

## CARDIOVASCULAR MODELING WITH ADAPTED PARAMETRIC INFERENCE

DIDIER LUCOR & OLIVIER P. LE MAÎTRE<sup>1</sup>

**Abstract.** Computational modeling of the cardiovascular system, promoted by the advance of fluid-structure interaction numerical methods, has made great progress towards the development of patient-specific numerical aids to diagnosis, risk prediction, intervention and clinical treatment. Nevertheless, the reliability of these models is inevitably impacted by rough modeling assumptions. A strong integration of patient-specific data into numerical modeling is therefore needed in order to improve the accuracy of the predictions through the calibration of important physiological parameters. The Bayesian statistical framework to inverse problems is a powerful approach that relies on posterior sampling techniques, such as Markov chain Monte Carlo algorithms. The generation of samples requires many evaluations of the cardiovascular parameter-to-observable model. In practice, the use of a full cardiovascular numerical model is prohibitively expensive and a computational strategy based on approximations of the system response, or surrogate models, is needed to perform the data assimilation. As the support of the parameters distribution typically concentrates on a small fraction of the initial prior distribution, a worthy improvement consists in gradually adapting the surrogate model to minimize the approximation error for parameter values corresponding to high posterior density. We introduce a novel numerical pathway to construct a series of polynomial surrogate models, by regression, using samples drawn from a sequence of distributions likely to converge to the posterior distribution. The approach yields substantial gains in efficiency and accuracy over direct prior-based surrogate models, as demonstrated via application to pulse wave velocities identification in a human lower limb arterial network.

---

<sup>1</sup> LIMSI, CNRS, Université Paris-Saclay, Campus universitaire bât 508, Rue John von Neumann, F-91405 Orsay cedex, France

© EDP Sciences, SMAI 2018

**Résumé.** En s'appuyant notamment sur les progrès réalisés dans le domaine des méthodes numériques pour l'interaction fluide-structure, la modélisation du système cardiovasculaire progresse fortement. Elle est aujourd'hui en passe de proposer des outils numériques personnalisés pour l'aide au diagnostic, à la prédiction des risques, à l'intervention et au traitement clinique. Cependant la fiabilité des simulations est inévitablement impactée par les approximations liées à une modélisation encore trop grossière ou partielle. L'intégration directe de données spécifiques au patient est donc souhaitable à l'amélioration de la précision des prédictions par le biais de l'inférence des paramètres physiologiques importants. Le cadre statistique du formalisme bayésien se prête naturellement à la résolution des problèmes inverses. Il s'appuie sur des techniques d'échantillonnages *a posteriori*, telles que les méthodes de Monte Carlo par chaînes de Markov. La génération de la chaîne requiert de multiples évaluations du modèle cardiovasculaire reliant les paramètres aux l'observables. En pratique, le recours à un modèle cardiovasculaire complet reste trop coûteux. Dans ce cas, le choix d'une stratégie de calcul reposant sur des approximations par un modèle de substitution de la réponse du système rend l'assimilation des données plus abordable. De plus, comme le support de la distribution des paramètres *a posteriori* tend à se concentrer sur une petite portion de la distribution *a priori*, une amélioration possible consiste à graduellement adapter le modèle de substitution de manière à minimiser l'erreur d'approximation pour les valeurs des paramètres ayant les plus grandes densité *a posteriori*. Nous proposons une nouvelle approche numérique générant une suite d'approximations polynomiales du modèle, construite par régression à partir d'échantillons tirés aléatoirement selon une séquence de distributions convergeant vers la distribution *a posteriori*. Cette approche permet d'obtenir un gain numérique substantiel en termes d'efficacité et de précision par rapport à une approximation polynomiale directement basée sur un échantillonnage selon la distribution *a priori* des paramètres. La méthode est appliquée au cas de l'assimilation de données d'un modèle hémodynamique humain de la propagation d'ondes de pouls dans le réseau artériel du membre inférieur.

## INTRODUCTION

Inverse problems, encountered in the parametric identification, data assimilation or shape optimization are ubiquitous in life sciences and related interdisciplinary fields such as biomedical, biomechanics and bioengineering applications. Being generally ill-posed, inverse identification often entails very large computational efforts as it requires to iteratively solve the equations characterizing the biological system (i.e. the forward problem) numerous times; the idea behind the process being to adjust some parameters in order to lower the discrepancy between the numerical prediction and some indirect clinical observations of the system. The computational burden of repeatedly solving the forward problem in order to get some outputs of interest is further amplified because the model (e.g. patient-specific cardiovascular model) is often nonlinear and brings in unknown high-dimensional parameter spaces, and potentially large but noisy datasets. Model inversion in the presence of measurement errors must typically take advantage of some – type of regularization (e.g., Tikhonov regularization) in order to recover the existence and uniqueness of solutions or – robust optimization method [13]. A potentially more natural setting is the Bayesian statistics.

Bayesian statistics can be comprehended as a systematic use of probability to decision making in the face of uncertainty. It is a completely probabilistic approach to inference. Probability models are set up for the unknowns and the data. The inference is made from the conditional distribution of the unknowns given the data: the so-called posterior distribution. Specification of a full probability model involves the determination of the likelihood function as well as the specification of a prior distribution which expresses probabilistically what is known about the unknowns before observing the data. The Bayesian framework has been successful in numerous fields of life sciences such as biomedical applications, biostatistics, biomedical imaging, neurosciences/brain, gene expression and proteomics, clinical trials, epidemiology and biomechanics modeling. Nevertheless, very few developments associated with cardiovascular mathematics and modeling are reported in the literature.

The accuracy of patient-specific cardiovascular simulation tools together with medical images and segmentation algorithms, has significantly increased in recent years, allowing for tremendous progress in the understanding and treatment of various medical conditions. One example of such achievements is the simulation of the full-scale fluid-structure interaction between the blood flow and the arterial walls of the systemic circulation network, thanks to the coupling of Navier-Stokes equations with elastic parietal models [49]. Nevertheless, reliability and sensitivity of this type of numerical predictions with respect to the model bio-physiological and biomechanical parametric uncertainties and errors are seldom reported. There exists for instance numerous uncertainties related to geometry (e.g., arterial wall thickness, lumen diameter in the case of stenosis model, lesion length), material properties (e.g., permeability, elasticity, compliance or Young's modulus), blood rheology/viscosity, micro-vasculature peripheral resistance (e.g., the boundary conditions), *etc.*... associated with cardiovascular simulations. When available, patient-specific data are often scarce and indirect observations of the quantity of interest (the model parameters) and are altered by measurements noise and/or averaging acquisition procedures (e.g., clinical and medical imaging). Otherwise, population-averaged values are commonly used in place of patient-specific data. In addition, accurate full-scale direct numerical simulations remain generally out of reach and are not of effective use in support to the clinician's diagnostics and interventions. It is then a common usage to rely on reduced-order approximations (ROM), which are computationally more efficient but also more prone to model errors. After parametric calibration, the carefully crafted low-order model requires validation. In the case of vascular hemodynamics simulations, some exhaustive comparison between 1D and 3D hemodynamics formulations have for instance been published, e.g. [8, 30]. Several studies have shown the ability of the 1D formulation to capture the main features of pressure, flow and area waveforms in large human arteries using *in vivo* measurements [9, 31, 48], or *in vitro* experiments or 3D numerical data [8, 30]. Nevertheless, deterministic ROM constructed with optimized parameters value thanks to *in vivo* measurements or 3D numerical data often have limited value in predicting unobserved scenarios. Moreover, they do not provide an easy access to global parametric sensitivity analyses, that are essential to guide the modeling effort. Several works, e.g., [10, 12, 36–38], have demonstrated the interest of incorporating generic inter- and intra-patient variability in the form of uncertainty in the modeling of the cardiovascular system, since many aleatory and epistemic uncertainties remain due to its complexity, diversity and multiscale nature. These studies have shown how to propagate parametric uncertainties into the model in order to determine confidence intervals and statistics on the simulation predictions.

When the level of uncertainty is overwhelming and one holds clinical data/measurements for a given patient, the uncertainty measure of his physiological parameters may be narrowed. Provided a prior distribution of the parameters to be calibrated is available, the Bayesian inference (BI) is a rigorous approach to this estimation problem and delivers, after inference, a complete probabilistic characterization of the parameters. Within this approach, the parameters are considered as random quantities which can be sampled according to their posterior distribution using, in particular, a Markov-Chain Monte-Carlo (MCMC) sampler. This technique is straightforward to implement, but it becomes computationally intractable when the forward model (FM) is expensive to solve. A simple way of alleviating this problem is to substitute the FM with a cheaper surrogate model (SM), see e.g. [29]. One of the main limitations of the surrogate acceleration for BI is that the construction of the surrogate usually aims at minimizing the approximation error for the whole prior distribution, and as such is blind to the data. In cardiovascular simulations, the nonlinearity of the FM and the dimensionality and range of the parameters make very costly the construction of an accurate SM over the whole prior distribution. When the observations are informative enough, the posterior distribution concentrates on a small fraction of the prior support and the SM needs to be accurate essentially only over the support of the posterior, as low posterior areas are not interesting. Local adaptive surrogates [27] are good candidates for the purpose of accelerating MCMC samplers, but these approximations are “memoryless” and the resulting sampling strategy may lack performance. The present work addresses some of these issues and proposes a novel formulation designed to accurately approximate the FM over the whole support of the posterior distribution of the parameters. A sequence of global, data-driven, polynomial SM is derived and used for the MCMC sampling of the posterior. Once the SM converges, in the sense of the measure of the (approximated) posterior, no subsequent call to the

full FM is required.

The proposed inverse modeling approach focuses on the efficient characterization of sensitive reduced-order model parameters based on clinical data. In addition to the obvious asset of uncertainty reduction, thanks to the calibration, the stochastic modeling provides patient-specific quantitative information about the confidence in simulation predictions of unobserved cardiovascular states and points to the most influential parameters, thanks to calibrated global sensitivity analysis.

## 1. ADAPTIVE CARDIOVASCULAR SURROGATE

### 1.1. Parametric Bayesian inference based on surrogate reduced-order models

Here, we introduce our notations and recall the basics of Bayesian inference for parameter identification. Then we show how the formulation may be loosened to speed up the inference. We consider a forward model that maps some unknown parameters  $\boldsymbol{\theta}$  consider random to some observations, or the quantities of interest,  $\mathbf{z}$  derived from the forward model solution  $\mathbf{y}$ . The vector of random parameters to be inferred is written as  $\boldsymbol{\theta} \equiv \boldsymbol{\theta}(\omega) = (\theta^{(1)}, \dots, \theta^{(n_\theta)}) \in \mathcal{I}_\theta \subseteq \mathbb{R}^{n_\theta}$ , The FM solution (considered discrete) is a  $n_\theta$ -variate functional the parameters,  $\mathbf{y} : \mathcal{I}_\theta \rightarrow \mathcal{I}_\mathbf{y} \subseteq \mathbb{R}^{n_\mathbf{y}}$ , so that the predicted quantities of interest are also random functionals:  $\mathbf{z} : \mathcal{I}_\mathbf{y} \subseteq \mathbb{R}^{n_\mathbf{y}} \rightarrow \mathcal{I}_\mathbf{z} \subseteq \mathbb{R}^{n_d}$ . Denoting  $\mathbf{d} \in \mathbb{R}^{n_d}$  the dataset of observations (or measurements), the Bayes' formula writes:

$$\pi_{\text{post}}(\boldsymbol{\theta}|\mathbf{d}) \propto \pi_\ell(\mathbf{d}|\boldsymbol{\theta})\pi_{\text{prior}}(\boldsymbol{\theta}), \quad (1)$$

where prior information about  $\boldsymbol{\theta}$  is encoded in the prior density  $\pi_{\text{prior}}(\boldsymbol{\theta})$ ,  $\pi_\ell(\mathbf{d}|\boldsymbol{\theta})$  is the likelihood function that incorporates both the data and the forward model and  $\pi_{\text{post}}(\boldsymbol{\theta}|\mathbf{d})$  the sought parameters posterior density. The likelihood results from the combination of the costly deterministic forward model  $\mathcal{M} : \mathbb{R}^{n_d} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_\mathbf{y}}$ , an observation operator:  $\mathcal{G} : \mathbb{R}^{n_\mathbf{y}} \rightarrow \mathbb{R}^{n_d}$  and statistical models for measurement noise and model error. Assuming additive measurement noise, mutually independent from the parameters, we have:

$$\mathbf{d} = \mathbf{z} + \boldsymbol{\varepsilon} = \mathcal{G}(\mathbf{y}) + \boldsymbol{\varepsilon} = \mathcal{G}(\mathcal{M}(\boldsymbol{\theta}, \mathbf{x})) + \boldsymbol{\varepsilon}, \quad (2)$$

where the distribution  $\pi_\varepsilon$  of  $\boldsymbol{\varepsilon}$  is prescribed. In this case, we will assume that the noise model involves no hyperparameters, so the likelihood function becomes:

$$\pi_\ell(\mathbf{d}|\boldsymbol{\theta}) = \pi_\varepsilon(\mathbf{d} - \mathcal{G}(\mathcal{M}(\boldsymbol{\theta}, \mathbf{x}))). \quad (3)$$

The likelihood function therefore contains a stochastic source term whose representation must encompass the response of the deterministic forward model over the support of  $\pi_{\text{prior}}(\boldsymbol{\theta})$ . In practice, no closed form analytical expression exists. Posterior moments, expectations or maximum *a posteriori* values must be estimated via sampling methods such as MCMC, which require many evaluations of  $\mathcal{M}(\boldsymbol{\theta}, \mathbf{x})$  and can range from a few thousands to a few millions depending on the problem. One obvious way to accelerate this computation is to substitute a faster and cheaper reduced-order model (ROM)  $\mathcal{M}_{\text{rom}}$  to the full deterministic forward model. Here we refer to numerical reductions with respect to the  $\mathbf{x}$  quantity, which denotes the variables or parameters of the system whom quantified representation is known (e.g. spatial coordinates).

Previous works have relied on ROMs to solve more efficiently computational inverse problems, exploiting either the deterministic, the frequentist [20], or the Bayesian approach [16]. For instance, deterministic reductions can be introduced using a Proper Orthogonal Decomposition [41] or a Reduced Basis method [32], dimension reduction, simpler geometry, etc...). A complementary approach consists in introducing a stochastic approximation that will alleviate the systematic sampling of the (costly) forward model, e.g. [29]. In our case, we build a stochastic approximation of the ROM, i.e. a surrogate model noted  $\hat{\mathcal{M}}_{\text{rom}}$ , this time approximated in terms of its dependence to the uncertain parameters  $\boldsymbol{\theta}$ . In this setup, the surrogate-based posterior distribution now

takes the following form:

$$\hat{\pi}_{\text{post}}(\boldsymbol{\theta}|\mathbf{d}) \propto \pi_{\boldsymbol{\varepsilon}}\left(\mathbf{d} - \mathcal{G}\left(\hat{\mathcal{M}}_{\text{rom}}(\boldsymbol{\theta}, \mathbf{x})\right)\right) \pi_{\text{prior}}(\boldsymbol{\theta}). \quad (4)$$

The surrogate model should approximate the quantity of interest as accurately as possible. Classically, it is constructed from a finite set of model predictions for selected parameter values in a “off-line” stage, and subsequently used “in-line” during the sampling stage. The construction and querying of the surrogate model should be computationally efficient with sufficient predictive capabilities. Different methods are available to construct such surrogate models; examples are Gaussian processes [33], support vector machines [43], stochastic interpolation [44], or stochastic spectral methods such as polynomial-based representations. The choice of the surrogate model should be made to guarantee the accuracy of the approximate posterior distribution incurring from the substitution of the model prediction in Equation (4). Previous works have shown that the approximate posterior distribution error (in Kullback-Liebler divergence norm) is bounded by the surrogate model error measured in the mean square sense induced by the prior measure [7, 29]. Therefore, polynomial chaos surrogate models [17, 25, 50] are well suited as they indeed minimize the error norm  $\|\mathbf{z} - \hat{\mathbf{z}}\|_{\ell_2(\pi_{\text{prior}})} = (\mathbb{E}_{\pi_{\text{prior}}}\{(\mathbf{z} - \hat{\mathbf{z}})^2\})^{\frac{1}{2}}$ .

In this paper we will focus also on the use of polynomial expansions, that are well-suited for smooth dependences, constructed by discrete least-squares type approach. The surrogate model then consists in a linear combination of prescribed multivariate polynomials in  $\boldsymbol{\theta}$ . We shall assume that  $\mathbf{z}(\boldsymbol{\theta})$  is a second order random vector for  $\boldsymbol{\theta} \sim \pi_{\text{prior}}(\boldsymbol{\theta})$ ; its polynomial surrogate will write:

$$\mathcal{G}(\mathbf{y}(\boldsymbol{\theta})) = \mathbf{z}(\boldsymbol{\theta}) \approx \hat{\mathbf{z}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{\gamma} \in \Lambda_p} a_{\boldsymbol{\gamma}} \psi_{\boldsymbol{\gamma}}(\boldsymbol{\theta}), \quad (5)$$

where  $\mathbf{a}_{\boldsymbol{\gamma}}$  are the unknown expansion coefficients,  $\Lambda_p$  is an multi-index set (to be defined) for multi-index  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_{n_{\boldsymbol{\theta}}}) \in \mathbb{N}^{n_{\boldsymbol{\theta}}}$ . The polynomial approximation space is  $\mathbb{P}_{\Lambda_p} \equiv \text{span}\{\psi_{\boldsymbol{\gamma}} \mid \boldsymbol{\gamma} \in \Lambda_p\}$ , where the multivariate polynomials are often constructed as product of one uni-variate polynomials,  $\psi_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) = \prod_{i=1}^{n_{\boldsymbol{\theta}}} \psi_{\gamma_i}^{(i)}(\theta^{(i)})$ , with  $\psi_{\gamma_i}^{(i)}$  having degree  $\gamma_i$ . We restrict ourselves to isotropic tensor-product spaces total degree (TD)  $p$  leading to  $\Lambda_p^{\text{TD}} = \{\boldsymbol{\gamma} \in \mathbb{N}^{n_{\boldsymbol{\theta}}} : \|\boldsymbol{\gamma}\|_1 \leq p\}$ . We shall assume that these polynomials set forms a basis for any subsequent measures associated to  $\boldsymbol{\theta}$ , in particular the successive approximations of the posterior density. For simplicity in the construction of the original polynomial basis, we assume prior independence of the prior parameters. Posterior parameters however may be dependent or at least correlated. In this latter case, a numerical conditioning of the parameters samples (e.g. Cholesky-type decomposition), mapping the parameters into centered and normalized uncorrelated coordinates will facilitate the polynomial regression.

### 1.2. Adaptive weighted regression construction

The construction of accurate polynomial surrogates of general functions over large (prior) supports remains very challenging and demanding. Moreover, it is very inefficient in the context of Bayesian inference when the data are informative, i.e. when the posterior density of the inferred parameters highly concentrates from the initial prior density. A more efficient and accurate approximation can be obtained by constructing a “posterior-adapted” surrogate model that would minimize the  $\|\mathbf{z} - \hat{\mathbf{z}}\|_{\ell_2(\pi_{\text{post}})}$ , that is the error norm induced by the posterior. A legitimate concern is the one of the construction of a surrogate in the important region of the posterior distribution before actually characterizing the posterior. As this measure is not known from start, we propose a sequential strategy where a sequence of surrogates  $\hat{\mathcal{M}}_{\text{rom}}^{(k)}$ , for  $k = 1, 2, \dots$ , is iteratively adapted until convergence, on the basis of model evaluations drawn from a sequence of posterior measure approximations  $\hat{\pi}_{\text{post}}^{(k)}(\boldsymbol{\theta}|\mathbf{d})$ .

Several approaches are available for building polynomial approximations: e.g. interpolation, projection, regression, compressive sensing, etc... that only require observations of the mapping  $\boldsymbol{\theta} \mapsto \mathbf{z}(\boldsymbol{\theta})$ . All of these approaches have a computational cost that is dominated by the FM evaluation. Therefore, the goal is to limit

the total number  $N$  of discrete samples of the mapping needed to construct the surrogate model. Because the parametric inference is driven by random sampling, we choose to rely on iterated polynomial regressions over sets of randomly generated points. Moreover, we make the choice of numerically emphasizing the samples for which the current surrogate  $\hat{\mathbf{z}}^{(k)}$  predicts well  $\mathbf{z}$ . This is ensured by associating scalar weights  $\rho^j > 0$  to the samples in the resolution of polynomial regression problem. In the following, we explain how the samples are incorporated, weighted and used in the construction of the surrogates.

Let us assume, for the sake of comprehension, that we dispose of a current surrogate polynomial model  $\hat{\mathcal{M}}_{\text{rom}}^{(k)}$  at some step ( $k$ ) of our sequential process. This surrogate, constructed on the basis of a set of previous weighted model simulations  $\mathcal{S}^{(k)}$ , is fast to evaluate and serves the purpose of the MCMC sampler that produces a new set of parameters drawn from  $\hat{\pi}_{\text{post}}^{(k)}(\boldsymbol{\theta}|\mathbf{d})$  according to the Bayes' rule of Equation (4). One may then consider a few number  $n^{(k)}$  of well-mixed parameters samples  $\boldsymbol{\theta}^{(k), 1 \leq j \leq n^{(k)}}$  of this chain, and performs the corresponding simulations using the ROM FM to compute  $\mathbf{z}(\boldsymbol{\theta}^{(k), j})$ . For each sample, one can then quantify the current surrogate error using a discrepancy measure  $\Delta^j = \left\| \mathcal{G} \left( \mathcal{M}_{\text{rom}}(\boldsymbol{\theta}^{(k), j}, \cdot) \right) - \mathcal{G} \left( \hat{\mathcal{M}}_{\text{rom}}^{(k)}(\boldsymbol{\theta}^{(k), j}, \cdot) \right) \right\|$ . There exists different ways of deriving sample weights from the discrepancy measure, but the main idea is to make it inversely proportional with some numerical bounding capability, i.e.  $\rho^j \sim (\max(\epsilon, \Delta^j))^{-1}$  [23]. If all  $\rho^j$  are large, it means that the surrogate is quite accurate in the sampled approximate posterior region, implying in turn that  $\hat{\pi}_{\text{post}}^{(k)}$  approximates well  $\pi_{\text{post}}$ . On the contrary, when  $\rho^j$  is low, it indicates that the sample  $\boldsymbol{\theta}^{(k), j}$  is not a reliable observation point as it is drawn from an inaccurate posterior. The sample weight completes the observation triplets constituting the set  $\mathcal{A}^{(k)} = \{(\mathbf{z}^j, \boldsymbol{\theta}^j, \rho^j)\}_{j=1,2,\dots,n^{(k)}}$  of observations collected at iteration  $k$ . This set is added to the current set  $\mathcal{A}$  of all observations collected so far:  $\mathcal{A} = \{(\mathbf{z}^j, \boldsymbol{\theta}^j, \rho^j)\}_{j \in \mathcal{J} = \{1,2,\dots,n=1+2+\dots+n^{(k-1)}+n^{(k)}\}}$ .

At this stage, a choice needs to be made in order to *select* among this growing database  $\mathcal{A}$ , the subset of samples that will be used to construct the next polynomial surrogate  $\hat{\mathcal{M}}_{\text{rom}}^{(k+1)}$  at step ( $k+1$ ). Again, there is no unique bandwidth selection choice, and one may only consider selecting among the more recently generated samples, select samples with larger trust indexes across multiple previous iterations, ... Irrespective of the selection strategy, let us denote  $\mathcal{S}^{(k+1)} \equiv \{(\mathbf{z}^j, \boldsymbol{\theta}^j, \rho^j)\}_{j \in \mathcal{J}^{(k+1)} \subseteq \mathcal{J}}$  the selected subset of  $\mathcal{A}$ . Depending on the number and quality of the samples in  $\mathcal{S}^{(k+1)}$ , the approximation space must be tailored. A simple option consists in relating the total order  $p$  of the polynomial approximation to the number of elements in  $\mathcal{S}^{(k+1)}$ .

In this work we used the following strategy. First, to construct the initial ( $k=0$ ) surrogate model we set  $p^{(0)} = 1$  and generate  $n_0$  samples  $\boldsymbol{\theta}^j$  drawn directly from the prior  $\pi_{\text{prior}}(\boldsymbol{\theta})$  to obtain the initial set of triplets using arbitrarily  $\rho^j = 1$ . The size of the initial set is  $n_0 = 3 \times n_{\text{pol}}(p^{(0)})$  where  $n_{\text{pol}}(p)$  is the dimension of the TD polynomial space at degree  $p$ . The whole set is used to compute the initial surrogate,  $\mathcal{S}^{(0)} = \mathcal{A}$ . Then, at each iteration ( $k$ ), we generate  $n_f \times n_{\text{pol}}(p^{(k)})$  new samples from the current surrogate posterior, for some  $n_f \geq 1$  integer. The polynomial degree of the next surrogate is selected according to the rule

$$p^{(k+1)} = \begin{cases} p^{(k)} + 1 & \text{if } n_f \times n_{\text{pol}}(p^{(k)} + 1) \leq |\mathcal{A}|/2, \\ p^{(k)} & \text{otherwise.} \end{cases}$$

For the construction, we attempt to use asymptotically the last half of the collected samples in  $\mathcal{A}$  using the definition  $\mathcal{J}^{(k+1)} = \{n - \max(n/2, n_f \times n_{\text{pol}}(p^{(k+1)}), \dots, n)\}$ . Here, we choose  $n_f = 2$  to ensure at least twice as many samples as polynomial coefficients. The value of  $n_f$  can be increased to improve the stability of the regression in particular when  $p^{(k+1)}$  becomes large.

The surrogate model  $\hat{\mathcal{M}}_{\text{rom}}^{(k+1)}$  is finally determined solving (component-wise) the following weighted least squared residual minimization problem:

For each component, cf. Equation (5), the new model at the next iteration  $\hat{\mathcal{M}}_{\text{rom}}^{(k+1)}$  is therefore obtained by solving a weighted least squared residual minimization problem:

$$\mathbf{a}^{(k+1)} = \underset{\mathbf{b} \in \mathbb{R}^{P^{(k+1)}}}{\operatorname{argmin}} \sum_{j \in \mathcal{J}^{(k+1)}} \rho^j \left| z^j - \sum_{\gamma \in \Lambda_p^{(k+1)}} b_\gamma \psi_\gamma(\boldsymbol{\theta}^j) \right|^2. \quad (6)$$

The surrogate model is then incrementally optimized to minimize the response error in the posterior norm. If  $\mathbf{z}$  presents sufficient regularity with respect to the parameters, one may expect fast convergence with a reasonably low polynomial degree. This translates into significant reductions in the number of FM evaluations. The convergence may be monitored numerically, for instance observing the samples weight distribution. Finally, the iterations are continued until convergence is reached or a maximum number of iterations ( $k_{\max}$ ) or simulations  $|\mathcal{A}| > N$  is attained.

## 2. PULSE WAVE VELOCITIES IDENTIFICATION IN CARDIOVASCULAR SYSTEMIC ARTERIAL NETWORKS

In the following, we propose to apply our technique to the characterization of a subject-specific hemodynamic model of the pulse wave propagation in a vascular network. Quantifying the relationship between the physical properties of the cardiovascular system and the shape of the arterial pulse wave is important because the latter carries valuable information for the diagnosis and treatment of cardiovascular diseases. The coupled interaction of the blood flow with the network of compliant arteries is a (nonlinear) fluid-structure interaction (FSI) problem with a closed distribution network, pulsatile fluid flow with complex internal dynamics and multiple reflections leading to relatively complicated pulse wave patterns [46].

### 2.1. Arterial stiffness

Cardiovascular diseases, such as atherosclerosis, aneurysm, and dissection all involve significant degeneration of the arterial wall tissue, e.g., deposition of calcified materials, reduction in elastin content, plaque formation [45]. These degenerations induce changes in the arterial stiffness, i.e., the capability of the vessel to accommodate and damp the pressure waves generated by the sudden change in blood pressure due to the heart systolic contraction. Arterial stiffening plays a key role in the development of cardiovascular diseases. Even if the exact mechanism by which it leads to increase in pulse pressure and systolic hypertension is still a controversial topic [34], it remains a valid predictor of cardiovascular morbidity and mortality. The problem is those stiffness properties of the aortic walls are not well known and subject to large (spatial) variabilities among individuals [35]. In fact, pulse wave velocities not only fluctuate during a cardiac cycle but also vary along the arterial tree due to the natural stiffening of the arterial walls toward distal locations. Arterial stiffness is hard to measure *in vivo* [47] and require numerical substitutes [1]. Common medical practice is to infer it from an indirect measurement of pressure (averaged) pulse wave velocity in the large vessels of the arterial tree – carotid-femoral or brachial-ankle assessments are common – thanks to medical imaging (Doppler Ultrasound, MRI, CTComputed Tomography) and the Moens-Korteweg equation [22]. For instance, the ankle-brachial arterial pressures index which reflects peripheral arterial obstruction due to atherosclerosis, is an established vascular marker for the diagnosis of peripheral arterial disease which affects more than 20% of the elderly population. In this case, accurate modeling of the arterial pulse wave velocities in the limbs are critical to assessing the revascularization of the lower extremities and its impact on a major cardiovascular risk [18].

Some UQ studies exist that have confirmed that the inherent aortic stiffness uncertainty (either due to biological variability or age-related) had a strong influence in terms of standard cardiovascular outputs such as pulse pressure, brachial to radial pulse pressure amplification index and some waves transit times across the arterial tree, e.g. [10, 12, 24, 28, 37, 51]. Nevertheless, in these studies, the system under consideration is sometimes too generic and the uncertainty modeling too simple and idealized, e.g. disregarding the correlations

and physical constraints among parameters. Other studies have considered the assimilation of patient-specific clinical geometrical or functional data/measurements together with the use of a reduced-order model of the arterial circulation by relying on a deterministic or statistical formulation [3–6, 11, 13, 26]. When measurement error is included in the observations and uncertainty considered in the *forward model* describing the effect of the sought-after parameters onto the outputs, it results in a stochastic inverse problem whose solution is a probabilistic description of the parameters [3], in contrast with a point-wise estimate as in inverse problems.

We propose to apply our approach to the calibration of an arterial hemodynamic model. We treat pulse wave velocities and arterial lumen area of the arteries at rest as piecewise-constant but unknown quantities modeled *a priori* as independent random variables (one for each artery component considered among  $\{A_i\}_{i=1\dots d=7}$  arteries). With the help of a simplified hemodynamic model together with patient-specific non-invasive local measurements of the vessel motion and blood flow velocity, we then apply our iterative approach in a Bayesian framework in order to efficiently calibrate these pulse wave velocities.

## 2.2. Reduced-order model of the distributed blood circulation

### 2.2.1. Governing equations

We consider a portion of the systemic arterial tree containing a finite number of arteries: we have a network of thin, deformable, and axisymmetric arterial segments filled with blood, taken as an incompressible Newtonian fluid [40]. The formulation, for each arterial segment, based on the conservation of mass and momentum laws and on Young-Laplace equation, is:

$$\begin{aligned} \frac{\partial A}{\partial t} + \frac{\partial Au}{\partial x} &= 0 \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} &= -\frac{1}{\rho} \frac{\partial p}{\partial x} + \frac{f}{\rho A} \\ p &= p_0 + \beta(\sqrt{A} - \sqrt{A_0}), \end{aligned} \tag{7}$$

where  $t$  denotes time,  $x \in \mathcal{D}$  is the axial coordinate along the arterial centerline,  $A(x, t)$  is the circular cross-sectional area of the lumen,  $u(x, t)$  and  $p(x, t)$  are average velocity and internal pressure, respectively,  $\rho$  is blood density. The term  $f$  is the friction force per unit length and is related to the velocity profile through  $f = -2\mu\pi \frac{\alpha}{\alpha - 1} u$ , where  $\mu$  is blood dynamic viscosity and  $\alpha \in [0, 1[$  is a correction factor accounting for the nonlinear integration of radial velocities in each cross-section [42]. The underscript  $_0$  denotes quantities at rest. The  $\beta$  parameter is a measure of the arterial wall stiffness related to its mechanical behavior:  $\beta = \sqrt{\pi} h_0 E / (1 - \nu^2) A_0$ , where  $h_0$  is the reference arterial wall thickness,  $E$  is the Young's modulus and  $\nu = 1/2$  is the constant Poisson's ratio. It is also related to the pulse wave velocity  $c$  through the following relation:

$$\beta = 2\rho \frac{c^2(x, t)}{\sqrt{A(x, t)}} = 2\rho \frac{c_0^2}{\sqrt{A_0}}. \tag{8}$$

Note that  $c$  and  $c_0$  increase with increasing elastic modulus and wall thickness, and decreasing luminal area. At some point in the following, due to the lack of knowledge of the pulse wave velocity at rest  $c_0$ , it will be modeled as a random quantity which implicitly denotes the uncertainty in Young's modulus. We refer the reader to [15] for more details about the mathematical form of this system of equations.

### 2.2.2. Numerical solver

In this work, we rely on a discontinuous Galerkin (DG) method with a Legendre spectral/ $hp$  spatial discretization [21] and a Riemann solver of Roe for the calculation of the upwind fluxes, to solve the system of Equations (7). It is a very efficient scheme for high-order discretization of convection-dominated flows as it propagates waves at different frequencies without suffering from the excessive dispersion and diffusion errors.



Detailed mathematical formulation and analysis about the hyperbolic system and the numerical scheme may be found elsewhere, e.g. [9].

We use an explicit second-order Adams-Bashforth scheme with a time step dictated by the CFL condition [21]. In the case of polynomial approximation of order  $p = 4$  within each cell, pulse pressure convergence is reached for an average mesh resolution of 20 cells per unit meter. Dirichlet-type time-dependent boundary conditions at the inlet of the domain are imposed thanks to subject-specific echo-tracking clinical data. At each exit of the simulated arterial tree, a simple terminal reflection is imposed via a scalar coefficient  $R \in [0, 1]$  that relates the backward to the forward characteristic information. It enables the simulation of reflected waves induced by the resistances of the missing peripheral arterioles and capillaries. This coefficient may be interpreted as an avatar of a lumped parameter model made of a single resistance [2, 13].

### 3. NUMERICAL RESULTS

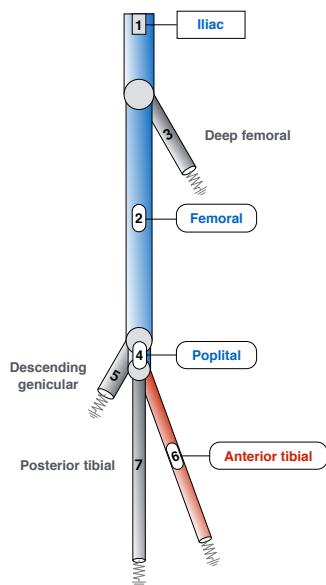


FIGURE 1. Network of the seven main human lower limb arteries. Resistance symbols represent partially reflecting outflow boundary conditions. Some medical imaging data relative to local temporal evolutions of cross-sectional area and blood flow velocity are available in the arteries in blue; only flow velocity data are available for the artery in red; nothing is measured in the arteries in grey color.

In the following, we are interested in a subject-specific data assimilation on an arterial human left lower limb model. Here we consider the same seven-artery simple network model described in details in reference [13], where – only the largest arteries are retained and physiological properties at rest are considered piecewise constant, see Figure (1) and for some of which – we hold some medical imaging data. The non-invasive measurements are collected on the male subject thanks to ultrasound echo-tracking technique [14]. The data are cross-sectional lumen area  $A(x_{\text{center}}, t)$  changes measured at the *approximate* center of the iliac, femoral and popliteal arteries  $\{A_i\}_{i=1,2,4}$ , and blood velocity changes  $u(x_{\text{center}}, t)$  measured at the same locations and at the center of the anterior tibial artery  $\{A_6\}$  as well, see Figures (1,5). These data are collected non-simultaneously for the same subject during two to three cardiac cycles at a time. They exhibit variabilities both due to the imaging acquisition tool and practitioner technique varying accuracy and the subject physiological changes (e.g. heart

beat) during the clinical examination. In order to reduce these uncertainties and also to make the application more challenging, we condense the temporal information into scalar quantities. Contrarily to the study of El Bouti [13], where entire time signals were put to use, here the data are utilized in the form of scalar relative arterial changes measured over time:  $\mathbf{d} \equiv \{d_l\}_{l=1:n_d} \in \mathbb{R}^{n_d=7}$  with:

$$d_l \equiv \left| \frac{\phi_{\{A_{i(l)}\}}(x_{\text{center}}, t^{\text{sys}}) - \phi_{\{A_{i(l)}\}}(x_{\text{center}}, t^{\text{dia}})}{\phi_{\{A_{i(l)}\}}(x_{\text{center}}, t^{\text{dia}})} \right|, \quad (9)$$

where  $\phi_{\{A_{i(l)}\}}$  is the  $l$ -th measured quantity at the center of the artery  $\{A_{i(l)}\}$ , and  $t^{\text{sys}}$  and  $t^{\text{dia}}$  relate to the cardiac systolic (diastolic) time at which the measured quantity reaches its maximum (minimum) respectively. For the data we have access to, the measured quantity is either the lumen area or the blood flow velocity. Due to the lack of information, measurements statistical description is evaluated from the collected time signals and modeled as  $\epsilon = \mathcal{N}(\mu_d, \Sigma_d)$ . In particular,  $\Sigma_d$  is directly approximated from the signals temporal deviation which is a very crude assumption that does not account for instance for measurements errors due to approximate spatial acquisition sites and non-simultaneous acquisition times. The statistical description is summarized in Table 3. The chosen values are such that the coefficient of variations of these measurements exhibit a 40% dispersion.

$l$	$A_{i(l)}$	$\phi$	$\mu_d$	$\Sigma_d$	$c_{v_d}$ (%)
1	A <sub>1</sub>	$A$	$1.2 \cdot 10^{-1}$	$6 \cdot 10^{-3}$	5
2	A <sub>1</sub>	$u$	$3.5 \cdot 10^{-2}$	$4.4 \cdot 10^{-3}$	12.6
3	A <sub>2</sub>	$A$	$4.5 \cdot 10^{-2}$	$5.6 \cdot 10^{-3}$	12.4
4	A <sub>2</sub>	$u$	2.66	0.13	4.9
5	A <sub>4</sub>	$A$	3.65	0.37	10.1
6	A <sub>4</sub>	$u$	2.39	0.24	10
7	A <sub>6</sub>	$u$	2.66	0.13	4.9

TABLE 1. Measurements dataset description and statistics (means, standard deviations and coefficient of variations), cf. Eq. (9).

The main parameters to be inferred are the pulse wave velocities at rest  $\mathbf{c}_0 \in \mathbb{R}^{d=7}$  in all arteries  $\{A_i\}_{i=1,\dots,d=7}$  of the network. Meanwhile, we also infer the resistance parameters at each outflow boundary condition, i.e.  $\mathbf{R} \in \mathbb{R}^{d=4}$  in arteries  $\{A_i\}_{i=3,5,6,7}$  and the lumen cross-section of the arteries at rest that have not been measured, i.e.  $\mathbf{A}_0 \in \mathbb{R}^{d=4}$  in arteries  $\{A_i\}_{i=3,5,6,7}$ . All the parameters,  $\boldsymbol{\theta} \in \mathbb{R}^{d=15}$ , are expressed *a priori* as independent normal random variables with  $p(\boldsymbol{\theta}) \sim \mathcal{N}(\mu_{\boldsymbol{\theta}}, \Sigma_{\boldsymbol{\theta}})$ . Mean values  $\mu_{\boldsymbol{\theta}}$  and standard deviation  $\Sigma_{\boldsymbol{\theta}}$  are chosen from the literature [39], and correspond to coefficient of variations of the order of  $c_{v_{\boldsymbol{\theta}}} \sim 15\%$ . In practice, these statistics insure positivity and satisfy the hyperbolicity condition of the system [51] almost surely.

The iterative surrogate model of the system response is adaptively constructed based on the following numerical tools: – a deterministic parallelized DG 1D fluid-structure interaction hemodynamic solver (here with a typical resolution:  $h \sim 0.2$  cell/cm; and a Legendre polynomial approximation basis of order  $p = 4$  in each mesh cell); – an automated data samples selector and preconditioner; – a set of orthonormal polynomial (here Hermite) libraries of arbitrary degree  $p$  and cardinality  $n_{\text{pol}}$  (here relying on tensor-products constructed to satisfy a *total degree* expansion) and – a weighted least-square  $w$ -LS solver; the posterior sampling of the parameters is handled thanks to parallel MCMC chains (Metropolis-Hastings scheme) of  $3.2 \cdot 10^5$  samples.

For this application, the forward model is too complex to make possible the existence of an exact solution mapping the quantities of interest to the uncertain inputs. Due to this lack of reference solution, the convergence of the numerical approach may be apprehended by monitoring the evolution of the surrogate model as well as the posterior characteristics of the parameters to be inferred. In Figure 2-(a), we illustrate the convergence

of the adaptive surrogate construction for  $N = 2500^1$  and  $p^{(0)} = 1$ , reporting the evolution of the samples trust-index value (colored dots) and their average (red solid line) over the new data completing at each iteration ( $k$ ) the total data set  $\mathcal{A}$ . The results are reported as a function of the cumulated number of exact model solves; also reported is the evolution of the surrogate model order  $p^{(k)}$  and the corresponding set of selected  $\mathcal{J}^{(k)}$  samples in Figures 2–(b-c). We see from the plot that the averaged trust-index globally increases, denoting the progressive improvement of the surrogate. Vertical black dotted lines materialize the points at which the surrogate order is automatically incremented. We notice patent increases when the surrogate order is augmented. Posterior statistics convergence (not displayed here) show that mean values are nicely calibrated and reach asymptotically statistically stable values after about  $\sim 500$  model evaluations. Concerning the std results, the data are informative enough so that the uncertainty ranges are often reduced.

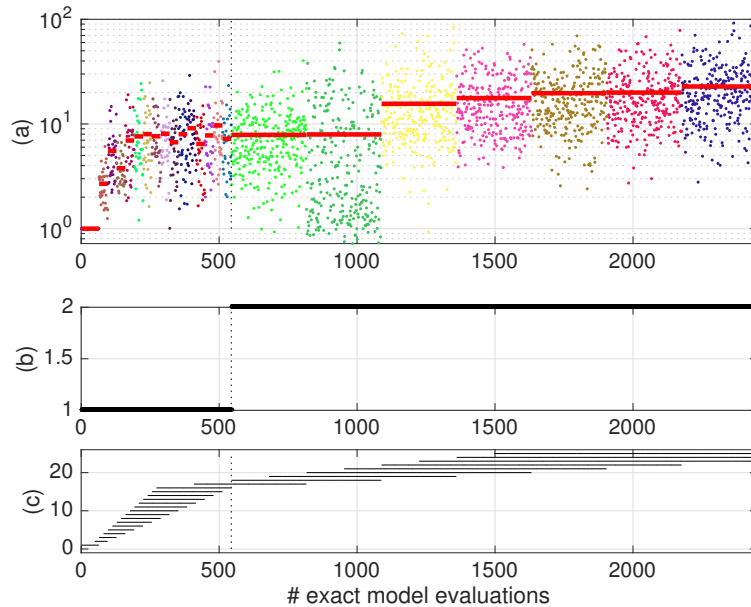


FIGURE 2. Evolutions of the sample trust-indices (colored dots) and their averages (solid red lines) (a), selected polynomial order  $p^{(k)}$  (b) and corresponding selected  $\mathcal{J}^{(k)}$  samples (c) used to build the successive surrogates vs. the total number of hemodynamic model evaluations.

Figure 3 presents one- and (joint) two-dimensional marginal distributions of the pulse wave velocities at rest  $\mathbf{c}_0$  obtained from kernel density estimates at the final iteration of the numerical approximation procedure. Other inferred parameters distributions are not represented as they are less influent. Globally we can say that the measurements in  $\{A_{2,4,6}\}$  arteries are informative enough in the sense that the stiffness uncertainty of the material system is lowered except in distal arteries  $\{A_{3,5,7}\}$ . It is particularly true for the top  $\{A_1\}$  artery. Located upstream of the network and contiguous of the deterministic inflow boundary condition, the system response in this artery is not prone to large uncertainties. Feeding on the downstream measurements, the arterial stiffness of this vessel is therefore sharply calibrated. This directly benefits the calibration of the connected downstream main arteries, in particular,  $\{A_{2,4,6}\}$ . In fact, we observe very noticeable correlations between the stiffness of arteries  $\{A_{1,2}\}$ ,  $\{A_{1,4}\}$  and also  $\{A_{2,4}\}$ . This makes sense due to the topology of the network, see Figure 1. We also notice that joint measures between the stiffness of  $\{A_6\}$  and each of  $\{A_{1,2,4}\}$  get nicely informed from the bayesian updating. The stiffness of artery  $\{A_7\}$  is somewhat updated but one

<sup>1</sup>This computational budget was chosen to match the total number of solver evaluations of the optimization algorithm in reference [13].

Artery #	Nominal			CMA-ES optimization			Bayesian inference (MAP)		
	$A_0$ ( $cm^2$ )	$c_0$ ( $m/s$ )	$R$	$A_0$ ( $cm^2$ )	$c_0$ ( $m/s$ )	$R$	$A_0$ ( $cm^2$ )	$c_0$ ( $m/s$ )	$R$
1 <sup>(A,u)</sup>	<b>0.53</b>	7.20	-	<b>0.53</b>	4.99	-	<b>0.53</b>	5.35	-
2 <sup>(A,u)</sup>	<b>0.47</b>	9.15	-	<b>0.47</b>	11.96	-	<b>0.47</b>	11.28	-
3	0.3	8.32	0.65	0.27	7.39	0.48	0.29	<i>8.86</i>	0.65
4 <sup>(A,u)</sup>	<b>0.4</b>	9.50	-	<b>0.40</b>	10.51	-	<b>0.40</b>	10.9	-
5	0.3	9.95	0.65	0.32	9.93	0.42	0.33	8.7	0.52
6 <sup>(u)</sup>	0.2	11.12	0.65	0.23	11.61	0.95	0.20	<i>8.97</i>	0.75
7	0.2	14.03	0.65	0.22	9.91	0.73	0.20	<i>13.85</i>	0.67

TABLE 2. Nominal and calibrated hemodynamic model parameters values estimated from optimization methods [13] and statistical inference (bold figures are obtained from measurements and are fixed). Calibrated pulse wave velocities relative error between genetic optimization and Bayesian inference is about 15% in average;  $c_0$  values represented in italic exceed this threshold.

should be cautious as previous results have shown that the surrogate model might not be fully stabilized for this vessel. Finally, there still exists significant residual uncertainty in the arterial stiffness of  $\{A_3\}$  and  $\{A_7\}$ . Joint marginal distributions between distal outflow arteries without measurements (i.e.  $\{A_{3,5,7}\}$ ) do not get informed and remain very similar to their prior distributions. We have noticed that the final level of uncertainty

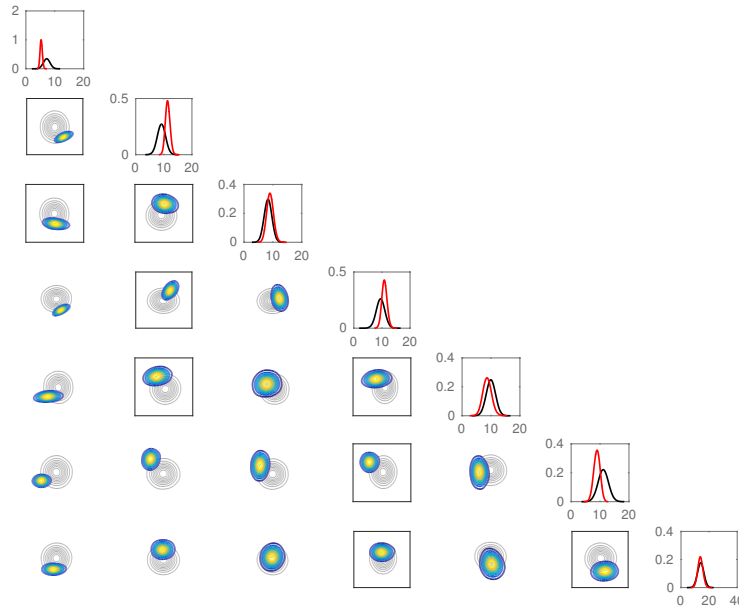


FIGURE 3. One- and two-dimensional pulse wave velocities  $c_0$  posterior (colored lines) vs. prior (gray lines) marginal distributions. Posterior distributions are estimated from the final iteration of the posterior-adapted surrogate method proposed. Results from artery  $\{A_1\}$  to  $\{A_7\}$  are ordered from top to bottom row and from left to right column. Results for arteries sharing a bifurcation node are framed with an axis box.

is considerably reduced for some of the parameters. In particular, the posterior distributions are sometimes far distant from the prior measures, e.g.  $\hat{\pi}_{\text{post}}^{(k_{\text{final}})}(c_{0_1}, c_{0_2})$  or  $\hat{\pi}_{\text{post}}^{(k_{\text{final}})}(c_{0_2}, c_{0_6})$ . Results of one-dimensional marginals in Figure 4 are an example of the capability of the method to gradually reach over remote posterior regions

of interest. Once the right region is reached, more available samples render possible the construction of a higher-order polynomial approximation of the response, which makes the surrogate-based posterior sampling more efficient.

The Bayesian formulation also gives access to a valuable point-wise estimate of the inferred parameters: the

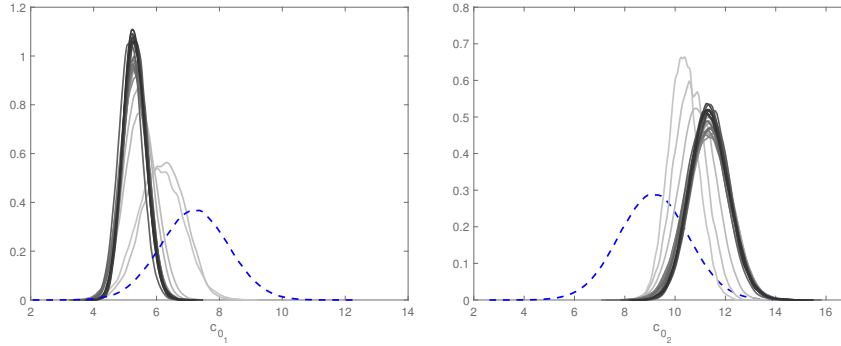


FIGURE 4. Pulse wave velocities  $\mathbf{c}_0$  marginal distribution in the iliac (left) and femoral (right) arteries. Dashed lines represent prior distributions. Gradual color shading illustrates the sequential progression of posterior distributions toward convergence (from light to dark colors) in regions of low prior probability.

maximum *a posteriori* probability (MAP) estimate that is defined as the mode of the posterior distribution:  $\hat{\boldsymbol{\theta}}_{\text{MAP}} = \text{argmax}_{\boldsymbol{\theta}} \hat{\pi}_{\text{post}}(\boldsymbol{\theta}|\mathbf{d})$ . In the following, we compare the results of the numerical model relying on the MAP hemodynamic parameters with those calibrated with a different approach. Another interesting subsequent study would have been to perform an uncertainty quantification of the effect of the entire probabilistic description of the calibrated parameters onto the quantity of interest.

Finally, the adaptively weighted regression approach yielded substantial gains in efficiency and accuracy over methods using direct prior-based surrogate models. A practical comparison – not presented here – was made for a similar but simpler version of the lower limb system introduced previously, i.e. with only seven stiffness parameters to calibrate. Our adaptive surrogate model approach was about one order of magnitude more efficient in terms of number of calls to the deterministic solver compared to a calibration based on a *global* surrogate model constructed over the prior parametric domain from a Gauss-Hermite Smolyak-designed sparse grid.

### 3.1. Comparison with an evolutionary optimization approach

A different pathway to the calibration of hemodynamic model parameters has been used in a previous work [13]. In that study, a Covariance Matrix Adaptation Evolution Strategy (CMA-ES) had been chosen. Is it an efficient stochastic, derivative-free evolution strategy with an auto-adaptive covariance matrix for numerical optimization of non-linear or non-convex continuous optimization problems, based on the principle of biological evolution [19]. Table 2 summarizes the different calibrated parametric results and compares them with nominal values. We mention that the CMA-ES-optimized parameters are averaged values obtained over ten reiterated optimization procedures. The first column designating the artery number also indicates what type of subject-specific data was measured for each artery (either cross-sectional area and/or blood velocity). We see that the MAP parameters obtained from the Bayesian inference are close to the ones obtained via the optimization procedure, especially in the upper part of the network where data are available. Discrepancies do occur in the lowest section, i.e. the tibial arteries. These differences may be interpreted in the light of several factors. First, it should be noted that optimized parameters in this section are subject to larger variabilities due to the lack of measured data [13]. Moreover, one should keep in mind that the information content of the data provided to the statistical inference is more limited. Indeed, it only provides relative magnitude scales with no information

about phasing nor timing.

Finally, figure 5 compares the predictions based on the nominal and MAP parameters values with the subject-specific measurements. The improvement is patent with the calibrated parameters, except for the cross-sectional lumen area in the iliac and the tibial arteries. Indeed, in the first case, the imposition of the upstream boundary condition data in the form of the measured area, is sufficient to drive the dilatation of the iliac artery to the right profile, even with parameters nominal values. In the second case, the measurements are not informative enough to induce real changes in the lumen area profile of the tibial arteries. Predictions obtained from the optimization procedure, see Fig. 9 in [13] look very similar, in particular for the first, second and third generations of arteries of the bifurcating network.

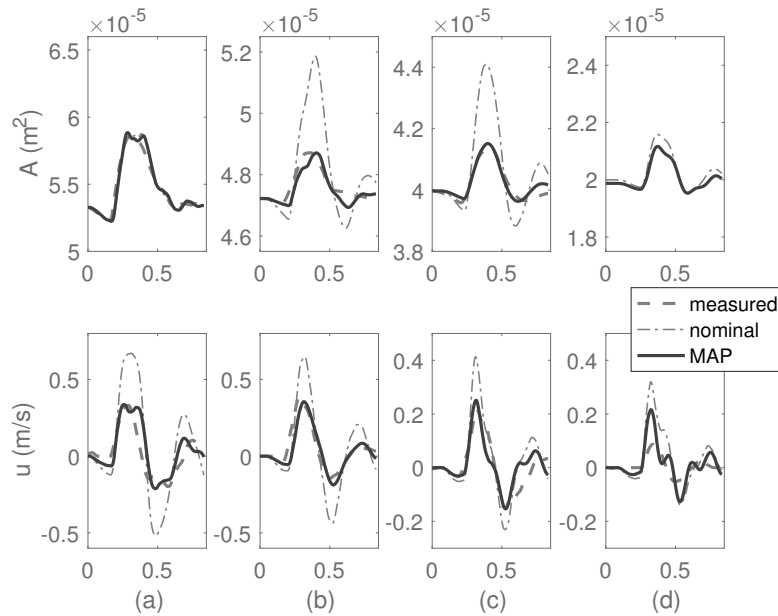


FIGURE 5. Comparison between measured, nominal and calibrated cross-sectional lumen area (top row) and blood flow velocity (bottom row) waveforms at the midpoint of some of arteries of the lower limb network in Figure 1, with (a): iliac, (b): femoral, (c): popliteal and (d): anterior tibial arteries.

#### 4. CONCLUSION

We have developed a computational approach in which computationally intensive cardiovascular models with unknown physiological parameters can be approximated and then calibrated via sample-based Bayesian tools and subject-specific data. To make the procedure computationally tractable, the full forward cardiovascular model or parameter-to-observable map are approximated at the deterministic and stochastic levels. More specifically, in this work, we introduce a novel adaptive pathway to construct a series of parametric polynomial surrogate models gradually adapting to minimize the approximation with respect to the posterior measure of the parameters. In our experiments, the approach yields substantial gains in efficiency and accuracy over prior-based surrogate models. The approach is applied to subject-specific pulse wave velocities identification in a human lower limb arterial network. Pulse propagation in a vascular network is a challenging multi-physics problem dominated by numerous high-speed pressure waves traveling forward and backward in the network, due to multiple bifurcations and terminal reflections. Despite this complexity, the numerical method succeeds in calibrating the physiological parameters adequately and the most probable parameter values compare well

with the ones obtained using an alternative approach based on an evolutionary optimization algorithm. In addition, our approach provides richer information in the form of a complete statistical characterization of the inferred parameters. This result is remarkable considering that: a) the deterministic reduced-order model is quite simplified compared to a more sophisticated full three-dimensional fluid-structure interaction model, both in terms of network geometry, scales, and numerical modeling assumptions; b) the subject-specific *real* clinical data of different nature are scarce and incomplete; and c) the computational budget is relatively modest regarding the high-dimensional parametric space to explore.

The first author is greatly thankful to Prof. L. Dumas from the Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay, France, for the organization and the invitation at the “Mathematical models in biology and medicine” school in Mauritius in 2016. The authors are also thankful to Prof. P. Boutouyrie from the Paris-Cardiovascular research Center (PARCC) at the Georges-Pompidou European Hospital (HEGP), Paris, France, for providing the echo-tracking measurements.

## REFERENCES

- [1] J. ALASTRUEY, *Numerical assessment of time-domain methods for the estimation of local arterial pulse wave speed*, Journal of Biomechanics, 44 (2011), pp. 885–891.
- [2] J. ALASTRUEY, K. H. PARKER, J. PEIRO, AND S. J. SHERWIN, *Lumped parameter outflow models for 1-d blood flow simulations: Effect on pulse waves and parameter estimation*, Communications in Computational Physics, 4 (2008), pp. 317–336.
- [3] F. AURICCHIO, M. CONTI, A. FERRARA, AND E. LANZARONE, *A clinically applicable stochastic approach for noninvasive estimation of aortic stiffness using computed tomography data*, IEEE Trans Biomed Eng, 62 (2015), pp. 176–187.
- [4] L. BERTAGNA, M. D’ELIA, M. PEREGO, AND A. VENEZIANI, *Data Assimilation in Cardiovascular Fluid–Structure Interaction Problems: An Introduction*, Springer Basel, Basel, 2014, pp. 395–481.
- [5] C. BERTOGLIO, D. BARBER, N. GADDUM, I. VALVERDE, M. RUTTEN, P. BEERBAUM, P. MOIREAU, R. HOSE, AND J.-F. GERBEAU, *Identification of artery wall stiffness: In vitro validation and in vivo results of a data assimilation procedure applied to a 3d fluid–structure interaction model*, Journal of Biomechanics, 47 (2014), pp. 1027 – 1034.
- [6] J. BIEHLER, S. KEHL, M. W. GEE, F. SCHMIES, J. PELISEK, A. MAIER, C. REEPS, H.-H. ECKSTEIN, AND W. A. WALL, *Probabilistic noninvasive prediction of wall properties of abdominal aortic aneurysms using bayesian regression*, Biomechanics and Modeling in Mechanobiology, 16 (2017), pp. 45–61.
- [7] A. BIROLLEAU, G. POËTTE, AND D. LUCOR, *Adaptive Bayesian inference for discontinuous inverse problems, application to hyperbolic conservation laws*, Communications in Computational Physics, 16 (2014), pp. 1–34.
- [8] E. BOILEAU, P. NITHIARASU, P. J. BLANCO, L. O. MULLER, F. E. FOSSAN, L. R. HELLEVIK, W. P. DONDEERS, W. HUBERTS, M. WILLEMET, AND J. ALASTRUEY, *A benchmark study of numerical schemes for one-dimensional arterial blood flow modelling*, International Journal for Numerical Methods in Biomedical Engineering, 31, p. e02732.
- [9] E. BOLLACHE, N. KACHENOURA, A. REDHEUIL, F. FROUIN, E. MOUSSEAU, P. RECHO, AND D. LUCOR, *Descending aorta subject-specific one-dimensional model validated against in vivo data*, Journal of Biomechanics, 47 (2014), pp. 424–431.
- [10] A. BRAULT, L. DUMAS, AND D. LUCOR, *Uncertainty quantification of inflow boundary condition and proximal arterial stiffness-coupled effect on pulse wave propagation in a vascular network*, International Journal for Numerical Methods in Biomedical Engineering, (2017). e2859 cnm.2859.
- [11] A. CAIAZZO, F. CAFORIO, G. MONTECINOS, L. O. MULLER, P. J. BLANCO, AND E. F. TORO, *Assessment of reduced-order unscented kalman filter for parameter identification in 1-dimensional blood flow models using experimental data*, International Journal for Numerical Methods in Biomedical Engineering, (2017), pp. n/a–n/a. cnm.2843.
- [12] P. CHEN, A. QUARTERONI, AND G. ROZZA, *Simulation-based uncertainty quantification of human arterial network hemodynamics*, International Journal for Numerical Methods in Biomedical Engineering, 29 (2013), pp. 698–721.
- [13] L. DUMAS, T. EL BOUTI, AND D. LUCOR, *A robust and subject-specific hemodynamic model of the lower limb based on non-invasive arterial measurements*, J. Biomech. Eng., doi: 10.1115/1.4034833 (2016).
- [14] T. EL BOUTI, L. DUMAS, AND D. LUCOR, *Blood flow modeling and application to non-invasive determination of arterial stiffness*, in ECCOMAS CFD 2014 Proceedings of the VI European Conference on Computational Fluid Dynamics, A. Huerta, ed., Barcelona, Spain, July 2014.
- [15] L. FORMAGGIA, D. LAMPONI, AND A. QUARTERONI, *One-dimensional models for blood flow in arteries*, Journal of Engineering Mathematics, (2003), pp. 251–276.
- [16] M. FRANGOS, Y. MARZOUK, K. WILLCOX, AND B. VAN BLOEMEN WAANDERS, *Surrogate and Reduced-Order Modeling: A Comparison of Approaches for Large-Scale Statistical Inverse Problems*, John Wiley & Sons, Ltd, 2010, pp. 123–149.
- [17] R. GHANEM AND P. SPANOS, *Stochastic finite elements: a spectral approach*, Dover Pubns, 2003.

- [18] G. GIUGLIANO, L. D. SERAFINO, C. PERRINO, V. SCHIANO, E. LAURENZANO, S. CASSESE, M. D. LAURENTIS, G. G. SCHIATTARELLA, L. BREVETTI, A. SANNINO, G. GARGIULO, A. FRANZONE, C. INDOLFI, F. PISCIONE, B. TRIMARCO, AND G. ESPOSITO, *Effects of successful percutaneous lower extremity revascularization on cardiovascular outcome in patients with peripheral arterial disease*, International Journal of Cardiology, 167 (2013), pp. 2566 – 2571.
- [19] N. HANSEN AND A. OSTERMEIER, *Completely derandomized self-adaptation, evolution strategies*, Evolutionary Computation, 9 (2001), pp. p.159–195.
- [20] D. HUYNH, D. KNEZEVIC, AND A. PATERA, *Certified reduced basis model validation: A frequentistic uncertainty framework*, Computer Methods in Applied Mechanics and Engineering, 201–204 (2012), pp. 13 – 24.
- [21] G. E. KARNIADAKIS AND S. SHERWIN, *Spectral/hp element methods for CFD*, Oxford University Press, New York, second edition (2005).
- [22] D. KORTEWEG, *Über die Fortpflanzungsgeschwindigkeit des Schalles in elastischen Röhren*, Ann. Phys. Chem. (NS), 5 (1878), pp. 525–527.
- [23] J. V. LANGENHOVE, D. LUCOR, AND A. BELME, *Robust uncertainty quantification using preconditioned least-squares polynomial approximations with  $l_1$ -regularization*, International Journal for Uncertainty Quantification, 6 (2016), pp. 57–77.
- [24] T. LASSILA, A. MANZONI, A. QUARTERONI, AND G. ROZZA, *A reduced computational and geometrical framework for inverse problems in hemodynamics*, International journal for numerical methods in biomedical engineering, 29 (2013), pp. 741–776.
- [25] O. P. LE MAÎTRE AND O. M. KNIO, *Spectral Methods for Uncertainty Quantification with applications to Computational Fluid Dynamics*, Scientific Computation, Springer Netherlands, 2010.
- [26] C. A. LEGUY, E. M. BOSBOOM, H. GELDERBLUM, A. P. HOEKS, AND F. N. VAN DE VOSSE, *Estimation of distributed arterial mechanical properties using a wave propagation model in a reverse way*, Med Eng Phys, 32 (2010), pp. 957–67.
- [27] J. LI AND Y. MARZOUK, *Adaptive construction of surrogates for the bayesian solution of inverse problems*, SIAM J. Sci. Comp., 36 (2015), pp. A1163–A1186.
- [28] D. LOMBARDI, *Inverse problems in 1D hemodynamics on systemic networks: A sequential approach*, International Journal for Numerical Methods in BioMedical Eng., 34 (2013).
- [29] Y. MARZOUK AND D. XIU, *A stochastic collocation approach to bayesian inference in inverse problems*, Comm. Comput. Phys., 6 (2009), pp. 826–847.
- [30] X. N., A. J., AND F. C., *A systematic comparison between 1-D and 3-D hemodynamics in compliant arterial models*, International Journal of Numerical Methods in Biomedical Engineering, 30 (2014), pp. 204–231.
- [31] M. S. OLUFSEN, C. S. PESKIN, W. Y. KIM, E. M. PEDERSEN, A. NADIM, AND J. LARSEN, *Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions*, Annals of Biomedical Engineering, 28 (2000), pp. 1281–99.
- [32] J. S. PETERSON, *The reduced basis method for incompressible viscous flow calculations*, SIAM Journal on Scientific and Statistical Computing, 10 (1989), pp. 777–786.
- [33] C. E. RASMUSSEN AND C. K. I. WILLIAMS, *Gaussian Processes for Machine Learning*, The MIT Press, 2006.
- [34] P. REYMOND, N. WESTERHOF, AND N. STERGIOPULOS, *Systolic hypertension mechanisms: Effect of global and local proximal aorta stiffening on pulse pressure*, Ann. Biomed. Eng., 40 (2012), pp. p.742–749.
- [35] S. ROCCABIANCA, C. FIGUEROA, G. TELLIDES, AND J. HUMPHREY, *Quantification of regional differences in aortic stiffness in the aging human*, Journal of the mechanical behavior of biomedical materials, 29 (2014), p. 10.1016/j.jmbbm.2013.01.026.
- [36] S. SANKARAN, L. GRADY, AND C. A. TAYLOR, *Impact of geometric uncertainty on hemodynamic simulations using machine learning*, Computer Methods in Applied Mechanics and Engineering, 297 (2015), pp. 167 – 190.
- [37] S. SANKARAN AND A. L. MARSDEN, *A stochastic collocation method for uncertainty quantification and propagation in cardiovascular simulations*, Journal of Biomechanical Engineering, 133 (2011), pp. 1–12.
- [38] D. E. SCHIAVAZZI, G. ARBIA, C. BAKER, A. M. HLAVACEK, T. Y. HSIA, A. L. MARSDEN, I. E. VIGNON-CLEMENTEL, AND T. M. OF CONGENITAL HEARTS ALLIANCE (MOCHA) INVESTIGATORS, *Uncertainty quantification in virtual surgery hemodynamics predictions for single ventricle palliation*, International Journal for Numerical Methods in Biomedical Engineering, 32 (2016), pp. e02737–n/a. e02737 cnm.2737.
- [39] S. J. SHERWIN, L. FORMAGGIA, AND J. PEIRÓ, *Computational modelling of 1D blood flow with variable mechanical properties*, in ECCOMAS CFD, Swansea, UK, September 2001.
- [40] S. J. SHERWIN, L. FORMAGGIA, J. PEIRO, AND V. FRANKE, *Computational modelling of 1D blood flow with variable mechanical properties and its application to the simulation of wave propagation in the human arterial system*, International Journal for Numerical Methods in Fluids, 43 (2003), pp. 673–700.
- [41] L. SIROVICH, *Turbulence and the dynamics of coherent structures part i: Coherent structures*, Quarterly of Applied Mathematics, 45 (1987), pp. 561–571.
- [42] N. SMITH, A. J. PULLAN, AND P. J. HUNTER, *An anatomically based model of coronary blood flow and myocardial mechanics*, SIAM J. Appl. Math., 62 (2001).
- [43] A. J. SMOLA AND B. SCHÖLKOPF, *A tutorial on Support Vector Regression*, Tech. Rep. TR-98-030, NeuroCOLT, 1998.
- [44] M. TATANG, W. PAN, R. PRINN, AND G. MCRAE, *An efficient method for parametric uncertainty analysis of numerical geophysical models.*, Journal of Geophysical Research, 102 (1997), pp. 21925–21932.



- [45] M. THIRIET, *Biology and Mechanics of Blood Flows. Part II: Mechanics and Medical Aspects*, CRM Series in Mathematical Physics, Springer, 2007.
- [46] F. N. VAN DE VOSSE AND N. STERGIOPULOS, *Pulse wave propagation in the arterial tree*, Annual Review of Fluid Mechanics, 43 (2011), pp. 467–499.
- [47] J. VAPPOU, J. LUO, AND E. E. KONOFAGOU, *Pulse wave imaging for noninvasive and quantitative measurement of arterial stiffness in vivo*, American Journal of Hypertension, 23 (2010), pp. 393–398.
- [48] M. WILLEMET, V. LACROIX, AND E. MARCHANDISE, *Validation of a 1D patient-specific model of the arterial hemodynamics in bypassed lower-limbs: Simulations against in vivo measurements*, Medical Engineering and Physics, 35, pp. 1573–1583.
- [49] N. XIAO, J. D. HUMPHREY, AND C. A. FIGUEROA, *Multi-scale computational model of three-dimensional hemodynamics within a deformable full-body arterial network*, Journal of Computational Physics, 244 (2013), pp. 22 – 40. Multi-scale Modeling and Simulation of Biological Systems.
- [50] D. XIU AND G. KARNIADAKIS, *The Wiener-Askey polynomial chaos for stochastic differential equations*, SIAM Journal on Scientific Computing, 24 (2002), pp. 619–644.
- [51] D. B. XIU AND S. J. SHERWIN, *Parametric uncertainty analysis of pulse wave propagation in a model of a human arterial network*, Journal of Computational Physics, 226 (2007), pp. 1385–1407.