Using parametric model order reduction for inverse analysis of large nonlinear cardiac simulations

Pfaller, M. R.; Cruz Varona, M.; Lang, J.; Bertoglio, C.; Wall, W. A.

Published in:
International journal for numerical methods in biomedical engineering

DOI:
10.1002/cnm.3320

Publication date:
2020

Received: 17 October 2018 | Revised: 29 October 2019 | Accepted: 25 January 2020

DOI: 10.1002/cnm.3320

WILEY

RESEARCH ARTICLE - FUNDAMENTAL

Using parametric model order reduction for inverse analysis of large nonlinear cardiac simulations

M. R. Pfaller[1] | M. Cruz Varona[2] | J. Lang[1] | C. Bertoglio[3] | W. A. Wall[1]

[1]Institute for Computational Mechanics, Technical University of Munich, Garching b. München, Germany

[2]Chair of Automatic Control, Technical University of Munich, Garching b. München, Germany

[3]Bernoulli Institute, University of Groningen, AG Groningen, The Netherlands

Correspondence
M. R. Pfaller, Institute for Computational Mechanics, Technical University of Munich, Boltzmannstr. 15, 85748 Garching b. München, Germany.
Email: martin.pfaller@tum.de

Abstract

Predictive high-fidelity finite element simulations of human cardiac mechanics commonly require a large number of structural degrees of freedom. Additionally, these models are often coupled with lumped-parameter models of hemodynamics. High computational demands, however, slow down model calibration and therefore limit the use of cardiac simulations in clinical practice. As cardiac models rely on several patient-specific parameters, just one solution corresponding to one specific parameter set does not at all meet clinical demands. Moreover, while solving the nonlinear problem, 90% of the computation time is spent solving linear systems of equations. We propose to reduce the structural dimension of a monolithically coupled structure-Windkessel system by projection onto a lower-dimensional subspace. We obtain a good approximation of the displacement field as well as of key scalar cardiac outputs even with very few reduced degrees of freedom, while achieving considerable speedups. For subspace generation, we use proper orthogonal decomposition of displacement snapshots. Following a brief comparison of subspace interpolation methods, we demonstrate how projection-based model order reduction can be easily integrated into a gradient-based optimization. We demonstrate the performance of our method in a real-world multivariate inverse analysis scenario. Using the presented projection-based model order reduction approach can significantly speed up model personalization and could be used for many-query tasks in a clinical setting.

KEYWORDS

cardiac mechanics, inverse analysis, parametric model order reduction, proper orthogonal decomposition

1 | INTRODUCTION

Cardiac solid mechanics simulations consist of solving large-deformation, materially nonlinear, elastodynamic coupled boundary-value problems. There exist different approaches to incorporate blood flow into the computational model. Three-dimensional fluid-structure interaction is resolved for example in References 1 and 2. As the exact fluid dynamics of blood within the heart are, however, usually not needed, the structural model is commonly coupled to

lumped-parameter fluid models which provide the pressure to the endocardial wall.[3] These so-called 0D Windkessel models are then coupled to cardiac solid mechanics. For a comprehensive review of models currently utilized in cardiac mechanics the reader is referred to Reference 4. Different 3D-0D structure-Windkessel coupling approaches are available in literature. A partitioned scheme was used for example in Reference 5. Due to its superior robustness, we use a monolithically coupled model in this work, as for example also used previously in References 6-9.

1.1 | Model order reduction

The needed huge number of degrees of freedom (DOFs) and other challenges of solving nonlinear problems make the solution of cardiac models computationally expensive and limit the models' use in clinical practice. For example, using a single node with 24 cores, a simulation of one heartbeat, which takes about 1 s in reality, takes about 1 d to compute with our high-fidelity four chamber model.[9] The potential to reduce computation time motivates the use of reduced order models (ROMs). In this work, we solely consider model order reduction (MOR) of time-dependent parametric problems. In the following, different strategies in reduced order modeling are reviewed.

An important category of cardiac ROMs is made up by simplified modeling. For these models, the same system of differential equations as for the full order model (FOM) is solved, but on a simplified analytical geometry. The displacements are commonly parameterized by only one scalar DOF. These models are thus referred to as 0D models. Examples in this category include monoventricular cylindrical,[10] spherical,[11] or prolate spheroid[12] or biventricular[13] geometries. These models allow extremely fast evaluation, with computation times less than 1 s. Their results are, however, only lumped quantities which usually need an extra correction step in order to predict the solution of a corresponding patient-specific 3D model.

Another approach of MOR in biomechanics is the use of coarsely discretized geometries; see, for example, References 8 and 14. Coarsely discretized models are easy to implement, since the computational framework is identical to the one of the FOM. The disadvantage of using coarsely discretized geometries is that there is no exact control over the approximation quality and important features of the FOM might not be preserved by the ROM.

A third category of cardiac ROMs makes a model computationally less expensive by reducing the dimension of the problem, starting from the FOM. These ROMs are utilized in this work. For example in cardiac electrophysiology, approximated lax pairs for propagating wave fronts were proposed in References 15 and 16. A local reduced basis method for parameterized cardiac electrophysiology was recently introduced in Reference 17. Reduced basis methods were proposed for general large deformation, material nonlinear finite element simulations.[18,19] A framework for linear coupled multiphysics problems was introduced in Reference 20.

Using our code,[21] for a large-scale cardiac finite element simulation, about 90% of the time is spent iteratively solving linear systems of equations. This proportion motivates the use of model order reduction by projection, where the full linear system is projected onto a much smaller dimensional subspace while preserving the model's most relevant features. The solution of the FOM is then approximated by a solution in the reduced space with a ROM. A popular method to generate such subspaces is proper orthogonal decomposition (POD), which is purely observation-based and independent of the underlying physics of the model. The snapshots, in our case transient observations of displacements, can be obtained from a FOM simulation of one heartbeat.

There are only few examples where POD has been applied to cardiac problems. A quasi-static cardiac model was reduced using POD and combined with hyper-reduction techniques in Reference 22. However, analysis was only carried out using an idealized ellipsoidal left ventricular geometry with few DOFs. While this is very instructive, results for speedup and accuracy of the ROM are not conclusive for real-world cardiac problems. The reduction of a patient-specific biventricular cardiac model using POD is described in Reference 23. In their work, a general approach is presented and analyzed mathematically, before being applied to the example of a patient-specific beating heart model. Parameter estimation is performed based on medical images. A ROM of blood dynamics in coronary arteries is used in Reference 24.

Cardiac models rely on a large set of patient-specific parameters, describing constitutive behavior, hemodynamics, boundary conditions, or local fiber orientation. In order not to rely on a FOM simulation for each new ROM simulation, which would render the ROM simulation useless, the reduced subspace must be able to adapt to a changing parameter set. This adaption requires parametric model order reduction (pMOR) of the time-dependent problem. Among many global and local pMOR techniques, various subspace interpolation methods have been proposed in the past.[25] Specifically, a popular method using a Grassmann manifold was proposed in Reference 26 and illustrated with a large coupled aeroelastic model of a fighter jet. The method proposed in Reference 23 uses a so-called “multi-POD” approach. A parameter-weighted variant of this approach is also used in this work and described in section 3.3.2 as weighted

concatenation of snapshots method. Furthermore, a global pMOR approach using a global basis over the whole parameter range is employed in Reference 22.

The performance of POD in realistic simulations of cardiac contraction is yet unknown. We demonstrate in this work the performance of POD applied to a patient-specific cardiac geometry with about 850 000 structural DOFs. In this work, we consider for the first time the case of a POD-reduced 3D structural model that is monolithically coupled to a 0D Windkessel model, where we only reduce the structural dimension of the problem. Additionally, we compare several subspace interpolation methods for cardiac contraction. In these parametric simulations, we vary the contractility parameter controlling maximum active tension of the myofibers in our model, as it is the most influential parameter for cardiac function and commonly calibrated to experiments.

1.2 | Inverse analysis

Many of the cardiac model parameters depend on a patient's physiology and are a priori unknown, as invasive experiments cannot be carried out on living human subjects. A predictive patient-specific cardiac model is thus subject to an iterative process termed *inverse analysis*. In this context, the simulation of one heartbeat with given parameters can be regarded as the forward problem. The reverted task of matching the parameters to given observations from the patient-specific heartbeat is then the inverse problem. Common clinical measurements of cardiac kinematics are displacement data extracted from cine or tagged magnetic resonance imaging (MRI), representing an Eulerian and Lagrangian description of motion, respectively. Other measurements include blood pressure or electrocardiograms. As cardiac mechanics simulations pose an expensive forward problem, repeated evaluation during inverse analysis has incredible computational demands. Furthermore, algorithms for inverse analysis commonly scale linearly with the number of parameters. Inverse analysis is thus a promising application of reduced order modeling.

During gradient-based optimization, the adjoint method offers computationally inexpensive gradient calculation when combining it with advanced methods such as automatic differentiation and checkpointing.[27] For example in Reference 28, regional contractility was estimated from short axis cine MRI. Using the adjoint method, ischemic regions in cardiac electrophysiology were identified in Reference 29. Most recently in Reference 30, passive material parameters and active fiber shortening was estimated for a biventricular geometry from ventricular volume and regional strain.

Gradient-free inverse analysis for cardiac problems was demonstrated in Reference 31, where regional cardiac contractility was estimated from cine MRI using the unscented Kalman filter (UKF). The reduced order UKF was further applied in References 32 and 33 to estimate boundary condition parameters of the aorta for a fluid-structure interaction problem. Other examples of gradient-free inverse analysis include,[34] where left-ventricular active and passive material parameters were estimated from 3D tagged MRI using a parameter sweep.

A good overview of using pMOR for inverse analysis is given in References 25 and 35. One example is to optimize over the ROM within a regularly updated trust region around the FOM; see, for example, Reference 36. There are some examples, where reduced order modeling has been combined with inverse analysis in biomechanics. For arterial hemodynamic fluid-structure interaction problems, an inverse analysis with uncertainty quantification was performed in Reference 37 using a reduced basis method. There are however few references for cardiac solid models. In Reference 23, the reduced order UKF was applied to estimate cardiac contractility in a healthy and an infarcted region. The forward simulations were carried out using POD. In Reference 8 a multifidelity approach was proposed to calibrate hemodynamical and structural parameters of a cardiac model to ventricular pressure measurements. Here, a Levenberg-Marquardt-based optimization uses evaluations switching between a 3D FOM, a coarsely discretized version of the 3D FOM, and a 2D surrogate model. Another multifidelity approach was used in Reference 38 between a 3D FOM and a 0D surrogate model. An evolutionary algorithm was used in Reference 39 using a ROM with a pre-computed POD-basis from a single FOM to identify four parameters of an electrophysiological cardiac model from a synthetic electrocardiogram. As in Reference 23, a reduced order UKF was used in Reference 40 to estimate parameters of a one-way coupled electromechanical cardiac model from synthetic data. Here, as well, a POD basis was constructed a priori from, in this case, four pre-computed parameter sample sets.

Using coarsely discretized or surrogate models does however not guarantee that the most important features of the FOM are preserved. These surrogate models further lack the ability of pMOR to inherently "learn" from evaluations of the FOM to become more precise throughout the optimization. Instead, they require an additional mapping between FOM and surrogate model solutions. Most importantly, using 2D or 0D surrogate models during inverse analysis, the heart can only be tuned to scalar measurements. However, a calibration to spatial measurements from cine or tagged MRI might be desired in many applications, for example, when detecting infarcted regions.[31] Furthermore, an a priori generation of the

ROM might not be suitable for non-convex approximation problems with several parameters as it is computationally expensive and the parameter range might not be known a priori. In this work, we thus propose a novel method how an automatically updated ROM can be integrated into any optimization-based inverse analysis leading to considerable savings in CPU time, and demonstrate its performance in a real-world multivariate inverse analysis scenario.

The remainder of this work is structured as follows. In section 2, we introduce the full order elastodynamic and hemodynamic models. We derive a reduced formulation for the monolithically coupled system in section 3 and review several pROM subspace interpolation methods. In numerical experiments in section 4, we demonstrate the accuracy and speedup of our ROM and show its response to parametric variations. Furthermore, we propose in section 5 a pMOR-based method for inverse analysis and analyze its performance with our four-chamber heart model. We close with a conclusion and future perspectives in section 6.

2 | 3D-0D COUPLED CARDIOVASCULAR MODELING

In this section we give a brief overview of our full order model (FOM) which is composed of a 3D elastodynamical model (see section 2.1) coupled to a 0D hemodynamical model (see section 2.2). We further provide insights into the numerical solution of the model in section 2.3. For a more detailed description, the reader is referred to Reference 9.

2.1 | 3D elastodynamical model

We follow the classic approach of nonlinear large deformation continuum mechanics to model the elastodynamic problem of 3D cardiac contraction. We define the reference configuration $\boldsymbol{X}$ and the current configuration x, which are connected by the displacements $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{X}$. We calculate the deformation gradient $\boldsymbol{F}$, the right Cauchy-Green tensor $\boldsymbol{C}$, and the Green-Lagrange strain tensor $\boldsymbol{E}$,

$$\boldsymbol{F} = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}}, \quad \boldsymbol{C} = \boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}, \quad \boldsymbol{E} = \frac{1}{2}(\boldsymbol{C} - \boldsymbol{I}). \tag{1}$$

Balance of momentum, a Neumann Windkessel coupling condition with left ventricular pressure p_v, omnidirectional spring-dashpot boundary conditions, and pericardial boundary conditions yield the weak form of the 3D elastodynamic boundary value problem in the reference configuration

$$\int_{\Omega_0} \left[\rho_0 \ddot{\mathbf{u}} \cdot \delta\boldsymbol{u} + \boldsymbol{S} : \delta\boldsymbol{E}\right] \mathrm{d}V + \int_{\Gamma_0^{\mathrm{endo}}} p_v J \boldsymbol{F}^{-\mathrm{T}} \cdot \boldsymbol{N} \cdot \delta\boldsymbol{u} \mathrm{d}A + \int_{\Gamma_0^{\mathrm{vess}}} \left[k_v \boldsymbol{u} + c_v \dot{\mathbf{u}}\right] \cdot \delta\boldsymbol{u} \mathrm{d}A + \int_{\Gamma_0^{\mathrm{epi}}} \boldsymbol{N}\left[k_e \boldsymbol{u} \cdot \boldsymbol{N} + c_e \dot{\mathbf{u}} \cdot \boldsymbol{N}\right] \cdot \delta\boldsymbol{u} \mathrm{d}A = 0, \tag{2}$$

with density ρ_0, accelerations $\ddot{\boldsymbol{u}}$, virtual displacements and strains $\delta\boldsymbol{u}$ and $\delta\boldsymbol{E}$, respectively, the Jacobian $J = \det \boldsymbol{F}$ of the deformation gradient, the second Piola-Kirchhoff stress tensor $\boldsymbol{S}$, the reference surface normal $\boldsymbol{N}$, and spring stiffness k_v, k_e and viscosity c_v, c_e for vessel and epicardial surface, respectively. We define three surfaces for the imposition of boundary conditions Γ_0^{endo}, Γ_0^{vess}, and Γ_0^{epi} at the left endocardium, the outside of the great vessels, and the epicardium respectively. At the cut-offs of the great vessels we apply homogeneous Dirichlet boundary conditions. We define different nonlinear materials

$$\text{Adipose tissue}: \quad \boldsymbol{S} = \frac{\partial}{\partial \boldsymbol{E}}(\psi_{\mathrm{NH}} + \psi_{\mathrm{vol}}) + \frac{\partial}{\partial \dot{\mathbf{E}}}\psi_{\mathrm{visco}}, \tag{3}$$

$$\text{Aorta, pulmonary, atrial myocardium}: \quad \boldsymbol{S} = \frac{\partial}{\partial \boldsymbol{E}}(\psi_{\mathrm{MR}} + \psi_{\mathrm{vol}}) + \frac{\partial}{\partial \dot{\mathbf{E}}}\psi_{\mathrm{visco}}, \tag{4}$$

$$\text{Ventricular myocardium}: \quad \boldsymbol{S} = \frac{\partial}{\partial \boldsymbol{E}}(\psi_{\mathrm{MR}} + \psi_{\mathrm{vol}}) + \frac{\partial}{\partial \dot{\mathbf{E}}}\psi_{\mathrm{visco}} + S_{\mathrm{act}}, \tag{5}$$

depending on the (pseudo-)potentials

$$\psi_{\text{NH}} = \frac{\mu}{2}(\bar{I}_1 - 3), \quad \psi_{\text{MR}} = C_1(\bar{I}_1 - 3) + C_2(\bar{I}_2 - 3), \quad \psi_{\text{vol}} = \kappa\left(J + \frac{1}{J} - 2\right), \quad \psi_{\text{visco}} = \frac{\eta}{8}\text{tr}\left(\dot{\bar{\mathbf{C}}}2\right),$$
$$\bar{\mathbf{C}} = J^{-2/3}\boldsymbol{C}, \quad \bar{I}_1 = \text{tr}(\bar{\mathbf{C}}), \quad \bar{I}_2 = \frac{1}{2}\left[\text{tr}^2(\bar{\mathbf{C}}) - \text{tr}\left(\bar{\mathbf{C}}^2\right)\right], \tag{6}$$

with isochoric invariants $\bar{I}_1, \bar{I}_2$, material parameters μ, C_1, C_2, η, and penalty parameter κ. The geometry and the different materials are shown in Figure 1. Each is composed of a hyperelastic potential ψ_{NH} or ψ_{MR} and a viscous contribution ψ_{visco}, depending on the rate of the right Cauchy-Green tensor $\dot{\boldsymbol{C}}$. Quasi-incompressibility is enforced by the penalty potential ψ_{vol}. The passive material behavior of myocardial tissue is known to be orthotropic; see, for example, References 41 and 42 and was also modeled via an orthotropic model in our previous work.[9] In this paper, however, the passive myocardium is simply modeled as isotropic, as it was also done for example in References 6 and 31. However, this should only have a minor influence on MOR results presented in this work.

The ventricular myocardium has an additive active stress component S_{act} modeling the contraction of myofiber bundles in reference fiber direction f_0

$$\boldsymbol{S}_{\text{act}} = \tau(t) \cdot f_0 \otimes f_0, \quad \text{with } \dot{\tau}(t) = -\,|\,u(t)\,|\,\tau(t) + \sigma|u(t)|_+, \quad \tau(0) = 0, \tag{7}$$

with active stress $\tau \in [0, \sigma[$ and the function $|u(t)|_+ = \max(u(t); 0)$. The contractility parameter σ controls the upper limit of the active stress component. The prescribed activation function $u(t)$ is

$$u(t) = \alpha_{\max} \cdot f(t) + \alpha_{\min} \cdot [1 - f(t)], \quad f(t) = S^+\left(t - t_{\text{sys}}\right) \cdot S^-(t - t_{\text{dias}}), \quad S^{\pm}(\Delta t) = \frac{1}{2}\left[1 \pm \tanh\left(\frac{\Delta t}{\gamma}\right)\right], \tag{8}$$

with steepness $\gamma = 0.005$ s and descending and ascending sigmoid functions S^+ and S^-, respectively. The indicator function $f \in]0, 1[$ indicates ventricular systole. The times t_{sys} and t_{dias} model the onset of systole and diastole, respectively. The times t_{sys} and t_{dias}, and maximum and minimum myocyte activation rates $\alpha_{\max}$ and $\alpha_{\min}$, respectively, are calibrated to match the timing of ventricular systole as observed in cine MRI, as demonstrated in section 5.

We discretize displacements u and virtual displacements δu arising in the weak form 2 using quadratic basis functions on each tetrahedral finite element. Assembly of the discretized problem leads to the matrix notation of the spatially semi-discrete residual of the full order structural model

$$\mathbf{R}^{\text{S}}_{\text{semi}} = \mathbf{M}\ddot{\mathbf{d}} + \mathbf{F}\left(\dot{\mathbf{d}}, \mathbf{d}, p_v\right) \stackrel{!}{=} 0, \tag{9}$$

with mass matrix $\mathbf{M}$, nonlinear force vector $\mathbf{F}$, and nodal displacements, velocities, and accelerations $\mathbf{d}$, $\dot{\mathbf{d}}$, and $\ddot{\mathbf{d}}$, respectively. We discretize the boundary value problem implicitly in time using a combination of Newmark's method[43]

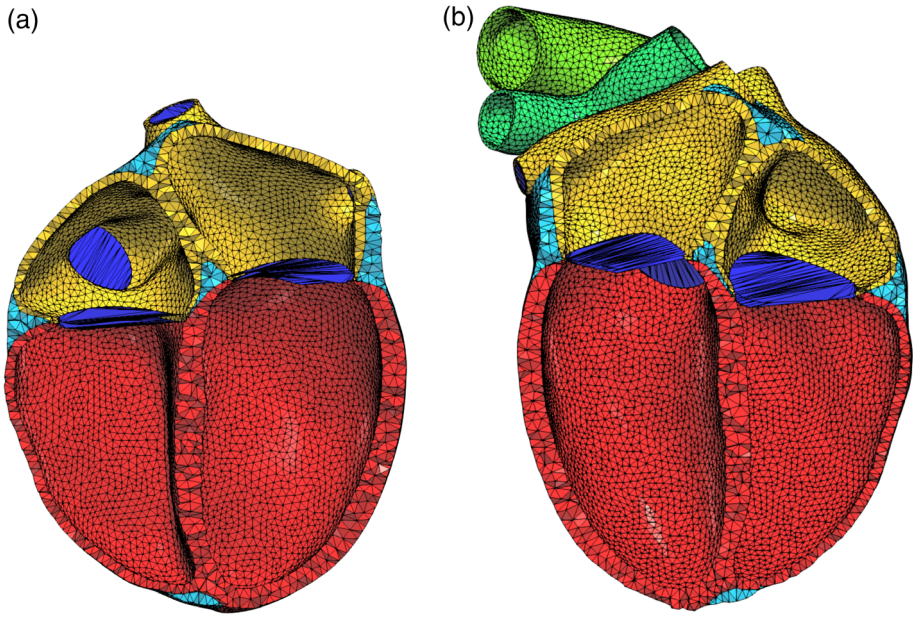

FIGURE 1 Computational mesh with quadratic tetrahedral elements cut in four-chamber view, colored by different materials: adipose tissue (cyan), atrial myocardium (yellow), ventricular myocardium (red), aorta and pulmonary artery (green), valve planes (blue). (a) Posterior view. (b) Anterior view

and the generalized-α scheme[44] to obtain the time and space discrete structural residual $\mathbf{R}^{\mathrm{S}}(\mathbf{d}_{j+1}, p_{v,\,j+1})$ at time step $j+1$. All considerations in this work are based on an implicit time integration scheme, as is it the default choice in cardiac mechanics for its improved time-stability properties.

2.2 | 0D hemodynamical model

We couple the left ventricular 3D structural model to a 0D lumped-parameter Windkessel model of the circulatory system. We utilize in this work a four element Windkessel, using resistances R, compliances C, and an inertance L_p. Pressures at different parts of the model are denoted by p_i. We distinguish between a proximal part p and a distal part d of the aorta. The atrial pressure $p_{at}(t)$ is prescribed to simulate atrial systole. The venous pressure p_{ref} is kept constant. We model the atrioventricular and semilunar valves with a smooth diode-like behavior by nonlinear resistances $R_{av} := R(p_{at}, p_v)$ and $R_{sl} := R(p_v, p_p)$, respectively, depending on a sigmoid function R. This yields the set of differential equations

$$\frac{p_v - p_{at}}{R_{av}} + \frac{p_v - p_p}{R_{sl}} + \dot{V}(u) = 0, \quad q_p - \frac{p_v - p_p}{R_{sl}} + C_p \dot{p}_p = 0, \quad q_p + \frac{p_d - p_p}{R_p} + \frac{L_p}{R_p}\dot{q}_p = 0, \quad \frac{p_d - p_{\mathrm{ref}}}{R_d} - q_p + C_d \dot{p}_d = 0, \tag{10}$$

which are coupled to the 3D structural model by the ventricular pressure p_v and the change in ventricular volume $\dot{V}$, depending on the structural displacements u. The vector of primary variables yields $\mathbf{p} = [p_v, p_p, p_d, q_p]^{\mathrm{T}}$, including the flux q_p through the inertance L_p. We discretize the set of Windkessel Equations 10 in time with the one-step-θ scheme.[45] This yields the discrete Windkessel residual $\mathbf{R}^{0\mathrm{D}}(\mathbf{d}_{j+1}, \mathbf{p}_{j+1})$ at time step $j+1$.

2.3 | Solving the coupled problem

We solve the coupled 3D-0D model with the structural and Windkessel residuals $\mathbf{R}^{\mathrm{S}}$ and $\mathbf{R}^{0\mathrm{D}}$, respectively, for the displacements $\mathbf{d}$ and Windkessel variables $\mathbf{p}$ at time step $j+1$ with the Newton-Raphson method

$$\begin{bmatrix} \frac{\partial \mathbf{R}^{\mathrm{S}}}{\partial \mathbf{d}} & \frac{\partial \mathbf{R}^{\mathrm{S}}}{\partial \mathbf{p}} \\ \frac{\partial \mathbf{R}^{0\mathrm{D}}}{\partial \mathbf{d}} & \frac{\partial \mathbf{R}^{0\mathrm{D}}}{\partial \mathbf{p}} \end{bmatrix}^{i}_{j+1} \cdot \begin{bmatrix} \Delta \mathbf{d} \\ \Delta \mathbf{p} \end{bmatrix}^{i+1}_{j+1} = - \begin{bmatrix} \mathbf{R}^{\mathrm{S}} \\ \mathbf{R}^{0\mathrm{D}} \end{bmatrix}^{i}_{j+1}, \quad \text{with } \mathbf{d}, \Delta\mathbf{d}, \mathbf{R}^{\mathrm{S}} \in \mathbb{R}^{846864} \text{ and } \mathbf{p}, \Delta\mathbf{p}, \mathbf{R}^{0\mathrm{D}} \in \mathbb{R}^{4}, \tag{11}$$

linearizing the residuals in iteration i. The solution is converged if

$$\left\|\mathbf{R}^{\mathrm{S}}\right\|_{\infty} < tol^{\mathrm{S}}_{\mathrm{res}}, \quad \|\Delta\mathbf{d}\|_{\infty} < tol^{\mathrm{S}}_{\mathrm{inc}}, \quad \left\|\mathbf{R}^{0\mathrm{D}}\right\|_{2} < tol^{0\mathrm{D}}_{\mathrm{res}}, \quad \|\Delta\mathbf{p}\|_{2} < tol^{0\mathrm{D}}_{\mathrm{inc}}, \tag{12}$$

with the structural and Windkessel residual and increment tolerances $tol^{\mathrm{S}}_{\mathrm{res}}$, $tol^{0\mathrm{D}}_{\mathrm{res}}$, $tol^{\mathrm{S}}_{\mathrm{inc}}$, and $tol^{0\mathrm{D}}_{\mathrm{inc}}$, respectively. Note that the coupled model in Equation 11 is independent of the concrete formulation of the structural and Windkessel models from sections 2.1 and 2.2, respectively. It is valid for any arbitrary residuals $\mathbf{R}^{\mathrm{S}}$ and $\mathbf{R}^{0\mathrm{D}}$. For details of the model used here, again see Reference 9.

3 | NONLINEAR PARAMETRIC MODEL ORDER REDUCTION

The 3D-0D cardiovascular model described above represents a large-scale, nonlinear, parametrized, and monolithically coupled model, featuring multiple sources of nonlinear system behavior and depending on several model parameters. First, our model contains geometric nonlinearities, due to the use of the Green-Lagrange strain tensor $E(u)$. Second, the utilized material laws for myocardial tissue induce material nonlinearity. The third and last source of nonlinearity is given by the nonlinear coupling between the structural and the hemodynamical model due to the Neumann Windkessel boundary condition, acting in direction of the current normal vector of the endocardium. Furthermore, our

model depends on many parameters $\boldsymbol{\mu} = \left[\mu_1, ..., \mu_{n_p}\right]^{\mathrm{T}} \in \Omega \subset \mathbb{R}^{n_p}$, classified in different categories. For instance, there exist parameters describing the constitutive behavior of the used materials (stiffness, viscosity, and incompressibility parameters), the additive active stress component $\boldsymbol{S}_{\mathrm{act}}$ (eg, the contractility σ, $\alpha_{\max}$, $\alpha_{\min}$, t_{sys}, t_{dias}), the hemodynamics (eg, resistances R, compliances C, inertance L_p), as well as the boundary conditions for the outside of the great vessels and the epicardium (spring stiffnesses k_v, k_e and dashpot viscosities c_v, c_e). Thus, our discrete nonlinear parametrized FOM reads

$$\mathbf{R}(\mathbf{d}, \mathbf{p}, \boldsymbol{\mu}) = \begin{bmatrix} \mathbf{R}^{\mathrm{S}}(\mathbf{d}, p_v, \boldsymbol{\mu}) \\ \mathbf{R}^{\mathrm{0D}}(\mathbf{d}, \mathbf{p}, \boldsymbol{\mu}) \end{bmatrix}_{j+1} \overset{!}{=} \mathbf{0}. \tag{13}$$

The use of a 0D lumped-parameter Windkessel model, instead of, for example, a 3D fluid dynamics model of the heart chambers and arteries, already simplifies the computational complexity of the coupled system. However, the numerical analysis of the present model still demands a high computational effort due to the large number of structural DOFs. While this is no problem for a few standard forward simulations, it is extremely challenging—and might even prohibit—fast model calibration, inverse analysis and clinical applications. Therefore, our aim is to employ projection-based model order reduction to obtain a cardiovascular reduced order model (ROM) that accurately approximates the original model with substantially less DOFs and, consequently, less numerical effort. To this end, in section 3.1 we first apply the classical projection-based model order reduction framework to the structural component of our cardiovascular problem and further describe the numerical solution of the coupled ROM. A suitable strategy to compute the required projection matrix $\mathbf{V}$ for a fixed parameter set is then explained in section 3.2. Afterwards, different subspace interpolation techniques are presented in section 3.3, in order to compute a parametric reduced order model (pROM) for any new parameter set. Finally, some implementation details are given in section 3.4.

3.1 | Cardiovascular reduced order model

As mentioned before, the high dimension of the FOM comes from the discretization of the 3D elastodynamical model, whereas the 0D model only contributes few, in our case four, Windkessel DOFs in $\mathbf{p}$. Consequently, we only apply model order reduction to the structural component while the 0D hemodynamical model remains unchanged. The process of reducing the FOM within the computational framework is outlined in Algorithm 1.

Algorithm 1 *Projection-based model order reduction.*

1. generate projection matrix $\mathbf{V}$ offline with methods from sections 3.2 (constant) or 3.3 (parametric)
2. **for** time step $j = 0, ..., n_s$ **do**
3. Newton iteration $i = 0$
4. **while** convergence criterion from (18) not fulfilled **do**
5. evaluate and assemble full Jacobian and residual (11)
6. reduce structural dimensions in Jacobian and residual (16)
7. solve reduced linear system (16)
8. approximate full displacements (14)
9. update solution (17)
10. $i \leftarrow i + 1$
11. **end while**
12. **return** solution $\mathbf{d}_j$ and $\mathbf{p}_j$
13. **end for**

Our aim is to approximate the high dimensional structural solution $\mathbf{d} \in \mathbb{R}^n$ using a linear combination of $q \ll n$ basis vectors, with n being the number of DOFs of the full elastodynamical model. With the set of basis vectors $\mathbf{v}_i \in \mathbb{R}^n$ contained in a projection matrix $\mathbf{V} = [\mathbf{v}_1 ... \mathbf{v}_q] \in \mathbb{R}^{n \times q}$ and the vector of the reduced model's DOFs $\mathbf{d}_{\mathrm{r}} \in \mathbb{R}^q$, the approximation is

$$\mathbf{d} \approx \mathbf{V}\mathbf{d}_\mathrm{r}. \tag{14}$$

To obtain a square system, we project the structural residual onto the space spanned by $\mathbf{V}^\mathrm{T}$ yielding the spatially semi-discrete reduced residual $\mathbf{R}^\mathrm{S}_{\mathrm{semi,r}} \in \mathbb{R}^q$,

$$\mathbf{R}^\mathrm{S}_{\mathrm{semi,r}} = \mathbf{V}^\mathrm{T}\mathbf{R}^\mathrm{S}_{\mathrm{semi}}\left(\ddot{\mathbf{V}\mathbf{d}}_\mathrm{r}, \dot{\mathbf{V}\mathbf{d}}_\mathrm{r}, \mathbf{V}\mathbf{d}_\mathrm{r}, p_v, \boldsymbol{\mu}\right). \tag{15}$$

This represents the Galerkin projection of the full structural residual $\mathbf{R}^\mathrm{S}_{\mathrm{semi}}$ onto the subspace $\mathcal{V}$ spanned by the vectors in $\mathbf{V}$. After discretization in time, we obtain the Newton-Raphson update $i+1$ at time step $j+1$ of the coupled reduced order model,

$$\begin{bmatrix} \mathbf{V}^\mathrm{T}\dfrac{\partial \mathbf{R}^\mathrm{S}}{\partial \mathbf{d}}\mathbf{V} & \mathbf{V}^\mathrm{T}\dfrac{\partial \mathbf{R}^\mathrm{S}}{\partial \mathbf{p}} \\ \dfrac{\partial \mathbf{R}^{0\mathrm{D}}}{\partial \mathbf{d}}\mathbf{V} & \dfrac{\partial \mathbf{R}^{0\mathrm{D}}}{\partial \mathbf{p}} \end{bmatrix}^i_{j+1} \cdot \begin{bmatrix} \Delta\mathbf{d}_\mathrm{r} \\ \Delta\mathbf{p} \end{bmatrix}^{i+1}_{j+1} = -\begin{bmatrix} \mathbf{V}^\mathrm{T}\mathbf{R}^\mathrm{S} \\ \mathbf{R}^{0\mathrm{D}} \end{bmatrix}^i_{j+1}, \tag{16}$$

by linearizing the spatially and time discrete form of the residual. Through the chain rule of differentiation, the left column entries of the block Jacobian matrix are right-multiplied by $\mathbf{V}$. The structural block $\mathbf{V}^\mathrm{T}\partial\mathbf{R}^\mathrm{S}/\partial\mathbf{d}\mathbf{V}$ is now of reduced dimension $\mathbb{R}^{q\times q}$. Note that the Windkessel Jacobian $\partial\mathbf{R}^{0\mathrm{D}}/\partial\mathbf{p}$ and Windkessel degrees of freedom $\mathbf{p}$ remain unchanged. For the off-diagonal coupling blocks $\partial\mathbf{R}^{0\mathrm{D}}/\partial\mathbf{d}$ and $\partial\mathbf{R}^\mathrm{S}/\partial\mathbf{p}$ only the structural dimension is right and left multiplied with the projection matrix $\mathbf{V}$ or its transpose, respectively. The original 11 and reduced 16 block-Jacobian are visualized in Figure 2.

The update step of a Newton iteration is carried out by extrapolating the reduced DOFs $\mathbf{d}_\mathrm{r}$ to the full order displacement vector $\mathbf{d}$ with the projection matrix $\mathbf{V}$ by

$$\begin{bmatrix} \mathbf{d} \\ \mathbf{p} \end{bmatrix}^{i+1}_{j+1} = \begin{bmatrix} \mathbf{d} \\ \mathbf{p} \end{bmatrix}^i_{j+1} + \begin{bmatrix} \mathbf{V}\Delta\mathbf{d}_\mathrm{r} \\ \Delta\mathbf{p} \end{bmatrix}^{i+1}_{j+1}. \tag{17}$$

Note that the full order structural residual $\mathbf{R}^\mathrm{S}$ and the full order block-Jacobian are always evaluated using the full order displacements $\mathbf{d}$. It is only after their full evaluation and assembly that their dimensions are reduced by projection. The convergence check is carried out with the reduced residual and reduced displacement increment

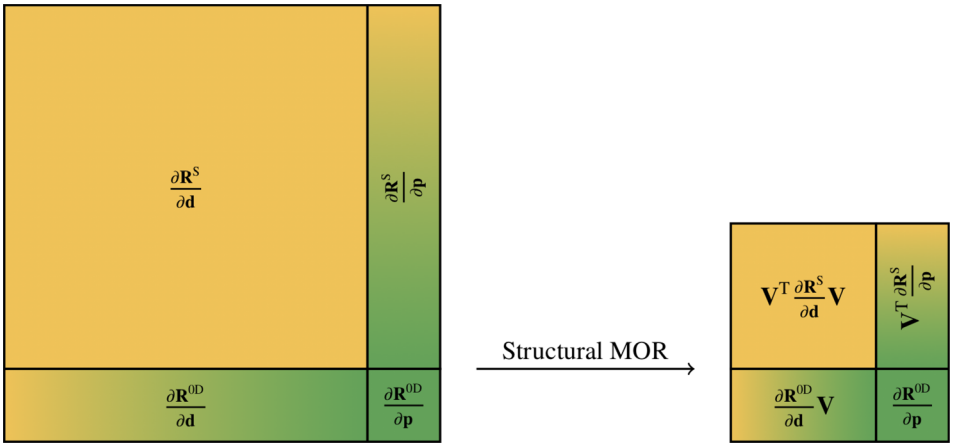

FIGURE 2 Visualization of the Jacobian during projection-based model order reduction of the structural dimension of the block matrix system in Equations 11 to 16. The diagonal structural and Windkessel blocks are colored yellow and green, respectively. Off-diagonal coupling blocks are shaded. Note that the dimension of the diagonal Windkessel block remains unchanged

$$\left\|\mathbf{V}^{\mathrm{T}}\mathbf{R}^{\mathrm{S}}\right\|_{\infty} < tol^{\mathrm{S}}_{\mathrm{res}}, \quad \left\|\Delta\mathbf{d}_{\mathrm{r}}\right\|_{\infty} < tol^{\mathrm{S}}_{\mathrm{inc}}. \tag{18}$$

The convergence criteria for the 0D Windkessel model remain unchanged. As the coupled full order model 11 in section 2.3, the coupled ROM in Equation 16 is valid for any full order structural and Windkessel residual $\mathbf{R}^{\mathrm{S}}$ and $\mathbf{R}^{\mathrm{0D}}$, respectively.

3.2 | Subspace computation via POD

In this work, we use the method of Proper Orthogonal Decomposition (POD) to compute the reduced basis $\mathbf{V}$ required for the projection-based reduction of our full problem. POD[46,47] is a straightforward and very well-known nonlinear model reduction approach, which relies on so-called *snapshots*, that is, discrete-time observations of the solution of our FOM for a fixed parameter set $\boldsymbol{\mu}$, to construct the basis $\mathbf{V}$. Given $n_{\mathrm{s}} \ll n$ snapshots $\mathbf{d}_i$ gained from a numerical simulation of the FOM sample point, we obtain the snapshot matrix

$$\mathbf{D} = [\mathbf{d}_1, ..., \mathbf{d}_{n_{\mathrm{s}}}] \in \mathbb{R}^{n \times n_{\mathrm{s}}}. \tag{19}$$

Typically, the number of snapshots n_s corresponds to the number of time steps in the FOM simulation. The goal of POD is to construct a basis for an optimal approximation of the *solution manifold* spanned by the snapshot matrix. In other words, the aim is to generate a basis that optimally approximates the information gathered in the snapshots. Therefore, we perform a singular value decomposition (SVD) of the snapshot matrix

$$\mathbf{D} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{T}^{\mathrm{T}}, \tag{20}$$

with the orthogonal matrices $\mathbf{U} \in \mathbb{R}^{n \times n}$ and $\mathbf{T} \in \mathbb{R}^{n_{\mathrm{s}} \times n_{\mathrm{s}}}$ containing the left and right singular vectors, respectively, stored column-wise. The diagonal matrix

$$\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, ..., \sigma_{n_{\mathrm{s}}}) \in \mathbb{R}^{n \times n_{\mathrm{s}}}, \quad \text{where } \sigma_1 \geq ... \geq \sigma_{n_{\mathrm{s}}} \geq 0, \tag{21}$$

features all n_{s} singular values σ_i sorted in descending order on its main diagonal. We now select the first q singular vectors $\mathbf{u}_i$ from the columns of the left singular matrix $\mathbf{U}$ corresponding to the q largest singular values σ_i in $\boldsymbol{\Sigma}$ to obtain the basis vectors

$$\mathbf{v}_i = \mathbf{u}_i, \quad \forall i \in \{1, ..., q\} \tag{22}$$

of the projection matrix $\mathbf{V}$. The singular values σ_i are frequently used to define the Relative Information Content (RIC),

$$RIC(q) = \frac{\sum_{i=1}^{q} \sigma_i^2}{\sum_{i=1}^{n_{\mathrm{s}}} \sigma_i^2} \in [0, 1]. \tag{23}$$

This measure allows to select an appropriate basis dimension q such that $RIC(q) \geq 1 - \varepsilon_{\mathrm{POD}}$ for a given small tolerance $\varepsilon_{\mathrm{POD}}$.[22] The approximation error made by selecting $q < n_{\mathrm{s}}$ basis vectors can be quantified by the sum of the squared truncated singular values

$$\mathrm{e}(q) = \sum_{i=q+1}^{n_{\mathrm{s}}} \sigma_i^2. \tag{24}$$

Note that this technique provides an optimal basis for the approximation of the snapshot matrix in a least-squares sense.[48-50] Thus, the efficiency of POD and the basis quality crucially depends on the selection of snapshots, which is

required to represent the model's dynamics behavior sufficiently. Further note that POD requires the expensive numerical simulation of the full forward model, in general for many parameter sets, to collect representative snapshots. Nevertheless, this data-driven approach is very well applicable for the reduction of any nonlinear system.

3.3 | Interpolation of subspaces

The cardiac model described in section 2 relies on many patient-specific parameters, describing, for example, constitutive behavior, hemodynamics, boundary conditions, or local fiber orientation. Consequently, a repeated model evaluation and simulation for many different values of the parameters is indispensable to personalize the model. The aim of parametric model order reduction (pMOR) is to find a reduced cardiovascular model that preserves the parameter-dependency, thus allowing a variation of any of the parameters directly in the reduced model without having to repeat the whole reduction process each time. The parametric reduced model can be then used, for example, for patient-specific parameter estimation or uncertainty quantification purposes. Note that a parametric solution of the ROM still follows the process as outlined in Algorithm 1, since pMOR subspace interpolation only influences the (offline) generation of the projection matrix.

To efficiently reduce the parametric cardiovascular model, we decompose the pMOR procedure into an offline and online stage. In the *offline phase*, the parametrized full order model with n_p parameters $\boldsymbol{\mu} = \left[\mu_1, ..., \mu_{n_p}\right]^{\mathrm{T}} \in \mathbb{R}^{n_p}$ is first simulated for several parameter sample points $\boldsymbol{\mu}_k$, $k = 1$, ..., K, and then corresponding local projection bases $\mathbf{V}(\boldsymbol{\mu}_k)$ are computed via POD from the data obtained. In the *online phase*, the projection matrix $\mathbf{V}(\boldsymbol{\mu}^*)$ for a new parameter value $\boldsymbol{\mu}^*$ is generated by interpolating between the precomputed subspaces. Note that the selection of suitable parameter samples is highly problem-specific, depending mainly on desired accuracy and the parameter set, and can be challenging especially for high dimensional parameter spaces. In Appendix A, we sample the (one-dimensional) parameter space uniformly. For a given interval, a uniform distribution maximizes information entropy. It is thus the least-biased distribution and should always be selected if there is no prior information about the system to be reduced. Further note that the inverse analysis method proposed in section 5 does not require a pre-sampling of the parameter space. It rather collects a minimal amount of samples automatically, scaling only linearly with the number of iterations of the optimization algorithm and not the dimension of the parameter space.

In this paper, different subspace interpolation techniques are examined, which will be explained in the following. To do so, we suppose that local basis matrices $\mathbf{V}_1 = \mathbf{V}(\boldsymbol{\mu}_1)$, ..., $\mathbf{V}_K = \mathbf{V}(\boldsymbol{\mu}_K) \in \mathbb{R}^{n \times q}$ spanning the subspaces $\mathcal{V}(\boldsymbol{\mu}_1), ..., \mathcal{V}(\boldsymbol{\mu}_K)$ have been computed in the offline phase from the snapshot matrices $\mathbf{D}(\boldsymbol{\mu}_1), ..., \mathbf{D}(\boldsymbol{\mu}_K) \in \mathbb{R}^{n \times n_s}$ at the sample points $\boldsymbol{\mu}_1$, ..., $\boldsymbol{\mu}_K$. Each basis matrix $\mathbf{V}(\boldsymbol{\mu}_k)$ is composed of the vectors $\{\mathbf{v}_i(\boldsymbol{\mu}_k)\}_{i=1}^{q}$. For the interpolation, appropriate weighting functions $w_k(\boldsymbol{\mu}^*)$ should be selected to compute the interpolated basis $\mathbf{V}(\boldsymbol{\mu}^*)$ in the online phase. Basically, any multivariate interpolation method could be used for this purpose: e.g. polynomial interpolation (Lagrange polynomials), piecewise polynomial interpolation (splines), radial basis functions (RBF), Kriging interpolation (Gaussian regression), inverse distance weighting (IDW) based on nearest-neighbor interpolation or even sparse grid interpolation.[25] For simplicity, in this paper we consider the special case of piecewise linear interpolation. We compare in this work four interpolation methods: weighted concatenation of bases (CoB), weighted concatenation of snapshots (CoS), adjusted direct basis interpolation, and basis interpolation on a Grassman manifold. These methods were chosen as they already have been used previously in similar interpolation settings; see Reference 26 (Grassmann manifold) and Reference 23 (CoS). The other methods, CoB and adjusted direct basis interpolation, can be seen as straight-forward adaptions thereof.

3.3.1 | Weighted concatenation of bases

A common and straightforward approach to obtain a global basis matrix $\mathbf{V}$ from the precomputed local bases $\mathbf{V}(\boldsymbol{\mu}_1)$, ..., $\mathbf{V}(\boldsymbol{\mu}_K)$ is given by the method "concatenation of bases." With this technique, the local bases are at first concatenated side-by-side, followed by a SVD of the resulting matrix to compute the global basis $\mathbf{V}$. This technique can be extended by introducing the weighting functions $w_k(\boldsymbol{\mu}^*)$ in the concatenation of bases, in order to compute a *parameter-dependent* interpolated basis $\mathbf{V}(\boldsymbol{\mu}^*)$ which takes the distance of the new query point $\boldsymbol{\mu}^*$ with respect to the sample points $\boldsymbol{\mu}_1$, ..., $\boldsymbol{\mu}_K$ into account. To this end, the matrices $\mathbf{V}(\boldsymbol{\mu}_1)$, ..., $\mathbf{V}(\boldsymbol{\mu}_K)$ are first weighted with weights $w_k(\boldsymbol{\mu}^*)$ and concatenated afterwards. Then, the SVD of the concatenated matrix $\tilde{\mathbf{V}}(\boldsymbol{\mu}^*)$,

$$\tilde{\mathbf{V}}(\boldsymbol{\mu}^*) = \left[w_1(\boldsymbol{\mu}^*)\mathbf{V}(\boldsymbol{\mu}_1), ..., w_K(\boldsymbol{\mu}^*)\mathbf{V}(\boldsymbol{\mu}_K)\right] = \tilde{\mathbf{U}}(\boldsymbol{\mu}^*)\tilde{\boldsymbol{\Sigma}}(\boldsymbol{\mu}^*)\tilde{\mathbf{T}}(\boldsymbol{\mu}^*)^{\mathrm{T}} \in \mathbb{R}^{n \times K \cdot q}, \tag{25}$$

is performed. The interpolated basis $\mathbf{V}(\boldsymbol{\mu}^*)$ is finally constructed by considering the first q left singular vectors $\{\tilde{\mathbf{u}}_i(\boldsymbol{\mu}^*)\}_{i=1}^{q}$ that best represent the weighted and concatenated matrix $\tilde{\mathbf{V}}(\boldsymbol{\mu}^*)$:

$$\mathbf{V}(\boldsymbol{\mu}^*) = \left[\tilde{\mathbf{u}}_1(\boldsymbol{\mu}^*), ..., \tilde{\mathbf{u}}_q(\boldsymbol{\mu}^*)\right] \in \mathbb{R}^{n \times q}. \tag{26}$$

Please note that the described weighting procedure is purely optional. The advantage of the weighted approach is that subspaces near the interpolation point $\boldsymbol{\mu}^*$ are favored and stronger considered than subspaces describing the dynamics for far-distant sample points. However, this extended technique requires more computational effort than the classical concatenation approach, since a SVD has to be performed for every new $\boldsymbol{\mu}^*$ to compute the parameter-dependent interpolated basis $\mathbf{V}(\boldsymbol{\mu}^*)$.

3.3.2 | Weighted concatenation of snapshots

The concatenation of bases approach explained in the previous section provides a basis $\mathbf{V}(\boldsymbol{\mu}^*)$ comprising the most important directions among the (weighted) basis vectors from all local bases. Note, however, that the bases $\mathbf{V}(\boldsymbol{\mu}_k) = \mathbf{U}(\boldsymbol{\mu}_k)(:, 1:q)$ for $k = 1, ..., K$ are calculated in our case by means of the SVD-based technique of POD and they, therefore, essentially approximate the snapshot matrices $\mathbf{D}(\boldsymbol{\mu}_k)$,

$$\mathbf{D}(\boldsymbol{\mu}_k) = \mathbf{U}(\boldsymbol{\mu}_k)\boldsymbol{\Sigma}(\boldsymbol{\mu}_k)\mathbf{T}(\boldsymbol{\mu}_k)^{\mathrm{T}}. \tag{27}$$

Since we are actually most interested in finding a basis that optimally approximates the system dynamics over a range of parameters, in our concrete POD-case it seems very reasonable to construct the interpolated basis $\mathbf{V}(\boldsymbol{\mu}^*)$ from a (weighted) concatenation of snapshots rather than from a (weighted) concatenation of bases. With the former technique, the matrix $\mathbf{V}(\boldsymbol{\mu}^*)$ is therefore constructed by considering the first q left singular vectors of the (weighted and) concatenated snapshot matrix $\tilde{\mathbf{D}}(\boldsymbol{\mu}^*)$:

$$\tilde{\mathbf{D}}(\boldsymbol{\mu}^*) = \left[w_1(\boldsymbol{\mu}^*)\mathbf{D}(\boldsymbol{\mu}_1), ..., w_K(\boldsymbol{\mu}^*)\mathbf{D}(\boldsymbol{\mu}_K)\right] = \tilde{\mathbf{U}}_{\tilde{\mathbf{D}}}(\boldsymbol{\mu}^*)\tilde{\boldsymbol{\Sigma}}_{\tilde{\mathbf{D}}}(\boldsymbol{\mu}^*)\tilde{\mathbf{T}}_{\tilde{\mathbf{D}}}(\boldsymbol{\mu}^*)^{\mathrm{T}} \in \mathbb{R}^{n \times K \cdot n_s}. \tag{28}$$

Remark: Connection between the concatenation methods

It can be shown that the just described (weighted) concatenation of snapshots approach corresponds to a *modified* (weighted) concatenation of bases, where each vector $\mathbf{v}_i(\boldsymbol{\mu}_k)$ is (further) weighted with the corresponding singular value $\sigma_i(\boldsymbol{\mu}_k)$ for all non-zero singular values. Thus, the first q left singular vectors of the—towards Equation 25—modified matrix $\tilde{\mathbf{V}}(\boldsymbol{\mu}^*)$

$$\tilde{\mathbf{V}}(\boldsymbol{\mu}^*) = \left[w_1(\boldsymbol{\mu}^*)\mathbf{V}(\boldsymbol{\mu}_1)\boldsymbol{\Sigma}_q(\boldsymbol{\mu}_1), ..., w_K(\boldsymbol{\mu}^*)\mathbf{V}(\boldsymbol{\mu}_K)\boldsymbol{\Sigma}_q(\boldsymbol{\mu}_K)\right] = \tilde{\mathbf{U}}(\boldsymbol{\mu}^*)\tilde{\boldsymbol{\Sigma}}(\boldsymbol{\mu}^*)\tilde{\mathbf{T}}(\boldsymbol{\mu}^*)^{\mathrm{T}}, \tag{29}$$

where $\boldsymbol{\Sigma}_q(\boldsymbol{\mu}_k) = \boldsymbol{\Sigma}(\boldsymbol{\mu}_k)(1:q, 1:q) \in \mathbb{R}^{q \times q}$, span the same interpolated subspace $\mathcal{V}(\boldsymbol{\mu}^*)$ than the q leading vectors in $\tilde{\mathbf{U}}_{\tilde{\mathbf{D}}}(\boldsymbol{\mu}^*)$.

3.3.3 | Adjusted direct basis interpolation

It is well-known that a straightforward interpolation of the basis vectors comprised in the local projection matrices $\mathbf{V}(\boldsymbol{\mu}_1)$, ..., $\mathbf{V}(\boldsymbol{\mu}_K)$ does generally not yield a basis. This is due to the fact that the basis vectors $\{\mathbf{v}_i(\boldsymbol{\mu}_k)\}_{i=1}^{q}$ for different sample points span diverse subspaces, thus having a distinct physical interpretation and possibly pointing even in opposite directions in space. Therefore, the basis vectors should be first arranged to point in similar directions, thus spanning

similar subspaces, before their entries are interpolated. This adjustment is performed using the Modal Assurance Criterion (MAC).[51,52]

$$\mathrm{MAC}(\mathbf{v}_i,\mathbf{v}_j)=\frac{\left|\mathbf{v}_i^{\mathrm{T}}\cdot\mathbf{v}_j\right|^2}{\|\mathbf{v}_i\|_2^2\cdot\|\mathbf{v}_j\|_2^2}\in[0,1], \tag{30}$$

which provides a measure for the similarity or linear dependence between the vectors $\mathbf{v}_i$ and $\mathbf{v}_j$. Using the symmetry of the MAC, Equation 30 needs to be evaluated $q(q+1)/2$ times. The maximum value of the MAC is 1, which corresponds to linear dependent vectors, whereas orthogonal vectors take the minimal value 0. Hence, the idea is to only interpolate vectors which are strongly correlated to each other and maximize the MAC. To do so, we first have to select a reference subspace with respect to which the adjustment of the bases should be performed. The reference subspace, spanned by the columns of $\mathbf{R}_\mathbf{V}\in\mathbb{R}^{n\times q}$, should ideally comprise the most important dynamics among all parameter sample points and be representative for all local bases. The simplest way to select $\mathbf{R}_\mathbf{V}$ is to take one particularly important local basis $\mathbf{R}_\mathbf{V}=\mathbf{V}_{k_0}$ with $k_0\in\{1,...,K\}$. Another possibility is to construct the reference subspace similarly as described in section 3.3.1, that is, using the (weighted) concatenation of bases approach, yielding $\mathbf{R}_\mathbf{V}=\tilde{\mathbf{U}}(:,1:q)$ or $\mathbf{R}_\mathbf{V}(\boldsymbol{\mu}^*)=\tilde{\mathbf{U}}(\boldsymbol{\mu}^*)(:,1:q)$. Once the reference subspace has been selected, the vectors $\mathbf{v}_{i^*(j,k)}(\boldsymbol{\mu}_k)$ that fulfill

$$i^*(j,k)=\mathrm{argmax}_i\mathrm{MAC}(\mathbf{v}_i(\boldsymbol{\mu}_k),\mathbf{R}_\mathbf{V}(:,j))\quad \text{for}\, j=1,...,q\ \text{ and }\ k=1,...,K \tag{31}$$

are taken to be interpolated. Furthermore, the orientation of the vectors $\mathbf{v}_{i^*(j,k)}(\boldsymbol{\mu}_k)$ and $\mathbf{R}_\mathbf{V}(:,j)$ is equalized by adapting the sign, in order to avoid that an interpolation between vectors pointing in (almost) opposite directions results in a mutual cancellation. Finally, the interpolation of the vectors is given by

$$\bar{\mathbf{v}}_j(\boldsymbol{\mu}^*)=\sum_{k=1}^{K}w_k(\boldsymbol{\mu}^*)\cdot\left[\pm\mathbf{v}_{i^*(j,k)}(\boldsymbol{\mu}_k)\right]\ \text{ with }\ \sum_{k=1}^{K}w_k(\boldsymbol{\mu}^*)=1. \tag{32}$$

The interpolation of orthonormal vectors does not necessarily yield a set of orthonormal vectors. Therefore, the interpolated vectors $\{\bar{\mathbf{v}}_j(\boldsymbol{\mu}^*)\}_{j=1}^q$ are subsequently orthonormalized by employing the SVD of $\bar{\mathbf{V}}(\boldsymbol{\mu}^*)$,

$$\bar{\mathbf{V}}(\boldsymbol{\mu}^*)=\left[\bar{\mathbf{v}}_1(\boldsymbol{\mu}^*),...,\bar{\mathbf{v}}_q(\boldsymbol{\mu}^*)\right]=\bar{\mathbf{U}}(\boldsymbol{\mu}^*)\bar{\boldsymbol{\Sigma}}(\boldsymbol{\mu}^*)\bar{\mathbf{T}}(\boldsymbol{\mu}^*)^{\mathrm{T}}, \tag{33}$$

and considering the first q left singular vectors $\{\bar{\mathbf{u}}_j(\boldsymbol{\mu}^*)\}_{j=1}^q$ for the interpolated basis $\mathbf{V}(\boldsymbol{\mu}^*)\in\mathbb{R}^{n\times q}$.

Special case of two precomputed bases and one parameter

In order to make the afore explained method more clear, we now briefly present the special case of two precomputed bases ($K=2$) and one single parameter ($n_p=1$). Let us assume that bases $\mathbf{V}(\mu_1)$ and $\mathbf{V}(\mu_2)$ have been computed at the parameter sample points μ_1 and μ_2, and that the new parameter value μ^* lies between these two samples. Suppose that we choose the reference basis, for example, as $\mathbf{R}_\mathbf{V}=\mathbf{V}(\mu_2)$. Then, the vectors $\mathbf{v}_{i^*(j)}(\mu_1)$ that fulfill

$$i^*(j)=\mathrm{argmax}_i\mathrm{MAC}\left(\mathbf{v}_i(\mu_1),\mathbf{v}_j(\mu_2)\right)\quad \text{for}\, j=1,...,q, \tag{34}$$

are selected to be combined with the vectors $\mathbf{v}_j(\mu_2)$. The interpolation reads

$$\bar{\mathbf{v}}_j(\mu^*)=w(\mu^*)\cdot\left[\pm\mathbf{v}_{i^*(j)}(\mu_1)\right]+\left(1-w(\mu^*)\right)\cdot\mathbf{v}_j(\mu_2), \tag{35}$$

with the weight

$$w(\mu^*)=\frac{\mu^*-\mu_2}{\mu_1-\mu_2}\in[0,1]\quad \text{for}\ \mu^*\in[\mu_1,\mu_2], \tag{36}$$

providing that a linear interpolation is employed.

3.3.4 | Basis interpolation on a Grassmannian manifold

As discussed before, a direct interpolation of the local bases is not meaningful, since they span different subspaces. In addition to the afore explained adjustment of the bases before interpolation, one may also interpolate the underlying subspaces on a tangent space of a manifold. The method proposed by Amsallem and Farhat[26] constructs a basis matrix $\mathbf{V}(\boldsymbol{\mu}^*)$ for a new parameter point $\boldsymbol{\mu}^*$ by interpolating the subspaces corresponding to the bases $\{\mathbf{V}(\boldsymbol{\mu}_k)\}_{k=1}^K$ on the tangent space to the Grassmannian manifold $\mathcal{G}_q(\mathbb{R}^n)$.

The first step of the approach consists in choosing a local basis matrix $\mathbf{V}_{k_0}$ for the reference point $\mathcal{V}_{k_0} \in \mathcal{G}_q(\mathbb{R}^n)$, at which the tangent space $\mathcal{T}_{\mathcal{V}_{k_0}}$ to the manifold $\mathcal{G}_q(\mathbb{R}^n)$ is constructed. Afterwards, all subspaces $\mathcal{V}(\boldsymbol{\mu}_k)$ spanned by the local bases $\mathbf{V}(\boldsymbol{\mu}_k)$ are mapped onto this tangent space by the so-called logarithmic mapping: $\mathrm{span}(\boldsymbol{\Gamma}_k) = \mathrm{Log}_{\mathcal{V}_{k_0}}(\mathcal{V}_k) \in \mathcal{T}_{\mathcal{V}_{k_0}}$. This is done basically by computing K thin SVDs

$$\left(\mathbf{I} - \mathbf{V}_{k_0}\mathbf{V}_{k_0}^{\mathrm{T}}\right)\mathbf{V}(\boldsymbol{\mu}_k)\left(\mathbf{V}_{k_0}^{\mathrm{T}}\mathbf{V}(\boldsymbol{\mu}_k)\right)^{-1} = \mathbf{U}(\boldsymbol{\mu}_k)\boldsymbol{\Sigma}(\boldsymbol{\mu}_k)\mathbf{T}(\boldsymbol{\mu}_k)^{\mathrm{T}} \quad \text{for } k = 1, \ldots, K, \tag{37}$$

and then calculating

$$\boldsymbol{\Gamma}(\boldsymbol{\mu}_k) = \mathbf{U}(\boldsymbol{\mu}_k)\arctan(\boldsymbol{\Sigma}(\boldsymbol{\mu}_k))\mathbf{T}(\boldsymbol{\mu}_k)^{\mathrm{T}}. \tag{38}$$

In order to compute the orthonormal basis $\mathbf{V}(\boldsymbol{\mu}^*)$ for a new parameter point $\boldsymbol{\mu}^*$, the matrices $\{\boldsymbol{\Gamma}(\boldsymbol{\mu}_k)\}_{k=1}^K$ are first interpolated using the weights $w_k(\boldsymbol{\mu}^*)$ to obtain

$$\boldsymbol{\Gamma}^* = \boldsymbol{\Gamma}(\boldsymbol{\mu}^*) = \sum_{k=1}^{K} w_k(\boldsymbol{\mu}^*)\boldsymbol{\Gamma}(\boldsymbol{\mu}_k). \tag{39}$$

The interpolated subspace $\mathrm{span}(\boldsymbol{\Gamma}^*) \in \mathcal{T}_{\mathcal{V}_{k_0}}$ is then mapped back to the original manifold $\mathcal{G}_q(\mathbb{R}^n)$ by the so-called exponential mapping: $\mathcal{V}(\boldsymbol{\mu}^*) = \mathrm{Exp}_{\mathcal{V}_{k_0}}(\mathrm{span}(\boldsymbol{\Gamma}^*)) \in \mathcal{G}_q(\mathbb{R}^n)$. The back-mapping step is numerically achieved by computing a thin SVD

$$\boldsymbol{\Gamma}(\boldsymbol{\mu}^*) = \mathbf{U}(\boldsymbol{\mu}^*)\boldsymbol{\Sigma}(\boldsymbol{\mu}^*)\mathbf{T}(\boldsymbol{\mu}^*)^{\mathrm{T}}, \tag{40}$$

followed by

$$\mathbf{V}(\boldsymbol{\mu}^*) = \mathbf{V}_{k_0}\mathbf{T}(\boldsymbol{\mu}^*)\cos(\boldsymbol{\Sigma}(\boldsymbol{\mu}^*)) + \mathbf{U}(\boldsymbol{\mu}^*)\sin(\boldsymbol{\Sigma}(\boldsymbol{\mu}^*)). \tag{41}$$

The special case of two precomputed bases ($K = 2$) and one single parameter ($n_p = 1$) is extensively described in Reference 26.

3.4 | Implementation details

The coupled FOM and ROM in Equations 11 and 16, respectively, are solved using our in-house parallel high-performance finite element software package BACI.[21] The code is implemented in C++ making use of the Trilinos library.[53] To solve the FOM's large linear system in Equation 11 we use a parallel iterative GMRES solver with 2×2 block SIMPLE-like preconditioning. For the ROM's small linear system in Equation 16 we use a serial direct solver. All preliminary calculations, that is, singular value decompositions and the interpolation of subspaces, are performed in MATLAB (Release 2017b, The MathWorks, Inc., Natick, MA).

4 | RESULTS MODEL ORDER REDUCTION

In this section we present results for the approximation of the FOM simulation with ROM simulations. We distinguish here between model order reduction and parametric model order reduction. For a fixed parameter set, using the contractility $\sigma = 280$ kPa, we explore the approximation qualities of POD. Following analysis of the heart's eigenmodes in section 4.1, we investigate the approximation quality of POD using a varying number of modes $q \in \{10, 50, 100, 200, 300, 400, 500\}$ in section 4.2. The model ROM10 was the model with the smallest mode number, where the cardiac simulation converged to a result in all time steps. The highest mode number for ROM500 was chosen, since there is a plateau around $q = 500$ in the decay of singular values in Figure 4a. We further demonstrate the computational speedup achieved by using POD in section 4.3 using again varying mode numbers. To explore accuracy in a single-parametric case, we analyze the approximation quality with respect to a changing contractility in Appendix A.

We employ a four chamber geometry obtained in vivo from a 33 year old healthy female volunteer. The imaging data was acquired at King's College London, UK using a Philips Achieva 1.5T magnetic resonance imaging (MRI) scanner. The geometry was meshed using Gmsh[54] with a resolution of 2 mm, yielding 282 288 nodes and 167 232 quadratic tetrahedral elements, totaling $n = 846\,864$ structural degrees of freedom (DOFs). As snapshots, we use all time steps of the simulation of a single heartbeat with $n_s = 874$. Additionally, we have four Windkessel DOFs. The cut geometry is displayed in Figure 3a. Note that evaluation of FOM and ROM with identical parameters would not yield any computational advantage in practice, as the solution of the FOM is already known and there is no need for a ROM solution. Nevertheless, approximation accuracy in the nonparametric case provides an upper bound for the parametric case. Furthermore, computational speedup is independent of the choice of parametric vs nonparametric MOR. The parameters enter the system solely via the projection matrix $\mathbf{V}$, which is generated offline. Online ROM evaluation is identical in both cases.

4.1 | POD-modes of the heart

To study the reducibility of our cardiac model, we analyze the decay of the singular values compared to the first one. This gives a measure of relative importance of the modes selected by POD. In Figure 4a we show the normalized singular value σ_i/σ_1 of mode i. For modes $i < 50$ there is a fast decay in relative importance, indicating good reducibility. There is a plateau for $250 < i < 700$, indicating that not much new information is gained by including those modes in the ROM.

The first modes of the heart are visualized in Figure 3, where the heart is cut in four-chamber-view. The simulation in reference configuration and at end-systole are shown in Figure 3a,b, respectively. Mode $i = 1$ in Figure 3c exhibits great similarity to the solution at end-systole and is characterized by a movement of the atrioventricular plane towards the apex with negligible change in outer shape of the heart. Mode $i = 2$ in Figure 3d consists of a more radial displacement of the outer walls of the ventricles and a pendulum motion of the intraventricular septum. Mode $i = 3$ in Figure 3e displays a rotating motion of the ventricles together with a large left-to-right movement of the intraventricular septum.

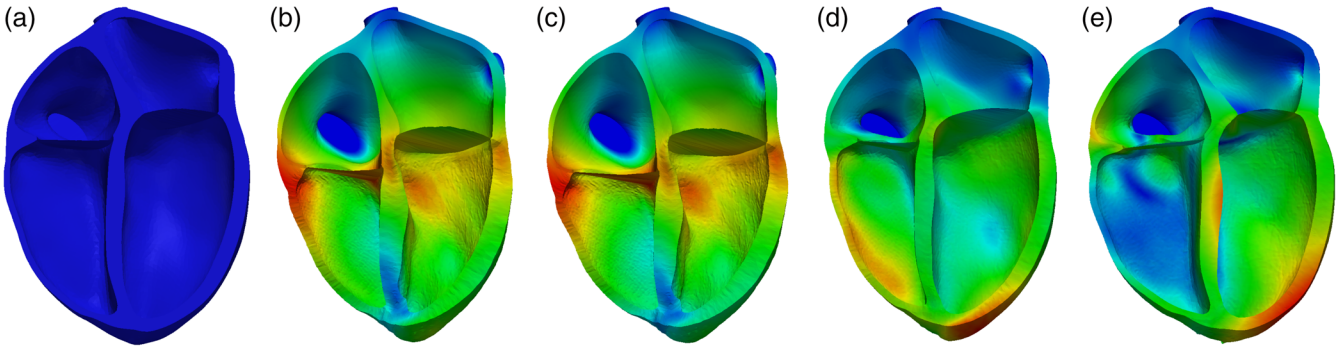

FIGURE 3 Visualized displacements in four-chamber-view. Displacements increase from blue to red regions. Reference configuration. (b) FOM at end-systole. (c) Mode $i = 1$. (d) Mode $i = 2$. (e) Mode $i = 3$

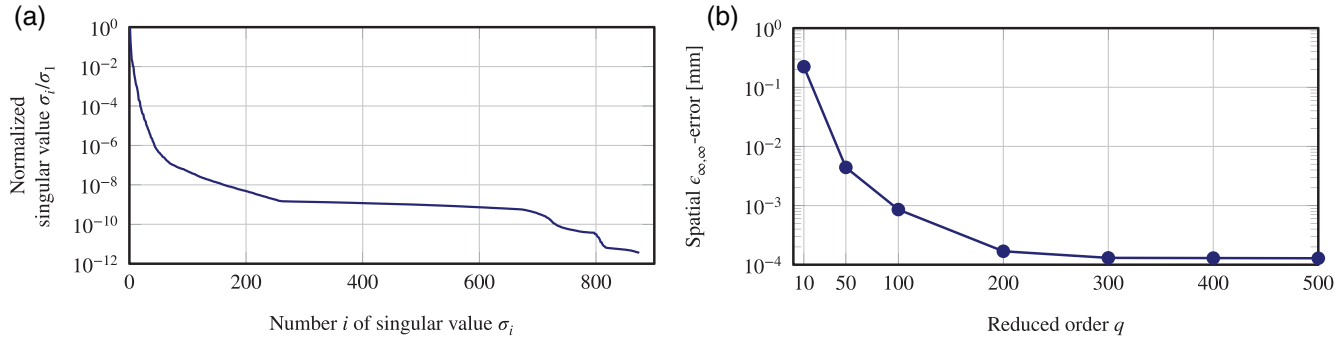

FIGURE 4 Accuracy of ROM. (a) Decay of normalized singular values. (b) Spatial $\varepsilon_{\infty,\infty}$-error compared to FOM depending on reduced order q

4.2 | Approximation quality

To quantify the overall approximation quality of ROM simulations of a full heartbeat, we calculate a spatial error compared to the FOM solution. We define here the spatial $\epsilon_{\infty,\,\infty}$ error

$$\epsilon_{\infty,\infty} = \max_{t_j}\left[\max_k \left\|\mathbf{d}^k_{\mathrm{ROM}}(t_j) - \mathbf{d}^k_{\mathrm{FOM}}(t_j)\right\|\right], \tag{42}$$

with $\mathbf{d}^k_{\mathrm{ROM}}(t_j)$ and $\mathbf{d}^k_{\mathrm{FOM}}(t_j)$ as nodal displacements at node k at time step t_j of ROM and FOM, respectively. The spatial $\epsilon_{\infty,\,\infty}$-error thus gives the highest displacement error at any node at any time step and is an upper bound for all spatial approximation errors. The $\epsilon_{\infty,\,\infty}$-error is shown in Figure 4b depending on the number of reduced modes q. It is clearly evident that the approximation error strongly decreases, when more modes are used for the approximation. Remarkably, even for the very low number of 10 modes we obtain a solution whose largest approximation error at any node at any time step is below 1 mm, which is the order of magnitude of our MRI resolution from which the geometry was obtained. Furthermore, using ROM simulations with a reduced order of $q > 300$ does not yield significant improvements in terms of accuracy. This is in agreement with the decay of the normalized singular values in Figure 4a, where modes $q > 300$ contain little more information than the preceding ones.

For many medical applications, it is not necessary to calculate an accurate spatial displacement field. There are rather a couple of scalar quantities which are used in clinical practice, for example, as a cardiac performance indicator or for the prediction of disease progression. Such a quantity is the ejection fraction

$$\mathrm{EF} = \frac{\max V - \min V}{\max V}, \tag{43}$$

which is calculated from left or right ventricular volume. To evaluate the approximation of the EF by a ROM simulation, we compare in Figure 5a the left ventricular (LV) volume curves of the FOM simulation with ROM simulations of various reduced orders q. It shows that minimum and maximum volume are approximated well and the time curves are almost indistinguishable. We further compare left ventricular pressure over time for all simulations in Figure 5b. Again, key features such as maximum pressure are approximated well. Minor oscillations occur for ROM10 and ROM50 after the closure of the mitral valve at $t \approx 0.2$. Furthermore, the closure of the aortic valve at $t \approx 0.5$ is delayed slightly for simulation ROM10.

4.3 | Speedup

For performance measurements, we ran all FOM and ROM simulations on a single node of our Linux cluster. One node features 64 GB of RAM and two Intel Xeon E5-2680 "Haswell" processors, each equipped with 12 cores operating at a frequency of 2.5 GHz. In Figure 6 we give a brief overview of the numerical performance of the FOM simulation. We show in Figure 6A the number of Newton iterations in each time step. The number of Newton iterations is between three and nine. It is elevated to five during ventricular systole and rises to nine at end-systole, where the aortic valve

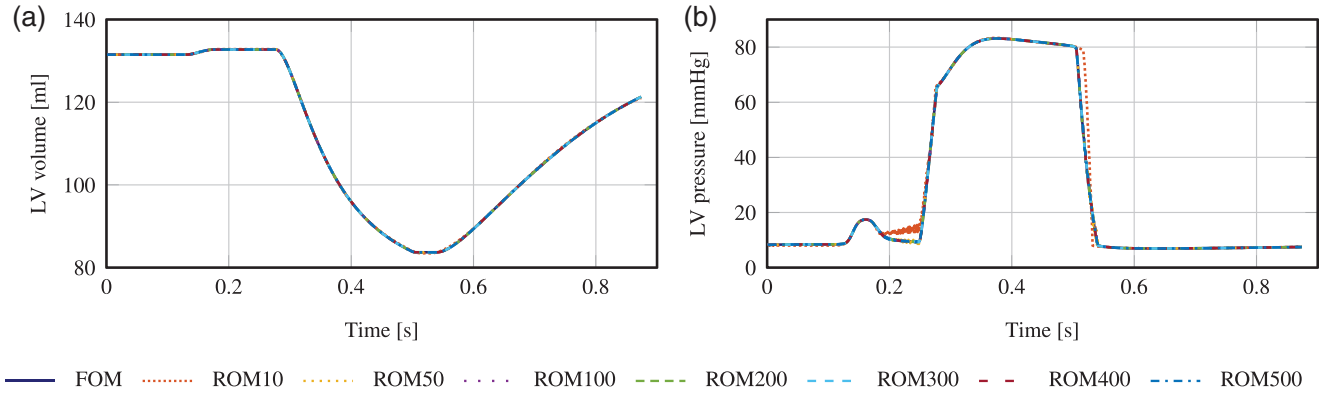

FIGURE 5 Scalar outputs of FOM and various ROMs over time. (a) Left ventricular volume. (b) Left ventricular pressure

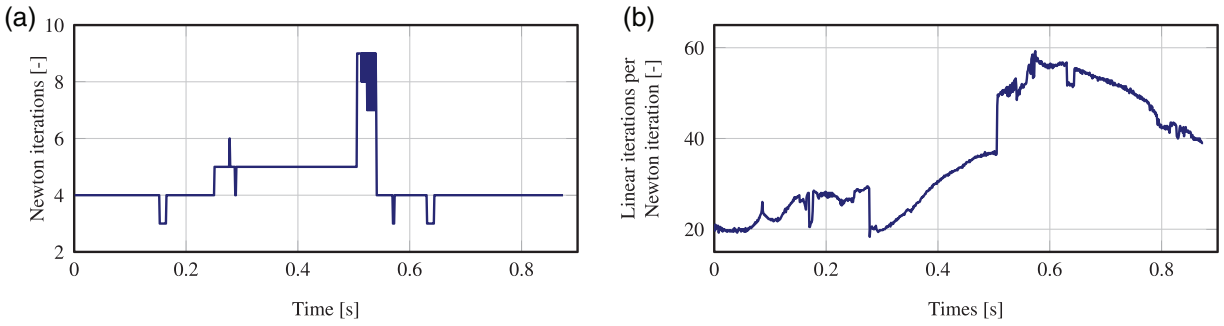

FIGURE 6 Solver performance of FOM in each time step. (a) Number of Newton iterations. (b) Number of linear solver iterations per Newton iteration

closes. The number of linear solver iterations of each Newton iteration at a given time step is shown in Figure 6B. The number of linear iterations is between 20 and 60 and shows similar trends as the number of Newton iterations. This performance is reasonable and assures a good basis to which ROM simulations can be compared to. In the following, we compare exclusively simulation time and exclude time for creating the projection matrix $\mathbf{V}$. It is calculated once in a preliminary step using the same hardware and requires only about 1 min. Due to the repeated evaluation of the MAC, the direct interpolation method is in this example about three times as expensive as the other interpolation methods, which perform equally.

The computation time of ROM simulations with various reduced orders q is compared to the total FOM simulation time in Figure 7. Firstly, we show in Figure 7a the speedup factor α of ROM over FOM simulations. The effect of POD is evident for ROM simulations between $q = 500$ and $q = 10$, where we achieve a speedup of $\alpha \approx 5$ and $\alpha \approx 13$ over the FOM simulation, respectively. Note that while we achieve high speedups we did not lower hardware demands. The RAM consumption has actually increased slightly for ROM simulations, since we need to additionally store the projection matrix $\mathbf{V}$. Hyper-reduction might be used in future studies to lower RAM consumption as here the residual and Jacobian are only assembled partially.

We distinguish in Figure 7b between three components of total computation time. Component "Linear system" includes the time required for the multiplications of the projection matrix $\mathbf{V}$ with the blocks of the Jacobian matrix in Equation 16 as well as the time to solve the reduced linear system. This component strongly depends on the reduced order q as it scales with the complexity of the matrix-matrix multiplications, which is the main time contributor. The solution time of the reduced linear system itself is negligible. Component "Evaluate elements" contains time spent during element evaluation to assemble the block Jacobian matrix and the right-hand side. As expected, this component is independent of q, since we still build the full system before projecting it to the q-dimensional subspace. Component "Other" sums up all other time spent during the simulation, for example, file input and output or general overhead, and is also independent of q.

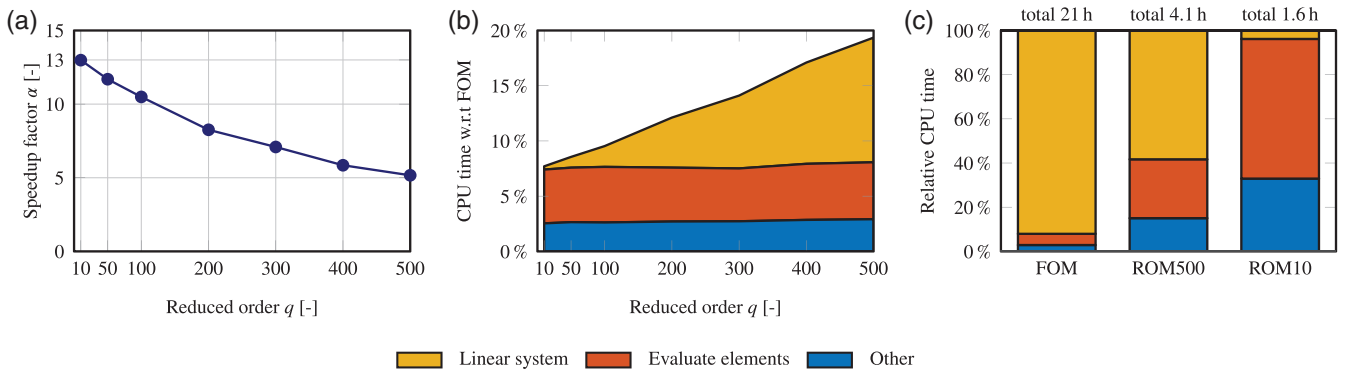

FIGURE 7 Simulation times of FOM and ROM. (a) Speedup factor. (b) Absolute components of computation time. (c) Relative components of computation time

In Figure 7c we show the relative distribution of simulation time for FOM, ROM500, and ROM10. For the 21 h spent during a FOM simulation, 92% of simulation time are spent solving the linear system. This large proportion shows the potential for savings using MOR with POD and is in part a consequence of the efficient parallelization of our finite element code.[21] Reducing and solving the linear system in ROM10 only makes up 4% of the simulation time. However, the new bottleneck is now the element evaluation at 63% of the simulation time. This, again, motivates the use of hyper-reduction methods for ROMs with very few degrees of freedom.

5 | PARAMETRIC MODEL ORDER REDUCTION FOR INVERSE ANALYSIS

Models of cardiac elasto-hemodynamics commonly depend on a large set of parameters which need to be calibrated to patient-specific measurements using inverse analysis. This procedure is however computationally very expensive, as it requires many evaluations of the model 11, which will be denoted forward model in the following. We propose in this section a novel approach for gradient-based inverse analysis utilizing the parametric reduced order model (pROM) developed in section 3.3. We replace the FOM forward evaluations typically required for the finite differences to obtain the gradient by pMOR forward evaluations (Equation 16). We illustrate this method using a simple Levenberg-Marquardt (LM) algorithm[55,56] in section 5.1. The method however is applicable to any gradient-based optimization. In section 5.2 we demonstrate its performance on a typical inverse analysis scenario using the cardiac forward problem introduced in section 2.

5.1 | Parameter identification based on reduced models

Given m normalized model outputs $\mathbf{f}(\boldsymbol{\mu}) \in \mathbb{R}^m$ of a FOM depending on n_p normalized parameters $\boldsymbol{\mu} \in \mathbb{R}^{n_p}$ and $m \geq n_p$ normalized measurements $\mathbf{y}$, we aim to minimize the squared sum S of residuals $\mathbf{r}$

$$\hat{\boldsymbol{\mu}} = \operatorname{argmin}_{\boldsymbol{\mu}} S(\boldsymbol{\mu}), \quad \text{with } S(\boldsymbol{\mu}) = \frac{1}{2}\|\mathbf{r}(\boldsymbol{\mu})\|_2^2, \quad \mathbf{r}(\boldsymbol{\mu}) = \mathbf{y}\text{-FOM}\,\mathbf{f}(\boldsymbol{\mu}), \quad \mathbf{J}(\boldsymbol{\mu}) = \frac{\partial \mathbf{r}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}, \tag{44}$$

to obtain the optimal set of parameters $\hat{\boldsymbol{\mu}}$. The Jacobian of the residual vector with respect to the parameter vector is $\mathbf{J} \in \mathbb{R}^{m \times n_p}$. At the optimum $\hat{\boldsymbol{\mu}}$, the gradient $\nabla S = \mathbf{J}^{\mathrm{T}}\mathbf{r} = \mathbf{0}$ vanishes and the Hessian $\nabla^2 S > \mathbf{0}$ is positive definite. Using the LM algorithm, we obtain the iterative procedure

$$\text{update} \quad \boldsymbol{\mu}^{i+1} = \boldsymbol{\mu}^{i} + \Delta\boldsymbol{\mu}^{i+1}, \tag{45}$$

$$\text{with} \quad \left[\mathbf{J}^{\mathrm{T}}\mathbf{J} + \lambda \operatorname{diag}\left(\mathbf{J}^{\mathrm{T}}\mathbf{J}\right)\right]^{i} \cdot \Delta\boldsymbol{\mu}^{i+1} = -\left[\mathbf{J}^{\mathrm{T}}\mathbf{r}\right]^{i}, \quad \lambda^{i} = \lambda^{i-1} \cdot \left\|\left[\mathbf{J}^{\mathrm{T}}\mathbf{r}\right]^{i}\right\|_2 / \left\|\left[\mathbf{J}^{\mathrm{T}}\mathbf{r}\right]^{i-1}\right\|_2, \tag{46}$$

$$\text{until} \quad \left\| \left[\mathbf{J}^{\mathrm{T}}\mathbf{r}\right]^{\mathrm{i}} \right\|_2 < \mathrm{tol}^{\boldsymbol{\mu}}_{\mathrm{grad}} \quad \text{and} \quad \left\| \Delta\boldsymbol{\mu}^{\mathrm{i}+1} \right\|_2 < \mathrm{tol}^{\boldsymbol{\mu}}_{\mathrm{inc}}, \tag{47}$$

at iteration $i+1$ with damping parameter λ. The LM algorithm approximates the Hessian as $\nabla^2 S \approx \mathbf{J}^{\mathrm{T}}\mathbf{J}$. The damping parameter λ should tend to zero as the parameter set $\boldsymbol{\mu}$ approaches the optimal solution $\hat{\boldsymbol{\mu}}$. For $\lambda \to \infty$ we approach the steepest descent method, for $\lambda = 0$ we obtain the Gauss-Newton method. In general, for a nonlinear model the analytical derivatives of the model evaluations $\mathbf{f}$ with respect to the parameters $\boldsymbol{\mu}$ required for the Jacobian matrix are not easily available. The n_p columns $\mathbf{J}^i_p$ of the Jacobian matrix $\mathbf{J}^i$ are thus typically approximated by finite differences

$$\mathbf{J}^i_p \approx \pm \frac{\mathrm{pROM}\,\mathbf{f}\left(\boldsymbol{\mu}^{\mathrm{i}}_{\Delta\mathrm{p}}\right) - \mathrm{pROM}\,\mathbf{f}\left(\boldsymbol{\mu}^{\mathrm{i}}\right)}{\epsilon_p}, \quad \text{with}\, \boldsymbol{\mu}^{\mathrm{i}}_{\Delta\mathrm{p}} = \boldsymbol{\mu}^{\mathrm{i}} \pm \epsilon_{\mathrm{p}}\mathbf{e}_{\mathrm{p}}, \quad \forall \mathrm{p} \in \left[1, \ldots, \mathrm{n_p}\right]. \tag{48}$$

The gradient evaluation vector $\boldsymbol{\mu}^i_{\Delta p}$ is built from the pth component ϵ_p of a finite distance vector $\boldsymbol{\epsilon} \in \mathbb{R}^p$ and the pth unit vector $\mathbf{e}_p \in \mathbb{R}^p$ in the direction of each parameter. The sign in Equation 48 is chosen for each parameter p so that the evaluation with parameter set $\boldsymbol{\mu}^i_{\Delta p}$ is within the range of all previously evaluated parameter sets.

Calculating the approximated Jacobian matrix requires $n_p + 1$ evaluations of our forward model which is computationally expensive in case of a large number of parameters n_p. The pROM introduced in section 3.3 is very accurate for parameters in the proximity of the sampled parameter sets, as was shown in section A1 for the cardiac contractility parameter σ. As demonstrated in section A2, using pMOR with two FOM sample points greatly improves the approximation of the tangent with respect to a changing contractility over using a single sample point. We therefore propose to use pROM evaluations of $\mathbf{f}$ in Equation 48 instead of FOM evaluations. The iterative procedure for the inverse analysis is sketched in Algorithm 2. Note that while using this approach, we still find a local minimum of the objective function S in Equation 44 with respect to the FOM.

Algorithm 2 *Inverse analysis with pMOR-gradient*

1. initialize $\boldsymbol{\mu}^0$, λ^0
2. $i = 0$
3. **while** convergence criterion from (47) not fulfilled **do**
4. evaluate FOM $\mathbf{f}(\boldsymbol{\mu}^i)$ and calculate residual $\mathbf{r}^i$ from (44)
5. store snapshots $\mathbf{D}(\boldsymbol{\mu}^i)$
6. **for** $p = 0, \ldots, n_p$ **do**
7. build reduced basis $\mathbf{V}\left(\boldsymbol{\mu}^i_{\Delta p}\right)$ from (49)
8. evaluate pROM $\mathbf{f}\left(\boldsymbol{\mu}^i_{\Delta p}\right)$
9. **end for**
10. calculate Jacobian $\mathbf{J}^i$ from (48)
11. update parameter vector $\boldsymbol{\mu}^{i+1}$ from (46)
12. $i \leftarrow i + 1$
13. **end while**
14. **return** $\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}^i$

Further note that the strategy introduced in Algorithm 2 requires in each iteration of the optimization that the FOM simulation is evaluated before the n_p gradients can be evaluated in parallel using the pROM simulations. Thus, considering a scenario of infinite available computing resources, this strategy would actually slightly increase computation time over the standard approach of using the FOM for all evaluations. Here, all $n_p + 1$ model evaluations can be run in parallel. However, considering the more likely scenario where computing resources are just sufficient to calculate one or few FOM simulations at a time, the strategy outline in Algorithm 2 leads to considerable time savings especially for a large number of parameters n_p.

Algorithm 2 can be combined with any subspace interpolation method in step 7. We use here the weighted concatenation of snapshots method (CoS) introduced in section 3.3.2, as it performed best in the experiments in Appendix A

and is easily applicable to high dimensions of n_p. For the weights of the snapshot matrix for gradient evaluation p, we use a simple inverse distance weighting between two evaluation points

$$\begin{aligned}\tilde{\mathbf{D}}\left(\boldsymbol{\mu}^i_{\Delta p}\right) &= \left[w_1\mathbf{D}\left(\boldsymbol{\mu}^i\right), w_2\mathbf{D}\left(\boldsymbol{\mu}^k\right)\right],\quad w_1 = \frac{1/d_1}{1/d_1+1/d_2},\quad w_2 = 1-w_1,\\ d_1 &= \left\|\boldsymbol{\mu}^i_{\Delta p}-\boldsymbol{\mu}^i\right\|_2,\quad d_2 = \left\|\boldsymbol{\mu}^i_{\Delta p}-\boldsymbol{\mu}^k\right\|_2,\quad k = \operatorname{argmin}_{j\in[0,\dots,i-1[}\left\|\boldsymbol{\mu}^i_{\Delta p}-\boldsymbol{\mu}^j\right\|_2,\end{aligned} \tag{49}$$

with normalized distances d_1, d_2 and weights w_1, w_2 for the current evaluation i and the next closest evaluation k, respectively. Since at the beginning of the iteration $\epsilon_p \ll |\mu^i_p - \mu^k_p|$, the weight w_1 of the current snapshot matrix $\mathbf{D}(\boldsymbol{\mu}^i)$ is always close to one, whereas the weight w_2 is close to zero. This can be interpreted that we "enrich" the snapshots of the current iteration with snapshots from a previous iteration to represent parametric dependence. As the optimization converges and the changes in parameters are close to the step size of the finite differences, the weights w_1 and w_2 equalize. For the first iteration of the optimization we rely on standard MOR evaluations using the constant projection matrix from the first FOM evaluation.

In the following, we give an equation for the speedup of gradient-based inverse analysis achieved by using pROM evaluations for the calculation of the Jacobian with respect to CPU time. Note that actual computation time depends on parallelization of model evaluations. We compare CPU time required to achieve convergence after n_i iterations for a model with gradients calculated from pROM and FOM forward model evaluations, denoted by superscript pROM and FOM, respectively. We do not include the time spent during subspace generation, as it is negligibly small compared to pROM and FOM evaluation time. The total CPU times T are

$$T^{\mathrm{FOM}} = n_i^{\mathrm{FOM}}\left(n_p+1\right)t^{\mathrm{FOM}}, \tag{50}$$

$$T^{\mathrm{pROM}} = n_i^{\mathrm{pROM}}\left[t^{\mathrm{FOM}} + \left(n_p+1\right)t^{\mathrm{pROM}}\right], \tag{51}$$

where t is the time required for a single forward evaluation. It can be observed from Equation 51 that the number of parameters only scales the pROM evaluation time but not the FOM evaluation time. Using the speedup α of a single pROM evaluation over a FOM evaluation, we obtain the total speedup β for the inverse problem with respect to CPU time as

$$\beta = \frac{T^{\mathrm{FOM}}}{T^{\mathrm{pROM}}} = \underbrace{\frac{n_i^{\mathrm{FOM}}}{n_i^{\mathrm{pROM}}}}_{\approx 1}\cdot\underbrace{\frac{1}{\frac{1}{\alpha}+\frac{1}{1+n_p}}}_{\to\alpha \text{ for } \mathrm{n_p}\to\infty},\quad \text{with } \alpha = \frac{\mathrm{t}^{\mathrm{FOM}}}{\mathrm{t}^{\mathrm{pROM}}}. \tag{52}$$

As will be shown in section 5.2, the first factor is close to one as the number of iterations is comparable for both approaches when using a reasonably large number of reduced modes q. The second factor approaches α in the case of many parameters. Note that in practice there is a trade-off between the two factors. Choosing a very low-dimensional reduced model with few degrees of freedom q results in a high single call speedup α but may increase the number of iterations for the inverse analysis, as the tangents are now approximated worse than with a higher q. Further note that, after their respective number of iterations, both approximations achieve the same convergence criterion, which is always evaluated using the FOM.

Other variants of Algorithm 2 are feasible, for example, replacing all FOM evaluations by pROM approximations as the inverse analysis algorithm converges closer to the optimum. Such algorithms however require more advanced strategies to switch between both model evaluations. The algorithm presented here demonstrates the most simple and straightforward approach of including a pROM within a finite difference gradient-based inverse analysis.

5.2 | Numerical results

We demonstrate the ability of the inverse analysis method proposed in section 5.1 to accurately and efficiently estimate parameters for a real-world cardiac identification problem. We consider the case of a cardiac simulation which we want to calibrate to a given volume curve, that is, measurements of left ventricular volume over time during one cardiac

cycle. We have no prior solutions of our FOM and thus need to build our projection matrices from scratch starting at the first iteration of Algorithm 2.

We choose the solution displayed in Figure 5a of a forward FOM simulation as our ground truth. As parameters we choose contractility σ from Equation 7 and myofiber activation rate α_{max}, myofiber deactivation rate α_{min}, onset of ventricular systole t_{sys}, and onset ventricular diastole t_{dias} from Equation 8. We thus estimate all parameters necessary to determine the shape of the input function of our model, that is, the active stress over time $\tau(t)$. The parameters σ, α_{max}, and α_{min} control cardiac output. However, due to their large variation they are commonly calibrated to a given patient.[31] These parameters are interconnected with the timing parameters t_{sys} and t_{dias}. The non-normalized parameters at the start of the inverse analysis and of the ground truth are listed in Table 1. We initialize the damping parameter $\lambda^0 = 0.1$. We further choose the number of reduced modes $q = 300$ as it offers a good trade-off between accuracy and speedup.

In Figure 8 we display the performance of the pROM inverse analysis using Algorithm 2 compared to the standard approach where the gradients are evaluated using the FOM only. Figure 8a shows the decay of the objective function S from Equation 44. We define a convergence criterion $S^i/S^0 < 10^{-5}$, which is achieved at $n_i^{\text{FOM}} = n_i^{\text{pROM}} = 7$. In Figure 8b we compare the development of the gradient of the objective function with respect to the parameters. As we consider a synthetic case in the absence of noise, both objective function and gradient should approach zero as $i \to \infty$.

	σ [kPa]	α_{max} $\left[\frac{1}{\text{s}}\right]$	α_{min} $\left[\frac{1}{\text{s}}\right]$	t_{sys} [s]	t_{dias} [s]
Initial	200	15	−15	0.35	0.60
Ground truth	280	10	−30	0.25	0.50

TABLE 1 Initial values and ground trouth of estimated parameters during inverse analysis ($n_p = 5$)

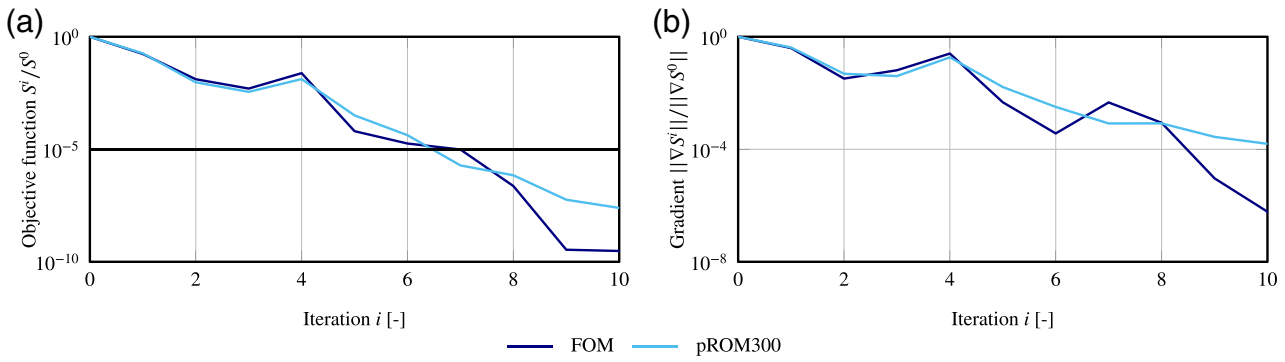

FIGURE 8 Convergence behavior during gradient-based inverse analysis with finite differences for gradient calculation. Shown are objective function and gradient for each iteration, comparing the use of FOM and pROM for gradient calculation. (a) Objective function. (b) Gradient

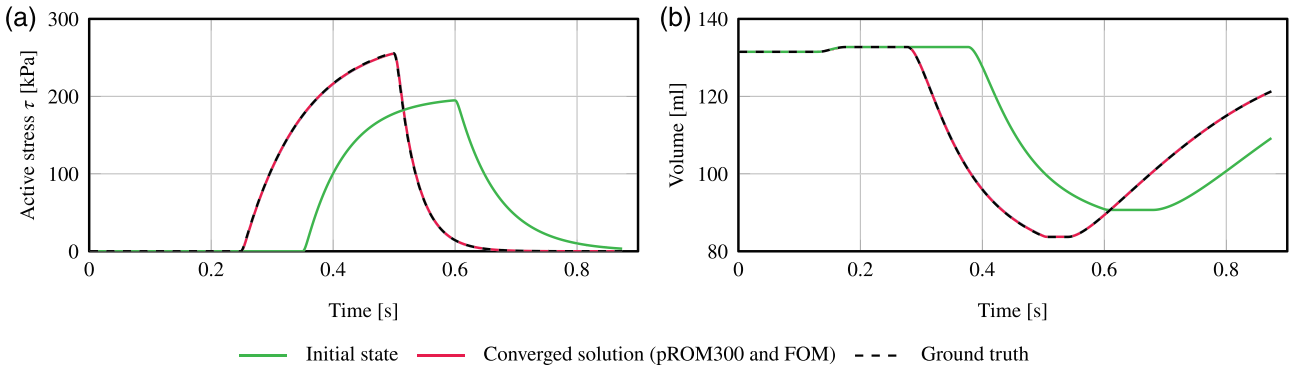

FIGURE 9 Initial state, converged solution, and ground truth of inverse analysis for model input (active stress) and model output (volume) over time. (a) Active stress (model input). (b) Volume (model output)

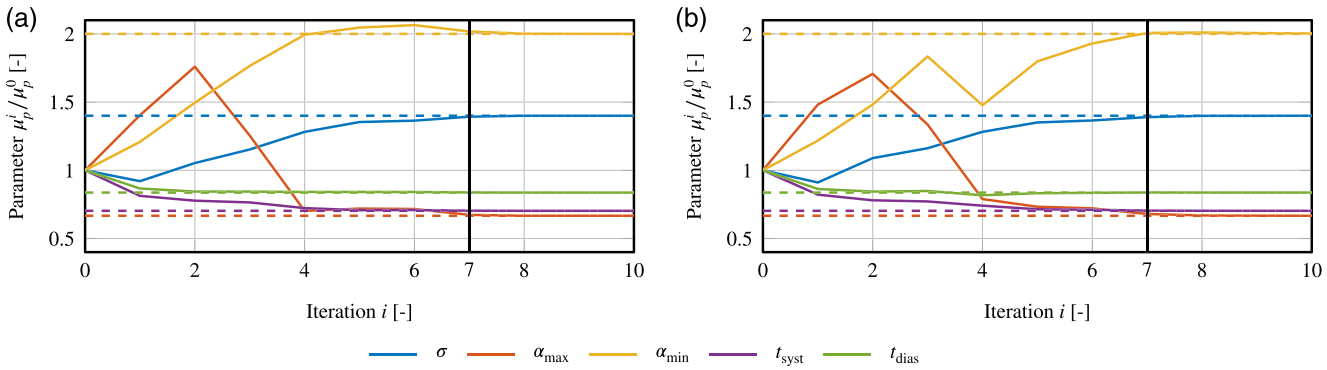

FIGURE 10 Convergence of parameters in inverse analysis. Dotted lines indicate ground truth of parameter. (a) FOM parameters. (b) pROM300 parameters

As both measures are non-monotonically decreasing, this indicates a non-convex optimization problem. However, the pROM300-gradient optimization is in excellent agreement with the FOM-gradient optimization.

The start, ground truth, and the converged solutions of both methods after seven iterations are shown in Figure 9. Here, the activation function, that is, the input of our model, and the volume, that is, the output of our model, are shown in Figure 9a,b, respectively. It can be observed that both optimization methods match well with ground truth data for the given convergence criterion. The convergence of the five parameters relative to their initial values is shown in Figure 10 for both methods. Additionally, the iteration where the convergence criterion is achieved is indicated. Both methods show a similar trend towards the optimal parameters. With a single evaluation speedup of $\alpha \approx 7.1$, we obtain an overall speedup in CPU time of the pROM300 method over the FOM method of $\beta \approx 3.3$, since $n_i^{\text{FOM}}/n_i^{\text{ROM}} = 1$ and $n_p = 5$ in our case.

6 | CONCLUSION AND OUTLOOK

6.1 | Model order reduction

In this work we proposed a new projection-based reduced order model for coupled structure-Windkessel cardiac models, where we solely reduced the large structural dimension. Specifically, we used a nonlinear large deformation cardiac finite element model with pericardial boundary conditions. For subspace generation, we employed proper orthogonal decomposition (POD) applied to displacement snapshots of the full order model (FOM). We demonstrated the accuracy and speedup of the reduced order model (ROM) for a range of reduced dimensions $q \in \{10, ..., 500\}$. In that range, the approximation error was found to be between $2 \cdot 10^{-1}$ mm and $1 \cdot 10^{-4}$ mm, which is well below the resolution of state of the art cardiac imaging employed in current clinical practice. For these simulations, we achieved speedups between 13 and 5 over our FOM. For highly reduced models, it was shown that the new bottleneck in simulation time is element evaluation. This motivates the inclusion of hyper-reduction methods, such as the discrete empirical interpolation method (DEIM)[57] or the energy conserving mesh sampling and weighting method (ECSW),[58] in future research. Due to the nonlinear structural finite element nature, we expect ECSW to perform better than DEIM for the cardiac problem. As the kinematics of a patient-specific heart can already be observed in motion MRI, it might be conceivable to incorporate this displacement information in the reduced space.

There exist many potential applications of MOR in cardiac many-query settings. One example is the task of obtaining a physiological periodic state, that is, where left and right ventricular output per cardiac cycle match. In these scenarios, a cardiac simulation with constant parameters is run for multiple cycles, until the change from one cycle to the next is below a given tolerance. In Reference 8 it was reported that in some cases more than ten cycles were necessary until converge to a periodic state was achieved. After simulating one FOM cycle and calculating the projection matrix, all preceding cycles could be run using a ROM since the shape of the cardiac contraction will be similar to the first cycle. In this use case, MOR can lead to drastic time savings, especially as the individual cardiac cycles cannot be run in parallel.

6.2 | Parametric model order reduction

We further compared four different methods of parametric model order reduction (pMOR) to allow ROM evaluations at parameter sets without prior FOM knowledge. The pMOR methods were evaluated by varying cardiac contractility, an important determinant of cardiac performance. The weighted concatenation of snapshots method was found to approximate the displacements of the FOM best for this example. Additionally, we showed that the clinically important scalar cardiac quantities ejection fraction, maximum left ventricular pressure, and left atrioventricular plane displacement are also well approximated using pMOR. Next to model calibration and design exploration, a possible application of cardiac pMOR could be multifidelity uncertainty quantification.[14]

6.3 | Inverse analysis

Finally, we introduced a novel method to include pMOR into a finite difference gradient-based inverse analysis. Using the Levenberg-Marquardt algorithm as an example, we proposed to use the FOM for all objective function evaluations and pMOR for all gradient evaluations, based on snapshots from the current and previous iterations. Using synthetic data in a real-world inverse analysis scenario, we demonstrated that pMOR-gradient-based optimization shows the same convergence properties as FOM-gradient-based optimization while achieving considerable CPU time savings. This method can be incorporated easily into existing optimization frameworks and could even be combined with commercial solvers for the structural problem. Using the inverse analysis approach proposed here has the advantage that we still calculate a full displacement field in each evaluation of the forward model. We can thus evaluate any spatial quantity, which is not possible when using 2D, 1D, or 0D surrogate models for the 3D structural model. In future research, this will allow us to compute a spatial approximation error with respect to cine or tagged MRI to estimate patient-specific parameters from clinical observations.

To the best of our knowledge, there is only one previous report about the application of parametric POD in the context of inverse problems in cardiac biomechanics. In Reference 23, the parameter space was sampled with four parametric FOM solutions. These snapshots were concatenated to form a global POD basis. These modes were used to build a ROM of a cardiac binventricular model, where two contractility values were estimated using a Kalman filter approach. As the authors state themselves, the main issue with such an approach is that the number of (offline) FOM evaluations grows exponentially with the number of parameters. Although these evaluations can be run in parallel, this still might cause problems in practice since in real world problems the FOMs are already very large, that is, one FOM already uses a large part of the available computational resources. Moreover, it is assumed that the parameter to be estimated is within the sampled parameter space.

We believe that our approach has the following advantages: (a) the number of FOM evaluations does not depend directly on the number of parameters but on the number of iterations in the optimization, which is usually independent on the number of parameters, and (b) there is no need to constrain the parameter space. It is noteworthy that at each optimization step higher order information of the least-squares functional (e.g. the Hessian) can be approximated using the POD modes at the same parameter, as it was done for the gradient. This should reduce the number of FOM evaluations to achieve a minimum. Furthermore, it is conceivable to replace some of the FOM evaluations of the objective function with pROM evaluations, depending on an error measure depicting the current quality of the pROM approximation.

ORCID

M. R. Pfaller https://orcid.org/0000-0001-5760-2617
C. Bertoglio https://orcid.org/0000-0001-5049-1707
W. A. Wall https://orcid.org/0000-0001-7419-3384

REFERENCES

1. Nordsletten D, Niederer S, Nash M, Hunter P, Smith N. Coupling multi-physics models to cardiac mechanics. *Prog Biophys Mol Biol*. 2011;104(1):77-88.
2. Laadhari A, Ruiz-Baier R, Quarteroni A. Fully Eulerian finite element approximation of a fluid-structure interaction problem in cardiac cells. *Int J Numer Methods Eng*. 2013;96(11):712-738.
3. Westerhof N, Lankhaar J-W, Westerhof BE. The arterial Windkessel. *Med Biol Eng Comput*. 2008;47(2):131-141.

4. Chabiniok R, Wang VY, Hadjicharalambous M, et al. Multiphysics and multiscale modelling, data–model fusion and integration of organ physiology in the clinic: ventricular cardiac mechanics. *Interface Focus*. 2016;6(2):20150083.
5. Kerckhoffs RC, Neal ML, Gu Q, Bassingthwaighte JB, Omens JH, McCulloch A. Coupling of a 3D finite element model of cardiac ventricular mechanics to lumped systems models of the systemic and pulmonic circulation. *Ann Biomed Eng*. 2007;35(1):1-18.
6. Sainte-Marie J, Chapelle D, Cimrman R, Sorine M. Modeling and estimation of the cardiac electromechanical activity. *Comput Struct*. 2006;84(28):1743-1759.
7. Fritz T, Wieners C, Seemann G, Steen H, Dössel O. Simulation of the contraction of the ventricles in a human heart model including atria and pericardium. *Biomech Model Mechanobiol*. 2013;13(3):1-15.
8. Hirschvogel M, Bassilious M, Jagschies L, Wildhirt S, Gee M. A monolithic 3D-0D coupled closed-loop model of the heart and the vascular system: experiment-based parameter estimation for patient-specific cardiac mechanics. *Int J Numer Methods Biomed Eng*. 2016;33(8):e2842.
9. Pfaller M. R, Hörmann J. M, Weigl M, Nagler A, Chabiniok R, Bertoglio C, Wall W. A. The importance of the pericardium for cardiac biomechanics: from physiology to computational modeling. *Biomech Model Mechanobiol*. 2019;18(2):503-529.
10. Arts T, Reneman RS, Veenstra PC. A model of the mechanics of the left ventricle. *Ann Biomed Eng*. 1979;7(3):299-318.
11. Caruel M, Chabiniok R, Moireau P, Lecarpentier Y, Chapelle D. Dimensional reductions of a cardiac model for effective validation and calibration. *Biomech Model Mechanobiol*. 2014;13(4):897-914.
12. Moulton MJ, Hong BD, Secomb TW. Simulation of left ventricular dynamics using a low-order mathematical model. *Cardiovasc Eng Technol*. 2017;8(4):480-494.
13. Lumens J, Delhaas T, Kirn B, Arts T. Three-Wall segment (TriSeg) model describing mechanics and hemodynamics of ventricular interaction. *Ann Biomed Eng*. 2009;37(11):2234-2255.
14. Biehler J, Gee MW, Wall WA. Towards efficient uncertainty quantification in complex and large-scale biomechanical problems based on a Bayesian multi-fidelity scheme. *Biomech Model Mechanobiol*. 2015;14(3):489-513.
15. Gerbeau J-F, Lombardi D. Approximated lax pairs for the reduced order integration of nonlinear evolution equations. *J Comput Phys*. 2014;265:246-269.
16. Gerbeau J-F, Lombardi D, Schenone E. Reduced order model in cardiac electrophysiology with approximated lax pairs. *Adv Comput Math*. 2015;41(5):1103-1130.
17. Pagani S, Manzoni A, Quarteroni A. Numerical approximation of parametrized problems in cardiac electrophysiology by a local reduced basis method. *Comput Methods Appl Mech Eng*. 2018;340:530-558.
18. Krysl P, Lall S, Marsden J. E. Dimensional model reduction in non-linear finite element dynamics of solids and structures. *Int J Numer Methods Eng*. 2001;51(4):479-504.
19. Milani R, Quarteroni A, Rozza G. Reduced basis method for linear elasticity problems with many parameters. *Comput Methods Appl Mech Eng*. 2008;197(51):4812-4829.
20. Schilders WHA, Lutowska A. *A Novel Approach to Model Order Reduction for Coupled Multiphysics Problems*. Cham: Springer; 2014:1-49.
21. Wall, WA, et al (2018) Baci: a parallel multiphysics simulation environment. Institute for Computational Mechanics, Technische Universität München; Technical report.
22. Bonomi D, Manzoni A, Quarteroni A. A matrix DEIM technique for model reduction of nonlinear parametrized problems in cardiac mechanics. *Comput Methods Appl Mech Eng*. 2017;324:300-326.
23. Chapelle D, Gariah A, Moireau P, Sainte-Marie J. A Galerkin strategy with proper orthogonal decomposition for parameter-dependent problems–analysis, assessments and applications to parameter estimation. *ESAIM: Math Modell Numer Anal*. 2013;47(6):1821-1843.
24. Ballarin F, Faggiano E, Manzoni A, et al. Numerical modeling of hemodynamics scenarios of patient-specific coronary artery bypass grafts. *Biomech Model Mechanobiol*. 2017;16(4):1373-1399.
25. Benner P, Gugercin S, Willcox K. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Rev*. 2015;57(4):483-531.
26. Amsallem D, Farhat C. Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA J*. 2008;46(7):1803-1813.
27. Griewank A. Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. *Optim Methods Software*. 1992;1(1):35-54.
28. Sermesant M, Moireau P, Camara O, et al. Cardiac function estimation from MRI using a heart model and data assimilation: advances and difficulties. *Med Image Anal*. 2006;10(4):642-656. (special issue).
29. Chávez CE, Zemzemi N, Coudière Y, Alonso-Atienza F, Álvarez D. Inverse problem of electrocardiography: estimating the location of cardiac ischemia in a 3D realistic geometry. In: Assen H, Bovendeerd P, Delhaas T, eds. *Functional Imaging and Modeling of the Heart*. Cham: Springer; 2015:393-401.
30. Finsberg H, Xi C, Tan JL, et al. Efficient estimation of personalized biventricular mechanical function employing gradient-based optimization. *Int J Numer Methods Biomed Eng*. 2018;34(7):e2982.
31. Chabiniok R, Moireau P, Lesault PF, Rahmouni A, Deux JF, Chapelle D. Estimation of tissue contractility from cardiac cine-MRI using a biomechanical heart model. *Biomech Model Mechanobiol*. 2012;11(5):609-630.
32. Bertoglio C, Moireau P, Gerbeau J.-F. Sequential parameter estimation for fluid-structure problems: application to hemodynamics. *Int J Numer Methods Biomed Eng*2012;28(4):434–455.

33. Moireau P, Bertoglio C, Xiao N, et al. Sequential identification of boundary support parameters in a fluid-structure vascular model using patient image data. *Biomech Model Mechanobiol*. 2013;12(3):475-496.
34. Asner L et al. Estimation of passive and active properties in the human heart using 3D tagged MRI. *Biomech Model Mechanobiol*. 2015;15(9):1121-1139.
35. Peherstorfer B, Willcox K, Gunzburger M. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Rev*. 2018;60(3):550-591.
36. Yue Y, Meerbergen K. Accelerating optimization of parametric linear systems by model order reduction. *SIAM J Optim*. 2013;23(2):1344-1370.
37. Lassila T, Manzoni A, Quarteroni A, Rozza G. A reduced computational and geometrical framework for inverse problems in hemodynamics. *Int J Numer Methods Biomed Eng*. 2013;29(7):741-776.
38. Molléro R, Pennec X, Delingette H, Garny A, Ayache N, Sermesant M. Multifidelity-CMA: a multifidelity approach for efficient personalisation of 3D cardiac electromechanical models. *Biomech Model Mechanobiol*. 2018;17(1):285-300.
39. Boulakia M, Schenone E, Gerbeau J-F. Reduced-order modeling for cardiac electrophysiology. Application to parameter identification. *Int J Numer Methods Biomed Eng*. 2012;28(6–7):727-744.
40. Corrado C, Gerbeau J-F, Moireau P. Identification of weakly coupled multiphysics problems. Application to the inverse problem of electrocardiography. *J Comput Phys*. 2015;283:271-298.
41. Holzapfel GA, Ogden RW. Constitutive modelling of passive myocardium: a structurally based framework for material characterization. *Philos Trans R Soc A Math Phys Eng Sci*. 2009;367(1902):3445-3475.
42. Gültekin O, Sommer G, Holzapfel GA. An orthotropic viscoelastic model for the passive myocardium: continuum basis and numerical treatment. *Comput Methods Biomech Biomed Eng*. 2016;19(15):1647-1664.
43. Newmark NM. A method of computation for structural dynamics. *J Eng Mech Div*. 1959;85(3):67-94.
44. Chung J, Hulbert G. A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-α method. *J Appl Mech*. 1993;60(2):371-375.
45. Crank J, Nicolson P. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Math Proc Cambridge Philos Soc*. 1947;43(1):50-67.
46. Kunisch K, Volkwein S. Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition. *J Optim Theory Appl*. 1999;102(2):345-371.
47. Chatterjee A. An introduction to the proper orthogonal decomposition. *Curr Sci*. 2000;78(7):808-817.
48. Kunisch K, Volkwein S. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J Numer Anal*. 2002;40(2):492-515.
49. Liang Y et al. Proper orthogonal decomposition and its applications—part I: theory. *J Sound Vib*. 2002;252(3):527-544.
50. Kerschen G, Golinval J-C, Vakakis AF, Bergman LA. The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview. *Nonlin Dyn*. 2005;41(1):147-169.
51. Allemang RJ, Brown DL. A correlation coefficient for modal vector analysis. *Proc Int Modal Anal Conf*. 1982;1:110-116.
52. Allemang RJ. The modal assurance criterion–twenty years of use and abuse. *Sound Vibr*. 2003;37(8):14-23.
53. Heroux MA, Phipps ET, Salinger AG, et al. An overview of the Trilinos project. *ACM Trans Math Software*. 2005;31(3):397-423.
54. Geuzaine C, Remacle J-F. Gmsh: a 3-D finite element mesh generator with built-in pre-and post-processing facilities. *Int J Numer Methods Eng*. 2009;79(11):1309-1331.
55. Levenberg K. A method for the solution of certain non-linear problems in least squares. *Quar Appl Math*. 1944;2(2):164-168.
56. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *J Soc Ind Appl Math*. 1963;11(2):431-441.
57. Chaturantabut S, Sorensen DC. Nonlinear model reduction via discrete empirical interpolation. *SIAM J Sci Comput*. 2010;32(5):2737-2764.
58. Farhat C, Chapman T, Avery P. Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models. *Int J Numer Methods Eng*. 2015;102(5):1077-1110.

How to cite this article: Pfaller MR, Cruz Varona M, Lang J, Bertoglio C, Wall WA. Using parametric model order reduction for inverse analysis of large nonlinear cardiac simulations. *Int J Numer Meth Biomed Engng*. 2020;e3320. https://doi.org/10.1002/cnm.3320

APPENDIX A.

Results single-parametric model order reduction

In this Appendix we provide a quantitative comparison of several subspace interpolation methods introduced in section 3.3 for parametric model order reduction (pMOR) to demonstrate the ability to evaluate the ROM simulations at parameter sets without prior FOM knowledge. We vary the contractility σ in Equation 7, controlling the upper limit

of the myocardium's active stress in fiber direction. It is a key parameter of cardiac contraction and has a major influence on elastodynamics as well as on several scalar cardiac measures. It is commonly calibrated to match the end-systolic volume of the left ventricle as measured in cine MRI. In this work, we vary the contractility $\sigma \in [280\ \text{kPa}, 430\ \text{kPa}]$, as this range produces FOM simulation results that are in agreement with cine MRI. We use $q = 300$ for all ROM simulations in this section, since it was shown in Figure 4B that no further improvements are made in approximation quality for $q > 300$. In section A1 we demonstrate the approximation quality with respect to the spatial displacement field. However, in many clinical applications a full solution of the displacements is not needed. We therefore show the approximation quality of pMOR with respect to scalar cardiac quantities of clinical significance in section A2.

Approximation of displacements

In Figure 1 we compare the spatial $\epsilon_{\infty,\,\infty}$-error for a varying contractility. We compare pMOR simulations using snapshots of FOM simulations from one, two, and four σ-sample points in Figure 11A-C, respectively. For comparison, we additionally compute a non-parametric ROM, denoted as direct ROM simulations in the following. Here, for every σ we evaluate the FOM to generate the corresponding projection matrix $\mathbf{V}(\sigma)$. This information is however not used in the

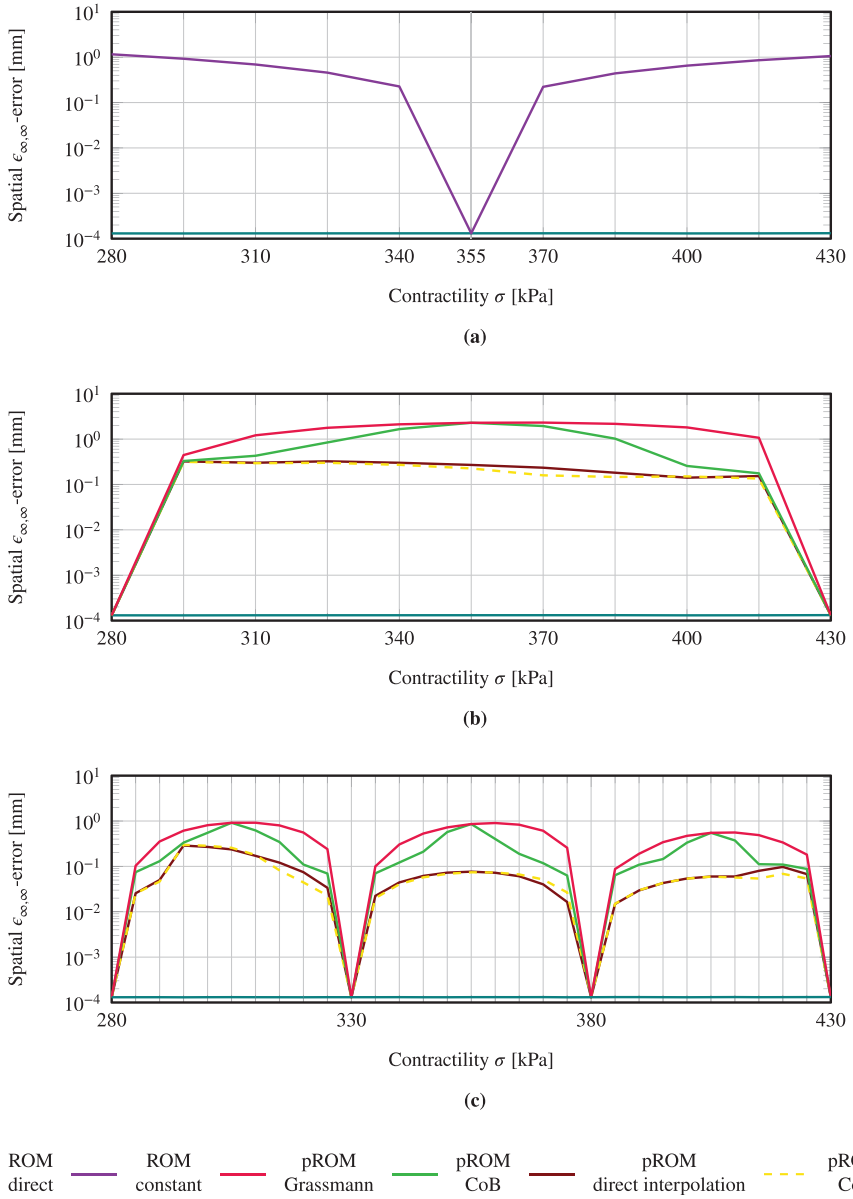

FIGURE 11 Spatial $\epsilon_{\infty,\,\infty}$-error for direct ROM300, constant ROM300, and pROM300 with different interpolation techniques for a varying number of sample points. (a) 669-EQN-335INEQ-END [kPa] increment $\Delta\sigma = 15$ kPa. (b) 673-EQN-337INEQ-END [kPa] increment $\Delta\sigma = 15$ kPa. (c) 677-EQN-339INEQ-END [kPa] increment $\Delta\sigma = 5$ kPa

pMOR solutions displayed here and would not be available in a typical MOR scenario, as it would render MOR useless. The direct ROM approximation error is mostly independent of the choice for parameter σ.

We show in Figure 1A the MOR approximation with varying σ and a constant projection matrix $\mathbf{V}(\sigma_1)$ which was obtained from a single FOM simulation with sample point $\sigma_1 = 355$ kPa. Technically, this would not be considered in pMOR, since the projection matrix is not adapted to the parameter set. It can be observed that MOR simulations with $\sigma \neq \sigma_1$ provide reasonable results with a spatial error below 1 mm using $\mathbf{V}(\sigma_1)$. However, with an increasing range of the parameter interval, the additional effort of subspace interpolation becomes advantageous. The approximation error of MOR simulations using two and four sample points are displayed in Figure 11B,C, respectively. Both studies show similar results. The error is highest between sample points and approaches the error of the direct ROM simulations close to the sample points. The pMOR approximation errors are reduced when using a finer resolution of sample points. The subspace interpolation with the largest spatial error is obtained by the Grassmannian manifold and concatenation of bases (CoB) methods, coinciding in the middle between two sample points. An error of one order of magnitude smaller is achieved by using the concatenation of snapshots (CoS) and the adjusted direct interpolation methods, staying well below a spatial $\epsilon_{\infty,\,\infty}$-error of 1 mm.

The adjusted direct interpolation method performs reasonably well in this example due to the POD of both snapshot matrices of the left and right sample point yielding modes that allow for a distinctive pairing according to the Modal Assurance Criterion. However, the same subspace can be spanned by two sets of orthogonal basis vectors which are not necessarily linearly dependent on each other. If the information of a mode related to the left sample point is scattered among various modes corresponding to the right sample point, the direct interpolation method will most probably yield considerably worse results.

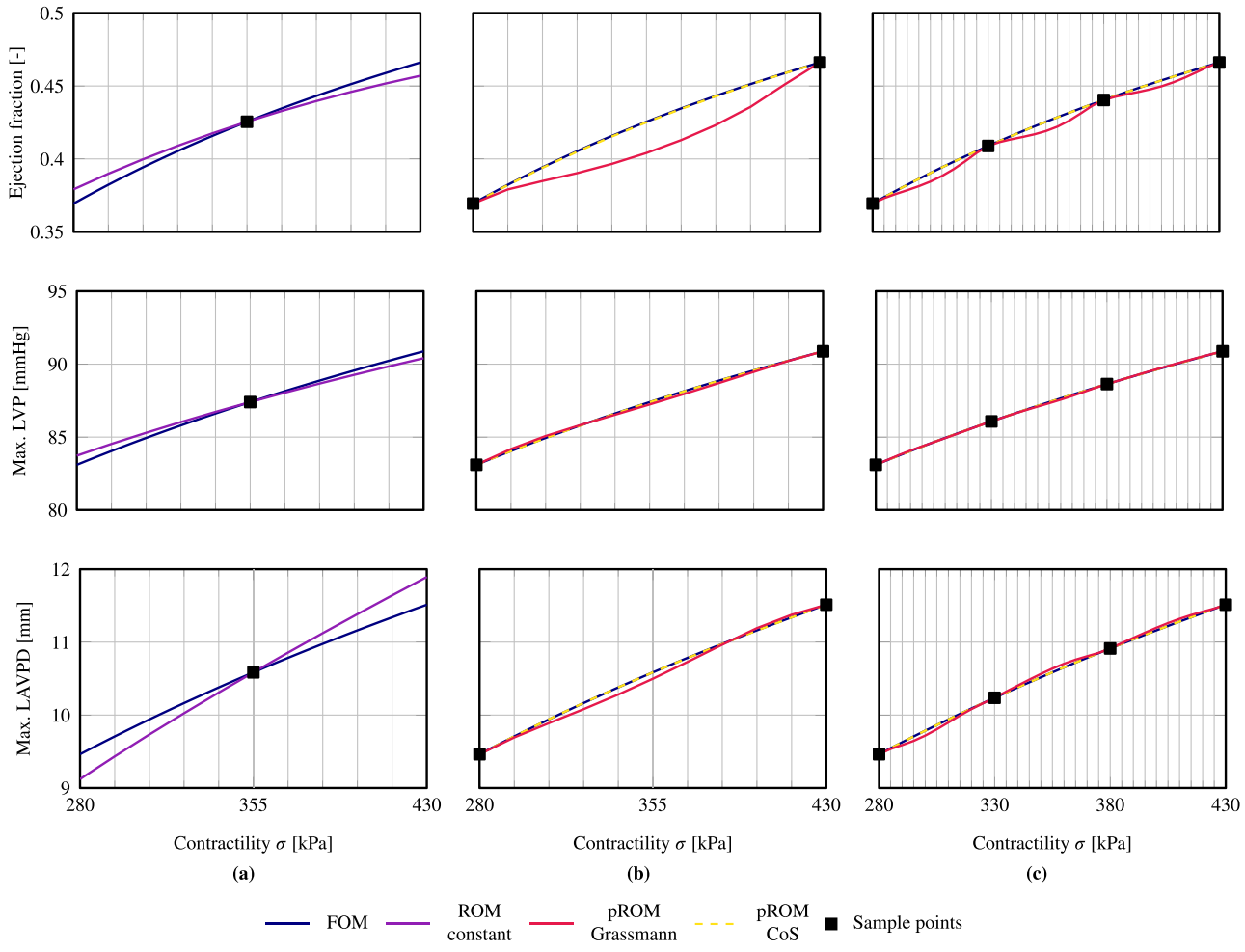

FIGURE 12 Scalar cardiac quantities ejection fraction (top), maximal left ventricular pressure (middle), and maximal left atrioventricular plane displacement (bottom) for varying contractility σ in FOM and ROM300. (a) One sample point. (b) Two sample points. (c) Four sample points

The CoS method is in our case far superior to the CoB method. As outlined in section 3.3, the CoS method allows for a direct usage of the snapshots at the sample points, while the CoB method uses the projection matrices computed at the sample points. The projection matrices, as compared to the snapshots, contain no information about the relative importance of the modes. The snapshots thus contain more information about the dynamics of the system, enabling the CoS method to select a more suitably interpolated subspace than the CoB method. Note, however, that the CoS method can only be applied when using it in combination with an observation-based reduction technique like POD.

Approximation of scalar cardiac quantities

In Figure 2 we show scalar output quantities of our cardiac model, evaluated for the same MOR simulations as in Figure 1 using one, two, and four sample points in Figure 12A-C, respectively. For each study we evaluate ejection fraction (EF), maximum left ventricular pressure (LVP), and maximum left atrioventricular plane displacement (LAVPD) and compare them to the FOM results. All three output quantities are important determinants of cardiac viability. They are also chosen because they allow us to study the approximation of different outputs of our coupled 3D-0D elasto-hemodynamical model with pMOR. EF, as defined in Equation (43), is an integral value of the spatial displacement field, LVP is an output of the 0D Windkessel model, and LAVPD is the average of a small subset of nodal directional displacements.

For one sample point we again compare the FOM solution in Figure 2A to the solution of a ROM using a constant projection matrix. The parameter dependence of the three cardiac quantities on σ is reproduced well by the constant ROM simulations. As expected, the deviations from the FOM solution are largest at evaluations furthest away from the sample point. However, the accuracy might still be sufficient for many applications. The cardiac quantities using two and four sample points are shown in Figure 2B,C, respectively. For clarity, we show here only the results of the subspace interpolation methods which performed best and worst in Figure 1, that is, the CoS and Grassmann method, respectively. The outputs oscillate visibly between sample points when using the Grassmann interpolation method, improving as the resolution of sample points is refined. As in Figure 1, the CoS method performs well, leading to a good approximation of the scalar cardiac quantities between sample points. Furthermore, comparing Figure 2B,C to Figure 2A, it is visible that using pROM with CoS interpolation yields not only a good approximation of the function *values* but also improves the *tangent* with respect to the contractility, compared to using constant ROM with a single sample point.