

**Modelling and control of hemodialysis systems for better treatment
management of Chronic Kidney Disease patients**

by

Aristarchus Gnanasekar

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Process Control

Department of Chemical and Materials Engineering

University of Alberta

© Aristarchus Gnanasekar, 2021

Abstract

Kidneys are essential organs located on either side of the vertebral column which perform several essential bodily functions. When there is a gradual, permanent loss of basic kidney functions, a person is said to have Chronic Kidney Disease (CKD). CKD has been identified to be a global public health issue affecting millions of people every year. CKD can progress to an end-stage, and the patient's life would be at stake without artificial filtering (dialysis) or a kidney transplant. Hemodialysis is a life sustaining treatment for End Stage Renal Disease (ESRD) patients. Though being the most frequently used treatment modality, there are numerous clinical complications, while the most common include Intradialytic Hypotension (IDH) and Dialysis Disequilibrium Syndrome (DDS), arising during the quick extra-corporeal depuration of blood in an external device called 'hemodialyzer', which is sometimes referred to as 'artificial kidney'. This thesis starts off with a literature review in chapter 1 followed by a technical preliminaries review in chapter 2, to help the readers understand the research background and the problem better.

The perturbations caused by hemodialysis in a patient's body are complex, though the underlying phenomenon is a simple bidirectional mass transfer. The use of a mathematical model can enable a quantitative analysis of perturbations (cardiovascular response, fluid and solute kinetics) induced by hemodialysis taking place within the patient's body in different hemodialysis treatment settings and can help in understanding the intricate physiological mechanisms. In chapter 3, the mathematical models selected for representing each of the hemodialysis subsystems are presented along with some derivations, assumptions, control relevant modifications along with some simulations representing the hemodynamics of different classes of CKD patients.

Each patient behaves differently to hemodialysis and the challenge is to achieve meaningful predictions for each individual patient. Chapter 4 talks about the design of a simultaneous state and parameter estimation algorithm, specifically intended to identify individualized virtual patient simulators, based on synthetic clinical data, which could aid prediction of important state variables like Mean Arterial Pressure (MAP), Heart Period (HP), etc.,. For consistent estimation, the observability of the system has to be ensured and the nonlinear system observability test is not as straightforward like it would be for linear models. In our proposed approach, a sensitivity-based local observability test shall be conducted. The sensitivity equations should be solved in parallel with the original model equations to obtain the sensitivity matrix. Then a singular value decomposition is done to obtain the observability signature graph. A clear drop in the graph indicates a lack of observability. If such clear drops are encountered the user has to identify a subset of observable variables from the total variable set for estimation. For this purpose, a sequential orthogonalization algorithm was applied, to forward select the non-correlated variables one at a time until the terminating conditions are met, starting from the most sensitive and least correlated variable. The returned subset would be the decision variables during the simultaneous state and parameter estimation routine.

Traditionally, the hemodialysis treatments are done in open-loop fashion where the treatments are stopped when clinical complications occur and started again after the patient returns to normalcy. The model thus identified from chapter 4 could be used to design ‘individualized optimal treatments’ using advanced model based controllers, like a Batch Zone Model Predictive Controller (BZMPC) with a built in nonlinear state estimator, with feedback implementation while taking the treatment objectives and safety constraints into account as discussed in chapter 5, thus paving the way for continuous optimal safer treatments. Finally, the future research directions are narrated in chapter 6.

Acknowledgments

First and foremost, I thank the LORD Almighty, for blessing me with this wonderful opportunity through Dr. Jinfeng Liu and Dr. Zukui Li to undertake postgraduate research in the Department of Chemical and Materials Engineering, University of Alberta and for providing me with phenomenal strength each new day.

I still could not forget the day when Dr. Jinfeng Liu and Dr. Zukui Li offered me the opportunity to work on this project under their supervision. The offer came during the toughest times in my professional career and it has been an incredible journey with them so far. My sincerest gratitude goes to my supervisors, Dr. Jinfeng Liu and Dr. Zukui Li, for their continuous support and guidance during the entire duration of my graduate study. Many a time when the project experienced a roadblock, they always had innovative ideas and suggestions and I have benefited a great deal from their advice in the process of creating this thesis. I feel privileged for having worked with them and having learnt a lot from them in this short period of two years and their dedication and hard work never stops to amaze me. I wish that our professional paths cross again in future.

I am also extremely grateful to Dr. Zukui Li for making me the teaching assistant, not just once but twice, for the undergraduate course ‘CHE446 - Process Dynamics and Control’. The Indian literature portrays knowledge as the only indestructible wealth in this universe. I strongly believe that knowledge grows exponentially when it is spread, but teaching is more than just propagating wisdom and as a TA, I also pruned and developed my fundamental knowledge in every session I taught while preparing myself for the greater good.

I also owe a great deal of thanks to the individuals and organizations that provided financial support for this research project. I thank Cybernius Medical Ltd. not only for their monetary contribution, but also for sharing their valuable clinical expertise and domain-specific knowledge. It was immensely motivating when I received the

Captain Thomas Farrell Greenhalgh Memorial Graduate Scholarship from the Faculty of Graduate Studies and Research (FGSR).

It would be incomplete if I miss to acknowledge the support and help day in and day out from my colleagues in the Process Systems And Control Engineering (PSACE), Dr. Zukui Li's and Dr. Biao Huang's research group. Special thanks to Benjamin Decardi-Nelson, Soumya Sahoo, Jianbang Liu, Song Bo, Rui Nian, Xunyuan Yin, Jannatun Nahar, An Zhang, Yi Zhang, Erfan Orouskhani, Sarupa Debnath, Zhiyinan Huang, Sanjula Kammammettu, Faraz Amjad, Arun Senthil Sundaramoorthy, Dr. Rahul Raveendran and Dr. Nabil Magbool Jan.

Last but not least, I would like to express my deepest love and gratitude to all my teachers, mentors, friends and to my family for their unconditional love and support and for resting their trust and belief in me at all times.

I wish everyone good luck in all walks of life.

Contents

Abstract	ii
Acknowledgements	iv
List of Tables	x
List of Figures	xi
1 Introduction	1
1.1 Physiology of kidneys	2
1.1.1 Anatomy of kidneys	2
a. Location	2
b. Structure	3
c. Blood supply	5
1.1.2 Functions of kidneys	5
1.2 Literature review	7
1.2.1 Chronic Kidney Disease (CKD)	8
1.2.2 The uremic syndrome	11
1.2.3 Renal Replacement Therapy (RRT)	12
a. Peritoneal Dialysis (PD)	13
b. Hemodialysis (HD)	14
1.2.4 Clinical challenges during hemodialysis	16
a. Intradialytic hypotension (IDH)	17
b. Intradialytic hypertension	18
c. Dialysis Disequilibrium syndrome (DDS)	18
d. Technical complications	18
1.2.5 Attempts to mathematically model the elements of a hemodialysis system	18

1.2.6	Selection and estimation of variables of a mathematical model	21
1.2.7	Feedback control in hemodialysis	23
1.3	Research background	24
1.3.1	Motivation	24
1.3.2	Problem statement	26
1.3.3	Objectives	26
1.4	Thesis organization	27
1.4.1	Outline	27
1.4.2	Contributions	28
2	Preliminaries	29
2.1	Terms and definitions	29
2.2	Representation of the HD system model	30
2.2.1	Augmentation of the HD system model	31
2.3	Observability of a system	31
2.4	Sensitivity analysis	32
2.5	Singular Value Decomposition (SVD)	34
2.6	Orthogonalization	34
2.7	Design of a nonlinear state estimator	35
2.7.1	Extended Kalman Filter (EKF)	35
2.8	Design of feedback control algorithms	37
2.8.1	Proportional, Integral, Derivative (PID) controller	37
2.8.2	Zone Model Predictive Controller (ZMPC)	38
2.9	Summary	40
3	Mathematical modelling of the hemodialysis system	41
3.1	Guide to kinetic modelling	42
3.1.1	Example derivation: Single compartment model	42
a.	Derivation of Single Pool Fixed Volume (SPFV) ki- netic model	42
3.2	Hemodialysis system model	46
3.2.1	Patient subsystem model: Solute and fluid dynamics	48
a.	Solute balance model (2C)	48
b.	Fluid balance model (3C)	50
3.2.2	Patient subsystem model: Cardiovascular dynamics	52
a.	The heart pressures (6C)	52
b.	Baroflex regulation (3 MVs)	54

3.2.3	Hemodialyzer subsystem model	57
a.	Simple model: Linear equations giving exit concentration	57
b.	Complex model: Partial differential equations giving spatio-temporal concentration profile	58
3.2.4	Model simulation and validation	61
a.	Simulation results: HD system with simple model of hemodialyzer	61
b.	Simulation results: HD system with complex model of hemodialyzer	64
c.	Leveraging the power of the HD system model: Does UFR play an important role in decreasing the concentration of solutes?	66
3.2.5	Summary of the HD system model	68
3.3	Summary	69
4	Simultaneous state and parameter estimation of the hemodialysis system	70
4.1	Initial attempts towards building an individualized virtual patient simulator	71
4.2	Proposed method	76
4.2.1	Augmentation of the HD system model	78
4.2.2	Nominal values of variables	78
4.2.3	Observability analysis framework	79
a.	Sensitivity analysis	79
b.	SVD analysis	80
c.	Orthogonalization	80
4.2.4	Optimization problem formulation for simultaneous estimation of states and parameters	82
a.	Analysis of convexity	83
b.	Optimization solution strategy: Sequential Coordinate Block Descent (SCBD)	83
c.	Selection of optimization solver	85
4.2.5	Validation of identified model	86
4.3	Design of simulation experiments	86
4.3.1	Generation of synthetic clinical data	87
4.3.2	Building of test cases	88

4.4	Simulation results and discussion	89
4.4.1	Class 3 patient: Severe IDH & weak hemodynamic stability	90
4.4.2	Class 2 patient: Mild IDH & acceptable hemodynamic stability	99
4.4.3	Class 1 patient: No IDH & good hemodynamic stability	102
4.5	Summary	104
5	Towards optimal control of the hemodialysis system	105
5.1	Application of traditional bio-feedback control techniques in HD	106
5.1.1	Simulation results and discussion	107
5.2	Layout of the feedback control system for application in HD	112
5.3	Design of a nonlinear state estimator for application in HD: Extended Kalman Filter (EKF)	114
5.3.1	Simulation results and discussion	116
5.4	Design of a nonlinear Batch Zone Nonlinear Model Predictive Controller (BZNMPC) for application in HD	121
5.5	BZNMPC: Simulation results and discussion	126
5.5.1	Class 3 patient: Severe IDH & weak hemodynamic stability	128
	Hypothetical control problem	128
5.5.2	Class 2 patient: Mild IDH & acceptable hemodynamic stability	131
5.5.3	Class 1 patient: No IDH & good hemodynamic stability	133
5.6	Determination of optimal HD treatment time	134
5.7	Summary	136
6	Conclusions	137
6.1	Conclusion	137
6.2	Major findings of this current work	139
6.3	Future research directions	140
	Bibliography	143

List of Tables

1.1	Stages of Chronic Kidney Disease (CKD) based on Glomerular Filtration Rate (GFR) [57]	8
1.2	Albuminuria categories in CKD [57]	9
3.1	Simulation settings for single and double pool fixed volume urea kinetic models	44
3.2	Simulation settings for a HD system with the complex model of hemodialyzer	65
4.1	Summary of model variables of the chosen HD system model with a simple model of the hemodialyzer	71
4.2	Description of notations involved in the optimization problem formulation	82
4.3	Bounds for the state variables of the HD system model	84
4.4	HD treatment settings for 7 different patients for synthetic clinical data generation. Data adopted from [30]	87
4.5	Optimization results summary - Case 1 & 2	89
4.6	Optimization results summary - Case 3 & 4	90
4.7	Ranking of uncorrelated variables by orthogonalization algorithm for Patient 4	95
4.8	Summary of chapter 4 analysis	104
5.1	PID tuning parameters	109
5.2	Tuned diagonal elements of the initial guess of the process covariance matrix obtained from Eq. 2.17	117
5.3	Realization of the BZNMPC problem for application in HD	125
5.4	Configuration of manipulated inputs in BZNMPC	126
5.5	Summary of BZNMPC results	127

List of Figures

1.1	Structure of kidneys. Taken from [51]	3
1.2	Major vessels that supply the blood flow to the kidney and a schematic of the microcirculation of each nephron. Taken from [51]	4
1.3	Prognosis of Chronic Kidney Disease (CKD). Taken from [76]	9
1.4	A pictorial representation of Peritoneal Dialysis (PD). Taken from National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (www.kidney.niddk.nih.gov)	13
1.5	A simple layout of Hemodialysis (HD) setup. Taken from National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (www.kidney.niddk.nih.gov)	14
1.6	The mass transfer processes in HD. Taken from [60]	15
1.7	Percentages of patients in Renal Replacement Therapy (RRT) across different treatment modalities in Europe 2018 (HD: Hemodialysis, PD: Peritoneal dialysis, T: Transplant). Data from [6]	16
1.8	Fluid overload chart of a patient undergoing HD 3 times a week. Taken from [110]	17
1.9	A schematic representation of the HD system	19
1.10	Pictorial representation of the 3 compartment model of the human body	20
1.11	Feedback control of hemodialysis. Taken from (www.uninet.edu)	23
2.1	Block diagram of a state estimation algorithm	35
2.2	Block diagram of a PID controller (forum.dronebotworkshop.com)	37
2.3	State trajectory in a typical zone control setting	39
3.1	Single Pool Fixed Volume (SPFV) urea kinetic model. Taken from [36]	43
3.2	Simulation of the Single Pool Fixed Volume (SPFV) and Double Pool Fixed Volume (DPFV) urea kinetic model	45
3.3	Double Pool Fixed Volume (DPFV) urea kinetic model. Taken from [36]	46

3.4	Body compartment model for solute (2C - dotted line) and fluid exchange (3C - solid line) kinetics. Adopted from [140]	49
3.5	A schematic of the 6 compartment model of the cardiovascular system. Adopted from [43]	53
3.6	Block diagram for the control of systemic resistance. Adopted from [140]	55
3.7	Concentration profile diagram for counter current flow of blood and dialysate in a hemodialyzer	57
3.8	Representation of hollow fiber hemodialyzer used in HD. Taken from [16]	59
3.9	Representation of stages : Discretization along z direction	60
3.10	Python work flow	61
3.11	Simulated hemodynamic variables of Class 1 patient: No hypotension	62
3.12	Simulated hemodynamic variables of Class 2 patient: Mild hypotension	63
3.13	Simulated hemodynamic variables of Class 3 patient: Severe hypotension	63
3.14	Determination of number of stages	64
3.15	Clinical vs Model predicted results - Patient concentration profile	64
3.16	Spatiotemporal concentration profile of urea - Hemodialyzer concentration profile	66
3.17	Spatiotemporal concentration profile of creatinine - Hemodialyzer concentration profile	66
3.18	Comparison of UFR profiles	67
3.19	Summary of the chosen HD system mathematical model	68
4.1	Observability analysis of the HD system using PBH test	72
4.2	Observability analysis using PBH test for the augmented system with all states and the first sensitive parameter	74
4.3	Estimation of top three sensitive parameters using three measurements	75
4.4	Proposed method for simultaneous state and parameter estimation	77
4.5	Orthogonalization method work flow	81
4.6	Comparison of simulations with nominal (blue) and true (yellow) values of states and parameters	91
4.7	Relative sensitivity trajectories of measurements for Patient 4	92
4.8	Observability signature graph: SVD analysis of the sensitivity matrix for Patient 4	93
4.9	Euclidean norm of the ranked variables by orthogonalization	94
4.10	Patient 4: Training (Case 1: $MSE_{M1} = 20.897$, $MSE_{M2} = 6.63E-04$, Case 2: $MSE_{M1} = 4.175$, $MSE_{M2} = 1.14E-04$)	96

4.11 Patient 4: Validation (Case 1: $MSE_{M1} = 52.044$, $MSE_{M2} = 1.22E-03$, Case 2: $MSE_{M1} = 6.102$, $MSE_{M2} = 3.66E-04$)	97
4.12 Patient 4: Training (Case 3: $MSE_{M1} = 27.642$, $MSE_{M2} = 1.39E-03$, Case 4: $MSE_{M1} = 13.349$, $MSE_{M2} = 1.12E-03$)	98
4.13 Patient 4: Validation (Case 3: $MSE_{M1} = 32.217$, $MSE_{M2} = 2.233E-03$, Case 4: $MSE_{M1} = 20.644$, $MSE_{M2} = 1.34E-03$)	99
4.14 Patient 1: Training (Case 1: $MSE_{M1} = 0.035$, $MSE_{M2} = 2.55E-05$, Case 2: $MSE_{M1} = 0.016$, $MSE_{M2} = 8.16E-06$)	100
4.15 Patient 1: Validation (Case 1: $MSE_{M1} = 0.033$, $MSE_{M2} = 3.48E-05$, Case 2: $MSE_{M1} = 0.019$, $MSE_{M2} = 1.02E-05$)	100
4.16 Patient 1: Training (Case 3: $MSE_{M1} = 0.16$, $MSE_{M2} = 1.00E-04$, Case 4: $MSE_{M1} = 0.191$, $MSE_{M2} = 1.15E-04$)	101
4.17 Patient 1: Validation (Case 3: $MSE_{M1} = 0.569$, $MSE_{M2} = 2.81E-04$, Case 4: $MSE_{M1} = 0.522$, $MSE_{M2} = 2.75E-04$)	102
4.18 Patient 7: Training (Case 1: $MSE_{M1} = 8.56E-05$, $MSE_{M2} = 9.62E-09$, Case 2: $MSE_{M1} = 8.80E-06$, $MSE_{M2} = 1.24E-09$)	102
4.19 Patient 7: Validation (Case 1: $MSE_{M1} = 0.003$, $MSE_{M2} = 3.05E-07$, Case 2: $MSE_{M1} = 3.12E-04$, $MSE_{M2} = 4.13E-08$)	103
4.20 Patient 7: Training (Case 3: $MSE_{M1} = 0.009$, $MSE_{M2} = 1.51E-06$, Case 4: $MSE_{M1} = 0.008$, $MSE_{M2} = 1.40E-06$)	103
4.21 Patient 7: Validation (Case 3: $MSE_{M1} = 0.103$, $MSE_{M2} = 1.30E-05$, Case 4: $MSE_{M1} = 0.094$, $MSE_{M2} = 1.30E-05$)	104
5.1 PID case 1: When initial states are 5% higher	108
5.2 PID case 2: When initial states are 5% lower	109
5.3 Class 1 patient under treatment in a BVT system	110
5.4 Class 2 patient under treatment in a BVT system	111
5.5 Class 3 patient under treatment in a BVT system	111
5.6 Block diagram of the proposed feedback control system with process and sensor noise	113
5.7 Extended Kalman filter (EKF) with 1 measurement (MAP)	116
5.8 Observer RSE dynamics with 1 measurement (MAP)	117
5.9 Extended Kalman filter (EKF) with 2 measurements (MAP, HP)	118
5.10 Observer RSE dynamics with 2 measurements (MAP, HP)	118
5.11 Comparison of estimation performance with the help of normalized RMSE dynamics	119

5.12	Estimation performance of EKF for different patients in table 4.4 in the presence of process and measurement noises (both $N(0, 0.0005^2)$)	120
5.13	Conflicting objectives posed by the HD problem statement	123
5.14	Shrinking horizon: Progression of the prediction horizon	124
5.15	Patient 4: Treatment objectives (3.5 h)	128
5.16	Patient 4: Optimal treatment profiles from the BZNMPC (3.5 h) . . .	129
5.17	Patient 4: Failed - Treatment objectives (3 h)	130
5.18	Patient 4: Failed - Optimal treatment profiles from the BZNMPC (3 h)	131
5.19	Patient 1: Treatment objectives (3 h)	132
5.20	Patient 1: Optimal treatment profiles from the BZNMPC (3 h) . . .	132
5.21	Patient 7: Treatment objectives (3 h)	133
5.22	Patient 7: Optimal treatment profiles from the BZNMPC (3 h) . . .	134
5.23	Chart for finding the optimal treatment time based on re-numeration	135
6.1	Assimilation of patient's clinical data from multiple HD treatments .	141

Chapter 1

Introduction

Chronic Kidney Disease (CKD) has been identified as a pressing public worldwide health issue affecting the livelihood of millions of people. In 2017, the Global Burden of Disease (GBD) study estimated that the global prevalence of CKD is around 9.1% (697.5 million cases) of the total global population [19]. One of the sustainable development goals of the United Nations (UN) is to reduce the mortality from Non Communicable Diseases (NCD) by one third by 2030 [25], and on those lines targeting CKD would be of prime importance for achieving that goal. The swift increase in the total CKD patient numbers worldwide could be because of a lot of factors like diabetes mellitus, hypertension, obesity, and aging [84].

Millions of people worldwide who are classified as ESRD patients among the CKD patient pool, undergo dialysis (mainly hemodialysis) treatment or kidney transplantation as a result [33]. The large number of deaths on account of CKD, is mainly due to poor access to renal replacement therapies and timely medical care. Over 80% of patients who receive good renal replacement therapy are in wealthy countries which have access to universal healthcare and have large elderly populations [61]. With the CKD population growing at an alarming rate, even the most wealthy countries will face a substantial economic burden because of CKD. At first, the reader is introduced to the basic physiology of kidneys, classification of CKD, followed by a vast array of information, related to the problem under consideration, from literature which will prompt the reader to critically analyze and appreciate this work's research background. The final section in this chapter is dedicated to explain the organization of this thesis.

1.1 Physiology of kidneys

Life of an organism is dependant on several biochemical processes working in co-ordination with each other. The process of excretion is as important as any other processes in the body of an organism responsible for the sustenance of life. During life activities such as cellular respiration, chemical reactions known as ‘metabolism’ take place in the body. These chemical reactions produce waste products such as carbon dioxide, water, salts, urea and uric acid. Accumulation of these wastes beyond a certain level inside the body is harmful to the body. Excretion is a process in which metabolic wastes are removed from a living organism and there is a specialised system of organs called ‘excretory organs’ for this purpose in human beings. Although the mode of excretion differs among organisms depending on their habitat and food habit, in fully evolved vertebrates, excretion is mainly carried out through lungs, kidneys and skin [17].

For instance, ammonia is the excretory product in aquatic animals, while birds and insects excrete mainly uric acid. Humans produce urea as the major excretory product and there are several parts of the body that are involved in the excretory process, such as sweat glands in the skin, the liver, the lungs and the kidney system. As this work primarily revolves around the kidneys and their associated phenomena, let us discuss more about the physiology of kidneys to understand and appreciate the motivation of this research better.

1.1.1 Anatomy of kidneys

a. Location

In humans, the excretory or more specifically the ‘urinary system’ includes the kidneys, a pair of ureters, a urinary bladder and an urethra. The kidneys are ‘bean-shaped organs’ located on either side of the vertebral column, inside the abdominal cavity, more specifically in the paravertebral gutter and lie in a retroperitoneal position at a slightly oblique angle [111] as shown in figure 1.1. There are two kidneys in total and each of the kidneys are about 4 to 5 inches long, 6 cm wide and 4 cm thick [131]. The position of the liver on the right side of the vertebral column creates an asymmetry within the abdominal cavity and this results in the left kidney being slightly upper than right kidney, and the left kidneys being located more medial than the right. At the vertebral level, the left kidney being slightly larger than the right, is positioned just about T12 to L3, and the right kidney slightly smaller and lower than the left. Both the kidneys sit below the diaphragm while the right is posterior to the

liver and the left is posterior to the spleen. The adrenal gland rests on top of each of the kidneys. The eleventh and twelfth ribs, protect the upper part of the organs to some extent and each whole kidney and adrenal gland are surrounded by two layers of fat namely, the perirenal and pararenal fat and the renal fascia. In a fully grown adult, each kidney weighs between 125 and 170 grams in males and between 115 and 155 grams in females [111].

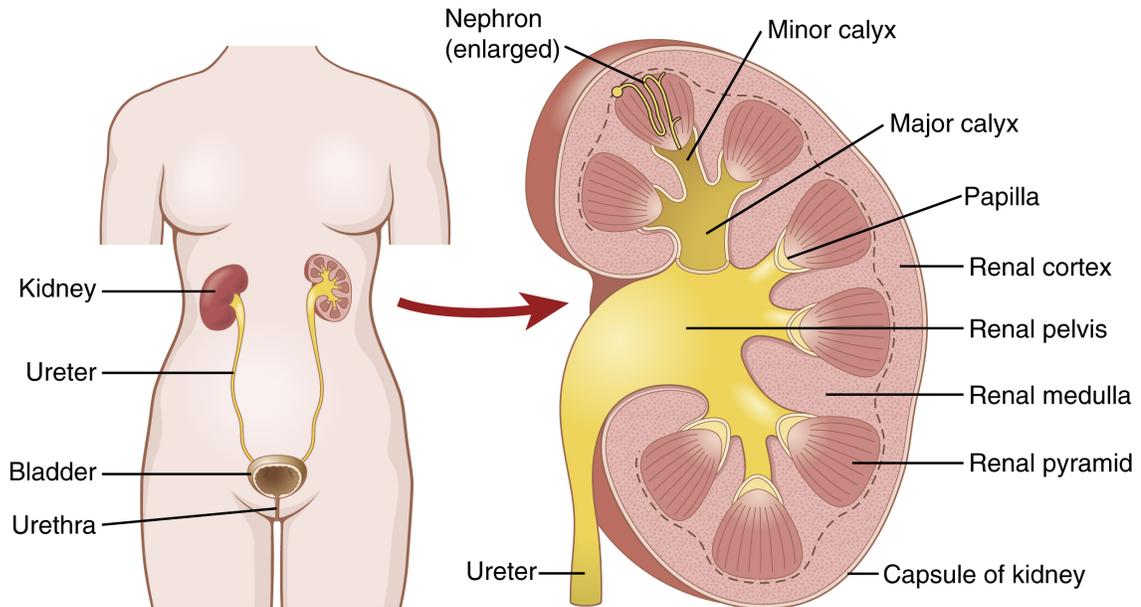


Figure 1.1: Structure of kidneys. Taken from [51]

b. Structure

If we examine the geometry of the kidneys, each kidney has a convex and a concave surface. If one looks closely at figure 1.2, a renal artery and a renal vein enters and exits each kidney respectively, at the medial indentation or the concave surface called ‘renal hilum’ and this notch gives the kidneys the shape of a bean [131]. A ureter (yellow coloured tube going out of kidneys) which is a smooth muscle walled tube also exits at the concave surface and this serves as the passage pathway for urine into the single urinary bladder. A thick fibrous tissue called renal capsule covers both kidneys. The renal capsule is surrounded by perinephric fat, renal fascia (Gerota) and paranephric fat. The anterior (front) border of these tissues is the peritoneum, while the posterior (rear) border is the transversalis fascia [111].

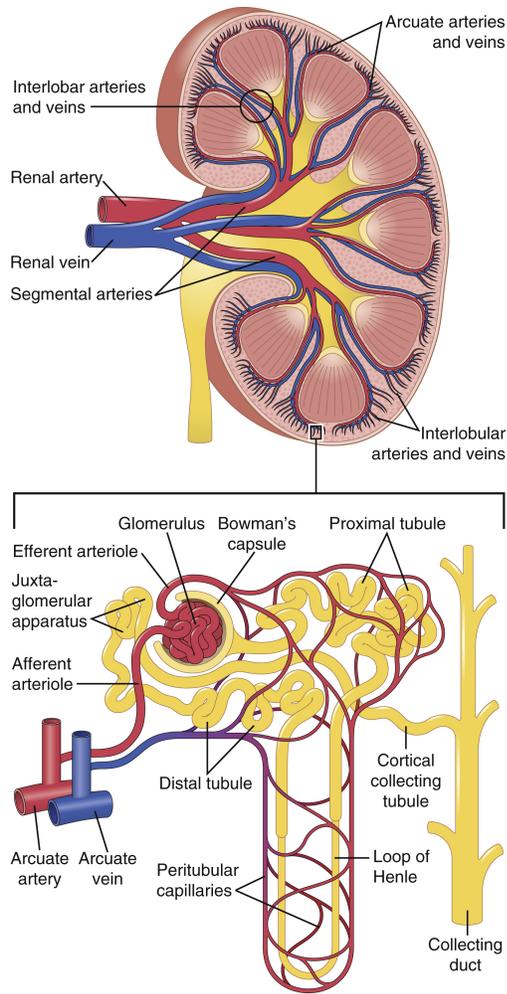


Figure 1.2: Major vessels that supply the blood flow to the kidney and a schematic of the microcirculation of each nephron. Taken from [51]

The superior border of left and right kidneys are adjacent to spleen and liver respectively. As the kidneys are located below the diaphragm, both of the kidneys move down during inhalation of the respiratory process. The functional tissue area of the kidney called ‘parenchyma’, is divided into two major structures: the renal cortex on the superficial side and the renal medulla in the interior. As shown in figure 1.1, these structures take the shape of roughly 8 to 18 cone-shaped renal lobes. Each renal lobe contains renal cortex surrounding a portion of medulla called as a renal pyramid. The finger shaped projections of renal cortex between the renal pyramids are known as renal columns. The basic functional unit of the kidneys are called nephrons. There are over a millions nephrons in each of the kidneys and they span across the entire cortex and medulla area [63], [150]. The filtration of metabolic wastes from the blood plasma begins in the renal corpuscle, situated in the renal cortex, which is followed by

renal tubules that passes from the cortex deep into the renal pyramids. A collection of renal tubules called as a medullary ray then drains into a single collecting duct. The tip of each renal pyramid called papilla, then drains urine into a minor calyx, which in turn empties it into major calyces. Finally these major calyces, empty urine into the renal pelvis which then becomes the ureter.

c. Blood supply

Every organ in a human body needs blood supply for the supply of essential nutrients and kidneys are no exception. The blood irrigation in the kidneys occurs through right and left renal arteries. These arteries branch out of abdominal aorta. Though the blood circulation system to the kidneys is comparatively smaller in size, it attracts 22% of the total cardiac output or 1100 ml/min [51].

The renal arteries which enter the kidneys at the medial position split into segmental arteries, which in turn branch into interlobar arteries. The interlobar arteries pass through the renal columns between the renal pyramids and transfix into the renal cortex. The interlobar arteries then branch into several arcuate arteries which irrigate the boundary between the cortex and the medulla. The glomerulus, where all the filtration takes place, is supplied with blood through afferent arterioles which branch out from the arcuate arteries. The functional void space in the kidneys between the individual glomerulus called the 'interstitium' reabsorbs fluid recovered from urine. The interstitium is rich in blood vessels too. One can understand by looking at figure 1.2, that blood is taken out of the kidneys for transfusion and purification through a similarly named system of veins and venules [55]. The only difference with the vein system is that the direction of blood flow is towards the heart.

1.1.2 Functions of kidneys

The first thing which pops out in one's mind when thinking about the functions of kidneys is that they are organs responsible for removing the toxic substances which are produced by ingestion or metabolism from the body. The second thought would be that kidneys help remove waste materials from the blood plasma in the form of urine and help in regulation of total body fluid volume. Kidneys maintain the electrolyte composition of the blood within strict physiological limits and they have several other quintessential functions too. Thus kidneys are responsible for maintaining a stable internal environment essential of the human body for proper functioning of many other cellular activities. The kidneys perform complex processes within the nephron (glomerular filtration, tubular secretion and tubular reabsorption of water,

electrolytes and metabolic waste products) depending on the needs of the body and for achieving homeostasis [79].

In overall, the kidneys clear toxic substances from the filtrate, which in turn is obtained from the blood, by excreting them in the urine while reabsorbing the essential nutrients and sending them back to the blood. The list of some of the important homeostatic functions performed by healthy kidneys [51], [55], [68], [111] are given below:

- **Excretion of metabolic waste products and foreign chemicals:** Kidneys are the main organs responsible for the excretion of most of the waste products of metabolism (urea, creatinine, uric acid, bilirubin and metabolites of various hormones) and other foreign substances that are taken by diet or synthesized by the body (pesticides, drugs, food additives).
- **Regulation of water and electrolyte balances:** For maintaining equilibrium within the human body or homeostasis, the kidneys excrete water and electrolytes adapting to one's eating and drinking habits.
- **Regulation of body fluid osmolality and electrolyte concentrations:** The regulatory response of the kidneys to aberrant fluctuations in not only sodium levels in the blood plasma but also to water and other electrolytes like such as chloride, potassium, calcium, hydrogen, magnesium, and phosphate ions is phenomenal.
- **Regulation of arterial pressure:** Kidneys play a vital role in both long-term and short-term regulation of arterial blood pressure by removing water and electrolytes and by secreting vasoactive substances and hormones (renin, aldosterone) respectively.
- **Regulation of acid-base balance:** The kidneys work along with lungs and body fluid buffers for the maintenance of acid-base balance. It is interesting to note that kidneys are the only means of elimination of some acids from the body like sulfuric acid and phosphoric acid.
- **Regulation of erythrocyte production:** The kidneys secrete almost all of a hormone called erythropoietin, which is responsible for the production of red blood cells by hematopoietic stem cells in the bone marrow. End Stage Renal Disease (ESRD) patients who are in a hemodialysis regime also face severe

anemia as a result of abnormality in this function.

- **Regulation of calcitrol production:** The main organs behind calcium homeostasis are kidneys. Kidneys produce 1,25-dihydroxyvitamin D₃ (Calcitrol), which is the active form of vitamin D. Calcitrol is quintessential for the normal deposition of calcium in the bones and reabsorption of calcium in the gastrointestinal tract.
- **Gluconeogenesis:** During starvation for a long period of time, kidneys play an essential role in the production of glucose from amino acids, by removing the amino group as ammonia waste and make the rest of the compound available for glucose production.
- **Secretion, metabolism, and excretion of hormones:** The hormones required for the functions mentioned above are secreted and excreted by kidneys. The maintenance of life depends on the homeostasis.

1.2 Literature review

Now that we have developed an understanding of the basic physiology of kidneys and its functions, we can dig deep into the available sources in literature to understand the motivation and the formulation of the problem statement in this thesis work better. In a human being with a healthy functioning ecosystem of organs, approximately 1500 litres of blood is circulated through the kidneys each day [110]. If one assumes that the average blood volume in a human being is around 5 litres [132], then the kidneys clean the whole blood volume around 300 times a day. The urine excretion rate heavily depends on the individual's intake of water and other factors and can vary from 0.5 L/day for a dehydrated person to upto 20 L/day for a body fluid overloaded person [51]. So kidneys play an essential role in maintaining the good health of a human body. But the performance of the kidneys degrades over time and they cannot always work with 100% efficiency always. There are several clinical and pathological conditions which result in the degradation of kidney function. A person can survive though life if the kidney function is more than 25% though the person needs medical attention in parallel for such low levels of kidney function, but if the kidney function drops below this point, the kidneys fail to perform its essential functions and would not be able to maintain homeostasis. The degradation of kidney function could be over a period of time or abrupt in onset. Also, the loss of kidney function could be reversible or irreversible. Any reversible loss of kidney function is

termed as ‘Acute Renal Failure (ARF)’ and if a person experiences irrecoverable loss of kidney function then he is said to have a Chronic Kidney Disease (CKD).

Table 1.1: Stages of Chronic Kidney Disease (CKD) based on Glomerular Filtration Rate (GFR) [57]

GFR category	GFR (ml/min/1.73m²)	Residual kidney function	Description of kidney function
G1	>90	>90%	Normal or high
G2	60 to 89	60 to 89%	Mildly decreased
G3	30 to 59	30 to 59%	Mild to severely decreased
G4	15 to 29	15 to 29%	Severely decreased
G5	<15	<15%	Kidney failure

1.2.1 Chronic Kidney Disease (CKD)

In the course of this work, more interest is diverted to patient groups with kidney diseases, most importantly Chronic Kidney Disease (CKD) patients. Therefore it is essential for the reader to understand the classification of CKD too. CKD refers to the gradual loss of kidney function and in the functioning of kidney tissue (nephrons). The US National Kidney Foundation defines CKD as abnormalities of kidney structure or function, present for over 3 months, with implications for health [57]. It is not easy to diagnose a patient with CKD until and unless the function of the kidneys are lost in an irreversible fashion and are significantly impaired. There are several symptoms associated with CKD namely, nausea, loss of appetite, fatigue, weakness, sleep cycle problems, reduced mental sharpness, muscle cramps, swelling of ankles or itching [38], [54]. But the condition of the kidneys are typically diagnosed by clinical practitioners through a series of tests of the patient’s blood, urine and other examinations [66], [80]. US National Kidney Foundation-Kidney Disease Outcomes Quality Initiative (NKF-KDOQI) classifies CKD into five stages based on Glomerular Filtration Rate (GFR) as shown in table 1.1. Glomerular Filtration Rate (GFR) is an indicator of kidney’s efficiency. In specific, it means the volume of blood passing every minute through the filters in the kidneys called glomeruli, where the waste materials are separated from the blood. A doctor finds the GFR based on the results of the blood sample test, age, body size and gender of the patient. One other marker used by the doctors is the quantity of a type of protein called albumin in the urine sample. The phenomenon of having albumin in the urine is called albuminuria, and the presence of protein in the urine can be taken as an initial sign of CKD and that the functions

of the kidneys are seriously disturbed. The higher the albumin concentration in the urine, the higher the risk of having a CKD. The category classification of albuminuria, according to the US National Kidney Foundation, is outlined in table 1.2. AER and ACR in the table stand for Albumin Excretion Rate and Albumin to Creatinine Ratio respectively.

Table 1.2: Albuminuria categories in CKD [57]

Category	AER (mg/day)	ACR (approx. equiv.) (mg/mmol)	ACR (approx. equiv.) (mg/g)	Description
A1	<30	<3	<30	Normal or mildly increased
A2	30-300	3-30	30-300	Moderately increased
A3	>300	>30	>300	Severely increased

				Albuminuria stages, description, and range (mg/g)				
				A1		A2	A3	
				Optimum and high-normal		High	Very high and nephrotic	
				<10	10-29	30-299	300-1999	≥2000
GFR stages, description, and range (mL/min per 1.73m ²)	G1	High and optimum	>105					
			90-104					
	G2	Mild	75-89					
			60-74					
	G3a	Mild-moderate	45-59					
	G3b	Moderate-severe	30-44					
	G4	Severe	15-29					
G5	Kidney failure	<15						

Figure 1.3: Prognosis of Chronic Kidney Disease (CKD). Taken from [76]

The doctor will continue to critically investigate the cause of CKD and check the function of the kidneys so that a better treatment can be planned. The final stage of a CKD in progression (Stage 5), when the GFR falls below 15 mL/min/1.73 m², is called the stage of Chronic Renal Failure (CRF) or End Stage Renal Disease (ESRD) [57].

At this stage, the kidneys cannot effectively remove excess fluids or maintain the electrolyte balance and filter out the waster materials from the blood and the patients need immediate medical support for survival. Though the GFR and higher levels of albuminaria are independently related to morbidity, ESRD, greater levels of albuminaria could be present at all stages of GFR [31], [77]. Hence, integrating both the categories of GFR and albuminaria in a single chart, as shown in figure 1.3, gives the clinical practitioner a better understanding and an enhanced prognosis of the CKD condition. The patients who fall in the categories marked with red boxes have a very high risk CKD and need immediate medical care.

Out of several factors resulting in renal failure, the most prominent causes would be heart attacks, kidney damage, decreased blood flow and complications from certain medications. Chronic renal failure could be brought about by diabetes, chronic high blood pressure (hypertension), lupus, chronic nephritis or polycystic kidney disease and kidney disease [38], [60]. Age, obesity, smoking, family history of CKD could amplify the occurrence of CKD too. Following are some of the consequences of ESRD in a patient:

- Uraemia or uremic syndrome caused by abnormal levels of waste products in the body. This needs more attention and is discussed in the next subsection.
- Accumulation of water in the tissues due to fluid overload, a condition called ‘oedema’.
- Hypertension due to increased circulating blood volume.
- Electrolyte imbalance.
- Hormonal imbalance. For example, diminished production of erythropoietin causes severe anaemia. Application of advanced process control strategies in addressing this issue was the focus in some of our research group’s previous works [13], [95].
- Abnormal enzyme production.
- Decalcification of bones.

1.2.2 The uremic syndrome

Some of the approaches taken to reverse the signs of renal failure include manipulation of the diet, transplantation and various types of artificial filtration schemes. ‘Uraemia’ means urine in the blood. It refers to a general class of complaints and signs that the patients exhibit even when the kidney function falls by 5% from the normal values. GFR and albuminuria which are identified through clinical tests, could serve as measurements of renal damage but not uraemia in general. There is huge biovariability in the occurrence of uraemia among different patients. The clinical practitioners and investigators have been in search of a toxin that could be responsible for this syndrome. The failure of over a century of research to identify one specific or a group of toxins responsible for this syndrome has made the researchers settle down with some alternate toxin theories [36]. If the readers are prompted to know more about the uremic toxins, they could look at some of these well accepted literature sources [39], [107].

Ever since advanced renal replacement therapies became widespread across the globe, the incidence of patients with severe levels of uraemia has been drastically reduced. Better treatment management of ESRD patients can be achieved, if the doctors are able to identify even minor symptoms and signs for the need for advanced treatment modalities. If there is a one stop solution for the measurement of uraemia, then this problem would have been addressed long back. Sadly, only marker solutes that correlated with uremic toxicity have been identified. These marker solutes are not toxic by themselves in nature. These include the end products of protein metabolism, urea and creatinine. But there are several other marker solutes which are toxic and accumulate in a patient’s body without any correlation with uremic toxicity [15], [62], [94].

Creatinine levels correlate fairly well with GFR and it is the most popular marker to represent the patient’s intrinsic renal function. But as CKD progresses, the tubular section of the kidneys account of a higher excretion of creatinine and so the GFR is overestimated. Moreover, creatinine is produced by the muscle and it is highly dependent on the muscle mass of the patient and there is not an easy way to apply a correction factor for the muscle mass in the calculations. For an ESRD patient, high blood creatinine levels could either mean a high production of creatinine by the muscles or inadequate treatment. So creatinine comes as the second choice, but it could be used as a marker for low molecular weight substances that are to be removed by the treatment modality.

On the other hand, the blood urea concentration, generally expressed as Blood Urea Nitrogen (BUN), steeply rises in patients with CRF. Urea fails several benchmark tests when used to quantify mild or moderate reduction in kidney function. However, as the patient goes into ESRD and encounters CRF, there are several benefits of using urea as the marker solute [107] and it correlates well with uremic toxicity. According to the National Cooperative Dialysis Study (NCDS), the average urea levels correlate with the treatment outcome as well [88]. This makes urea the primary marker solute choice of uremia for clinical practitioners and researchers.

1.2.3 Renal Replacement Therapy (RRT)

When the kidneys of the patients completely stop functioning or if the patient is in a ESRD (Stage 5 CKD), the doctors resort to Renal Replacement Therapy (RRT) to sustain the life of the patient. One could think that the easy way out would be to do a kidney transplantation (third treatment modality), but not everyone is lucky enough to find a donor and the queues are so long even in first world countries. This is because the donors are always lesser than the number of CKD patients worldwide. Only around 20% of the CKD patients who are in the organ reception waitlist, actually get a kidney transplantation while the other 80% relies on other treatment modalities [65]. Even after all these challenges, if one successfully finds a donor, there could be instances of newly implanted organ rejection too.

Dialysis which means ‘separation’, refers to the process in which the blood from the human body undergoes therapeutic purification artificially. Some of the functions of dialysis [110] are given below:

- Toxin management
- Fluid level management
- Electrolyte balance
- Correcting metabolic acidosis
- Arterial blood pressure management

The reader understands from the above list that dialysis is an essential life sustaining therapy for ESRD patients, but it does not replace all the basic functions of the kidneys. The underlying phenomenon occurring in a dialyzer is a simple bidirectional mass transfer (diffusion, convection or both). Normally, blood flows on one side of

the semi-permeable membrane and the cleaning fluid called ‘dialysate’ flow on the other side in a direction opposite to the direction of blood flow.

The patients might still need additional doses of synthetic hormones, vitamin D and other medicines. In addition to this , the doctors might advice patients to limit the intake of fluids, electrolytes and protein to minimize the weight gain due to fluid overload between dialysis sessions [2], [34]. The two broad classifications of dialysis are defined below, but more attention is given to hemodialysis as it is the focus of study of this work.

a. Peritoneal Dialysis (PD)

Peritoneal Dialysis (PD) is an intracorporeal (within body) technique in which a sterile cleansing fluid is injected into abdominal cavity through a tube (catheter). The lining of the abdomen (peritoneum) acts as the filter and removes excess water and metabolic wastes from the blood. After the prescribed period of time, the fluid with toxins is removed from the abdomen and is discarded and this kind of dialysis can be done at home, at work or even while travelling and giving the patient a greater degree of flexibility. This type of dialysis is carried out more frequently (4 to 5 times) in a day or usually overnight [78], [150].

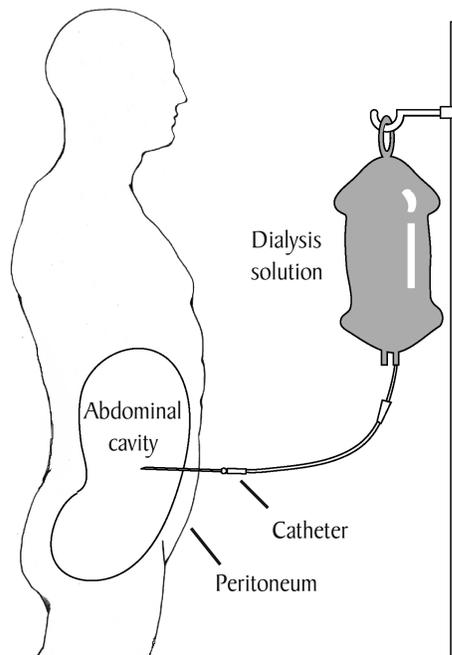


Figure 1.4: A pictorial representation of Peritoneal Dialysis (PD). Taken from National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (www.kidney.niddk.nih.gov)

b. Hemodialysis (HD)

Hemodialysis (HD) is an extracorporeal (outside body) technique in which the blood to be purified is taken to an external circuit element called as a ‘hemodialyser’, or simply a ‘dialyzer’. The toxic wastes and the overloaded fluid from the blood is removed in the hemodialyzer and then sent into the patient’s body. Dialysis was first described in 1854 by Thomas Graham, known as the ‘Father of dialysis’, and the milestones in the development of the HD process are well documented in literature [10], [40], [99].

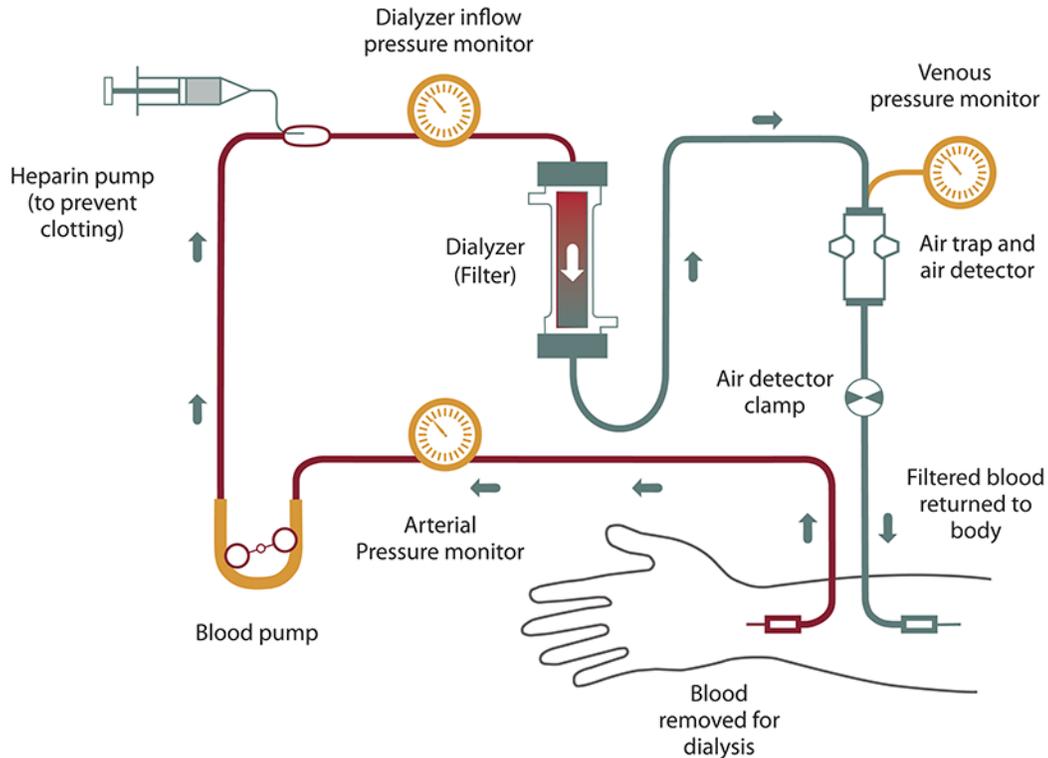


Figure 1.5: A simple layout of Hemodialysis (HD) setup. Taken from National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (www.kidney.niddk.nih.gov)

As depicted in figure 1.5, an anticoagulant (Heparin) is added to the blood side to prevent the blood from clotting during the treatment process. The dialysate is free of toxins and it should be strictly in the right specifications [20] as guided by the governing medical authority. The concentration of important electrolytes in the dialysate is set according to the required direction of flow (high for addition and zero to low for removal). The waste products being at a higher concentration in the blood, diffuses through the semipermeable membrane in the hemodialyzer to the dialysate side due to a concentration gradient. During the entire filtration process, blood

flows in and out of the hemodialyzer continuously, while most often the dialysate is used only once or fresh dialysate is sent through the hemodialyzer every time. This ensures that there is always a gradient for clearance of substances from the blood. Presence of a blood pump on the blood side circuit increases the pressure on the blood side and excess water from the blood flows into the dialysate side because of the transmembrane difference in pressure. Toxin and fluid management targets are achieved by a combination of both diffusion and convection (ultrafiltration). The reader can look at figure 1.6 to get a visual understanding of how the mass transfer processes work.

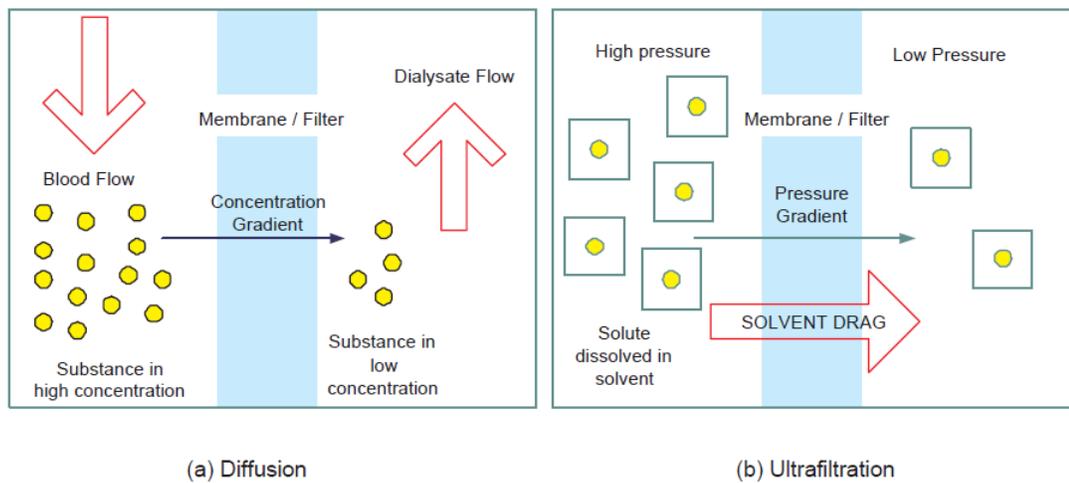


Figure 1.6: The mass transfer processes in HD. Taken from [60]

On a global perspective, the number of dialysis patients is constantly rising at a rate of 6% annually [65] and HD is the most common treatment modality with almost 90% of the dialysis patients undergoing HD and so it is the focus of this research work. To support our argument, let us look at the data from the European Renal Association (ERA) for the year 2018 [6]. Roughly over 80,000 patients began their RRT for ESRD (an overall unadjusted incidence rate of 129 per million population). By the end of the same year, approximately 569,000 patients were on RRT (unadjusted prevalence rate of 897 per million population) in total, of which the majority were men. Most of the new ESRD patients were primarily kept on a dialysis regime for at least 3 months, while waiting for a transplantation. At the end of 2012, if the reader looks on a broader perspective examining figure 1.7, 57% of the ESRD patients were on HD, 5% on PD and the rest 38% were living with a transplanted kidney. Most of the developed countries could follow a similar pattern and so advanced scientific understanding of

the hemodialysis treatment modality which contributes to a major chunk of the RRT becomes quintessential.

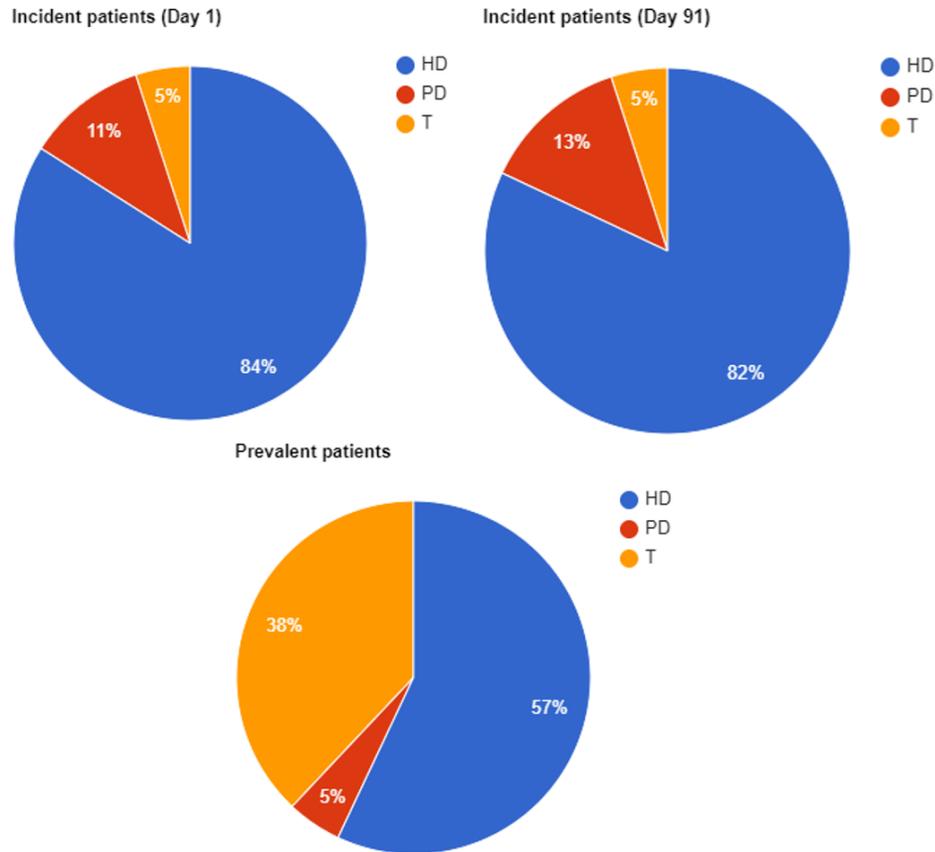


Figure 1.7: Percentages of patients in Renal Replacement Therapy (RRT) across different treatment modalities in Europe 2018 (HD: Hemodialysis, PD: Peritoneal dialysis, T: Transplant). Data from [6]

1.2.4 Clinical challenges during hemodialysis

The patients on HD receive intermittent treatments, around 3 to 4 times a week, removing a prescribed amount of overloaded fluid from the blood plasma by ultra-filtration for meeting the dry weight targets. The clinical prescription of dialysis has the information about the flow rates of blood, dialysate, composition of the dialysate, frequency and the length of the treatment [110]. Though the patient receives good renal treatment therapy, they are subjected to severe cardiovascular complications [37], [119] because of the quick removal of solutes and fluids during the short treatment windows (typically 3 to 5 hours). One can see from figure 1.8, that ESRD patients are always in a state of fluid overload between treatments as

the excretory function of the kidneys is completely lost. This is one of the reasons for cardiovascular complications in HD patients. Additionally, the fluid removed in the dialyzer is directly from the circulatory blood volume, and so the hemodynamic stability depends on how quick vascular refilling happens [103], [106], [128] and the regulatory action of the autonomic nervous system.

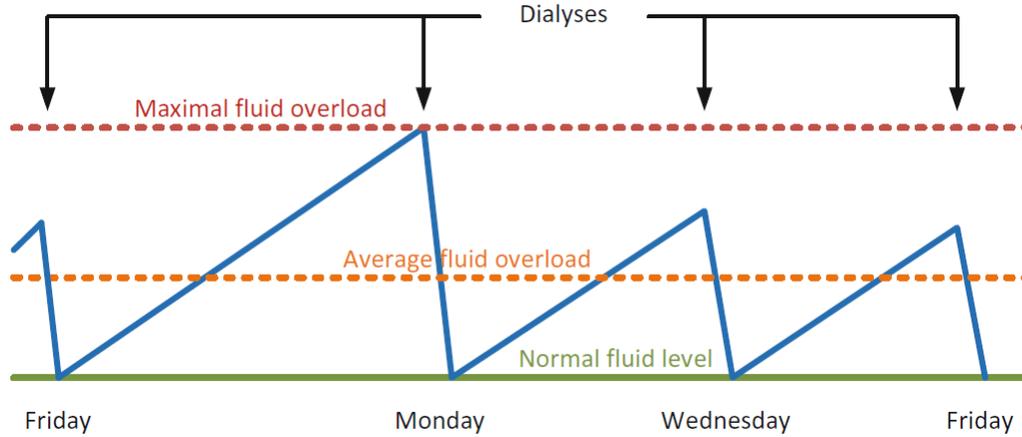


Figure 1.8: Fluid overload chart of a patient undergoing HD 3 times a week. Taken from [110]

Also the HD procedure could induce burden on the hemodynamic stability of the patient within or after the treatment session. A majority of the patients can handle the sudden drop in blood volume by HD without exhibiting fluctuations in blood pressure [27], [136] while for some the hemodynamic stability collapses. Some of the clinical problems encountered during HD deserve an explanation.

a. Intradialytic hypotension (IDH)

This is most frequently occurring hemodialysis complication and it is encountered in almost 30% of the HD patient population [159]. The US National Kidney Foundation's Kidney Disease Outcomes Quality Initiative (KDOQI) defines IDH as a decrease of Systolic Blood Pressure (SBP) of at least 20 mmHg or a decrease of mean arterial pressure (MAP) of at least 10 mmHg and this can even lead to the collapse of the circulatory system [151]. It is defined not only based on the drop in blood pressure but also based on the frequency of blood pressure drops by some researchers [8]. The causes of IDH are complex, sometimes asymptomatic [23] and very patient specific and so it stands as the major problem in the dialysis treatment centres [24].

b. Intradialytic hypertension

Intradialytic hypertension is the exact opposite to IDH in definition and it occurs less frequently, around 15%, than IDH [58]. One cannot find a textbook definition for intradialytic hypertension but some other authors say that it has occurred when the MAP has risen to greater than or equal to 15 mmHg or an increase in SBP of at least 10 mmHg [59] from the pre dialysis values. Similar to IDH, some investigators define intradialytic hypertension based on the frequency of occurrence over a time period (over 6 months) [142], rather than defining it based on a single occurrence. The causes of intradialytic hypertension are not very well understood in a clinical sense, yet this has resulted in more adverse outcomes and higher mortality rates [143].

c. Dialysis Disequilibrium syndrome (DDS)

The Dialysis Disequilibrium Syndrome (DDS) is a less frequent but serious complication of HD, which still needs to be investigated a lot [157]. This could lead to seizures, coma or even death. The possible explanation of DDS could be the HD induced sudden drop in blood plasma osmolarity (concentration), leading to a shift of water from plasma to the intracellular compartment (brain tissue and cells) resulting in cerebral oedema. It occurs more frequently in new ESRD patients on HD, if they are dialysed too fast [34], [157].

d. Technical complications

HD system involves not only the patient but also a list of other mechanical devices like pumps, tubes. There could be a lot of technical complications on the treatment side, completely out of the patient's purview, like clotting of blood in the extracorporeal circuit, air embolism, blood access complications, blood line leaks, blood line disconnections and dislodging of dialysis needles [135]. Some of these complications could be avoided if the dialysis equipment is built with appropriate safety systems and kept in a well maintained condition and proper checks are made before starting and while the treatment is in progress. Continuous training of the clinical personnel involved in giving dialysis care is mandatory to avoid these issues.

1.2.5 Attempts to mathematically model the elements of a hemodialysis system

A hemodialysis system consists of the patient who is undergoing the treatment, the hemodialysis machine (data processing system, water purification system, control

system, safety monitoring system), the hemodialyzer and several other minor components. For the purpose of modelling, if one looks at the hemodialysis system, there are only a few subsystems which need to be modelled to study the integrated dynamics and interactions between different subsystems shown in figure 1.9.

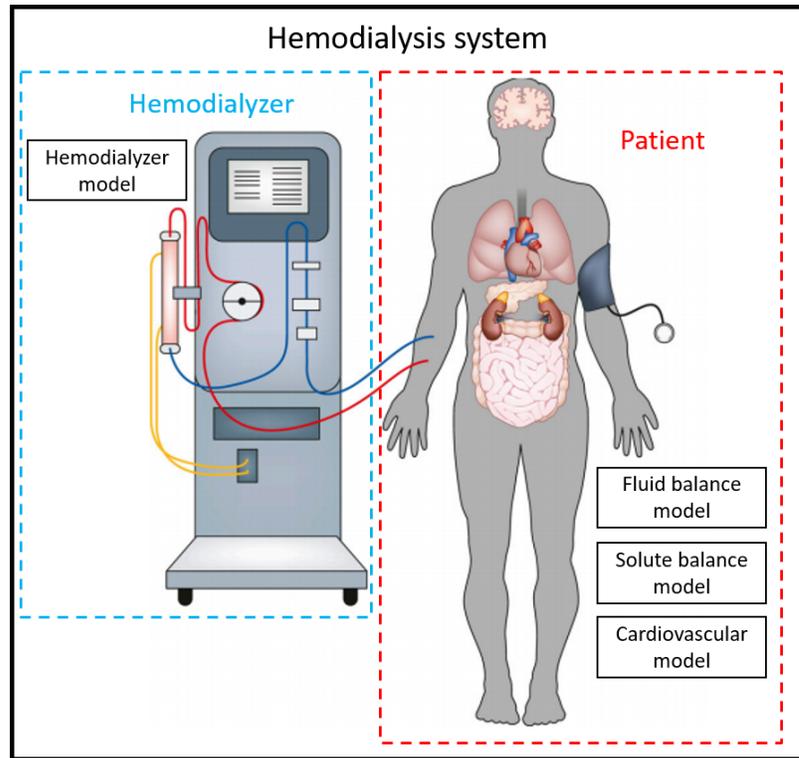


Figure 1.9: A schematic representation of the HD system

One might wonder that the underlying phenomenon behind HD is just a simple bidirectional mass transfer process, a combination of diffusion and convection (ultrafiltration), but the fluctuations induced by HD in the body of a patient disturbs the complete harmony between different body fluid compartments and the whole cardiovascular system and are quite complex to predict, analyse and comprehend. HD is nothing but an example of an external systemic perturbation, which causes the patient's body to deviate from a pre-treatment state and tries to bring the body back to equilibrium (homeostasis) before the end of the treatment. But with the help of a mathematical model capable of quantitatively describing the perturbations (cardiovascular response, fluid and solute kinetics) in an effective way, the HD process can be well studied. With the aid of a versatile mathematical model of the HD system, the researcher will not only be able to investigate different HD treatment settings but also will be able to understand the intricate physiological mechanisms taking place

within the patient's body and find answers to several unanswered questions in the context of HD. Even from a systems engineering perspective, a mathematical model (white, grey or black box) becomes quintessential if the reader wishes to look at possible opportunities for the application of advanced process control strategies on a HD system.

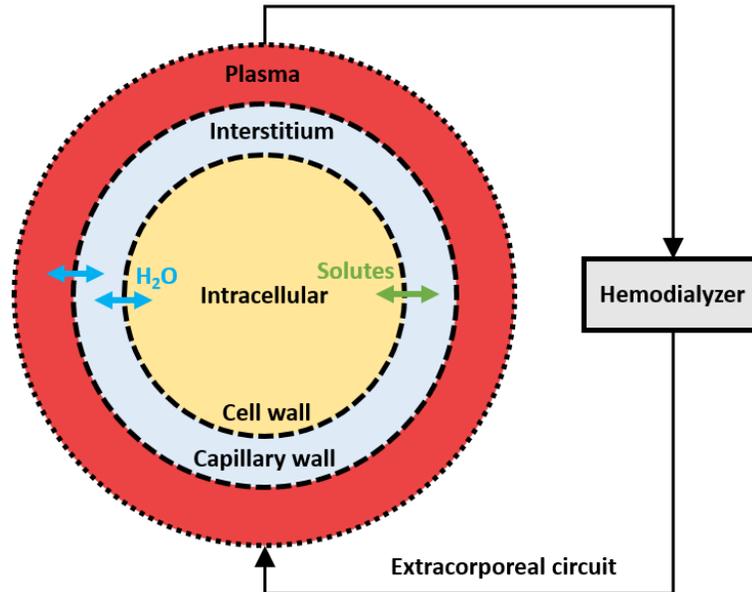


Figure 1.10: Pictorial representation of the 3 compartment model of the human body

There are several models in literature to study the solute and fluid (water) dynamics, derived to critically analyse the blood volume changes during infusion therapy [18], [49], [149] or to study the HD induced perturbations [90], [104], [105], [138], [148]. Most of these models look at the human body as multiple compartments as shown in figure 1.10. For instance, all the cells and tissues are lumped into one intracellular compartment and anything exterior to the cells are lumped together into an extracellular compartment. The transport of solutes and fluid between different compartments arrive at ordinary differential equations if one starts deriving them from scratch using mass and volume balance principles.

The use of mathematical models to represent hemodialysis therapy has come a long way [34], [39]. The first HD adequacy index KT/V (Dialyzer clearance multiplied by the treatment duration divided by the total fluid distribution volume) was introduced in the year 1985 [47], after analysis of clinical data using a one compartment kinetic model of urea [121]. With the scientific advancement in the design and utilization of better hemodialyzers, higher clearance rates were achievable in shorter duration of

time. The perturbations induced by faster clearance were more strong and necessitated the introduction of a two compartment model for the marker solutes, urea and creatinine [87], [158]. Urea was found to be perfused differently in different organs of the body and so an alternative to the two compartment solute models was proposed by some authors [126], [129].

If one looks at figure 1.10 again, the splitting of the extracellular compartment into interstitium and plasma is mainly because of the need to understand the cardiovascular dynamics of a patient undergoing HD. Plasma is the circulating fluid volume through the vessels in the body. There are several models in literature to study the cardiovascular dynamics and the patient's inherent regulatory mechanisms for different cases [52], [100], [101], [109], some particularly with a focus on HD [53], [110], [140]. In those models, similar vascular compartments are grouped together. The number of vascular compartment groupings equals the number of compartments of the cardiovascular model. For example, all arteries are grouped into one arterial compartment and all veins are grouped into one venous compartment. Depending on the necessity and the problem under consideration, a researcher has the freedom to choose between simple and more advanced models with pulsatile blood pressure output [35], [114].

Some authors have focused exclusively on deriving the equations for mass transfer dynamics in the hemodialyzer. The researcher has the choice to select a simple model [86], [122], [147] to compute the exit concentration of the solutes from the dialyzer at every time instant or a complex model [16], [82] (Partial differential equations) to compute the spatio temporal concentration profile of the solutes at every time instant during HD. The choice of the nature, type and complexity of the mathematical models to represent the individual elements of the HD system is solely dependent on the requirements set by the problem statement.

1.2.6 Selection and estimation of variables of a mathematical model

Mathematical models are usually composed of relationships and variables, written down in the form of equations. The accuracy and flexibility of the mathematical model determines whether the mathematical model will be used confidently by the researchers and investigators. The accuracy of the model in turn depends on the accuracy of the variables (parameters and states). We focus on studying HD and its associated symptoms and impacts on a patient in this work and so in our case, the patient subsystem models should be initialized with meaningful physiological values

of variables (parameters and states) to give meaningful model predictions. There are nominal values of physiological variables of patient groups available in literature [51], but this might not work for all. Each patient undergoing HD behaves differently, and this makes the problem even more interesting.

State and parameter estimation problems arise in several areas of advanced process systems engineering (modelling, control, process monitoring and fault diagnosis). These concepts have been applied in a wide range of industries, starting from a simple batch chemical plant to the most advanced aerospace, power, petroleum and petrochemical industries [7], [69], [74], [108]. In literature, the reader can find many attempts by researchers to improve the performance of the existing state and parameter estimation algorithm and to develop new efficient algorithms [108], [141], [155]. In general, there can be two different approaches to solve the state and parameter estimation problem. The first one being a sequential, separate estimation of parameters and states [108], [153]. In this approach, the parameter estimation is carried out first and updated when new data comes into the system. The second style of tackling this problem would be to do simultaneous, joint estimation of parameters and states. This approach has caught the attention of a lot of researchers and offers superior performance too [22], [50], [56], [64], [134]. A common way of solving the state and parameter estimation problem using the second approach, is to augment the parameters as additional states of the system [21], [81], [146] and looking at the problem as a state estimation problem.

Observability of a system plays a crucial role in the estimation of states and parameters of a system. In practice, it is not very easy to test the observability of a nonlinear system model, like for the one we have adopted in this work to represent the HD system. The test involves computationally demanding steps like the calculation of higher order Lie derivatives and the test results are sensitive to noise [93], [133], [144]. The observability test of a nonlinear system can be also done using approximate, alternative approaches through linearization of the nonlinear system, sensitivity analysis and structural observability [85], [145]. Each of these methods come with both advantages and disadvantages. If the original system model under consideration fails the observability test, then only a subset of variables can be estimated from the original list of variables (states and parameters). There are several articles focusing on variable (parameter) selection and estimation [42], [48], [96], but this work [85] shall be adopted for building a sensitivity analysis based framework, for the selection and estimation of hemodialysis system model variables in this thesis work.

1.2.7 Feedback control in hemodialysis

The scientific advancement of the tools used for intradialytic monitoring, with an ambition to prevent clinical complications induced by HD and to continuously measure hemodynamic process variables, has led to a significant improvement in the way the HD therapy is delivered to CKD patients over the last few decades. These tools were not considered seriously for clinical applications until recently. Traditionally, the clinical care give would determine the time of dialysis (T), clearance (K) for the estimated patient volume (V), using ‘ KT/V ’ as the dosage prescription for dialysis. Due to several process disturbances and treatment complications, there is always a difference between the delivered and the prescribed dose, resulting in the treatment objectives not being met fully [12]. The traditional therapy has always worked in an open loop fashion, with no information fusion of the physiological condition of the patient. Researchers have looked at the application of the closed-loop strategy in HD, by integrating different techniques into the dialysis computer machine over the last few decades. These closed loop systems continuously monitor the patient’s physiological conditions and automatically adjust the parameters of the treatment. Some of the feedback systems (blood volume, blood temperature, arterial pressure, ionic dialysance) in HD, which are currently in use, are reviewed here.

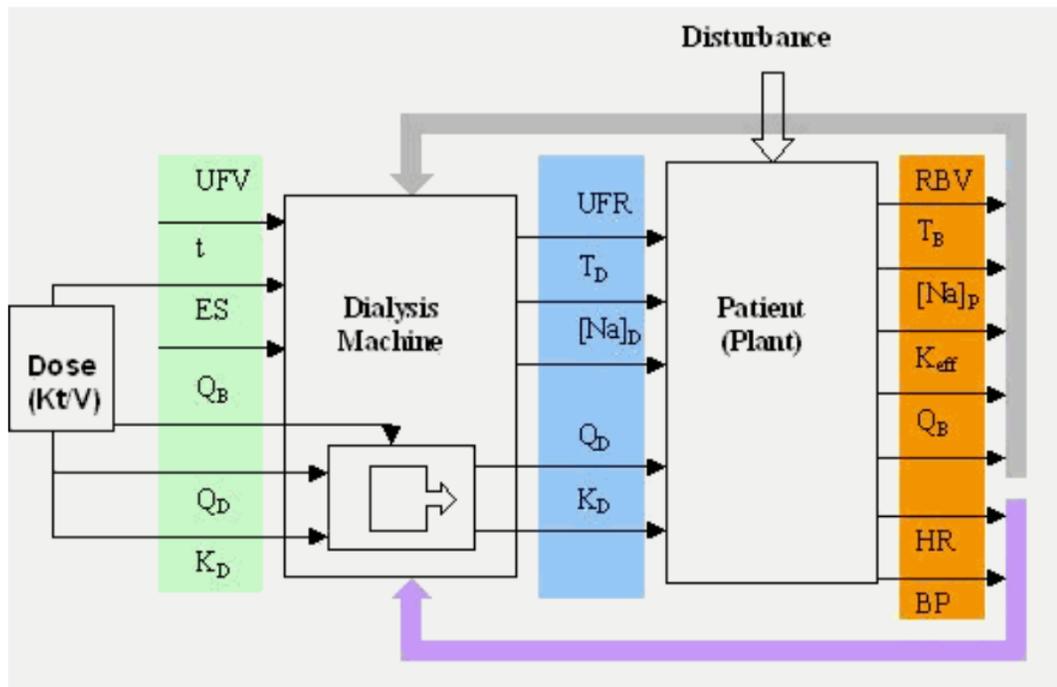


Figure 1.11: Feedback control of hemodialysis. Taken from (www.uninet.edu)

A group of investigators along with Gambro-Hospal research group, put forth their automatic Blood Volume (BV) control system which varies the ultrafiltration in a continuous fashion [118]. This system was capable of guiding the BV along a predefined trajectory by continuously changing two inputs, Ultrafiltration Rate (UFR) and Dialysate Sodium Concentration (DSC) [116], [117]. This kind of strategy has got the support of some other authors too [41], [115]. In spite of active research, there are only two feedback control systems which are commercially available, the Hospal and the Fresenius systems [12]. Temperature control has also been investigated by some researchers with a motive of preventing heat accumulation, which rises the body temperature of the patient during HD [125], [127]. The dialysate temperature is also one other variable to be individualised. The thermal balance and the patient's body temperature can be controlled by the Blood Temperature Monitor (BTM), available in Fresenius machines. The commercially used controllers for blood volume and temperature control are of Proportional, Integral, Derivative (PID) type. Active research has been done in arterial blood pressure feedback control systems based on fuzzy logic systems too [91], [98], [124]. These fuzzy systems work on a set of rules, rather than having a mathematical model of the system, mimicking how a doctor would make decisions in the event of occurrence of any clinical complications. Ionic dialysance and the plasma conductivity of the patient can be computed easily from on-line measurements at two different steps of dialysate conductivity. A feedback system has been designed to identify the plasma conductivity of the patient and modulate the conductivity of the dialysate continuously in order to achieve a desired patient plasma conductivity corresponding to a desired plasma sodium concentration at the end of the treatment [12], [70], [102].

The number of literature works focusing on the application of advanced feedback control strategies like model predictive controllers are very few and can be hand numbered [60], [156]. In conclusion, the reader has to understand that the adaptive control of the hemodialysis treatment session using feedback systems will enhance the process of renal replacement therapy and make it more physiological and safe but we still have a long way to go.

1.3 Research background

1.3.1 Motivation

According to the Kidney Foundation of Canada, approximately 10% of the Canadian population have kidney disease and millions more are at risk. The number of people

living with kidney disease has grown 36% from 2007 to 2018. The most common form of treatment modality for kidney failure in Canada is dialysis. Nearly 48,000 Canadians are being treated for kidney failure and 58.4% of those patients are treated on hemodialysis (HD). Across all provinces in Canada, HD remains the most frequently used treatment modality for new patients who require dialysis [97]. In 2013, the rate of patients with End Stage Renal Disease (ESRD) initiated on HD varied from 91% in Newfoundland and Labrador to 71% in Manitoba. Moreover, in the same year, most Canadian dialysis patients (76%) received in-centre HD, which describes HD performed in an institution such as a hospital, satellite unit, or a dialysis facility, with the assistance of a health care professional.

Based on the literature review presented, the reader is prompted to understand that there have been several attempts to model the hemodialysis system to study the integrated dynamics. But there has not been any progress towards building a systematic framework for data (patient history: clinical measurement data) based identification of the first principles model, with an agenda of building an individualized virtual patient simulator (grey box modelling approach). If done, that would be the first step towards achieving the so-called ‘Precision medicine’ in HD. As presented before, the identification of an individualized patient model is essentially a simultaneous state and parameter estimation problem and the observability of the nonlinear system model has to be ensured for consistency and so we have adopted a sensitivity analysis framework, which could be easily implemented and is computationally more friendly than finding Lie derivatives. On top of that, majority of the feedback control systems explored by researchers, are PID controllers which do not have the ability to take optimal control actions and handle process constraints and they follow a reference trajectory (BV: exponential profile) defined to be clinically safe by doctors after conducting randomized control trials. It is quite interesting to note that, even in the works on the application of MPC to HD, the model-based controllers were forced to follow clinically determined reference trajectories and also the control inputs were forced into taking only certain values towards the end of the treatment (funnel shaped constraints). It can be argued that the clinically determined reference trajectories are always not optimal.

In addition to all the reasons stated above, HD is often considered the most expensive medical intervention that society will pay for on an ongoing basis [72]. According to the Kidney Foundation of Canada, dialysis costs the health care system between \$56,000 and \$107,000 per patient per year. The cost to the health care system for

chronic kidney disease is approximately \$50 billion per year. This situation is getting more severe due to the limited treatment resources, qualified doctors (Total 715 registered nephrologists in Canada, equates to roughly 1.9 nephrologists per 100,000 population [5]) and increasing number of CKD patients. Also, people living with kidney failure (ESRD) frequently face significant financial challenges related to increased medication costs due to longer treatment time. Hence, it is necessary to increase the dialysis equipment utilization efficiency from a resource optimization perspective.

1.3.2 Problem statement

Hemodialysis, although being a life-sustaining therapy for ESRD patients, is found to induce several clinical complications, like IDH, DDS, in patients. Each patient behaves differently to HD and so the need for identifying an individualized treatment regime gets amplified. In the past, the hemodialysis treatments were done in an open-loop or semi automated fashion where the treatments were stopped when clinical complications occurred and started again after the patient returns to normalcy. These kinds of treatments might not be optimal and safe from a clinical point of view and often result in under treatment. The problem preventing optimal control techniques to be applied in HD is the unavailability of a comprehensive framework of modeling, control and optimization for the treatment process. This research work is an attempt to address this multifaceted problem in a systems engineering perspective and will provide benefits not only to the large population of CKD patients through improvements in the quality and safety of the delivery of treatment and clinical care, but also to the scientists and researchers who are trying to develop a scientific understanding supported by first principles models of this clinical phenomena and its associated symptoms. It will also provide benefits to the nephrology health care providers through improvement in operating efficiency, which will result in improved business and financial performance, and a more attractive economic return.

1.3.3 Objectives

With the motivation of this research presented already, this thesis work aims to accomplish the following objectives:

- To develop a control relevant first principles model of the HD system with integrated dynamics (cardiovascular, solute and fluid dynamics).
- To devise a systematic framework for building an individualized virtual patient simulator from limited clinical data.

- To design a feedback control strategy to automatically control the treatment parameters, thereby achieving clinical treatment objectives, while ensuring superior safety and hemodynamic stability of the patient undergoing HD.

If the above objectives are met, a comprehensive framework for tailoring HD treatments according to the individual patient's needs will emanate.

1.4 Thesis organization

1.4.1 Outline

This thesis is organised as follows:

- Chapter 1 provides the reader with the necessary background information to appreciate this piece of work and a concise literature review of CKD, RRT modalities and some complications associated with it, previously developed mathematical models of the HD system, parameter and state estimation algorithms and feedback control techniques applied in HD treatments.
- Chapter 2 introduces the reader to the basic concepts of observability, sensitivity analysis, Singular Value Decomposition (SVD), orthogonalization, design of an Extended Kalman Filter (EKF), Proportional Integral Derivative (PID) controller and Zone Model Predictive Controller (ZMPC).
- Chapter 3 acquaints the reader with the mathematical model used for representing the HD system in the course of this study, along with a few modifications done to make it control relevant. Additionally, a mass transport model (Partial differential equations) of the hemodialyzer is explored for practical implementation in future.
- Chapter 4 focuses on the identification of an individualized virtual patient simulator from synthetic clinical data, which is essentially a simultaneous state and parameter estimation problem, solved using the techniques outlined in chapter 2. The results of the sensitivity based observability test performed on the HD model are also presented here.
- Chapter 5 talks about the application of optimal control strategies as discussed in chapter 2, a model based controller with process constraints integrated with an estimator, with feedback implementation for a HD treatment setting. Then

the proposed control algorithm is tested on the virtual patient simulator and the results are presented. The proposed algorithm is also compared with traditionally prevalent feedback control strategies in HD.

- Chapter 6 summarizes the work presented in this thesis and presents some possible improvements and future research directions which can be pursued to tailor the treatment according to each individual patient, making HD treatments even better, safer, comfortable and more quicker than it is today.

1.4.2 Contributions

The contributions of this thesis are as follows:

- A well-accepted nonlinear mathematical model of Ursino and Innocenti [140] for studying the integrated fluid, solute, cardiovascular and mass transfer dynamics during HD is adopted from literature and modified in such a way to give the investigators and HD treatment providers the freedom of having upto 6 control inputs or even more. This is one of the first attempts in HD research, where a first principles model of HD is realized in a optimal control sense with practical feasibility of implementation in the near future.
- A framework for nonlinear system observability based on sensitivity analysis, with a variable selection algorithm to ensure consistent estimation (state and parameters simultaneously) even for systems which show signs of unobservability. This is done for the first time with respect to HD and this paves the way for the development of individualized grey box patient simulator models in HD with the available non invasive clinical measurement data.
- Development of a computer controlled HD system constructed out of a novel control methodology based on Batch Zone Nonlinear Model Predictive Control (BZNMPC) with physiological constraints and treatment objectives. This is the first time a NMPC of this kind, where the model in the controller is completely based on first principles with cardiovascular dynamics predictive capability, is used in HD.

Chapter 2

Preliminaries

2.1 Terms and definitions

Firstly, in this section, the important terms and definitions that will be frequently used in this research work are defined. After this section, the reader will be introduced to all the fundamental technical concepts which are utilized for building the proposed methodologies in this research work.

- **Hemodialysis system:** If one neglects most of the intricate mechanical equipment as shown in figure 1.9, the hemodialysis system consists of two subsystems namely, the patient and the hemodialyzer.
- **Patient:** The person with CKD undergoing HD treatment.
- **Hemodialyzer:** The fibre module where the extracorporeal cleansing of the blood occurs. It is also called as ‘artificial kidney’.
- **Dialysate or dialysis fluid:** The cleaning solution consisting of water and chemicals (electrolytes) which passes through the hemodialyzer to remove excess fluids and waste products from the blood.
- **Solute:** A solute is a substance dissolved in a solution. In our problem, the solutes could be toxins (Urea) or essential nutrients (Sodium, Potassium) dissolved in blood plasma (solvent).
- **Fluid overload:** The condition in which excess fluid (water) in the body causes edema.

- **Dry weight:** It refers to the original weight (ideal) of the patient when overloaded fluid has been removed.
- **Ultrafiltration:** The migration of fluid across a semipermeable membrane because of transmembrane pressure gradient. In a dialysis prescription, the total fluid volume to be extracted is called ‘Ultrafiltration Volume (UFV)’.
- **Dialysance:** The number of milliliters of blood completely cleared of any substance by a hemodialyzer in a unit of time. Mathematically, it could be expressed as shown below [122].

$$\begin{aligned} \text{Dialysance} &= \frac{\text{Change in solute content of incoming blood}}{\text{Concentration driving force}} \\ &= \frac{Q_{bi}(C_{bi} - C_{bo})}{C_{bi} - C_{di}} \end{aligned}$$

where Q_{bi} is the blood flow rate into the hemodialyzer, C_{bi} is the inlet concentration of a solute in blood, C_{bo} is the exit concentration of the solute in blood and C_{di} is the inlet concentration of the solute in dialysate.

- **Clearance:** It is the dialysance when C_{di} becomes zero.
- **Urea Reduction Ratio (URR):** It is the HD treatment induced reduction in the blood concentration of urea over time. The target is to reach an URR of great than 65%. It can be computed from the expression shown below.

$$URR = 100\% \times \frac{C_{urea}^{pre} - C_{urea}^{post}}{C_{urea}^{pre}}$$

where C_{urea}^{pre} is the predialysis blood urea concentration and C_{urea}^{post} is the post-dialysis blood urea concentration.

- **Infusion rate:** It is the amount of fluid administered to the patient intravenously or through the blood stream. It is also called as ‘injection rate’ or ‘saline infusion rate’ if a hypertonic saline solution is infused.

2.2 Representation of the HD system model

The adopted mathematical model of the hemodialysis system can be written as a set of nonlinear ordinary differential equations in the general form as shown below,

$$\dot{x}(t) = f(x(t), u(t), \theta) + \omega(t), \quad x(t_0) = x_0 \quad (2.1)$$

$$y(t) = h(x(t), u(t), \theta) + \nu(t) \quad (2.2)$$

In the outlined compact representation, $x(t) \in R^{N_x}$, $u(t) \in R^{N_u}$, $y(t) \in R^{N_y}$, $\theta \in R^{N_\theta}$ denote the state, input, output and parameters at time t , respectively. $\omega(t) \in R^{N_x}$ and $\nu(t) \in R^{N_y}$ represent the model disturbance and measurement noise at time t respectively. $f(\cdot)$ and $h(\cdot)$ denote the nonlinear state and output equations respectively. x_0 is the state of the system at the initial time t_0 . In this work, the parameters of the HD system (θ) are assumed to be constant and time invariant. Also, the total number of parameters of the HD system is a sum of patient-specific parameters (θ_p), which are to be estimated, and other known system parameters (θ_k) as shown in Eq. (2.3).

$$N_\theta = N_{\theta_p} + N_{\theta_k} \quad (2.3)$$

2.2.1 Augmentation of the HD system model

Out of the parameter vector of the HD system, the patient-specific parameters were removed and augmented as additional states of the system with zero dynamics [145], as by our assumption all parameters are time invariant. For the sake of presentation simplicity, the time index, disturbance and noise are omitted.

$$\begin{aligned} \begin{cases} \dot{x} = f(x, u, \theta) \\ y = h(x, u, \theta) \end{cases} &\Rightarrow \begin{cases} \dot{x} = f(x, u, \theta_k) \\ \dot{\theta}_p = 0 \\ y = h(x, u, \theta_k) \end{cases} &\Rightarrow \begin{cases} \dot{X}_a = f_a(X_a, u, \theta_k) \\ y = h_a(X_a, u, \theta_k) \end{cases} \end{aligned} \quad (2.4)$$

In Eq. (2.4), $X_a = [x^T \ \theta_p^T]^T \in R^{N_x + N_{\theta_p}} \in R^{N_{x_a}}$ is the augmented state vector. $f_a(\cdot)$ and $h_a(\cdot)$ denote the augmented nonlinear state and output equations respectively.

2.3 Observability of a system

In control theory, the observability and controllability of a linear system are mathematical duals. Let us consider a discrete time system, as shown below, to define the notion of observability.

$$x(k+1) = Ax(k) + Bu(k) \quad (2.5a)$$

$$y(k) = Cx(k) \quad (2.5b)$$

The system is said to be observable if, for any possible sequence of state and control vectors, the current states can be determined uniquely in finite time using only the outputs. We can check if the system is observable by checking whether the following observability matrix (O), built by using the system matrices A and C , is full rank or

not [112]. In Eq. (2.6), n is the size of the state vector x .

$$O = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad (2.6)$$

If the above observability matrix is of full rank, we can uniquely determine the initial states of the system based on the input and output data, and we say that the system is observable. If the observability matrix is rank deficient, all system states cannot be estimated simultaneously. For a nonlinear system like the HD system model, it is in general challenging to check the observability, as it involves the computation of Lie-derivatives and Lie-brackets [133]. In literature, there are some alternative approaches to approximate the nonlinear system's observability based on the linearization of the nonlinear system successively along typical trajectories and checking the observability of the linearized models (using Popov–Belevitch–Hautus test) [22], or using sensitivity equations [85], [133].

2.4 Sensitivity analysis

By definition, sensitivity analysis aims to study the uncertainty in the output of a mathematical model or a system through a rigorous numerical approach by splitting the uncertainty into different possible input sources. It is widely used in the field of systems engineering and there are two broad classifications of sensitivity analysis namely, local and global sensitivity analysis. A local structural observability test could be formulated based on a rank-test of the sensitivity matrix. A full rank sensitivity matrix is a sufficient condition for observability [85], [133]. In this section, the fundamental equations which are required to construct the aforementioned sensitivity matrix are outlined. Omitting θ_k from Eq. (2.4) for simplicity, the dynamics of the augmented HD system model can be represented by the following set of equations.

$$\dot{X}_a = f_a(X_a, u, X_{a0}), \quad X_a(t_0) = X_{a0} \quad (2.7)$$

$$y = h_a(X_a, u, X_{a0}) \quad (2.8)$$

In the above equations, the initial states have been explicitly mentioned as variables. The sensitivity variables are defined as follows,

$$S_{X_a X_{a0}} = \frac{\partial X_a}{\partial X_{a0}}, \quad S_{y X_{a0}} = \frac{\partial y}{\partial X_{a0}} \quad (2.9)$$

The sensitivity ODEs which give the sensitivities of the initial states to the output (or measurements) are given below [48],

$$\dot{S}_{X_a X_{a0}} = \frac{\partial f_a}{\partial X_a} S_{X_a X_{a0}} + \frac{\partial f_a}{\partial X_{a0}}, \quad S_{X_a X_{a0}}(t_0) = I \quad (2.10)$$

$$S_{y X_{a0}} = \frac{\partial h_a}{\partial X_a} S_{X_a X_{a0}} + \frac{\partial h_a}{\partial X_{a0}} \quad (2.11)$$

$$\frac{\partial f_a}{\partial X_{a0}} = \frac{\partial h_a}{\partial X_{a0}} = 0 \quad (2.12)$$

The sensitivity equation (2.10) shall be solved in parallel with the original augmented system model equation (2.7). The initial state to output sensitivities ($S_{y X_{a0}}$) can be obtained by numerically integrating equations (2.7), (2.10) and substituting the results in equation (2.11). Then, the sensitivity matrix O_r , which is an approximation of the observability matrix of a nonlinear system, can be constructed as shown below.

$$O_r = \begin{bmatrix} S_{y X_{a0}}|_{t_0} \\ S_{y X_{a0}}|_{t_1} \\ \vdots \\ S_{y X_{a0}}|_{t_N} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial X_{a0}^1}|_{t_0} & \frac{\partial y_1}{\partial X_{a0}^2}|_{t_0} & \cdots & \frac{\partial y_1}{\partial X_{a0}^{N X_a}}|_{t_0} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{N_y}}{\partial X_{a0}^1}|_{t_0} & \frac{\partial y_{N_y}}{\partial X_{a0}^2}|_{t_0} & \cdots & \frac{\partial y_{N_y}}{\partial X_{a0}^{N X_a}}|_{t_0} \\ \frac{\partial y_1}{\partial X_{a0}^1}|_{t_1} & \frac{\partial y_1}{\partial X_{a0}^2}|_{t_1} & \cdots & \frac{\partial y_1}{\partial X_{a0}^{N X_a}}|_{t_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{N_y}}{\partial X_{a0}^1}|_{t_N} & \frac{\partial y_{N_y}}{\partial X_{a0}^2}|_{t_N} & \cdots & \frac{\partial y_{N_y}}{\partial X_{a0}^{N X_a}}|_{t_N} \end{bmatrix}_{N_y(N+1) \times N X_a} \quad (2.13)$$

In the matrix given above, N is the total simulation time steps. Then scale the sensitivity matrix using the nominal values (from [140]) of the variables (\hat{X}_{a0}^j) and this scaled sensitivity matrix (for simplicity O_r again) shall be used for further analysis [96], [154]. In addition to that, the user should feel free to adopt other scaling approaches (for instance, using the bounds of variables or by simply using a reasonable guess values of variables) as deemed to be suitable for the problem under consideration.

$$\frac{\partial y_i}{\partial X_{a0}^j}|_{t_k} \Rightarrow \frac{\hat{X}_{a0}^j}{y_i|_{t_k}} \frac{\partial y_i}{\partial X_{a0}^j}|_{t_k} \quad (2.14)$$

where $i \in \{1, 2, \dots, N_y\}$, $j \in \{1, 2, \dots, N_{X_a}\}$ and $k \in \{0, 1, 2, \dots, N\}$.

2.5 Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) is a general eigen decomposition (factorization) of a real or complex matrix and it is a well known approach to find the rank of a matrix [44]. The SVD of a real matrix (M) of shape $m \times n$ is given by,

$$M_{m \times n} = U_{m \times m} \cdot \Sigma_{m \times n} \cdot V_{n \times n}^T \quad (2.15)$$

Here U and V are real orthogonal matrices and the columns of U and V are called the left and right singular vectors of M respectively. The diagonal entries of Σ are the singular values of the matrix M and the number of non zero singular values determine the rank of matrix M .

2.6 Orthogonalization

The sensitivity based selection methods are widely used in parameter selection and estimation [3], [96], [154]. We can directly borrow them for the implementation of state and parameter selection process in this research work. The basic ideas and methods of variable (state and parameter) selection will not be discussed here in detail. The reader is motivated to look into some interesting works in literature to understand the nitty gritty of the variable selection process [28], [69], [96]. The revised forward selection procedure to sequentially choose the most important and estimable variables based on the orthogonalization method [89], [154] is given below.

- Step 1: Calculate the magnitude (i.e. the two norm) of each column of the relative sensitivity matrix O_r , and set $k = 1$;
- Step 2: Select the column with the largest magnitude as X_k , and mark its corresponding state or parameter as the first estimable variable;
- Step 3: Estimate the optimal effect of the selected states and parameters using X_k based on forward selection: $\hat{Z}_k = X_k(X_k^T X_k)^{-1} X_k^T O_r$;
- Step 4: Eliminate the effect of selected states and parameters from the original matrix, and calculate the residual matrix R_k : $R_k = O_r - \hat{Z}_k$;
- Step 5: Calculate the magnitude of each column in the residual matrix R_k , and choose the state or parameter with the largest magnitude as the next estimable variable;

- Step 6: Select the corresponding column of the next estimable variable in O_r , and augment X_k to X_{k+1} by including the new column;
- Step 7: Advance the iteration counter, and repeat steps 3 to 6 until one of the prescribed termination conditions is satisfied;
 1. All states and parameters have been ranked.
 2. The largest magnitude of the columns in R_k is smaller than a prescribed cut-off value.
 3. It is impossible to carry on the optimal effect estimation of X_k in step 3 due to matrix singularity (inverse computation).

2.7 Design of a nonlinear state estimator

2.7.1 Extended Kalman Filter (EKF)

In industrial practice, one has to measure a large number of physical variables to ensure the required level of state estimation performance. However, some physical quantities cannot be directly measured or measuring them is not desired in a cost reduction motive. For example in chemical process industries, there is a huge lag in the receipt of concentration measurements from fluid samples. In that case, one has to estimate the states of the system from the available online measurements and a generic block diagram of a state estimation algorithm is shown in figure 2.1.

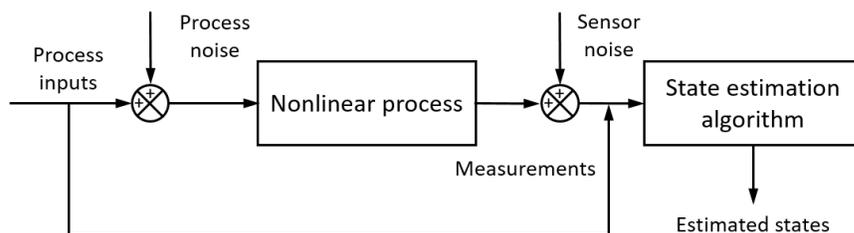


Figure 2.1: Block diagram of a state estimation algorithm

One of the methods used for the estimation of states of a nonlinear system is the Extended Kalman Filter (EKF), which is widely used in several practical applications, like sensorless control, fault-tolerant control of AC drives, energy systems, robotics, signal processing and industrial control systems, owing to this computational simplicity [9]. It is based on the successive linearization of the original nonlinear system.

In a broader perspective, the EKF algorithm can be divided into two fundamental steps, the first being a prediction step and the next being an updation step. In the prediction step, the state (x) of the system under consideration and the state covariance matrix (P) are predicted. New measurements come into the system from the sensors and when a new measurement is available, the Kalman gain (K) is calculated first and then x and P are updated. The Kalman gain is the relative weight given to the measurements and the current state estimate. With a high gain, the filter places more weight on the most recent measurements, and thus follows them more responsively. With a low gain, the filter follows the model predictions more closely. The steps involved in an EKF algorithm are summarized below.

1. Initialisation:

- (a) The filter is initialized as follows,

$$E(x_0) = \hat{x}_0 \quad (2.16)$$

$$E((x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T) = P(0|0) \quad (2.17)$$

2. Prediction step:

- (a) State prediction:

$$\hat{x}(t|t-1) = f(\hat{x}(t-1|t-1), u(t-1)) \quad (2.18)$$

The model disturbance is included in the state covariance prediction explicitly.

- (b) State covariance prediction:

$$P(t|t-1) = A(t)P(t-1|t-1)A(t)^T + Q \quad (2.19)$$

where $A(t) = \frac{\partial f}{\partial x} \Big|_{\hat{x}(t-1|t-1)}$ and Q is the covariance matrix of the model disturbance ω .

3. Update step:

- (a) Kalman gain calculation:

$$K(t) = P(t|t-1)C(t)^T [C(t)P(t|t-1)C(t)^T + R]^{-1} \quad (2.20)$$

where $C(t) = \frac{\partial h}{\partial x} \Big|_{\hat{x}(t|t-1)}$ and R is the covariance matrix of the measurement noise ν .

(b) State update:

$$\hat{x}(t|t) = \hat{x}(t|t-1) + K(t) \cdot (y(t) - h(\hat{x}(t|t-1))) \quad (2.21)$$

The state vector x of the system is updated when new measurement data $y(t)$ is available.

(c) State covariance update:

$$P(t|t) = (I - K(t)C(t)) \cdot P(t|t-1) \quad (2.22)$$

State covariance matrix P is updated. I is the identity matrix with dimension N_x .

2.8 Design of feedback control algorithms

In this section a brief overview of the feedback control strategies tested in this research work will be elucidated.

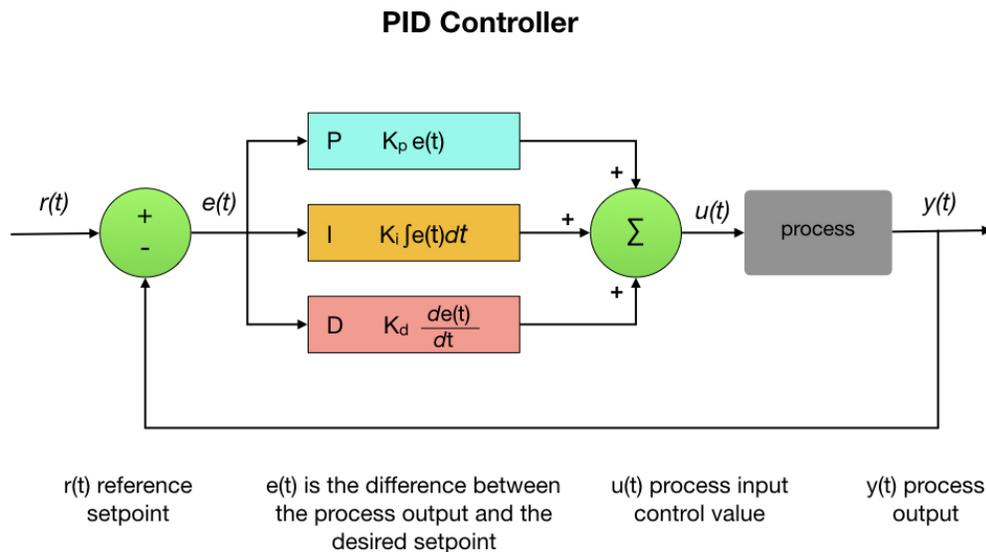


Figure 2.2: Block diagram of a PID controller (forum.dronebotworkshop.com)

2.8.1 Proportional, Integral, Derivative (PID) controller

PID controllers have been in existence for many decades and it is the most commonly used form of feedback control strategy. It was initially used in governors and it became quintessential in process control by the middle of 20th century. Most of the

controllers in industries are of PI type. It is also an integral part of the Distributed Control Systems (DCS) package sold by automation vendors. PID controllers today often come with built in logics, sequential functions and function blocks and are used in a wide range of industries. Even in hierarchical control layout, PID controllers work at the lower level to keep the system near the set points provided by upper multivariable control layers. Though there are several forms of PID, the parallel form of PID control algorithm is shown in figure 5.1. The textbook definition of a PID control law is given below,

$$u(t) = K_c \left[e(t) + \frac{1}{\tau_I} \int_{t_1}^{t_2} e(t).dt + \tau_D \frac{de(t)}{dt} \right] \quad (2.23)$$

From figure 5.1, it is clear that $y(t)$ is the measured process variable, $r(t)$ is the reference signal or set point, $u(t)$ is the control signal and $e(t)$ is the error signal ($e = y_{sp} - y$, as it is negative feedback). The control signal is therefore a sum of three terms: the P-term (that is proportional to the error), the I-term (that is proportional to the integral of the error), and the D-term (that is proportional to the derivative of the error). The parameters of the controllers which are to be tuned are proportional gain K_c , integral time τ_I , and derivative time τ_D . The integral, proportional and derivative parts can be visualised as past, the present and the future control actions respectively.

2.8.2 Zone Model Predictive Controller (ZMPC)

In HD, the outputs are more in number than the control inputs (i.e., more objectives than degrees of freedom) and so zone control could be considered as a natural choice. Also many of the physiological variables are defined in ranges (with upper and lower bounds as shown in figure 2.3) rather than a strict clinically accepted point value because of bio-variability among individuals. Also zone control can be used when there are conflicting objectives. If there are conflicting objects, the set point based MPC often suffers from parameter tuning difficulties. One strategy to reduce this burden is to specify reasonable zone targets. Another factor that comes into play is the process uncertainty and noise. This adds another dimension of objectives to the control system (i.e., variances and disturbance rejection). Zone Model Predictive Controllers (ZMPC) have gained momentum in medical applications like diabetes [45] and many more recently. In this particular work, the researchers have clinically tested the efficiency of the ZMPC algorithm under plant model mismatch and measurement noises.

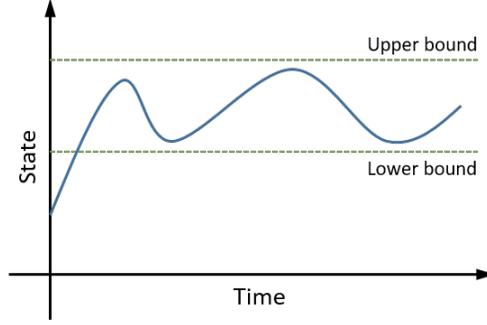


Figure 2.3: State trajectory in a typical zone control setting

The mathematical problem formulation of a ZMPC with a prediction horizon N_p is shown below.

$$\min_{x(i), \epsilon^L, \epsilon^U, \Delta u} \sum_{i=1}^{N_p} \left(\epsilon^L Q^L \epsilon^L + \epsilon^U Q^U \epsilon^U \right) + \sum_{i=0}^{N_p-1} \left(R \cdot \Delta u_i^2 \right) \quad (2.24a)$$

$$\text{s.t.} \quad x_{i+1} = f(x_i, u_i) + \omega_i \quad i = 1, 2, \dots, N_p \quad (2.24b)$$

$$y_i = h(x_i, u_i) + \nu_i \quad i = 1, 2, \dots, N_p \quad (2.24c)$$

$$x_{min} \leq x_i \leq x_{max} \quad i = 1, 2, \dots, N_p \quad (2.24d)$$

$$u_{min} \leq u_i \leq u_{max} \quad i = 0, 1, \dots, N_p - 1 \quad (2.24e)$$

$$\Delta u_{min} \leq \Delta u_i \leq \Delta u_{max} \quad i = 0, 1, \dots, N_p - 1 \quad (2.24f)$$

$$B^L - \epsilon^L \leq y_i \leq B^U + \epsilon^U \quad (2.24g)$$

$$\epsilon^L \geq 0, \epsilon^U \geq 0 \quad (2.24h)$$

In the above formulation, zone control is realized with the help of slack variables ϵ^L , ϵ^U for the lower (B^L) and upper (B^U) target zone bounds respectively, along with its associated constraints (states, inputs and rate of input change). The use of a slack variable eliminates the cost as long as the states are within the specified zone bounds. If the states are outside the specified target zone, the cost associated with that point is computed by taking the product of the squared distance from the point to the nearest zone boundary and the tuning parameter Q . R is the tuning parameter associated with the rate of change of input. Both Q and R are adjusted according to the user's requirements. In HD, the measurements are subjected to huge disturbances and the same patient behaves differently on different days, which results in plant model mismatch. Therefore, a ZMPC can also be applied to HD treatment settings and it will be discussed in detail in chapter 5.

2.9 Summary

To facilitate this research work to be understood by people from a wide array of professions and to promote a healthy dialogue between the heterogeneous community of engineers and physicians and systems engineers the fundamentals have to be outlined first. Therefore, in this chapter, the reader is first introduced to the fundamental definitions which are quite frequently used in the context of HD and aid in understanding the language used in a clinical HD treatment setting better. Finally, the technical concepts required for building the proposed frameworks and methods for simultaneous state and parameter estimation in chapter 4 and for optimal control of HD treatments in chapter 5 are presented to the reader.

Chapter 3

Mathematical modelling of the hemodialysis system

In the broad context, modelling aims to achieve a realistic reproduction of the dynamics of the process through mathematical relations. A model can either be physical, which is a replica of the real world object, or conceptual like a mathematical model. The mathematical model built should reproduce the dynamics of the system from which it was inspired, with a user acceptable degree of accuracy and consistency. Analysis of process characteristics and intervariable relationships are of paramount importance in prediction, control, monitoring, design and innovation of process systems. In our study, mathematical modeling of the hemodialysis system helps in understanding the process intricacies, patient's response to HD, designing and prescribing optimal treatments, tracking the course of the patient's treatment and paves the way for the application of Model Predictive Controls (MPCs) [83]. Process models can either be developed from fundamental laws of science (white box) or developed from input-output data in an empirical approach (black box). Quite often, developing white box models becomes a cumbersome task as it requires in-depth knowledge of the system under consideration and we propose using a grey box approach, a combination of both fundamental laws of science and synthetic input-output data. The very first objective of this research is to build a first principles model of the hemodialysis system which could be used for control relevant applications. Firstly, the reader is briefly introduced to a HD kinetic modelling guide, and the derivation of a single compartment model is shown as an example. Then the HD system model equations considered in this research are presented to the reader along with some derivations, model assumptions and modifications.

3.1 Guide to kinetic modelling

The structured approach for building a mathematical model for any physiological system [120] is as follows:

1. The initial step is to formulate the system under consideration in to a diagram with inputs and outputs to systems represented by arrows going into and out of the system respectively.
2. To write down the balance (mass, volume, solute) equations of the system, subsystem or compartment under consideration.

$$\begin{aligned}(\textit{Accumulation}) &= \sum(\textit{Inputs}) - \sum(\textit{Outputs}) \\ \frac{d(\textit{Content})}{dt} &= \sum(\textit{Inputs}) - \sum(\textit{Outputs})\end{aligned}$$

3. The unit consistency of the inputs and outputs must be ensured. Inputs and the outputs must be in similar units and should be going into and out of the system respectively.
4. In most of the cases, we end up with first order differential equations. Solve the equations either analytically or numerically using the classical techniques of applied mathematics.
5. Based on the user's requirement, the final expression can be rearranged to calculate the required parameters.

3.1.1 Example derivation: Single compartment model

a. Derivation of Single Pool Fixed Volume (SPFV) kinetic model

If the entire body fluid is considered to be a single pool of constant volume, then we can derive the expression for the blood toxins in the patient's body at any time instant. The single pool fixed volume model assumes that the total body water is constant and that there is only a single volume for urea distribution (no urea generation occurs here). The urea enters the compartment only from the liver and it is generated from amino acid catabolism. The urea is continuously removed through the patient's kidneys with clearance (K_r) and through the dialyzer with clearance (K_d) during HD treatment. Therefore, the overall clearance(K) is the sum of renal clearance and dialyzer clearance during dialysis and equal to the patient's renal clearance alone

between treatments. However, for chronic kidney patients if the native kidneys have completely failed (ESRD), then the renal clearance (K_r) can be taken as zero.

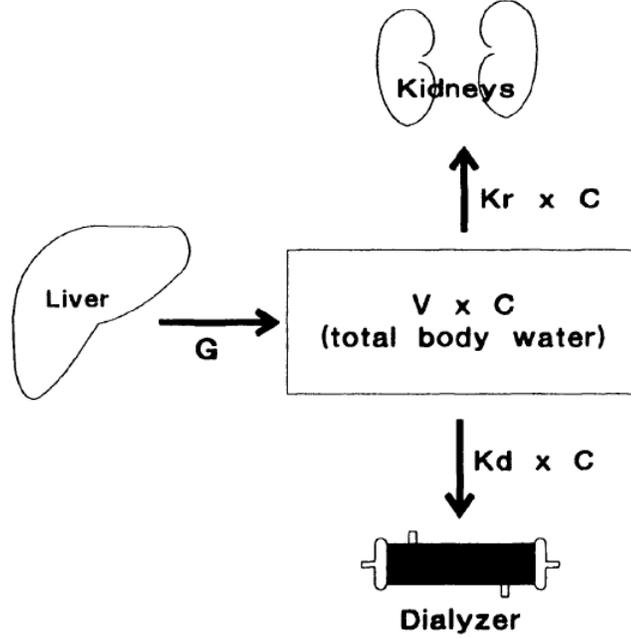


Figure 3.1: Single Pool Fixed Volume (SPFV) urea kinetic model. Taken from [36]

We assume that the volume change occurring during dialysis is negligible ($\frac{dV}{dt} = 0$) and if we write down the urea mass balance over the single compartment we get,

$$\frac{d(VC)}{dt} = G - KC \quad (3.1)$$

Here, G is the urea generation rate expressed in (mg/min) and K is the total clearance expressed in (ml/min). Applying product rule over the left hand side of Eq. (3.1) we get,

$$C \frac{dV}{dt} + V \frac{dC}{dt} = G - (K_r + K_d)C \quad (3.2)$$

$$V \frac{dC}{dt} = G - (K_r + K_d)C \quad (3.3)$$

The analytical solution for Eq. (3.3) can be obtained by separating the variables and by the use of integration formulae. Upon integration and applying limits (At time $t = 0$, $C = C_0$ and at time $t = t$, $C = C$) we get,

$$\frac{dC}{G - (K_r + K_d)C} = \frac{dt}{V} \quad (3.4)$$

$$\int_{C_0}^C \frac{dC}{G - (K_r + K_d)C} = \int_0^t \frac{dt}{V}$$

$$\ln \left[\frac{G - (K_r + K_d)C}{G - (K_r + K_d)C_0} \right] = \frac{-(K_r + K_d)t}{V}$$

$$G - (K_r + K_d)C = (G - (K_r + K_d)C_0) \cdot e^{\frac{-(K_r + K_d)t}{V}} \quad (3.5)$$

We can get an expression for the concentration of a toxin (as obtained by authors [46]) at a time instant ‘ t ’ by rearranging Eq. (3.5) as shown below,

$$C = C_0 \left[e^{\frac{-(K_r + K_d)t}{V}} \right] + \frac{G}{K_r + K_d} \left[1 - e^{\frac{-(K_r + K_d)t}{V}} \right] \quad (3.6)$$

where C_0 and C are the concentrations of toxin at time $t = 0$ and $t = t$ respectively in (mg/ml), V is the urea distribution volume (ml). The above equation can be used to compute the solute removal and if one examines Eq. (3.6) it is understood that solute removal is achieved by either dialyzer or renal clearance and the concentration of the solute is a function of both $\frac{Kt}{V}$ and the urea generation rate (G). Urea concentration drops during each hemodialysis session (falls exponentially during dialysis, not linearly) and increases between consecutive hemodialysis sessions (due to urea generation and because of reduced renal function) as shown in figure 3.2. The simulation settings are given in table 3.1.

Table 3.1: Simulation settings for single and double pool fixed volume urea kinetic models

S.No	Model variable name	Notation	Value	Unit
1	Urea generation rate	G	0.1	<i>mmol/min</i>
2	Renal clearance	K_r	0	<i>ml/min</i>
3	Dialyzer clearance	K_d	150	<i>ml/min</i>
4	Urea distribution volume	V	34800	<i>ml</i>
5	Extracellular volume	V_e	11600	<i>ml</i>
6	Intracellular volume	V_i	23200	<i>ml</i>
7	Intercompartmental mass transfer coefficient	K_c	300	<i>ml/min</i>
8	Initial blood urea concentration (BUC)	C_0	35	<i>mmol/L</i>
9	Initial extracellular BUC	$C_e(0)$	35	<i>mmol/L</i>
10	Initial intracellular BUC	$C_i(0)$	35	<i>mmol/L</i>
11	Treatment time	t	210	<i>min</i>

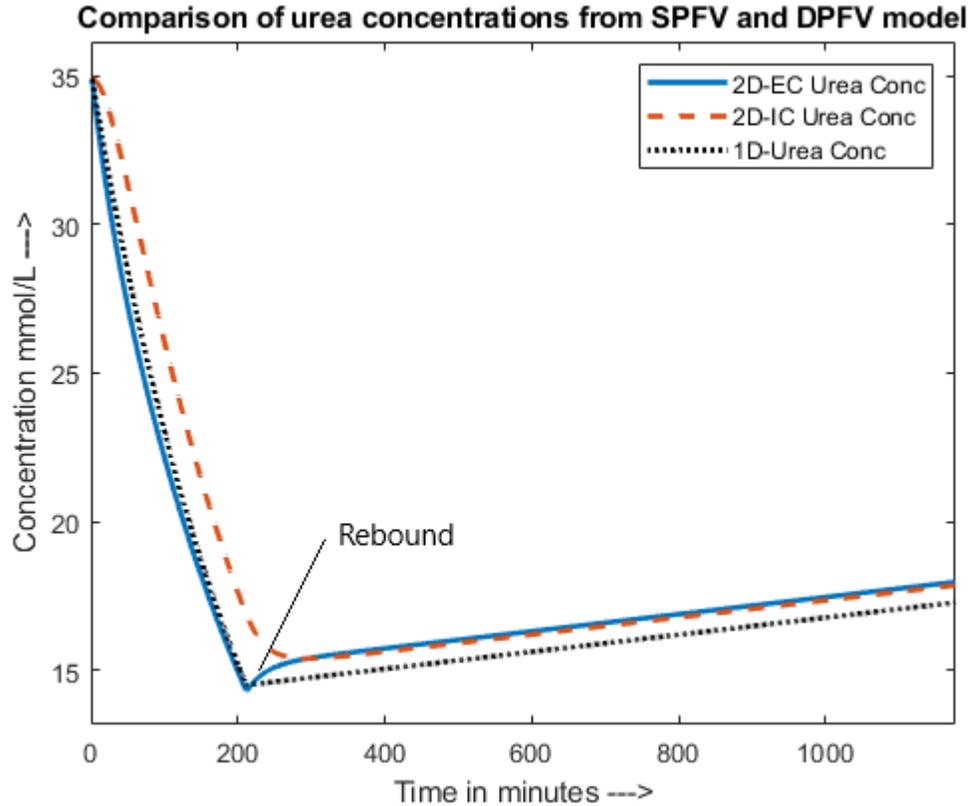


Figure 3.2: Simulation of the Single Pool Fixed Volume (SPFV) and Double Pool Fixed Volume (DPFV) urea kinetic model

The derivation and simulation results for the single pool urea kinetic model are shown here to educate the reader about the first approaches which were taken previously by researchers to study the dynamics. However, many of the assumptions of the single pool urea kinetic model might not hold true. We have assumed that the volume distribution remains constant during HD, but this is not true in real HD treatments. The patient loses weight during dialysis (due to ultrafiltration) and gains weight between dialysis treatments (due to fluid loading). With the development of more efficient dialyzers, we could remove urea 4 to 7 times quicker than normally functioning human kidneys [36]. This causes a steep change in the concentration of urea in the body in a short duration of time through a small funnel in the human body, i.e, the Arteriovenous (AV) graft. One more assumption while deriving the single pool model is that urea attains equilibrium rapidly. This means that the cell membrane and vascular permeability to urea is infinite, allowing us to consider the entire fluid space as one single pool. However, scientific findings suggest that urea permeability is finite. The mass transfer of urea across tissues and cell walls cannot

keep up with the rate at which urea is removed in the dialyzer. Therefore, a noticeable concentration gradient develops [130]. This gradient is believed to be developed at the cell wall between intracellular (IC) and extracellular (EC) pools. Therefore, we should consider that the body consists of two compartments as shown in figure 3.3. To get a visual understanding, two-pool fixed volume model equations from [158] are solved numerically (as analytical solutions are hard to obtain by hand) and the results are compared with the single pool model in figure 3.2. The double pool model can be used to explain the clinically significant ‘rebound’ phenomenon, which refers to the sudden jump in the extracellular concentration of urea immediately after stopping dialysis (right after 210 minutes in figure 3.2).

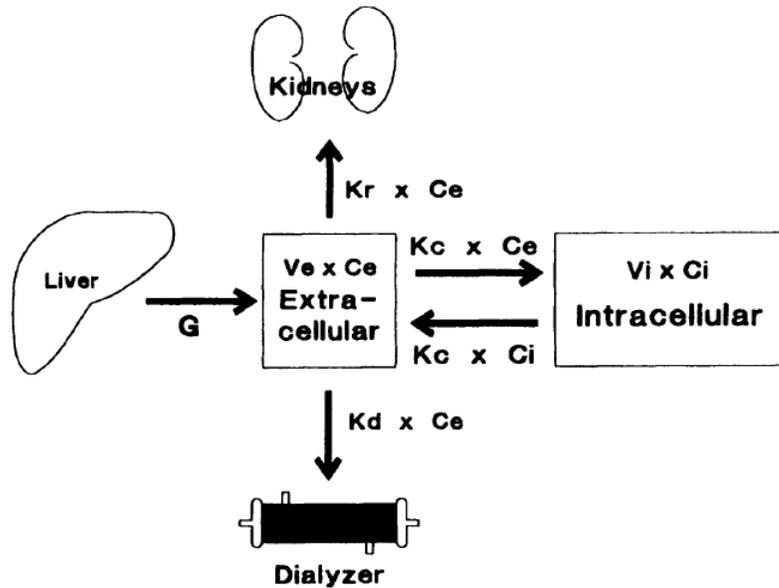


Figure 3.3: Double Pool Fixed Volume (DPFV) urea kinetic model. Taken from [36]

3.2 Hemodialysis system model

Intradialytic hypotension (IDH) and Dialysis Disequilibrium Syndrome (DDS) are the most frequently occurring HD complications because of the removal of overloaded fluid in a short period of time. Various superimposing nonlinear complex factors help in blood pressure stability of the patient. Some of these factors include, vascular refilling from the interstitial fluid space and the impact of the internal cardiovascular control mechanisms in the patient’s body. As a result, each patient behaves differently to the same HD treatment. For instance, some can tolerate quicker treatments while some have very poor hemodynamic stability. One cannot understand the com-

plex relationships between different HD variables very easily, and so a comprehensive mathematical model is necessary. The model utilized in this research work [140] contains the equations to represent the fluid transfer dynamics between compartments with the related solute dynamics, a simple compartment model of the human heart and cardiovascular dynamics (systemic and pulmonary circulation) along with the action of pressoreceptors (low pressure cardiopulmonary baroreceptors and high pressure arterial baroreceptors), bidirectional solute and fluid exchange dynamics (diffusion and convection) across the hemodialyzer. The HD system is divided into the following subsystems as shown here for comprehensible presentation.

1. Patient subsystem

- (a) Solute balance model (2 compartment)
- (b) Fluid balance model (3 compartment)
- (c) Cardiovascular dynamics model
 - i. The heart pressures (6 compartment)
 - ii. Baroflex regulation (3 manipulated variables)

2. Hemodialyzer subsystem

- (a) Simple model: Exit concentration (Used in chapters 3, 4)
- (b) Complex model: Spatio-temporal concentration profile

Although this model is a comprehensive framework, it comes with some assumptions to make the model simple enough for easy implementation in limited computing resources while explaining all integrated dynamics without fail. The most important assumptions of the model are outlined next while the other assumptions will be presented in the subsequent subsections.

1. For studying fluid exchange dynamics, the patient's body is split into three compartments namely, intracellular fluid space, interstitial fluid space and plasma.
2. For studying solute dynamics, the interstitial fluid space and plasma are summed into an 'extracellular' compartment as the concentration of any solute in the interstitial fluid space and plasma is nearly equal. Thus two compartments are enough to explain solute dynamics.

3. The main solutes considered are Urea (marker solute for HD), Sodium and Potassium. All other solutes that do not cross the cellular membrane and hemodialyzer membrane are clubbed into one and then considered for the simulations.
4. This model does not differentiate the hemodynamics and solute kinetics of different organs in the body. Instead, each compartment is considered to be a representation of the entire patient body.
5. The cardiovascular system is considered to be a culmination of two cardiac (left and right) and four vascular compartments (systemic arteries, systemic veins, pulmonary arteries and pulmonary veins).
6. The blood pressure output is not pulsatile, instead it is an average. This is because the cardiac cycle (roughly 0.83 seconds) is very short when compared to the HD treatment (roughly 4 hours).
7. It is assumed that the cardiovascular reflex is mainly due to the activity of two groups of baroreceptors (low and high pressure).
8. Only bicarbonate dialysis is considered for the purpose of modelling. This is because sodium bicarbonate is most commonly used buffer solution in HD.

In the next few subsections, the fundamental equations representing different subsystems of the biotechnological HD system (figure 1.9) are presented.

3.2.1 Patient subsystem model: Solute and fluid dynamics

a. Solute balance model (2C)

In this work, only 3 main solutes ($s = Na, K, U$) are considered. They are sodium (Na) and potassium (K), which are the primary substances responsible for osmolarity at the cellular membrane, and urea (U), which is the marker solute for HD. The bidirectional flow of solutes between the intracellular and extracellular pools because of the concentration gradient is assumed to be linear. Active and passive transport are accounted for sodium and potassium, while urea is transported only by diffusion. If the readers look at figure 3.4, they can see that the pale yellow boxes enclosed within dotted lines represent the two compartments and that the change in the concentration of any solute in the extracellular pool is because of two positive (from infusion fluid, from intracellular pool) and two negative contributions (going out from

the extracellular pool, crossing dialyzer). According to some authors' findings [67], the solute concentrations in plasma (C_s^{pl}) can be computed from the interstitial fluid solute concentrations (C_s^{is}).

$$C_s^{pl} = \frac{F_p}{\alpha_s} \cdot C_s^{is} \quad (3.7)$$

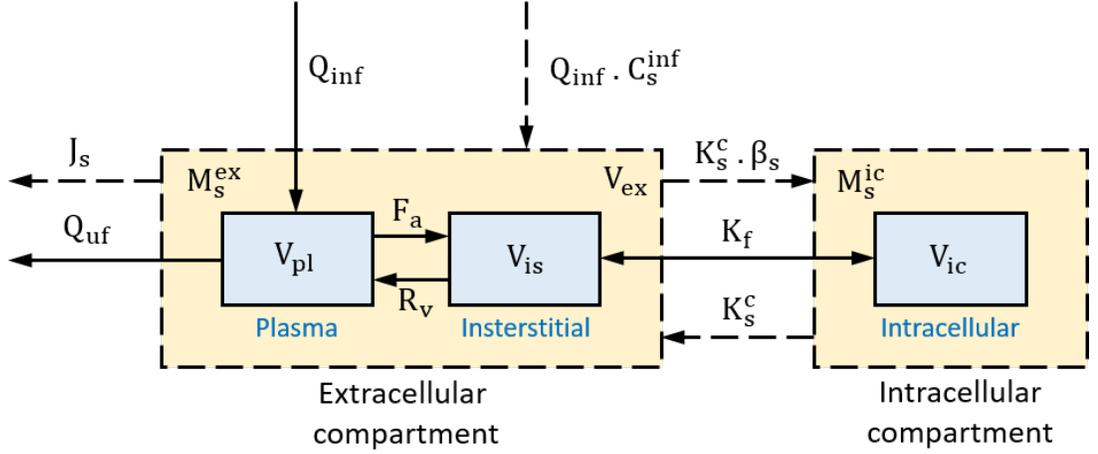


Figure 3.4: Body compartment model for solute (2C - dotted line) and fluid exchange (3C - solid line) kinetics. Adopted from [140]

The Donnan-Gibbs ratio for a solute (α_s) in the interstitial fluid is approximately equal to 0.95 for sodium and potassium while the plasma water fraction (F_p) is approximately equal to 0.94 [122]. So the concentration of sodium and potassium is almost equal in the plasma and in the interstitial fluid and hence a two compartment model could be used for describing solute kinetics with fair accuracy.

$$C_s^{ex} = C_s^{is} \simeq C_s^{pl} \quad (3.8)$$

There are different solutes in the patient's body that should be within safe limits for the normal functioning of the body. The general mass balance equations for any particular solute (s) over the two compartments namely, intracellular and extracellular, considering both active and passive transport can be written as:

$$\frac{dM_s^{ic}}{dt} = K_s^c(\beta_s C_s^{ex} - C_s^{ic}) \quad (3.9)$$

$$\frac{dM_s^{ex}}{dt} = G_s - K_s^c(\beta_s C_s^{ex} - C_s^{ic}) - J_s + Q_{inf} \cdot C_s^{inf} \quad (3.10)$$

$$C_s^{ic} = \frac{M_s^{ic}}{V_{ic}} \quad (3.11)$$

$$C_s^{ex} = \frac{M_s^{ex}}{V_{pl} + V_{is}} \quad (3.12)$$

In the above equations, M_s^{ic} , M_s^{ex} , M_s^{is} , M_s^{pl} represent the mass of solutes (s) in intracellular, extracellular, interstitial and plasma compartment respectively. The letter V along with a subscript is used to represent the volume of any compartment. K_s^c , β_s are the intercompartmental mass transfer coefficients of the solute (s). G_s symbolises the generation rate of the solute (s) in the patient and J_s is the mass flux of the solute (s) across the hemodialyzer. The solute concentration difference between the intracellular and the extracellular pool is the main driving forcing for mass transfer.

b. Fluid balance model (3C)

Approximately 60% of the body weight of the patient comprises of fluids and there are different formulae to arrive at the total fluid distribution volume (V_{tot}) of a patient [158]. The total fluid volume is assumed to be distributed in three compartments namely, the intracellular pool (V_{ic}), interstitial pool (V_{is}) and plasma (V_{pl}), as represented in figure 3.4 by solid outlined blue boxes. The blood volume (V_b) is the sum of plasma volume and Red Blood Cell (RBC) volume (V_{rc}), and hematocrit (H) is the volume percentage of RBCs in blood. The set of fluid balance equations given in this subsection describe the dynamics of transcellular fluid shifts due to fluctuations in plasma osmolarity induced by the hemodialysis treatment.

$$V_{ex} = V_{pl} + V_{is} \quad (3.13)$$

$$V_{tot} = V_{pl} + V_{is} + V_{ic} = V_{ex} + V_{ic} \quad (3.14)$$

$$V_b = V_{pl} + V_{rc} \quad (3.15)$$

$$H = \frac{V_{rc}}{V_b} = \frac{V_b - V_{pl}}{V_b} = 1 - \frac{V_{pl}}{V_b} \quad (3.16)$$

The fluid balance over the three compartments are derived next and one can see that the sum of equations (3.17), (3.18), (3.19) is same as equation (3.20), which means that the total change in the body fluid volume is because of Ultrafiltration rate (UFR) and Infusion Rate (IR). In a simpler sense, UFR is the rate of removal of fluid from the blood plasma and IR is the rate of addition of fluid into the blood plasma. The

equations governing the intercompartmental fluid transfer are outlined next.

$$\frac{dV_{pl}}{dt} = -Q_{uf} + Q_{inf} - F_a + R_v \quad (3.17)$$

$$\frac{dV_{is}}{dt} = -K_f \cdot (C^{ic} - C^{is}) + F_a - R_v \quad (3.18)$$

$$\frac{dV_{ic}}{dt} = K_f \cdot (C^{ic} - C^{is}) \quad (3.19)$$

$$\frac{dV_{tot}}{dt} = -Q_{uf} + Q_{inf} \quad (3.20)$$

$$C^{ic} = \frac{M_K^{ic} + M_{Na}^{ic} + M_U^{ic} + M_{eq}^{ic}}{V_{ic}} \quad (3.21)$$

$$C^{is} = C^{ex} = \frac{M_K^{ex} + M_{Na}^{ex} + M_U^{ex} + M_{eq}^{ex}}{V_{pl} + V_{is}} \quad (3.22)$$

where C^{ic} , C^{is} are the concentrations of all the osmotically effective solutes in the intracellular and the interstitial fluid respectively while M_{eq} represents the mass of all osmotically substances that remain constant throughout HD. Q_{uf} represents UFR, Q_{inf} denotes IR and K_f is the water transfer coefficient. F_a is the filtration rate at the arterial capillaries and R_v is the absorption rate at the venous capillaries. Additionally, the following algebraic equations are necessary for solving the ordinary differential equations of fluid balance given above.

$$F_a = L_a \cdot (P_{ac} - P_{is} - \pi_{pl} + \pi_{is}) \quad (3.23)$$

$$R_v = L_v \cdot (\pi_{pl} - \pi_{is} - P_{vc} + P_{is}) \quad (3.24)$$

In the above equations, L_a and L_v are the permeability coefficients of the arterial capillaries and venous capillaries respectively. P_{is} , P_{ac} , P_{vc} are the pressures at interstitial fluid space, arterial capillaries, venous capillaries respectively. P_{ac} , P_{vc} are derived by imposing mass preservation principles at the arterial and venous capillary nodes [140]. π_{pl} , π_{is} are the oncotic pressures in the plasma and interstitial fluid respectively. Landis-Pappenheimer equations [71] are used to compute the oncotic pressure in the interstitial fluid and plasma compartments.

$$\pi_{is} = 2.8 \cdot C_p^{is} + 0.18 \cdot (C_p^{is})^2 + 0.012 \cdot (C_p^{is})^3 \quad (3.25)$$

$$\pi_{pl} = 2.1 \cdot C_p^{pl} + 0.16 \cdot (C_p^{pl})^2 + 0.009 \cdot (C_p^{pl})^3 \quad (3.26)$$

The protein content in the plasma (C_p^{pl}) and in the interstitial fluid (C_p^{is}) is assumed to be constant throughout the hemodialysis treatment. Their concentrations fluctuate because of the changing compartment volumes but the total mass of protein in a

compartment will remain the same. The subscript (n) represents the basal or nominal value of a physiological quantity.

$$C_p^{is} = C_p^{isn} \cdot \frac{V_{isn}}{V_{is}} \quad (3.27)$$

$$C_p^{pl} = C_p^{pln} \cdot \frac{V_{pln}}{V_{pl}} \quad (3.28)$$

Finally, the hydraulic pressure in the interstitial fluid space is a linear function of the interstitial fluid volume by assuming a constant elastance of the interstitial place (E_{is}).

$$P_{is} = E_{is} \cdot (V_{is} - V_{isn}) + P_{isn} \quad (3.29)$$

3.2.2 Patient subsystem model: Cardiovascular dynamics

a. The heart pressures (6C)

A patient's cardiovascular system is modelled as six compartments as shown in figure 3.5. It was assumed that there are four vascular compartments and two cardiac compartments. Each compartment has a hydraulic resistance which accounts for energy dissipation and loss of pressure, and a compliance that represents the total blood volume in that compartment. The pressure (P) changes in each compartment depend on the compliance (C) of the compartment, pressures in the adjacent compartments or cardiac outputs from left and right heart (q_l , q_r) and the vascular resistances (R) [43]. As the heart was represented by an equivalent electric circuit by the original authors of this model [140], ohm's law for electric current could be applied here with blood flow, pressure and vascular resistance resembling current, voltage and electric resistance respectively. Thus, the pressure equations can be derived by mass preservation principles at each of the compartment as given below.

$$\frac{dP_{sa}}{dt} = \frac{1}{C_{sa}} \cdot \left(q_l - \frac{P_{sa} - P_{ac}}{R_{s1}} \right) \quad (3.30)$$

$$\frac{dP_{ra}}{dt} = \frac{1}{C_{ra}} \cdot \left(\frac{P_{sv} - P_{ra}}{R_{sv}} - q_r \right) \quad (3.31)$$

$$\frac{dP_{pa}}{dt} = \frac{1}{C_{pa}} \cdot \left(q_r - \frac{P_{pa} - P_{pv}}{R_{pa}} \right) \quad (3.32)$$

$$\frac{dP_{pv}}{dt} = \frac{1}{C_{pv}} \cdot \left(\frac{P_{pa} - P_{pv}}{R_{pa}} - \frac{P_{pv} - P_{la}}{R_{pv}} \right) \quad (3.33)$$

$$\frac{dP_{la}}{dt} = \frac{1}{C_{la}} \cdot \left(\frac{P_{pv} - P_{la}}{R_{pv}} - q_l \right) \quad (3.34)$$

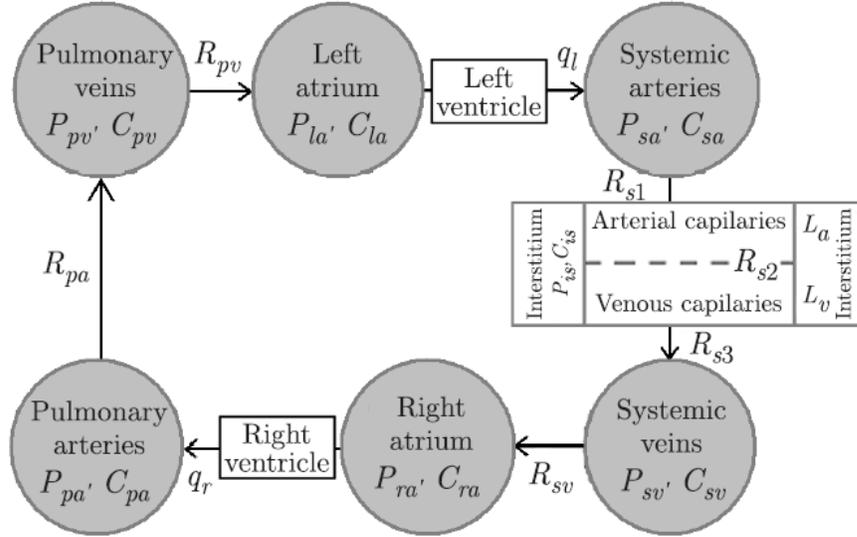


Figure 3.5: A schematic of the 6 compartment model of the cardiovascular system. Adopted from [43]

The subscripts sa , sv , pa , pv , la , ra represent systemic arteries, systemic veins, pulmonary arteries, pulmonary veins, left and right atrium respectively. The pressure in the systemic veins has been computed by considering that the volume in a compartment must be equal to the difference between the total blood volume (V_b) and the remaining compartment volumes. Because blood vessels are elastic in nature, we have unstressed volumes (V_u) and filling volumes ($C_c P_c$) and so we can write,

$$P_{sv} = \frac{1}{C_{sv}} \cdot (V_b - V_u - C_{sa}P_{sa} - C_{pa}P_{pa} - C_{pv}P_{pv} - C_{ra}P_{ra} - C_{la}P_{la}) \quad (3.35)$$

$$V_u = V_{usa} + V_{usv} + V_{upa} + V_{upv} + V_{ura} + V_{ula} \quad (3.36)$$

$\delta(t)$ is a sympathoinhibitory signal, with a gain G_δ and a time constant τ_δ , normally equal to zero. The dynamics of this sympathoinhibitory signal is as follows,

$$\frac{d\delta}{dt} = \left\{ \begin{array}{ll} 0, & \text{if } P_{la} \geq P_{lat} \\ \frac{G_\delta \cdot (P_{lat} - P_{la}) - \delta}{\tau_\delta}, & \text{if } P_{la} < P_{lat} \end{array} \right\} \quad (3.37)$$

The initial value of $\delta(t)$ is always zero and so we can assume that we know this state before running the state estimation algorithm. There are some more expressions that have to be specified to enable the user to compute the pressure dynamics. The cardiac output is equal to the product of the stroke volume (SV) and the cardiac frequency

(f), therefore for the left (q_l) and right (q_r) ventricles we can write,

$$q_l = SV_l \cdot f \quad (3.38)$$

$$q_r = SV_r \cdot f \quad (3.39)$$

Frank-Starling mechanism says that stroke volume depends on the end diastolic volume and therefore upon upstream atrial pressure [140]. A linear dependence is used in this derivation. Furthermore, the stroke volume decreases if the downstream atrial pressure rises above normal. To replicate this physiological cardiovascular phenomenon, a square root dependence upon arterial pressure has been used [137]. Hence, the stroke volumes from the left and the ventricles can be written as follows,

$$SV_l = k_l \cdot (P_{la} - P_{la0})/a_l \quad (3.40)$$

$$SV_r = k_r \cdot (P_{ra} - P_{ra0})/a_r \quad (3.41)$$

If $P_{sa} \leq P_{san}$, then $a_l = 1$ and if $P_{sa} > P_{san}$ then $a_l = \sqrt{P_{sa}/P_{san}}$. Similarly for the right ventricle we can say, if $P_{pa} \leq P_{pan}$, then $a_r = 1$ and if $P_{pa} > P_{pan}$ then $a_r = \sqrt{P_{pa}/P_{pan}}$, where k_l, k_r are the slopes of the stroke volume versus atrial pressure relationships (or cardiac effectiveness) for the left and right heart respectively, P_{la0}, P_{ra0} are the x-axis intercepts of the same relationships, the quantities a_l, a_r describe the effect of afterload on the stroke volume and the subscript ‘ n ’ denotes that a physiological quantity is in its nominal or basal condition [140].

b. Baroflex regulation (3 MVs)

Based on our assumptions, the short term baroflex (pressure) regulation is primarily due to the activity of two groups of baroreceptors namely, arterial (high pressure) and cardiopulmonary (low pressure) baroreceptors. The mathematical equations for describing the inherent control mechanisms for maintaining hemodynamic stability (blood pressure) of the patient undergoing HD by adjusting cardiovascular system states (Systemic resistance, heart period or heart rate, unstressed venous volume) both during the sympathoexcitatory phase and the sympathoinhibitory phase have been taken from the article [140]. From a systems engineering perspective, these three cardiovascular states can be called as manipulated variables while blood pressure can be called as the controlled variable. To get a good understanding, the mechanism of system resistance manipulation is explained here. The afferent signals coming from both the baroreceptors are weighted by their respective gain values and then integrated at the level of the central nervous system into an efferent signal and then sent

out to adjust the systemic resistance value. The presence of physiological saturation limits for the manipulated variables is brought in by assuming a sigmoidal characteristic curve for the effector response and because of this, the low pressure and the high pressure baroreceptors interact in a highly nonlinear way. Finally, a first order transfer function is included to simulate the time required for this particular mechanism to occur.

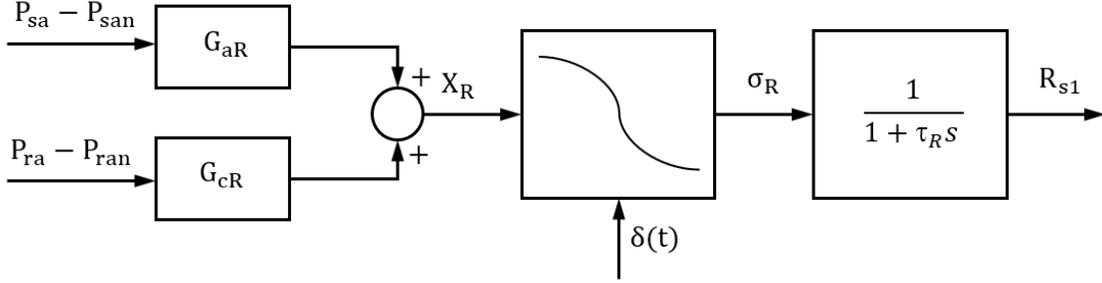


Figure 3.6: Block diagram for the control of systemic resistance. Adopted from [140]

The systemic resistance (R_{s1}) control can be represented by figure 3.6 as shown above and the set of equations are as follows,

$$\frac{dR_{s1}}{dt} = \frac{\sigma_R - R_{s1}}{\tau_R} \quad (3.42)$$

$$\sigma_R = \frac{R_{s1 \max} + R_{s1 \min} \cdot \exp(X_R/K_R)}{1 + \exp(X_R/K_R)} \quad (3.43)$$

$$K_R = \frac{\Delta\sigma_R}{4} \quad (3.44)$$

$$X_R = G_{aR} \cdot (P_{sa} - P_{san}) + G_{cR} \cdot (P_{ra} - P_{ran}) \quad (3.45)$$

$$R_{s1 \min} = \sigma_{Rn} - \Delta\sigma_R/2 \quad (3.46)$$

$$R_{s1 \max} = R_{s1 \min} + \Delta\sigma_R \cdot (1 - \delta(t)) \cdot \epsilon(t) \quad (3.47)$$

In the above set of equations, τ_R is the time constant of the systemic resistance feedback mechanism, σ_R is the sigmoidal static characteristic, K_R is a parameter which determines the slope of the sigmoidal characteristic at its central point, X_R is a linear combination of pressure changes at the level of high pressure(arterial) and low pressure(cardiopulmonary) baroreceptors, σ_{Rn} is the central value of the sigmoidal characteristic, $\Delta\sigma_R$ is the amplitude of the sigmoidal characteristic, G_{aR} and G_{cR} are the maximum open loop gains of the arterial and cardiopulmonary baroreceptor mechanisms respectively, $R_{s1 \min}$ is the lower threshold of the resistance control baroreceptor mechanism and $R_{s1 \max}$ is the upper threshold of the resistance control

baroreceptor mechanism. The variable $\epsilon(t)$ is always maintained at 1 during a standard bicarbonate dialysis, which means it has no effect on the systemic resistance. In our research work, only a standard bicarbonate dialysis is considered.

Heart period (T) or the heart rate control can be represented with a set of equations similar to the systemic resistance control as shown below,

$$\frac{dT}{dt} = \frac{\sigma_T - T}{\tau_T} \quad (3.48)$$

$$\sigma_T = \frac{T_{min} + T_{max} \cdot \exp(X_T/K_T)}{1 + \exp(X_T/K_T)} \quad (3.49)$$

$$K_T = \frac{\Delta\sigma_T}{4} \quad (3.50)$$

$$X_T = G_{aT} \cdot (P_{sa} - P_{san}) + G_{cT} \cdot (P_{ra} - P_{ran}) \quad (3.51)$$

$$T_{max} = \sigma_{Tn} + \Delta\sigma_T/2 \quad (3.52)$$

$$T_{min} = T_{max} - \Delta\sigma_T \cdot (1 - \delta(t)) \quad (3.53)$$

The control of venous capacity or systemic venous unstressed volume (V_{usv}) is represented by the following equations,

$$\frac{dV_{usv}}{dt} = \frac{\sigma_V - V_{usv}}{\tau_V} \quad (3.54)$$

$$\sigma_V = \frac{V_{usv min} + V_{usv max} \cdot \exp(X_V/K_V)}{1 + \exp(X_V/K_V)} \quad (3.55)$$

$$K_V = \frac{\Delta\sigma_V}{4} \quad (3.56)$$

$$X_V = G_{aV} \cdot (P_{sa} - P_{san}) + G_{cV} \cdot (P_{ra} - P_{ran}) \quad (3.57)$$

$$V_{usv max} = \sigma_{Vn} + \Delta\sigma_V/2 \quad (3.58)$$

$$V_{usv min} = V_{usv max} - \Delta\sigma_V \cdot (1 - \delta(t)) \quad (3.59)$$

During the sympathoinhibitory phase, the heart period and venous unstressed volume were presumed to be driven to the upper saturation limit and the explanations of the notations used are very similar to the systemic resistance control. These three reflex mechanisms will help us get a reasonable replication of the actual cardiovascular reflex mechanisms taking place in a human body.

3.2.3 Hemodialyzer subsystem model

a. Simple model: Linear equations giving exit concentration

The mass flux (J_s) of a solute can be computed by the equations given in ([122]), if the dialysance of that solute (Dia_s) through the hemodialyzer membrane is known. This will enable the computation of the solute concentrations at the exit of the hemodialyzer at any given time instant.

$$J_s = [Dia_s \cdot (1 - Q_{uf}/Q_{e,s}) + Q_{uf}] \cdot C_s^{ex} - Dia_s \cdot (1 - Q_{uf}/Q_{e,s}) \cdot C_s^d \quad (3.60)$$

$$Q_{e,s} = Q_b \cdot [F_p \cdot (1 - H) + F_r \cdot \gamma_s \cdot R_s^D] \quad (3.61)$$

Here, C_s^d is the concentration of a solute in dialysate, $Q_{e,s}$ is the effective blood flow rate of solute (s) in the dialyzer, F_p and F_r are the fractions of plasma and RBCs respectively, γ_s is the fraction of RBC water that participates in the transfer during a single pass in the hemodialyzer, R_s^D is the Donnan ratio of solute (s) and H is the hematocrit. However, the dialysance of a solute is constant, only if the blood flow rate, and dialysate flow rate and the membrane properties are held constant. In a practical setting, the membrane for HD is chosen before hand, but the blood flow rate and dialysate flow rate are varied continuously to achieve the desired levels of treatment. So, the author's assumption of constant dialysance [140] might not hold, if one is interested in looking at the feedback control of blood and dialysate flow rates.

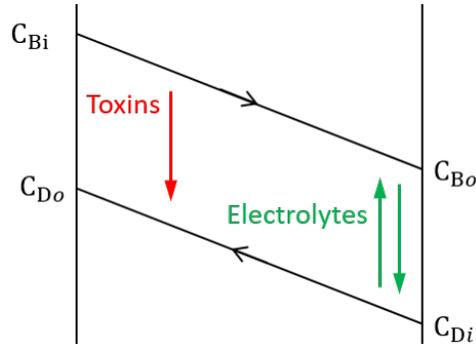


Figure 3.7: Concentration profile diagram for counter current flow of blood and dialysate in a hemodialyzer

In figure 3.7, the direction of flow of blood and dialysate are considered to be in counter-current fashion like in a standard hollow fibre dialyzer. For simplicity, let us name the hemodialyzer inlet solute concentration of the solutes on the blood side as C_{Bi} and the exit concentration after a single pass through the hemodialyzer as C_{Bo} .

Similarly, the hemodialyzer inlet and exit solute concentration of the solutes on the dialysate side are C_{Di} and C_{Do} respectively. It is clearly evident from figure 3.7, that the toxins are transferred into the dialysate because of the concentration gradient and so their concentration drops during a single pass. Quite interestingly, the dialysance of a solute can be derived just based on flow rates (blood Q_b and dialysate Q_d) and the membrane characteristics (K_0A : Mass transfer coefficient times the surface area of the membrane) using a log mean concentration difference approach as shown in this author's work [99] dedicated to designing and modelling a portable hemodialysis system. The dialysance equations have been adopted from [99], and the readers can look into the derivations if they are curious to understand the mathematical fundamentals. The equation for the dialysance of a solute (Dia_s) is given below,

$$Dia_s = Q_b \times \frac{e^{\frac{\kappa_0 A \left(1 - \frac{Q_b}{Q_d}\right)}{Q_b}} - 1}{e^{\frac{\kappa_0 A \left(1 - \frac{Q_b}{Q_d}\right)}{Q_b}} - \frac{Q_b}{Q_d}} \quad (3.62)$$

Eq. (3.62) can be substituted into Eq. (3.60), and this enables us to arrive at a control relevant model of the hemodialyzer mass flux equation, where Q_b , Q_d and C_s^d can be varied continuously to achieve treatment targets. The power of Eq. (3.62) lies in the fact that the dialysance can be computed without the concentrations of solutes in blood and dialysate.

b. Complex model: Partial differential equations giving spatio-temporal concentration profile

If the reader is interested in studying the intricate mass transfer dynamics within a hemodialyzer, the mathematical model should be derived in such a way which includes solute transfer, hemodialyzer design specifications, parameters of the hollow fibres and the properties of the process and service fluids as done by these authors [16]. Although there are several geometries and flow configurations, the hemodialyzer is considered to be of cylindrical geometry, having a lot of hollow semi permeable membrane fibres within the cylindrical structure. This arrangement is widely used in commercial hemodialyzers because of a low priming volume of blood while offering a large surface area for bidirectional mass transfer. Pressure gradients develop on both the blood and dialysate channels in a hollow fibre type hemodialyzer and local convection becomes unavoidable even if the net zero filtration is achieved. Blood normally flows through the tube side (along z direction) and the dialysate fluid flows in a direction opposite to blood (along $-z$ direction) in the void space between the fibres,

for higher mass transfer efficiency. A single hollow fibre of the hemodialyzer is shown in figure 3.8. The authors start by writing down the continuity equation for a solute. Continuity equation in cylindrical coordinates are utilized because the geometry of the hemodialyzer is also cylindrical and one can obtain a natural easy representation of the phenomena considering the nature of this problem. Additionally, both diffusion (Fick's law) and convection (Darcy's law) equations are plugged into the derivations.

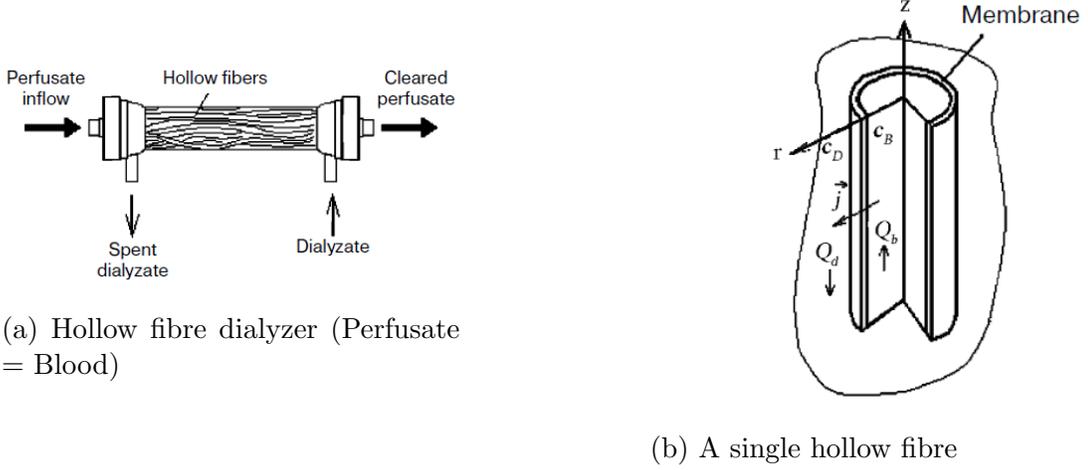


Figure 3.8: Representation of hollow fiber hemodialyzer used in HD. Taken from [16]

The final equations for computing the spatio temporal concentrations of solutes in the hemodialyzer are presented here, but the reader is encouraged to look into the fundamental intermediate steps in derivations [16] to get a clear picture.

$$\frac{\partial C_{sB}(t, z)}{\partial t} = -\frac{1}{(1-H)} \frac{\partial(C_{szB} \cdot u_b)}{\partial z} - \frac{2 \cdot C_{sB}}{R_m(1-H)} \left[\frac{D_s}{d} + f_s \right] + \frac{2 \cdot C_{sD} \cdot D_s}{R_m(1-H)d} \quad (3.63)$$

$$\frac{\partial C_{sD}(t, z)}{\partial t} = \frac{R^2}{(R^2 - R_m^2 \cdot N)} \frac{\partial(C_{szD} \cdot u_d)}{\partial z} + \frac{2 \cdot C_{sB} \cdot R}{R^2 - R_m^2 \cdot N} \left[\frac{D_s}{d} + f_s \right] - \frac{2 \cdot C_{sD} \cdot R \cdot D_s}{(R^2 - R_m^2 \cdot N)d} \quad (3.64)$$

In the equations given above, C_{sB} is the concentration of solute (s) in Blood, C_{sD} is the concentration of solute (s) in Dialysate, H is the hematocrit, R_m is the radius of the fibre, R is the radius of the dialyzer, D_s is the diffusivity coefficient of solute (s) through the membrane, d is the membrane thickness, N is the total number of fibres, f_s is the filtration rate of solute (s), u_b is the velocity of blood and u_d is the velocity of dialysate. These simultaneous partial differential equations could be solved using the method of characteristics, but we resorted to leverage the power of

a computer. To solve this problem numerically using the easily implementable finite difference method, the first step would be to discretize these equations along one variable (z) so that they become a set of ordinary differential equations with respect to the other variable (t). There are different discretization schemes (forward, backward, central) available in literature [26], but we are most interested in putting all these equations in feedback control application and so we resorted to a backward discretization scheme which skips the requirement of future state values. Therefore, the above set of equations were discretized along the spatial direction using backward difference approximation formula to convert them into a set of simultaneous ordinary differential equations with time as the independent variable. Finally, the discretized partial differential equations were solved numerically for predicting the exit concentrations of different solutes in the hemodialyzer by supplying the necessary initial conditions and inlet conditions ($z = 0$ and $z = L$) at every time instant. A visual representation of the discretization along z direction is presented in figure 3.9. Discretization along z direction results in the concept of imaginary stages in the hemodialyzer, and anything leaving a stage (n) takes the superscript (n). Depending on the number of stages in the hemodialyzer, the number of ordinary differential equations on the blood and dialysate side (2 ODEs for 1 stage for 1 solute) could be easily determined.

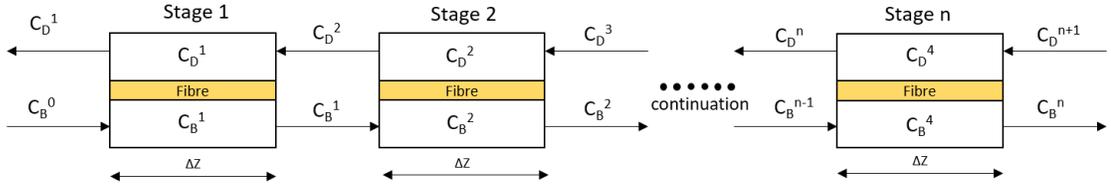


Figure 3.9: Representation of stages : Discretization along z direction

The discretized simultaneous ordinary differential equations are given below,

$$\frac{dC_{sB}^n}{dt} = -\frac{u_b}{(1-H)} \left[\frac{C_{sB}^n - C_{sB}^{n-1}}{\delta z} \right] - \frac{2.C_{sB}^n}{R_m(1-H)} \left[\frac{D_s}{d} + \frac{f_s}{n_z} \right] + \frac{2.C_{sD}^n.D_s}{R_m(1-H)d} \quad (3.65)$$

$$\frac{dC_{sD}^n}{dt} = \frac{u_d R^2}{(R^2 - R_m^2.N)} \left[\frac{C_{sD}^{n+1} - C_{sD}^n}{\delta z} \right] + \frac{2.C_{sB}^n.R}{R^2 - R_m^2.N} \left[\frac{D_s}{d} + \frac{f_s}{n_z} \right] - \frac{2.C_{sD}^n.R.D_s}{(R^2 - R_m^2.N)d} \quad (3.66)$$

Here, n_z is the total number of stages (all other notations explained previously).

encounter in a HD clinical setting. The simulations were carried out for a total treatment time of 4 hours, where the ideal 70 kg patient lost 3 kilograms of overloaded fluid and the results were compared with previously available clinical data (marked with red crosses in figure 3.12, 3.13) from literature [14], [32]. The values of all other parameters of the HD system have been taken from literature [140].

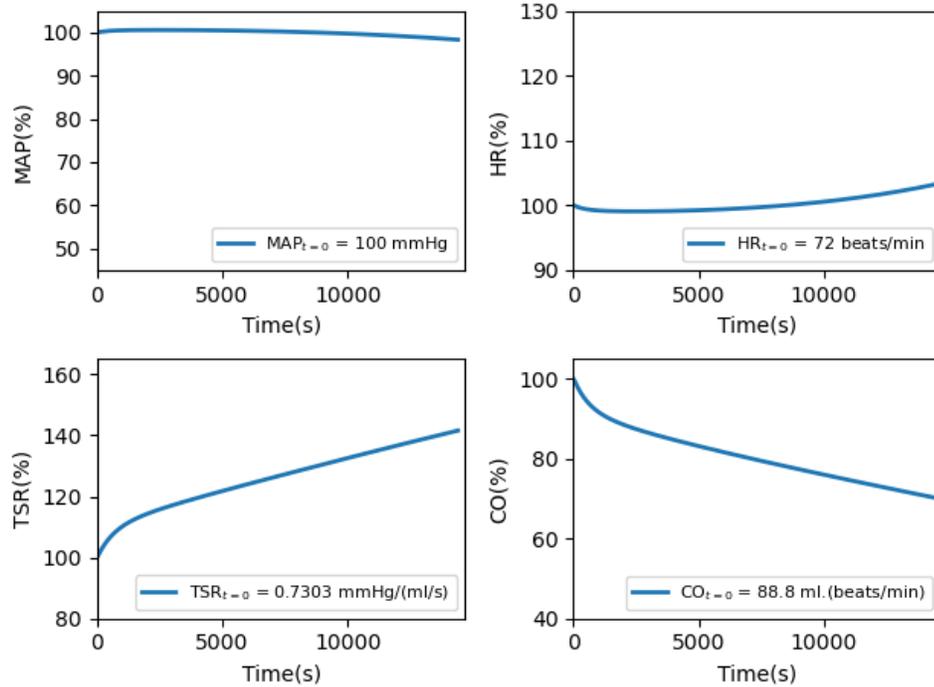


Figure 3.11: Simulated hemodynamic variables of Class 1 patient: No hypotension

Percentage changes of the most important hemodynamic variables namely, Mean Arterial Pressure (MAP), Heart Rate (HR), Total Systemic Resistance (TSR), Cardiac Output (CO), are presented for the 3 patient classes. In figure 3.11, one can see that a Class 1 patient can handle the removal of 3 litres of fluid within the treatment time without a considerable change in MAP or HR. Good hemodynamic stability is because of vasoconstriction (increasing TSR) and CO falls significantly because of ultrafiltration in HD. Figure 3.12 shows that for Class 2 patients, there is a reduction in MAP and an increase in HR in the second half of the treatment as TSR reaches saturation midway, leading to a mild hypotension. Normally, ESRD patients (Class 3) are characterized by poor hemodynamic stability and vascular refilling because of uremia, resulting in the sympathoinhibitory phase and in the exhaustion of both baroreflex mechanisms (high pressure and low pressure) towards the end of HD. Figure 3.13 replicates this phenomenon well and one can see a drastic drop in all

hemodynamic variables leading to syncope (fainting), for the same treatment settings.

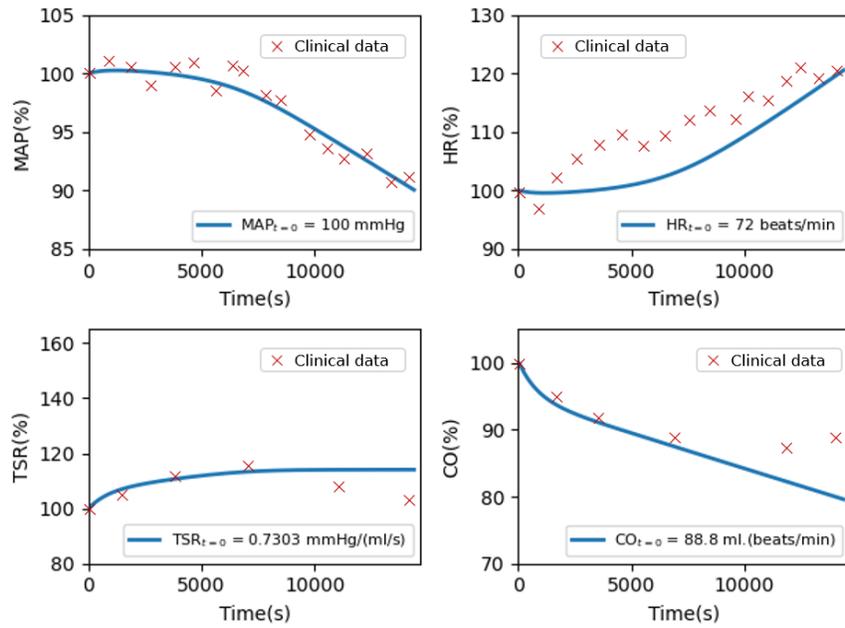


Figure 3.12: Simulated hemodynamic variables of Class 2 patient: Mild hypotension

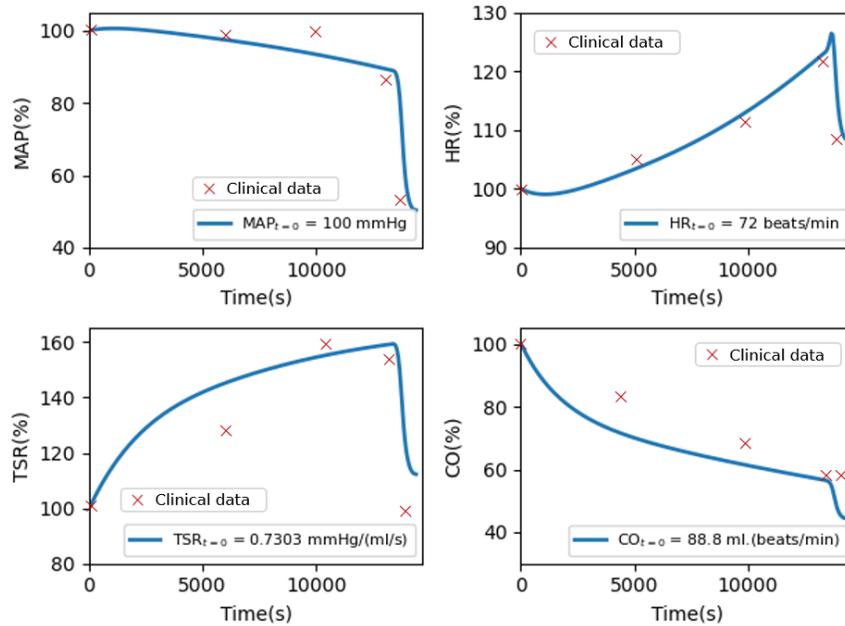


Figure 3.13: Simulated hemodynamic variables of Class 3 patient: Severe hypotension

b. Simulation results: HD system with complex model of hemodialyzer

In this section, the efficacy of the PDE mass transfer model of the hemodialyzer shall be tested. The hemodialyzer PDE equations were solved repeatedly for the same initial and boundary conditions but by changing the number of stages (nz) at each trial. It is evident from figure 3.14 that, if the hemodialyzer is split into 15 or more stages, the Root Mean Square Errors (RMSE) of the solute concentration predictions on the blood and dialysate side were almost zero. Therefore, 15 (computationally effective but accurate) was set as the number of discretization stages.

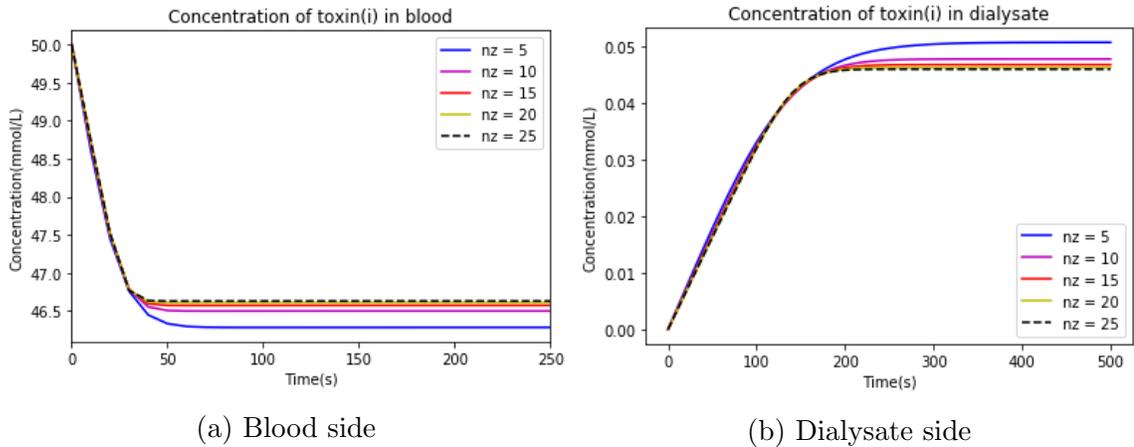


Figure 3.14: Determination of number of stages

Then the PDE model of hemodialyzer was plugged into Eq. (3.10) and slight modifications were done to the structure of Eq. (3.10) (following [1]). Then the simulation results were compared with raw clinical data and the model from Ziolkó's work [158].

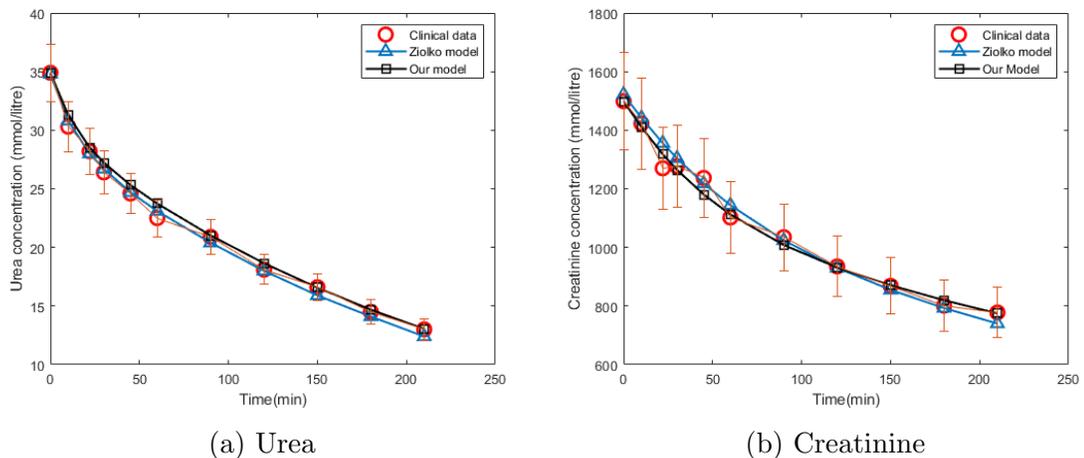


Figure 3.15: Clinical vs Model predicted results - Patient concentration profile

The simulation settings and the parameter values of the hemodialyzer are outlined in table 3.2 and they were adopted from literature [16], [158].

Table 3.2: Simulation settings for a HD system with the complex model of hemodialyzer

S.No	Symbol	Name	Range/Value	Unit
1	DW	Dry weight	58	kg
2	V_{tot}	Total fluid distribution	0.03364	m^3
3	V_{ex}	Extracellular volume	0.011213	m^3
3	V_{ic}	Intracellular volume	0.022426	m^3
4	Q_b	Blood flow rate	250	ml/min
5	Q_d	Dialysate flow rate	500	ml/min
6	Q_{uf}	Ultrafiltration rate	0 - 35	ml/min
7	G_s	Generation rate of solute(s)	0	$mmol/(Litre \cdot min)$
8	R_{gas}	Universal gas constant	$45 \cdot 10^{-6}$	$L \cdot atm/(mol \cdot K)$
9	T_{abs}	Absolute temperature	298	K
10	H	Hematocrit	0.44	
11	R_m	Radius of fibre	$45 \cdot 10^{-3}$	m
12	R	Radius of dialyzer	$220 \cdot 10^{-6}$	m
13	D_{urea}	Diffusivity coefficient of urea	10^{-11}	m^2/sec
14	$D_{creatinine}$	Diffusivity coefficient of creatinine	$0.365 \cdot D_s^{urea}$	m^2/sec
15	d	Membrane thickness	$8 \cdot 10^{-6}$	m
16	N	Number of fibres	12000	
17	n_z	Number of stages	15	

As it can be observed from figure 3.15, our HD system model with complex model of the hemodialyzer was able to satisfactorily explain the clinical data and it is at par with Ziolkó's model for both urea and creatinine. Error bars were drawn with Absolute Percentage Errors (APE) in successive clinical measurements (at the primary tube) of urea (APE = 7) and creatinine (APE = 11) [73]. Both the models were well within errors bars with our model (RMSE = 25.17) slightly performing better than Ziolkó's model (RMSE = 34.36) for creatinine. Our goal here was not to choose a superior model but to showcase to the reader that the complex HD system model could also replicate actual clinical data with acceptable levels of accuracy. The spatiotemporal concentration profile distribution of urea and creatinine (figure 3.16, 3.17) can be obtained by plotting the concentrations at all stages across the hemodialyzer in three-dimensional plots. During HD, blood enters at stage 1 always and loses toxins to the dialysate before leaving the dialyzer at stage 15. It can be observed that the dialysate always enters at stage 15 with zero concentrations of toxins and picks up toxins from the blood along the length of the hemodialyzer. One other factor which

could not be ignored in the plots is the dropping concentration gradient which drives the mass transfer process. This is because of the continuous purification of blood during HD thereby resulting in a fall in the concentration of toxins in the blood and also the total concentration gradient ($C_{Bi} - C_{Di}$).

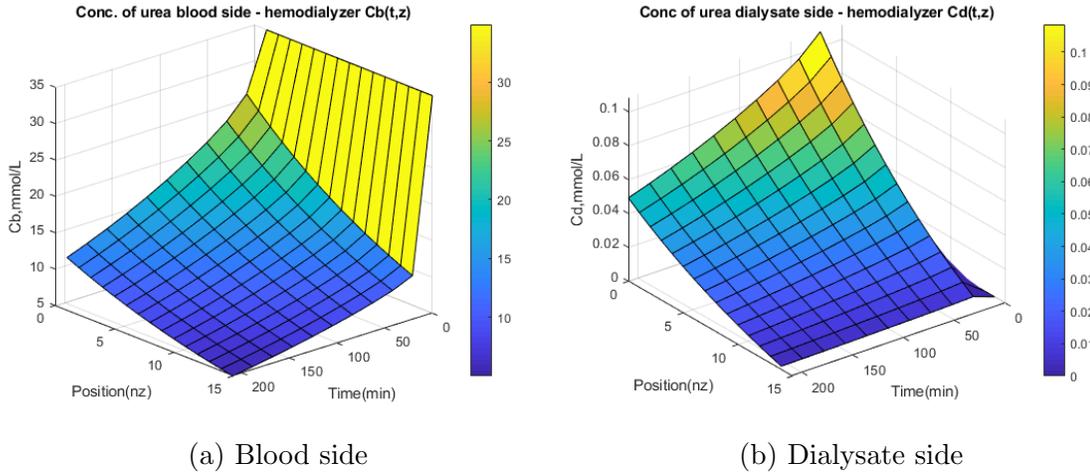


Figure 3.16: Spatiotemporal concentration profile of urea - Hemodialyzer concentration profile

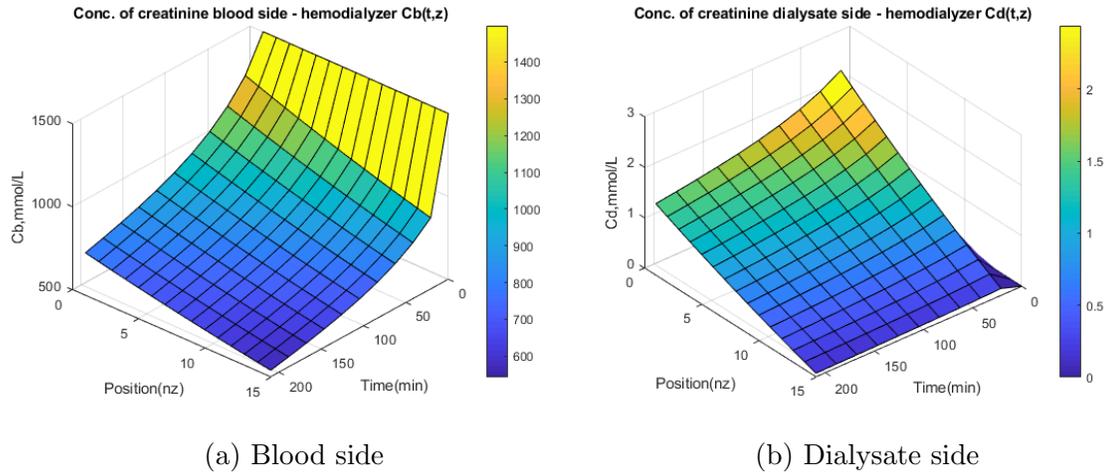


Figure 3.17: Spatiotemporal concentration profile of creatinine - Hemodialyzer concentration profile

c. Leveraging the power of the HD system model: Does UFR play an important role in decreasing the concentration of solutes?

During HD, hemodynamic stability is disturbed because overloaded fluid is removed at a much faster rate than the native human kidneys. The rate of vascular refilling

is often not able to compensate the swift decrease in body fluids by ultra filtration, thus resulting in hypovolemia. If the patient has a severe hypovolemic episode, treatment is stopped temporarily and this often results in inadequate HD. Several articles show the impact of profiling of UFR for better cardiovascular stability of the patients undergoing HD. To understand the significance of UFR in the final blood concentration of urea, it was assumed that 3 litres of fluid were removed from the patient approximately following different profiles (zero, constant, linearly decreasing, exponentially decreasing, step, pulse). It was noted that the final concentration is more or less closer to each other irrespective of the UFR profile chosen. The slight differences are due to manual errors in UFR profile design. By removing fluid from the patient, both mass and volume are removed, hence the concentration (ratio of mass and volume) remains the same while the total mass of the solute in the system decreases. The simple model-based simulation results (figure 3.18) confirm that the change in extracellular solute concentration is primarily because of diffusion across the semipermeable membrane and not because of UFR. So UFR should be seen as a variable which is free to be manipulated for achieving good weight management in HD.

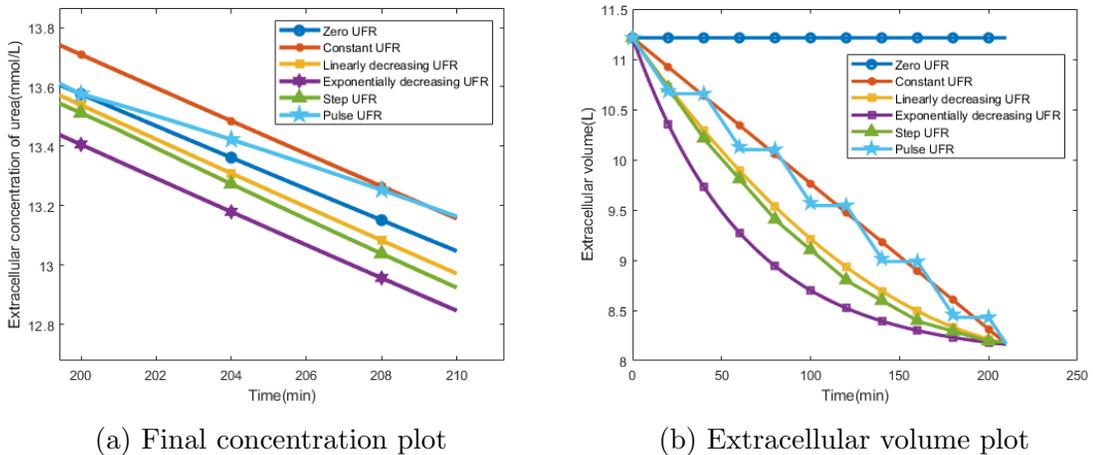


Figure 3.18: Comparison of UFR profiles

The above example was introduced to the reader to appreciate the benefits of having a mathematical model, based on first principles, of the physiological process (HD in our case) in hand. The user could answer some research questions and get a good understanding without in-vivo clinical experiments. Even from a systems engineering perspective, only a researcher with an indepth knowledge of the underlying phenomena can design better strategies to enhance the efficiency of a process.

3.2.5 Summary of the HD system model

There are several models in literature for representing the different subsystems in a HD system, but the work of Ursino and Innocenti [137] is the first attempt to put all HD subsystem dynamics (solute kinetics, fluid exchange, cardiovascular dynamics and mass transport in a hemodialyzer) equations in one single framework. Although this model has few limitations as outlined in [110], it offers the researcher an advantage to study the cardiovascular response (MAP, HP) of different classes of patients to different HD treatment settings. Our main objective in this research work is to build a comprehensive framework for HD treatment control, optimization and individualisation. As a first step, we plugged in the equation for dialysance based on flow rates of the process streams into the simple model, to make it control relevant. The ordinary differential equations (3.9), (3.10), (3.17), (3.18), (3.19), (3.30), (3.31), (3.32), (3.33), (3.34), (3.37), (3.42), (3.48), (3.54) and their necessary algebraic equations constitute the HD system model used in the coming chapters. Here, the algebraic equations can be substituted into the ODEs and so the HD system model is a Differential Algebraic Equations (DAE) system of index 0. A brief summary of the HD system model equations is shown in figure 3.19.

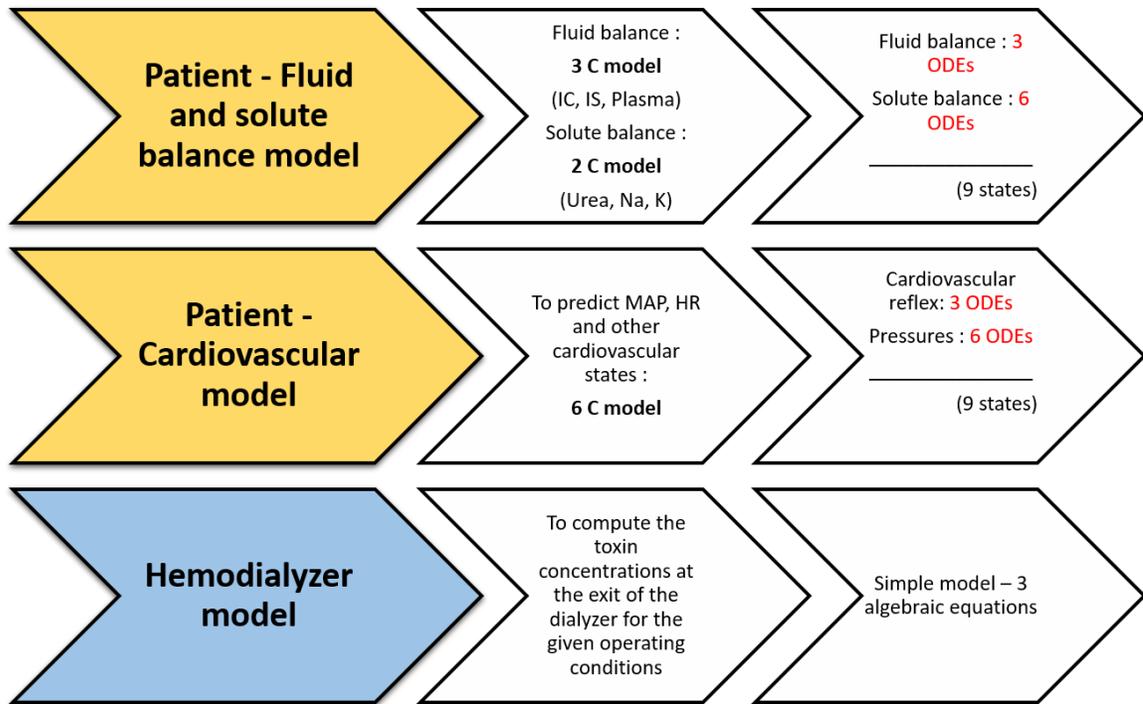


Figure 3.19: Summary of the chosen HD system mathematical model

If one opts to use the PDE equations in the hemodialyzer instead of the simple model, there will be 2 ODEs (one for blood side and one for dialysate side) for 1 solute for each stage. We have split the hemodialyzer into 15 stages and we need to study the dynamics of 3 solutes. Therefore, we will end up with $2 \times 3 \times 15 = 90$ extra ODEs just for the hemodialyzer alone. This might come as addition burden during state estimation and feedback control. Furthermore, only the exit concentrations are necessary to compute the adequacy of the HD treatment. Thus, only the simple model with 18 states will be considered for the future course of this research work.

3.3 Summary

In the beginning of this chapter, the guide to kinetic modelling as developed by the first few researchers who attempted to model the HD process kinetics was presented along with an example derivation and analytical solution of Single Pool Fixed Volume (SPFV) model. With the needs of the HD investigators and clinicians growing, the models grew more and more complex and getting an analytical solution with the help of a pen and paper would be cumbersome and resorting to numerical integration with the help of a computer would be the ideal choice. After getting into the context of modelling, the reader was introduced to all the fundamental mathematical model equations of different subsystems (Patient, hemodialyzer) of the HD system used in the course of this research work. It was proven by simulations that the chosen HD system model can represent the hemodynamics of different classes of patients. Furthermore, partial differential equations representing the mass transfer phenomena in a hemodialyzer were also explored and the performance was tested against actual clinical data from literature.

Chapter 4

Simultaneous state and parameter estimation of the hemodialysis system

Across the globe, every individual differs from every other individual and this bio-variability feature makes us the most unique among several millions of species which dwell on planet Earth. Although biovariability is speculated to be because of genetic plasticity, it is often quite interesting to note that even two identical twins behave completely different in different scenarios. The difference is due to the fact that biological processes are inherently variable. Technically, HD treatment is a biotechnological process which disturbs the patient's homeostasis in a short period of time with an objective of reaching treatment targets. Therefore, in HD treatments too each patient can behave differently to the same HD treatment setting. These complex variations can be quantified only with the help of a mathematical model. We have adopted a comprehensive mathematical model from literature to represent the HD system and that model has the capacity to replicate the most important hemodynamic variables of different classes of patients routinely encountered in a HD treatment centre. Although some test simulations for different classes of patients were shown in chapter 2, the model was simulated by making meaningful changes in physiological parameters instead of trying to find the best fit (by minimizing an error criterion function) between the clinical data and model results [140]. This chapter kicks off by explaining the initial attempts taken by the author to address the model individualization problem. Then later in this chapter, a framework for the simultaneous estimation of states and parameters based on the HD system model will be discussed in detail and the inferences made by the author are backed by extensive simulation results.

4.1 Initial attempts towards building an individualized virtual patient simulator

Our goal in this chapter was to individualize the HD system model (presented in chapter 2) which was adopted from literature, and can be used to represent the hemodynamics of different classes of patients by changing a few physiological parameter values. In the context of modelling, each patient's characteristics and behaviour can be represented by a unique set of parameter values pertaining to that particular patient. There are some sets of nominal parameter values given in literature [51], but every patient might not fit perfectly into the nominal value range and deviation from the nominal behaviour is not totally unavoidable. So to make this HD system model more user friendly and as a tool for optimal control of HD, an algorithm for identifying the state and parameter values of the patient based on clinical measurement data has to be devised. To achieve our goal, all variables of the HD system model are categorized into different groups as shown in table 4.1.

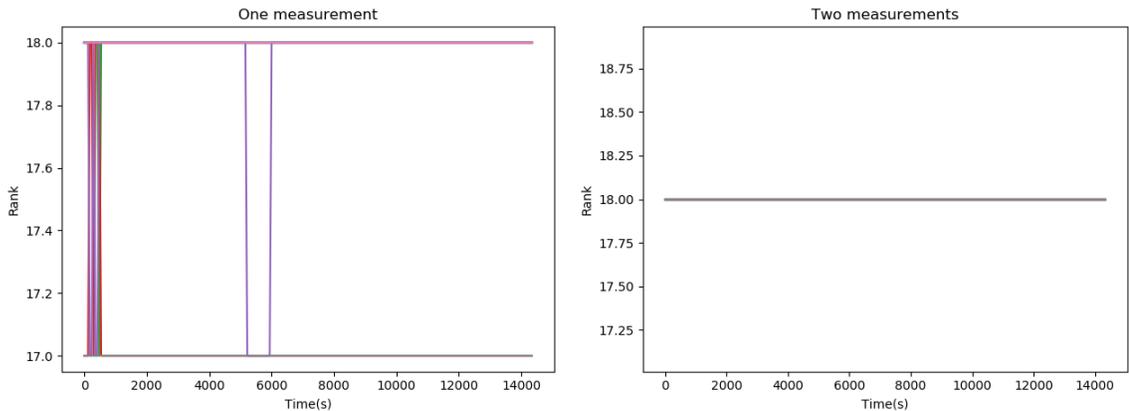
Table 4.1: Summary of model variables of the chosen HD system model with a simple model of the hemodialyzer

S.No	Description	Notation	List
1	States	x	$V_{pl}, V_{is}, V_{ic}, M_U^{ex}, M_U^{ic}, M_{Na}^{ex}, M_{Na}^{ic}, M_K^{ex}, M_K^{ic}, P_{sa}, P_{ra}, P_{pa}, P_{pv}, P_{la}, \delta, R_{s1}, T, V_{usv}$
2	Manipulated variables	u	$Q_{uf}, C_{Na}^D, Q_b, Q_d, Q_{inf}, C_{Na}^{inf}$
3	Parameters - Patient specific	θ_p	$C_{sa}, C_{sv}, C_{pa}, C_{pv}, C_{ra}, C_{la}, R_{s2}, R_{s3}, R_{sv}, R_{pa}, R_{pv}, L_a, L_v, H, V_{usa}, V_{upa}, V_{upv}, V_{ura}, V_{ula}, k_l, k_r, P_{la0}, P_{ra0}, M_{eq}^{ic}, M_{eq}^{ex}, C_p^{pln}, C_p^{isn}, E_{is}, P_{isn}, \Delta\sigma_R, \tau_R, G_{aR}, G_{cR}, \Delta\sigma_T, \tau_T, G_{aT}, G_{cT}, \Delta\sigma_V, \tau_V, G_{aV}, G_{cV}, P_{lat}, \sigma_{Rn}, \sigma_{Tn}, \sigma_{Vn}$
4	Parameters - Known	θ_k	$K_U^c, K_{Na}^c, K_K^c, \beta_U, \beta_{Na}, \beta_K, K_f, C_U^{inf}, C_K^{inf}, C_U^D, C_K^D, F_p, F_r, \gamma_U, \gamma_{Na}, \gamma_K, R_U^D, R_{Na}^D, R_K^D, \tau_\delta, G_\delta, \tau_\epsilon, G_\epsilon$

A first glance at table 4.1 will reveal that the HD system model has 18 state variables ($N_x = 18$). The reader can look into chapter 2, to get an understanding of how the ordinary differential equation representing each state variable was developed. Furthermore, the author has identified 42 patient-specific parameters ($N_{\theta_p} = 42$, excluding last three for proposed approach) which characterize an individual patient. The classification and grouping of patient-specific and known parameters shown here

is not unique. However, all parameters pertaining to the cardiovascular system were grouped into patient-specific parameters as these values differ for every individual. The list of parameters which describe the intercompartmental mass transfer dynamics of solutes were assumed to be known apriori and grouped under known parameters (θ_k). These parameters were fixed at their nominal values taken from literature. A dynamic system is said to be observable, if the initial states of the system can be uniquely determined from the knowledge of the available measurements over a finite time interval. This holds well for a linear time invariant system. But in our case, we have a nonlinear model to represent the HD system. Moreover, individualisation of the model can occur only if we estimate both the states and the patient-specific parameters. Firstly, we started by testing the observability of the system and then we examined if more variables (patient-specific parameters) can be estimated along with the states of the system with the available clinical measurement data.

Initially for testing the system observability, an observability matrix as described in Eq. (2.6) was built by using the linearized system matrices. The observability matrix built was severely ill-conditioned and the rank computation of the matrix ran into numerical issues (in MATLAB and Python). This was due to the fact that many entries in the observability matrix were below the numerical machine tolerance values and so the algorithms rounded off the entries to zero when the entries themselves are not perfect zeroes. In addition to that, the results might not be reliable because of the computation of higher powers of the ill-conditioned system matrix (A) while constructing the observability matrix. This comes as a serious consequence and resulted in a reduced rank of 2 for the observability matrix always. Increasing the number of measurements had no impact on the system observability.



(a) One measurement is available

(b) Two measurements are available

Figure 4.1: Observability analysis of the HD system using PBH test

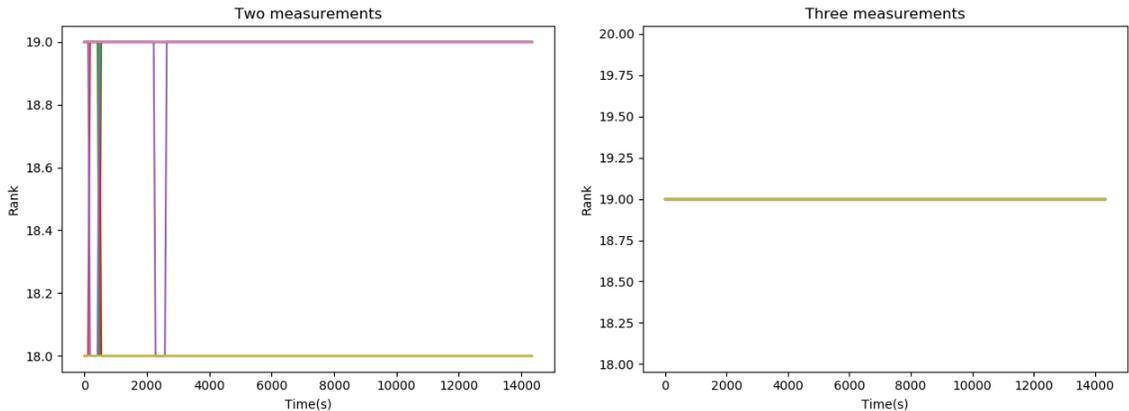
Then, we resorted to the Popov-Belevitch-Hautus (PBH) observability theory, which was found to be more numerically robust than the Kalman's observability matrix rank testing condition for our HD system model as it skipped the higher order matrix computation step. Since the HD system equations are nonlinear, the model should be linearized before PBH test could be applied. In literature, it is recommended that instead of linearizing the system at a single point, it should be linearized at different points along a typical treatment setting or operating trajectory [22]. In our case, the total number of states of the HD system model are 18. During PBH test, for the system to be observable, each of the nodes or eigen values of the linearized system matrix ($A(t)$) should have a rank of 18. The results from the PBH test are summarized below.

- Availability of one measurement (P_{sa} only): We observe in figure 4.1(a) that the rank of some of the eigen values fluctuates between 17 and 18 during simulation time and so some of the states of the system are unobservable. This means that all states of the system cannot be recovered from just one measurement in finite time.
- Availability of two measurements (P_{sa}, T): We see in figure 4.1(b) that the rank of all eigen values or nodes of the system are 18 during simulation time and so the system is observable. This means that all states of the system can be recovered from two measurements in finite time.

For any two measurements chosen from the set of 18 states, the system was observable. However, with just one measurement, the system was unobservable. Mean Arterial Pressure (MAP) and Heart Period (HP) or Heart Rate (HR) are two non-invasive clinical measurements which are generally available during HD treatments and so they were considered as available measurements for our analysis. Other states like the extracellular concentrations of urea, sodium and potassium could be considered as measurements but they are only available by blood sampling and the frequency of these states being measured is aberrant. Furthermore, these measurements are not available in a continuous fashion in clinical settings and for state estimation we need a continuous supply of sensor information. We understood that we needed atleast two measurements to estimate the initial states of the HD system model uniquely from measurement data, and it would be much better if the measurements were non-invasive. Although PBH test might have emerged to be a more numerically robust approach than the Kalman test condition to check the observability of our HD system, we designed a nonlinear state estimator (Extended Kalman Filter) to

estimate the states based on two measurements (MAP and HP) and verified the system observability results. If the readers are curious to skim through the Extended Kalman filter (EKF) results, they can quickly jump to chapter 4.

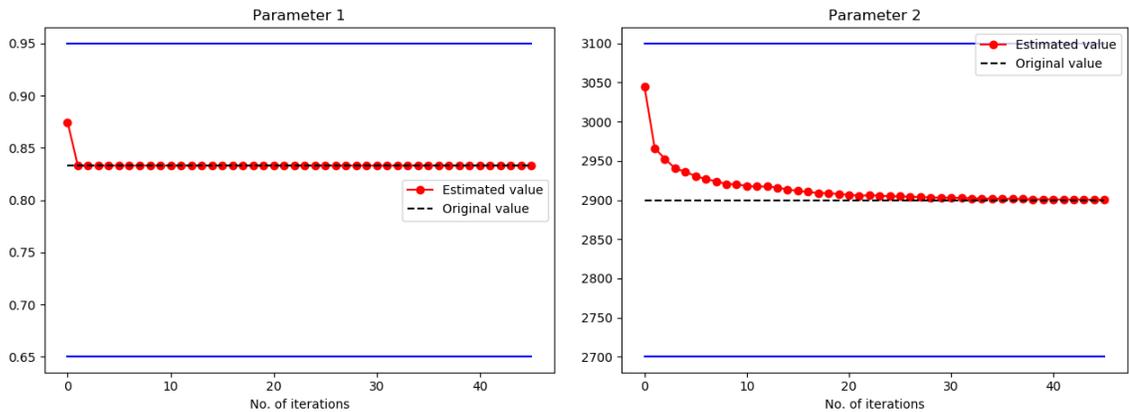
In the most practical scenario, some of the initial states of the patient and the patient-specific parameters are unknown. Hence, one has to estimate both the initial states and parameters simultaneously to identify the individualized virtual patient simulator if we do not have any blood sampling results in hand before HD treatment. As we know that our original system is observable if we have two clinical measurements (MAP and HP), we augmented the state vector with parameters (θ_p) to be estimated. For consistent estimation, the observability of the augmented system has to be ensured. We saw that the system was becoming unobservable when a parameter is added as an additional state as depicted in figure 4.2(a) by the rank fluctuations between 18 and 19 for some of the eigen values. The total number of state variables will be 19 because of augmentation, and the first sensitive parameter according to orthogonalization algorithm was the newly augmented state. The system could be made observable again, by increasing the number of measurements to three. Then, when the second sensitive parameter was added, four measurements were needed to make the system observable. This trend (typical linear system behaviour) continued for every new augmented state and the practical limitation now would be measuring the number of outputs required to make the system observable. The three measurements considered in figure 4.2(b) were plasma volume (V_{pl}), MAP (P_{sa}) and HP (T). V_{pl} cannot be obtained directly, but it can be obtained by measuring the hematocrit level at the start of dialysis and at any particular time during HD.



(a) Two measurements are available (b) Three measurements are available

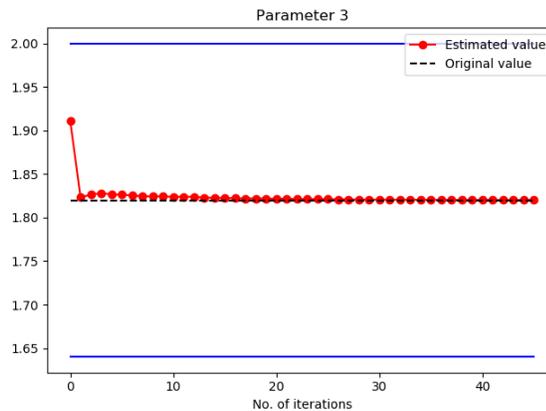
Figure 4.2: Observability analysis using PBH test for the augmented system with all states and the first sensitive parameter

Then a simulation study was conducted to check if the observations from the PBH test were accurate. Initially, we started with one augmented state and at every trail one additional parameter was added as an augmented state. An initial offset of 5% was added to every parameter (augmented state) and all the original states were initialized with a 3% offset. Based on the understanding we had from PBH test, we regarded that the user will be only able to estimate one extra parameter in addition to the original states of the HD system from three measurements. To our surprise, we were able to estimate more parameters than suggested by the PBH test. The results were the same, even in the case of noiseless and noisy measurements. The estimation problem was formulated as a nonlinear least squares problem, and the solution trajectory values of the augmented states at every iteration of the optimization solver for a 3 parameter case are shown in figure 4.3. The blue solid lines indicate the upper and lower bounds of the augmented states. We observe that all the augmented states converge to the true values before the end of the simulation.



(a) First augmented state (σ_{Tn})

(b) Second augmented state (σ_{Vn})



(c) Third augmented state (P_{ra0})

Figure 4.3: Estimation of top three sensitive parameters using three measurements

The simulation results cannot be well explained by PBH system observability test, which suggested that one could only estimate one extra parameter along with the states from 3 measurements and so we resorted to other modified nonlinear system observability tests in our proposed approach, as elaborated in the next section. Also in clinical practice, the number of non-invasive measurements are scarce and the number of data points are limited and so only MAP and HP are considered as available measurements in the user’s hand in the upcoming sections in chapter 3, and in chapter 4.

4.2 Proposed method

In this section, we propose a systematic procedure that can be used for the identification of the individualised patient model while ensuring that the entire system under consideration is theoretically observable. In practice, it would be hard for the modeller to determine the accurate values of the states and parameters of the HD system without any experimental analysis, and many of the states and parameters cannot be measured directly. We see that the original HD system is observable with just two measurements (MAP and HP) and this signifies that the original states can be recovered. However, we were not able to theoretically justify and quantify the number of parameters which could be estimated along with the original states of the HD system with the aid of PBH test. In our proposed approach, we aim to estimate the patient-specific parameters by augmenting them as additional states of the system [145]. It is mandatory to ensure that the newly augmented states are also observable.

Our proposed approach (figure 4.4) works towards checking the observability of the augmented system through sensitivity analysis followed by singular value decomposition. This enables the construction of the sensitivity matrix which is an equivalent form of the system observability matrix. Rank deficiency is examined by SVD analysis instead of using the rank computation algorithm in Python, which runs into numerical tolerance issues. In case of rank deficiency, the modeller has to identify the list of observable variables, thereby eliminating the list of unobservable variables by a forward sequential variable selection algorithm. The sequential variable selection method used in our approach is ‘orthogonalization’ and it takes into account the overall magnitude of the local sensitivities and the possible correlations between the original and the newly augmented states of the HD system. The orthogonalization method ranks the variables in the order of their importance. The cut-off value for the

orthogonalization algorithm is not based on a rigorous approach, instead the cut-off value is given in such a way which will ensure the selection of the largest possible observable subset of variables from the total list of variables. Our method gives the user a more robust estimability potential based classification of model variables and provides more insights into the HD system model, the quality of the available clinical data and measurements. During this method, the user will arrive at the largest subset of observable variables that could be estimated accurately from the available measurement data and the user will find if additional data collection (new data collection experiments or additional sensor information) is necessary to make the variables more estimable. Our ultimate goal is to use a scientific approach towards finding the largest optimal subset of observable model variables that guarantee full model reliability.

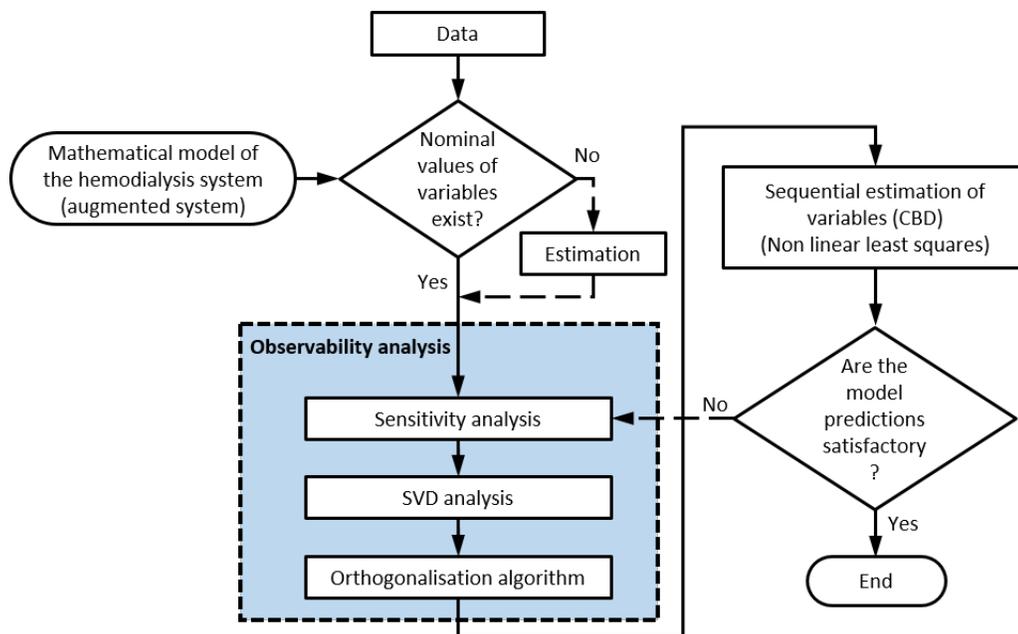


Figure 4.4: Proposed method for simultaneous state and parameter estimation

The proposed observability analysis framework aims at identifying the optimal variables of model variables from synthetic clinical data through a nonlinear optimization algorithm. For the steps outlined in the observability analysis framework, the user will need the vector of nominal values of the model variables, commonly assigned with the aid of prior information about the process under consideration. If the user does not have access to the nominal values of the model variables, then the nominal values of all variables can be simply obtained by running an optimization algorithm before starting the observability analysis [42]. Although there is a vast array of approaches for identifying the parameter values of a complex chemical process in literature, non-

linear optimization techniques have been widely employed owing to their efficiency and accuracy, [75], [113]. After analyzing the observability of the model, we have employed a Sequential Coordinate Block Descent (SCBD) framework for the nonlinear least squares estimation of decision variables. The total number of variables were split into blocks, thereby reducing the large optimization problem into a sequence of subroutines that can be run sequentially to arrive at the optimal solutions. As a consequence of SCBD, the variable importance ranking by orthogonalization can be infused into the coordinate selection rule, eliminating the need to identify the threshold values or cut-off values for determining the number of decision variables from the largest optimal subset of variables. The coordinate blocks will be filled with the top sensitive and least correlated variables first, and the last coordinate block will have the least sensitive variables. Finally, the estimated model is put to test. If the user is satisfied with the results, the method ends. However, if the user is not satisfied with the model prediction results, the observability analysis framework is repeated with the estimated values of variables and the entire process is repeated until the model prediction results are user satisfactory. A flow diagram of the proposed procedure is shown in figure 4.4. The key steps involved in this procedure for the practical implementation of the proposed method will be elaborated next.

4.2.1 Augmentation of the HD system model

We form an augmented HD system model by including the list of patient-specific parameters (θ_p) as additional states of the system (exactly in the same order mentioned in table 4.1) with zero dynamics, thereby forming a new augmented state vector. There arises some practical difficulties with this kind of system augmentation. The augmented system model becomes more nonlinear. This resulted in the under performance of PBH test and showed the deviation from a typical linear system as pointed out earlier. The last three θ_p parameters denoted by σ_{Rn} , σ_{Tn} and σ_{Vn} are assumed to be equal to the initial state values of R_{s1} , T , V_{usv} respectively and so they were removed from θ_p list. Now, our HD system model has 18 original states and 42 patient-specific parameters. After augmentation, the new state vector (X_a) will be of size 60 ($X_a = [x \ \theta_p]^T \in R^{N_x+N_{\theta_p}} \in R^{N_{X_a}} \in R^{60}$).

4.2.2 Nominal values of variables

Nominal values of all variables (X_a , θ_k) for a 70 kilogram patient belonging to any particular class (Class 1: Healthy, Class 2: Mild IDH and Class 3: Severe IDH) can be adopted from literature [110], [140]. The nominal values of patient variables of

any other body weight can be generated by weight scaling [140]. The list of model variables for which the nominal values can be generated by weight scaling includes the parameters associated with vascular compliances, elastance and unstressed volumes. For model applications, where the nominal values of variables are not available in literature, they can be estimated utilizing available experimental data [42].

4.2.3 Observability analysis framework

Once the user has identified the nominal values of all variables, observability analysis on the HD system model is commenced. Our model is nonlinear and the augmentation of the patient-specific parameters as extra states increases the nonlinearity further because of the position of the patient-specific parameters in the mathematical structure of the original HD system ODE equations. The nonlinear control theories were understood satisfactorily only in the late 20th century after the rise of computational power and resources. Generally, nonlinear system observability test includes the computation of Lie-derivatives and Lie-brackets. This step often puts the computing resources to a toss and is very cumbersome, occupying a huge amount of memory. We resort to a modified observability test based on sensitivity analysis as this approach has been found to answer the nonlinear biological system observability in literature [133], with a fair degree of accuracy. The minor details involved in each of the substeps of the observability analysis framework are outlined next.

a. Sensitivity analysis

The sensitivity equations (Eq. (2.9) to Eq. (2.12)) should be solved in parallel to the original HD system equations. There will be one sensitivity equation for each state of the augmented system (as shown in Eq. (2.10)), and the total number of sensitivity equations will be equal to N_{x_a} . Then, the sensitivity matrix is constructed. The sensitivity matrix can be thought of as a series of snapshot entries of the sensitivity dynamics that have been stacked vertically for each time instant [133]. In our case, we assumed that we have two clinical measurements (MAP and HP) in hand. So at each time instant, there will two rows of sensitivity dynamics entries, which each row corresponding to one measurement. To scale this approach well for a system like ours with a large number of variables, one has to ensure very accurate computation of the original model and sensitivity ODE equations leading to the construction of a numerically accurate and precise sensitivity matrix thereby minimizing potential errors in the computation of the singular values in the next step. To achieve this, the model was formulated in CasADi framework [4] for the accurate computation of

Jacobians and the numerical integration of all ODEs is done using ‘odeint’ imported from the Scientific Python (SciPy) library with the lowest possible absolute ($atol = 3 \times 10^{-14}$) and relative ($rtol = 3 \times 10^{-14}$) tolerances for maximum accuracy. Finally, the constructed sensitivity matrix is scaled as outlined in Eq. (2.14) and used for further analysis.

b. SVD analysis

Usually for a nonlinear system, the conditioning number of the sensitivity matrix might be very high making it very hard to accurately compute the rank of the matrix. During the analysis of Patient 4’s (a typical ESRD patient) data in table 4.4, the conditioning number of the sensitivity matrix was 1.2×10^{16} . The steps involved in the detection of rank deficiency are given below.

- Locating zero singular values is an important part of the process of finding the observability [133]. Presence of any potential zero singular values indicate the lack of observability.
- Plot the observability signature graph.
 - Plot singular values in log scale. A drop of 3.5 decibels or higher on log scale indicates the presence of singular values that can be considered to be a zero.
 - Plot the right singular vectors corresponding to the identified potential zero singular values.

The observability signature graph will help the user identify the total number of non zero singular values, which can be taken as the rank of the sensitivity matrix. Plotting the right singular vectors corresponding to the potential zero singular values will help the user to get an understanding of the null space of the sensitivity matrix.

c. Orthogonalization

For our HD system model, there are 60 variables to be estimated but we have only two measurements. One of the goals of this approach is to identify the largest subset of uncorrelated variables that can be estimated from the two available clinical measurements. In this work, the author resorts to orthogonalization algorithm, which is found to be superior to the SVD null space detection [123], for the forward selection of all the observable variables in a sequential manner. The scaled sensitivity matrix is

fed to this algorithm (as shown in figure 4.7) and the set number of iterations of the orthogonalization algorithm is equal to the number of non zero singular values from the previous step, but the algorithm could exit if other terminating conditions are satisfied. The Euclidean norm of a column indicates the degree of sensitivity, and the extent of linear dependence between two columns indicates the similarity of the effects of model variables [75]. At the very first step, the column with the highest 2 norm (as 2 norm is not sensitive to the direction of change) value from the scaled sensitivity matrix (corresponding to the most sensitive X_a) is selected and then an orthogonal projection is done to remove the effect of the selected column and a residual matrix is obtained. From the second iteration to algorithm termination, the column selection based on highest two norm is done on the residual matrix.

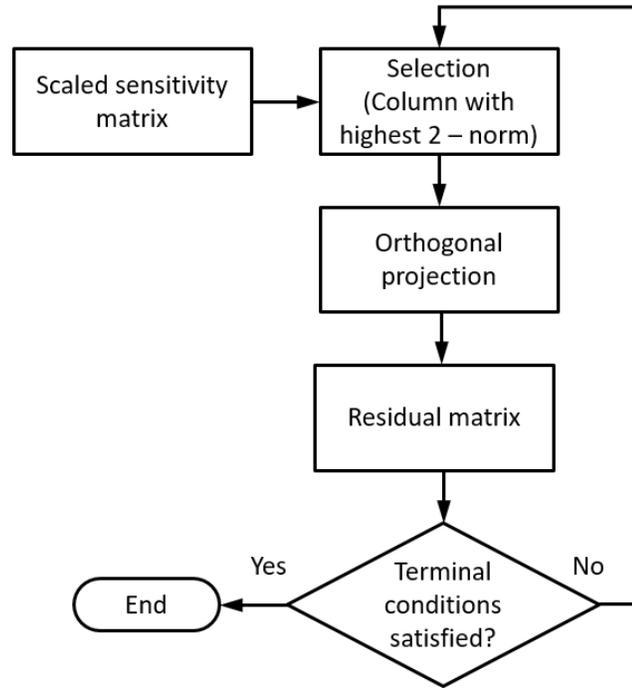


Figure 4.5: Orthogonalization method work flow

The cut-off value prescribed for the orthogonalization is 1×10^{-5} . The significance of setting a cut-off value is to break the algorithm if the highest two norm value of the remaining columns is lesser than the supplied cut-off value. The cut-off is set based on the nature of the system model with different authors using different cut-off values for their systems [85], [154]. In our work, the cut off value was set to ensure the selection of maximum number of observable variables from the total list of variables in the orthogonalization routine. The subset of variables returned by the orthogonalization algorithm shall become the decision variable set in the next optimization step.

4.2.4 Optimization problem formulation for simultaneous estimation of states and parameters

To carry out the estimation of variables, a nonlinear least squares problem was formulated. In the nonlinear least squares problem shown here, we have a model and actual (or synthetic) clinical measurement data points (MAP and HP). The idea is to run the model with different values of the decision variables to arrive at the optimal values of those variables which produces the least error of prediction. The prediction error is the difference between the actual (or synthetic) measurements and the model predicted values. The objective is to minimize the sum of squared prediction errors for both the measurements. The decision variables includes the set of states and the parameters of the HD system model, which are deemed to be observable following the results of the proposed observability analysis framework. All the other variables which turn out to be unobservable, are removed from the decision variable set and are fixed at their nominal values from literature. The mathematical description of the nonlinear least squares problem, for our application, is shown below.

$$\min_D J = \frac{1}{2} \sum_{i=1}^t (y_{mi} - \hat{y}_{mi})^2 \quad \forall m = 1, 2, \dots, N_y \quad (4.1)$$

$$\text{st. } \dot{\hat{X}}_a(t) = f(\hat{X}_a(t), u(t), \theta_k(t)) \quad (4.2)$$

$$\hat{y}(t) = h(\hat{X}_a(t), u(t), \theta_k(t)) \quad (4.3)$$

$$\hat{X}_a^c = X_a^{\hat{nom}} \quad \forall c \notin \{\bar{O}\} \quad (4.4)$$

$$D = \{\hat{X}_a^r\} \quad \forall r \in \{\bar{O}\} \quad (4.5)$$

$$LB(\hat{X}_a^r) \leq \hat{X}_a^r \leq UB(\hat{X}_a^r) \quad (4.6)$$

Table 4.2: Description of notations involved in the optimization problem formulation

S.No	Notation	Description
1	y_m	True output value
2	\hat{y}	Model predicted output value
3	t	Total number of measurements
4	D	Set of all decision variables
5	θ_k	Constant parameters of the HD system model
6	\bar{O}	A set containing the orthogonalization rank index
7	nom	Nominal value of a variable
8	$LB(\hat{X}_a^r)$	Lower Bound of a decision variable
9	$UB(\hat{X}_a^r)$	Upper Bound of a decision variable

a. Analysis of convexity

Norms are convex functions and the square of a norm is also a convex function. In general, a least squares problem is convex and it has a unique solution if the constraints are all convex too. But in the case of a HD system model, the constraints are non linear and non convex because the model itself appears in the constraints. Because of non convexity, one cannot find the global optimal solutions easily and one has to resort to global optimization algorithms, which are computationally very demanding, to find the true solutions for large scale problems. The initial guess given to kick start the optimization problem also affects the solution. In our case, the initial guess of all the decision variables were taken from literature and used with minor modifications. In our method, we propose to use a coordinate block descent algorithm instead of a global optimization algorithm to arrive at optimal solutions.

b. Optimization solution strategy: Sequential Coordinate Block Descent (SCBD)

Coordinate descent is an optimization algorithm that successively minimizes along coordinate directions to find the minimum of a function. At each iteration, the algorithm determines a coordinate or coordinate block through a coordinate selection rule, then exactly or inexactly minimizes over the corresponding coordinate hyperplane while fixing all other coordinates or coordinate blocks [152]. At every iteration, a line search along that particular coordinate direction can be conducted to find the optimal step size. One more advantage of this algorithm is that it can be applied in both differentiable and derivative free contexts.

In our test simulations, we discovered that splitting the original optimization problem into smaller sub problems was able to give us a better optimal solution. In addition to that, because the problems are solved sequentially in CBD, the chances of getting stuck in a local optimal solution could be minimized as the solver perturbs the previous solution by a small magnitude at the start of the next iteration, and if in case there is a decrease in cost function the solver will drive to a better, new optimal solution. Also, we had to find a way to infuse the order of the sensitivity ranking into the optimization algorithm and the sensitivity ranking from the orthogonalization algorithm can be naturally supplied as the coordinate selection rule in SCBD. The SCBD framework seemed to work well based on test simulations for our HD system model, and so it was adopted for this work. But the reader should not comprehend that this kind of a solution strategy will work for all problems.

The SCBD optimisation procedure can be executed by following the steps outlined below.

1. Optimization is done in a scaled domain so the movement is uniform along each direction and equal importance is given to both the measurements. This step is very important because the magnitude of both the measurements differ a lot and also there is huge difference in the magnitude of the decision variables.
 - Measurement data was scaled using the upper and lower bounds from their own data set.
2. Initial guess of all variables for different classes of patients are adopted from literature. Weight scaling is done wherever necessary and the initial guess values are supplied to the optimization solver.
3. Specify reasonable bounds for all variables.
 - The lower and upper body weight limits were assumed to be 60 and 100 respectively. So the lower and upper bounds for the weight scaled parameters can be generated from the nominal values of a 70 kilogram patient.
 - All the cardiovascular reflex parameters were assumed to vary between 0.2 to 3.5 times the nominal values.
 - For all the other parameters which are not included in the above categories, the variation is assumed to be $\pm 10\%$ from the nominal values in literature [140].
 - The bounds for the original state variables of the model are given below.

Table 4.3: Bounds for the state variables of the HD system model

	V_{pl}	V_{is}	V_{ic}	C_{U}^{ex}	C_{U}^{ic}	C_{Na}^{ex}	C_{Na}^{ic}	C_{K}^{ex}	C_{K}^{ic}	P_{sa}	P_{ra}	P_{pa}	P_{pw}	P_a	δ	R_{s1}	T	V_{usv}
LB	3375	9675	21750	10	10	137	5	4	137	95	3.8	16.15	6.65	6.175	0	0.693	0.79	2486
UB	5625	16125	36250	50	50	150	20	7	150	105	4.2	17.85	7.35	6.825	0	0.766	0.87	4142.85
Unit	ml	ml	ml	$\frac{mmol}{L}$	$\frac{mmol}{L}$	$\frac{mmol}{L}$	$\frac{mmol}{L}$	$\frac{mmol}{L}$	$\frac{mmol}{L}$	mmHg	mmHg	mmHg	mmHg	mmHg	No unit	$\frac{mmol \cdot s}{ml}$	s	ml

The first three states deserve special attention. The total fluid distribution volume (in litres) is assumed to be 0.58 times the body weight (in kilogram). The ratio of the extracellular to the intracellular fluid volume is taken to be 3:5 [158]. The patient is assumed to contain 75 millilitres of

blood per kilogram of body weight.

4. The nonlinear least squares problem is solved in SCBD framework.
 - In our case, we had 60 variables in total and each of the coordinate block is taken to be of size 10 variables. In total, there will be 6 coordinate blocks in total. For instance, if the orthogonalization algorithm finds 10 unobservable variables, then these variables will be removed from the decision variable set and there will be only 5 coordinate blocks in the optimization problem.
 - For example, first coordinate block = top 10 sensitive variables as per orthogonalization ranking, second coordinate block = next 10 sensitive variables and so on.
 - Minimization of the multivariate cost function (J) takes place along one particular coordinate block direction at a time.
5. When the value of the objective function changes in a magnitude smaller than ϵ ($J_{t-1} - J_t < \epsilon$), exit the loop. In our work, $\epsilon = 5 \times 10^{-4}$.

c. Selection of optimization solver

The numerical integration of the model equations and the optimization algorithms were implemented in Python using the ‘SciPy’ library. Least squares module (with 3 solver options) was imported from ‘scipy.optimize’ library. ‘TRF’ solver was used, when solving the optimization sub-routines because of its ability to work with constrained, large nonconvex problems. TRF (trust-region optimization method incorporated in the interior reflective Newton algorithm) is a simple, yet powerful approach to solve bound constrained nonlinear minimization problems. It is an algorithm of choice if one has a nonconvex problem with only bounds or with only linear equalities [29]. The trust region method has a different approach than the general gradient descent methods. Let’s assume $f(x)$ is the objective function, with x as a vector of decision variables, in which each decision variable is constrained by lower and upper bounds. The trust region algorithm approximates $f(x)$ with a quadratic function $q(s)$, which reflects the behaviour of $f(x)$ in a neighbourhood Δ , which is called the trust region around the current point x_k . The trust region sub problem of the method is to compute a trail step s by minimizing the area Δ . If $f(x_k + s) < f(x_k)$, the current point is updated to be $x_k + s$. This is a successful step and the trust region Δ can remain

the same for the next step. Otherwise, if the step turns out to be unsuccessful, and $x_{k+1} = x_k$, the trust region will be reduced for the next step. The idea of interior reflective method is to generate iterations x^k which is within the interior F , defined by the lower and upper bound constraints, using a reflective line search and to ensure convergence locally and globally. Since the interior F is bounded by constraints, an iteration will be reflected into the interior if that iteration lies on the boundary, and hence the name reflective line search. These enhancements help to avoid making steps directly into bounds and efficiently explore the whole space of variables. To obey the theoretical requirements, the algorithm keeps iterations strictly feasible. The TRF algorithm is written in MINPACK, a library of FORTRAN subroutines which efficiently solves the least squares minimisation of a residual of a set of linear or nonlinear equations. Dogleg (‘Dogbox’) algorithm with rectangular trust regions is typically used for solving small problems with bounds and it is not recommended for problems with rank deficient Jacobians. Levenberg-Marquardt (‘LM’) algorithm is the most efficient for small unconstrained problems, but does not handle bounds and so TRF was the author’s choice for solving the SCBD optimization sub-routines.

4.2.5 Validation of identified model

Inputs of different profiles and magnitudes were used for validation and the performance of the identified model is studied using the Mean Square Error (MSE) criteria. The MSE is given by the following expression,

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4.7)$$

Here, n is the total number of data points, y_i is the true output and \hat{y}_i is the estimated output from the model.

4.3 Design of simulation experiments

The fundamental raw material for our approach to be tested is a good data set. Due to the limitation of a vast array of HD data sets, simulation experiments have been designed to test and validate our proposed approach designed for identifying the individualized virtual patient simulator model. Existing clinical data of 7 different patients was taken from literature [30], and it has been enhanced with additional features pertaining to cardiovascular stability. The original data did not classify the patients according to their cardiovascular health. Patients 1, 3, 5, 7 were assumed to be hemodynamically stable patients with patient 1 displaying mild symptoms of IDH.

Patients 2, 4, 6 were prone to severe IDH, possess weak hemodynamic stability, and represent the ESRD patient group. We had to make this kind of a random assignment of cardiovascular health to the adopted patient data set because there were not many HD data sets in literature with hemodynamic data. Table 4.4 contains the data entries (body weight, cardiovascular health, initial blood concentration of urea and sodium and the bounds of DSC) for 7 patients as shown below.

Table 4.4: HD treatment settings for 7 different patients for synthetic clinical data generation. Data adopted from [30]

Patient Index	Body weight (kg)	Cardiovascular health	Initial urea concentration (mmol/L)	Initial sodium concentration in plasma (mEq/L)	Dialysate Sodium Concentration (DSC) (mEq/L)
1	72	H (Class 2)	24.5	140	143-152
2	74.6	U (Class 3)	32.5	140	147-152
3	90.3	H (Class 1)	20	140	145-152
4	70.8	U (Class 3)	26.8	142	148-155
5	76	H (Class 1)	36	143	140-149
6	75.8	U (Class 3)	30.2	142	145-152
7	76.4	H (Class 1)	29.5	140	144-153

Some of the characteristics of the chosen data set would attract the reader’s special attention. There are 5 men and 2 women in the patient pool, with a mean age of 62.0 ± 9.0 years and a mean dry body weight of 76.5 ± 6.4 kg. The residual renal function and residual diuresis of all the patients are 0 ml/min and 0 ml/24 h respectively. The systolic blood pressure at the beginning of the treatment session is 100 ± 5 mmHg.

4.3.1 Generation of synthetic clinical data

Firstly, the nominal values of all variables of the patient are generated. There are some blood concentration results of solutes in table 4.4 which could be factored in while generating the concentration values. The extracellular and intracellular concentrations of other solutes are assigned in a way that ensures osmotic equilibrium before the start of the HD treatment session. Weight scale factor is multiplied to the variables which are dependant on the patient’s body weight. Then based on the class of the patient, the corresponding cardiovascular reflex parameters from literature [140] are chosen as nominal values. Then suitable random noise sequences (high noise of 20% standard deviation and zero mean for all cardiovascular reflex parameters and low noise of 2% standard deviation and zero mean for all other variables) are added

to all variables to generate a completely new ideal patient. As the variables were of different magnitudes, they were scaled (with the help of the lower and upper bounds mentioned in subsection 3.2.4.b.) before adding the noise sequences and then they were unscaled to get the true values. True MAP, HP data (synthetic clinical data) was collected at a sampling time of 1 minute, by simulating the model with the true values of variables. Then these true values of the variables of that particular patient were treated as unknown values (no prior information) and our proposed approach is employed to arrive at the optimal values starting from the nominal values of all variables which will closely the ideal patient’s hemodynamic behaviour. All patients came to the treatment centre with a fluid overloaded state and we assumed that 3 litres of fluid were removed by ultrafiltration from each patient for reaching the weight targets. Each of the HD treatment sessions ran for 4 hours. In table 4.1, the author identified 6 manipulated variables that can perturb the HD system. Only the first two (UFR and DSC) will be used for this chapter’s analysis and the others shall be fixed at the nominal operating values outlined in literature. All θ_k parameters shall be held at the nominal values mentioned in literature [140].

4.3.2 Building of test cases

We have created 4 different case settings to test the efficiency of our proposed approach. All these cases will use only two available clinical measurements (MAP and HP) for the identification of the individualized virtual patient simulator.

- Case 1:
 - Noise free measurements
 - Coordinate selection rule = Random
 - Block size = 10
- Case 2:
 - Noise free measurements
 - Proposed observability analysis method’s information infused
 - Coordinate selection rule = Orthogonalization ranking order
 - Block size = 10

- Case 3:
 - Noise (Mean = 0, Standard deviation = 0.1) corrupted measurements
 - Coordinate selection rule = Random
 - Block size = 10

- Case 4:
 - Noise (Mean = 0, Standard deviation = 0.1) corrupted measurements
 - Proposed observability analysis method’s information infused
 - Coordinate selection rule = Orthogonalization ranking order
 - Block size = 10

4.4 Simulation results and discussion

In this section, the results obtained from simulation experiments for the four different case settings will be presented to the reader. We have three different classes of patients who come to the HD treatment centre and so one representative patient shall be selected from each patient class and the results of our proposed approach will be applied on that particular patient data and explained with supporting figure plots wherever necessary. As this thesis work is dedicated to the better treatment management of ESRD patients, let us start by analyzing the results for a patient belonging to class 3 (ESRD) patient with weak hemodynamic stability) first in much detail. Tables 4.5, 4.6 contain the optimization and model validation results and quite often the reader will be directed here while explaining the results for the different classes of patients.

Table 4.5: Optimization results summary - Case 1 & 2

Patient Index	Initial cost	Case 1						Case 2				
		Final cost	Training MSE		Validation MSE		Final cost	Training MSE		Validation MSE		
			M1	M2	M1	M2		M1	M2	M1	M2	
1	142.59	0.604	0.035	2.55E-05	0.033	3.48E-05	0.235	0.016	8.16E-06	0.019	1.02E-05	
2	674.56	0.744	0.363	7.71E-05	1.302	2.91E-04	0.352	0.043	1.88E-05	0.04	1.96E-05	
3	76.59	0.113	5.91E-04	8.90E-08	9.99E-04	2.31E-07	0.057	2.69E-04	4.95E-08	8.55E-04	1.59E-07	
4	54.111	3.618	20.897	6.63E-04	52.044	1.22E-03	0.7	4.175	1.14E-04	6.102	3.66E-04	
5	11.72	0.007	6.66E-05	9.04E-09	0.007	8.97E-07	0.002	2.25E-05	3.14E-09	0.001	1.39E-07	
6	65.703	2.346	8.551	2.85E-05	32.31	2.53E-04	0.226	0.648	4.20E-06	2.809	1.85E-05	
7	38.45	0.021	8.56E-05	9.62E-09	0.003	3.05E-07	0.002	8.80E-06	1.24E-09	3.12E-04	4.13E-08	

Table 4.6: Optimization results summary - Case 3 & 4

Patient Index	Initial cost	Case 3					Case 4				
		Final cost	Training MSE		Validation MSE		Final cost	Training MSE		Validation MSE	
			M1	M2	M1	M2		M1	M2	M1	M2
1	100.23	1.816	0.16	1.00E-04	0.569	2.81E-04	2.119	0.191	1.15E-04	0.522	2.75E-04
2	503.49	1.88	0.418	8.42E-05	0.758	1.43E-04	3.282	0.9469	2.07E-04	0.79	1.46E-04
3	42.09	1.481	0.013	2.70E-06	0.108	1.70E-05	1.323	0.011	2.43E-06	0.088	1.46E-05
4	28.65	3.074	27.642	1.39E-03	32.217	2.23E-03	1.749	13.349	1.12E-03	20.644	1.34E-03
5	8.2	1.35	0.022	3.64E-06	0.201	2.81E-05	1.35	0.021	3.63E-06	0.2	2.79E-05
6	38.26	4.447	29.988	8.31E-05	46.169	1.47E-04	1.998	10.155	6.61E-05	13.231	6.67E-05
7	22.88	1.478	0.009	1.51E-06	0.103	1.30E-05	1.342	0.008	1.40E-06	0.094	1.30E-05

The tables presented above contain the initial (initial cost) and final (final cost) values of the objective function, computed at the starting and ending of the optimization routine respectively. In addition to that, the MSE between the true data set and the estimated data set is presented for both the training and validation profiles. The input profiles which were used to train the virtual simulator were called as training profiles, and the input profiles which were not exposed to the model during training are called as validation profiles. The use of a validation profile here is to test performance of the estimated HD system model under completely new treatment settings. By this way, the user can be assured that the estimated patient simulator will work under a wide range of settings. Different profiles (constant, step) of manipulated inputs were chosen for validation because our ultimate goal was to embed this nonlinear model into a controller and the optimal trajectory computed by the model predictive controller could take any profile as long as the input and process constraints are satisfied.

4.4.1 Class 3 patient: Severe IDH & weak hemodynamic stability

Let us first discuss the results for a class 3 patient from table 4.4, following the complete workflow of our proposed approach. The representative patient chosen for explaining the results is patient 4. The body weight of patient 4 is 70.8 kg and if 3 litres of overloaded fluid are removed in a 4 hour treatment time, then the final dry body weight would be 67.8 kg. The nominal values of all the variables of the mathematical model was calculated using the initial body weight. The weight scale factor in this case would be 70.8/70. This factor was multiplied to all variables pertaining to body vascular compliances, elastances and unstressed volumes. The total fluid distribution volume can be calculated by using the simple formula from literature [158], and it would be $0.58 \times 70.8 \times 1000$ ml which is equal to 41064 ml. We

have assumed that the ratio of the extracellular to the intracellular fluid is 3:5 and so the initial fluid compartment volumes can be computed accordingly. The nominal values of other variables were directly adopted from literature [140]. Then we would have generated the nominal values of all the variables in the augmented state vector and we need this to start the observability analysis of the HD system model. The following figure 4.6 represents the input output data set used for training the HD system model to make it an individualized virtual patient simulator. Clearly, there is a difference between the true values (yellow trajectory) and the nominal values (blue trajectory) of variables used to simulate the model as the trajectories are completely different. Our objective is to test if our proposed method will help identify the optimal values of variables that will closely resemble the true patient’s hemodynamic behaviour.

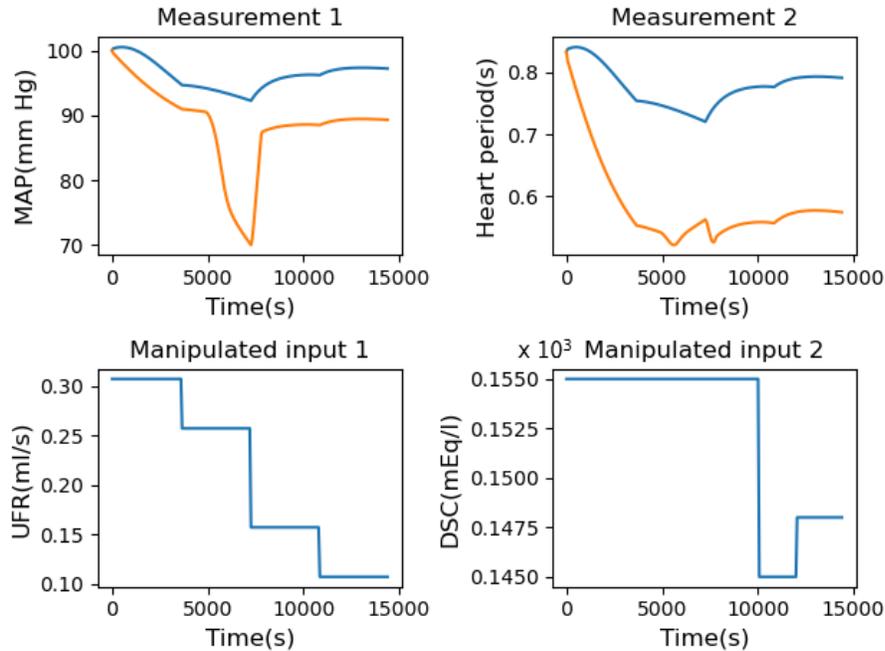
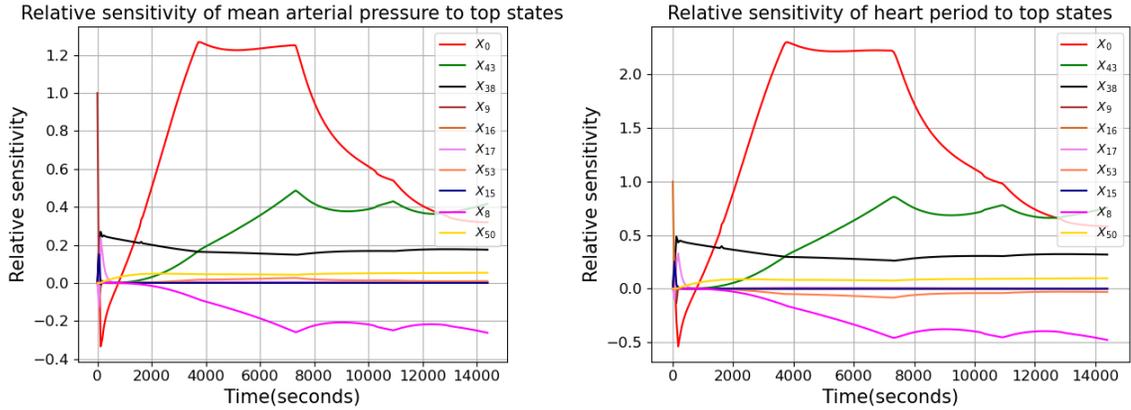


Figure 4.6: Comparison of simulations with nominal (blue) and true (yellow) values of states and parameters

As the first substep in the observability analysis framework, sensitivity analysis shall be done around the nominal values of the variables by solving the sensitivity equations in parallel with the original model equations. The sensitivity dynamics seeds from the original model equations. The authors of the adopted mathematical model have attempted sensitivity analysis investigation of the model by manually perturbing a few parameters of the model, collecting the hemodynamic data at intervals of 1 hour,

and analyzing them [139]. But our approach of solving the sensitivity equations of all variables together accounts for both the direct and indirect (interactions) impact of the variables on the measurement data and gives a more comprehensive understanding of the physiological phenomena. After solving the model and sensitivity ODEs (with a sampling time of 1 minute), the user can construct the sensitivity matrix as expressed in Eq. (2.12). For clarity, the relative sensitivity trajectories of the top 10 most sensitive states with respect to both measurements (MAP and HP) are plotted in figure 4.7.



(a) Measurement 1 (MAP) to top 10 states (b) Measurement 2 (HP) to top 10 states

Figure 4.7: Relative sensitivity trajectories of measurements for Patient 4

At the initial few times instants, the reader will observe that the relative sensitivities of most of the states fluctuate much faster, indicating that the initial states affect the model to a greater extent and that reasonable initial states are to be supplied for fair model prediction. However, the relative sensitivity of only one state (X_a^0) goes beyond 1.0. Most of the other relative sensitivities are below 0.5, indicating an acceptable level of robustness of the model. The plot shows that the relative sensitivities continue to grow or decrease for all the states throughout the entire treatment session. Towards the end of the treatment, the reader will observe that none of the relative sensitivities are above 1.0, pointing out that none of the states affect the model outputs by more than the actual change in state values (in terms of percentage). If the reader examines figure 4.7 carefully, X_a^0 and X_a^{43} emerges out as the states to which both measurements are most sensitive. X_a^0 and X_a^{43} in our case, denote the plasma volume and the plasma protein concentration of the patient. HD removes overloaded fluid directly from the circulatory system and there is a drop in blood plasma volume. All the cardiovascular compensatory responses

(vasoconstriction and the rising of heart rate) are triggered because of the drop in the arterial pressure. Arterial pressure drop is directly proportional to the drop in plasma volume. There is no wonder that blood plasma comes out as the state to which MAP and HP are most sensitive. Plasma protein concentration is responsible for vascular refilling. If fluid from plasma is removed at a faster rate than vascular refilling, then there is a drop in blood pressure and a rise in heart rate until the reflex mechanisms reach their saturation limits.

Then the next step in our work flow is to do the SVD analysis on the sensitivity matrix. This will enable us to plot the observability signature graph (figure 4.8). The singular values are plotted on the upper half of the image and the right singular vectors corresponding to any potential zero singular values are plotted on the bottom half of the image. The right singular vectors corresponding to the zero singular values provides detailed insights into the variables of the model which are linearly related and cause the overall system to be unobservable. In supporting literature [133], any drop of more than 3.5 decibels on the log scale of singular values can be taken as a clear sign of unobservability. We encounter a drop of 4.6 decibels after the 52nd singular value. This indicates that the rank of the sensitivity matrix is 52 and there are 52 observable variables out of the total 60 variables. The user has to identify the 8 unobservable states (or nodes) in the HD system during the optimization routine.

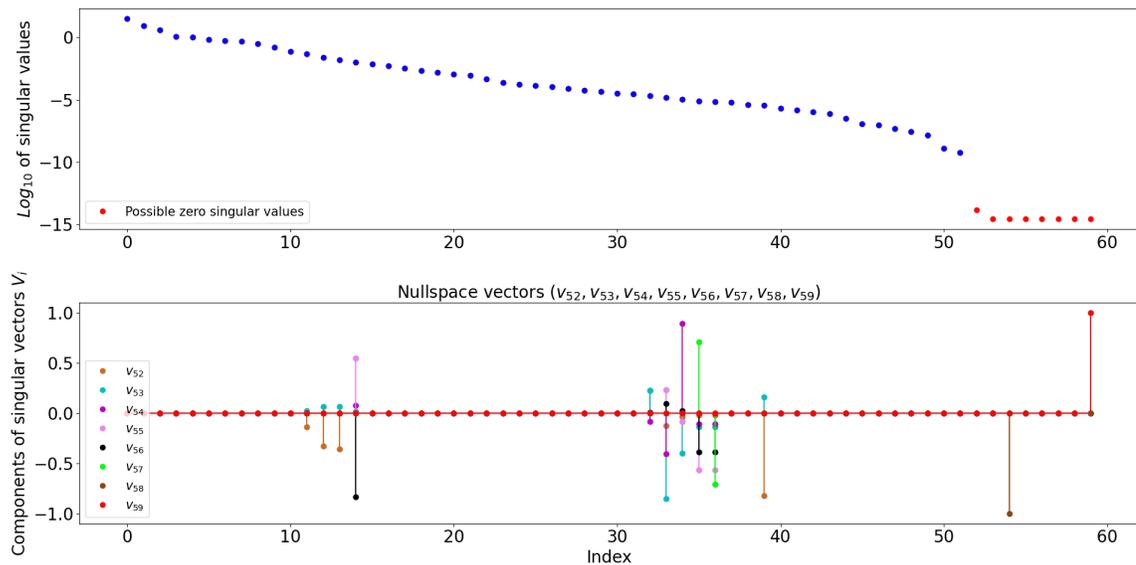


Figure 4.8: Observability signature graph: SVD analysis of the sensitivity matrix for Patient 4

Looking at the right singular vectors plot corresponding to all the identified zero singular values, give us an idea of all the variables which are involved in total correlation and does not give any information about pairwise correlation between two variables. $X_a^{11}, X_a^{12}, X_a^{13}, X_a^{14}, X_a^{32}, X_a^{33}, X_a^{34}, X_a^{35}, X_a^{36}, X_a^{39}, X_a^{54}, X_a^{59}$ are the variables involved in total correlation. States with indices 11 to 13 refer to the cardiovascular pressure states and states with indices 32 to 36 refer to the unstressed volumes of the vascular compartments. States with indices 14, 54 are always zero at the beginning of the simulation and has no impact on the measurements. The reader should understand that there are only 8 unobservable variables among these 12 variables. We could remove all these 12 variables from the total variable list, but our goal is to obtain the largest subset of observable variables. This is why we resorted to a sequential selection algorithm which would forward select the variables one by one until the terminating conditions are satisfied.

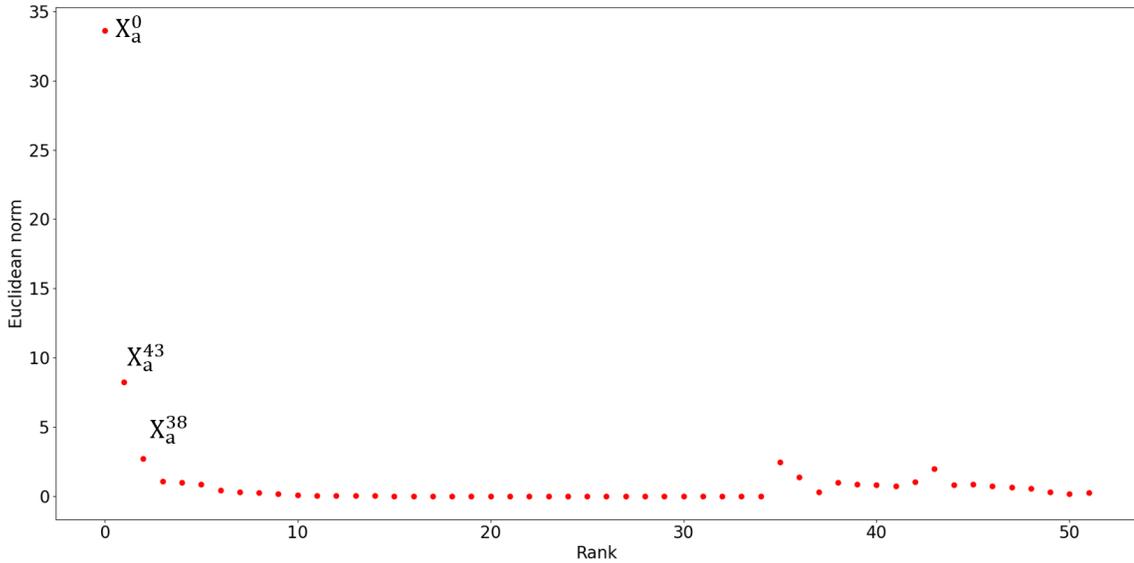


Figure 4.9: Euclidean norm of the ranked variables by orthogonalization

We know there are 52 observable variables from the previous step and so the iteration counter for the orthogonalization algorithm is set at 52. The algorithm runs selecting one variable at a time starting from the most sensitive and least correlated variable to the least sensitive and most correlated variable. The 52 variables returned by the algorithm are automatically arranged in the decreasing order of sensitivities (estimability potential based on the two available measurements in hand) as shown in table 4.7. The reader can notice that most of the cardio vascular reflex variables come up in the ranking as reasoned by the original authors of the HD system model [139].

The eight variables removed by the orthogonalization algorithm include X_a^6 , X_a^{10} , X_a^{13} , X_a^{14} , X_a^{29} , X_a^{31} , X_a^{54} and X_a^{59} . The algorithm has clearly removed all the variables which had zero impact on the measurements namely, X_a^{14} and X_a^{54} . We can observe from figure 4.9 that a substantial number of variables show sensitivity values closer to zero. This clearly indicates that the model is over-parameterized. However, commenting on this aspect of the model is beyond the scope of this research work and we assume that all the variables (as introduced by the original authors of the HD system model) are necessary to reproduce the actual physiological phenomenon. The subset of observable variables returned by the orthogonalization algorithm becomes the the decision variable set in the optimization routine.

Table 4.7: Ranking of uncorrelated variables by orthogonalization algorithm for Patient 4

Identifier	Ranking based on estimability potential
Notation	$V_{pl}, C_p^{pln}, k_r, P_{sa}, T, V_{usv}, G_{aT}, R_{s1}, M_K^{ic}, G_{cR}, P_{pv}, \Delta\sigma_V, C_{sv}, V_{ic}, P_{pa}, \Delta\sigma_T, M_{Na}^{ex}, k_l, G_{cV}, E_{is}, \Delta\sigma_R, C_{sa}, G_{aV}, \tau_T, M_U^{ex}, G_{aR}, \tau_R, \tau_V, L_v, R_{s2}, P_{ra0}, V_{usa}, V_{is}, P_{la0}, C_{pv}, V_{upa}, V_{upv}, P_{isn}, C_{ra}, C_{la}, C_p^{isn}, C_{pa}, R_{sv}, R_{s3}, M_U^{ic}, M_{eq}^{ex}, R_{pa}, V_{ura}, V_{ula}, M_{eq}^{ic}, M_K^{ex}, R_{pv}$
Index	0, 43, 38, 9, 16, 17, 53, 15, 8, 50, 12, 55, 19, 2, 11, 51, 5, 37, 58, 45, 47, 18, 57, 52, 3, 49, 48, 56, 30, 24, 40, 32, 1, 39, 21, 33, 34, 46, 22, 23, 44, 20, 26, 25, 4, 42, 27, 35, 36, 41, 7, 28

The rank index and the list of observable variables are passed on to the optimization block where the nonlinear least squares problem is solved in Sequential Coordinate Block Descent (SCBD) framework with an objective to identify the optimal values for all observable variables returned by the orthogonalization algorithm. In SCBD framework design, each coordinate block should accommodate a maximum of 10 variables and the rank index array is supplied as the coordinate selection rule. Therefore, the coordinate blocks are filled in the order of the orthogonalization ranking and the list of variables in each coordinate block (for cases 2 and 4) is shown below.

1. First block = $[V_{pl}, C_p^{pln}, k_r, P_{sa}, T, V_{usv}, G_{aT}, R_{s1}, M_K^{ic}, G_{cR}]$
2. Second block = $[P_{pv}, \Delta\sigma_V, C_{sv}, V_{ic}, P_{pa}, \Delta\sigma_T, M_{Na}^{ex}, k_l, G_{cV}, E_{is}]$
3. Third block = $[\Delta\sigma_R, C_{sa}, G_{aV}, \tau_T, M_U^{ex}, G_{aR}, \tau_R, \tau_V, L_v, R_{s2}]$
4. Fourth block = $[P_{ra0}, V_{usa}, V_{is}, P_{la0}, C_{pv}, V_{upa}, V_{upv}, P_{isn}, C_{ra}, C_{la}]$

5. Fifth block = $[C_p^{isn}, C_{pa}, R_{sv}, R_{s3}, M_U^{ic}, M_{eq}^{ex}, R_{pa}, V_{ura}, V_{ula}, M_{eq}^{ic}]$
6. Sixth block = $[M_K^{ex}, R_{pv}]$

For cases 1 and 3, all the 60 variables are taken as the decision variable set. The unobservable variables are not removed. Furthermore, the coordinate blocks are filled in random order and the orthogonalization ranking information has no role to play in cases 1 and 3. This will help us compare the results obtained from different case settings and justify if our proposed method was a good step towards solving the model individualization problem. Let us examine the raw (synthetic) clinical data of Patient 4 plotted in figure 4.10. MAP and HP measurements are always prone to measurement errors and so errorbars of 5% were drawn over the reading given by the clinician (over the model simulated data in our case), indicating that the true value of the measurements could be anywhere within the range. MAP is the most important physiological variable which directly correlates with the hemodynamic stability of the patient and HP (or HR) is an indicator pointing out if the patient’s cardiovascular reflex mechanism has been triggered or not.

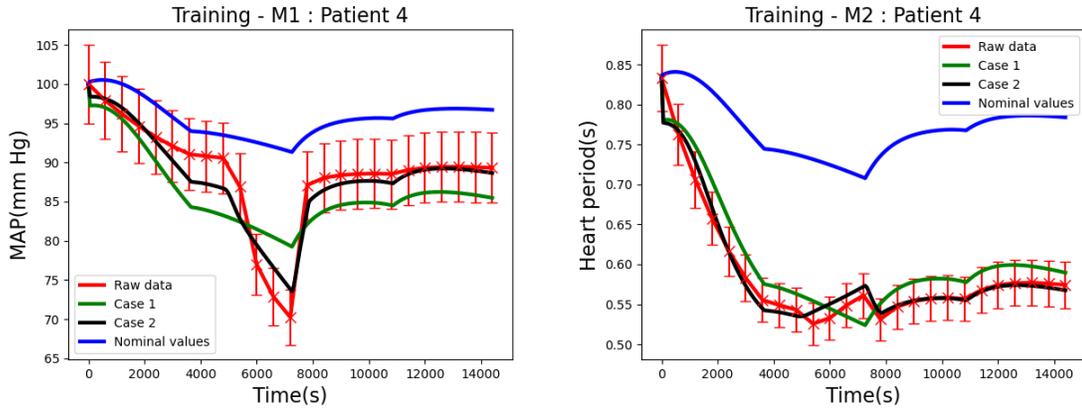


Figure 4.10: Patient 4: Training (Case 1: $MSE_{M1} = 20.897$, $MSE_{M2} = 6.63E-04$, Case 2: $MSE_{M1} = 4.175$, $MSE_{M2} = 1.14E-04$)

In figure 4.10, we observe that Patient 4’s MAP drops ever since the treatment started and displays symptoms of severe IDH right in the middle of the second hour of treatment. It takes almost an hour for the patient to come back to normalcy if the HD treatment is continued. In clinical practice, if severe IDH occurs, the treatment is stopped and the patient is kept in Trendelenburg position for a while. IDH could be due to a lot of complex physiological factors, but IDH in a hemodialysis context mainly occurs because a uremic patient’s reflex mechanism fails to cope with the rapid

reduction in plasma volume by HD. There is a steady decrease in heart period too, indicating that the heart rate of the patient rises until the saturation level of the reflex mechanism. The model is trained on this data set and we clearly see that the model simulations (for MAP and HP) with the nominal values of variables are nowhere close to capturing the actual hemodynamics of the patient. SCBD optimization is observed to improve the prediction results substantially. Although the initial guess given to the optimization solver in both cases 1 and 2 is the same (initial cost = 54.111), in the absence of measurement noise, case 2 (final cost = 0.7) with our proposed method has reached a better optimal solution than case 1 (final cost = 3.618) with no observability analysis information. Case 2 succeeds in capturing the severe IDH dynamics, while Case 1 fails. Furthermore, the model simulations with the estimated values of variables from case 2 always stay within the error bounds. Also, case 2 model predicts both MAP and HP with a lesser MSE as outlined in figure 4.10. A clear distinction between the two cases emanates even at the training stage.

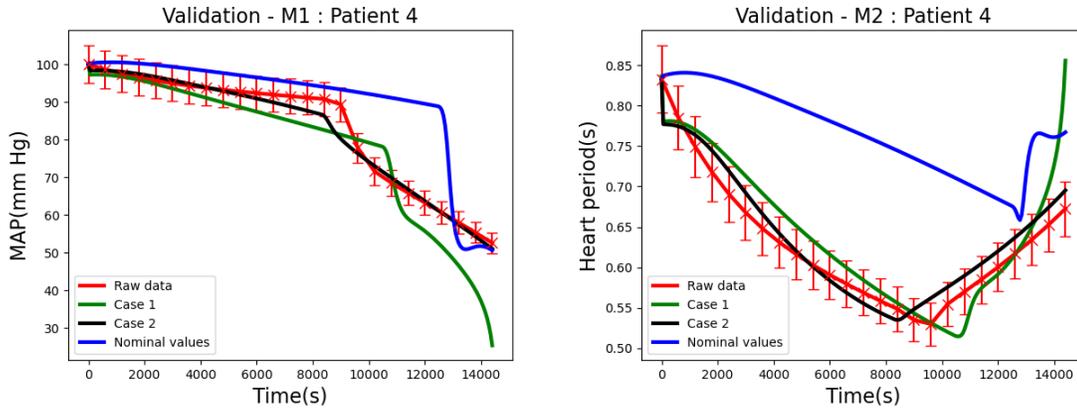


Figure 4.11: Patient 4: Validation (Case 1: $MSE_{M1} = 52.044$, $MSE_{M2} = 1.22E-03$, Case 2: $MSE_{M1} = 6.102$, $MSE_{M2} = 3.66E-04$)

Now, the estimated values of variables from both the cases are tested by supplying a random treatment profile which is meaningful and could be used in a actual HD treatment setting. Constant UFR (0.2083 ml/s, which ensures 3 litres fluid removal) and DSC (142 mEq/l, which ensures the exit concentration of sodium in the blood plasma to be within physiologically acceptable ranges) are given as validation profiles. From figure 4.11, the reader can observe that the nominal values of variables from literature predicts the occurrence of severe IDH with a considerable time lag and cannot capture the actual cardiovascular behaviour of Patient 4 under HD. Case 1

model predicts the severe IDH late as well and over predicts HP. The main motto of using a model to predict hemodynamic behaviour is to capture the occurrence of clinical complications well ahead of time so that clinical practitioners (or the optimal model-based controller in chapter 4) can take necessary precautions. Case 2 model outperforms other candidates, under predicting only one data point in terms of MAP, and the predictions closely resemble the actual patient’s behaviour. Furthermore, the increase (from the training scenario) in MSE of prediction is lesser in case 2 than case 1 and this signifies that the case 2 model is fairly robust and gives reasonable predictions even for HD treatment settings which are completely different than the input profiles in the training data set. Case 2 model shines out as a clear better performer and in the absence of noise our proposed method improves the identified grey box model prediction results.

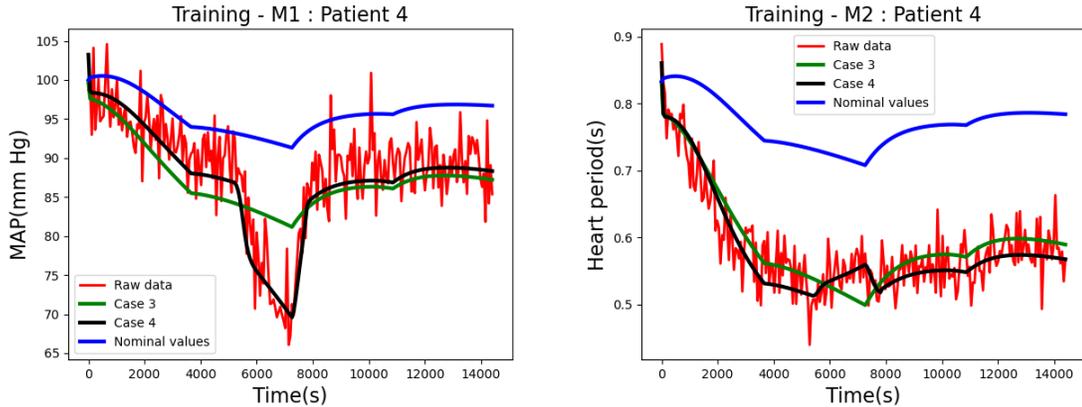


Figure 4.12: Patient 4: Training (Case 3: $MSE_{M1} = 27.642$, $MSE_{M2} = 1.39E-03$, Case 4: $MSE_{M1} = 13.349$, $MSE_{M2} = 1.12E-03$)

The test would be incomplete, if the proposed method is not put under a scenario which resembles the actual clinical setting. Clinical measurements are noisy and so random noise sequences were added to corrupt both the measurements. The same routines were repeated, but the cases were renumbered to differentiate them from the noise free counterparts. Although the optimization solvers started from the same point (initial cost = 28.65), case 4 with our proposed method (final cost = 1.749) slightly outperforms case 3 with no observability analysis information (final cost = 3.074) by finding a better optimal solution of all the variables (characterized by lower MSE of prediction for both measurements in case 4). From figure 4.12, the reader can clearly see that case 4 captures most of the hemodynamics of Patient 4, including the

sudden drop in MAP in the middle of the second hour of treatment. Case 3 model and the model with nominal values of variables do not seem to replicate the actual patient’s hemodynamic behaviour well. Then the identified models from cases 3 and 4 were put to validation tests. The reader can observe a similar kind of performance (as shown in figure 4.13) as seen in the absence of noise, with the case 4 model predicting severe IDH fairly close to the true model with a lesser MSE among the two cases, while the case 3 model predicts the pressure drop with a substantial lag and the model with nominal values completely fails to capture the pressure drop towards the end of the treatment. The simulation results support the fact that our proposed method succeeds in identifying the individualized virtual patient simulator model of a class 3 patient, while offering robustness even in the presence of measurement noises and changing treatment conditions. The reader is prompted to understand that the fixing of unobservable variables at nominal values could restrict the other observable variables from reaching their true values, but we are more focussed on identifying the optimal values of variables which closely capture the true dynamics.

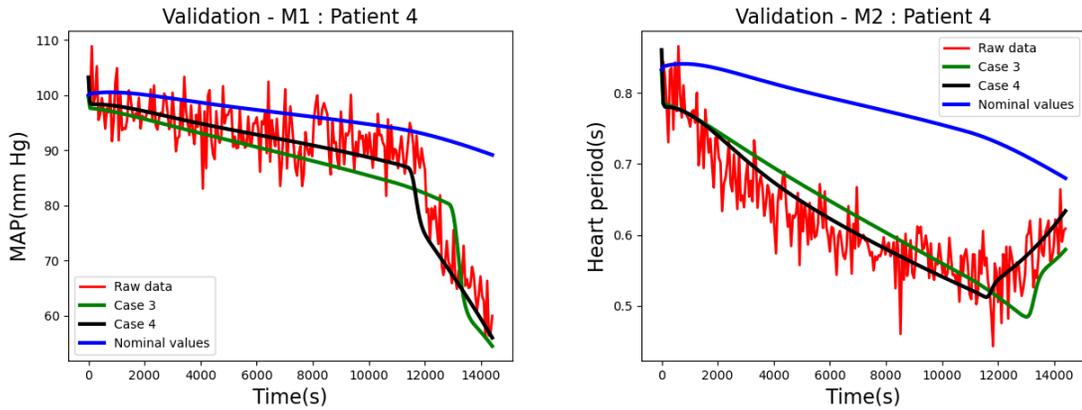


Figure 4.13: Patient 4: Validation (Case 3: $MSE_{M1} = 32.217$, $MSE_{M2} = 2.233E-03$, Case 4: $MSE_{M1} = 20.644$, $MSE_{M2} = 1.34E-03$)

4.4.2 Class 2 patient: Mild IDH & acceptable hemodynamic stability

In this section, the proposed method shall be tested on a class 2 patient from the table 4.4 data set. The representative sample for class 2 patients with acceptable levels of hemodynamic stability is patient 1. The true values of the cardiovascular reflex variables of patient 1 were generated in such a way that Patient 1 displays signs of mild IDH, which develops after the commencement of HD and continues until the

end of the treatment. However, patient 1’s reflex was responsive enough to prevent sudden MAP drops which result in syncope. All steps of the proposed method will not be explained in much detail like it was done for the class 3 patient example, but the result highlights and the inferences made by the author will be presented.

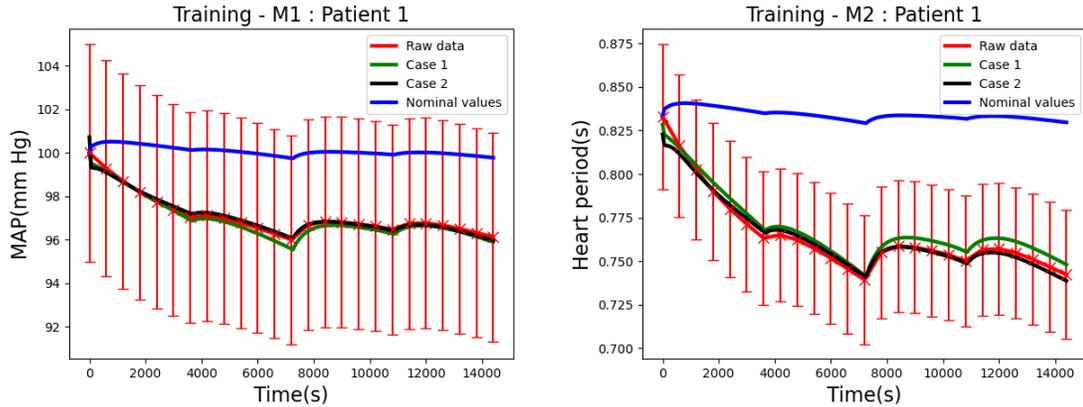


Figure 4.14: Patient 1: Training (Case 1: $MSE_{M1} = 0.035$, $MSE_{M2} = 2.55E-05$, Case 2: $MSE_{M1} = 0.016$, $MSE_{M2} = 8.16E-06$)

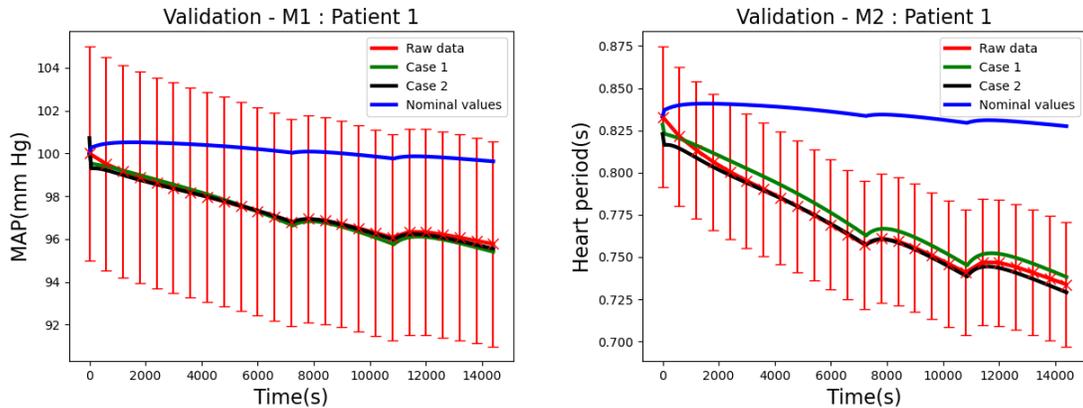


Figure 4.15: Patient 1: Validation (Case 1: $MSE_{M1} = 0.033$, $MSE_{M2} = 3.48E-05$, Case 2: $MSE_{M1} = 0.019$, $MSE_{M2} = 1.02E-05$)

Firstly, let us look at cases 1 and 2 with noise free measurements. Looking at figure 4.14, the reader will immediately notice that the model simulated with the estimated values of variables from both the cases do a fairly good job in capturing the hemodynamics of the actual patient. But the model simulated with the nominal values of the variables of a class 2 patient from literature, clearly has an offset (within

error bounds for MAP and out of error bounds for HP). Although the initial guess given to both cases was the same (initial cost = 142.59), the solver has found a better optimal solution in case 2 (final cost = 0.235) than in case 1 (final cost = 0.604). In terms of MSE of prediction, case 2 model offers superior predictions of both MAP and HP. During model validation, both case 1 and case 2 models succeed in predicting the hemodynamics of patient 1 even when introduced to arbitrary treatment inputs.

There was an interesting revelation when the same routines were repeated with noisy measurements. The cost was 100.23, but it fell down to 1.816 for case 3 and 2.119 for case 4 after SCBD optimization. One cannot vaguely conclude that the proposed approach did not work better here. Instead, if we analyze the MSE prediction errors during training and validation, we see that case 3 has lesser MSE than case 4. However, in validation, the MSE of case 3 model predictions gets higher than the prediction errors of the case 4 model (for both MAP and HP). This clearly indicates that there are signs of overfitting in case 3. Even though a good solution was found in case 3, the identified model failed to maintain consistency during validation. As case 3 involves finding the optimal solution of all 60 variables, the solver has more degrees of freedom to drive the system to a solution which unknowingly extracts some of the residual variation or noises in the data as if that variation represented the underlying true model structure during the training stage. In a simpler sense, the case 3 model has more variables than can be justified by the available data, and it has failed to predict future observations more reliably than the case 4 model, as indicated by higher MSE than case 4, during the validation stage.

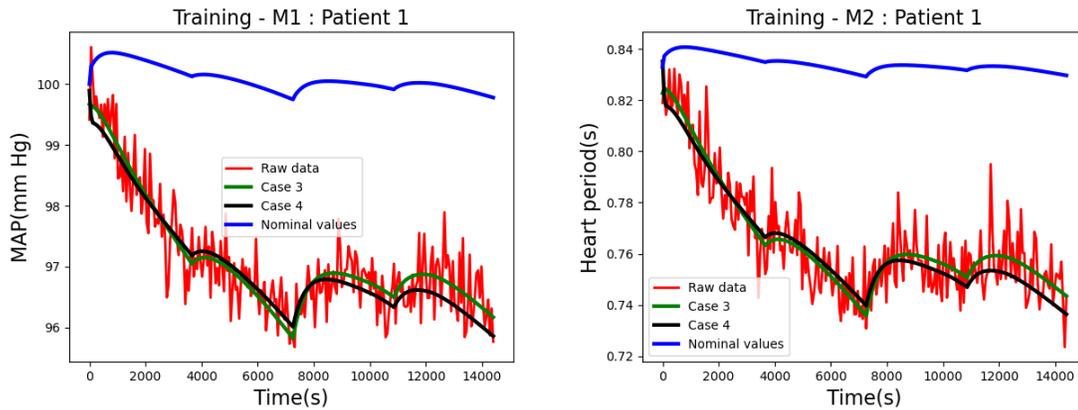


Figure 4.16: Patient 1: Training (Case 3: $MSE_{M1} = 0.16$, $MSE_{M2} = 1.00E-04$, Case 4: $MSE_{M1} = 0.191$, $MSE_{M2} = 1.15E-04$)

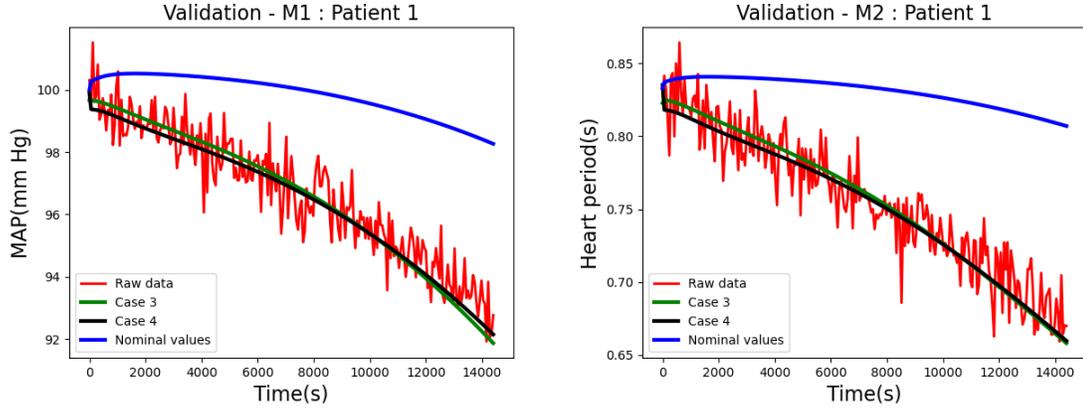


Figure 4.17: Patient 1: Validation (Case 3: $MSE_{M1} = 0.569$, $MSE_{M2} = 2.81E-04$, Case 4: $MSE_{M1} = 0.522$, $MSE_{M2} = 2.75E-04$)

4.4.3 Class 1 patient: No IDH & good hemodynamic stability

Finally, to be convinced with the inferences made from the previous examples, the proposed approach is also tested on a class 1 patient. From the random generation of true values of variables, patient 7 fell into the category of an individual who could handle different treatment settings of HD with good cardiovascular stability. The cardiovascular reflex mechanisms of patient 7 are able to maintain the hemodynamic variables without any considerable changes. Most of the patients on HD treatment regime do not exhibit this kind of ideal hemodynamic pattern.

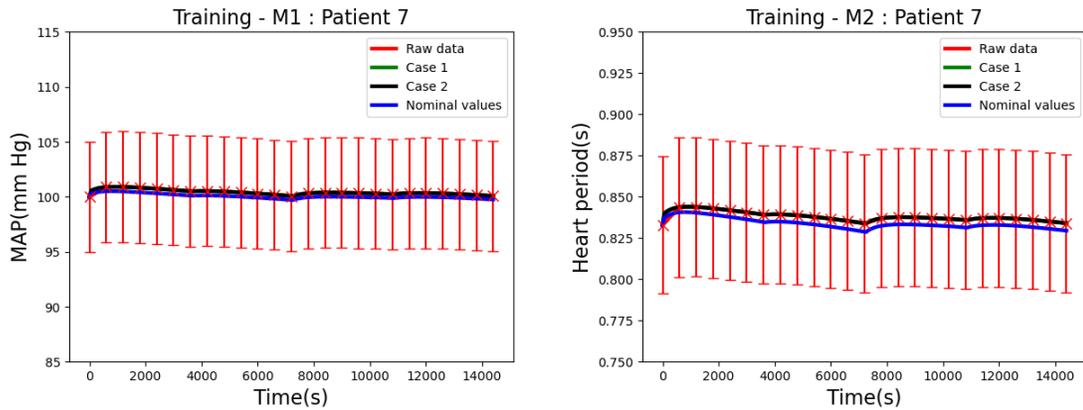


Figure 4.18: Patient 7: Training (Case 1: $MSE_{M1} = 8.56E-05$, $MSE_{M2} = 9.62E-09$, Case 2: $MSE_{M1} = 8.80E-06$, $MSE_{M2} = 1.24E-09$)

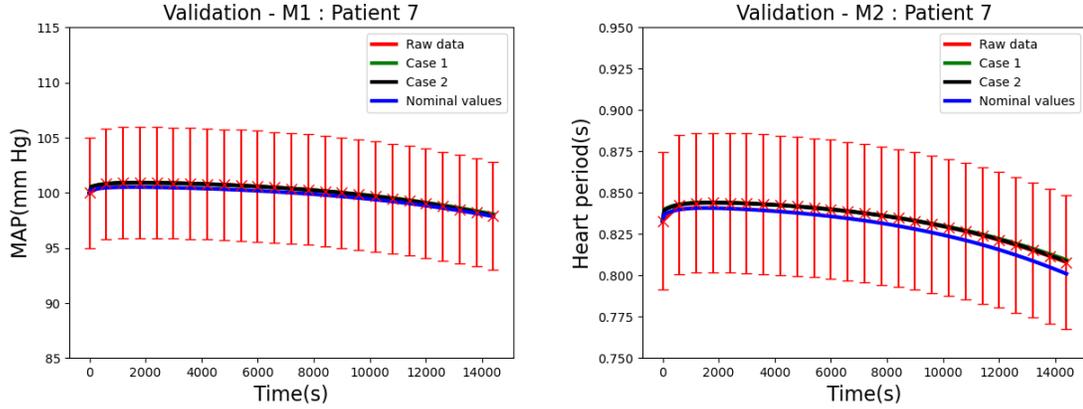


Figure 4.19: Patient 7: Validation (Case 1: $MSE_{M1} = 0.003$, $MSE_{M2} = 3.05E-07$, Case 2: $MSE_{M1} = 3.12E-04$, $MSE_{M2} = 4.13E-08$)

The trajectories obtained from the model with nominal values of variables and the estimated values of variables from both cases 1 and 2 are all within error bounds and explain the hemodynamics of patient 7 well both during training and validation. The results indicate that even the nominal values of variables can be used to simulate the hemodynamic behaviour of a class 1 patient. However, in the absence of noise, the final solution obtained from case 2 (final cost = 0.002) is better than case 1 (final cost = 0.021) and the nominal values (initial cost = 38.45). Moreover, if the reader looks at the training (figure 4.20 and validation (fig 4.21) test results when the measurements are noise corrupted, the same observations hold true.

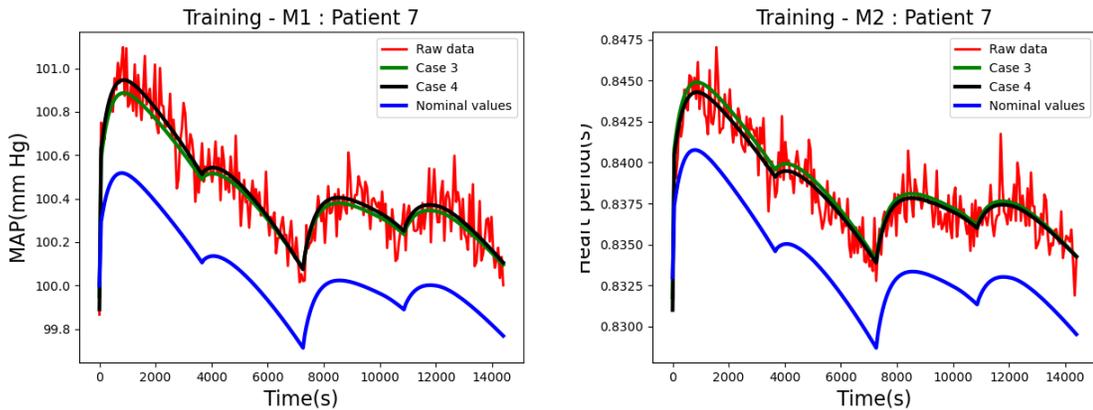


Figure 4.20: Patient 7: Training (Case 3: $MSE_{M1} = 0.009$, $MSE_{M2} = 1.51E-06$, Case 4: $MSE_{M1} = 0.008$, $MSE_{M2} = 1.40E-06$)

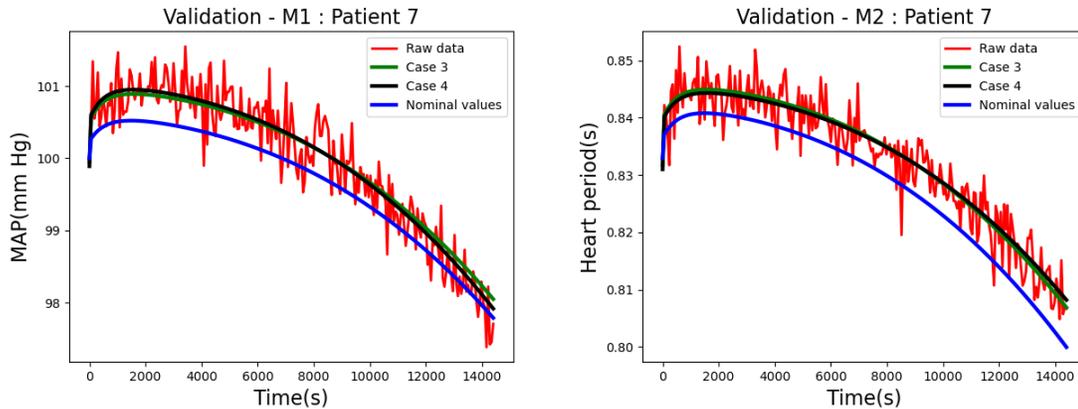


Figure 4.21: Patient 7: Validation (Case 3: $MSE_{M1} = 0.103$, $MSE_{M2} = 1.30E-05$, Case 4: $MSE_{M1} = 0.094$, $MSE_{M2} = 1.30E-05$)

4.5 Summary

In this chapter, the reader was initially educated about the initial attempts taken by the author in tackling the individualized model identification problem. Then the proposed method for nonlinear system observability was explained and tested on synthetic clinical data of different classes of patients. Our understanding after extensive simulations is that observability indeed has a crucial role to play in identifying the best model candidate. Furthermore, observability dictates the maximum extent of information that can be extracted from the available data. A brief summary of the simulation results is shown in table 4.8 and the claim is that the individualized virtual patient simulator built through our proposed method works fairly robust under different HD treatment settings and in the presence of measurement uncertainties for any patient class and could be used as an embedded model in an optimal control framework.

Table 4.8: Summary of chapter 4 analysis

Patient class	Nominal values		Cases without proposed method		Cases with proposed method	
	No Noise	With noise	No Noise (Case 1)	With noise (Case 3)	No Noise (Case 2)	With noise (Case 4)
Class 1	Works but not best	Works but not best	Works but not best	Works but not best	Works and best	Works and best
Class 2	Fails	Fails	Works but not best	Works but not best	Works and best	Works and best
Class 3	Fails	Fails	Fails	Fails	Works and best	Works and best

Chapter 5

Towards optimal control of the hemodialysis system

Ever since civilization started, human beings have been trying to establish power and take control of the things happening in their surroundings. In engineering, although control is a term used to denote feedback control (which involves the reception and transmission of process information), today's usage of the term contains a much broader perspective. For example, it can refer to the control of man-made machines, artificial devices, regulation of activities in a social sphere, such as cost and business process optimizations, the control of a country's economy by means of legislative policies. The original authors of the adopted HD system model envisioned the model being used for the identification of patient-specific parameters and optimal control of ultrafiltration rate and sodium profile in the dialysate. The main issue in delivering an optimal HD treatment for CKD patients is the efficient management of overloaded fluid volume. Removing too much of fluid can result in an acute depletion of intravascular volume or blood plasma volume causing Intradialytic Hypotension (IDH), while removing too little can result in chronic volume expansion, hypertension, left ventricular hypertrophy and increased cardiovascular morbidity. In addition, stability of the patient is the most important factor above all. In the earlier days, HD practice was based on quantifying the fluid amount to be removed and prescribing a fluid removal rate, choice of dialyzer, blood and dialysate flow rates, dialysate sodium concentration and the temperature of the dialysate. The above parameters are not modified until a HD complication occurs. If there are hypotensive episodes, the clinicians would react by manually adjusting the prescription and by post-event medical care. Although there have been several instances of application of feedback control

in HD, there are only a few research works based on a comprehensive dynamic model of all the components (patient and hemodialyzer) of the HD system based on first principles. Therefore, in cases where dynamic models are not employed, the clinicians do not get insights into the hemodynamic perturbations caused by HD. The author is thus motivated in this chapter to develop strategies to identify individualised optimal treatment plans, taking the patient's anthropometric physiological data into consideration, with the help of advanced process control techniques.

5.1 Application of traditional bio-feedback control techniques in HD

The bio-feedback concept is analogous to a feedback control system in systems engineering except that the controlled and measured variables are physiological variables of an actual patient. We can just call the bio-feedback system as a closed loop control system of biological variables. The closed-loop system has controllers, sensors and a range of actuators through which they can directly or indirectly act on the physiological variables. There are different commercial feedback control products, trying to address the feedback control of blood volume, pressure and temperature, which one can buy and utilize today. There has been a gradual increase in the global average age of HD patients, alarming conditions of diabetes and other pathological conditions, and this decreased the patient's tolerance to HD. Also on the other hand, there has been a huge ask in the reduction of the HD treatment time. When treatments are shortened, the traditional management of HD treatments no longer holds good as there are a lot of complex physiological factors that could cause cardiovascular instability in patients and quite usually uremic patients are prone to cardiovascular stress. Advancement of continuous (invasive and noninvasive) monitoring techniques in HD, for measuring blood volume, heart rate, blood pressure and solute concentrations, would be the initial step towards building better feedback control systems. Development of continuous monitoring techniques is beyond the scope of this research work, but the readers are motivated to look at a few research works [11], [60], to get an understanding of the progress made in that genre.

The second step would be the testing and practical implementation of the built feedback control system. HD is undoubtedly a perfect example representing both a servo-control problem where the treatment objectives can be considered as set points specified to the controller and a regulatory control problem with an ambition to keep the patient in equilibrium and reduce any treatment abnormalities. Some of the set

points could be static but some are dynamic. In current clinical practice, there are three commercial routine feedback technologies from manufacturers for application in HD. They include blood volume feedback, blood temperature feedback and blood pressure feedback systems. However, the development of a comprehensive feedback controller is not very because the designer needs to understand the significance of all physiological variables to be controlled with a deep comprehension of the HD process dynamics. Let us understand the working of a blood volume tracking system with the help of a few test simulations in the next section and later dive into the optimal control framework for HD.

5.1.1 Simulation results and discussion

The blood volume dynamics have been investigated by a lot of researchers extensively and several factors have been found to influence the blood volume changes during the HD sessions [128]. Out of the several factors identified, ultrafiltration rate and dialysate sodium concentration emerge as the most important influential factors affecting the blood volume changes. The cardiovascular stability of the patient and the occurrence of hypovolemia are primarily based on these two dialysis variables. Furthermore, profiling of ultrafiltration has been found to significantly reduce HD induced complications by ensuring a stable blood behaviour, although there could be a few possible exceptions. Then it was discovered that different patients have different plasma refilling capacities and a feedback control module based on ultrafiltration alone could not adapt well to the tune of plasma water removal as ultrafiltration is of several orders higher in magnitude than the plasma refilling rate.

The body weight targets at the end of dialysis cannot be achieved just by specifying the ultrafiltration volume. There was a need to identify a variable that would be able to mobilize fluid across different body compartments, thereby compensating for the rapid reduction in the circulating plasma volume by ultrafiltration. Dialysate sodium concentration (DSC) was found to be the perfect fit for this job based on extensive clinical studies. The investigators realized that by elevating the extracellular sodium concentration during dialysis, more fluid can be internally mobilized from the intracellular compartment to the extracellular compartment (flow of water due to osmosis). This helps in achieving the desired body weight targets by the end of the treatment. DSC not only increases plasma refilling but also triggers the activity of the autonomic nervous system, inducing a slightly better hemodynamic reflex response behaviour in patients. The Blood Volume Tracking (BVT) control system was born out of collaborative research efforts from the Gambro-Hospital research groups [118].

It was indeed a great improvement from the traditional hemodialysis treatment delivery. The blood volume control system was capable of manipulating UFR and DSC based on the supplied blood volume trends (volemia).

The core of the BVT system has Proportional Integral Derivative (PID) controllers and so two parallel PID controllers were designed to duplicate the performance of an actual BVT system. This study would reveal to us if we could really benefit from implementing feedback control. There are complex interactions between the controlled and manipulated variables, but they are assumed to be independent of each other and so parallel PIDs are used to control each state variable. The parameters of the two PIDs are given in the table 5.1. A patient of 70 kg and a treatment time of 240 minutes are considered for this trail and the ideal profiles (dotted red line) for the extracellular volume - UFR (U_1) pair and the extracellular sodium concentration - DSC (U_2) were designed based on clinical heuristics. A step-wise profile, which closely results in an exponential reduction in blood volume, was used for ultrafiltration rate for the removal of 3 litres of overloaded fluid (corresponding to a 3 kg body weight reduction) from the patient. The dialysate sodium concentration profile was designed in such a way that the peak is reached within the first half of the treatment and the second half is just for regulating the extracellular sodium concentration to be within safe limits. The initial value of the concentration of all solutes was set in a way such that the osmotic potential is zero at the beginning of the treatment. The performance of the BVT, in the presence of initial state estimation errors, was showcased with the help of two simulation settings.

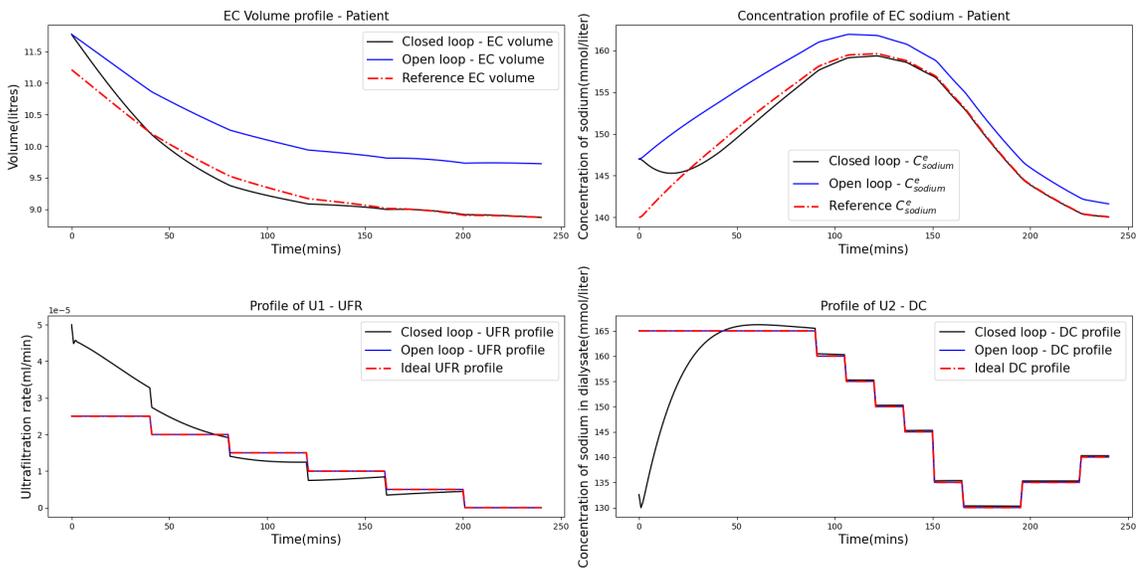


Figure 5.1: PID case 1: When initial states are 5% higher

Table 5.1: PID tuning parameters

Parameters	PID 1 (UFR control)	PID 2 (DSC control)
States feedback	Plasma Volume or Blood Volume	Extracellular sodium concentration
Gain	-0.045	5
Integration time (<i>min</i>)	40	60
Derivative time (<i>min</i>)	5	10

The nominal values of all the variables of the HD system model were taken from literature [140]. It was ensured that the PID controllers would compute realizable values for both the manipulated variables (from Eq. (2.23)) by supplying suitable saturation limits. The extracellular volume which is the sum of the interstitial fluid volume and the blood plasma is shown in the plots as we have considered only two compartments with respect to the concentration of solutes. If we look at figures 5.1 and 5.2, we can observe that the final targets (extracellular volume and extracellular sodium concentration) could be achieved, although there are errors in the initial state values only when a closed loop system (BVT) is employed. The ideal treatment trajectories, if followed in an open loop scheme, fail to reach the final targets and ends with an offset resulting in over (initial states are lower) or under (initial states are higher) treatments.

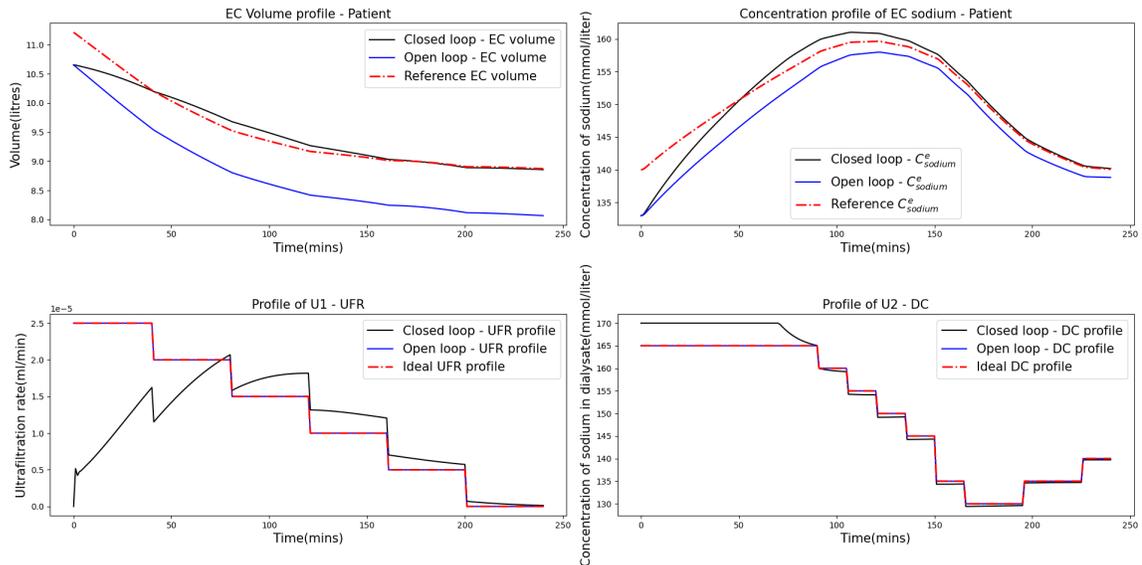


Figure 5.2: PID case 2: When initial states are 5% lower

The results clearly show that feedback control is quintessential in reaching treatment targets in the presence of errors (or disturbances). The PID controllers have com-

puted the trajectories (away from ideal profiles) that would be necessary to reach the final targets, based on the feedback information (or error between the desired and the actual values). For the purpose of feedback, the controlled variables should be continuously measured with the help of sensors. The main advantage of a BVT system is that the user does not need a comprehensive dynamic model of the system in this kind of feedback implementation. A simple understanding of the relationship between the input-output pair is sufficient [11]. Following exponential volemia trajectories have been tested in actual clinical practice and it was found that it has reduced the occurrence of HD induced complications. Then a question arises if this BVT system will always ensure cardiovascular stability and safety of the patient. To study this, we experimented the ideal treatment profiles on different classes of HD patients (as categorized in chapter 3). The same simulation settings from the previous study in this section have been adopted here and it was assumed that treatment targets were met in all cases.

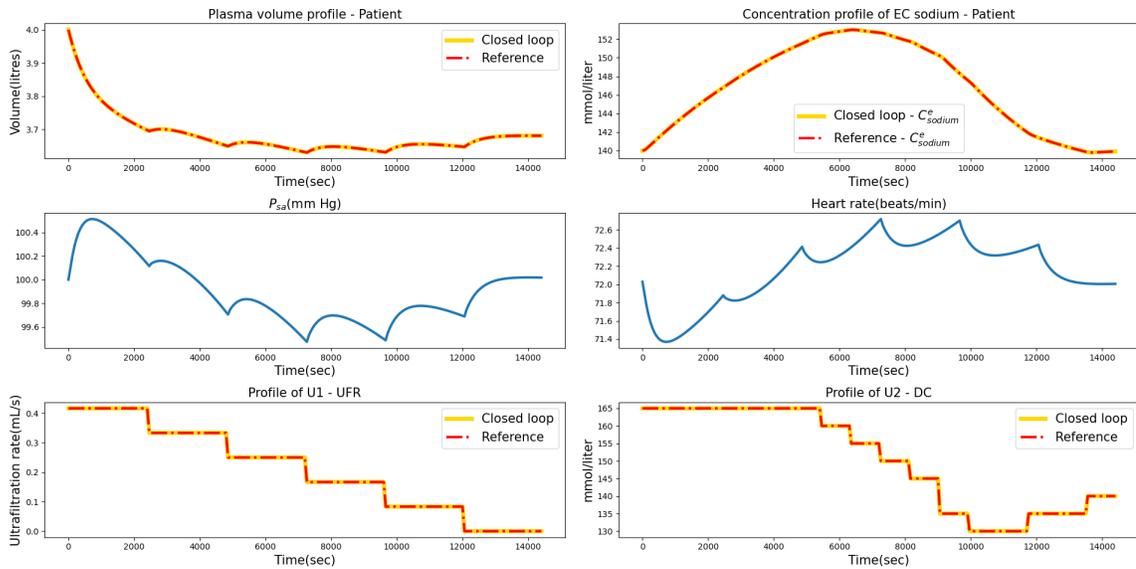


Figure 5.3: Class 1 patient under treatment in a BVT system

Figure 5.3 depicts that a class 1 patient can handle this profiled HD treatment without any substantial drop or increase in blood pressure and heart rate. Therefore, the BVT system could be sufficient for those patients with an excellent cardiovascular reflex mechanism. However, a class 2 patient exhibits mild signs of IDH in the middle of the treatment as shown in figure 5.4. The patient’s reflex system works acceptable enough to bring the patient back to hemodynamic stability towards the last hour of HD, when both the manipulated inputs are maintained closer to their lower levels. We can

observe from figure 5.5 that even when the BVT system reaches all treatment targets following an ideal reference trajectory based on clinical heuristics, there would be a class of ESRD patients who experience sudden and severe IDH during the treatment. The patient's reflex could not compensate with the rapid reduction in plasma volume and so the patient experienced syncope and took almost 25 minutes to return to normalcy. If such complications occur, the treatments have to be manually stopped and carried out in a semi-automated fashion.

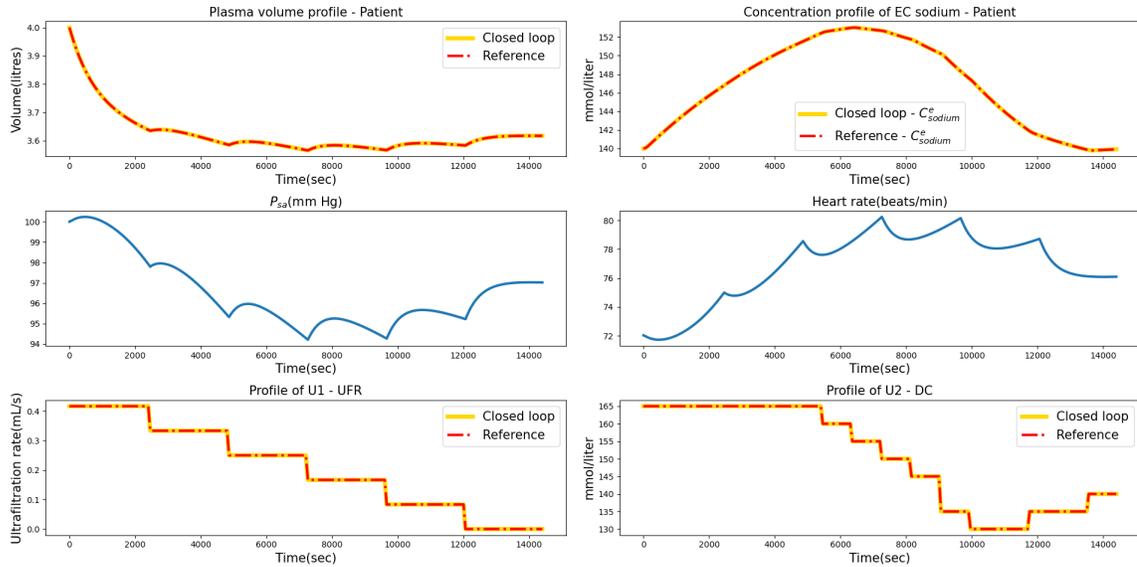


Figure 5.4: Class 2 patient under treatment in a BVT system

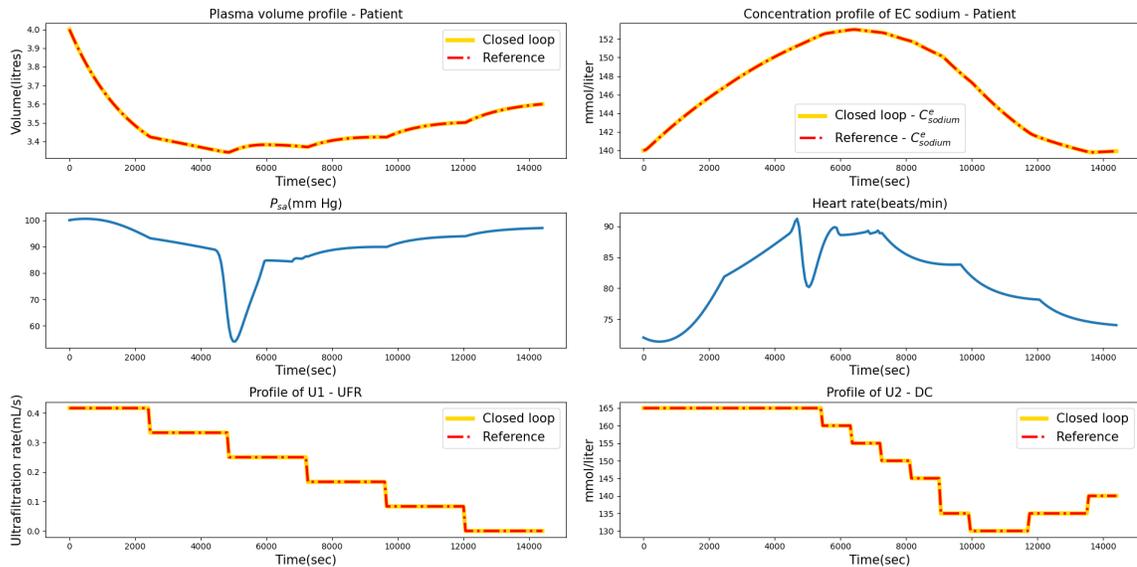


Figure 5.5: Class 3 patient under treatment in a BVT system

We understand from our simulations that a BVT system is in general much robust to external noises and to internal system abnormalities, because the deviations from the desired values are first measured and then compensated for by the action of the controller on the manipulated inputs. However, the BVT system does not guarantee safety of the patient always and there is a need for a more advanced comprehensive control strategy which would work on all different classes of patients.

5.2 Layout of the feedback control system for application in HD

Modern dialysis has undergone major developments since the earlier days when feedback control in HD just meant control of flow and pressure. In the view of clinicians, the arrival of the bicarbonate dialysate and volumetric control of ultrafiltration seemed like major advancements in their potential to deliver more reliable treatments to the patients who come to the HD treatment centre. However, Intradialytic Hypotension (IDH) still occurs in 30% of the HD patient population leading to inadequate treatments [159]. This calls the need for a a new strategy which will involve the use of more online information regarding the condition of the patient during the treatment. The clinical information that can be used for enhanced treatment includes blood volume, hemodynamic variables, therapy adequacy and blood access monitoring. The researchers in HD have reached a consensus that the treatments have to be tailored for every individual according to their personal clinical needs set by their doctors.

Semi-automated treatments by manual adjustment of the dialysis machine variables like Ultrafiltration Rate (UFR) and Dialysate Sodium Concentration (DSC) are becoming increasingly difficult because of the complex dynamic interactions between different hemodynamic variables and the perturbations induced by HD. This prompts a necessity for the dialysis machine variables to be adjusted continuously rather than maintaining constant, predefined settings by the clinical practitioner. On those lines, this chapter is an attempt by the authors to develop a comprehensive optimal control framework, with the objective of providing completely automated treatments for actual clinical realization in the near future, and in this section, the overall layout of our proposed optimal control framework, as shown in figure 5.6, for potential practical implementation in an actual HD treatment setting will be explained. The most important components namely, the observer (Extended Kalman Filter (EKF)) and the nonlinear controller (Shrinking Horizon Batch Zone Nonlinear Model Predictive

Controller (SHBZNMPC)), will be elucidated in much detail in separate sections later.

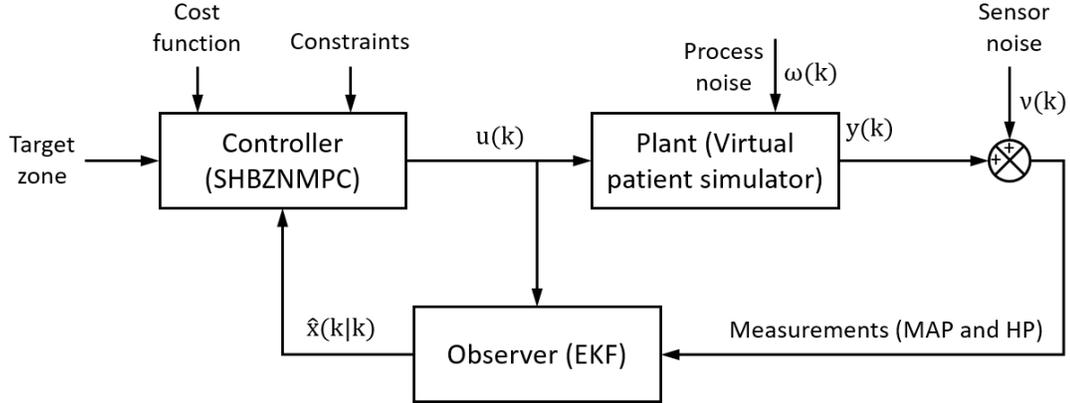


Figure 5.6: Block diagram of the proposed feedback control system with process and sensor noise

The feedback control system, shown above in figure 5.6, is a gross layout and does not include all the micro components in the system. The HD control system is a network of several components and configurations that will drive the system to the desired target state while maintaining the stability of the patient. Most of the feedback control systems might have a similar layout, however, some features could be ignored or added depending on the user's requirements. The plant in our case, is the patient who is undergoing the HD treatment. The patient is the most important physical component in the HD control system. It is the patient's output (total body water, toxin concentrations and many more) which are to be guided to physiological targets. This research work is completely a simulation study and so a dynamic mathematical model of the plant (refer chapter 3 for all dynamic equations of the plant) is quintessential for understanding and designing an advanced feedback control strategy. The dialysis machine is the actuator which receives the control signal from the controller and executes it on the patient to force the patient to reach the specified targets.

There are several disturbances which are external to the HD control system and the model structure (poor patient lifestyle leading to abnormal physiological changes than usual, infusion before treatment) and they affect the process. All these disturbances have been grouped under process noise (ω) in our design. Although, the nature of noise is purely stochastic, it is considered to be of additive nature in our design. The actual response of the patient has to be measured to know the impact of the controller's actions and also for comparing the actual states with the desired state.

The outputs of the plant are measured with the help of sensors. A sensor is a device which is used for measuring the actual system output and the measured states are not always accurate and identical with the true states. To replicate the inherent uncertainty in the measurements, sensor noise (ν) of additive nature is considered. The process ($N(0, 0.005^2)$) and sensor ($N(0, 0.025^2)$) noise are white noise signals and they are drawn from normal distributions of zero mean and the variances mentioned. To compute the optimal treatment profiles (u), the dynamic model in the controller, which has several states to predict the physiological condition of the plant, has to be run many times. To do so, the controller needs the initial states of the plant at every sampling instant. In real practice, it is quite not practical or it is impossible to measure all the actual dynamic states of the plant and so a state estimation algorithm which will work with the available noninvasive clinical measurements (MAP and HP) to reconstruct the dynamic state vector (\hat{x}) is used. The state estimation algorithm used in our design is an Extended Kalman Filter.

The controller is the heart of the HD control system and it takes in all information from the state estimator and the clinical practitioners (targets and constraints) and computes the optimal values of all manipulated variables based on the embedded replica dynamic model of the plant. The controller used here is a Shrinking Horizon Batch Zone Nonlinear Model Predictive Controller (SHBZNMPC) and it uses an optimization-based approach (minimizing an objective function) to arrive at the optimal input sequences for all manipulated variables. The computed optimal inputs are fed to the dialysis machine which acts on the patient in turn. This loop is continued until the end of the treatment. In the next few sections, the design of the state estimator and the controller for this particular application will be elaborated in much detail with test simulation results.

5.3 Design of a nonlinear state estimator for application in HD: Extended Kalman Filter (EKF)

In control theory, the Extended Kalman Filter (EKF) is the nonlinear version of the well-known Kalman Filter. It works by successively linearizing the system around the current estimate of the states and its covariance. The mathematical equations related to the design of the EKF were outlined in chapter 2 and the reader could look back at the list of equations whenever necessary. Also on the other hand, the EKF would be sufficient to verify the results of the PBH observability test mentioned in the first half of chapter 4. In chapter 4, we discovered that any 2 measurements

will make the HD system model observable. The 18 states of the original system model include the volumes of the body compartments, the concentration of solutes in each of the body compartments, the pressures of the vascular compartments and the short-term baroreflex regulation states. Although a blood sample is taken before starting the treatment to find the solute concentrations in the blood plasma, the Mean arterial Pressure (MAP), from the pressures of the vascular compartments, and the Heart Period (HP) or Heart Rate (HR), from the internal reflex mechanism states, are the two easily available noninvasive clinical measurements during HD. Because these measurements are available in discrete time intervals, a discrete time EKF would be a good choice for this purpose. The design settings of the discrete time EKF are described below:

- The state covariance matrix (P) matrix was initialized as explained in Eq. (2.17). In that initial guess equation, x_0 is the vector of actual initial state values and \hat{x}_0 is the user specified vector of the guess values of states.
- x_0 is fixed at the nominal values available in literature [140]. The offset in the initial guess is assumed to be 5%.
- The parameters of the original system model are fixed at the nominal values from literature for the initial test case and then they are fixed at the true values of each patient when they are tested on the patient data set from chapter 4.
- The process covariance matrix (Q) and the measurement covariance (R) matrices are diagonal matrices of their respective covariances. The process covariance is assumed to be 0.0001 and the measurement covariance is assumed to be 0.00001 for our initial test case and later on they will be increased to 0.005 and 0.00005 respectively when the EKF design is tested on the patient data set in table 4.4.
- The simulation time is set to 4 hours.
- It is assumed that the measurements were noise free in our initial test case and then additive noise, drawn from a normal distribution $N(0, 0.0005^2)$, is added to the measurements (noise corruption = multiplying the noise values with the true values and then adding them to the true values).
- New measurements of the hemodynamic variables are available every minute and so $\Delta T = 60$ seconds.

5.3.1 Simulation results and discussion

The designed EKF was tested on multiple case settings so that it could be confidently embedded into the optimal control framework for the observation of states with an acceptable level of accuracy. All the simulations were done in Python writing the individual blocks of the EKF method in separate functions. To answer the question of system observability, the performance of the EKF is tested with the availability of both one and two clinical measurements. This shall be our first test case. Like mentioned earlier, an initial offset error of 5% was assumed in the guess values of the states. In figure 5.7, the blue line is the true state trajectory, the yellow line is the EKF estimated state trajectory and the green line is the open loop simulation with the wrong initial states. The EKF and the open loop will start at the same initial point. In our simulations, the Root Mean Square Error (RMSE) will be used to evaluate the performance of the EKF and the error trajectories of the individual states are plotted by computing the Root Square Error (RSE) at a particular time instant (t). The squared errors were scaled and then used in the equations.

$$RMSE_x(t) = \sqrt{\frac{\sum_{k=1}^{N_x} (\hat{x}_k(t) - x_k(t))^2}{N_x}} \quad (5.1)$$

$$RSE_x(t) = \sqrt{\sum_{k=1}^{N_x} (\hat{x}_k(t) - x_k(t))^2} \quad (5.2)$$

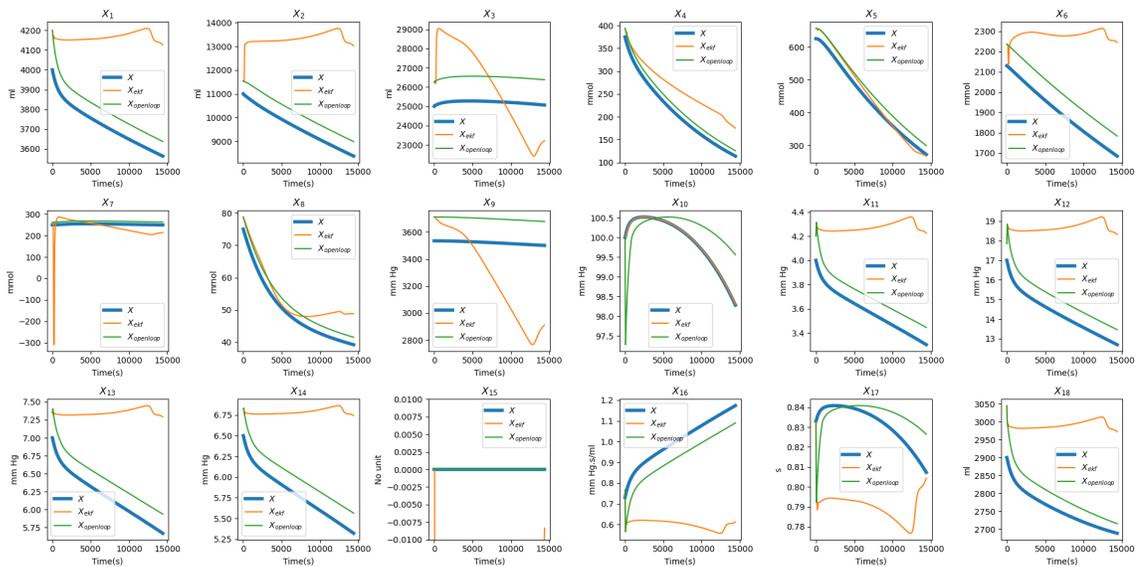


Figure 5.7: Extended Kalman filter (EKF) with 1 measurement (MAP)

The simulation time was set to 4 hours because the EKF will be run in parallel to the controller if implemented in a real clinical setting. For practical realization, the EKF should also converge within the standard treatment time of 4 hours. However, running the EKF for a longer duration would allow us to comment on the convergence results better. In our test cases, the performance of the EKFs will be evaluated within the treatment time of 4 hours. In the case where only one measurement (MAP) is available, the reader can easily observe from figure 5.7 that the EKF fails to recover the true state information. Most of the EKF estimated states trajectories are found to diverge away from the true state trajectories. Also if one looks at the RSE plots of states (figure 5.8), most of the estimated errors at the end of 4 hours are well away from zero with the exception of only a few states ($X_7, X_{10}, X_{15}, X_{17}$).

Table 5.2: Tuned diagonal elements of the initial guess of the process covariance matrix obtained from Eq. 2.17

Position index	1,1	2,2	3,3	4,4	6,6	7,7	rest
Value	50	1E8	100	300	1E7	75000	500

The diagonal entries of the initial guess of the process covariance matrix were fine-tuned to check if the true states were actually recoverable from 1 measurement, but the attempt was a failure and so we infer that the system is unobservable when we have only one measurement. Only the final tuning settings are shown, but a rigorous effort (multiple tries) was made while fixing the diagonal entries of the P matrix.

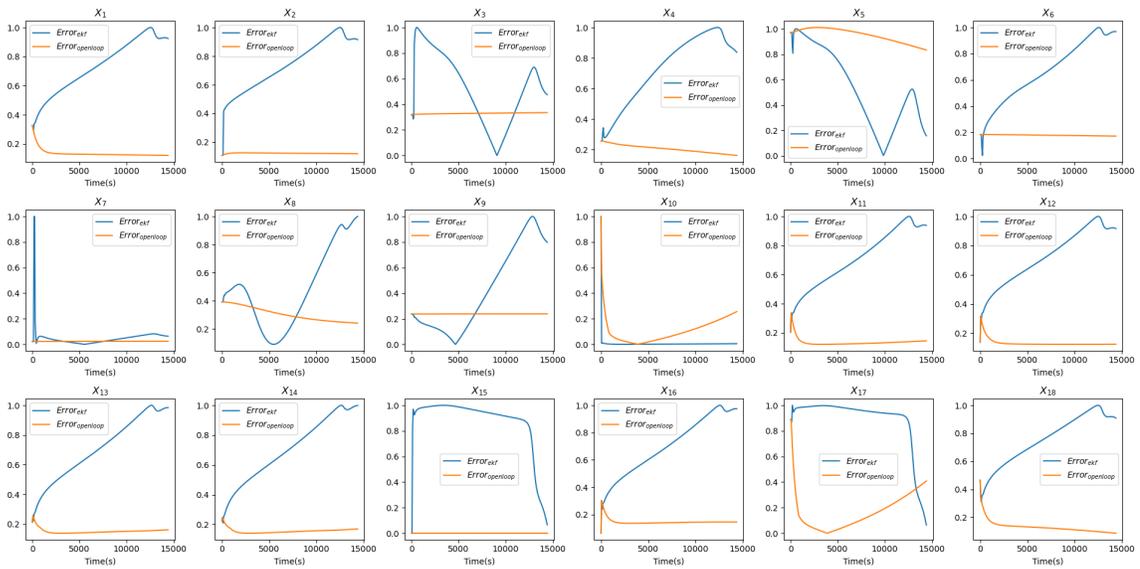


Figure 5.8: Observer RSE dynamics with 1 measurement (MAP)

Now, let us look at the case (figure 5.9) where we have two measurements (MAP and HP) in hand. The reader will easily observe from the figure that most of the states have reached the true state values within the 4 hour treatment time with an exception of only state (mass of urea in the extracellular pool). But the RSE trajectory of that particular state displays an asymptotic behaviour and so it can be considered as an observable state as well. The same tuning employed for the previous case (given in table 5.2) has been used here and we can see an aggressive EKF performance.

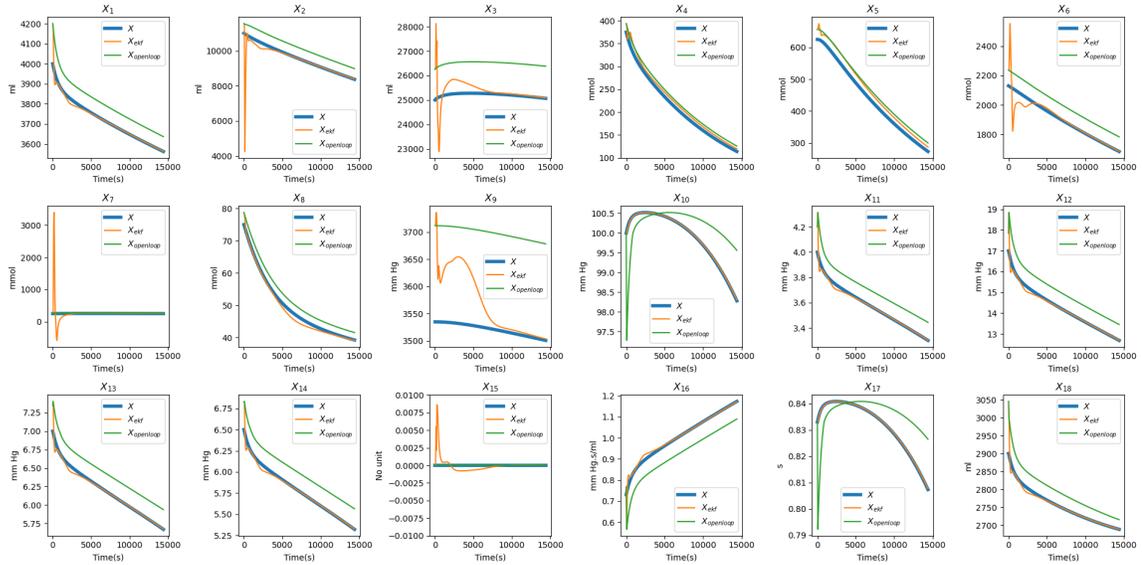


Figure 5.9: Extended Kalman filter (EKF) with 2 measurements (MAP, HP)

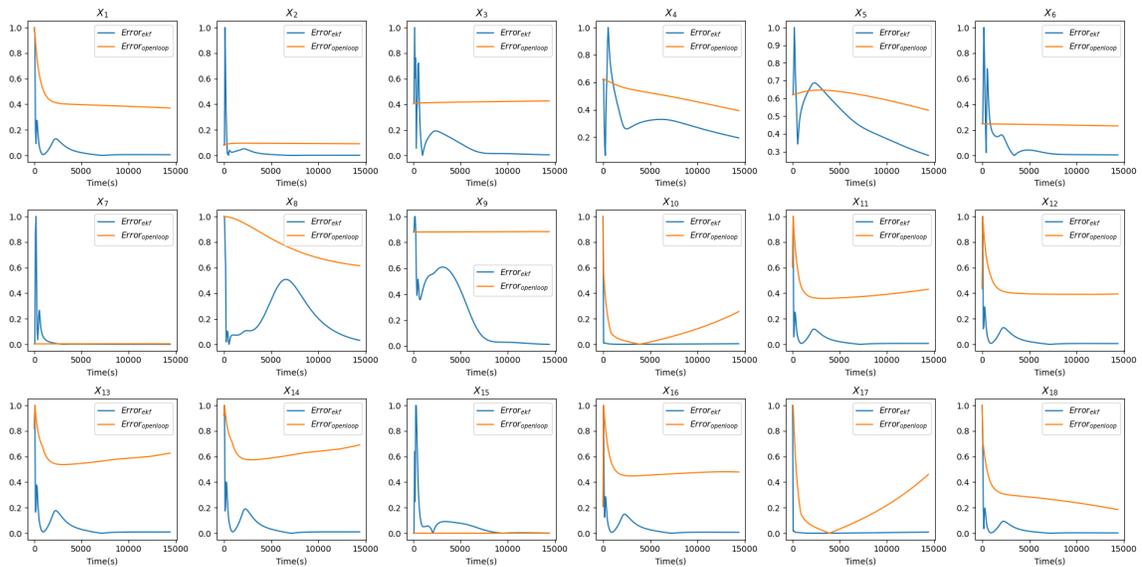


Figure 5.10: Observer RSE dynamics with 2 measurements (MAP, HP)

The aggressive EKF performance seen in figure 5.9 comes as a favourable feature for our case because in the HD control system design we expect the EKF to converge faster to ensure that the controller can implement moves based on the state feedback and drive the system to the desired target by the end of the treatment session. But all our states have a physiological meaning and so we supplied suitable saturation limits so that the results would be meaningful. For instance, the states pertaining to the mass and volume cannot go negative.

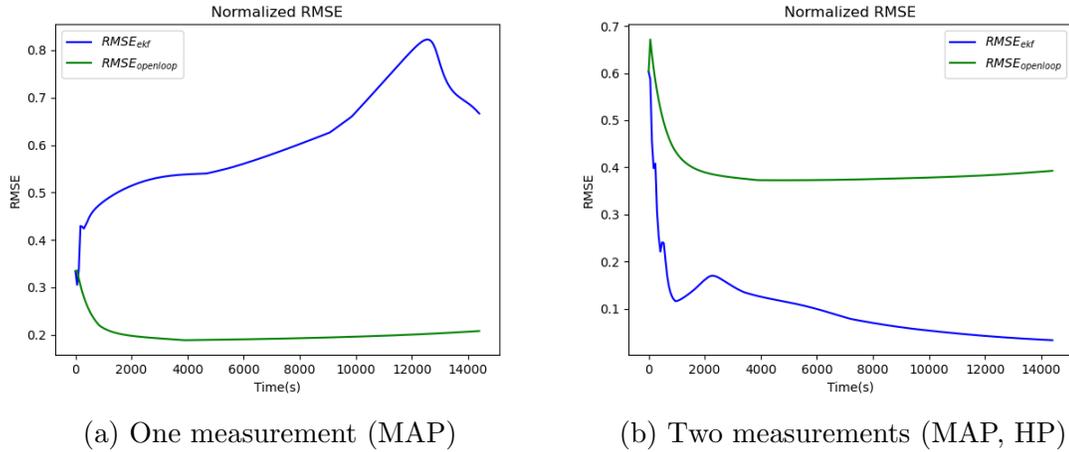


Figure 5.11: Comparison of estimation performance with the help of normalized RMSE dynamics

Finally, to get a conclusion based on our simulation results, we look at the overall normalized RMSE plots. For a state observer to be estimably stable [112], the observer error dynamics should follow an asymptotic decaying trajectory. This means that the state estimation error will decay to zero over the extended finite time horizon. In figure 5.11, we observe that in the case where we have only one measurement, the EKF is not able to achieve asymptotic error tracking while in the case where we have two measurements, the estimation error falls rapidly from the initial values and gradually keeps decreasing towards the end of the treatment. With the results from the PBH test and from the EKF estimations (1 & 2 measurements), we conclude that the system is observable with two measurements and the observer will be able to reconstruct the full state vector from these two measurements with an acceptable range of accuracy.

Then we applied the designed EKF for two measurements, to the synthetic clinical patient information from 4.4. The same EKF design settings were used for all patients. This would enable us to analyse if our original design is flexible enough to work across

different patients in the data set. The nominal values of the state variables (adopted from literature) were given as the initial guess for both the EKF and the open-loop simulations. We clearly see that the RMSE of estimation falls rapidly, than the open-loop simulations, when an EKF is employed for state estimation. For patients 2, 3, 5, 6 and 7 the EKF results are noticed to converge within the end of the treatment. However, for patients 1 and 4, the state estimation errors have still not converged by the end of the treatment. The user will be able to arrive at better designed EKFs for patients 1 and 4 if the EKF parameters are tuned. Although tuning the EKF for each patient would guarantee even better estimation results, the large number of patients on HD makes it cumbersome for the user to design the EKF individually for every patient. It would be ideal to adopt a EKF design that will work across different patients for user-friendly practical realization.

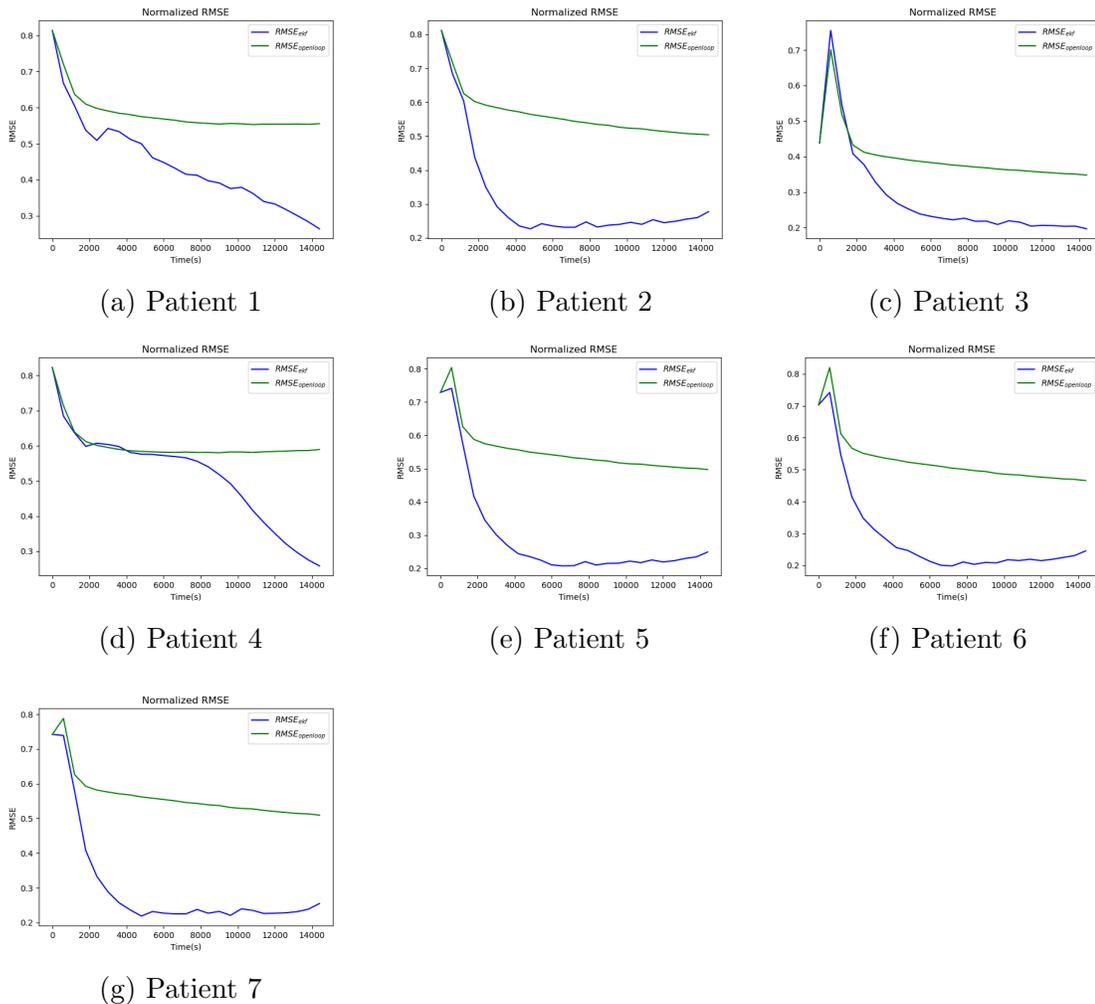


Figure 5.12: Estimation performance of EKF for different patients in table 4.4 in the presence of process and measurement noises (both $N(0, 0.0005^2)$)

5.4 Design of a nonlinear Batch Zone Nonlinear Model Predictive Controller (BZNMPC) for application in HD

Model Predictive Control (MPC) as a control strategy emerged as a commercial advanced process control package in the later half of the 20th century. The term MPC does not mean a single control methodology, but rather it could refer to any control strategy which extensively uses a model of the process to compute the optimal control inputs by minimizing an objective function. Usually, a MPC consists of two segments, the system model block and the optimizer block. In our case, the prime agenda of a MPC would be to use a dynamic model of the HD system to predict the hemodynamic variables like MAP and HP and optimize the future system dynamics. At each sampling time, the MPC will solve an open-loop control problem over a finite time interval by taking all process and physiological constraints (both current and future) into account. The optimizer block strives to select the control inputs which have the least objective function value. The objective function is designed by the user and it is completely problem-specific there is no set standard. Only the first value in the optimal control sequence is implemented while the rest is discarded and the same procedure is repeated for the following control intervals. It might look as an open-loop control strategy, but it is converted to a closed-loop strategy by featuring in the measurements as the current state values. In HD, MPC control strategy gives the user a more natural way of formulating the optimal control problem in time domain and allows the user to explicitly specify the physiological and process constraints of HD [60]. For an indepth understanding of the MPC control algorithm, the reader is motivated to look at excellent literature sources [112].

There are several drawbacks of using a traditional MPC in HD because of its aggressiveness when using the same design (tuning parameters) for several groups of patients. Moreover, when controlling around a set point, the controller will make aggressive moves around the set point because of the inherent integral action and also makes extra moves to keep the system closer to the set point. The global patient pool of HD is increasing at an alarming rate and it might not be ideal to sit and tune the controller repeatedly for every patient before the start of the treatment. Each patient will have a different hemodynamic response to the treatment and the controller should be intelligent enough to work across different patients. In addition to that, the clinical targets encountered in HD settings are not strict set points but rather are acceptable physiological ranges. When there are no strict clinical set points, Zone

MPC (ZMPC) would be the perfect fit for the problem and this technology is quickly emerging in medical field as a potential advanced control strategy of choice (diabetes, anemia management) [95]. It has also been successfully applied to high precision agricultural systems for soil moisture regulation [92]. ZMPC is realized by setting upper and lower bounds for the outputs rather than strict point values of targets. Zone control is essentially a MPC strategy where all the constraints are softened by the introduction of slack variables and then penalizing the slack variables in the objective function. The basic formulation of a ZMPC is given in chapter 2 and it can be modified to fit our problem statement as given below.

$$\min_{\hat{x}(i), \epsilon, \Delta u} \sum_{i=1}^{N_p} Q \cdot \epsilon_s^2 + \sum_{i=0}^{N_p-1} R \cdot \Delta u_i^2 + P \cdot y_f^2 + Q_f \cdot \epsilon_f^2 \quad (5.3a)$$

$$\text{s.t. } \hat{x}_{i+1} = f(\hat{x}_i, u_i) \quad i = 1, 2, \dots, N_p \quad (5.3b)$$

$$\hat{y}_i = h(\hat{x}_i, u_i) \quad i = 1, 2, \dots, N_p \quad (5.3c)$$

$$x_{min} \leq \hat{x}_i \leq x_{max} \quad i = 1, 2, \dots, N_p \quad (5.3d)$$

$$u_{min} \leq u_i \leq u_{max} \quad i = 0, 1, \dots, N_p - 1 \quad (5.3e)$$

$$\Delta u_{min} \leq \Delta u_i \leq \Delta u_{max} \quad i = 0, 1, \dots, N_p - 1 \quad (5.3f)$$

$$B^L - \epsilon_k^L \leq \hat{y}_i \leq B^U + \epsilon_k^U \quad k = s \text{ (or) } f \quad (5.3g)$$

$$\epsilon = \{\epsilon_s, \epsilon_f\} \quad (5.3h)$$

$$\epsilon_s \geq 0, \epsilon_f \geq 0 \quad (5.3i)$$

Here, ϵ_s denotes the list of slack variables for the state (or outputs) constraints and appear in the stage cost formulation and ϵ_f denotes the list of slack variables for the terminal states (or outputs) and appears in the terminal cost formulation. All the slack variables used in the problem formulation are non negative. Q and Q_f are the penalty matrices for the slack variables in stage cost and terminal cost formulation respectively. P matrix penalizes the terminal states (or outputs) and R is the penalty matrix for the rate of change of input. The diagonal entries of all the penalty matrices are tuned according to the user's requirements.

Let us recall the main objectives of a HD treatment, based on our understanding of the problem statement from chapter 1. We can deduce the following objectives from the basic functions of the kidneys and our clinical acumen of HD treatments.

- Weight (or) fluid management: Removal of the overloaded fluid in the patient's body.

- Toxin management (balance): Removal of toxic substances from the patient's body.
- Electrolyte balance: Regulation of blood sodium and potassium levels within physiological limits.
- Hemodynamic stability: Regulation of blood pressure within tight bounds.
- Optimal minimum time: The minimum time taken to reach all the objectives of the treatment.

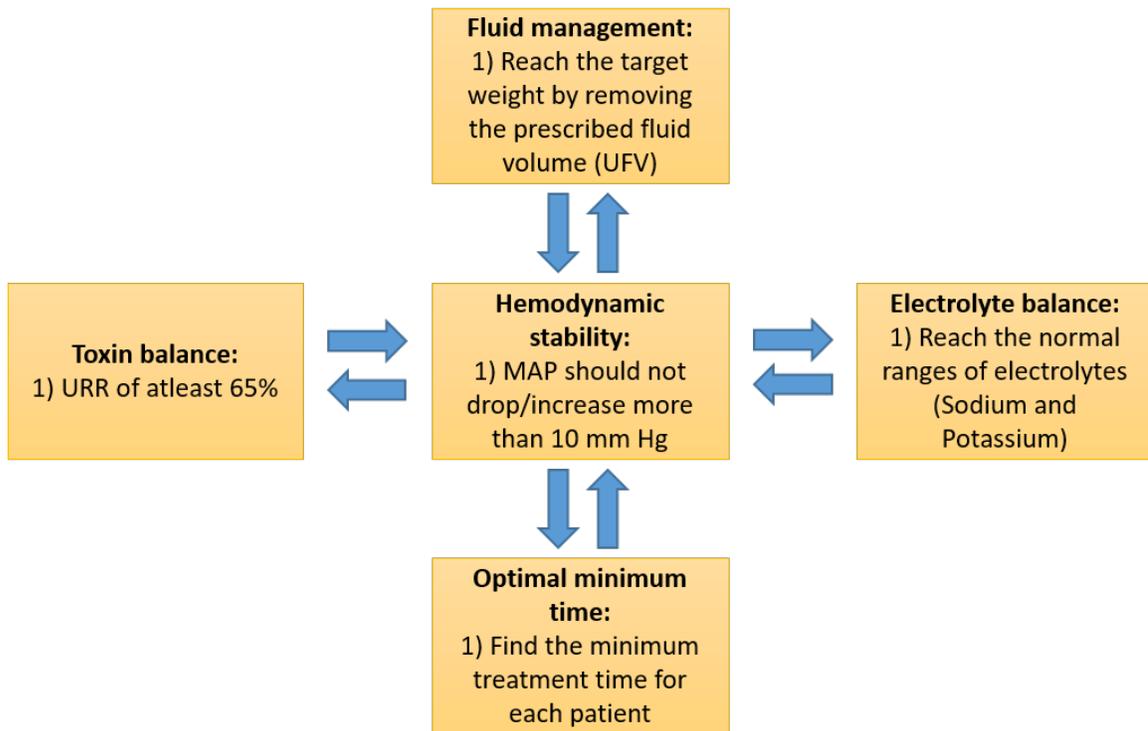


Figure 5.13: Conflicting objectives posed by the HD problem statement

We can observe from figure 5.13 that each of these objectives are interconnected and achieving one occurs at the cost of the other. For instance, a quick treatment would induce a rapid reduction in the circulating volume and results in a considerable drop in MAP. The quick removal of urea in the hemodialyzer in the first half of the treatment causes a drop in the extracellular osmolarity and this results in the transfer of fluid to intracellular fluid space due to osmosis. This aggravates the risk of cardiovascular instability. UFR on the other hand, does not influence the concentrations of any of the solutes but plays a crucial role in weight management. The volume of fluid

removed during UF is directly drawn from the circulating blood volume and affects the hemodynamic stability of the patient. In our particular problem application, we see that there are more objective (or outputs) than degrees of freedom (or inputs). Zone MPC works well in striking a balance in problems with conflicting objectives and this supports our choice of Zone MPC for HD application.

In HD, clinicians are more interested in the terminal states of the patient, which means that the patient should reach the given HD targets by the end of the treatment. The only exception to this is the blood pressure (MAP), which should be within safe limits at all times. Moreover, as the final time is fixed, our ZMPC problem essentially becomes an end point optimization problem. Putting all this together, HD is a batch problem where the patients have to reach some clinical targets by the end of the treatment. A more natural way for tackling the end point optimization problem is the conversion of the classical MPC receding horizon approach to shrinking horizon as shown in figure 5.14, to restate to the controller that the targets have to be reached at the final time no matter whatever trajectory the states take as long as they satisfy the safety and process constraints. The control interval is 600 seconds, which is sufficiently large enough for the optimal control problem to be solved and ideal enough to reduce excessive wear and tear on the actuators in the dialysis machine. If the total simulation time is 4 hours, the prediction horizon at the initial time would be 24. Then it decreases by one after every iteration as shown in figure 5.14. In addition to all this, the ZMPC has the nonlinear dynamic model of the plant embedded into it. All these put together, our ZMPC can be called as a Shrinking Horizon Batch Zone Nonlinear Model Predictive Controller (SHBZNMPC) or simply Batch ZNMPC (BZNMPC).

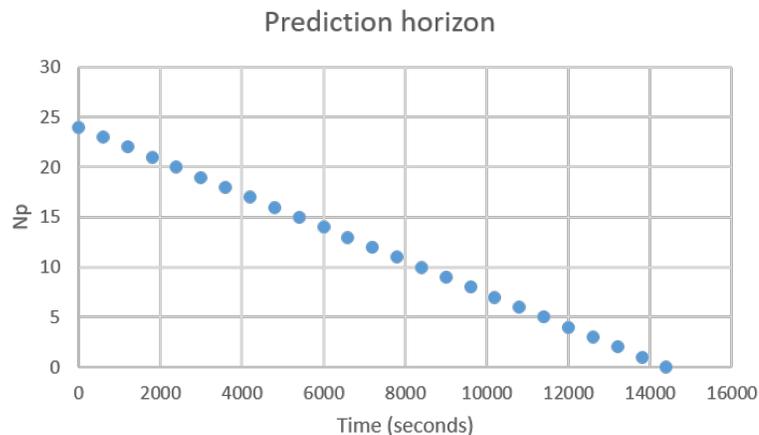


Figure 5.14: Shrinking horizon: Progression of the prediction horizon

Table 5.3: Realization of the BZNMPC problem for application in HD

S.No	Goal	Realization	Mathematical formulation	Penalty matrix	Penalty value	Zone bounds (if any)	
						Lower	Upper
1	Toxin management	Terminal cost (set point)	$M_U^{ex} - 0.3 \times UC \times (V_{pl} + Vis)$	P	1E3	$0.3 \times UC$	$0.3 \times UC$
2	Weight management	Terminal cost (zone)	$V_{pl} + Vis + V_{ic}$	Q_f	1E3	$V_{total}(t_0) - 3500$	$V_{total}(t_0) - 3000$
		Terminal cost (zone)	V_{ic}	Q_f	1E3	$V_{ic}(t_0) - 1000$	$V_{ic}(t_0) - 500$
3	Electrolyte balance	Terminal cost (set point)	$M_{Na}^{ex} - 0.1425 \times (V_{pl} + Vis)$	P	1E4	0.1425	0.1425
		Terminal cost (zone)	C_{Na}^{ex}	Q_f	1E2	0.140	0.145
		Terminal cost (zone)	C_K^{ex}	Q_f	1E2	0.0036	0.0052
4	Hemodynamic stability	Stage cost (zone)	P_{sa}	Q	1E8	90	100
5	Optimal treatment time	Renumeration	Fixed the final time at different values starting from 2 hours (2, 2.5, 3, 3.5, 4) and the BZNMPC problem was solved repeatedly.				

There are a few problems with a general ZMPC formulation too. The cost function is designed in such a way that there will be no penalties when it is within the specified zone. Most of the time, the optimal solution will allow the system to settle at either the upper or lower zone boundaries. This comes as an undesirable feature as the presence of plant model mismatch and other noises will often maneuver the system out of the desired zone region. To overcome this, a mixture of setpoints are also used in the terminal cost formulation. For instance, the extracellular urea concentration at the end of the treatment should have been reduced by 70% (atleast 65% according to National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) which corresponds to a ‘KT/V’ factor of 1.2). The extracellular sodium concentration should be within the zone bounds and also close to 142.5 mmol/litre. So to realize this, it is included both as a set point (middle point of the zone boundaries) and as a target zone. This ensures that the extracellular sodium concentration will stay within bounds and not fluctuate often to noises. The blood pressure stability is enforced by including the MAP within zone bounds by supplying very high penalties for violation. By definition, an increase or decrease of 10 mmHg from the initial blood pressure value is a HD complication. In our data set, it was assumed that the initial MAP of all the patients was 100 mmHg. This justifies the selection of 90 and 110 mmHg as the bounds for MAP. Moreover, high penalty for leaving the MAP safe zone will prevent the controller from choosing input trajectories that could cause rapid pressure changes. The bounds mentioned in table 5.3 are not rigid and the user is free to construct any of the bounds based on their problem requirements and the data set in hand. The method in which the optimal minimum treatment time was found is discussed more in detail in a separate section.

Finally, the reader is presented with the configuration of manipulated inputs considered in our BZNMPC problem in table 5.4. Most of the bounds were adopted from other works in literature [60], [156]. Infusion fluid has not been explored as a degree of freedom before and so it was set at logical bounds with a high penalty for utilization. In clinical practice, saline injection is administered when there is an occurrence of severe IDH or in other treatment modalities. Hypertonic saline solutions are available in 10% or 20% concentrations commercially. We have employed a 10% hypertonic solution as the infusion fluid and the concentration of it has been adopted from Lavoisier Chemical’s Material Safety Data Sheet (MSDS). No strict state constraints are enforced in the problem formulation. The lower bounds of all states were set at 0 and the upper bounds were at infinity. The introduction of slack variables and the presence of the model itself in the constraints make sure that the model predictions are meaningful.

Table 5.4: Configuration of manipulated inputs in BZNMPC

S.No	Manipulated variable	Unit	Bounds		Rate of change bounds		Penalty for rate of change
			Lower	Upper	Lower	Upper	
1	Ultrafiltration Rate (UFR)	ml/s	0	0.3472	-0.05	0.05	1E3
2	Dialysate Sodium Concentration (DSC)	mmol/ml	0.130	0.165	-0.01	0.01	1E3
3	Blood flow rate	ml/min	0.3472	400	-25	25	1
4	Dialysate flow rate	ml/min	400	800	-25	25	1
5	Infusion Rate (IR)	ml/s	0	0.05	$-\infty$	$+\infty$	1E5
6	Infusion Fluid Concentration (IFC)	mmol/ml	1.711	1.711	0	0	1E3

5.5 BZNMPC: Simulation results and discussion

Before starting the control experiments, the controllability of the system was checked using the PBH test and the system was found to be stabilizable. In this section, the results obtained by applying the BZNMPC algorithm to the patients (in table 4.4) are explained. A weight reduction of at least 3 kilograms was targeted during HD. A URR of at least 65% was targeted and the extracellular sodium concentration at the end of the treatment was targeted to be within acceptable physiological ranges [110]. The BZNMPC and the EKF work at the same sampling interval of 600 seconds. Clinical measurements (MAP and HP) are taken every 600 seconds. The identified model for each of the patients in chapter 4 was set as the embedded model in the controller and the estimator. The variables of the plant model were set at the true values. This by default, introduces plant model mismatch, although the estimation performance of the identified model closely resembled the true patient behaviour. The process noise ($N(0, 0.005^2)$) and measurement noise ($N(0, 0.025^2)$) are white noise sequences and

they are drawn from normal distributions. They are assumed to be of additive nature. In HD clinical protocol, a blood sampling is done to get an idea of the blood urea levels, other solute levels before dialysis and so it was assumed that the concentrations of all solutes were known before the treatment started. The nominal values used for each patient in chapter 4’s analysis were supplied as the initial guess to the EKF. The EKF reconstructs the full state vector based on new measurements, and supplies it as the current state to the BZNMPC. The controller then uses this as the starting point and computes the optimal value of inputs based on an optimization framework. The user has to specify the type of dialyzer used for HD and we have considered K_0A to be 900 ml/min in our simulation study. This closely matches with Xenium H11 polyethersulfone membrane dialyzer from Baxter International Incorporation. The K_0A values can be set at the values mentioned in the manufacturer’s design specification sheet.

Different methods were attempted and finally the mathematical models were implemented in CasADi framework [4] and the numerical integration was solved by using Orthogonal Collocation on Finite Elements (OCFE) with 3 collocation points for every sampling interval. The OCFE method minimises the difference between the Lagrange interpolation and the actual solution at predetermined collocation points [156]. The OCFE method is used to simulate the HD process and solve the optimal control problem based on the embedded mathematical model. IPOPT solver was used to solve the optimal control problem as the nature of the problem was very sparse and large. According to the founders of CasADi, IPOPT works best for very large and sparse Non Linear Programming (NLP) problems. The tolerance for the problem was set at 1×10^{-4} and the maximum number of iterations for each iteration was set at 5000. The BZNMPC optimal control framework was realized by using the ‘MPC-Tools’ open source package developed in the research group of James B. Rawlings. ‘MPCTools’ works with CasADi in the back-end. The summary of the results from the BZNMPC simulations are outlined in table 5.5. The reader should look into the table results in parallel when reading the next three subsections.

Table 5.5: Summary of BZNMPC results

Patient index	Target UFV (mL)	Actual UFV (mL)	Target blood urea levels (mmol/L)	Actual final blood urea concentration (mmol/L)	Target blood sodium levels (mmol/L)	Actual final blood sodium concentration (mmol/L)	Optimal minimum treatment time (hours)
1	3000 to 3500	3600.44	8.58	7.16	140 to 145	141.57	3
2	3000 to 3500	3623.32	11.38	8.18	140 to 145	141.66	3.5
3	3000 to 3500	3683.99	7	6.99	140 to 145	141.77	3
4	3000 to 3500	3600.91	9.38	6.16	140 to 145	141.66	3.5
5	3000 to 3500	3624.82	12.6	12.6	140 to 145	144.39	3
6	3000 to 3500	3628.23	10.57	9.71	140 to 145	142.88	3.5
7	3000 to 3500	3619.57	10.33	9.63	140 to 145	142.34	3

5.5.1 Class 3 patient: Severe IDH & weak hemodynamic stability

To make our case study more interesting let us create a hypothetical clinical case as mentioned below. The reader should note that if a trajectory is within the green shaded area, it is within the specified zone targets.

Hypothetical control problem A ESRD patient (Patient 4) enters the clinic with an increase of roughly around 3 kilograms in body weight from the previous hemodialysis treatment because of poor lifestyle (eating and drinking) choices. Patient 4's kidneys have completely failed to work and the patient has been in the queue for a very long time for kidney transplantation. Meanwhile, HD being the life saving therapy for patient 4, patient 4's nephrologist wants 3 litres of overloaded fluid to be removed and a reduction of atleast 65% in his blood urea concentration and his blood sodium and potassium levels to be within 140 to 145 mmol/litres and 3.6 to 5.2 mmol/litres respectively, so that the patient can be assured to be at homeostasis while exiting the treatment centre. Patient 4's medical records say his cardiovascular stability during hemodialysis treatment is very poor and the nurses had to stop the treatment several times and then restart after keeping him in the Trendelenburg position position (TP) for a while (or by administering hypertonic saline injections). How do we help Patient 4 get a continuous safe treatment so that he does not have to spend long hours in the treatment center and the clinical care givers and the patient can be at less risk?

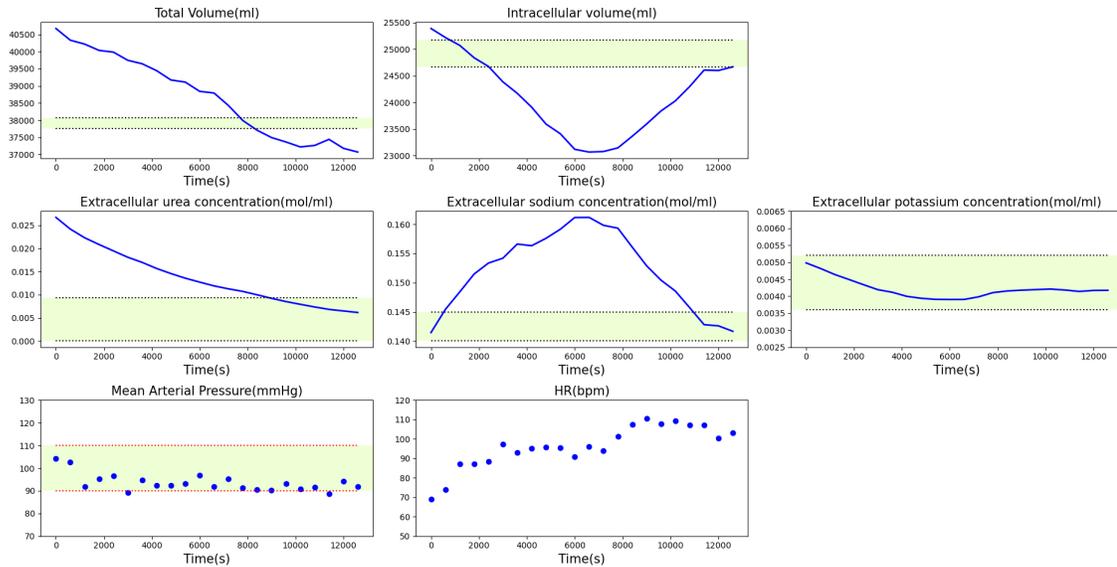


Figure 5.15: Patient 4: Treatment objectives (3.5 h)

The hypothetical problem posted involves solving our original BZNMPC problem. All factual information mentioned in the problem statement has been formulated as target zones and constraints in the BZNMPC problem. The optimal profiles calculated for each of the inputs will be the optimal tailored HD treatment for patient 4. We observe from figure 5.15 that, all the treatment objectives are met. Let us understand the objectives one after the other. The images on the first row cater to the weight management objective. The total volume removed by UF from patient 4 is 3600.91 ml. This is 2.8% higher than the lower bound (3500 ml) supplied to the BZNMPC. As long as the BZNMPC meets all treatment targets it could be taken as a successful treatment. Under treating the patient could be worse than over treatment. The hemodynamic stability of patient 4 is not disturbed by removing slightly more fluid than required. The potential cause of over treatment could be plant model mismatch, EKF estimation errors and the process noise itself. The second zone imposed on the intracellular volume keeps a check on the amount of fluid that can be inter transferred between the compartments. There are higher chances of DDS if the intracellular volume is excessively removed during HD.

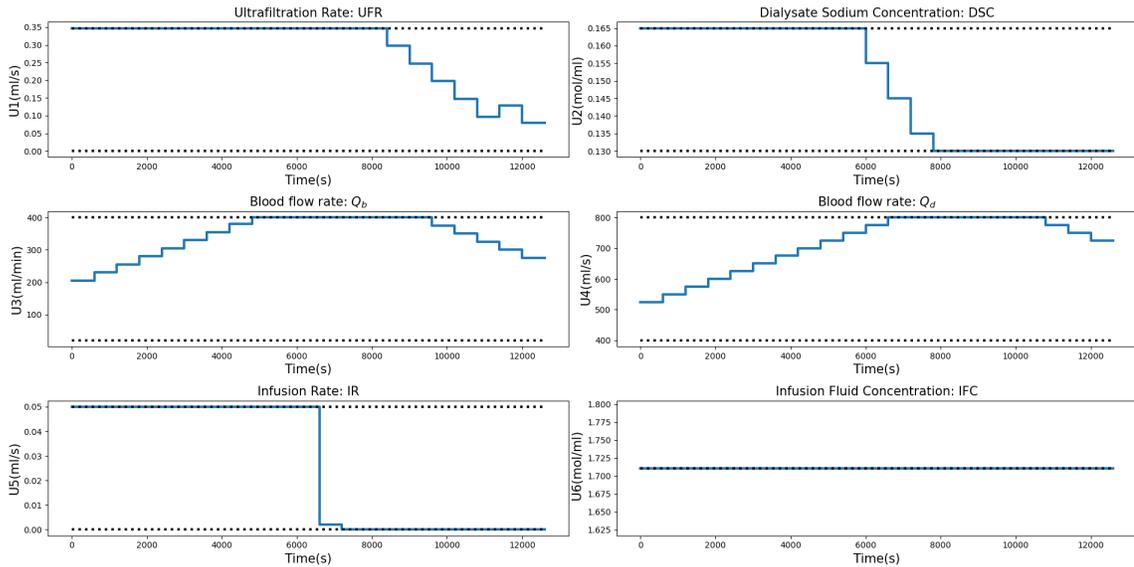


Figure 5.16: Patient 4: Optimal treatment profiles from the BZNMPC (3.5 h)

DSC profiling increases the extracellular osmolarity and pulls more fluid out of the intracellular pool during the first half of the treatment. In the second half, the extracellular sodium concentration has to be brought back to physiological limits and so DSC runs at the lower bounds. The blood and dialysate flow rates were optimally adjusted to reach the urea targets before the end of the treatment. The dialysate has

constant concentration of potassium and so the extracellular potassium concentration is driven closer to the potassium levels in the dialysate. An interesting observation is that this particular patient 4, whose cardiovascular stability has always been notorious, has successively survived the treatment without a considerable drop of MAP more than 10 mm Hg. Few points outside the MAP zone could be potentially because of measurement errors and there is no evidence of rapid blood pressure drop. The heart rate of the patient rises right after the commencement of treatment, indicating that the patient's reflex mechanism is working hard to maintain the hemodynamic stability. The concentration of the infusion fluid was fixed and so the controller can only determine whether to use the infusion fluid or not through U_5 . However, the user can set the concentration of the infusion fluid equal to 10% or 20% concentration. The mixed use of set point based tracking and zone control has succeeded in maintaining the sodium levels close to the midpoint between the upper and lower zone boundaries. We can see observe from figure 5.15 that all the treatment objectives have been met, while satisfying all constraints and so the optimal treatment profiles given by the BZNMPC (shown in figure 5.16) can be considered as the chosen treatment regime for patient 4. The simulations were started by fixing the treatment time as 4 hours and as it was successful the treatment time was reduced by 30 minutes to 3.5 hours. The results presented above are for a total treatment time of 3.5 hours.

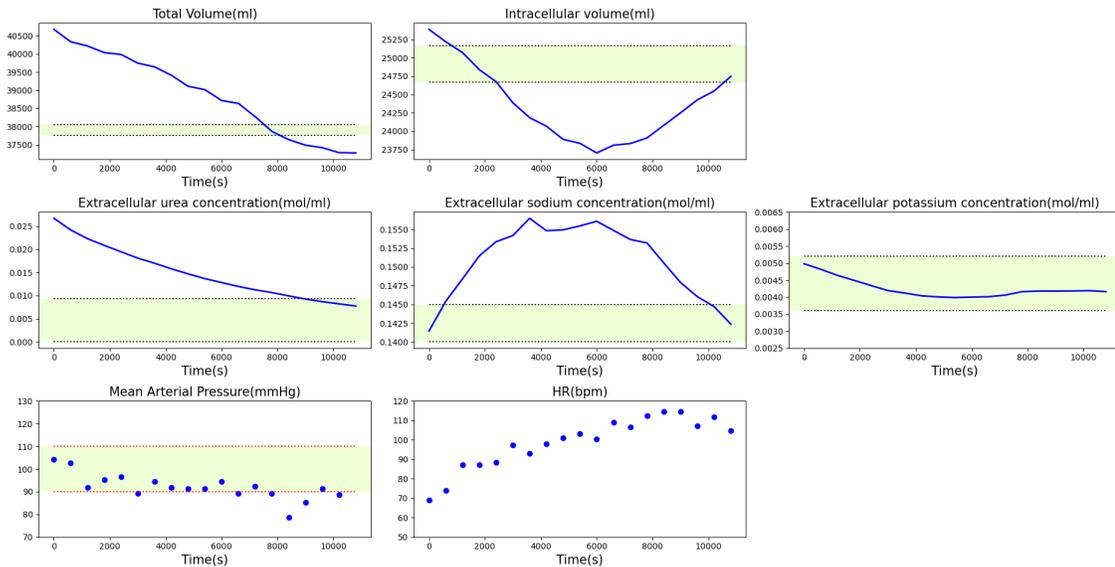


Figure 5.17: Patient 4: Failed - Treatment objectives (3 h)

After the successful trial, let us again reduce the treatment time to 3 hours and then resolve the BZNMPC problem again. We observe from figure 5.17 that, all the treat-

ment objectives are met except MAP stability for a 3 hour treatment. The patient's MAP starts falling rapidly right after the second hour of treatment. This kind of a rapid pressure drop is not favourable in the context of HD and so we deem this scenario as a fail. Only if all the objectives are met, the corresponding treatment regime is chosen to be successful. Therefore, the treatment profile (figure 5.18) computed for the 3 hour case is rejected and a 3 hour treatment is not possible for this patient if the clinician wants to meet all the treatment objectives without compromising on anything. The treatment were repeated for 2.5 and 2 hours and the results of the treatment were recorded. This will help us comment on the minimum optimal treatment time.

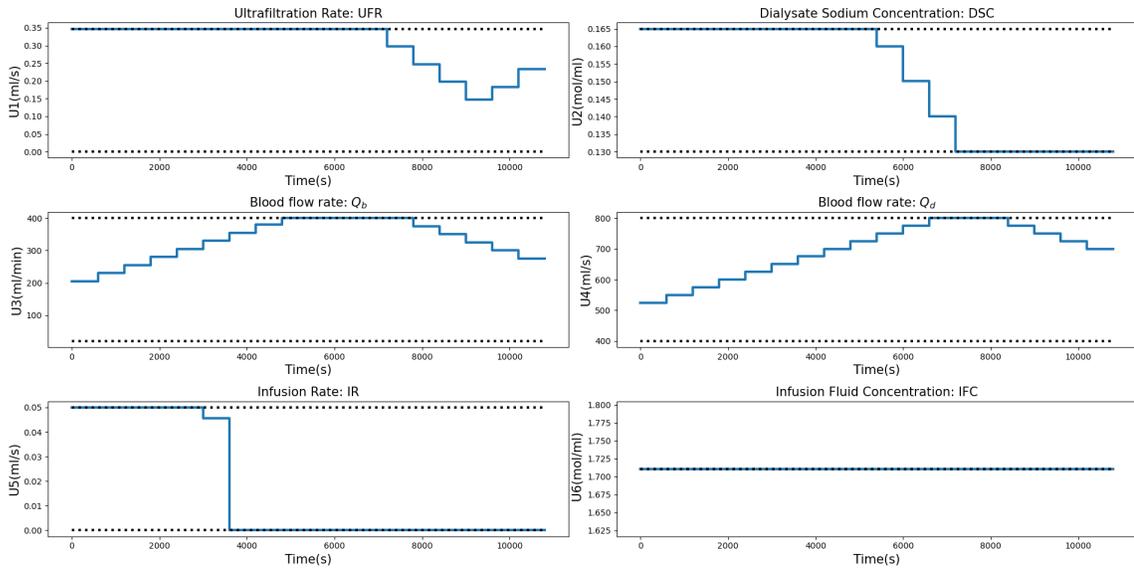


Figure 5.18: Patient 4: Failed - Optimal treatment profiles from the BZNMPC (3 h)

5.5.2 Class 2 patient: Mild IDH & acceptable hemodynamic stability

Let us look into the BZNMPC results obtained for a class 2 patient in this section. Patient 1 is the representative of the CKD patient group who exhibit mild hypotension during HD. At first, the BZNMPC problem was solved for a treatment time of 4 hours and after every successful treatment the treatment time was reduced by 30 minutes. The first glance at figure 5.19 indicated that all treatment objectives have been met for this patient in 3 hours in contrast to a standard 4 hour treatment. However, there are signs of mild pressure drop toward the end of the treatment. This is coupled with a gradual increase in heart rate towards the end of the treatment. The actual volume removed by the controller is 100.44 ml more than the specified lower bound, possibly

because of process disturbances and a EKF estimate offset. The bounds could be adjusted for reducing the amount of fluid removed by ultrafiltration. However, in our design we tried to keep all settings of the BZNMPC across different patient groups uniform to comment on the robustness of the BZNMPC tuning parameters selected.

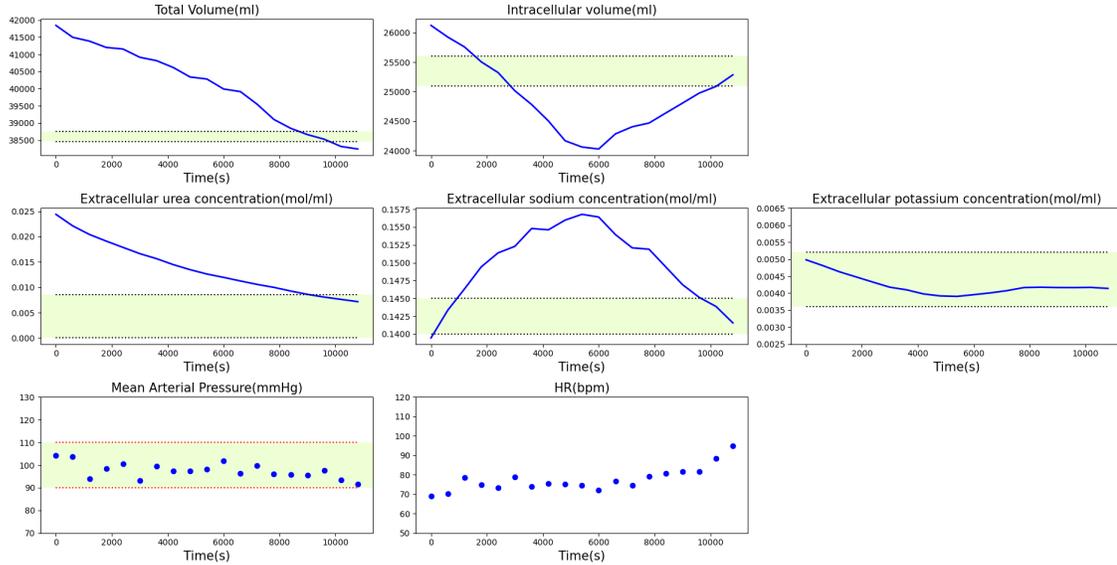


Figure 5.19: Patient 1: Treatment objectives (3 h)

We can associate the BZNMPC's ability to run at UF rates closer to the upper bounds, as portrayed in figure 5.20, to the increased cardiovascular stability of a class 2 patient when compared with a class 3 patient.

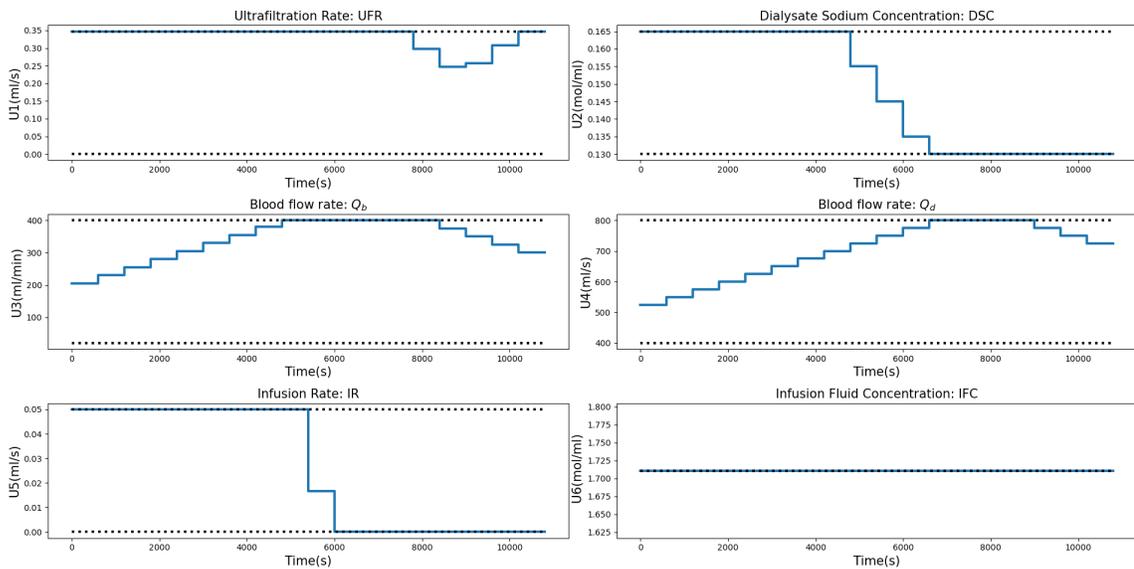


Figure 5.20: Patient 1: Optimal treatment profiles from the BZNMPC (3 h)

5.5.3 Class 1 patient: No IDH & good hemodynamic stability

As the final class candidate, let us examine the results from the BZNMPC for a class 1 patient. Patient 7's results are presented here. However, the other patients (3 & 5) in this category displayed similar behaviour. One can observe from figure 5.21 that all the targets could be met without the patient undergoing signs of hemodynamic instability. Though is a slight extra removal of fluid by ultrafiltration, the patient's heart rate remains flat indicating that this patient can handle a fluid removal of upto 3.6 litres at ease without experiencing any complications. One other reason for the extra removal of total fluid volume, in addition to the process disturbance, for all the test cases revealing a similar behaviour could be because of the fact that the same EKF settings were used for state estimation.

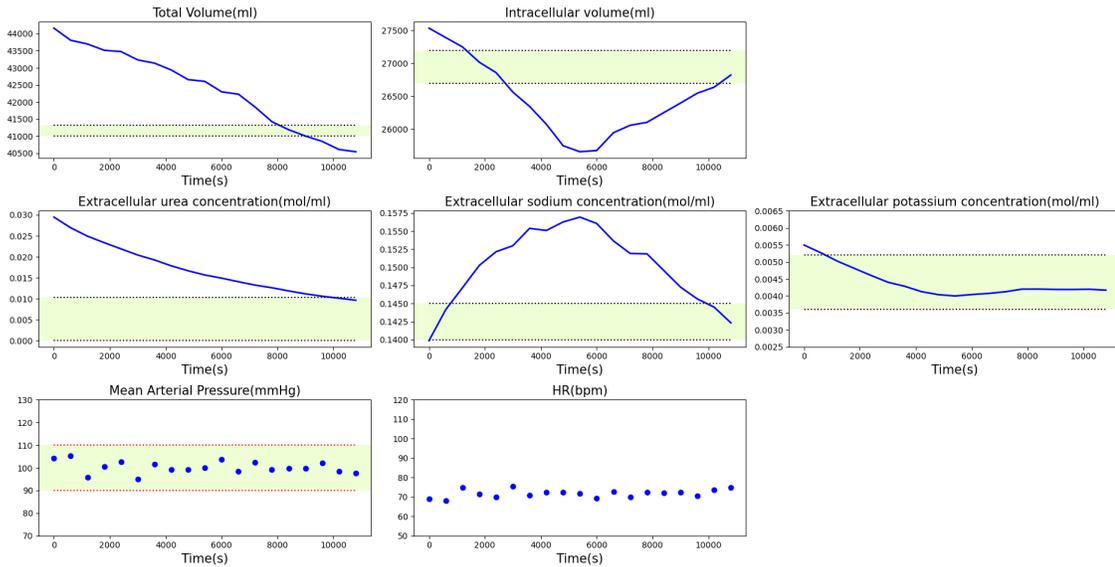


Figure 5.21: Patient 7: Treatment objectives (3 h)

The superior feature of our approach in determining the optimal profiles in comparison with other research works in literature [30], [138] is that the BZNMPC has the independence to choose any profile trajectory between the lower and upper bounds of inputs. In the cited literature, the DSC profiles were forced along a polynomial curve to help the controller choose a DSC profile which will have a peak at the first half of the treatment. In our design, we plugged in the cardiovascular predictive capability into the model and let the controller choose the optimal trajectory of inputs. The BZNMPC has naturally chosen higher UF rates for class 1 and class 2 patients, while for class 3 patients the UF profile looks like a decreasing step towards the end of

the treatment. This shows that the controller has the acumen to visualize that a class 3 patient will experience more frequent pressure drops in the last hour of HD treatment. The optimal treatment profiles for a class 1 patient (patient 7) are shown in figure 5.22. These treatment profiles are the personalized HD therapy regimes for patient 7.

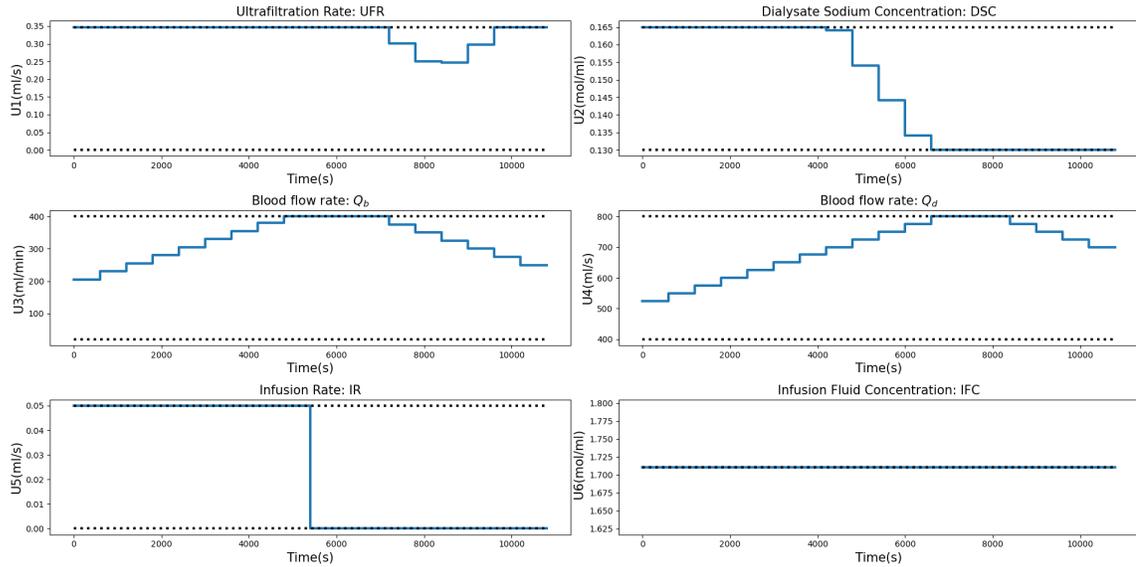


Figure 5.22: Patient 7: Optimal treatment profiles from the BZNMPC (3 h)

5.6 Determination of optimal HD treatment time

As a final exercise for this chapter, the results from the BZNMPC results were plotted in a chart with the different treatment objectives of HD as shown in figure 5.23. The objectives of the HD treatment include Toxin Management (TM), Weight Management (WM), Electrolyte Balance (EB), Safety Constraints (SC). The HD treatment times could be anywhere between 2 to 6 hours [40]. Our goal in this section is to find the minimum optimal treatment time in which all treatment objectives would be met without any compromise. All the objectives are given equal importance when determining the success or failure of a BZNMPC computed treatment profile regime. If a particular objective is met during HD it is marked in green and if it is not satisfied the box is coloured red. Our trials taught us that all the treatment objectives were met for any treatment time above 4 hours and so they were omitted for clarity. The longer the dialysis, the more stable the patient will be and the dialyzer has enough time to clear all toxins and regulate the electrolytes within tight limits with ease. The BZNMPC simulations were repeated for every patient recursively only by changing

the final treatment time at every iteration. We see that all patients were able to handle a treatment time reduction of 30 minutes from the standard 4 hour dialysis. The hemodynamically unstable patients in the data set (Patients 2, 4, 6) displayed signs of safety constraints violation (MAP decreased by more than 10 mmHg) when subjected to a 3 hour HD removing roughly around 3.5 litres of overloaded fluid. However, all patients with acceptable and good cardiovascular stability (Patients 1, 3, 5, 7) were able to handle upto a reduction of 1 hour in the standard treatment time of 4 hours, while meeting all treatment objectives. Even at 2.5 (and 2) hours HD, the healthy patients did not exhibit signs of cardiovascular instability. The limiting factors preventing the healthy patients from undergoing successful 2.5 hour treatments are purely because of mass transfer (only toxin management and electrolyte balance are not met). We chose a low flux membrane from Baxter for all our simulations but if the simulations were repeated with the membrane characteristics of a high flux dialyzer allowing more clearance of solutes, the 2.5 hour target could be reached for healthy patients.

Patient index	Cardiovascular health	Treatment time in hours																			
		2				2.5				3				3.5				4			
		TM	WM	EB	SC	TM	WM	EB	SC	TM	WM	EB	SC	TM	WM	EB	SC	TM	WM	EB	SC
1	H	Red	Red	Red	Green	Red	Green														
2	U	Red	Red	Red	Green	Red	Green														
3	H	Red	Red	Red	Green	Red	Green														
4	U	Red	Red	Red	Green	Red	Green														
5	H	Red	Red	Red	Green	Red	Green														
6	U	Red	Red	Red	Green	Red	Green														
7	H	Red	Red	Red	Green	Red	Green														

Note : H - Healthy, U - Unhealthy, TM - Toxin Management, WM - Weight Management, EB - Electrolyte Balance, SC - Safety Constraints

Figure 5.23: Chart for finding the optimal treatment time based on re-numeration

In addition to the above analysis, the bounds on inputs like ultrafiltration rate have also influenced the 2 hour treatment results for healthy patients. The upper bound for UFR (1.25 L/h) was taken from literature [60]. In addition to the choice of dialyzer, the relaxing of upper bounds could enable 2 hour HD treatments for patients with excellent cardio vascular stability (Class 1). Although the re-numeration method is a more crude approach to find the optimal minimum time for each patient, our comprehensive control framework with the nonlinear system model has given more insights into process dynamics. Based completely on our simulation settings and results from our data set, we conclude that BZNMPC is able to reduce the treatment

time by 30 minutes for a class 3 patient and by an hour for a patient belonging to classes 1 and 2.

5.7 Summary

In this chapter, the reader was first explained about the working of a commercially available feedback control strategy called Blood Volume Tracking (BVT) system, which consists of PID controllers. PID could be the gold standard for process industries but as the subjects involved in HD are live human beings, the problem statement calls for a comprehensive optimal control framework that would have the ability to compute optimal treatments taking the patient's safety and other treatment objectives into factor. To make the process of renal replacement therapy more physiological, a more accurate and complete monitoring and adaptive control of the dialysis machine parameters are required. Our proposed framework for optimal control of HD consists of a Batch Zone Nonlinear Model Predictive Controller (BZNMPC), embedded with a dynamic nonlinear model of the HD system, along with an Extended Kalman Filter (EKF) for state estimation from noninvasive clinical measurements (MAP and HP). The tuning parameters of the BZNMPC and EKF were left unchanged during the entire simulation study. The designed optimal control framework was tested on synthetic clinical data and the BZNMPC was able to successfully compute optimal treatment profiles which were continuous and safe even for a class 3 patient (characterized by weak hemodynamic stability) in the presence of disturbances (process noise and plant model mismatch) and measurement uncertainties. Finally, the minimum optimal HD treatment times were computed based on a sequential simulation approach.

The reader might have understood by this chapter that automatic control of HD, based on advanced model-based process control strategies, has the potential to deliver better treatments to the galloping number of ESRD patients who are increasingly being diagnosed with complicated co-morbid conditions. The mission of introducing automation control is not very easy from an engineering and medical point of view, as the knowledge base about advanced process control application in HD is only booming. Moreover, the advanced feed control strategies are primarily designed by control engineers. Traditional medical practitioners may be demeaned by such a complex engineering intervention in HD. However, the success of personalized precision medicine in HD would be largely due to a synergy between medicine and engineering.

Chapter 6

Conclusions

6.1 Conclusion

In this chapter, the author concludes all of the research investigations that has been done in this thesis, which includes the design of new strategies to improve the safety, quality and efficiency of the hemodialysis treatments delivered today, with the application of well understood systems and control engineering principles. The prime agenda was to build new strategies to improve today's hemodialysis treatments, backed by fundamental scientific principles and with a huge potential of practical implementation. This work contains the design and application of an individualized virtual patient simulator that can be embedded in an optimal control framework in hemodialysis (HD) with the objective of achieving safe and continuous HD treatments even for an End Stage Renal Disease (ESRD) patient, characterized by poor cardiovascular stability (severe intradialytic hypotension and poor vascular refilling). The proposed approaches were tested on synthetic clinical data generated from the original nonlinear HD system model and with the help of extensive simulation experiments. The main body of the thesis consisted of three main parts as outlined below.

The first part (chapter 3) primarily was an attempt to identify the most comprehensive model available in literature to represent all the components (patient and hemodialyzer) and intricate physiological phenomena happening in a HD system. The adopted model not only had the ability to simulate the solute, fluid kinetics and bidirectional mass transfer dynamics of solutes across the hemodialyzer membrane, but also it had the inherent feature to predict the hemodynamics of different classes of patients who are typically encountered in a practical HD clinical setting. A few

modifications were done to the some of the equation parameters of the HD system model to make it more control application friendly.

The second part (chapter 4) revolved around the customization of the adopted mathematical model of HD system with the objective of converting it into a tailor-made virtual patient simulator that could closely predict the hemodynamics of the actual patient under consideration. This was essentially realised by formulating a simultaneous state and parameter estimation problem as a nonlinear least squares full information estimation problem. Our proposed approach to solve the resulting simultaneous state and parameter estimation problem included all information from a modified sensitivity-based observability analysis and a sequential optimization based solution strategy. We understood from our simulations that system observability played a quintessential role in arriving at meaningful estimations of the states and parameters and governed the maximum extent of information that could be recovered from available measurement data. Based on experiments, our proposed approach turned out to be successful in identifying the custom virtual patient simulator, although the available non invasive clinical measurements were noisy. Furthermore, the custom virtual patient simulator predicted the actual patient hemodynamics (Mean Arterial Pressure (MAP) and Heart Period (HP)) reasonably well for all different classes of HD patients, even for a wide range of HD treatment profile settings.

Fundamentally, HD causes an external perturbation of the patient from the initial physiological steady state and the patients have to be brought back to homeostasis before exiting the treatment centre. The quick removal of overloaded fluid and extracorporeal cleansing of blood induces clinical complications in patients undergoing HD. The third part (chapter 5) focused on the development and application of an advanced feedback control strategy based on the hemodynamic responses of the patients undergoing treatment. The custom virtual patient simulator (identified from chapter 4) was proposed to be used as the embedded model of the patient in the controller and a computer-controlled HD system was developed that would help meet all treatment objectives while ensuring superior patient safety. The designed computer-controlled HD system was validated through simulation experiments and its performance was demonstrated with the help of test cases. This was the first step towards developing a new feedback control strategy, based on a identified grey box patient model, capable of preventing clinical complications, like Intradialytic Hypotension (IDH) and Dialysis Disequilibrium syndrome (DDS), which are usually encountered in HD.

6.2 Major findings of this current work

The main contributions and findings of this research work are summarized with greater detail in this section.

- **Chapter 3:** In this chapter, the existing models available in literature for representing the different components of the HD system was studied. Both simple equations and complex Partial Differential Equations (PDEs) to represent the bidirectional mass transfer process taking place in a hemodialyzer were explored and validated against actual clinical data from literature. The PDEs were realized in a control application framework but only the simple equations were chosen to represent the hemodialyzer, owing to the computational simplicity and the requirements set by the problem statement of this research work. Modifications were done to the way the dialysance of solutes were computed in the simple equations, based on a log mean concentration difference approach for counter current flow of process and service fluid streams. This enabled the model to have more degrees of freedom (6 manipulated variables) namely, Ultrafiltration Rate (UFR), Dialysate Sodium Concentration (DSC), blood flow rate, dialysate flow rate, Infusion Rate (IR), Concentration of the infusion fluid, that could be adjusted by the clinicians or the dialysis machine during the course of the treatment.
- **Chapter 4:** Here, our proposed method converted the adopted model from literature to represent the HD system, to an individualized virtual patient simulator through a grey box model identification approach, in terms of system identification. In this chapter, we found that we needed atleast two clinical measurements (MAP and HP considered as they are available non-invasively) to make the original HD system model fully observable. We then augmented the system with patient-specific parameters, considering them as additional states of the system with zero dynamics. For consistent estimation, observability of the augmented system has to be ensured. The augmented system was highly nonlinear and so a modified observability test was conducted to pick the largest subset of observable variables. Then a Sequential Coordinate Block Descent (SCBD) optimization framework was employed to solve the full information state estimation problem formulated in a nonlinear least squares sense and to efficiently integrate the information from the observability analysis procedure. The model thus identified worked for all different classes of HD patients and proved its practical implementation potential.

- **Chapter 5:** The most important contributions of this chapter include the realization and validation of a model-based feedback control strategy. The HD treatment had conflicting objectives and so zone control was the ideal choice. We formulated and tested a Batch Zone Nonlinear Model Predictive Controller (BZNMPC) with a built in nonlinear state estimator (Extended Kalman Filter (EKF)), with feedback implementation which takes the treatment objectives and safety constraints into consideration. The clinicians were more interested in the physiological conditions of the patient only towards the end of the treatment rather than focussing on the treatment objectives at every time instant, with the only exception of the patient’s safety conditions which were monitored throughout the treatment sessions. Therefore, the BZNMPC was built in a shrinking horizon framework for end point treatment optimization. The BZNMPC was experimentally tested in the presence of measurement noise and model uncertainty, and it succeeded in meeting the clinical treatment objectives, while ensuring continuous optimal treatments that satisfied all the process, safety and input constraints. The virtual treatment simulation results demonstrated that this kind of feedback control strategy is practically applicable and feasible for all different classes of patients. Finally, the optimal treatment time for each patient in a data set from literature was computed by following a re-numeration based approach.

6.3 Future research directions

The reader is navigated through the possible research extensions of this thesis work in this section. The suggestions outlined below are completely based on the results and inferences from this research work.

- **Chapter 3:**
 - Although the adopted model from literature for this research work had several assumptions and simplifications, it quantified all HD induced dynamics with an acceptable level of accuracy for this research work. Future analysis could be done with even more comprehensive models in literature [110], that give the user the capacity to understand the dynamics of several other solutes, along with pulsatile blood pressure curves and dialysate temperature dynamics.
 - The PDE model of mass transfer of the hemodialyzer was built by utilizing

all the design specifications and characteristics of a hemodialyzer. Sometimes the hemodialyzer used for a particular patient could be over specified or under specified. Thus the PDE model of hemodialyzer could be integrated into the HD system model, and by treating the hemodialyzer design specification parameters as decision variables, the most optimal choice of hemodialyzer can be selected from a wide range of commercially available options.

- **Chapter 4:**

- The results from the sensitivity analysis can be used to re-parameterize the model equations and the observability results could be validated using nonlinear system observability test involving the computation of Lie-derivatives to get a sense of global system observability.
- Treatment data from multiple sessions (3 to 4) over a week, can be used to identify the patient-specific parameters of the model and a clinical track record of the patient-specific parameters could be maintained.

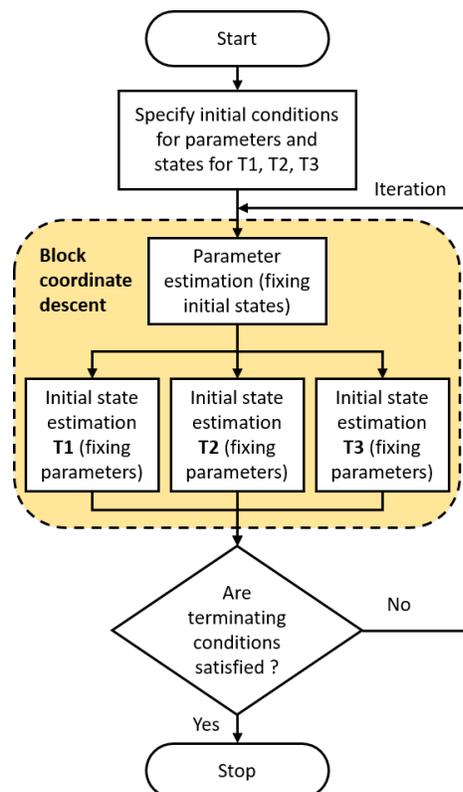


Figure 6.1: Assimilation of patient’s clinical data from multiple HD treatments

- A combination of a global optimization solver and a local gradient based solver can be used when searching for the optimal values of the decision variables during the individualized model identification problem. Furthermore, improvements could be brought about in the optimization solution strategy itself.

- **Chapter 5:**

- The designed computer-controlled system was tested on a small number of patients from a data set in literature. Only additive noises (process and sensor) and plant model mismatch was included in the formulation of the stochastic controllers in this research work. Inclusion of patient-specific parameter uncertainty in the optimal control problem formulation would be the most appropriate way to obtain a robust controller. Then, actual clinical studies could be conducted on a large population of patients, with different physiological characteristics, in a practical HD setting.
- As our original system model is nonlinear, a rigorous optimal estimator like a 'Nonlinear Moving Horizon estimator (NMHE)', which is more robust and accurate than a EKF, could be tested in future studies for state estimation also because of its ability to enforce physiological bounds on state variables.
- Optimal treatment time was found out by using a re-numeration based approach in chapter 5. In future, it could be computed more efficiently by formulating a Mixed Integer Nonlinear Program (MINLP) with final treatment time as an added decision variable in the optimal control problem solved in this work.
- The conventional HD schedule of 3 to 4 times a week has been used as the standard modality to treat ESRD patient. However, the optimal HD schedule that suits each patient may be different. Reduced dialysis-induced myocardial stunning with frequent HD, compared to conventional HD, was reported and improvements in hyperphosphatemia and blood pressure control were also reported with frequent HD. With all the above information, the next step could be the optimal HD treatment scheduling for a single patient (over a week) and it could be possibly extended to the resource optimization of a complete HD treatment facility.

Bibliography

- [1] Hasan Abaci and Sacide Altinkaya. Modeling of hemodialysis operation. *Annals of biomedical engineering*, 38:3347–62, 11 2010.
- [2] Suhail Ahmad. *Manual of clinical dialysis*. Springer Science & Business Media, 2009.
- [3] Kese Alberton, André Luís Alberton, Jimena Maggio, Maria Diaz, and Argimiro Secchi. Accelerating the parameters identifiability procedure: Set by set selection. *Computers Chemical Engineering*, 55:181–197, 08 2013.
- [4] Joel A E Andersson, Joris Gillis, Greg Horn, James B Rawlings, and Moritz Diehl. CasADi: A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, In Press, 2018.
- [5] Canadian Medical Association. Nephrology Profile. 2019.
- [6] European Renal Association et al. ERA-EDTA registry: ERA-EDTA registry annual report 2018. *Academin Medical Center, Department of Medical Informatics, Amsterdam*, 2020.
- [7] Richard C Aster, Brian Borchers, and Clifford H Thurber. *Parameter estimation and inverse problems*. Elsevier, 2018.
- [8] Shahnaz Atabak. Change in definition could change results. *Nephrology Dialysis Transplantation*, 19(10):2676–2676, 2004.
- [9] François Auger, Mickael Hilairat, Josep M Guerrero, Eric Monmasson, Teresa Orłowska-Kowalska, and Seiichiro Katsura. Industrial applications of the kalman filter: A review. *IEEE Transactions on Industrial Electronics*, 60(12):5458–5471, 2013.
- [10] Ahmad Taher Azar. *Modelling and Control of Dialysis Systems Volume 1: Modeling Techniques of Hemodialysis Systems (Vol. 1)*. Berlin, Heidelberg: Springer, 2013.
- [11] Ahmad Taher Azar. *Modelling and Control of Dialysis Systems Volume 2: Biofeedback Systems and Soft Computing Techniques of Dialysis (Vol. 2)*. Berlin, Heidelberg: Springer, 2013.
- [12] Ahmad Taher Azar et al. Biofeedback systems and adaptive control hemodialysis treatment. *Saudi Journal of Kidney Diseases and Transplantation*, 19(6):895,

2008.

- [13] Nirwair Singh Bajwa. Hemoglobin modeling and simulation for anemia management in chronic kidney disease. 2018.
- [14] CA Baldamus, W Ernst, U Frei, and KM Koch. Sympathetic and hemodynamic response to volume removal during different forms of renal replacement therapy. *Nephron*, 31(4):324–332, 1982.
- [15] Piero L Balestri, Marco Biagini, Paolo Rindi, and Sergio Giovannetti. Uremic toxins. *Archives of internal medicine*, 126(5):843–845, 1970.
- [16] NA Bazaev, VM Grinval'd, and SV Selishchev. Simulation of mass transfer in a hemodialyzer. *Biomedical engineering*, 42(6):307, 2008.
- [17] B.S. Beckett. *Biology: A Modern Introduction*. Modern Introduction: Gcse Edition. Oxford University Press, 1986.
- [18] JL Bert, CC Gyenge, BD Bowen, RK Reed, and T Lund. A model of fluid and solute exchange in the human: validation and implications. *Acta physiologica scandinavica*, 170(3):201–209, 2000.
- [19] Boris Bikbov, Caroline A Purcell, Andrew S Levey, Mari Smith, Amir Abdoli, Molla Abebe, Oladimeji M Adebayo, Mohsen Afarideh, Sanjay Kumar Agarwal, Marcela Agudelo-Botero, et al. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 395(10225):709–733, 2020.
- [20] LA Bland and MS Favero. Microbiologic aspects of hemodialysis systems. *AAMI Standards and Recommended Practices*, 3:257–265, 1993.
- [21] Song Bo and Jinfeng Liu. A decentralized framework for parameter and state estimation of infiltration processes. *Mathematics*, 8(5):681, 2020.
- [22] Song Bo, Soumya R Sahoo, Xunyuan Yin, Jinfeng Liu, and Sirish L Shah. Parameter and state estimation of one-dimensional infiltration processes: A simultaneous approach. *Mathematics*, 8(1):134, 2020.
- [23] Wendi Bradshaw and Paul N Bennett. Asymptomatic intradialytic hypotension: The need for pre-emptive intervention. *Nephrology Nursing Journal*, 42(5):479–485, 2015.
- [24] Wendi Bradshaw et al. Intradialytic hypotension: a literature review. *Renal Society of Australasia Journal*, 10(1):22, 2014.
- [25] Ellen F Carney. The impact of chronic kidney disease on global health. *Nature Reviews Nephrology*, 16(5):251–251, 2020.
- [26] Steven C Chapra and Raymond P Canale. *Numerical methods for engineers*, volume 2. Mcgraw-hill New York, 2011.
- [27] K-J Chou, P-T Lee, C-L Chen, C-W Chiou, C-Y Hsu, H-M Chung, C-P Liu, and H-C Fang. Physiological changes during hemodialysis in patients with

- intradialysis hypertension. *Kidney international*, 69(10):1833–1838, 2006.
- [28] Yunfei Chu and Juergen Hahn. Generalization of a parameter set selection procedure based on orthogonal projections and the d-optimality criterion. *AIChE journal*, 58(7):2085–2096, 2012.
- [29] Thomas F Coleman and Yuying Li. On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds. *Mathematical programming*, 67(1):189–224, 1994.
- [30] Luigi Coli, Mauro Ursino, Vittorio Dalmastrì, Fabio Volpe, Gaetano La Manna, Guido Avanzolini, Sergio Stefoni, and Vittorio Bonomini. A simple mathematical model applied to selection of the sodium profile during profiled haemodialysis. *Nephrology, dialysis, transplantation: official publication of the European Dialysis and Transplant Association-European Renal Association*, 13(2):404–416, 1998.
- [31] Chronic Kidney Disease Prognosis Consortium et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *The Lancet*, 375(9731):2073–2081, 2010.
- [32] Richard L Converse, Tage N Jacobsen, CM Jost, Robert D Toto, Paul A Grayburn, Troy M Obregon, Fetnat Fouad-Tarazi, Ronald G Victor, et al. Paradoxical withdrawal of reflex vasoconstriction as a cause of hemodialysis-induced hypotension. *The Journal of clinical investigation*, 90(5):1657–1665, 1992.
- [33] William G Couser, Giuseppe Remuzzi, Shanthi Mendis, and Marcello Tonelli. The contribution of chronic kidney disease to the global burden of major non-communicable diseases. *Kidney international*, 80(12):1258–1270, 2011.
- [34] John T Daugirdas, PG Blake, and TS Ing. *Handbook of Dialysis. 5th edn, Vol. 2015*. New York, USA: Wolters Kluwer Health, 2015.
- [35] Aurelio A de los Reyes V. Dynamics of a cardiovascular model obtaining measurable pulsatile pressure output. *World Journal of Modelling and Simulation*, 11(1):20–32, 2015.
- [36] Thomas A Depner. *Prescribing hemodialysis: a guide to urea modeling*, volume 29. Springer Science & Business Media, 2012.
- [37] Luca Di Lullo, Andrew House, Antonio Gorini, Alberto Santoboni, Domenico Russo, and Claudio Ronco. Chronic kidney disease and cardiovascular complications. *Heart failure reviews*, 20(3):259–272, 2015.
- [38] Paul Drawz and Mahboob Rahman. Chronic kidney disease. *Annals of internal medicine*, 162(11):ITC1–ITC16, 2015.
- [39] William Drukker, Frank M Parsons, and John Francis Maher. *Replacement of renal function by dialysis: a textbook of dialysis*. Springer Science & Business Media, 2012.
- [40] Sunny Eloit. *Experimental and numerical modeling of dialysis*. PhD thesis,

Ghent University, 2004.

- [41] Casper FM Franssen, Judith J Dasselaar, Paulina Sytsma, Johannes GM Burgerhof, Paul E de Jong, and Roel M Huisman. Automatic feedback control of relative blood volume changes during hemodialysis improves blood pressure stability during and after dialysis. *Hemodialysis International*, 9(4):383–392, 2005.
- [42] Dimitrios Fysikopoulos, Brahim Benyahia, Akos Borsos, Zoltan K Nagy, and Chris D Rielly. A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models. *Computers & Chemical Engineering*, 122:275–292, 2019.
- [43] M. Gałach and Andrzej Werynski. Development of ”virtual patient” model for simulation of solute and fluid transport during dialysis. 2005.
- [44] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU press, 2013.
- [45] Ravi Gondhalekar, Eyal Dassau, and Francis J Doyle III. Periodic zone-mpc with asymmetric costs for outpatient-ready safety of an artificial pancreas to treat type 1 diabetes. *Automatica*, 71:237–246, 2016.
- [46] Frank A Gotch, JA Sargent, ML Keen, and M Lee. Individualized, quantified dialysis therapy of uremia. In *Proceedings of the Clinical Dialysis and Transplant Forum*, number 4, pages 27–35, 1974.
- [47] Frank A Gotch and John A Sargent. A mechanistic analysis of the national cooperative dialysis study (ncds). *Kidney international*, 28(3):526–534, 1985.
- [48] Nik LM Grubben and Karel J Keesman. Controllability and observability of 2D thermal flow in bulk storage facilities using sensitivity fields. *International Journal of Control*, 91(7):1554–1566, 2018.
- [49] Christina C Gyenge, Bruce D Bowen, Rolf K Reed, and Joel L Bert. Transport of fluid and solutes in the body i. formulation of a mathematical model. *American Journal of Physiology-Heart and Circulatory Physiology*, 277(3):H1215–H1227, 1999.
- [50] David Haessig and Bernard Friedland. A method for simultaneous state and parameter estimation in nonlinear systems. In *Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041)*, volume 2, pages 947–951. IEEE, 1997.
- [51] John E Hall and Michael E Hall. *Guyton and Hall textbook of medical physiology e-Book*. Elsevier Health Sciences, 2020.
- [52] Thomas Heldt. *Computational models of cardiovascular response to orthostatic stress*. PhD thesis, Massachusetts Institute of Technology, 2004.
- [53] Nat Hemasilpin. *Toward Optimal Adaptive Control of Hemodialysis*. PhD thesis, University of Cincinnati, 2013.
- [54] Phillip Hoffsten and Saulo Klahr. *The kidney and body fluids in health and*

- disease*. Springer, 1983.
- [55] Sue E Huether and Kathryn L McCance. *Pathophysiology: The Biologic Basis for Disease in Adults and Children, 6e*. Elsevier, 2014.
 - [56] Salim Ibrir. Joint state and parameter estimation of non-linearly parameterized discrete-time nonlinear systems. *Automatica*, 97:226–233, 2018.
 - [57] Lesley A Inker, Brad C Astor, Chester H Fox, Tamara Isakova, James P Lash, Carmen A Peralta, Manjula Kurella Tamura, and Harold I Feldman. KDOQI US commentary on the 2012 KDIGO clinical practice guideline for the evaluation and management of CKD. *American Journal of Kidney Diseases*, 63(5):713–735, 2014.
 - [58] Jula K Inrig. Intradialytic hypertension: a less-recognized cardiovascular complication of hemodialysis. *American Journal of Kidney Diseases*, 55(3):580–589, 2010.
 - [59] Jula K Inrig, Uptal D Patel, Robert D Toto, and Lynda A Szczech. Association of blood pressure increases during hemodialysis with 2-year mortality in incident hemodialysis patients: a secondary analysis of the dialysis morbidity and mortality wave 2 study. *American Journal of Kidney Diseases*, 54(5):881–890, 2009.
 - [60] Faizan Javed. *Monitoring, Analysis and Control of Haemodynamic Variables During Haemodialysis*. PhD thesis, The University of New South Wales, Australia 2010.
 - [61] Vivekanand Jha, Guillermo Garcia-Garcia, Kunitoshi Iseki, Zuo Li, Saraladevi Naicker, Brett Plattner, Rajiv Saran, Angela Yee-Moon Wang, and Chih-Wei Yang. Chronic kidney disease: global dimension and perspectives. *The Lancet*, 382(9888):260–272, 2013.
 - [62] WJ Johnson, WW Hagge, RD Wagoner, RP Dinapoli, and JW Rosevear. Effects of urea loading in patients with far-advanced renal failure. In *Mayo Clinic Proceedings*, volume 47, pages 21–29, 1972.
 - [63] Judith Z Kallenbach. *Review of hemodialysis for nurses and dialysis personnel-e-book*. Elsevier Health Sciences, 2020.
 - [64] Rushikesh Kamalapurkar. Simultaneous state and parameter estimation for second-order nonlinear systems. pages 2164–2169. IEEE, 2017.
 - [65] Fresenius Medical Care AG & Co. KGaA. Fresenius medical care: Annual report. 2020.
 - [66] Kwan E Kim, Martin Neff, Barry Cohen, Michael Somerstein, Joel Chinitz, Gaddo Onesti, and Charles Swartz. Blood volume changes and hypotension during hemodialysis. *ASAIO Journal*, 16(1):508–514, 1970.
 - [67] Genjiro Kimura, John C. Van Stone, John H. Bauer, and Prakash R. Keshaviah. A simulation study on transcellular fluid shifts induced by hemodialysis. *Kidney international*, 24(4):542–548, 1983.

- [68] Saulo Klahr. *The kidney and body fluids in health and disease*. Springer Science & Business Media, 2013.
- [69] Costas Kravaris, Juergen Hahn, and Yunfei Chu. Advances and selected recent developments in state and parameter estimation. *Computers & chemical engineering*, 51:111–123, 2013.
- [70] Uwe Kuhlmann, Rainer Goldau, Nader Samadi, Thomas Graf, Malte Gross, Giancarlo Orlandini, and Harald Lange. Accuracy and safety of online clearance monitoring based on conductivity variation. *Nephrology Dialysis Transplantation*, 16(5):1053–1058, 2001.
- [71] EM Landis. Exchange of substances through the capillary walls. *Handbook of Physiology, circulation II*, pages 961–1034, 1963.
- [72] Andreas Laupacis, David Feeny, Allan S Detsky, and Peter X Tugwell. How attractive does a new technology have to be to warrant adoption and utilization? tentative guidelines for using clinical and economic evaluations. *CMAJ: Canadian Medical Association Journal*, 146(4):473, 1992.
- [73] Michael A Lazzari. Errors in specimen processing and the potential misdiagnosis of acute renal failure in a 12-year-old girl. *Laboratory Medicine*, 43(4), 2012.
- [74] Bangjun Lei, Guangzhu Xu, Ming Feng, Yaobin Zou, Ferdinand Van der Heijden, Dick De Ridder, and David Tax. *Classification, Parameter Estimation and State Estimation*. Wiley Online Library, 2017.
- [75] Juha T Leppavuori, Michael M Domach, and Lorenz T Biegler. Parameter estimation in batch bioreactor simulation using metabolic models: Sequential solution with direct sensitivities. *Industrial & engineering chemistry research*, 50(21):12080–12091, 2011.
- [76] Andrew S Levey and Josef Coresh. Chronic kidney disease. *The lancet*, 379(9811):165–180, 2012.
- [77] Andrew S Levey, Paul E De Jong, Josef Coresh, Meguid El Nahas, Brad C Astor, Kunihiro Matsushita, Ron T Gansevoort, Bertram L Kasiske, and Kai-Uwe Eckardt. The definition, classification, and prognosis of chronic kidney disease: a KDIGO controversies conference report. *Kidney international*, 80(1):17–28, 2011.
- [78] Jeremy Levy, Edwina Brown, and Anastasia Lawrence. *Oxford handbook of dialysis*. Oxford university press, 2016.
- [79] Sharon Mantik Lewis, Margaret McLean, and Shannon R Dirksen. *Medical-Surgical nursing: Assessment and management of clinical problems*. Vol. 1. 2000.
- [80] John K Leypoldt. *The artificial kidney: physiological modeling and tissue engineering*. RG Landes, 1999.
- [81] Chao Li and Li Ren. Estimation of unsaturated soil hydraulic parameters using the ensemble kalman filter. *Vadose Zone Journal*, 10(4):1205–1227, 2011.

- [82] Zhijie Liao, Churn K Poh, Zhongping Huang, Peter A Hardy, William R Clark, and Dayong Gao. A numerical and experimental study of mass transfer in the artificial kidney. *J. Biomech. Eng.*, 125(4):472–480, 2003.
- [83] Robert M Lindsay and Daniel Schneditz. Online monitoring and feedback-control. In *Replacement of renal function by dialysis*, pages 555–584. Springer, 2004.
- [84] Bi-Cheng Liu, Hui-Yao Lan, and Lin-Li Lv. *Renal Fibrosis: Mechanisms and Therapies*. Springer, 2019.
- [85] Jianbang Liu, Aristarchus Gnanasekar, Yi Zhang, Song Bo, Jinfeng Liu, Jingtao Hu, and Tao Zou. Simultaneous state and parameter estimation: the role of sensitivity analysis. *Industrial & Engineering Chemistry Research*, 60(7):2971–2982, 2021.
- [86] NW Loney, CR Huang, and L Simon. Mathematical model of a countercurrent flow multi-fibre dialyser. *World Journal of Microbiology and Biotechnology*, 21(6):791–796, 2005.
- [87] Edmund G Lowrie. The kinetic behaviors of urea and other marker molecules during hemodialysis. *American Journal of Kidney Diseases*, 50(2):181–183, 2007.
- [88] EG Lowrie, NM Laird, TF Parker, and JA Sargent. Effect of the hemodialysis prescription on patient morbidity: Report from the national cooperative dialysis study. *New England Journal of Medicine*, 305(20):1176–1181, 1981.
- [89] Berit Floor Lund and Bjarne A Foss. Parameter ranking by orthogonalization—applied to nonlinear mechanistic models. *Automatica*, 44(1):278–281, 2008.
- [90] M Maasrani, MY Jaffrin, M Fischbach, and B Boudailliez. Urea, creatinine and phosphate kinetic modeling during dialysis: application to pediatric hemodialysis. *The International journal of artificial organs*, 18(3):122–129, 1995.
- [91] Elena Mancini, Emanuele Mambelli, Mina Irpinia, Danila Gabrielli, Carmelo Cascone, Ferruccio Conte, Gina Meneghel, Fosco Cavatorta, Alessandro Antonelli, Giuseppe Villa, et al. Prevention of dialysis hypotension episodes using fuzzy logic control system. *Nephrology dialysis transplantation*, 22(5):1420–1427, 2007.
- [92] Yawen Mao, Su Liu, Jannatun Nahar, Jinfeng Liu, and Feng Ding. Soil moisture regulation of agro-hydrological systems using zone model predictive control. *Computers and Electronics in Agriculture*, 154:239–247, 2018.
- [93] Riccardo Marino. Nonlinear control design: Geometric. *Adaptive and robust*, 1995.
- [94] MF Mason, H Resnik, AS Minot, J Rainey, C Pilcher, and TR Harrison. Mechanism of experimental uremia. *Archives of Internal Medicine*, 60(2):312–336, 1937.
- [95] Jayson CW McAllister. Modeling and control of hemoglobin for anemia man-

- agement in chronic kidney disease. 2017.
- [96] Kevin A. P. McLean and Kim B. McAuley. Mathematical modelling of chemical processes—obtaining the best model predictions and parameter estimates using identifiability and estimability procedures. *The Canadian Journal of Chemical Engineering*, 90(2):351–366, 2012.
 - [97] Kristen Moulton, Eftyhia Helis, Janet Crain, and Melissa Severn. Dialysis programs in Canada: Implementation considerations and funding practices. Technical Report (Environmental scan; no.60), March 2017.
 - [98] M Nordic, S Giove, S Lorenzi, P Marchini, and E Saporiti. A new approach to blood pressure and blood volume modulation during hemodialysis: an adaptive fuzzy control module. *The International journal of artificial organs*, 18(9):513–517, 1995.
 - [99] Jeffrey Carter Olson. *Design and modeling of a portable hemodialysis system*. PhD thesis, Georgia Institute of Technology, 2009.
 - [100] Mette S Olufsen, Johnny T Ottesen, Hien T Tran, Laura M Ellwein, Lewis A Lipsitz, and Vera Novak. Blood pressure and blood flow variation during postural change from sitting to standing: model development and validation. *Journal of applied physiology*, 99(4):1523–1537, 2005.
 - [101] Johnny T Ottesen, Mette S Olufsen, and Jesper K Larsen. *Applied mathematical models in human physiology*. SIAM, 2004.
 - [102] T Petitclerc and B Coevoet. Ionic dialysance and quality control in hemodialysis. *Nephrologie*, 22(5):191–197, 2001.
 - [103] Mauro Pietribiasi, Krassimir Katzarski, Magda Galach, Joanna Stachowska-Pietka, Daniel Schneditz, Bengt Lindholm, and Jacek Waniewski. Kinetics of plasma refilling during hemodialysis sessions with different initial fluid status. *Asaio Journal*, 61(3):350–356, 2015.
 - [104] Mauro Pietribiasi, Jacek Waniewski, Alicja Wójcik-Załuska, Wojciech Załuska, and Bengt Lindholm. Model of fluid and solute shifts during hemodialysis with active transport of sodium and potassium. *PloS one*, 13(12):e0209553, 2018.
 - [105] Mauro Pietribiasi, Jacek Waniewski, Alicja Załuska, Wojciech Załuska, and Bengt Lindholm. Modelling transcapillary transport of fluid and proteins in hemodialysis patients. *PLoS One*, 11(8):e0159748, 2016.
 - [106] Mauro Pietribiasi, Alicja Wójcik-Załuska, Wojciech Załuska, and Jacek Waniewski. Does the plasma refilling coefficient change during hemodialysis sessions? *The International journal of artificial organs*, 41(11):706–713, 2018.
 - [107] David Powell, Jonas Bergstrom, Rastislav Dzurik, Paul Gulyassy, Dean Lockwood, and Lawrence Phillips. Toxins and inhibitors in chronic renal failure. *American Journal of Kidney Diseases*, 7(4):292–299, 1986.
 - [108] Anggoro Primadianto and Chan-Nan Lu. A review on distribution system state estimation. *IEEE Transactions on Power Systems*, 32(5):3875–3883, 2016.

- [109] Leszek Pstras, Karl Thomaseth, Jacek Waniewski, Italo Balzani, and Federico Bellavere. Mathematical modelling of cardiovascular response to the valsalva manoeuvre. *Mathematical medicine and biology: a journal of the IMA*, 34(2):261–292, 2017.
- [110] Leszek Pstras and Jacek Waniewski. *Mathematical Modelling of Haemodialysis: Cardiovascular Response, Body Fluid Shifts, and Solute Kinetics*. 2019.
- [111] Mallikarjun Raghavendra and M Vidya. Functions of kidney & artificial kidneys. *International Journal of Innovative Research*, 1(11):1–5, 2013.
- [112] James B Rawlings and David Q Mayne. Model predictive control theory and design. 2009. *Madison, WI: Nob Hill Publishing, LLC*, 2009.
- [113] James B Rawlings, Stephen M Miller, and Walter R Witkowski. Model identification and control of solution crystallization processes: a review. *Industrial & Engineering Chemistry Research*, 32(7):1275–1296, 1993.
- [114] De Boer Roel W, Karemaker John M, et al. Beat-to-beat variability of heart interval and blood pressure. 1983.
- [115] Claudio Ronco, Alessandra Brendolan, Massimo Milan, Mariapia P Rodeghiero, Monica Zanella, and Giuseppe La Greca. Impact of biofeedback-induced cardiovascular stability on hemodialysis tolerance and efficiency. *Kidney international*, 58(2):800–808, 2000.
- [116] Antonio Santoro and Elena Mancini. Biofeedback in blood volume regulation during hemodialysis. *Nefrología*, 17:56–60, 1997.
- [117] Antonio Santoro, Elena Mancini, Francesco Paolini, Giovanni Cavicchioli, Antonio Bosetto, and Pietro Zucchelli. Blood volume regulation during hemodialysis. *American journal of kidney diseases*, 32(5):739–748, 1998.
- [118] Antonio Santoro, Elena Mancini, Francesco Paolini, Marco Spongano, and Pietro Zucchelli. Automatic control of blood volume trends during hemodialysis. *ASAIO Journal (American Society for Artificial Internal Organs: 1992)*, 40(3):M419–22, 1994.
- [119] Antonio Santoro and Marcora Mandreoli. Chronic renal disease and risk of cardiovascular morbidity-mortality. *Kidney and Blood Pressure Research*, 39(2-3):142–146, 2014.
- [120] John A Sargent. Mathematic modeling of dialysis therapy. *Kidney Int.*, 10:2–10, 1980.
- [121] John A Sargent. Control of dialysis by a single-pool urea model: The national cooperative dialysis study. 1983.
- [122] John A Sargent and Frank A Gotch. Principles and biophysics of hemodialysis. *Replacement of Renal Function by dialysis. Kluwer Academic Publishers: Dordrecht, The Netherlands*, pages 35–102, 1996.
- [123] Klaus Schittkowski. Experimental design tools for ordinary and algebraic dif-

- ferential equations. *Industrial & Engineering Chemistry Research*, 46(26):9137–9147, 2007.
- [124] Reinhard Schmidt, Ottfried Roeher, Heiko Hickstein, and Steffen Korth. Prevention of haemodialysis-induced hypotension by biofeedback control of ultrafiltration and infusion. *Nephrology Dialysis Transplantation*, 16(3):595–603, 2001.
- [125] Daniel Schneditz. Temperature and thermal balance in hemodialysis. In *Seminars in dialysis*, volume 14, pages 357–364. Wiley Online Library, 2001.
- [126] Daniel Schneditz, Dieter Platzter, and John T Daugirdas. A diffusion-adjusted regional blood flow model to predict solute kinetics during haemodialysis. *Nephrology Dialysis Transplantation*, 24(7):2218–2224, 2009.
- [127] Daniel Schneditz, Claudio Ronco, and Nathan Levin. Renal research institute symposium: Temperature control by the blood temperature monitor. In *Seminars in dialysis*, volume 16, pages 477–482. Wiley Online Library, 2003.
- [128] Daniel Schneditz, Johannes Roob, Martina Oswald, Helmuth Pogglitsch, Maximilian Moser, Thomas Kenner, and Ulrich Binswanger. Nature and rate of vascular refilling during hemodialysis and ultrafiltration. *Kidney international*, 42(6):1425–1433, 1992.
- [129] Daniel Schneditz, John C Van Stone, and JOHN T Daugirdas. A regional blood circulation alternative to in-series two compartment urea kinetic modeling. *ASAIO journal (American Society for Artificial Internal Organs: 1992)*, 39(3):M573–7, 1993.
- [130] Ralph Shackman, GD Chisholm, Angela J Holden, and RW Pigott. Urea distribution in the body after haemodialysis. *British medical journal*, 2(5301):355, 1962.
- [131] Lauralee Sherwood. *Introduction to human physiology*. Brooks/Cole, 2013.
- [132] Cecie Starr, Ralph Taggart, Christine Evers, and Lisa Starr. *Biology: The unity and diversity of life*. Cengage Learning, 2015.
- [133] Hans Stigter, Dominique Joubert, and Jaap Molenaar. Observability of complex systems: Finding the gap. *Scientific reports*, 7(1):1–9, 2017.
- [134] Jonathan R Stroud, Matthias Katzfuss, and Christopher K Wikle. A bayesian adaptive ensemble kalman filter for sequential state and parameter estimation. *Monthly Weather Review*, 146(1):373–386, 2018.
- [135] Hilarie Surrena. *Handbook for Brunner and Suddarth’s textbook of medical-surgical nursing*. Lippincott Williams & Wilkins, 2010.
- [136] Marcello Tonelli, Paula Astephen, Pantelis Andreou, Stephen Beed, Peter Lundrigan, and Kailash Jindal. Blood volume monitoring in intermittent hemodialysis for acute renal failure. *Kidney international*, 62(3):1075–1080, 2002.
- [137] M Ursino, M Antonucci, and E Belardinelli. Role of active changes in venous capacity by the carotid baroreflex: analysis with a mathematical model. *American*

- Journal of Physiology-Heart and Circulatory Physiology*, 267(6):H2531–H2546, 1994.
- [138] M Ursino, L Coli, C Brighenti, L Chiari, A De Pascalis, and G Avanzolini. Prediction of solute kinetics, acid-base status, and blood volume changes during profiled hemodialysis. *Annals of biomedical engineering*, 28(2):204–216, 2000.
- [139] Mauro Ursino and Marina Innocenti. Mathematical investigation of some physiological factors involved in hemodialysis hypotension. *Artificial organs*, 21(8):891–902, 1997.
- [140] Mauro Ursino and Marina Innocenti. Modeling arterial hypotension during hemodialysis. *Artificial organs*, 21(8):873–890, 1997.
- [141] Jayaram Valluru and Sachin C Patwardhan. An integrated frequent rto and adaptive nonlinear mpc scheme based on simultaneous bayesian state and parameter estimation. *Industrial & Engineering Chemistry Research*, 58(18):7561–7578, 2019.
- [142] Peter Noel Van Buren and Jula K Inrig. Mechanisms and treatment of intradialytic hypertension. *Blood purification*, 41(1-3):188–193, 2016.
- [143] Peter Noel Van Buren and Jula K Inrig. Special situations: Intradialytic hypertension/chronic hypertension and intradialytic hypotension. In *Seminars in dialysis*, volume 30, pages 545–552. Wiley Online Library, 2017.
- [144] Alejandro F Villaverde. Observability and structural identifiability of nonlinear biological systems. *Complexity*, 2019, 2019.
- [145] Alejandro F Villaverde, Nikolaos Tsiantis, and Julio R Banga. Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models. *Journal of the Royal Society Interface*, 16(156):20190043, 2019.
- [146] Dingbao Wang, Yuguo Chen, and Ximing Cai. State and parameter estimation of hydrologic models using the constrained ensemble kalman filter. *Water resources research*, 45(11), 2009.
- [147] Jacek Waniewski. Mathematical modeling of fluid and solute transport in hemodialysis and peritoneal dialysis. *Journal of Membrane Science*, 274(1-2):24–37, 2006.
- [148] AV Wolf, Donald G Remp, John E Kiley, Gordon D Currie, et al. Artificial kidney function: Kinetics of hemodialysis. *The Journal of clinical investigation*, 30(10):1062–1070, 1951.
- [149] Matthew B Wolf. Whole body acid-base and fluid-electrolyte balance: a mathematical model. *American Journal of Physiology-Renal Physiology*, 305(8):F1118–F1131, 2013.
- [150] Keng Thyee Woo. *Clinical Nephrology*. World Scientific, 1999.
- [151] KDOQI Workgroup. KDOQI clinical practice guidelines for cardiovascular disease in dialysis patients. *Am J Kidney Dis*, 45:S1–S153, 2005.

- [152] Stephen J Wright. Coordinate descent algorithms. *Mathematical Programming*, 151(1):3–34, 2015.
- [153] Ling Xu and Feng Ding. Parameter estimation for control systems based on impulse responses. *International Journal of Control, Automation and Systems*, 15(6):2471–2479, 2017.
- [154] K Zhen Yao, Benjamin M Shaw, Bo Kou, Kim B McAuley, and DW Bacon. Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design. *Polymer Reaction Engineering*, 11(3):563–588, 2003.
- [155] Xunyuan Yin and Jinfeng Liu. State estimation of wastewater treatment plants based on model approximation. *Computers & Chemical Engineering*, 111:79–91, 2018.
- [156] Tianhao Yu and Vivek Dua. Nonlinear model predictive control of haemodialysis. *Computer Aided Chemical Engineering*, 46:1285–1290, 2019.
- [157] Diana Zepeda-Orozco and Raymond Quigley. Dialysis disequilibrium syndrome. *Pediatric nephrology*, 27(12):2205–2211, 2012.
- [158] Mariusz Ziółko, Jacek A Pietrzyk, and Joanna Grabska-Chrzastowska. Accuracy of hemodialysis modeling. *Kidney international*, 57(3):1152–1163, 2000.
- [159] Pietro Zucchelli and Antonio Santoro. Dialysis-induced hypotension: A fresh look at pathophysiology. *Blood purification*, 11(2):85–98, 1993.