Bayesian Multiobjective Optimisation With Mixed Analytical and Black-Box Functions: Application to Tissue Engineering

Simon Olofsson, Mohammad Mehrian, Roberto Calandra, Liesbet Geris, Marc Peter Deisenroth, and Ruth Misener

***Abstract*—Tissue engineering and regenerative medicine looks at improving or restoring biological tissue function in humans and animals. We consider optimising neotissue growth in a three-dimensional scaffold during dynamic perfusion bioreactor culture, in the context of bone tissue engineering. The goal is to choose design variables that optimise two conflicting objectives, first, maximising neotissue growth and, second, minimising operating cost. We make novel extensions to Bayesian multiobjective optimisation in the case of one analytical objective function and one black-box, i.e. simulation based and objective function. The analytical objective represents operating cost while the black-box neotissue growth objective comes from simulating a system of partial differential equations. The resulting multiobjective optimisation method determines the tradeoff between neotissue growth and operating cost. Our method exhibits better data efficiency than genetic algorithms, i.e. the most common approach in the literature, on both the tissue engineering example and standard test functions. The multiobjective optimisation method applies to real-world problems combining black-box models with easy-to-quantify objectives such as cost.**

***Index Terms*—Bayesian optimisation, black-box optimisation, multi-objective optimisation, tissue engineering.**

I. Introduction

MODERN developments in obtaining patient-specific genetic characteristics [1] open up the possibilities of personalised healthcare, which is expected to greatly improve the quality and cost-effectiveness of healthcare, as it offers more personalised and targeted therapies [2]. Bone tissue engineering develops methods for healing, improving or replacing damaged bone tissue. Engineered skeletal tissue has the potential to treat bone fractures or regenerate tissue after surgical bone removal [3].

The keys for successful *in vivo* or *in vitro* formation of bone tissue [4], [5] are (i) a vibrant cell population with predictable *in vitro* and *in vivo* behaviour, (ii) carriers and scaffolds for the cells to grow on, (iii) growth factors that control what cell types are created, (iv) proper mechanical conditions that do not put too much stress on the cells, (v) sufficient blood supply to carry oxygen to the cells and remove waste products. Ensuring that–or understanding when–all key elements are in place is a major difficulty in tissue engineering. Neotissue (cells and their extracellular matrix) generated *in vitro* has been shown to hold great potential for use as scaffold material in tissue engineering [6].

Bioreactors play an important role in clinical applications of tissue engineering because bioreactors provide controllable environments for large-scale production of cells or neotissue. Standardising bioreactor processes may reduce the cost of neotissue production, a necessary condition for the future commercial viability of tissue-engineered products [3], [7].

In silico models help address reproducibility and quality control issues in *in vitro* experiments [8]–[15] by revealing biological insight into neotissue growth inside the bioreactor, and predicting bioreactor settings optimising those processes. But *in silico* models often consist of solving systems of differential equations with implementations that may be specialised to specific hardware or software. We treat these *in silico* models as black boxes for the purpose of optimisation when we can simulate reactor conditions but cannot efficiently derive gradient or Hessian information. Relevant black-box optimisation methods include genetic algorithms [16], trust-region methods [17], and Bayesian optimisation [18]–[21]. Black-box methods give flexibility with respect to the black-box contents. If there are no restrictions on the black box, the optimisation methods are equally relevant to *in vitro* experiments as legacy computer code.

It is often difficult to define what constitutes an optimal solution to a problem as there may be multiple, conflicting objectives. The goal of the optimisation procedure is to find a good trade-off between costs, e.g. cost of materials, time and negative side-effects, and rewards, e.g. successful neotissue growth.

Manuscript received April 4, 2018; accepted June 30, 2018. Date of publication July 19, 2018; date of current version February 18, 2019. This work was supported in part by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant 675251, in part by the EPSRC Research Fellowship EP/P016871/1 to R.M., and in part by the EPSRC EP/M028240/1 *(Corresponding author: Ruth Misener.)*

S. Olofsson and M. P. Deisenroth are with the Department of Computing, Imperial College London.

M. Mehrian and L. Geris are with the Biomechanics Research Unit, GIGA-In silico Medicine, Université de Liège, and also with Prometheus, the Division of Skeletal Tissue Engineering, KU Leuven.

R. Calandra is with the Department of Electrical Engineering, University of California, Berkeley.

R. Misener is with the Department of Computing, Imperial College London, London SW7 2AZ, U.K. (e-mail: r.misener@imperial.ac.uk).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2018.2855404

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/

Multi-objective optimisation (MOO) methods explicitly find the trade-offs between conflicting objectives [22], [23].

This paper considers optimising design variables for neotissue growth on a three-dimensional scaffold in a bioreactor setting. The problem has two conflicting objectives: (i) maximising neotissue growth and (ii) minimising operating cost. The first objective function is an expensive-to-evaluate black box. The second objective function is analytical and cheap to evaluate. Combining analytical and black-box methods is an active area of research [24], [25]. This manuscript introduces a method for finding solutions to the MOO problem using a novel extension of Bayesian MOO methods that exploits the mixed analytical and black-box natures of the objective functions. We compare this method to random search and a variety of genetic algorithms (the most common approach in literature to solving MOO problems), and demonstrate that our method performs better (in terms of the required number of function evaluations) for a collection of test problems as well as for the tissue engineering application. We show how the MOO results can guide tissue engineering practice.

A limited study on the acquisition function in Sec. II-A was previously published at a conference [26]. The contribution of this article is the study of additional acquisition functions, more significant experimental results, and discussions of the applications and implications of our method.

II. Background

In the following section, we provide background on (A) the tissue engineering application and multi-objective problem, (B) Gaussian process regression and Bayesian optimisation, and (C) multi-objective optimisation.

A. Objective Functions: Tissue Engineering Application

We consider maximising the percentage of a scaffold filled with neotissue, i.e. the *filling*, while minimising the amount of material used by the bioreactor, i.e. the *cost*.

The filling objective function is a black box, i.e. we cannot write down an analytical, differentiable expression of the filling with respect to the inputs. We have an expensive-to-evaluate *in silico* partial differential equations (PDE) model of neotissue growth on a three-dimensional scaffold in a bioreactor setting (see Fig. 1) [10]–[12].

The PDE model assumes curvature-driven neotissue growth kinetics [10] with local shear stress effects on the growth rate [12]. The scaffold geometry and the neotissue filling level determine the curvature of the neotissue surface. As illustrated in Fig. 1(c), greater curvature yields a greater growth rate and zero curvature yields zero neotissue growth velocity [12]. The medium flow through the bioreactor exerts shear stress on the neotissue. Some shear stress encourages neotissue growth, but too much shear stress inhibits it. The Navier-Stokes equation models the flow in the scaffold void space. The neotissue, modelled as a homogeneous porous medium with a given permeability, exhibits Darcy, i.e. creeping, flow. The shear stress distribution on the neotissue interface and the flow of nutrients and waste products change as the scaffold neotissue filling level changes, which in turn affects the neotissue growth kinetics.

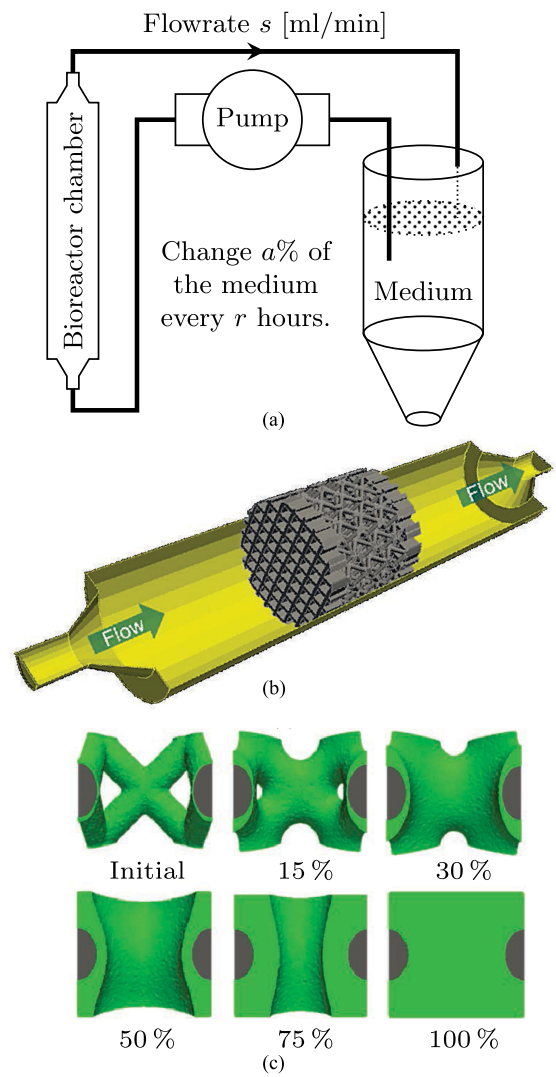

Fig. 1. (a) The experimental bioreactor setup. (b) Cross-section of the bioreactor chamber, with the scaffold placed in the centre [11]. (c) A scaffold pore under different neotissue filling levels illustrates how the neotissue growth rate increases with greater neotissue curvature [10].

The PDE model implementation currently only runs on one machine. Parallel model evaluations are impossible due to hardware constraints. Legacy *in silico* models are difficult to maintain [27], especially in research settings where long-term code maintenance typically brings few benefits to the original developer. We could replicate the existing model, but we estimate six person-months for recreating and validating the new code. Instead, we repurpose the existing model.

The PDE model considers several design variables:

- **Refreshment period:** length of time r between changing the bioreactor medium.
- **Refreshment amount:** percentage a of the medium changed every time r hours.
- **Flow rate:** rate s at which the medium is pumped through the bioreactor.

Table I summarises bounds and units for the design variables. See also Fig. 1(a). We fix the scaffold parameters, e.g. the scaffold shape and pore size [10].

TABLE I
DESIGN VARIABLES

Description	Symbol	Min.	Max.	Unit
Refreshment period	r	12	96	hours
Refreshment amount	a	0	100	per cent
Flow rate	s	0.5	2.5	ml/min

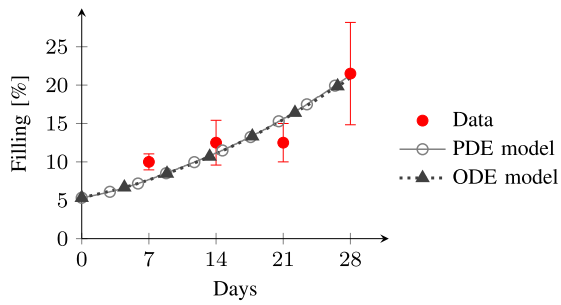

Fig. 2. Experimental data and model predictions [28].

TABLE II
HUMAN GROWTH FACTORS, THEIR APPROXIMATE COSTS (SIGMA ALDRICH WEBSITE, 8 MARCH 2017) AND APPROXIMATE CONCENTRATIONS NEEDED IN THE MEDIUM, BASED ON LITERATURE

Name	Cost	Conc.	Ref.
Transforming growth factor-β1	\$70/μg	10 ng/ml	[31]
Platelet-derived growth factor	\$25/μg	40 ng/ml	[31]
Bone Morphogenetic Protein	\$41/μg	40 ng/ml	[32]
Vascular Endothelial Growth Factor	\$29/μg	25 ng/ml	[33]
Fibroblast Growth Factor	\$17/μg	1 ng/ml	[34]

We also have a reduced-order, cheap-to-evaluate *in silico* ordinary differential equations (ODE) model [28]. The reduced model only takes the refreshment period and refreshment amount as design variables, and assumes a fixed flow rate s. One model evaluation takes on the order of milliseconds. The reduced-order model will be used for testing our optimisation method and comparing its performance to competing methods. The full PDE and reduced-order ODE models show agreement with each other and the experimental data (see Fig. 2).

The cost objective function (the cost of the bioreactor medium) scales linearly with the amount of medium used. We assume that the medium has an associated cost c_ℓ per ml, with $V \approx 10$ ml of the medium in the system at any time, and that the cost of an experiment is given solely by the amount of medium used. The medium contains growth factors inherently present in the foetal bovine serum [29], which, with its cost of around \$0.11/ml (Sigma Aldrich website, 8 March 2017), is the most expensive component.

However, foetal bovine serum is an animal product, and therefore problematic in the context of clinical applications. Chemically defined media are animal-free alternative growth media consisting of cocktails of various well-known (recombinant versions of) growth factors under well-defined concentrations, generally leading to a cost increase by at least one order of magnitude. Table II lists some of the required human growth factors [30], their approximate costs (Sigma Aldrich website, 8 March 2017) and approximations of the concentrations needed, based on growth factor concentrations studied in [31]–[34]. Often the growth factor concentrations used in *in vitro* experiments will be at least a factor of 2 higher. The scaffold shown in Fig. 1 is 6 mm in diameter and 6 mm in length [11], and one experiment can require more than 100 ml of medium. Thus, the cost of materials for growing a human bone graft of less than 1 cm^3 could be hundreds of dollars. This expense clearly motivates using optimisation to mitigate high costs.

The PDE and ODE models simulate 504-hour (21 days) experiments. Given a refreshment period r, the medium is changed a discrete number ($\lfloor 504/r \rfloor$) of times. The cost of one experiment is the sum of the initial cost $V \cdot c_\ell$ and the cost to change the medium by the refreshment amount a. In total, $c(r, a) = V \cdot c_\ell \cdot (1 + \lfloor 504/r \rfloor \cdot a)$. For simplicity, the constant $V \cdot c_\ell$ is ignored, such that $c(r, a) = 1 + \lfloor 504/r \rfloor \cdot a \in [1, 43]$.

We could create a single-objective optimisation problem by setting a monetary value on the neotissue produced in the bioreactor, and optimise the design variables by maximising the difference between the monetary reward of the neotissue and the production cost, e.g. as in growing red blood cells [35], [36]. This would be equivalent to scalarisation, which is described later in this section. However, the commercial viability of bone tissue engineering depends on growing high-quality bone grafts. Therefore, we also want to maximise the filling of the scaffold. Hence, an MOO approach using a Pareto frontier is more appropriate.

B. Bayesian Optimisation and Gaussian Processes

We begin by considering the minimisation of a single objective function f. Bayesian optimisation is a global black-box optimisation method [18]. It uses function evaluations to construct a probabilistic objective function surrogate model, which exhibits a higher level of certainty in well-explored regions than in unexplored regions. Bayesian optimisation does not require the objective function gradient with respect to its input. Using a surrogate model, we can compute the optimal choice for the next function evaluation point. This makes Bayesian optimisation more data efficient, i.e. it requires fewer function evaluations, than many other optimisation methods. It is therefore practical to use with objective functions that are expensive to evaluate. Bayesian optimisation has proven very useful, e.g. in robotics [37], sensor networks [38] and biology [39].

The typical surrogate model is a Gaussian process. A Gaussian process is a collection of random variables, any finite subset of which are jointly Gaussian distributed [40]. The random variables are the values of the objective function. We place a Gaussian process prior $\mathcal{GP}(\mu(\cdot), k(\cdot,\cdot))$ on the unknown function f, where $\mu(\cdot)$ and $k(\cdot,\cdot)$ are the mean and covariance functions, respectively. The Gaussian process prior is fully specified by $\mu(\cdot)$ and $k(\cdot,\cdot)$.

The Gaussian process prior implies that any finite number of function values $\boldsymbol{f} = [f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_n)]^\top$ at points $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top$ are jointly Gaussian distributed with the function value $f(\boldsymbol{x}_*)$ at a query point $\boldsymbol{x}_*$:

$$\begin{bmatrix} \boldsymbol{f} \\ f(\boldsymbol{x}_*) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu} \\ \mu(\boldsymbol{x}_*) \end{bmatrix}, \begin{bmatrix} \boldsymbol{K} & \boldsymbol{k}_* \\ \boldsymbol{k}_*^\top & k(\boldsymbol{x}_*, \boldsymbol{x}_*) \end{bmatrix} \right),$$

where $[\boldsymbol{\mu}]_i = \mu(\boldsymbol{x}_i)$, $[\boldsymbol{k}_*]_i = k(\boldsymbol{x}_i, \boldsymbol{x}_*)$ and $[\boldsymbol{K}]_{ij} = k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. Assume we have observed the function values $\boldsymbol{f}$ at locations $\boldsymbol{X}$. Given training data $\boldsymbol{f}$, $\boldsymbol{X}$ and an unseen test location $\boldsymbol{x}_*$, Gaussian process regression uses Bayes' theorem to compute the posterior predictive distribution $f(\boldsymbol{x}_*)|\boldsymbol{f}, \boldsymbol{X} \sim \mathcal{N}(\mu_*, \sigma_*^2)$, where $\mu_* = \mu(\boldsymbol{x}_*) + \boldsymbol{k}_*^\top \boldsymbol{K}^{-1}(\boldsymbol{f} - \boldsymbol{\mu})$ and $\sigma_*^2 = k(\boldsymbol{x}_*, \boldsymbol{x}_*) - \boldsymbol{k}_*^\top \boldsymbol{K}^{-1}\boldsymbol{k}_*$. If we have a noisy process, such that we observe $y_i = f(\boldsymbol{x}_i) + \eta_i$, where $\eta_i \sim \mathcal{N}(0, \sigma_\eta^2)$ are i.i.d. noise terms, then the posterior predictive mean and variance are $\mu_* = \mu(\boldsymbol{x}_*) + \boldsymbol{k}_*^\top(\boldsymbol{K} + \sigma_\eta^2 \boldsymbol{I})^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$ and $\sigma_*^2 = k(\boldsymbol{x}_*, \boldsymbol{x}_*) - \boldsymbol{k}_*^\top(\boldsymbol{K} + \sigma_\eta^2 \boldsymbol{I})^{-1}\boldsymbol{k}_*$, respectively.

We assume the black-box objective functions are smooth. A common choice of covariance function k for smooth latent functions is the radial basis function with automatic relevance detection (RBF-ARD) [40]:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \rho^2 \exp\left(-\tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{x}')^\top \boldsymbol{\Lambda}^{-1}(\boldsymbol{x} - \boldsymbol{x}')\right) ,$$

where ρ^2 is the signal variance and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1^2, \ldots, \lambda_D^2)$, $\boldsymbol{x} \in \mathbb{R}^D$, is a diagonal matrix of squared length scales λ_d. The *hyperparameters* $(\rho^2, \boldsymbol{\Lambda}, \sigma_\eta^2)$ are commonly learnt by maximising the marginal likelihood $p(\boldsymbol{y} \mid \boldsymbol{X}, \rho^2, \boldsymbol{\Lambda}, \sigma_\eta^2)$.

One computational bottleneck of Gaussian process models lies in reoptimising the hyperparameters after new observations, and inverting the matrix $\boldsymbol{K}$ (or $\boldsymbol{K} + \sigma_\eta^2 \boldsymbol{I}$), e.g. by using its Cholesky decomposition [40]. Retraining the model scales as $\mathcal{O}(n^3)$. Only adding an additional observation without reoptimising the hyperparameters scales as $\mathcal{O}(n^2)$. Computing the Gaussian process posterior predictive distribution for M test points also requires several matrix-vector multiplications and scales as $\mathcal{O}(Mn^2)$. Sec. IV will discuss how our proposed methodology depends on the trade-off between evaluating the black-box function, evaluating the posterior distribution, and updating the Gaussian process surrogate model.

In Bayesian optimisation, the optimal next point to query (in the search for the minimum of the objective function) is determined by *maximising* an acquisition function [41]. The acquisition function has a built-in trade-off between exploration and exploitation, i.e. trying to find potential optima in unexplored regions versus exploiting already existing information about where a minimum might be.

Assume the conditional distribution $f(\boldsymbol{x}_*)|\boldsymbol{y}, \boldsymbol{X} \sim \mathcal{N}(\mu_*, \sigma_*^2)$ is given for a point $\boldsymbol{x}_*$ and let $\tilde{\mu} = (y_{\min} - \mu_*)/\sigma_*$, where $y_{\min} = \min_i y_i$ is the best value of f observed so far. Common acquisition functions for single-objective Bayesian optimisation are confidence bounds [42], probability of improvement [18] $P(f(\boldsymbol{x}_*) < y_{\min}) = \Phi(\tilde{\mu})$, and expected improvement [43] $\mathbb{E}[y_{\min} - f(\boldsymbol{x}_*)] = \sigma_*(\tilde{\mu}\Phi(\tilde{\mu}) + \phi(\tilde{\mu}))$, where $\phi(\cdot)$ is the zero-mean Gaussian probability density function with unit variance and $\Phi(\cdot)$ the corresponding cumulative distribution function. Fig. 3 shows an example of the expected improvement acquisition function, where the posterior predictive distributions are computed given three observations of an unknown objective function.

We optimise the acquisition function through multi-start gradient descent: The acquisition function is evaluated at a large number, in our case 10^6, of random initial locations. Out of the initial set of locations, the locations that generated the best acquisition function values form a smaller subset of locations that are chosen as starting locations for gradient descent. The objective function is evaluated at the location of the optimum found through optimising the acquisition function, after which we update the Gaussian process surrogate model with the new data point.

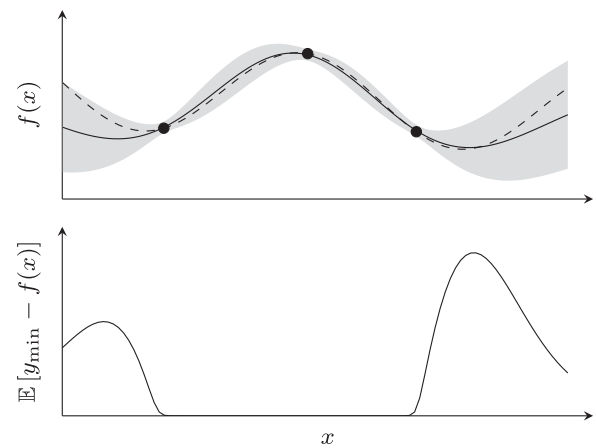

Fig. 3. The top graph shows the posterior predictive distribution at each point, with posterior mean (solid line) and two standard deviations, given three observations (dots) of an unknown function (dashed line). The bottom graph shows the expected improvement acquisition function, which is maximised in order to find the optimal next query point in the search for the minimum of the objective function.

C. Multi-Objective Optimisation

MOO aims to compromise multiple conflicting objectives. Assume a D-dimensional input space and n_f conflicting objective functions $f_i : \mathbb{R}^D \to \mathbb{R}$, $i = 1, \ldots, n_f$, where we are interested in finding the optimal input $\arg\min_{\boldsymbol{x}}\{f_i(\boldsymbol{x})\}$. There are two different ways of finding an optimal trade-off between the objective functions [22]: Scalarisation, where the trade-off is made *a priori*, and the Pareto method, where the trade-off is made *a posteriori*.

1) Scalarisation: The first method for dealing with multiple, conflicting objectives is to create a single, aggregated objective function f_s. A common example is a weighted sum[1] $f_s(x) = \sum_i \omega_i f_i(x)$ [44]. This function can then be minimised to find a trade-off solution to the original problem. Scalarisation is easy to implement. But choosing the weight coefficients $\{\omega_i\}$ introduces a new problem as they have to be chosen *a priori*, often with incomplete understanding of the system. Additionally, an optimal solution for a specific set of weights yields little or no information about other possible optimal solutions for different sets of weights. This approach has been used in previous MOO work integrating analytical and black-box functions [45].

2) Pareto Method: This method aims to finds the frontier of Pareto-optimal points in objective space. Pareto-optimality is a state in which one objective function cannot be improved without impairing another [46]. The Pareto frontier is commonly

[1]Also known as a linear combination, or convex combination if $\sum_i \omega_i = 1$, $\forall_i \omega_i \geq 0$).

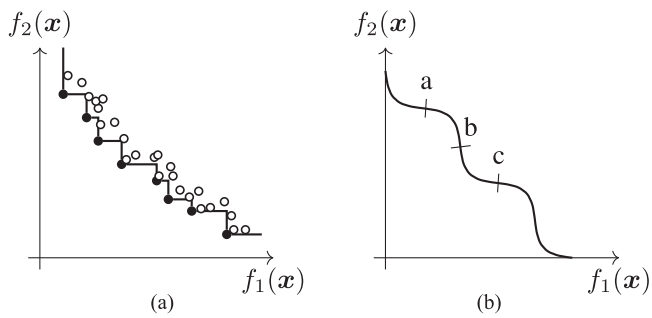

Fig. 4. (a) Approximate Pareto frontier (line) given by the non-dominated observations (black dots). (b) Convexity and concavity in a Pareto frontier. The section a–b is concave, and the section b–c is convex.

approximated using the set of non-dominated[2] function observations (see Fig. 4(a)). Once an approximate Pareto frontier has been computed, the user decides which optimum, i.e. what trade-off, best suits the particular application.

Weighted sum scalarisation suffers from the limitation that it can only converge on trade-offs lying on convex[3] sections of the Pareto frontier (see Fig. 4(b)) [47]. If the entire Pareto frontier is concave, only the extreme points at the boundaries can be found using weighted sum scalarisation.

A problem with the Pareto method is the difficulty to visualise the Pareto frontier for more than three objective functions ($n_f > 3$), which makes it difficult for the user to choose an optimal trade-off. An advantage of the Pareto method over scalarisation is that restrictions imposed on the range of one objective function immediately translate into restrictions on the range of optimal values that can be achieved for the other objective(s). This paper uses the Pareto method.

III. Method

The neotissue growth application has two conflicting objectives: maximising filling, i.e. minimising *negative* filling, and minimising cost. We can compute the probability of an input $\boldsymbol{x}$ resulting in an observation $\boldsymbol{y} = (y_1, y_2)$. The cost function $f_1(\boldsymbol{x})$ is deterministic, whereas the filling function $f_2(\boldsymbol{x}) \sim \mathcal{N}(\mu_2(\boldsymbol{x}), \sigma_2^2(\boldsymbol{x}))$ is modelled by a Gaussian process. This means the joint probability is $p(\boldsymbol{y}|\boldsymbol{x}) = \delta(y_1 - f_1(\boldsymbol{x})) \cdot p(y_2|\boldsymbol{x})$, where $\delta(\cdot)$ is the Dirac delta function.

For notational convenience we drop the dependency on $\boldsymbol{x}$ of f_1, μ_2 and σ_2^2 from here on. We also define the functions

$$\hat{\phi}(\gamma_1, \gamma_2) = \phi\left(\frac{\gamma_1 - \mu_2}{\sigma_2}\right) - \phi\left(\frac{\gamma_2 - \mu_2}{\sigma_2}\right),$$

$$\hat{\Phi}(\gamma_1, \gamma_2) = \Phi\left(\frac{\gamma_1 - \mu_2}{\sigma_2}\right) - \Phi\left(\frac{\gamma_2 - \mu_2}{\sigma_2}\right),$$

[2]Given two observations $\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)} \in \mathbb{R}^{n_f}$, we say that $\boldsymbol{y}^{(1)}$ is dominated by $\boldsymbol{y}^{(2)}$ if for all $j = 1, \ldots, n_f$ it holds that $y_j^{(1)} \geq y_j^{(2)}$, and for at least one j it holds that $y_j^{(1)} > y_j^{(2)}$.

[3]For the case of two objective functions ($n_f = 2$), the Pareto frontier is convex for sections with a positive second derivative $\mathrm{d}^2 f_2 / \mathrm{d} f_1^2 > 0$, and concave for sections with $\mathrm{d}^2 f_2 / \mathrm{d} f_1^2 < 0$.

TABLE III
Summary of Notation

Symbol	Description	Symbol	Description
f_i	Objective function i	n_f	No. of objective functions
$\boldsymbol{x}$	Input, point in variable space	$\boldsymbol{y}$	Observation, point in objective space
$\boldsymbol{r}_j$	Non-dominated observation j	$r_{j,i}$	Dim. i of non-dominated observation j
$\mathcal{P}_\alpha$	Approximated Pareto frontier	n_p	No. of non-dominated observations, $\lvert\mathcal{P}_\alpha\rvert$
μ	Mean	σ^2	Variance
$\mathcal{N}$	Gaussian distribution	$p(\boldsymbol{y}\|\boldsymbol{x})$	Probability of $\boldsymbol{y}$ given $\boldsymbol{x}$
$\phi(\cdot)$	Zero-mean, unit variance Gaussian probability density function	$\Phi(\cdot)$	Zero-mean, unit variance Gaussian cumulative distribution function
$\delta(\cdot)$	Dirac delta function	$\mathbb{1}_S$	1 if S is true, 0 otherwise

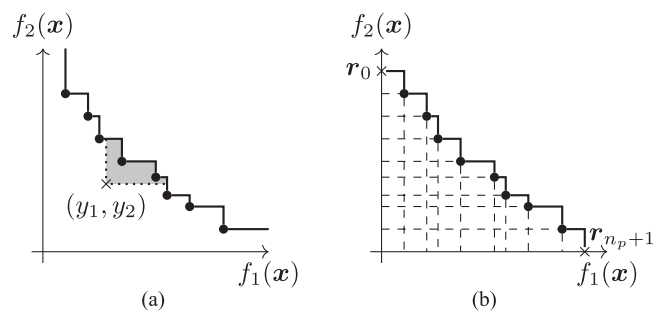

Fig. 5. Expected hyper-volume improvement, EHVI. (a) Area improvement (grey region) a new observation (y_1, y_2) would yield. (b) Division of region below the approximated Pareto frontier into rectangles C_{ij} to integrate over in Eq. (2).

as differences between zero-mean, unit variance probability density and cumulative distribution functions, respectively.

We assume that an approximated Pareto frontier $\mathcal{P}_a = \{\boldsymbol{r}_j\} = \{(r_{j,1}, r_{j,2})\}$ from n_p non-dominated observations is given, and that $\mathcal{P}_\alpha$ is sorted, i.e. $r_{j,1} < r_{j+1,1}$ and $r_{j,2} > r_{j+1,2}$ for all $j = 1, \ldots, n_p$.

There are many different acquisition functions for Bayesian MOO, e.g. expected hyper-volume improvement (EHVI) [48], expected maximin improvement (EMmI) [49], probability of improvement and expected Euclidean improvement [50], predictive entropy search [51], and uncertainty reduction [52]. This section presents novel closed-form expressions, given one black-box and one analytical objective function, for the EHVI and EMmI acquisition functions. Similar expressions may be derived for other acquisition functions. Table III summarises some mathematical notation.

A. Expected Hyper-Volume Improvement (EHVI)

The EHVI [48] $\mathbb{E}_{\mathrm{HI}}(\boldsymbol{x})$ for input $\boldsymbol{x}$ is defined as:

$$\mathbb{E}_{\mathrm{HI}}(\boldsymbol{x}) = \int I_V(\boldsymbol{y}, \mathcal{P}_\alpha) p(\boldsymbol{y}|\boldsymbol{x}) \mathrm{d}\boldsymbol{y}, \tag{1}$$

where the function $I_V(\boldsymbol{y}, \mathcal{P}_\alpha) = \mathrm{Vol}(\mathcal{P}_\alpha \cup \boldsymbol{y}) - \mathrm{Vol}(\mathcal{P}_\alpha)$ is the Lebesgue volume improvement that a point $\boldsymbol{y}$ would yield to the area bounded by $\mathcal{P}_\alpha$ (see Fig. 5(a)), calculated with respect to

some reference point. Note that $I_V(\boldsymbol{y}, \mathcal{P}_\alpha) = 0$ if $\boldsymbol{y}$ is dominated by any point in $\mathcal{P}_\alpha$. For the two-dimensional case, Eq. (1) is the expected area improvement.

We divide the region below $\mathcal{P}_\alpha$ into rectangles C_{ij} defined by corners $(r_{i,1}, r_{j,2})$ and $(r_{i+1,1}, r_{j+1,2})$, as in Fig. 5(b). Surrogate points $\boldsymbol{r}_0$ and $\boldsymbol{r}_{n_p+1}$ act as rectangle corner points on the available objective space. The reference point for calculating $\mathrm{Vol}(\cdot)$ becomes $(r_{n_p+1,1}, r_{0,2})$.

Now the EHVI in Eq. (1) can be rewritten as a sum over all rectangles, i.e.:

$$\mathbb{E}_{\mathrm{HI}}(\boldsymbol{x}) = \sum_{i=0}^{n_p} \sum_{j=i}^{n_p} \int_{\boldsymbol{y} \in C_{ij}} I_V(\boldsymbol{y}, \mathcal{P}_\alpha) p(\boldsymbol{y}|\boldsymbol{x}) \mathrm{d}\boldsymbol{y} . \tag{2}$$

Since $\boldsymbol{y} \in C_{ij}$ is in the region below $\mathcal{P}_\alpha$, it will never be dominated by any point in $\mathcal{P}_\alpha$, but will instead dominate existing observations. The set of non-dominated points in $(\mathcal{P}_\alpha \cup \boldsymbol{y})$ is $\{\boldsymbol{r}_0, \ldots, \boldsymbol{r}_i, \boldsymbol{y}, \boldsymbol{r}_{j+1}, \ldots, \boldsymbol{r}_{n_p+1}\}$. Now let $V_n = \sum_{k=1}^{n} (r_{n_p+1,1} - r_{k,1})(r_{k-1,2} - r_{k,2})$ denote the area bounded by the first n observations in $\mathcal{P}_\alpha$, with $V_0 = 0$, and $V_{n_p} = V_{n_p+1} = \mathrm{Vol}(\mathcal{P}_\alpha)$. This way, $\mathrm{Vol}(\mathcal{P}_\alpha \cup \boldsymbol{y})$ becomes:

$$\begin{aligned}\mathrm{Vol}(\mathcal{P}_\alpha \cup \boldsymbol{y}) = \mathrm{Vol}(\mathcal{P}_\alpha) &+ V_i - V_{j+1} \\ &+ (r_{n_p+1,1} - y_1)(r_{i,2} - y_2) \\ &+ (r_{n_p+1,1} - r_{j+1,1})(y_2 - r_{j+1,2}) ,\end{aligned}$$

and we see that the area improvement $I_V(\boldsymbol{y}, \mathcal{P}_\alpha)$ for $\boldsymbol{y} \in C_{ij}$ can be rewritten as:

$$\begin{aligned}I_V(\boldsymbol{y}, \mathcal{P}_\alpha) = V_i - V_{j+1} &+ (r_{n_p+1,1} - y_1)(r_{i,2} - y_2) \\ &+ (r_{n_p+1,1} - r_{j+1,1})(y_2 - r_{j+1,2}) .\end{aligned} \tag{3}$$

In our application, one objective function is deterministic, hence $p(\boldsymbol{y}|\boldsymbol{x}) = 0$ for all $y_1 \neq f_1$. Let h be the integer index such that $0 \leq h \leq n_p$ and $r_{h,1} < f_1 \leq r_{h+1,1}$, which means Eq. (2) simplifies to:

$$\mathbb{E}_{\mathrm{HI}}(\boldsymbol{x}) = \sum_{j=h}^{n_p} \int_{r_{j+1,1}}^{r_{j,1}} I_V((f_1, y_2), \mathcal{P}_\alpha) p(y_2|\boldsymbol{x}) \mathrm{d}y_2 . \tag{4}$$

Inserting the expression in Eq. (3) into Eq. (4) and integrating out y_2 yields the final expression:

$$\begin{aligned}\mathbb{E}_{\mathrm{HI}}(\boldsymbol{x}) = \sum_{j=h}^{n_p} \Big[&\hat{\Phi}(r_{j,2}, r_{j+1,2}) \Big(V_h - V_{j+1} + \mu_2 (f_1 - r_{j+1,1}) \\ &+ r_{h,2}(r_{n_p+1,1} - f_1) - r_{j+1,2}(r_{n_p+1,1} - r_{j+1,1}) \Big) \\ &- \sigma_2 (f_1 - r_{j+1,1}) \hat{\phi}(r_{j,2}, r_{j+1,2}) \Big] .\end{aligned} \tag{5}$$

Eq. (5) is differentiable with respect to f_1, μ_2 and σ_2, so we can use gradient-based optimisation to maximise the EHVI.

B. Expected Maximin Improvement (EMmI)

The modified maximin fitness function [53], developed for MOO using genetic algorithms, is defined as:

$$I_M(\boldsymbol{x}) \triangleq - \max_{j=1,\ldots,n_p} \min_{i=1,\ldots,n_f} (f_i(\boldsymbol{x}) - r_{j,i}) . \tag{6}$$

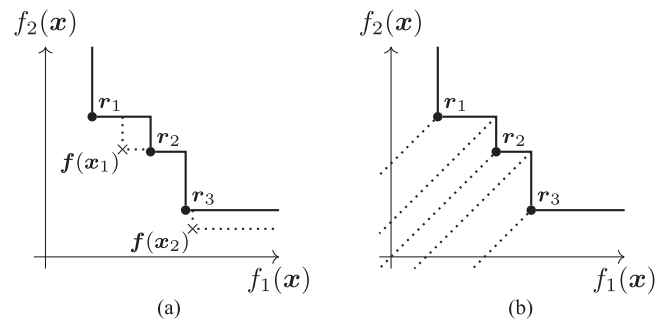

Fig. 6. Expected maximin improvement, EMmI. (a) $\mathcal{P}_\alpha$ is given by observations $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, and $\boldsymbol{r}_3$. The point $\boldsymbol{f}(\boldsymbol{x}_1) = (f_1(\boldsymbol{x}_1), f_2(\boldsymbol{x}_1))$ has a larger smallest distance to the frontier than $\boldsymbol{f}(\boldsymbol{x}_2)$, which means that $\boldsymbol{f}(\boldsymbol{x}_1)$ yields a larger maximin improvement. (b) Division of regions to integrate over in Eq. (8). The dotted lines correspond to the integration limits in Eq. (9).

It is a measure of the distance from the point $\boldsymbol{f}(\boldsymbol{x})$ to the approximated Pareto frontier, and for best performance requires that the scales of the different objective functions are comparable. Fig. 6(a) illustrates the modified maximin improvement function. For Bayesian MOO, Svensson and Santner [49] propose using the (truncated) EMmI:

$$\mathbb{E}_{\mathrm{MI}}(\boldsymbol{x}) \triangleq \begin{cases} \mathbb{E}\left[I_M(\boldsymbol{x}) | \mathcal{P}_a\right] , & \text{if } I_M(\boldsymbol{x}) \geq 0 , \\ 0 , & \text{otherwise,} \end{cases} \tag{7}$$

as the acquisition function. They show that for the case of two objective functions ($n_f = 2$), Eq. (7) is equivalent to:

$$\mathbb{E}_{\mathrm{MI}}(\boldsymbol{x}) = \sum_{i=1}^{2} \sum_{j=1}^{n_p} I(\boldsymbol{x}, i, j) , \tag{8}$$

where the function $I(\boldsymbol{x}, i, j)$ is the integral:

$$I(\boldsymbol{x}, i, j) = \int_{y_i = -\infty}^{y_i = r_{j,i}} \int_{y_{\mathrm{low}}(i)}^{y_{\mathrm{up}}(i)} (r_{j,i} - y_i) p(\boldsymbol{y}|\boldsymbol{x}) \mathrm{d}\boldsymbol{y} . \tag{9}$$

The integration limits, illustrated in Fig. 6(b), are $y_{\mathrm{up}}(i) = r_{h(i,j),k(i)} - r_{j,i} + y_i$ and $y_{\mathrm{low}}(i) = r_{j,k(i)} - r_{j,i} + y_i$, where $k(1) = 2$, $k(2) = 1$, $h(1,j) = j - 1$, $h(2,j) = j + 1$ and $r_{n_p+1,1} = r_{0,2} = \infty$.

In our setting of one analytical and one black-box objective function, the integral $I(\boldsymbol{x}, 1, j)$, where $i = 1$, simplifies to:

$$\begin{aligned}I(\boldsymbol{x}, 1, j) &= \mathbb{1}_{f_1 \leq r_{j,1}} \int_{y_{\mathrm{low}}^{(1)}}^{y_{\mathrm{up}}^{(1)}} (r_{j,1} - f_1) p(y_2|\boldsymbol{x}) \mathrm{d}y_2 \\ &= \mathbb{1}_{f_1 \leq r_{j,1}} (r_{j,1} - f_1) \hat{\Phi}(y_{\mathrm{up}}^{(1)}, y_{\mathrm{low}}^{(1)}) ,\end{aligned} \tag{10}$$

where $y_{\mathrm{up}}^{(1)} = r_{j-1,2} - r_{j,1} + f_1$ and $y_{\mathrm{low}}^{(1)} = r_{j,2} - r_{j,1} + f_1$, and $\mathbb{1}_S$ is 1 if S is true and 0 otherwise. Similarly, the

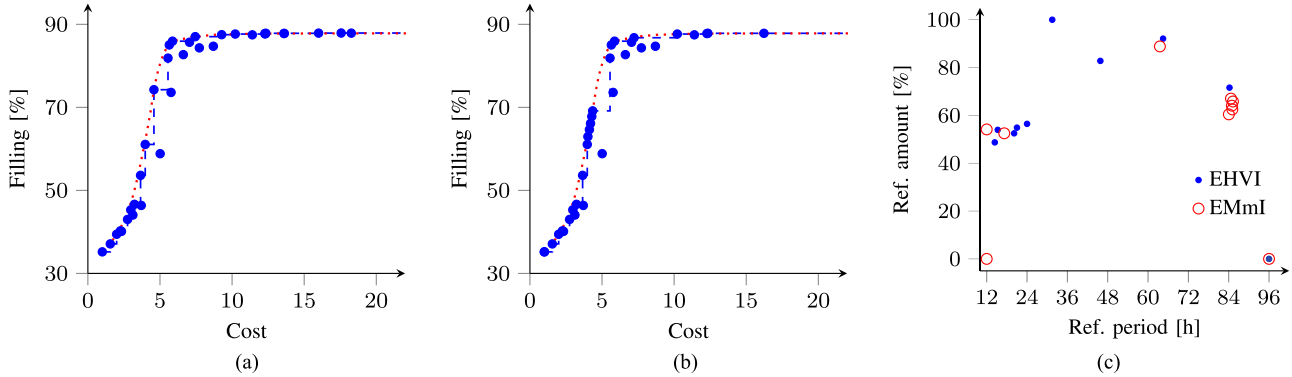

Fig. 7. Reduced-order ODE model. Plots (a) and (b) show the approximated Pareto frontiers (blue line) from observations (blue dots) using the EHVI and EMmI acquisition functions, respectively, and the true Pareto frontier (dotted red line). The same 10 random initial points were used for (a) and (b), after which 25 additional points were chosen by the respective acquisition functions. Plot (c) shows the additional 25 chosen query points in variable space.

integral $I(\boldsymbol{x}, 2, j)$, where $i = 2$, simplifies to:

$$\begin{aligned} I(\boldsymbol{x}, 2, j) &= \mathbb{1}_{f_1 \leq r_{j+1,1}} \int_{y_{\text{low}}^{(2)}}^{y_{\text{up}}^{(2)}} (r_{j,2} - y_2) p\left(y_2 | \boldsymbol{x}\right) \mathrm{d}y_2 \\ &= \mathbb{1}_{f_1 \leq r_{j+1,1}} (r_{j,2} - \mu_2) \\ &\quad \cdot \left(\hat{\Phi}(y_{\text{up}}^{(2)}, y_{\text{low}}^{(2)}) - \sigma_2^2 \hat{\phi}(y_{\text{up}}^{(2)}, y_{\text{low}}^{(2)}) \right) , \end{aligned} \tag{11}$$

where $y_{\text{up}}^{(2)} = \min(r_{j,2}\,,\, r_{j,2} - r_{j,1} + f_1)$ and $y_{\text{low}}^{(2)} = r_{j,2} - r_{j+1,1} + f_1$. Thus, we have derived a closed-form expression for the EMmI, differentiable with respect to f_1, μ_2 and σ_2.

C. Other Acquisition Functions

Similar expressions to Eq. (5), (10) & (11) can be derived for other acquisition functions, e.g. the probability of improvement and expected Euclidean improvement [50]. For two objective functions, the probability of improvement $\mathrm{PI}(\boldsymbol{x})$ is:

$$\mathrm{PI}(\boldsymbol{x}) = \sum_{i=0}^{n_p} \sum_{j=i}^{n_p} \int_{\boldsymbol{y} \in C_{ij}} p(\boldsymbol{y}|\boldsymbol{x}) \mathrm{d}\boldsymbol{y} \,.$$

For one analytical objective function f_1 and one black-box objective function $f_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$, using a similar but simplified approach to the one described in Sec. II-A, $\mathrm{PI}(\boldsymbol{x})$ becomes:

$$\mathrm{PI}(\boldsymbol{x}) = \sum_{j=h}^{n_p} \hat{\Phi}(r_{j,2}, r_{j+1,2}) = \hat{\Phi}(r_{h,2}, r_{n_p+1,2}) \,. \tag{12}$$

The probability of improvement may not encourage uniform Pareto frontier exploration since $\mathrm{PI}(\boldsymbol{x})$ does not consider the expected amount of improvement. One way to try to overcome this is to use the expected Euclidean improvement [50]:

$$\mathbb{E}_{\text{Euc.}}(\boldsymbol{x}) = \mathrm{PI}(\boldsymbol{x}) \|\hat{\boldsymbol{y}}(\boldsymbol{x}) - \boldsymbol{r}_*\|_2 \,, \tag{13}$$

where $\boldsymbol{r}_* = \arg\min_{\boldsymbol{r} \in \mathcal{P}_\alpha} \|[f_1(\boldsymbol{x}_*), \mu_2(\boldsymbol{x}_*)]^\top - \boldsymbol{r}\|_2$, with $\boldsymbol{x}_* = \arg\max \mathrm{PI}(\boldsymbol{x})$, and:

$$\hat{\boldsymbol{y}}(\boldsymbol{x}) = \frac{1}{\mathrm{PI}(\boldsymbol{x})} \sum_{i=0}^{n_p} \sum_{j=i}^{n_p} \int_{\boldsymbol{y} \in C_{ij}} \boldsymbol{y}\, p(\boldsymbol{y}|\boldsymbol{x}) \mathrm{d}\boldsymbol{y} \,.$$

For one analytical and one black-box objective function, Eq. (13) can be simplified using $\hat{y}_1(\boldsymbol{x}) = f_1/\mathrm{PI}(\boldsymbol{x})$ and:

$$\hat{y}_2(\boldsymbol{x}) = \mu_2 - \sigma_2 \left(\hat{\phi}(r_{h,2}, r_{n_p+1,2}) / \hat{\Phi}(r_{h,2}, r_{n_p+1,2}) \right) \,.$$

For predictive entropy search [51], the acquisition function is in general intractable, even for single-objective optimisation. But similar expressions combining analytical and black-box objectives may be developed for other acquisition functions.

IV. Results

This section presents the results of MOO for the neotissue growth problem with both the ODE and the PDE model. It also introduces performance metrics enabling a comparison of the EHVI and EMmI acquisition functions and competing MOO methods, for a set of MOO test problems.

A. Results for the Reduced-Order ODE Model

The reduced-order ODE model is not guaranteed to mimic the full PDE model but it is useful for testing possible algorithms before moving to the expensive-to-evaluate model [28]. Recall that the objective is to maximise filling and minimise cost. We compute approximate Pareto frontiers using the EHVI (Fig. 7(a)) and EMmI (Fig. 7(b)) acquisition function, with 10 random initialisation points after which 25 additional points are chosen by the optimisation algorithm.

Fig. 7 shows that the query points chosen by EHVI and EMmI favour different variable space regions. EHVI selects more points with higher filling and higher cost (flat part of the frontier), whereas EMmI selects more points with a medium filling and low cost (steeply increasing part of the frontier). EHVI chooses points along the flat part of the frontier because a small filling improvement may yield a large area improvement.

We compute a probabilistic Pareto frontier [54]–[56] by mapping a grid of variable points through the cost function and the filling function posterior distribution. We determine the non-dominated sampled point set by letting a point with cost $f_1^{(1)}$

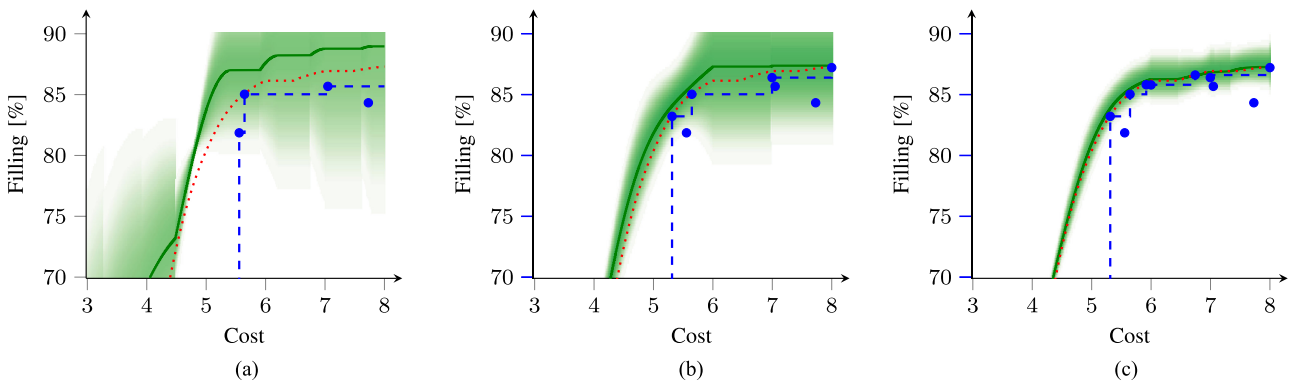

Fig. 8. Reduced-order ODE model: The true Pareto frontier (dotted red line), approximated Pareto frontier (dashed blue line) from observations (blue dots), and probabilistic approximation of the Pareto frontier with mean (green line) and some measure of its uncertainty, with (a) the initial 10 data points, and (b), (c) after additional points to query have been chosen by using the EMmI acquisition function.

and mean filling $\mu_2^{(1)}$ dominate a point with cost $f_1^{(2)}$ and mean filling $\mu_2^{(2)}$ if $f_1^{(1)} < f_1^{(2)}$ and $\mu_2^{(1)} > \mu_2^{(2)}$.

Fig. 8 shows an example evolution of a subregion of the probabilistic Pareto frontier. In this example, the optimisation starts with 10 initial data points, and further points are chosen using the EMmI acquisition function. The probabilistic Pareto frontier approaches the true Pareto frontier. As the number of data points grows, the probabilistic frontier approximates the true Pareto frontier better than the approximation from observations alone.

B. Performance Metrics

To compare algorithm performance, we use three performance metrics. The first metric is the Euclidean generational distance (GD) [57]:

$$\mathrm{GD}(\mathcal{P}_a) = \frac{1}{n_p} \sum_{i=1}^{n_p} \min_{\boldsymbol{r}^{(t)} \in \mathcal{P}_{\mathrm{true}}} \|\boldsymbol{r}_i - \boldsymbol{r}^{(t)}\|_2 \,, \tag{14}$$

which measures the average Euclidean distance between points $\boldsymbol{r}_i \in \mathcal{P}_\alpha$ to the true Pareto frontier $\mathcal{P}_{\mathrm{true}}$.

The second performance metric is the maximum Pareto frontier error (MPFE) [57]:

$$\mathrm{MPFE}(\mathcal{P}_a) = \max_{\boldsymbol{p}^{(t)} \in \mathcal{P}_{\mathrm{true}}} \min_i \|\boldsymbol{r}_i - \boldsymbol{r}^{(t)}\|_2 \,, \tag{15}$$

a measure of the largest approximation error introduced by using $\mathcal{P}_\alpha$ to approximate $\mathcal{P}_{\mathrm{true}}$, defined as the greatest Euclidean distance between any point in $\mathcal{P}_\alpha$ to the true Pareto frontier $\mathcal{P}_{\mathrm{true}}$.

The third metric is the volume ratio (VR):

$$\mathrm{VR}(\mathcal{P}_a) = -\log\left(1 - \frac{\mathrm{Vol}(\mathcal{P}_a)}{\mathrm{Vol}(\mathcal{P}_{\mathrm{true}})}\right) . \tag{16}$$

between (i) the area (or hypervolume) bounded by the approximated Pareto frontier $\mathcal{P}_\alpha$ and (ii) the area bounded by the true Pareto frontier $\mathcal{P}_{\mathrm{true}}$.

The logarithm in Eq. (16) distinguishes the methods' performance for volume ratios close to 1, which frequently occur in our tests. Increasing the ratio $\mathrm{Vol}(\mathcal{P}_a)/\mathrm{Vol}(\mathcal{P}_{\mathrm{true}})$ from 99% and 99.9% is more significant than increasing it from 19% and 19.9%. Good optimisation method performance, in terms of producing a good approximate Pareto frontier $\mathcal{P}_a$, is indicated by small GD and MPFE metrics and large VR metric. We approximate the true Pareto frontier $\mathcal{P}_{\mathrm{true}}$ for each problem through exhaustive grid search.

C. Test Problems

We additionally consider five test problems based on the following criteria:

- There should be exactly two objective functions for each problem, and at least one of them has to be analytical.
- The optimisation problems should be unconstrained, but with finite (hyper-rectangular) search domains.
- The Pareto frontiers should be different: concave, convex, both concave and convex, and discontinuous.

The first test problem is Fonzeca & Fleming's problem [58], where the conflicting objective functions are:

$$\begin{cases} f_1(\boldsymbol{x}) &= 1 - \exp\left(-\sum_{i=1}^{D}\left(x_i - 1/\sqrt{D}\right)^2\right) , \\ f_2(\boldsymbol{x}) &= 1 - \exp\left(-\sum_{i=1}^{D}\left(x_i + 1/\sqrt{D}\right)^2\right) , \end{cases} \tag{17}$$

where $x_i \in [-4, 4]$ for $i = 1, \ldots, D$, and $f_1(\boldsymbol{x}), f_2(\boldsymbol{x}) \in [0, 1]$, and we choose $D = 2$. Fonzeca & Fleming's test problem has a concave Pareto frontier. The output function bounds are used to compute hyper-volume improvement and performance metrics. The second test problem is Schaffer's two-objective problem [59]:

$$\begin{cases} f_1(x) = x^2 \,, \\ f_2(x) = (x-2)^2 \,, \end{cases}$$

where $x \in [-3, 3]$, which implies output bounds $f_1(x) \in [0, 9]$ and $f_2(x) \in [0, 25]$. Schaffer's test problem has a convex Pareto frontier. The third test problem is Kursawe's problem [60]:

$$\begin{cases} f_1(\boldsymbol{x}) = \sum_{i=1}^{D-1}\left[-10\exp\left(-0.2\sqrt{x_i^2 + x_{i+1}^2}\right)\right] , \\ f_2(\boldsymbol{x}) = \sum_{i=1}^{D}\left[|x_i|^{0.8} + 5\sin\left(x_i^3\right)\right] , \end{cases} \tag{18}$$

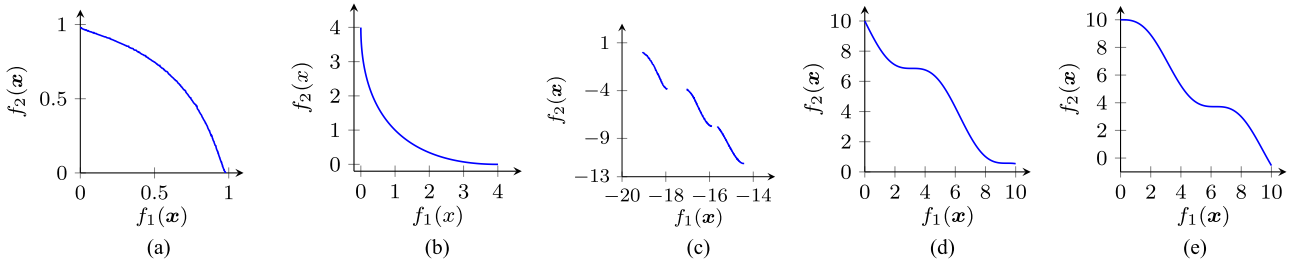

Fig. 9. True Pareto frontiers for the different test problems ($D = 2$ except where noted otherwise). (a) Fonzeca & Fleming ($D \geq 2$). (b) Schaffer. (c) Kursawe ($D = 3$). (d) S^-. (e) S^+.

which has a discontinuous Pareto frontier. The input bounds are $x_i \in [-5, 5]$ for $i = 1, \ldots, D$ and we choose $D = 3$, which yields approximate output bounds $f_1(\boldsymbol{x}) \in [-20, -4]$ and $f_2(\boldsymbol{x}) \in [-12, 25]$. The fourth and fifth test problems, constructed with the aim of a continuous Pareto frontier with both convex and concave section, are defined as:

$$\begin{cases} f_1(\boldsymbol{x}) = x_1 \,, \\ f_2(\boldsymbol{x}) = 10 - x_1 + x_2 \pm \sin(x_1) \,, \end{cases}$$

and called S^- and S^+ depending on the sign in front of the sine-term in f_2. Problem S^- has two convex sections with a concave section in-between, whereas S^+ has two concave sections with a convex section in-between. For both S^- and S^+, the two-dimensional inputs have ranges $x_1 \in [0, 10]$ and $x_2 \in [0, 2]$, which yields outputs with approximate domains $f_1(\boldsymbol{x}) \in [0, 10]$ and $f_2(\boldsymbol{x}) \in [-1, 12]$. Fig. 9 shows the Pareto frontiers for all test problems.

The acquisition functions derived in Sec. II require one analytical and one black-box objective function. Therefore, for Bayesian MOO of each test problem, we assume that f_1 is an analytical function and f_2 is a black box.

D. Comparison to Competing Methods

This section compares the EHVI and EMmI acquisition functions and competing MOO methods. The first competing method is random search, where we uniformly sample input values for which to evaluate the objective functions. The second competing method is genetic algorithms, commonly used for MOO in biomedical engineering, e.g. [61], [62]. We consider the algorithms: NSGA-II, NSGA-III, EpsMOEA, GDE3, SPEA2, MOEAD, OMOPSO, SMPSO, CMAES, IBEA, PAES, PESA2, as implemented in the Python package Platypus [63]. We use default parameter values for all genetic algorithms, with 10 outer divisions for NSGA-III and $\boldsymbol{\epsilon} = [1, 1]$ for EpsMOEA and OMOPSO.

For each test problem, as well as for the optimisation of the ODE model for neotissue growth, the Bayesian optimisation algorithms are given 10 random initial observations after which the acquisition functions run for 50 iterations. The genetic algorithms and random search run for 60 function evaluations. Performance metrics are computed after each function evaluation. This is repeated 25 times, with new random seeds for all methods, after which the average performance metrics are computed for each algorithm.

Table IV compares the average values (with standard error) of the performance metrics at initialisation (after 10 function evaluations) and after 10, 25 and 50 additional function evaluations. For each performance metric, the first two columns are the average performance metrics for EHVI and EMmI, respectively. The third column for each performance metric is the best average value (with standard error) achieved by any of the competing genetic algorithms and random search.

We see from Table IV that, with the exception of some instances in the MPFE column, at least one (and often both) of the Bayesian optimisation methods perform better than all competing methods (genetic algorithms and random search). In Fig. 10, we display the learning curves for the tissue engineering application using the reduced-order ODE model. Learning curves for the test problems can be found in the supplementary material. The curves in Fig. 10 are typical for the difference between the different MOO algorithms, in that the Bayesian MOO methods perform better than the competition for the GD and VR metrics, and similar to (or better than) the competing methods for the MPFE metric.

E. Scaling to Higher Dimensions

This section compares the Bayesian MOO methods versus the genetic algorithms for test problems of higher dimension D. Fonzeca & Fleming's Eq. (17) and Kursawe's Eq. (18) are defined for $D \geq 2$, so we compare using these test problems.

Consider $D \in \{5, \ldots, 20, 25, 30\}$. The data needed for a "good" Pareto frontier approximation may increase with D. Therefore, we choose to begin each Bayesian MOO experiment with $N_0 = 10 + D$ initial, uniformly-sampled data points and then select $N = 30 + 10D$ additional data points. The genetic algorithms and random search run for $N_0 + N$ iterations. Defining $\mathcal{P}^{\mathrm{A}}_{a,j,i}$ as the approximate Pareto frontier produced by algorithm A in experiment j using the $N_0 + i$ first observations, we compute the average metrics:

$$\overline{\mathrm{M}}_{\mathrm{A},i} = \frac{1}{J} \sum_{j=1}^{J} \mathrm{M}(\mathcal{P}^{\mathrm{A}}_{a,j,i}) \,,$$

of $J = 100$ experiments, with $\mathrm{M} \in \{\mathrm{GD}, \mathrm{MPFE}, \mathrm{VR}\}$ defined in Eqs. (14)–(16). To present the results, we redefine the metrics as the difference between the average performances of

TABLE IV
COMPARISON OF THE BAYESIAN MOO METHOD WITH EHVI AND EMmI FUNCTIONS, AND COMPETING MOO METHODS, AT THE START OF THE TESTS (10 FUNCTION EVALUATIONS) AND AFTER 10, 25, AND 50 ADDITIONAL FUNCTION EVALUATIONS. AVERAGE VALUES (WITH STANDARD ERROR) ARE SHOWN FOR THREE DIFFERENT PERFORMANCE METRICS: GD, MPFE, AND VR

		GD [×100]			**MPFE**			**VR**		
Problem	**Eval.**	EHVI	EMmI	Comp.	EHVI	EMmI	Comp.	EHVI	EMmI	Comp.
Fonzeca & Fleming [55]	10	25.82 ± 1.8	25.82 ± 1.8	15.7 ± 0.2	0.96 ± 0.0	0.96 ± 0.0	0.79 ± 0.0	0.06 ± 0.0	0.06 ± 0.0	0.15 ± 0.0
	20	$\mathbf{0.88} \pm 0.0$	1.01 ± 0.0	11.45 ± 0.1	0.77 ± 0.1	$\mathbf{0.43} \pm 0.0$	0.65 ± 0.0	0.72 ± 0.1	$\mathbf{1.02} \pm 0.1$	0.22 ± 0.0
	35	$\mathbf{0.29} \pm 0.0$	0.36 ± 0.0	7.97 ± 0.3	0.67 ± 0.1	$\mathbf{0.4} \pm 0.1$	0.49 ± 0.0	1.00 ± 0.2	$\mathbf{1.42} \pm 0.3$	0.56 ± 0.0
	60	$\mathbf{0.14} \pm 0.0$	0.18 ± 0.0	1.37 ± 0.0	0.6 ± 0.1	0.39 ± 0.1	$\mathbf{0.31} \pm 0.0$	1.23 ± 0.3	$\mathbf{1.63} \pm 0.5$	1.22 ± 0.0
Schaffer [56]	10	26174	26174	16.45 ± 0.6	265 ± 140338	265 ± 140338	1.86 ± 0.1	0.16 ± 0.3	0.16 ± 0.3	3.87 ± 0.1
	20	$\mathbf{0.4} \pm 0.0$	1.5 ± 0.2	5.67 ± 0.1	1.01 ± 0.0	$\mathbf{0.82} \pm 0.1$	1.24 ± 0.0	5.21 ± 0.0	$\mathbf{5.25} \pm 0.1$	4.63 ± 0.0
	35	$\mathbf{0.16} \pm 0.0$	0.27 ± 0.0	0.34 ± 0.0	0.52 ± 0.0	$\mathbf{0.49} \pm 0.0$	0.84 ± 0.0	6.47 ± 0.0	$\mathbf{6.50} \pm 0.1$	5.31 ± 0.0
	60	$\mathbf{0.11} \pm 0.0$	0.15 ± 0.0	0.21 ± 0.0	0.38 ± 0.0	$\mathbf{0.37} \pm 0.0$	0.54 ± 0.0	7.18 ± 0.1	$\mathbf{7.28} \pm 0.2$	6.24 ± 0.1
Kursawe [57]	10	342 ± 139	342 ± 139	391 ± 26.7	9.4 ± 1.4	9.4 ± 1.4	9.48 ± 0.8	0.73 ± 0.0	0.73 ± 0.0	0.74 ± 0.0
	20	125 ± 86.9	$\mathbf{107} \pm 67.6$	304 ± 24.9	7.76 ± 4.0	$\mathbf{7.61} \pm 2.4$	8.29 ± 0.4	1.86 ± 0.1	$\mathbf{1.88} \pm 0.0$	0.85 ± 0.0
	35	$\mathbf{26.43} \pm 6.3$	51.79 ± 25.8	240 ± 16.8	$\mathbf{4.16} \pm 3.1$	5.12 ± 3.1	7.34 ± 0.4	$\mathbf{2.58} \pm 0.3$	2.29 ± 0.1	1.02 ± 0.0
	60	$\mathbf{9.82} \pm 0.6$	18.19 ± 9.0	108 ± 14.4	$\mathbf{1.76} \pm 1.3$	3.35 ± 3.0	5.37 ± 0.6	$\mathbf{4.38} \pm 1.0$	2.92 ± 0.6	1.45 ± 0.1
S^-	10	12.39 ± 0.1	12.39 ± 0.1	24.22 ± 0.1	2.65 ± 0.6	2.65 ± 0.6	2.58 ± 0.1	1.67 ± 0.1	1.67 ± 0.1	1.47 ± 0.0
	20	$\mathbf{4.4} \pm 0.1$	6.77 ± 0.2	16.45 ± 0.0	1.57 ± 0.1	$\mathbf{1.21} \pm 0.1$	1.65 ± 0.0	$\mathbf{3.01} \pm 0.0$	2.85 ± 0.0	1.84 ± 0.0
	35	$\mathbf{1.13} \pm 0.0$	2.11 ± 0.0	12.0 ± 0.0	1.09 ± 0.0	$\mathbf{0.88} \pm 0.1$	1.36 ± 0.0	$\mathbf{4.12} \pm 0.0$	3.91 ± 0.0	2.15 ± 0.0
	60	$\mathbf{0.71} \pm 0.0$	0.99 ± 0.0	8.2 ± 0.2	$\mathbf{0.56} \pm 0.0$	0.61 ± 0.0	1.09 ± 0.0	$\mathbf{5.33} \pm 0.0$	5.08 ± 0.0	2.45 ± 0.0
S^+	10	15.44 ± 0.1	15.44 ± 0.1	23.15 ± 0.0	3.8 ± 3.5	3.8 ± 3.5	2.98 ± 0.2	1.48 ± 0.0	1.48 ± 0.0	1.48 ± 0.0
	20	$\mathbf{4.39} \pm 0.0$	6.07 ± 0.1	16.45 ± 0.0	1.55 ± 0.0	$\mathbf{1.26} \pm 0.0$	2.1 ± 0.1	$\mathbf{2.90} \pm 0.0$	2.27 ± 0.1	1.79 ± 0.0
	35	$\mathbf{0.61} \pm 0.0$	0.82 ± 0.0	12.31 ± 0.0	0.93 ± 0.0	$\mathbf{0.87} \pm 0.0$	1.53 ± 0.0	$\mathbf{4.00} \pm 0.0$	2.77 ± 0.0	2.07 ± 0.0
	60	$\mathbf{0.14} \pm 0.0$	0.17 ± 0.0	8.78 ± 0.0	0.76 ± 0.0	$\mathbf{0.64} \pm 0.0$	1.21 ± 0.0	$\mathbf{5.15} \pm 0.0$	3.00 ± 0.0	2.34 ± 0.0
Neotissue growth (ODE model [27])	10	20.77 ± 0.6	20.77 ± 0.6	14.21 ± 0.1	15.18 ± 6.6	15.18 ± 6.6	14.07 ± 1.5	3.36 ± 0.1	3.36 ± 0.1	3.52 ± 0.0
	20	$\mathbf{7.64} \pm 0.2$	8.63 ± 0.2	9.47 ± 0.0	11.85 ± 2.6	12.22 ± 6.0	$\mathbf{11.09} \pm 0.8$	$\mathbf{4.43} \pm 0.1$	4.28 ± 0.1	4.02 ± 0.0
	35	3.14 ± 0.0	$\mathbf{3.01} \pm 0.0$	6.56 ± 0.0	10.87 ± 3.9	$\mathbf{10.37} \pm 3.2$	10.5 ± 0.8	$\mathbf{5.00} \pm 0.0$	4.91 ± 0.0	4.45 ± 0.0
	60	1.85 ± 0.0	$\mathbf{1.6} \pm 0.0$	4.14 ± 0.0	11.18 ± 2.1	$\mathbf{9.99} \pm 3.9$	10.29 ± 1.2	$\mathbf{5.58} \pm 0.0$	5.48 ± 0.1	4.82 ± 0.0

Bold font denotes the best average performance.

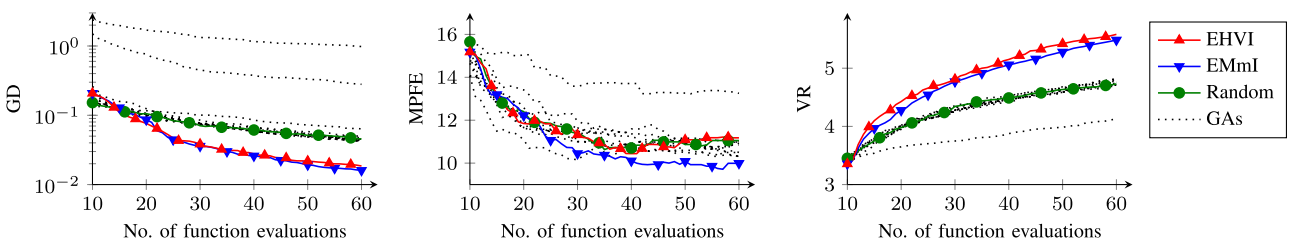

Fig. 10. Reduced-order ODE model: Performance metric learning curves for MOO using Bayesian optimisation with the EHVI and EMmI acquisition functions, random search, and genetic algorithms (GAs). A good approximate Pareto frontier has small GD and MPFE metrics and high VR metric.

algorithm A compared to random search R, summed over all iterations:

$$\underline{\mathrm{GD}}_{\mathrm{A}} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{\mathrm{GD}}_{\mathrm{R},i} - \overline{\mathrm{GD}}_{\mathrm{A},i}\right), \tag{19}$$

$$\underline{\mathrm{MPFE}}_{\mathrm{A}} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{\mathrm{MPFE}}_{\mathrm{R},i} - \overline{\mathrm{MPFE}}_{\mathrm{A},i}\right), \tag{20}$$

$$\underline{\mathrm{VR}}_{\mathrm{A}} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{\mathrm{VR}}_{\mathrm{A},i} - \overline{\mathrm{VR}}_{\mathrm{R},i}\right), \tag{21}$$

Note: larger, positive values for the metrics in Eqs. (19)–(21) indicate better performance than random search.

As we scale up the dimensionality for Fonzeca & Fleming's test problem, performance deteriorates for all algorithms. The performance of the Bayesian optimisation methods and the genetic algorithms are equally poor, and effectively indistinguishable from that of the random search.

For Kursawe's test problem (Fig. 11), the Bayesian methods outperform both random search and the genetic algorithms on metrics $\underline{\mathrm{GD}}_{\mathrm{A}}$ and $\underline{\mathrm{VR}}_{\mathrm{A}}$. Digging deeper into Kursawe's test problem, the Bayesian methods' performance stays relatively stable on $\underline{\mathrm{GD}}_{\mathrm{A}}$ and $\underline{\mathrm{VR}}_{\mathrm{A}}$ for large D, while random search and the genetic algorithms degrade with increasing dimension D. Metric $\underline{\mathrm{MPFE}}_{\mathrm{A}}$ has large standard errors, i.e. the mean fluctuates a lot, making it difficult to derive conclusions about Kursawe's test problem from $\underline{\mathrm{MPFE}}_{\mathrm{A}}$.

F. Result for PDE Model

Sec. III-D shows no large performance difference between the EHVI and EMmI acquisition functions for the test problems. We expect the full PDE model to have a concave Pareto frontier resembling that of the reduced-order ODE model in Figs. 7 & 8. Our interpretation is that EMmI performs better for the reduced-order ODE model, and that it chooses to explore less in the region with a high operating cost. EMmI also out-performed EHVI for the concave Fonzeca & Fleming test problem, so we choose the EMmI acquisition function for optimising the full, expensive-to-evaluate PDE.

The full PDE model takes the refreshment period, refreshment amount and flow rate as design parameters. There are 8 data points from past model evaluations (used e.g. for validating

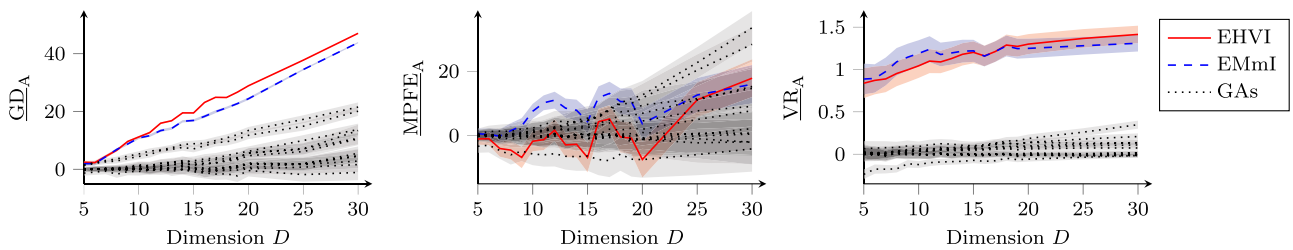

Fig. 11. Kursawe's test problem in high dimension D. Plotted are the means with standard error. Well-performing methods have larger metrics.

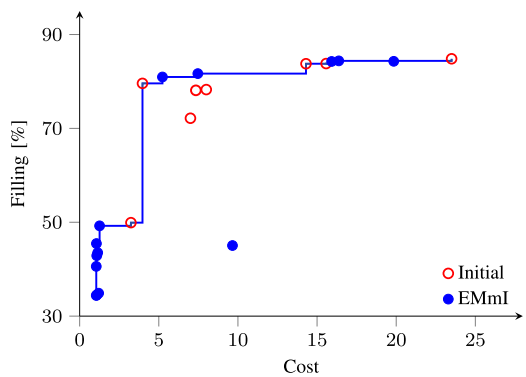

Fig. 12. Approximated Pareto frontier resulting from Bayesian MOO of the full, expensive-to-evaluate PDE model, using the EMmI acquisition function.

the reduced-order model) that we use as our initial data set. The EMmI acquisition function chooses 14 additional points to evaluate, with computations running from 12 December 2015 until 19 February 2016. As mentioned in Sec. I-A, despite our best efforts, parallel model evaluations are not possible (the server runs out of memory and the computations crash).

Fig. 12 shows the resulting approximated Pareto frontier. Our initial data set already contained a point (top left, roughly 4 units' cost and 80 % filling) on the Pareto frontier yielding a good trade-off between cost and scaffold filling. In the bottom left of Fig. 12, the surrogate model has also explored regions where it overestimated the expected filling at a low cost.

V. Discussion

Single-objective scaffold filling maximisation would lead to frequent medium replacement and therefore unrealistically high operating costs. Incorporating cost minimisation shows the trade-off between scaffold filling and culture cost. Fig. 12 clearly indicates that there is a refreshment regime beyond which the amount of filling is hardly increasing whilst the cost is increasing dramatically. Based on these *in silico* results, new experiments should be set up investigating the culture regimes that are close to the optimal point in the top left of the Pareto frontier and compare them to (new or historic) experiments on other Pareto frontier regimes.

This manuscript develops novel extensions of acquisition functions EHVI and EMmI for Bayesian MOO of one black-box and one analytical objective function. The Bayesian optimisation method outperforms twelve state-of-the-art genetic algorithms for the tissue engineering application as well as for a set of test problems. These results demonstrate successful exploitation of problems combining both black-box and analytical objective functions.

Our positive results depend on the data-efficient nature of Bayesian optimisation, which requires relatively few function evaluations. Our working assumption is that the limiting step is the computational cost of evaluating the black-box objective function, i.e. that the computational cost of training the Gaussian process surrogate model and maximising the acquisition function is comparatively negligible. When maximising the acquisition function, the major computational cost comes from the large number M of Gaussian process predictions. Maximising the acquisition function scales as $\mathcal{O}(Mn^2)$, with n the number of training data points and $M \gg n$, whereas training the surrogate model scales as $\mathcal{O}(n^3)$. Our working assumption is valid for the full, expensive-to-evaluate PDE model which inspires this paper. However, for problems where training the surrogate model and/or evaluating the posterior predictive distribution is more expensive than evaluating the black-box function, genetic algorithms and random search will be able to explore the variable space much quicker (in terms of wall clock time) than Bayesian optimisation, giving them an advantage.

Gaussian processes typically do not scale well to higher dimensions due to the curse of dimensionality – the training data needed for surrogate models may grow exponentially with the number of dimensions. But Sec. III-E shows that the Bayesian optimisation methods scaled either (i) as well (or poorly) as or (ii) better than random search and genetic algorithms with default parameter settings for two test problems with dimensions up to $D = 20$.

The Pareto frontier of the neotissue growth problem with the reduced-order ODE model has both convex and concave sections, making optimisation through scalarisation, e.g. by using a weighted sum, unattractive for the full PDE model, since scalarisation may not be able to capture all optima. The Pareto frontier also yields an understanding of the interaction between the conflicting objectives, making an *a posteriori* trade-off decision preferable to an *a priori* one.

We showed in Sec. III-A that constructing a probabilistic surrogate model of the expensive black box also enables us to construct a probabilistic Pareto frontier. This provides a measure of confidence for selecting optima in unexplored regions of the variable space. Other MOO methods, e.g. genetic algorithms

and random search, do not provide any measure of confidence for selecting unexplored optima.

Combining expensive-to-evaluate black-box models with easy-to-quantify objectives like cost is common in the life sciences and engineering. The benefit of Gaussian processes is that they act as cheap simulators and allow us to treat black boxes almost like analytical functions, whilst providing us with confidence bounds. Note that the choice of Gaussian process prior is important: Most off-the-shelf covariance functions make explicit assumptions about the smoothness of the underlying function [40], but there exist methods to overcome these assumptions, e.g. [64].

Tissue engineering has come a long way in the past decade, and tissue-engineered products are finding clinical use [65]. However, scaling up production introduces new problems, many of which have in common that they require optimisation of conflicting black-box and easy-to-quantify, analytical objectives. The Bayesian MOO method presented in this paper puts no restrictions on the contents of the black box, which makes the method equally useful for problems involving *in vitro* experiments as legacy computer code.

VI. Conclusions

We have proposed a method for exploiting one black-box and one analytical objective for Bayesian MOO, where the black-box function is expensive to evaluate. The Bayesian method outperforms competing MOO methods for a diverse set of test problems and for a tissue engineering application. Given the Gaussian process surrogate model, which we construct of the black box during the Bayesian optimisation, we infer a probabilistic Pareto frontier that discriminates optimal solutions all across the variable space, with some measure of their uncertainty. The results show that the Bayesian multi-objective optimisation method is highly applicable to real-world problems combining black-box models with easy-to-quantify objectives like cost. The acquisition functions have been implemented in python and made available on GitHub [66].

References

[1] R. Saha *et al.*, "Recent advances in the reconstruction of metabolic models and integration of omics data," *Current Opinion Biotechnol.*, vol. 29, pp. 39–45, 2014.

[2] M. Fuentes-Garí *et al.*, "A systematic framework for the design, simulation and optimization of personalized healthcare: Making and healing blood," *Comput. Chem. Eng.*, vol. 81, pp. 80–93, 2015.

[3] E. Salter *et al.*, "Bone tissue engineering bioreactors: A role in the clinic?" *Tissue Eng. B Rev.*, vol. 18, no. 1, pp. 62–75, 2012.

[4] P. V. Giannoudis *et al.*, "Fracture healing: The diamond concept," *Injury*, vol. 38S4, pp. S3–6, 2007.

[5] L. Geris, "Regenerative orthopaedics: In vitro, in vivo... in silico," *Int. Orthop.*, vol. 38, no. 9, pp. 1771–1778, 2014.

[6] H. Lu *et al.*, "Autologous extracellular matrix scaffolds for tissue engineering," *Biomaterials*, vol. 32, pp. 2489–2499, 2011.

[7] I. Martin *et al.*, "The role of bioreactors in tissue engineering," *Trends Biotechnol*, vol. 22, no. 2, pp. 80–86, 2004.

[8] C. M. Bidan *et al.*, "Geometry as a factor for tissue growth: Towards shape optimization of tissue engineering scaffolds," *Adv. Healthcare Mater.*, vol. 2, no. 1, pp. 186–194, 2013.

[9] C. M. Bidan *et al.*, "A three-dimensional model for tissue deposition on complex surfaces," *Comput. Methods Biomech. Biomed. Eng.*, vol. 16, no. 10, pp. 1056–1070, 2013.

[10] Y. Guyot *et al.*, "A computational model for cell/ECM growth on 3D surfaces using the level set method: A bone tissue engineering case study," *Biomech. Model. Mechanobiol.*, vol. 13, pp. 1361–1371, 2014.

[11] Y. Guyot *et al.*, "A three-dimensional computational fluid dynamics model of shear stress distribution during neotissue growth in a perfusion bioreactor," *Biotechnol Bioeng.*, vol. 112, pp. 2591–2600, 2015.

[12] Y. Guyot *et al.*, "Coupling curvature-dependent and shear stress-stimulated neotissue growth in dynamic bioreactor cultures: A 3D computational model of a complete scaffold," *Biomech. Model. Mechanobiol.*, vol. 15, pp. 169–180, 2016.

[13] Y. Guyot *et al.*, "Immersed boundary models for quantifying flow-induced mechanical stimuli on stem cells seeded on 3D scaffolds in perfusion bioreactors," *PLoS Comput. Biol.*, vol. 12, pp. 1–21, 2016.

[14] A. Boccaccio *et al.*, "A mechanobiology-based algorithm to optimize the microstructure geometry of bone tissue scaffolds," *Int. J. Biol. Sci.*, vol. 12, no. 1, pp. 1–17, 2016.

[15] M. A. Alias and P. R. Buenzli, "Modeling the effect of curvature on the collective behavior of cells growing new tissue," *Biophys J.*, vol. 112, no. 1, pp. 193–204, 2017.

[16] D. Ashlock, *Evolutionary Computation for Modeling and Optimization*. Berlin, Germany: Springer, 2006.

[17] C. Cartis *et al.*, "Adaptive cubic regularisation methods for unconstrained optimization. Part I: Motivation, convergence and numerical results," *Math. Prog.*, vol. 127, no. 2, pp. 245–295, 2011.

[18] H. J. Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise," *J. Basic. Eng.*, vol. 86, no. 1, pp. 97–106, 1964.

[19] J. Mockus, *Bayesian Approach to Global Optimization: Theory and Applications*. Norwell, MA, USA: Kluwer, 1989.

[20] F. Boukouvala and M. G. Ierapetritou, "Surrogate-based optimization of expensive flowsheet modeling for continuous pharmaceutical manufacturing," *J. Pharmaceutical Innovations*, vol. 8, no. 2, pp. 131–145, 2013.

[21] J. Nogueira *et al.*, "Unscented Bayesian optimization for safe robot grasping," in *Proc. Int. Conf. Intell. Robots Syst.*, 2016, pp. 1967–1972.

[22] C.-L. Hwang and A. S. M. Masud, *Multiple Objective Decision Making—Methods and Applications: A State-of-the-Art Survey*. Berlin, Germany: Springer, 1979.

[23] J. R. Banga, "Optimization in computational systems biology," *BMC Syst. Biol.*, vol. 2, no. 47, pp. 1–7, 2008.

[24] F. Boukouvala *et al.*, "Global optimization advances in mixed-integer nonlinear programming, MINLP, and constrained derivative-free optimization, CDFO," *Eur. J. Oper. Res.*, vol. 252, pp. 701–727, 2016.

[25] S. Olofsson *et al.*, "Design of experiments for model discrimination hybridising analytical and data-driven approaches," in *Proc. Int. Conf. Mach. Learn.*, Stockholm, Sweden, 2018, pp. 3905–3914.

[26] S. Olofsson *et al.*, "Bayesian multi-objective optimisation of neotissue growth in a perfusion bioreactor set-up," in *Proc. Eur. Symp. Comput. Aided Process Eng.*, 2017, pp. 2155–2160.

[27] C. I. Siettos *et al.*, "Enabling dynamic process simulators to perform alternative tasks: A time-stepper-based toolkit for computer-aided analysis," *Ind. Eng. Chem. Res.*, vol. 42, pp. 6795–6801, 2003.

[28] M. Mehrian *et al.*, "Maximizing neotissue growth kinetics in a perfusion bioreactor: An *in silico* strategy using model reduction and Bayesian optimization," *Biotechnol. Bioeng.*, vol. 115, no. 3, pp. 617–629, 2018.

[29] Y. Guyot, "A multiphysics multiscale computational framework for the simulation of perfusion bioreactor processes in bone tissue engineering," Ph.D. dissertation, Dept. Aerosp. Mech. Eng., Université de Liège, Liège, Belgium, 2015.

[30] M. E. Gomes *et al.*, "In vitro localization of bone growth factors in constructs of biodegradable scaffolds seeded with marrow stromal cells and cultured in a flow perfusion bioreactor," *Tissue Eng*, vol. 12, no. 1, pp. 177–188, 2006.

[31] T. J. Anderson *et al.*, "Effects of transforming growth factor-β and platelet-derived growth factor on human gingival fibroblasts grown in serum-containing and serum-free medium," *J. Clin. Periodontol.*, vol. 25, pp. 48–55, 1998.

[32] G. Mundy *et al.*, "Stimulation of bone formation in vitro and in rodents by statins," *Science*, vol. 286, no. 5446, pp. 1946–1949, 1999.

[33] Y. S. Chang *et al.*, "Effect of vascular endothelial growth factor on cultured endothelial cell monolayer transport properties," *Microvascular Res.*, vol. 59, pp. 265–277, 2000.

[34] T. L. Hebert *et al.*, "Culture effects of epidermal growth factor (EGF) and basic fibroblast growth factor (bFGF) on cryopreserved human adipose-derived stromal/stem cell proliferation and adipogenesis," *J. Tissue Eng. Regenerative Med.*, vol. 3, no. 7, pp. 553–561, 2009.

[35] R. Misener *et al.*, "Global superstructure optimisation of red blood cell production in a parallelised hollow fibre bioreactor," *Comput. Chem. Eng.*, vol. 71, pp. 532–553, 2014.

[36] R. Misener *et al.*, "Stem cell biomanufacturing under uncertainty: A case study in optimizing red blood cell production," *AIChE J.*, vol. 64, pp. 3011–3022, 2017.

[37] R. Calandra *et al.*, "Bayesian optimization for learning gaits under uncertainty," *Ann. Math. Artif. Intell.*, vol. 76, no. 1, pp. 5–23, 2016.

[38] R. Garnett *et al.*, "Bayesian optimization for sensor set selection," in *Proc. Int. Conf. Inf. Process. Sensor Netw.*, 2010, pp. 209–219.

[39] D. Ulmasov *et al.*, "Bayesian optimisation with dimension scheduling algorithm: Application to biological systems," in *Proc. Eur. Symp. Comput. Aided Process Eng.*, 2016, pp. 1051–1056.

[40] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[41] B. Shahriari *et al.*, "Taking the human out of the loop: A review of Bayesian optimization," *Proc IEEE*, vol. 104, no. 1, pp. 148–175, Jan. 2016.

[42] P. Auer, "Using confidence bounds for exploitation-exploration trade-offs," *J. Mach. Learn. Res.*, vol. 3, pp. 397–422, 2002.

[43] J. Mockus *et al.*, "The application of Bayesian methods for seeking the extremum," *Towards Global Optim.*, vol. 2, pp. 117–129, 1978.

[44] L. Zadeh, "Optimality and non-scalar-valued performance criteria," *IEEE Trans. Autom. Control*, vol. AC-8, no. 1, pp. 59–60, Jan. 1963.

[45] B. Beykal *et al.*, "Optimal design of energy systems using constrained grey-box multi-objective optimization," *Comput. Chem. Eng.*, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S009813541830084X

[46] V. Pareto, *Manuale di Economia Politica*. Bologna, Italy: Societa Editrice, 1906.

[47] A. Messac *et al.*, "Ability of objective functions to generate points on nonconvex Pareto frontiers," *AIAA J.*, vol. 38, no. 6, pp. 1084–1091, 2000.

[48] M. Emmerich, "Single- and multi-objective evolutionary design optimization assisted by Gaussian random field metamodels," Ph.D. dissertation, Dept. Comput. Sci., Technical University of Dortmund, Dortmund, Germany, 2005.

[49] J. Svenson and T. Santner, "Multiobjective optimization of expensive-to-evaluate deterministic computer simulator models," *Comput. Statist. Data Anal.*, vol. 94, pp. 250–264, 2016.

[50] A. J. Keane, "Statistical improvement criteria for use in multiobjective design optimization," *AIAA J.*, vol. 4, pp. 879–891, 2006.

[51] D. Hernández-Lobato *et al.*, "Predictive entropy search for multi-objective Bayesian optimization," in *Proc. Int. Conf. Mach. Learn.*, New York, USA, 2016, pp. 1492–1501.

[52] V. Picheny, "Multiobjective optimization using Gaussian process emulators via stepwise uncertainty reduction," *Statist. Comput.*, vol. 25, pp. 1265–1280, 2015.

[53] R. Balling, "The maximin fitness function; multi-objective city and regional planning," in *Proc. Int. Conf. Evol. Multicrit. Optim.*, 2003, pp. 1–15.

[54] M. Binois *et al.*, "Quantifying uncertainty on Pareto fronts with Gaussian process conditional simulations," *Eur. J. Oper. Res.*, vol. 243, no. 2, pp. 386–394, 2015.

[55] R. Calandra *et al.*, "Pareto front modeling for sensitivity analysis in multi-objective Bayesian optimization," in *Proc. NIPS BayesOpt*, 2014, pp. 1–5.

[56] R. Calandra, "Bayesian modeling for optimization and control in robotics," Ph.D. dissertation, Dept. Comput. Sci., Technical University of Darmstadt, Dortmund, Germany, 2017.

[57] D. A. van Veldhuizen, "Multiobjective evolutionary algorithms: Classifications, analyses, and new innovations," Ph.D. dissertation, School of Engineering, Air Force Institute of Technology, Dayton, Ohio, USA, 1999.

[58] C. M. Fonzeca and P. J. Fleming, "An overview of evolutionary algorithms in multiobjective optimization," *Evol. Comput.*, vol. 3, no. 1, pp. 1–16, 1995.

[59] J. D. Schaffer, "Some experiments in machine learning using vector evaluated genetic algorithms," Ph.D. dissertation, Dept. Elect. Eng., Vanderbilt University, Nashville, TN, USA, 1984.

[60] F. Kursawe, "A variant of evolution strategies for vector optimization," in *Proceedings of the 1st Workshop on Parallel Problem Solving from Nature (Comput Science, 496)*. Berlin, Germany: Springer-Verlag, 1991, pp. 193–197.

[61] F. P. Vidal *et al.*, "Tuning of patient-specific deformable models using an adaptive evolutionary optimization strategy," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 1, pp. 2942–2949, Oct. 2012.

[62] S. M. K. Heris and H. Khaloozadeh, "Open- and closed-loop multiobjective optimal strategies for HIV therapy using NSGA-II," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1678–1685, Jun. 2011.

[63] D. Hadka, "Platypus," GitHub, 2015. [Online]. Available: https://github.com/Project-Platypus/Platypus. Accessed on: Jan. 18, 2017.

[64] R. Calandra *et al.*, "Manifold Gaussian processes for regression," in *Proc. Int. Jt. Conf. Neural. Netw.*, 2016, pp. 3338–3345.

[65] Y. Y. Lipsitz *et al.*, "Quality cell therapy manufacturing by design," *Nature Biotechnol.*, vol. 34, no. 4, pp. 393–400, 2016.

[66] S. Olofsson, "HBMOO," 2018. [Online]. Available: https://github.com/scwolof/HBMOO