Estimation of dialysis treatment efficiency by means of system identification

Johan Larsson

LUND
UNIVERSITY

Department of Automatic Control

Abstract

When treating patients suffering from renal failure with hemodialysis, an obvious point of interest is the actual blood cleaning efficiency of the dialyzer (artificial kidney). This efficiency is called clearance or dialysance.

The method currently used for estimating clearance is based on doing a step-change on the process. Due to the nature of the process, this method is slow and has a relatively large output spread.

This master thesis investigates a new method of finding clearance by means of system identification. The dialyzer is modelled as a discrete-time system, and perturbed by use of a pseudo-binary random sequence. The input/output-data is then fed into an optimal Kalman filter for parameter estimation. The gain and offset of the identified system is directly related to the dialyzer clearance of the treatment.

The method shows promising results, usually converging to good parameters within 15 minutes, and then tracking changes continuously for the rest of the treatment. It also provides better accuracy, with a considerable reduction in spread compared to the old method.
Main obstacles stem from variable time-delays in the system and measurement offsets.

Keywords: System identification, pseudo-random binary sequence, PRBS, Kalman filter, hemodialysis, dialyzer, modelling

3

Acknowledgements

I would like to extend my warmest gratitude to the software team at Gambro, who welcomed me into their fold. Special thanks to my desk neighbour Pierre El-khoury, for your friendship, advice, and for convincing me to start training again.

A warm thank you to Sture Hobro, for your cheery mood, your "sowing", and your dialysis bible. It is always a pleasure to speak with you!

Thank you Eva-Maria, for your proofreading and linguistic suggestions. It is indeed very convenient to have a teacher of English proofread my report, since this completely absolves me from any blame concerning spelling or grammatical errors!

Of course, none of this had been possible were it not for my advisor Jan Sternby. Thank you for offering me this thesis, and for your superb advice and guidance during the project. I thoroughly enjoyed your lectures on control and system theory. It has truly been a pleasure to work with you and learn from you.

Last but not least, a big thank you to my examiner and advisor at the automatic control department, Tore Hägglund. Thank you for introducing me in 2010 to the wonderful field of control theory, and for always having my back and supporting me during the rather long course of this master thesis.

Contents

Variables used

Physical quantities

C_{Bin}	conductivity of blood entering the dialyzer (equals overall blood conductivity)	[mS/cm]
C_{Bout}	conductivity of blood exiting the dialyzer	[mS/cm]
C_B	overall blood conductivty short form for C_{Bin}	[mS/cm]
$\hat{C}_B$	estimated overall blood conductivity from Kalman filter	[mS/cm]
C_{in}	conductivity of dialysate entering the dialyzer	[mS/cm]
C_{out}	conductivity of dialysate exiting the dialyzer	[mS/cm]
C_{in}^*	conductivity set-point for C_{in}	[mS/cm]
D	dialysance, a measure of blood volume cleaned per minute for solute present in dialysate (usually electrolytes). Equal to C in case of urea.	[ml/min]
$\hat{D}$	estimated dialysance from Kalman filter	[ml/min]
Dr	relative dialysance, defined as $\frac{D}{Q_D}$	
$\hat{Dr}$	estimated relative dialysance from Kalman filter	
K	clearance, a measure of blood volume cleaned per minute for solute not present in dialysate	[ml/min]
Q_D	dialysate flow	[ml/min]
Q_B	blood flow	[ml/min]

Model variables

$\varnothing$	time constant	[seconds]
T	hydraulic time-delay in dialyzer model	[seconds]
p	number of a-parameters	
q	number of b-parameters	
h	sampling interval	[seconds]

Kalman filter

$\boldsymbol{\theta}$	parameter vector to be identified
$\hat{\boldsymbol{\theta}}$	estimated parameter vector
ϕ	regressor vector
$\boldsymbol{P}$	error covariance matrix, a measure of the estimated accuracy of the parameter estimates $\hat{\theta}$
$\boldsymbol{P_0}$	initial value of P
$\boldsymbol{K}$	optimal Kalman gain
$\boldsymbol{Q}$	parameter covariance
$\boldsymbol{R}$	observation covariance
$\boldsymbol{tc}$	time-constant of pre-filter

Glossary and abbrevations

clearance: blood volume cleaned per minute for a specific solute

dialysance: blood volume cleaned per minute for a specific solute, when diffusing through a dialyzer

dialyzer: the fundamental component of a hemodialysis system, emulating the kidney function

nephron: the functional unit in the kidneys. It is in the nephrons that the blood is filtered through a series of processes.

semi-permeable membrane: membrane which blocks solutes selectively

solution: a homogeneous mixture consisting of one or more solutes dissolved in a solvent

solute: a substance that creates a solution when dissolved in a solvent

solvent: a substance that dissolves a solute (e.g. water)

GFR: glomerular filtration rate

MSE: mean squared error

PBRS: pseudo-random binary sequence

LFSR: linear-feedback shift register

1

Introduction

1.1 The company

Gambro is a global medical technology company developing, manufacturing and supplying products and therapies for Kidney and Liver dialysis, Myeloma Kidney Therapy, and other extracorporeal therapies for Chronic and Acute patients.

Gambro develops and delivers complete solutions to dialysis clinics and intensive care units around the world.

Gambro was founded in 1964 and has today 8000 employees, 13 production facilities in 9 countries, and sales in more than 90 countries.[1]

---

[1] Gambro. *About gambro*. Oct. 2013. URL: `http://www.gambro.com/en/uk/About-Gambro/Gambro-in-brief/`.

1.2 Problem description

Background

When treating patients with hemodialysis, an obvious parameter of interest is the actual efficiency of the dialysis process. This efficiency is referred to as dialysance (D) or clearance (K). D and K are for most intents and purposes identical, but differ slightly in definition. They both give a measure of blood volume cleaned per minute and are defined in the unit [ml/min], i.e. a flow.

In a hemodialysis system, the blood cleaning is done in a component called a *dialyzer*, which functions as an artificial kidney. See figs. 1.1 and 1.2.
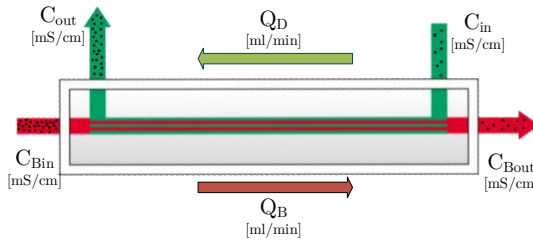

Figure 1.1   Picture of a typical dialyzer

Figure 1.2   Illustration of a dialyzer with physical quantities of importance

The solute removal rate $\left[\frac{\text{mmol}}{\text{min}}\right]$ in the dialyzer is defined as:

$$\text{Removal rate} = \text{blood inflow of solute} - \text{blood outflow of solute}$$
$$= Q_B \cdot (C_{Bin} - C_{Bout})$$

where Q_B[ml/min] is blood flow in the dialyzer and $C_{B_{in}}$, $C_{B_{out}}$ are blood conductivities in and out of the dialyzer respectively. Conductivity [mS/ml] is directly proportional to concentration [mmol/ml] in dialysance calculations and is the main unit of input/output signals in this master thesis.[2]

---

[2] W. H. Hörl et al. *Replacement of renal function by dialysis.* 5th ed. Kluwer academic publishers, 2004, p. 347.

Due to hygiene constraints, conductivity measurements can not be made directly on the blood. However, the solutes removed from the blood are transferred to the dialysate (or dialysis fluid), which is available for measurement. Therefore, the following equal relationship proves more useful:

Removal rate = dialysate outflow of solute − dialysate inflow of solute

$$\tag{1.1}$$

$$= Q_D \cdot (C_{out} - C_{in}) \tag{1.2}$$

$$= D \cdot (C_{Bin} - C_{in}) \tag{1.3}$$

Where C_{in} is the dialysate conductivity at the inlet of the dialyzer and C_{out} the dialysate conductivity at the outlet of the dialyzer.

In eq. (1.3), D acts as a gain over the concentration difference between C_{Bin} and C_{in}, motivating its use as an indicator of efficiency. Reformulating the equations to have D on the left-hand side we get[3]

$$D = Q_D \cdot \frac{C_{out} - C_{in}}{C_B - C_{in}} \tag{1.4}$$

This yields an equation with two unknowns, D and C_B, both of which are of interest to find.

In the current lines of dialysis machines from Gambro, this is simply solved by providing two sets of equations, through making a step change on the process and measuring the stationary values (see fig. 1.3 for an example). This method of dialysance estimation is called Diascan®.

While this method gives a satisfactory estimation of the dialysance, it has some drawbacks:

- can take up to 15 minutes to get a value

- the estimated value has a large variance

- the method is not continuous

- the method can not be used often (usually two times per hour), due to various events in the machine

Purpose of master thesis

For the future generations of dialysis machines, there is a desire to replace the Diascan® step-change method by a continuous method. The idea is to set up a discrete-time model for how the dialyzer outlet conductivity (C_{out}) changes as a function of the inlet conductivity (C_{in}) and identify the parameters of this model, using a recursive least squares Kalman filter.

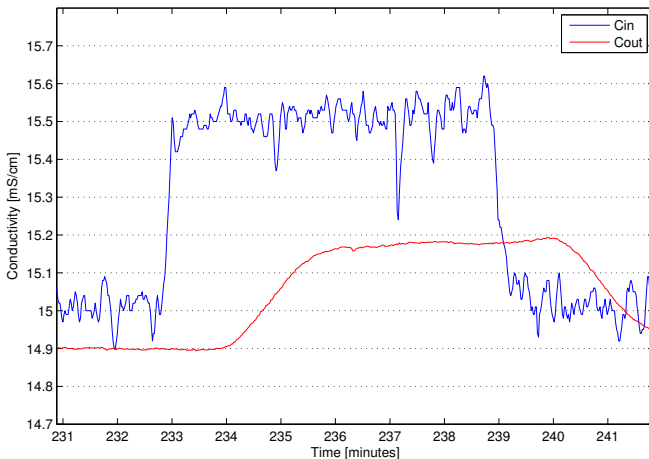[3] Hörl et al., *Replacement of renal function by dialysis*, p. 415.

Figure 1.3 An example of an actual Diascan® step. A step change on the inlet conductivity C_{in} is made to generate a second set of stationary values. The step lasts for over six minutes.

Expressed in the $\mathcal{Z}$-transform the model looks like

$$A(z^{-1})C_{out} = B(z^{-1})C_{in} + c \tag{1.5}$$

with

$$A(z^{-1}) = 1 - a_1 z^{-1} - ... - a_n z^{-n} \tag{1.6}$$

$$B(z^{-1}) = b_0 z^{-T} + b_1 z^{-T-1} - ... - b_m z^{-T-m} \tag{1.7}$$

T is a hydraulic time-delay which varies depending on choice of dialyzer, the flows Q_B and Q_D. Thus, it needs to be investigated and determined before identification is possible.

The gain and offset of this discrete-time model is directly related to D and C_B. Therefore, if the parameters in eq. (1.5) can be identified, D and C_B can be calculated.

The purpose of this master thesis is to investigate the feasibility of this idea, by conducting experiments on an actual hemodialysis machine and processing the data in MATLAB®.

2

The need for dialysis

This chapter will briefly describe the concept of renal failure, and why dialysis treatment is necessary. This overview is based on internal teaching material provided by Gambro.[1]

2.1 The blood

The primary function of the blood is to work as a transportation network in the body, such as providing oxygen (picked up in the lungs) and nutrients (picked up in the gastrointestinal system) to the cells. The blood is also responsible for extraction of the waste products from the metabolism. Because of this, all substances that enter or exit the body can be found in the blood.

2.2 Some waste products of metabolism

Carbon dioxide, CO_2

Molecular weight: 44
By-product when fuel from food such as proteins, carbohydrates and fats are converted into energy. Produced in large quantities.[2]

Creatinine

Molecular weight: 113
By-product of muscular metabolism. Produced from creatine, which is an important molecule for energy production in muscles. Rising creatinine levels

---

[1] Gambro. *Gambro BASICS 1, Njurfunktion*. Internal teaching material. Jan. 2002.
[2] Gambro, *Gambro BASICS 1, Njurfunktion*, p. 9.

is very useful as an indicator of impaired kidney functions, as they may occur before the patient notices any symptoms.[3,4]

Urea

Molecular weight: 60
A nitrogen-containing substance produced when proteins are metabolised. Accumulates rapidly during renal failure.

2.3   Homeostasis

> *The coordinated physiological processes which maintain most of the steady states in the organism are so complex and so peculiar to living beings - involving, as they may, the brain and nerves, the heart, lungs, kidneys and spleen, all working cooperatively - that I have suggested a special designation for these states, homeostasis.*
>
> *The word does not imply something set and immobile, a stagnation. It means a condition - a condition wh2ich may vary, but which is relatively constant.*
>
> **- Walter B. Cannon, 1929**[5]

To maintain proper function the cells of the human body need a stable environment. This stability needs to be maintained at all times, and a self-regulating system doing this is defined to be in a state of *homeostasis*. In the human body, homeostasis is achieved by continuous regulation of properties such as body temperature, electrolyte concentrations, water levels and pH levels.

To maintain this state of homeostasis, there has to be balance in what enters and exits the body, in sense that the waste products of metabolism must be removed from the body.
As mentioned earlier, CO_2 is one such prominent waste product, and it is a well-known fact that it is excreted by our lungs through exhalation. However, most waste products do not go the path of the lungs, but need another way to be excreted.[6]

---

[3] MedicineNet. *Definition of creatinine*. Oct. 2013. URL: http://www.medterms.com/script/main/art.asp?articlekey=12550.

[4] Gambro, *Gambro BASICS 1, Njurfunktion*, p. 11.

[5] Today in science history. *Science quotes by Walter Bradford Cannon*. Jan. 2014. URL: http://todayinsci.com/C/Cannon_Walter/CannonWalter-Quotations.htm.

[6] Gambro, *Gambro BASICS 1, Njurfunktion*, p. 9.

2.4 The kidneys

The main route for waste products to be excreted go through the kidneys, which are supplied with unclean blood by the renal arteries.

Through a series of filtering, absorption and secretion processes, the waste products of metabolism and superfluous water are separated from the blood. The main functional unit in the kidney is called the nephron, and is where the urine is formed. The amount of nephrons in each kidney are estimated to about one million.[7]

Description of the renal filtering process

A simplified description of the formation of urine follows on the next page. A more advanced treatment of this complex physiological function is not relevant for the scope of this thesis.
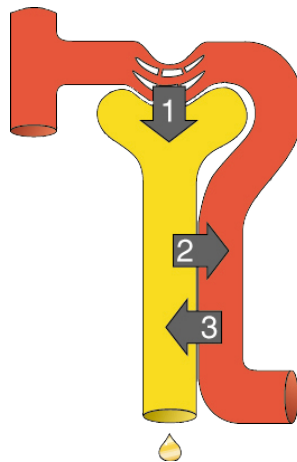

Figure 2.1 Illustration of urine formation in the nephrons. (Picture courtesy of Gambro)

1. **Glomerular filtering**: blood is channelled from the renal artery into the nephron. Through a largely non-selective filtering process a primary urine is formed from water and small molecules. Larger molecules (e.g. proteins) stay in the blood.

2. **Re-absorption**: the primary urine is led through the **tubules** (the yellow area in the picture). In the tubules the majority of the fluid (up

---

[7] Gambro, *Gambro BASICS 1, Njurfunktion*, p. 13.

to 99%) and small substances are reabsorbed into the blood. Contrary to the glomerular filtering, this process is selective:

- The waste-product urea is only reabsorbed in small quantities.
- Creatinine is not reabsorbed at all, which is a fact being utilized in evaluation of renal performance (discussed later).
- Electrolytes are absorbed in varying quantities to maintain electrolyte balance and homeostasis
- Substances useful to the body (e.g. glucose and vitamins) are normally completely reabsorbed to the blood.

3. **Secretion**: some specific substances (often with high molecular mass) are transported from the blood back to the urine.

The end product of the filtering process is urine, a slightly acidic liquid containing roughly 96% water, 2% urea and 2% other substances.

2.5   Clearance — the efficiency of the kidneys

Clearance (K, unit: [ml/min]) is a measurement of the blood volume that the kidneys completely clean from waste per time unit.[8] The mathematical expression is as follows:

$$K = \frac{\text{Removal rate[mmol/min]}}{\text{Concentration[mmol/ml]}} \tag{2.1}$$

The most common way of studying a patient's renal function is by estimating the GFR (glomerular filtration rate). This is done by study of creatinine clearance. As mentioned above, creatinine is filtered only through glomerular filtering, and is not re-absorbed in the tubules. This means that measuring creatinine clearance gives a direct measure of the GFR and therefore the renal function. Well-performing kidneys have a creatinine clearance of about 120 ml/min.[9]
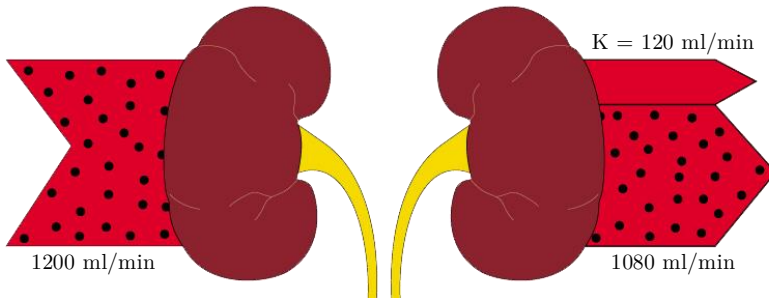

Figure 2.2   Graphical depiction of clearance. Blood flows through the kidneys at 1200 ml/min, and 120 ml/min of this leaves completely cleaned, corresponding to the clearance K. (Picture courtesy of Gambro)

---

[8] Hörl et al., *Replacement of renal function by dialysis*, p. 602.
[9] Gambro, *Gambro BASICS 1, Njurfunktion*, p. 17.

2.6 Renal failure and treatment

Concerning renal failure, a distinction is made between **acute kidney failure** and **chronic kidney disease**. Acute failure may only cause a temporary problem, and can be possible to cure with a limited treatment period. Causes for acute failure can be temporary decreases in blood flow to the kidneys or direct kidney trauma.

Chronic kidney disease can develop slowly or be a direct consequence of acute kidney failure. In this case, the kidneys are irreversibly damaged and will never be able to reach full functionality again. Treatment is generally aimed at slowing the development of the chronic kidney disease. It is common to prescribe diets with low protein contents, since this would reduce production of metabolic waste products and renal strain. The patient also needs to reduce water intake.

However, if the GFR decreases below 5 ml/min (see fig. 2.3), this is referred to as **advanced chronic kidney disease**. Active treatment is necessary to ensure survival of these patients, either by means of kidney transplantation or dialysis.[10]
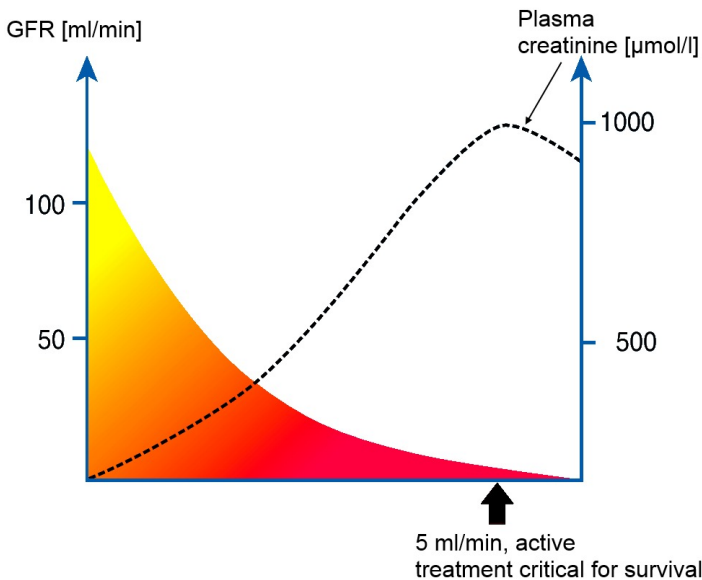

Figure 2.3   Creatinine levels in blood plasma vs. GFR. (Picture courtesy of Gambro)

---

[10] Gambro, *Gambro BASICS 1, Njurfunktion*, p. 19.

3

The hemodialysis system

3.1 Principles of hemodialysis

Overview of the dialyzer

The dialyzer is essentially the core of the hemodialysis system, since it is in the dialyzer that the blood is actually cleaned.

The dialyzer is fundamentally divided into two opposing chambers by a *semi-permeable membrane*, (see **??**). The contaminated blood flows in one of these chambers, and a solution called dialysate flows in the other. By utilizing the principles of *diffusion* and *ultra filtration*, the blood is cleaned and superfluous fluid is extracted.
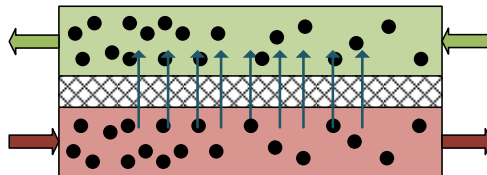

Figure 3.1 Fundamental function of the dialyzer. A semi-permeable membrane in the middle separates the blood and dialysate. The blood solute concentration decreases as it flows through the dialyzer, as the solutes cross the membrane to the dialysate.

Diffusion

Definition: transport of solvents caused by a concentration gradient.

Diffusion is the strive of freely moving molecules to distribute evenly in a solvent or in a gas. In the context of solutions (e.g. blood) this results in solutes "wandering" from an area with high concentration to an area with lower concentration, until equilibrium is reached (see fig. 3.2). The magnitude of this movement J (diffusive flux) is directly proportional to the concentration gradient $\frac{\partial \phi}{\partial x}$ and a diffusion constant D according to Fick's first law of diffusion:[1]

$$J = -D\frac{\partial \phi}{\partial x} \tag{3.1}$$

The dialysate in the dialyzer contains no waste products in order to maximize the concentration gradient ($\frac{\partial \phi}{\partial x}$ in eq. (3.1)), so that the waste products in the blood will diffuse to the dialysate. Due to superfluous concentrations of electrolytes in the patient, the dialysate is also mixed in such a way that it contains optimal levels of these concentrations, so that the patient's concentration is normalized.[2] See fig. 3.3 for a graphical representation of this.
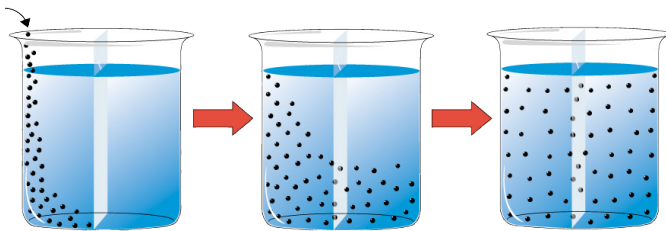

Figure 3.2   The principle of diffusion. The solutes are poured into the fluid and gradually diffuse to form a solution. (Picture courtesy of Gambro)

Semi-permeable membrane and osmosis

The semi-permeable membrane is constructed in such a way that it lets through small molecules easily, but acts as an obstruction for larger molecules. Medium-sized molecules can pass through but at a slower rate.

A consequence of using a semi-permeable membrane is the physical phenomenon of *osmosis*. Assuming that the two solutions differ in composition, there will be one side with higher concentration of large substances, which can not diffuse through the membrane. This is equivalent to having a lower

---

[1] Hörl et al., *Replacement of renal function by dialysis*, p. 122.

[2] Gambro. *Gambro BASICS 3, Principer för Hemodialys*. Internal teaching material. Jan. 2002, p. 9.
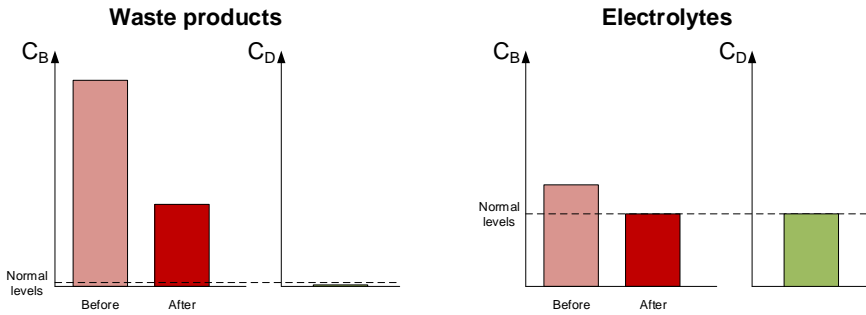
Figure 3.3   Concentrations levels of blood before and after treatment vs. dialysate concentration levels. The blood can not be completely cleaned from waste products but C_D contains none of these in order to maximize the concentration gradient. (Picture courtesy of Gambro)

water concentration. Because of this, equilibrium can only be achieved if the solvent (water) wanders from the opposing side. The solvent will continue to wander until the hydrostatic pressure of the resulting water pillar equals the osmotic pressure.[3]
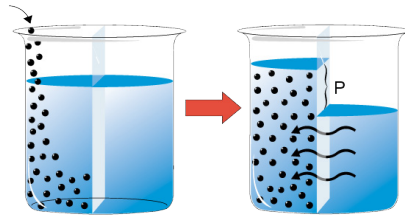

Figure 3.4   Since the molecules on the left side can not permeate through the semi-permeable membrane, water instead flows from the right compartment in an attempt to equalize the concentration levels. The process halts when the hydrostatic pressure P from the resulting pillar equals the osmotic pressure. (Picture courtesy of Gambro)

---

[3] Gambro. *Gambro BASICS 2, Transportprinciper*. Internal teaching material. Jan. 2002, pp. 10-11.

Ultra filtration

Definition: transport of fluid through a membrane caused by a pressure gradient.[4]
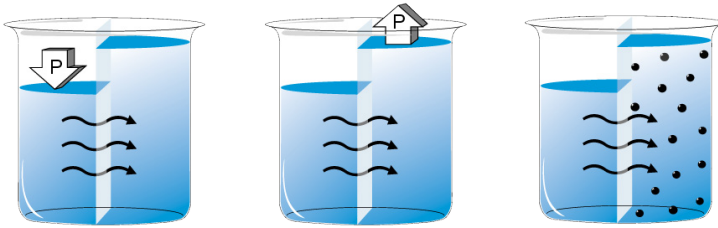

Figure 3.5 Three different methods of achieving ultra filtration. The first is through applying positive pressure (pushing), the second by applying negative pressure (suction) and the third is by means of osmosis. (Picture courtesy of Gambro)

In contrast to diffusion, which concerns the movements of solutes, ultra filtration concerns movement of the fluid (or solvent) itself, and specifically through a semi-permeable membrane. As described above, a solvent may flow through a membrane as a result of osmotic pressure. By applying external pressure to the fluid, it is possible to alter the pressure gradient and cause the fluid to wander in a desired direction (see fig. 3.5).

---

[4] Gambro, *Gambro BASICS 2, Transportprinciper*, pp. 12-13.

3.2   Dialyzer clearance

The effect of increasing Q_B and Q_D

As mentioned in section 3.1, small molecules such as urea and creatinine diffuse easily through a semi-permeable membrane. This means that the clearance of these solutes scales up greatly with increasing blood flow (Q_B), until reaching a saturating point. This does not hold true for larger molecules, which can not permeate the membrane easily. An increase of blood flow will have a very small impact on the clearance of these solutes (see fig. 3.6).
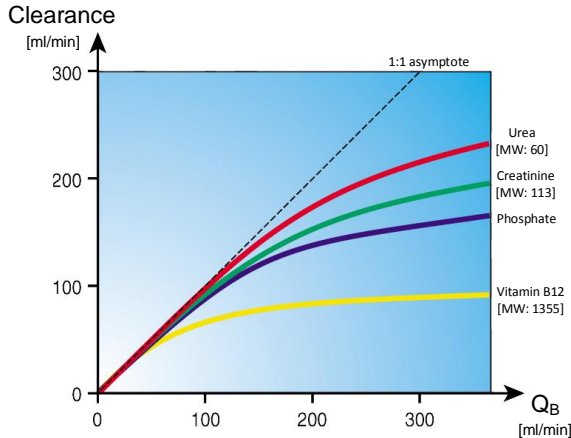

Figure 3.6   Clearance vs. Q_B (blood flow) for different solutes. For increasing molecular weight (MW), the clearance falls off earlier vs. blood flow. A 1:1 ratio asymptote is included for reference. (Picture courtesy of Gambro)

The same general principle holds for the dialysate flow Q_D, which is responsible for disposing of the solutes. Thus Q_D must not bottleneck the blood flow. A general consensus is that Q_D should be roughly double that of Q_B. The most common value for Q_D is 500 ml/min, which in practice is enough for Q_B in the range of 300 to 350 ml/min.[5]

---

[5] Gambro, *Gambro BASICS 3, Principer för Hemodialys*, p. 21.

Estimating urea clearance through electrolyte clearance

NaCl and urea diffusion constants are almost equal, and therefore NaCl clearance is a good indicator of urea clearance.[6] Dialysate also primarily consists of NaCl, as seen in fig. 3.7.
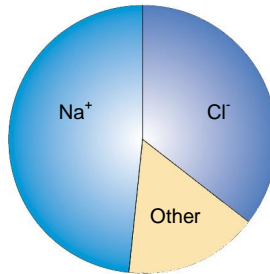

Figure 3.7 Composition of electrolytes in dialysate. Note the large contents of Na^+ and Cl^-. (Picture courtesy of Gambro)

Electrolyte concentration in the dialysate is directly proportional to the conductivity of the dialysate, which is a physical quantity that can be measured accurately with conductivity cells. By measuring variations in conductivity between the inlet and outlet of the dialyzer, it is therefore possible to get an estimation of the urea clearance. This is also the main method used today in clearance estimations.[7]

The conductivity readings at the inlet and outlet of the dialyzer (C_{in} and C_{out} respectively) are the primary signals used in system identification in this master thesis.

Deriving the dialyzer dialysance and clearance

We repeat the mathematical expression for renal clearance:

$$K = \frac{\text{Removal rate}}{\text{Concentration}} \tag{3.2}$$

If the solute to be dialysed is present in the dialysate, which is the case for electrolytes, we do not use the term clearance but *dialysance*, D. The removal rate is then defined as:

Removal rate = dialysate outflow of solute − dialysate inflow of solute

---

[6] H. Polaschegg. "Automatic, noninvasive intradialytic clearance measurement". *The International Journal of Artificial Organs* **16**:4 (1993), p. 1.

[7] Hörl et al., *Replacement of renal function by dialysis*, p. 347.

Which gives the expression for D:

$$D = Q_D \cdot \frac{C_{out} - C_{in}}{C_{Bin} - C_{in}} \tag{3.3}$$

The denominator is the concentration difference between incoming blood C_{Bin} and dialysate C_{in}. See fig. 3.8 for a graphical description of the parameters.
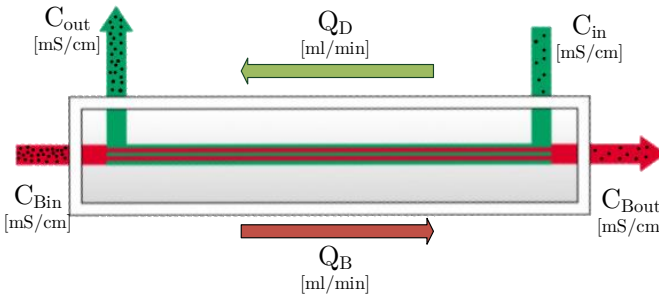

Figure 3.8   The different parameters of interest in the dialyzer. (Picture courtesy of Gambro)

Clearance   Clearance is in fact a special case of dialysance. If the solute to be dialysed is not present in the dialysate (e.g. urea), $C_{in} = 0$ and eq. (3.3) becomes:

$$K = Q_D \frac{C_{out}}{C_{Bin}} \tag{3.4}$$

Because the dialysance of NaCl is used to model urea clearance, these two terms are sometimes used interchangeably, which might cause some confusion.[8]

> **Note**
>
> These formulas do not take ultra filtration into consideration, since it is not used in the experiments in this thesis.

---

[8] Hörl et al., *Replacement of renal function by dialysis*, pp. 280-281.

3.3   The complete hemodialysis system

It is quite natural to divide the hemodialysis system into the *blood circuit* and the *fluid circuit*, with the intermediate dialyzer connecting the two, see fig. 3.9.
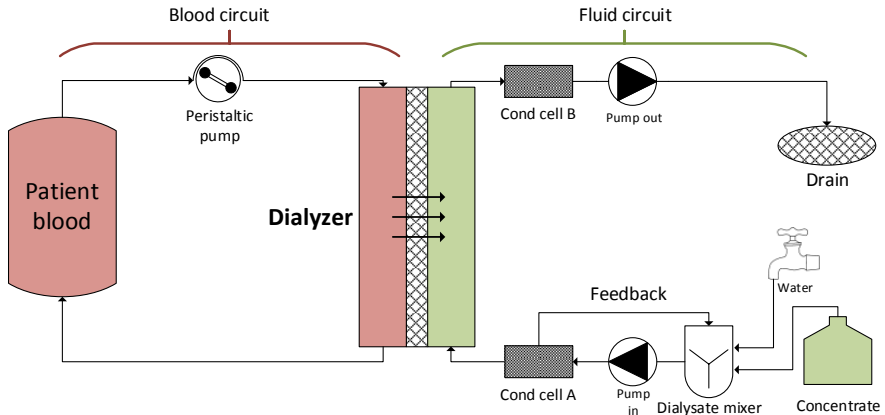

Figure 3.9   Simplified schematic of the hemodialysis machine. The blood circuit is to the left of the dialyzer and the dialysate circuit on the right.

Blood circuit

In the blood circuit, it is crucial that the patient's blood is not contaminated. Hence, all the piping is disposable and one-time-use only. The same principle applies to the dialyzer, which is discarded after use. To achieve flow in the blood circuit, a peristaltic pump is used. This kind of pump is not in direct contact with the blood, but applies external pressure to the piping so that the fluid is pushed forward (see fig. 3.10).
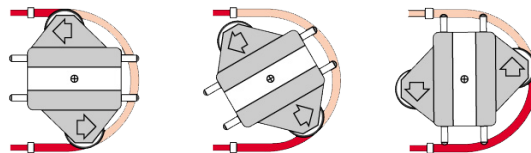

Figure 3.10   The peristaltic pump rotates counter-clockwise and pushes the blood forward. (Picture courtesy of Gambro)

Fluid circuit

The fluid circuit is a far more complex construction, and only the parts directly relevant to the thesis are described here. The first step on the fluid side is the dialysate mixer: the user sets a reference conductivity C_{in}^* which C_{in} tracks through a control system, utilizing conductivity cell `A` after the mixer as feedback. The dialysate then flows through the dialyzer, extracting waste and electrolytes from the blood. After this, the dialysate passes conductivity cell `B`, which is used for the current Diascan® feature and is used as out signal for the model in this thesis. The used dialysate is then discarded through the drain.

3.4   Example of a full run

In fig. 3.11 is a graph containing values for how a typical treatment session can look in concept . In this case the patient is actually a wheelie bin containing 40 litres of dialysis fluid, but electrolyte-wise it emulates a patient quite well.

C_{in} is in this case set to a relatively constant set-point, at about $14.1\,\mathrm{mS/cm}$. Due to the blood having a conductivity of $C_B = 14.7\,\mathrm{mS/cm} > C_{in}$, electrolytes will be extracted from the blood through the membrane and into the dialysate. Because of the electrolyte increase, C_{out} will have a higher conductivity than C_{in}. Since electrolytes are continuously drawn from the blood, C_B will gradually decrease during the course of the treatment. In the example, C_B goes from $14.7\,\mathrm{mS/cm}$ to $14.3\,\mathrm{mS/cm}$.
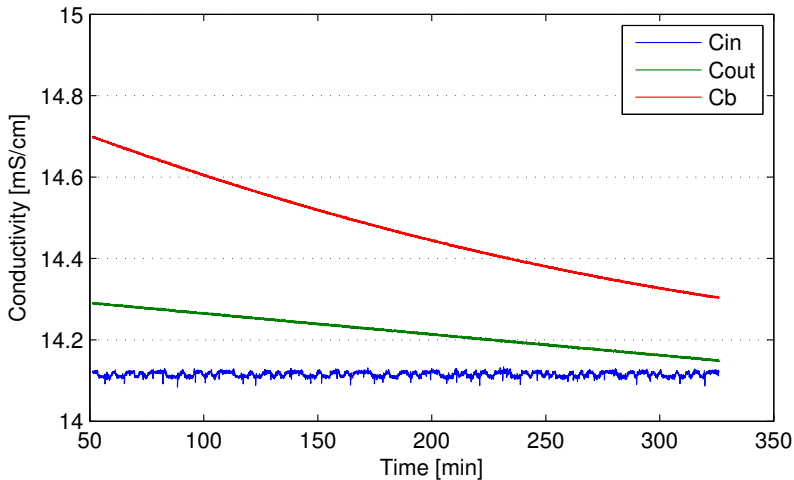

Figure 3.11   Example conductivity values for a typical treatment.

4

Method

4.1 First look at system characteristics

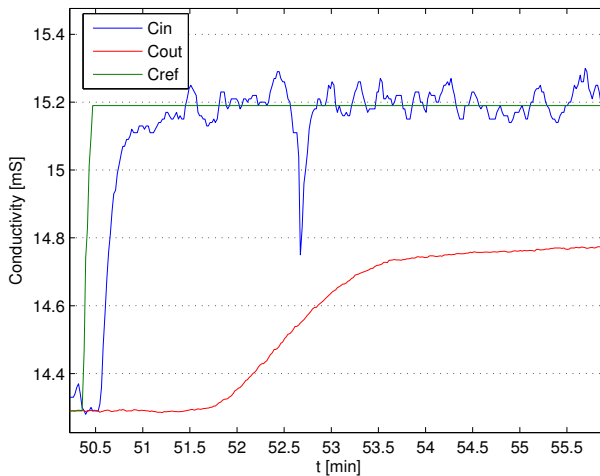To give a general idea of the characteristics of the system, a step response is presented in fig. 4.1.

Figure 4.1   Step response at $Q_D = 500\,\text{ml/min}$. C_{in} and C_{out} are initially stable at a conductivity of $14.29\,\text{mS/cm}$. At $t = 50.5\,\text{min}$, C_{in}^* is increased to $15.19\,\text{mS/cm}$. After a slight delay the dialysate mixing controller kicks in, and C_{in} finally reaches C_{in}^* at roughly $t = 51.5\,\text{min}$.

Input-output time delay

The effect of the time delay in the system is obvious, considering how long it takes for C_{out} to start tracking the change in C_{in}. In this particular case

C_{out} starts to rise at about $t = 51.75\,\text{min}$, which is more than 60 seconds after the change in C_{in}.

Time constant

The time constant τ of the system is in the magnitude of minutes, as is readily apparent in the graph.

C_{in} noise with respect to C_{out}

C_{in} is of a quite noisy nature, with an especially noticeable drop at $t = 52.75\,\text{min}$. This is due to the mixer pumps working in a pulsed manner. In the long run, the mean value of the conductivity is very close to the reference value.

Because of the physical low-pass dynamics of the dialyzer, the conductivity fluctuations on C_{in} are averaged out which leads to C_{out} looking very clean in comparison. The noisy behaviour of C_{in} implies that some pre-filtering of the data before it is processed might be necessary.

Conclusion of first look at system

The system of concern in this thesis is a relatively slow first-order process, with both time delays and time constants ranging in minutes. Pre-filtering could be necessary due to the noisy conductivity of C_{in} .

4.2   Mathematical model

The continuous-time mathematical model for transfer of an ionized substance through the membrane of a dialyzer is as follows:[1]

$$\frac{dC_{out}(t)}{dt} = \frac{1}{\varnothing}[-C_{out}(t) + (1 - Dr) \cdot C_{in}(t - T) + Dr \cdot C_B(t)] \qquad (4.1)$$

- $\varnothing$ is the system time constant, proportional to patient fluid volume

- T is the hydraulic time delay

- C_{in} is the dialysate inlet conductivity

- C_{out} is the dialysate outlet conductivity

- C_B is the patient's overall blood conductivity

- $Dr = \frac{D}{Q_D}$ is the unknown *relative dialysance*

---

[1] A. Bosetto and F. Paolini. "Method for determining a parameter indicative of an extracorporeal blood treatment". Pat. US 6691040 B2. Feb. 2004.

In stationarity the equation reduces to:

$$C_{out} = (1 - Dr) \cdot C_{in} + Dr \cdot C_B \tag{4.2}$$

In this equation, the factor $(1 - Dr)$ acts as a gain from C_{in} to C_{out}, and the term $Dr \cdot C_B$ an offset. It is of interest to find the values of Dr and C_B.

4.3   Current method, Diascan®

The current method of finding the dialysance and solving eq. (4.2) is quite simple in its approach. Let's begin by restructuring the equation so that Dr is on the left-hand side:

$$Dr = \frac{C_{out} - C_{in}}{C_B - C_{in}} \tag{4.3}$$

This equation is equal to that of eq. (3.3), with the only difference being the new variable $Dr = \frac{D}{Q_D}$. The equation contains two unknowns: Dr and C_B. Basic linear algebra gives that two equations are required to calculate both of these unknowns. The current Diascan® feature solves this by making a step-change in the dialysate inlet conductivity C_{in} of $50\,\text{mS/cm}$. With two sets of stationary values the relative dialysance can be calculated as:[2]

$$Dr = 1 - \frac{C_{out}(2) - C_{out}(1)}{C_{in}(2) - C_{in}(1)} \tag{4.4}$$

Note that we have made the assumption that C_B is constant at the time of the measurements, which is feasible due to the large time constant of the patient's fluid pool.[3]

Method drawbacks

While Diascan® is appealing in its simplicity in implementation and calculation, it does have some disadvantages. The method is very slow, because of the long time constants and delays in the system. Also, due to the noisy behaviour of C_{in}, the results can vary greatly. See fig. 4.2 for a graph of a Diascan® step. The output from the Diascan® feature can be seen in fig. 4.3.

---

[2] Polaschegg, "Automatic, noninvasive intradialytic clearance measurement", p. 186.
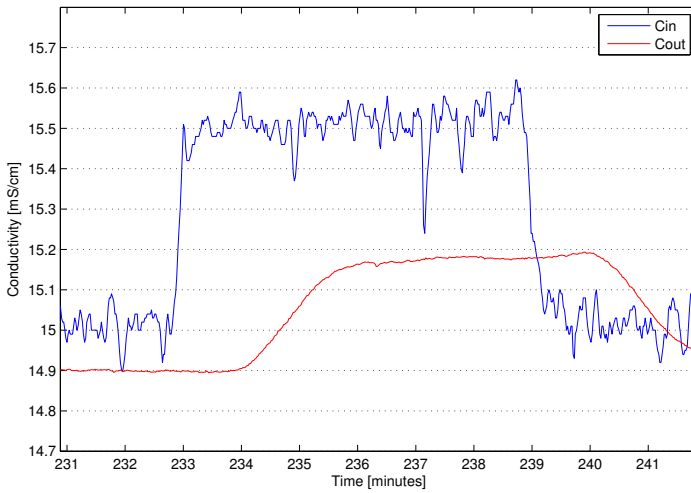[3] Polaschegg, "Automatic, noninvasive intradialytic clearance measurement", p. 188.

Figure 4.2 An example of an actual Diascan® step. The step lasts for over six minutes.
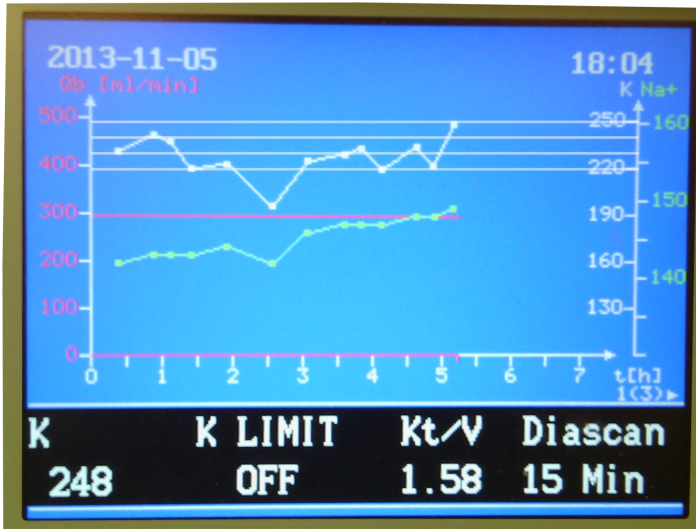

Figure 4.3 The white line shows Diascan® output during an experiment. Note the large deviations, with clearance ranging from 190 ml/min to 250 ml/min.

4.4 New method: system identification

The model

We start by setting up a discrete dynamic model of how the outlet conductivity C_{out} changes as a function of the inlet conductivity C_{in}:

$$C_{out_{k+1}} = a_1 C_{out_k} + ... + a_n C_{out_{k-n}} + b_1 C_{in_{k-T}} + ... + b_m C_{in_{k-T-m}} + c \quad (4.5)$$

Expressed in the $\mathcal{Z}$-transform we get:

$$A(z^{-1})C_{out} = B(z^{-1})C_{in} + c \quad (4.6)$$

with

$$A(z^{-1}) = 1 - a_1 z^{-1} - ... - a_n z^{-n} \quad (4.7)$$
$$B(z^{-1}) = b_0 z^{-T} + b_1 z^{-T-1} - ... - b_m z^{-T-m} \quad (4.8)$$

Rearranged:

$$C_{out} = \frac{B(z^{-1})}{A(z^{-1})} C_{in} + \frac{c}{A(z^{-1})} \quad (4.9)$$

Now, let us compare the static discrete-time model to the stationary equation from eq. (4.2):

$$C_{out} = \frac{B(1)}{A(1)} C_{in} + \frac{1}{A(1)} c \quad (4.10)$$

$$C_{out} = (1 - Dr)C_{in} + Dr \cdot C_B \quad (4.11)$$

From this we can identify the relationships

$$\frac{B(1)}{A(1)} = (1 - Dr) \quad (4.12)$$

$$\frac{c}{A(1)} = Dr \cdot C_B \quad (4.13)$$

From this, we see that if we are able to identify the coefficients in the discrete-time model, we can calculate the corresponding values of Dr and C_B.

> **Note**
>
> It's important to remember that we are identifying the *relative dialysance* $Dr = \frac{D}{Q_D}$ using this model. D is easy to calculate from this since Q_D is always known!

The general Kalman filter

The general **state-predicting** Kalman filter[4] is based on the following state-space model:

$$x_{k+1} = \Phi x_k + \Gamma u_k + v_k$$
$$y_k = C x_k + D u_k + e_k$$

(4.14)

where v_k and e_k is zero-mean Gaussian noise with $cov\{v_k\} = Q$ and $cov\{e_k\} = R$ respectively.

The noisy input-output data $\{y_k\}$ and $\{u_k\}$ is the only data available, and the goal is optimal estimation of the state vector x_k.

This can be solved by optimization of the criterion

$$J(\hat{x}) = \mathbb{E}\{(x_{k+1} - \hat{x}_{k+1|k})^2\}$$

(4.15)

In other words, minimizing the MSE (mean squared error) of the expected value between the predicted state and actual state.

In this case, the dynamics of the system (Φ and Γ) are given a priori. However, for this thesis, identification of an unknown parameter vector θ is of interest.

Kalman filter for system identification

By reformulating the model in eq. (4.14) we get a suitable model for system identification:

$$\theta_{k+1} = \theta_k + v_k$$
$$y_k = \phi_k^T \theta_k + e_k$$

(4.16)

where

$$\phi_k = \begin{pmatrix} C_{out}(k-1) & \cdots & C_{out}(k-n) & C_{in}(k-T) & \cdots & C_{in}(k-T-m) & 1 \end{pmatrix}$$

is the regressor vector and

$$\theta_k = \begin{pmatrix} a_1 & \ldots & a_n & b_0 & \ldots & b_m & c \end{pmatrix}$$

the parameter vector.

v_k and e_k are zero-mean Gaussian noise with $cov\{v_k\} = Q$ and $cov\{e_k\} = R$ respectively.

Q and R are both important design parameters, and the diagonal elements of these should match the variations of the parameters in θ and the observation noise e_k. In the case of our model in eq. (4.6), it is certain that the offset parameter c should be allowed to vary (since c is related to C_B which changes during treatment).

---

[4] R. Johansson. *System modeling & identification.* Prentice Hall, 1993. ISBN: 0-13-482308-7, p. 120.

It is possible that the parameters in $A(z^{-1})$ and $B(z^{-1})$ can be static, since these relate to the dialysance D, which should not change notably during treatment. This means that the diagonal elements of Q can be set to 0 for these parameters.

The optimization criterion is now formulated as minimizing the MSE of the parameter estimation accordingly:

$$J(\hat{\theta}) = \mathbb{E}\{(\theta_{k+1} - \hat{\theta}_{k+1|k})^2\} = P_k \tag{4.17}$$

where P_k is the error covariance matrix, which is a measure of the estimated accuracy of the parameter estimates. The error term is:

$$\varepsilon^2(\hat{\theta}) = \{(y_k - \phi_k^T \hat{\theta}_{k+1|k})^2\} \tag{4.18}$$

The recursive update equations for the filter are

$$\hat{\theta}_{k+1} = \hat{\theta}_k + K_k(y_k - \phi_k^T \hat{\theta}_k) \tag{4.19}$$

$$K_k = P_k \phi_k (R + \phi_k^T P_k \phi_k)^{-1} \tag{4.20}$$

$$P_{k+1} = P_k + Q - K_k \phi_k^T P_k \tag{4.21}$$

In eq. (4.20), K_k denotes the optimal Kalman gain, which in every step minimizes the optimization criterion $J(\hat{\theta})$ in eq. (4.17). A very important parameter in tuning of the Kalman filter is the initial value of the P matrix: P_0. Setting this to large values might increase the initial speed of convergence when the algorithm starts.

An excellent derivation of the Kalman filter can be found online at `http://mpdc.mae.cornell.edu/Courses/UQ/kf1.pdf`.[5]

4.5 Excitation of the system

The below discussion is entirely based on Rolf Johansson's *System modeling & identification*, pp. 173-180, unless otherwise stated.

The success of the identification is very much dependent on how the system is perturbed. In general, the input signal $U(s)$ (C_{in} in the context of this thesis) should be chosen so that there is much input energy for frequencies around the bandwidth frequency of the investigated system $G(s)$ (the dialyzer).

Pseudo-random binary sequences

Considering that the system is ill-fitted for sinusoidal signals (due in part to the pulsed and noisy behaviour of C_{in}), the excitation method chosen

---

[5] C. U. Tony Lacey. *Tutorial: the kalman filter*. Oct. 2013. URL: `http://mpdc.mae.cornell.edu/Courses/UQ/kf1.pdf`.

for this thesis is PRBS (pseudo-random binary sequence). A PRBS is easy to implement through linear feedback shift registers and has large energy content in a large frequency range.[6]

An example of a PRBS polynomial is:

$$P(z^{-1}) = 1 \oplus z^{-3} \oplus z^{-5} \tag{4.22}$$

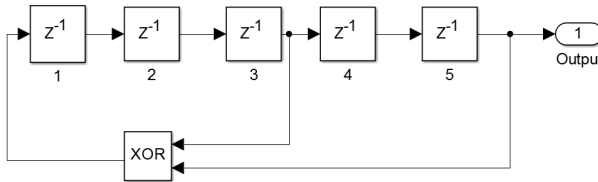Which has a shift register realization as seen in fig. 4.4.

Figure 4.4   Linear feedback shift register implementation to generate PRBS described in eq. (4.22).

The PRBS has a period of $2^n - 1$, where n denotes the number of shift registers in the circuit. So in the case of the PRBS generated in eq. (4.22), the period is

$$2^n - 1 = 2^5 - 1 = 31$$

The output from this PRBS for two periods can be seen in fig. 4.5.
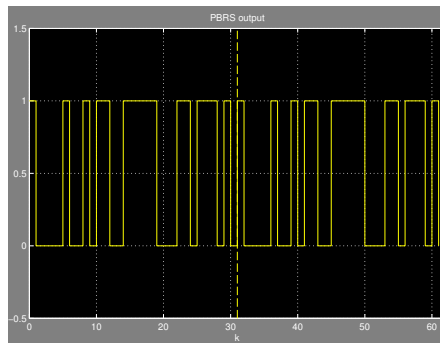
Figure 4.5   Output from the PRBS described in eq. (4.22). The dashed line marks the periodicity.

---

[6] L. Andersson et al. *A manual for system identification.* Dec. 2013. URL: `http://www.control.lth.se/media/Education/EngineeringProgram/FRT041/2011/manuallab.pdf`.

4.6   Conclusion

The system to be identified is a first-order process, and the method that will be used is a Kalman filter that minimizes the MSE of the parameters estimates. In order to excitate the signals, a linear feedback shift register that generates PRBS is used.

There are a great number of parameters to be investigated in order to optimize the identification process. Optimizing in this context means arriving at a correct and feasible model as quick as possible.

- **p, number of a-parameters**
  How many (if any) a-parameters are needed?

- **q, number of b-parameters**
  How many b-parameters are needed to handle the unknown and flow-dependent hydraulic delay T?

- **h, sampling interval**
  What sampling interval should be used in the Kalman filter?

- **tc, time constant of pre-filter**
  A 2nd-order Butterworth filter is used to pre-filter C_{in}. How should this be set?

- **P$_0$, initial value of P**
  How high should this be set, in order to maximize speed of convergence?

- **Q, parameter covariance matrix**
  Should the parameters in the A and B polynomials be allowed to vary?

- **R, observation covariance**
  Should match the observation noise.

- **PRBS amplitude**
  How large variations of C_{in} are needed to ensure identifiability?

- **PRBS period and bit length**
  How fast should the variations of C_{in} be?

The results of this is covered later on in chapter 7.

5

Lab set-up and MATLAB® implementation

5.1 Experiment set-up and collection of data

The dialysis machine used is the AK200 Ultra series from Gambro. A graphical comparison between the hemodialysis schematic previously shown in section 3.3 and a photograph of the lab set-up can be seen in figs. 5.1 and 5.2.

Gambro Log System GLS

Data is collected during experiments over serial communication interface RS-232 by the program GLS.

LogTool

LogTool is a Gambro in-house developed tool primarily used for converting log files from GLS into more "useful" formats such as excel (**.xls**) and MATLAB (**.mat**). Obviously the latter is what is of interest in this thesis.

MATLAB® Data GUI

A data GUI (see fig. 5.3)was implemented to reduce the vast amounts of time that would otherwise be spent inputting data into MATLAB®, plotting etc. The main function of the GUI is to load data sets, plot signals of interest and extract specific data into MATLAB® workspace for further analysis and use in the Kalman algorithms.

The functionality of automatically extracting data from specific time intervals has proved very useful in this thesis, since the alternative of manual manipulation of data would be extremely time consuming.
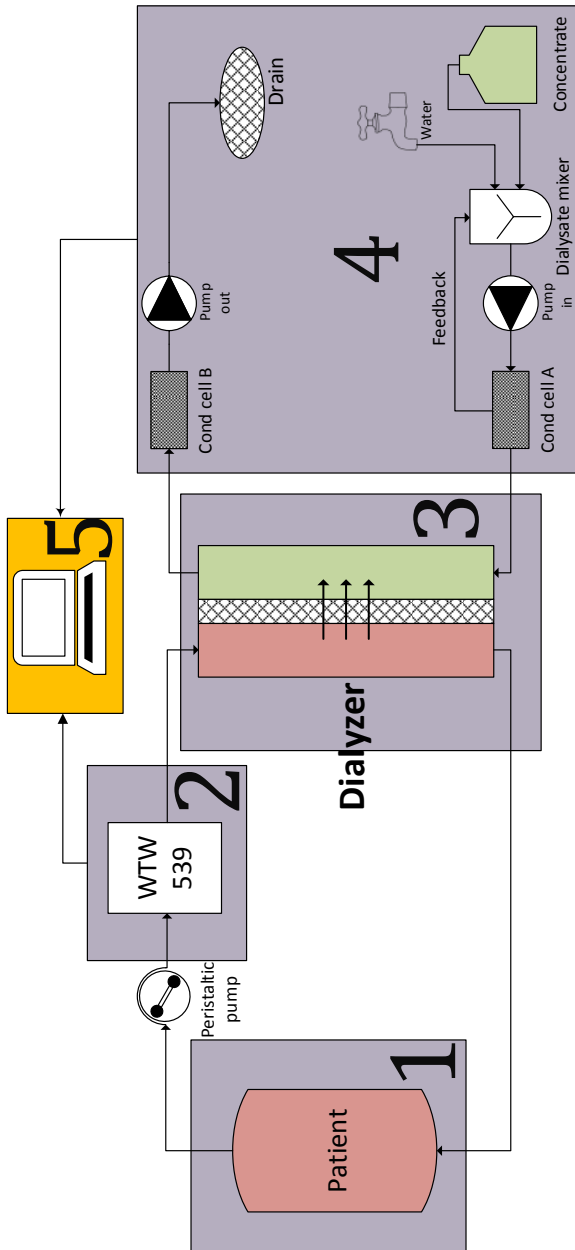
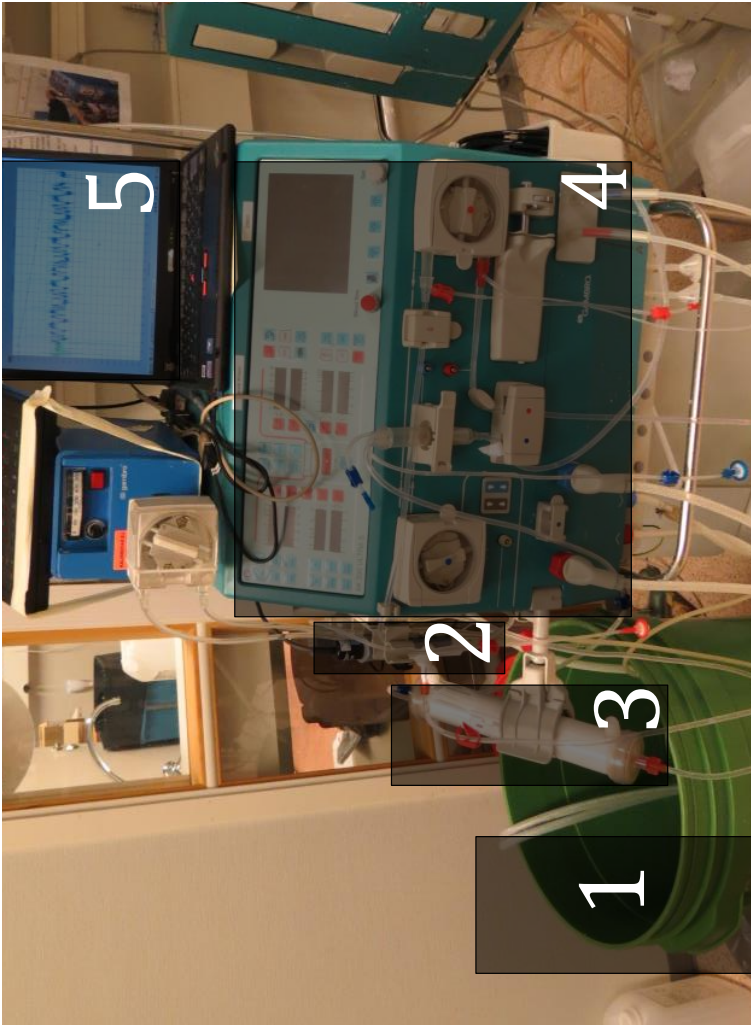**Figure 5.1**   Lab set-up outline

Figure 5.2   The actual lab set-up with relevant parts highlighted

1. "Patient" container
2. Extra conductivity meter to get readings of "patient" C_B
3. Dialyzer
4. Hemodialysis machine, AK200 ULTRA S
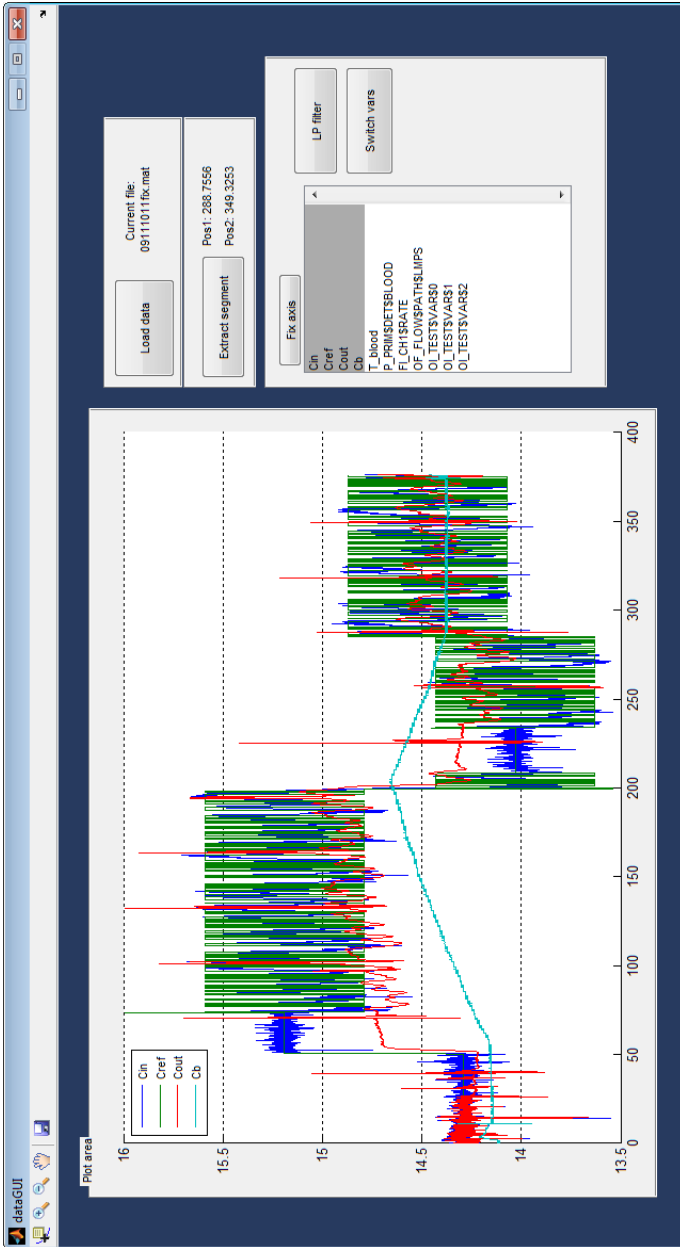5. PC collecting measurement data

Figure 5.3   The data GUI used for analysing and extracting data

5.2   Taring

Every 30 minutes, the dialysis machine will go into a so called taring mode for roughly 60 seconds. While in this mode, the dialysate is not routed to the dialyzer but through a special bypass pipe. The purpose of this is to calibrate pumps in the system to ensure that ultra filtration works as expected. Since the dialyzer in bypass, the Kalman filter must be halted during these periods, as data values not relevant to the model would otherwise be processed. See fig. 5.4 for a graph showing a taring interval.
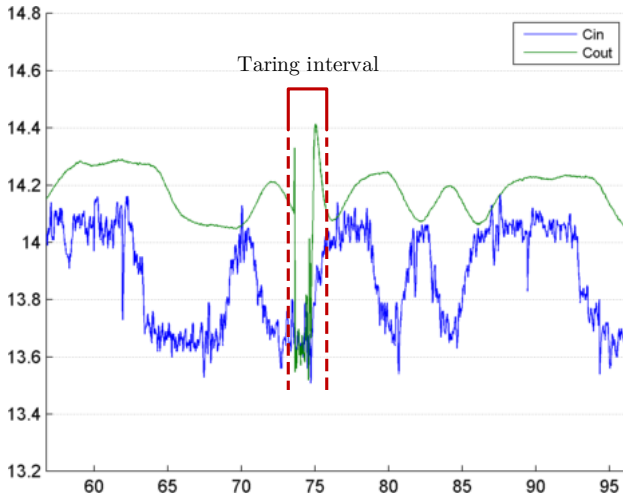
Figure 5.4   The taring interval is marked in the picture. During this time C_{out} is equal to C_{in}, due to the dialysate being in bypass mode.

5.3   Main Kalman loop

The full code for the Kalman filter loop and for a Kalman filter step can be found in appendices A.1 and A.2.

Function declaration

```
function [out_error, out_theta, P_diag,
 Cin_model, Cout_model, q] =
kalmanmain(t, Cin, Cout, alarms, Qparams, Qoffset,
   R, Pinit, h, p, T, tc, reset)
```

Description of input variables:

- **t**: time array
- **Cdin**: C_{in} array
- **Cout**: C_{out} array
- **alarms**: array that contains information about when the system is in a non-treatment state, such as taring intervals
- **Qparams**: covariance for a- and b-parameters.
- **Qoffset**: covariance for offset parameter.
- **R**: observation covariance
- **Pinit**: initial diagonal values of the P-matrix
- **h**: model sampling interval, specified as integer
- **p**: number of a-parameters
- **T**: time-delay spectrum, specified as a sample range (e.g. [30 150])
- **tc**: low-pass filter time constant, specified in seconds
- **reset**: specifies if P-matrix should be reset after alarms

Description of output variables:

- **out_error**: output error array
- **out_theta**: predicted parameters array
- **P_diag**: diagonal elements of the P-matrix for all filter iterations
- **Cin_model**: C_{in} values input to the filter
- **Cout_model**: C_{out} values input to the filter
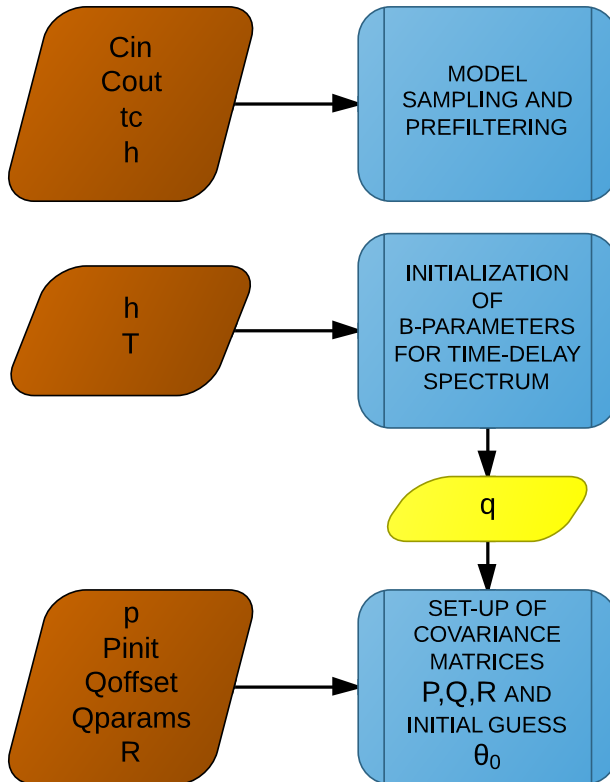- **q**: number of b-parameters used in filter
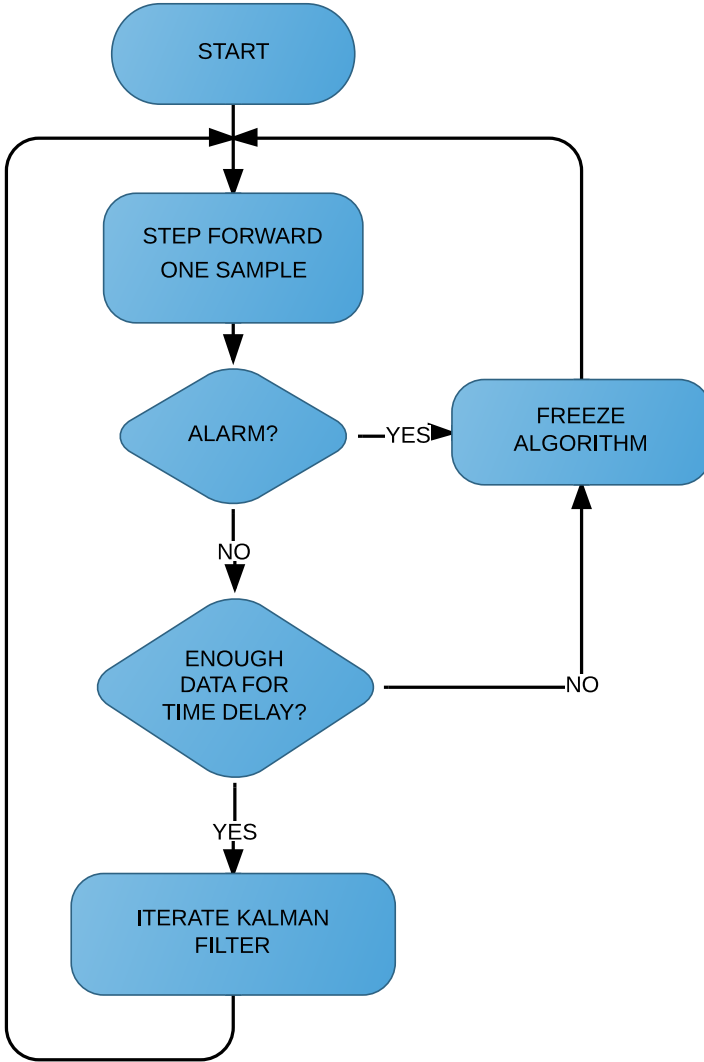
Figure 5.5   Initialization of Kalman filter

Figure 5.6　Main Kalman filter loop

Overview of algorithm

The main Kalman method (amply named **kalmanmain.m**) is divided into a number of sub processes. The algorithm has an initialization phase where the data is sampled and filtered according to user input (fig. 5.5). Depending on the time-delay spectrum (which is explained in the next chapter) and sampling time an appropriate number of b-parameters (q) is decided, according to the following formula:

$$q = \lceil \frac{T_{end} - T_{start}}{h} \rceil$$

After this, the covariance matrices P, Q, R and the initial guess $\hat{\theta}$ are set-up.

The main loop (fig. 5.6) continuously checks for alarms (includes taring) events. The Kalman algorithm halts when these occur, since running with faulty values would corrupt the identification process. Because the model includes considerable time delays (up to 300 seconds), the Kalman filter may not run immediately at start-up or after an alarm event, but has to wait for non-faulty values to buffer up before continuing.

Pre-filtering

There is an option to use a second-order Butterworth filter to filter C_{in} and C_{out}.

IIR or FIR

By setting the input parameter `p` to 0, a pure FIR-model can be used.

6

Investigation of time-delays

An unknown time delay T is part of the system to be identified. It is of importance to investigate the properties of this delay, so that the Kalman filter parameters may be set accordingly.

Since the machine has a variable dialysate flow rate Q_D between $300\,\text{ml/min}$ to $700\,\text{ml/min}$, T will vary depending on this setting. Running the machine with Q_D at $700\,\text{ml/min}$ will obviously reduce the time delay since the dialysate flows faster, and vice versa.

6.1   Method

Cross-correlation

Since C_{out} is fundamentally a time-shifted and low-pass filtered version of C_{in} looking at the cross-correlation with respect to sample lag should give a good idea of the time delays in the system. The formula for the cross-correlation is:

$$(C_{in} \star C_{out})[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} C_{in} * [m]C_{out}[n+m]. \tag{6.1}$$

Size of b-coefficients

After an estimated time delay has been found through cross-correlation, the Kalman filter is configured accordingly, to make sure that the spectrum with large cross-correlation is covered by the b-parameters.

After running the filter, a sample of the vector of b-parameters is extracted and analysed. By looking at the size of the parameters, a suitable spectrum to be used in the final implementation of the filter can be decided. It is crucial that this spectrum is well-tuned, since losing out on dynamics would worsen the model.

6.2 Results

Note

Data collected with a sampling interval of $1.1358\,$s.

Q_D at $300\,$ml/min

From fig. 6.2, the time-delay spectrum for $Q_D = 700\,$ml/min ranges from 80 to 230 samples, which corresponds to about $90\,$s to $260\,$s.
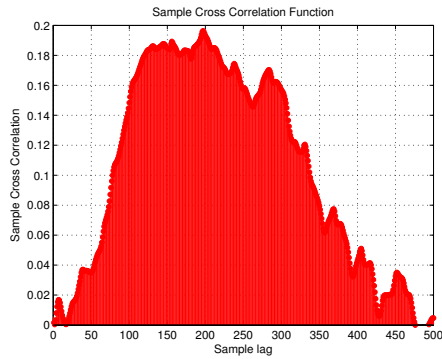

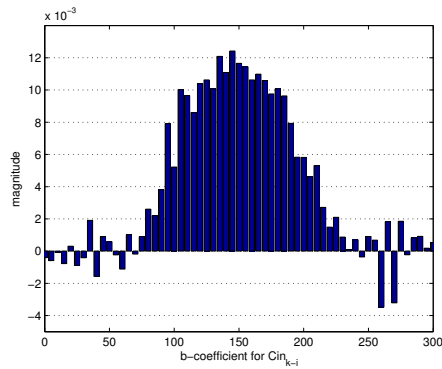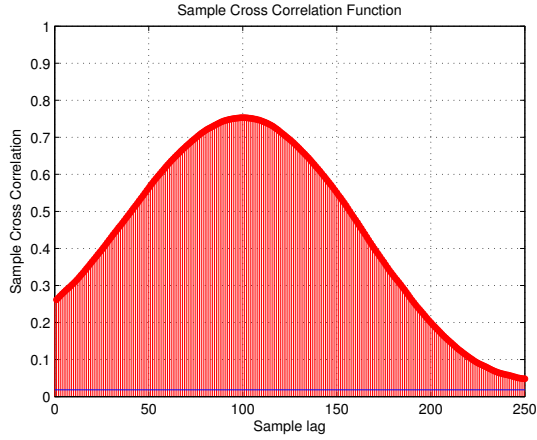**Figure 6.1**   Cross-correlation for $Q_D = 300\,$ml/min. The peak is located at $\tau = 197$.

Figure 6.2   b-parameter coefficients for $Q_D = 300\,$ml/min. The spectrum is centred around $\tau = 197$.

Q_D *at* $500 \,\mathrm{ml/min}$

Figure 6.3   Cross-correlation for $Q_D = 500 \,\mathrm{ml/min}$. The peak is located at $\tau = 100$.
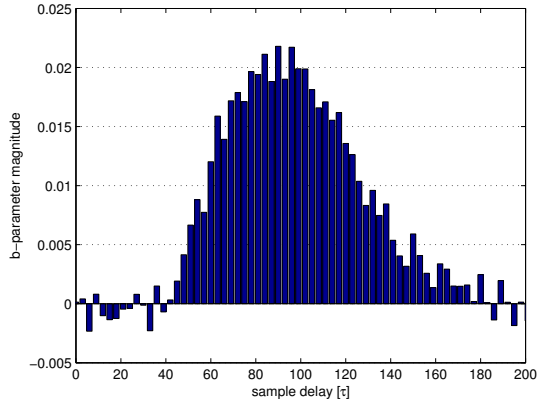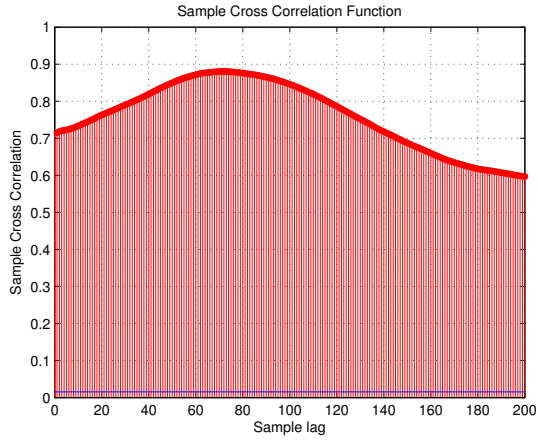

Figure 6.4   b-parameter coefficients for $Q_D = 500 \,\mathrm{ml/min}$. The spectrum has its centre around $\tau = 100$

From fig. 6.4, the time-delay spectrum for $Q_D = 500 \,\mathrm{ml/min}$ ranges from 40 to 180 samples, which corresponds to about $45\,\mathrm{s}$ to $205\,\mathrm{s}$.

Q_D *at* $700\,\text{ml/min}$

Sample Cross Correlation Function

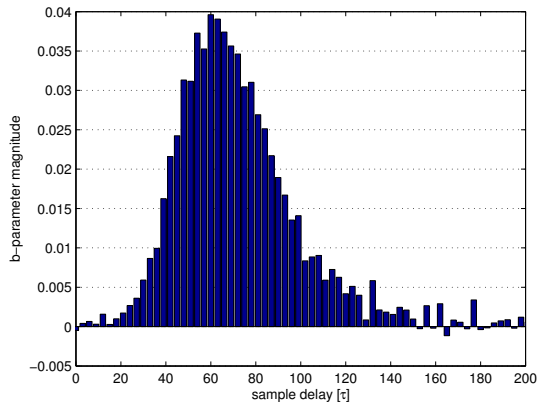Figure 6.5 Cross-correlation for $Q_D = 700\,\text{ml/min}$. The peak is located at $\tau = 72$.

Figure 6.6 b-parameter coefficients for $Q_D = 700\,\text{ml/min}$. The spectrum is centred around $\tau = 65$.

From fig. 6.6, the time-delay spectrum for $Q_D = 700\,\text{ml/min}$ ranges from 20 to 150 samples, which corresponds to about $25\,\text{s}$ to $170\,\text{s}$.

Conclusion

Of interest in this investigation are the shapes of the b-parameter coefficients, which resemble mean-centred distributions. This actually has a physical explanation. The dialyzer is in reality not a simple two-chamber construct as depicted in simplified schematics, but actually consists of thousands of fibres[1] (see fig. 6.7). This means that both the blood and dialysate will take different fluid paths through the dialyzer, yielding a distribution of flows.

The results of the analysis are collected in table 6.1.

Table 6.1   Collected results from lag analysis

Q_D,[ml/min]	$[T_{start} \ T_{end}]$ (samples)	$[T_{start} \ T_{end}]$ (seconds)
300	[80 230]	[90 260]
500	[40 180]	[45 205]
700	[20 150]	[25 170]

There are seemingly big differences in time-delays depending on flow, especially between Q_D at 300 ml/min to 700 ml/min, where the difference is 90 s. Considering the Kalman filter has to buffer up values at start and after taring events, the speed of convergence should be worse for low choices of Q_D.

Figure 6.7   Cross-section of a dialyzer, showing the fibres through which the blood flows

---

[1] Gambro. *Gambro BASICS 4, Dialysatorn.* Internal teaching material. Jan. 2002.

7

Kalman filter tuning

This chapter will go through what Kalman filter and LFSR parameters were found to give the best results. Even though there is a lot of cross-dependence between the parameters I will try to generalize and go through them one by one.

No claim is made that these parameters are "optimal", but these parameters do work very well for all experiments conducted so far and do note pose any extreme implementation costs.

For convenience, the parameters are listed again below.

- **p**, number of a-parameters

- **q**, number of b-parameters

- **h**, sampling interval

- **tc**, time constant of pre-filter

- $\mathbf{P_0}$, initial value of P

- **Q**, parameter covariance matrix

- **R**, observation covariance

- PRBS amplitude

- PRBS period and bit length

7.1 p, number of a-parameters (and choice of FIR or IIR)

When I first set out with this thesis the general idea was to model the system as an IIR-model, meaning that the output would be a function of old outputs and inputs, according to:

$$C_{out_{k+1}} = a_1 C_{out_k} + ... + a_n C_{out_{k-n}} + b_1 C_{in_{k-T}} + ... + b_m C_{in_{k-T-m}} + c$$

However, it would show that when using a-parameters, the coefficients for these would not converge, no matter how many or few (see fig. 7.1). Therefore, I decided to use a FIR-model instead, which yielded superior performance (see fig. 7.2) resulting in the discrete-time model:

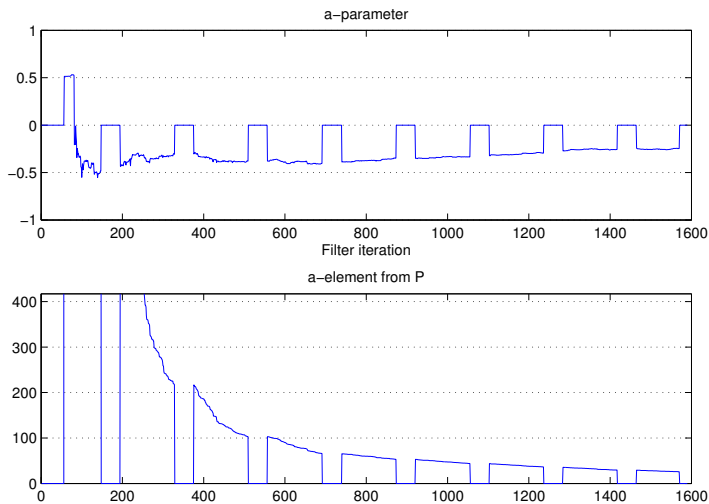$$C_{out} = B(z^{-1})C_{in} + c \tag{7.1}$$

Figure 7.1 The a-parameter is obviously not very stable and seems to be slowly approaching zero. The P-matrix shows that the accuracy of the estimation is not very good.

7.2   q, number of b-parameters

As discussed in a previous chapter, q depends on the width of the time-delay spectrum, which increases with decreasing Q_D. It also depends on the sampling interval h, in the sense that a shorter sampling interval leads to more b-parameters, and vice versa. When the start and end of the time-delay spectrum has been chosen, q is calculated from:

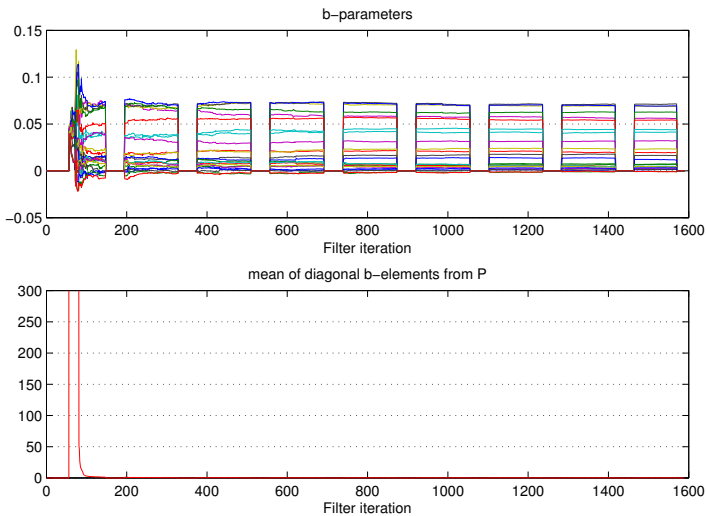$$q = \lceil \frac{T_{end} - T_{start}}{h} \rceil$$

Figure 7.2   The b-parameters show a completely different result, with very fast convergence and accurate estimations of the parameters.

7.3   h, sampling interval

The sampling interval of the Kalman filter is the parameter with largest impact on cost (computational power).

Since there aren't any fast dynamics of interest in the system, a short sampling interval (i.e. high sampling rate) is not needed. I ended up with a sampling interval of approximately $6\,\mathrm{s}$. This has to be paired with a suitable choice of filter time constant to avoid aliasing.

7.4   tc, time constant of pre-filter

The purpose of the Butterworth filter is primarily to mitigate the effect of outliers and aliasing. If the filtering is too aggressive, loss of information may occur which will slow down the identification process. It may also cause phase delays on C_{in} and C_{out}, which leads to lagging of the time-delay spectrum.

In my tests, I found tc to have quite some impact on performance. If it is set too high, important dynamics are filtered out and the parameters do not converge. As a result of this, the estimations $\hat{C}_B$ and $\hat{D}r$ suffer. I chose to settle for $tc = 1.5\,\text{s}$. See figs. 7.3 to 7.5 for a comparison between three different choices of tc.
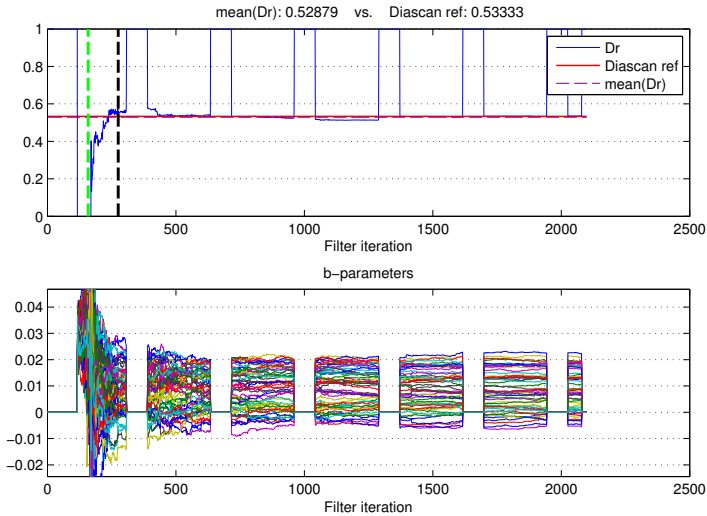

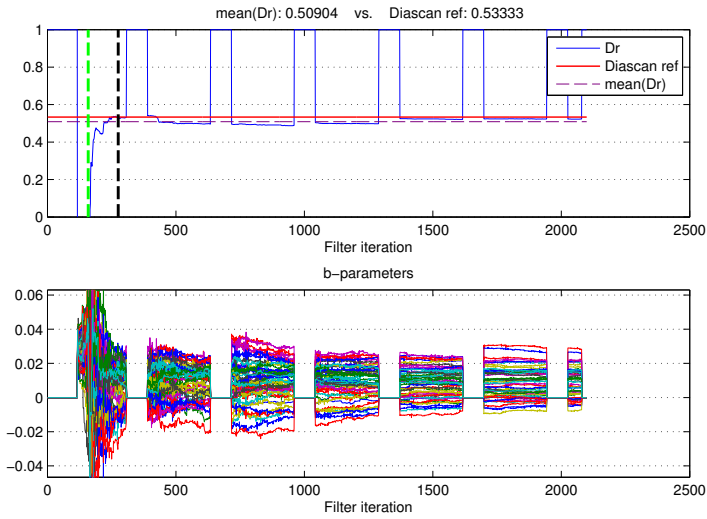**Figure 7.3** $tc = 1.5\,\text{s}$. The parameters converge quickly and are stable.

Figure 7.4 $tc = 10.5\,\mathrm{s}$. Speed of convergence and stability of b-parameters have decreased. Also note that the mean of $\hat{Dr}$ has decreased from 0.52870 to 0.50904 compared to fig. 7.3
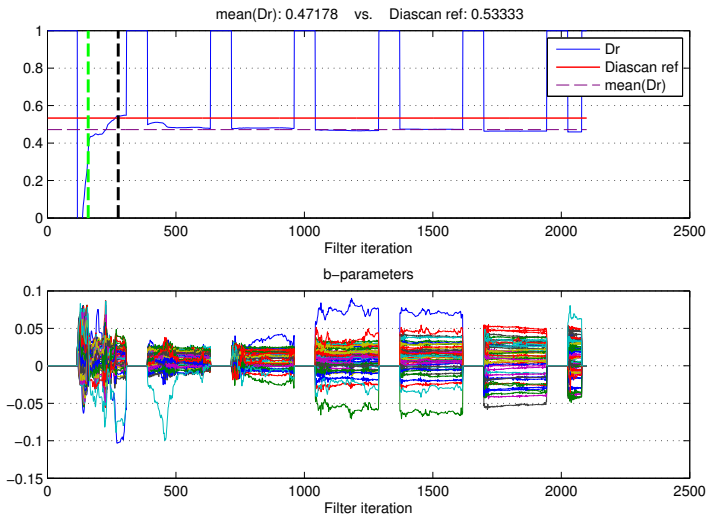

Figure 7.5 $tc = 100.5\,\mathrm{s}$. b-parameters are completely unstable, and the mean of $\hat{Dr}$ has fallen further

7.5 P_0, initial value of P

Setting a high initial value of the diagonal elements of P tells the Kalman filter to put more emphasis on new data than the initial guess $\hat{\theta}_0$. In this thesis $\hat{\theta}$ is completely unknown and $\hat{\theta}_0$ is chosen arbitrarily. This implies P_0 should be set to a large value. I chose to set it to 10000. Setting it higher did not yield any noticeable performance gains. For a comparison between comparison different values of P_0, see fig. 7.6.
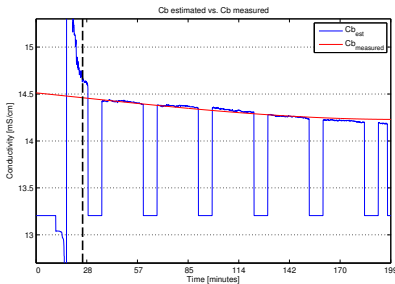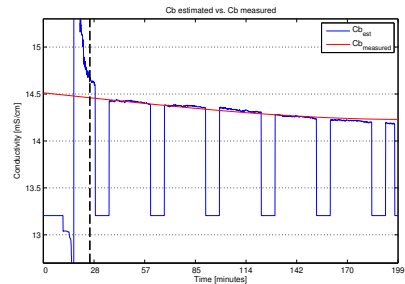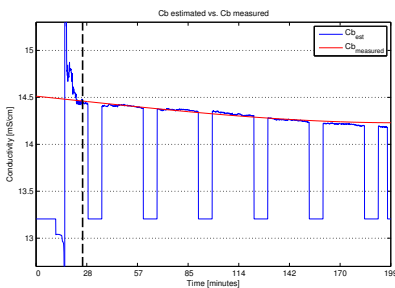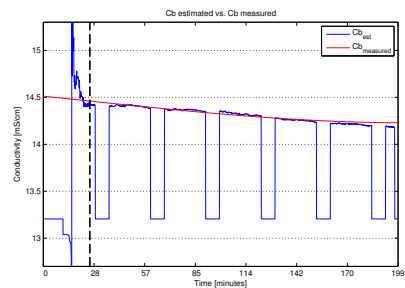

(a) $P_0 = 100$.

(b) $P_0 = 1000$

(c) $P_0 = 10000$

(d) $P_0 = 100000$

Figure 7.6   Comparison between different values of P_0. The dashed black line marks the 15th minute after start of the Kalman filter. The filter hardly converges for $P_0 = 100$. There is no discernible difference between $P_0 = 10000$ and $P_0 = 100000$.

7.6   Q, parameter covariance matrix

As discussed in section 4.4, it is possible that the Q-elements for $A(z^{-1})$ and $B(z^{-1})$ should be set to 0, since they should not vary substantially during treatment. This is also what I decided for. If something were to happen to the system which could change $A(z^{-1})$ and $B(z^{-1})$ and invalidate this
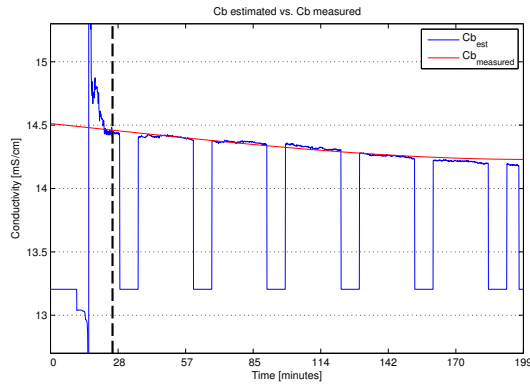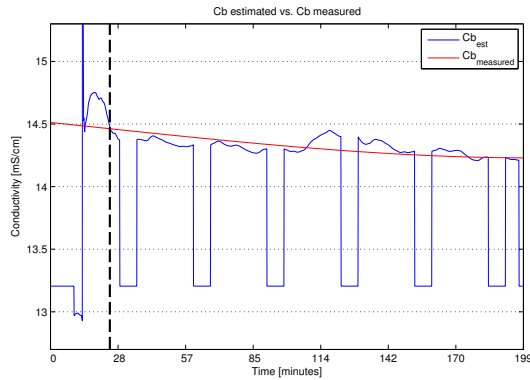
Figure 7.7 $Q_c = 0.1$. Normal convergence.

Figure 7.8 $Q_c = 0$. System can not be identified.

assumption, it makes more sense to reset P back to P_0 than to let $A(z^{-1})$ and $B(z^{-1})$ vary at all times. Setting P to P_0 instructs the Kalman filter to once again prioritize new data over old estimates.

However, the *c*-parameter must be allowed to vary, as this is related to the time-varying C_B. This element was set to 0.1. See fig. 7.8 to see what happens when it is set to 0.

7.7   R, observation covariance

This setting has negligible effect on the convergence of the filter, as long as it is set to a low value. I chose to set this to $1 \cdot 10^{-6}$, which is close the covariance of the Kalman filter output.

7.8   PRBS amplitude

A large amplitude in the excitations of C_{in} yields better parameter estimation, as stated in Rolf Johansson's *System modeling & identification*:

$$Cov(\hat{\theta}) \quad \text{proportional to} \quad \frac{1}{\text{input power}} \tag{7.2}$$

However, it is desirable to keep the amplitude as low as possible in a treatment context. A good rule of thumb for choosing a minimum amplitude is that the effect on the input should be perceptible to the eye when looking at a diagram.

I chose to vary C_{in} ±0.10 mS/cm around the set- point C_{in}^*, resulting in a total peak-to-peak amplitude of 0.20 mS/cm. Compared to the current Diascan® step-change amplitude of 0.50 mS/cm this is a vast improvement.

7.9   PRBS bit length

It is very important to set the bit length so that the variations have a relevant effect. Since the system has very long time constants, the excitations should match these, as otherwise the variations will not have a perceptible effect on C_{out} because of the low-pass dynamics.
According to Andersson et al., a rule of of thumb is[1]

$$\text{bit length} = \tau \tag{7.3}$$

where τ is the time constant of the system.
I chose to use a bit length of 90 s, which works well for all flows.

In fig. 7.9 the effect of the variations are clearly perceptible.

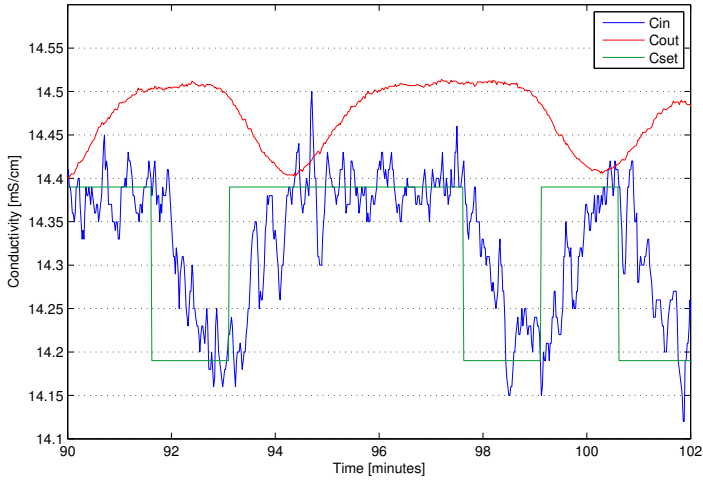[1] Andersson et al., *A manual for system identification.*

Figure 7.9   The effect of C_{in} variations on C_{out}

7.10   Summary of Kalman filter tuning

- **p** $= 0$. Resulting in the model

$$C_{out} = B(z^{-1})C_{in} + c$$

- **q** proportional to sampling rate and time-delay spectrum according to

- **h** $\approx 6\,\mathrm{s}$

- **tc** $= 1.5\,\mathrm{s}$

- **P$_0$** $= 10^4$

- **Q** $= 0$ for b-parameters, 0.1 for c-parameter

- **R** $= 10^{-7}$

- PRBS peak-to-peak amplitude $= 0.20\,\mathrm{mS/cm}$

- PRBS bit length $= 90\,\mathrm{s}$

8

Results

8.1 How to evaluate performance?

During this thesis, I had two ways of evaluating if my results were feasible:

1. comparing calculated $\hat{C}_B$ to C_B measurements from the extra conductivity meter

2. comparing calculated $\hat{D}$ to dialysance output from the current Diascan® feature

$\hat{C}_B$ comparison

The most important and precise accuracy indicator is the mean error and standard deviation of the blood conductivity estimation $\hat{C}_B$. This is due to the fact that a direct measurement of C_B is available. The mean error is defined as

$$\bar{\varepsilon}_{\hat{C}_B} = \frac{\sum_{k=1}^{n}(C_{B_k} - \hat{C}_{B_k})}{n}$$

The standard deviation is defined as

$$\sigma_\varepsilon = \sqrt{Var(\hat{C}_B - C_B)}$$

$\hat{C}_B$ *offset issue*   During my extensive testing with the machine, I noticed that the conductivity cells measuring C_{in} and C_{out} tend to drift from each other. When the machine is in bypass, C_{out} should be equal to C_{in}. But usually there is an offset between them, see fig. 8.1 for an example of this.

To understand the impact of this offset, let's formulate the FIR discrete-time model:

$$C_{out} = B(z^{-1})C_{in} + c \tag{8.1}$$

Let's reformulate c as
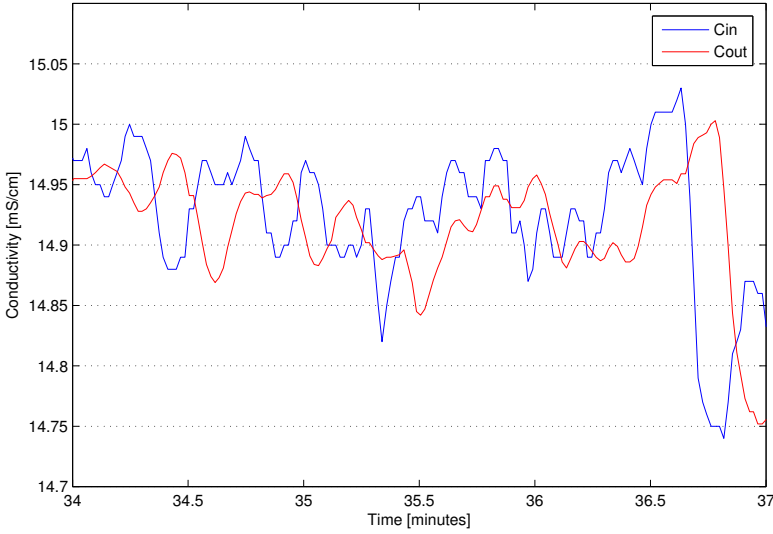
$$c = c_0 + c_{off} \tag{8.2}$$

Figure 8.1 Offset between C_{in} and C_{out}. In this case $C_{out} \approx C_{in} - 0.0144$

where c_0 is the ordinary model offset and c_{off} is the additional offset from conductivity cell bias. The model becomes

$$C_{out} = B(z^{-1})C_{in} + c_0 + c_{off} \tag{8.3}$$

and the $\hat{C}_B$ calculation

$$\hat{C}_B = \frac{c}{\hat{D}r} \tag{8.4}$$

$$\hat{C}_B = \frac{c_0 + c_{off}}{\hat{D}r} \tag{8.5}$$

Thus, if $C_{in} > C_{out}$ (as in fig. 8.1), the calculated $\hat{C}_B$ will decrease.

Example: Assume the relative dialysance Dr is 0.5 ($D = 250 \, \text{ml/min}$ at $Q_D = 500 \, \text{ml/min}$) and $c_{off} = -0.02 \, \text{mS/cm}$. $\hat{C}_B$ will have a resulting offset of $-0.04 \, \text{mS/cm}$, which is clearly noticeable.

Diascan® comparison

As already shown in section 4.3, the Diascan® output has large variations in its output. I still opted to use a mean of the Diascan® output for each experiment as a reference to compare $\hat{D}r$ to, mainly in order to spot any unexpectedly large differences.

Conclusion

$\hat{C}_B$ vs. C_B comparison is the main method used to evaluate performance, since C_B data is sampled. However, offset errors are present in $\hat{C}_B$ which can not be circumvented. The Diascan® output is only used as a reference.

8.2 Results

All these results are from using the Kalman filter with the parameters listed in section 7.10. The only parameter that is changed between experiments is the time-delay spectrum.

The results are presented with two graphs for every experiment:

1. $\hat{C}_B$ vs. measured C_B

2. $\hat{D}r$ vs. Diascan® output (for reference purposes)

Following these graphs are the calculated values of $\bar{\varepsilon}_{\hat{C}_B}$ and σ_ε (see section 8.1 for definitions).

Note

The discontinuities in the graphs are a result from the taring intervals, which happen every 30 minutes. During these, the Kalman filter must halt, as described in section 5.2.

Note

The black dashed lines denote when 15 minutes have passed since the Kalman filter started processing data.

$Q_D = 300 \, \text{ml/min}$
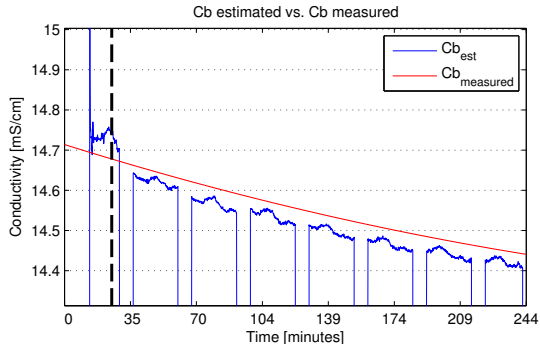

Note

Time-delay range: 70-250 samples.

Figure 8.2 $\hat{C}_B$ vs. C_B

Figure 8.3 $\hat{Dr}$ vs. Diascan$^{®}$ reference. Dashed purple line denotes mean($\hat{Dr}$).

$$\bar{\varepsilon}_{\hat{C}_B} = 0.0310 \, \text{mS/cm}$$
$$\sigma_\varepsilon = 0.0083 \, \text{mS/cm}$$

$Q_D = 500\,\text{ml/min}$

Figure 8.4 $\hat{C}_B$ vs. C_B

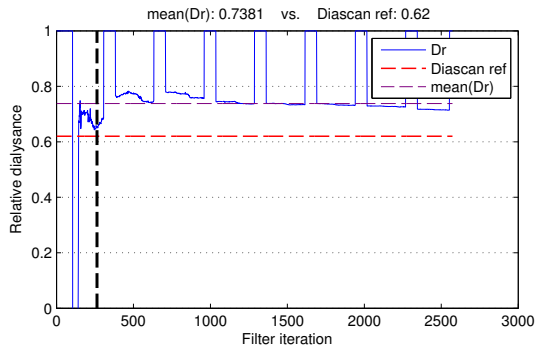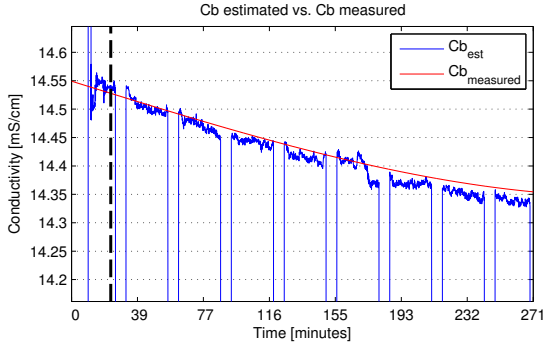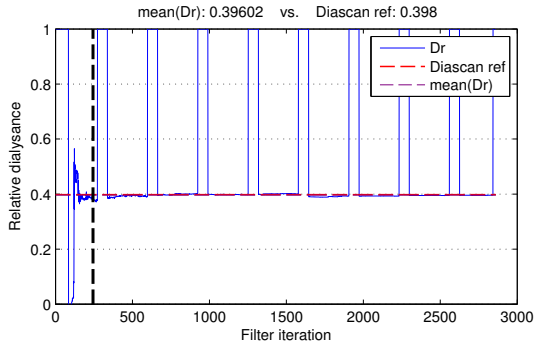**Figure 8.5** $\hat{Dr}$ vs. Diascan® reference. Dashed purple line denotes mean($\hat{Dr}$).

$$\bar{\varepsilon}_{\hat{C}_B} = 0.0126\,\text{mS/cm}$$
$$\sigma_\varepsilon = 0.0092\,\text{mS/cm}$$

$Q_D = 700 \, \text{ml/min}$

Note

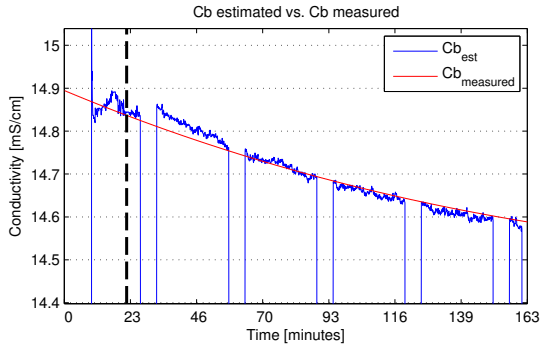Time-delay range: 10-170 samples.

Figure 8.6 $\hat{C}_B$ vs. C_B

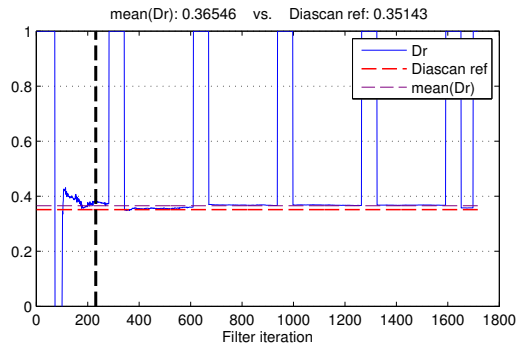**Figure 8.7** $\hat{Dr}$ vs. Diascan® reference. Dashed purple line denotes mean($\hat{Dr}$), and dashed red denotes the Diascan® reference.

$$\bar{\varepsilon}_{\hat{C}_B} = -0.0021 \, \text{mS/cm}$$
$$\sigma_\varepsilon = 0.0152 \, \text{mS/cm}$$

67

Spread of $\hat{D}$ compared with Diascan®

A way of comparing the spread of $\hat{D}$ from the Kalman filter with that of Diascan® is looking at the standard deviation of $\sigma_{\hat{D}}$ vs. $\sigma_{Diascan}$.

With data from the experiment at $Q_D = 500\,\text{ml/min}$, the results of this were:

$$\sigma_{\hat{D}} = 1.8577\,\text{ml/min}$$

$$\sigma_{Diascan} = 8.3919\,\text{ml/min}$$

The Diascan® standard deviation is about 4.5 times larger. For a visual comparison, see fig. 8.8.
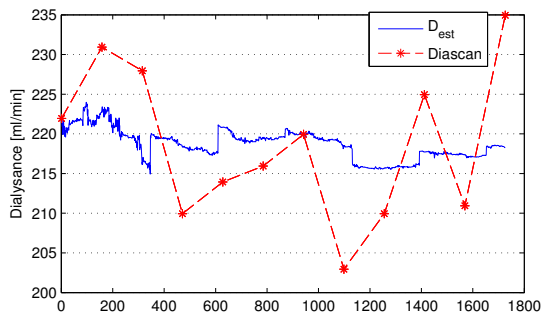

Figure 8.8 Visual comparison of estimated $\hat{D}$ and Diascan®. The discontinuities stem from taration breaks.

8.3 Discussion of results

Accuracy After a relatively short settling period, $\hat{C}_B$ tracks C_B during the rest of the experiment, showing that the dialyzer model has been successfully identified. The mean error as well as the standard deviation of the error are small, with magnitudes usually ranging around $0.01\,\text{mS/cm}$. However, in the case of $300\,\text{ml/min}$ (section 8.2), the mean error is $0.0310\,\text{mS/cm}$, yet the spread is still at less than $0.01\,\text{mS/cm}$. Thus, it is likely that the main component of this mean error is a result from the offset in the conductivity cells.

Speed of convergence The speed of convergence differs somewhat between the different dialysate flows. Comparing $Q_D = 300\,\text{ml/min}$ (fig. 8.2) with $Q_D = 500\,\text{ml/min}$ and $700\,\text{ml/min}$ (figs. 8.4 and 8.6), convergence is noticeably slower. This is expected, considering that the time-delay spectrum is larger, resulting in more b-parameters and a longer initialization time.

Time-delay ranges   The time-delay ranges are somewhat larger than the results from chapter 6.

I chose this to provide a small margin around those ranges.

Overall, the time-delay ranges can vary quite a lot depending on Q_D and Q_B.

Comparison with Diascan®   From the comparison with Diascan® (fig. 8.8), it is shown that the new method has a standard deviation almost five times smaller than that for Diascan®, which is a huge improvement.

9

Conclusion

The proposed method investigated in this thesis shows a lot of promise, with a large reduction of spread compared to Diascan®. If the implementation is not too complex or expensive, it could be a solid candidate to replace Diascan® in the future. Considering the current filter sampling interval is 6 s, implementation should not be too costly.

Suggestions for improvements and future analysis

Time-delay tuning In this thesis the time-delay parameters are separately identified and then input to the filter a priori. This obviously would not work in a real treatment setting, where different factors such as clogging, different flows and different dialyzers will alter the time-delay spectrum. Therefore, some sort of algorithm to automatically find a suitable spectrum is needed. A possible way to adaptively find the time-delay spectrum could be to look at the size of the b-parameters from the Kalman filter and resize the spectrum accordingly.

Improve the offset calibrations of conductivity cells Offsets between the conductivity cells had quite some effect on the calculated $\hat{C}_B$. Reducing these offsets would result in increased accuracy of $\hat{C}_B$.

Bibliography

Andersson, L., U. Jönsson, K. H. Johansson, and J. Bengtsson. *A manual for system identification.* 2013. URL: http://www.control.lth.se/media/Education/EngineeringProgram/FRT041/2011/manuallab.pdf (cit. on pp. 37, 60).

Bosetto, A. and F. Paolini. "Method for determining a parameter indicative of an extracorporeal blood treatment". Pat. US 6691040 B2. 2004 (cit. on p. 31).

Gambro. *Gambro BASICS 1, Njurfunktion.* Internal teaching material. 2002 (cit. on pp. 15–17, 19, 20).

Gambro. *Gambro BASICS 2, Transportprinciper.* Internal teaching material. 2002 (cit. on pp. 23, 24).

Gambro. *Gambro BASICS 3, Principer för Hemodialys.* Internal teaching material. 2002 (cit. on pp. 22, 25).

Gambro. *Gambro BASICS 4, Dialysatorn.* Internal teaching material. 2002 (cit. on p. 52).

Gambro. *About gambro.* 2013. URL: http://www.gambro.com/en/uk/About-Gambro/Gambro-in-brief/ (cit. on p. 11).

Hörl, W. H., K. M. Koch, R. M. Lindsay, C. Ronco, and J. F. Winchester. *Replacement of renal function by dialysis.* 5th ed. Kluwer academic publishers, 2004 (cit. on pp. 12, 13, 19, 22, 26, 27).

Johansson, R. *System modeling & identification.* Prentice Hall, 1993. ISBN: 0-13-482308-7 (cit. on pp. 35, 36, 60).

MedicineNet. *Definition of creatinine.* 2013. URL: http://www.medterms.com/script/main/art.asp?articlekey=12550 (cit. on p. 16).

Polaschegg, H. "Automatic, noninvasive intradialytic clearance measurement". *The International Journal of Artificial Organs* **16**:4 (1993) (cit. on pp. 26, 32).

Today in science history. *Science quotes by Walter Bradford Cannon*. 2014. URL: http : / / todayinsci . com / C / Cannon _ Walter / CannonWalter-Quotations.htm (cit. on p. 16).

Tony Lacey, C. U. *Tutorial: the kalman filter*. 2013. URL: http://mpdc.mae.cornell.edu/Courses/UQ/kf1.pdf (cit. on p. 36).

A

MATLAB® Code

A.1 kalmanmain.m

```
1  function [out_error, out_theta, P_diag, Cin_model, Cout_model,q] =...
2      kalmanmain(t, Cin, Cout, alarms, Qparams, Qoffset, R, Pinit, h, p, T,
         tc)
3
4  %% Print original data sampling interval and model sampling interval
5  dt = (t(end)−t(1))/(length(t)−1)∗60;
6
7  disp(['Data sampling time: ' num2str(dt)]);
8  disp(['Model sampling time: ' num2str(h∗dt)]);
9
10  %% model sampling and prefiltering
11  if (tc ~= 0) %tc = 0 yields no filtering
12      fc = 1/(2∗pi∗tc);
13      fs = 1/dt;
14      wc=fc/(fs/2); %normalized freq
15      [B,A]=butter(2,wc);
16      Cdinf = filter(B,A,Cin);
17      Cdoutf = filter(B,A,Cout);
18      u = Cdinf(1:h:end);  %resample with h
19      y = Cdoutf(1:h:end);
20  else
21      u = Cin(1:h:end);  %resample with h
22      y = Cout(1:h:end);
23  end
24  alarms = alarms(1:h:end); %resample with h
25
26  %% Set up time−delay spectrum
27  T_start = T(1);
28  T_end = T(2);
29  T_start_h = ceil(T_start/h);
30  T_end_h = ceil(T_end/h);
```

73

```matlab
31   q = length(T_start_h:phi_h:T_end_h);
32
33   %% covariance matrices set−up
34   P_init = Pinit;
35   P = diag(ones(1,p+q+1))*P_init;
36   Q = Qparams*diag(ones(1,p+q+1)); %parameter covariance
37   Q(end,end) = Qoffset; %set offset cov.
38   R = R; %output covariance
39
40   %% initial guess
41   %currently based on Dr at 0.5
42   init_guess_a = 0.5;
43   init_guess_b = 0.4*(1−init_guess_a*p)/q;
44   init_guess_c = 0.05;
45   theta_init = ones(p+q+1,1) * init_guess_b;
46   theta_init(1:p) = init_guess_a;
47   theta_init(end) = init_guess_c;
48   theta = theta_init;
49
50   %% sets up safety margin for taring periods
51   extra_margin = 5;
52   alarm_margin = T_end_h + 1 + extra_margin;
53   alarm_prox = alarm_margin;
54
55   %% initiation of out vectors
56   out_error = zeros(1,length(u));
57   out_theta = zeros(length(u),length(theta));
58   out_a = zeros(length(u),p);
59   out_b = zeros(length(u),q);
60   out_c = zeros(length(u),1);
61   P_diag = zeros(length(u),length(P));
62   Cout_model = zeros(length(u),1);
63   Cin_model = zeros(length(u),1);
64
65   %% kalman filtering loop
66   for i=(T_end_h):length(u)−20
67       if alarms(i) == 3 && alarms(i+10) == 3; %have to look in the "
                future" for alarms due to lag in alarms vector
68           alarm_prox = max(0,alarm_prox−1);
69       else
70           alarm_prox = alarm_margin;
71       end
72
73       if (alarm_prox == 0)
74
75           if (p > 0) %if we have a−parameters
```

```
76              phi = [flipud(y(i−p+1:i)) ; flipud(u(i−T_end_h:phi_h:i−
                   T_start_h)) ; 1];
77          else
78              phi = [flipud(u(i−T_end_h:phi_h:i−T_start_h)) ; 1];
79          end
80
81          [theta, P, e] = kalmanstep(y(i+1),phi,theta,P,Q,R); %do a
                   kalman iteration
82          %Save data
83          Cout_model(i) = y(i);
84          Cin_model(i) = u(i−T_start_h);
85          out_error(i) = e;
86          out_theta(i,:) = theta';
87          P_diag(i,:) = diag(P)';
88          %Separation of theta_est into a/b/c parameters
89          if (p > 0)
90              out_a(i,:) = theta(1:p);
91          end
92          out_b(i,:) = theta(p+1:q+p);
93          out_c(i,:) = theta(end);
94      else
95          %DO NOTHING
96      end
97  end
```

A.2   kalmanstep.m

```
1  function [thetanew, Pnew, K, error] = kalmanstep(y, phi, thetaold, Pold, Q
       , R)
2  %the standard update equations as described in report
3  K = Pold∗phi/(R+phi'∗Pold∗phi);
4
5  thetanew = thetaold + K∗(y−phi'∗thetaold);
6  error = (y − phi' ∗ thetaold);
7  Pnew = Pold − K∗phi'∗Pold + Q;
8  Pnew=(Pnew+Pnew')/2; %fix symmetry in case of rounding errors
9  end
```

| *Author(s)*
Johan Larsson | *Supervisor*
Jan Sternby, Gambro
Tore Hägglund, Dept. of Automatic Control, Lund University, Sweden (examiner) |
| | *Sponsoring organization* |

Title and subtitle

Estimation of dialysis treatment efficiency by means of system identification

Abstract

When treating patients suffering from renal failure with hemodialysis, an obvious point of interest is the actual blood cleaning efficiency of the dialyzer (artificial kidney). This efficiency is called clearance or dialysance.

The method currently used for estimating clearance is based on doing a step-change on the process. Due to the nature of the process, this method is slow and has a relatively large output spread.
This master thesis investigates a new method of finding clearance by means of system identification. The dialyzer is modelled as a discrete-time system, and perturbed by use of a pseudo-binary random sequence. The input/output-data is then fed into an optimal Kalman filter for parameter estimation. The gain and offset of the identified system is directly related to the dialyzer clearance of the treatment.

The method shows promising results, usually converging to good parameters within 15 minutes, and then tracking changes continuously for the rest of the treatment. It also provides better accuracy, with a considerable reduction in spread compared to the old method.

Main obstacles stem from variable time-delays in the system and measurement offsets.

Classification system and/or index terms (if any)

Supplementary bibliographical information

| *Language*
English | *Number of pages*
1-75 | *Recipient's notes* |
| *Security classification* | | |

http://www.control.lth.se/publications/