\|_2 \leq (b - a) \max_{\zeta \in [a, b]} \|f'(\zeta)\|_2$ .*

*Proof.* It suffices to apply the mean value theorem to the scalar function  $\phi(\zeta) = (f(b) - f(a)) \cdot f(\zeta)$ .  $\square$

Next Lemma introduces relevant quantities for the subsequent result.

**Lemma B.2.** *Let  $U$  be a three-dimensional bounded domain of class  $C^2$ . Then, there exist positive constants  $L, L_{\text{inv}}, r$  such that for any  $x \in \partial U$  there exists a bijective mapping  $\Phi_x : A \rightarrow U \cap \mathcal{B}_r(x)$  with  $A = \mathcal{B}_r(0) \cap \{x' : (x')_2 < 0\}$  with the following properties: (i)  $\Phi_x$  has Lipschitz constant  $L$ , (ii) the inverse of  $\Phi_x$  has Lipschitz constant  $L_{\text{inv}}$ , (iii)  $s \in \mathcal{C}_r(0) \mapsto \Phi_x([s, 0])$  is a regular parameterization of  $\partial U \cap \mathcal{B}_r(x)$  where  $\mathcal{C}_r(0)$  denotes the two-dimensional ball of radius  $r$  centered in the origin. Furthermore, there exists  $\kappa_{\text{max}} > 0$  such that the principal curvatures are uniformly bounded in absolute value by  $\kappa_{\text{max}}$ .*

*Proof.* If  $U$  is of class  $C^2$ , it is in particular a bi-Lipschitz domain. This implies that for any  $x \in \partial U$  there exists  $R = R(x)$  and a bijection  $\Phi_x : A \rightarrow U \cap \mathcal{B}_R(x)$  with  $A = \mathcal{B}_R(0) \cap \{x' : (x')_2 < 0\}$  with the following properties: (i)  $\Phi_x$  has Lipschitz constant  $L$ , (ii) the inverse of  $\Phi_x$  has Lipschitz constant  $L_{\text{inv}}$ , (iii)  $s \in \mathcal{C}_R(0) \mapsto \Phi_x([s, 0])$  is a regular parameterization of  $\partial U \cap \mathcal{B}_R(x)$ . We denote by  $R_{\text{max}} : \partial U \rightarrow \mathbb{R}_+$  the maximum radius for which the above conditions hold: it is easy to verify that  $R_{\text{max}}$  is a continuous function; therefore, since  $\partial U$  is compact, exploiting the Weierstrass theorem, we find that  $R_{\text{max}}$  attains a minimum at some  $x \in \partial U$ ; we thus find  $r := R_{\text{max}}(x) > 0$ .

If  $U$  is of class  $C^2$ , the principal curvatures are continuous over  $\partial U$  and thus uniformly bounded.  $\square$

Then, Lemma B.3 provides a refined version of Lemma 3.3. Note that the bound on  $\delta_0$  depends on the global property  $r$  of  $\partial U$ , and on the local properties  $\kappa_{\max}, L, L_{\text{inv}}$ . The proof involves the introduction of the so-called shape operator (see, e.g., [26, Chapter 5]). Given  $x \in \partial U$ , we denote by  $T_x$  the tangent space at  $x$  and we define the operator  $S[x]$  such that

$$S[x]\mathbf{t} = -\lim_{\epsilon \rightarrow 0} \frac{\mathbf{n}(x + \epsilon\mathbf{t}) - \mathbf{n}(x)}{\epsilon}, \quad \forall \mathbf{t} \in T_x.$$

It is possible to prove that  $S[x]$  is a linear symmetric operator from  $T_x$  to  $T_x$  whose eigenvalues are the principal curvatures of  $\partial U$  at  $x$   $\kappa_1(x), \kappa_2(x)$ . Therefore, we find that the dual norm of  $S[x]$  satisfies

$$\|S[x]\|_* := \sup_{\mathbf{t} \in T_x} \frac{\|S[x]\mathbf{t}\|_2}{\|\mathbf{t}\|_2} = \max\{|\kappa_1(x)|, |\kappa_2(x)|\}. \quad (48)$$

**Lemma B.3.** *Let  $L, L_{\text{inv}}, r, \kappa_{\max} > 0$  be the quantities introduced in Lemma B.2. Then, we find that  $U$  is  $\delta$ -regular for  $\delta < \min\{\frac{r}{2}, \frac{1}{LL_{\text{inv}}\kappa_{\max}}\}$ .*

*Proof.* Let  $x, y \in \partial U$  satisfy

$$x - \delta\mathbf{n}(x) = y - \delta\mathbf{n}(y). \quad (49)$$

Below, we prove that if  $\delta < r/2$  then (49) implies that  $\delta \geq \frac{1}{LL_{\text{inv}}\kappa_{\max}}$ . Recalling Lemma 3.2, the latter implies that  $U$  is guaranteed to be  $\delta$ -regular if  $\delta < r/2$  and  $\delta < \frac{1}{LL_{\text{inv}}\kappa_{\max}}$ .

If  $\delta < r/2$ , we have  $\|x - y\|_2 \leq \delta\|\mathbf{n}(x) - \mathbf{n}(y)\|_2 \leq 2\delta < r$ ; therefore, given  $A = \mathcal{B}_r(0) \cap \{x' : (x')_3 < 0\}$ , there exists a mapping  $\Phi : A \rightarrow U \cap \mathcal{B}_r(x)$  that satisfies the conditions of Lemma B.2. Then, if we denote by  $t, s \in \mathbb{R}^2$  the points satisfying  $x = \Phi([t, 0]), y = \Phi([s, 0])$  such that  $\|s\|_2, \|t\|_2 < r$ , we find

$$\|t - s\|_2 \stackrel{(i)}{\leq} L_{\text{inv}}\|x - y\|_2 \stackrel{(ii)}{=} (L_{\text{inv}}\|\mathbf{n}(x) - \mathbf{n}(y)\|_2)\delta.$$

Note that in (i) we used the fact that  $\Phi^{-1}$  is Lipschitz, while in (ii) we used (49).

Let  $\gamma : \xi \mapsto [(1 - \xi)t + \xi s, 0]$  be a path from  $t$  to  $s$  in  $A$  and define  $\mathbf{N} : \xi \mapsto \mathbf{n}(\Phi(\gamma(\xi)))$ . Clearly, we have  $\mathbf{N}(0) = \mathbf{n}(x)$ ,  $\mathbf{N}(1) = \mathbf{n}(y)$ , and we also have that  $\nabla\Phi(\gamma(\xi))\gamma'(\xi)$  belongs the tangent space  $T_x$  with  $x = \Phi(\gamma(\xi))$ . Then, applying Lemma B.1, the chain rule, and the definition in Eq. (48), we obtain

$$\|\mathbf{n}(x) - \mathbf{n}(y)\|_2 \leq \max_{\xi \in [0,1]} \|\mathbf{N}'(\xi)\| = \max_{\xi \in [0,1]} \|S[\Phi(\gamma(\xi))]\|_* \|\nabla\Phi(\gamma(\xi))\|_2 \|t - s\|_2 \leq \kappa_{\max} L \|t - s\|_2.$$

By combining the latter two equations, we obtain

$$\|t - s\|_2 \leq LL_{\text{inv}}\kappa_{\max} \|t - s\|_2 \delta.$$

By dividing both sides by  $\|t - s\|_2$ , we obtain  $\delta \geq (LL_{\text{inv}}\kappa_{\max})^{-1}$ .  $\square$

## B.5 Lemma 3.4

We first remark that

$$\text{dist}_{\text{H}}(U, V) \leq \text{dist}_{\text{H}}(\partial U, \partial V) \quad (50)$$

We have indeed that  $\sup_{x \in V} \text{dist}(x, U) = \sup_{x \in V \setminus U} \text{dist}(x, U) = \sup_{x \in \partial V \setminus U} \text{dist}(x, U) = \sup_{x \in \partial V \setminus U} \text{dist}(x, \partial U) \leq \text{dist}_{\text{H}}(\partial U, \partial V)$ . Similarly, we verify that  $\sup_{x \in U} \text{dist}(x, V) \leq \text{dist}_{\text{H}}(\partial U, \partial V)$ .

Let  $x \in \partial V$ . By contradiction, assume that  $\text{dist}(x, \partial U) > \delta$ , that is  $\mathcal{B}_\delta(x) \cap \partial U = \emptyset$ , which implies  $\partial U \subset \text{Neigh}_\delta(\partial V) \setminus \mathcal{B}_\delta(x)$ . Since  $V$  is  $\delta$ -regular, we must have that  $U \subset \text{Neigh}_\delta(\partial V)$  and thus  $|U|_{(d)} \leq \text{Neigh}_\delta(\partial V)$ , which contradicts Hypothesis (ii).

In conclusion, we find

$$\text{dist}_{\text{H}}(U, V) \stackrel{(50)}{\leq} \text{dist}_{\text{H}}(\partial U, \partial V) = \max \left\{ \underbrace{\max_{x \in \partial V} \text{dist}(x, \partial U)}_{\leq \delta}, \underbrace{\max_{x \in \partial U} \text{dist}(x, \partial V)}_{\leq \delta} \right\} \leq \delta. \quad \square$$

## C Further investigations for the three-point registration problem

We assess performance of the elasticity-based objectives introduced in Remark 2.2 for the model problem in section 4.1. We consider  $t = 0.7$ ,  $\delta = 10^{-6}$ ,  $n_{\text{lp}} = 25$ ; we set  $\lambda_1 = \frac{\nu}{(1-2\nu)(1+\nu)}$  and  $\lambda_2 = \frac{E}{1+\nu}$  with  $E = 1$  and  $\nu = 0.3$  in (18) and (19). Note that the choice of  $\lambda_1, \lambda_2$  corresponds to the plane strain assumption; note also that, since (18) and (19) are linear with respect to the Young's modulus  $E$ , results are independent of its choice.

Figure 12 shows the results of (10) with (a) linear-elasticity objective, (b) H2 objective, and (c) neo-hookean objective. We observe that both the linear-elasticity model and the neo-hookean model fail to deliver a proper deformed mesh for this test case, while the H2 objective does. The linear-elasticity model also leads to a non-bijective mapping — the minimum Jacobian determinant  $J_{\min}$  is equal to 0.012 for H2, to 0.47 for the neo-hookean model and to  $-0.09$  for linear elasticity. We also remark that the introduction of higher-order derivatives in the objective function has the effect of smoothening the mapping  $\Phi$ : we conjecture that this feature of the approach might be important in the pMOR framework to improve the compressibility of the solution manifold.

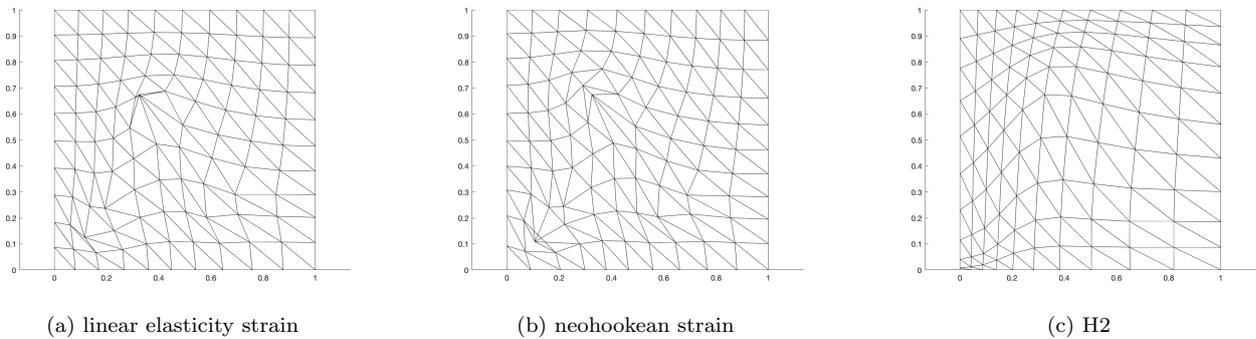


Figure 12: three-point registration; visualization of the deformed meshes for three choices of the objective function in (10).

## References

- [1] Adams, R. A., & Fournier, J. J. (2003). Sobolev spaces. Elsevier.
- [2] Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4). Springer.
- [3] Brenier, Y. (1987). Polar decomposition and increasing rearrangement of vector-fields. *Comptes Rendus de l'Academie Des Sciences Serie I-Mathematique*, 305(19), 805–808.
- [4] Farhat, C., Chapman, T., & Avery, P. (2015). Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models. *International Journal for Numerical Methods in Engineering*, 102(5), 1077–1110.
- [5] Farhat, C., Grimberg, S., Manzoni, A., & Quarteroni, A. (2021). Computational bottlenecks for PROMs: Pre-computation and hyperreduction. *Handbook on Model Order Reduction. Volume 2: Snapshot-Based Methods and Algorithms*. De Gruyter, 2, 181–244.
- [6] Froehle, B., & Persson, P.-O. (2015). Nonlinear elasticity for mesh deformation with high-order discontinuous Galerkin methods for the Navier-Stokes equations on deforming domains. In *Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2014* (pp. 73–85). Springer.
- [7] Gray, A. (2003). *Tubes* (Vol. 221). Springer Science & Business Media.
- [8] Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.
- [9] Hesthaven, J. S., Rozza, G., & Stamm, B. (2016). *Certified reduced basis methods for parametrized partial differential equations* (Vol. 590). Springer.
- [10] Horaud, R., Forbes, F., Yguel, M., Dewaele, G., & Zhang, J. (2010). Rigid and articulated point registration with expectation conditional maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3), 587–602.
- [11] Iollo, A., & Lombardi, D. (2014). Advection modes by optimal mass transfer. *Physical Review E*, 89(2), 022923.
- [12] Iollo, A., & Taddei, T. (2022). Mapping of coherent structures in parameterized flows by learning optimal transportation with Gaussian models. *Journal of Computational Physics*, 111671.

- [13] Ivanov, V. K., Vasin, V. V., & Tanana, V. P. (2013). Theory of linear ill-posed problems and its applications. In *Theory of Linear Ill-Posed Problems and its Applications*. De Gruyter.
- [14] Jacobs, M., & Léger, F. (2020). A fast approach to optimal transport: The back-and-forth method. *Numerische Mathematik*, 146(3), 513–544.
- [15] Knupp, P. M. (2001). Algebraic mesh quality metrics. *SIAM Journal on Scientific Computing*, 23(1), 193–218.
- [16] Lassila, T., Manzoni, A., Quarteroni, A., & Rozza, G. (2014). Model order reduction in fluid dynamics: challenges and perspectives. In *Reduced Order Methods for modeling and computational reduction* (pp. 235–273). Springer.
- [17] Løvgrén, A. E., Maday, Y., & Rønquist, E. M. (2006). A reduced basis element method for the steady Stokes problem. *ESAIM: Mathematical Modelling and Numerical Analysis*, 40(3), 529–552.
- [18] Ma, J., Qiu, W., Zhao, J., Ma, Y., Yuille, A. L., & Tu, Z. (2015). Robust  $L_2$   $E$  estimation of transformation for non-rigid registration. *IEEE Transactions on Signal Processing*, 63(5), 1115–1129.
- [19] Ma, J., Wu, J., Zhao, J., Jiang, J., Zhou, H., & Sheng, Q. Z. (2018). Nonrigid point set registration with robust transformation learning under manifold regularization. *IEEE Transactions on Neural Networks and Learning Systems*, 30(12), 3584–3597.
- [20] Maiseli, B., Gu, Y., & Gao, H. (2017). Recent developments and trends in point set registration methods. *Journal of Visual Communication and Image Representation*, 46, 95–106.
- [21] Manzoni, A., & Negri, F. (2017). Efficient reduction of PDEs defined on domains with variable shape. In *Model Reduction of Parametrized Systems* (pp. 183–199). Springer.
- [22] Manzoni, A., Quarteroni, A., & Rozza, G. (2012). Model reduction techniques for fast blood flow simulation in parametrized geometries. *International Journal for Numerical Methods in Biomedical Engineering*, 28(6–7), 604–625.
- [23] Meng, X.-L., & Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2), 267–278.
- [24] Myronenko, A., & Song, X. (2010). Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12), 2262–2275.
- [25] Myronenko, A., Song, X., & Carreira-Perpinan, M. (2006). Non-rigid point set registration: Coherent point drift. *Advances in Neural Information Processing Systems*, 19.
- [26] O’Neill, B. (2006). *Elementary differential geometry*. Elsevier.
- [27] Quarteroni, A., Manzoni, A., & Negri, F. (2015). *Reduced basis methods for partial differential equations: an introduction* (Vol. 92). Springer.
- [28] Rocha, H. (2009). On the selection of the most adequate radial basis function. *Applied Mathematical Modelling*, 33(3), 1573–1583.
- [29] Rozza, G., Huynh, D. B. P., & Patera, A. T. (2008). Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3), 229–275.
- [30] Staten, M. L., Owen, S. J., Shontz, S. M., Salinger, A. G., & Coffey, T. S. (2011). A comparison of mesh morphing methods for 3D shape optimization. In *Proceedings of the 20th international meshing roundtable* (pp. 293–311). Springer.
- [31] Taddei, T. (2020). A registration method for model order reduction: data compression and geometry reduction. *SIAM Journal on Scientific Computing*, 42(2), A997–A1027.
- [32] Taddei, T., & Zhang, L. (2021a). A discretize-then-map approach for the treatment of parameterized geometries in model order reduction. *Computer Methods in Applied Mechanics and Engineering*, 384, 113956.
- [33] Taddei, T., & Zhang, L. (2021b). Registration-based model reduction in complex two-dimensional geometries. *Journal of Scientific Computing*, 88(3), 79.

- [34] Tezduyar, T. E., Behr, M., Mittal, S., & Liou, J. (1992). A new strategy for finite element computations involving moving boundaries and interfaces—the deforming-spatial-domain/space-time procedure: II. Computation of free-surface flows, two-liquid flows, and flows with drifting cylinders. *Computer Methods in Applied Mechanics and Engineering*, 94(3), 353–371.
- [35] Tonon, P., Sanches, R. A. K., Takizawa, K., & Tezduyar, T. E. (2021). A linear-elasticity-based mesh moving method with no cycle-to-cycle accumulated distortion. *Computational Mechanics*, 67(2), 413–434.
- [36] Volkwein, S. (2011). Model reduction using proper orthogonal decomposition. Lecture Notes, Institute of Mathematics and Scientific Computing, University of Graz. See [Math.Uni-Konstanz.de/Numerik/Personen/Volkwein/Teaching/POD-Vorlesung.Pdf](http://Math.Uni-Konstanz.de/Numerik/Personen/Volkwein/Teaching/POD-Vorlesung.Pdf), 1025.
- [37] Wendland, H. (2004). *Scattered data approximation* (Vol. 17). Cambridge university press.
- [38] Yano, M. (2021). Model reduction in computational aerodynamics. *Handbook on Model Order Reduction. Volume 3: Applications*. De Gruyter, 201–236.
- [39] Zahr, M. J., & Persson, P.-O. (2018). An optimization-based approach for high-order accurate discretization of conservation laws with discontinuous solutions. *Journal of Computational Physics*, 365, 105–134.
- [40] Zitova, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11), 977–1000.