

## RESEARCH ARTICLE - FUNDAMENTAL

# Bayesian optimisation for efficient parameter inference in a cardiac mechanics model of the left ventricle

Agnieszka Borowska | Hao Gao | Alan Lazarus | Dirk Husmeier\*

<sup>1</sup>School of Mathematics and Statistics,  
University of Glasgow, Glasgow, UK

## Correspondence

Dirk Husmeier, School of Mathematics and  
Statistics, University of Glasgow, Glasgow  
G12 8QQ, UK. Email:  
Dirk.Husmeier@glasgow.ac.uk

## Summary

We consider parameter inference in cardio-mechanic models of the left ventricle, in particular the one based on the Holtzapfel-Ogden (HO) constitutive law, using clinical in vivo data. The equations underlying these models do not admit closed form solutions and hence need to be solved numerically. These numerical procedures are computationally expensive making computational run times associated with numerical optimisation or sampling excessive for the uptake of the models in the clinical practice. To address this issue, we adopt the framework of Bayesian optimisation (BO), which is an efficient statistical technique of global optimisation. BO seeks the optimum of an unknown black-box function by sequentially training a statistical surrogate-model and using it to select the next query point by leveraging the associated exploration-exploitation trade-off. To guarantee that the estimates based on the in vivo data are realistic also for high-pressures, unobservable in vivo, we include a penalty term based on a previously published empirical law developed using ex vivo data. Two case studies based on real data demonstrate that the proposed BO procedure outperforms the state-of-the-art inference algorithm for the HO law.

## KEYWORDS:

Global optimisation; statistical inference; Holtzapfel-Ogden law; biomechanical model calibration

## 1 | INTRODUCTION

The last decade has seen impressive progress in the mathematical modelling of soft-tissue mechanics<sup>1,2</sup>. In particular, the specific application to cardiac mechanics<sup>3</sup> promises to improve our understanding of cardiac physiology and pathology. A recent case-control study<sup>4</sup> involving patients diagnosed with myocardial infarction (MI) and healthy volunteers has demonstrated that the biophysical parameters of a cardio-mechanic model of the left ventricle (LV) of the heart can be non-invasively estimated from cardiac magnetic resonance (CMR) scans, and that the parameters thus inferred allow for an accurate classification of

subject disease status. However, the cardio-mechanic equations do not admit a closed-form solution and require computationally expensive numerical procedures based on finite element method (FEM)<sup>5</sup>. These procedures have to be repeated iteratively as part of a numerical optimization or sampling procedure, leading to potentially excessive computational run times that discourage uptake of the methodology in the clinical practice.

We aim to address this difficulty by adopting the framework of Bayesian optimization (BO), which has been particularly designed for optimization problems where each evaluation of the objective function is expensive and the number of such evaluations has to be kept as low as possible. The main idea is to sequentially approximate the unknown objective function by a statistical model and use this model to identify the next query point, typically according to some exploration-versus-exploitation trade-off criterion. In the literature the said statistical model is typically referred to as a *surrogate function*<sup>6,7</sup> and Gaussian processes<sup>8</sup> are commonly used in this context. (Surrogates are also called approximations, metamodels, response surface models or emulators<sup>9</sup>. We reserve the latter term for pre-trained approximations based on considerably larger numbers of points compared to those used in BO.) The trade-off criterion is quantified by the so-called *acquisition function*. Several strategies on how to specify the details of BO, in particular the acquisition function, have been proposed in the literature<sup>6</sup>. In this study, we use a popular criterion called Expected Improvement<sup>10</sup> as it admits a closed form expression for Gaussian process surrogates, is conceptually simple and performs well for a range of problems<sup>11</sup>.

The purpose of the present paper is to adapt the BO framework to the problem of inferring passive myocardial properties of healthy human LVs from in vivo measurements. The passive myocardial properties are characterised by an invariant-based hyperelastic material model proposed by Holzapfel and Ogden<sup>12</sup> (HO law), which has been widely used in cardiac modelling communities. The first attempt to address parameter estimation, also known as the inverse problem, in this context was made by Gao et al.<sup>13</sup> These authors proposed an algorithmic, multi-step procedure to infer the parameters of the HO law, in which in each step a subset of parameters is optimised with possibly different objective functions. The choice of particular subsets was based on expert knowledge of the myocardial properties of the LV. Our proposed BO approach is conceptually different, as BO iterations are run with respect to an *a priori* chosen parametrisation and objective function, thus requiring little expert knowledge.

To the best of our knowledge we are the first to introduce BO into cardiac mechanics. Our work contributes both to the the global optimisation literature and to the application specific literature. Our first main contribution is developing a BO framework for the challenging problem of parameter inference in the HO law. Specifically, we propose modifications of the two components of BO, the target function and the acquisition function. The former modification consists in augmenting the target function by incorporating prior knowledge of the end-diastolic volume-pressure relationship from an empirical law established by Klotz et al.<sup>14</sup> based on ex vivo data. This allows us to obtain more realistic myocardial properties not only at physiologically normal pressures but also at high pressures, usually not observed in vivo. Generally, the myocardium response to increasing pressure is nonlinear: it appears linear for typical low pressures observed in vivo for healthy subjects, and exponential for increasing

high pressures, such as those observed e.g. for patients with hypertension. It can be challenging to infer those parameters that characterise the nonlinear stiffening effects based solely on in vivo data.

Our second main contribution is adapting the standard Expected Improvement criterion<sup>10</sup> to the challenges of inference in the HO law. First, following previous studies<sup>15,16</sup>, we allow for unknown constraints in the parameter space by training a classifier based on a probit model with Gaussian process priors. Unknown constraints are related to those parameter regions that violate the physiological assumptions of cardio-mechanic models, which is then demonstrated as crashes of the associated forward simulators. Second, next to the standard approach of developing a surrogate for the objective function as a whole, we develop an alternative paradigm, based on a *partial error surrogate*, in which we develop surrogates for individual squared error terms of the target function. This is motivated by a previous study<sup>17</sup> in the context of parameter inference in the HO law, which reported a substantial improvement in accuracy in the emulation framework resulting from emulating a vector of outputs of the cardio-mechanic model. Our proposed approach based on using a surrogate for the squared error terms instead of outputs has an advantage of being directly applicable to the Expected Improvement acquisition function, without requiring any approximation.

Our third main contribution is related to two empirical studies. First, we demonstrate that our proposed BO framework outperforms the state-of-the-art algorithm of Gao et al.<sup>13</sup> for parameter estimation in cardiac mechanic models. Second, we show how incorporating the prior knowledge about volume-pressure relationship (based on an empirical law established by Klotz et al.<sup>14</sup> using ex vivo data) allows for physiologically more realistic parameter estimation consistent with the LV's characteristics in high pressure regimes, relevant e.g. for patients with hypertension or patients with diastolic heart failure<sup>18</sup>.

The structure of this paper is as follows. Section 2 provides necessary details of the cardio-mechanic model of interest and introduces the state-of-the-art inference method<sup>13</sup>. We discuss BO methodology, including the proposed extensions, in Section 3. Our three empirical studies are presented in Section 4. Section 5 concludes the paper and presents the outline for further research.

## 2 | CARDIO-MECHANIC MODEL OF THE LEFT VENTRICLE IN DIASTOLE

Biomechanical modelling of the LV in diastole is concerned with describing the process of diastolic filling, starting at early diastole and finishing at the end of diastole, which can be considered as a purely passive response. The nonlinear myocardial passive response means the myocardium is soft under lower pressure but becomes stiffer and stiffer under higher pressures because of gradually engaged collagen fibres<sup>12</sup>. A few models of the LV biomechanics have been proposed in the literature, see a recent review by Chabiniok et al.<sup>2</sup>. In this work, we follow previous studies<sup>19,13</sup> and focus on the constitutive law proposed by Holzapfel and Ogden<sup>12</sup> to describe passive myocardial response, in other words the myocardial stiffness. The HO law is an invariant-based hyperelastic constitutive law which takes into account the layered myofibre structure of the myocardium.

Various studies have demonstrated that it can realistically capture the key properties of the myocardium and at the same time is relatively easy to implement numerically. For these reasons it has been widely used in cardiac modelling communities<sup>20,21,22,4,3</sup>. The HO law is based on the following strain energy function

$$W = \frac{a}{2b} \{\exp[b(I_1 - 3)] - 1\} + \sum_{i \in \{f, s\}} \frac{a_i}{2b_i} \{\exp[b_i(I_{4i} - 1)^2] - 1\} + \frac{a_{fs}}{2b_{fs}} [\exp(b_{fs} I_{8fs}^2) - 1], \quad (1)$$

where  $I_i$ ,  $i \in \{1, 4f, 4s, 8fs\}$  are invariants describing the deformation (corresponding to the matrix and fibre structure of the myocardium), and  $a, b, a_f, b_f, a_s, b_s, a_{fs}, b_{fs}$  are the eight constitutive parameters, which we collect in vector  $\boldsymbol{\psi}$ . The invariants in (1) are given as

$$\begin{aligned} I_1 &= \text{tr}(\mathbb{C}), & I_{4f} &= \mathbf{f}_0 \cdot (\mathbb{C} \mathbf{f}_0), \\ I_{4s} &= \mathbf{s}_0 \cdot (\mathbb{C} \mathbf{s}_0), & I_{8fs} &= \mathbf{f}_0 \cdot (\mathbb{C} \mathbf{s}_0), \end{aligned} \quad (2)$$

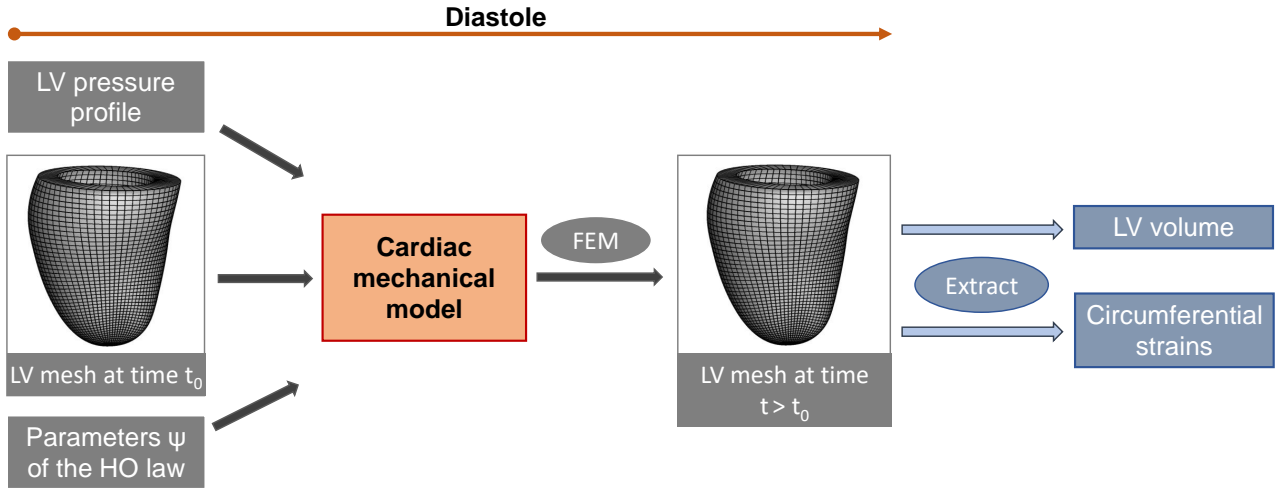
where  $\mathbb{C} = \mathbb{F}^T \mathbb{F}$  is the right Cauchy–Green deformation tensor, with  $\mathbb{F}$  being the deformation gradient, while  $\mathbf{f}_0$  and  $\mathbf{s}_0$  are the myocyte and sheet directions, respectively, determined via a rule based approach, detailed can be found in Wang et al.<sup>19</sup>. Note that  $\mathbf{f}_0$  and  $\mathbf{s}_0$  are initial values hence are known before running the forward simulator. The eight parameters in  $\boldsymbol{\psi}$  have the following interpretations:  $a$  and  $b$  describe the matrix response,  $a_f$  and  $b_f$  describe the contribution from the fibres along myocytes;  $a_s$  and  $b_s$  represent the contribution of the fibres along the sheet direction;  $a_{fs}$  and  $b_{fs}$  account for the shear effects in the fibre–sheet plane. Interested readers are referred to Holzapfel and Ogden<sup>12</sup> for a detailed explanation of the HO law.

The quasi-static cardio-mechanics model in diastole is given by<sup>23</sup>

$$\left. \begin{aligned} \nabla \cdot \boldsymbol{\sigma} + \mathbf{b} &= 0, & \text{in } \Omega, \\ \boldsymbol{\sigma} \cdot \mathbf{n} &= \mathbf{t}, & \text{in } \Gamma^N, \\ \mathbf{u} &= \bar{\mathbf{u}}, & \text{in } \Gamma^D, \end{aligned} \right\} \quad (3)$$

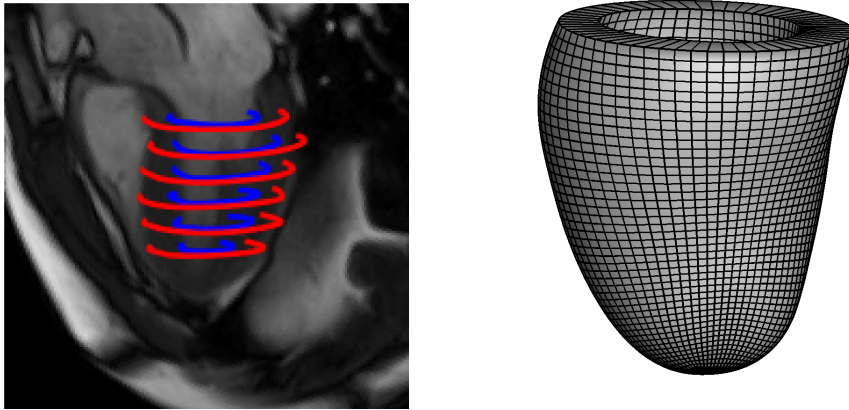
where  $\boldsymbol{\sigma}$  is the Cauchy stress, which can be derived from the HO law (1) with respect to the deformation state of myocardium,  $\mathbf{b}$  is the body force density per unit volume,  $\mathbf{t}$  is the traction force, and  $\bar{\mathbf{u}}$  is the vector of displacements at boundaries.  $\Omega$  is the computational domain defined by the LV mesh with the boundary denoted as  $\partial\Omega$ ,  $\mathbf{n}$  is the normal direction of  $\partial\Omega$ ,  $\Gamma^N$  and  $\Gamma^D$  are the Neumann and Dirichlet boundaries, respectively, with  $\Gamma^N \cup \Gamma^D = \partial\Omega$  and  $\Gamma^N \cap \Gamma^D = \emptyset$ . Solving (3) gives us the LV dynamics in response to the applied LV pressure, from which relevant quantities of interest can be readily extracted, such as the end-diastolic LV volume and circumferential strains. Figure 1 illustrates the essential components of the cardio-mechanic model in diastole.

The boundary value problem (3) does not admit a closed-form solution, hence numerical solutions are necessary in this context. The standard approach to solving biomechanical models is via the finite element method (FEM)<sup>5</sup>. In this work we use



**FIGURE 1** Inference pipeline in cardiac mechanic models.

the nonlinear finite element software Abaqus FEA (Simulia, Providence, RI, USA) to solve the discretised cardio-mechanics model. For a given parameter vector  $\psi$ , the LV mesh, i.e. the discretised LV geometry at early diastole, and the desired end diastolic LV pressure, the Abaqus forward simulator will determine the deformed mesh at the end of diastole by balancing the internal and external mechanical energies under prescribed boundary conditions, see Figure 1. The deformed mesh is then used to extract segmental circumferential strains at corresponding segments defined in four short-axial CMR cine images following clinical convention as specified by the American Heart Association<sup>24</sup>, as well as to compute the cavity volume at the end of diastole. Figure 2 illustrates an example of the reconstructed LV geometry derived from in vivo cine images with manually defined wall boundaries. Details of the LV geometry reconstruction can be found in Gao et al.<sup>4</sup>, including the definition of circumferential segmental strains.



**FIGURE 2** LV mesh reconstruction from a CMR image for a healthy volunteer. Left: segmented ventricular boundaries (blue: endocardium, red: epicardium) superimposed on a long axis CMR; right: the reconstructed LV geometry discretised with hexahedron elements.

The parameter vector  $\boldsymbol{\psi}$  describes the myocardial material properties and hence determines its mechanical responses through the HO law. Table 1 presents two sets of reference values of  $\boldsymbol{\psi}$  used in the literature. However, a subject-specific value of  $\boldsymbol{\psi}$  can neither be directly observed nor measured in vivo. Thus, we are interested in inferring it using routinely-available non-invasive in vivo data obtained from CMR images. Figure 2, left, presents an instance of such a scan together with manually annotated LV boundaries for the epicardium (red) and endocardium (blue). To the best of our knowledge, Gao et al.<sup>13</sup> were the first to address the issue of parameter inference in the HO law from in vivo clinical measurements (CMR scans). We discuss their algorithm, together with its updated version<sup>25</sup>, in Section 2.1. Still, parameter estimation for the HO law using in vivo data remains a challenging problem due to the potential non-identifiability of the constitutive parameters and their strong correlation<sup>13,17</sup>, as well as the high cost of solving the cardio-mechanics model numerically.

Work	$a$ [kPa]	$b$	$a_f$ [kPa]	$b_f$	$a_s$ [kPa]	$b_s$	$a_{fs}$ [kPa]	$b_{fs}$
Wang et al. <sup>19</sup>	0.236	10.810	20.037	14.154	3.724	5.164	0.411	11.300
Gao et al. <sup>4</sup>	0.18	2.6	3.34	2.73	0.69	1.11	0.31	2.58

**TABLE 1** Reference parameters  $\boldsymbol{\psi}_0$  proposed in the literature. Wang et al.<sup>19</sup> estimated their parameters based on ex vivo data, while Gao et al.<sup>4</sup> used in vivo data.

## 2.1 | Existing estimation methods

Gao et al.<sup>13</sup> proposed a three step algorithm based on matching the responses from the forward simulator associated with the HO law with the in vivo measurements obtained from CMR scans (circumferential segmental strains and the end-diastolic cavity volume). We refer to their approach as the HGO algorithm. The key idea behind the HGO algorithm is not estimating  $\boldsymbol{\psi}$  directly but via rescaling a reference value  $\boldsymbol{\psi}_0$ , taken from an earlier study<sup>19</sup>. Moreover, each step rescales a different subset of HO parameters.

---

### Algorithm 1: Original algorithm of Gao et al.<sup>13</sup>.

---

- 1 Optimise  $C_a, C_b$  with  $f_{O1}$  on a grid  $[0.1, 0.2, \dots, 1.0]^2$ .
  - 2 Optimise  $a_f, b_f$  with  $f_{O2}$ , s.t.  $a_f \geq 2a_s$  and  $b_f \geq 2b_s$ .
  - 3 Optimise  $C_3$  with  $f_{O1}$ .
- 

---

### Algorithm 2: Updated algorithm of Gao et al.<sup>25</sup>.

---

- 1 Optimise  $C_a, C_b$  with  $f_{O1}$  on a grid  $[0.1, 0.2, \dots, 1.0]^2$ .
  - 2 Refine  $C_a, C_b$  with  $f_{O2, \text{Klotz}}$ , s.t.  $a \geq 0.1, a_f \geq 0.1, a_s \geq 0.1$  and  $a_{fs} \geq 0.1$ .
  - 3 Optimise  $a_f, b_f$  with  $f_{O2}$ , s.t.  $a_f \geq 2a_s$  and  $b_f \geq 2b_s$ .
  - 4 Optimise  $a, b$  with  $f_{O2}$ , s.t.  $a \geq 0.1$ .
- 

Listing 1 presents steps of the HGO algorithm, where the scaling scalars  $C_a, C_b$  and  $C_3$  are defined in the following way. Scalars  $C_a$  and  $C_b$  are used to match the parameters in two parameter groups,  $a^{\text{group}} = \{a, a_f, a_s, a_{fs}\}$  and  $b^{\text{group}} = \{b, b_f, b_s, b_{fs}\}$ ,

respectively, as

$$\mathbf{a}^{\text{group}} = C_a \mathbf{a}_0^{\text{group}}, \quad \mathbf{b}^{\text{group}} = C_b \mathbf{b}_0^{\text{group}},$$

while scalar  $C_3$  matches  $a$  and  $a_{\text{fs}}$

$$a = C_3 a^{(2)}, \quad a_{\text{fs}} = C_3 a_{\text{fs}}^{(2)},$$

where  $a^{(2)}$  and  $a_{\text{fs}}^{(2)}$  are values of  $a$  and  $a_{\text{fs}}$ , respectively, optimised in the second step of the algorithm. The two objective functions matching the simulated values, depending on the constitutive parameter  $\boldsymbol{\psi}$  and geometry  $\mathcal{H}$ , to the measurement are defined as follows

$$f_{O1}(\boldsymbol{\psi}, \mathcal{H}) = (V(\boldsymbol{\psi}, \mathcal{H}) - V^*)^2 + \sum_{i=1}^K (\varepsilon_i(\boldsymbol{\psi}, \mathcal{H}) - \varepsilon_i^*)^2, \quad (4)$$

$$f_{O2}(\boldsymbol{\psi}, \mathcal{H}) = \frac{(V(\boldsymbol{\psi}, \mathcal{H}) - V^*)^2}{V^*} + \sum_{i=1}^K (\varepsilon_i(\boldsymbol{\psi}, \mathcal{H}) - \varepsilon_i^*)^2, \quad (5)$$

where  $V^*$  and  $\varepsilon_i^*$ ,  $i = 1, \dots, K = 24$  are the measurements of the LV volume and 24 circumferential strains, respectively, while  $V(\boldsymbol{\psi}, \mathcal{H})$  and  $\varepsilon_i(\boldsymbol{\psi}, \mathcal{H})$  are the corresponding values obtained from the forward simulator using the parameter vector  $\boldsymbol{\psi}$  and the LV geometry  $\mathcal{H}$ .

### Updated algorithm

Gao et. al<sup>25</sup> have developed an updated version of the original HGO algorithm<sup>13</sup>, which we present in Algorithm 2. We can see that the main difference between the two versions is that the new one includes an additional step based on refining  $C_a$  and  $C_b$  with  $f_{O2, \text{Klotz}}$ , which we discuss in Section 2.2. In short,  $f_{O2, \text{Klotz}}$  augments  $f_{O2}$  with an additional term which, for the given parameters  $\boldsymbol{\psi}$  and LV geometry  $\mathcal{H}$ , quantifies the discrepancy between the LV volume predicted for pressure of 30 mmHg and – since no in vivo data is observed at such a high pressure – the prediction from the so-called Klotz-curve<sup>14</sup>, which we discuss in Section 2.2. This aims at inferring parameters which not only agree with the in vivo data observed at low pressures, but also give realistic prediction for high pressure scenarios. Another difference between Algorithms 1 and 2 can be observed in the last step: the updated version matches  $a$  and  $b$  while the original algorithm optimises  $a$  and  $a_{\text{fs}}$ . There are two reasons for the latter modification. First, the sensitivity of  $a_{\text{fs}}$  is much lower compared to  $a$  and  $b$ <sup>13</sup> and it is more related to the shearing response, which does not directly link to the measured circumferential strains<sup>12</sup>. Second,  $a$  and  $b$  are two highly sensitive parameters and directly related to the extracellular matrix response. Thus, in the final step  $a$  and  $b$  are updated to match the measured data.

## Parameter reduction

The eight parameters in  $\boldsymbol{\psi}$  are highly correlated and not identifiable from limited data based on CMR scans<sup>13</sup>. Therefore, we follow Davies et al.<sup>17</sup> and use a parametrisation reduced to a 4-dimensional manifold

$$\left. \begin{aligned} a &= \theta_1 a_0, & b &= \theta_1 b_0, \\ a_f &= \theta_2 a_{f0}, & a_s &= \theta_2 a_{s0}, \\ b_f &= \theta_3 b_{f0}, & b_s &= \theta_3 b_{s0}, \\ a_{fs} &= \theta_4 a_{fs0}, & b_{fs} &= \theta_4 b_{fs0}, \end{aligned} \right\} \quad (6)$$

where  $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4)^T$  is the vector of scalings of the reference parameters  $\boldsymbol{\psi}_0 = (a, b, a_f, b_f, a_s, b_s, a_{fs}, b_{fs})^T$ . Wang et al.<sup>19</sup> proposed a popular reference parameter vector, used in several later studies<sup>13,17</sup>. However, those values were based on ex vivo data, making them potentially of limited relevance to our problem (inference based on in vivo data). Hence, we use different values for  $\boldsymbol{\psi}_0$  based on in vivo data population-wide averages<sup>4</sup>. The only exception is the implementation of the HGO algorithm<sup>13</sup>, which originally uses the earlier reference values<sup>19</sup>. We have implemented this algorithm using both reference parameter vectors<sup>19,4</sup>.

## 2.2 | Klotz-curve

*In vivo* data are collected at physiologically normal (low) ventricular pressures. These pressures, however, are not recorded in healthy volunteers due to the invasive nature of the procedure. We assume a population-based end-diastolic (ED) pressure of 8 mmHg<sup>13</sup>. Working with “low pressure data” implies we can only estimate those parameters of the HO law that describe myocardium properties at such low pressures. However, parameters leading to a similar low pressure behaviour (in terms of the implied stress-strain relationships) may result in a very different behaviour at high pressures, such as 30 mmHg. These high pressures are of interest as they may reveal LV stiffness with impaired relaxation, which characterises diastolic heart failure<sup>18</sup>. Hence, we want to avoid that parameters which have been inferred from in vivo pressures exhibit physiologically unrealistic behaviour at high pressures.

However, high pressure in vivo volume measurements are not available, therefore we propose to predict them using the empirical law found by Klotz et al.<sup>14</sup>, which we will refer to as the Klotz-curve. The normalised ED volume  $\tilde{V}^*$  is defined as<sup>14</sup>

$$\tilde{V}^* = \frac{V^* - V_0}{V_{30} - V_0}, \quad (7)$$

where  $V^*$  is the measured unnormalised volume at  $P^*$ ,  $V_0$  is the zero-pressure volume called the load-free volume. An empirical law based on population-wide ex vivo data is then established<sup>14</sup> relating the normalised LV volume to the corresponding



measurement pressure  $P^*$  with the specific form of

$$P^* = A (\tilde{V}^*)^B, \quad (8)$$

where  $A$  and  $B$  are parameters. In our case  $V^*$  is the in vivo measured volume at pressure  $P^* = 8$  mmHg. Combining (8) and (7) results in the following formula for the high-pressure ED volume

$$V_{30}^{KI} = V_0 + \frac{V^* - V_0}{\tilde{V}^*} = V_0 + \frac{V^* - V_0}{\left(\frac{P^*}{A}\right)^{1/B}}, \quad (9)$$

where  $V_{30}^{KI}$  denotes  $V_{30}$  predicted using the Klotz-curve.

Note that formula (9) depends on the load-free volume  $V_0$ , which can only be measured ex vivo. With only in vivo measurements available, we need to make some assumption regarding how to estimate it. We use the early diastolic volume as a proxy for the load-free volume, since it is known that the LV pressure is the lowest at early diastole<sup>13,26</sup>. Thus, we predict  $V_{30}^{KI}$  using the approximated  $\hat{V}_0$  and the in vivo measured volume  $V^*$  at  $P^* = 8$  mmHg as

$$\hat{V}_{30}^{KI} = \hat{V}_0 + \frac{V^* - \hat{V}_0}{(P^*/\hat{A})^{1/\hat{B}}},$$

where  $\hat{A} = 27.78$  and  $\hat{B} = 2.76$  are the estimates<sup>14</sup>.

We use the Klotz-curve predicted high-pressure volume  $\hat{V}_{30}^{KI}$  to add a high-pressure penalty term to the objective function (5) as follows

$$f_{O2,Klotz}(\boldsymbol{\psi}, \mathcal{H}) = \left( \frac{V_8(\boldsymbol{\psi}, \mathcal{H}) - V_8^*}{V_8^*} \right)^2 + \sum_{i=1}^K (\varepsilon_i(\boldsymbol{\psi}, \mathcal{H}) - \varepsilon_i^*)^2 + \left( \frac{V_{30}(\boldsymbol{\psi}, \mathcal{H}) - \hat{V}_{30}^{KI}}{\hat{V}_{30}^{KI}} \right)^2, \quad (10)$$

where  $V_8$  and  $V_{30}$  are LV cavity volumes from the forward simulator, at 8 mmHg and 30 mmHg, respectively;  $\varepsilon_i$ ,  $i = 1, \dots, K = 24$ , are regional circumferential strains from the forward simulator at 8 mmHg;  $V_8^*$  and  $\varepsilon_i^*$  are the in vivo measured LV cavity volume (at the normal pressure of 8 mmHg) and regional circumferential strains, respectively.

### 3 | BAYESIAN OPTIMISATION

Suppose we are interested in finding the global optimum of a smooth function  $f(\mathbf{x}) : \mathbb{R}^d \supset \mathcal{X} \rightarrow \mathbb{R}$ . We will refer to  $f$  as the objective or target function and without loss of generality assume it is a loss function to be minimised. Further, suppose computing  $f$  at any given  $\mathbf{x}$  is expensive and/or time-consuming so that we want to find the optimiser  $\mathbf{x}^* = \arg \min_{\mathbf{x}} f(\mathbf{x})$  in as few steps as possible. BO proceeds in an iterative fashion, at each iteration  $t$  selecting an input  $\mathbf{x}_t$  and querying the black-box function  $y_t = f(\mathbf{x}_t)$ . Input selection is done by optimising the so-called acquisition function (AF), which is based on an

approximation of  $f$  with a surrogate and models the desired trade-off between *exploration* (searching for new promising regions) and *exploitation* (increasing the resolution around beneficial regions already discovered).

### 3.1 | Gaussian process surrogates

Typically, the AF is based on a *surrogate model* of the objective function  $f$ , with a popular choice for such a surrogate being Gaussian processes (GPs). As opposed to alternative nonlinear methods, such as e.g. splines and neural networks, GPs allow exact interpolation rather than regression, which is what is conceptually required for deterministic objective functions, such as the one based on a deterministic cardiomechanic model. However, due to technical reasons discussed below, restating the interpolation problem as a regression problem is more practical, so that below we focus on the GP regression setting.

In BO GP regression is trained on a sequentially expanding set of inputs  $\mathbf{X}_t = [\mathbf{x}_j]_{j=1}^t$  and outputs  $\mathbf{Y}_t = [y_j]_{j=1}^t$ . We refer to Rasmussen and Williams<sup>8</sup> for an extensive treatment of GPs. GP regression is a nonparametric model in which a GP prior is put over the unknown function of interest to represent our initial beliefs about its smoothness. The target values at points  $\mathbf{x}_j$ ,  $j = 1, \dots, t$  observed up to iteration  $t$  are given by

$$y_j = f(\mathbf{x}_j) + \varepsilon_j, \quad \varepsilon_j \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2), \quad (11)$$

where  $\mathcal{N}(\mu, \sigma^2)$  denotes the Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$  (and *iid* stands for independently and identically distributed). Specification (11) is a general one, allowing for noisy observations. Our forward simulator of interest is deterministic, so mathematically there is no noise ( $\sigma^2 = 0$ ). However, including a non-trivial noise term  $\varepsilon_j$  is still convenient for at least two reasons. First, in practice there may be numerical noise, due to the finite precision of the numerical solution of the forward problem. Second, when working with GPs it is a common practice to allow for a small *jitter* (or nugget) term (e.g.  $10^{-9}$ ), which is added to the diagonal of the training covariance matrix for numerical stability of the matrix inversion. Effectively, jitter is imposed noise, which slightly dilutes the informativeness of our deterministic data.

In (11) the latent values  $f(\mathbf{x})$  follow a GP

$$f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')),$$

where

$$\mu(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})], \quad k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - \mu(\mathbf{x}))(f(\mathbf{x}') - \mu(\mathbf{x}'))]$$

are the prior mean and the covariance function (kernel) of the process  $f$ , respectively. Hence, [] it is assumed that the responses  $y_j$  are conditionally independent given the latent values  $f(\mathbf{x}_j)$ . Below we adopt a standard assumption that  $\mu(\mathbf{x}) = 0$  so the latent process  $f(\mathbf{x})$  is fully specified by its kernel function.

The kernel can be specified in many different ways, including numerous standard functional forms<sup>8</sup>. We consider the Matérn 3/2 kernel with automatic relevance determination (ARD) given as

$$k_0(\mathbf{x}, \mathbf{x}') = \sigma_k^2 \left(1 + \sqrt{3}r\right) \exp\left(-\sqrt{3}r\right), \quad r = \left(\sum_{i=1}^d \frac{(x_i - x'_i)^2}{l_i^2}\right)^{1/2},$$

where  $l_i$ ,  $i = 1, \dots, d$  is the length scale parameter for the  $i$ th input variable and ARD refers to this allowing for a different length scale for each dimension. Thus our GP regression model in (11) is parametrised by a vector of hyperparameters  $\boldsymbol{\phi} = (\sigma^2, \sigma_k^2, l_1, \dots, l_d)^T$ , so that we consider the observation noise variance  $\sigma^2$  a hyperparameter<sup>8</sup>. The Matérn 3/2 kernel produces continuous trajectories with continuous derivatives at the same time providing a considerable degree of roughness (it is once differentiable) allowing us to model functions with cliffs and plateaus. Although this choice of kernel rules out standard second-order optimisation methods, like quasi-Newton, we note that second-order methods are of limited use in the presence of extensive multimodality (which is a typical problem when optimising AFs, see Section 3.2.1). Moreover, based on our experimentation this low-level Matérn class lead to numerically more stable covariance matrix inversion than a higher-level Matérn class.

For the collected inputs  $\mathbf{X}_t$  we obtain the GP prior over function values  $p(\mathbf{f}_t | \mathbf{X}_t, \boldsymbol{\phi}) = \mathcal{N}(\mathbf{f}_t | \mathbf{0}, \mathbf{K}_t)$ , with  $\mathbf{f}_t = [f(x_j)]_{j=1}^l$  and  $\mathbf{K}_t = k(\mathbf{X}_t, \mathbf{X}_t)$ . The likelihood is given by  $p(\mathbf{Y}_t | \mathbf{f}) = \mathcal{N}(\mathbf{Y}_t | \mathbf{f}, \sigma^2 \mathbb{I})$ , where  $\mathbb{I}$  is the identity matrix, and marginalising over the latent variables  $\mathbf{f}_t$  gives the formula for the marginal likelihood

$$p(\mathbf{Y}_t | \mathbf{X}_t, \boldsymbol{\phi}) = \mathcal{N}(\mathbf{Y}_t | \mathbf{0}, \mathbf{K}_t + \sigma^2 \mathbb{I}).$$

Under the adopted Gaussian observation model (11) the conditional posterior distribution of the latent variables also becomes Gaussian, with the following form

$$p(\mathbf{f}_t | \mathbf{Y}_t, \mathbf{X}_t, \boldsymbol{\phi}) = \mathcal{N}(\mathbf{K}_t(\mathbf{K}_t + \sigma^2 \mathbb{I})^{-1} \mathbf{Y}_t, \mathbf{K}_t - \mathbf{K}_t(\mathbf{K}_t + \sigma^2 \mathbb{I})^{-1} \mathbf{K}_t).$$

### 3.2 | Acquisition function

The performance of BO algorithms depends on the particular choice of AF. Various AFs have been proposed in the literature<sup>6</sup>, but the Expected Improvement<sup>10</sup> criterion remains one of the most popular choices. EI has gained popularity as it is conceptually simple (as opposed to e.g. information theoretic AFs), it admits a closed form expression for GP surrogates and performs well for a range of problems<sup>11</sup>. As its name suggests, EI quantifies the expected improvement, i.e. the amount of improvement over the current *incumbent*  $f^*$ , i.e. the best (lowest) value of  $f$  found so far, weighted by the probability of it occurring. EI is defined as

$$\text{EI}(\mathbf{x}) = \mathbb{E}_{p(y|\mathbf{x}, D)}[\min(f^* - f(\mathbf{x}), 0)],$$

where  $\mathcal{D}$  denotes the set of inputs and outputs recorded so far and  $p(y|\mathbf{x}, \mathcal{D})$  is the predictive posterior output distribution given the chosen evaluation point  $\mathbf{x}$  and the previously observed  $\mathcal{D}$ . With a GP surrogate  $\mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}))$  EI can be expressed<sup>10,6</sup> as

$$\text{EI}(\mathbf{x}) = (f^* - \mu(\mathbf{x}))\Phi(z) + \sqrt{k(\mathbf{x})}\phi(z),$$

where  $z = (f^* - \mu(\mathbf{x}))/\sqrt{k(\mathbf{x})}$  and  $\Phi$  and  $\phi$  are the CDF and PDF of the standard normal distribution, respectively. The EI is made up of two terms. The first term is increased by decreasing the predictive mean  $\mu(\mathbf{x})$ , the second term is increased by increasing the predictive uncertainty  $k(\mathbf{x})$ . This shows how EI automatically balances exploitation and exploration.

### 3.2.1 | Inner optimisation

BO selects the next query point based on maximising the AF, which is often referred to as *inner optimisation* as it is within the *outer optimisation* (optimising the unknown objective function  $f$ ) of the BO algorithm. Even though inner optimisation does not require any computationally expensive forward simulations, it can be a challenging task<sup>7</sup>. This is often due to the observed multimodality of EI, which typically is picked around several points and flat between them. A popular approach to overcome this problem is the multi-starts method<sup>27</sup>, which performs standard numerical optimisation using e.g. quasi-Newton methods such as BFGS multiple times, from different initial points. However, a pilot study we conducted revealed that the multi-starts method performs poorly in our application, irrespective of a particular kernel choice (we tested several standard kernels, such as squared exponential, Matérn 5/2, Matérn 3/2 and neural network), often indicating suboptimal points and thus leading to unstable BO runs. That was the case even with an atypically large number of restarts (set to 1000 and drawn from a Latin hypercube design).

For these reasons we use a different approach based on a global optimisation algorithm called OQNLP<sup>28</sup> (or Global Search in its implementation in MATLAB's Global Optimization toolbox that we use), which led to very stable results. In short, the OQNLP is a heuristic algorithm also based on multiple starting points. Those starting points, however, are not randomly (or quasi-randomly) selected, but rather calculated from the scatter search algorithm<sup>29</sup>. They are then scored based on the corresponding value of the objective function and, if present, constraint satisfaction. Next, gradient-based nonlinear programming (NLP) solvers are run from the starting point with the highest score value. Thereafter, the remaining points are analysed and those that are unlikely to improve the best local minimum found so far are rejected (which is different to the standard multistart approach in which solvers are run for all the initial points). The filtering analysis takes into account the corresponding value of the objective function, constraint violations and basins of attraction. The latter means that the points in the same basins of attraction as previously analysed points are excluded from further NLP optimisation. This feature, i.e. analysis of basis of attraction, contributes to the efficiency of the OQNLP algorithm as it allows for discarding points which would converge to the already found local optima.

### 3.2.2 | Unknown constraints

It may happen during BO iterations that the proposed point, in spite of maximizing the AF based on the current surrogate, will violate the biomechanical assumptions behind the HO law. Such a situation will demonstrate itself as a crash of the associated forward simulator, which is inefficient as some computations will be carried out in vain. The problem of a simulator failing to terminate or crashing was addressed in the literature<sup>15,16</sup> by weighting EI by the probability of the constraint being satisfied

$$EI_{con}(\mathbf{x}) = EI(\mathbf{x})\mathbb{P}(\mathbf{x} \in C),$$

where  $\mathbb{P}(\mathbf{x} \in C)$  is the probability of  $\mathbf{x}$  being a valid point not leading to a crash of the forward simulator.

To model the probability of the unknown constraint being satisfied we use a GP classifier with a probit likelihood<sup>15</sup> with approximate inference carried out using expectation propagation<sup>30</sup>. Similarly as for the regression model, we use the Matérn 3/2 kernel. A probit likelihood is a conceptually sound approach, however it means that we lose the analytic tractability of the standard EI function as for non-Gaussian likelihoods approximate inference methods (such as Laplace approximation<sup>31</sup>, variational inference<sup>32</sup>, expectation propagation<sup>30</sup>) are necessary<sup>8</sup>. A simpler though heuristic classifier could be provided by the so-called label regression<sup>33</sup> (LR). LR is an ordinary linear regression model fitted to explain a binary outcome variable and its advantage is that  $\mathbb{P}(\mathbf{x} \in C)$  can be computed explicitly, yet based on a suboptimal probability model (a Gaussian, rather than a logit or probit model). LR has been successfully employed to account for hidden constraints in the context of BO applied to pulmonary circulation<sup>34</sup>. However, in our preliminary study we found that in our application LR led to unstable results, generally performing poorly.

We note that the probability of a crash happening during a BO run is LV geometry and material parameter dependent. For the investigated LV geometries crashes are not very frequent, see Section 4.1-4.2, however, we note that the reported numbers of crashes, if nonzero, are already based on training a classifier. In fact, our pilot study revealed that not allowing for any learning from past unsuccessful simulations can lead to repeatedly querying very similar “crash points”, resulting in much higher numbers of crashes.

### 3.2.3 | Partial error surrogates

Notice that both target functions,  $f_{O_2}$  and  $f_{O_2, Klotz}$ , are sums of individual squared error terms. Thus, instead of developing a surrogate for the target as a whole we propose to construct a surrogate for each error term separately. In the context of emulation<sup>17</sup> a similar approach led to a substantial improvement in accuracy by emulating the vector of outputs from the cardiac mechanic model. That approach is efficient in the context of emulation when the developed emulator is aimed to be used for a wide range of different subjects. However, in our context using surrogate functions for the outputs would require us to *approximate* EI, as implied by the predictive means from individual output surrogates. Considering surrogates for the individual error terms instead

allows us to obtain a closed form expression for EI. Note that considering partial error surrogates instead of output surrogates does not cause an efficiency loss as in our case optimisation is carried out for subject-specific measurements, known before running computations (which is different than for emulators).

We refer to the proposed approach as *partial error surrogates*.

Let

$$f(\boldsymbol{\psi}, \mathcal{H}) = \sum_{i=1}^{K^*} f^{(i)}(\boldsymbol{\psi}, \mathcal{H}),$$

where  $f$  is either  $f_{O2}$  in (5) so that  $K^* = K + 1 = 25$ , or  $f_{O2, \text{Klotz}}$  in (10) so that  $K^* = K + 2 = 26$ , and each  $f^{(i)}$  refers to the squared error terms on the right-hand side in those formulae. We assume that  $f^{(i)}$ s are mutually independent and model each  $f^{(i)}$  using a GP regression parametrised by hyperparameters  $\boldsymbol{\phi}^{(i)}$ . Allowing for correlated  $f^{(i)}$ s would require us to use computationally much more expensive multi-output GPs<sup>35,36,37</sup>, which would increase the computational overhead of each BO iteration, diminishing the posited “cheapness” of the surrogate compared to the forward simulator. Each partial GP regression has the conditional posterior mean  $\mu_t^{(i)}(\mathbf{x})$  and covariance  $k_t^{(i)}(\mathbf{x})$

$$\mathbf{f}_t^{(i)} | \mathbf{Y}_t^{(i)}, \mathbf{X}_t, \boldsymbol{\phi}^{(i)} \sim \mathcal{GP}(\mu_t^{(i)}(\mathbf{x}), k_t^{(i)}(\mathbf{x})),$$

where  $\mathbf{f}_t^{(i)} = [f^{(i)}(\mathbf{x}_j)]_{j=1}^t$  are the latent variables for the  $i$ th output from the cardiomechanic model and  $\mathbf{Y}_t^{(i)} = [y_j^{(i)}]_{j=1}^t$  are the observed values for the  $i$ th output. Then the conditional posterior mean  $\mu_t$  of the full target  $f$  is just the sum of conditional posterior means  $\mu_t^{(i)}$  from the partial surrogates and the conditional posterior variance  $k_t$  is the sum of the partial variances  $k_t^{(i)}$ :

$$\mathbf{f}_t | \mathbf{Y}_t^{(1)}, \dots, \mathbf{Y}_t^{(K^*)}, \mathbf{X}_t, \boldsymbol{\phi}^{(1)}, \dots, \boldsymbol{\phi}^{(K^*)} \sim \mathcal{GP}(\mu_t(\mathbf{x}), k_t(\mathbf{x})),$$

with

$$\mu_t(\mathbf{x}) = \sum_{i=1}^{K^*} \mu_t^{(i)} \quad \text{and} \quad k_t(\mathbf{x}) = \sum_{i=1}^{K^*} k_t^{(i)}. \quad (12)$$

## Standardisation

Standardising inputs and outputs is a common practice when working with GPs, aimed to provide numerical stability of the computations. In our case input standardisation is not an issue as all the partial surrogates use the same inputs. Predictions for the outputs, however, enter the formulae for the conditional posterior distribution (12) so we need to “de-standardise” the predictions to obtain correct partial moments. Let  $\tilde{\mathbf{Y}}_t^{(i)}$  denote the  $i$ th,  $i = 1, \dots, K^*$ , standardised partial squared error

$$\tilde{\mathbf{Y}}_t^{(i)} = \frac{\mathbf{Y}_t^{(i)} - m_0^{(i)}}{\sigma_0^{(i)}},$$

where  $m_0^{(i)}$  and  $\sigma_0^{(i)}$  are the mean and standard deviation, respectively, used to standardise the  $i$ th partial squared errors (e.g. calculated on the initial design). The reverse transformation is given as

$$\mathbf{Y}_t^{(i)} = g(\tilde{\mathbf{Y}}_t^{(i)}) = m_0^{(i)} + \sigma_0^{(i)} \tilde{\mathbf{Y}}_t^{(i)},$$

so if the  $i$ th GP surrogate predicts  $\mathcal{N}(\tilde{\mu}_t^{(i)}, \tilde{k}_t^{(i)})$  for the standardised  $i$ th partial squared error, this corresponds to predicting  $\mathcal{N}(g(\tilde{\mu}_t^{(i)}), (\sigma_0^{(i)})^2 \tilde{k}_t^{(i)})$  for the original  $i$ th partial squared error. Then the moments (12) used in the acquisition function are given by

$$\mu_t(\mathbf{x}) = \sum_{i=1}^{K^*} (m_0^{(i)} + \sigma_0^{(i)} \tilde{\mu}_t^{(i)}) \quad \text{and} \quad k_t(\mathbf{x}) = \sum_{i=1}^{K^*} (\sigma_0^{(i)})^2 \tilde{k}_t^{(i)}.$$

### 3.3 | Emulator for high-pressure volume

Computing  $f_{O2, \text{Klotz}}$  in (10) requires running the forward simulator twice, at 8 mmHg and 30 mmHg, which requires double computing time/power. Each of these runs may be unsuccessful, i.e. may lead to a crash. However, parameters leading to crashes may differ between the 8 mmHg and 30 mmHg cases. Hence, we would need two classifiers, one for the 8 mmHg case and one for the 30 mmHg case, and we could only update the GP surrogate for  $f_{O2, \text{Klotz}}$  when both cases terminate successfully.

To overcome these problems and to reduce the computational costs we propose to obtain the value for  $V_{30}$  using an emulator rather than by running the forward simulator. Emulators of cardiac mechanics models often include constitutive parameters as inputs<sup>17</sup>. However, the outputs from the forward simulator depend not only on the constitutive parameters but also on the LV geometry. Since we want the  $V_{30}$  emulator to approximate  $V_{30}$  for a broad range of LV geometries (as opposed to developing a geometry-specific emulator<sup>17</sup>) we want to include some information about the LV geometry as input. However, we cannot use the LV mesh directly as it is a high-dimensional (with thousands of coordinates) object and such a dimension is prohibitively high for any emulator. Our preliminary experimentation revealed that accurately predicting only the LV volume at 30 mmHg, and not the strains, does not require much details about a subject's specific shape of the LV and can be done based on a single "summary statistic" of the LV mesh – an approximation  $\hat{V}_0$  to the LV cavity load-free volume  $V_0$  (see Section 2.2) – without a substantial loss in the accuracy of the emulator. We note that if we were also after emulating circumferential strains at 30 mmHg, accounting for more subtle features of the LV geometry would be necessary. To sum up, we train the  $V_{30}$  emulator on the scaling parameters  $\theta$  and the approximate load-free volume  $\hat{V}_0$ , in which we approximate the true function  $V_{30}(\theta, \mathcal{H})$  by  $V_{30}(\theta, V_0)$ .

### 3.4 | Final algorithm

Listings 3 and 4 summarise the initialisation and iterations, respectively, of the BO algorithm with the constraint-weighted EI, Klotz-curve objective (10), partial error surrogates and  $V_{30}$  emulator.

---

**Algorithm 3:** Initialisation of Bayesian optimisation with partial error surrogates.

---

**Input** :  $d$  – input dimension (here:  $d = 4$ );

$K^*$  – number of outputs to be emulated ( $K^* = 26$ );

$P^*$  – measurement (low) pressure (8 mmHg);

$V_8^*$  and  $\varepsilon_i^*$ ,  $i = 1, \dots, K$  – volume and strain measurements at  $P^*$ ;

$\hat{V}_0$  – load free volume estimate.

**Output:**  $\hat{V}_{30}^{KI}$  – Klotz-curve predicted volume at 30 mmHg;

$\mathbf{X}(\tilde{\mathbf{X}})$  – explanatory variables (standardised) for regressions;

$\mathbf{Y}^{(i)}(\tilde{\mathbf{Y}}^{(i)})$ ,  $i = 1, \dots, K^*$  – response var. (standardised) for regressions;

$\phi^{(i)}$ ,  $i = 1, \dots, K^*$  – regression hyperparameters on the initial design;

$m_x, \sigma_x$  – inputs mean and st. dev. on the initial design;

$m_0^{(i)}, \sigma_0^{(i)}$ ,  $i = 1, \dots, K^*$  – outputs mean and st. dev. on the initial design;

$\mathbf{Z}$  (optional) – explanatory variables for classification on the initial design;

$\mathbf{U}$  (optional) – response variables for classification on the initial design;

$\phi^c$  (optional) – classification hyperparameters on the initial design.

1 Calculate  $\tilde{V}_8^* = (P^*/\hat{A})^{1/\hat{B}}$  (normalised volume at  $P^*$  from (8)).

2 Calculate  $\hat{V}_{30}^{KI} = \hat{V}_0 + (V_8^* - \hat{V}_0)/\tilde{V}_8^*$ .

3 Generate  $t_0 = 10d$  points  $\mathbf{X} = [\mathbf{x}_j]_{j=1}^{t_0}$  from an initial design scheme (e.g. Latin hypercube).

4 Run the forward simulator on  $\mathbf{X}$  to obtain partial squared errors  $\mathbf{Y}^{(i)}$ ,  $i = 1, \dots, K^*$ .

5 **if** *crashes* for  $\mathbf{X}_B = [\mathbf{x}_j]_{j \in B}$ ,  $B \subset \{1, \dots, t_0\}$  **then**

6     Set  $\mathcal{G} = \{1, \dots, t_0\} \setminus B$  (indices of good points).

7     Set  $\mathbf{Z}_G = \mathbf{X}_G = [\mathbf{x}_j]_{j \in \mathcal{G}}$ ,  $\mathbf{Z}_B = \mathbf{X}_B$  and  $\mathbf{Z} = [\mathbf{Z}_G^T, \mathbf{Z}_B^T]^T$ .

8     Set  $\mathbf{U}_G = \mathbf{1}_{|\mathcal{G}| \times 1}$ ,  $\mathbf{U}_B = \mathbf{0}_{|B| \times 1}$  and  $\mathbf{U} = [\mathbf{U}_G^T, \mathbf{U}_B^T]^T$ .

9     Set  $\mathbf{X} := \mathbf{X}_G$  (only good points for regression).

10    Fit an initial GP probit classifier on  $\mathbf{Z}$  and  $\mathbf{U}$  (obtain  $\phi^c$ ).

11 **end**

12 Calculate  $m_x$  and  $\sigma_x$  of  $\mathbf{X}^{(i)}$ .

13 Standardise  $\mathbf{X}$  with  $m_x, \sigma_x$  to obtain  $\tilde{\mathbf{X}}$ .

14 **for**  $i = 1$  **to**  $K^*$  **do**

15     Calculate  $m_0^{(i)}$  and  $\sigma_0^{(i)}$  of  $\mathbf{Y}_{t_0}^{(i)}$ .

16     Standardise  $\mathbf{Y}^{(i)}$  with  $m_0^{(i)}, \sigma_0^{(i)}$  to obtain  $\tilde{\mathbf{Y}}^{(i)}$ .

17     Fit an initial GP regression model on  $\tilde{\mathbf{X}}$  and  $\tilde{\mathbf{Y}}^{(i)}$  (obtain  $\phi^{(i)}$ ).

18 **end**

---



---

**Algorithm 4:** Iterations of Bayesian optimisation with partial error surrogates.

---

**Input :**  $t_0$  – number of valid points on the initial design;  
 $K^*$  – number of outputs ( $K^* = 26$ );  
 $V_8^*$  and  $\varepsilon_i^*$ ,  $i = 1, \dots, K$  – volume and strain measurements;  
 $\hat{V}_{30}^{KI}$  – Klotz-curve predicted volume at 30 mmHg;  
 $\mathbf{X}(\tilde{\mathbf{X}})$  – explanatory variables (standardised) for regressions;  
 $\mathbf{Y}^{(i)}(\tilde{\mathbf{Y}}^{(i)})$ ,  $i = 1, \dots, K^*$  – response var. (standardised) for regressions;  
 $\phi^{(i)}$ ,  $i = 1, \dots, K^*$  – regression hyperparameters on the initial design;  
 $m_x, \sigma_x$  – inputs mean and st. dev. on the initial design;  
 $m_0^{(i)}, \sigma_0^{(i)}$ ,  $i = 1, \dots, K^*$  – outputs mean and st. dev. on the initial design;  
 $\mathbf{Z}$  (optional) – explanatory variables for classification on the initial design;  
 $\mathbf{U}$  (optional) – response variables for classification on the initial design;  
 $\phi^c$  (optional) – classification hyperparameters on the initial design.

**Output:**  $\mathbf{X}$  – valid inputs;  
 $\mathbf{Y}^{(i)}$ ,  $i = 1, \dots, K^*$  – partial squared errors;  
 $\mathbf{Z}$  (optional) – crash points.

```

1 for  $t = 1$  to  $T$  do
2   if exists  $\mathbf{Z}$  then
3     Optimise  $\text{EI}_{con}$  with GlobalSearch and set  $\tilde{\mathbf{x}}^* = \arg \min (-\text{EI}_{con})$ .
4   else
5     Optimise EI with GlobalSearch and set  $\tilde{\mathbf{x}}^* = \arg \min (-\text{EI})$ .
6   end
7   Re-standardise  $\tilde{\mathbf{x}}^*$  to obtain  $\mathbf{x}^*$ .
8   Run the forward simulator with  $\psi$  implied by  $\mathbf{x}^*$ .
9   if crash for  $\mathbf{x}^*$  then
10    Add  $\mathbf{x}^*$  as the last row of  $\mathbf{Z}_B$ .
11    Set  $\mathbf{Z}_G := \mathbf{X}$  and  $\mathbf{Z} = [\mathbf{Z}_G^T, \mathbf{Z}_B^T]^T$ .
12    Set  $\mathbf{U}_B := [\mathbf{U}_B^T, 1]^T$ .
13    Denote the number of rows of  $\mathbf{X}$  by  $g$  and set  $\mathbf{U}_G = \mathbf{1}_{g \times 1}$  and  $\mathbf{U} = [\mathbf{U}_G^T, \mathbf{U}_B^T]^T$ .
14    Update the classifier using  $\mathbf{Z}$  and  $\mathbf{U}$  (obtain  $\phi^c$ ).
15  else
16    Add  $\mathbf{x}_*$  as the last row of  $\mathbf{X}$ .
17    Standardise  $\mathbf{x}_*$  with  $m_x, \sigma_x$  to obtain  $\tilde{\mathbf{x}}_*$ .
18    Add  $\tilde{\mathbf{x}}_*$  as the last row of  $\tilde{\mathbf{X}}$ .
19    Use the emulator to predict  $V_{30}$ .
20    for  $i = 1$  to  $K^*$  do
21      Calculate  $y^{(i)}$  using the  $i$ th output and the corresponding measurement/the Klotz-curve prediction  $\hat{V}_{30}^{KI}$ .
22      Add  $y^{(i)}$  as the last row of  $\mathbf{Y}^{(i)}$ .
23      Standardise  $y^{(i)}$  with  $m_0^{(i)}, \sigma_0^{(i)}$  to obtain  $\tilde{y}^{(i)}$ .
24      Add  $\tilde{y}^{(i)}$  as the last row of  $\tilde{\mathbf{Y}}^{(i)}$ .
25      Update the GP regression model on  $\tilde{\mathbf{X}}$  and  $\tilde{\mathbf{Y}}^{(i)}$  (obtain  $\phi^{(i)}$ ).
26    end
27  end
28 end

```

---

## 4 | APPLICATIONS

This section presents three applications of BO to the problem of parameter inference in the HO law of the LV. We start with comparing BO *without* the Klotz-curve prior with the original HGO algorithm<sup>13</sup>, which we refer to as the *basic study*. We demonstrate that BO attains a similar or a lower value of the objective function  $f_{O_2}$  in considerably fewer iterations (Section 4.1). Next, in the *Klotz-curve study* we compare the versions of both algorithms that include the Klotz-curve (Section 4.2). We show that also in this case BO typically outperforms the multi-step algorithm, even though the updated HGO algorithm<sup>25</sup> works better than its original version. In both these studies we consider four healthy volunteers, same in both applications, which we label HV A, HV B, HV C and HV D.

### Set-up

In all our applications we compare two variants of BO, with target and partial error surrogates. We use 500 iterations for BO, which are run after 40 iterations corresponding to an initial design obtained with Latin hypercube sampling<sup>38</sup>, in which we follow a commonly followed rule-of-thumb<sup>10</sup> to use  $10d$  points to initialise optimisation (for a  $d$ -dimensional problem). We use a recent value for  $\psi_0$  (defined below (6)) based on in vivo data<sup>4</sup> for both BO and the HGO algorithm, except the original HGO algorithm in the *basic study* for which we use the reference value from Wang et al.<sup>19</sup>, originally used in the HGO algorithm.

### Evaluation

In our result analysis we focus on two aspects. First, we consider the value of the objective function, as this is what both algorithms directly target. In the *basic study* this is  $f_{O_2}$  given in (5), in the two later studies this is  $f_{O_2, \text{Klotz}}$  given in (10). Visually, we analyse the convergence of the *incumbent*, i.e. the best value of the objective function recorded so far, which is standard in the BO literature<sup>6,16</sup>. The convergence is with respect to the invocations of the forward simulator, which are the main computational burden and which we want to limit. The physical computing time is a less reliable measure of computing costs as it necessarily depends on the machine used, number of task performed in parallel, etc., and does not generalise to other mechanistic models with different computational costs of the forward problem. Note that the HGO algorithm is based on optimisation solvers implemented as “black-boxes” (such as the trust-region-reflective algorithm or the sequential quadratic programming algorithm in MATLAB), so that tracking the incumbent over the individual iterations is not directly available. Moreover, each step in HGO has its own stopping criterion determining the number of iterations carried out in each step. Therefore, for HGO we just report the final value of the objective  $y_{\min}^{\text{HGO}}$  and the corresponding number of iterations  $i_{\min}^{\text{HGO}}$  (defined as the sum of iterations performed in all steps).

Quantitatively, we compare BO and HGO based on  $(y_{\min}, i_{\min})$  and  $(y_{\min}^{\text{HGO}}, i_{\min}^{\text{HGO}})$ . The former pair is concerned with the lowest value attained by BO over all 540 iterations  $y_{\min}$  and the number of iterations completed to obtain  $y_{\min}$  denoted

$i_{\min}$ . These values are compared with their HGO counterparts,  $y_{\min}^{\text{HGO}}$  and  $i_{\min}^{\text{HGO}}$ . The latter pair relates BO iterations to the final result from the HGO algorithm:  $y(i_{\min}^{\text{HGO}})$  is the value of a BO run after the number of iterations used by the HGO algorithm  $i_{\min}^{\text{HGO}}$ , while  $i(y_{\min}^{\text{HGO}})$  is the number of iterations required to achieve at least as good a performance as the HGO algorithm. If after 540 iterations the value of the objective obtained with BO is still higher than the final value from the HGO algorithm we set  $i(y_{\min}^{\text{HGO}}) = \emptyset$ .

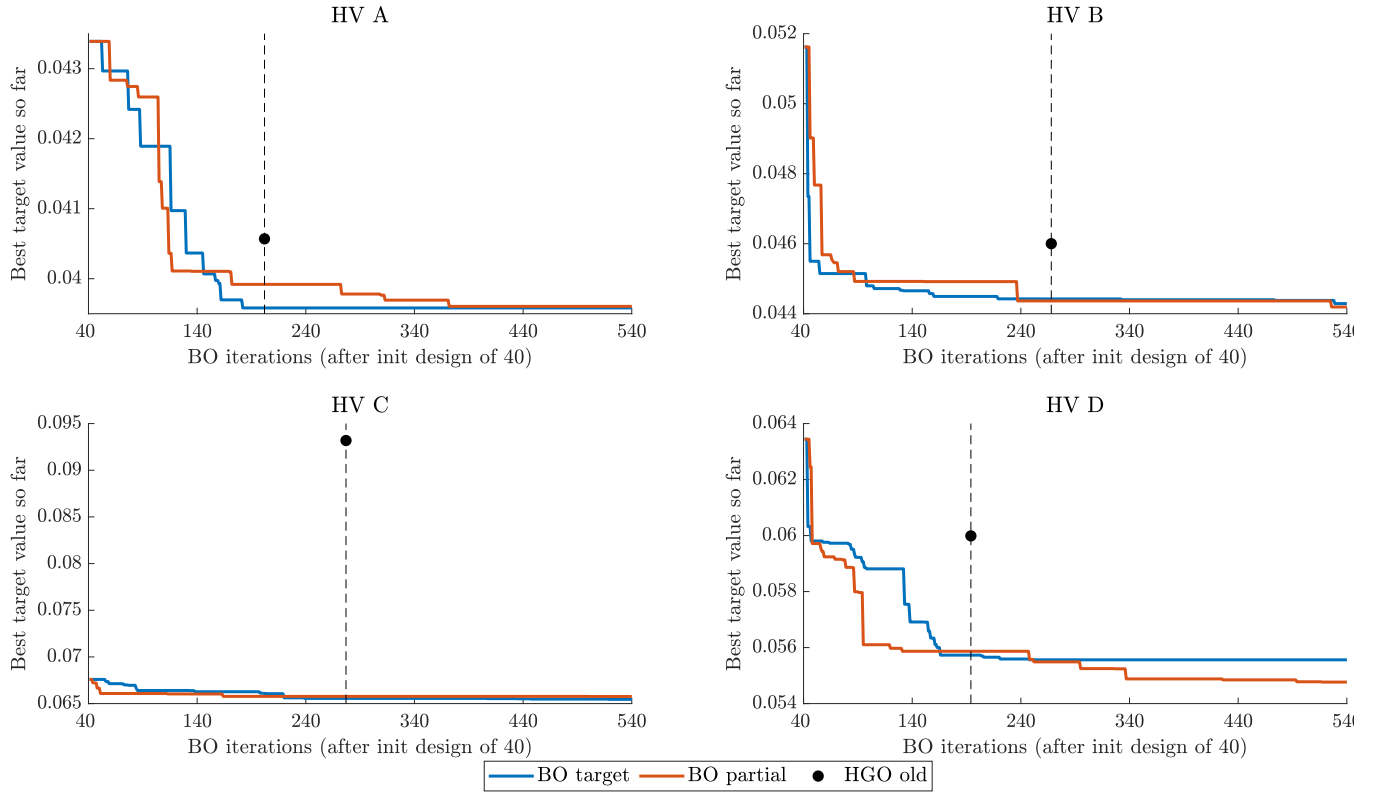
Second, we analyse the *stress-stretch curves* corresponding to the final parameter estimates. We consider two virtual uni-axial stretch experiments, one is along the myofibre direction (myocytes), and the other one is along the sheet direction (transmurally), often adopted in the literature for characterizing nonlinear myocardial property<sup>13,20,39</sup>. Stretch-stress curves present the relationship between stress and stretch along two axes in the myocardium, along the myocyte and along the sheet direction. They provide insights into the predicted stiffness of the myocardium, with higher values of stress for a given stretch value indicating a stiffer material. They are of interest to cardiac mechanics specialist and clinicians as the parameters of the HO model cannot always be uniquely identified using experimental data<sup>13</sup> and due to the nonlinear myocardium response. The latter makes the differences in the predicted parameters not fully informative about the implied output differences. For the reproducibility of stretch-stress curves, we report the underlying parameters in Appendix A.

## 4.1 | Basic study

Gao et al.<sup>13</sup> have established a benchmark for parameter inference in the HO model. Their algorithm (see Listing 1) does not involve the Klotz-curve prior discussed in Section 2.2, hence for comparability we first consider a version of BO without the Klotz-curve prior, with the objective function given by  $f_{O2}$  from (5).

Figure 3 presents the convergence of the *incumbent* for the four considered HVs. We observe that BO converges very quickly, typically in around 200 iterations after the initial design, so that the graph stays flat over the remaining iterations. The HGO algorithm typically requires more iterations to terminate. Moreover, in all the cases BO converges to a lower value of the objective function than HGO. In the most extreme case of HV C, the lowest value found on the BO initial design is lower than the final value for HGO. The two variants of BO, with target and partial error surrogates, generally perform comparably, except for HV D, for which BO with a partial error surrogate converges faster and to a better point than BO with a target surrogate.

The main quantitative insights from Figure 3 are summarised in Table 2. We can see that  $y_{\min}$  obtained with BO are always lower than their HGO counterparts. However, in most cases (except HV A and BO with a target surrogate) they are obtained using more iterations than HGO. A fairer comparison is hence provided by  $y_{\min}^{\text{HGO}}$ . Even after less iterations BO still provides lower values of the objective function. This better performance of BO can be alternatively assessed using  $i(y_{\min}^{\text{HGO}})$  which shows that for HV B, HV C and HV D we record at least as good performance of BO as for HGO using 4–7 times less iterations.



**FIGURE 3 Basic study:** convergence of the objective function  $f_{O2}$  for Bayesian optimization and the original HGO algorithm for four LV geometries (HV A, HV B, HV C, HV D). Horizontal axis: Bayesian optimization iterations starting from 40 iterations for the initial design. Vertical axis: best value of the objective function  $f_{O2}$  recorded so far. Black dot and horizontal dashed line: the final value of the objective function  $f_{O2}$  for the HGO algorithm and the associated number of iterations. Bayesian optimization with a target surrogate (target) and a partial error surrogate (partial) together with the old version of the HGO algorithm (HGO old).

As mentioned at the beginning of this section, we focus on analysing stretch-stress curves, presented in Figure 4, and not parameters estimates explicitly (the latter are nevertheless reported in Appendix A). We observe that the HGO algorithm predicts stiffer myocardium for high stretches along myocyte, while for low stretches the responses from BO and HGO are typically very comparable. The only exception is HV C for which HGO predicts much higher stiffness than BO along the whole range of stretches. On the other hand, for HV C both algorithms predict almost identical responses for stretches along the sheet direction. For the remaining subjects sheet direction stretches are associated with higher stiffness predicted by BO compared with HGO.

We conclude that the original HGO algorithm is easily outperformed by both variants of BO. Moreover, the mechanical response predicted by both BO and HGO for high stretches (i.e.  $> 1.15$ ) can differ substantially. This could be partially explained by the limitation of using in vivo measurements, which are less informative about stiffening effects under high stretch as the myocardium is mostly operating in the linear regime of the stress-stretch curve in vivo, in particular for healthy LVs with a maximum stretch below 1.2<sup>40</sup>. The latter confirms that inferring parameters characterising the nonlinear stiffening effects is

Case	$y_{\min}$	$i_{\min}$	$y(i_{\min}^{\text{HGO}})$	$i(y_{\min}^{\text{HGO}})$	#Cr.
HV02 BO targ.	0.0396	181	0.0396	129	1
HV02 BO part.	0.0396	371	0.0399	113	1
HV02 HGO old	0.0406	201	0.0406	201	0
HV B BO targ.	0.0443	528	0.0444	45	0
HV B BO part.	0.0442	525	0.0444	56	6
HV B HGO old	0.0460	267	0.0460	267	0
HV C BO targ.	0.0655	518	0.0656	40	3
HV C BO part.	0.0658	499	0.0658	40	0
HV C HGO old	0.0932	276	0.0932	276	0
HV D BO targ.	0.0556	247	0.0557	46	7
HV D BO part.	0.0548	516	0.0559	47	0
HV D HGO old	0.0600	193	0.0600	193	0

**TABLE 2 Basic study:** convergence of the objective function for Bayesian optimization and the original HGO algorithm for four LV geometries (HV A, HV B, HV C, HV D). Bayesian optimization with a target surrogate (targ.) and a partial error surrogate (part.) together with the old version of the HGO algorithm (HGO old). The lowest values of the  $f_{O_2}$  objective function ( $y_{\min}$ ), number of iterations leading to  $y_{\min}$  ( $i_{\min}$ ;  $y(i_{\min}) = y_{\min}$ ), the value of the objective after the number of iterations used by the HGO algorithm ( $y(i_{\min}^{\text{HGO}})$ ), the number of iterations required to achieve at least as good a performance as the HGO algorithm ( $i(y_{\min}^{\text{HGO}})$ ), and the total number of crashes of the forward simulator (#Cr.).

challenging when based on in vivo data only. Therefore, introducing extra information about pressures higher than the normal end-diastolic pressure becomes necessary in order to characterise nonlinear responses at higher pressures. In the clinic, this could be done through stress CMR<sup>41</sup>. In this study, in our next application we will rely on the Klotz-law discussed in Section 2.2, which provides an alternative way to capture the nonlinear relationship between the LV cavity volume and the loading pressure within a large range, i.e. up to as high as 30 mmHg.

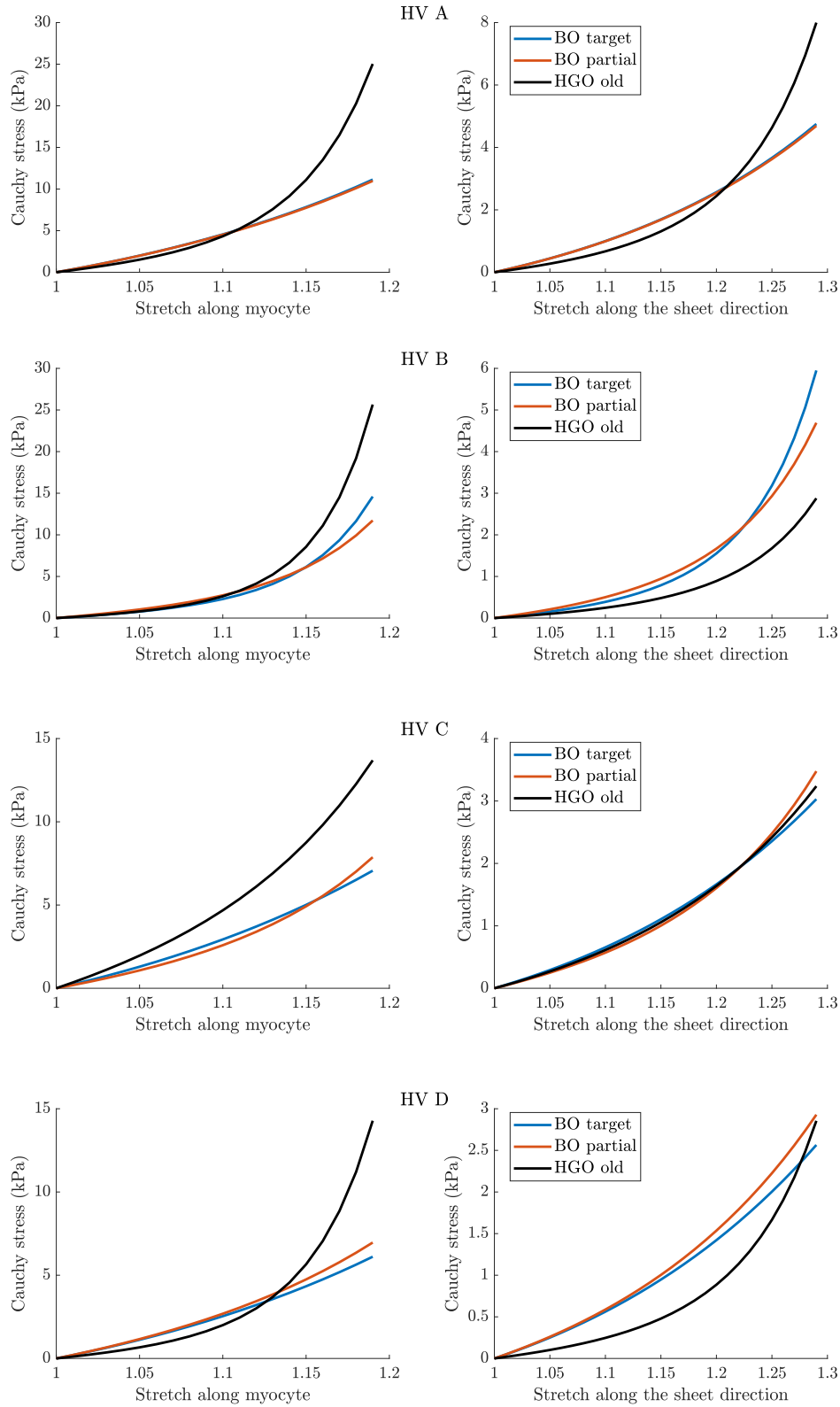
## 4.2 | Klotz-curve prior study

In our analysis of the basic study we have demonstrated that our proposed BO framework outperforms the original HGO algorithm<sup>13</sup> in terms of delivering lower values of the objective function using less invocations of the forward simulator. In this section we want to replace that benchmark with its updated version<sup>25</sup>. We can now use the full BO algorithm described in Listing 4 as the new HGO algorithm includes the Klotz-curve prior. This time we run each version of BO (with a target surrogate and a partial error surrogate) three times (with independent runs based on the same initial design) to verify the robustness and stability of the procedure.

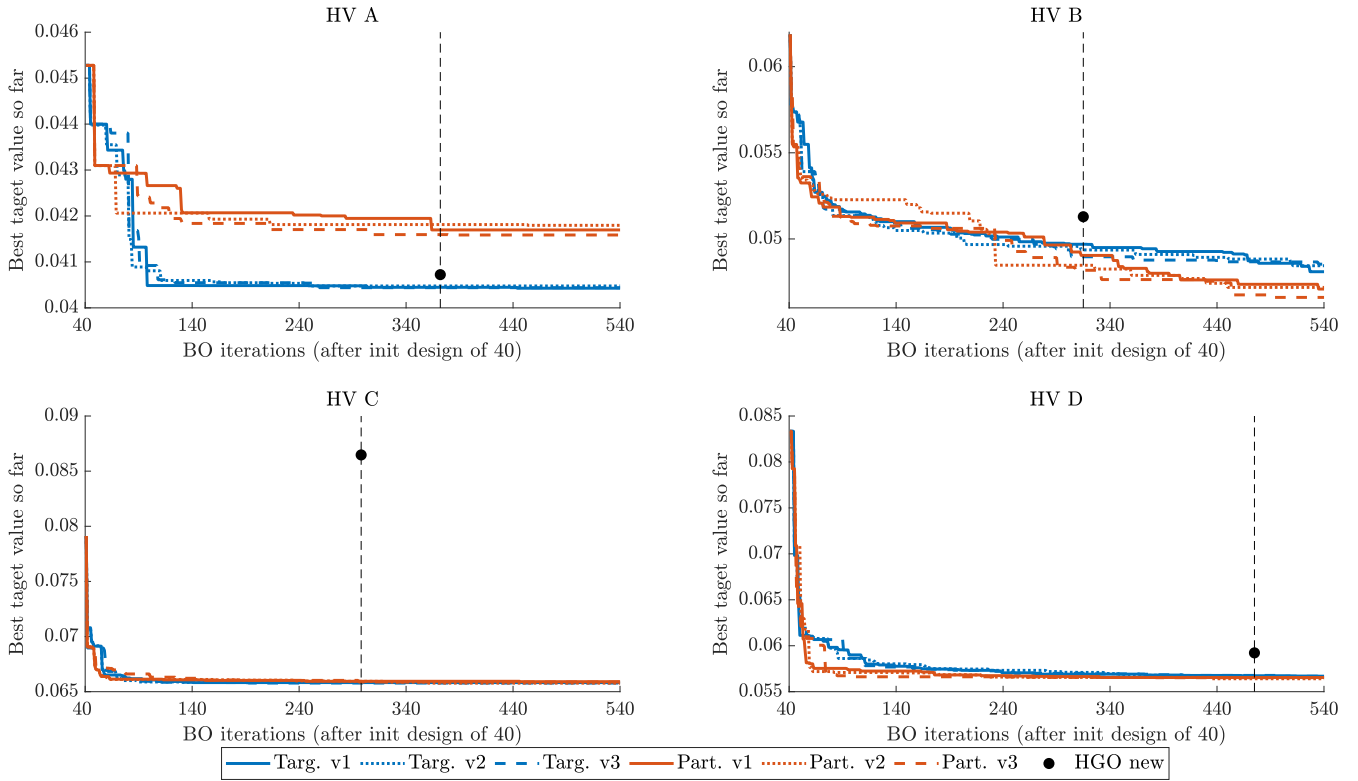
Figure 5 shows that BO again outperforms the HGO algorithm in terms of achieving lower values of the objective function  $f_{O_2, \text{Klotz}}$  in fewer iterations. We can see that the updated HGO algorithm performs better than the original one *relatively* to BO, e.g. for HV C and HV D the new HGO algorithm converges closer to BO than in the basic study.

Table 3 summarises the results related to the convergence of the objective function. Since in the present study the performance of the HGO algorithm is more comparable to the one of BO, we decided to investigate whether the low values of the objective function for BO were not a consequence of a potential bias in the  $V_{30}$  emulator. The latter is used as a part of our framework aimed at achieving computational savings as discussed in Section 3.3, however the new HGO algorithm still uses the forward simulator to obtain the volume at 30 mmHg. We thus run the forward simulator for each run of BO at points  $\theta$  corresponding to  $y_{\min}$  and  $y(i_{\min}^{\text{HGO}})$ , to obtain the true values of  $V_{30}$  at these points and hence true values of  $f_{O2, \text{Klotz}}$ . We report the volumes in the columns indicated with  $V_{30}$ , with the  $E(\cdot)$  values being the volumes from the emulator and the  $S(\cdot)$  values being the volumes from the simulator. Similarly, we have two sets of the objective function values, obtained with the simulator (with superscript  $S$ ) and the emulator. These results reveal that our  $V_{30}$  emulator provides very accurate predictions, with hardly any differences in the high pressure volumes and consequently in the values of the objective function compared to the simulator. Figure 6 illustrates this point using a Bland-Altman plot. The first four sets of points (from the left) show the differences between the simulator-based and emulator-based values of  $f_{O2, \text{Klotz}}$  against the average over both these values. Note that the scale of the y-axis is  $10^{-3}$ . We can see that all the differences are indeed negligible so comparing the HGO algorithm against emulator-based runs of BO is valid.

The stretch-stress curves in Figure 7 show that including the Klotz-curve prior leads to the results from BO and HGO being more comparable. For all the curves, except HV C along the sheet direction, the differences in predictions for small stretches (observable in vivo) observed in the basic study now with the Klotz-curve prior are considerably reduced. Moreover, comparing e.g. the curves along the myocyte direction for HV B with those from Figure 4 we observe that now the predicted stiffness for high stretches is noticeably reduced.



**FIGURE 4 Basic study:** stretch-stress curves for four LV geometries (HV A, HV B, HV C, HV D). Left: responses to stretches along the myocyte direction  $\mathbf{f}_0$ , right: responses to stretches along the sheet direction  $\mathbf{s}_0$  (see (2)). Bayesian optimization with a target surrogate (target) and a partial error surrogate (partial) together with the old version of the HGO algorithm (HGO old).



**FIGURE 5 Klotz-curve study:** convergence of the objective function  $f_{O2, \text{Klotz}}$  for Bayesian optimization and the original HGO algorithm for four LV geometries (HV A, HV B, HV C, HV D). Horizontal axis: Bayesian optimization iterations after 40 iterations for the initial design. Vertical axis: best value of the objective function  $f_{O2}$  recorded so far. Black dot and horizontal dashed line: the final value of the objective function  $f_{O2}$  for the HGO algorithm and the associated number of iterations. Bayesian optimization with a target surrogate (targ.) and a partial error surrogate (part.), three independent runs (v1, v2, v3) in each version, together with the new version of the HGO algorithm (HGO new). For HGO, the Klotz curve error was computed using the forward simulator (not the emulator).



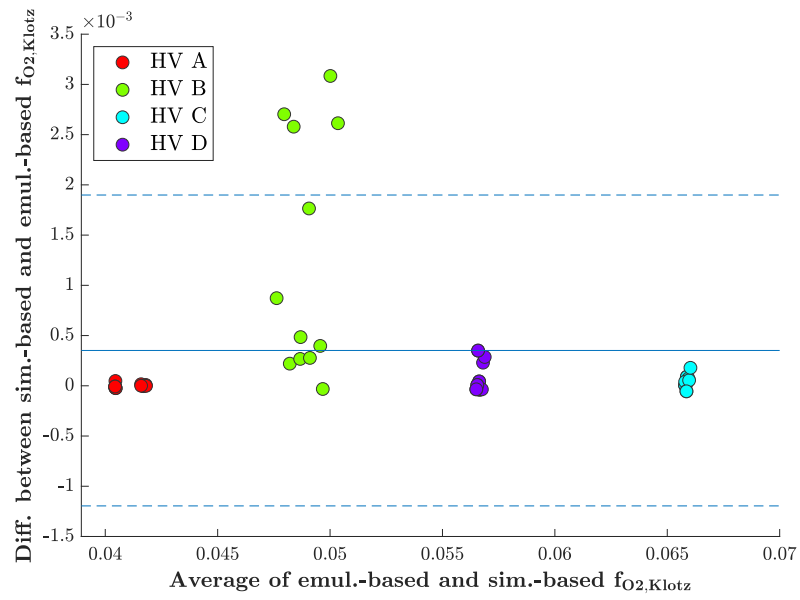
**TABLE 3 Klotz-curve study:** convergence of the objective function for Bayesian optimization and the updated HGO algorithm for four LV geometries (HV A, HV B, HV C, HV D). Three independent runs for BO (v1, v2, v3) together with the new version of HGO (HGO new). Table reports:  $y_{\min}$  – the lowest value of the  $f_{O_2, \text{Klotz}}$  objective function,  $y_{\min}^S$  –  $y_{\min}$  but with the Klotz error computed using the simulator (not emulator),  $i_{\min}$  – number of iterations leading to  $y_{\min}$  ( $y(i_{\min}) = y_{\min}$ ),  $y(i_{\min}^{\text{HGO}})$  – value of the objective after the number of iterations used by HGO,  $y^S(i_{\min}^{\text{HGO}})$  –  $y(i_{\min}^{\text{HGO}})$  but with the Klotz error computed using the simulator (not emulator),  $i(y_{\min}^{\text{HGO}})$  – number of iterations required to achieve at least as good a performance as the HGO algorithm ( $\emptyset$  if that did not happen within 540 iterations),  $V_{30}$  – LV volumes at 30 mmHg, with  $E(y_{\min})$  for  $y_{\min}$  computed with the emulator;  $S(y_{\min})$  for  $y_{\min}$  computed with the simulator;  $E(i_{\min}^{\text{HGO}})$  for  $y(i_{\min}^{\text{HGO}})$  computed with the emulator;  $S(i_{\min}^{\text{HGO}})$  for  $y(i_{\min}^{\text{HGO}})$  computed with the simulator; #Cr. – total number of crashes of the forward simulator. Missing values for simulator-based values ( $y_{\min}^S$ ,  $V_{30}$  with  $S(y_{\min})$  and  $S(i_{\min}^{\text{HGO}})$ ) when the simulator crashed.

(a) Target surrogate.

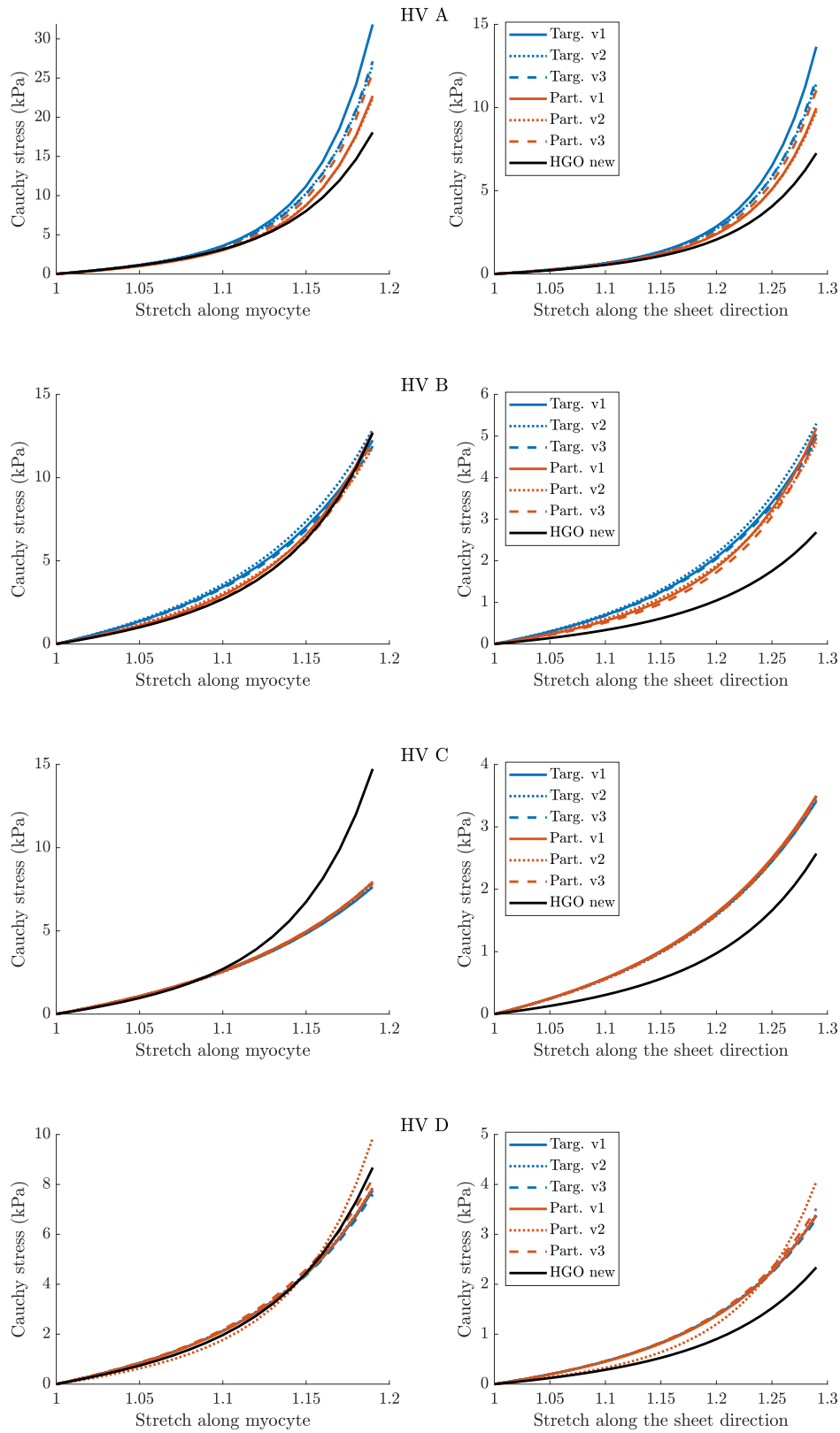
Case	$y_{\min}$	$y_{\min}^S$	$i_{\min}$	$y(i_{\min}^{\text{HGO}})$	$y^S(i_{\min}^{\text{HGO}})$	$i(y_{\min}^{\text{HGO}})$	$V_{30}$				#Cr.
							$E(y_{\min})$	$S(y_{\min})$	$E(i_{\min}^{\text{HGO}})$	$S(i_{\min}^{\text{HGO}})$	
HV A BO v1	0.040	0.040	443	0.040	0.040	97	223	222	224	223	3
HV A BO v2	0.040	0.040	248	0.040	0.040	110	225	224	225	224	2
HV A BO v3	0.040	0.040	258	0.040	0.040	107	224	223	224	223	3
HV A HGO new	0.041	0.041	371	0.041	0.041	371	–	224	–	224	0
HV B BO v1	0.048	0.048	526	0.050	0.050	116	189	187	188	188	0
HV B BO v2	0.048	0.049	510	0.049	0.050	114	189	186	188	186	0
HV B BO v3	0.049	0.049	534	0.049	0.049	111	189	187	187	186	0
HV B HGO new	0.051	0.051	314	0.051	0.051	314	–	179	–	179	0
HV C BO v1	0.066	0.066	140	0.066	0.066	40	229	227	229	227	1
HV C BO v2	0.066	0.066	327	0.066	0.066	40	230	227	230	227	1
HV C BO v3	0.066	0.066	192	0.066	0.066	40	230	228	230	228	0
HV C HGO new	0.086	0.086	297	0.086	0.086	297	–	200	–	200	0
HV D BO v1	0.057	0.057	534	0.057	0.057	95	152	149	152	149	0
HV D BO v2	0.057	0.057	464	0.057	0.057	85	153	149	153	149	2
HV D BO v3	0.057	0.057	500	0.057	0.057	92	153	150	153	150	1
HV D HGO new	0.059	0.059	474	0.059	0.059	474	–	144	–	144	0

(b) Partial error surrogate.

Case	$y_{\min}$	$y_{\min}^S$	$i_{\min}$	$y(i_{\min}^{\text{HGO}})$	$y^S(i_{\min}^{\text{HGO}})$	$i(y_{\min}^{\text{HGO}})$	$V_{30}$				#Cr.
							$E(y_{\min})$	$S(y_{\min})$	$E(i_{\min}^{\text{HGO}})$	$S(i_{\min}^{\text{HGO}})$	
HV A BO v1	0.042	0.042	363	0.042	0.042	$\emptyset$	223	223	223	223	1
HV A BO v2	0.042	0.042	453	0.042	0.042	$\emptyset$	224	224	222	222	2
HV A BO v3	0.042	0.042	379	0.042	0.042	$\emptyset$	224	225	223	223	1
HV A HGO new	0.041	0.041	371	0.041	0.041	371	–	224	–	224	0
HV B BO v1	0.047	0.050	534	0.049	0.052	97	185	179	182	178	1
HV B BO v2	0.047	0.048	449	0.048	0.052	208	188	184	183	177	1
HV B BO v3	0.047	0.049	486	0.048	0.050	88	185	179	184	180	0
HV B HGO new	0.051	0.051	314	0.051	0.051	314	–	179	–	179	0
HV C BO v1	0.066	0.066	446	0.066	0.066	40	229	227	229	227	5
HV C BO v2	0.066	0.066	357	0.066	0.066	40	229	227	228	226	1
HV C BO v3	0.066	0.066	295	0.066	0.066	40	231	228	231	228	0
HV C HGO new	0.086	0.086	297	0.086	0.086	297	–	200	–	200	0
HV D BO v1	0.057	0.057	330	0.057	0.057	55	152	150	152	150	0
HV D BO v2	0.056	0.057	434	0.056	0.057	58	151	148	151	148	2
HV D BO v3	0.057	0.056	382	0.057	0.056	73	153	150	153	150	3
HV D HGO new	0.059	0.059	474	0.059	0.059	474	–	144	–	144	0



**FIGURE 6 Klotz-curve:** Bland–Altman plot comparing the Klotz objective function with  $V_{30}$  computed using the emulator and the forward simulator.



**FIGURE 7 Klotz-curve study:** stretch-stress curves for four LV geometries (HV A, HV B, HV C, HV D). Left: responses to stretches along the myocyte direction  $\mathbf{f}_0$ , right: responses to stretches along the sheet direction  $\mathbf{s}_0$  (see (2)). Bayesian optimization with a target surrogate (targ.) and a partial error surrogate (part.), three independent runs (v1, v2, v3) in each version.

## 5 | DISCUSSION

We have proposed an efficient, Bayesian optimisation–based framework for parameter inference in a cardiac mechanic model of the left ventricle in diastole. We have demonstrated that BO converges to lower values of the two objective functions considered and requires less invocations of the associated forward simulator than the state-of-the-art multi-step algorithm of Gao et al.<sup>13</sup>. We have developed a new approach to minimising a target function given as a sum of error terms based on approximating each of these terms individually via partial error surrogates. The findings of our empirical studies suggest that this approach may outperform the standard one based on target surrogates. However, which of the two variants is preferred in a specific case could be subject-dependent (depending on e.g. the LV geometry).

The particular choice of the acquisition function is pivotal for the performance of BO. The Expected Improvement (EI) criterion applied in this paper has a number of advantages, most importantly it admits a closed form formula under Gaussian process surrogates and is widely used in the machine learning community. However, we note that the analytic tractability of EI becomes less advantageous as we do not use just EI as our AF, but  $EI_{con}$ , in which EI is augmented with the probit-based probability of unknown constraint satisfaction, making the whole AF not tractable. Moreover, EI might perform in a greedy way<sup>42</sup>.

The development of better specifications for acquisition functions is an active research area in the BO literature<sup>6</sup>. Two classes of alternative acquisition functions seem to be particularly relevant in our context. The first group consists of information-based policies, such as the Entropy Search<sup>43,44</sup> (ES) criterion, that aim to infer the location of the unknown minimum by focusing on the posterior distribution of the minimiser. When deciding where to query ES takes into account what we expect to *learn* about the unknown minimum from evaluating at a particular point. This is different to EI, which evaluates where we believe the minimum is, without taking into account the effect of such a query on learning. ES is thus conceptually very appealing, however it is more challenging to implement in practice than EI. The second group of potentially useful AF is related to the so-called portfolios of AFs<sup>45,6</sup>. As demonstrated by our empirical studies, it is unlikely to develop a globally preferable acquisition strategy, i.e. performing best regardless of the given LV geometry and measurements. The portfolio approach considers multiple AFs and uses a meta-criterion (a high level AF) to select the final query point from those proposed by individual AFs.

Despite considerable efficiency gains compared to the HGO algorithm, BO is still likely to be too time-consuming to provide a viable tool for the clinical practice if optimisation is supposed to be performed independently for each subject, starting from scratch. Multi-task BO<sup>46</sup> could address this issue by leveraging prior knowledge acquired from searches over comparable domains, in our case from optimisation performed for previous subjects. Multi-task BO uses the data coming from different searches by processing them simultaneously and constructing a multi-task GP<sup>37</sup> to capture correlation between related tasks. This approach would allow the user to focus on parameter regions that have been found to be promising.

## ACKNOWLEDGMENTS

This work was funded by the UK Engineering and Physical Sciences Research Council (EPSRC), grant numbers EP/N014642/1 and EP/R018634/1. Dirk Husmeier is supported by a grant from the Royal Society of Edinburgh, award number 62335.

## Conflict of interest

The authors declare no potential conflict of interests.

## References

1. Mangion K, Gao H, Husmeier D, Luo XY, Berry C. Advances in computational modelling for personalised medicine after myocardial infarction. *Heart* 2018; 104(7): 550–557.
2. 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.
3. Gao H, Mangion K, Carrick D, Husmeier D, Luo XY, Berry C. Estimating prognosis in patients with acute myocardial infarction using personalized computational heart models. *Scientific Reports* 2017; 7(13527): 1.
4. Gao H, Aderhold A, Mangion K, Luo X, Husmeier D, Berry C. Changes and Classification in Myocardial Contractile Function in the Left Ventricle following Acute Myocardial Infarction. *Journal of The Royal Society Interface* 2017; 14(132): 20170203.
5. Bathe KJ. *Finite element procedures, 2nd edition*. Englewood Cliffs, N. J.: Prentice-Hall . 2014.
6. Shahriari B, Swersky K, Wang Z, Adams RP, Freitas dN. Taking the human out of the loop: a review of Bayesian optimization. *Proceedings IEEE* 2016; 104(1): 148–175.
7. Wilson J, Hutter F, Deisenroth M. Maximizing acquisition functions for Bayesian optimization. In: *Advances in Neural Information Processing Systems 31*; 2018: 9884–9895.
8. Rasmussen CE, Williams CKI. *Gaussian Processes for Machine Learning*. MIT Press . 2006.
9. Wang C, Duan Q, Gong W, Ye A, Di Z, Miao C. An evaluation of adaptive surrogate modeling based optimization with two benchmark problems. *Environmental Modelling & Software* 2014; 60: 167–179.
10. Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. *Journal of Global optimization* 1998; 13(4): 455–492.

11. Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. In: *Advances in Neural Information Processing Systems* NeurIPS. ; 2012: 2951–2959.
12. Holzapfel GA, Ogden RW. Constitutive modelling of passive myocardium: a structurally based framework for material characterization. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 2009; 367(1902): 3445–3475.
13. Gao H, Li WG, Cai L, Berry C, Luo XY. Parameter estimation in a Holzapfel–Ogden law for healthy myocardium. *Journal of Engineering Mathematics* 2015; 95(1): 231–248.
14. Klotz S, Hay I, Dickstein ML, et al. Single-beat estimation of end-diastolic pressure-volume relationship: a novel method with potential for noninvasive application. *American Journal of Physiology-Heart and Circulatory Physiology* 2006; 291(1).
15. Snoek JR. *Bayesian optimization and semiparametric models with applications to assistive technology*. PhD thesis. University of Toronto, Toronto, Canada; 2013.
16. Gelbart MA, Snoek J, Adams RP. Bayesian optimization with unknown constraints. In: *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence*; 2014: 250–259.
17. Davies V, Noè U, Lazarus A, et al. Fast parameter inference in a biomechanical model of the left ventricle by using statistical emulation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2019; 68(5): 1555–1576.
18. Zile MR, Brutsaert DL. New concepts in diastolic dysfunction and diastolic heart failure: part I. *Circulation* 2002; 105(11): 1387–1393.
19. Wang HM, Gao H, Luo XY, et al. Structure based finite strain modelling of the human left ventricle in diastole. *International Journal for Numerical Methods in Biomedical Engineering* 2013; 29(1): 83–103.
20. Guan D, Ahmad F, Theobald P, Soe S, Luo XY, Gao H. On the AIC-based model reduction for the general Holzapfel–Ogden myocardial constitutive law. *Biomechanics and Modeling in Mechanobiology* 2019; 18(4): 1213–1232. doi: 10.1007/s10237-019-01140-6
21. Sack KL, Aliotta E, Ennis DB, et al. Construction and validation of subject-specific biventricular finite-element models of healthy and failing swine hearts from high-resolution DT-MRI. *Frontiers in Physiology* 2018; 9: 539.
22. Peirlinck M, Sahli Costabal F, Sack KL, et al. Using machine learning to characterize heart failure across the scales. *Biomechanics and Modeling in Mechanobiology* 2019; 18(6): 1987–2001.

23. Noè U, Lazarus A, Gao H, et al. Gaussian process emulation to accelerate parameter estimation in a mechanical model of the left ventricle: a critical step towards clinical end-user relevance. *Journal of the Royal Society Interface* 2019; 16(156): 20190114.
24. Cerqueira MD, Weissman NJ, Dilsizian V, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *Circulation* 2002; 105(4): 539–542.
25. Gao H, Mangion K, Berry C, Luo XY. Mathematical Modelling Acute Myocardial Infarction using in vivo Magnetic Resonance Imaging. *Proceeding of virtual physiological human conference, Zaragoza Spain* 2018.
26. Moisejko D, McGarvey JR, Dorsey SM, et al. Estimating passive mechanical properties in a myocardial infarction using MRI and finite element simulations. *Biomechanics and Modeling in Mechanobiology* 2015; 14(3): 633–647.
27. Kirschner J, Mutný M, Hiller N, Ischebeck R, Krause A. Adaptive and safe Bayesian optimization in high dimensions via one-dimensional subspaces. In: *Proceedings of the 36th International Conference on Machine Learning*; 2019: 3429–3438.
28. Ugray Z, Lasdon L, Plummer J, Glover F, Kelly J, Martí R. Scatter search and local NLP solvers: A multistart framework for global optimization. *INFORMS Journal on Computing* 2007; 19(3): 328–340.
29. Glover F. A template for scatter search and path relinking. In: *European Conference on Artificial Evolution* Springer Berlin Heidelberg; 1998: 1–51.
30. Minka TP. Expectation propagation for approximate Bayesian inference. In: *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*; 2001: 362–369.
31. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series B (statistical methodology)* 2009; 71(2): 319–392.
32. Jordan MI, Ghahramani Z, Jaakkola TS, Saul LK. An introduction to variational methods for graphical models. *Machine learning* 1999; 37(2): 183–233.
33. Nickisch H, Rasmussen CE. Approximations for binary Gaussian process classification. *Journal of Machine Learning Research* 2008; 9(Oct): 2035–2078.
34. Noè U, Chen W, Filippone M, Hill N, Husmeier D. Inference in a partial differential equations model of pulmonary arterial and venous blood circulation using statistical emulation. In: Bracciali A, Caravagna G, Gilbert D, Tagliaferri R., eds. *Computational Intelligence Methods for Bioinformatics and Biostatistics* Springer International Publishing; 2017.

35. Conti S, O'Hagan A. Bayesian emulation of complex multi-output and dynamic computer models. *Journal of Statistical Planning and Inference* 2010; 140(3): 640–651.
36. Conti S, Gosling JP, Oakley JE, O'Hagan A. Gaussian process emulation of dynamic computer codes. *Biometrika* 2009; 96(3): 663–676.
37. Bonilla EV, Chai KM, Williams C. Multi-task Gaussian process prediction. In: *Advances in Neural Information Processing Systems* 21; 2008: 153–160.
38. McKay MD, Beckman RJ, Conover WJ. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 1979; 21(2): 239–245.
39. Sommer G, Haspinger DC, Andrä M, et al. Quantification of shear deformations and corresponding stresses in the biaxially tested human myocardium. *Annals of biomedical engineering* 2015; 43(10): 2334–2348.
40. Mangion K, Gao H, McComb C, et al. A novel method for estimating myocardial strain: assessment of deformation tracking against reference magnetic resonance methods in healthy volunteers. *Scientific Reports* 2016; 6(1): 38774.
41. Hundley WG, Morgan TM, Neagle CM, Hamilton CA, Rerkpattanapipat P, Link KM. Magnetic resonance imaging determination of cardiac prognosis. *Circulation* 2002; 106(18): 2328–2333.
42. Qin C, Klabjan D, Russo D. Improving the expected improvement algorithm. In: *Advances in Neural Information Processing Systems*; 2017: 5381–5391.
43. Hennig P, Schuler CJ. Entropy search for information-efficient global optimization. *The Journal of Machine Learning Research* 2012; 13(1): 1809–1837.
44. Hernández-Lobato JM, Hoffman MW, Ghahramani Z. Predictive entropy search for efficient global optimization of black-box functions. In: *Advances in Neural Information Processing Systems* 27; 2014: 918–926.
45. Hoffman M, Brochu E, Freitas dN. Portfolio allocation for Bayesian optimization. In: *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*; 2011: 327–336.
46. Swersky K, Snoek J, Adams RP. Multi-task Bayesian optimization. In: *Advances in Neural Information Processing Systems*; 2013: 2004–2012.



**How to cite this article:** Borowska A., H. Gao, A. Lazarus, and D. Husmeier (2020), Bayesian optimisation for efficient parameter inference in a cardiac mechanics model of the left ventricle, *Int J Numer Meth Biomed Engng.*, 2020;XX:XX–XX.

## APPENDIX

### A RESULTS ON PARAMETERS

#### A.1 Basic study

**TABLE A1 Basic setting:** final optimized values of the eight parameters of the HO law for Bayesian optimization and the original HGO algorithm (HGO old) for four different LV geometries (HV A, HV B, HV C, HV D), Bayesian optimization with a target surrogate (targ.) and a partial error surrogate (part.).

Case	$a$	$b$	$a_f$	$b_f$	$a_s$	$b_s$	$a_{fs}$	$b_{fs}$
HV A BO targ.	0.257	3.715	8.636	0.409	1.784	0.166	0.045	0.371
HV A BO part.	0.258	3.728	8.581	0.360	1.773	0.146	0.074	0.619
HV A HGO old	0.100	5.405	6.172	7.114	1.117	2.582	0.123	5.650
HV B BO targ.	0.018	0.260	3.174	7.859	0.656	3.196	1.444	12.016
HV B BO part.	0.018	0.260	4.344	4.780	0.898	1.944	1.458	12.131
HV B HGO old	0.100	5.766	3.090	11.252	0.372	2.582	0.041	5.650
HV C BO targ.	0.207	2.991	5.567	0.294	1.150	0.119	1.546	12.864
HV C BO part.	0.245	3.532	4.481	2.158	0.926	0.877	1.550	12.900
HV C HGO old	0.100	1.960	8.447	1.811	1.117	0.516	0.123	1.130
HV D BO targ.	0.162	2.342	4.834	0.293	0.999	0.119	0.676	5.623
HV D BO part.	0.174	2.519	4.982	0.878	1.029	0.357	0.169	1.406
HV D HGO old	0.100	5.405	2.626	8.791	0.372	2.582	0.041	5.650

## A.2 Klotz-curve study

**TABLE A2 Klotz-curve study:** final optimized values of the parameters of the HO law for Bayesian optimization and the updated HGO algorithm for four LV different geometries (HV A, HV B, HV C, HV D). For each BO variant (target surrogate, partial error surrogate three independent runs are reported (v1, v2, v3).

(a) Target surrogate.								
Case	$a$	$b$	$a_f$	$b_f$	$a_s$	$b_s$	$a_{fs}$	$b_{fs}$
HV A v1	0.279	4.033	4.326	10.536	0.894	4.284	0.034	0.287
HV A v2	0.288	4.154	4.530	9.225	0.936	3.751	0.119	0.994
HV A v3	0.299	4.319	4.307	9.624	0.890	3.913	0.031	0.258
HV A HGO	0.205	6.404	4.411	7.123	0.819	2.912	0.368	6.767
HV B v1	0.166	2.395	5.727	3.382	1.183	1.375	0.139	1.156
HV B v2	0.166	2.398	5.977	3.432	1.235	1.395	0.039	0.324
HV B v3	0.177	2.554	5.605	3.329	1.158	1.353	0.120	0.999
HV B HGO	0.100	2.113	4.126	5.506	0.555	1.655	0.250	3.847
HV C v1	0.254	3.673	4.463	1.995	0.922	0.811	1.550	12.899
HV C v2	0.241	3.476	4.357	2.287	0.900	0.930	1.519	12.640
HV C v3	0.242	3.496	4.515	2.156	0.933	0.877	1.548	12.886
HV C HGO	0.126	4.098	3.897	6.689	0.481	1.750	0.216	4.067
HV D v1	0.213	3.075	3.458	3.591	0.714	1.460	0.087	0.727
HV D v2	0.207	2.989	3.479	3.641	0.719	1.480	0.062	0.515
HV D v3	0.213	3.070	3.473	3.446	0.718	1.401	0.061	0.511
HV D HGO	0.108	3.618	3.043	5.041	0.458	1.684	0.206	3.915

(b) Partial error surrogate								
Case	$a$	$b$	$a_f$	$b_f$	$a_s$	$b_s$	$a_{fs}$	$b_{fs}$
HV A v1	0.319	4.605	3.819	9.279	0.789	3.773	0.619	5.150
HV A v2	0.313	4.516	3.977	8.964	0.822	3.645	0.660	5.489
HV A v3	0.270	3.897	4.046	9.684	0.836	3.937	0.586	4.875
HV A HGO	0.205	6.404	4.411	7.123	0.819	2.912	0.368	6.767
HV B v1	0.105	1.521	4.564	4.942	0.943	2.009	0.515	4.287
HV B v2	0.139	2.007	4.958	4.018	1.024	1.634	0.452	3.764
HV B v3	0.076	1.101	4.274	5.112	0.883	2.078	0.628	5.227
HV B HGO	0.100	2.113	4.126	5.506	0.555	1.655	0.250	3.847
HV C v1	0.246	3.555	4.536	2.121	0.937	0.862	1.433	11.930
HV C v2	0.238	3.432	4.398	2.335	0.909	0.950	1.550	12.900
HV C v3	0.239	3.446	4.407	2.225	0.910	0.905	1.509	12.562
HV C HGO	0.126	4.098	3.897	6.689	0.481	1.750	0.216	4.067
HV D v1	0.201	2.897	3.421	3.734	0.707	1.518	0.175	1.457
HV D v2	0.087	1.258	2.534	6.858	0.523	2.789	0.304	2.531
HV D v3	0.194	2.797	3.534	3.821	0.730	1.554	0.130	1.080
HV D HGO	0.108	3.618	3.043	5.041	0.458	1.684	0.206	3.915