

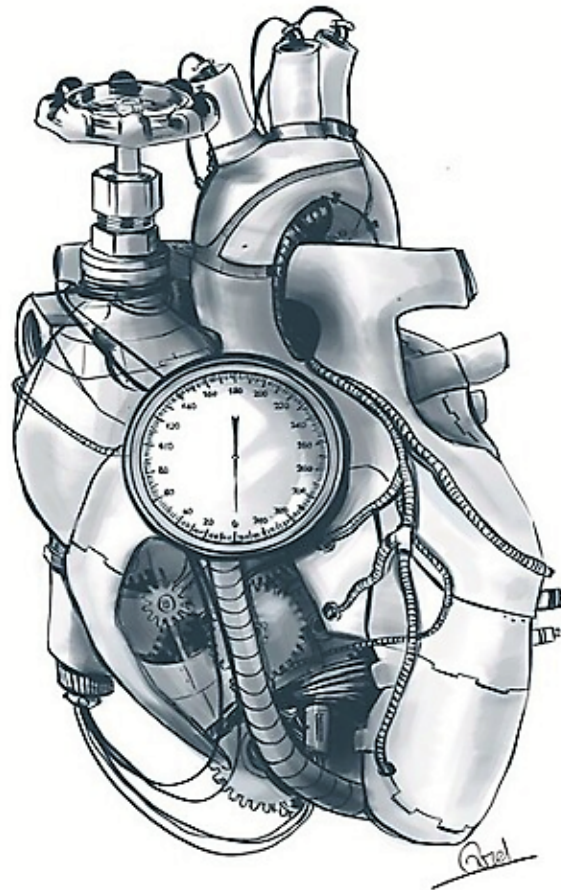
DEPARTMENT OF BIOMEDICAL ENGINEERING  
CARDIOVASCULAR BIOMECHANICS  
MSc Medical Engineering

# Real-time estimation of cardiac model parameters

*“Towards model-based support of weaning from Extracorporeal Membrane Oxygenation”*

S.C. Snijders

1038947



Committee:

DR. IR. W. HUBERTS (*Supervisor*)

DR. IR. R.G.P. LOPATA

DR. IR. M.C.M. RUTTEN (*Supervisor*)

DR. IR. M. VAN 'T VEER

PROF. DR. IR. F. VAN DE VOSSE

**TU/e** EINDHOVEN  
UNIVERSITY OF  
TECHNOLOGY

EINDHOVEN (THE NETHERLANDS), MAY 31, 2022

---

# Nomenclature

agPCE	Adaptive generalized polynomial chaos expansion
AV	Aortic valve
DA	Data assimilation
ECMO	Extracorporeal membrane oxygenation
EKF	Extended Kalman filter
hTEE	hemodynamic Transesophageal echocardiography
HR	Heart rate
KF	Kalman filter
LA	Left atrium
LV	Left ventricle
LVEF	Left ventricular ejection fraction
MLE	Maximum likelihood estimation
MV	Mitral valve
RMSE	Root-mean-square error
ROUKF	Reduced-order unscented kalman filter
RA	Right atrium
RV	Right ventricle
SVD	Singular value decomposition
PDF	Probability density function
PL	Profile likelihood
PA	Pulmonary artery
PV	Pulmonary valve
TEE	Transesophageal echocardiography
TV	Tricuspidal valve
Tr	Trace
UKF	Unscented Kalman filter
UT	Unscented transform
VA-ECMO	Veno-arterial Extracorporeal membrane oxygenation
VC	Venous compartment
VCI	Vena cava inferior
VV-ECMO	Veno-venous Extracorporeal membrane oxygenation

---

# Abstract

**Introduction:** Veno-arterial extra corporeal membrane oxygenation (VA-ECMO) is a complex salvage technique used to provide cardiac and respiratory support to critically ill patients. However, the mortality rates of patients undergoing VA-ECMO remains high and it is therefore desirable to establish a standardized protocol that could assist in determining when to initiate weaning and optimizing the weaning process itself. Consequently, a method is sought that can accurately predict in real-time if the heart is capable of independently providing the required perfusion, ideally based on minimal to non-invasive measurements. **Method:** An algorithm has been developed in the current study based on the reduced-order Unscented Kalman filter (ROUKF), which is capable of estimating a cardiac functionality parameter; left ventricular contractility. The proposed ROUKF-based algorithm uses a 0D lumped element model of the left ventricle to obtain simulations of the behavior of the cardiovascular system, which are subsequently merged with left ventricle pressure measurements. The measurements used in the current study originate from three types of data: synthetic, in vitro or in vivo. The output of the algorithm was evaluated by using a sensitivity and identifiability analysis, while a stability analysis was used to assess the stability of the mathematical model. **Results and discussion:** First of all, the proposed ROUKF-based algorithm was applicable to all three types of data. When applied to synthetic data, the algorithm was able to converge to the true parameter in real-time, while being independent of initial parameter conditions and being robust for fluctuating parameter conditions. Furthermore, the measurement was only sensitive to changes in the parameter during the contraction phase, with the parameter being identifiable from the measurement until a reasonable level of noise. Lastly, the used mathematical model was considered to be stable for a physiologically feasible range of the parameter. However, the results of the algorithm application to the in vitro data showed that when the model discrepancy increases, the accuracy of the parameter estimation reduces. Therefore, future research needs to be conducted to evaluate the performance of the algorithm when the complexity of the mathematical model increases or additional measurements are provided, with the holy grail being an algorithm that only uses non-invasive measurements to provide an accurate and real-time estimate. **Conclusion:** The proposed ROUKF-based algorithm was capable of providing a patient-specific estimation of the left ventricular contractility in real-time based on the left ventricular pressure data. Therefore, the proposed algorithm is considered to be promising for the development of a standardized weaning protocol, but needs further optimization to be applicable in clinical practice.

**Keywords:** *parameter estimation, real-time, ROUKF, VA-ECMO, weaning.*

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Extra corporeal membrane oxygenation . . . . .	1
1.2	Data assimilation . . . . .	3
1.3	Aim and outline . . . . .	4
<b>2</b>	<b>Data assimilation</b>	<b>5</b>
2.1	Kalman filter . . . . .	5
2.1.1	Assimilation scheme . . . . .	5
2.1.2	Kalman filter algorithm . . . . .	9
2.2	Unscented Kalman filter . . . . .	10
2.2.1	Unscented transform . . . . .	10
2.2.2	Unscented Kalman filter algorithm . . . . .	12
2.3	Reduced-order unscented Kalman filter . . . . .	14
2.3.1	Assimilation scheme . . . . .	15
2.3.2	Simplex sigma-points . . . . .	18
2.3.3	Reduced-order unscented Kalman filter algorithm . . . . .	19
<b>3</b>	<b>Model and measurements</b>	<b>21</b>
3.1	Mathematical model . . . . .	21
3.2	Measurements . . . . .	24
3.2.1	Synthetic data . . . . .	24
3.2.2	In vitro data . . . . .	25
3.2.3	In vivo data . . . . .	27
<b>4</b>	<b>Data processing</b>	<b>29</b>
4.1	Probability density function . . . . .	29
4.2	Moving time window . . . . .	31
<b>5</b>	<b>Data and model analysis</b>	<b>33</b>
5.1	Local sensitivity analysis . . . . .	33
5.2	Identifiability analysis . . . . .	34
5.3	Stability analysis . . . . .	36
5.3.1	First and second method of Lyapunov . . . . .	36
5.3.2	Lyapunov equation . . . . .	37
5.3.3	Krasovskii's method . . . . .	39

<b>6 Results</b>	<b>41</b>
6.1 Synthetic data . . . . .	41
6.1.1 Sensitivity analysis . . . . .	43
6.1.2 Identifiability analysis . . . . .	43
6.1.3 Stability analysis . . . . .	44
6.2 In vitro data . . . . .	44
6.3 In vivo data . . . . .	46
<b>7 Discussion</b>	<b>48</b>
7.1 Major findings . . . . .	48
7.2 Limitations . . . . .	49
7.3 Future work . . . . .	50
<b>8 Conclusion</b>	<b>52</b>
<b>References</b>	<b>52</b>
<b>A Log-normal distributed PDF</b>	<b>58</b>
<b>B Global sensitivity analysis</b>	<b>60</b>
<b>C Decision rule of profile likelihood</b>	<b>63</b>
<b>D Initialized <i>a posteriori</i> estimate</b>	<b>64</b>
<b>E Results identifiability analysis</b>	<b>65</b>
<b>F Results in vitro data</b>	<b>67</b>
<b>G Results in vivo data</b>	<b>70</b>

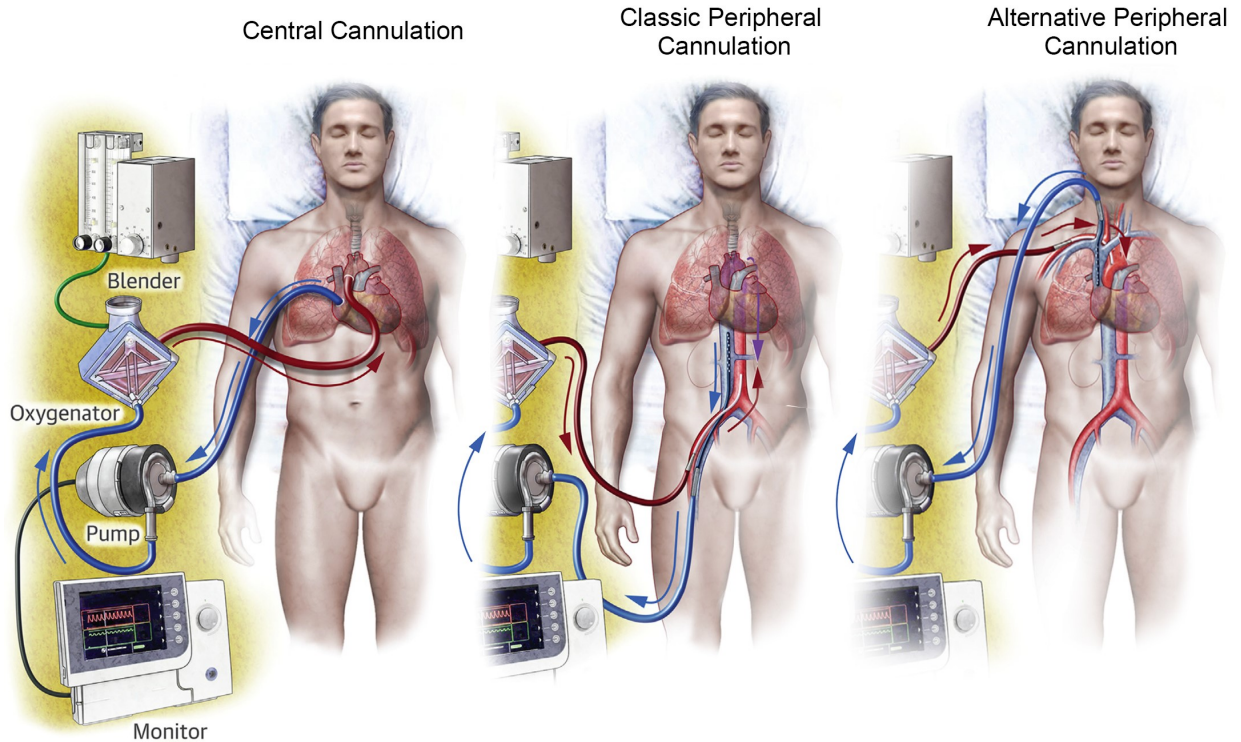
# 1. Introduction

## 1.1 Extra corporeal membrane oxygenation

Extracorporeal membrane oxygenation (ECMO) is a complex salvage technique used to provide cardiac and/or respiratory support to seriously ill patients in which maximal conventional medical treatment is considered to be inadequate [23]. This ECMO technique can be deployed in two distinct configurations: the veno-venous ECMO (VV-ECMO) and the veno-arterial ECMO (VA-ECMO) [47]. The VV-ECMO configuration is used to supply respiratory support and can therefore be considered as an alternative to classical mechanical ventilation [9]. In contrast, the VA-ECMO configuration is used to provide cardiac support, as well as to simultaneously supply both cardiac and respiratory support. [23, 47]. This VA-ECMO configuration can be utilized to establish a cardiopulmonary bypass, which can be useful when applying cardiac surgery [38]. Secondly, this configuration can also be used to allow cardiopulmonary recovery in order to gain time until certain therapeutic alternatives can be executed [56].

Generally, the main indication for deploying VA-ECMO in clinical practice remains cardiogenic shock [38, 47]. Cardiogenic shock is a patient state in which a low cardiac output results in hypoperfusion and hypoxia of critical organs. The most frequent cause of cardiogenic shock is an acute myocardial infarction combined with left ventricular dysfunction, but it could also result from myocarditis, acute decompensated heart failure, drug intoxication, hypothermia or arrhythmia [74, 38]. In cardiogenic shock patients, VA-ECMO provides cardiopulmonary support as a bridge to either myocardial recovery, implementation of a mechanical circulatory support device or a heart transplant [34].

The VA-ECMO circuit finds its origin in draining deoxygenated blood from the venous circulation and re-infusing oxygenated blood to the arterial circulation, both through a cannula [23]. When comparing different VA-ECMO circuits, a distinction can be made between the location of cannula insertion, which can be either peripheral or central [38]. Both of these cannulation approaches are depicted in Figure 1.1, in which also the different components of the ECMO circuit are illustrated.



**Figure 1.1:** Illustration of the three cannulation approaches used in VA-ECMO, with the components of the VA-ECMO circuit being depicted as well. Adapted from: ‘Venoarterial Extracorporeal Membrane Oxygenation in Cardiogenic Shock’ by M. E. Keebler et al, 2018, *Heart Failure*, 6, 503–516 [34].

The central configuration of VA-ECMO needs to be established surgically, involving blood drainage from a cannula inserted in the right atrium and re-infusing it through a cannula in the ascending aorta. In the classic peripheral configuration of VA-ECMO, blood is drained from a cannula located in the femoral vein and returned through a cannula in the femoral artery. [38]. Lastly, in the alternative peripheral configuration of VA-ECMO, blood is drained from a cannula inserted in the interval jugular vein and re-infused through a cannula in the axillary artery [34]. In both configurations, deoxygenated blood is drained from the venous circulation and drawn into a pump by a negative pressure. Next, the pump exerts a positive pressure to the blood and propels it to the oxygenator, where the blood is oxygenated and decarboxylated. This oxygenation and decarboxylation is achieved by flushing the blood over bundles of hollow fibers, which are fueled by an air/oxygen blender, allowing the gas exchange to occur. Eventually, the oxygenated and decarboxylated blood leaves the oxygenator under an appropriate positive pressure, driving it back to the arterial circulation [9].

The combination of the underlying disease and the duration of the VA-ECMO puts already critically ill patients at risk for developing serious complications. These complications can have different origins which can roughly be classified in the following categories: bleeding and coagulation, vascular problems, lower extremity ischemia, neurological disorders, systemic inflammation or severe infection, hemodynamic complications, cannulation complications and device failure [35]. Statistically, both the complication rates and the mortality rates of VA-ECMO remain high [35, 48, 56]. According to the study of Papadopoulos et al. (2015), the hospital survival rate after ECMO remains only between 20% to 40% [56]. Another study, Maruscalco et al. (2021), reported a similar survival rate of around 35 % [48]. However, these mentioned survival rates differ from the number of patients that are decoupled successfully from VA-ECMO, with decoupling referred to as weaning in clinical practice [48, 70]. According to the study of Maruscalco et al. (2021), 54% of the patients were weaned successfully from VA-ECMO and 35% of these successfully weaned patients died before discharge from the hospital [48]. In the study of Tohme et al. (2021) a successful weaning was reported in 31% to 76% of the patients, whereas 20% to 65% of the patients died before hospital discharge [70]. A possible explanation for the death of patients after successful weaning may be found in inadequate myocardial recovery, end-organ failure, neurological damage or other comorbidities [43, 57].

Weaning protocols for VA-ECMO differ substantially between centers of expertise. Given the previously mentioned relatively high mortality of patients undergoing VA-ECMO, it is considered important to establish a standardized protocol which helps to identify when to initiate weaning and, in addition, to optimize the weaning process itself [43]. Basically, weaning can be initiated when the heart is capable of providing the required perfusion independently. In the current clinical practice, a Transesophageal Echocardiography (TEE) is performed to identify when weaning can be initiated safely [57, 66]. Based on this TEE technique, the left ventricular ejection fraction (LVEF) can be evaluated, which is used as the main threshold parameter to determine when to initiate weaning [66]. However, this LVEF is considered to be a moderate parameter to verify if the heart is capable of providing the required perfusion independently, because it is based on cardiac volumes rather than cardiac functionality parameters. Moreover, another limitation of this TEE technique is that it can only present a single measurement in a certain time interval [2].

The hemodynamic TEE (hTEE) is an improved miniature version of the conventional TEE, which is capable of evaluating contractility and filling of both right ventricle (RV) and left ventricle (LV) at the bedside of the patient in real-time. The ability to assess certain cardiac functionality parameters in real-time implies that this technique could be used as a guide throughout the weaning process, which is promising for the development of a standardized weaning protocol [15]. However, the evaluation of contractility by using echocardiography is considered to be unreliable due to VA-ECMO-induced hemodynamic circumstances [57]. Another downside is that the administration of intravenous sedation to the patient is required throughout the total weaning attempt to ensure the toleration of the transesophageal probe [1]. Furthermore, successful weaning requires an adequate myocardial recovery and end-organ function, while echocardiographic monitoring appears to be unreliable for evaluation of this ventricular recovery. Therefore, hTEE is considered to be incapable of determining whether VA-ECMO can be weaned successfully or not [43].

Hence, there is a need for a method that can assess myocardial recovery in order to improve the timing of weaning initiation and consequently the success of weaning. This method must be able to predict whether the heart is capable of providing the required perfusion independently. Based on this capability, the method can be considered as a promising tool to determine the initiation of weaning, while it can also provide valuable information during the weaning process. Furthermore, this desired method should ideally be minimally invasive to non-invasive and accurate, applicable in real-time, to act as a guide throughout the weaning process. In case of fulfilling all mentioned properties, the method can be regarded as valuable for the development of a standardized weaning protocol, which can improve the weaning from VA-ECMO and consequently reduce the mortality rate of VA-ECMO.

## 1.2 Data assimilation

When a patient is connected to VA-ECMO, several signals like the ECMO flow are mean arterial pressure (MAP), are constantly monitored [16]. These available signals provide continuous information about the circulatory state of the patient, with information about the myocardial state of the patient hidden inside these signals. A mathematical model can be used to describe the myocardial state of a patient. However, a mathematical model is not sequentially adapted to actual patient-specific information and is therefore considered to be unsuitable to describe the myocardial state of a certain patient. A method called data assimilation (DA) could provide a solution, as it is based on merging sparse and possibly noisy measurements with the simulation of a mathematical model to obtain an optimal combination of these two aspects [3]. Given the intended application, the mathematical simulation should be a physics-based model that describes the behavior of relevant parts of the cardiovascular system. In this particular model, the circulatory system is considerably simplified with a fixed pulmonary pressure representing the preload and a transmission line of three-element Windkessel models representing the afterload of the left ventricle [49]. This mathematical model of the cardiovascular system is able to describe the hemodynamics of the LV, which will be combined with measurements to obtain unknown parameters through DA.

There are three different types of DA methods: variational, sequential and hybrid. All these DA methods seek for an optimal solution based on different principles. The variational form is derived from variational calculus and is based on numerically minimizing a cost function that represents the observational error, usually the difference between the model and the measurements [3]. Alternatively, the sequential form aims to minimize the variance of the model prediction of every time step by solving a direct algebraic function, while updating this prediction based on the between the real-time measurement and the prediction itself. [3, 6, 59]. Lastly, the hybrid form refers to a variational- and a sequential DA method being intertwined. In this form, these two DA methods run simultaneously and exchange information about uncertainties of the model-based estimates of states or parameters concerning the given system [5, 59].

The sequential DA is capable of estimating a system state or parameter when a measurement becomes available, which is considered appropriate for the real-time execution of the desired method [60]. In addition to its real-time capabilities, the sequential method is considered favorable because it also provides information about the probability distribution of the solution, which is crucial in clinical decision making. One of the most often used sequential DA methods is the Kalman filter, which provides a recursive solution based on discrete linear data [4, 33]. However, the evolution of real-life systems is typically governed by a set of non-linear equations, which significantly restrict the validity of the linear oriented Kalman filter [46]. This is also likely to apply to the intended application, which results in deficiency of the classical Kalman filter approach.

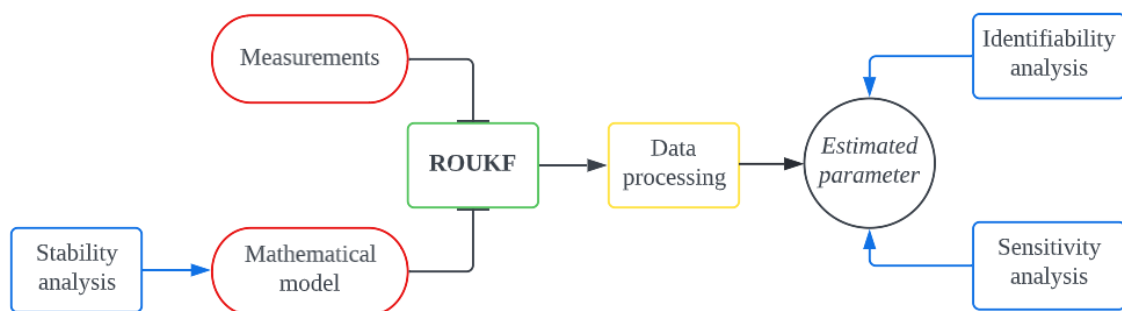
Several extensions for the Kalman filter that can deal with these non-linearities are available, one of them is the unscented Kalman filter (UKF) [46, 77]. However, a Kalman filter with non-linear extensions is known for its computational expensiveness when applied to problems with large amounts of data. An effective approach to circumvent this computational problem is to apply the suggested UKF only to estimate the unknown model parameters. Nevertheless, information about the states of the system should also be included in the used DA approach, because they contain valuable insights about the system at hand. Ideally, these system states should be addressed with a less computationally intensive DA method to be a potential solution to the computational problem [17]. The idea of combining two DA methods, one for parameter estimation and one for state estimation, into a hybrid DA approach to estimate states and parameters simultaneously is known as joint state-parameter estimation [52]. By combining a state estimator and the UKF through the joint state-parameter estimation approach, a reduced-order version of the UKF can be derived. This version is called the reduced-order unscented Kalman filter (ROUKF) and will be used in the current study to perform patient-specific parameter estimation [51].



### 1.3 Aim and outline

The aim of the current study is to provide a proof of principle for using the ROUKF method to establish a virtual patient-specific representation of an unknown cardiac functionality parameter of the cardiovascular system. This virtual patient-specific representation is called a digital shadow, a digital object with a unidirectional data flow between real-world object and the digital object. In such a digital shadow, a change in the state of the real-world object will result in a change in the state of the digital object, which is not true vice versa [37]. In the current study, this digital shadow needs to be capable of estimating the cardiac functionality parameter accurately, in real-time, and based on minimal to non-invasive data to be considered as the ideal tool for weaning. To perform this parameter estimation in real-time, the computational expensiveness of the used algorithm needs to be minimized. In addition, minimizing the number of different measurement types is favorable because it will minimize both the potential invasiveness and the number of actions required by the clinician. After the development of a digital shadow capable of achieving the elaborated objectives, it will be used to evaluate: synthetic, in vitro and in vivo data. The performance of the digital shadow will be quantified by performing three distinct analyses. First of all, a sensitivity analysis will be conducted to identify the sensitivity of the measurement to changes in the parameter of interest. Secondly, an identifiability analysis will be carried out to evaluate the possibility of identifying a unique parameter value from the available measurements. Lastly, a stability analysis will be executed to quantify the sensitivity of the system to its initial states. Ultimately, the goal of the current study is to develop a digital shadow based on ROUKF which is capable of estimating patient-specific cardiac functionality parameters in real-time with an acceptable confidence interval, using a minimal number of different measurement types to serve as the ideal guide during weaning.

The outline of the current study will be as follows. First, in Chapter 2 the Kalman filter theory will be elaborated, starting with the basic KF, then the UKF and eventually the ROUKF. In Chapter 3 the mathematical model of the cardiovascular system and gathering of measurements from synthetic, in vitro and in vivo data will be discussed. Subsequently, in Chapter 4 the data processing will be detailed, followed by the elaboration of data and model analysis in Chapter 5, which includes the sensitivity, identifiability analysis and stability analysis. The interactions among the topics in the chapters related to the method part of the current study, Chapter 2 through Chapter 5, are shown in Figure 1.2. Lastly, in Chapter 6 the results of the previously mentioned chapters will be illustrated, which will be discussed in Chapter 7 to eventually draw conclusions in Chapter 8.



**Figure 1.2:** Overview of the interaction between the subjects discussed in the Chapter 2 till Chapter 4, in which the subjects of Chapter 2 are depicted in green, Chapter 3 in red, Chapter 4 in yellow and Chapter 5 in blue.

## 2. Data assimilation

This chapter will be devoted to a comprehensive theoretical exposition of the Kalman filter theory [33]. The first section of the chapter discusses the Kalman filter (KF), which is a sequential DA method that provides a recursive solution based on merging measurements and model predictions [6]. The unscented Kalman filter (UKF) is a non-linear extension to the conventional KF, which is elaborated in the second section of this chapter. In the third and final section of this chapter the reduced-order unscented Kalman filter (ROUKF) is detailed, which is a computationally more efficient version of the UKF.

### 2.1 Kalman filter

The Kalman filter is able to estimate the states of a linear dynamic system by merging a mathematical model with time-discrete measurements. This sequential DA method basically consists of two iterative steps, namely, the prediction step based on the mathematical model and the analysis step based on the available measurement. This sequential method is considered to be very efficient because, once a measurement is available, the prediction can be updated easily in case it does not agree with the measurement [5]. The eventual goal of the KF is to minimize the variance of the estimate of every time step by solving an algebraic function [59].

The considered linear dynamic system with additive noise is governed by a set of linear stochastic differential equations, that after time discretisation can be written as:

$$\mathbf{x}_{k+1} = \mathbf{M}_{k+1}\mathbf{x}_k + \mathbf{u}_k, \quad \mathbf{x} \in \mathbb{R}^{n_s}, \quad k = 0, 1, 2, \dots, \tau, \quad (2.1)$$

with measurable output:

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \quad \mathbf{y} \in \mathbb{R}^{n_z}, \quad k = 0, 1, 2, \dots, \tau. \quad (2.2)$$

In the first phrase of the dynamic system description  $\mathbf{x}_k$  denotes the state column vector of discrete time step  $k$ , with a maximum of  $\tau$ , which contains a total of  $n_s$  states ( $\mathbf{x} \in \mathbb{R}^{n_s}$ ). In addition, the behavior of  $\mathbf{x}_{k+1}$  is governed by the state transition matrix  $\mathbf{M}_{k+1}$  and the process noise vector  $\mathbf{u}_k$  in this first phrase. In the second mathematical expression  $\mathbf{y}_k$  represents a vector of  $n_z$  measurements ( $\mathbf{y} \in \mathbb{R}^{n_z}$ ),  $\mathbf{H}_k$  the measurement matrix and  $\mathbf{v}_k$  the measurement noise vector. In the current dynamic system (Equations 2.1 and 2.2), the matrices  $\mathbf{M}_{k+1}$  and  $\mathbf{H}_k$  are both considered to be linear. In this system, it is assumed that both error vectors ( $\mathbf{u}_k$  and  $\mathbf{v}_k$ ) are uncorrelated and contain white noise with a probability distribution that is presumed to be Gaussian:

$$\mathbf{u}_k \sim \mathcal{N}(0, \mathbf{Q}_k) \quad \text{and} \quad \mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k). \quad (2.3)$$

In this expression, matrix  $\mathbf{Q}_k$  denotes the model error covariance and matrix  $\mathbf{R}_k$  the measurement error covariance. The model error covariance  $\mathbf{Q}_k$  is considered to be difficult to determine because it accounts for the model's deficiencies in representing underlying physical mechanisms as well as the cumulative effects of parameter errors. In contrast, the measurement error covariance  $\mathbf{R}_k$  can be determined empirically by estimating the trueness and precision of the used measurement equipment [69].

#### 2.1.1 Assimilation scheme

The sequential assimilation scheme of the KF consists of two distinct steps; the prediction step and analysis step. Starting with the prediction step, where the estimate of the current time step  $\hat{\mathbf{x}}_k$  will be used to estimate the state vector at the next time step  $\hat{\mathbf{x}}_{k+1}$  by using a mathematical model. The resulting estimate of the prediction step is called the *a priori* estimate  $\hat{\mathbf{x}}_{k+1}^p$ , which can mathematically be denoted as:

$$\hat{\mathbf{x}}_{k+1}^p = \mathbf{M}_{k+1}\hat{\mathbf{x}}_k, \quad (2.4)$$

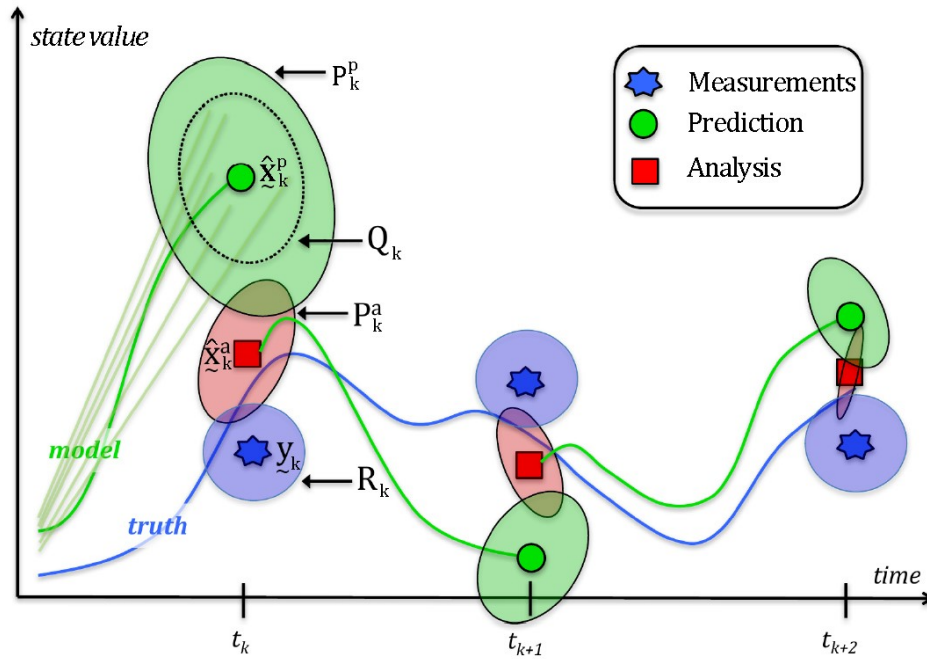
with the *a priori* covariance  $\mathbf{P}_{k+1}^p$  and associated model error covariance  $\mathbf{Q}_k = E[\underline{u}_k \underline{u}_k^T]$ :

$$\mathbf{P}_{k+1}^p = \mathbf{M}_{k+1} \mathbf{P}_{k+1} \mathbf{M}_{k+1}^T + \mathbf{Q}_k. \quad (2.5)$$

In the subsequent step of the Kalman filter, which is the analysis step, the *a priori* estimate will be improved by using the available measurement. This analysis step will result in the *a posteriori* estimate  $\hat{\underline{x}}_{k+1}^a$ , which is defined as:

$$\hat{\underline{x}}_{k+1}^a = \hat{\underline{x}}_{k+1}^p + \mathbf{K}_{k+1}(\underline{y}_{k+1} - \hat{\underline{z}}_{k+1}). \quad (2.6)$$

In this analysis step, the *a priori* estimate  $\hat{\underline{x}}_{k+1}^p$  is thus improved by the discrepancy between the available measurement  $\underline{y}_{k+1}$  and the measurement prediction  $\hat{\underline{z}}_{k+1}$  in order to obtain the *a posteriori* estimate  $\hat{\underline{x}}_{k+1}^a$ . The measurement prediction  $\hat{\underline{z}}_{k+1}$  follows from mapping the *a priori* estimate  $\hat{\underline{x}}_{k+1}^p$  to the measurement space by using matrix  $\mathbf{H}_{k+1}$ , which can formally be stated as:  $\hat{\underline{z}}_{k+1} = \mathbf{H}_{k+1} \hat{\underline{x}}_{k+1}^p$ . The discrepancy  $(\underline{y}_{k+1} - \hat{\underline{z}}_{k+1})$  will be weighted by using the Kalman gain  $\mathbf{K}_{k+1}$ , which will be chosen with the objective to minimize the *a posteriori* error covariance  $\mathbf{P}_{k+1}^a$ . Note that in the current study, the *a priori* elements resulting from the prediction step will be denoted with  $(\cdot)^p$  while the *a posteriori* elements resulting from the analysis step will be indicated with  $(\cdot)^a$ . Lastly, an illustrative representation of the elaborated sequential assimilation scheme of the Kalman filter is given in Figure 2.1.



**Figure 2.1:** A sketch of the sequential assimilation scheme of the Kalman filter in the measurement space of measurements  $\underline{y}_k$ , where matrix  $\mathbf{H}_k$  is omitted for simplicity. In this sketch, the prediction (*a priori*) estimate  $\hat{\underline{x}}_k^p$  and analysis (*a posteriori*) estimate  $\hat{\underline{x}}_k^a$  can be found in the middle of the two ellipses, which depict the corresponding prediction  $\mathbf{P}_k^p$  and analysis  $\mathbf{P}_k^a$  covariance. The *a priori* estimate  $\hat{\underline{x}}_k^p$  will be affected by the model error covariance  $\mathbf{Q}_k$ , while the measurements  $\underline{y}_k$  will be influenced by the measurement error covariance  $\mathbf{R}_k$ . Adapted from: ‘A review of innovation-based methods to jointly estimate model and observation error covariance matrices in ensemble data assimilation’ by P. Tandeo et al, 2020, *Monthly weather review*, 148, 3973-3994 [69].

In the remaining part of the current section, the optimal Kalman gain  $\mathbf{K}_{k+1}$  to fulfill the objective of minimizing the *a posteriori* error covariance  $\mathbf{P}_{k+1}^a$  will be mathematically derived. First, the errors of the *a priori* estimate (prediction step) and *a posteriori* estimate (analysis step) can be defined as the difference between the estimate and the unknown true state  $\mathbf{x}_k^t$ . Based on these two errors, the expected values  $E[\cdot]$  of the *a priori* and *a posteriori* covariance ( $\mathbf{P}_k^p$  and  $\mathbf{P}_k^a$ ) can be determined. Mathematically, these two error-based statements can be formulated as:

$$\underline{\epsilon}_k^p = \hat{\mathbf{x}}_k^p - \mathbf{x}_k^t, \quad \mathbf{P}_k^p = \text{cov}(\underline{\epsilon}_k^p) = E[\underline{\epsilon}_k^p \underline{\epsilon}_k^{pT}] = E[(\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t)(\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t)^T], \quad (2.7)$$

$$\underline{\epsilon}_k^a = \hat{\mathbf{x}}_k^a - \mathbf{x}_k^t, \quad \mathbf{P}_k^a = \text{cov}(\underline{\epsilon}_k^a) = E[\underline{\epsilon}_k^a \underline{\epsilon}_k^{aT}] = E[(\hat{\mathbf{x}}_k^a - \mathbf{x}_k^t)(\hat{\mathbf{x}}_k^a - \mathbf{x}_k^t)^T]. \quad (2.8)$$

The derivation of optimal Kalman gain  $\mathbf{K}_k$  starts with rewriting the expression of the *a posteriori* estimation error  $\underline{\epsilon}_k^a$  (Equation 2.8) by substituting: 1) the expression of the *a posteriori* estimate  $\hat{\mathbf{x}}_k^a$  (Equation 2.6), 2) the measurement estimation  $\hat{\mathbf{z}}_k = \mathbf{H}_k \hat{\mathbf{x}}_k^p$  and 3) the measurement expression (Equation 2.2). This results in:

$$\begin{aligned} \underline{\epsilon}_k^a &= \hat{\mathbf{x}}_k^a - \mathbf{x}_k^t, \\ &= \hat{\mathbf{x}}_k^p + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{z}}_k) - \mathbf{x}_k^t, \\ &= \hat{\mathbf{x}}_k^p + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^p) - \mathbf{x}_k^t, \\ &= \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^p) + (\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t), \\ &= \mathbf{K}_k(\mathbf{H}_k \mathbf{x}_k^t + \mathbf{v}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^p) + (\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t), \\ &= \mathbf{K}_k(-\mathbf{H}_k(\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t) + \mathbf{v}_k) + (\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t). \end{aligned} \quad (2.9)$$

Next, the above derived expression of *a posteriori* estimation error  $\underline{\epsilon}_k^a$  (Equation 2.9) will be substituted into the expression of *a posteriori* covariance  $\mathbf{P}_k^a$  (Equation 2.8). Subsequently, the number of terms will be reduced by substituting the *a priori* estimation error  $\underline{\epsilon}_k^p$  (Equation 2.7). Lastly, after rewriting the resulting equation, the *a priori* covariance  $\mathbf{P}_k^p$  (Equation 2.7) and the measurement error covariance  $\mathbf{R}_k = E[\mathbf{v}_k \mathbf{v}_k^T]$  will be used to obtain an alternative expression of the *a posteriori* covariance  $\mathbf{P}_k^a$ . These mathematical operations lead to:

$$\begin{aligned} \mathbf{P}_k^a &= E[\underline{\epsilon}_k^a \underline{\epsilon}_k^{aT}], \\ &= E[(\mathbf{K}_k(-\mathbf{H}_k(\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t) + \mathbf{v}_k) + (\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t))(\mathbf{K}_k(-\mathbf{H}_k(\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t) + \mathbf{v}_k) + (\hat{\mathbf{x}}_k^p - \mathbf{x}_k^t))^T], \\ &= E[(\mathbf{K}_k(-\mathbf{H}_k \underline{\epsilon}_k^p + \mathbf{v}_k) + \underline{\epsilon}_k^p)(\mathbf{K}_k(-\mathbf{H}_k \underline{\epsilon}_k^p + \mathbf{v}_k) + \underline{\epsilon}_k^p)^T], \\ &= E[(\mathbf{K}_k(\mathbf{v}_k - \mathbf{H}_k \underline{\epsilon}_k^p) + \underline{\epsilon}_k^p)(\mathbf{K}_k(\mathbf{v}_k - \mathbf{H}_k \underline{\epsilon}_k^p) + \underline{\epsilon}_k^p)^T], \\ &= E[(\mathbf{K}_k \mathbf{v}_k - \mathbf{K}_k \mathbf{H}_k \underline{\epsilon}_k^p + \underline{\epsilon}_k^p)(\mathbf{K}_k \mathbf{v}_k - \mathbf{K}_k \mathbf{H}_k \underline{\epsilon}_k^p + \underline{\epsilon}_k^p)^T], \\ &= E[\underline{\epsilon}_k^p \underline{\epsilon}_k^{pT} - (\mathbf{K}_k \mathbf{H}_k \underline{\epsilon}_k^p)(\mathbf{K}_k \mathbf{H}_k \underline{\epsilon}_k^p)^T + (\mathbf{K}_k \mathbf{v}_k)(\mathbf{K}_k \mathbf{v}_k)^T], \\ &= E[\underline{\epsilon}_k^p \underline{\epsilon}_k^{pT}] - (\mathbf{K}_k \mathbf{H}_k) E[\underline{\epsilon}_k^p \underline{\epsilon}_k^{pT}] (\mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k E[\mathbf{v}_k \mathbf{v}_k^T] \mathbf{K}_k^T, \\ &= \mathbf{P}_k^p - (\mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^p (\mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T, \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^p (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T. \end{aligned} \quad (2.10)$$

This expression of the *a posteriori* covariance  $\mathbf{P}_k^a$  can subsequently be used to calculate the Kalman gain  $\mathbf{K}_k$ . However, this Kalman gain  $\mathbf{K}_k$  is not necessarily the most optimal one, and therefore an expression to compute the optimal Kalman gain  $\mathbf{K}_k$  will be derived below. In retrospect, the goal of the Kalman filter is to minimize the variance of the *a posteriori* estimate  $\hat{\mathbf{x}}_{k+1}^a$ . The variances of all the states included in this estimate will be located on the diagonal of the *a posteriori* covariance  $\mathbf{P}_k^a$ , which implies that the minimal sum of these variances can be obtained by minimizing the trace ( $\text{Tr}$ ) of the matrix. Therefore, the optimal Kalman gain  $\mathbf{K}_k$  can be determined by taking the derivative of the trace of matrix  $\mathbf{P}_k^a$  with respect to  $\mathbf{K}_k$  and setting it equal to zero. To determine the derivative of the trace, the following differential matrix expressions are considered useful:

$$\frac{\partial}{\partial \mathbf{A}} \text{Tr}(\mathbf{A}\mathbf{B}) = \mathbf{B}^T, \quad \frac{\partial}{\partial \mathbf{A}} \text{Tr}(\mathbf{A}\mathbf{B}^T) = \mathbf{B}, \quad \frac{\partial}{\partial \mathbf{A}} \text{Tr}(\mathbf{A}\mathbf{C}\mathbf{A}^T) = 2\mathbf{A}\mathbf{C}. \quad (2.11)$$

In these matrix expressions, it is assumed that  $\mathbf{A}$  and  $\mathbf{B}$  are both square matrices and  $\mathbf{C}$  is a symmetric matrix. Before determining the derivative of the trace of the *a posteriori* covariance  $\mathbf{P}_k^a$ , Equation 2.10 will be expanded first:

$$\mathbf{P}_k^a = \mathbf{P}_k^p - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_k^p - \mathbf{P}_k^p \mathbf{K}_k^T \mathbf{H}_k^T + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_k^p \mathbf{H}_k^T + \mathbf{R}_k) \mathbf{K}_k^T. \quad (2.12)$$

Subsequently, the trace of Equation 2.12 will be differentiated with respect to the Kalman gain  $\mathbf{K}_k$  to acquire the most optimal gain. The differentiation of the trace can be denoted as:

$$\frac{\partial}{\partial \mathbf{K}_k} \text{Tr}(\mathbf{P}_k^a) = -2(\mathbf{H}_k \mathbf{P}_k^p)^T + 2\mathbf{K}_k (\mathbf{H}_k \mathbf{P}_k^p \mathbf{H}_k^T + \mathbf{R}_k). \quad (2.13)$$

When Equation 2.13 is set to zero and then solved algebraically, the expression to compute the optimal Kalman gain  $\mathbf{K}_k$  can finally be obtained:

$$\mathbf{K}_k = \mathbf{P}_k^p \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^p \mathbf{H}_k^T + \mathbf{R}_k)^{-1}. \quad (2.14)$$

This expression of the optimal Kalman gain  $\mathbf{K}_k$  (Equation 2.14) can be evaluated mathematically through observing the limits of this expression:

$$\lim_{\mathbf{R}_k \rightarrow 0} \mathbf{K}_k = \mathbf{H}_k^{-1} \quad \text{and} \quad \lim_{\mathbf{P}_k^p \rightarrow 0} \mathbf{K}_k = \mathbf{O}. \quad (2.15)$$

When evaluating the above stated results combined with the expression of the *a posteriori* covariance  $\mathbf{P}_k^a$  (Equation 2.8), it can be observed that when the measurement error covariance  $\mathbf{R}_k$  approaches zero, the Kalman gain  $\mathbf{K}_k$  will weigh the influence of the innovation ( $y_{k+1} - \hat{z}_{k+1}$ ) more heavily. Conversely, when the *a priori* covariance  $\mathbf{P}_k^p$  approaches zero, the Kalman gain  $\mathbf{K}_k$  will weigh the influence of the innovation almost negligibly. These mathematical observations intuitively agree with the expected behavior of a properly functioning sequential DA.

The expression of the optimal Kalman gain  $\mathbf{K}_k$  (Equation 2.14) can be simplified by introducing two additional covariance matrices, the innovation covariance  $\mathbf{P}_k^z$  and the cross covariance  $\mathbf{P}_k^{pz}$ . These two covariance matrices can be defined as:

$$\mathbf{P}_k^z = \mathbf{H}_k \mathbf{P}_k^p \mathbf{H}_k^T + \mathbf{R}_k \quad \text{and} \quad \mathbf{P}_k^{pz} = \mathbf{P}_k^p \mathbf{H}_k^T. \quad (2.16)$$

When using these two expressions (Equation 2.16) to simplify the formula of the optimal Kalman gain  $\mathbf{K}_k$  (Equation 2.14), the following equation for the Kalman gain  $\mathbf{K}_k$  can be obtained:

$$\mathbf{K}_k = \mathbf{P}_k^{pz} (\mathbf{P}_k^z)^{-1}. \quad (2.17)$$

Lastly, the derived expression of the *a posteriori* covariance  $\mathbf{P}_k^a$  (Equation 2.12) will be rewritten by substituting the expression of the Kalman gain  $\mathbf{K}_k$  (Equation 2.17). Furthermore, the expression of the cross covariance  $\mathbf{P}_k^{pz}$  (Equation 2.16) will be used to obtain an alternative expression for the *a posteriori* covariance  $\mathbf{P}_k^a$ , resulting in:

$$\begin{aligned} \mathbf{P}_k^a &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^p, \\ &= \mathbf{P}_k^p - \mathbf{P}_k^{pz} (\mathbf{P}_k^z)^{-1} (\mathbf{P}_k^{pz})^T, \\ &= \mathbf{P}_k^p - \mathbf{K}_k (\mathbf{P}_k^z)^{-1} \mathbf{K}_k^T. \end{aligned} \quad (2.18)$$

This eventual derived equation of the *a posteriori* covariance  $\mathbf{P}_k^a$  (Equation 2.18) will be used in the current study.

### 2.1.2 Kalman filter algorithm

The derivation of the optimal Kalman gain and the definition of all elements of the conventional KF are provided in the previous subsection, while in this subsection the different implementation steps of the KF algorithm will be outlined in chronological order. The KF consists of two recursive sections, the prediction step and the analysis step, with the initialization being carried out before the first iteration of these recursive steps. In this initialization, the *a posteriori* estimate  $\hat{x}_k^a$  and the *a posteriori* covariance  $\mathbf{P}_k^a$  will be obtained by conducting an initial guess. Subsequently, the two steps of the algorithm will be performed recursively, starting with the prediction step (resulting in a *a priori* estimate  $\hat{x}_k^p$  and covariance  $\mathbf{P}_k^p$ ), followed by the analysis step (resulting in a *a posteriori* estimate  $\hat{x}_k^a$  and covariance  $\mathbf{P}_k^a$ ). The initialization and the two recursive steps of the KF algorithm will be detailed below.

#### Initialization

Initial guess of:  $\mathbf{P}_0^a$ ,  $\mathbf{Q}_0$ ,  $\mathbf{R}_0$  and  $\hat{x}_0^a$ .

#### Prediction

Predict the *a priori* estimate:

$$\hat{x}_{k+1}^p = \mathbf{M}_{k+1} \hat{x}_k^a.$$

Predict the *a priori* covariance:

$$\mathbf{P}_{k+1}^p = \mathbf{M}_{k+1} \mathbf{P}_k^a \mathbf{M}_{k+1}^T + \mathbf{Q}_{k+1}.$$

#### Analysis

Map prediction to measurement space:

$$\hat{z}_{k+1} = \mathbf{H}_{k+1} \hat{x}_{k+1}^p.$$

Determine the innovation- and the cross covariance:

$$\mathbf{P}_{k+1}^z = \mathbf{H}_{k+1} \mathbf{P}_{k+1}^p \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1},$$

$$\mathbf{P}_{k+1}^{pz} = \mathbf{P}_{k+1}^p \mathbf{H}_{k+1}^T.$$

Calculate the Kalman gain:

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1}^{pz} (\mathbf{P}_{k+1}^z)^{-1}.$$

Determine the *a posteriori* estimate of the mean:

$$\hat{x}_{k+1}^a = \hat{x}_{k+1}^p + \mathbf{K}_{k+1} (y_{k+1} - \hat{z}_{k+1}).$$

Determine the *a posteriori* covariance:

$$\mathbf{P}_{k+1}^a = \mathbf{P}_{k+1}^p (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}).$$

## 2.2 Unscented Kalman filter

The conventional KF, discussed in the previous section of this chapter, has limited utility in the real world because it is only applicable to linear time-invariant systems. However, there are several extensions available that are suitable for the conventional KF to deal with non-linear systems. For example the extended Kalman filter (EKF), which generalizes the conventional KF approach by linearizing the non-linear functions of a given system using Taylor's theorem [4, 46]. Another available extension that is suitable for the conventional KF is the unscented Kalman filter (UKF), which is capable of handling non-linearities even better than the previously mentioned EKF [77]. In addition, the UKF outperforms the EKF because the UKF can make an estimation of the mean and covariance of the states and measurements with second-order accuracy, while the EKF can only achieve first-order accuracy [31, 73]. To describe the UKF in detail, the following non-linear time-invariant system is considered:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) + \mathbf{w}_k \quad (2.19)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \quad (2.20)$$

The non-linear time-invariant behaviour of this system can be described by the non-linear transition function  $\mathbf{f}(\mathbf{x}_k)$ . Furthermore, the measurements of this system will be selected and extracted by using the measurement function  $\mathbf{h}(\mathbf{x}_k)$ .

### 2.2.1 Unscented transform

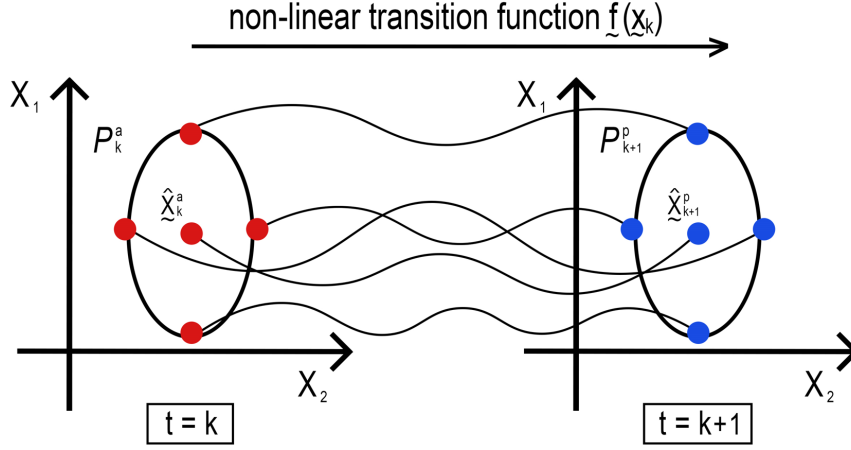
First of all, it is important to note that UKF retains the structure of the conventional Kalman filter, meaning that it contains a prediction step and an analysis step. These two recursive steps result in the *a priori* and the *a posteriori* estimate and corresponding covariance, respectively [50]. In addition to the similarity in structure, the UKF consists of additional components to accommodate non-linearities, such as unscented transform (UT). This UT is a mathematically driven technique that can be used to propagate the mean and covariance to the next time step by using a non-linear function [46]. The idea of this UT technique originates from the study of Julier et al. (1997), in which it is intuitively stated that it should be easier to approximate a probability distribution than to approximate a non-linear function [31]. In this UT approach, a set of sample points, called sigma-points, are obtained from *a posteriori* distribution of the current time step ( $\Upsilon(\hat{\mathbf{x}}_k^a, \mathbf{P}_k^a)$ ), which can contain any form and is denoted by  $\Upsilon$ . The sigma-points are chosen such that the mean and covariance of the sigma-points correspond to the mean and covariance of that particular *a posteriori* distribution  $\Upsilon$ . Subsequently, the non-linear function, which is the transition function  $\mathbf{f}(\mathbf{x}_k)$  in the current study, is used to propagate each of these sigma-points to the next time step. By calculating the mean and covariance of these propagated sigma-points, the *a priori* estimate  $\hat{\mathbf{x}}_{k+1}^p$  and the corresponding covariance  $\mathbf{P}_{k+1}^p$  are obtained [31, 32, 46]. Figure 2.2 below illustrates the propagation of sigma-points from the current to the next time step following the UT approach.

In order to calculate these sigma-points, a few constants ( $\alpha$ ,  $\beta$ , and  $\nu$ ) need to be determined first. The distribution of a certain number of sigma-points around the mean is governed by the constant  $\alpha$ , which can be calculated by:

$$\alpha = \frac{1}{n_s + 1}. \quad (2.21)$$

In this equation, the number of states in the considered non-linear system (Equation 2.19 and 2.20) is denoted by  $n_s$ . Furthermore, the constant  $\beta$  is a non-negative constant that relies on prior knowledge of the distribution of the estimate. In case of a Gaussian distribution, a value of 2 is considered to be optimal for  $\beta$  according to van der Merwe et al. (2001) [73]. Lastly, the constant  $\nu$  fulfils an important position in the calculation of the weights of the sigma-points, which will be detailed later on. The constant  $\nu$  can be determined by:

$$\nu = n_s(\alpha^2 - 1). \quad (2.22)$$



**Figure 2.2:** Here, the unscented transform (UT) is visualized considering a 2D state space. The *a posteriori* estimate  $\hat{x}_k^a$  and covariance  $\mathbf{P}_k^a$  are converted into 5 sigma-points, calculation of this amount will be elaborated later on, to describe the distribution of the current time step ( $k$ ). These sigma-points are then propagated to the next time step ( $k+1$ ) by using the non-linear transition function  $f(x_k)$ . Based on these propagated sigma-points, the *a priori* estimate  $\hat{x}_{k+1}^p$  and covariance  $\mathbf{P}_{k+1}^p$  of the next time step can be determined.

After determining all previously mentioned constants, the calculation of the sigma-points can be carried out. According to the currently discussed UKF approach, the number of sigma-points required to describe the distribution of a certain state is equal to  $2n_s + 1$ . Eventually, this will result in a sigma-points matrix  $\chi_k$  with dimensions  $n_s \times 2n_s + 1$ . The very first sigma-point of every state ( $\chi_{k,i=1}$ ), where  $i$  indexes the column in the sigma-point matrix  $\chi_k$ , represents the mean of that particular state. As a result, it can be stated that the first sigma-point of every state is equal to the *a posteriori* estimate of the current time step  $\hat{x}_k^a$ . Subsequently, the remaining sigma-point are distributed around the mean of that particular state to describe the rest of the distribution. These remaining sigma-points can be calculated by adding or subtracting the factor  $(\sqrt{(n_s + \nu)\mathbf{P}_k^a})_{(i-1)}$  to/from the *a posteriori* estimate  $\hat{x}_k^a$  [30, 31, 73]. This factor  $(\sqrt{(n_s + \nu)\mathbf{P}_k^a})_{(i-1)}$  represents the  $(i-1)$ -th column of the matrix square root of the manipulated *a posteriori* covariance matrix  $(n_s + \nu)\mathbf{P}_k^a$  [31]. This matrix square root of the covariance matrix  $((\sqrt{(n_s + \nu)\mathbf{P}_k^a})_{(i-1)})$  can be calculated by using the Cholesky decomposition in combination with the singular value decomposition (SVD), which are both detailed and substantiated in the study of Ma et al. (2010) [46]. The calculation of the sigma-points can be summarized as follows:

$$\chi_{k,i} = \hat{x}_k^a, \quad i = 1, \quad (2.23)$$

$$\chi_{k,i} = \hat{x}_k^a + (\sqrt{(n_s + \nu)\mathbf{P}_k^a})_{(i-1)}, \quad i = 2, \dots, n_s + 1, \quad (2.24)$$

$$\chi_{k,i} = \hat{x}_k^a - (\sqrt{(n_s + \nu)\mathbf{P}_k^a})_{((i-1)-n_s)}, \quad i = n_s + 2, \dots, 2n_s + 1. \quad (2.25)$$

The result of these sigma-point calculations are all column vectors, which will collectively form the sigma-point matrix  $\chi_k$ . Eventually, all these sigma-points of matrix  $\chi_k$  will be propagated to the next time step by using a non-linear transition function  $f(x_k)$  according to the UT approach [30]. In order to determine the *a priori* estimate and *a priori* covariance based on the propagated sigma-points, they have to be weighted accordingly:

$$\omega_1^m = \frac{\nu}{n_s + \nu}, \quad (2.26)$$

$$\omega_1^c = \frac{\nu}{n_s + \nu} + (1 - \alpha^2 + \beta), \quad (2.27)$$

$$\omega_i^m = \omega_i^c = \frac{\nu}{2(n_s + \nu)}, \quad i = 2, \dots, 2n_s + 1. \quad (2.28)$$



### 2.2.2 Unscented Kalman filter algorithm

The UKF algorithm can roughly be divided into three distinct recursive sections; the sigma-point determination, the prediction step and the analysis step. Before executing these recursive sections, the initialization needs to be completed first. The following elements remain constant throughout the iterations of the algorithm and will therefore be calculated in the initialization phase;  $\alpha, \beta, \nu$  and all weights ( $\omega_1^m, \omega_1^c, \omega_i^m$  and  $\omega_i^c$ ). The equations and values of the former four components are elaborated in the previous subsection, while the weights can be calculated by using Equation 2.26 - 2.28. The *a posteriori* estimate  $\hat{x}_k^a$  and the *a posteriori* covariance  $\mathbf{P}_k^a$  will be initially guessed in order to be able to execute the first iteration of the UKF algorithm.

The UT, which was discussed in the previous subsection, is used in the UKF algorithm to handle non-linearities that occur in the transition function  $f(x_k)$  and/or the measurement function  $h(x_k)$ . The possible non-linearities in these functions are addressed by making use of sigma-points according to the UT approach. Therefore, the determination of the sigma-point is considered to be the first step of the UKF algorithm. As mentioned before, the general structure of the UKF algorithm is fairly comparable to the conventional KF algorithm. Therefore, the second step of the algorithm is the prediction step, where the mathematical model is used to provide a prediction, which results in an *a priori* estimate  $\hat{x}_k^p$  and an *a priori* covariance  $\mathbf{P}_k^p$ . The third and last step is the analysis step, in which the available measurement is used to accomplish an *a posteriori* estimate  $\hat{x}_k^a$  and an *a posteriori* covariance  $\mathbf{P}_k^a$ . In this last step, the Kalman gain  $\mathbf{K}_k$  will be used to weigh the influence of innovation ( $y_{k+1} - \hat{z}_{k+1}$ ) with the objective of minimizing the *a posteriori* covariance  $\mathbf{P}_k^a$ . The initialization and three recursive steps will be presented on the next page in chronological order.

**Initialization**

Determine:  $\alpha, \beta, \nu, \omega_1^m, \omega_1^c, \omega_i^m$  and  $\omega_i^c$ .

Initialize:  $\mathbf{P}_0^a, \mathbf{Q}_0, \mathbf{R}_0$  and  $\hat{\mathbf{x}}_0^a$ .

**Sigma-points determination**

Determine the sigma-points of the current time step:

$$\begin{aligned} \chi_{k,i} &= \hat{\mathbf{x}}_k^a, & i &= 1, \\ \chi_{k,i} &= \hat{\mathbf{x}}_k^a + (\sqrt{(n_s + \nu)\mathbf{P}_k^a})_{(i-1)}, & i &= 2, \dots, n_s + 1, \\ \chi_{k,i} &= \hat{\mathbf{x}}_k^a - (\sqrt{(n_s + \nu)\mathbf{P}_k^a})_{((i-1)-n_s)}, & i &= n_s + 2, \dots, 2n_s + 1. \end{aligned}$$

**Prediction**

Propagate the sigma-points to the prediction space by using the transition function  $\underline{f}$ :

$$\chi_{k,i}^p = \underline{f}(\chi_{k,i}), \quad i = 1, \dots, 2n_s + 1.$$

Determine the *a priori* estimate:

$$\hat{\mathbf{x}}_{k+1}^p = \sum_{i=1}^{2n_s+1} \omega_i^m \chi_{k,i}^p.$$

Determine the *a priori* covariance:

$$\mathbf{P}_k^p = \sum_{i=1}^{2n_s+1} \omega_i^c (\chi_{k,i}^p - \hat{\mathbf{x}}_{k+1}^p)(\chi_{k,i}^p - \hat{\mathbf{x}}_{k+1}^p)^T + \mathbf{Q}_k.$$

**Analysis**

Propagate the sigma-points to the measurement space by using the measurement function  $\underline{h}$ :

$$\chi_{k,i}^z = \underline{h}(\chi_{k,i}^p), \quad i = 1, \dots, 2n_s + 1.$$

Determine the estimate in the measurement space:

$$\hat{\mathbf{z}}_{k+1} = \sum_{i=1}^{2n_s+1} \omega_i^m \chi_{k,i}^z.$$

Determine the covariance in the measurement space:

$$\mathbf{P}_{k+1}^z = \sum_{i=1}^{2n_s+1} \omega_i^c (\chi_{k,i}^z - \hat{\mathbf{z}}_{k+1})(\chi_{k,i}^z - \hat{\mathbf{z}}_{k+1})^T + \mathbf{R}_k.$$

Calculate the cross covariance:

$$\mathbf{P}_{k+1}^{pz} = \sum_{i=1}^{2n_s+1} \omega_i^c (\chi_{k,i}^p - \hat{\mathbf{x}}_{k+1}^p)(\chi_{k,i}^z - \hat{\mathbf{z}}_{k+1})^T.$$

Calculate the Kalman gain:

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1}^{pz} (\mathbf{P}_{k+1}^z)^{-1}.$$

Determine the *a posteriori* estimate:

$$\hat{\mathbf{x}}_{k+1}^a = \hat{\mathbf{x}}_{k+1}^p + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \hat{\mathbf{z}}_{k+1}).$$

Determine the *a posteriori* covariance:

$$\mathbf{P}_{k+1}^a = \mathbf{P}_{k+1}^p - \mathbf{K}_{k+1} \mathbf{P}_{k+1}^z \mathbf{K}_{k+1}^T.$$

## 2.3 Reduced-order unscented Kalman filter

The previously discussed UKF is an extension for the conventional KF, which can deal with non-linear systems [73]. However, a Kalman filter with non-linear extension, like this UKF, is known for its computational expensiveness when applied to problems with a considerably large amount of data. Therefore, a computationally more efficient method is desired, which is the reduced-order unscented Kalman filter (ROUKF) [17]. Before diving into the fundamentals and reasoning behind this ROUKF method, let us first describe the general form of the non-linear system to which the method will be applied. In the considered non-linear system, the states and parameters will be gathered into one collective state-parameter space, which is called the augmented state vector. Formally, this augmented state vector can be denoted by:

$$\underline{x}_{aug} = \begin{bmatrix} \underline{x} \\ \underline{\theta} \end{bmatrix}. \quad (2.29)$$

where  $\underline{x}$  represents the state vector and  $\underline{\theta}$  the parameter vector. The size of  $\underline{x}$  is determined by the number of states ( $n_s$ ) and the size of  $\underline{\theta}$  by the number of parameters ( $n_p$ ). As a result, the size of the augmented state vector is equal to  $n = n_s + n_p$ . The evolution of the augmented state vector will be governed by the transition function  $\underline{f}(\underline{x}_{aug,k})$ , which describes the behavior of the considered time-invariant non-linear system. In this particular system, the measurements will be selected and extracted by using measurement function  $\underline{h}(\underline{x}_k)$ . Therefore, the considered time-invariant non-linear system can be given as:

$$\underline{x}_{aug,k+1} = \underline{f}(\underline{x}_{aug,k}) + \underline{w}_k, \quad (2.30)$$

$$\underline{y}_k = \underline{h}(\underline{x}_k) + \underline{v}_k. \quad (2.31)$$

The basis of the reduction of computational expenses in the ROUKF method can be attributed to using two distinct DA methods into a collective method that is known as a hybrid DA approach. Explicitly, the estimation of the states  $\underline{x}$  and parameters  $\underline{\theta}$  of the augmented state vector  $\underline{x}_{aug,k}$  will be both be addressed with distinct DA methods. In the ROUKF method, these two DA methods are intertwined and will be executed simultaneously following the joint state-parameter estimation approach [51]. The mathematical proof of this joint state-parameter estimation approach is supplied in the study of Moireau et al. (2008) [52].

To accomplish the desired computational benefit, the state estimation is addressed with a simple and computationally cheap DA method, while the parameter estimation is conducted with an accurate and computationally more expensive DA method [17]. The parameter estimation will be executed by using the UKF, which is able to estimate the mean and covariance of the parameters with second-order accuracy [73]. A DA method that is considered to be suitable for state estimation is the Luenberger observer, because it is a computationally inexpensive and stable method [17, 42, 51]. Consequently, the Luenberger observer will be responsible for the state estimation and the UKF for the parameter estimation, which will be executed collectively following the joint state-parameter approach to eventually result in a hybrid DA method called ROUKF [17, 51]. Conceptually, the ROUKF method can be given as:

$$\underline{x}_{aug} = \left. \begin{array}{l} \underline{x} \rightarrow \text{Luenberger observer} \\ \underline{\theta} \rightarrow \text{unscented Kalman filter} \end{array} \right\} \text{reduced-order unscented Kalman filter.}$$

The Luenberger observer is considered to be computationally inexpensive and stable because it is based on a forward nudging approach [51]. Nudging is a relatively simple method which adjusts the solution of the model towards the available measurements, without depending on an optimal criterion. This adjustment is accomplished by introducing a feedback term that captures the discrepancy between the model and measurement [3].

### 2.3.1 Assimilation scheme

In the ROUKF method, the state estimation is thus performed by the Luenberger observer, while the parameter estimation is executed by the UKF. The Luenberger observer also fulfils another important role in addition to the state estimation, which is the allocation of the estimation error to the parameter space. In essence, this is equivalent to confining the model uncertainty to the parameter space [17]. As a result, the rank of the *a posteriori* covariance matrix  $\mathbf{P}_{k+1}^a$  will be equal to the number of parameters  $n_p$ . This rank reduction forms the foundation for the computationally efficient reduced-order filtering, because it allows a rank decomposition of  $\mathbf{P}_{k+1}^a$ . Based on the rank decomposition, the matrix  $\mathbf{P}_{k+1}^a$  can be denoted in its factorized form [51]:

$$\mathbf{P}_{k+1}^a = \mathbf{L}_{k+1} \mathbf{U}_{k+1}^{-1} \mathbf{L}_{k+1}^T. \quad (2.32)$$

The *a posteriori* covariance matrix  $\mathbf{P}_{k+1}^a$  is a square matrix with dimensions  $n \times n$ , containing the covariance between each pair of elements of the augmented state vector  $\mathbf{x}_{aug,k}$ . The two matrices that resulted from the rank decomposition contain smaller dimensions, namely, the Luenberger matrix  $\mathbf{L}_{k+1}$   $n \times n_p$  and the uncertainty matrix  $\mathbf{U}_{k+1}$   $n_p \times n_p$ . The usefulness of this rank decomposition lies in the fact that manipulations and calculations can be performed on matrices with smaller dimensions ( $\mathbf{L}_{k+1}$  and  $\mathbf{U}_{k+1}$ ), instead of on the conventional covariance matrix ( $\mathbf{P}_{k+1}^a$ ). Consequently, the calculations involving the covariance matrix become computationally tractable without influencing the eventual dimensions of the *a posteriori* covariance matrix  $\mathbf{P}_{k+1}^a$  [51].

The mathematical proof of the factorized form of the *a posteriori* covariance matrix  $\mathbf{P}_{k+1}^a$  will be derived below, starting with some definitions. First, two different forms of the Luenberger matrix are introduced, of which one is located in the prediction space ( $\mathbf{L}_{k+1}$ ) and the second in the measurement space ( $\mathbf{B}_{k+1}$ ). The Luenberger matrix in the prediction space  $\mathbf{L}_{k+1}$  depends on the sigma-points located in the prediction space  $\chi_{k+1}^p$ , whereas the Luenberger matrix in the measurement space  $\mathbf{B}_{k+1}$  depends on the sigma-points in the measurement space  $\chi_{k+1}^z$ . The definitions of these Luenberger matrices are given by [51]:

$$\begin{aligned} \mathbf{L}_{k+1} &= f(\chi_{k+1}) \mathbf{D}_\varepsilon \mathbf{S}^T \\ &= \chi_{k+1}^p \mathbf{D}_\varepsilon \mathbf{S}^T, \end{aligned} \quad (2.33)$$

$$\begin{aligned} \mathbf{B}_{k+1} &= h(\chi_{k+1}^p) \mathbf{D}_\varepsilon \mathbf{S}^T \\ &= \chi_{k+1}^z \mathbf{D}_\varepsilon \mathbf{S}^T. \end{aligned} \quad (2.34)$$

Both expressions of the Luenberger matrix contain an equal constant part, which consists of a diagonal matrix  $\mathbf{D}_\varepsilon$  and a simplex sigma-point matrix  $\mathbf{S}$  (see Section 2.3.2). The diagonal matrix  $\mathbf{D}_\varepsilon$  has dimensions  $n_p + 1 \times n_p + 1$  and contains the scalar value of the weight  $\varepsilon = \frac{1}{n_p + 1}$  on all elements of the diagonal. Second, the uncertainty matrix  $\mathbf{U}_{k+1}$  is defined as [51]:

$$\mathbf{U}_{k+1} = \mathbf{P}_\varepsilon + \mathbf{B}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{B}_{k+1}, \quad (2.35)$$

in which,  $\mathbf{P}_\varepsilon$  is the constant part of the covariance matrix with dimensions  $n_p \times n_p$  and is being expressed by [51]:

$$\mathbf{P}_\varepsilon = \mathbf{S} \mathbf{D}_\varepsilon \mathbf{S}^T. \quad (2.36)$$

Finally, a few expressions of different covariance matrices are given [51]:

$$\mathbf{P}_k^p = \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{L}_k^T, \quad (2.37)$$

$$\mathbf{P}_k^z = \mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k, \quad (2.38)$$

$$\mathbf{P}_k^{pz} = \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T. \quad (2.39)$$

Here,  $\mathbf{P}_k^p$  denotes the *a priori* covariance matrix,  $\mathbf{P}_k^z$  the innovation covariance matrix and  $\mathbf{P}_k^{pz}$  the cross-covariance matrix. Note that for the ease of reading, time step  $k$  is used instead of time step  $k + 1$ .

Now that some key matrices have been introduced, the mathematical proof regarding the factorized form of the *a posteriori* covariance matrix  $\mathbf{P}_k^a$  can be supplied. Starting with the expression of the *a posteriori* covariance  $\mathbf{P}_k^a$  of the conventional Kalman filter (Equation 2.18) and substituting the equation of the optimal Kalman gain  $\mathbf{K}_k$  (Equation 2.17):

$$\begin{aligned}\mathbf{P}_k^a &= \mathbf{P}_k^p - \mathbf{K}_k \mathbf{P}_k^z \mathbf{K}_k^T, \\ &= \mathbf{P}_k^p - \mathbf{P}_k^{pz} (\mathbf{P}_k^z)^{-1} \mathbf{P}_k^z (\mathbf{P}_k^z)^{-1} \mathbf{P}_k^{pzT}, \\ &= \mathbf{P}_k^p - \mathbf{P}_k^{pz} (\mathbf{P}_k^z)^{-1} (\mathbf{P}_k^{pz})^T.\end{aligned}\tag{2.40}$$

Next, the previously stated expressions of the *a priori* covariance  $\mathbf{P}_k^p$  (Equation 2.37), the innovation covariance  $\mathbf{P}_k^z$  (Equation 2.38) and the cross-covariance  $\mathbf{P}_k^{pz}$  (Equation 2.39) are substituted, which results in:

$$\begin{aligned}\mathbf{P}_k^a &= \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{L}_k^T - \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (\mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T)^T, \\ &= \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{L}_k^T - \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} \mathbf{L}_k^T \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k, \\ &= \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{L}_k^T - \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k \mathbf{L}_k^T, \\ &= \mathbf{L}_k (\mathbf{P}_\varepsilon^{-1} - \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k) \mathbf{L}_k^T.\end{aligned}\tag{2.41}$$

Subsequently, this last expression will be simplified by using the Woodbury formula, which in general terms reads [18]:

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} \mathbf{A}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{D} \mathbf{A}^{-1}.\tag{2.42}$$

In this equation, matrix  $\mathbf{A}$  and matrix  $\mathbf{C}$  are both square invertible matrices with a dimension of  $n_A \times n_A$  and  $n_C \times n_C$ , respectively. Further, matrix  $\mathbf{B}$  and matrix  $\mathbf{D}$  have dimensions  $n_A \times n_C$  and  $n_C \times n_A$ , respectively. According to the derivation that is currently being detailed, the matrices referred in Equation 2.41 can be substituted by;  $\mathbf{A} = \mathbf{P}_\varepsilon$ ,  $\mathbf{B} = \mathbf{B}_k^T$ ,  $\mathbf{C} = \mathbf{R}_k^{-1}$  and  $\mathbf{D} = \mathbf{B}_k$ . This Woodbury formula will be used to rewrite Equation 2.41. Subsequently, the expression of the uncertainty matrix  $\mathbf{U}_{k+1}$  (Equation 2.35) will be used to simplify the resulting expression. Formally, this can be given as:

$$\begin{aligned}\mathbf{P}_k^a &= \mathbf{L}_k (\mathbf{P}_\varepsilon + \mathbf{B}_k^T \mathbf{R}_k^{-1} \mathbf{B}_k)^{-1} \mathbf{L}_k^T, \\ &= \mathbf{L}_k (\mathbf{P}_\varepsilon + \mathbf{B}_k \mathbf{R}_k^{-1} \mathbf{B}_k^T)^{-1} \mathbf{L}_k^T, \\ &= \mathbf{L}_k (\mathbf{U}_k)^{-1} \mathbf{L}_k^T, \\ &= \mathbf{L}_k \mathbf{U}_k^{-1} \mathbf{L}_k^T.\end{aligned}\tag{2.43}$$

As a result, it is proven that the factorized form of the *a posteriori* covariance  $\mathbf{P}_k^a$  is mathematically valid.

Another essential element of the ROUKF approach is the *a posteriori* estimate  $\hat{x}_k^a$ , which will be mathematically derived below. This derivation starts with the expression of the *a posteriori* estimate of the conventional Kalman filter (Equation 2.6), in which the expression of the Kalman gain  $\mathbf{K}_k$  (Equation 2.17) will be substituted. Next, the mathematical definitions of the innovation covariance  $\mathbf{P}_k^z$  (Equation 2.38) and cross covariance  $\mathbf{P}_k^{pz}$  (Equation 2.39) will be used to rewrite the equation, leading to:

$$\begin{aligned}\hat{x}_k^a &= \hat{x}_k^p + \mathbf{K}_k (y_k - \hat{z}_k), \\ &= \hat{x}_k^p + \mathbf{P}_k^{pz} (\mathbf{P}_k^z)^{-1} (y_k - \hat{z}_k), \\ &= \hat{x}_k^p + \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k).\end{aligned}\tag{2.44}$$

Subsequently, the second term in this expression will be multiplied with the identity matrix  $\mathbf{I}$ , followed by a substitution of: 1)  $\mathbf{I} = \mathbf{U}_k^{-1}\mathbf{U}_k$  and 2) the expression of the uncertainty matrix  $\mathbf{U}_k$  (Equation 2.35). Lastly, the resulting expression will be rewritten by taking  $\mathbf{P}_\varepsilon^{-1}$  and  $\mathbf{B}_k^T$  inside the first pair of parentheses. Explicitly, this can be given as follows:

$$\begin{aligned}
\hat{\mathbf{x}}_k^a &= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{I} \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} \mathbf{U}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} (\mathbf{P}_\varepsilon + \mathbf{B}_k^T \mathbf{R}_k^{-1} \mathbf{B}_k) \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} (\mathbf{P}_\varepsilon \mathbf{P}_\varepsilon^{-1} + \mathbf{B}_k^T \mathbf{R}_k^{-1} \mathbf{B}_k \mathbf{P}_\varepsilon^{-1}) \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} (\mathbf{I} + \mathbf{B}_k^T \mathbf{R}_k^{-1} \mathbf{B}_k \mathbf{P}_\varepsilon^{-1}) \mathbf{B}_k^T (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} (\mathbf{B}_k^T + \mathbf{B}_k^T \mathbf{R}_k^{-1} \mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T) (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k). \tag{2.45}
\end{aligned}$$

Finally, the Luenberger matrix located in the measurement space  $\mathbf{B}_k^T$  and the inverse of the measurement noise  $\mathbf{R}_k^{-1}$  are both drawn outside the parentheses, which can be given as:

$$\begin{aligned}
\hat{\mathbf{x}}_k^a &= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} (\mathbf{B}_k^T + \mathbf{B}_k^T \mathbf{R}_k^{-1} \mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T) (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} \mathbf{B}_k^T (\mathbf{I} + \mathbf{R}_k^{-1} \mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T) (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} \mathbf{B}_k^T \mathbf{R}_k^{-1} (\mathbf{R}_k + \mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T) (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} \mathbf{B}_k^T \mathbf{R}_k^{-1} (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k) (\mathbf{B}_k \mathbf{P}_\varepsilon^{-1} \mathbf{B}_k^T + \mathbf{R}_k)^{-1} (y_k - \hat{z}_k), \\
&= \hat{\mathbf{x}}_k^p + \mathbf{L}_k \mathbf{U}_k^{-1} \mathbf{B}_k^T \mathbf{R}_k^{-1} (y_k - \hat{z}_k). \tag{2.46}
\end{aligned}$$

After deriving both the *a posteriori* covariance  $\mathbf{P}_k^a$  (Equation 2.43) and the *a posteriori* estimate  $\hat{\mathbf{x}}_k^a$  (Equation 2.46), still one key component of the ROUKF algorithm is left uncovered.

This key component is the allocation of the estimation error to the parameter space, which can be realized choosing appropriate initial conditions. Specifically, the state positions of the initial Luenberger matrix  $\mathbf{L}_0^x$  will be equaled to a zero matrix  $\mathbf{O}$ , while the parameter positions of the initial Luenberger matrix  $\mathbf{L}_0^\theta$  will be equaled to the identity matrix  $\mathbf{I}$ . When equaling the initial uncertainty matrix  $\mathbf{U}_0$  to the identity matrix  $\mathbf{I}$  and applying Equation 2.32, this results in the initial *a posteriori* covariance matrix  $\mathbf{P}_0^a$  packed with zeros, while containing a value of one on the diagonal of the parameter positions. As a result, the estimation error is concentrated to the parameter space, which proves the validity of the previous mentioned notion matrix  $\mathbf{P}_0^a$  is reduced rank [6, 17]. Formally, the initialization can be given as follows:

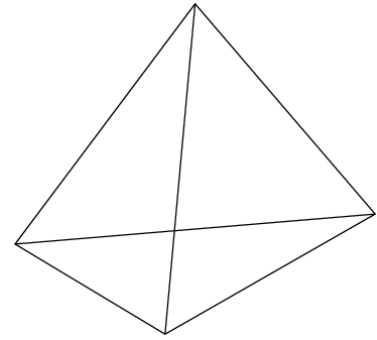
$$\mathbf{L}_0^x = \mathbf{O}, \quad \mathbf{L}_0^\theta = \mathbf{I}, \quad \mathbf{U}_0 = \mathbf{I}. \tag{2.47}$$

Conclusively, it can be stated that the application of the initial Luenberger observer matrix  $\mathbf{L}_0$  does confine the estimation error to the parameter space, which results in the rank reduction of the initial *a posteriori* covariance matrix  $\mathbf{P}_0^a$ . Consequently, a rank decomposition can be applied to the *a posteriori* covariance matrix  $\mathbf{P}_0^a$  which results in its factorized form, containing matrices with smaller dimensions ( $\mathbf{L}_k$  and  $\mathbf{U}_k$ ). This decline in size will eventually result in the desired reduction of computational expensiveness, especially in larger dimensional systems where  $n_s \gg n_p$ .

### 2.3.2 Simplex sigma-points

The use of sigma-points to describe the distribution of a certain state or parameter, as well as the use of the non-linear transition function  $f(x_k)$  to propagate these points to the next time step is equivalent for both the UKF and the ROUKF. However, some significant differences can be identified when comparing the ROUKF to the UKF: first, the approach to deterministically selection sigma-points is different, and second, the number of sigma-points is different. The ROUKF method makes use of the simplex approach to determine the sigma-points, which is based on using the smallest possible number of sigma-points necessary to truthfully represent a given distribution. Furthermore, the *a posteriori* covariance matrix  $\mathbf{P}_k^a$  is considered to be reduced rank in the ROUKF approach, which allows rank decomposition of  $\mathbf{P}_k^a$  into its factorized form. Because the dimensions of the matrices in this factorized form are much smaller compared to the conventional covariance matrix, which also results in a reduction of the number of sigma-points. Consequently, only  $n_p + 1$  sigma-points are required to represent the distribution of a certain state or parameter in the ROUKF approach [51]. In the earlier discussed UKF,  $2n + 1$  sigma-points are used to represent the distribution of the certain states or parameters, which is significantly larger than the  $n_p + 1$  sigma-points used in the ROUKF approach.

The currently discussed simplex sigma-point approach is an intuitive way to decrease the computational expensiveness even further, as the expensiveness was previously limited by the factorization of the covariance matrix. In this approach, a simplex sigma-point matrix  $\mathbf{S}$  is used in which all sigma-points are located on a regular polyhedron with radius  $\sqrt{n_p}$ . A regular polyhedron is a polyhedron which is composed of a few identical polygonal faces. These faces have equal lengths and equal angles between them, whereas the number of faces determine the name of the polyhedron [39]. Eventually, this simplex sigma-point matrix  $\mathbf{S}$  is used to compose the sigma-point matrix  $\chi_k$ , which is used to propagate the mean and covariance. This sigma-point matrix  $\chi_k$  has dimensions  $n \times n_p + 1$ , because  $n_p + 1$  sigma-points are used to describe the distribution of every state or parameter of the augmented state vector. The simplex sigma-point matrix  $\mathbf{S}$  can be determined as follows:



**Figure 2.3:** An intuitive example of a regular tetrahedron, which holds for  $n_p=3$ .

$$\mathbf{S}^1 = \begin{bmatrix} -\frac{1}{\sqrt{2\varepsilon}} & \frac{1}{\sqrt{2\varepsilon}} \end{bmatrix} \quad (2.48)$$

$$\mathbf{S}^d = \begin{bmatrix} & & & 0 \\ & & & \vdots \\ & \mathbf{S}^{d-1} & & 0 \\ \frac{1}{\sqrt{\varepsilon d(d+1)}} & \cdots & \frac{1}{\sqrt{\varepsilon d(d+1)}} & -d \end{bmatrix}, 2 \leq d \leq n_p. \quad (2.49)$$

In this expression, the weight that is used in this simplex sigma-points matrix is denoted by  $\varepsilon$  (see Section 2.3.1), which is equal for all sigma-points. The reason for this equality is that every simplex sigma-point is located on the regular polyhedron and therefore has an equal distance to the center, which is the mean. Since the simplex sigma-points are located on the regular polyhedron, the mean of every column vector of the sigma-point matrix is equal to zero because all simplex sigma-points are centered around the mean. Lastly, the considered simplex sigma-point matrix has an identity covariance matrix, which implies that the matrix is completely unconstrained and the individual components of the matrix are uncorrelated [51]. Concerning the current research, only one unknown parameter will be approximated, which results in a weight of  $\varepsilon = \frac{1}{2}$  with a simplex sigma-point matrix of  $\mathbf{S}^1 = \begin{bmatrix} -1 & 1 \end{bmatrix}$ .

### 2.3.3 Reduced-order unscented Kalman filter algorithm

In the first three subsections of this ROUKF section, the main principles behind the ROUKF algorithm were elaborated. Based on the main principles discussed in these subsections, the total structure of the ROUKF algorithm can finally be defined. The ROUKF algorithm can roughly be divided into four sections: 1) initialization, 2) sigma-point determination, 3) prediction and 4) analysis.

#### Initialization

In the initialization phase, a few elements will be initialized as well as a few elements will be determined that remain constant throughout the recursive steps of the algorithm. The elements that remain constant are the following: the simplex sigma-point matrix  $\mathbf{S}$ , the weight  $\varepsilon$ , the diagonal matrix  $\mathbf{D}$ , the constant part of the covariance matrix  $\mathbf{P}_\varepsilon$  and the measurement error covariance  $\mathbf{R}$ . The definitions of  $\mathbf{S}$ ,  $\varepsilon$ ,  $\mathbf{D}$  and  $\mathbf{P}_\varepsilon$  are already given in the previous subsections. The measurement error covariance  $\mathbf{R}$  has dimensions  $n_z \times n_z$ , where  $n_z$  represents the number of states in the measurement space. For simplicity, the measurement error covariance  $\mathbf{R}$  is assumed to be known and time-independent. Therefore, the measurement error covariance  $\mathbf{R}$  is arbitrarily set to  $100 * \mathbf{I}$  in the current study. Furthermore, the *a posteriori* estimate  $\hat{\mathbf{x}}_0^a$  will be guessed, while the Luenberger matrix in the prediction space  $\mathbf{L}_0$  and the uncertainty matrix  $\mathbf{U}_0$  will both be initialized such that the estimation error is concentrated to the parameters space (Equation 2.47). The discrete time steps in this algorithm are denoted by  $k$ , where  $k = 1$  represents the first time step and  $k = 0$  the initialization phase.

#### Sigma-point determination

After the initialization phase is completed, the sigma-points determination can be conducted. In this sigma-point determination, the Cholesky matrix is denoted by  $\mathbf{C}_k$ , which will be determined based on performing the Cholesky decomposition. The sigma-points of the current time step can be determined by using the following equations, where the  $i$ -th column vector of the sigma-point matrix  $\chi_k$  is denoted by  $\chi_{k,i}$ :

$$\mathbf{C}_k = \sqrt{\mathbf{U}_k^{-1}} \quad (2.50)$$

$$\chi_{k,i} = \hat{\mathbf{x}}_k^a + \mathbf{L}_k \mathbf{C}_k^T \mathbf{S}_i, \quad 1 \leq i \leq n_p + 1. \quad (2.51)$$

#### Prediction

After the sigma-points are determined, the prediction phase of the algorithm can be executed. In this phase, the sigma-points of the current time step  $\chi_k$  will be propagated to the prediction space of the next time step by using the non-linear transition function  $f$ . This propagation will eventually result in a sigma-point matrix in the prediction space  $\chi_{k+1}^p$ , with the  $i$ -th column vector of this sigma-point matrix being denoted by  $\chi_{k+1,i}^p$ :

$$\chi_{k+1,i}^p = f(\chi_{k,i}), \quad 1 \leq i \leq n_p + 1. \quad (2.52)$$

The *a priori* estimate of the next time step  $\hat{\mathbf{x}}_{k+1}^p$ , can be determined from the sigma-points located in the prediction space  $\chi_{k+1}^p$ . Explicitly, the *a priori* estimate  $\hat{\mathbf{x}}_{k+1}^p$  can be calculated as follows:

$$\hat{\mathbf{x}}_{k+1}^p = \sum_{i=1}^{p+1} \varepsilon \chi_{k+1,i}^p. \quad (2.53)$$

The last step of the prediction phase is the calculation of the Luenberger matrix in the prediction space  $\mathbf{L}$ , which is calculated by:

$$\mathbf{L}_{k+1} = \chi_{k+1}^p \mathbf{D}_\varepsilon \mathbf{S}^T. \quad (2.54)$$



### Analysis

The final phase of the ROUKF algorithm is the analysis phase. In this phase the *a priori* estimate  $\hat{x}_{k+1}^p$  will be updated by utilizing the measurements  $y_{k+1}$ , which eventually results in a *a posteriori* estimate  $\hat{x}_{k+1}^a$ . First, the sigma-points located in the prediction space  $\chi_{k+1}^p$  will be propagated to the measurement space by utilizing the measurement function  $h$ . This will result in a sigma-point matrix located in the measurement space  $\chi_{k+1}^z$ , where the  $i$ -th column vector of the sigma-point matrix is denoted by  $\chi_{k+1,i}^z$ :

$$\chi_{k+1,i}^z = h(\chi_{k+1,i}^p), \quad 1 \leq i \leq n_p + 1. \quad (2.55)$$

Then, the estimate in the measurement space ( $\hat{z}_{k+1}$ ) will be calculated conventionally by taking the sum of the weighted sigma-points, as follows:

$$\hat{z}_{k+1} = \sum_{i=1}^{p+1} \varepsilon \chi_{k+1,i}^z. \quad (2.56)$$

Next, the Luenberger matrix in the measurement space  $\mathbf{B}$  can be determined based on the sigma-points in the measurement space  $\chi_{k+1}^z$ , which can be calculated by:

$$\mathbf{B}_{k+1} = \chi_{k+1}^z \mathbf{D}_\varepsilon \mathbf{S}^T. \quad (2.57)$$

Subsequently, the uncertainty matrix  $\mathbf{U}_{k+1}$  can be determined by:

$$\mathbf{U}_{k+1} = \mathbf{P}_\varepsilon + \mathbf{B}_{k+1}^T \mathbf{R}_{k+1}^{-1} \mathbf{B}_{k+1}. \quad (2.58)$$

The former to last step in the algorithm is the calculation of the *a posteriori* covariance matrix  $\mathbf{P}_{k+1}^a$ , which can be given as:

$$\mathbf{P}_{k+1}^a = \mathbf{L}_{k+1} \mathbf{U}_{k+1}^{-1} \mathbf{L}_{k+1}^T. \quad (2.59)$$

Lastly, the *a posteriori* estimate  $\hat{x}_{k+1}^a$  is calculated. In this step the *a priori* estimate  $\hat{x}_{k+1}^p$  will be updated by a weighted difference between the actual measurement  $y_{k+1}$  and the estimate of the measurement space  $\hat{z}_{k+1}$ , which is basically the innovation factor. The weight of this innovation factor is based on various matrices, which can be seen in the following expression:

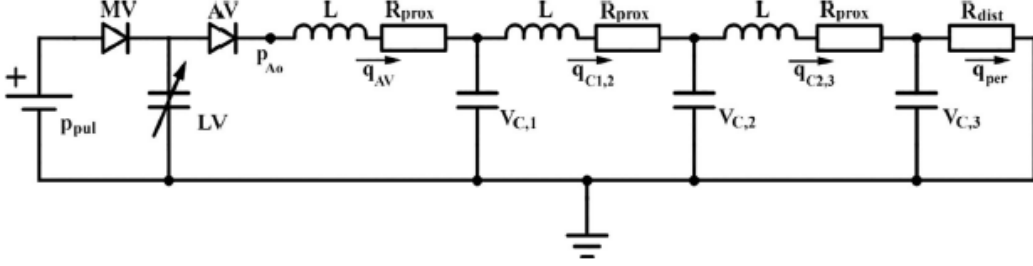
$$\hat{x}_{k+1}^a = \hat{x}_{k+1}^p + \mathbf{L}_{k+1} \mathbf{U}_{k+1}^{-1} \mathbf{B}_{k+1} \mathbf{R}_{k+1}^{-1} (y_{k+1} - \hat{z}_{k+1}). \quad (2.60)$$

After completion of the last step of the algorithm, the subsequent discrete time step will be initiated  $k = k + 1$ . In this subsequent timestep, the algorithm will start from the prediction phase, since the initialization phase is only executed in the first iteration.

## 3. Model and measurements

### 3.1 Mathematical model

In the current study, a mathematical model, which originates from the study of Meiburg et al. (2020), will be used to describe the behavior of the cardiovascular system [49]. The mathematical model, which is depicted in Figure 3.1 below, is a 0D lumped element model of the LV. In this model, the behavior of the LV is described by the single-fibre heart model, as proposed in the study of Bovendeerd et al. (2006) [8]. Furthermore, the vascular behavior is described by a transmission line of three-element Windkessel models, representing the arterial network.



**Figure 3.1:** Representation of the 0D lumped element model of the left ventricle, which describes the behavior of the cardiovascular system. In this model the preload is described as a fixed pulmonary pressure, while the afterload is modeled as a transmission line of three-element Windkessel models. This lumped element model originates for the study of Meiburg et al. (2020) [49].

In the single-fibre heart model, the LV is approximated as an incompressible thick-walled sphere consisting of multiple thin spherical shells, with muscle fibres oriented in circumferential direction. It is assumed that the stretches and stresses in each spherical shell are equal, leading to a homogeneous distribution of these two quantities throughout the ventricular wall. As a result of this homogeneous distribution, global properties (volume and pressure) can be related to local properties of a myocardial fibre (stress and strain). Therefore, the pressure inside the LV ( $p_{LV}$ ) can be related to the fibre stress in direction of the fibre itself ( $\sigma_f$ ) and radial direction ( $\sigma_r$ ) in association with the ratio between LV wall volume ( $V_w$ ) and LV cavity volume ( $V_{LV}$ ), which implies that:

$$p_{LV} = \frac{1}{3}(\sigma_f - 2\sigma_r)\ln\left(1 + \frac{V_w}{V_{LV}}\right). \quad (3.1)$$

The stress in the direction of the fibre ( $\sigma_f$ ) is a combination of an active as well as a passive component, whereas the stress in radial direction ( $\sigma_r$ ) is only dependent on a passive component. Formally, this can be stated as follows:

$$\sigma_f = \sigma_{f,a} + \sigma_{f,p} \quad \text{and} \quad \sigma_r = \sigma_{r,p}. \quad (3.2)$$

Before describing these passive and active components, the fibre stretch ( $\lambda_f$ ) will be determined first. The fibre stretch ( $\lambda_f$ ) is dependent on the following volume parameters; the wall volume ( $V_w$ ), the ventricular cavity volume ( $V_{LV}$ ) and the unloaded cavity volume ( $V_{LV,0}$ ), as given by:

$$\lambda_f = \left( \frac{V_{LV} + \frac{1}{3}V_w}{V_{LV,0} + \frac{1}{3}V_w} \right)^{\frac{1}{3}}. \quad (3.3)$$

The myocardial tissue is assumed to be incompressible, leading to the following relation between the fibre stretch ( $\lambda_f$ ) and the radial stretch ( $\lambda_r$ ):

$$\lambda_r = \lambda_f^{-2}. \quad (3.4)$$

Furthermore, it is generally assumed that the passive compressive stiffness is negligible. Therefore, the following equation can be used to describe passive stress:

$$\sigma_{i,p} = \begin{cases} 0, & \lambda_i < 1, \\ \sigma_{i,0} e^{(c_i(-\lambda_i-1)-1)}, & \lambda_i \geq 1. \end{cases} \quad (3.5)$$

The subscript  $i$  denotes that all elements of this equation can be regarded to either the fibre or radial direction. This separation is considered necessary because both orientations do contain different material parameters ( $\sigma_{i,0}$  and  $c_i$ ) due to anisotropy. Furthermore, the active component can be calculated as follows:

$$\sigma_{f,a} = c_{LV} \sigma_{a,ref} l_{rel}(l_s) t_{act}(t_a) v_{s,rel}(v_s). \quad (3.6)$$

The above stated active stress component ( $\sigma_a$ ) is dependent on the left ventricular contractility ( $c_{LV}$ ), the stress of the reference state ( $\sigma_{a,ref}$ ), the relative sarcomere length ( $l_{rel}(l_s)$ ), the shape of the fibre activation curve ( $t_{act}(t_a)$ ) and the relative shortening velocity of the sarcomere ( $v_{s,rel}(v_s)$ ). The last three mentioned dependencies are represented by the mathematical representations below:

$$l_{rel}(l_s) = \begin{cases} 0, & l_s < l_{s,0}, \\ \frac{l_s - l_{s,0}}{l_{s,ref} - l_{s,0}}, & l_s \geq l_{s,0}, \end{cases} \quad (3.7)$$

$$t_{act}(t_a) = \begin{cases} 1 - \frac{\cosh(t_{sharp} \left( \frac{2t_a}{t_{max}} - 1 \right) - 1)}{\cosh(t_{sharp}) - 1}, & 0 \leq t_a \leq t_{max}, \\ 0, & t_{max} < t_a < 1, \end{cases} \quad (3.8)$$

$$v_{s,rel}(v_s) = \frac{1 - \frac{v_s}{v_{s,0}}}{1 + s_v \frac{v_s}{v_{s,0}}}. \quad (3.9)$$

In the first equation, the sarcomere lengths of the unloaded state ( $l_{s,0}$ ), the reference state ( $l_{s,ref}$ ) and the actual sarcomere length ( $l_s$ ) enable determination of the relative sarcomere length ( $l_{rel}(l_s)$ ). Next, the activation curve ( $t_{act}(t_a)$ ) is dependent on timing parameters: the time since activation ( $t_a$ ), the relative contraction duration ( $t_{max}$ ) and the shape of the activation curve ( $t_{sharp}$ ). The above stated equation of the activation curve differs from the study of Bovendeerd et al. (2006) [8], to render a more physiologically representative approximation of the LV pressure signal shape, as given by Guyton and Hall (2006) [27]. Lastly, the calculation of the relative shortening velocity of the sarcomere ( $v_{s,rel}(v_s)$ ) is governed by the shortening velocity of the sarcomere ( $v_s$ ) and the shortening velocity of the sarcomere in unloaded state ( $v_{s,0}$ ). In this last equation, the exact shape of the stress-velocity relation is determined by  $s_v$ . Eventually, the behavior of the LV pressure can be described based on the discussed equations.

The complex vascular hemodynamics will be simplified by modelling the circulation by a transmission line of three-element Windkessel models, which consists of a resistor ( $R$ ), a capacitor ( $C$ ) and an inductor ( $L$ ). The pressure difference across the resistor ( $\Delta p_R$ ), the pressure drop across the inductor ( $\Delta p_L$ ) and the pressure drop across the capacitor ( $\Delta p_C$ ) can be formulated by describing the mentioned elements of the Windkessel model in relation to the flow ( $q$ ) or volume ( $V$ ), which can be given as follows:

$$\Delta p_R = Rq, \quad (3.10)$$

$$\Delta p_L = L \frac{dq}{dt}, \quad (3.11)$$

$$\Delta p_C = \frac{V}{C}. \quad (3.12)$$

Furthermore, the currently discussed mathematical model contains two valves; the mitral and the aortic valve. Both of these valves are described as a non-linear Bernoulli resistor, which results in a pressure drop across the valve ( $\Delta p_v$ ). This pressure drop can be described as:

$$\Delta p_v = \begin{cases} B_f q |q|, & q < 0, \\ B_b q |q|, & q \geq 0. \end{cases} \quad (3.13)$$

In this equation,  $B_f |q|$  denotes the resistance that will be experienced by the forward flow, whereas  $B_b |q|$  represents the resistance experienced by the backward flow. After describing the LV behavior and determining the behavior of the remaining model components, the total set of equations that describe the behavior of the cardiovascular system can be determined. This set of equations can be determined using the element relations described in Equation 3.10 to 3.12 and Kirhhoff's current and voltage law. Explicitly, the mathematical model can be given as:

$$f_{p_{LV}} = p_{LV}(t) - p_{LV}(t+1), \quad (3.14)$$

$$f_{p_{ao}} = p_{ao}(t) - \frac{V_{c1}}{C_{art}} - q_{av}(t)R_{prox} - \frac{L_{art}}{dt}q_{av}(t), \quad (3.15)$$

$$f_{V_{LV}} = V_{LV}(t) - V_{LV}(t-1) - q_{mv}(t)dt + q_{av}(t)dt, \quad (3.16)$$

$$f_{q_{mv}} = q_{mv}(t) \frac{B_{mv} + L_{mv}}{dt} - q_{mv}(t-1) \frac{L_{mv}}{dt} - p_{pul} + p_{LV}(t), \quad (3.17)$$

$$f_{q_{av}} = B_{av}q_{av}(t)|q_{av}(t)| + R_{prox}q_{av}(t) + \frac{L_{art}}{dt}q_{av}(t) - \frac{L_{art}}{dt}q_{av}(t-1) - p_{LV} + \frac{V_{c1}}{C_{art}}, \quad (3.18)$$

$$f_{V_{c1}} = V_{c1}(t) - V_{c1}(t-1) - q_{av}(t)dt + q_{c1}(t)dt, \quad (3.19)$$

$$f_{q_{c1-c2}} = q_{c1-c2}(t) \frac{R_{prox} + L_{art}}{dt} - q_{c1-c2}(t-1) \frac{L_{art}}{dt} - \frac{V_{c1}(t) - V_{c2}(t)}{C_{art}}, \quad (3.20)$$

$$f_{V_{c2}} = V_{c1}(t) - V_{c1}(t-1) - q_{c1-c2}(t)dt + q_{per}(t)dt, \quad (3.21)$$

$$f_{q_{per}} = q_{per}(t) - \frac{V_{c2}(t)}{C_{art}R_{dist}} - \frac{p_{per}}{R_{dist}}. \quad (3.22)$$

The above stated mathematical expressions are constructed in such a way that they all represent the residual error of a certain state. This residual error needs to be minimized as much as possible to be considered as an accurate description of the next time step. Therefore, this set of equations is being solved by using the Newton-Raphson method, an algorithm that attempts to iteratively improve the approximation of the roots of a given set of functions. This method starts with an initial estimate of the roots and improves the estimate based on the tangent of the set of functions [25]. The tangent can be represented by the Jacobian matrix ( $\mathbf{J}_f(\mathbf{x})$ ) when considering the current set of functions, which result in the following iterative approximation of the root based on the Newton-Raphson method:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{J}_f^{-1}(\mathbf{x}_k) \mathbf{f}(\mathbf{x}_k). \quad (3.23)$$

The iterative approximation of the root is accepted if the absolute sum of the minimization function solutions is lower than a certain predefined threshold. Furthermore, the majority of model input parameters will be adopted from the study of Meiburg et al. (2020) [49]. One alteration that is applied to the model input parameters is the lowering of the aortic valve resistance  $B_{av}$  to  $10^{-4}$  [mmHg s<sup>2</sup> / mL<sup>2</sup>]. This alteration is conducted because the mathematical model of Meiburg et al. (2020) was originally used to describe stenotic aortic valves, whereas in the current study the aortic valve is assumed to be non-stenotic. Furthermore, the left ventricular contractility ( $c_{LV}$ ) will become variable because this is the parameter that will be approximated in the current study.

## 3.2 Measurements

Recapitulating, a data assimilation approach is used in the current study to combine the information of model-based predictions with available measurements. Specifically, the objective of data assimilation is to sequentially improve the model-based predictions by making use of the available measurements [3]. The DA method used in the current study is ROUKF, which is described in Section 2.3, whereas the model-based predictions are accomplished by executing the mathematical model described in Section 3.1. The measured quantity used in the proposed ROUKF-based algorithm is the left ventricular pressure  $p_{LV}$ , which can be obtained through using various methods. In the current study, three types of data are considered: 1) synthetic data generated by the mathematical model, 2) in vitro data generated by a material twin and 3) in vivo data measured in patients. In the current section, all these three types of data will be elaborated separately in the given order.

### 3.2.1 Synthetic data

The use of synthetic data enables the simulation of multiple patient scenarios without the need of actually measuring them. The synthetic data are generated by executing the mathematical model that was described in the previous section (Section 3.1), which omits model discrepancy because the behavior of measurements is exactly described by the used mathematical model. As a result, the generated synthetic data are considered to be ideal to evaluate the capabilities of the proposed ROUKF-based algorithm, since the parameter of interest  $\theta$  is exactly known. In the current study, the parameter of interest  $\theta$  is the left ventricular contractility  $c_{LV}$ , which will be varied in the simulated patient scenarios. The range of simulated  $c_{LV}$  values used in the current study is determined based on the theoretical background with clinical experience. The mathematical model of Meiburg et al. (2020), which has been described in Section 3.1, utilizes the single fibre heart model of Bovendeerd et al. (2006) to describe the behaviour of the LV. In the study of Bovendeerd et al. (2006) is stated that the  $c_{LV}$  is a scaling factor for the active stress in the myocardial fibre that can vary between 0 and 1 [8]. However, this model parameter appears to be higher as 1 according to clinical data. Concretely, the left ventricular contractility  $c_{LV}$  will be varied between 0.4 to 1.4 in the current study.

A summary of all scenarios that will be simulated in the current study are given in Tabel 3.1. In the first six simulations, different initialization values of the  $c_{LV}$  will be simulated to investigate if the proposed ROUKF-based algorithm can deal with different values of  $c_{LV}$ . The following three simulations will be executed with an equally initialized  $c_{LV}$  of 0.8, whereas noise will be added to these simulations following a logarithmic increment. Based on this noise addition, it can be investigated if the proposed ROUKF-based algorithm can deal with noisy measurements. The added noise will be Gaussian distributed with a mean of zero and a simulation-specific standard deviation. More specifically, the standard deviation of the noise is defined as a percentage of the maximum measured  $p_{LV}$ , whereas different percentages are chosen for all three simulations. All simulations mentioned until now will be executed for a time span of 10 seconds, whereas the final two simulations, which will be specified next, will be executed for a time span of 30 seconds. These final two simulations contain multiple  $c_{LV}$  values, with a new  $c_{LV}$  being initiated every 10 seconds. These final three simulations are conducted to evaluate if the proposed ROUKF-based algorithm is capable of following changes in  $c_{LV}$  over time, which determines the robustness of the algorithm.

Test	$C_{LV}$ [-]	Noise [%]	Time [s]
1	0.4	-	10
2	0.6	-	10
3	0.8	-	10
4	1.0	-	10
5	1.2	-	10
6	1.4	-	10
7	0.8	10	10
8	0.8	20	10
9	0.8	50	10
10	0.6, 0.8, 1.0	-	30
11	0.6, 0.8, 1.0	20	30
12	1.4, 0.4, 1.2	20	30

**Table 3.1:** Simulations of various scenarios to generate the synthetic data that will be used to evaluate the capabilities of the proposed ROUKF-based algorithm.

### 3.2.2 In vitro data

In the current study, the in vitro data is generated by using a material twin, which is a device that is capable of mimicking the behavior of the human cardiovascular system. More specifically, this material twin mimics the cardiac function, coupled to the systemic and pulmonary circulation, including various auto-regulation mechanisms. These auto-regulation mechanisms will be discussed in the current section and can be switched on or off independently, making the material twin an appropriate device for the evaluation of model discrepancy. The difference between the reality and the mathematical model is known as model discrepancy, which can be used to evaluate the effect on the algorithm's ability to estimate the true parameter when such a discrepancy arises. Model discrepancy is considered as an important component of the uncertainty in the model output, where ignoring model discrepancy can lead to a biased and over-confident parameter estimation [10]. Consequently, the in vitro data will be generated using different configurations of the material twin, which will be used to evaluate the influence of model discrepancy. However, the in vitro data produced by the material twin are considered less perfect than the synthetic data, because it originates from a real-world physical object. The current section will be divided into two parts, the first part of which is dedicated to the elaboration of the different components of the material twin, and the second to the different configurations of the twin.

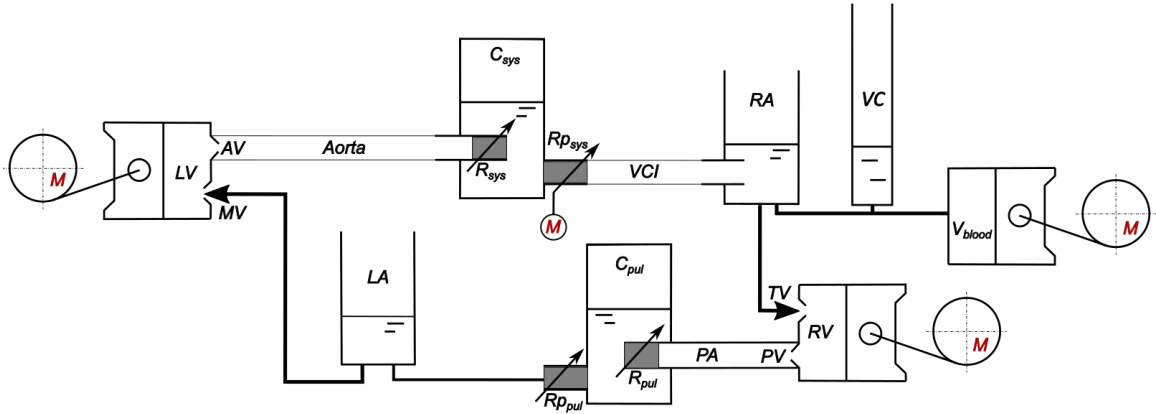
#### Material twin system

The material twin can be decomposed into two fundamental parts, namely a mechanical system combined with a real-time controllable system. The mechanical system consists of different mechanical components that together replicate the behavior of systemic and pulmonary circulation. In this material twin, the behavior of these two circulations is described by three-element Windkessel models, with the different elements of this Windkessel models being represented by mechanical components. The capacitor of the Windkessel model is represented by a container filled with both fluid and air, with the compliance being determined by the amount of pressure required to compress the air in the container and consequently allow an increased volume of fluid. The resistor of the Windkessel model is represented by the dimensions of the outflow aperture, whereby the resistance is controlled by obstructing or opening this aperture. Lastly, the inductor of the Windkessel model is a non-regulatable component in the material twin, where the inductance being dependent on the density of the fluid in combination with the dimensions of the connecting tube that facilitates the inflow into the three-element Windkessel model.

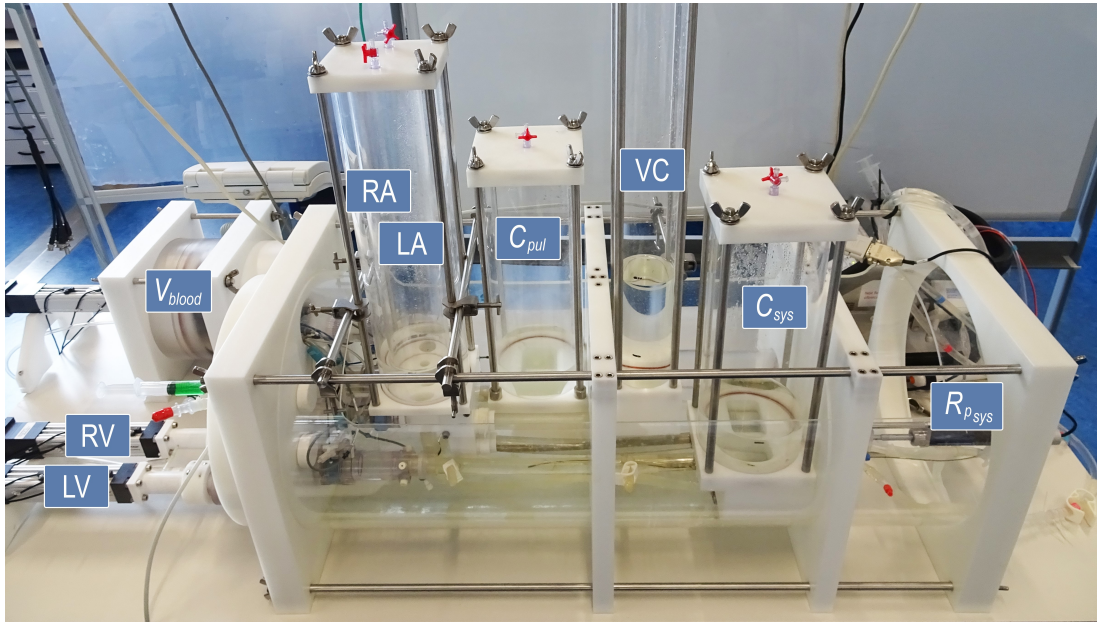
The second fundamental part of this material twin is the real-time controllable system. This controllable system consists of four linear actuators, which enable the movement of certain components. Two of these linear actuators are connected to piston pumps that represent the left and right ventricle, a third linear actuator controls the systemic peripheral resistance and a fourth linear actuator is connected to a piston pump which controls the fluid volume in the circulation. All of these linear actuators are driven by various mathematical models, which collaboratively enable the real-time control of the material twin. The two actuators that enable the pump function of both ventricles are controlled by the single fibre heart model, which originates from the study of Bovendeerd et al. (2006) [8]. The actuator that regulates the systemic peripheral resistance is driven by a mathematical model that describes the acute cardiovascular response, which was proposed in the study of Ursino and Magosso (2000) [72]. The actuator that controls the fluid volume in the circulation is driven by a mathematical model that describes the fluid exchange in the cardiovascular system, which originates from the study of Rosalina et al. (2019) [64]. Furthermore, this material twin contains a sophisticated auto-regulatory mechanism that regulates the blood pressure, which is controlled by a mathematical model that was proposed in the study of van Loo (2020) [41]. This auto-regulatory mechanism affects different aspects of the material twin; the systemic peripheral resistance ( $R_{p_{sys}}$ ), the heart rate ( $HR$ ) and the left ventricular contractility ( $c_{LV}$ ), whereas it can manually be adjusted which of these aspects are auto-regulated.

All mechanical and real-time controllable components are assembled to ultimately form a material twin capable of mimicking the behavior of the human cardiovascular system. A schematic representation of how all these components are connected to each other is given in Figure 3.2, while a real-world representation of the material twin is given in Figure 3.3. In both of these figures, the link between the material twin components and the human cardiovascular system is indicated to clarify their correspondence. Starting with the left ventricle (LV), where the linear actuator drives the piston of the pump that propels the fluid through the aortic valve (AV) into the aorta. Subsequently, the fluid arrives at a three-element Windkessel model that represents the systemic circulation, which consists of a systemic resistance ( $R_{sys}$ ) and a systemic compliance ( $C_{sys}$ ). Furthermore, the outflow of this systemic three-element Windkessel model

is regulated by the systemic peripheral resistance ( $R_{p_{sys}}$ ), which is controlled by a linear actuator. Then, the fluid flows from this systemic Windkessel model through the vena cava inferior (VCI) into a reservoir that represents the right atrium (RA), which in turn is connected to two different compartments. Firstly, to the venous compartment (VC) that acts as a reservoir for the total fluid volume, which is connected to a large piston pump that controls the total volume of fluid in the entire circulation ( $V_{blood}$ ). Secondly, to the right ventricular (RV) compartment, where the fluid that enters this compartment is passed through the tricuspidal valve (TV). The fluid in the right ventricular compartment is then propelled forward by the piston pump through the pulmonary valve (PV) into the pulmonary artery (PA) and towards a three-element Windkessel model that represents the pulmonary circulation. This Windkessel model consists of a pulmonary resistance ( $R_{pul}$ ) and a pulmonary compliance ( $C_{pul}$ ), combined with a pulmonary peripheral resistance ( $R_{p_{pul}}$ ) that controls the outflow of this Windkessel model. After leaving the pulmonary three-element Windkessel model, the fluid is guided to a reservoir that represents the left atrium (LA). Lastly, the fluid flows back into the left ventricle (LV) through the mitral valve (MV), which completes the material twin layout.



**Figure 3.2:** Schematic representation of the material twin, with the correspondence between components and the human cardiovascular system being explained in the text above. In this schematic representation, the motorized components are indicated with a red M.



**Figure 3.3:** Real-world representation of the material twin, with the correspondence between components and the human cardiovascular system being explained in the text above.

### Material twin configurations

However, this material twin is designed to mimic a far more complicated cardiovascular behavior than the one described by the mathematical model of the current study. To obtain a fair starting point of the evaluation of the proposed ROUKF-based algorithm, the material twin will be extensively simplified. Specifically, the fluid-filled reservoir that represents the right atrium (RA) will be dismissed and the linear actuators that control the right ventricular (RV), total fluid volume ( $V_{blood}$ ) and systemic peripheral resistance ( $R_{p_{sys}}$ ) will be deactivated. The result of these simplifications is a material twin that matches the cardiovascular behavior described by the mathematical model used the current study (Section 3.1), which is called Configuration 1. This configuration will be used to simulate a variety of  $c_{LV}$  scenarios, ranging from 0.4 to 1.4, in steps of 0.2. All these  $c_{LV}$  scenarios will be simulated for a duration of 30 seconds. Subsequently, this simplistic configuration (Configuration 1) will be made more complex by switching on auto-regulation mechanisms. Since the mathematical model will not be altered, the increased complexity of the material twin will result in an increase in model discrepancy.

In Configuration 2, the material twin will become a bit more complicated by allowing the auto-regulation of the systemic peripheral resistance ( $R_{p_{sys}}$ ). This second configuration will be simulated for the same variety of  $c_{LV}$  scenarios as Configuration 1. Subsequently, in Configuration 3, the auto-regulation of both the systemic peripheral resistance ( $R_{p_{sys}}$ ) and the heart rate ( $HR$ ) will be enabled, which will also be simulated for the same variety of  $c_{LV}$  scenarios as Configuration 1. This complexity build-up is then continued to create additional configurations which contain an increasing model discrepancy. Concretely, these additional configurations are obtained by adding the auto-regulation of blood volume ( $V_{blood}$ ), the auto-regulation of  $c_{LV}$  and lastly adding the pulmonary circulation. The application of this complexity build-up will eventually result in material twin Configuration 1 to 6, which are summarized in Table 3.2 below.

<i>Material twin configurations:</i>						
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
Simulated $c_{LV}$	0.4 $\rightarrow$ 1.4	0.4 $\rightarrow$ 1.4	0.4 $\rightarrow$ 1.4	0.4 $\rightarrow$ 1.4	0.4 $\rightarrow$ 1.4	0.4 $\rightarrow$ 1.4
Auto-regulation $R_{p_{sys}}$	×	✓	✓	✓	✓	✓
Auto-regulation $HR$	×	×	✓	✓	✓	✓
Auto-regulation $V_{blood}$	×	×	×	✓	✓	✓
Auto-regulation $c_{LV}$	×	×	×	×	✓	✓
Pulmonary circulation	×	×	×	×	×	✓

**Table 3.2:** Summary of the different configurations of the material twin that will be used to evaluate the following  $c_{LV}$  values; 0.4, 0.6, 0.8, 1.0, 1.2 and 1.4 (0.4  $\rightarrow$  1.4), for a duration of 30 seconds. The different configurations are ordered by degree of sophistication, starting with a simplistic Configuration 1 and ending with a complex Configuration 6.

### 3.2.3 In vivo data

The in vivo data used in the current study originates from the study of Johnson et al. (2018) [29]. In that particular study, it was aimed to describe the pressure loss ( $\Delta p$ ) vs. flow ( $q$ ) relationships in patients with a stenotic aortic valve. In that study, invasive measurements were performed before and after application of a transcatheter aortic valve implementation (TAVI). One of the quantities that was measured in that study is the left ventricular pressure  $p_{LV}$ , which was measured by using a catheter with a pressure wire and placing it in the left ventricle. The inserted pressure wire was calibrated using a recording system (Quantien analyzer, St. Jude Medical), whereas this system was also used to capture the pressure signal. While performing these pressure measurements, a step-wise intravenous administration of dobutamine was conducted [29]. Dobutamine is a pharmaceutical positive inotropic agent which leads to an increase of cardiac output due to an increase in stroke volume, which is caused by the enhancement of the left ventricular contractility  $c_{LV}$  [65].



The measured  $p_{LV}$  data of the study of Johnson et al. (2018) will be used as an input for the proposed ROUKF-based algorithm, allowing the validation of the proposed algorithm on in vivo data. It is suspected that the proposed algorithm will estimate an increase in left ventricular contractility  $c_{LV}$  over time due to the step-wise dobutamine administration. Furthermore, the maximum pressure gradient of the left ventricle (LV  $\frac{dp}{dt}_{max}$ ) will be determined for every cardiac cycle. Subsequently, the mean of these maximum pressure gradients will be determined for every moving time window (see Section 4.2). As a result, the LV  $\frac{dp}{dt}_{max}$  will be determined on the same interval as the  $c_{LV}$ , which allows comparison of these two parameters to determine whether they provide the same cardiovascular information. This comparison will be conducted based on visual trend evaluation.

## 4. Data processing

This chapter will be dedicated to the data processing that will be applied to sequentially estimated statistics, which are obtained by using the ROUKF algorithm (Section 2.3). The first part of this chapter will be devoted to the description of the probabilistic nature of these sequentially estimated statistics, which will be quantified by making use of a probability density function (PDF). In the second part of this chapter, the length of the analysed time interval, the re-initialization of certain elements and the moving average of the cardiac cycle time will be described in detail.

### 4.1 Probability density function

When there is complete and perfect knowledge of a certain system combined with the equations that govern its evolution over time, perfect predictions of this certain system can be obtained. However, a perfect prediction does not exist and therefore Bayes' theorem aims to describe conditional probability, in other words the probability that event  $A$  will occur, knowing that event  $B$  has occurred [4]. Suppose that  $\theta$  represents the parameter of interest and  $y$  the data or measurable quantity, then Bayes' theorem can be given as:

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)}. \quad (4.1)$$

Based on this theorem of Bayes, the *a posteriori* probability density function (PDF) of parameter  $\theta$  given data  $y$ , which is denoted as  $P(\theta|y)$ , can be determined. In this theorem,  $P(\theta)$  represents the *a priori* PDF that represents all the knowledge about parameter  $\theta$ ,  $P(y|\theta)$  denotes the likelihood function of the data  $y$  assuming parameter  $\theta$  and  $P(y)$  represents the marginal distribution of  $y$ . Subsequently, Bayes' theorem suggests to describe the parameter of interest in terms of a PDF, which takes into consideration the stochastic nature of the system under consideration [12]. However, the DA methods arising from this Bayesian theorem pursue a complete knowledge of the *a posteriori* PDF, which is impossible to determine in the real world [4]. Moreover, it is known that when the considered system is high-dimensional, only approximations of the *a posteriori* PDF can be obtained. In case of such a high-dimensional system, it is commonly assumed that the *a posteriori* PDF is Gaussian distributed. Based on this assumption, the PDF can be computed analytically based on sequentially estimated statistics, which can be obtained using a Kalman filter-oriented approach [13]. Explicitly, the Gaussian *a posteriori* PDF can be computed by using the following function:

$$P(\psi_i|\hat{\theta}_k, \sigma_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2}(\frac{\psi_i - \hat{\theta}_k}{\sigma_k})^2}, \quad \text{with } i = 1, 2, \dots, n_{bins}. \quad (4.2)$$

In this Gaussian distributed *a posteriori* PDF expression, the estimated parameter  $\hat{\theta}_k$  and corresponding standard deviation  $\sigma_k$ , both at a certain time step  $k$ , can be obtained from the statics estimated by the ROUKF method. First of all, the estimated parameter  $\hat{\theta}_k$  can directly be obtained from the *a posteriori* estimate  $\hat{x}_{k+1}^a$  (Equation 2.60). Second, the variances belonging to this *a posteriori* estimate  $\hat{x}_{k+1}^a$  are located on the diagonal of the *a posteriori* covariance  $\mathbf{P}_{k+1}^a$  (Equation 2.59). By taking the square root of these variance, the standard deviation  $\sigma_k$  of the estimated parameter  $\hat{\theta}_k$  can be obtained. Lastly, the parameter value is denoted by  $\psi_i$ , which will be used to evaluate the estimated parameter  $\hat{\theta}_k$ . The parameter value  $\psi_i$  originates from a vector  $\psi$  with possible parameter values, which contains the same number of parameter values as the number of bins ( $n_{bins}$ ) in the PDF. The number of bins in the PDF is determined by two aspects; the bin width ( $b$ ) and the interval of the PDF  $[\delta^-, \delta^+]$ . In the current study, the bin width ( $b$ ) will be calculated by using the Freedmann-Diaconis rule, which was originally designed to minimise the discrepancy between the histogram and the density of the theoretical probability distribution [24]. According to the Freedmann-Diaconis rule, the bin width ( $b$ ) can be calculated by:

$$b = 2 \frac{\Phi_{\hat{\theta}}}{\sqrt[3]{N_k}}. \quad (4.3)$$

Here,  $N_k$  denotes the number of time steps included in the *a posteriori* PDF and  $\Phi_{\hat{\theta}}$  denotes the interquantile range of the estimated parameter vector  $\hat{\theta}$ . This estimated parameter vector  $\hat{\theta}$  contains all estimates of the parameter of interest  $\theta$  ( $\hat{\theta}_k$ ) from  $N_k$  time steps. As mentioned before, the second aspect that affects the number of bins ( $n_{bins}$ ) in the PDF is the interval of the PDF  $[\delta^-, \delta^+]$ , which can be calculated by:

$$\delta^- = \min(\hat{\theta}) - 3 \frac{\sum_{k=1}^{N_k} \sigma_k}{N_k} \quad \text{and} \quad \delta^+ = \max(\hat{\theta}) + 3 \frac{\sum_{k=1}^{N_k} \sigma_k}{N_k}. \quad (4.4)$$

These expressions represent the lower limit  $\delta^-$  and upper limit  $\delta^+$  of the PDF interval, with  $\min(\hat{\theta})$  and  $\max(\hat{\theta})$  indicating the minimum and maximum value of the estimated parameter vector  $\hat{\theta}$ , respectively. Next, three times the mean of the standard deviation  $\sigma_k$  is subtracted from the minimal value and added to the maximal value, resulting in boundaries of the PDF interval  $[\delta^-, \delta^+]$ . In essence, the width of this PDF interval is determined by the min-max range of the estimated parameter vector  $\hat{\theta}$  with a three times the mean standard deviation  $\sigma_k$  to both sides of the range. The idea of adding three times the standard deviation to both sides originates from a statistical point of view, where it is known that 99.7% of the observed data can be found within three standard deviations from the mean. Conversely, three times the standard deviation is added to the min-max range in the current study, which is too wide according to the stated statistical rule. However, this wider PDF interval is chosen to incorporate the sometimes volatile nature of the estimated parameter  $\hat{\theta}$  in case the observed hemodynamic behavior of the cardiovascular system changes rapidly.

After determining the number of bins ( $n_{bins}$ ) and the PDF interval  $[\delta^-, \delta^+]$ , the vector of possible parameter values  $\psi$  can be composed by distributing the number of bins evenly over the PDF interval  $[\delta^-, \delta^+]$ . Every possible parameter value in this vector  $\psi$  will be evaluated with respect to the parameter estimate  $\hat{\theta}_k$  and standard deviation  $\sigma_k$  of a certain time step  $k$  by using Equation 4.2. After completing this evaluation for every element in  $\psi$ , the cumulative product of all these evaluations finally represents the PDF of a certain time step  $k$ , which can be denoted as  $P(\psi|\hat{\theta}_k, \sigma_k)$ . The eventual *a posteriori* PDF will be constructed by summing all  $P(\psi|\hat{\theta}_k, \sigma_k)$  of  $N_k$  time steps, whereas the contribution of every time step will be weighted by using weights  $\omega_k$  (Equation 5.2). The weighted *a posteriori* PDF can be determined by:

$$P(\psi|\hat{\theta}, \sigma) = \frac{\sum_{k=1}^{N_k} \omega_k P(\psi|\hat{\theta}_k, \sigma_k)}{\sum_{k=1}^{N_k} \omega_k}. \quad (4.5)$$

However, to ensure a positive value of the parameter of interest  $\theta$ , a logarithmic transformation will be applied in the current study. This logarithmic transformation can be illustrated based on the parameter value  $\psi_i$ , which can be represented logarithmically as:  $\phi_i = \phi_{ref} 2^{\psi_i}$ . In this transformation, the normalization factor  $\phi_{ref}$  is used to obtain a dimensionless logarithmic parameter value. Consequently, the applied logarithmic transformation will result in an altered distribution of the *a posteriori* PDF, which can be defined as a log-normal distribution. Explicitly, the log-normal distributed PDF can be formulated as:

$$P(\phi_i|\mu_k, s_k) = \frac{1}{\phi_i \sqrt{2\pi} s_k} e^{-\frac{1}{2} \left( \frac{\ln(\phi_i) - \mu_k}{s_k} \right)^2}, \quad \text{with } i = 1, 2, \dots, n_{bins}. \quad (4.6)$$

In this expression, the logarithmic transformed parameter value is denoted by  $\phi_i$ , with the logarithmic parameter estimate  $\mu_k$  and corresponding logarithmic standard deviation  $s_k$ . The logarithmic parameter estimate  $\mu_k$  and logarithmic standard deviation  $s_k$  can be calculated by using:

$$\mu_k = \hat{\theta}_k \ln(2) + \ln(\theta_{ref}) \quad \text{and} \quad s_k = \sigma_k \ln(2). \quad (4.7)$$

The total derivation of the log-normal distributed PDF (Equation A.8) and its associated components (Equation 4.7) can be found in Appendix A. The log-normal distributed PDF of a certain time step  $k$  will be constructed in the same fashion as the normal distributed PDF, by evaluating every element of the logarithmic parameter vector  $\phi$  with respect to logarithmic parameter estimate  $\mu_k$  and logarithmic standard deviation  $s_k$ . The cumulative product of these evaluations will result in the log-normal distributed PDF of a certain time step  $k$ , which can be denoted as  $P(\phi|\mu_k, s_k)$ . The eventual log-normal distributed *a posteriori* PDF can be obtained by performing this for  $N_k$  time steps and subsequently weighting the result:

$$P(\phi|\mu, s) = \frac{\sum_{k=1}^{N_k} \omega_k P(\phi|\mu_k, s_k)}{\sum_{k=1}^{N_k} \omega_k}. \quad (4.8)$$

The final parameter prediction of  $N_k$  time steps is represented as the mean  $\mu$  of this *a posteriori* PDF  $P(\phi|\mu, s)$ , with the corresponding 95% confidence interval being calculated by subtracting/adding two times the standard deviation  $s$  from/to the mean  $\mu$ . The number of time steps  $N_k$  included in this weighted log-normal distributed *a posteriori* PDF  $P(\phi|\mu, s)$  depends on the length of the analysed time interval in combination with the duration of a single time step. The length of the analysed time interval length will be detailed in the next subsection, whereas the duration of a single time step depends on the used data.

Lastly, the root-mean-square error (RMSE) is a measure which will be used to quantify the discrepancy between the estimated parameter value  $\hat{\theta}_k$  of a certain time step  $k$  and the final parameter prediction  $\mu$ . This measure will be based on the same number of time steps  $N_k$  as included in the weighted log-normal distribution and can be calculated by using the following equation:

$$RMSE = \sqrt{\sum_{k=1}^{N_k} \frac{(\mu - \hat{\theta}_k)^2}{N_k}}. \quad (4.9)$$

## 4.2 Moving time window

In the current study, a moving time window is used to provide a semi real-time update of the parameter of interest  $\theta$ . The length of this moving time window needs to be narrow enough to provide a semi real-time update of parameter  $\theta$ , whereas it needs to be wide enough to enable the ROUKF algorithm to converge sufficiently. Therefore, the length of the moving time window is arbitrarily chosen at 10 seconds for application intended in the current study, where the time interval of one moving time window can be stated as:  $[t_k, t_{k+N_k}]$ . However, at the first iteration of every moving time window, which is basically at time step  $t_k$ , certain elements of the ROUKF algorithm will re-initialized. This re-initialization is conducted to ensure the convergence of the estimated parameter  $\hat{\theta}$  over time. The re-initialization is considered necessary for convergence because certain elements of the ROUKF algorithm do contain a certain memory, which withholds the convergence of the parameter of interest  $\theta$  in case of a constantly changing environment in time. The two elements of the ROUKF algorithm (Section 2.3.3) that will be re-initialized at the first iteration of every moving time window are the Luenberger matrix in the prediction space  $\mathbf{L}_{k=N_k}$  (Equation 2.54) and the uncertainty matrix  $\mathbf{U}_{k=N_k}$  (Equation 2.58), explicitly:

$$\mathbf{L}_{k=N_k}^\theta = \mathbf{I}, \quad \mathbf{L}_{k=N_k}^x = \mathbf{O}, \quad \mathbf{U}_{k=N_k} = \mathbf{I}. \quad (4.10)$$

All other elements of the ROUKF algorithm will be left untouched during this re-initialization. However, not the entire time interval of the moving time window  $[t_k, t_{k+N_k}]$  will be used to construct the log-normal distributed *a posteriori* PDF due to inlet effects.

The inlet effects are caused by the re-initialization of certain elements of the ROUKF algorithm and contain no truthful information regarding the parameter of interest  $\theta$ , as it takes a certain time for the algorithm to converge. Therefore, the first 5 seconds of every moving time window will be dismissed from the analysis due to inlet effects. Concretely, the analysed time interval of the moving time window can be expressed as:  $[t_{k+N_{start}} = 5\tau[s], t_{k+N_k} = 10\tau[s]]$ , where  $\tau$  indexes the number of the moving time window. In this time interval expression, the number of time steps before the start of the analysed interval of the  $\tau$ -th moving time window is denoted by  $N_{start}$  and the total number of time steps included in the  $\tau$ -th moving time window are indicated by  $N_k$ . Note that both variables,  $N_{start}$  and  $N_k$ , are dependent on the used data generation method because the duration of a single time step will vary between different methods.

However, one input parameter of the mathematical model will also be redefined throughout the execution of the proposed ROUKF-based algorithm and that is the duration of one cardiac cycle ( $T_0$ ). The duration of the cardiac cycle is decisive for the relative increment of a single time step within the cardiac cycle, which is considered to be essential for the synchronization between the mathematical model and the patient. Therefore, the average duration of the cardiac cycle  $T_{0,av}$  will be determined by using a moving average. This moving average will use the information about the duration of the last 10 cardiac cycles to determine the average duration of the cardiac cycle  $T_{0,av}$ . When a new cardiac cycle is initiated, the information about the duration of the previous cardiac cycle is added to the moving average, whereas the information about the last cardiac cycle in the moving average is dismissed. As a result, the average duration of the cardiac cycle  $T_{0,av}$ , based on the last 10 cardiac cycles, will be used to determine the relative time increment of a single time step at the initiation of the contraction phase. Moreover, the determination of the initiation of the contraction phase is considered to be the other essential component for the time synchronization between the mathematical model and the patient. However, the determination of the contraction initiation is dependent on the used type of data.

## 5. Data and model analysis

The data and model analysis that will be applied to either the sequentially estimated statistics of the ROUKF method (Section 2.3) or the mathematical model (Section 3.1), will be elaborated in the current chapter. More specific, the first two sections will be devoted to data analysis, whereas the last section will be attributed to model analysis. The sensitivity analysis will be discussed in the first section of this chapter, which will be used to evaluate the sensitivity of the measured states to changes in the parameter of interest throughout the heart cycle. In the second section of this chapter the identifiability analysis will be detailed, which will be used to identify if the right type of data are collected regarding a certain model structure. The third and final section of will be devoted to the stability analysis, which aims to quantify the sensitivity of a non-linear system to its initial states.

### 5.1 Local sensitivity analysis

In the current study, the local sensitivity analysis will be used to identify the sensitivity of the measured states to changes in the parameter of interest. When the measured states are not sensitive to changes in the parameter of interest, it is considered to be unlikely that these states can be valuable for a truthful estimation of the parameter of interest [55]. Therefore, the contribution of every time step to the eventual *a posteriori* PDF will be weighted accordingly. Furthermore, this local sensitivity analysis can provide insights in which parts of the cardiac cycle are most influential to estimate the parameter of interest by using the measured states [49].

A local sensitivity method is based on evaluating the effect on the measured states when an infinitesimal change is applied to one of the input parameters, while all other input parameters remain unchanged. This infinitesimal change will be applied around a certain reference value of the input parameter in question, thereby obtaining only local information of this input parameter [21]. The local sensitivity can be calculated by the following equation:

$$S_{loc,i,k} = \frac{1}{z_{i,ref}} \frac{\partial h_{i,k}}{\partial \psi_k}. \quad (5.1)$$

Here, the local sensitivity ( $S_{loc,i,k}$ ) represents the variations in the measured states  $h_{i,k}$  with respect to a infinite small change of parameter value  $\psi$ . The measured states follows from the observation vector  $h_{i,k}$ , where the  $i$ -th element of this vector indicates the current state of interest. This local sensitivity  $S_{loc,i,k}$  will be calculated for each measurable state  $i$  and each time step  $k$ . Lastly, the local sensitivity will be normalized by reference value  $z_{i,ref}$ , which is the maximal measured value of state  $i$  of the current moving time window. This normalization is performed to obtain a non-dimensional local sensitivity, which is necessary to correctly compare the sensitivity of different measurement states regarding the parameter of interest. The information provided by  $S_{loc,i,k}$  will be used to weight the contribution of time step  $k$  to the *a posteriori* PDF. The weight can be calculated by:

$$w_k = \sqrt{\sum_{i=1}^{n_z} (S_{loc,i,k})^2}. \quad (5.2)$$

A small (i.e. close to zero) value of the weight of the parameter estimate at time step  $k$  ( $w_k$ ) would indicate that the measured states are not sensitive to the parameter at this time step. In other words, this would mean that this time step ( $k$ ) contains a negligible amount of information about the parameter of interest [55]. However, this local sensitivity method includes only a small part of the input parameter space, namely only very local information around a certain reference value. Thereby, the relation between the input parameter and model output is often non-additive or non-linear in certain parts of the input space. Consequently, the outcome of the local sensitivity analysis can be misleading when used to quantify the uncertainty in the model output that is caused uncertain input parameters [21]. The global sensitivity analysis is considered to be more appropriate to determine the relative contribution of uncertainty input parameters and their interaction to the total amount of uncertainty [61]. A suitable method to conduct the global sensitivity analysis is the adaptive generalized polynomial chaos expansion (agPCE), which is cover in Appendix B.

## 5.2 Identifiability analysis

Let us take a step back and imagine an ideal world in which the data and knowledge about a particular cardiovascular system are so extensive that the mathematical model can be defined in such a way that it always returns the correct answer. In such an ideal world, the uncertainty could either be reduced or even nullified and biological variability could be quantified perfectly. However, in the real world, the knowledge and data of a specific system are usually insufficient to conceptualize and simulate it flawlessly. And even if the behavior of a certain system is understood to a significant degree, its complexity and temporal variability can only be approximated [26]. Besides, it is a tough challenge to define the level of model complexity that describes the complexity and variability of a system in such a way that an accurate model output can be obtained. In general, the more accurate the model is able to describe the reality, the more complex it becomes. Therefore, it can be stated that the definition of the model structure affects the degree of discrepancy between the model output and the real world [21]. When an appropriate model structure is selected to describe a particular cardiovascular system, more data will result in a smaller uncertainty of model output. However, this convention is only valid when the right type of data regarding this particular system are collected. The technique that can be used to investigate if the right data are collected, given a certain model structure, is parameter identifiability analysis. This identifiability analysis evaluates the possibility to identify a unique vector of parameter values from the measured data, given a certain model structure [26].

Identifiability analysis is considered an inverse problem and refers to the possibility of determining the input parameters from a given model output, which is the opposite of the earlier mentioned forward problem [26]. However, there is a fundamental mathematical difference between the forward and the inverse problem; forward problems are almost always well-posed, whereas inverse problems are often ill-posed. Two of the well-posed properties, the existence and uniqueness of a solution, are collectively known as the identifiability of a mathematical problem. Conversely, an ill-posed mathematical problem allows an infinite number of possible solutions and can therefore be considered as non-identifiable [3].

In the current research, the profile likelihood method will be used to exploit the identifiability analysis. This profile likelihood method makes use of a cost function, which represents the agreement of measured states with the output predicted by the model [75]. The cost function is represented as the weighted sum of squared residuals, which is stated by the following equation:

$$J(\theta) = \sum_{i=1}^{n_z} \frac{1}{N_c} \frac{1}{N_k} \sum_{k=1}^{N_k} \left( \frac{z_{i,k} - y_i(t_k|\theta)}{\sigma(e_{i,k}^z)} \right)^2. \quad (5.3)$$

The  $J(\theta)$  represents the cost function in which the agreement between the measured states ( $z_{i,k}$ ) and the predicted model output ( $y_i(t_k|\theta)$ ) is expressed. In the term  $z_{i,k}$ , the  $i$  denotes the measured state and  $k$  denotes the time step. The  $y_i(t_k|\theta)$  reflects the predicted model output of the measured state  $i$  at time step  $k$  concerning the parameter of interest  $\theta$ . The difference between  $z_{i,k}$  and  $y_i(t_k|\theta)$  is scaled with the standard deviation of the measurement error  $\sigma(e_{i,k}^z)$  of measured state  $i$  at time step  $k$ . This scaled difference is squared and averaged over the number of time steps ( $N_k$ ), which is subsequently normalized to the number of included cardiac cycles ( $N_c$ ). This process is repeated for every measured state, where the eventual cost function  $J(\theta)$  will consist of information from all measured states ( $n_z$ ). Finally, the parameter of interest  $\theta$  can be estimated numerically by minimizing the cost function  $J(\theta)$ , as given by:

$$\hat{\theta} = \operatorname{argmin}[J(\theta)]. \quad (5.4)$$

The confidence interval  $[\sigma_{\theta}^-, \sigma_{\theta}^+]$  of this estimated parameter  $\hat{\theta}$  will be determined based on the chosen confidence level  $\gamma$ . In the current research, the confidence level  $\gamma$  is chosen at 0.95, which implies that the true value of parameter  $\theta$  will be located within interval  $[\sigma_{\theta}^-, \sigma_{\theta}^+]$  with a probability of  $\gamma$ . In the paper of Raue et al. (2009) two distinct profile likelihood-oriented methods are suggested to determine the confidence interval of the estimated parameter; the asymptotic method and the finite sample method. The asymptotic method is dependent on the amount of analysed data, the measurement noise and the linearity of the dependence between  $z_{i,k}$  and  $\theta$  [62]. The finite sample method does not contain these dependencies and is therefore considered to be superior over the asymptotic method for determining the confidence interval [53]. In this finite sample method, a threshold is applied to the cost function, with the

intersection points of the cost function and the threshold representing the boundaries of the confidence interval [62]. This method will be used in the current study and is given by the following mathematical expression:

$$\{(\theta | J(\theta) - J(\hat{\theta})) < \Delta_\gamma\} \text{ with } \Delta_\gamma = \chi^2_\gamma(df = n_p). \quad (5.5)$$

In this equation, the threshold ( $\Delta_\gamma$ ) is determined by the defining the  $\gamma$  quantile of the  $\chi^2_{df}$ -distribution, with the degrees of freedom (df) being equal to the number of parameters ( $n_p$ ). To define the intersection points between the cost function and the threshold, the actual shape of the cost function needs to be approximated first. The currently discussed profile likelihood method will start the approximation at the estimated parameter ( $\hat{\theta}$ ) and subsequently trace a path through the parameter space. This path will be computed by forcing the parameter of interest to change incrementally, while optimizing all other parameters. This process will continue until the threshold is reached, in other words the confidence interval for this particular parameter estimate  $\hat{\theta}$  is determined. Concretely, the decision rule of the currently discussed PL method can be based on the maximum likelihood estimation (MLE) combined with the threshold of the confidence interval. The complete derivation of this decision rule is given in Appendix C, which can finally be defined mathematically as follows:

$$-2\log\left(\frac{L(\theta_{PL})}{L(\theta_{opt})}\right) \leq \chi^2_{1-\gamma,1}. \quad (5.6)$$

Here, the maximum likelihood estimation (MLE) is used to find model parameters that result in a probability density function that most likely reflects the measured data. These model parameter values are found by maximizing the likelihood function  $L(\theta)$  such that the measured data is most probable. In the equation above, the  $L(\theta_{PL})$  represents the likelihood for the incrementally changing parameter of interest  $\theta$ , while optimizing the other parameters. However, note that in the current study only one parameter ( $\theta$ ) is estimated, which results in the fact that  $L(\theta_{PL})$  only represents the likelihood for the incrementally changing parameter  $\theta$ . The  $L(\theta_{opt})$  term denotes the likelihood of the optimal parameter ( $\hat{\theta}$ ), which was already determined by minimizing the cost function. These two likelihood terms are the result of two model circumstances ( $Model(\theta_{PL})$  and  $Model(\theta_{opt})$ ) that can be transformed into each other by imposing a specific linear constrain to the parameter [75].

Finally, the identifiability analysis can be conducted after the determining the estimated parameter  $\hat{\theta}$  and the corresponding confidence interval  $[\sigma_\theta^-, \sigma_\theta^+]$ . The parameter of interest  $\theta$  can be considered as identifiable if the confidence interval of the estimated parameter is finite. However, if the confidence interval of the estimated parameter is unbounded in either one direction or both directions, the parameter of interest  $\theta$  can be considered as non-identifiable. Two distinct types of non-identifiability can be distinguished; structural- and practical non-identifiability, which are illustrated in Figure 5.1 below.



**Figure 5.1:** Three plots of  $J(\theta)$  for a one-dimensional parameter space are depicted regarding three different cases of identifiability. On the left a structural non-identifiable parameter, in the middle a practical non-identifiable parameter and on the right an identifiable parameter. Adapted from: ‘Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood.’ by A. Raue et al, 2009, *Bioinformatics*, 25, 1923–1929 [62].



Structural non-identifiability is illustrated on the left of Figure 5.1, which can be observed as an unbounded confidence interval in both directions. This structural non-identifiability is invariably related to model structure and results in a non-uniqueness of the estimated parameter  $\hat{\theta}$ . The middle image of Figure 5.1 depicts practical non-identifiability, which can be observed by the confidence interval being unbounded in only one direction. This practical non-identifiability is mostly related to the absence of a particular effect in the measured data or to observation errors. Lastly, the right illustration of Figure 5.1 shows an identifiable parameter in which the confidence interval is bounded in both directions [26, 62].

### 5.3 Stability analysis

The initial states of a cardiovascular system are never known exactly because the available measurements and used data assimilation method are both subjected to a finite accuracy and precision. The implication of this finite precision can be evaluated by perturbing initial states, which generates an anomalous trajectory in the system's input space compared to a certain reference trajectory [71]. To evaluate and predict the reaction of the system to perturbed initial states, the sensitivity of a system to its initial states needs to be quantified. There are certain mathematical theories available that determine the degree of instabilities in the given system, which can be used to quantify the sensitivity to its initial states [13]. The evaluation of the stability of a system is crucial to the overall performance of the used data assimilation method. Moreover, it is also essential for the convergence of the sequential estimates to the actual evolution of the considered physical system and thus ultimately for obtaining a truthful representation of the system [14]. Using certain mathematical approaches, the stability of a system can be examined without knowing the solutions of its differential equations for a given initial condition [54]. The mathematical approach used to quantify the stability in the current study is based on the Lyapunov method.

#### 5.3.1 First and second method of Lyapunov

Lyapunov's method was described in 1890 by Aleksandr Mikhailovich Lyapunov and can be used to evaluate the stability of a system that is governed by ordinary differential equations [45]. The original Lyapunov method, in other words the first method of Lyapunov, is applicable to a system consisting of linear time-invariant ordinary differential equations, given as:

$$\dot{\underline{x}} = \underline{f}(\underline{x}) = \mathbf{A}\underline{x}. \quad (5.7)$$

Explicitly, a time-invariant dynamic system is a system that is not varying over time, which is also known as an autonomous dynamic system [54]. The stability analysis of this linear autonomous system (5.7) can be obtained by evaluating the eigenvalues  $\lambda_i$  of  $\mathbf{A}$ . Following the first method of Lyapunov, the equilibrium point  $\underline{x}^-$  is asymptotically stable if and only if all eigenvalues of  $\mathbf{A}$  satisfy  $Re(\lambda_i) < 0$ . Intuitively can be stated that asymptotic stability requires all eigenvalues of  $\mathbf{A}$  to be located on the left-half of the complex plane. If one or more eigenvalues of  $\mathbf{A}$  are located on the right-half of the complex plane ( $Re(\lambda_i) > 0$ ), then the equilibrium point  $\underline{x}^-$  is unstable. Lastly, if at least one eigenvalue of  $\mathbf{A}$  lies on the imaginary axis ( $Re(\lambda_i) = 0$ ), no conclusions concerning stability can be drawn. However, this first method of Lyapunov is only applicable to linear autonomous systems, whereas in the real world most systems are governed by non-linear differential equations. Other drawbacks of this first method are that the stability approximation is only valid near the equilibrium point  $\underline{x}^-$  and that for some scenarios ( $Re(\lambda_i) = 0$ ) no conclusions can be drawn concerning stability [67].

The second method of Lyapunov is often referred to as the direct method of Lyapunov and can be used for dynamic systems that consist of non-linear differential equations. This method generates a scalar energy-like function  $V(\underline{x})$  from which conclusions can be drawn concerning the stability of the set of differential equations without an explicit knowledge of corresponding solutions [54, 67]. This energy-like function  $V(\underline{x})$  is a useful substitute for the true energy function of a given system, because it is very complicated and sometimes even impossible to construct the true energy function for a complex dynamic system [54]. If such a function  $V(\underline{x})$  exists for the given system, then this function can be called a Lyapunov function. The first two characteristics of the Lyapunov function  $V(\underline{x})$  originate from the property of being positive definite ( $V(\underline{x}) > 0$ ) in the region  $\Omega$ . Positive definiteness of  $V(\underline{x})$  implies that  $V(\underline{x}) = 0$  at the equilibrium point  $\underline{x}^-$ , while  $V(\underline{x}) > 0$  for all other values of  $\underline{x}$  in  $\Omega$ . Mathematically this can be stated as:

$$V(x) = 0, \quad x = x^- = \operatorname{argmin}(V(x)), \quad (5.8)$$

$$V(x) > 0, \quad x \in \Omega, \quad x \neq x^-. \quad (5.9)$$

The third and fourth characteristics of the Lyapunov function  $V(x)$  are associated with the dynamics of the Lyapunov function. These dynamics will be evaluated by taking the Lie derivative of  $V(x)$ , which is mathematically denoted as  $\dot{V}(x)$ . If  $\dot{V}(x)$  is negative semi-definite ( $\dot{V}(x) \preceq 0$ ), it would indicate that  $\dot{V}(x) \leq 0$  for all  $x$  in  $\Omega$ , which can be formulated as:

$$\dot{V}(x) = \nabla V(x)f(x) \leq 0, \quad x \in \Omega. \quad (5.10)$$

If  $V(x)$  is positive definite ( $V(x) \succ 0$ ) and  $\dot{V}(x)$  is negative semi-definite ( $\dot{V}(x) \preceq 0$ ), in other words Equations 5.8 - 5.10 hold, the equilibrium point  $x^-$  is *locally stable*. However, if  $\dot{V}(x)$  is strictly negative definite ( $\dot{V}(x) \prec 0$ ), this would imply that  $\dot{V}(x) < 0$  for all  $x$  in  $\Omega$ , with exception of  $x^-$ . Formally, this can be denoted as follows:

$$\dot{V}(x) = \nabla V(x)f(x) < 0, \quad x \in \Omega, \quad x \neq x^-. \quad (5.11)$$

When  $V(x)$  is positive definite ( $V(x) \succ 0$ ) in  $\Omega$  and  $\dot{V}(x)$  is negative definite ( $\dot{V}(x) \prec 0$ ) in  $\Omega$ , in other words Equations 5.8 - 5.11 are valid, then the equilibrium point  $x^-$  can be considered as *locally asymptotically stable*. Now suppose that the region where both  $V(x) \succ 0$  and  $\dot{V}(x) \prec 0$  hold is equal to the total input parameter space ( $\Omega = \mathbb{R}^n$ ), then the Lyapunov function  $V(x)$  can be considered as unbounded, which implies that  $V(x) \rightarrow \infty$  if  $\|x\|_2 \rightarrow \infty$ . Summarize, the unboundness of  $V(x)$  can be given as:

$$\Omega = \mathbb{R}^n, \quad V(x) \rightarrow \infty \quad \text{if} \quad \|x\|_2 \rightarrow \infty. \quad (5.12)$$

If Equations 5.8 - 5.12 are all applicable to the considered dynamic system, then the equilibrium point  $x^-$  can be considered as *globally asymptotically stable*. In general, these asymptotic statistics are preliminary determined by the ability of the considered system to control the expansion of the error in its initial conditions. However, due to the fact that real-life systems will evolve over time, the mathematical model used to describe the dynamics will be inadequate to represent the reality. Strictly speaking, the discrepancy between reality and the mathematical model should be incorporated in the used DA method by introducing additional errors throughout the forecast cycle [13]. For sake of simplicity, the additional errors to reflect this discrepancy will not be introduced in the current study.

The main difficulty concerning the applicability of the second method of Lyapunov is the construction of the Lyapunov function  $V(x)$ , since there are no standardized methods available for constructing or selecting a suitable Lyapunov function for a given system. However, there are several methods available to generate a Lyapunov function for certain systems, although these methods are only applicable to a limited number of systems and all have their own advantages and disadvantages [54]. Some of these methods use an additional Lyapunov-based mathematical expression to assess stability, namely, the Lyapunov equation. A detailed explanation and derivation of the Lyapunov equation will be supplied in the next subsection.

### 5.3.2 Lyapunov equation

The Lyapunov equation differs from the earlier discussed Lyapunov function, but can be considered a fundamental part of several Lyapunov-based methods for proving stability of a dynamic system. To derive the Lyapunov equation, let us consider a linearized version of a non-linear autonomous system, given as:

$$\dot{x} = f(x) = \mathbf{A}x + g(x). \quad (5.13)$$

In this non-linear autonomous system, the linear part is denoted by  $\mathbf{A}\underline{x}$  and the non-linear part by  $\underline{g}(\underline{x})$  with the equilibrium point  $\underline{x}^-$  at the origin ( $\underline{x} = 0$ ). To evaluate the stability of this system in region  $\Omega$ , the following quadratic Lyapunov function candidate is postulated:

$$V(\underline{x}) = \underline{x}^T \mathbf{Z} \underline{x}. \quad (5.14)$$

In this expression, matrix  $\mathbf{Z}$  is symmetric and positive definite ( $\mathbf{Z} \succ 0$ ). When evaluating the stability criteria according to the second method of Lyapunov, it can be stated that the first two criteria (Equations 5.8 and 5.9) both hold due to the positive definiteness of the postulated quadratic Lyapunov function  $V(\underline{x})$ . To prove that the third and fourth criterion (Equations 5.10 and 5.11) also hold for the linearized non-linear autonomous system (Equation 5.13), the Lie derivative  $\dot{V}(\underline{x})$  need to be evaluated. Starting with the definition of the Lie derivative  $\dot{V}(\underline{x})$  (Equation 5.11) and subsequently substituting the definition of the considered system (Equation 5.13) and the gradient of  $V(\underline{x})$  ( $\nabla V(\underline{x})$ ), which results in:

$$\begin{aligned} \dot{V}(\underline{x}) &= \nabla V(\underline{x}) \underline{f}(\underline{x}), \\ &= \mathbf{Z} \underline{x} (\mathbf{A} \underline{x} + \underline{g}(\underline{x})), \\ &= \underline{x}^T \mathbf{Z}^T \mathbf{A} \underline{x} + \underline{x} \mathbf{Z}^T \underline{g}(\underline{x}) + \mathbf{Z} \underline{x} \underline{g}(\underline{x}). \end{aligned} \quad (5.15)$$

When normalizing all the  $\underline{x}$ -components in the above derived equation, with respect to the length of  $\underline{x}$  ( $\|\underline{x}\|_2$ ), this results in:

$$\begin{aligned} \dot{V}(\underline{x}) &= \underline{I}^T \mathbf{Z}^T \mathbf{A} \underline{I} + \mathbf{A}^T \underline{I}^T \mathbf{Z} \underline{I} + \mathbf{Z} \underline{I} \frac{\underline{g}(\underline{x})}{\|\underline{x}\|_2}, \\ &= \underline{I} (\mathbf{Z}^T \mathbf{A} + \mathbf{A}^T \mathbf{Z}) \underline{I}^T + \mathbf{Z} \underline{I} \frac{\underline{g}(\underline{x})}{\|\underline{x}\|_2}, \\ &= \mathbf{Z}^T \mathbf{A} + \mathbf{A}^T \mathbf{Z} + \mathbf{Z} \underline{I} \frac{\underline{g}(\underline{x})}{\|\underline{x}\|_2}. \end{aligned} \quad (5.16)$$

After normalization, the relative contribution of certain components in the expression can be evaluated with respect to each other. When assessing the scenario of  $\underline{x} \rightarrow \infty$ , it can be stated that the last term in the expression will approach zero. Therefore, it can be argued that the last term contributes less to  $\dot{V}(\underline{x})$  compared to the first term. According to the earlier stated stability criteria, the derivative of the Lyapunov function ( $\dot{V}(\underline{x})$ ) needs to be negative in order to satisfy Equations 5.10 and 5.11. Because the first term of Equation 5.16 has the largest relative contribution to  $\dot{V}(\underline{x})$ , it can be stated that this term has to be negative in order to satisfy Equations 5.10 and 5.11. Formally, this can be formulated as:

$$-\mathbf{W} = \mathbf{Z}^T \mathbf{A} + \mathbf{A}^T \mathbf{Z}. \quad (5.17)$$

Whereas, due to the fact that  $\mathbf{Z}$  is a symmetric matrix:

$$-\mathbf{W} = \mathbf{Z} \mathbf{A} + \mathbf{A}^T \mathbf{Z}. \quad (5.18)$$

This expression is known as the Lyapunov equation and can be used to evaluate the stability of a dynamic system. The equilibrium state can be considered as *locally asymptotically stable* if there exists a symmetric positive definite  $\mathbf{Z}$  for a given symmetric positive definite  $\mathbf{W}$ . However, it is important to note that this stability requirement does not hold vice versa. Therefore, the first step in proving stability by using the Lyapunov equation is choosing a symmetric positive definite  $\mathbf{W}$ , then solve the Lyapunov equation (Equation 5.18) for  $\mathbf{Z}$  and eventually check if  $\mathbf{Z}$  is positive definite. As mentioned before, this Lyapunov equation is often used in various Lyapunov-based stability approaches, but this equation does not appear in every approach.

### 5.3.3 Krasovskii's method

In the current study, the Krasovskii method will be used to determine the Lyapunov function  $V(\underline{x})$  and subsequently evaluate the stability of the considered dynamic system. Specifically, Krasovskii's method can be used to prove the stability of a non-linear autonomous dynamic system of the following form:

$$\dot{\underline{x}} = \underline{f}(\underline{x}). \quad (5.19)$$

Consequently, the Jacobian matrix of the the above defined dynamic system is denoted as:

$$\mathbf{J}_f(\underline{x}) = \frac{\partial \underline{f}(\underline{x})}{\partial \underline{x}}. \quad (5.20)$$

Following Krasovskii's method, the equilibrium point  $\underline{x}^-$  can be considered as *locally asymptotically stable* in region  $\Omega$  if  $\mathbf{F}(\underline{x})$  is negative definite ( $\mathbf{F}(\underline{x}) \prec 0$ ) in region  $\Omega$ . The definition of  $\mathbf{F}(\underline{x})$ , which is basically Krasovskii's equation, is given as:

$$\mathbf{F}(\underline{x}) = \mathbf{J}_f(\underline{x}) + \mathbf{J}_f(\underline{x})^T. \quad (5.21)$$

In Krasovskii method, a simple Lyapunov function is proposed for the considered non-linear autonomous system (Equation 5.19) of the following form;

$$V(\underline{x}) = \underline{f}(\underline{x})^T \underline{f}(\underline{x}). \quad (5.22)$$

To substantiate the proposed method of Krasovskii, let us first of all prove why  $\mathbf{F}(\underline{x}) \prec 0$  implies that  $\underline{f}(\underline{x}) \neq 0$  for any  $\underline{x} \neq 0$ . It can be argued that if square matrix  $\mathbf{F}(\underline{x})$  is negative definite for any  $\underline{x} \neq 0$ , then the Jacobian matrix  $\mathbf{J}_f(\underline{x})$ , of which  $\mathbf{F}(\underline{x})$  is composed, will be invertible. The Jacobian matrix  $\mathbf{J}_f(\underline{x})$  is considered to be invertible because if  $\mathbf{J}_f(\underline{x})$  and its transpose are both negative definite, then the sum of these two (Equation 5.21) will be negative definite as well. Due to this negative definiteness of  $\mathbf{J}_f(\underline{x})$ , it can be stated that matrix  $\mathbf{J}_f(\underline{x})$  contains no eigenvalues that are equal to zero. Consequently, the invertibility of matrix  $\mathbf{J}_f(\underline{x})$  can be guaranteed due to its squared dimensions and its non-zero eigenvalues. The invertibility of  $\mathbf{J}_f(\underline{x})$  combined with the continuity of  $\mathbf{J}_f(\underline{x})$  ensures that  $\underline{f}(\underline{x})$  is uniquely invertible. This implies that  $\underline{f}(\underline{x})$  has only one equilibrium point  $\underline{x}^-$  in the region  $\Omega$ , which guarantees that  $\underline{f}(\underline{x}) \neq 0$  for  $\underline{x} \neq 0$ . As a result, the Lyapunov function  $V(\underline{x})$  postulated by Krasovskii (Equation 5.22) is considered to be positive definite, which means that Equations 5.8 and 5.9 both hold.

After proving that  $\mathbf{F}(\underline{x}) \prec 0$  implies  $V(\underline{x}) \succ 0$ , let us prove that  $\mathbf{F}(\underline{x}) \prec 0$  also implies  $\dot{V}(\underline{x}) \prec 0$ , which explicitly means that Equations 5.8 - 5.11 hold if  $\mathbf{F}(\underline{x}) \prec 0$ . First, the derivative of  $V(\underline{x})$  can be written as:

$$\dot{V}(\underline{x}) = \underline{f}(\underline{x})^T \dot{\underline{f}}(\underline{x}) + \dot{\underline{f}}(\underline{x})^T \underline{f}(\underline{x}). \quad (5.23)$$

Then, the expression of  $\dot{\underline{f}}(\underline{x})$  can be rewritten when using the definition of the dynamic system (Equation 5.19) and the expression of the Jacobian matrix  $\mathbf{J}_f(\underline{x})$  (Equation 5.20), which will result in:

$$\dot{\underline{f}}(\underline{x}) = \frac{\partial \underline{f}(\underline{x})}{\partial \underline{x}} \frac{\partial \underline{x}}{\partial t} = \mathbf{J}_f(\underline{x}) \dot{\underline{x}} = \mathbf{J}_f(\underline{x}) \underline{f}(\underline{x}). \quad (5.24)$$

The above derived expression of  $\dot{\underline{f}}(\underline{x})$  (Equation 5.24) in combination with the Krasovskii's equation (Equation 5.21) can subsequently be used to reformulate Equation 5.23. This reformulation can be given as:

$$\begin{aligned} \dot{V}(\underline{x}) &= (\mathbf{J}_f(\underline{x}) \underline{f}(\underline{x}))^T \dot{\underline{f}}(\underline{x}) + \dot{\underline{f}}(\underline{x})^T \mathbf{J}_f(\underline{x}) \underline{f}(\underline{x}), \\ &= \mathbf{J}_f(\underline{x})^T \underline{f}(\underline{x})^T \dot{\underline{f}}(\underline{x}) + \dot{\underline{f}}(\underline{x})^T \mathbf{J}_f(\underline{x}) \underline{f}(\underline{x}), \\ &= \underline{f}(\underline{x})^T (\mathbf{J}_f(\underline{x}) + \mathbf{J}_f(\underline{x})) \underline{f}(\underline{x}), \\ &= \underline{f}(\underline{x})^T \mathbf{F}(\underline{x}) \underline{f}(\underline{x}). \end{aligned} \quad (5.25)$$

According to the derived equation (Equation 5.25), it can be stated that the  $\mathbf{F}(\underline{x}) \prec 0$  will result in  $\dot{V}(\underline{x}) \prec 0$ . Therefore, it can finally be concluded that Equations 5.8 - 5.11 hold if  $\mathbf{F}(\underline{x}) \prec 0$ , which implies that the equilibrium point  $\underline{x}^-$  can be considered as *locally asymptotically stable*. If additionally Equation 5.12 also hold for the considered dynamic system, then the equilibrium point  $\underline{x}^-$  can be considered as *globally asymptotically stable*.

However, Krasovskii's method is only applicable to a limited number of dynamic systems, because in many real world cases the sum of the Jacobian matrices does not satisfy the criteria of negative definiteness [67]. If this is the case for the system assessed in the current study, alternatives should be sought to classify the stability of the system. Besides, if the considered dynamic system contains high-order equations, it is hard to evaluate the negative definiteness of  $\mathbf{F}(\underline{x})$  for every possible value of  $\underline{x}$  [67]. However, this difficulty with high-order equation is unlikely to apply to the dynamic system considered in the current research because the used mathematical model to describe the cardiovascular system, elaborated in Chapter 3, contains only first-order differential equations. Consequently, Krasovskii's method can be considered appropriate to evaluate the stability of the mathematical model used in the current research.

## 6. Results

In this chapter, the results regarding the application of the proposed ROUKF-based algorithm to different types of measurement data will be presented. These different types of measurement data have been detailed in Section 3.2, whereas the proposed ROUKF-based algorithm was covered in Section 2.3 and the used mathematical model in Section 3.1. In the first section of this chapter, the results of the application of the algorithm to the synthetic data will be discussed. Based on these results, various analysis methods will be applied, more specific, the sensitivity, identifiability and stability analysis. Subsequently, in the second section of this chapter, the results of applying the algorithm to the in vitro data will be elaborated. Lastly, the results of the application of the algorithm to the in vivo data will be covered in the final section of the current chapter.

### 6.1 Synthetic data

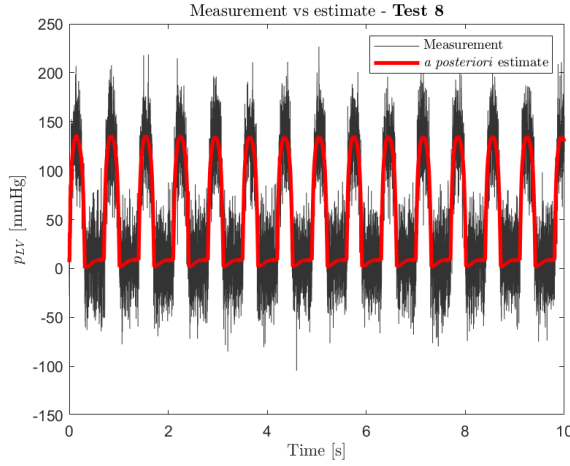
The synthetic data were generated based on the simulation of multiple test scenarios involving different values of the left ventricular contractility ( $c_{LV}$ ) and varying levels of measurement noise. Based on these simulated test scenarios, which have been detailed in Section 3.2.1, the capabilities of the proposed ROUKF-based algorithm were evaluated. Recapitulating, in tests 1 through 6 different values of  $c_{LV}$  were simulated to determine whether the algorithm can deal with different  $c_{LV}$  values. Subsequently, in tests 7 through 9, the simulated  $c_{LV}$  value was kept constant, while different levels of measurement noise were simulated to determine if the algorithm can handle noisy measurements. Lastly, in tests 10 through 12 the  $c_{LV}$  value was varied over the simulated time span to evaluate the robustness of the algorithm. When executing the proposed algorithm on all these mentioned test scenarios, the initialization values of the *a posteriori* estimate  $\hat{x}_0^a$  are kept identical for all tests. The exact initialization values of the *a posteriori* estimate  $\hat{x}_0^a$  are supplied in Appendix D, whereas it can be emphasised that the initialized  $c_{LV}$  value is equal to 0.8 for all tests.

The results of the application of the proposed ROUKF-based algorithm to the synthetic simulated test scenarios 1 through 9 are presented in Table 6.1. Based on these results, it can be stated that tests 1 through 6 depict a narrow confidence interval (CI  $c_{LV}$ ) combined with a low root-mean-square error (RMSE) and a low difference of  $c_{LV}$  (Diff  $c_{LV}$ ). Specifically, the difference of  $c_{LV}$  is calculated by subtracting the true  $c_{LV}$  from the estimated  $c_{LV}$ , which consequently indicates whether the algorithm under- or overestimates the  $c_{LV}$ . Then, the results of tests 7 through 9 depict that the deviation of the estimated  $c_{LV}$  from the true  $c_{LV}$  increases slightly when the noise level increases. In addition, the magnitude of the  $c_{LV}$  difference points in fluctuating direction because the behavior of the Gaussian distributed noise is purely stochastic in the simulated test scenarios.

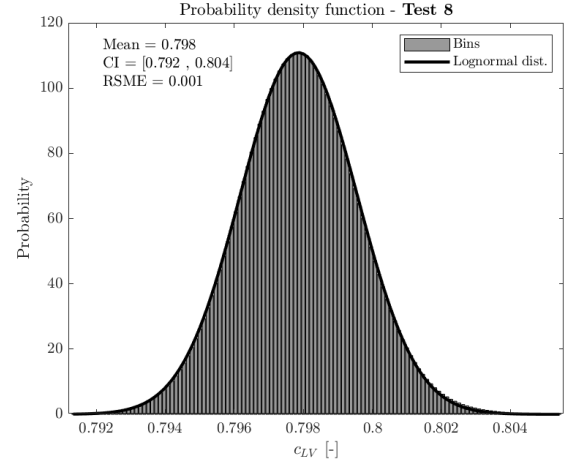
Test	True $c_{LV}$ [-]	Noise [%]	Estimated $c_{LV}$ [-]	CI $c_{LV}$ [-]	RMSE	Diff. $c_{LV}$ [-]
1	0.4	-	0.400	[0.390 , 0.410]	0	0
2	0.6	-	0.600	[0.594 , 0.606]	0	0
3	0.8	-	0.801	[0.795 , 0.807]	0	+0.001
4	1.0	-	1.001	[0.997 , 1.005]	0	+0.001
5	1.2	-	1.201	[1.197 , 1.205]	0	+0.001
6	1.4	-	1.401	[1.397 , 1.405]	0	+0.001
7	0.8	10	0.802	[0.796 , 0.808]	0.001	+0.002
8	0.8	20	0.798	[0.792 , 0.804]	0.001	-0.002
9	0.8	50	0.806	[0.796 , 0.816]	0.002	+0.006

**Table 6.1:** Results of the application of the proposed ROUKF-based algorithm to the synthetic data of tests 1 through 9. All tests included in this table are simulated for a duration of 10 seconds. The added noise is Gaussian distributed with a mean of zero and a test-specific standard deviation.

Figures 6.1 and 6.2 illustrate the performance of the proposed ROUKF-based algorithm concerning test 8, in which 20% noise was added to the synthetic measurement. In Figure 6.1, the measured  $p_{LV}$  is plotted against the *a posteriori* estimated  $p_{LV}$ , where it can be observed that the *a posteriori* estimated  $p_{LV}$  is not significantly influenced by the noise-containing  $p_{LV}$  measurement. The probability density function of the *a posteriori* estimated  $c_{LV}$  belonging to this particular test is presented in Figure 6.2, in which a log-normal distributed PDF can be observed.

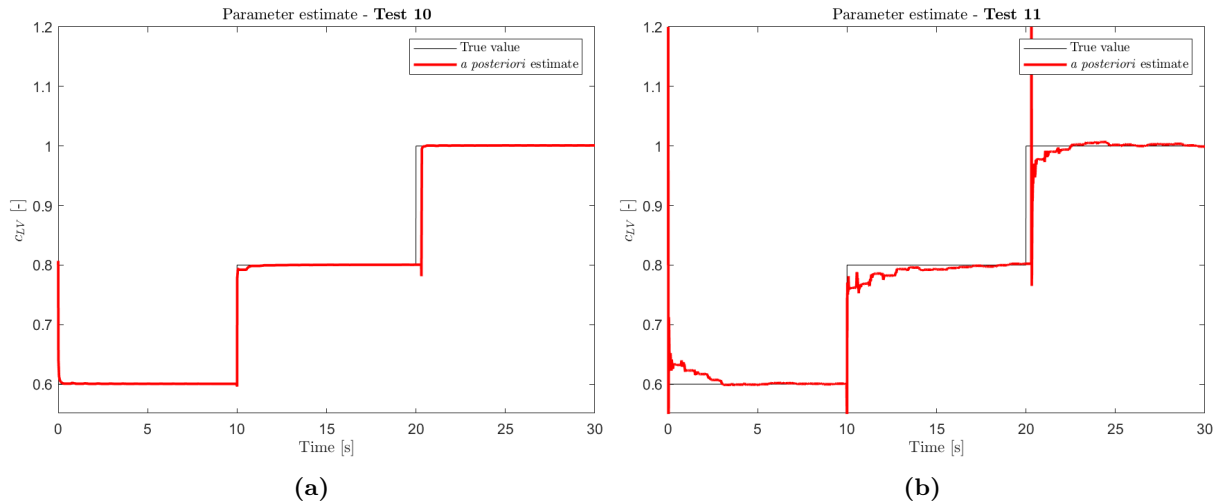


**Figure 6.1:** Comparison of the measured  $p_{LV}$  and the *a posteriori* estimated  $p_{LV}$ , both originating from test 8.



**Figure 6.2:** Illustration of the probability density function of the *a posteriori* estimated  $c_{LV}$  of test 8.

The test scenarios 10 through 12 were all simulated for a time interval of 30 seconds, which is equal to the duration of three moving windows, and therefore results in three  $c_{LV}$  estimations for each test. At the beginning of every moving window, certain elements of the proposed ROUKF-based algorithm are re-initialized, as explained in Section 4.2. This re-initialization can result in a so-called inlet effects, which can be regarded as a transient phenomenon with considerable fluctuations in the estimated  $c_{LV}$ . In tests 10 and 11, the true  $c_{LV}$  was varied incrementally between physiologically feasible ranges with a step size of 0.2 every 10 seconds. Specifically, the distinction between these two tests is that in test 10 no noise was added to the  $p_{LV}$  measurement, whereas in test 11 20% noise was added to the  $p_{LV}$  measurement. Figures 6.3a and 6.3b depict the *a posteriori* estimated  $c_{LV}$  of tests 10 and 11 over time, respectively. Based on these figures, it can be observed that the inlet effects is more prominent when measurement noise is present (test 11) compared to when measurement noise is absent (test 10). However, the presence of the inlet effects does not influence the estimated  $c_{LV}$ , due to the fact that only the last part of the moving window is used for PDF calculation.



**Figure 6.3:** Illustration of the *a posteriori* estimated  $c_{LV}$  plotted against the true  $c_{LV}$  concerning (a) test 10 and (b) test 11 over time.

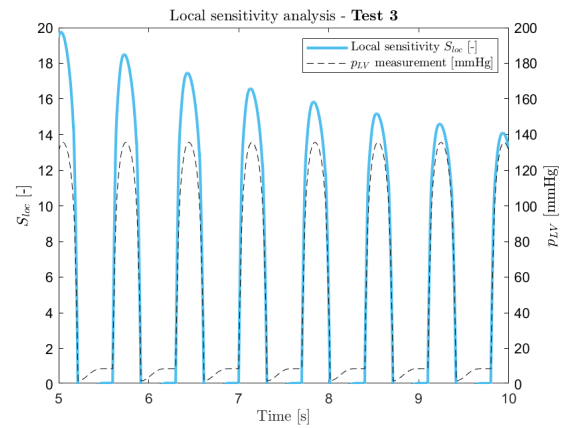
The results regarding the application of the proposed algorithm to the synthetic simulated test scenarios 10 through 12 are presented in Table 6.2. It can be stated that all three tests contain small values of RMSE and minor  $c_{LV}$  difference for each  $c_{LV}$ -step. Tests 10 and 11, whose estimated  $c_{LV}$  over time is shown in Figures 6.3a and 6.3b, simulate an uncommon but more or less physiologically feasible scenario. However, in test 12 a physiologically unrealistic scenario is simulated to evaluate whether the proposed algorithm can handle extreme  $c_{LV}$  changes. In all three tests, and even in the extreme scenario of test 12, the algorithm was able to converge to the true  $c_{LV}$  and follow the simulated changes in  $c_{LV}$ .

Test	True $c_{LV}$ [-]	Noise [%]	Estimated $c_{LV}$ [-]	CI $c_{LV}$ [-]	RMSE	Diff. $c_{LV}$ [-]
10	0.6	-	0.601	[0.595 , 0.607]	0	+0.001
	0.8	-	0.800	[0.794 , 0.806]	0	0
	1.0	-	1.001	[0.997 , 1.005]	0	+0.001
11	0.6	20	0.602	[0.592 , 0.612]	0.002	+0.002
	0.8	20	0.796	[0.790 , 0.802]	0.001	-0.004
	1.0	20	1.002	[0.994 , 1.010]	0.002	+0.002
12	1.4	20	1.403	[1.399 , 1.407]	0.002	+0.003
	0.4	20	0.398	[0.374 , 0.422]	0.004	-0.002
	1.2	20	1.205	[1.199 , 1.211]	0.003	+0.005

**Table 6.2:** Results of the application of the proposed ROUKF-based algorithm to the synthetic data of test 10 to 12. All tests included in this table are simulated for a duration of 30 seconds. The added noise is Gaussian distributed with a mean of zero and a test-specific standard deviation.

### 6.1.1 Sensitivity analysis

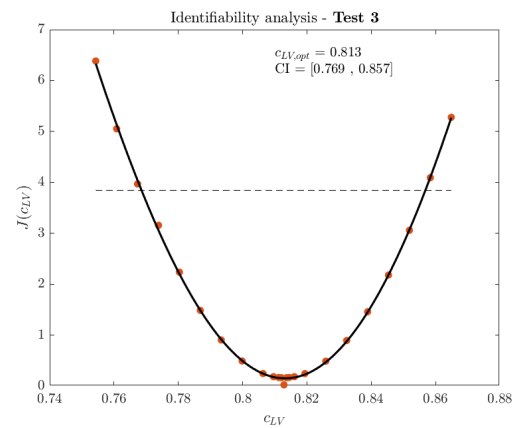
The local sensitivity analysis is used to identify the sensitivity of the measured states to changes in the parameter of interest. Consequently, the relative importance of the uncertain input parameter ( $p_{LV}$ ) regarding the estimation of the parameter of interest ( $c_{LV}$ ) within a heart cycle can be quantified. In Figure 6.4, the synthetic  $p_{LV}$  data is plotted against the local sensitivity  $S_{loc}$ . Based on these results, it can be stated that the local sensitivity  $S_{loc}$  has exactly the same shape as the measured  $p_{LV}$ . Lastly, it can be observed that local sensitivity  $S_{loc}$  decreases when the relative time in the moving window increases.



**Figure 6.4:** Visual representation of the local sensitivity  $S_{loc}$  plotted against the measured  $p_{LV}$  based on the synthetic data of test 3.

### 6.1.2 Identifiability analysis

The identifiability analysis is used to determine if the right type of data is collected, given a certain system and model structure, to identify a unique parameter value based on the measured data. This analysis was conducted by using the profile likelihood method, which is based on the evaluation of a cost function that represents the agreement between the measured  $p_{LV}$  and the *a posteriori* estimated  $p_{LV}$ . In Figure 6.5, several evaluations of the cost function are depicted by orange dots, which are subsequently used to determine the actual shape of the cost function. According to the profile likelihood method, the optimal  $c_{LV}$  ( $c_{LV,opt}$ ) is located at the minimum of the cost function. The boundaries of the CI belonging to this  $c_{LV,opt}$  are represented by the intersection points of the cost function and threshold, with the threshold being illustrated by the dotted line in Figure 6.5.



**Figure 6.5:** Illustration of the identifiability analysis based on the synthetic data of test 3.



The results of the identifiability analysis are presented in Table 6.3. Based on these results, it can be said that the CI is bounded in both directions for tests 1 through 8, while it is unbounded in one direction of test 9. Moreover, the confidence interval obtained by using the PL method (Table 6.3) is slightly wider compared to the confidence interval determined based on the PDF (Table 6.1) for tests 1 through 6 and significantly wider for tests 7 through 9. Besides, it can be observed that the  $c_{LV,opt}$  of the (Table 6.3) also differs from the estimated  $c_{LV}$  (Table 6.1) when comparing the synthetic simulated test scenarios.

Test	True $c_{LV}$ [-]	Noise [%]	$c_{LV,opt}$ [-]	CI $c_{LV}$ [-]	Diff. $c_{LV}$ [-]
1	0.4	-	0.406	[0.370 , 0.442]	+0.006
2	0.6	-	0.608	[0.567 , 0.649]	+0.008
3	0.8	-	0.813	[0.769 , 0.857]	+0.013
4	1.0	-	1.019	[0.973 , 1.066]	+0.019
5	1.2	-	1.227	[1.185 , 1.275]	+0.027
6	1.4	-	1.450	[1.417 , 1.485]	+0.050
7	0.8	10	0.733	[0.405 , 1.321]	-0.067
8	0.8	20	0.670	[0.280 , 2.130]	-0.130
9	0.8	50	0.483	[0.113 , -]	+0.317

**Table 6.3:** Results of the identifiability analysis conducted on the synthetic data originating from tests 1 through 9. The identifiability analysis is executed based on the profile likelihood method. The optimal  $c_{LV}$  according to this method is denoted by  $c_{LV,opt}$  with the belonging confidence interval (CI  $c_{LV}$ ).

### 6.1.3 Stability analysis

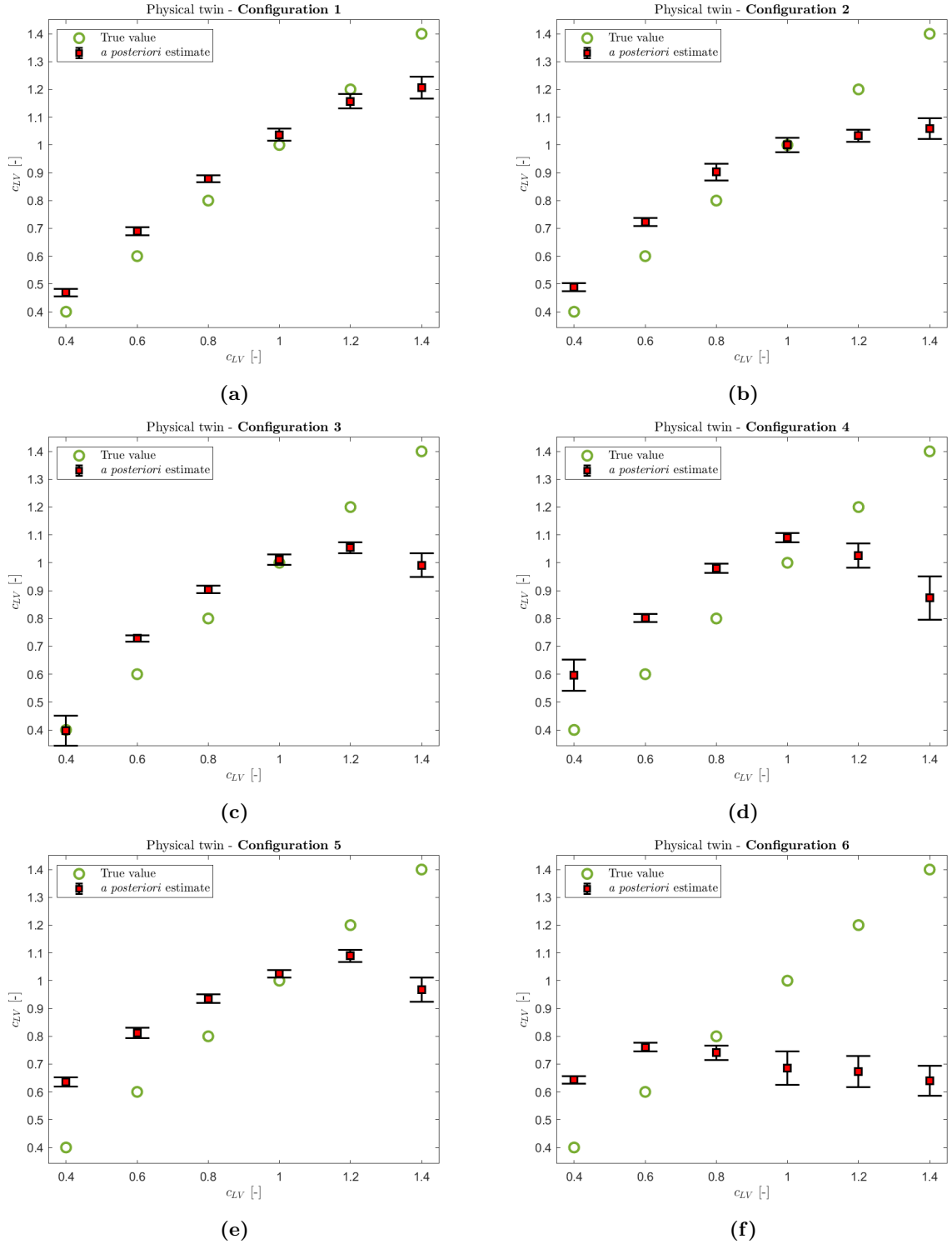
The stability analysis is used to quantify the sensitivity of the mathematical model to its initial states, whereas, the stability of the model is also considered to be crucial for the convergence of the sequential estimates of the proposed ROUKF-based algorithm. This analysis was conducted by using Krasovskii's method, which claims that the equilibrium point  $\bar{x}$  can be considered as *locally asymptotically stable* if Krasovskii's equation  $\mathbf{F}(\bar{x})$  is negative definite ( $\mathbf{F}(\bar{x}) \prec 0$ ) in a certain region  $\Omega$  of the parameter space. To determine the region  $\Omega$  in the parameter space where this holds, the negative definiteness of Krasovskii's equation  $\mathbf{F}(\bar{x})$  was evaluated for a range of  $c_{LV}$  values. The results of this evaluation are illustrated in Table 6.4, in which the equilibrium point  $\bar{x}$  of the mathematical model is considered to be stable if Krasovskii's equation  $\mathbf{F}(\bar{x})$  was negative definite and unstable if Krasovskii's equation  $\mathbf{F}(\bar{x})$  was positive definite. Based on the results depicted in Table 6.4, it can be stated the equilibrium point  $\bar{x}$  of the mathematical model can be considered as *local asymptotically stable* in the region  $\Omega$  that ranges from a  $c_{LV}=0.1$  to  $c_{LV}=2.0$ .

$c_{LV}$ [-]	0	0.1	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0	2.1	2.2
Stable?	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No

**Table 6.4:** Results of the stability analysis carried out based on Krasovskii's method. In this table, a stable system is considered as a negative definite  $\mathbf{F}(x)$ , while an unstable system is defined as a positive definite  $\mathbf{F}(x)$ .

## 6.2 In vitro data

The in vitro data is generated based on the material twin, which is a device that is capable of mimicking the behavior of the human cardiovascular system. When generating the in vitro data, this material twin was used in different configurations of increasing complexity to evaluate the effect of model discrepancy on the ability of the proposed ROUKF-based algorithm to estimate the true  $c_{LV}$ . The different configurations that were used in current study have been elaborated in Section 3.2.2, where for each configuration the same variety of  $c_{LV}$  values was being simulated. The exact results of the application of the proposed algorithm to all configurations included in the current study are provided in Appendix F. A summarized representation of these exact results is supplied per configuration in Figure 6.6, in which the true  $c_{LV}$  value is being plotted against the *a posteriori* estimated  $c_{LV}$  with an error-bar that indicates the corresponding confidence interval. In this summarized representation, the *a posteriori* estimated  $c_{LV}$  and confidence interval are calculated by averaging the statistics of three time windows.



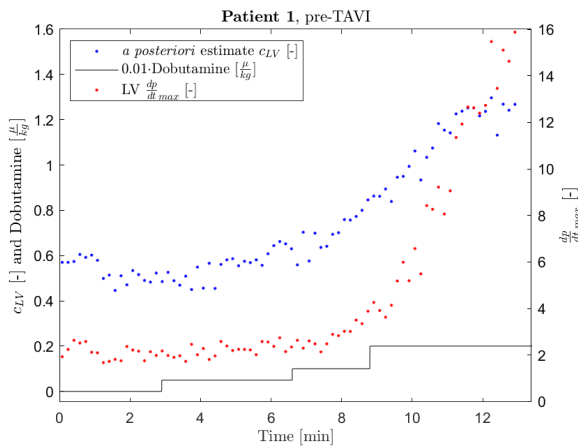
**Figure 6.6:** Results of the application of the proposed ROUKF-based algorithm to the following material twin configurations: (a) configuration 1, (b) configuration 2, (c) configuration 3, (d) configuration 4, (e) configuration 5, (f) configuration 6. The  $a posteriori$  estimated  $c_{LV}$ , visually expressed in the figures above, is obtained by averaging the three estimates that belong to a specific simulated  $c_{LV}$  value. The same holds for the standard deviation, which is used to determine the confidence interval (CI) of the  $a posteriori$  estimated  $c_{LV}$ .

When observing the results displayed in Figure 6.6, it can be stated that for Configurations 1 through 5 the proposed algorithm is overestimating for  $c_{LV}$  values between 0.4 and 1, whereas the algorithm is underestimating for  $c_{LV}$  values of 1.2 and 1.4. Moreover, it can roughly be stated that the severity of under- or overestimation increases when the complexity of the material twin configuration increases. The only configuration that illustrates a different pattern compared to the other configurations is the most advanced configuration, which is Configuration number 6. When examining the error-bars belonging to the *a posteriori* estimates, it can be said that the CI is quite small. Finally, the size of the CI does not correlate with the difference between the *a posteriori* estimated  $c_{LV}$  and the true  $c_{LV}$ .

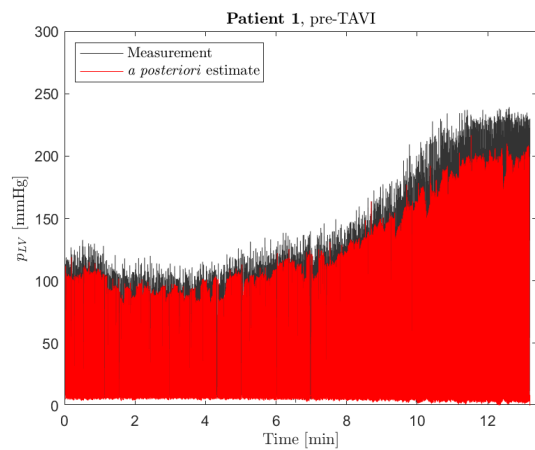
### 6.3 In vivo data

In the current study, the proposed ROUKF-based algorithm was applied to the in vivo data originating from patient-specific measurements performed in the study of Johnson et al. (2018) [29]. However, the data set used in the current study does not exactly correspond to the data set used in the original study. There are roughly two reasons that explain the absence of certain patient-specific analysis in the current study. The first reason is the absence of  $p_{LV}$  measurements in certain patients, the second reason is a significant deviation of the  $p_{LV}$  measurement compared to a normal physiological  $p_{LV}$  curve. This significantly deviated  $p_{LV}$  measurement was considered as non-representative and was therefore not included in the data set used in the current study.

The results of the application of the proposed ROUKF-based algorithm to the data set with patient-specific measurements are supplied in Appendix G. In this appendix, two types of graphs are presented for every patient-specific measurement. One of these types is depicted in Figure 6.7, in which the *a posteriori* estimated  $c_{LV}$ , the LV  $\frac{dp}{dt}_{max}$  and the step-wise Dobutamine administration are illustrated over time. This type of graph is used to evaluate whether the proposed ROUKF-based algorithm is able to reflect the suspected increase in  $c_{LV}$  due to the step-wise administration of Dobutamine. In addition, this type of graph is also used to identify a possible relationship between the *a posteriori* estimated  $c_{LV}$  and the LV  $\frac{dp}{dt}_{max}$ . The other type of graph that is supplied in Appendix F is illustrated in Figure 6.8, in which the measured  $p_{LV}$  is plotted against the *a posteriori* estimated  $p_{LV}$  over time. This type of graph is used to determine if the *a posteriori* estimated  $p_{LV}$  follows the behaviour of the measured  $p_{LV}$ . When comparing the results depicted in Figures 6.7 and 6.8, which both belong to patient 1 before the TAVI procedure, it can be observed that the  $c_{LV}$  surprisingly follows the same trend as the maximum  $p_{LV}$ .

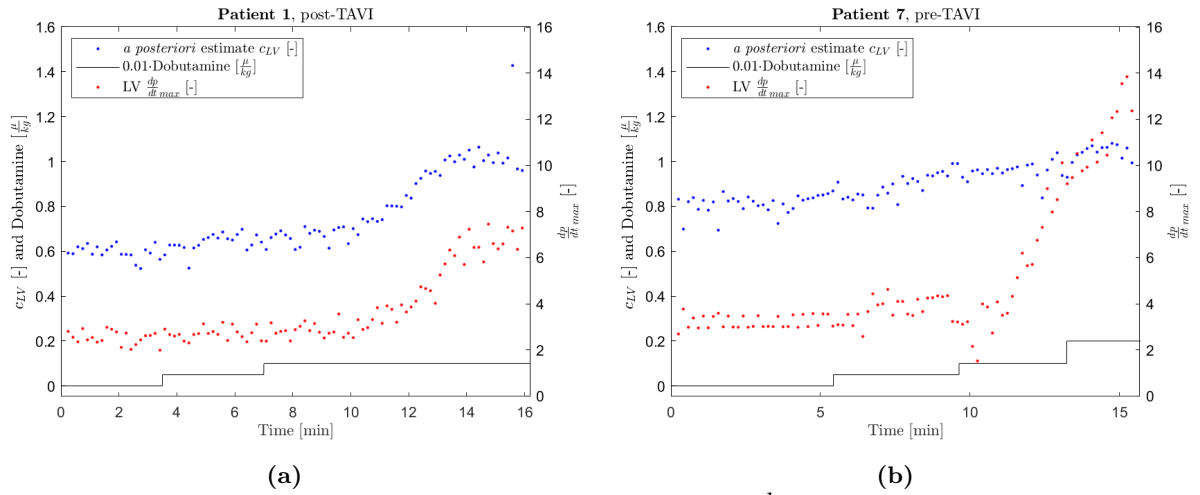


**Figure 6.7:** Illustration of the *a posteriori* estimated  $c_{LV}$ , the LV  $\frac{dp}{dt}_{max}$  and the Dobutamine administration of **patient 1**, obtained *before* conducting the TAVI procedure.



**Figure 6.8:** Comparison of the measured  $p_{LV}$  and the *a posteriori* estimated  $p_{LV}$  of **patient 1**, obtained *before* conducting the TAVI procedure.

When evaluating the results of the application of the proposed ROUKF-based algorithm to the data set with patient-specific measurements, which are presented in Appendix G, a few potentially valuable aspects can be noted. First, the expected increase in  $c_{LV}$  caused by the step-wise administration of Dobutamine could not be identified for every patient in the data set. Second, when comparing the *a posteriori* estimated  $c_{LV}$  to the  $LV \frac{dp}{dt}_{max}$ , two more or less distinguishable trends could be identified across the data set, which are depicted in Figure 6.9a and 6.9b. In Figure 6.9a, the *a posteriori* estimated  $c_{LV}$  and the  $LV \frac{dp}{dt}_{max}$  illustrate the same behavior over time. On the other hand, in Figure 6.9b, the *a posteriori* estimated  $c_{LV}$  and the  $LV \frac{dp}{dt}_{max}$  depict the same behavior in the first half of the observed time interval. However, after a certain amount of Dobutamine has been administered to the patient, the  $LV \frac{dp}{dt}_{max}$  suddenly increases significantly while the *a posteriori* estimated  $c_{LV}$  remains more or less the same. Lastly, based on the comparison between the measured  $p_{LV}$  and the *a posteriori* estimated  $p_{LV}$ , it can be stated that the proposed ROUKF-based algorithm is able to follow the  $p_{LV}$  behavior of the measured data. However, the *a posteriori* estimated  $p_{LV}$  is consistently underestimating the measured  $p_{LV}$ , with this underestimation ranging between approximately 10 to 60 [mmHg]. Furthermore, it can be observed that the level of underestimation depends on the height of the measured  $p_{LV}$ , which implies that the underestimation is more significant for increasing values of  $p_{LV}$ .



**Figure 6.9:** Illustrations of the *a posteriori* estimated  $c_{LV}$ , the  $LV \frac{dp}{dt}_{max}$  and the step-wise administered Dobutamine in (a) **patient 1** after conducting the TAVI procedure and (b) **patient 7** before conducting the TAVI procedure.

## 7. Discussion

To develop a standardized weaning protocol, a method is sought that can accurately and in real-time predict if the heart is capable of independently providing the required perfusion, ideally based on minimal to non-invasive measurements. Therefore, the aim of the current study is to supply the basis for such a method by providing a proof of principle for utilizing ROUKF to establish a digital shadow, which is capable of estimating the left ventricular contractility ( $c_{LV}$ ). The proposed ROUKF-based algorithm uses a 0D lumped element model of the left ventricular to obtain simulations of the behavior of the cardiovascular system, which are subsequently merged with left ventricular pressure ( $p_{LV}$ ) measurements originating from: synthetic, in vitro or in vivo data. In the current section, the major findings, the limitations and the potential future work will be elaborated regarding the current study.

### 7.1 Major findings

In the current study, a number of different evaluations and analyses were conducted to examine the capability and feasibility of the proposed ROUKF-based algorithm regarding the estimation of the  $c_{LV}$ . Starting with the results of the synthetic data, which have been supplied in Section 6.1. First of all, the ability of the proposed algorithm to converge to the true  $c_{LV}$  is considered to be independent of the initialized  $c_{LV}$ . This can be stated on the basis of the ability of the proposed algorithm to converge to the true  $c_{LV}$  in all synthetic tests, in which the initialized  $c_{LV}$  was identical. As a result, it can be argued that the performance of the algorithm is independent of the initialized  $c_{LV}$  conditions. Secondly, the proposed algorithm is considered to be capable of handling noisy measurements, since no significant deviations in the estimated  $c_{LV}$  were encountered when noise was added to the  $p_{LV}$  measurement. Therefore, it can be said that the proposed algorithm seems to be insensitive to noisy measurements, which strengthens its potential applicability in clinical practice. Thirdly, the proposed algorithm can be considered as robust, since it was able to converge to the true  $c_{LV}$  in case the  $c_{LV}$  changed during the simulated time. Even in the case of physiologically infeasible and extreme  $c_{LV}$  steps, the proposed algorithm was able to converge to the true  $c_{LV}$  regardless of the presence of noise. Following a changing  $c_{LV}$  over time is essential to become a digital shadow of the patient, because the  $c_{LV}$  of the patient will change over time. Therefore, this finding also proves the potential of the algorithm to act as a potentially valuable tool in the clinical practice.

Subsequently, the results of the various analysis methods, which have been performed on the basis of the synthetic data (Sections 6.1.1 through 6.1.3), will be discussed first. The results of the local sensitivity analysis (Section 6.1.1) illustrate that the  $c_{LV}$  is only sensitive for the  $p_{LV}$  measurement during the contraction phase of the cardiac cycle. The results of the identifiability analysis (Section 6.1.2) showed a bounded CI in both directions for almost all synthetic simulations (tests 1 through 8), with exception of one synthetic simulation (test 9). This implies that, given the current model structure, the  $c_{LV}$  is identifiable from a  $p_{LV}$  measurement that contains no noise or a limited level of noise, whereas the  $c_{LV}$  is practically non-identifiable if the level of noise in the measurement becomes too severe. These identifiability observations are consistent with the theoretical background, because the measurement error gives rise to parameter uncertainty and can therefore cause practical non-identifiability [26]. Furthermore, the estimated  $c_{LV}$  and associated CI determined by using the identifiability analysis differ significantly from those determined on the basis of the PDF. This significant difference could be attributed to the divergent approaches of these two methods, since the identifiability analysis uses a cost function that describes the discrepancy between measured and predicted  $p_{LV}$ , while the PDF relies on the sequentially estimated  $c_{LV}$ . Lastly, according to the results of the stability analysis (Section 6.1.3), the equilibrium point of the mathematical model can be considered as *locally asymptotically stable* in the region of the parameter space that ranges from  $c_{LV}=0.1$  to  $c_{LV}=2.0$ . As a result, it can be concluded that the mathematical model used in the current study is stable in a physiologically plausible range of the  $c_{LV}$ . This range is considered physiologically plausible because the  $c_{LV}$  is a scaling factor for the active stress in the myocardial fibre which cannot be lower than approximately 0.2/0.3 to be viable and is only higher than approximately 1.5/1.6 in well trained athletes or under inotropic support conditions. As a result, it can be argued that the mathematical model is stable in the  $c_{LV}$  range that can occur in patients, which proves that the proposed algorithm can be applied in clinical practice.

The in vitro data generated by the material twin was used to evaluate the effect of model discrepancy on the ability of the proposed ROUKF-based algorithm to estimate the true  $c_{LV}$  (Section 6.2). First of all, when the cardiovascular behavior of the material twin was identical to the one described by the mathematical model of the current study, even then the proposed algorithm was not able to estimate the true  $c_{LV}$ . In addition, it can be stated that deviation from the true  $c_{LV}$  increased when the complexity of the material twin configuration increased. These two observations illustrate that the proposed ROUKF-based algorithm is not capable of estimating the true  $c_{LV}$  in a real world setting, based on the currently used algorithm architecture. Furthermore, the CI belonging to the estimated  $c_{LV}$  is considered to be relatively narrow in all configurations and does not correlate with the discrepancy between the estimated  $c_{LV}$  and true  $c_{LV}$ . This reflects that the proposed algorithm can be considered as over-confident, as well as that the uncertainty of the estimated  $c_{LV}$  does not incorporate the model discrepancy.

Finally, the results of the application of the proposed algorithm to the in vivo data, which has been elaborated in Section 6.3. First, it is considered to be promising that the proposed algorithm is able to deal with  $p_{LV}$  measurements that were performed in patients. This is considered to be promising in terms of the robustness of the proposed algorithm, because patient measurements contain several imperfections compared to synthetic data. These imperfections could be attributed to the measurement equipment, the environment of the measurement, the patient itself or the clinician that performs the measurement. Therefore, the fact that the proposed algorithm is applicable to patient data strengthens its possible applicability in clinical practice. Besides, the patient itself consists of an even more complex cardiovascular behavior than described by the most advanced configuration of the material twin. Therefore, it can be considered as promising that the proposed algorithm seems to be able to deal with such a level of model discrepancy. However, the exact magnitude of the discrepancy between the estimated  $c_{LV}$  and true  $c_{LV}$  is unknown, since the true  $c_{LV}$  is not known. Furthermore, it can be argued that the  $c_{LV}$  did not increase with the step-wise administration of Dobutamine, which is considered to be inconsistent with the predefined expectation. Besides, it can be observed that the LV  $\frac{dp}{dt}_{max}$  does not present the same information as the  $c_{LV}$  in some patients, which can be said based on the identification of two distinct trends in the data set. Both of these observations could be regarded as possible valuable information concerning certain specific patients, or either as a deficiency of the proposed algorithm to describe the true  $c_{LV}$ . Furthermore, the fact that the estimated  $c_{LV}$  did follow the same trend as the  $p_{LV}(max)$  can be considered as a surprising observation, which can be regarded as unfavorable for the potential of the algorithm because it is not able to supply additional information. Lastly, based on the comparison of the measured and estimated  $p_{LV}$  it can be stated that the proposed algorithm is not able to follow the exact behavior of the measurement. Besides, it can be stated that if the magnitude of the  $p_{LV}$  measurement increases, the amount of discrepancy increases.

## 7.2 Limitations

A promising finding in the current study is the fact that the proposed ROUKF-based algorithm is able to converge to the true  $c_{LV}$  when being applied to various simulated scenarios of the synthetic data. However, the proposed algorithm was not able to estimate the true  $c_{LV}$  in case of application to the material twin, which could be attributed to different aspects. First of all, the  $p_{LV}$  measurement signal of the material twin used as input for the proposed algorithm was not filtered. As a result, the signal did contain a significant level of noise, which could possibly affect the accuracy of the estimated  $c_{LV}$  or the identifiability of the  $c_{LV}$ . Second, the complexity of the mathematical model that is used to describe the behavior of the cardiovascular system is most likely attributable for the deviating  $c_{LV}$  estimation. This can be said based on the evaluation of the model discrepancy and it is therefore suggested that increasing the complexity of the mathematical model will result in a more truthful representation of the  $c_{LV}$ . Third, the number of measured quantities could be regarded as too small to compose a truthful estimation of the  $c_{LV}$  in the real world. An increase in the number of measured quantities leads to an increase in the amount of patient-specific characteristics included in the proposed algorithm, which could result in a more truthful representation of the  $c_{LV}$ . Another aspect that could be accountable for the failure of the proposed algorithm to estimate the true  $c_{LV}$  are the input parameters of the mathematical model. These input parameters are currently fixed to certain parameter values, which mostly originate from the study of Meiburg et al. (2020) [49]. However, some of these fixed input parameters could contain valuable information regarding the  $c_{LV}$ . Therefore, the input parameters that are most influential for  $c_{LV}$  estimation need to be identified first. Subsequently, these input parameters could be specified per patient to obtain a more truthful estimation of the  $c_{LV}$  in a real world scenario.

A important consolation is that the proposed ROUKF-based algorithm is considered to be over-confident and does not incorporate the model discrepancy into the CI of the estimated  $c_{LV}$ . Consequently, it can be stated that the incorporation of the model discrepancy in the proposed algorithm will lead to a more realistic representation of the uncertainty in the estimated  $c_{LV}$ . Furthermore, the fact that the algorithm is over-confident could be attributed to the rank reduction of the covariance matrix. In the current algorithm, the rank of the covariance matrix is equal to the dimension of the parameter space, which can be considered as insufficient to truthfully describe the uncertainty of the estimated  $c_{LV}$ . Increasing the rank of the covariance matrix to allow a certain amount of uncertainty on the states of the augmented space could result in a more realistic representation of the CI of the estimated  $c_{LV}$ . However, it is considered to be essential for the ROUKF method to retain a reduced rank covariance matrix, because it allows a rank decomposition of this matrix, which is fundamental for the reduced-order approach. Eventually, the computational expensiveness of the proposed ROUKF-based algorithm will increase when increasing the rank of the covariance matrix, which could influence the real-time capabilities of the algorithm.

In the proposed ROUKF-based algorithm, the measurement error covariance matrix  $\mathbf{R}$  is arbitrarily chosen. However, this covariance matrix depends on the trueness and precision of the used measurement equipment [69]. Ideally, this covariance matrix should be adapted to measurement equipment which is used to perform the  $p_{LV}$  measurement. An accurate description of the noise statistics is considered to be crucial for the performance of the algorithm, as it is known that an incorrect description can result in significant worsening of the estimation or even failure of the algorithm [20]. Furthermore, when these noise statistics are known, they can be used for calculating the cost function of the identifiability analysis. Based on applying this analysis in a clinical setting, the identifiability of the parameter of interest can be evaluated, given a certain level of noise. Another arbitrarily choice in the proposed ROUKF-based algorithm is that the re-initialization conditions are equal to the initialization conditions, which are chosen to concentrate the initial covariance matrix to the parameters only. The rank reduction of the covariance is considered to be insufficient to truthfully estimate the parameter, which may also call into question the correctness of the initialization and re-initialization choices. Lastly, the initialized *a posteriori* estimate was also determined arbitrarily (Appendix D), but only the influence of the initialized  $c_{LV}$  regarding the convergence of the proposed algorithm to the true  $c_{LV}$  is investigated. The effect of the initialization of the states on the convergence of the proposed algorithm is unknown.

The PDF of the  $c_{LV}$  is assumed to be lognormal distributed, which can be doubted in some cases. For example, consider a  $p_{LV}$  measurement of the in vivo data that is used in the current study, which can constantly change in amplitude and period length, while containing various imperfections. As a result, the estimated parameter  $c_{LV}$  will fluctuate over the time of the analyzed interval, which may result in various forms of distributions and consequently in a violation of the lognormal distributed assumption.

### 7.3 Future work

Increasing the number of measured quantities used in the proposed algorithm increases the amount of included patient-specific characteristics, which could lead to a more veracious  $c_{LV}$  estimation. The study of Bunt (2021) could serve as a starting point to identify which additional measured quantities could be added to the proposed algorithm. In this particular study, an UKF-based algorithm was used to estimate the  $c_{LV}$  based on five measured quantities, while using the same mathematical model as the current study [11]. However, future research needs to be conducted to identify which additional measurements could be valuable in terms of convergence of the proposed ROUKF-based algorithm to the true  $c_{LV}$  in a real-world environment. Furthermore, a global sensitivity analysis can be used to determine the relative contribution of uncertain model input parameters and their interactions to the amount of uncertainty in the model output. Based on this sensitivity analysis, the input parameters that contributes the most to the uncertainty in the model output can be identified, which can be used to conduct parameter prioritization. In future research, this global sensitivity analysis could be applied based on the adaptive generalized polynomial chaos expansion (agPCE), which is elaborated in Appendix B.

Based on the results of the application to the material twin, it can be observed that the proposed algorithm is incapable of following the behavior of the  $c_{LV}$  when the right side of the heart was included in the material twin configuration. As a result, it can be argued that describing the behavior of the right side of the heart and the pulmonary circulation by a fixed pulmonary pressure is considered to be inadequate. Therefore, it can be suggested that substituting the fixed pulmonary pressure by a more advanced description of the right side of the heart and pulmonary circulation could be a promising option to increase the complexity of the mathematical model. However, various other modifications of the mathematical model could be valuable options as well. Future research needs to be conducted to determine which modifications of the mathematical model allow the proposed algorithm to provide a more truthful estimation.

In the current study, it is aimed to provide a proof of principle of using the ROUKF method to establish a digital shadow, which could be used to support the initiation of weaning and the weaning process itself. Consequently, the ECMO circuit needs to be incorporated in the mathematical model to describe the cardiovascular behavior of a patient that is undergoing VA-ECMO. Future research needs to be carried out to identify how to accomplish this incorporation, whereas this incorporation could also supply possible beneficial alternatives for the invasiveness of the proposed method. Concretely, the pressure sensors in the ECMO device could be used to calculate the flow of the ECMO, which is detailed in the study of Pennings et al. (2013) [58]. Subsequently, the flow of the ECMO could be used as input in the proposed algorithm to estimate the  $c_{LV}$ .

The model discrepancy is considered to be an important component for the uncertainty in the model output. Therefore, including the model discrepancy in the determination of CI of the estimated parameter, could result in a more truthful representation of the uncertainty of the estimate. Further research needs to be conducted on how the model discrepancy can be incorporated in the proposed algorithm. For this incorporation, the study of Brynjarsdóttir and O'Hagen (2014) can be used, which elaborates on the importance of discrepancy and challenges concerning its incorporation [10]. Besides, a more representative CI could also be obtained by increasing the rank of the covariance matrix and consequently allow a certain amount of uncertainty on the states. However, this rank should not be increased too much to limit the increase in computational expensiveness, whereas the rank should be increased sufficiently to allow for an appropriate amount of uncertainty on the states that is required to accurately describe the uncertainty of the estimated  $c_{LV}$ . Additional research needs to be conducted to identify the optimal ratio between these two aspects.

Lastly, a method is desired that is able to describe every possible distribution of the PDF. This method could lead to a more truthful determination of the estimated  $c_{LV}$  and belonging CI in clinical practice. A method that can be used to approximate every possible distribution is called Gaussian Mixture Models. This method approximates the shape of the PDF by the weighting sum of multiple Gaussian distributed functions, to eventually represent the actual shape of PDF [63]. However, future research needs to be conducted to evaluate the applicability to the current ECMO application. Furthermore, a method that is capable of estimating the measurement error covariance matrix could increase the performance of the proposed ROUKF-based algorithm, because proper knowledge of the noise statistics is required for optimal performance of an estimator [40]. As mentioned before, the covariance matrix incorporates various stochastic components of the considered system, like instrumental error and the error in the measurement function  $h(x)$  [20]. However, these stochastic components can vary over time, and therefore an estimator is required that can describe the evolution of this noise covariance matrix over time. In order to find a suitable estimator for the proposed ROUKF-based algorithm, the study of Dunik et al. (2017) can be used [20]. This study elaborates various feedback and feedback free methods, which are based on different assumptions and are therefore suitable being applicable to selective problems. To identify the most suitable method for the envisioned application, future research needs to be conducted.



## 8. Conclusion

The goal of the current study was to develop a digital shadow using ROUKF, which is capable of accurately estimating left ventricular contractility ( $c_{LV}$ ) in real-time based on minimal to non-invasive measurements. The measurement that was used to conduct the  $c_{LV}$  estimation is the left ventricular pressure, which originated from various different data types: synthetic, in vitro and in vivo data. The proposed ROUKF-based algorithm was applicable to all three mentioned data types and was able to perform  $c_{LV}$  estimations in real-time. In addition, the algorithm presented some promising results when being applied to synthetic and in vivo data. As a result, the algorithm can be regarded as a promising method for the development of a digital shadow of the  $c_{LV}$ . However, the evaluation of model discrepancy proved that the accuracy of the proposed algorithm decreased when the model discrepancy increased. Therefore, future research needs to be conducted to identify the right increase in complexity of the mathematical model to obtain a more truthful  $c_{LV}$  estimation. Furthermore, additional research also needs to be performed to evaluate the performance of the proposed algorithm when adding or substituting input measurements. Eventually, the holy grail remains a method that can perform an accurate estimation of the  $c_{LV}$  in real-time based on only non-invasive measurements to be the ideal tool to conduct weaning from VA-ECMO.

# Bibliography

- [1] Nadia Aissaoui, Aly El-Banayosy, and Alain Combes. How to wean a patient from veno-arterial extracorporeal membrane oxygenation. *Intensive Care Medicine*, 41(5):902–905, 2015. 2
- [2] Nadia Aissaoui, Pascal Leprince, and Philippe Le. Predictors of successful extracorporeal membrane oxygenation ( ECMO ) weaning after assistance for refractory cardiogenic shock. *Intensive Care Med*, 37:1738–1745, 2011. 2
- [3] M. Asch, M. Bocquet, and M. Nodet. Chapter 1: Introduction to data assimilation and inverse problems. In *Data assimilation: methods, algorithms, and applications*, pages 3–25. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, United States of America, 2016. 3, 14, 24, 34
- [4] M. Asch, M. Bocquet, and M. Nodet. Chapter 3: Statistical estimation and sequential data assimilation. In *Data assimilation: methods, algorithms, and applications*, pages 90–121. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, United States of America, 2016. 3, 10, 29
- [5] M. Asch, M. Bocquet, and M. Nodet. Chapter 4: Nudging methods. In *Data assimilation: methods, algorithms, and applications*, pages 122–134. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, United States of America, 2016. 3, 5
- [6] Cristóbal Bertoglio, Philippe Moireau, and Jean-frederic Gerbeau. Sequential parameter estimation for fluid – structure problems : Application to hemodynamics. (November 2011):434–455, 2012. 3, 5, 17
- [7] Graud Blatman and Bruno Sudret. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering and System Safety*, 95(11):1216–1229, 2010. 60, 61
- [8] Peter H.M. Bovendeerd, Petra Borsje, Theo Arts, and Frans N. Van De Vosse. Dependence of intramyocardial pressure and coronary flow on ventricular loading and contractility: A model study. *Annals of Biomedical Engineering*, 34(12):1833–1845, 2006. 21, 22, 24, 25
- [9] Daniel Brodie, Arthur S. Slutsky, and Alain Combes. Extracorporeal life support for adults with respiratory failure and related indications: A review. *JAMA - Journal of the American Medical Association*, 322(6):557–568, 2019. 1, 2
- [10] J. Brynjarsdóttir and A. O’Hagen. Learning about physical parameters : the importance of model discrepancy. *Inverse problems*, 30(11), 2014. 25, 51
- [11] R Bunt. *Improving ECLS weaning outcome using real-time patient parameter estimation*. [Master-thesis, Eindhoven University of Technology], Eindhoven University of Technology research portal, 2021. 50
- [12] Alberto Carrassi, Marc Bocquet, Laurent Bertino, and Geir Evensen. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):1–50, 2018. 29
- [13] Alberto Carrassi, Marc Bocquet, Jonathan Demaeyer, Colin Grudzien, Patrick Raanes, and Stéphane Vannitsem. Data Assimilation for Chaotic Dynamics. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*, 4(1):1–42, 2022. 29, 36, 37
- [14] Alberto Carrassi, Michael Ghil, Anna Trevisan, and Francesco Uboldi. Data assimilation as a non-linear dynamical systems problem: Stability and convergence of the prediction-assimilation system. *Chaos*, 18(2):023112, 2008. 36
- [15] Nicholas C Cavarocchi, Harrison T Pitcher, Qiong Yang, Pawel Karbowski, Joseph Miessau, Harold M Hastings, and Hitoshi Hirose. Weaning of extracorporeal membrane oxygenation using continuous hemodynamic transesophageal echocardiography. *The Journal of Thoracic and Cardiovascular Surgery*, 146(6):1474–1479, 2013. 2

- [16] Mabel Chung, Ariel L. Shiloh, and Anthony Carlese. Monitoring of the adult patient on venoarterial extracorporeal membrane oxygenation. *The Scientific World Journal*, 2014:1–10, 2014. 3
- [17] Cesare Corrado, Jean Frédéric Gerbeau, and Philippe Moireau. Identification of weakly coupled multiphysics problems. Application to the inverse problem of electrocardiography. *Journal of Computational Physics*, 283:271–298, 2015. 3, 14, 15, 17
- [18] Chun Yuan Deng. A generalization of the Sherman – Morrison – Woodbury formula. *Applied Mathematics Letters*, 24(9):1561–1564, 2011. 16
- [19] Wouter P. Donders, Wouter Huberts, Frans N. van de Vosse, and Tammo Delhaas. Personalization of models with many model parameters: an efficient sensitivity analysis approach. International Journal for Numerical Methods in Biomedical Engineering. *International Journal for Numerical Methods in Biomedical Engineering*, 31(10):18, 2015. 60
- [20] Jindřich Duník, Ondřej Straka, Oliver Kost, and Jindřich Havlík. Noise covariance matrices in state-space models: A survey and comparison of estimation methods—Part I. *International Journal of Adaptive Control and Signal Processing*, 31(11):1505–1543, 2017. 50, 51
- [21] Vinzenz Eck, Wouter Donders, Jacob Sturdy, Jonathan Feinberg, Tammo Delhaas, Leif Hellevik, and Wouter Huberts. A guide to uncertainty quantification and sensitivity analysis for cardiovascular applications. *International Journal for Numerical Methods in Biomedical Engineering*, 2755(November 2015):31, 2016. 33, 34, 60, 61
- [22] Noura Fajraoui, Stefano Marelli, and Bruno Sudret. Sequential design of experiment for sparse polynomial chaos expansions. *SIAM-ASA Journal on Uncertainty Quantification*, 5(1):1061–1085, 2017. 61
- [23] J F Fraser, K Shekar, S Diab, K Dunster, S R Foley, C I Mcdonald, M Passmore, G Simonova, and J A Roberts. ECMO – the clinician ’ s view. *ISBT Science series*, 7:82–88, 2012. 1
- [24] David Freedman and Persi Diaconis. On the histogram as a density estimator:L2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57(4):453–476, 1981. 29
- [25] S.J. Garret. Chapter 13: Introductory Numerical Methods. In *Introduction to Actuarial and Financial Mathematical Methods*, pages 411–463. Elsevier, Leicester, United Kingdom, 2015. 23
- [26] Joseph H.A. Guillaume, John D. Jakeman, Stefano Marsili-Libelli, Michael Asher, Philip Brunner, B. Croke, Mary C. Hill, Anthony J. Jakeman, Karel J. Keesman, S. Razavi, and Johannes D. Stigter. Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. *Environmental Modelling and Software*, 119(July 2018):418–432, 2019. 34, 36, 48, 60
- [27] A.C. Guyton. and J.E. Hall. *Textbook of Medical Physiology*. Elsevier Saunders, Philadelphia, United States of America, 2006. 22
- [28] W. Huberts, C. de Jonge, W. P.M. van der Linden, M. A. Inda, J. H.M. Tordoir, F. N. van de Vosse, and E. M.H. Bosboom. A sensitivity analysis of a personalized pulse wave propagation model for arteriovenous fistula surgery. Part A: Identification of most influential model parameters. *Medical Engineering and Physics*, 35(6):810–826, 2013. 60
- [29] Nils P Johnson, Jo M Zelis, Pim A L Tonino, Patrick Houthuizen, R Arthur Bouwman, Guus R G Brueren, Daniel T Johnson, Jacques J Koolen, Hendrikus H M Korsten, Inge F Wijnbergen, Frederik M Zimmermann, Richard L Kirkeeide, Nico H J Pijls, and K Lance Gould. Pressure gradient vs . flow relationships to characterize the physiology of a severely stenotic aortic valve before and after transcatheter valve implantation. *European Heart Journal*, 39(28):2646–2655, 2018. 27, 46
- [30] Simon J Julier. The Spherical Simplex Unscented Transformation. *Proceedings of the 2003 American Control Conference*, 3:2430–2434, 2003. 11
- [31] Simon J. Julier and Jeffrey K. Uhlmann. New extension of the Kalman filter to nonlinear systems. *Signal Processing, Sensor Fusion, and Target Recognition VI*, 3068:182, 1997. 10, 11
- [32] Simon J. Julier and Jeffrey K. Uhlmann. Corrections to "Unscented filtering and nonlinear estimation". *Proceedings of the IEEE*, 92(12):1958, 2004. 10

- [33] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Journal of Fluids Engineering, Transactions of the ASME*, 83(1):95–108, 1961. 3, 5
- [34] Mary E. Keebler, Elias V. Haddad, Chun W. Choi, Stuart McGrane, Sandip Zalawadiya, Kelly H. Schlendorf, D. Marshall Brinkley, Matthew R. Danter, Mark Wigger, Jonathan N. Menachem, Ashish Shah, and Jo Ann Lindenfeld. Venoarterial Extracorporeal Membrane Oxygenation in Cardiogenic Shock. *JACC: Heart Failure*, 6(6):503–516, 2018. 1, 2
- [35] Michael M. Koerner, Michael D. Harper, Christopher K. Gordon, Douglas Horstmannshof, James W. Long, Michael J. Sasevich, James D. Neel, and Aly El Banayosy. Adult cardiac veno-arterial extracorporeal life support (VA-ECMO): Prevention and management of acute complications. *Annals of Cardiothoracic Surgery*, 8(1):66–75, 2019. 2
- [36] Katerina Konakli and Bruno Sudret. Polynomial meta-models with canonical low-rank approximations: Numerical insights and comparison to sparse polynomial chaos expansions. *Journal of Computational Physics*, 321:1144–1169, 2016. 60
- [37] Werner Kritzinger, Matthias Karner, Georg Traar, Jan Henjes, and Wilfried Sihn. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11):1016–1022, 2018. 4
- [38] Arthur Le Gall, Arnaud Follin, Bernard Cholley, Jean Mantz, Nadia Aissaoui, and Romain Pirracchio. Veno-arterial-ECMO in the intensive care unit: From technical aspects to clinical practice. *Anaesthesia Critical Care and Pain Medicine*, 37(3):259–268, 2018. 1, 2
- [39] Timothy W Leishman, Sarah Rollins, Heather M Smith, Timothy W Leishman, Sarah Rollins, and Heather M Smith. An experimental evaluation of regular polyhedron loudspeakers as omnidirectional sources of sound. *Acoustical Society of America*, 120(3):1411–1422, 2006. 18
- [40] Fernando V Lima, Murali R Rajamani, Tyler A Soderstrom, and James B Rawlings. Covariance and State Estimation of Weakly Observable Systems : Application to Polymerization Processes. *IEEE Transactions on Control Systems Technology*, 21(4):1249–1257, 2013. 51
- [41] C Van Loo. *Blood pressure regulation in patients on extracorporeal life support*. [Masterthesis, Eindhoven University of Technology], Eindhoven University of Technology research portal, 2020. 25
- [42] David G Luenberger. An Introduction to Observers. *IEEE Transactions on Automatic Control*, 16(6):592–602, 1971. 14
- [43] Enzo Lüsebrink, Christopher Stremmel, Konstantin Stark, Dominik Joskowiak, Thomas Czermak, Frank Born, Danny Kupka, Clemens Scherer, Mathias Orban, Tobias Petzold, Patrick Von Samsonhimmelstjerna, Stefan Kääb, Christian Hagl, Sven Peterss, and Martin Orban. Update on Weaning from Veno-Arterial Extracorporeal Membrane Oxygenation. *Journal of Clinical Medicine*, 9(4), 2020. 2
- [44] Nora Lüthen, Stefano Marelli, and Bruno Sudret. Sparse polynomial chaos expansions: Literature survey and benchmark. *SIAM-ASA Journal on Uncertainty Quantification*, 9(2):593–649, 2021. 60
- [45] A.M. Lyapunov. The General Problem of the Stability of Motion. Technical report, Kharkov Mathematical Society, Kharkov, 1892. 36
- [46] Yulong Ma, Zhiqian Wang, Xingang Zhao, Jianda Han, and Yuqing He. A UKF algorithm based on the singular value decomposition of state covariance. *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*, (60705028):5830–5835, 2010. 3, 10, 11
- [47] Silvana F. Marasco, George Lukas, Michael McDonald, James McMillan, and Benno Ihle. Review of ECMO (Extra Corporeal Membrane Oxygenation) Support in Critically Ill Adult Patients. *Heart Lung and Circulation*, 17(4):41–47, 2008. 1
- [48] Giovanni Mariscalco, Zein El-dean, Hakeem Yusuff, Thomas Fux, Angelo M Dell Aquila, Sigurdur Ragnarsson, Antonio Fiore, Magnus Dal, Dario Perna, Giuseppe Gatti, Tatu Juvonen, Svante Zipfel, Andrea Perrotti, Karl Bounader, Khalid Alkhamees, Antonio Loforte, Andrea Lechiancole, Marek Pol, Cristiano Spadaccio, Matteo Pettinari, Dieter De Keyzer, Henryk Welp, Daniele Maselli, Artur

- Lichtenberg, and Vito G Ruggieri. Duration of Venoarterial Extracorporeal Membrane Oxygenation and Mortality in Postcardiotomy Cardiogenic Shock. *Journal of Cardiothoracic and Vascular Anesthesia*, 35:2662–2668, 2021. 2
- [49] Roel Meiburg, Wouter Huberts, Marcel C.M. Rutten, and Frans N. van de Vosse. Uncertainty in model-based treatment decision support: Applied to aortic valve stenosis. *International Journal for Numerical Methods in Biomedical Engineering*, 36(10):1–21, 2020. 3, 21, 23, 33, 49, 64
- [50] Henrique M.T. Menegaz, João Y. Ishihara, Geovany A. Borges, and Alessandro N. Vargas. A Systematization of the Unscented Kalman Filter Theory. *IEEE Transactions on Automatic Control*, 60(10):2583–2598, 2015. 10
- [51] Philippe Moireau and Dominique Chapelle. Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems. *ESAIM - Control, Optimisation and Calculus of Variations*, 17(2):380–405, 2011. 3, 14, 15, 18
- [52] Philippe Moireau, Dominique Chapelle, and Patrick Le Tallec. Joint state and parameter estimation for distributed mechanical systems. *Computer Methods in Applied Mechanics and Engineering*, 197:659–677, 2008. 3, 14
- [53] Michael C. Neale and Michael B. Miller. The use of likelihood-based confidence intervals in genetic models. *Behavior Genetics*, 27(2):113–120, 1997. 34
- [54] Seyed Kamaledin Yadavar Nikraves. Chapter 2: Stability Analysis of Autonomous Systems. In *Nonlinear Systems Stability Analysis, Lyapunov-Based Approach*, pages 11–119. CRC Press, Tehran, Iran, 2013. 36, 37
- [55] S. Pant, B. Fabregas, J. F. Gerbeau, and I.E. Vignon-Clementel. A methodological paradigm for patient-specific multi-scale CFD simulations: from clinical measurements to parameter estimates for individual analysis. *International Journal for Numerical Methods in Biomedical Engineering*, 30(12):1614–1648, 2014. 33
- [56] N. Papadopoulos, S. Marinos, A. El-Sayed Ahmad, H. Keller, P. Meybohm, K. Zacharowski, A. Moritz, and A. Zierer. Risk factors associated with adverse outcome following extracorporeal life support: Analysis from 360 consecutive patients. *Perfusion (United Kingdom)*, 30(4):284–290, 2015. 1, 2
- [57] Federico Pappalardo, Marina Pieri, Blanca Arnaez Corada, Silvia Ajello, Giulio Melisurgo, Michele De Bonis, and Alberto Zangrillo. Timing and Strategy for Weaning From Venoarterial ECMO are Complex Issues. *Journal of Cardiothoracic and Vascular Anesthesia*, 29(4):906–911, 2015. 2
- [58] K.A.M.A. Pennings, J.R. Martina, B.F.M. Rodermans, J.R. Lahpor, F. N. van de Vosse, and M.C.M. Rutten. Pump Flow Estimation From Pressure Head and Power Uptake for the HeartAssist5 , HeartMate II , and HeartWare VADs. *ASAIO Journal*, 59(4):420–426, 2013. 51
- [59] Stephen G Penny. Mathematical foundations of hybrid data assimilation from a synchronization perspective perspective. *Chaos: An interdisciplinary Journal of Nonlinear Science*, 27(12):126801, 2017. 3, 5
- [60] D. T. Pham. Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly Weather Review*, 129(5):1194–1207, 2001. 3
- [61] Sjeng Quicken, Wouter P. Donders, Emiel M.J. Van Disseldorp, Kujtim Gashi, Barend M.E. Mees, Frans N. Van De Vosse, Richard G.P. Lopata, Tammo Delhaas, and Wouter Huberts. Application of an Adaptive Polynomial Chaos Expansion on Computationally Expensive Three-Dimensional Cardiovascular Models for Uncertainty Quantification and Sensitivity Analysis. *Journal of Biomechanical Engineering*, 138(12):1–11, 2016. 33, 60, 61, 62
- [62] Andreas Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 2009. 34, 35, 36
- [63] Douglas Reynolds. Gaussian Mixture Models. *Encyclopedia of Biometrics*, 741:827–832, 2015. 51

- [64] Tilai T. Rosalina, R. Arthur Bouwman, Marc R.H.M. van Sambeek, Frans N. van de Vosse, and Peter H.M. Bovendeerd. A mathematical model to investigate the effects of intravenous fluid administration and fluid loss. *Journal of Biomechanics*, 88:4–11, 2019. 25
- [65] B Y Robert R Ruffolo. Review : The Pharmacology of Dobutamine. *The American Journal of the Medical Sciences*, 294(4):244–248, 1987. 27
- [66] Gianluca Santise, Giovanna Panarello, Cettina Ruperto, Marco Turrisi, Gerlando Pilato, Andrea Giunta, Sergio Sciacca, and Michele Pilato. Extracorporeal membrane oxygenation for graft failure after heart transplantation: A multidisciplinary approach to maximize weaning rate. *International Journal of Artificial Organs*, 37(9):706–714, 2014. 2
- [67] Jean-Jacues E. Slotine and Weiping Li. Fundamentals of Lyapunov Theory. In *Applied Nonlinear Control*, pages 40–100. Prentice-Hall Inc., New Jersey, United States of America, 1991. 36, 40
- [68] Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, 2008. 61, 62
- [69] Pierre Tandeo, Pierre Ailliot, Marc Bocquet, Alberto Carrassi, Takemasa Miyoshi, Manuel Pulido, and Yicun Zhen. A Review of Innovation-Based Methods to Jointly Estimate Model and Observation Error Covariance Matrices in Ensemble Data Assimilation. *Monthly Weather Review*, 148(10):3973–3994, 2020. 5, 6, 50
- [70] Joanna Tohme, Camille Piat, Nadia Aissat, Guillaume Lebreton, Baptiste Duceau, and Adrien Bougl. Weaning-Related Shock in Patients With ECMO : Incidence , Mortality , and Predisposing Factors. *Journal of Cardiothoracic and Vascular Anesthesia*, 35(1):41–47, 2021. 2
- [71] Maxime Tondeur, Alberto Carrassi, Stephane Vannitsem, and Marc Bocquet. On Temporal Scale Separation in Coupled Data Assimilation with the Ensemble Kalman Filter. *Journal of Statistical Physics*, 179(5-6):1161–1185, 2020. 36
- [72] Mauro Ursino and Elisa Magosso. Acute cardiovascular response to isocapnic hypoxia. I. A mathematical model. *American Journal of Physiology - Heart and Circulatory Physiology*, 279(1 48-1):166–175, 2000. 25
- [73] R. Van Der Merwe and E. A. Wan. The square-root unscented Kalman filter for state and parameter-estimation. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 6:3461–3464, 2001. 10, 11, 14
- [74] Sean Van Diepen, Jason N. Katz, Nancy M. Albert, Timothy D. Henry, Alice K. Jacobs, Navin K. Kapur, Ahmet Kilic, Venu Menon, E. Magnus Ohman, Nancy K. Sweitzer, Holger Thiele, Jeffrey B. Washam, and Mauricio G. Cohen. *Contemporary Management of Cardiogenic Shock: A Scientific Statement from the American Heart Association*, volume 136. 2017. 1
- [75] J. Vanlier, C. A. Tiemann, P. A.J. Hilbers, and N. A.W. van Riel. An integrated strategy for prediction uncertainty analysis. *Bioinformatics*, 28(8):1130–1135, 2012. 34, 35
- [76] Pengfei Wei, Zhenzhou Lu, and Jingwen Song. Regional and parametric sensitivity analysis of Sobol’ indices. *Reliability Engineering and System Safety*, 137:87–100, 2015. 61
- [77] K. Xiong, H. Y. Zhang, and C. W. Chan. Performance evaluation of UKF-based nonlinear filtering. *Automatica*, 42(2):261–270, 2006. 3, 10

## A. Log-normal distributed PDF

In the current study, the logarithmic parameter transformation is applied to ensure a positive value of the parameter of interest. According to this logarithmic transformation, the parameter value  $\psi_i$  can be expressed logarithmic as follows:

$$\phi_i = \phi_{ref} 2^{\psi_i} \quad (\text{A.1})$$

Next, this equation is rewritten in order to express the parameter value  $\psi_i$  in terms of the logarithmic parameter value  $\phi_i$  and normalization factor  $\phi_{ref}$ , which results in:

$$\begin{aligned} \phi_i &= \phi_{ref} 2^{\psi_i}, \\ \frac{\phi_i}{\phi_{ref}} &= 2^{\psi_i}, \\ \ln\left(\frac{\phi_i}{\phi_{ref}}\right) &= \ln(2^{\psi_i}), \\ \ln\left(\frac{\phi_i}{\phi_{ref}}\right) &= \ln(2)\psi_i, \\ \psi_i &= \frac{1}{\ln(2)} \ln\left(\frac{\phi_i}{\phi_{ref}}\right). \end{aligned} \quad (\text{A.2})$$

When assuming that the parameter of interest is normally distributed, the formula for the Gaussian distributed PDF (Equation 4.2) can be used to describe the *a posteriori* PDF of the parameter of interest. Formally, this can be given as:

$$P(\psi_i|\theta_k, \sigma_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2}\left(\frac{\psi_i - \theta_k}{\sigma_k}\right)^2}, \quad \text{with } i = 1, 2, \dots, n_{bins}. \quad (\text{A.3})$$

To derive an expression for the log-normal distributed PDF, the normal distributed PDF has to be scaled accordingly. This scaling relies on the normal expressed parameter  $\psi_i$  relative to the logarithmic expressed parameter  $\phi_i$ , resulting in:

$$\begin{aligned} P(\phi_i|\theta_k, \sigma_k) &= \left| \frac{d\psi_i}{d\phi_i} \right| P(\psi_i|\theta_k, \sigma_k), \\ P(\phi_i|\theta_k, \sigma_k) &= \left| \frac{d\left(\frac{1}{\ln(2)} \ln\left(\frac{\phi_i}{\phi_{ref}}\right)\right)}{d\phi_i} \right| P(\psi_i|\theta_k, \sigma_k), \\ P(\phi_i|\theta_k, \sigma_k) &= \left| \frac{d}{d\phi_i} \ln\left(\frac{\phi_i}{\phi_{ref}}\right) \right| \frac{1}{\ln(2)} P(\psi_i|\theta_k, \sigma_k). \end{aligned} \quad (\text{A.4})$$

When choosing a value of 1 for the normalization factor  $\phi_{ref}$ , this expression can be reduced further, resulting in:

$$\begin{aligned} P(\phi_i|\theta_k, \sigma_k) &= \left| \frac{d}{d\phi_i} \ln(\phi_i) \right| \frac{1}{\ln(2)} P(\psi_i|\theta_k, \sigma_k), \\ P(\phi_i|\theta_k, \sigma_k) &= \left| \frac{1}{\phi_i} \right| \frac{1}{\ln(2)} P(\psi_i|\theta_k, \sigma_k). \end{aligned} \quad (\text{A.5})$$

Subsequently, the definition of the Gaussian distributed PDF (Equation A.3) and the logarithmically expressed parameter (Equation A.2) will be substituted in the above supplied expression, which leads to:

$$\begin{aligned}
 P(\phi_i|\theta_k, \sigma_k) &= \left| \frac{1}{\phi_i} \frac{1}{\ln(2)} \right| \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2} \left( \frac{\psi_i - \theta_k}{\sigma_k} \right)^2}, \\
 P(\phi_i|\theta_k, \sigma_k) &= \frac{1}{\phi_i} \frac{1}{\sqrt{2\pi}\sigma_k \ln(2)} e^{-\frac{1}{2} \left( \frac{\psi_i - \theta_k}{\sigma_k} \right)^2}, \\
 P(\phi_i|\theta_k, \sigma_k) &= \frac{1}{\phi_i} \frac{1}{\sqrt{2\pi}\sigma_k \ln(2)} e^{-\frac{1}{2} \left( \frac{\frac{1}{\ln(2)} \ln\left(\frac{\phi_i}{\phi_{ref}}\right) - \theta_k}{\sigma_k} \right)^2}, \\
 P(\phi_i|\theta_k, \sigma_k) &= \frac{1}{\phi_i} \frac{1}{\sqrt{2\pi}\sigma_k \ln(2)} e^{-\frac{1}{2} \left( \frac{\ln\left(\frac{\phi_i}{\phi_{ref}}\right) - \theta_k \ln(2)}{\sigma_k \ln(2)} \right)^2}, \\
 P(\phi_i|\theta_k, \sigma_k) &= \frac{1}{\phi_i} \frac{1}{\sqrt{2\pi}\sigma_k \ln(2)} e^{-\frac{1}{2} \left( \frac{\ln(\phi_i) - \theta_k \ln(2) - \phi_{ref}}{\sigma_k \ln(2)} \right)^2}, \\
 P(\phi_i|\theta_k, \sigma_k) &= \frac{1}{\phi_i} \frac{1}{\sqrt{2\pi}\sigma_k \ln(2)} e^{-\frac{1}{2} \left( \frac{\ln(\phi_i) - (\theta_k \ln(2) + \ln(\phi_{ref}))}{\sigma_k \ln(2)} \right)^2}. \tag{A.6}
 \end{aligned}$$

This expression will be simplified by defining the logarithmic estimated parameter  $\mu_k$  and the corresponding logarithmic standard deviation  $s_k$ . Both of these logarithmic definitions can be given as:

$$\mu_k = \theta_k \ln(2) + \ln(\theta_{ref}) \quad \text{and} \quad s_k = \sigma_k \ln(2). \tag{A.7}$$

The above stated definitions of the estimated parameter  $\mu_k$  and standard deviation  $s_k$  will be used to simplify the derived expression of the log-normal distributed PDF (Equation A.6). Eventually, this will result in the equation that will be used to calculate the log-normal distributed PDF in the current study, explicitly:

$$P(\phi_i|\mu_k, s_k) = \frac{1}{\phi_i} \frac{1}{\sqrt{2\pi}s_k} e^{-\frac{1}{2} \left( \frac{\ln(\phi_i) - \mu_k}{s_k} \right)^2}, \quad \text{with } i = 1, 2, \dots, n_{bins}. \tag{A.8}$$



## B. Global sensitivity analysis

Adjusting the model input to patient-specific conditions is not a straightforward process, but rather a hard challenge. The quantification of these model inputs can be performed in various ways, like mathematical derivation or measurements [21]. However, the measurements needed to obtain information about physical parameters are hampered by measurement uncertainty as well as biological variability. These factors will result in sparse and uncertain model input leading to dissociation of model outputs, in other words uncertainty in the model output [19]. Roughly two distinct sources of uncertainty can be distinguished; the first is a lack of knowledge, which is called epistemic uncertainty, whereas the second relies on an association of intrinsic variability [7]. Regarding the application discussed in the current study, appropriate quantification of uncertainty in the model output is considered to be crucial for a decision-making in clinical practice. In order to determine the relative contribution of uncertain input parameters and their interactions to the total amount of uncertainty in the model output, sensitivity analysis can be used. The outcome of the sensitivity analysis can provide guidance to identify which input parameters can be fixed based on literature values, which is called parameter fixing [19]. Additionally, the result of the sensitivity analysis can help to identify which input parameters contribute the most to the total amount of uncertainty in the model outputs. By analyzing this relative uncertainty contribution, the development of more accurate methods for the assessment of certain input parameters can be prioritized over others, which is called parameter prioritization [28]. Furthermore, sensitivity analysis can be defined as a forward problem, which consists of predicting a model response based on a given set of input parameters [26].

Global sensitivity analysis can be used to quantify the degree of uncertainty in the output arising from variations in input parameters [7]. This global analysis includes the entire input parameter space, as well as the interactions between different input parameters in the uncertainty quantification, while the earlier discussed local method only includes certain parts of the input space. Therefore, this global method can be regarded as a more accurate method to quantify the uncertainty in the model output compared to the local method [61]. Consequently, this global method can be used to identify the contribution of uncertain input parameters to the total uncertainty in the model output.

However, a thorough analysis of the model response requires that all uncertainties of the model and its fundamental phenomena are taken into consideration, which necessitates repeated simulation with different configurations of input parameters [36]. As a result, many model simulations are required which would make a thorough analysis of the model response hardly applicable for a computational demanding model [7]. To solve this computational problem, a popular approach in engineering is to construct a surrogate model, which is called a meta-model [7, 44]. This meta-model is an analytical approximation of the original model, which is typically computationally inexpensive to evaluate and properly approximates the input-output relation of the original model. Since the evaluation of the meta-model is computationally inexpensive, the required number of simulations can be performed at low computational cost, making the execution of uncertainty analysis feasible [44].

The global sensitivity analysis can be performed by using a method called adaptive generalized polynomial chaos expansion (agPCE). By employing this method, a meta-model will be constructed based on the output of original model, which is accomplished by expanding the model output by using a basis of orthogonal polynomials [61]. Formally, an explicit representation of model function  $f$  is established by expanding the model output  $\mathbf{Y}$  into a series of orthogonal polynomials:

$$\mathbf{Y} = f(\mathbf{X}) = f_{\text{PCE}}(\mathbf{X}) = \sum_{\rho \in \mathbf{G}^\infty} c_\rho \Phi_\rho(\mathbf{X}). \quad (\text{B.1})$$

In this equation, model output  $\mathbf{Y}$  is expanded into orthogonal polynomials  $\Phi_\rho$ , which are functions of the model input  $\mathbf{X}$  with associated expansion coefficients  $c_\rho$ . The set of orthogonal polynomials  $\Phi_\rho$  can be obtained by making use of univariate polynomials  $\phi_{\rho_i}$  with order  $\rho_i$ , which can be derived from the multi-index vector  $\rho = [\rho_1 \ \rho_2 \ \dots \ \rho_D]$ . Mathematically, this can be stated as follows:

$$\Phi_\rho(\mathbf{X}) = \prod_{i=1}^D \phi_{\rho_i}(\mathbf{X}_i). \quad (\text{B.2})$$

These expansion coefficients  $c_\rho$  belonging to the orthogonal polynomials  $\Phi_\rho$  are computed by using a

non intrusive method, which implies that the agPCE method can be applied to any model without modifying the model code itself [7, 22]. Furthermore, the extent of the series orthogonal polynomials  $\Phi_\rho$  with corresponding coefficients  $c_\rho$  that together represent the meta-model increases exponentially with number of input parameters and polynomial degree. As a result, the required number of model evaluation will still be large in case of a model with a substantial amount of input parameters. Therefore, the agPCE method aims to reduce the number of expansion terms by only adding those that contribute significantly to the quality of the meta-model. This reductive approach is justifiable in the majority of cases because the model response is dominated by low order interactions between parameters and therefore the high order interactions can be neglected. In the agPCE method, this reduction of expansion terms is achieved by iteratively applying forward- and backward steps. During the forward steps, polynomials are added that increase the quality of the meta-model, whereas during the backward steps polynomials are removed that initially seemed to be relevant but can be eliminated without reducing the quality of the meta-model [61]. Ultimately, the primary reason why this agPCE method is considered suitable for global sensitivity analysis is that the constructed meta-model can be used to analytically compute the Sobol' indices [68].

Sobol' indices are global, variance-based sensitivity indices that can be used for uncertainty quantification [21]. There are two distinct types of Sobol' indices that will be considered for every input parameter of the considered system, namely the main- and total sensitivity index. The main sensitivity index of a particular input parameter represents the individual contribution of this input parameter concerning the total model output variance, examined by fixing this particular input parameter [76]. This main sensitivity index is also considered as a first-order Sobol' sensitivity index because it only includes direct effects of a certain input parameter, which is defined as:

$$S_i = \frac{\mathbb{V}[\mathbb{E}(\mathbf{Y}|\Theta_i)]}{\mathbb{V}[\mathbf{Y}]}.$$
 (B.3)

In this expression the main sensitivity index  $S_i$  represents the expected reduction in the total variance of the model output  $\mathbb{V}[(\mathbf{Y})]$  when the uncertain input parameter  $\Theta_i$  would be fixed. Because the main sensitivity index reflects the direct effects of a particular input parameter, it is considered appropriate for identifying which uncertain input parameters contribute significantly to the total variance of the model output and are therefore eligible for parameter prioritization [21]. Furthermore, the second-order Sobol' sensitivity index will represent the portion of the total variance in the model output  $\mathbb{V}[(\mathbf{Y})]$  that can be contributed to interactions between parameter  $\Theta_i$  and  $\Theta_j$ . This second-order index is defined as following:

$$S_{i,j} = \frac{\mathbb{V}[\mathbb{E}(\mathbf{Y}|\Theta_i, \Theta_j)]}{\mathbb{V}[\mathbf{Y}]}.$$
 (B.4)

However, this separate estimation of each second-order sensitivity index, which is basically every possible combination of input parameters, will take a considerable amount of time. Luckily, the second mentioned type of Sobol' indices, the total sensitivity index, includes information about interaction effects without the need of estimating all second-order sensitivity index separately [21]. In fact, the total sensitivity index incorporates both the individual effect of the input parameter and the interaction effect of the input parameter with all the remaining parameters [76]. The total sensitivity index of parameter  $\Theta_i$  is given by:

$$S_{T,i} = \frac{\mathbb{V}[\mathbf{Y}] - \mathbb{V}[\mathbb{E}(\mathbf{Y}|\Theta_{-i})]}{\mathbb{V}[\mathbf{Y}]} = 1 - \frac{\mathbb{V}[\mathbb{E}(\mathbf{Y}|\Theta_{-i})]}{\mathbb{V}[\mathbf{Y}]}.$$
 (B.5)

Here,  $\Theta_{-i}$  represents a set of all uncertain input parameters except the input parameter of interest  $\Theta_i$ . This expression enables the calculation of the total sensitivity index  $S_{T,i}$  of input parameter  $\Theta_i$ , which reflects the variance of the model output that is attributable to both the direct effect of  $\Theta_i$  and all interaction effects of  $\Theta_i$  with other input parameters. Therefore, the difference between  $S_{T,i}$  and  $S_i$  will only reflect the interaction effect, whereas these terms will be equal if there are no interaction effects present. The total sensitivity index is considered useful for determining which input parameters are suitable for parameter fixing within their uncertainty domain because this index considers both direct- and interaction effects [21].

Both Sobol' sensitivity indices will be determined through evaluation of the meta-model, which is constructed by using the agPCE method. As a result, both indices can be calculated analytically [68]. To establish this calculation, first the variance of the meta-model will be calculated by the following expression:

$$\mathbb{V}[\mathbf{Y}] \approx \sum_{\rho \in \mathbf{G}} c_{\rho}^2 \mathbf{N}_{\rho}. \quad (\text{B.6})$$

Where,  $\mathbf{N}_{\rho}$  represents the normalization factor and  $c_{\rho}$  the expansion coefficient with order  $\rho$ , which can be obtained from the multi-index vector  $\rho = [\rho_1 \ \rho_2 \ \dots \ \rho_D]$ . The normalization factor  $\mathbf{N}_{\rho}$  depends on the type of polynomials used to construct the meta-model. The type of polynomials that will be used is adopted from the study of Quicken et al. (2016), in which orthogonal Legendre polynomials are used to construct the meta-model [61]. The use of Legendre polynomials results in a normalization factor  $\mathbf{N}_{\rho}$  given by:

$$\mathbf{N}_{\rho} = \prod_{i=1}^D \frac{1}{2\rho_i + 1}. \quad (\text{B.7})$$

After calculating the normalization factor  $\mathbf{N}_{\rho}$ , the main- and total sensitivity index of input parameter  $\Theta_i$  can eventually be calculated based on the meta-model. First, the main sensitivity index  $S_i$  of input parameter  $\Theta_i$  can be obtained from evaluating the meta-model according to:

$$S_i \approx \frac{1}{\mathbb{V}[\mathbf{Y}]} \sum_{\rho \in \mathbf{G}_i} c_{\rho}^2 \mathbf{N}_{\rho}. \quad (\text{B.8})$$

In this equation, the term  $\mathbf{G}_i$  contains all multi-indices of vector  $\rho$  in which the index of input parameter  $\Theta_i$  ( $\rho_i$ ) is positive while all other terms are equal to zero. Basically, this convention is equal to stating that only direct effects of input parameter  $\Theta_i$  will be included in this main sensitivity index. Lastly, the total sensitivity index of input parameter  $\Theta_i$  can be calculated from the meta-model by using:

$$S_{T,i} \approx \frac{1}{\mathbb{V}[\mathbf{Y}]} \sum_{\rho \in \mathbf{G}_{T,i}} c_{\rho}^2 \mathbf{N}_{\rho}. \quad (\text{B.9})$$

In this final mathematical expression,  $\mathbf{G}_{T,i}$  represents all multi-indices in which the index of input parameter  $\Theta_i$  ( $\rho_i$ ) is positive. Formally, the resulting set of multi-indices  $\mathbf{G}_{T,i}$  incorporate both direct- and interaction effects of input parameter  $\Theta_i$ .

## C. Decision rule of profile likelihood

As described in Section ??, the probability density function of a certain parameter of interest  $\theta$  will describes the conditional probability of parameter value  $\psi$  when knowing parameter estimate  $\hat{\theta}$  and corresponding standard deviation  $\sigma$ . When assuming a Gaussian distribution of the *a posteriori* PDF, the likelihood of a certain parameter  $L(\theta_{PL})$  by evaluating parameter value  $\psi_{PL}$ , which can be described by using Equation 4.2:

$$L(\theta_{PL}) = P(\psi_{PL}|\hat{\theta}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{\psi_{PL}-\hat{\theta}}{\sigma})^2}. \quad (C.1)$$

Subsequently, the maximum likelihood estimation (MLE) is used to find the parameter value  $\psi$  which will most likely reflect the measured data. The maximum likelihood of the parameter of interest  $\theta$ , or in other words the most probable parameter value  $\psi_{opt}$ , will occur when parameter value  $\psi$  is exactly matching parameter estimate  $\hat{\theta}$ . In this particular case, the difference between the parameter value  $\psi$  and the parameter estimate  $\hat{\theta}$  will be equal to 0, which implies that the last of Equation C.1 will be equal to 1. Therefore, the maximum likelihood of the parameter of interest  $L(\theta_{opt})$  will be located at the most probable parameter value  $\psi_{opt}$ , which can be stated as:

$$L(\theta_{opt}) = P(\psi_{opt} = \hat{\theta}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \quad (C.2)$$

After determining the definition of the profile likelihood (Equation C.1) and the maximum likelihood (Equation C.2), the maximum likelihood ratio  $\Lambda$  can be derived. The derivation of the maximum likelihood ratio  $\Lambda$  will start by supplying the logarithmic defined ratio  $\Lambda$ , following by substitution of Equations C.1 and C.2:

$$\begin{aligned} \Lambda(\theta_{PL}) &= -2\ln \frac{L(\theta_{PL})}{L(\theta_{opt})} \\ &= -2\ln \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{\psi-\hat{\theta}}{\sigma})^2}}{\frac{1}{\sqrt{2\pi}\sigma}} \\ &= -2\ln e^{-\frac{1}{2}(\frac{\psi-\hat{\theta}}{\sigma})^2} \\ &= \frac{-2}{-2} \left( \frac{\psi - \hat{\theta}}{\sigma} \right)^2 \\ &= \left( \frac{\psi - \hat{\theta}}{\sigma} \right)^2 \\ &= \frac{(\psi - \hat{\theta})^2}{\sigma^2} \end{aligned} \quad (C.3)$$

This eventually derived expression does contain the structure of the Chi-squared formula. Therefore, the Chi-squared distribution with a certain amount of degrees of freedom ( $df = 1$ ) and chosen confidence level  $\gamma$  will be used to determine the boundaries of the confidence interval  $[\sigma_{\theta}^-, \sigma_{\theta}^+]$ . As a result, the decision rule for the profile likelihood method is formulated as:

$$-2\ln \left( \frac{L(\theta_{PL})}{L(\theta_{opt})} \right) \leq \chi_{1-\gamma, df=1}^2. \quad (C.4)$$

## D. Initialized *a posteriori* estimate

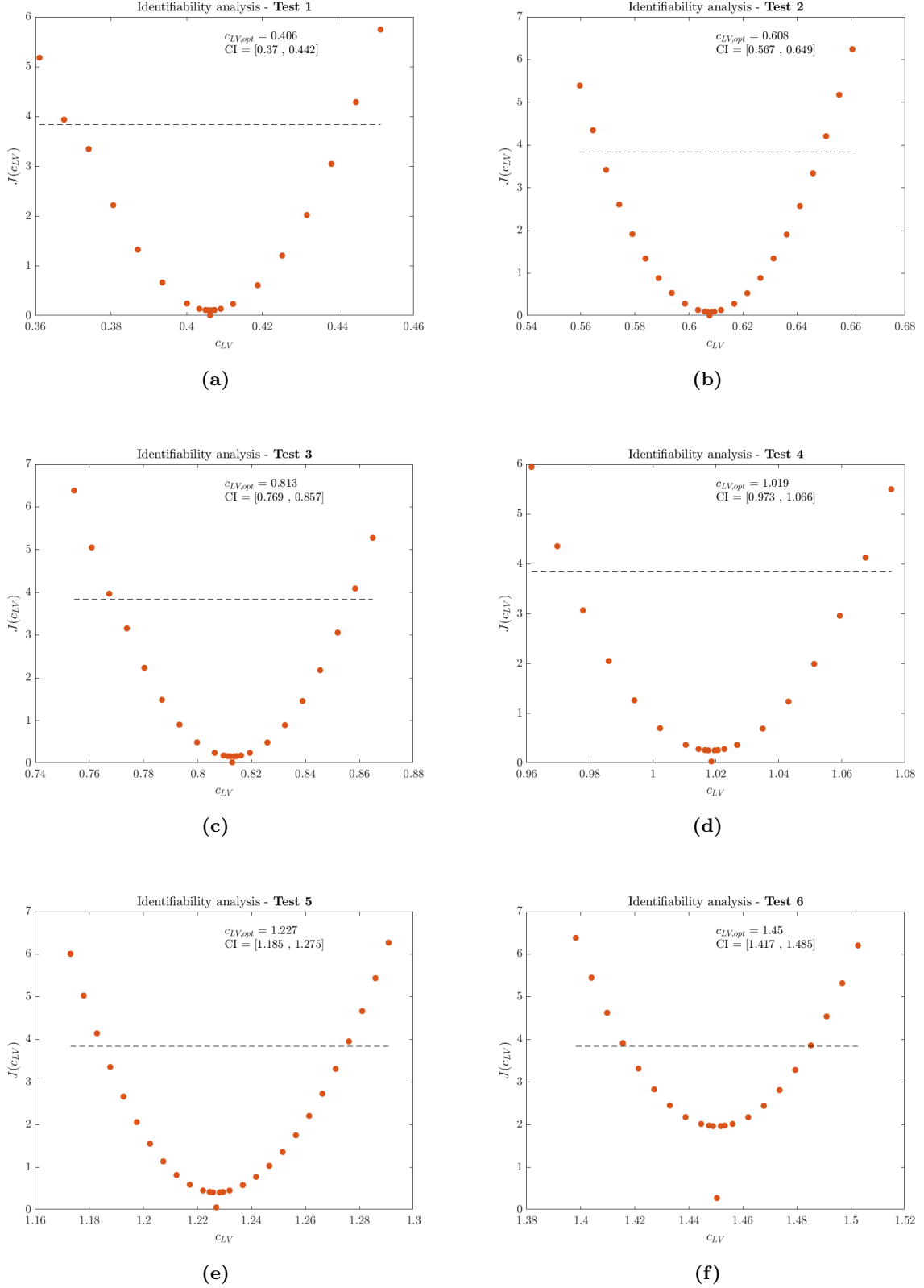
The *a posteriori* estimate  $\hat{x}_k^a$  needs to be initialized to be able to perform the first iteration of the proposed ROUKF algorithm. In this *a posteriori* estimate  $\hat{x}_k^a$  the states ( $x$ ) and parameters ( $\theta$ ) are gathered in one collective state-parameter space, which is called the augmented state vector. Therefore, the *a posteriori* estimate can be denoted as  $\hat{x}_k^a = [\hat{x}_k \ \hat{\theta}_k]^T$ . The size of this *a posteriori* estimate is determined by the number of states ( $n_s$ ) and the number of parameters ( $n_p$ ), explicitly the size of  $\hat{x}_k^a$  is equal to  $n_s + n_p$ . In the current study, the number of states ( $n_s$ ) is equal to 9, with their explicit definition being elaborated below. The number of parameters ( $n_p$ ) is equal to 1, because the parameter of interest regarding the ECMO application is the left ventricular contractility  $c_{LV}$ . The initial *a posteriori* estimate  $\hat{x}_k^a$  is arbitrarily determined, which can be given as:

$$\hat{x}_0^a = \begin{bmatrix} \hat{x}_0 \\ \hat{\theta}_0 \end{bmatrix} = \begin{bmatrix} p_{LV} \\ p_{ao} \\ V_{LV} \\ q_{mv} \\ q_{av} \\ V_{c_1} \\ q_{c_1-c_2} \\ V_{c_2} \\ q_{per} \\ c_{lv} \end{bmatrix} = \begin{bmatrix} 8 & [mmHg] \\ 80 & [mmHg] \\ 2V_{LV,0} & [mL] \\ 0 & [\frac{mL}{s}] \\ 0 & [\frac{mL}{s}] \\ p_{ao}C_{art} & [mL] \\ 0 & [\frac{mL}{s}] \\ p_{ao}C_{art} & [mL] \\ 0 & [\frac{mL}{s}] \\ 0.8 & [-] \end{bmatrix} \quad (D.1)$$

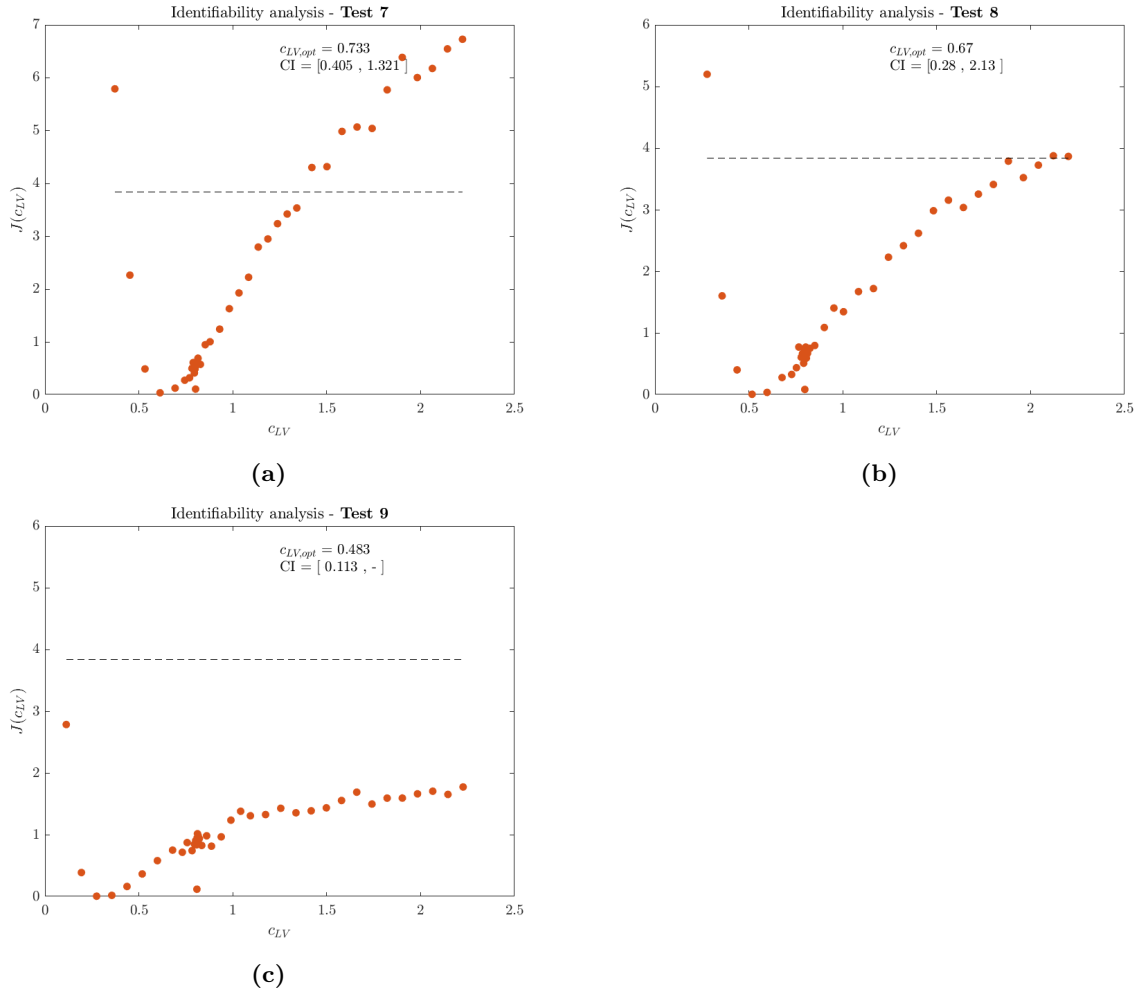
Some elements of this initial *a posteriori* estimate  $\hat{x}_0^a$  are correlated to the input parameters of the mathematical model. In the current study, almost all input parameter are adopted from the study of Meiburg et al. (2020) [49]. According to this study, the unloaded cavity volume  $V_{LV,0}$  is set to 80 [mL] and the arterial compliance  $C_{art}$  to 1 [ $\frac{mL}{mmHg}$ ]. This results in the follow representation of the *a posteriori* estimate  $\hat{x}_0^a$ :

$$\hat{x}_0^a = \begin{bmatrix} 8 & [mmHg] \\ 80 & [mmHg] \\ 160 & [mL] \\ 0 & [\frac{mL}{s}] \\ 0 & [\frac{mL}{s}] \\ 80 & [mL] \\ 0 & [\frac{mL}{s}] \\ 80 & [mL] \\ 0 & [\frac{mL}{s}] \\ 0.8 & [-] \end{bmatrix} \quad (D.2)$$

## E. Results identifiability analysis



**Figure E.1:** Illustration of the identifiability analysis based on the synthetic data of (a) Test 1, (b) Test 2, (c) Test 3, (d) Test 4, (e) Test 5 and (f) Test 6.



**Figure E.2:** Illustration of the identifiability analysis based on the synthetic data of (a) Test 7, (b) Test 8 and (c) Test 9

## F. Results in vitro data

$True\ c_{LV}\ [-]$	$Estimated\ c_{LV}\ [-]$	$CI\ c_{LV}\ [-]$	$RMSE$	$Diff.\ c_{LV}\ [-]$
0.4	0.469	[0.455 , 0.483]	0.001	+0.069
0.4	0.467	[0.451 , 0.483]	0.002	+0.067
0.4	0.470	[0.456 , 0.484]	0.001	+0.070
0.6	0.693	[0.679 , 0.707]	0.002	+0.093
0.6	0.683	[0.667 , 0.699]	0.001	+0.083
0.6	0.693	[0.679 , 0.707]	0.002	+0.093
0.8	0.878	[0.866 , 0.890]	0.003	+0.078
0.8	0.879	[0.867 , 0.891]	0.003	+0.079
0.8	0.878	[0.868 , 0.880]	0.002	+0.078
1.0	1.054	[1.022 , 1.086]	0.009	+0.054
1.0	1.024	[1.010 , 1.038]	0.004	+0.023
1.0	1.032	[1.012 , 1.052]	0.006	+0.032
1.2	1.191	[1.149 , 1.233]	0.014	-0.009
1.2	1.152	[1.132 , 1.172]	0.008	-0.048
1.2	1.129	[1.115 , 1.143]	0.008	-0.071
1.4	1.196	[1.158 , 1.238]	0.014	-0.204
1.4	1.239	[1.189 , 1.289]	0.017	-0.161
1.4	1.184	[1.150 , 1.218]	0.013	-0.216

**Table F.1:** Results of the application of the proposed ROUKF-based algorithm to the in vitro data that were generated by the material twin in **Configuration 1**.

$True\ c_{LV}\ [-]$	$Estimated\ c_{LV}\ [-]$	$CI\ c_{LV}\ [-]$	$RMSE$	$Diff.\ c_{LV}\ [-]$
0.4	0.484	[0.470 , 0.498]	0.001	+0.084
0.4	0.488	[0.472 , 0.504]	0.001	+0.088
0.4	0.491	[0.477 , 0.505]	0.001	+0.091
0.6	0.728	[0.712 , 0.744]	0.003	+0.128
0.6	0.721	[0.709 , 0.733]	0.002	+0.119
0.6	0.721	[0.709 , 0.733]	0.002	+0.119
0.8	0.924	[0.868 , 0.988]	0.015	+0.124
0.8	0.893	[0.881 , 0.905]	0.003	+0.093
0.8	0.892	[0.880 , 0.904]	0.003	+0.092
1.0	1.080	[0.978 , 1.038]	0.009	+0.008
1.0	0.985	[0.969 , 1.010]	0.005	-0.015
1.0	1.007	[0.973 , 1.041]	0.010	+0.007
1.2	1.071	[1.047 , 1.095]	0.008	-0.129
1.2	1.020	[1.000 , 1.040]	0.008	-0.180
1.2	1.008	[0.986 , 1.030]	0.008	-0.192
1.4	1.066	[1.018 , 1.114]	0.015	-0.334
1.4	1.070	[1.034 , 1.106]	0.012	-0.330
1.4	1.042	[1.012 , 1.072]	0.011	-0.358

**Table F.2:** Results of the application of the proposed ROUKF-based algorithm to the in vitro data that were generated by the material twin in **Configuration 2**.



$True\ c_{LV}\ [-]$	$Estimated\ c_{LV}\ [-]$	$CI\ c_{LV}\ [-]$	$RMSE$	$Diff.\ c_{LV}\ [-]$
0.4	0.373	[0.343 , 0.403]	0.005	-0.027
0.4	0.398	[0.328 , 0.468]	0.025	-0.002
0.4	0.422	[0.358 , 0.486]	0.008	+0.022
0.6	0.729	[0.717 , 0.741]	0.002	+0.129
0.6	0.731	[0.717 , 0.745]	0.003	+0.131
0.6	0.725	[0.713 , 0.737]	0.002	+0.125
0.8	0.903	[0.891 , 0.915]	0.003	+0.103
0.8	0.905	[0.893 , 0.917]	0.003	+0.105
0.8	0.903	[0.885 , 0.921]	0.003	+0.103
1.0	1.017	[0.997 , 1.037]	0.006	+0.017
1.0	1.008	[0.990 , 1.026]	0.005	+0.008
1.0	1.007	[0.991 , 1.023]	0.005	+0.007
1.2	1.057	[1.039 , 1.075]	0.007	-0.143
1.2	1.044	[1.022 , 1.066]	0.009	-0.156
1.2	1.060	[1.040 , 1.080]	0.008	-0.140
1.4	0.998	[0.960 , 1.036]	0.015	-0.402
1.4	0.976	[0.928 , 1.024]	0.016	-0.424
1.4	0.999	[0.959 , 1.039]	0.013	-0.401

**Table F.3:** Results of the application of the proposed ROUKF-based algorithm to the in vitro data that were generated by the material twin in **Configuration 3**.

$True\ c_{LV}\ [-]$	$Estimated\ c_{LV}\ [-]$	$CI\ c_{LV}\ [-]$	$RMSE$	$Diff.\ c_{LV}\ [-]$
0.4	0.564	[0.540 , 0.588]	0.004	+0.164
0.4	0.629	[0.549 , 0.709]	0.015	+0.229
0.4	0.599	[0.533 , 0.665]	0.012	+0.199
0.6	0.801	[0.787 , 0.815]	0.003	+0.201
0.6	0.804	[0.790 , 0.818]	0.003	+0.204
0.6	0.801	[0.789 , 0.813]	0.003	+0.201
0.8	0.990	[0.976 , 1.004]	0.004	+0.190
0.8	0.978	[0.956 , 1.000]	0.005	+0.178
0.8	0.972	[0.960 , 0.984]	0.003	+0.172
1.0	1.114	[1.090 , 1.138]	0.007	+0.114
1.0	1.073	[1.059 , 1.087]	0.005	+0.073
1.0	1.084	[1.072 , 1.096]	0.004	+0.084
1.2	1.080	[1.010 , 1.150]	0.020	-0.120
1.2	0.993	[0.965 , 1.021]	0.010	-0.207
1.2	1.005	[0.973 , 1.037]	0.011	-0.195
1.4	0.879	[0.793 , 0.965]	0.023	-0.521
1.4	0.916	[0.818 , 1.014]	0.025	-0.484
1.4	0.826	[0.776 , 0.876]	0.015	-0.574

**Table F.4:** Results of the application of the proposed ROUKF-based algorithm to the in vitro data that were generated by the material twin in **Configuration 4**.

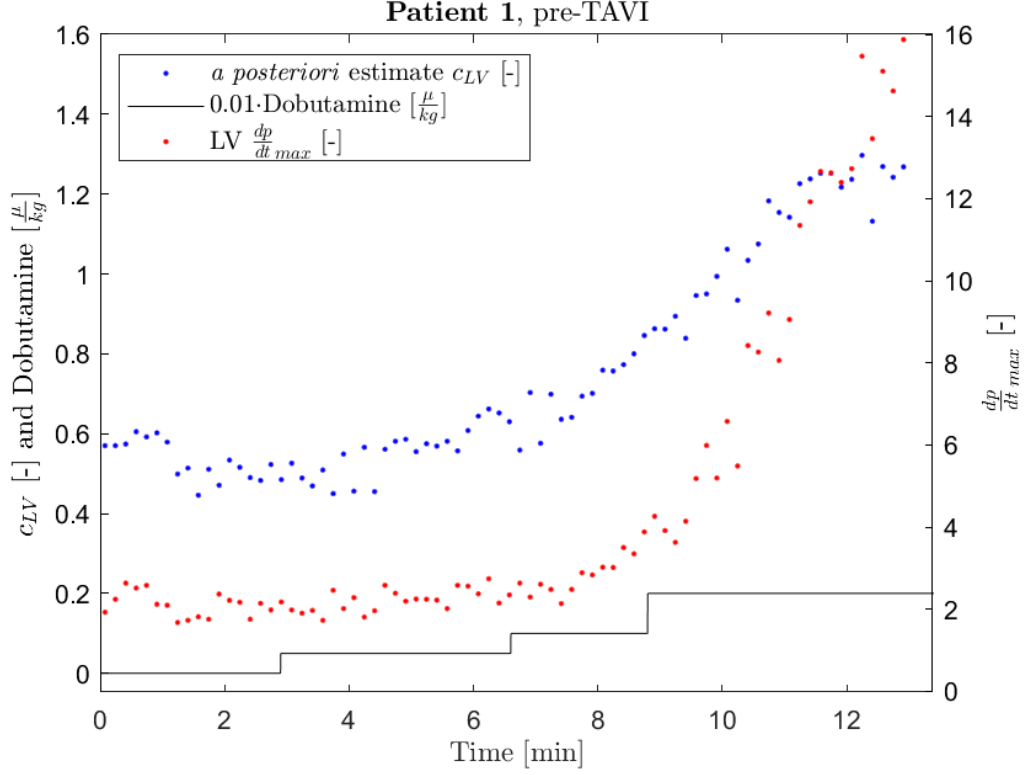
$True\ c_{LV}\ [-]$	$Estimated\ c_{LV}\ [-]$	$CI\ c_{LV}\ [-]$	$RMSE$	$Diff.\ c_{LV}\ [-]$
0.4	0.636	[0.624 , 0.648]	0.002	+0.236
0.4	0.633	[0.615 , 0.651]	0.002	+0.233
0.4	0.639	[0.623 , 0.655]	0.002	+0.239
0.6	0.824	[0.800 , 0.848]	0.006	+0.224
0.6	0.800	[0.788 , 0.812]	0.002	+0.200
0.6	0.812	[0.794 , 0.830]	0.004	+0.212
0.8	0.935	[0.921 , 0.949]	0.004	+0.135
0.8	0.932	[0.918 , 0.946]	0.004	+0.132
0.8	0.938	[0.918 , 0.958]	0.005	+0.138
1.0	1.043	[1.023 , 1.063]	0.006	+0.043
1.0	1.020	[1.008 , 1.032]	0.003	+0.020
1.0	1.013	[1.003 , 1.023]	0.003	+0.013
1.2	1.128	[1.088 , 1.168]	0.013	-0.072
1.2	1.072	[1.060 , 1.084]	0.005	-0.128
1.2	1.067	[1.053 , 1.081]	0.005	-0.133
1.4	0.976	[0.928 , 1.024]	0.015	-0.424
1.4	0.980	[0.936 , 1.024]	0.013	-0.420
1.4	0.949	[0.911 , 0.987]	0.011	-0.451

**Table F.5:** Results of the application of the proposed ROUKF-based algorithm to the in vitro data that were generated by the material twin in **Configuration 5**.

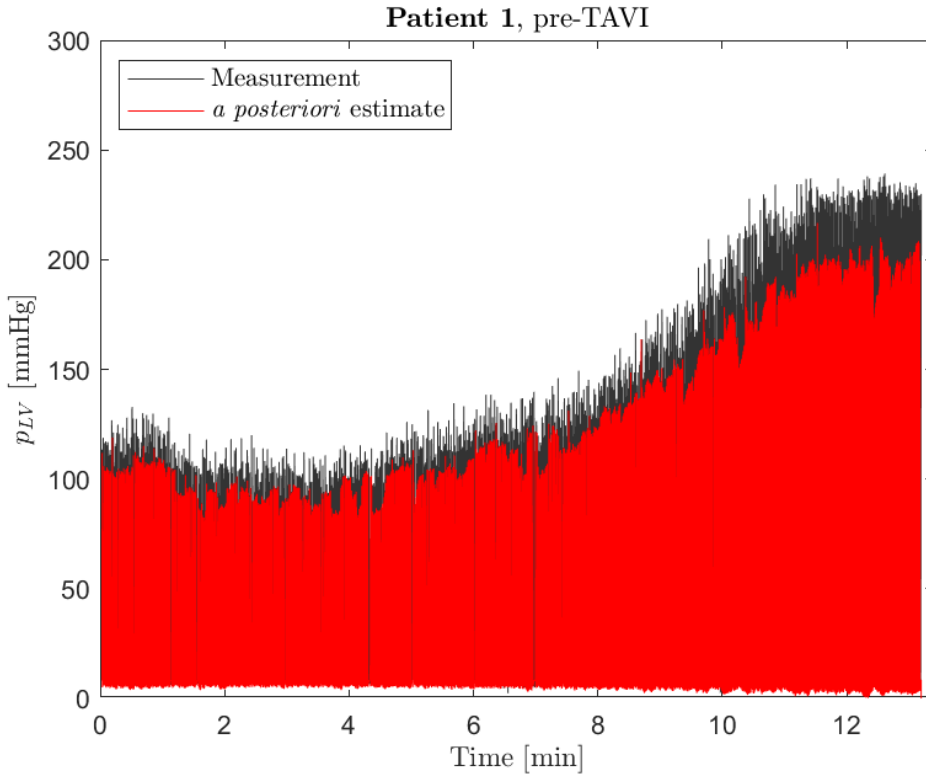
$True\ c_{LV}\ [-]$	$Estimated\ c_{LV}\ [-]$	$CI\ c_{LV}\ [-]$	$RMSE$	$Diff.\ c_{LV}\ [-]$
0.4	0.631	[0.613 , 0.649]	0.003	+0.231
0.4	0.648	[0.636 , 0.660]	0.002	+0.248
0.4	0.650	[0.638 , 0.662]	0.002	+0.250
0.6	0.760	[0.744 , 0.776]	0.003	+0.160
0.6	0.761	[0.745 , 0.770]	0.003	+0.161
0.6	0.762	[0.744 , 0.780]	0.004	+0.162
0.8	0.754	[0.732 , 0.776]	0.006	-0.046
0.8	0.710	[0.676 , 0.774]	0.005	-0.090
0.8	0.758	[0.734 , 0.782]	0.010	-0.042
1.0	0.719	[0.637 , 0.801]	0.016	-0.281
1.0	0.695	[0.655 , 0.735]	0.009	-0.305
1.0	0.640	[0.584 , 0.696]	0.017	-0.360
1.2	0.661	[0.615 , 0.707]	0.010	-0.539
1.2	0.729	[0.655 , 0.803]	0.015	-0.471
1.2	0.632	[0.584 , 0.680]	0.013	-0.568
1.4	0.654	[0.602 , 0.706]	0.012	-0.746
1.4	0.666	[0.612 , 0.720]	0.011	-0.734
1.4	0.597	[0.541 , 0.635]	0.019	-0.803

**Table F.6:** Results of the application of the proposed ROUKF-based algorithm to the in vitro data that were generated by the material twin in **Configuration 6**.

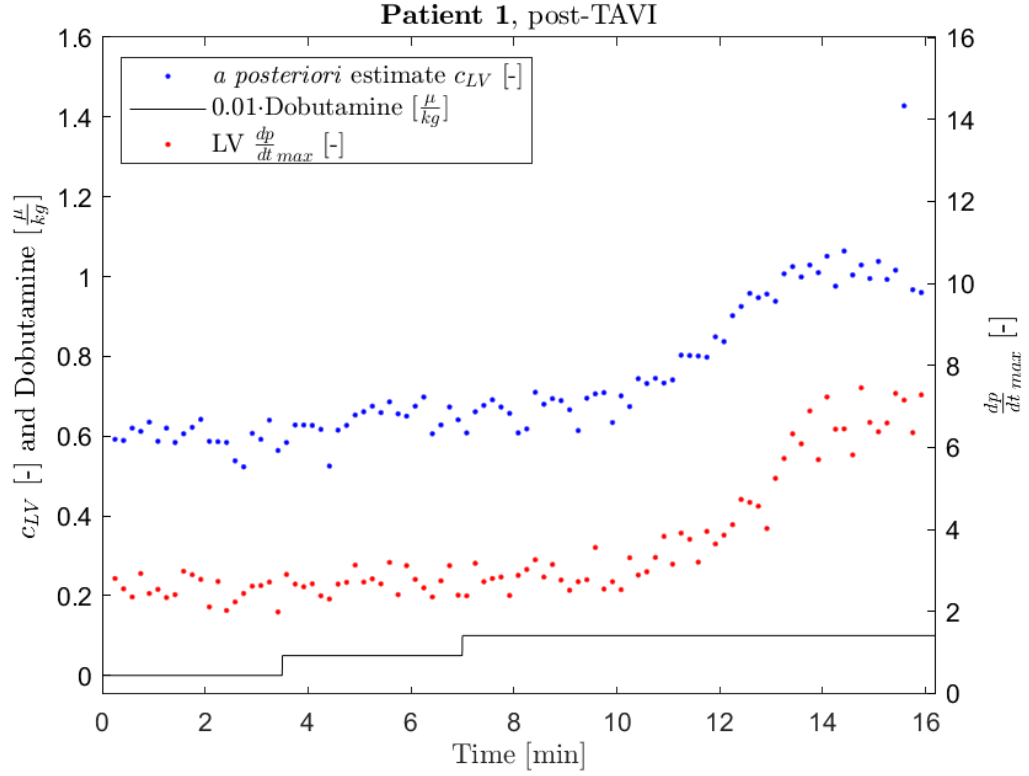
## G. Results in vivo data



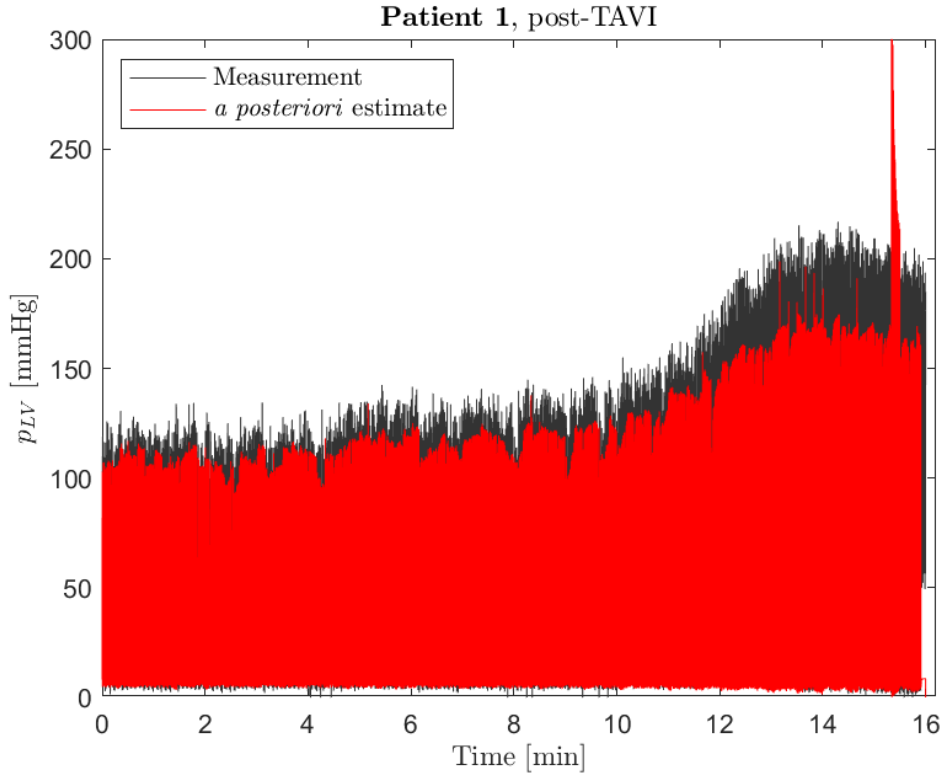
**Figure G.1:** Data belonging to **patient 1**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



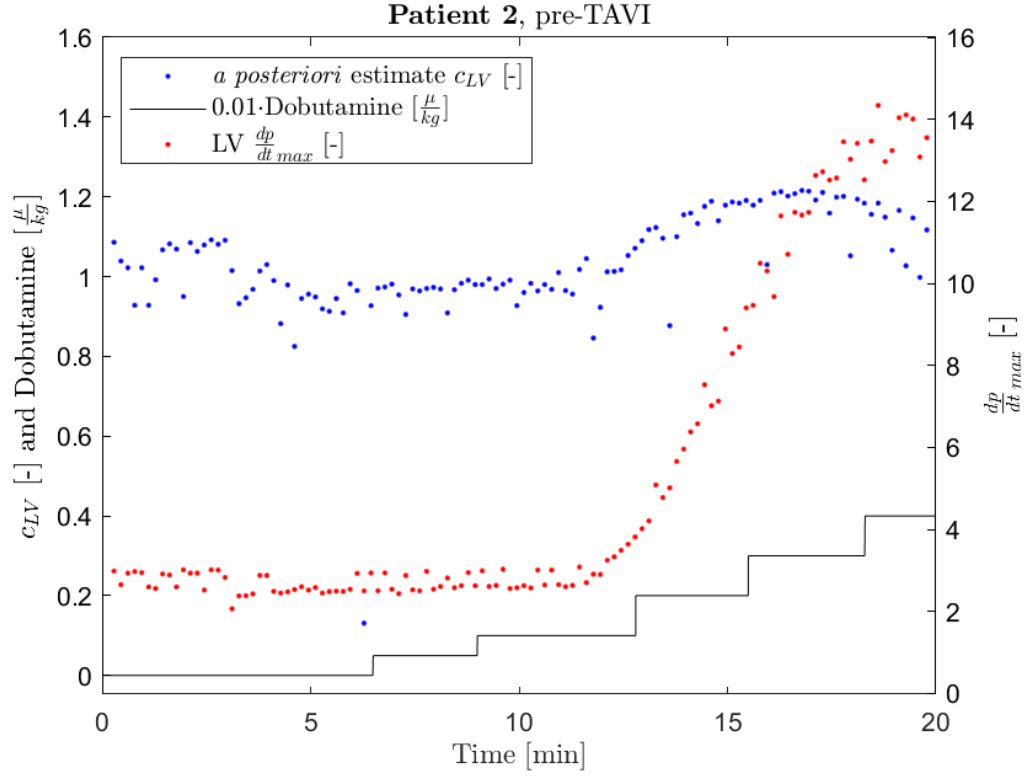
**Figure G.2:** Data belonging to **patient 1**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the  $a posteriori$  estimate of the  $p_{LV}$ .



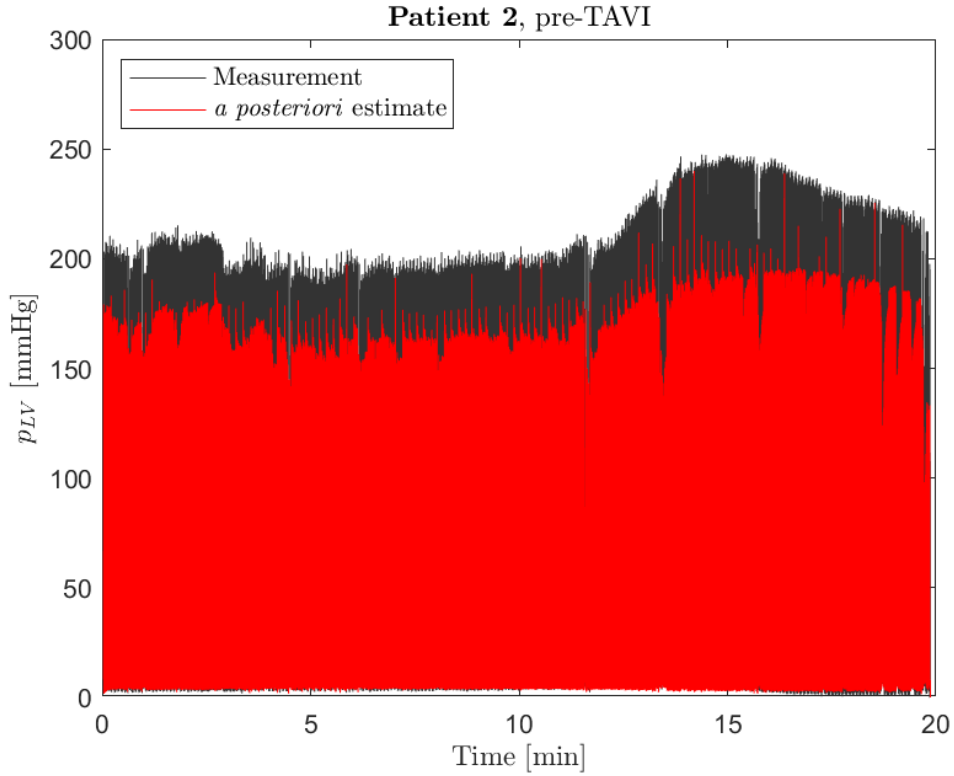
**Figure G.3:** Data belonging to **patient 1**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



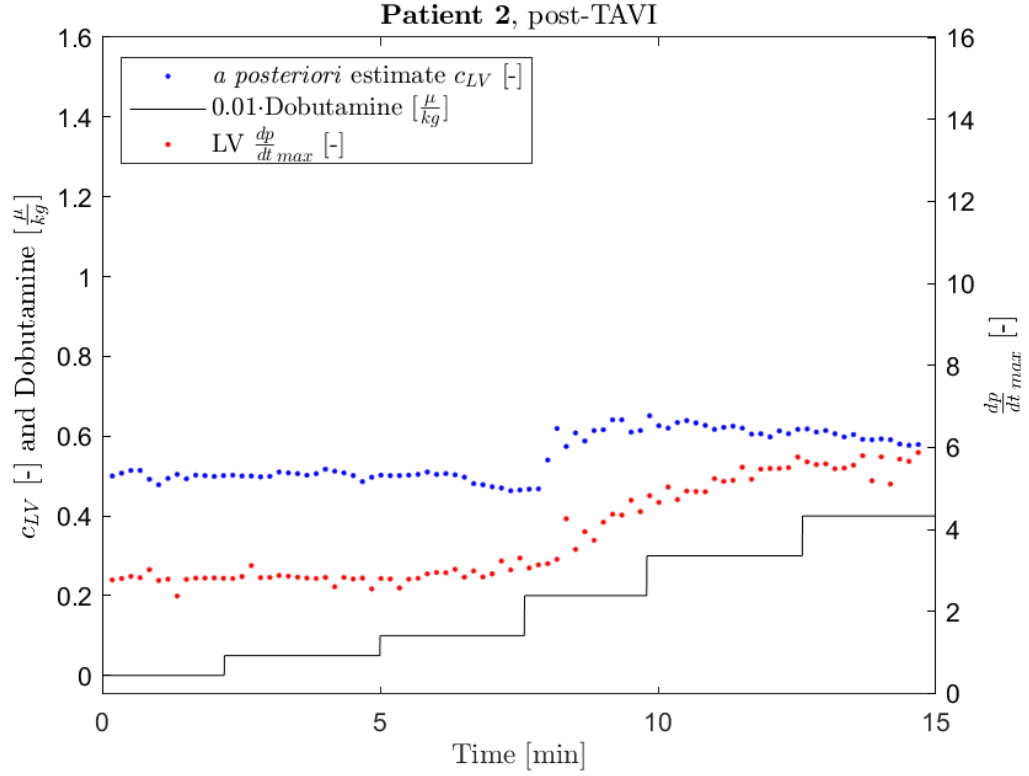
**Figure G.4:** Data belonging to **patient 1**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



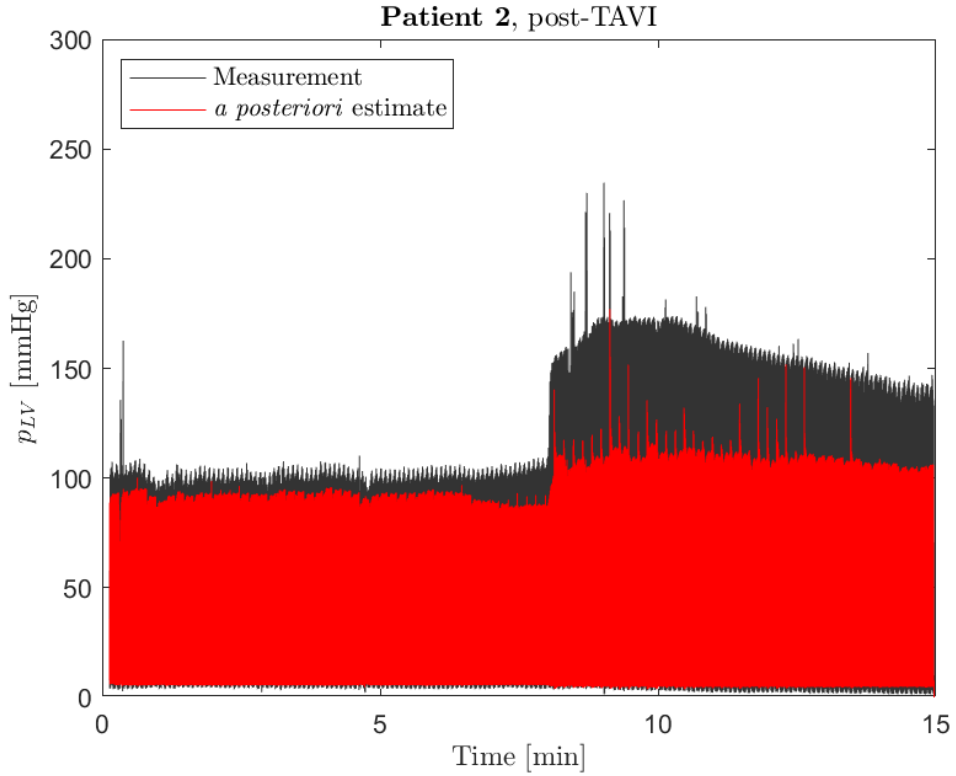
**Figure G.5:** Data belonging to **patient 2**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



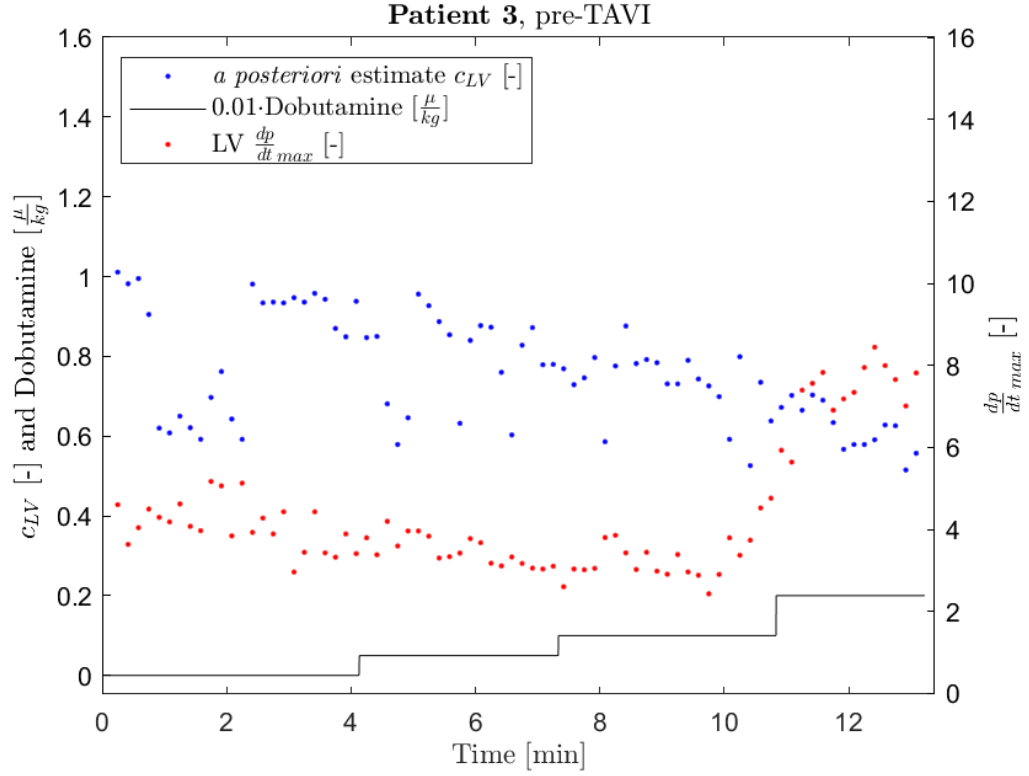
**Figure G.6:** Data belonging to **patient 2**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



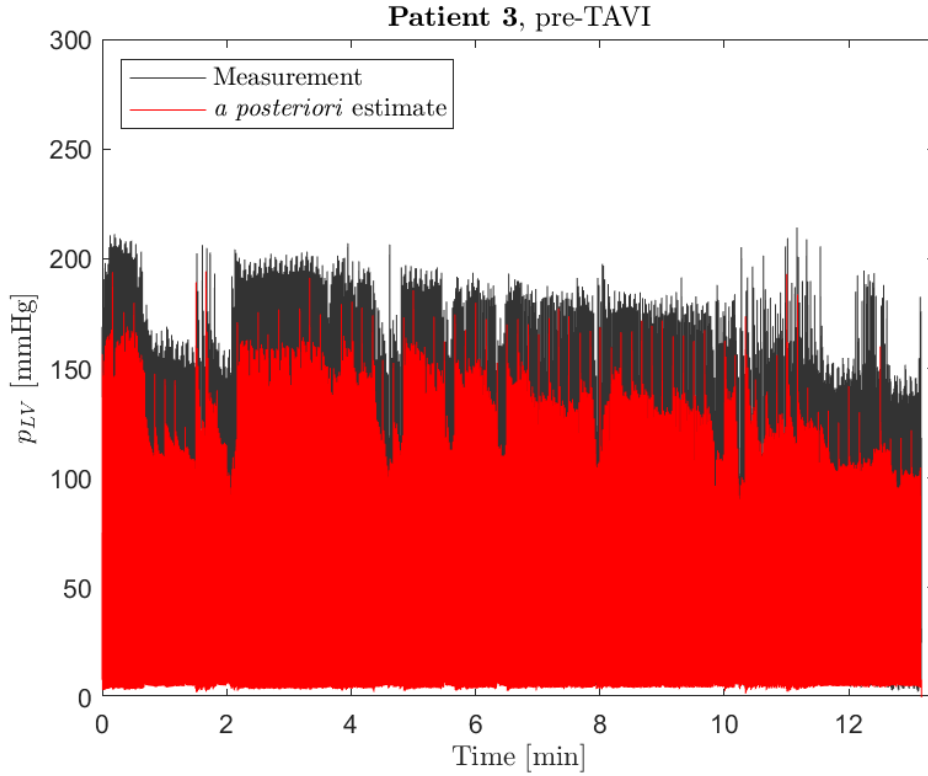
**Figure G.7:** Data belonging to **patient 2**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



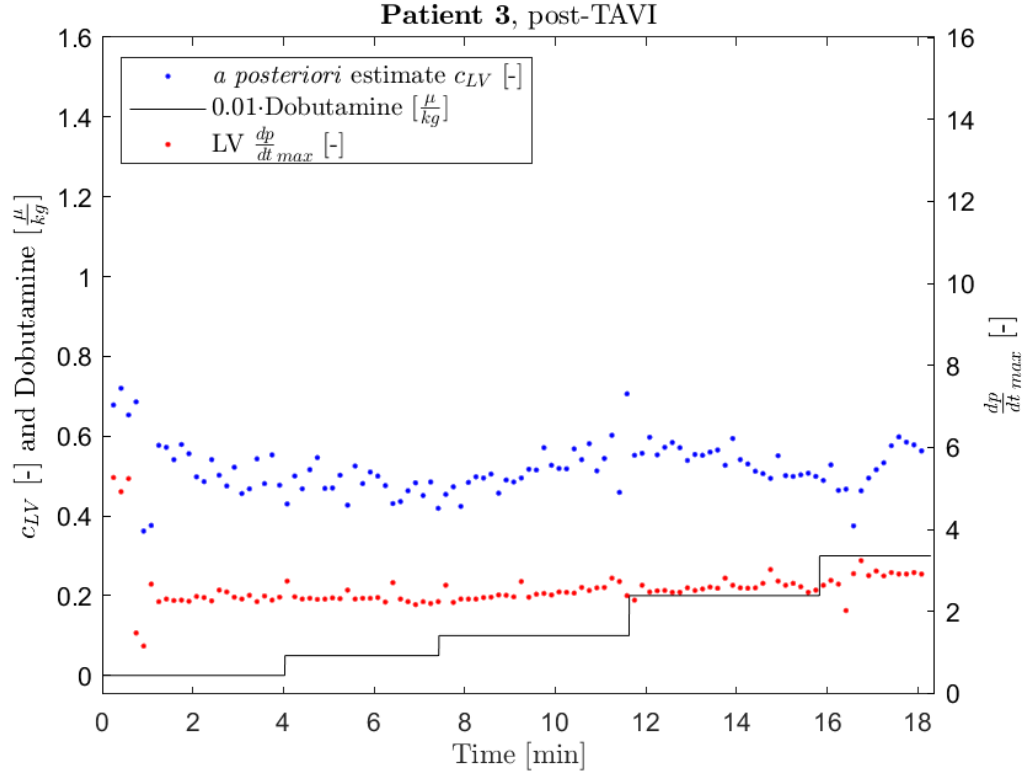
**Figure G.8:** Data belonging to **patient 2**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



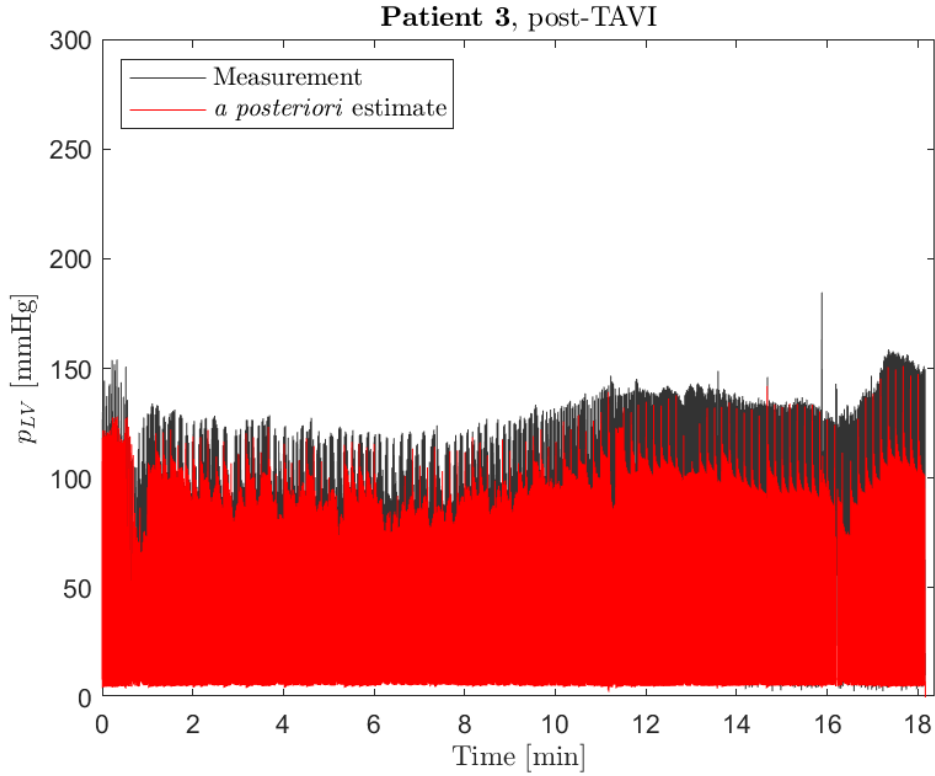
**Figure G.9:** Data belonging to **patient 3**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



**Figure G.10:** Data belonging to **patient 3**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .

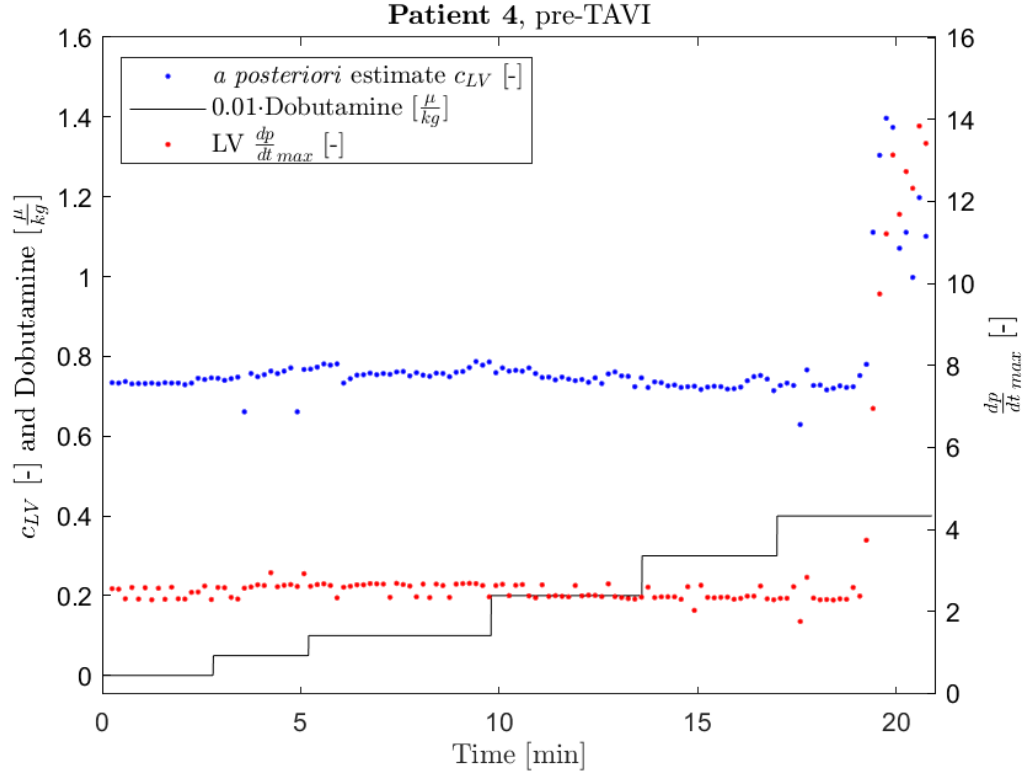


**Figure G.11:** Data belonging to **patient 3**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.

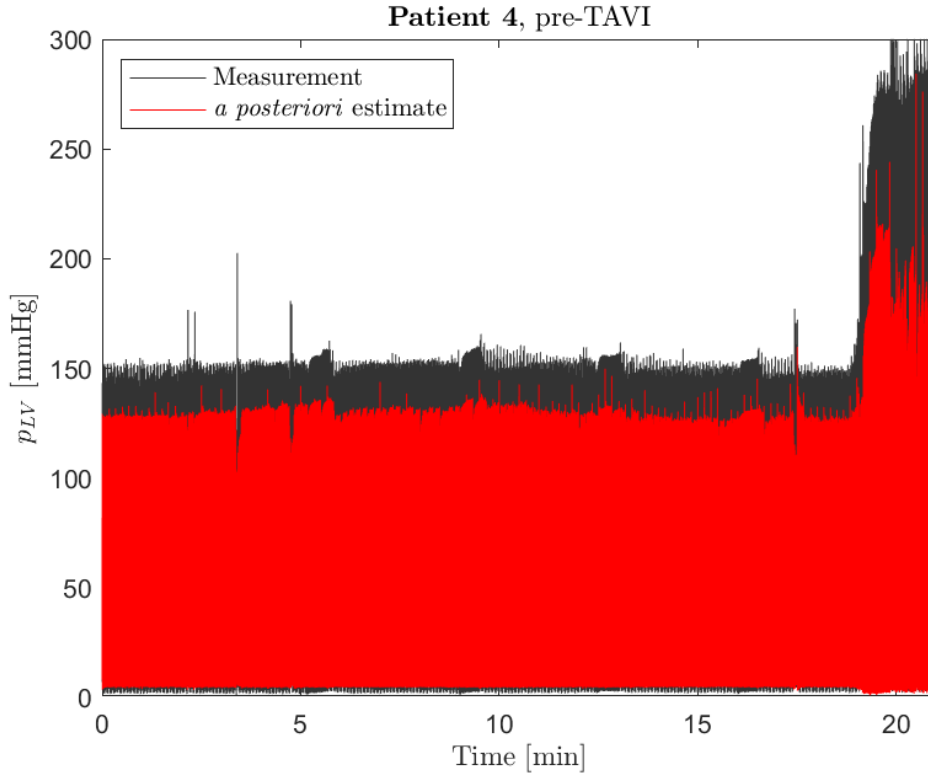


**Figure G.12:** Data belonging to **patient 3**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .

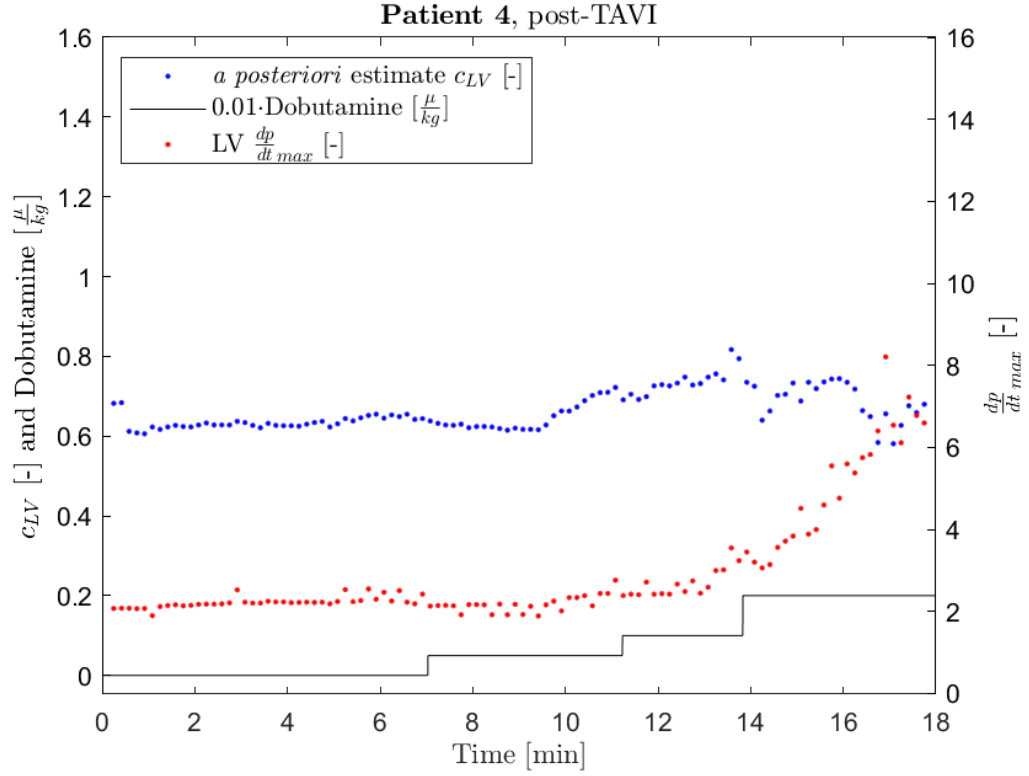




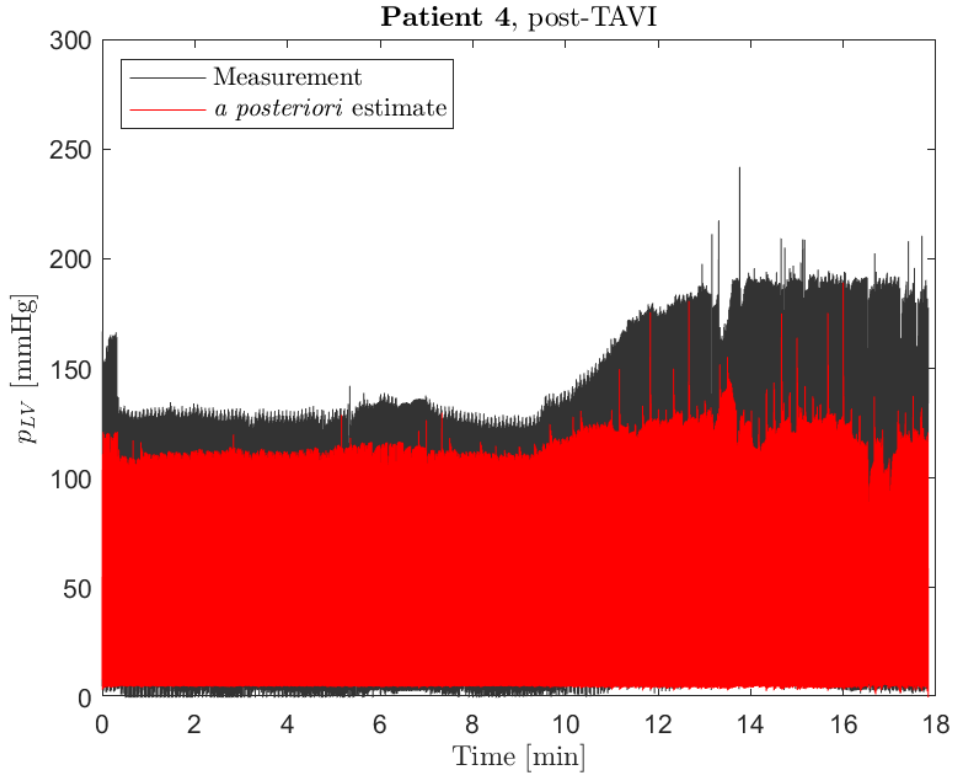
**Figure G.13:** Data belonging to **patient 4**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



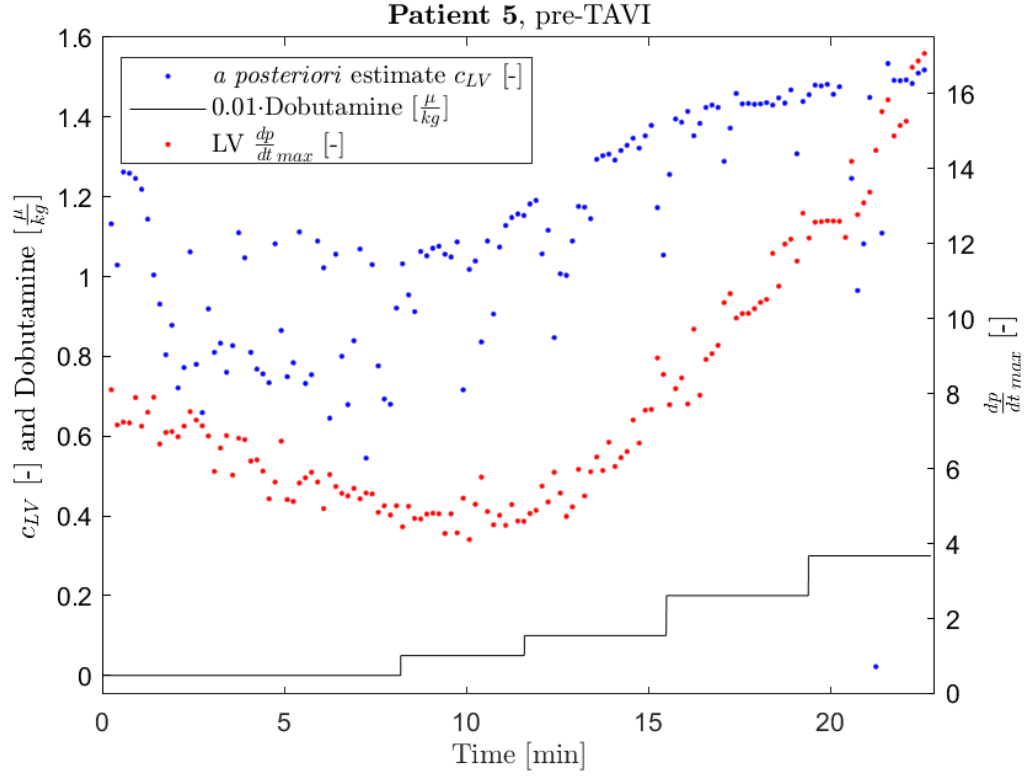
**Figure G.14:** Data belonging to **patient 4**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



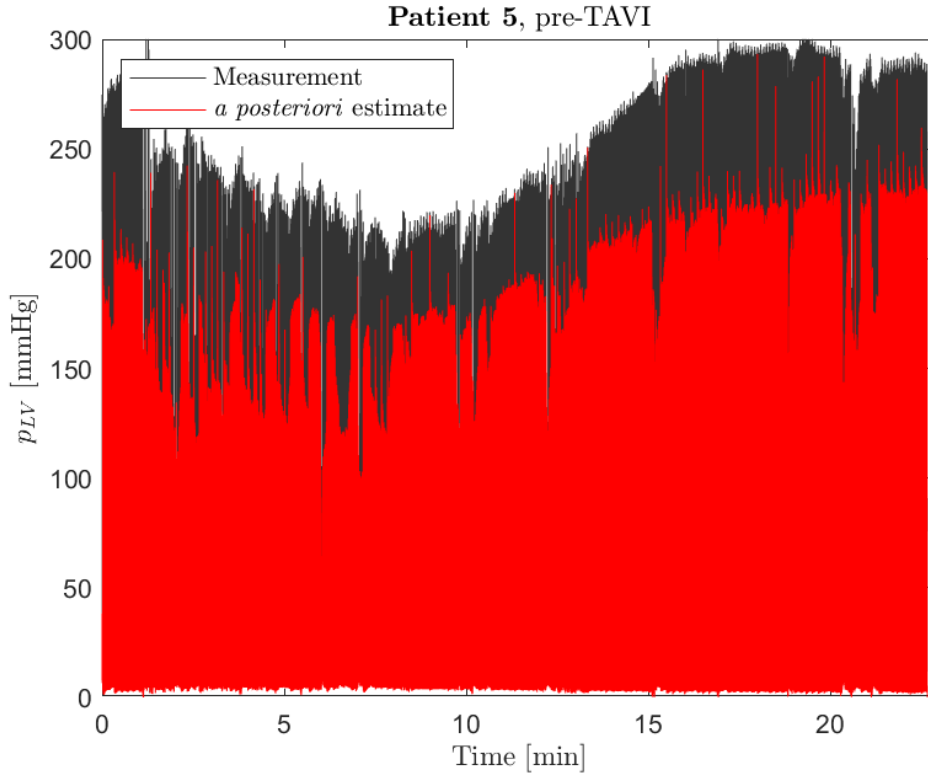
**Figure G.15:** Data belonging to **patient 4**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



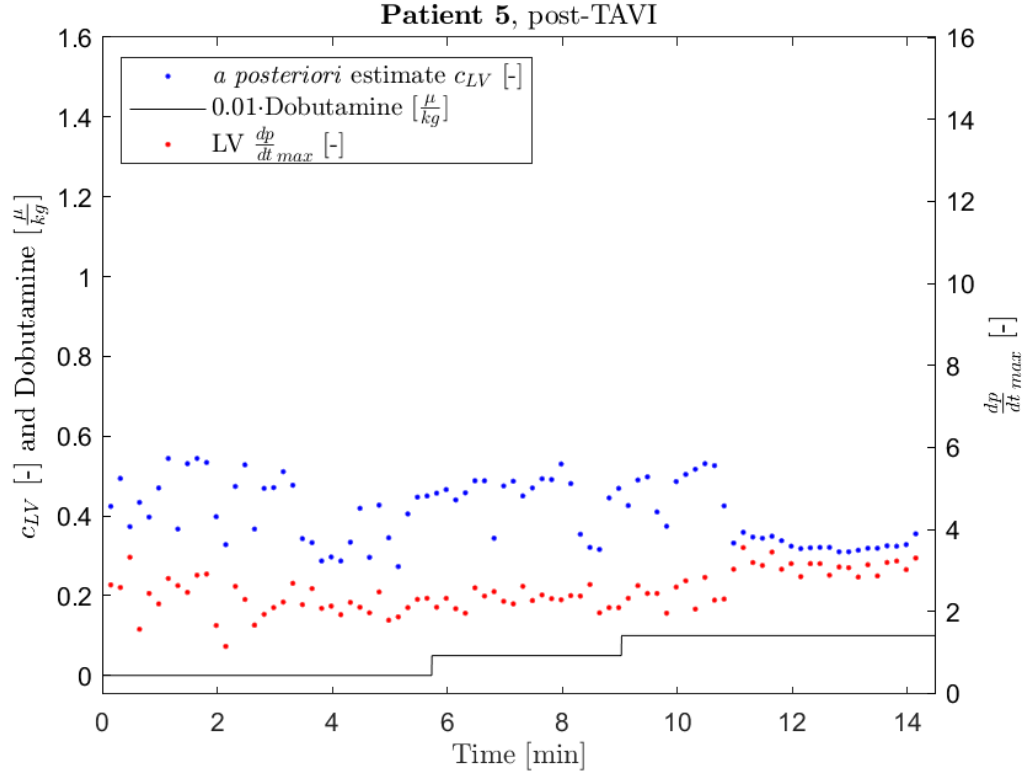
**Figure G.16:** Data belonging to **patient 4**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the  $a posteriori$  estimate of the  $p_{LV}$ .



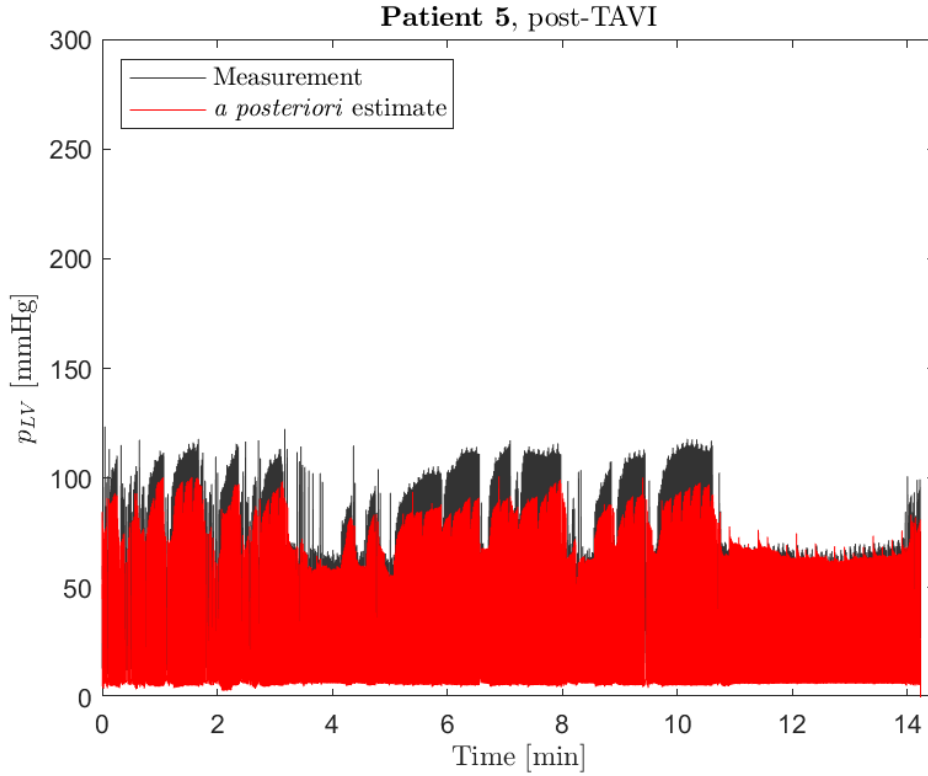
**Figure G.17:** Data belonging to **patient 5**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



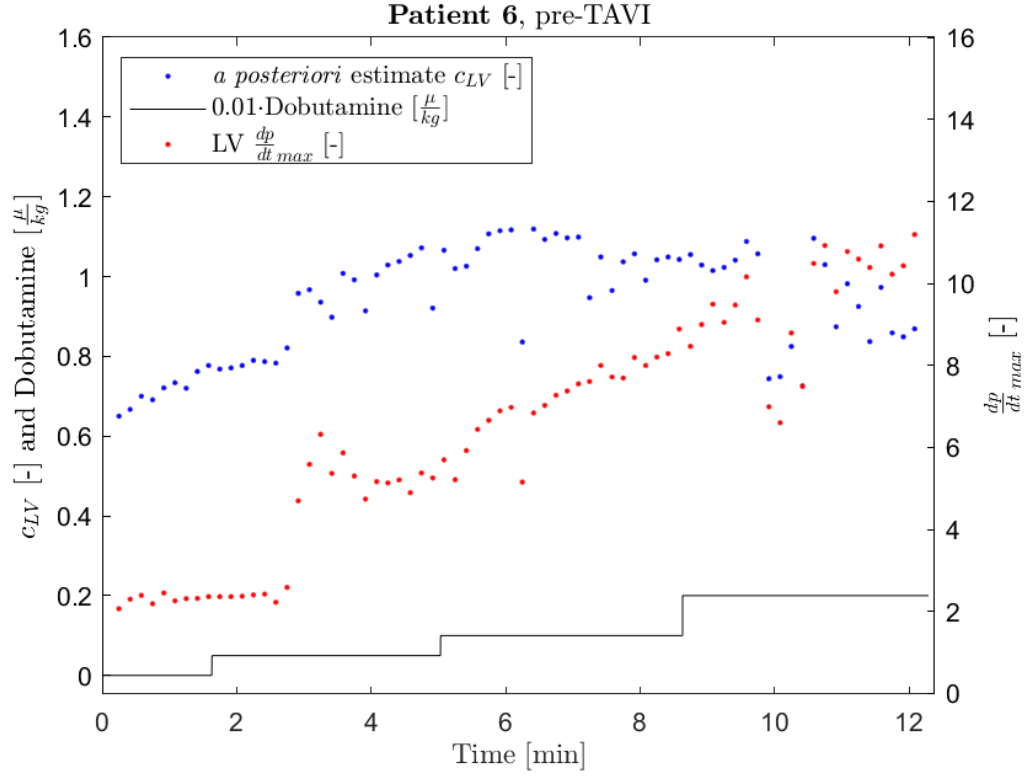
**Figure G.18:** Data belonging to **patient 5**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



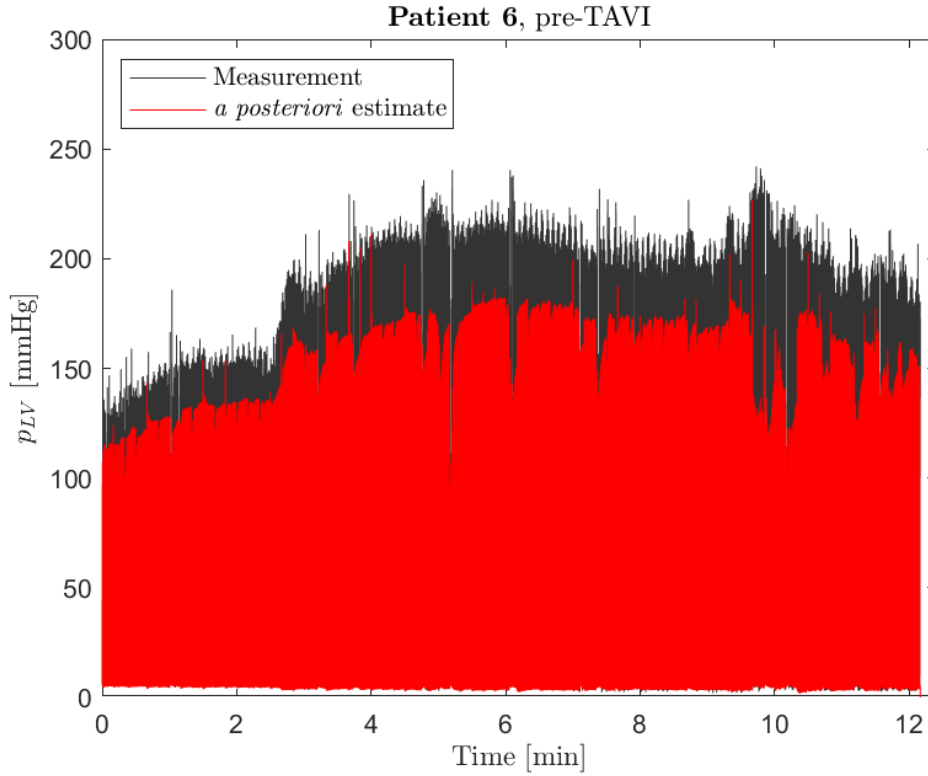
**Figure G.19:** Data belonging to **patient 5**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



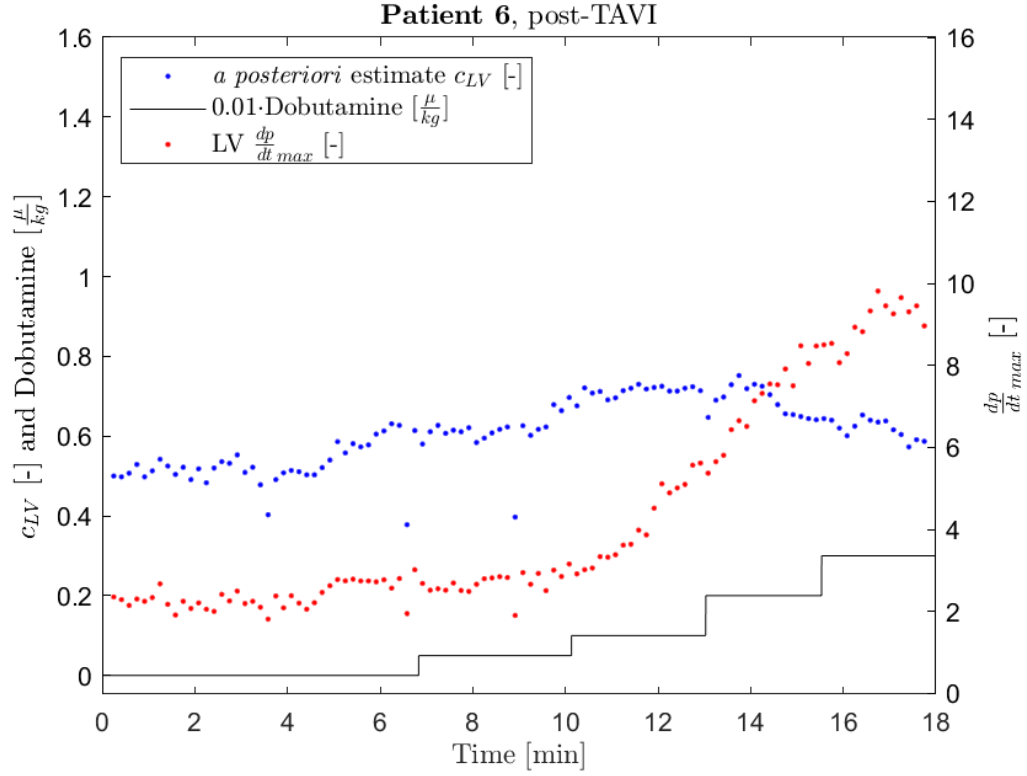
**Figure G.20:** Data belonging to **patient 5**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



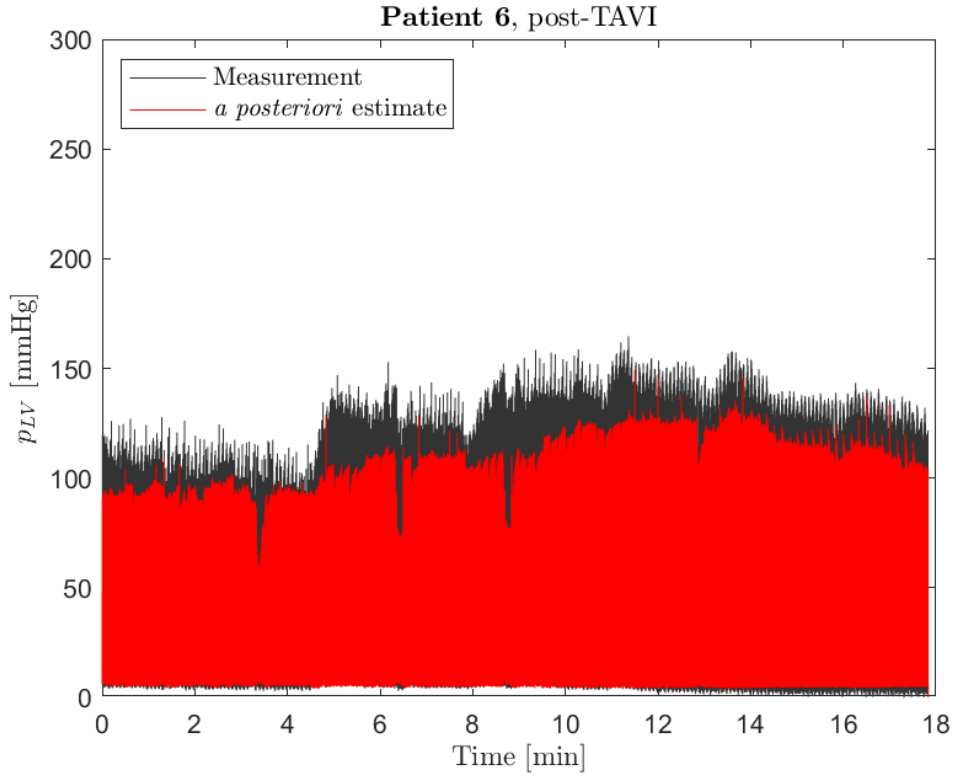
**Figure G.21:** Data belonging to **patient 6**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



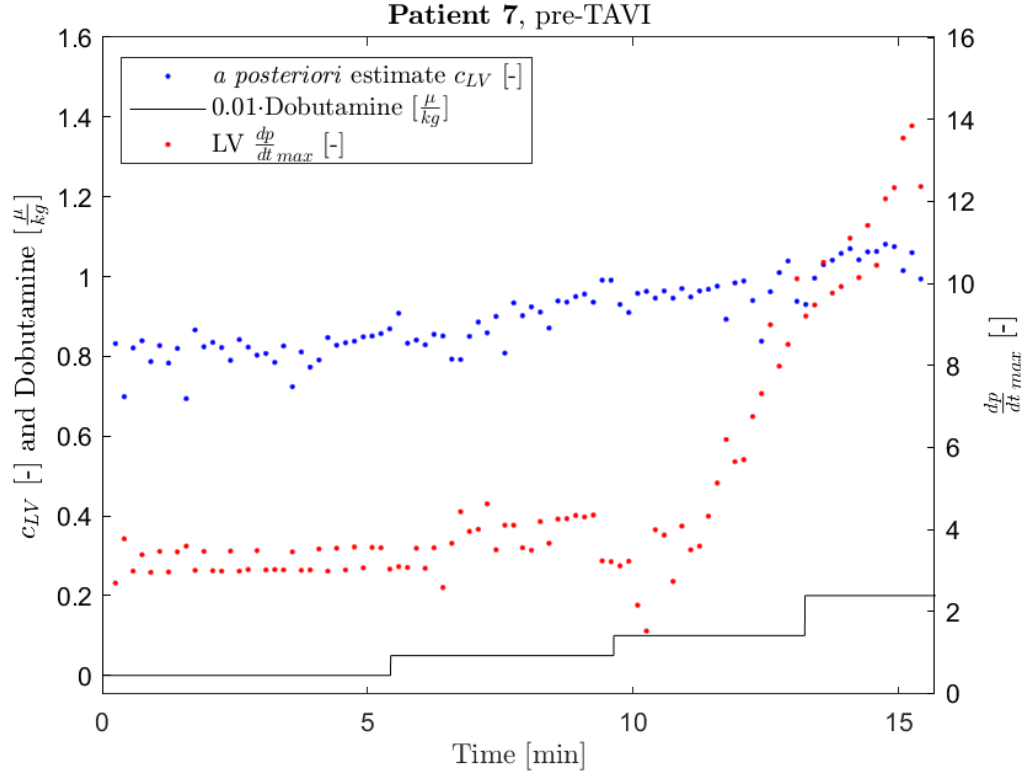
**Figure G.22:** Data belonging to **patient 6**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



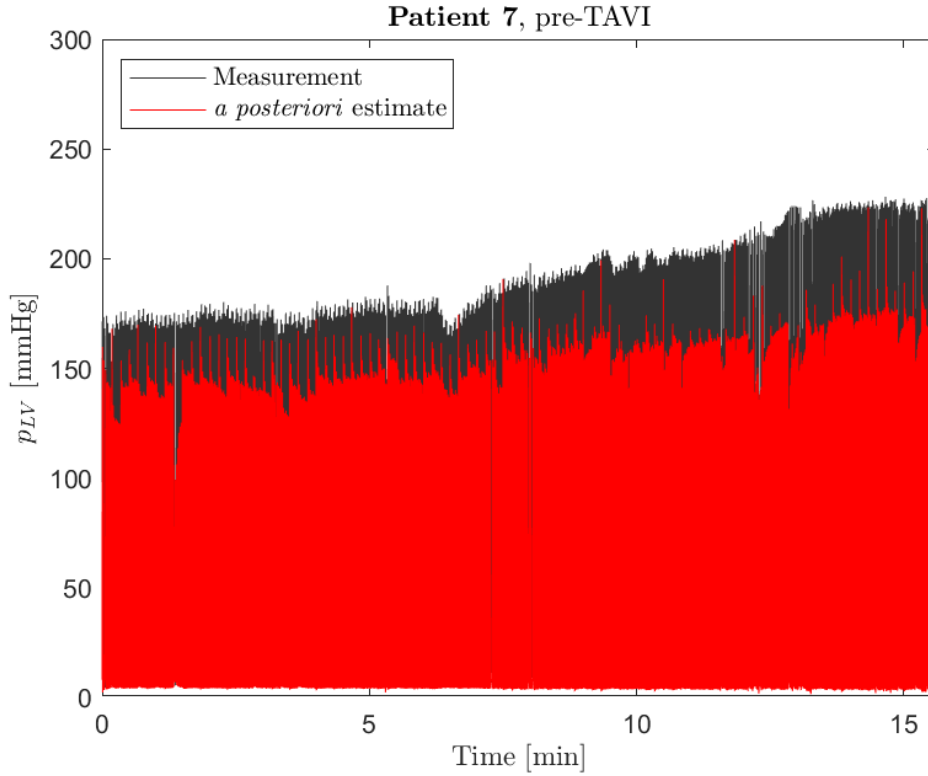
**Figure G.23:** Data belonging to **patient 6**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



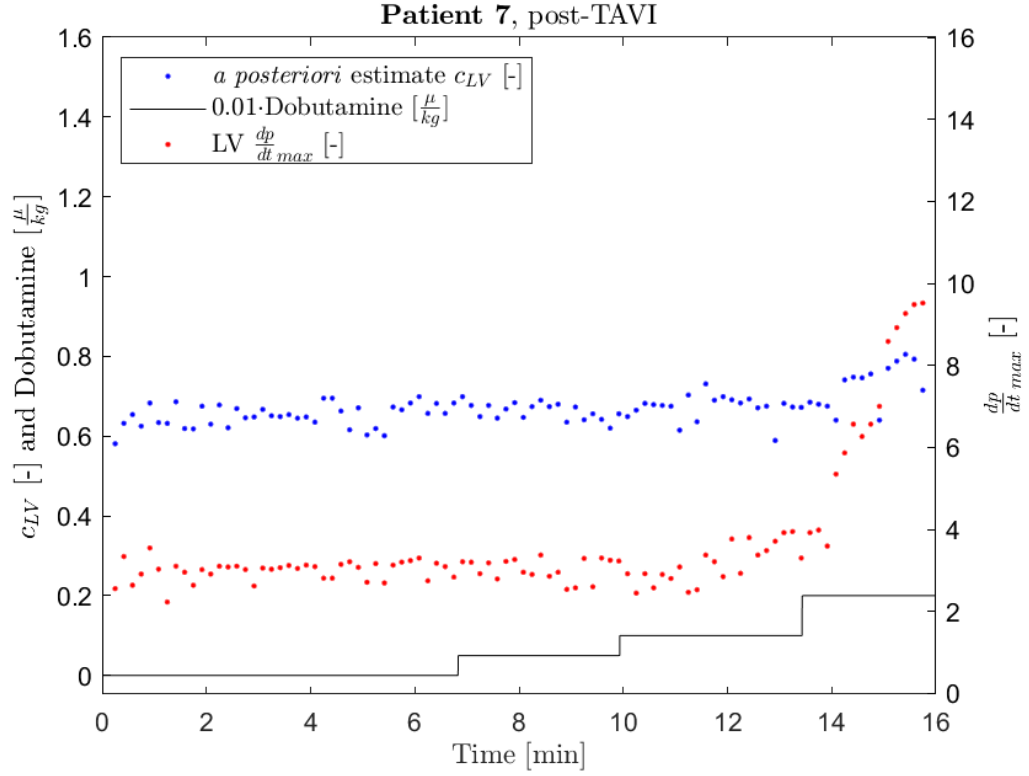
**Figure G.24:** Data belonging to **patient 6**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the  $a posteriori$  estimate of the  $p_{LV}$ .



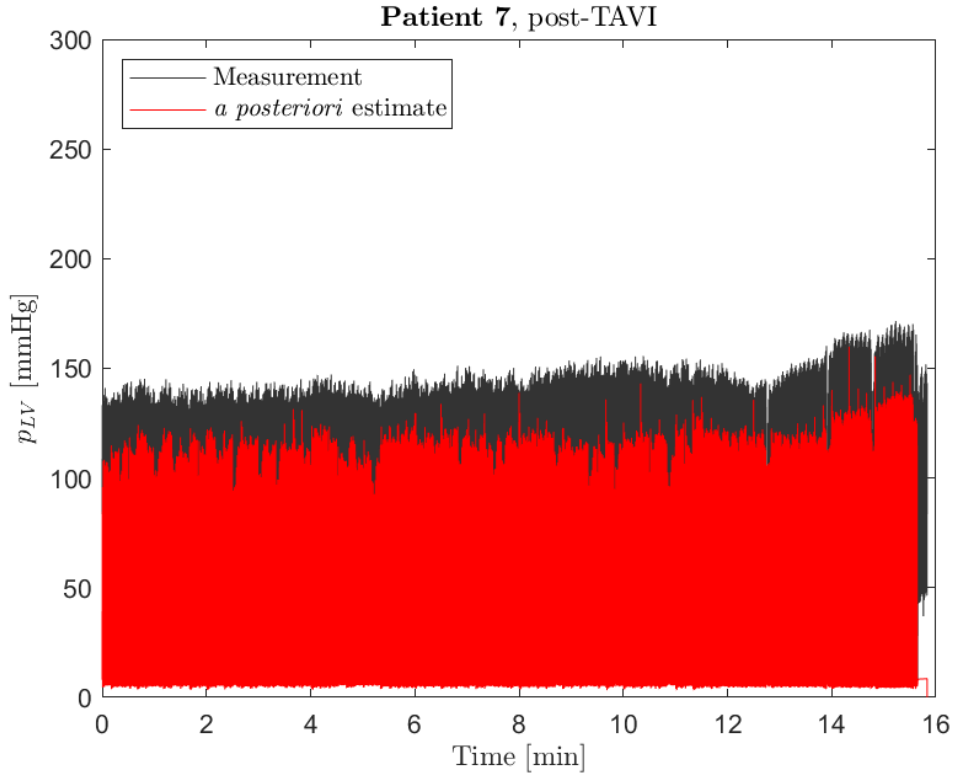
**Figure G.25:** Data belonging to **patient 7**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



**Figure G.26:** Data belonging to **patient 7**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the  $a posteriori$  estimate of the  $p_{LV}$ .

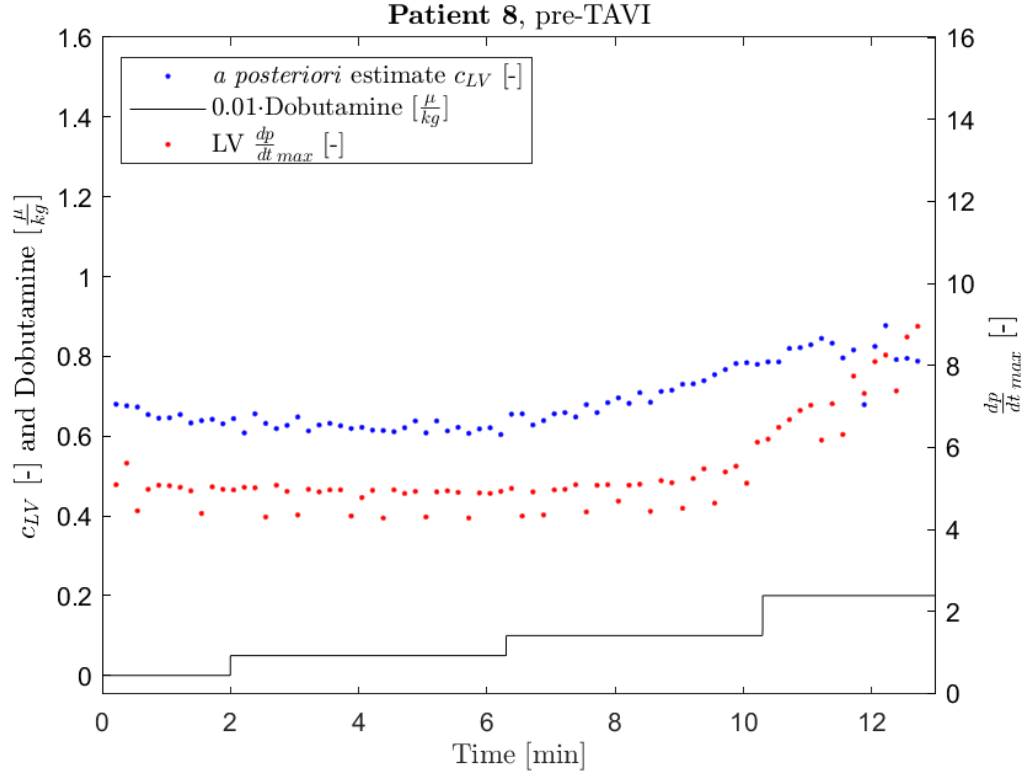


**Figure G.27:** Data belonging to **patient 7**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.

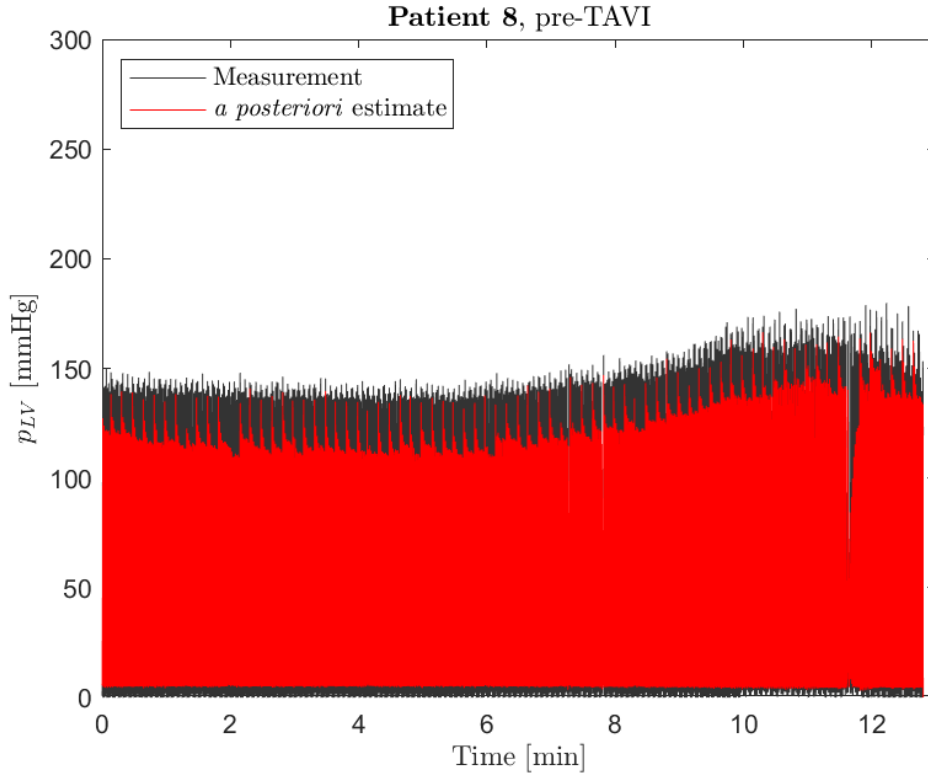


**Figure G.28:** Data belonging to **patient 7**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .

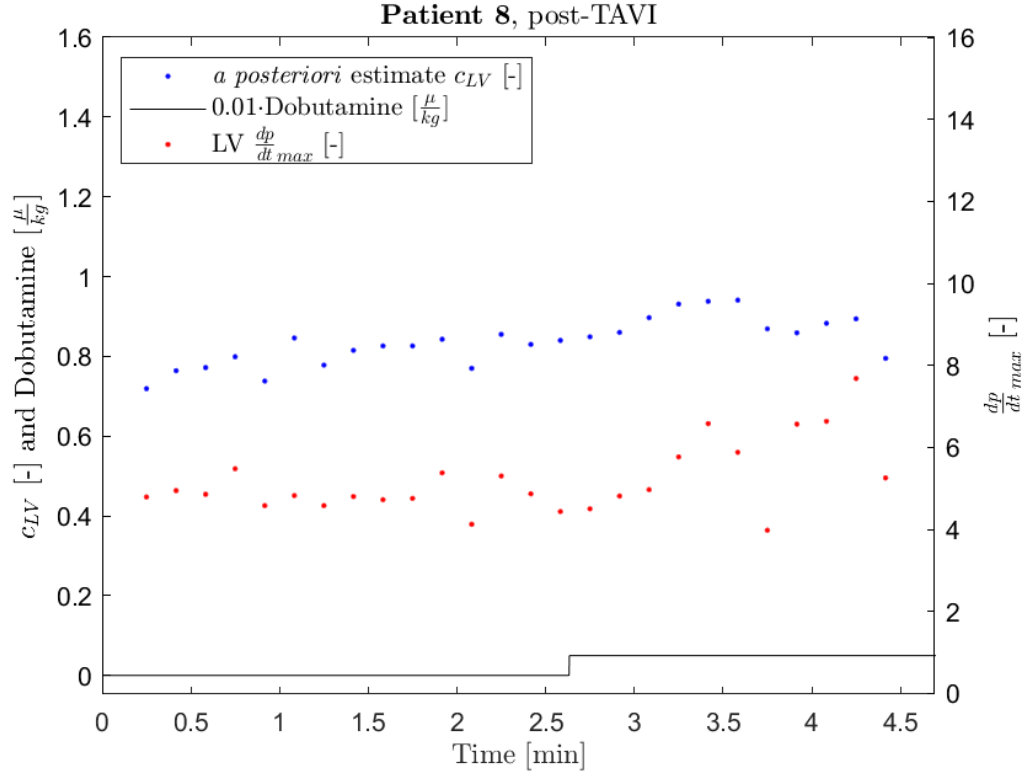




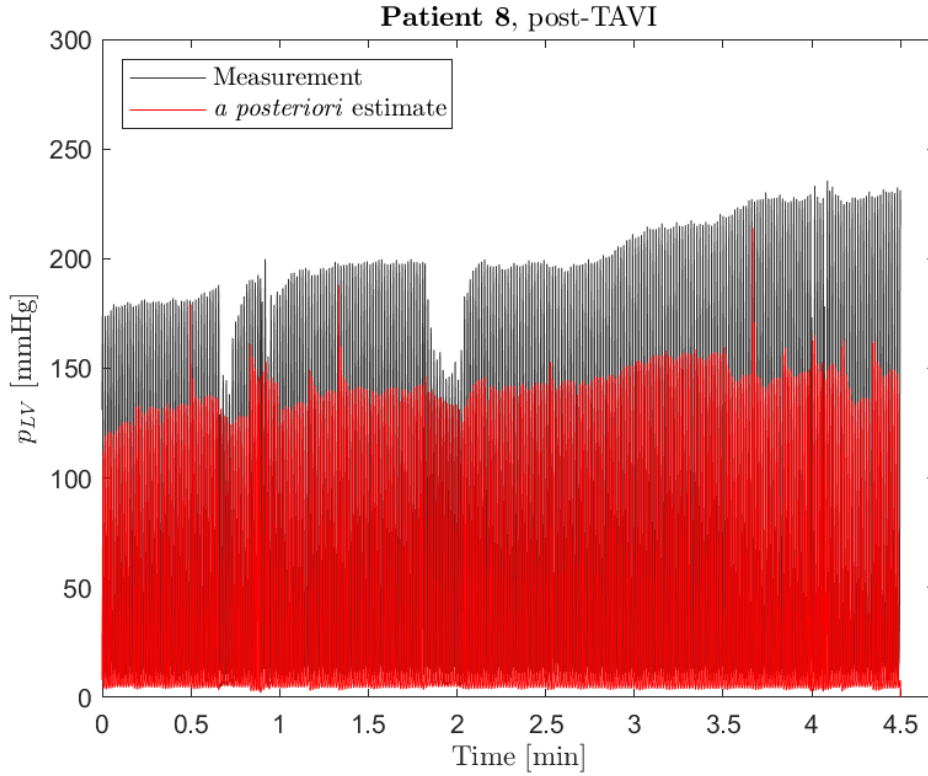
**Figure G.29:** Data belonging to **patient 8**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



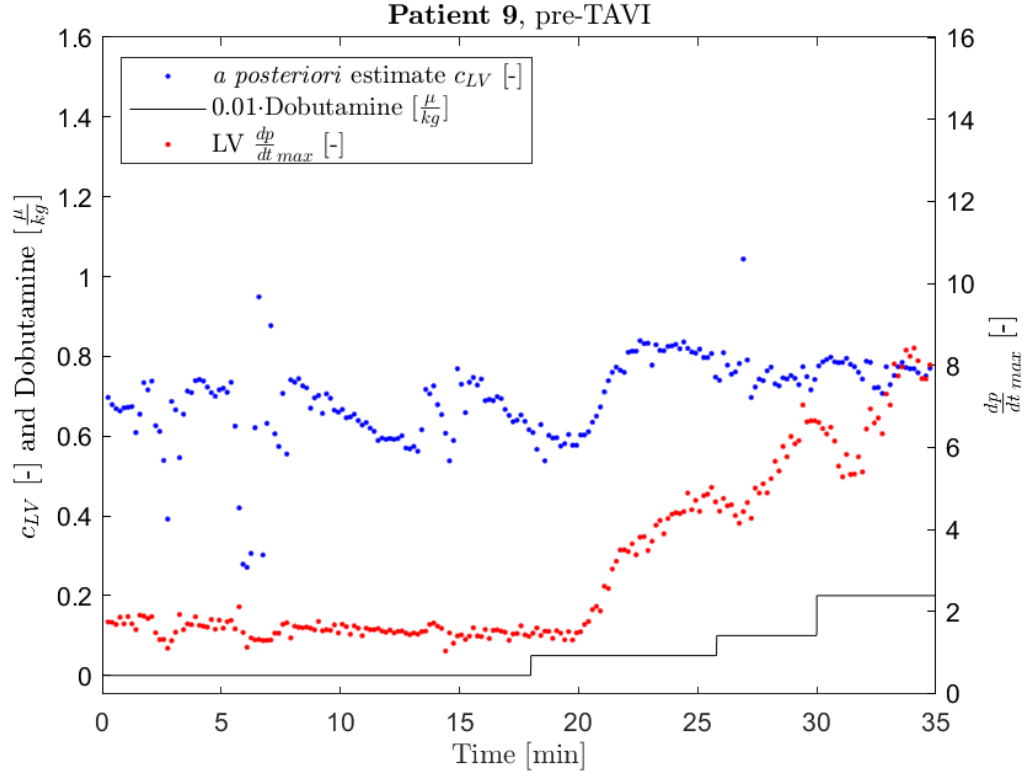
**Figure G.30:** Data belonging to **patient 8**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



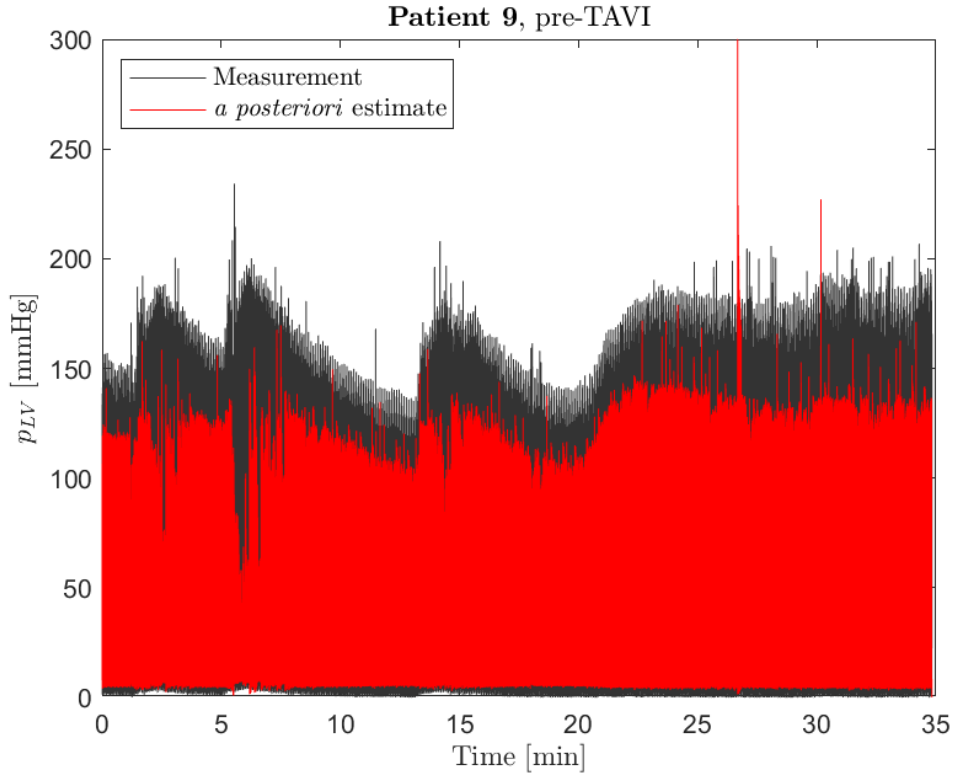
**Figure G.31:** Data belonging to **patient 8**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



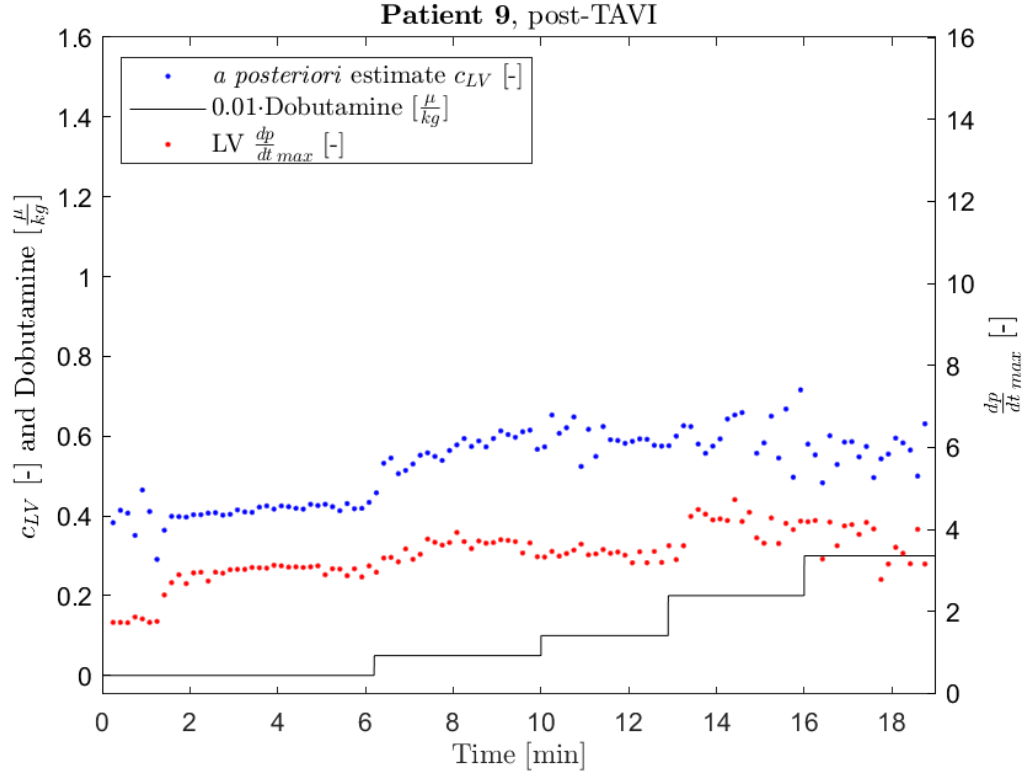
**Figure G.32:** Data belonging to **patient 8**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



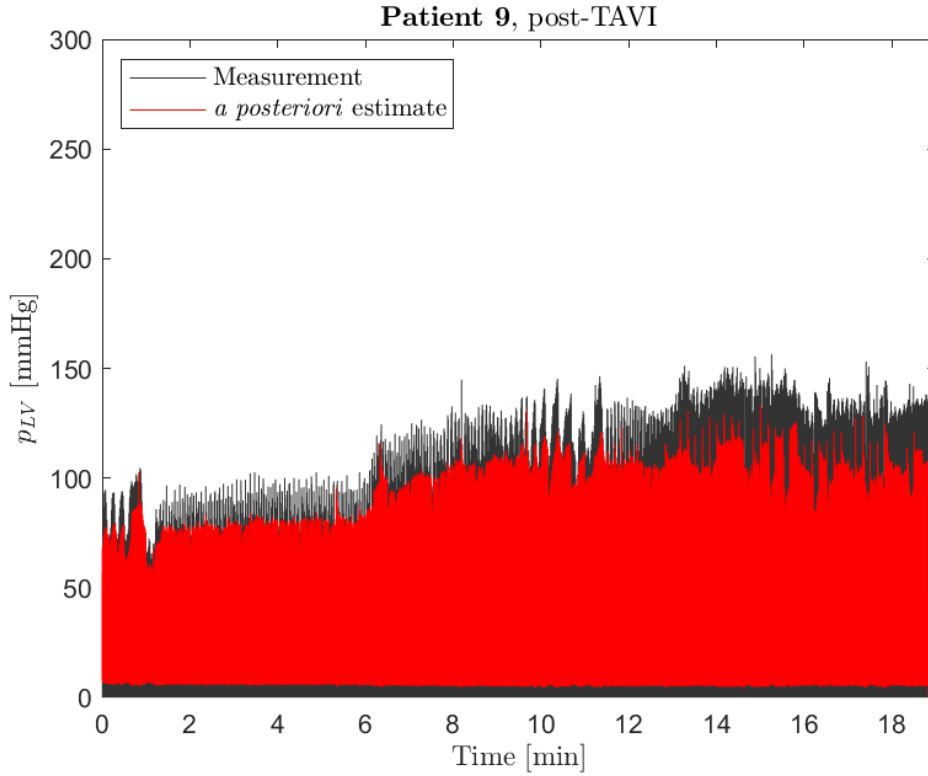
**Figure G.33:** Data belonging to **patient 9**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



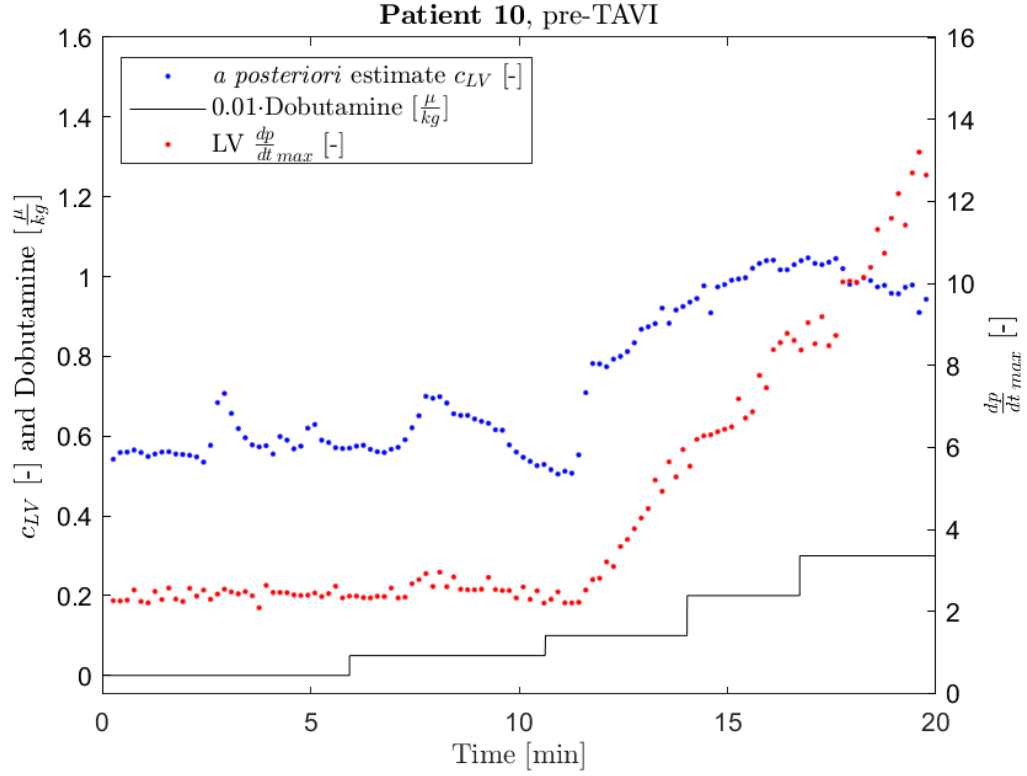
**Figure G.34:** Data belonging to **patient 9**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



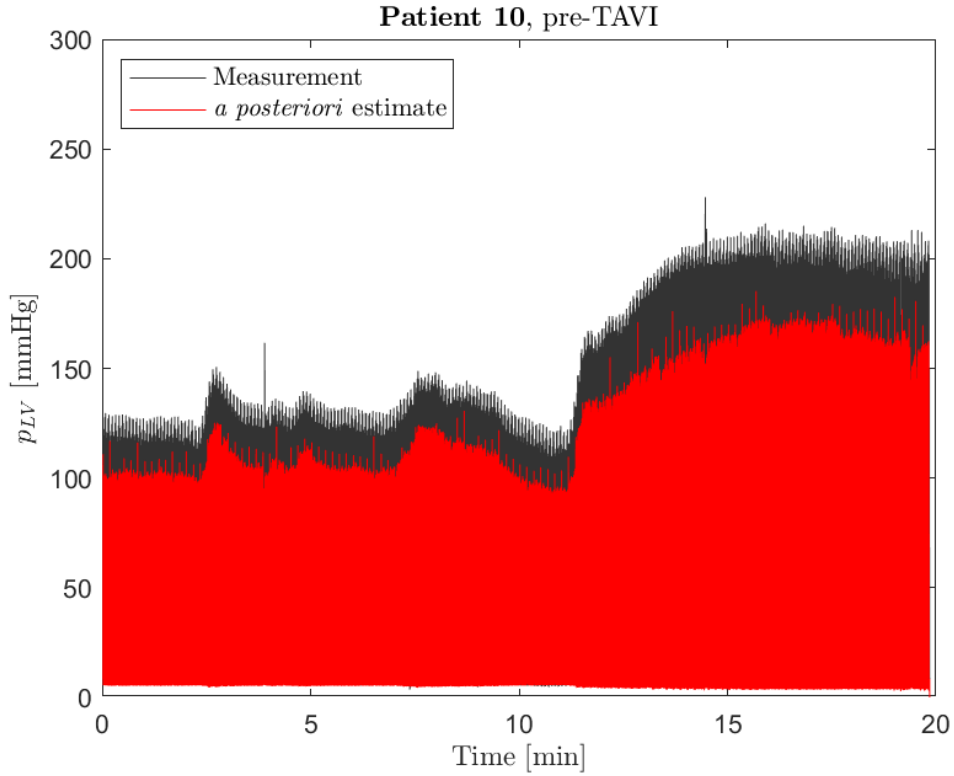
**Figure G.35:** Data belonging to **patient 9**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



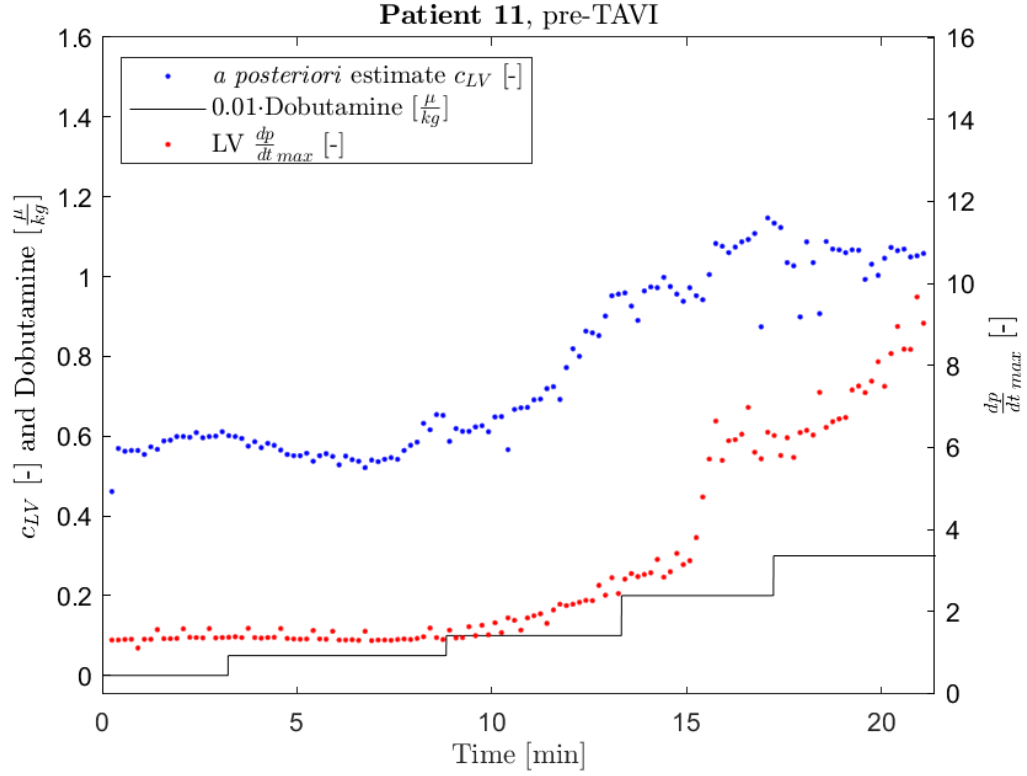
**Figure G.36:** Data belonging to **patient 9**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



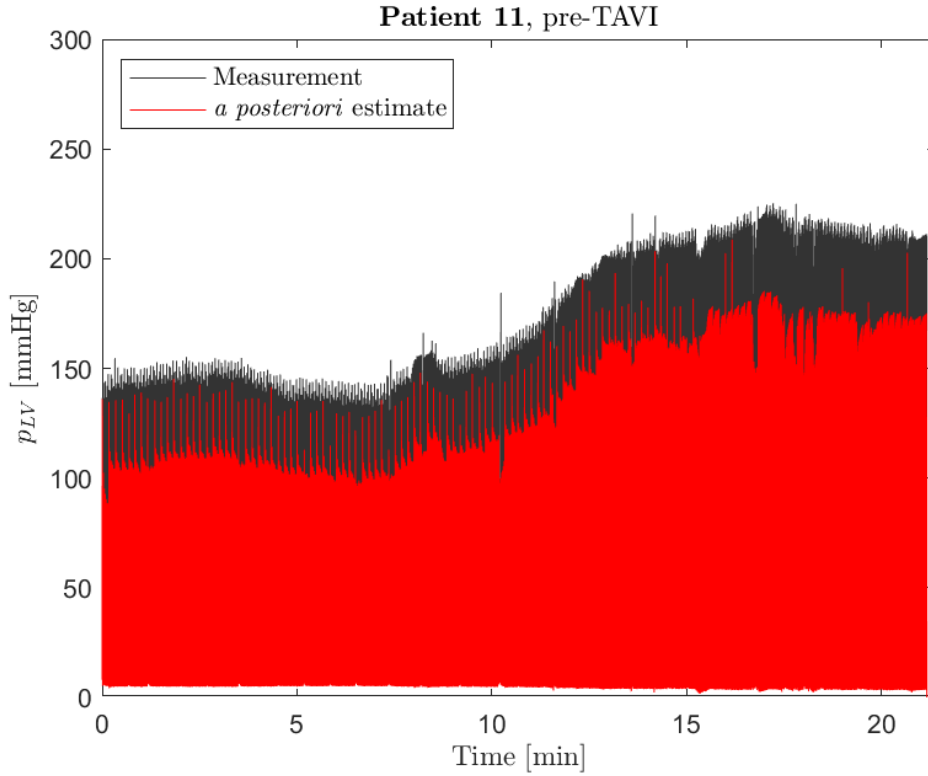
**Figure G.37:** Data belonging to **patient 10**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



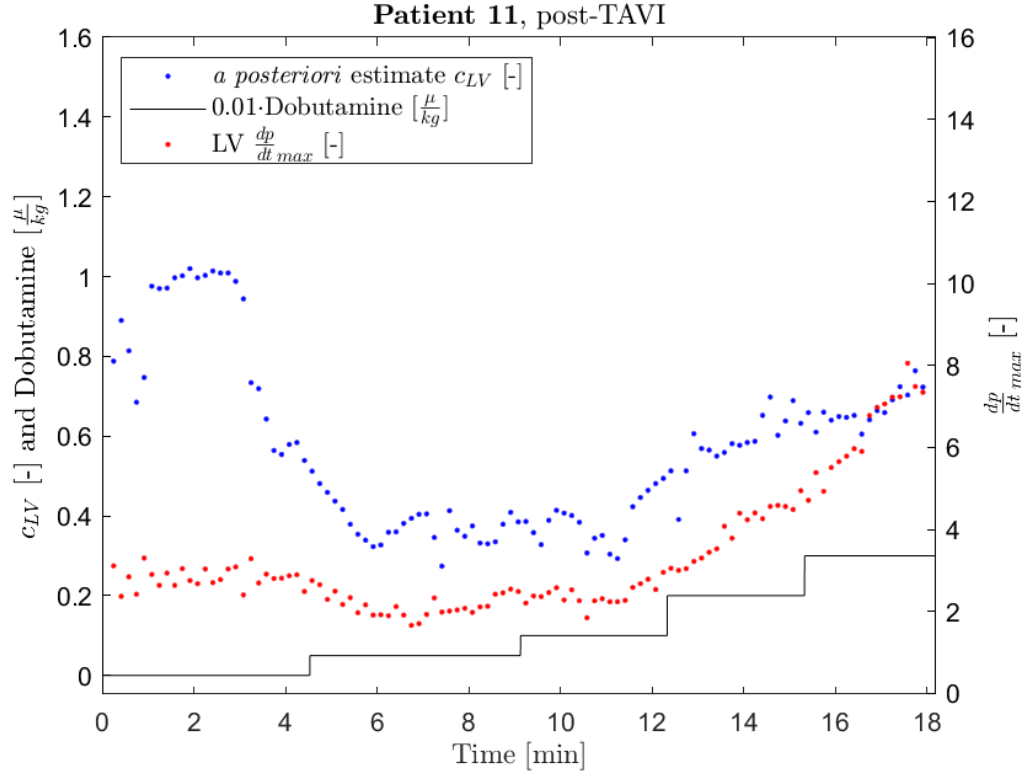
**Figure G.38:** Data belonging to **patient 10**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the  $a posteriori$  estimate of the  $p_{LV}$ .



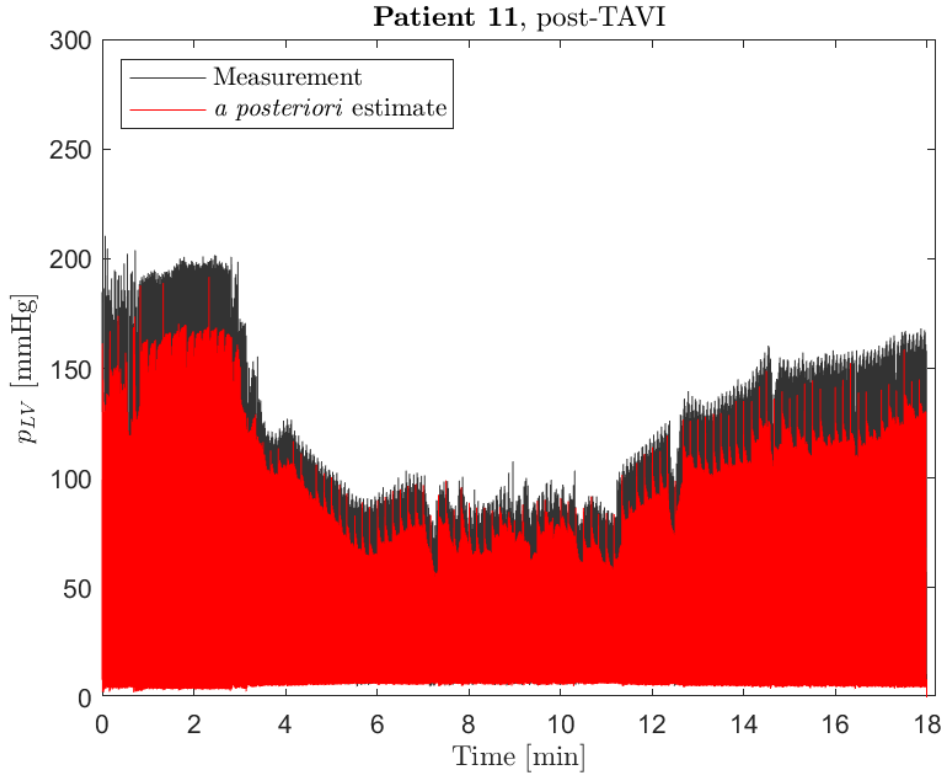
**Figure G.39:** Data belonging to **patient 11**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



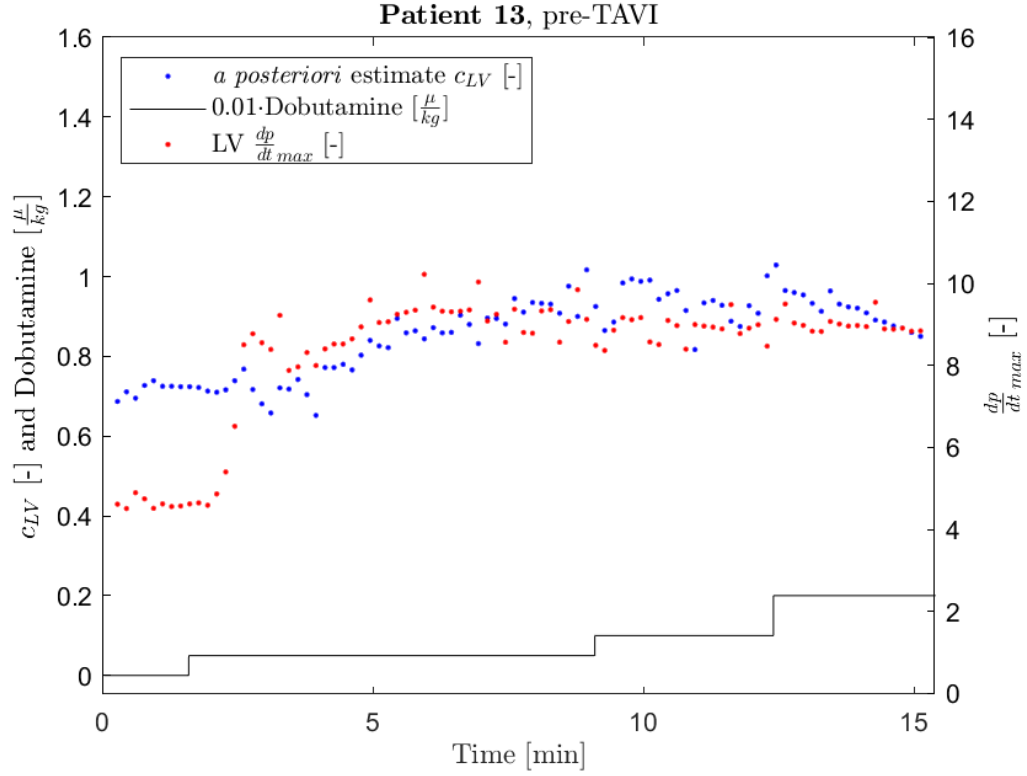
**Figure G.40:** Data belonging to **patient 11**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



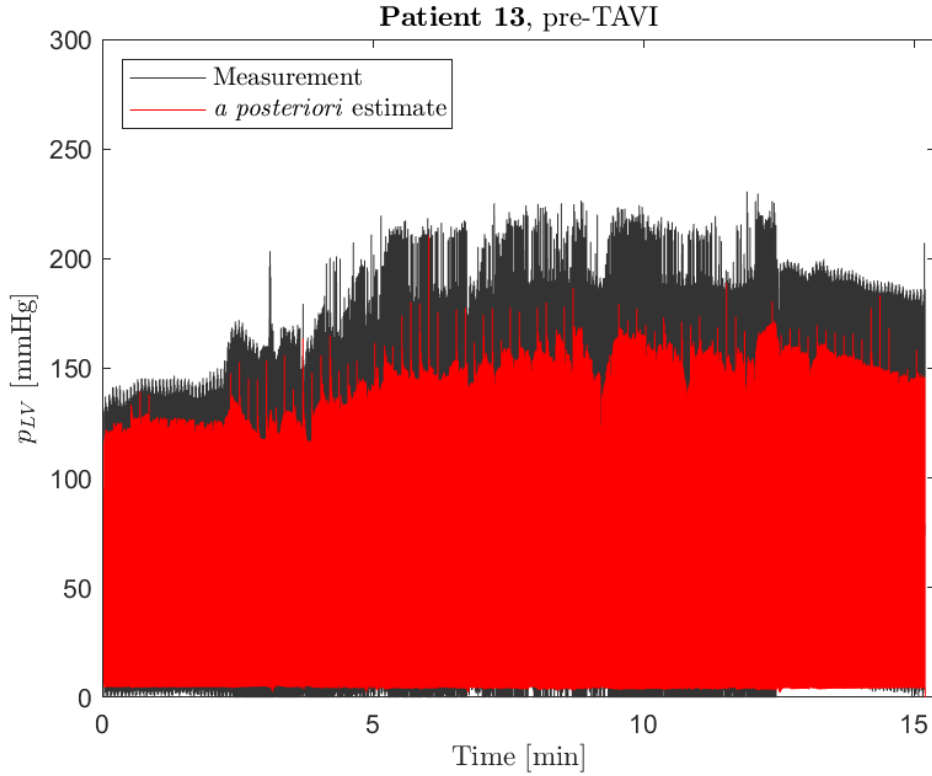
**Figure G.41:** Data belonging to **patient 11**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



**Figure G.42:** Data belonging to **patient 11**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .

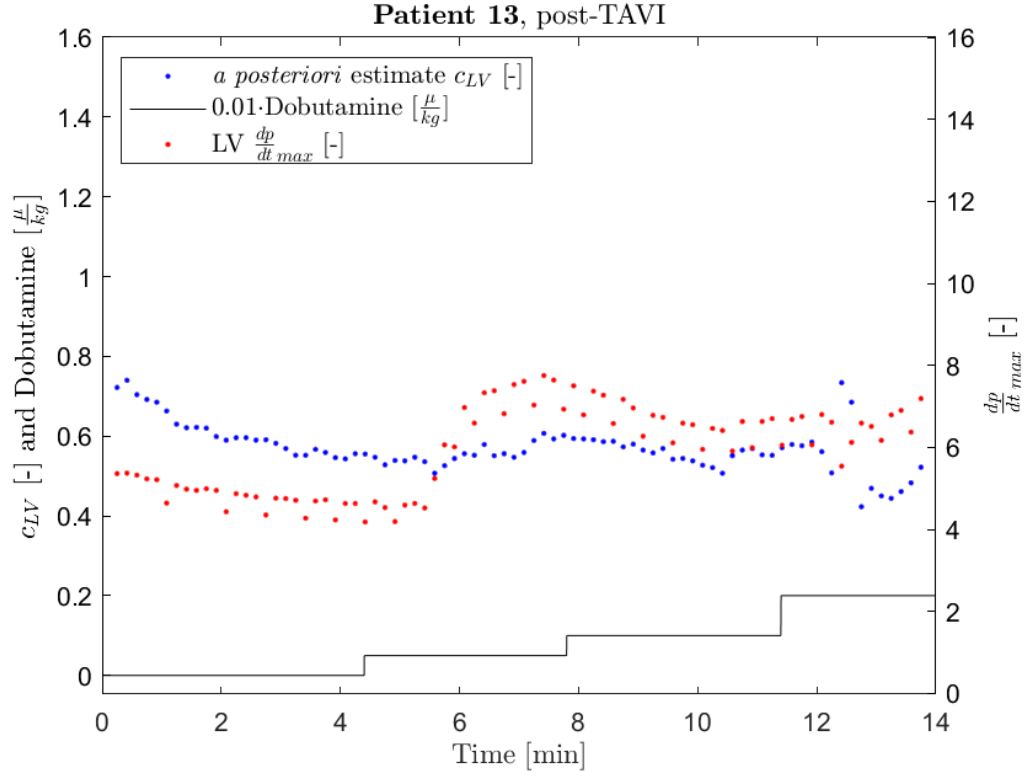


**Figure G.43:** Data belonging to **patient 13**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.

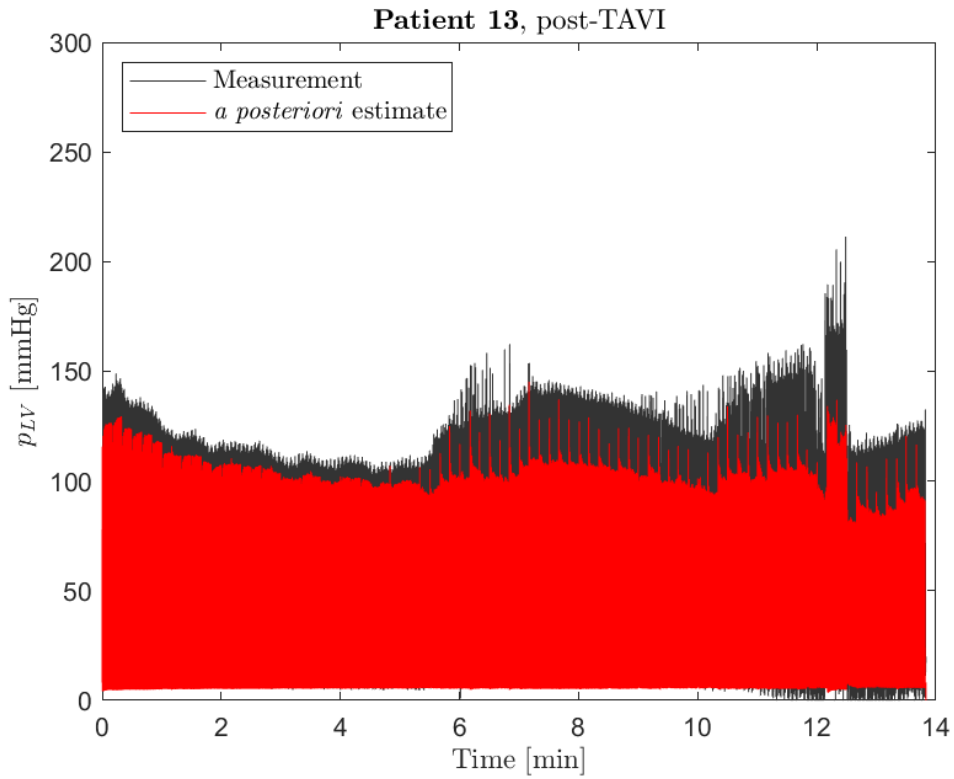


**Figure G.44:** Data belonging to **patient 13**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .

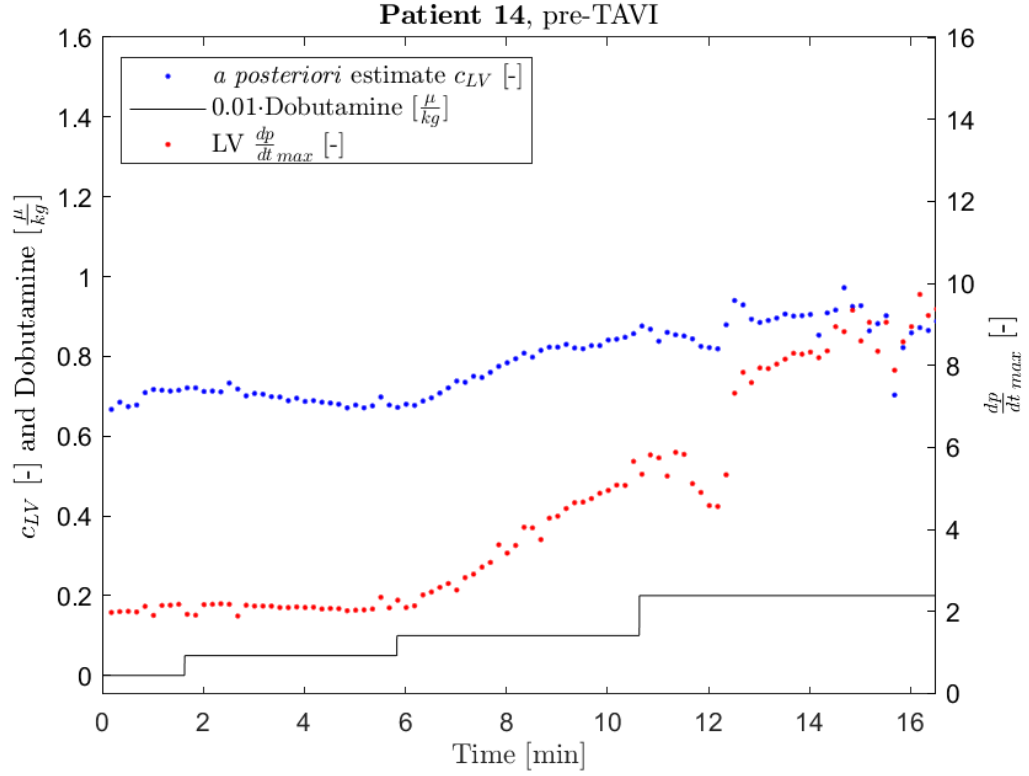




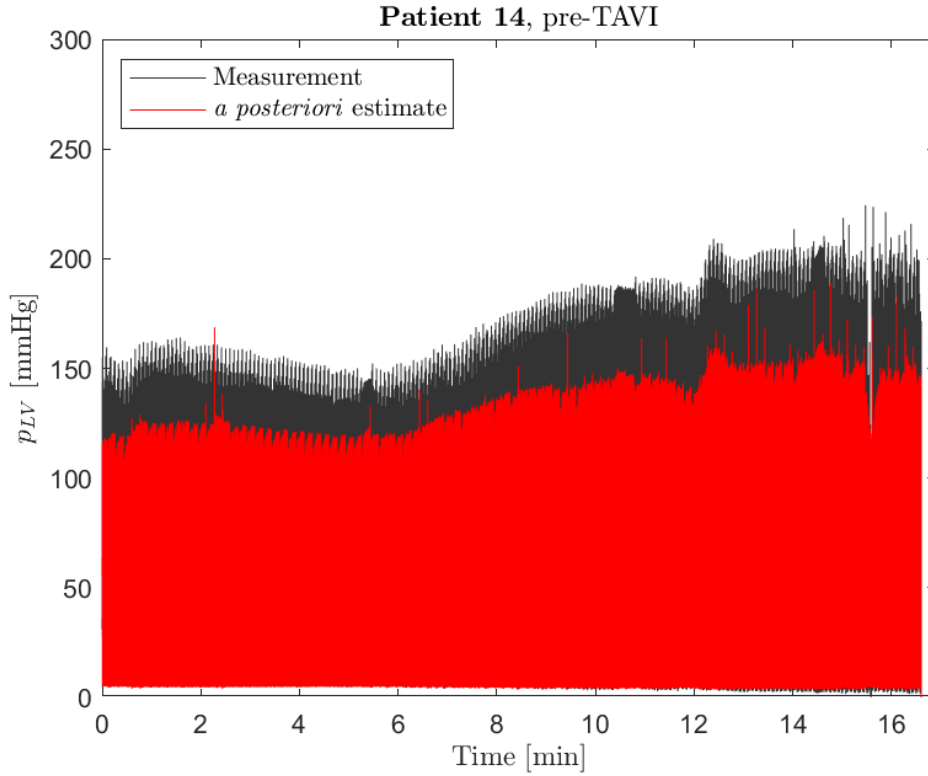
**Figure G.45:** Data belonging to **patient 13**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and LV  $\frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



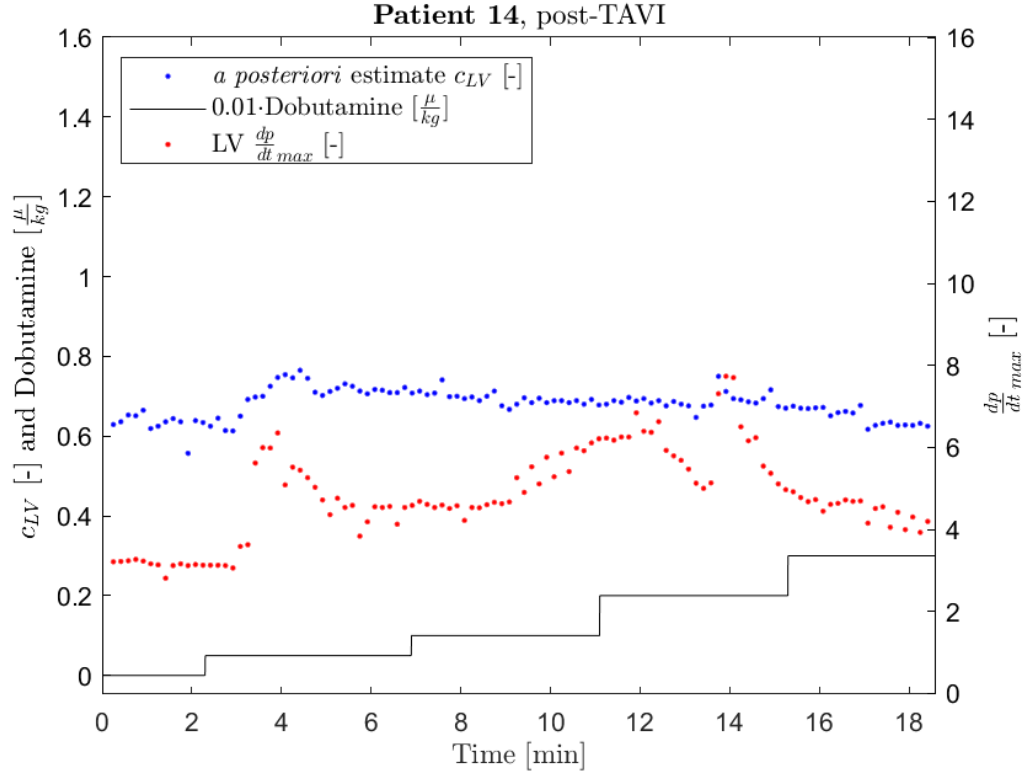
**Figure G.46:** Data belonging to **patient 13**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



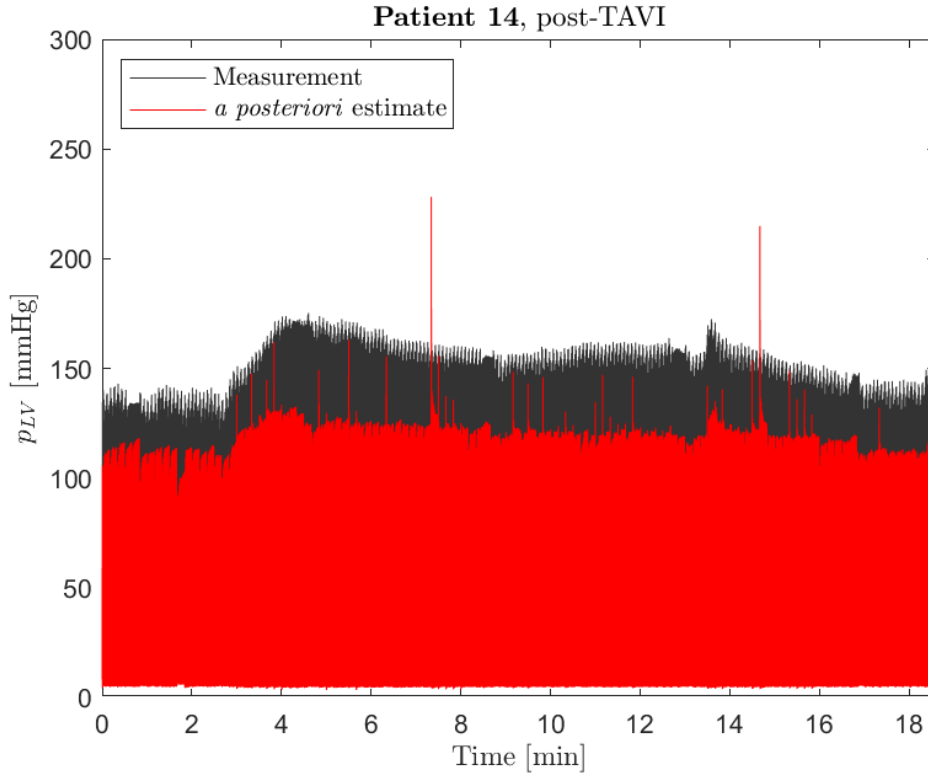
**Figure G.47:** Data belonging to **patient 14**, obtained *before* conducting the TAVI procedure. Comparison of estimated  $C_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



**Figure G.48:** Data belonging to **patient 14**, obtained *before* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the *a posteriori* estimate of the  $p_{LV}$ .



**Figure G.49:** Data belonging to **patient 14**, obtained *after* conducting the TAVI procedure. Comparison of estimated  $c_{LV}$  and  $LV \frac{dp}{dt}_{max}$  with step-wise Dobutamine administration.



**Figure G.50:** Data belonging to **patient 14**, obtained *after* conducting the TAVI procedure. Comparison of the measured  $p_{LV}$  and the  $a posteriori$  estimate of the  $p_{LV}$ .