Inverse problems in reduced order models of cardiovascular haemodynamics: aspects of data-assimilation and heart-rate variability
Sanjay Pant, Chiara Corsini, Catriona Baker, Tain-Yen Hsia, Giancarlo Pennati, Irene Vignon-Clementel

To cite this version:

HAL Id: hal-01413446
https://inria.hal.science/hal-01413446
Submitted on 21 Dec 2016
Inverse problems in reduced order models of cardiovascular haemodynamics: aspects of data-assimilation and heart-rate variability

Sanjay Pant, Chiara Corsini, Catriona Baker, Tain-Yen Hsia, Giancarlo Pennati, and Irene E. Vignon-Clementel

Abstract—Inverse problems in cardiovascular modelling have become increasingly important to address computational efficiency and uncertainty consideration in such problems. This work highlights and presents solutions to several challenges of this method pertinent to models of cardiovascular haemodynamics. These include methods to a) avoid ill-conditioning of covariance matrix; b) handle a variety of measurement types; c) include a variety of prior knowledge in the method; and d) incorporate measurements acquired at different heart-rates, a common situation in the clinic where patient-state differs between various clinical acquisitions. Results are presented for two patient-specific cases of congenital heart disease. To illustrate and validate data-assimilation with measurements at different heart-rates, results are presented on synthetic data-set under different sets of known parameters or different sets of conditions (open or closed-loop) to higher order models (1D or 3D) in the geometrical multiscale method [1], [2]. For a patient-specific analysis, however, these models must be adapted to each patient individually. This means that the number of parameters can be large even in relatively simplistic models representing the whole circulation. The inverse problem has been addressed in many previous studies, see for example [3]–[16] and [17], [18] for their advantages and disadvantages. A recent Bayesian approach to parameter estimation in cardiovascular models is based on markov chain monte carlo methods with mean or maximum target quantities [19]. In recent years, the data-assimilation approach has gained popularity as it can take advantage of full-time varying measurements as opposed to just scalar mean/minimum/maximum indices. It has been employed in a range of cardiovascular problems such as the estimation of Windkessel parameters [18], [20]–[22], closed loop lumped model parameters for single-ventricle physiology [17], artery wall stiffness parameters [22], [23] models, and tissue support parameters for fluid-structure interaction models [24]. Recently, it has also been used for parameter estimation in electrocardiography/electromechanical models [25]–[28], cardiac motion recovery [29], [30], inverse electrophysiology applications [31], [32], and for baroreflex regulation modelling [33]. This study focuses on some aspects of such a data assimilation method, namely the unscented Kalman filter (UKF), towards successful estimation and interpretation in closed-loop lumped parameter models of cardiovascular haemodynamics. The first contribution of this study is the management of Sigma Point collapse in the sequential UKF framework which can lead to ill-conditioning of the empirical covariance matrix, and consequently numerical problems. The second contribution of this study is to present a method to include the measurements of minimum (min) and maximum (max) values of model outputs in the method in an efficient manner. This is relevant as sometimes only min/max values, as opposed to full time-varying measurements, are available/measurable for certain quantities in the clinic. The third contribution is the presentation of different forms of prior knowledge that can be incorporated in the UKF setting to ensure that a reasonable solution of the inverse problem is obtained. The final contribution is a method to include measurements at
different heart-rates which is a challenge in the basic UKF method.

The first two contributions are presented in two patient specific cases of congenital heart disease, namely stage-1 single ventricle physiology (one with normal valves and one with a regurgitant atrioventricular valve). The last contribution of assimilating measurements at variable heart-rates is demonstrated first for a synthetic data-set, to verify the approach, and then on a patient-specific case. This method can be readily extended to other dynamical systems that exhibit different responses under: i) differing values of a single (or a set of) known parameters (i.e. parameters that are not to be estimated), for example, estimation of medium diffusivity and conductivity in an advection-diffusion problem where species concentrations are measured at two different known advection velocities; and ii) different inputs, for example different input signals or different forcing/exciting frequencies. The latter case is especially relevant as, in practice, certain models can be particularly sensitive to some parameters in certain time-zones (or space-time zones in case of partial differential equations) [34]. It is necessary to gain information either by having very precise information in these zones, or to visit them several times sequentially, so that they are revisited when knowledge on the other parameters has been gained. Assimilating data over several quasi-periodic signals is such an example that is naturally encountered in physiology.

It should be noted that single-ventricle physiology presents a challenging treatment as only one heart pump functions well at birth, usually performed in three stages of open-heart surgery where important decisions must be made at each stage [35], [36]. In recent years, the adoption of computational fluid dynamics (CFD) models, large parts of which comprise lumped parameter models, has been employed to aid clinical-decision making and development of novel surgical strategies [36], [37]. For individual case-by-case clinical-decision making in single-ventricle patients, however, the estimation of lumped model parameters while taking into account the uncertainty in clinical measurements is necessary. It is this problem that the methods developed in this manuscript cater to.

The model of single-ventricle physiology adopted in this study is presented in [17] and employs the one-fibre model [38], [39] to describe the heart chambers and an extension to the valve model of [40] presented in [17]. The main motivation for such model choices is that they are based on physically meaningful parameters such as wall-volumes of the heart chambers, effective areas of the valves, sarcomere properties, etc., for which clinical/experimental estimates are available. This reduces the black-box nature of the lumped models and routine measurements such as four-chamber electrocardiography, Doppler velocimetry, etc., can be employed, at least qualitatively, for validation of some of the parameter estimates. This moreover allows the estimated lumped model parameters of the heart to serve as a first good guess for myocardium material properties in larger 3D models.

\[ \sigma_t/p = (1 + 3V/V_w) \]  

\[ \lambda = \frac{l}{l_0} = \left( \frac{1 + (3V/V_w)}{1 + (3V_0/V_w)} \right)^{1/3}. \]

Consequently, the sarcomere shortening velocity, \( v_s \), can be computed as \(-dl/dt\). The active component of the stress is described as

\[ \sigma_a = T_{\sigma_0} f(l) g(t_a) h(v_s) \]
activation (see Fig. 2).

2) Valve description: The pressure drop across a valve is given by the following Bernoulli relation

\[
\Delta p = B q |q| + L \frac{dq}{dt}, \quad B = \frac{\rho}{2 A_{\text{eff}}}^2, \quad \text{and} \quad L = \frac{\rho l_{\text{eff}}}{A_{\text{eff}}}, \tag{8}
\]

where \( \rho \) is the density of blood, and \( A_{\text{eff}} \) and \( l_{\text{eff}} \) denote the effective opening area and effective length of the valve, respectively. The valve state is described by a single variable, \( \xi(t) \), that relates to the effective area at time \( t \) as follows

\[
A_{\text{eff}}(t) = \begin{cases} 
A_{\text{eff}}^\text{max} \xi(t) & \text{normal valve} \\
(A_{\text{eff}}^\text{max} - A_{\text{eff}}^\text{min}) \xi(t) + A_{\text{eff}}^\text{min} & \text{regurgitant} \\
A_{\text{eff}}^\text{max} \xi(t) & \text{incomplete closure} \\
-A_{\text{eff}}^\text{max} \xi(t) & \text{regurgitant} \\
-A_{\text{eff}}^\text{max} \xi(t) & \text{prolapse} 
\end{cases} \tag{9}
\]

where \(-1 \leq \xi \leq 1\) and is negative only if the valve is regurgitant due to prolapse, and the parameters \( A_{\text{eff}}^\text{max} \), \( A_{\text{eff}}^\text{min} \), and \( A_{\text{eff}}^\text{max} \) denote the maximum, minimum, and maximum regurgitant area under prolapse, respectively. The valve dynamics is described by ordinary differential equations such that the rate of valve opening/closing is proportional to the

Fig. 1. Schematic of a closed-loop model for single-ventricle circulation. Measurements selected as observations to the UKF are shown in red. The pressure-flow \( p-q \) relationships for various model components are shown in the right. SA denotes single atrium; SV denotes single ventricle; subscripts LB, UB, LU, and SH correspond to components of lower body, upper body, lungs, and shunt, respectively. In each of the LB, UB, and LU segments, parameters \( R_1, K, L_1, C_1 \), and \( C_2 \) represent the large arteries, the parameter \( R_2 \) represents the resistance of the smaller vessels and the vascular bed, and the parameters \( C_2 \) and \( R_3 \) correspond to the veins.

\[
f(l) = \begin{cases} 
0, & \text{if } l < l_{a0} \\
(l - l_{a0})/(l_{am} - l_{a0}), & \text{if } l_{a0} < l \leq l_{am} \\
1.0, & \text{if } l_{am} < l \leq l_{ae} \\
(l_{af} - l)/(l_{af} - l_{ae}), & \text{if } l > l_{ae}
\end{cases} \tag{4}
\]

\[
g(t_a) = \begin{cases} 
\frac{1}{2} \left(1 - \cos \left(\frac{t_a}{t_{\text{max}}} \right)\right)^E_a, & \text{if } t_a < t_{\text{max}} \\
0, & \text{otherwise}
\end{cases} \tag{5}
\]

\[
h(v_a) = \frac{1 - (v_a/v_0)}{1 + c_v (v_a/v_0)} \tag{6}
\]

where \( l_{a0}, l_{am}, l_{af}, l_{ae} \) describe the dependence of sarcomere length \( l \) on active stress \([17]\), \( t_a \) is the time elapsed since contraction activation, \( t_{\text{max}} \) is the activation duration of the chamber in one cardiac cycle, \( v_0 \) is the initial sarcomere shortening velocity, \( T_{a0} \) is the maximum active sarcomere stress, and \( c_v \) is a shape parameter. The passive stress is given by

\[
\sigma_p = \begin{cases} 
0, & \text{if } \lambda < 1 \\
T_{p0} \exp \{c_p (\lambda - 1)\} - 1, & \text{if } \lambda \geq 1
\end{cases} \tag{7}
\]

where \( T_{p0} \) and \( c_p \) are sarcomere material properties.

Similar to \([17]\), \( c_v = 0 \) is assumed and \( g(t_a) \) is parameterised by three parameters: \( t_{a0}, t_{a0} \), the activation duration of the heart chambers, namely the single atrium, \( t_{a0} \), the activation duration of the single ventricle, and \( t_1 \), the overlap between the end of atrium activation and the beginning of ventricle
favourable pressure difference (see [17] for details)
\[
\begin{cases}
\dot{x} = (1 - \xi) K_{vo} \Delta p & \text{if } \Delta p \geq 0 \\
\dot{x} = K_{vc} \Delta p & \text{if } \Delta p < 0 \text{ and } \xi \geq 0 \\
\dot{x} = (1 + \xi) K_{vg}^{r} (\Delta p - \Delta p_{rg}) & \text{if } \Delta p \leq \Delta p_{rg} \\
\dot{x} = -\xi K_{vc}^{r} (\Delta p - \Delta p_{rg}) & \text{if } \Delta p > \Delta p_{rg} \text{ and } \xi < 0
\end{cases}
\]

where \(\Delta p_{rg}\) is the pressure gradient beyond which the valve prolapses, and \(K_{(\cdot)}\) are proportionality parameters.

\[\text{B. Overview of the unscented Kalman filter}\]

After having described the hemodynamics model above, let us explain how its parameters are estimated based on dynamics measurements of blood flow and pressure. The unscented Kalman filter is an extension to the Kalman filter for non-linear problems [41], [42]. Let \(y \in \mathbb{R}^{d}\) be the state-vector. The so called forward model that relates the state \(y_{n}\) at time \(t_{n}\) to that at \(t_{n+1}\) can be written as

\[y_{n+1} = F(y_{n}, \theta, \Delta t_{n}),\]

where \(\theta \in \mathbb{R}^{q}\) represents the parameter vector, \(\Delta t_{n}\) represents the time-step, and \(F\) represents the forward operator. For example, in the model represented in Fig. 1, the state vector is represented as \(y = [V_{SA}, V_{SV}, q_{AVV}, q_{AOV}, \xi_{AVV}, \xi_{AOV}, p_{AO}, p_{1UB}, p_{2UB}, p_{1LB}, p_{2LB}, p_{AO}, q_{UB}, q_{LB}, q_{SH}]\). The operator \(F\) is then defined by the discretisation of the ODE system resulting from the heart and valve models described in sections II-A1 and II-A2 and the individual component equations shown in Fig. 1 (right). The measurements are described by the observation operator as

\[z_{n} = H(y_{n}, t_{n}) + \varepsilon(t_{n}),\]

where \(z_{n} \in \mathbb{R}^{m}\) is the observation vector, i.e. the subset of measurements chosen to estimate parameters, at time \(t_{n}\). \(H\) represents the observation operator that relates the model state to the measurements, and \(\varepsilon(t_{n})\) represents the measurement noise. For parameter estimation through sequential filtering methods, the state vector \(y_{n}\) is combined with the parameter vector to yield an augmented state vector \(x_{n}\), and the forward, \(F\), and observation, \(H\), operators, are written as

\[x_{n+1} = \begin{bmatrix} y_{n+1} \\ \theta_{n+1} \end{bmatrix} = \begin{bmatrix} F(y_{n}, \theta_{n}, \Delta t_{n}) \\ \theta_{n} \end{bmatrix} = F(x_{n}, \Delta t_{n}),\]

\[z_{n} = H(x_{n}, t_{n}) + \varepsilon(t_{n}).\]

The goal of sequential methods is to provide an estimate for the state \(x_{n}\), by taking all the measurements \(z_{i}\) to \(z_{n}\) into account. This is achieved recursively by providing an initial estimate \(\hat{x}_{0}\) with a covariance matrix \(P_{0}^{x}\). The recursion is as follows: first, the state \(\hat{x}_{n}\) and covariance \(P_{n}^{x}\) are propagated to an intermediate state \(\hat{x}_{n+1}^{-}\) with covariance \(P_{n+1}^{-}\) through the forward model of equation (13), and then the estimate \(\hat{x}_{n+1}\) with covariance \(P_{n+1}^{x}\) is obtained by correcting the intermediate state, by accounting for (assimilating or probabilistically conditioning on) the observation \(z_{n+1}\) through equation (14). If \(N\) \((n = 1, 2, \ldots N)\) observations are available, then an estimate of the parameters is provided by the corresponding components of \(x_{N}\), and associated variances by the corresponding diagonal elements of \(P_{N}^{x}\).

If the forward and observation operators for the augmented state \(F\) and \(H\) are linear, and \(\varepsilon(t_{n})\) is white noise, then the Kalman filter [43] provides an optimal state estimate and both the propagation and assimilation steps can be analytically computed. For non-linear problems, various extensions such as the extended Kalman filter (EKF), ensemble Kalman filter (EnKF), unscented Kalman filter (UKF), etc. have been proposed, which differ primarily in the forward propagation step. For a discussion on such choices, the reader is referred to [18]. In what follows the forward propagation and assimilation steps of the UKF are briefly described.

1) Forward propagation of uncertainty: In UKF, forward propagation is performed by propagating a set of deterministically chosen particles, called Sigma Points, through the forward model of equation (13). The Sigma Points, denoted by \(X_{k}^{i}, k = 0, \ldots, 2(d + s)\) are generated as follows

\[X_{n}^{0} = \hat{x}_{n}, \]

\[X_{n}^{i} = \hat{x}_{n} + \sqrt{L P_{n}^{x}} \zeta_{n}^{i}, \quad i = 1, \ldots, L \]

\[X_{n}^{i} = \hat{x}_{n} - \sqrt{L P_{n}^{x}} \zeta_{n}^{i}, \quad i = L + 1, \ldots, 2L \]

where \(L = d + s\) and \((\sqrt{\cdot})\) represents the \(i^{th}\) row of the matrix square-root of \((\cdot)\). Each \(X_{n}^{i}\) is then propagated as

\[X_{n+1}^{i} = F(X_{n}^{i}, \Delta t_{n}), \quad i = 0, \ldots, 2L \]

where \(X_{n+1}^{i}\) denotes the propagated particle. \(\hat{x}_{n+1}^{-}\) and \(P_{n+1}^{-}\) are calculated from the propagated Sigma Points as

\[\hat{x}_{n+1}^{-} = \sum_{i=0}^{2L} W_{m}^{i} X_{n+1}^{i}, \]

\[P_{n+1}^{-} = \sum_{i=0}^{2L} W_{c}^{i} [X_{n+1}^{i} - \hat{x}_{n+1}^{-}][X_{n+1}^{i} - \hat{x}_{n+1}^{-}]^{T}, \]

where \(W_{m}^{i}\) and \(W_{c}^{i}\) represent the weights associated with \(i^{th}\) Sigma Points for empirically calculating mean and covariance, respectively [18]. In a similar manner, the intermediate (prior) mean and covariance of the observation vector are calculated by propagating \(X_{n+1}^{i}\) through equation (14) and calculating the empirical statistics

\[Z_{n+1}^{i} = H(X_{n+1}^{i}, t_{n+1}), \quad i = 0, \ldots, 2L \]

\[\hat{z}_{n+1}^{i} = \sum_{i=0}^{2L} W_{m}^{i} Z_{n+1}^{i}, \]

\[P_{n+1}^{z} = \sum_{i=0}^{2L} W_{c}^{i} [Z_{n+1}^{i} - \hat{z}_{n+1}^{i}][Z_{n+1}^{i} - \hat{z}_{n+1}^{i}]^{T} + \Sigma_{n+1}, \]

where \(\Sigma_{n+1}\) represents the covariance matrix of the measurement noise \(\varepsilon(t_{n+1})\). Finally, the cross-covariance matrix between \(\hat{x}_{n+1}^{-}\) and \(P_{n+1}^{z}\) is

\[P_{n+1}^{xz} = \sum_{i=0}^{2L} W_{c}^{i} [X_{n+1}^{i} - \hat{x}_{n+1}^{-}][Z_{n+1}^{i} - \hat{z}_{n+1}^{i}]^{T}. \]
2) Assimilating the observations: Through the measurement \( z_{n+1} \), the above calculated intermediate state is corrected to yield an estimate of the state \( \hat{x}_{n+1} \) with covariance \( P^x_{n+1} \) at \( t_{n+1} \) by minimising the following cost function \( J(\hat{x}_{n+1}) \)

\[
J = \frac{1}{2} \| \hat{x}_{n+1} - \bar{x}_{n+1} \|^2 (P^x_{n+1})^{-1} + \frac{1}{2} \| z_{n+1} - \mathcal{H}(\hat{x}_{n+1}) \|^2 (\Sigma_{n+1})^{-1}
\]

The solution to the above in the UKF setting is given by

\[
\hat{x}_{n+1} = \bar{x}_{n+1} + K_{n+1} (z_{n+1} - \hat{z}_{n+1})
\]

\[
P^x_{n+1} = P^x_{n+1} - K_{n+1} P^z_{n+1} K_{n+1}^T
\]

where the Gain matrix \( K_{n+1} \) is given by

\[
K_{n+1} = P^x_{n+1} (P^z_{n+1})^{-1}
\]

Remark 1: It should be noted that if the observation operator \( \mathcal{H} \) is linear then the propagation of the Sigma Points through the observation operator, equation (19), is not necessary. Instead, \( P^z_{n+1} \) and \( P^x_{n+1} \) in equations (21) and (22), respectively, can be analytically computed. For ODE models, in certain cases, this may induce a choice of formulation. Consider the heart model as an example. This model, see Fig. 1 and section II-A1, is highly non-linear where, for each heart chamber, the ODE is formed naturally for volume \( V \) as follows

\[
\dot{V} = dV / dt = q_{in} - q_{out}
\]

and the pressure, \( p \), is calculated by equations (1)–(7). However, it is the pressure, \( p_{SA} \) or \( p_{SV} \), that is typically observed. If one were to reformulate the ODE in terms of pressure (by differentiating equations (1) and (27)) rather than the volume, then pressure would be a part of the state \( x \) and hence a linear observation manager would suffice. However, given the nature of equations (1)–(7), it is much easier to keep the ODE state variable as \( V \) and employ a non-linear observation operator.

C. Open-loop vs. closed-loop

In the context of estimating parameters for haemodynamic systems, such as the one presented in Fig. 1, a natural question that arises is whether one should perform parameter estimation in the entire closed-loop, or if one can perform simpler estimation in sub-segments of the closed-loop circulation, i.e. in an open-loop setting. In terms of computational cost, solving multiple smaller open-loop segments is cheaper than one larger closed-loop model, and this may be a driving motive for open-loop estimation. However, in ODE systems such as that of Fig. 1, the cost of solving the full closed-loop model is not exorbitantly high, and hence the computational gain would be of little practical significance. Another consequence of estimation in open-loop segments is related to the manner in which uncertainty in the measurements is accounted for. As an example, consider the lower body segment in Fig. 1 (the region between nodes N-1 and N-2). Since at both ends, the flow-rates, i.e. \( q_{LB} \) and \( q_{IVC} \), and the pressures, i.e. \( p_{AO} \) and \( p_{SA} \), are available as measurements, one may consider an open-loop model just for this segment. However, in order to solve the forward model one must impose two forcing conditions, in this case say at the ends N-1 and N-2. For this, one may choose either \( p \) or \( q \) at each end and employ the remaining two measurements as observations to the UKF. While a reasonable approach, in doing so two of the measurements at the boundaries have been deterministically imposed and hence are considered devoid of any uncertainty. This might lead to erroneous results if such an assumption, depending on the quality of the measurements, cannot be justified. Furthermore, there is little assurance that when the estimated parameters for each segment are plugged into the closed-loop model with no such forcing conditions, the model output will agree with the measurements. On the other hand, there is a distinct advantage in an open loop setting. In certain cases, the measurements are not synchronised in time with respect to each other. This is especially true if simultaneous electrocardiogram (ECG) measurements are not taken or are unreliable, or due to different techniques for the pressure (typically catheterisation) and flow-rates (typically magnetic resonance imaging (MRI)) measurements implying that their simultaneous measurement is rarely possible. In such cases, some of the time-shifts can be posé as parameters in the open-loop models, and can be estimated by the UKF. As an example, consider the aforementioned lower body open-loop segment. Assume that the pressure measurement \( p_{SAO}(t) \) is synchronised with \( p_{SA}(t) \) and the flow-rate measurement \( q_{LB}^m(t) \) is synchronised with \( q_{IVC}^m(t) \), but the pressure and flow-rate are not synchronised with respect to each other. In the UKF one may impose the following as forcing conditions (FCs) and measurements while accounting for the phase/time-shift between \( q^m \) and \( p^m \) by a parameter \( \delta t_{p,q} \):

\[
FCs: \begin{cases} p_{SA}(t) = p_{SA}^m(t) \\ q_{LB}(t) = q_{LB}^m(t + \delta t_{p,q}) \end{cases}
\]

\[
Observations: \begin{cases} p_{AO}(t) = p_{AO}^m(t) \\ q_{IVC}(t) = q_{IVC}^m(t + \delta t_{p,q}) \end{cases}
\]

Using the above formulation, and by including \( \delta t_{p,q} \) as one of the parameters to be estimated, the discrepancy due to asynchronous measurements can be taken into account. Note that the parameters must affect the forward model in order to be estimated; hence only in an open-loop setting where \( \delta t_{p,q} \) appears in a FC and this affects the forward model, can it be estimated. Such an application of estimating the asynchronisation is presented in [18]. On the contrary, in a closed-loop setting the asynchronisation between the measurements cannot be automatically accounted for.

Preliminary tests suggested that parameter estimation on individual open-loop segments of Fig. 1 may not result in as good reproduction of all the measurements when substituted in the closed-loop model (also see [44] where in an open-loop configuration the heart-parameters estimated for patient-A are considerably different than those estimated in a closed-loop setting in [17]). Thus in what follows, estimation is performed directly on the closed-loop model. Such an observation suggests that uncertainty in the measurements is important to be considered.
D. Sigma Point collapse

The valve model presented in section II-A2 is essentially described by the valve dynamics parameter ξ. In haemodynamics problems, the valve state remains closed or open for long periods of time in a cardiac cycle. This implies, that during these periods the valve state ξ remains 0, 1, or -1, and all perturbations (i.e. Sigma Points) to this variable at a time \( t_n \) may result in the same final state (0, 1, or -1) at time \( t_{n+1} \). In two dimensions, this phenomenon is depicted in Fig. 3, where at \( t_{n+1} \) all Sigma Points have the same value for the ξ variable. This phenomenon makes the empirical covariance matrix \( P_{n+1}^{x} \) not positive definite, which may result in a non-positive definite \( P_{n+1}^{x} \) after the data-assimilation step is performed. Note that the matrix square-root in equation (15) is typically calculated via Cholesky factorisation, and hence the non positive definiteness of \( P_{n+1}^{x} \) presents problems. This issue can be remedied by adding a small diagonal term to \( P_{n+1}^{x} \) in equation (18) as follows

\[
P_{n+1}^{x} = \delta_r I + \sum_{i=0}^{2L} W_i \begin{bmatrix} \mathcal{X}^{x}_{i,n+1} - \hat{x}_{n+1} \end{bmatrix} \begin{bmatrix} \mathcal{X}^{x}_{i,n+1} - \hat{x}_{n+1} \end{bmatrix}^T , \tag{30}
\]

where \( I \) is an \( L \times L \) identity matrix, and \( \delta_r \) is a scaling factor. The added matrix \( \delta_r I \) can be interpreted as a regularisation term, or as the covariance matrix of error in model forward propagation. In the latter view, the discretised model can be seen, in lieu of equation (13), as

\[
x_{n+1} = F(x_n, \Delta t_n) + \varepsilon_m, \tag{31}
\]

where \( \varepsilon_m \) is the discretised model propagation error with zero mean and covariance \( \delta_r I \). The end result is that even though the Sigma Points collapse in one or more dimensions, the propagated variance in these directions is maintained at a low value of \( \delta_r \). In the Bayesian perspective of UKF, the prior on \( x_{n+1} \) can be seen as a Normal distribution with mean \( \hat{x}_{n+1} \) and covariance \( P_{n+1}^{x} \). Then, the data assimilation step consists in computing the posterior distribution of \( x_{n+1} | z_{n+1} \). When collapse of Sigma Points occurs, the prior on the collapsed states becomes a dirac delta function, i.e. the prior has no uncertainty associated with these states, and hence cannot be corrected in the posterior. The term \( \delta_r I \) avoids such a case by ensuring that no state variables become entirely certain at any step of the UKF.

E. Minimum and maximum of a time-varying output

While the UKF can easily incorporate time-varying measurements by time discretisation, i.e. a measurement \( z(t) \) is provided at discrete time-instants \( t_k \) as observations, it is not straightforward to include measurements that correspond to minimum or maximum of a time-varying output. This is a common measurement in haemodynamics problems, when only min/max volumes, pressures, or flow-rates are known. If the time instants when the minima or maxima occur is known \textit{a priori}, then inclusion of such measurements is straightforward by only providing these observations at the known time-instants. However, the location of maxima or minima depends on the parameters and the state variables, which are continuously being modified by the UKF. Therefore, methods to include such measurements are needed. One way, in a non-linear observation operator setting, is to consider the following observation operator that returns the max. or min. of an output, say the \( i^{th} \) component of the state, as follows

\[
z_n = H(x_n) = \min_{0 \leq k \leq K} (\hat{x}_n^{(i)}), \quad K = T/\Delta t, \tag{32}
\]

where \( \hat{x}_n^{(i)} \) represents the \( i^{th} \) component of \( \hat{x} \), \( \Delta t \) is the time-step, \( T \) is the time-period (one cardiac cycle), and \( \hat{x}_{n+(k+1)} = F(\hat{x}_{n+k}, \Delta t) \); with initial value \( \hat{x}_{n} = x_{n} \). (33)

In the above method, a forward model needs to be solved for all the Sigma Points \( \mathcal{X}^{x^{(i)}}_{n+1} \) for at least one cardiac cycle to obtain the min/max values of the observable output. In standard UKF, there are \( 2(d+s)+1 \) Sigma Points, and consequently \( 2(d+s)+1 \) forward simulations are required. If the above method is utilised to include the min/max observations then at each time when observations are assimilated, \( 2(d+s)+1 \) additional forward models need to be evaluated for at least one cardiac cycle. This blows up the computational cost associated with UKF. As an example, if the UKF is run for two cardiac cycles and observations are assimilated at a total of \( N \) discrete time instants, then \( 2(d+s)+1 \) forward simulations (for two cardiac cycles) need to be run for the standard UKF, and additional \( 2N(d+s)+N \) forward simulations (for one cardiac cycle) to include the min/max observations. Alternatively, in certain cases, the time instants where the maxima/minima (\( t_{\text{min/max}} \)) for a particular output occur can be predetermined and are usually a function of the parameters. In such cases the observation operator can simply be appended whenever \( t = t_{\text{min/max}} \) to include the min/max measurement. As an example, consider that for the ventricle, only the end systolic volume (ESV) and end diastolic volume (EDV) measurements are available. Typically, these will correspond to the maximum and minimum chamber volumes during a cardiac cycle. In the model the contraction of the ventricle is dictated by the activation function parametrised by the parameters \( t_{\text{aSV}} \). Hence, given the parameter \( t_{\text{aSV}} \), the time instants where the ventricle will be at min/max volumes can be determined. This is shown in Fig. 4: the EDV (max) occurs just before the ventricle begins to contract, i.e. \( t_c = 0 \) where \( t_c \) is the time elapsed since the beginning of the current cardiac cycle.
cycle, and the ESV occurs when the contraction phase of the ventricle ends, i.e. at \( t_c = t_{\text{asv}} \). Hence, if the EDV and ESV measurements are available then one may consider a small interval of width \( \pm \delta_{\text{minmax}} \) around \( t_c = 0 \) and \( t_c = t_{\text{asv}} \), and provide the EDV and ESV measurements as the observations of \( V_{SV} \) only when the UKF time \( t \) is in these intervals, the intervals in turn being determined by the current state (which determines \( t_{\text{asv}} \)).

Fig. 4. Relation of EDV and ESV times with respect to ventricle activation function: ventricular activation is shown in red and its volume in blue, both as a function of time. The large dash line indicates beginning of ventricular contraction.

F. Prior knowledge: dynamical system and parameters

As the number of parameters to be estimated increases, the inverse problem of estimating parameters becomes harder. This is particularly true if the associated cost function between the model and the measurements, i.e. \( L = \frac{1}{2} \sum_{n=0}^{N} \| \mathcal{H}(x_n, t_n) - z_n \|_{\Sigma_n} \) is multimodal or non-convex (i.e. it has either multiple minima or no minimum at all). In classical least-squares minimisation framework, identifiability is partly improved by adding a regularisation term as follows

\[
L = \frac{1}{2} \| x_0 - \hat{x}_0 \|_{P_0} + \frac{1}{2} \sum_{n=0}^{N} \| \mathcal{H}(x_n, t_n) - z_n \|_{\Sigma_n}.
\]  

As is evident from the above notation, the regularisation term can be seen as the prior, i.e. the initial estimate for the state with mean \( \hat{x}_0 \) and covariance \( P_0 \), employed in the UKF method. As demonstrated in [18], a bad choice of \( \hat{x}_0 \) and \( P_0 \) can lead to failure of the UKF method even in relatively simple dynamical systems. This aspect becomes more important when the size of the dynamical system and the number of parameters increase. Consequently, for such systems, prior knowledge on the parameters is important for the success of the method. The alternative is to run many instances of the UKF method with different priors, but this is a brute force approach that, although realisable, requires significantly more computational effort. Prior information may be classified into the following categories:

1) Parameter values: Certain parameter values are known to be close to values determined clinically, through previous studies, or through experiments. For example, estimates of the wall volume of the ventricles, maximum valvular areas, actuation durations of the heart chambers, are available through clinical imaging, echocardiography, or past clinical studies. In essence, this knowledge is a measure of what is expected in a population of patients of similar age and physiology, and especially helpful when model parameters have a physiological meaning, such as effective valve areas, wall volumes of the heart chambers, etc. Such knowledge can be easily integrated in the UKF initial state \( \hat{x}_0 \). A priori there is generally no reason to assume correlation between the elements of \( \hat{x}_0 \) and hence \( P_0 \) can be assumed to be diagonal with respective entries corresponding to the respective variances (inverse of confidences arising from the population-based knowledge) in the chosen prior estimates.

2) Parameter ranges: A common form of prior knowledge is about parameter values in terms of known ranges of variation. Even more common is the knowledge that a certain parameter cannot be negative. For example all the resistances, capacitances, and inductances in Fig. 1 are known a priori to be positive. Similarly, the volumes cannot be negative and activation durations cannot be larger than the cardiac cycle. While such information cannot be included in the form of a prior, one can include appropriate parameterisations to enforce such constraints. For positive parameters

\[
\Psi = 2^\psi,
\]

and for a parameter constrained between \( \Psi_{\text{max}} \) and \( \Psi_{\text{min}} \)

\[
\Psi = \Psi_{\text{min}} + \frac{\Psi_{\text{max}} - \Psi_{\text{min}}}{1 + \exp(-\psi)}
\]

may be chosen, where \( \Psi \) is the real parameter and the \( \psi \) is the form employed in the state \( x \). The parameterisation ensures that the variable \( \psi \) can be manipulated by the UKF to take values from \( -\infty \) to \( +\infty \) in the Sigma Points while yielding reasonable values of the real parameter to be employed in the forward model. The parameterisation of equation (35) can also be employed for state variables such as pressure if they are required to be positive. This also makes sense from a conceptual viewpoint in the sense that in the UKF the measurement errors (variances diagonal elements of \( \Sigma_n \)) are seen as Gaussian, and such assumption leaves a finite non-zero probability that the pressure might be negative. Instead, the parameterisation of equation (35) transfers the Normal assumption to the exponent \( \psi \), thus making the real pressure \( \Psi \) log-normally distributed, which is a much more reasonable assumption due to only positive support of the log-normal distribution.

In the models used in this study, all pressures and volumes are parameterised by equation (35) to enforce positivity constraints. The state variables \( \xi_{\mathrm{AVV}} \) and \( \xi_{\mathrm{AOV}} \) are parameterised by equation (36) to be constrained between 0 and 1 or between -1 and 1 depending on whether regurgitation is included. Similarly, the parameters \( t_{\text{asv}} \) and \( t_{\text{agv}} \) are constrained between 0 and \( T \) (the time-period of the cardiac cycle) and the overlap parameter \( t_1 \) is constrained between 0 and 0.05s through equation (36).

3) Prior knowledge on the range of variation of a state variable: The case when min/max values of certain state variables are known a priori with reasonable degree of confi-

\[
\begin{align*}
\xi_{\mathrm{AVV}} &\sim \mathcal{N}(m_{\xi_{\mathrm{AVV}}}, \sigma_{\xi_{\mathrm{AVV}}}^2) \\
\xi_{\mathrm{AOV}} &\sim \mathcal{N}(m_{\xi_{\mathrm{AOV}}}, \sigma_{\xi_{\mathrm{AOV}}}^2) \\
t_{\text{asv}} &\sim \mathcal{U}(0, 1) \\
t_{\text{agv}} &\sim \mathcal{U}(0, 1) \\
t_1 &\sim \mathcal{U}(0, 0.05s)
\end{align*}
\]

\[
\begin{align*}
\xi_{\mathrm{AVV}} &\sim \mathcal{N}(m_{\xi_{\mathrm{AVV}}}, \sigma_{\xi_{\mathrm{AVV}}}^2) \\
\xi_{\mathrm{AOV}} &\sim \mathcal{N}(m_{\xi_{\mathrm{AOV}}}, \sigma_{\xi_{\mathrm{AOV}}}^2) \\
t_{\text{asv}} &\sim \mathcal{U}(0, 1) \\
t_{\text{agv}} &\sim \mathcal{U}(0, 1) \\
t_1 &\sim \mathcal{U}(0, 0.05s)
\end{align*}
\]
dence is presented in section II-E. Here, the case when such ranges are not precisely known in a particular patient and yet there is some belief about what this range should not be is presented. For example, in stage-I single-ventricle physiology, the ventricle volumes may vary from 5-25 ml during end-systole (end of the heart contraction period) to 15-50 ml during end-diastole (end of the cardiac chamber filling), depending on the age and physiology of the patient. A stronger belief is that these should not vary between 40 ml at end diastole to 70 ml at end systole, as these are too high for a 3–6 months old patient. This information may be seen as an imprecise knowledge about the mean value of ventricle volume. From equations (1) and (2) it is evident that in the heart model, the ratios $V/V_w$ and $V_0/V_w$ determine the pressure. If only a noisy measurement of pressure in aorta (reflective of pressure in the ventricle) is available, the parameters $V_0$ and $V_w$ are weakly identifiable: i.e. for a higher $V_w$ and a correspondingly higher $V_0$, the aforementioned mean value of $V$ might increase while yielding a similar value of pressure. The UKF, in such a case, may show a constant drift in both the parameters $V_0$ and $V_w$ and in the volume $V$. A natural question then is how to constrain the mean value from drifting too far from a known value, say $V_k$. One way to include such a constraint is to provide the following measurement for the volume $V(t)$

$$z_n = V_k = V(t_n) + \varepsilon_n \ ; \quad \varepsilon_n \sim \mathcal{N}(0, \sigma^2_k), \quad (37)$$

where $\sigma^2_k$ is a relatively high value in comparison to the real measurement variances. Essentially, for a time-varying quantity $V$ this provides a pseudo constant measurement of $V_k$ at each time of data assimilation. It results in a pulling force towards the value $V_k$ proportional to the distance of $V$ from $V_k$, and if appropriately chosen affects the solution only when $V(t)$ drifts too far from $V_k$. For example, assume that the true $V(t)$ varies between 8ml and 32 ml. If one provides a constant pseudo measurement of 20 ml with a variance 400 ml$^2$ (i.e. 0–40 ml with one standard deviation), the effect on the solution is negligible when the volume is in the true range, and will pull the solution towards the 20 ml significantly when $V(t)$ drifts far from 20 ml.

4) Initial value of the state variables: A high degree of confidence can be associated with the initial values of some of the state variables based on the model physics. For example, in the heart model note that the driving functions are the activation functions shown in Fig. 2, and hence $t = 0$ corresponds to the time just before the activation of the ventricle begins. Furthermore, a priori one may assume that the parameter $t_1$ is zero implying no overlap between the atrium and ventricle activations and hence at $t = 0$ the atrium activation has just finished. Under these assumptions, confidence is high that at $t = 0$ both the valves are closed, consequently both the valvular flows are zero, the aortic pressure is at end-diastole, the atrium is at end-systole, and the ventricle is at end-diastole.

In a sense, all the above points are about choosing appropriate regularisation for the inverse problem and how to include this regularisation in the UKF method. These priors/regularisations are essential when estimating a high number of parameters and make the inverse problem better posed. Without appropriate prior information, one may see $x_0$ and the diagonal elements of $P_0$ as parameters of the UKF method. These are $2d + 2p$ parameters, and searching for the
right parameters to utilise in a method to estimate $p$ parameters of the inverse problem beats the purpose.

G. Heart-rate variability in measurements

A common problem associated with haemodynamics measurements is heart-rate variability. In particular, due to different techniques of measurements—and consequently different patient states during measurements (sometimes a difference of days)—the pressure (obtained by catheterisation) and the flow-rate (obtained by MRI) measurements are typically at different heart rates. This poses problems in the basic UKF setting as the forward model is run only at one heart rate. A method to cope with such heart rate variability in the UKF is presented here. Consider the following two models at different heart rates

$$\begin{align*}
\begin{cases}
 y^{(1)}_{n+1} &= F^{(1)} \left( y^{(1)}_n, \theta_n, \Delta t_n \right) \quad \text{at heart rate } HR_1 \\
 y^{(2)}_{n+1} &= F^{(2)} \left( y^{(2)}_n, \theta_n, \Delta t_n \right) \quad \text{at heart rate } HR_2 
\end{cases}
\end{align*}$$

(38)

where $y^{(1)}_n$ and $y^{(2)}_n$ represent the state variables at time $t_n$ for the models $F^{(1)}$ and $F^{(2)}$ with heart-rates $HR_1$ and $HR_2$, respectively. Assume that the pressures are observed at $HR_1$ and the flow-rates at $HR_2$ as follows

$$\begin{align*}
\begin{cases}
 z^{(1)}_{n+1} &= H^{(1)} \left( y^{(1)}_{n+1}, t_n \right) + \varepsilon^{(1)} (t_{n+1}) \\
 z^{(2)}_{n+1} &= H^{(2)} \left( y^{(2)}_{n+1}, t_n \right) + \varepsilon^{(2)} (t_{n+1})
\end{cases}
\end{align*}$$

(39)

where $H^{(1)}$ are the observation operators, $z^{(1)}$ the observations, and $\varepsilon^{(1)} (t_{n+1})$ the corresponding measurement errors. For parameter estimation, the following augmented state models for forward and observation are considered

$$\begin{align*}
 x_{n+1} =
 \begin{bmatrix}
 y^{(1)}_{n+1} \\
 y^{(2)}_{n+1} \\
 \theta_{n+1}
\end{bmatrix}
 =
 \begin{bmatrix}
 F \left( y^{(1)}_n, \theta_n, \Delta t_n \right) \\
 F \left( y^{(2)}_n, \theta_n, \Delta t_n \right)
\end{bmatrix}
 =
 F (x_n, \Delta t_n),
\end{align*}$$

$$\begin{align*}
 z_n =
 \begin{bmatrix}
 z^{(1)}_n \\
 z^{(2)}_n
\end{bmatrix}
 =
 \begin{bmatrix}
 H^{(1)} \left( y^{(1)}_{n+1}, t_n \right) + \varepsilon^{(1)} (t_{n+1}) \\
 H^{(2)} \left( y^{(2)}_{n+1}, t_n \right) + \varepsilon^{(2)} (t_{n+1})
\end{bmatrix}
\end{align*}$$

(40)

(41)

Denoting the RHS of equation (41) by $H (x_n, t_n) + \varepsilon (t_n)$, the UKF is performed on the state $x_n$ with forward operator $F$ of equation (40) and observation operator $H$ of equation (41).

In the above formulation the parameters $\theta$ are corrected by both the observations at $HR_1$ and $HR_2$, but each state $y^{(1)}$ and $y^{(2)}$ is only individually corrected by the observations $z^{(1)}$ and $z^{(2)}$, respectively. Since most ODE systems for haemodynamics models are initial state forgetting, in the sense that irrespective of the initial state $y_0$, all runs eventually converge to the same state depending only on the parameters, it should be possible to estimate complete states $x^{(1)}$ and $x^{(2)}$ by correcting the parameters $\theta$ through observations at both heart-rates. It should be noted that the above formulation requires $4d + 2s + 1$ forward and observation model simulations as opposed to $2d + 2s + 1$ in the basic UKF. Finally, the proposed method can be easily extended to cases where measurements at more than two different heart-rates are available in a straightforward manner, albeit at a linearly increasing computational cost.

All the computations in this study are implemented in-house in the Python/Cython programming language. All simulations are run on an HPC node (RAM 48 GB) with 2 Intel Xeon X5650 processors, each with 6 cores and 12 threads.
III. RESULTS AND DISCUSSION

The results are divided into two categories: a) when only a single-heart rate is considered; and b) when measurements are available at differing heart-rates. For quick reference, a nomenclature of all the estimated parameters is presented in Table I in the appendix.

A. Inverse problems considering a single heart-rate

The model and methods for parameter estimation are applied to two patients: patient-A with normal atrioventricular valve (AVV) and patient-B with regurgitant AVV valve due to prolapse. Details of clinical data acquisition can be found in the appendix. Both patients correspond to stage-I surgery for the hypoplastic left/right heart syndrome, a form of single-ventricle physiology. The parameter estimation results for these two patients are presented in [17] and the reader is referred to this study for the choice of model, time evolution of parameter estimates with the UKF method, validation of the parameter estimates, and the clinical significance of the results. Here, the method to set up the UKF and the results of parameter variances, that are both not included in [17], are presented. Parameter variances in particular can provide an indication of the identifiability of the system.

1) Typical patient-specific case: For patient-A, a non-linear observation manager (see Remark 1) is employed, Sigma Point collapse is prevented for the valve states (section II-D), and the prior knowledge is considered as described in section II-F. In particular, $\varepsilon_m$ in equation (31) is assumed to be $\delta_r I$ with $\delta_r$ equal to 1e-08. Additionally, to constrain the atrium and ventricle volumes, pseudo measurements of $V_k = 12$ml and $V_k = 15$ ml are respectively provided with $\sigma_2 = 50$ ml², see equation (37). Fig. 5 shows the evolution of $\log_2(\sigma^2/\sigma_0^2)$ in the UKF for the 33 parameters. In this figure, which complements Figure 3 of [17], $\sigma_0^2$ is the initial prior variance (diagonal elements of $P_0$) and $\sigma^2$ is the UKF estimated variance (diagonal elements of $P_n$). Note that a negative value of $-n$ in this plot corresponds to an uncertainty reduction by a factor of $1/2^n$ relative to the prior variance. The rate of decrease corresponds to the amount of information obtained from the measurements at each time: in general, a faster decreasing variance corresponds to a relatively easily identifiable parameter. Fig. 5 shows that within 4 cardiac cycles, $t \approx 2s$, the variance of all the parameter estimates, except $R_{1LU}$ and $\nu_{SH}$, is decreased by at least a factor of four relative to the prior variance. This trend of consistently decreasing variance continues as more measurements are added. On the contrary, the parameters $R_{1LU}$ and $\nu_{SH}$, appear to be weakly identifiable in the current setting as evidenced by the fact that in 20 cardiac cycles their variance has been less than halved. This may be related to small sensitivities of the measured model outputs to the parameters around the estimated values (see Table 3 of [17] where the estimated values for these parameters are close to negligible). Small sensitivities and close to zero estimates indicate that the parameters are not critical to model output. On the contrary, if they are critical to model output and appear to be weakly identifiable during estimation, it indicates a need for more measurements to improve identifiability. For example, the identifiability in the pulmonary block in Fig. 1 may be improved by providing an added measurement of
Fig. 8. Parameter evaluation for the case with heart-rate variability (synthetic data): horizontal lines indicate true parameter values. Estimated parameters over time: Volumes (V) are in ml; areas (A) are in cm²; linear resistances (R) are in mmHg.s/ml; quadratic resistances (K) are in mmHg.s²/ml²; compliances (C) are in ml/mmHg; inductances (L) are in mmHg.s²/ml; and t_{AVV} is in seconds.

Solid lines labelled on the primary left y-axis show the parameter evolution and dashed lines labelled on the secondary right y-axis show the evolution of \( \log_2(\sigma^2/\eta_0^2) \) for the \( \psi \)-parameterisation (see section II-F2) of the model parameters. Volumes are in ml; areas (A) are in cm²; linear resistances (R) are in mmHg.s/ml; quadratic resistances (K) are in mmHg.s²/ml²; compliances (C) are in ml/mmHg; inductances (L) are in mmHg.s²/ml; t_{AVV} and t_1 are in seconds.

Fig. 10. Parameter evolution over time (in s) when considering heart-rate variability of patient-B.
the pulmonary artery pressure, $p_{\text{LVR}}$. It should be noted that the idea of first identifying most influential parameters to the model output (or the measurements) and then proceeding with the estimation procedure has also been advocated in literature; see for example [45]–[47].

Lastly, the parameter estimates obtained by the proposed method are clinically reasonable: for example, the ventricular wall volume measured by MRI for this patient is 34.2 ml, whereas that estimated by the method is 33.9 ml [17]. Similarly, the SV volume variation was measured via MRI between 10 ml and 29 ml, whereas that reproduced by the model is between 10.7 ml and 24.5 ml [17]. Measurements such as ventricular pressure which are not employed for parameter estimation also show a good agreement with the model output [17]. The parameter estimates and model output are further validated through other qualitative measures, such as echocardiography, Doppler velocimetry, E/A ratio for ventricular filling, and pulmonary venous wedge pressures; for details see [17].

2) Enforcing ventricular volume extrema to model a case of valve regurgitation: Patient-B includes atrioventricular valve regurgitation due to prolapse. In order to correctly estimate heart model parameters with a regurgitant atrioventricular valve, measurements pertaining to heart-chamber volumes are necessary. In the absence of direct measurements of regurgitation, such as area and flow-rate measurements at the atrioventricular valve, if measurements of chamber volumes are not provided, the UKF has no information pertaining to the valve being regurgitant, and consequently, a parameter estimation run yields a non-regurgitant valve where $\int_0^T q_{\text{AO}}(t)dt = \text{EDV} - \text{ESV}$. In fact, even in the clinic regurgitant blood volume is assessed via the difference between EDV-ESV and $\int_0^T q_{\text{AO}}(t)dt$. Consequently, it is critical to include measurements of ventricular volume along with measurements for $q_{\text{AO}}(t)$ (or alternatively $q_{\text{AVV}}(t)$) to model regurgitation. While conductance catheters are the gold standard for measurement of time-varying ventricular volume information, this is rare in routine clinical practice. Commonly, the EDV and ESV are estimated from echocardiography, but more accurate measurements are possible from 3D echocardiography and MRI. These result in volumes at specific points of the cardiac cycle, such as end-diastole and end-systole, and not volume changes over the full cycle. Consequently, in addition to the methods employed for patient-A, the EDV and ESV measurements are included in the UKF to provide information on valve regurgitation through the approach presented in section II-E. With these methods, the evolution of parameter estimate means for the 34 model parameters are shown in Figure 6 of [17]. The associated decrease in the variance of the parameter estimates is shown in Fig. 5. For this patient, all variances of the parameter estimates are reduced by at least a factor of four when the UKF is run for 40 cardiac cycles. Varying degrees of decrease in the parameter variances, however, is apparent, which is likely to be a result of difference in sensitivities. Due to the inclusion of EDV and ESV volume measurements, the clinically measured regurgitation fraction of 25.3% is well captured in the model which produces a regurgitation fraction of 22.7%. For further validation of the parameter estimates, through direct MRI measurements and indirect electrocardiogram and wedge pressure measurements, see [17] and the discussion in section III-B2.

B. Inverse problems with heart-rate variability

Results for heart-rate variability are first presented for a synthetic data-set followed by a patient-specific case.

1) Synthetic data-set: To demonstrate the efficacy of the proposed UKF method in taking measurements at different heart rates into account (see section II-G), a further reduced system of circulation is constructed. This system lumps all the systemic and pulmonary circulations into a single block and is shown in Fig. 6. For the heart activation functions, it is assumed that the ventricle activation duration $t_{\text{AV}}$ and the ratio $t_{\text{AVA}}/T = t_{\text{AVA}}/60$ remain invariant with changes in heart-rate. This implies that at higher heart rates, the duration of ventricle activation remains constant while the decrease in heart-beat time-period is reflected in a decrease in $t_{\text{AVA}}$. Such an assumption is justified by clinical studies, for example see [48], which report that at higher heart-rates, the decrease in systolic time-interval is significantly smaller than the decrease in diastolic time-interval. Synthetic observations are generated by running the forward model with known parameter values at two different heart rates of 120 and 150 bpm. Pressure output is stored from the lower HR model and flow-rate output for the higher HR model. Small noise (2% of maximum values) is added to these two sets of outputs to generate measurements for parameter estimation. The combined set of measurements corresponds to $[p_{\text{AO}}, p_{\text{LB}}, p_{\text{SA}}, p_{\text{SV}}, q_{\text{AVV}}, q_{\text{AOV}}, q_{\text{SA}}]$. The 16 parameters to be estimated are $[V_{\text{SA}}, V_{\text{SA}}, V_{\text{SV}}, V_{\text{SV}}, V_{\text{AVV}}, V_{\text{AVV}}, A_{\text{eff,AVV}}, A_{\text{eff,AVV}}, C_{\text{AO}}, R_{1c}, R_{2c}, R_{3c}, K_{\text{LC}}, L_{1c}, C_{1c}, C_{2c}, t_{\text{ASA}}, t_{\text{AVA}}]$. Fig. 7 shows how the two model states at different heart rates are corrected by the UKF. As expected, in each heart-rate, the model state that is observed is corrected much faster than the unobserved state. For example, in the lower-heart rate model, since only pressures are observed, the flow-rates are corrected much slower than the pressures. Nonetheless, within 20 cardiac cycles (of lower HR model), the UKF parameters converge very close to the true parameters. This is shown in Fig. 8. Finally, the evolution of UKF parameter variances for two cardiac cycles of the lower HR model are shown in Fig. 9. The results display a consistent decrease in the parameter variances with the observation of measurements. Here, within 2 cardiac cycles the reduction in the variance for all parameters is, at least, more than a factor of $2^4$, with the exception of $K_{\text{LC}}$ where this factor is, at least, $2^8$.

2) Patient-specific case: In [17] it is noted that in patient-B the pressure and flow-rate measurements were acquired at different heart rates of 106 and 140 bpm, respectively. In that study the heart beat-time-period of the pressure measurements was artificially shortened to be consistent with the flow-rate measurements. Here, to palliate this shortcoming, the method to deal with HR variability presented in section II-G is applied to model the circulation of patient-B. The aforementioned hypothesis that $t_{\text{ASA}} = t_{\text{AVA}}/T = t_{\text{AVA}}/60$ remains constant
between different heart rates is made and $t_{aSA}$ is estimated as opposed to $t_{aSA}$ in [17]. The results of the UKF estimation are presented in Fig. 10. The evolution of estimate means and variances shows that all parameters are converged and that their variances decrease by a factor of at least $2^5$. A comparison of model output (with UKF-estimated parameters) with the measurements is shown in Fig. 11. Comparing this figure to the results of [17] (reproduced in this figure in solid grey), it is observed that the measured venous flows, particularly $q_{SVC}$ and $q_{IVC}$ are significantly better reproduced when HR variability is taken into account. It is also observed that the agreement between ventricular pressure in model output and the measurements is improved, particularly in relation to the duration of ventricular mechanical activation. The change of PV-loop from lower HR (solid line) to higher heart rate (dashed-line) is consistent with the assumption of $t_{aSA}$ and $t_{rSV}$ remaining constant when HR changes. It appears that at higher heart rates there is restricted ventricular filling time, resulting in lower EDV. Consequently, there is lower ventricular preload and hence peak systolic pressure generation is reduced. Overall, due to better modelling of physics it is observed that consideration of HR variability improves the parameter estimates and the agreement between model output and the measurements, albeit at a higher computational cost (see section II-G). It should be noted that the force-frequency relationship [49] which suggests that cardiac muscles produce higher forces at higher frequencies is not included in the model, and it is likely that inclusion of this phenomenon will result in further improvement of the results. This, however, is an area of future investigation.

Validation of parameter estimation procedure can be performed in two ways: first, to consider clinical estimates of certain measurable parameters such as ventricular wall volumes (measured by MRI), valve annulus areas, and mechanical activation times of heart chambers (qualitatively assessed by ECG readings); and second, by considering measurements of model output that were not employed in the parameter estimation procedure. The latter can include qualitative measures such as pulmonary venous wedge pressures and measures derived through echocardiography, and quantitative measures such as ventricular pressure and volume variations. For patient-B, a two-fold qualitative validation is performed: first, by the range of variation of pulmonary venous wedge pressure (measured to be $\sim 5.5$ mmHg vs. $\sim 4$ mmHg produced by the model); and second by mechanical activation times of the single atrium and single ventricle which show an overlap between their contractions (implying that there is a significant time-period when both the atrium and ventricle are contracting, which is indeed qualitatively observed in the ECG readings depicting the merging of the T and P waves). Quantitative measurement is also performed in two ways: first, by the measurement of ventricular pressure tracing that was not used for parameter estimation (see Figure 10, second last plot on the right column, which shows this measurement in green and the model output in solid blue); and second by the measurements of end systolic and end diastolic ventricle volumes measured by MRI to be 13 ml and 30 ml, respectively, vs. model output of 12.7 ml and 30.3 ml, respectively.

When only a single heart-rate is considered, see [17], with 20 parallel processes the UKF takes approximately 140 s for 200 discrete observations in one cardiac cycle. The same computation time when HR variability is considered is approximately 200 s.

While the proposed method works quite well, in certain cases it may not be possible to write a closed system of equations for the dynamical system without using a noisy measurement (for instance, as a forcing/boundary condition). This method may fail in such cases. The method may also not be applicable when only min./max./mean values are available as measurements without any full time-varying measurement at all. In such cases, it may be advantageous to consider other Bayesian methods for parameter estimation, such as Markov chain Monte Carlo (MCMC), see for example [19]. Note that methods such as MCMC may also be used in general estimation problems, such as those studied in this manuscript, albeit at a higher associated computational cost. Lastly, if the measurement noise is coloured (perhaps, due to a correlated structure), then appropriate modifications, see for example [50], may be needed to the proposed method.

IV. Conclusion

Several challenges of the UKF method for parameter estimation in haemodynamics problems are tackled. These include the prevention of Sigma Point collapse to avoid ill-conditioning of the empirical covariance matrix, inclusion of minimum/maximum measurements in the method, and the choice of appropriate prior knowledge. Furthermore, an extension to the basic UKF method to assimilate measurements at different heart rates is presented. The methods are demonstrated on two patient-specific cases (one with and one without atrioventricular valve regurgitation). The efficacy of the proposed extension to the UKF method for assimilating measurements at different heart-rates is first shown for a synthetic dataset as a proof-of-concept where the true parameters are known, and then on a patient-specific case with atrioventricular valve regurgitation. In the latter case 34 parameters are estimated and the model predictions are shown to be significantly improved when heart-rate variability is taken into account. The proposed methods enable analysis of patient-specific physiologies in a computationally efficient manner while taking measurement uncertainty into account. Lastly, these methods can be readily applied to other pathophysologies.

COMPETING INTERESTS

We have no competing interests.

AUTHORS’ CONTRIBUTIONS

SP carried out the numerical experiments and drafted the manuscript. SP, CC, GP, and IVC developed the mathematical models. SP, GP, and IVC conceived the study. CB and TYH provided medical measurements, clinical input, and interpretation of results. All authors gave final approval for publication.
Fig. 11. Forward model with the UKF-estimated parameters when considering heart-rate variability for patient-B presented in [17]. Unlabelled x-axis: time (s), higher heart-rate: 140 bpm, lower heart-rate 106 bpm, model output at higher heart-rate: dashed blue, model output at lower heart-rate: solid blue, measurements provided to UKF: red, measurements not provided to UKF: green, pressures (p) are in mmHg, flow-rates (q) are in ml/s, volumes (V) are in ml, flow-rate measurements are at higher heart-rate, pressure measurements are at lower heart-rate, model output from single heart-rate study [17]: grey.
TABLE I
NOMENCLATURE OF THE PARAMETERS ESTIMATED

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>Single atrium (see Fig. 1)</td>
</tr>
<tr>
<td>SV</td>
<td>Single ventricle (see Fig. 1)</td>
</tr>
<tr>
<td>AVV</td>
<td>Atrioventricular valve (see Fig. 1)</td>
</tr>
<tr>
<td>AOV</td>
<td>Aortic valve (see Fig. 1)</td>
</tr>
<tr>
<td>LB</td>
<td>Lower body (see Fig. 1)</td>
</tr>
<tr>
<td>LU</td>
<td>Lungs (see Fig. 1)</td>
</tr>
<tr>
<td>UB</td>
<td>Upper body (see Fig. 1)</td>
</tr>
<tr>
<td>LC</td>
<td>Lumped circulation (see Fig. 6)</td>
</tr>
<tr>
<td>$R_{1(\cdot)}$</td>
<td>Resistance (linear dependence of pressure on flow-rate) of large arteries in segment (\cdot)</td>
</tr>
<tr>
<td>$R_{2(\cdot)}$</td>
<td>Resistance (linear dependence of pressure on flow-rate) of the vascular bed in segment (\cdot)</td>
</tr>
<tr>
<td>$R_{3(\cdot)}$</td>
<td>Resistance (linear dependence of pressure on flow-rate) of veins in segment (\cdot)</td>
</tr>
<tr>
<td>$K_{(\cdot)}$</td>
<td>Resistance (quadratic dependence of pressure on flow-rate) of large arteries in segment (\cdot)</td>
</tr>
<tr>
<td>$L_{1(\cdot)}$</td>
<td>Inductance (due to blood inertia) of the larger arteries in segment (\cdot)</td>
</tr>
<tr>
<td>$C_{1(\cdot)}$</td>
<td>Compliance of the large arteries in segment (\cdot)</td>
</tr>
<tr>
<td>$C_{2(\cdot)}$</td>
<td>Compliance of veins in segment (\cdot)</td>
</tr>
<tr>
<td>$C_{SH}$</td>
<td>Compliance of the systemic-to-pulmonary shunt</td>
</tr>
<tr>
<td>$V_{0(\cdot)}$</td>
<td>Volume of chamber cavity (\cdot) at zero transmural pressure</td>
</tr>
<tr>
<td>$V_{w(\cdot)}$</td>
<td>Wall volume of chamber (\cdot)</td>
</tr>
<tr>
<td>$A_{\text{max}}^{\text{eff}(\cdot)}$</td>
<td>maximum effective area of valve (\cdot) in normal function</td>
</tr>
<tr>
<td>$A_{\text{max}}^{\text{eff}(\cdot)}$</td>
<td>maximum effective area of valve (\cdot) during regurgitation due to prolapse</td>
</tr>
<tr>
<td>$C_{AO}$</td>
<td>Aortic compliance</td>
</tr>
<tr>
<td>$E_{a(\cdot)}$</td>
<td>Shape controlling exponent for activation of chamber (\cdot), see equation (5)</td>
</tr>
<tr>
<td>$t_{a(\cdot)}$</td>
<td>Activation duration of chamber (\cdot)</td>
</tr>
<tr>
<td>$t_{a_{SA}}^r$</td>
<td>ratio of activation duration of the atrium to the heart time period, $T$: $t_{a_{SA}}^r = t_{a_{SA}} / T$</td>
</tr>
<tr>
<td>$t_1$</td>
<td>Overlap between the activations of single ventricle and single atrium, see Fig. 2</td>
</tr>
<tr>
<td>$T_{p0}$</td>
<td>Scaling factor for passive fibre stress, see equation (7)</td>
</tr>
</tbody>
</table>
REFERENCES


