Laila Kristin Gabriel

A Mathematical Model for Estimating Drug Concentrations in Human Brain Fluid

Master Thesis
September 20, 2016

Supervisor:
Dr. Vivi Rottschäfer

Mathematisch Instituut,
Universiteit Leiden

Abstract

The treatment of neurological diseases requires knowledge of the drug contration at the central-nervous-system (CNS) target site, which, in most cases, is the extracellular fluid (ECF) of the brain. Determining the drug concentration in the brain ECF is a major challenge in the development of CNS-targeted drugs since, for ethical reasons, it is not possible to measure concentrations in human brain ECF. It has therefore been suggested to use the concentrations in cerebrospinal fluid (CSF) as an indication for ECF drug concentrations. The advantage of using CSF concentrations is that CSF is easily accessible, since we can extract CSF in a human by a lumbar puncture.

The aim of this thesis is to find a mathematical model that helps to predict CNS target-site concentrations when only CSF concentrations in the lumbar region are known. To this end, we will use an approach based on partial differential equations (PDE), in order to improve the existing models, which are based on ordinary differential equations. A one-dimensional model consisting of an inverse problem will be developed and presented. The solvability of this inverse problem will be investigated, and a method for solving it will be proposed.

Contents

1 Introduction

The treatment of central-nervous-system (CNS) disorders such as Alzheimer's disease, multiple sclerosis and Parkinson's disease is a major problem. Due to ageing societies and other factors, such as widespread stress and anxiety, neurological diseases become more and more common. Most existing drugs treat symptoms rather than the diseases. Therefore new drugs and new methods of treatment are needed. In drug development, however, many new drugs fail in clinical studies while having a significant effect in laboratory species. Reasons for this failure may be that drugs are given at the wrong dose or schedule, or that they do not reach the CNS target site (that is the site the drug needs to reach in order to be effective) in the right concentration because they cannot cross the blood-brain barrier (BBB). These problems make it diffcult to predict effects of the drugs on the human CNS. Also, some drugs need a minimal concentration in the brain in order to be effective. Of the chemotherapeutic agent methotrexate, for example, it is known that in order to effectively kill tumor cells, it needs a concentration of about 0.45μg/ml at the target site [37]. While it is important to reach this concentration, overdoses must be avoided since these might have negative effects on healthy cells in different parts of the body. It is thus necessary to be able to predict the drug concentration at the CNS target site, which is the brain extracellular fluid (ECF), in order to give the right drug at the right time and dose. In humans, however, the only accessable fluid is the cerebrospinal fluid (CSF), which can be obtained by a spinal tap in the lumbar region located in the lower back. Data obtained in animals indicates that CSF concentration gives a rather good indication of brain-ECF concentrations, but not a reliable prediction.

The aim of this thesis is to derive a simple model which can help to predict human-brain target-site drug concentrations when only CSF concentrations in the lumbar region are known. To this end, we will use an approach based on partial differential equations (PDEs) to model the change in the concentration over time along the CSF system, in order to improve the existing multi-compartment models presented by Westerhout et al. in [38] and [39], which are based on ordinary differential equations (ODEs). One advantage of using a PDE instead of ODEs is that the former describes the concentration in any point of the CSF system at any time in one single equation. In compartmental models, a new equation has to be found for every compartment, which yields a system of ODEs that must be satisfied simultaneously. Multi-compartment models give only overall concentrations for the modelled compartments, thus neglecting local variation within a given compartment. Therefore, we expect a PDE to be more accurate. Once a PDE has been found, we can impose initial

and boundary conditions based on observed rat data and solve the resulting initial-boundary value problem to understand how the concentration profiles change according to the PDE. The results can then be compared to those obtained by means of the compartmental models. In the last step of modelling, we want to reconstruct brain-ECF concentrations from measurements in the lumbar region using the PDE. This leads to a so-called inverse problem. We will investigate the solvability of the inverse problem, and a method for solving it will be introduced.

Before we can start modelling, a basic understanding of the processes involved in drug distribution is needed. Therefore the following chapter will give a short introduction to the CNS, the CNS system, and the brain barriers. It should be mentioned that the CNS and the CSF system are very complex, and that many anatomical and physiological parameters influence the drug distribution. These processes are subject to ongoing reasearch. Since what interests us is the mathematical model, the chapter on clinical and physiological background is kept as simple as possible and only those mechanisms, barriers and processes which are necessary to understand the model equations and parameters are discussed.

2 Clinical background: Anatomy and physiology of the central nervous system

This chapter aims to give an overview of the anatomy and physiology of the CNS and the mechanisms involved in drug distribution, in order to provide a basis for the modelling.

The central nervous system (CNS) can roughly be divided into three main compartements:

1. the extracellular fluid (ECF) of the brain,
2. the brain cells,
3. the ventricular system, consisting of the left and right lateral ventricles plus the third and fourth ventricles, the cerebral aqueduct, the cisterna magna, and the subarachnoid space (SAS).

The localizations of these compartments are shown in figure 1. The ventricular system works, in a manner of speaking, as a communication network [23] and is filled with cerebrospinal fluid (CSF). In humans, the subarachnoid space is of great interest since it includes the lumbar region in the lower back, where CSF can be accessed easily. (Remember that we want to use CSF concentrations in order to reconstruct brain-ECF concentrations.) Within each of the ventricles there is a choroid plexus, which is a network of cells forming an interface between blood and CSF. The exchange of substances between blood, brain ECF, and CSF is regulated by the blood-brain barrier (BBB) and the blood-CSF barrier (BCSFB). These will be introduced in chapter 2.2. CSF is continuously produced and eliminated.

2.1 Cerebrospinal fluid (CSF) and CSF flow

CSF cushions the CNS against trauma, removes metabolic waste and foreign substances from the brain, and helps to maintain an optimal environment in the extracellular space, that is the liquid-filled interstice between the brain cells and around the brain. CSF is mainly produced in the choroid plexuses of the four ventricles. A certain portion of CSF is also contributed by the ECF across the ependyma, which is a cellular layer lining the ventricles [36]. Forming a barrier between brain ECF and CSF, the ependymal layer plays an important role in the exchange of substances and drug compounds between brain ECF and CSF. From the ventricles, the CSF flows through the subarachnoid space (SAS) where it can be accessed by a spinal tap in the lower back. CSF is mainly eliminated by the arachnoid granulations, hereinafter called *arachnoid villi*, where it flows back into the blood. This

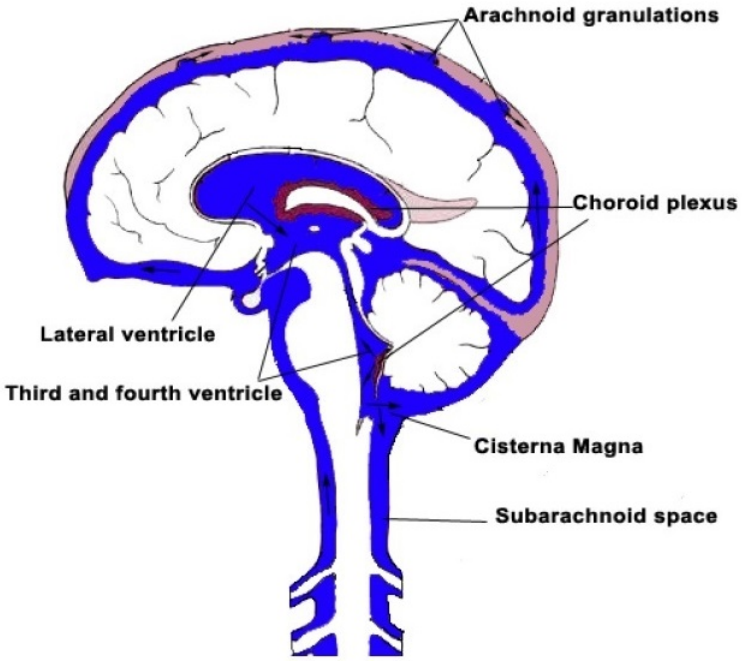

Figure 1: Localization of the compartments of the brain and CSF flow. Blue= CSF, pink=venous blood, white=spine/brain. Reproduction of figure 4 in [14].

process is driven by hydrostatic pressure differences between CSF and the cerebral veins. These pressure differences make the CSF flow a one-way stream. See figures 1 and 2 for the localizations of the brain compartments and for a schematic representation of the CSF flow.

It has been shown in several studies that a small portion of CSF is reabsorbed into the blood along the spine, but it is not clear yet to what extend this has to be considered. Small portions of CSF may also be absorbed by the cervical lymphatics, that is by the lymph nodes in the neck (see [8] and [27]). Again, it is not clear to what extend the drainage into the lymphatics has to be considered in our model.

In an adult human, approximately 500 ml of CSF is produced every day. The total amount of CSF present in an adult human at a given time is about 150 ml. This implies a CSF turnover of about 4 times per day, or, in other words, CSF is replaced every 5 to 6 hours. These amounts and rates, however, may vary somewhat depending on a person's health, age, and gender. Use of anesthesia and time of day may also have an impact on the CSF flow and the drug concentration in the SAS [14]. By using the turnover rate and the length

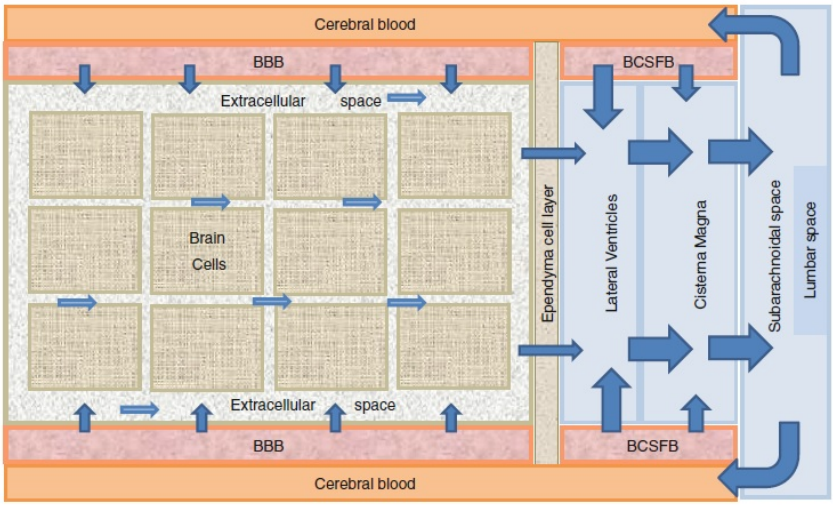

Figure 2: Simplified schematic one-dimensional CSF flow. Reproduction of figure 2 in [23].

of the CSF space, it is also possible to compute an average flow velocity. The CSF space is approximately twice as long as the spine. On average, female human spines are 61 cm and male human spines 71 cm long. The spine has a varying width of 6.4 mm in the thoracic area and 13 mm in the cervical and lumbar regions. The 71 cm of a male human spine are divided as follows: The cervical region measures about 12.5 cm, the thoracic 28 cm, and the lumbar region 18 cm. The last 12.5 cm form the sacrum and coccyx. The locations of the different regions are shown in figure 3.[1]

2.2 Barriers of the brain

One of the first things to consider in testing a new drug is its way of delivery into the brain. Some substances might be suitable for the treatment of certain neurological diseases, but if they fail to cross the barriers to the brain, they fail in clinical studies as they do not reach the brain ECF. Hence, when testing a new drug it is important to also investigate the pathways into the brain.

The blood-brain barrier (BBB) is the largest and most important barrier to be considered. It is the barrier between the blood and the brain ECF. As such, its main task is to prevent harmful substances from entering the brain

[1]Graphic from
www.compelvisuals.com/compel_blog/normal-anatomy-of-the-human-vertebral-column/

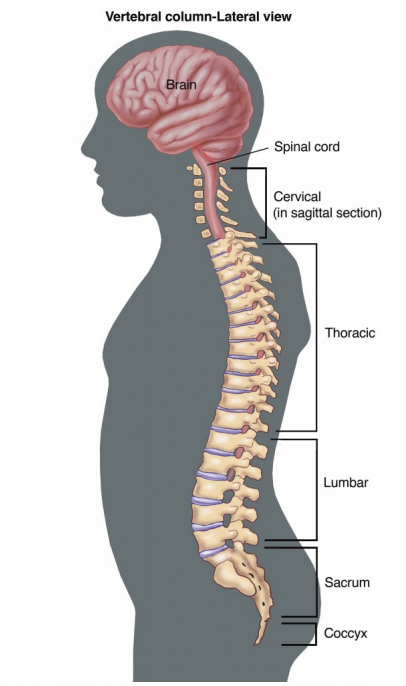

Figure 3: The vertebral column (spine).

while allowing passage to other substances, such as certain proteins or water. The BBB is formed by tight junctions between the endothelial cells.

The surface area of the BBB is large: somewhere between 12 m^2 and 18 m^2 in a human adult [1], which indicates the importance of the BBB in drug delivery to the brain.

The blood-CSF barrier (BCSFB) is situated inside the brain around the ventricles and is formed by tight junctions between the epithelial cells of the choroid plexuses. Its permeability and modes of transport are similar to those of the BBB, but the BCSFB is much smaller: only about 0.2 m^2 [37], which implicates that the BCSFB plays a minor role in drug delivery to the brain. The main task of the choroid plexuses and the BCSFB barrier is to secrete CSF and to restrict free diffusion in the brain.

2.3 Drug characteristics and physiological parameters

In order to become effective, any kind of drug must be transported across the BBB. The smaller and more lipophilic molecules are, the more likely are they to cross the BBB by passive diffusion, which is the simplest mode of transport. Alongside, there are several possible modes of active transport (e.g. through transporter proteins); by these, larger and/or more hydrophilic

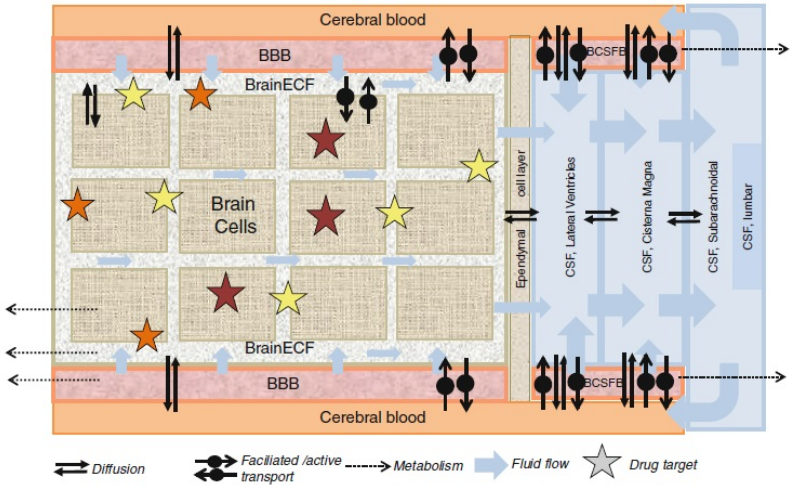

Figure 4: Simplified schematic representation of drug transport, fluid flow, and metabolism, which govern the concentration-time profiles of the drug in different parts of the CNS. Reproduction of figure 3 in [23].

molecules may be able to enter the brain ECF, too. This means that drug characteristics such as molecular size and weight, lipophilicity, affinity for active transporters, and others, have an influence on what amount of a dose of drug actually reaches the brain ECF and CSF. Some drug characteristics also affect how the molecules further disperse in the CSF.

Apart from characteristics of the drug itself, certain physiological parameters have been shown to have an impact on drug distribution. These include, for example, cerebrospinal blood flow, pH of blood or tissue, and capillary surface area. Many diseases, such as Parkinson's disease, Alzheimer's disease, multiple sclerosis, or tumors, can cause or contribute to BBB dysfunction, which also affects the drug transport into the brain [1]. Other relevant physiological parameters are a persons's age, gender, health, and even diet [24]. We must consider this when comparing data obtained in different studies.

In our model, we will not distinguish between different modes of transport specifically.[2] Drug transport along the CNS will be reflected in three coefficients: the diffusion coefficient, κ; the CSF flow velocity, v; and the fraction of brain-ECF concentration reaching the CSF, c_{in}. The diffusion coefficient describes the ability of a drug to diffuse in the CSF: the larger the coefficient, the faster will the drug diffuse. The diffusion coefficient is

[2]Note, however, that the compartmental model in [39], which will be presented in section 3, does distinguish between active and passive transport.

equivalent to the square over the travelled distance of the drug molecules per time unit; it is dependant on drug characteristics including molecular weight and lipophilicity. The second coefficient is equivalent to the velocity of the CSF flow, which depends on a person's age, gender, and health. The third coefficient, c_{in}, depends on drug characteristics as well as physiological parameters. It is obvious that these coefficients need to be estimated or determined experimentally for every drug and person individually. Their computation will not be treated in this thesis.

3 Previous Models

The compartmental models introduced by Westerhout et al. in [38] and [39] are based on research in rats on the distribution of the drugs acetaminophen (paracetamol), methotrexate, and quinidine as examples of different modes of transport. Methotrexate is a chemotherapeutic agent used to treat certain kinds of cancer, such as breast or skin cancer, but it is also used in the treatment of rheumatoid arthritis. Quinidine is mainly used to treat irregular heart rates. While acetaminophen is transported by passive diffusion only, for methotrexate and quninidine several modes of transport are relevant. The models by Westerhout et al. in [38] and [39] are multi-compartment models. They divide the CNS into the following compartments:

1. the blood (plasma),
2. the brain ECF,
3. the lateral ventricles (CSF_{LV}),
4. the third and fourth ventricles (CSF_{TFV}),
5. the cisterna magna (CSF_{CM}),
6. the SAS (CSF_{SAS}),
7. the periphery.

The periphery represents tissues to which drugs can equilibrate, for example muscle or fat tissue. Depending on the drug, we can distinguish between fast and slow periphery, based on whether the equilibration is relatively fast or slow in a given kind of tissue. (Note that for quinidine, yet another compartment has been considered, the deep brain compartment. We will not treat this model in detail, since for quinidine no extrapolation for the human case is available.) Schematic representations of the models are shown in figure 5. In the case of acetaminophen, we can see an additional input F_{abs} into the plasma compartment besides the intravenous injection (i.v.). This is due to the fact that for acetaminophen we have to consider the so-called enterohepatic recirculation. It means that a certain portion of acetaminophen is not eliminated but reabsorbed into the blood, from where it enters the brain ECF and CSF again. For each of the compartments, an ordinary differential equation describing the change of the concentration over time is derived. The ODE for a given compartment in the models can be determined by means of the schematic representations shown in figure 5, using the following method:

Let N be the number of compartments and denote c_i the concentration in compartment i, $i = 1, \ldots, N$. Let c_k be the concentration in the compartment

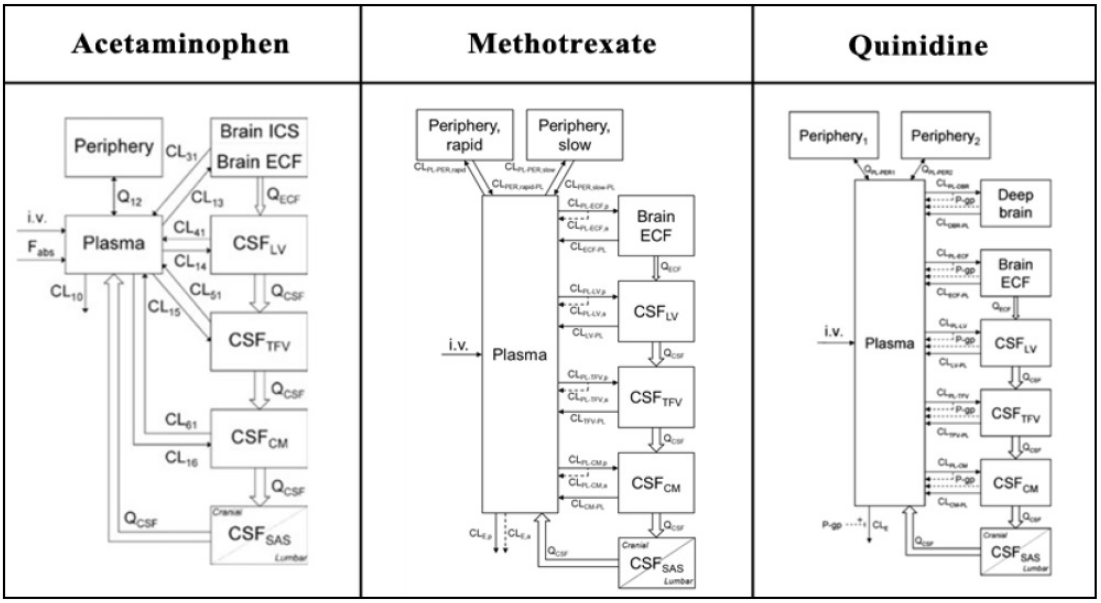

Figure 5: Schematic representation of the multiple-compartment models for acetaminophen, methotrexate, and quinidine. CL=clearence; striped arrows=active transport; CSF_{LV}=CSF in lateral ventricles; CSF_{TFV}=CSF in third and fourth ventricles; CSF_{CM}=CSF in cisterna magna; CSF_{SAS}=CSF in SAS; Q=flow rate of CSF and ECF respectively; i.v.=intravenous injection; F_{abs}=fraction of the dose reabsorbed by enterohepatic recirculation (acetaminophen). Human value of $F_{abs} = 0.06$ %/minute [38]. Reproduction of figure 6 in [14].

in question. The change in the concentration c_k is then given by

$$\frac{\partial c_k}{\partial t} = \sum_{i=1,\dots,N;i\neq k} (c_i p_{ik}) - \sum_{i=1,\dots,N;i\neq k} (c_k p_{ki}) \,, \tag{3.1}$$

where p_{ik} denotes the percentage of the concentration c_i that enters compartment k, and p_{ki} is the percentage of the concentration c_k that is given off to compartment i. The p_{ik} thus correspond to the inward pointing arrows while the p_{ki} correspond to the arrows pointing out of the compartment in question. Finding the ODE for each compartment yields a system of interdependent ODEs that need to be solved simultaneously.

The main problem with compartmental models is the underlying assumption that the concentration is the same in every point within the compartment. While this simplification may be justified for intra-brain compartments, which are quite small, neglecting the flow within the much larger SAS decreases

the accuracy of the general solution significantly. Taking the SAS as one compartment and neglecting inner local variation in the drug concentration means that the model predicts the same concentration-time profile for the neck and the lower back as both are part of the SAS compartment. This is inaccurate since it takes some time before the CSF containing the drug reaches the lumbar region in the lower back. Therefore it is desirable to find a more realistic solution which takes into account not only compartments but also the location within a given compartment. This is possible when instead of a system of ODEs we use a PDE, by means of which we can model the concentration in any point at any time. Moreover, the PDE can be applied to most drugs by substituting the corresponding coefficients, while the ODEs for compartmental models need to be derived for each drug individually due to the fact that the number of relevant compartments, the number of clearences that need to be considered for each compartment, and other factors may differ.

4 Modelling

We will now develop a simple model describing how the drug concentration in the CSF changes over time. We will restrict our model to the 1D case, since the width of the CSF space is quite small compared to its length and the CSF flow is unidirectional. First, we need to find the governing equation. To this end, we will use the information on CSF flow given in chapter 2. It is important to understand how the drugs enter the CSF and which pathways they take. In practice, drugs will be given orally or by venous injection. In both cases, a drug enters the brain via the blood across the BBB and the BCSFB. It then enters the CSF in the lateral ventricles by ECF bulk flow and diffusion across the ependymal layer, which separates brain ECF and CSF (see chapter 2.1). In this, bulk flow has a much stronger effect than diffusion. Moreover, drugs can enter the CSF directly from blood by transport across the BCSFB. Since the BCSFB is rather small and is assumed to play a minor role in drug transport into the brain, the input from transport across the BCSFB will be neglected. Thus, mathematically speaking, the input from brain ECF will be the source in our model. The PDE will describe the transport of the drug, which occurs not only by flowing along with the CSF but also by diffusion. By imposing suitable initial and boundary conditions and by assuming that the brain ECF concentration is known, we can then solve the resulting initial-boundary value problem and observe how the concentrations change over time along the CSF space. We will call this initial-boundary value problem our *forward* or *direct* problem. Subsequently, we are interested in the so-called *backward* equation or *inverse* problem, since, in fact, we do not know the initial concentrations in the brain ECF but rather seek to reconstruct them by using known lumbar-CSF values measured in the lower back. An introduction to inverse problems will be given in section 4.4.

4.1 Model equation

In order to find the governing equation, we will make simplified assumptions. In reality, the CSF space is U-shaped since it starts and ends in the head. However, we will assume it to be a straight pipe, which, in one dimension, can be modelled as an interval $[0, x_L]$. The lower bound of the interval now corresponds to the intra-brain compartments, while the upper bound corresponds to the arachnoid villi. The drug concentration at time t and place x is modelled by a function $c(x, t)$. We assume that the drug enters the CSF space at time $t = 0$ and that this exchange of brain ECF and CSF is only possible in the lateral ventricles, that is for small values of x. Moreover, it is presumed that the duration of the experiment is $T > 0$. If we assume that

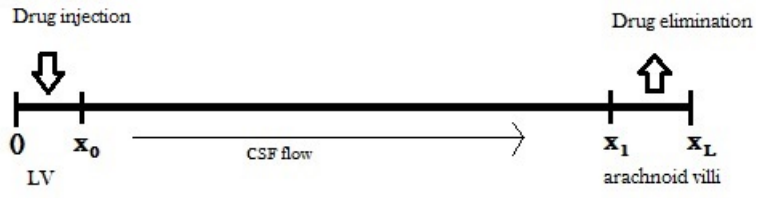

Figure 6: The interval $[0, x_L]$ and its subintervals. Thick arrows indicate injection or elimination of the drug. The thin arrow indicates the direction of the CSF flow. LV=lateral ventricles.

there is no elimination nor injection, then for each subinterval $[a, b] \subset [0, x_L]$ the change in concentration should be equal to the inflow minus the outflow on the boundaries. That is

$$\frac{d}{dt}\int_a^b c(x,t)dx = c(a,t)v(a,t) - c(b,t)v(b,t) = -\int_a^b \frac{\partial}{\partial x}(cv)(x,t)dx,$$

where $v(x,t)$ is the flow velocity. In practice, however, the equation must also describe the changes in concentration when drug is injected and when drug is eliminated by flowing back into the blood. Therefore, we must additionally consider a source term (injection) and a sink term (elimination) which will be modelled by a function $I(x,t)$. Exchange of brain ECF and CSF is not possible elsewhere than in the lateral ventricles, which means that the drug can enter the CSF space in the lateral ventricles only. The source term will thus be of the form $c_{in}c_{ecf}(t)$ in a subinterval $[0, x_0]$ of $[0, x_L]$ and equal to 0 everywhere else. Here, x_0 is small compared to x_L. The coefficient c_{in} describes the fraction of the brain-ECF concentration c_{ecf} that enters the CSF space, and the brain-ECF concentrations are assumed to be the same in any point of the extracellular space. Similarly, elimination is assumed to take place in the arachnoid villi only. Therefore the sink term should be of the form $c_{out}c(x,t)$ in another subinterval $[x_1, x_L]$ of $[0, x_L]$ and equal to 0 everywhere else. Here, x_1 is close to x_L so that the interval $[x_1, x_L]$ describes the aracnoid villi. The interval $[0, x_L]$ and its subintervals are shown in figure 6. Consequently, the function $c_{in}c_{ecf}(t)$ describes the amount of drug that can pass from brain ECF to CSF, and c_{out} is the fraction of the drug concentration $c(x,t)$ that leaves the system when CSF is eliminated by the arachnoid villi. Of course, c_{ecf}, c_{in} and c_{out} depend on the drug characteristics, as well as gender, age, health, and other parameters of the subject, and need to be determined for each case individually. We presume that all functions are smooth enough so we can switch integration and differentiation. We then

have

$$\int_a^b (\partial_t c + \partial_x(cv) - I)dx = 0.$$

By adding the above source and sink functions, we find I to be of the form

$$I(x,t) = \begin{cases} c_{in}c_{ecf}(t) & \text{for } 0 \le x < x_0 \\ 0 & \text{for } x_0 \le x \le x_1 \\ -c_{out}c(x,t) & \text{for } x_1 < x \le x_L. \end{cases} \tag{4.1}$$

Here, the subintervals of $[0, x_L]$ correspond to the lateral ventricles, the spinal SAS, and the arachnoid villi, respectively. Since $a,\ b$ were arbitrary, this can only be achieved for all intervals $[a, b]$ if

$$\partial_t c + \partial_x(cv) = I \quad \text{in } (0, x_L) \times (0, T).$$

We will neglect elimination in our model. This is not an unreasonable simplification since the flow is assumed to be unidirectional and the fluid passes the lumbar region before it reaches the arachnoid villi. We have presumed that there is no elimination in the spinal region of the SAS. Thus elimination, which only takes place in the arachnoid villi (that is in the head), does not affect the concentration in the lower back. Then the function I reduces to a function F, which is obtained by setting the coefficient c_{out} equal to 0. Note that, technically, the resulting equation describes the concentration in those parts of the system in which no elimination occurs, that is in the interval $[0, x_1]$. However, we will continue using x_L as the upper bound since the difference $|x_L - x_1|$ is small. Moreover, there can be diffusion, which is assumed to have a rather small impact on the concentration compared to the effect of advection, such that no drug diffuses against the flow direction. In order to incorporate the effect of diffusion, we can add to the equation a diffusion term according to Fick's second law:

$$\frac{\partial c}{\partial t} = \kappa \frac{\partial^2 c}{\partial x^2}.$$

So the whole equation becomes

$$\partial_t c + \partial_x(cv) - \kappa\, \partial_{xx} c = I \quad \text{in } (0, x_L) \times (0, T), \tag{4.2}$$

where κ is the diffusion coefficient, which we assume to be constant. The diffusion coefficient depends on properties of the CSF, but also on the drug in question. We will assume that the flow velocity v remains constant in the whole interval. Then

$$\frac{\partial c}{\partial t} = -v\frac{\partial c}{\partial x} + \kappa \frac{\partial^2 c}{\partial x^2} + I(x,t). \tag{4.3}$$

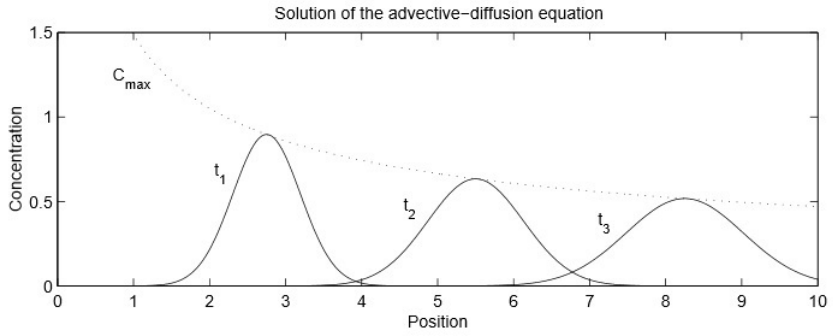

Figure 7: Schematic solution of the advection-diffusion equation in the (c, x)-plane for different values of t. The dotted line represents the maximum concentration as the wave moves downstream. Reproduction of figure 2.2 in [35].

Since the spine is wider in some parts, the velocity may vary in reality, but for now we assume the velocity v to be constant, using the average flow velocity obtained from the CSF turnover rate. In equation (4.3), the first term on the right hand side is called the advection term, while the second is called the diffusion or dispersion term. These represent the following effects on the concentration:

- Advection: This term represents the effect of the CSF flow. It means that a given concentration moves along with the CSF with velocity v.
- Diffusion: This term describes that the drug tends to diffuse to places where the concentration is lower. If this process is slow compared to the advection process, then κ is small.

The effects of advection and diffusion are depicted in figure 7 (=figure 2.2 in [35]), in which a schematic solution to the advection-diffusion equation is plotted in the (c, x)-plane for different values of t. The initial and boundary conditions for the direct problem will be given in section 4.3.1. Before we continue developping the model, let us have a look at the dimensions and scales that appear in the equation.

4.2 Dimensions and scales

So far, we have done the modelling without specifying the dimensions and scales of the terms appearing in the equation. We simply assumed that all terms in the PDE could be compared with each other and that no unreasonable

comparisons between different scales, for example length scale and time scale, took place. It is common to transform all parameters and variables into dimensionless variables. This reduces the number of variables and simplifies the computations since one need not worry about matching dimensions. In this model, however, we will choose to work with the original variables and parameters. Hence we must check if all terms in the equation are in the same dimension. So consider the PDE

$$\frac{\partial c}{\partial t}(x,t) = -v\frac{\partial c}{\partial x}(x,t) + \kappa\frac{\partial^2 c}{\partial x^2}(x,t). \tag{4.4}$$

In the following, we will use $[\cdot]$ to indicate dimension. Length scale is denoted by L, time scale by T, and mass by M.

First, consider the right hand side of (4.4). The concentration is given in mass per volume. Since the model is 1D, the volume scale L^3 reduces to L. Hence,

$$[c(x,t)] = \frac{[mass]}{[volume]} = \frac{[M]}{[L]}.$$

This implies

$$\left[\frac{\partial c}{\partial x}(x,t)\right] = \left[\frac{\partial}{\partial x}\right][c(x,t)] = \frac{[1]}{[L]}\frac{[M]}{[L]} = \frac{[M]}{[L^2]}.$$

The velocity v is given in length per time $[v] = \frac{[L]}{[T]}$, hence

$$\left[v\frac{\partial c}{\partial x}(x,t)\right] = \frac{[L]}{[T]}\frac{[M]}{[L^2]} = \frac{[M]}{[L][T]}.$$

The diffusion coefficient κ is given in $[\kappa] = \frac{[L^2]}{[T]}$, so

$$\left[\kappa\frac{\partial^2 c}{\partial x^2}\right] = [\kappa]\left[\frac{\partial^2}{\partial x^2}\right][c] = \frac{[L^2]}{[T]}\frac{[1]}{[L^2]}\frac{[M]}{[L]} = \frac{[M]}{[L][T]}.$$

We see that both terms on the right hand side are given in the same dimension. Now consider $\frac{\partial c}{\partial t}(x,t)$:

$$\left[\frac{\partial c}{\partial t}\right] = \frac{[1]}{[T]}\frac{[M]}{[L]} = \frac{[M]}{[L][T]}.$$

Thus, all terms in (4.4) are indeed given in the same dimension.

4.3 The direct problem

Before we define the inverse problem, let us have a closer look at the direct problem to get a better understanding of the behaviour of the solution. Consider the advection-diffusion equation

$$\frac{\partial c}{\partial t} = -v\frac{\partial c}{\partial x} + \kappa\frac{\partial^2 c}{\partial x^2} + F(x,t),$$

where $F(x,t)$ is the source term

$$F(x,t) = \begin{cases} c_{in}c_{ecf}(x,t) & \text{for } 0 \leq x < x_0 \\ 0 & \text{otherwise.} \end{cases}$$

In order to find a unique solution if possible, we need to impose some conditions on the function c.

4.3.1 Initial and boundary conditions

We have assumed that the drug first reaches the CSF space at time $t = 0$. It is therefore reasonable to assume that no drug is yet present throughout the CSF system at $t = 0$. This corresponds to the initial condition

$$c(x,0) = 0; \quad x \in (0, x_L). \tag{4.5}$$

Since the problem is a first-order problem with respect to t, one initial condition is enough. Since it is second order with respect to x, we have to impose two boundary conditions. If we take the simplest case, that is exchange between ECF and CSF is only possible in one point located in the lateral ventricles, then we can shift this point to be at $x = 0$. In this case, the source term $F(x,t)$ can be set to 0 and the first boundary condition becomes

$$c(0,t) = c_{in}c_{ecf}(t); \quad 0 < t < T. \tag{4.6}$$

The source is now incorporated as a boundary condition at $x = 0$. For the second boundary condition at $x = x_L$ note that the drugs tend to diffuse from regions of higher concentrations to those of lower concentrations. In a manner of speaking, diffusion has an equalizing effect on the drug concentration. It can be seen in figure 7 that the concentration-position profiles flatten for later times. This implies that for large values of x the concentration difference with respect to the space variable should be neglectable for all times: If t is small, the wave will not have reached the point yet; if t is large, the waves passing the point are very flat. Therefore, the concentration difference with respect

to the space variable x will be neglectable in $x = x_L$ if x_L is sufficiently large. (Remember that we neglect elimination, so the drug does not leave the CSF system.) Then, the second boundary condition is

$$\partial_x c(x_L, t) = 0; \quad 0 < t < T \tag{4.7}$$

due to the distance of the point x_L to the source.

4.3.2 Solution to the direct problem for $\kappa = 0$

In order to gain a rough idea about how c_{ecf} influences the concentration in the CSF system, we can solve the direct problem for a general initial profile $G(t) = c_{in}c_{ecf}(t)$ so that we obtain a general expression. Then we can plug in different functions for $c_{ecf}(t)$ according to the available data. Let us first study the simplest case possible. For this, we assume that there is no diffusion, the velocity v is constant, and exchange between brain ECF and CSF is only possible at one point in the lateral ventricles. Then we can use the above initial and boundary conditions, and the PDE reduces to the linear transport equation. Thus, the problem becomes

$$\frac{\partial c}{\partial t} = -v\frac{\partial c}{\partial x}; \qquad t > 0;\ x \in (0, x_L) \tag{4.8}$$

$$c(x, 0) = 0; \qquad x \in (0, x_L) \tag{4.9}$$

$$c(0, t) = c_{in}c_{ecf}(t) =: G(t); \qquad 0 < t < T \tag{4.10}$$

where v is the velocity of the CSF and where the source is now included as an initial condition at $x = 0$. This system can be solved using the method of characteristics. Suppose that K is a parametrized curve $(x(s), t(s))$ in the (x, t)-plane. For K we have that

$$c(x, t) = c(x(s), t(s)).$$

Differentiation on K with respect to s yields that

$$\frac{dc}{ds} = \frac{\partial c}{\partial x}\frac{dx}{ds} + \frac{\partial c}{\partial t}\frac{dt}{ds}.$$

Now, comparing this with (4.8), we see that if

$$\frac{dx}{ds} = v, \quad \frac{dt}{ds} = 1, \tag{4.11}$$

we also have

$$\frac{dc}{ds} = 0 \text{ on } K.$$

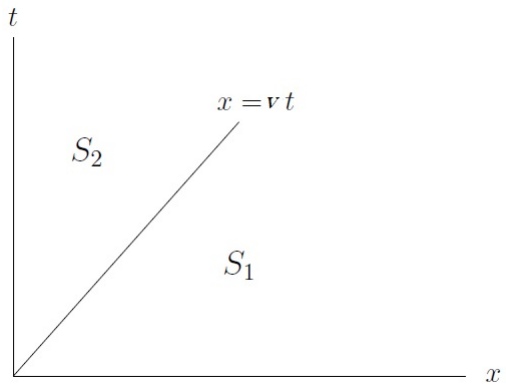

Figure 8: The areas S_1 and S_2.

K is thus indeed a characteristic curve of equation (4.8). Along a characteristic, the concentration c is constant. From (4.11) we derive that

$$x(s) = vs + x_0; \quad t(s) = s + t_0; \tag{4.12}$$

where $x_0 = x(0),\ t_0 = t(0)$. Also, since $\frac{dc}{ds} = 1$, we find that $dt = ds$, so we can write $x = vt + x_0$. Since the solution c is constant along each characteristic, the solution only depends on the value of $x - vt$ or $t - \frac{x}{v}$. So $c(x,t)$ is of the form

$$c(x,t) = h(x - vt) \text{ or } c(x,t) = H\left(t - \frac{x}{v}\right),$$

where h, H are arbitrary differentiable functions. These functions can be determined by the initial and boundary conditions (4.9), (4.10). By (4.9), the function h is only known for positive arguments, that is for $x > vt$, while H is determined by (4.10) and the condition only gives information on H for $0 < x < vt$, that is for $t - \frac{x}{v} > 0$. So we can define $S_1 = \{(x,t) \mid x > vt\},\ S_2 = \{(x,t) \mid 0 < x < vt\}$ as shown in figure 8. In S_1, the solution is determined by (4.9), yielding $h(t) = 0$ in S_1. In S_2, the solution is determined by (4.10), yielding $H(t) = c_{in}c_{ecf}(t - \frac{x}{v})$ in S_2.

Thus, the solution is

$$c(x,t) = \begin{cases} 0; & x > vt \\ c_{in}c_{ecf}(t - \frac{x}{v}); & 0 < x < vt \end{cases}. \tag{4.13}$$

From this, we see that the drug moves along with the fluid with velocity v. For $x > vt$, no concentration can be measured since the fluid containing the drug has not passed this point yet. The solutions behave like travelling

waves. We can see this simply by using some function for c_{ecf} and plotting the concentration-time profile defined by (4.13) for different values of x. Note that we get the same solution by solving the inhomogeneous transport equation

$$\begin{aligned} &\partial_t c + v\partial_x c = c_{in}c_{ecf}(t)\delta(x); \quad 0 < t < T,\ x \in (0, x_L) \\ &c(x, 0) = 0; \quad x \in (0, x_L). \end{aligned}$$

Here, $\delta(x)$ is the delta distribution, which is zero everywhere except in $x = 0$. This indicates that the source is inactive everywhere except in the point $x = 0$. We call this kind of source a point source. More properties of the delta distribution are given in appendix B. Using the method of characteristics, we find that

$$\begin{aligned} c(x,t) &= \int_0^t c_{in}c_{ecf}(s)\delta(x + v(s - t))\, dt \\ &= c_{in}c_{ecf}(\{s \mid x + v(s - t) = 0\}) \\ &= c_{in}c_{ecf}(t - \frac{x}{v}) \end{aligned}$$

by the integral property of the delta distribution.

Now let us find a suitable function $G(t)$ with $G(t) = c_{in}c_{ecf}(t)$ as defined above. Westerhout et al. [38] plotted rat data for acetaminophen distribution, which results in figure 9. Note that only plasma and intra-brain ECF and CSF concentrations are plotted here, so the effects of the advection term are almost invisible due to the short distances and coarse time scale. Therefore, there is no travelling wave in this figure. We can see that all graphs have a similar shape, which is almost Gaussian. Hence, we can use the function

$$G(t) = c_{in}c_{ecf}(t) = c_{in}\frac{1}{\sqrt{4\pi Dt}} \exp\left(\frac{t - A}{4Dt}\right)$$

with $A = 0.8$ as initial profile. By choosing different values for A, it is, of course, possible to modify the shape a bit. Also, by choosing different values for D and c_{in} or by multiplying the function $G(t)$ by a suitable constant, we can make the plots look more similar to the actual data. But for simplicity, and because we are interested in the shape of the graph rather than the actual values, we take $c_{in} = v = 1$, $A = 0.8$ and $D = 0.1$. Then, according to (4.13), the solution is

$$G(t - \frac{x}{v}) = \frac{1}{\sqrt{4\pi D(t - \frac{x}{v})}} \exp\left(\frac{(t - \frac{x}{v}) - 0.8}{4D(t - \frac{x}{v})}\right). \tag{4.14}$$

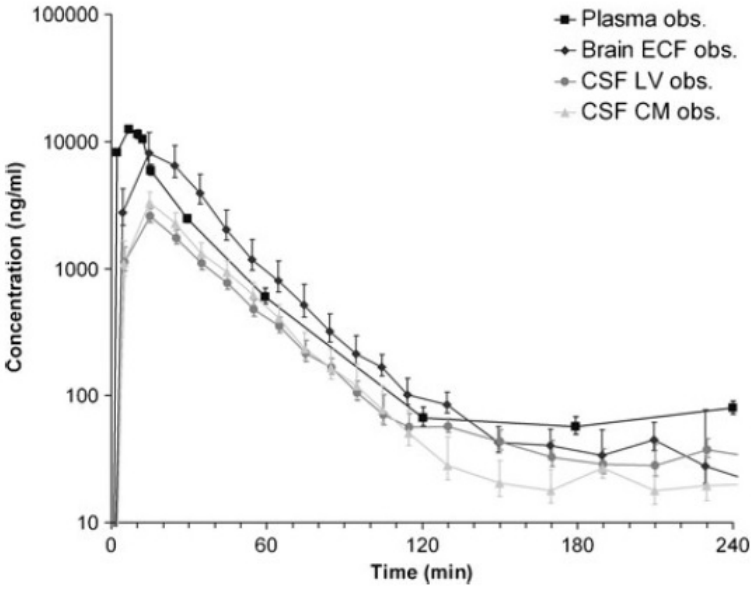

Figure 9: Observed rat data for acetaminophen. Average free drug concentration-time profiles for plasma (n=10), brain ECF (n=10), CSF LV (n=14), and CSF CM (n=8); n=number of subjects (sample size). Reproduction of figure 1 in [38].

The function $G(t)$ and the solutions for $v = 1,\ D = 0.1$, and $x = 1, 5$ are shown in figure 10. We can see that if no diffusion, reaction, or elimination takes place, the initial profile simply travels along with the fluid without changing in shape, so that in each point x we can see the same profile simply shifted to the right. This is reasonable, since without diffusion or elimination present in the system the drug does not spread nor leave the system.

In practice, the concentration-time profile obtained by plotting the available rat data has a similar shape in each point x, increasing rather fast over a short period in the beginning and then decaying possibly exponentially over a longer period of time. However, the speed at which this happens and the intensity of the peak differ. For the case of rats, SAS data is not available, but there are predictions based on the model presented in [38]. The predicted concentration-time profiles for rats obtained by means of the model in [38] are shown in figure 11. The predictions for brain ECF and intra-brain CSF suggest that the model is quite accurate. For this compare figures 9 and 11. We therefore expect the concentration-time profile in the SAS to be similar to that shown in figure 11. The shape of the curve of the predicted CSF concentrations in the SAS differs from the other curves. This is probably due

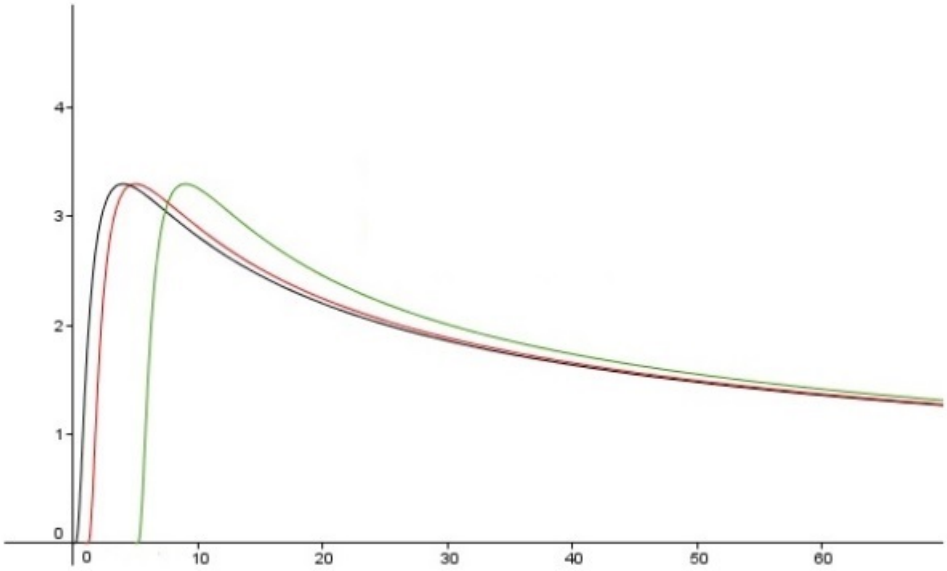

Figure 10: Travelling-wave solutions (4.14) for $x = 0$ (black), $x = 1$ (red), and $x = 5$ (green). Parameters are $v = 1$, $D = 0.1$.

to effects of recirculation or diffusion, or possibly reaction or elimination. If the effects of diffusion are responsible for the change of the shape, we should see a similar shape for larger values of x when plotting the solution for $\kappa \neq 0$.

In case of other drugs, for which no recirculation takes place, the drug concentration tends to 0 for larger values of t. The observed data obtained from different studies and in different species as well as the predictions obtained by the model for methotrexate introduced in [39] are shown in figure 12, which is a reproduction of figure 7 in [39]. The curves for methotrexate representing brain-ECF and SAS concentrations in different species look more similar to each other than in the case of acetaminophen. Compare for this the second and last rows in figure 12. This might indicate that the recirculation has an impact on the shape of the SAS curve, or it might be due to the different modes of transport across the BBB and BCSFB or other drug properties such as molecular size, or a different diffusion coefficient. If diffusion is responsible for the change in the shape of the acetaminophen concentration-time profiles for larger values of x, we should see a similar effect when plotting solutions for $\kappa \neq 0$ (as above). We can see the expected travelling-wave effects in the SAS concentrations in figures 12D ("healthy dogs") and 12E ("diseased children").

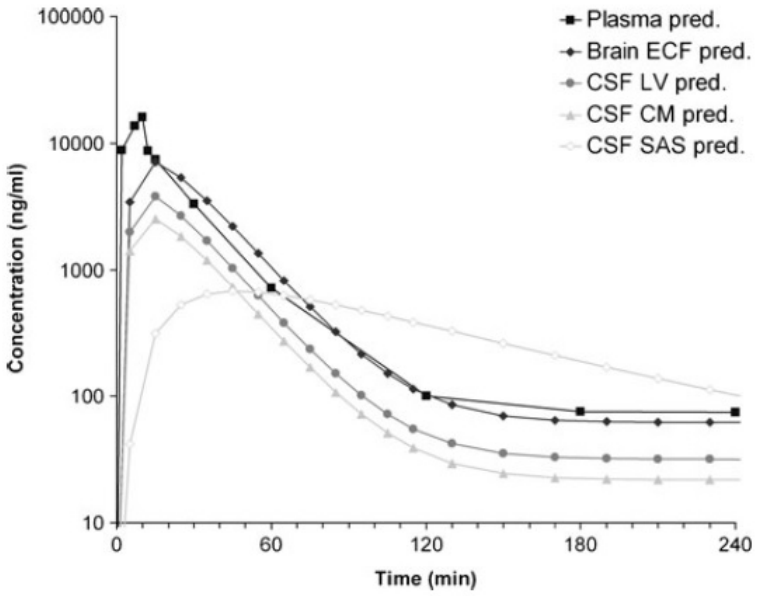

Figure 11: Predicted acetaminophen concentrations in plasma, brain compartments, and SAS. Predictions are based on the multi-compartment model in [38] and a 10-minute infusion of 15 mg/kg in a rat weighing 250 g. Reproduction of figure 4 in [38].

4.3.3 Solution to the direct problem for $\kappa \neq 0$

Now let us include the diffusion term into the model and see what effect it has on the behaviour of the solution. We wish to find a solution to the problem

$$\partial_t c(x,t) + v\partial_x c(x,t) - \kappa \partial_{xx} c(x,t) = 0; \quad x \in (0, x_L), t \in (0,T) \tag{4.15}$$
$$c(x,0) = 0; \quad x \in (0, x_L) \tag{4.16}$$
$$c(0,t) = G(t),\ \partial_x c(x_L,t) = 0; \quad t \in (0,T). \tag{4.17}$$

Using the transformation

$$c(x,t) = \Gamma(x,t) \exp\left(\frac{vx}{2\kappa} - \frac{v^2 t}{4\kappa}\right) \tag{4.18}$$

as proposed by Ogata and Banks in [30], equation (4.15) with respect to Γ can be written as

$$\partial_t \Gamma = \kappa \partial_x^2 \Gamma. \tag{4.19}$$

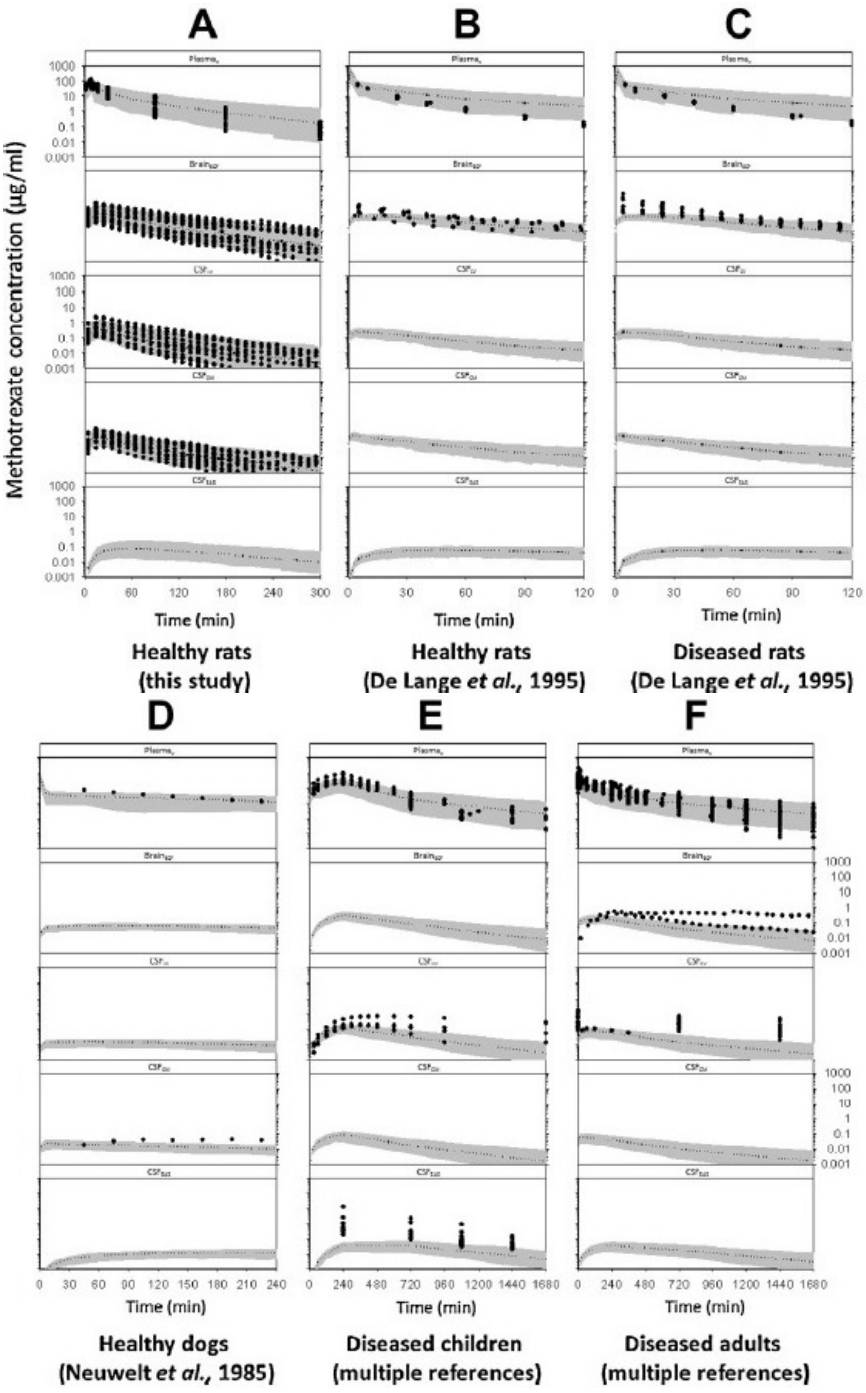

Figure 12: Observed and predicted methotrexate plasma, brain ECF, CSF LV, CSF CM, and CSF SAS concentration-time profiles in different species. Black dots are observed data. Grey areas show the 95% confidence interval. Predictions are based on the multi-compartment model presented in [39]. Note the differences in the time scales. Reproduction of figure 7 in [39].

The corresponding initial and boundary conditions are

$$\begin{aligned}
&\Gamma(x,0) = 0 \\
&\Gamma(0,t) = G(t)\exp\left(\frac{v^2 t}{4\kappa}\right) =: \phi(t) \\
&\partial_x\Gamma(x_L,t) + \frac{v}{2\kappa}\Gamma(x_L,t) = 0;\ x_L \to \infty.
\end{aligned} \tag{4.20}$$

Now we want to use the method of Laplace transformation as presented in chapter 12.3 in Carslaw and Jaeger [7] to solve the heat equation (4.19). Multiplying (4.19) by e^{-pt} and integrating from 0 to ∞ yields the ODE

$$\frac{d^2\bar{\Gamma}}{dx^2} = \frac{p}{\kappa}\bar{\Gamma}, \tag{4.21}$$

where

$$\bar{\Gamma}(x) = \int_0^\infty e^{-pt}\Gamma(x,t)\,dt.$$

The initial and boundary conditions are transformed in the same way, that is,

$$\bar{\Gamma}(x,0) = \int_0^\infty e^{-pt}\Gamma(x,0)\,dt = 0$$

$$\bar{\Gamma}(0,t) = \int_0^\infty e^{-pt}\Gamma(0,t)\,dt = \int_0^\infty e^{-pt}G(t)\exp\left(\frac{v^2 t}{4\kappa}\right)\,dt = \int_0^\infty G(t)\exp\left(\left(\frac{v^2}{4\kappa} - p\right)t\right)\,dt$$

$$\int_0^\infty e^{-pt}\left(\partial_x\Gamma(x_L,t) + \frac{v}{2\kappa}\Gamma(x_L,t)\right) = 0;\ x_L \to \infty.$$

The general solution to (4.21) is given by

$$\bar{\Gamma}(x) = Ae^{-qx} + Be^{qx},$$

where $q = \sqrt{\frac{p}{\kappa}}$. We want Γ and $\bar{\Gamma}$ to be bounded, so $B = 0$. Now we can use Duhamel's principle (see 1.14 (10) in [7]). It states that instead of solving problem (4.19) for $\Gamma(0,t) = \phi(t)$ as defined by (4.20), we may solve the problem for $\Gamma(0,t) = 1$ first. More precisely: Let $f(x,t)$ be the solution to (4.19) with $\Gamma(0,t) = 1$ and the other two conditions as in (4.20). Then, for $\Gamma(0,t) = \phi(t)$, the solution of (4.19) is given by

$$\Gamma = \int_0^t \phi(\tau)\frac{\partial}{\partial t}f(x,t-\tau)\,d\tau. \tag{4.22}$$

So let $\Gamma(0,t) = 1$. Since

$$\Gamma(0,t) = 1 \Rightarrow \bar{\Gamma}(0) = \int_0^\infty e^{-pt}\Gamma(0,t) = -\frac{1}{p}\left[e^{-pt}\right]_0^\infty = \frac{1}{p},$$

we find that

$$\bar{\Gamma}(x) = \frac{1}{p} e^{-\sqrt{\frac{p}{\kappa}}x}.$$

Using appendix V (8) in [7], we find Γ as

$$\Gamma(x,t) = \operatorname{erfc}\left(\frac{x}{2\sqrt{\kappa t}}\right) = \frac{2}{\sqrt{\pi}} \int_{\frac{x}{2\sqrt{\kappa t}}}^{\infty} e^{-\eta^2}\, d\eta.$$

Then, by Duhamel's principle (4.22), the solution to the original problem with $\Gamma(0,t) = \phi(t)$ is given by

$$\Gamma = \int_0^t \phi(\tau) \frac{\partial}{\partial t} \left[\frac{2}{\sqrt{\pi}} \int_{\frac{x}{2\sqrt{\kappa(t-\tau)}}}^{\infty} e^{-\eta^2}\, d\eta \right] d\tau. \tag{4.23}$$

Since $e^{-\eta^2}$ is continuous, we can differentiate under the integral, using the formula

$$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)\, dt\right) = f(x,b(x))b'(x) - f(x,a(x))a'(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)\, dt.$$

Then,

$$\frac{\partial}{\partial t} \frac{2}{\sqrt{\pi}} \int_{\frac{x}{2\sqrt{\kappa(t-\tau)}}}^{\infty} e^{-\eta^2}\, d\eta = \frac{x}{2\sqrt{\kappa\pi}(t-\tau)^{3/2}} e^{-\frac{x^2}{4\kappa(t-\tau)}}.$$

Thus,

$$\Gamma(x,t) = \frac{x}{\sqrt{\pi\kappa}} \int_0^t \phi(\tau) e^{\frac{-x^2}{4\kappa(t-\tau)}} \frac{d\tau}{(t-\tau)^{3/2}}.$$

Substitute

$$\xi = \frac{x}{2\sqrt{\kappa(t-\tau)}}.$$

Then,

$$\Gamma(x,t) = \frac{2}{\sqrt{\pi}} \int_{\frac{x}{2\sqrt{\kappa t}}}^{\infty} \phi\left(t - \frac{x^2}{4\kappa\xi^2}\right) e^{-\xi^2}\, d\xi. \tag{4.24}$$

We have thus derived a general formula for the solution of (4.19) with the initial and boundary conditions (4.20). In order to understand how concentration-time profiles behave for different choices of $c(0,t) = G(t)$, let us first discuss what types of functions are relevant in applications. If a drug is given orally or by a single injection into the blood, the concentration will increase over a short period of time until it reaches a maximum. With most drugs, the concentration will then decrease exponentially over a longer period of time [32]. This results in concentration-time profiles as depicted in figure 13. The

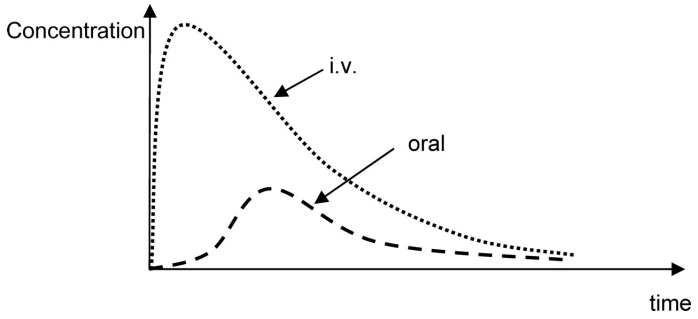

Figure 13: Schematic plasma concentration-time profile for intrevenous injection (i.v) and oral taking. Reproduction of figure 4 in [32].

curves for acetaminophen differ somewhat from the other ones. Due to the effects of recirculation, the concentration decreases so slowly for large values of t that, in the relevant time period, it seems to approach a limiting value > 0. But let us for now restrict the discussion to drugs whose concentration-time profiles in plasma look similar to the curves in figure 13. If a drug is given at a constant rate over a longer period of time, typically by infusion, the concentration in plasma will eventually reach a steady state. Once the infusion is stopped, the concentration will start decreasing exponentially (see figure 14).[3] Since $c(0,t)$ is of the form $c_{in}c_{ecf}(t)$ and the concentration-time profiles for brain ECF look similar to those for plasma in most studies, we can conclude from figures 13 and 14 that there are three cases to be considered:

1 the increase of drug concentration up to the maximum concentration (drug aggregation);

2 the steady state;

3 the elimination of the drug.

Of these, case 3 corresponds to $c(0,t) = c_0 \exp(-\lambda t)$ for a constant $c_0 > 0$. Case 2 corresponds to an initial concentration $c(0,t) = c_0$. For case 1, we may choose to use a function of the form $c_0 - c_0 \exp(-\lambda t)$. Let us first consider the simplest case, in which we assume a stable concentration level in the brain ECF. This corresponds to a situation where the steady state has been reached and a maintenance dose is given in order to keep the drug concentrations constant in the brain ECF. The question is how the concentration in the CSF changes along the CSF space if the concentration in brain ECF remains

[3]Graphic taken from slide 27 in
`http://de.slideshare.net/PallaviKurra/compartment-modelling`

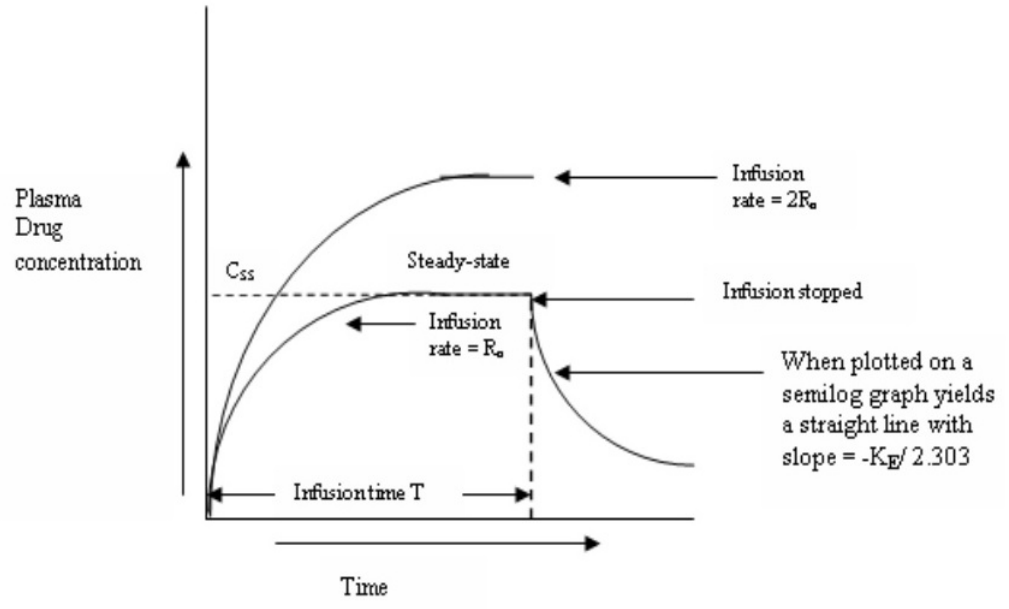

Figure 14: Schematic concentration-time profiles in plasma if steady-state concentration is reached.

constant. Let us assume the following problem.

$$\begin{aligned} &\partial_t c = -v\partial_x c + \kappa \partial_{xx} c \\ &c(0,t) = c_0; \quad 0 \le t \\ &c(x,0) = 0; \quad 0 < x < x_L \\ &\partial_x c(x_L,t) = 0; \quad t \ge 0. \end{aligned} \tag{4.25}$$

For x_L sufficiently large, the last condition is merely an approximation for $\partial_x c(\infty, t) = 0$. For simplicity, let us therefore first examine the above problem with the second boundary condition replaced by $\partial_x c(\infty, t) = 0$. Applying the evaluation of the integral used by Ogata and Banks in [30], we find that the solution can be given by

$$c(x,t) = \frac{c_0}{2}\left(\operatorname{erfc}\left(\frac{x-vt}{2\sqrt{\kappa t}}\right) + \exp\left(\frac{vx}{\kappa}\right)\operatorname{erfc}\left(\frac{x+vt}{2\sqrt{\kappa t}}\right)\right)$$

through substituting $\phi(t) = G(t)\exp\left(\frac{v^2 t}{4\kappa}\right) = c_0 \exp\left(\frac{v^2 t}{4\kappa}\right)$ in (4.24). Solutions for $\kappa = 0.7$, $v = 0.5$ and for different values of x are shown in figure 15. We can see that the curves flatten for larger values of x, which implies that the drug aggregation decelerates. After some time, the steady state-concentration

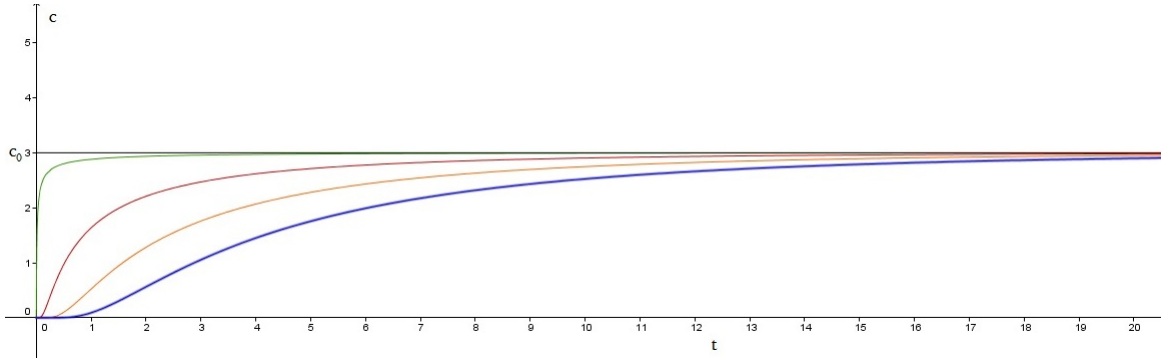

Figure 15: Solutions to problem 4.25 with $\kappa = 0.7$, $v = 0.5$, $c_0 = 3$ for $x = 0$ (black), $x = 0.1$ (green), $x = 1$ (red), $x = 2$ (orange), and $x = 3$ (blue).

in the CSF is reached for all values of x. We also see the travelling-wave effect: the curve remains at 0 for $x = 3$ and $t < 0.5$.

Now let us consider the case where $c(0,t) = c_0$ for $0 < t < t_0$, but $c(0,t) = 0$ for $t > t_0$. This case is similar to the former, but the maintenance dose is now given for $t < t_0$ only. The infusion is abruptly stopped at $t = t_0$ and the brain-ECF concentration, that is the input from brain ECF to CSF, instantly drops to 0. Then, the solution is given by

$$c(x,t) = \begin{cases} c_0 A(x,t) & 0 < t < t_0 \\ c_0 A(x,t) - c_0 A(x,t-t_0) & t > t_0, \end{cases} \tag{4.26}$$

where

$$A(x,t) = \frac{1}{2}\left(\operatorname{erfc}\left(\frac{x-vt}{2\sqrt{\kappa t}}\right) + \exp\left(\frac{vx}{\kappa}\right)\operatorname{erfc}\left(\frac{x+vt}{2\sqrt{\kappa t}}\right)\right).$$

This is obtained by solving the problem for $0 < t < t_0$ and $t_0 < t$ separately using the solution of 4.25 and matching both functions at $t = t_0$.

The solutions for $v = 0.5$, $\kappa = 1$, and $\kappa = 0.1$ for different values of x are shown in figures 16 and 17. We find that for smaller values of x (especially $x = 1$) the shape of the resulting solution looks quite similar to that shown in figure 14 (and the intravenous injection in figure 13). For $\kappa = 1$, a strong flattening or equalizing effect can be seen, while the travelling-wave effect is stronger for $\kappa = 0.1$ than for $\kappa = 1$.

In the elimination phase (case 3 above), exponential decay takes place. We will therefore use $c(0,t) = c_0 \exp(-\lambda t)$ with $c_0 = 3$, $\lambda = 1$. This case corresponds to the drug being present in the CSF in the lateral ventricles at some concentration c_0. At $t = 0$, the injection is stopped and the elimination

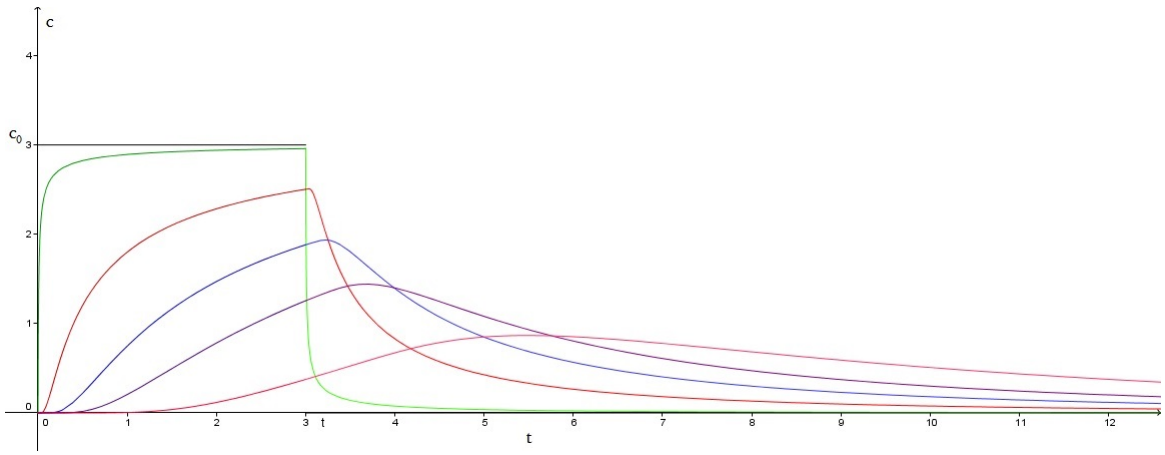

Figure 16: Solutions of problem (4.25) with $c(0,t) = 3$ only if $t < 3$ according to (4.26) for $x = 0$ (black), $x = 0.1$ (green), $x = 1$ (red), $x = 2$ (blue), $x = 3$ (purple), and $x = 5$ (pink). The coefficients are $v = 0.5, \kappa = 1$.

begins. This case is relevant for concentration-time profiles on coarse time scales. Since the drug aggregation happens quite fast in the case of intraveneous injection, it will not be visible on such a time scale and we will see only the elimination phase in the plot. Due to the initial condition $c(x,0) = 0$ for $x \in (0, x_L)$, the shape of profile i.v. in figure 13 is restored for all $x > 0$, that is $c(x,0) = 0$ for all $x > 0$. Substituting for $\phi\left(t - \frac{x^2}{4\kappa t}\right)$ yields

$$\begin{aligned}\Gamma &= \frac{2c_0}{\sqrt{\pi}} \int_\alpha^\infty \exp\left(\left(\frac{v^2}{4\kappa} - \lambda\right) t + \frac{\lambda x^2}{4\kappa\xi^2} - \frac{\epsilon^2}{\xi^2} - \xi^2\right) d\xi \\ &= \frac{2c_0}{\sqrt{\pi}} e^{\left(\frac{v^2}{4\kappa} - \lambda\right)t} \int_\alpha^\infty \exp\left(\frac{\lambda x^2}{4\kappa\xi^2} - \frac{\epsilon^2}{\xi^2} - \xi^2\right) d\xi,\end{aligned}$$

where $\alpha = \frac{x}{2\sqrt{\kappa t}}$, $\epsilon = \frac{vx}{4\kappa}$. Now, since the first two terms in the exponential both depend on $\frac{1}{\xi^2}$, we may introduce a new variable μ such that

$$\mu^2 = \epsilon^2 - \frac{\lambda x^2}{4\kappa}. \tag{4.27}$$

Then, the above integral can be written as

$$\int_\alpha^\infty \exp\left(-\xi^2 - \frac{\mu^2}{\xi^2}\right) d\xi. \tag{4.28}$$

According to [30], we find that

$$\Gamma(x,t) = 2\frac{c_0}{\sqrt{\pi}} e^{(\frac{v^2}{4\kappa} - \lambda)t} \int_\alpha^\infty \exp\left(-\xi^2 - \frac{\mu^2}{\xi^2}\right) d\xi$$

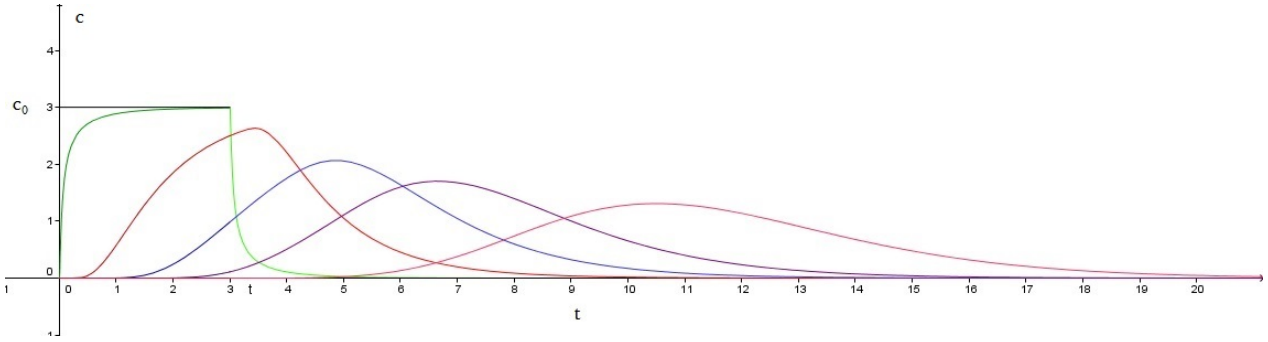

Figure 17: Solutions of the problem with $c(0,t) = 3$ if $t < 3$ according to 4.26 for $x = 0$ (black), $x = 0.1$ (green), $x = 1$ (red), $x = 2$ (blue), $x = 3$ (purple), and $x = 5$ (pink). The coefficients are $v = 0.5, \kappa = 0.1$.

$$
\begin{aligned}
&= 2\frac{c_0}{\sqrt{\pi}} e^{(\frac{v^2}{4\kappa}-\lambda)t} \left[\frac{\sqrt{\pi}}{2} e^{-2\mu} - \frac{1}{2} \left(e^{-2\mu} \int_{\frac{\mu}{\alpha}-\alpha}^{\infty} e^{-\beta}\, d\beta - e^{2\mu} \int_{\frac{\mu}{\alpha}+\alpha}^{\infty} e^{-\beta}\, d\beta \right) \right] \\
&= \frac{c_0}{2} e^{(\frac{v^2}{4\kappa}-\lambda)t} \left(e^{2\mu} \mathrm{erfc}\left(\alpha + \frac{\mu}{\alpha} \right) + e^{-2\mu} \mathrm{erfc}\left(\alpha - \frac{\mu}{\alpha} \right) \right).
\end{aligned}
$$

Since $c(x,t) = \Gamma(x,t) \exp\left(\frac{vx}{2\kappa} - \frac{v^2 t}{4\kappa}\right)$, we find that

$$
\begin{aligned}
c(x,t) &= \frac{c_0}{2} e^{-\lambda t} \left(e^{\frac{vx}{2\kappa}+2\mu} \mathrm{erfc}\left(\alpha + \frac{\mu}{\alpha} \right) + e^{\frac{vx}{2\kappa}-2\mu} \mathrm{erfc}\left(\alpha - \frac{\mu}{\alpha} \right) \right) \\
&= \frac{c_0}{2} e^{-\lambda t} \left(e^{\frac{(v-y)x}{2\kappa}} \mathrm{erfc}\left(\frac{x - yt}{2\sqrt{\kappa t}} \right) + e^{\frac{(v+y)x}{2\kappa}} \mathrm{erfc}\left(\frac{x + yt}{2\sqrt{\kappa t}} \right) \right),
\end{aligned}
$$

where $y = \sqrt{v^2 - 4\lambda\kappa}$. This is the same result as for problem A9 in [16]. Note that the expression for y is a restriction on the values of v, κ, and λ in order to give reasonable results. If $v^2 - 4\lambda\kappa < 0$, then y is imaginary.

The corresponding solutions for $\kappa = 0.2$, $v = 1$ are shown in figure 18 for different values of x. We can see that the resulting solution for $x = 0.1$ looks very similar to the expected concentration-time profiles in plasma or brain-ECF and intra-brain compartments as shown for intravenous injection in figure 13, while the plots for $x \geq 2$ already look similar to the predicted concentration-time profile in the SAS in figure 11, only without the pseudo-limiting value caused by the recirculation.

For case 1 (drug aggregation), we may now use $c(0,t) = c_0 - c_0 e^{-\lambda x}$. This case corresponds to the case in which the drug is given continuously over a longer period of time. The drug reaches the CSF at time $t = 0$ and no drug was present in the brain ECF and CSF before that. The injection is continued long enough in order to reach the steady-state concentration and is

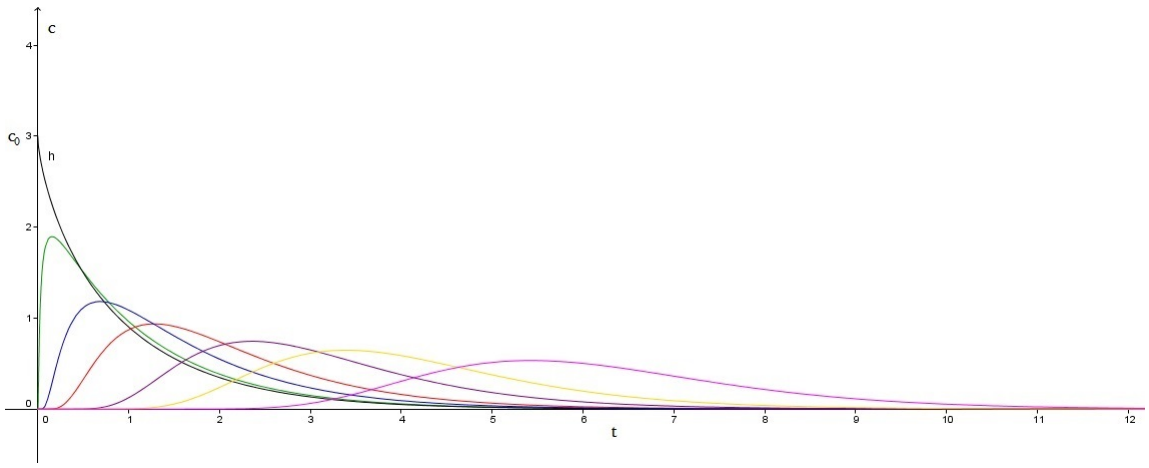

Figure 18: Solutions for drug elimination for $x = 0$ (black), $x = 0.1$ (green), $x = 0.5$ (blue), $x = 1$ (red), $x = 2$ (purple), $x = 3$ (yellow), $x = 5$ (pink). The coefficients are $v = 1$, $\kappa = 0.2$.

not stopped within the period of time considered. The solution is obtained by adding the solutions for $c(0, t) = c_0$ and for $c(0, t) = -c_0 e^{-\lambda t}$. The solutions for $c_0 = 3$, $\lambda = 1$, $\kappa = 0.2$, $v = 1$, and for different values of x are shown in figure 19. The solutions look similar to those for larger values of x in figure 15. This is reasonable because, once the steady state has been reached, the situation is the same as in figure 15.

Let us now look at the solution with the boundary condition $\partial_x c(x_L, t) = 0$ in place. In figures 16 and 17, we plotted the approximate solution for the boundary condition $\partial_x c(\infty, t) = 0$. When we use the formula for the approximate solution to problem A3 in [16] for the finite case with $v =$

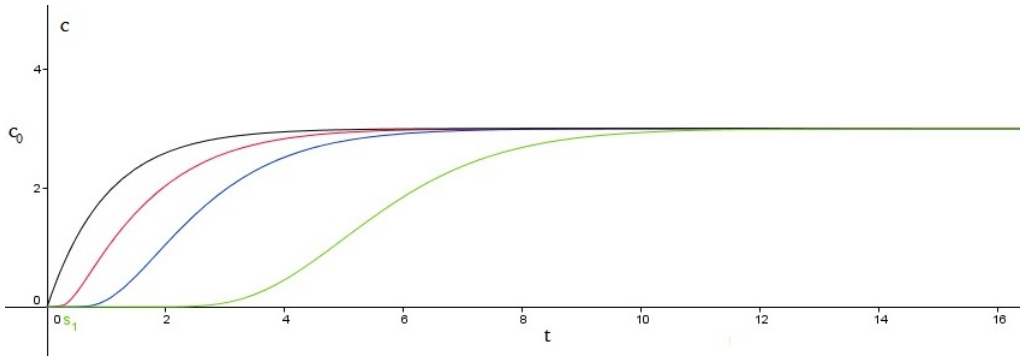

Figure 19: Solutions for drug aggregation for x=0 (black), x=1 (red), x=2 (blue), x=5 (green). The coefficients are $v = 1$, $\kappa = 0.2$

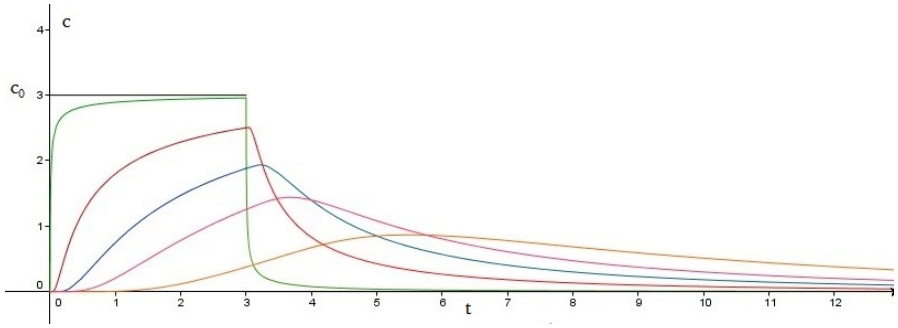

Figure 20: Solutions for finite domain ($x_L = 15$) in the case of $c(0,t) = 3$ for $0 < t < t_0$ for $x = 0$ (black), $x = 0.1$ (green), $x = 1$ (red), $x = 2$ (blue), $x = 3$ (pink), $x = 5$ (orange). The coefficients are $v = 0.5$ and $\kappa = 1$.

0.5, $\kappa = 1$, and $x_L = 15$, this yields the results shown in figure 20. A comparison with figure 16 indicates that the approximation obtained by using $\partial_x c(\infty, t) = 0$ instead of $\partial_x c(x_L, t) = 0$ is quite accurate in this case. See also figure 21 for the plots for x close to $x_L = 15$. We can see from part B of this figure that even if $x_L = 15$ the curve for $\kappa = 1$ and $x = 14$ is very flat. In practice, x_L will be larger than 14 in adults, so the approximation $\partial_x c(\infty, t) = 0$ instead of $\partial_x c(x_L, t) = 0$ seems sufficiently accurate.

For completeness' sake, let us look at the concentration-position profiles for different values of t. In figure 22, the solutions are plotted for $c(0,t) = 3$ for $\kappa = 0.2$ and $\kappa = 1$. By comparing A and B, we see that with a higher diffusion coefficient the curves become flatter as the drug spreads more. In A, it takes the wave longer to reach later positions (larger values of x): for $t = 15$ there is no drug present at $x = 14$, while in B the curve for $t = 15$ at $x = 14$ is clearly above 0. The concentration-position profiles for

$$c(0,t) = \begin{cases} 3, & 0 < t < t_0 = 3 \\ 0, & t > t_0 \end{cases}$$

and different values of t are shown in figure 23 for $v = 0.5$ and $\kappa = 0.2$ and $\kappa = 1$. We can see the jump from $c(0,t) = 3$ to $c(0,t) = 0$. If $\kappa = 1$, the curve for $t = 7$ is quite flat already, so it will be flat enough to justify the assumption $\partial_x c(x_L, t) = 0$ if x_L is large. If κ is quite small, the same effect will eventually occur provided that x_L is large enough. In the case of drug elimination, that is for $c(0,t) = 3e^{-t}$ as used above, we obtain the curves depicted in figure 24 for $v = 1$ and $\kappa = 0.2$. The effects of advection and diffusion are clearly visible in this figure: The wave travels along with the fluid and the drug spreads, which leads to a flattening of the curves. The concentration at $x = 0$ decreases from close to 3 (for $t = 0.1$) to 0 (e.g. for

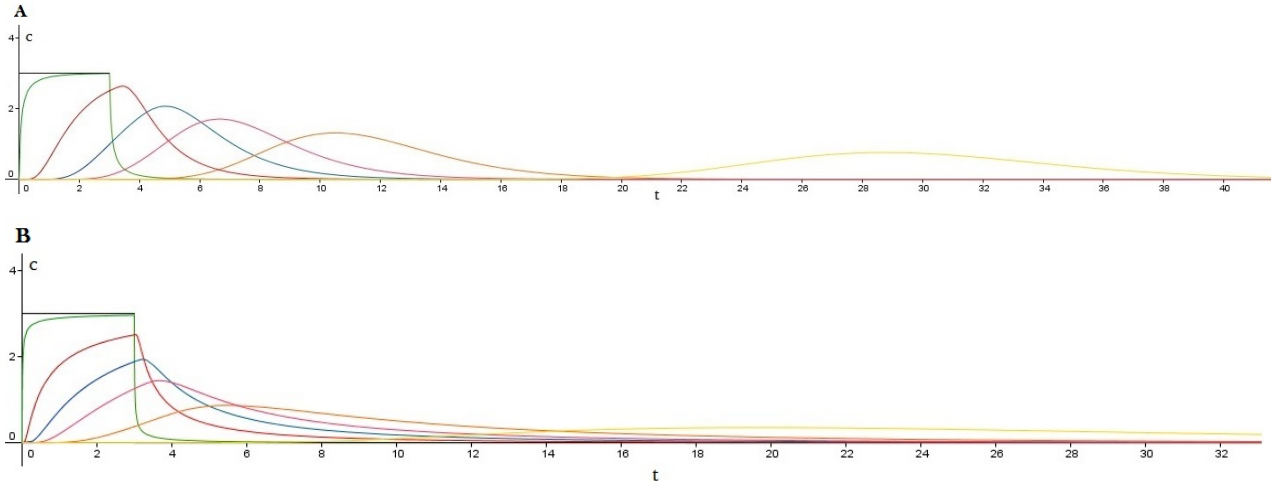

Figure 21: Solutions for $c(0,t) = 3$ if $0 < t < t_0 = 3$ with $\kappa = 1$ as in figure 20 (B) and for $\kappa = 0.2$ (A) for $x = 0.1$ (green), $x = 1$ (red), $x = 2$ (blue), $x = 3$ (pink), $x = 5$ (orange), and $x = 14$ (yellow). In both cases $v = 0.5$.

$t = 7$). Finally, the concentration-position profiles for $c(0,t) = 3 - 3e^{-t}$ and for $v = 1,\ \kappa = 0.2$ are shown in figure 25. We can see the curves rising from 0. At $t = 4$ the steady-state concentration $c_0 = 3$ is almost reached. When it has been reached, the curves for $t > 4$ look similar to those in figure 22.

If

$$\phi(t) = \frac{c_0}{\sqrt{4\kappa\pi t}} \exp\left(\frac{t - 0.8}{4\kappa t} + \frac{v^2 t}{4\kappa}\right),$$

which corresponds to the function $c(0,t) = G(t)$, which looks like the rat data for acetaminophen concentrations in brain ECF, equation (4.24) becomes

$$\frac{2}{\pi} c_0 e^{\frac{v^2 t}{4\kappa}} \int_\alpha^\infty \frac{1}{\sqrt{4\pi\kappa t - \frac{x^2\pi}{\xi^2}}} \exp\left(-\xi^2 - \frac{\epsilon^2}{\xi^2} + \frac{t - 0.8 - \frac{x^2}{4\kappa\xi^2}}{4\kappa t - \frac{x^2}{\xi^2}}\right) d\xi.$$

Here, $\alpha = \frac{x}{2\sqrt{\kappa t}}$ and $\epsilon = \frac{vx}{4\kappa}$ as defined ealier. We will not solve the above integral analytically. Figures 26 to 28 show the numerical solution of the advection-diffusion equation with $v = 0.5$, $\kappa = 1$ and the boundary condition

$$c(0,t) = \frac{1}{\sqrt{4\pi D t}} \exp\left(\frac{t - 0.8}{4tD}\right) \tag{4.29}$$

at $x = 0$ with $D = 0.1$ obtained in Mathematica. The concentration-time profiles in figures 26 and 27 show that the model describes the shapes of the concentration-time profiles seen in figure 11 quite well. The plots for

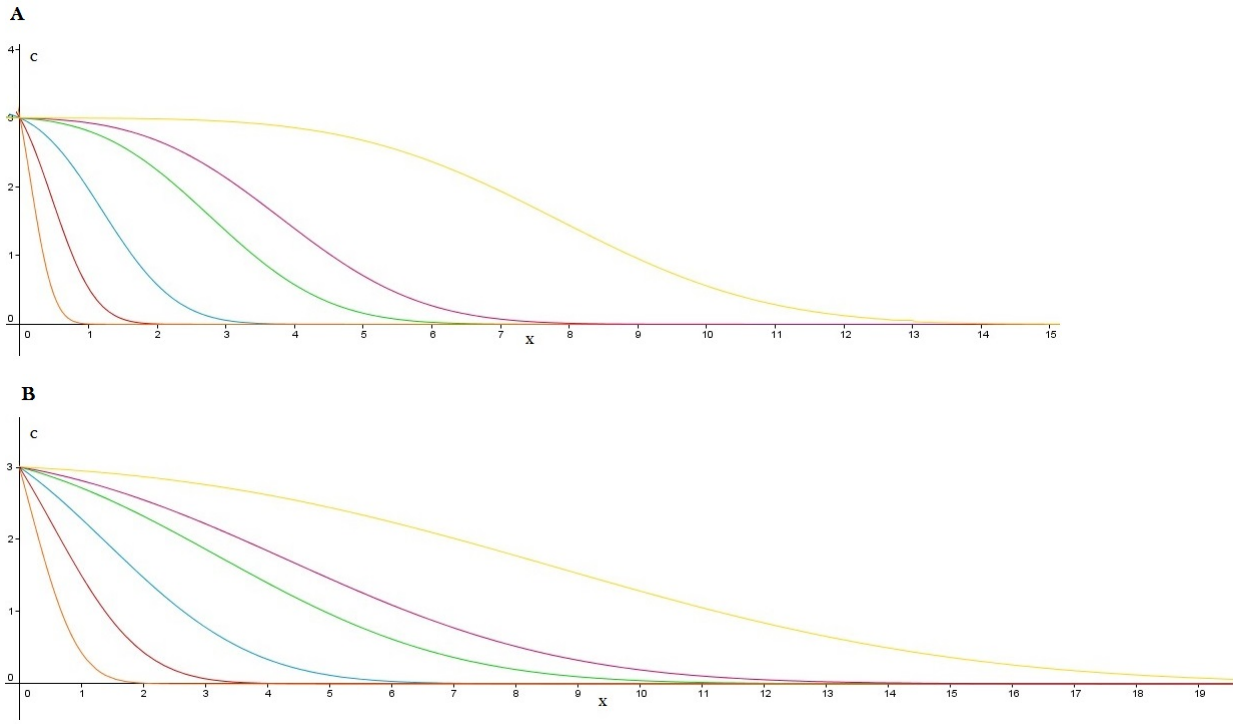

Figure 22: Solutions if $c(0,t) = 3$ for different values of t: $t = 0.2$ (orange), $t = 0.7$ (red), $t = 2$ (blue), $t = 5$ (green), $t = 7$ (pink), $t = 15$ (yellow). Coefficients are $v = 0.5$ and (A): $\kappa = 0.2$, (B): $\kappa = 1$.

small values of x look similar to the shapes of the plasma, brain-ECF and intra-brain concentration-time profiles, while the plots for larger values of x have shapes similar to the predicted SAS concentration-time profile in figure 11. Our equation seems able to describe the distribution of acetaminophen. The spatial plots for $x_L = 15$ and $t = 0, 2.5, 5, \ldots, 50$ are shown in figure 28. The concentration rises from 0 along the CSF system and the equalizing effect of the diffusion term is clearly visible. After the maximum concentration has been reached, the concentration decreases again at the beginning of the interval $[0, x_L]$, that is for small x, when t is increasing.

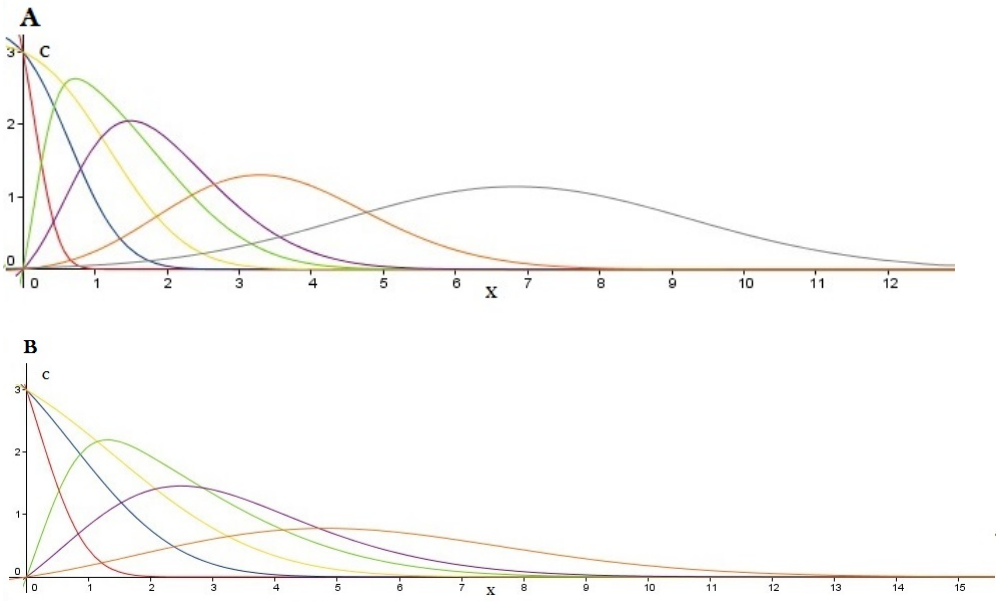

Figure 23: Solutions in the (c, x)-plane for $c(0, t) = 3$ if $0 < t < t_0$ for $t = 0.2$ (red), $t = 1$ (blue), $t = 2$ (yellow), $t = 3.2$ (green), $t = 4$ (purple), $t = 7$ (orange), and $t = 15$ (grey, only in (A)). Coefficients are $v = 0.5$ and (A): $\kappa = 0.2$, (B): $\kappa = 1$.

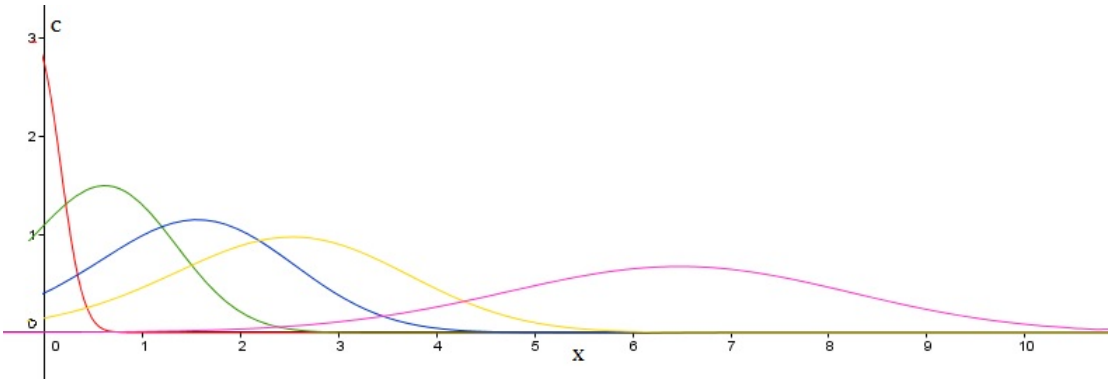

Figure 24: Solutions in (c, x)-plane for $c(0, t) = 3e^{-t}$ for different values of t: $t = 0.1$ (red), $t = 1$ (green), $t = 2$ (blue), $t = 3$ (yellow), $t = 7$ (pink). Coefficients are $v = 1$ and $\kappa = 0.2$.

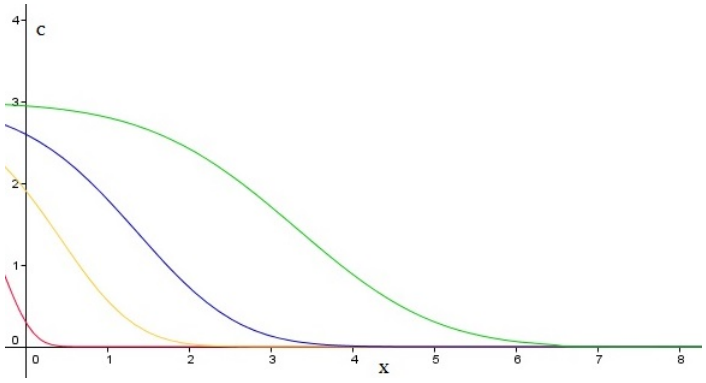

Figure 25: Solutions in (c, x)-plane for $c(0, t) = 3 - 3e^{-t}$ for different values of t: $t = 0.1$ (red), $T = 1$ (yellow), $t = 2$ (blue), $t = 4$ (green). Coefficients are $v = 1$, $\kappa = 0.2$.

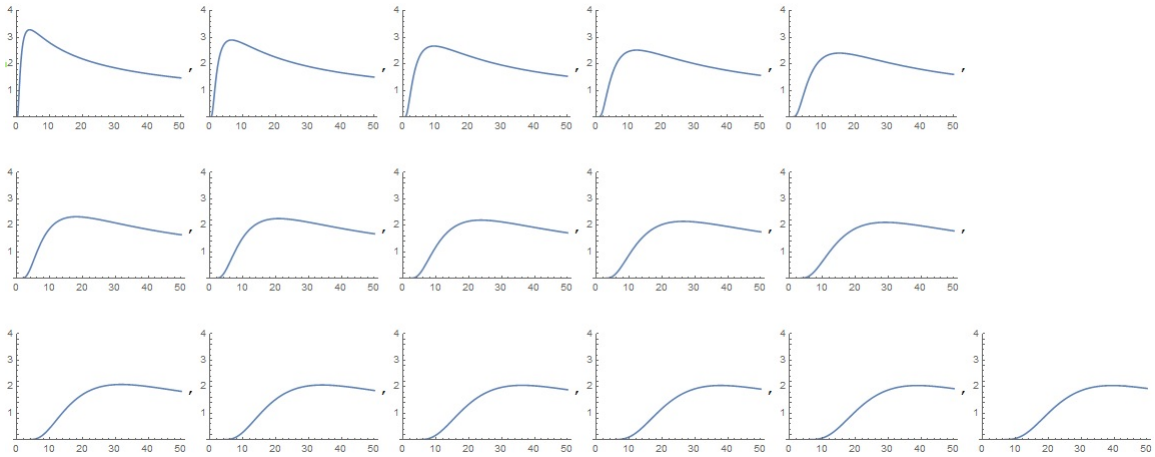

Figure 26: Numerical solutions of the advection-diffusion equation with $c(0, t)$ given by (4.29) with $v = 0.5$, $\kappa = 1$ and $T = 50$ for $x = 0, 1, 2, \ldots, 15$.

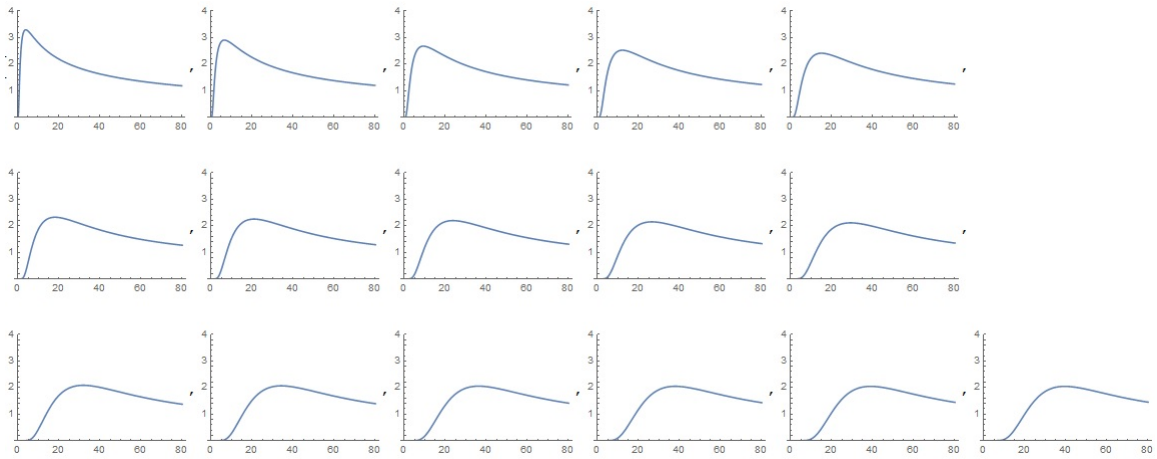

Figure 27: Numerical solutions of the advection-diffusion equation with $c(0,t)$ given by (4.29) with $v = 0.5$, $\kappa = 1$ and $T = 80$ for $x = 0, 1, 2, \ldots, 15$.

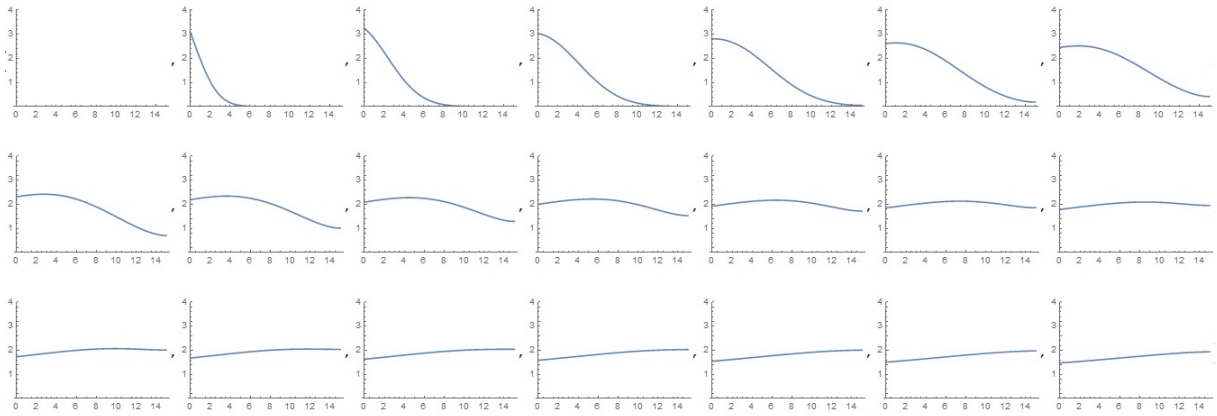

Figure 28: Numerical spatial solutions of the advection-diffusion equation with $c(0,t)$ given by (4.29) with $v = 0.5$, $\kappa = 1$ and $x_l = 15$ for $t = 0, 2.5, 5, \ldots, 50$.

4.4 Introduction to inverse problems

In the previous sections, we have derived a governing equation describing the change of the drug concentration in the CSF over time. In order to find a unique solution, we imposed initial and boundary conditions, assuming that the concentration in $x = 0$ was given by a known function $c(0, t) = G(t)$. We examined the behaviour of the solutions for different doses and different phases of drug distribution by solving the resulting problem

$$\begin{aligned} &\partial_t c = -v\partial_x c + \kappa\partial_x^2 c \\ &c(x, 0) = 0 \\ &c(0, t) = G(t);\ \partial_x c(x_L, t) = 0 \end{aligned}$$

for different functions $G(t)$. Unfortunately, this approach is not practicable: $G(t) = c_{in}c_{ecf}(t)$, while our aim is to determine brain-ECF concentrations. $G(t)$ is thus, in fact, the function we are looking for and not known to us. Our only available data are CSF concentrations, meaning that we do not know the source function or the initial concentration at $x = 0$ but only the concentration in the lumbar region (in the lower back). From this we aim to reconstruct the source function. Problems of this kind are called *inverse*, because instead of predicting effects from model inputs we seek to determine the causes for observed effects. In the terminology of inverse problems, the more standard problem of predicting effects from known model parameters and given initial and boundary conditions is called a *direct* problem (see figure 29). This section will provide the theoretical basis needed for formulating and solving our inverse problem. We will formulate it in section 4.5 based on the governing equation we derived in section 4.1 and the assumption that measurements in the lumbar region are available. We will then summarize the results obtained for a similar problem by El Badia et al. in [10], which will show that our inverse problem is solvable. The method proposed in [10] uses the Tikhonov regularization, to which we will give a brief introduction.

Solving inverse problems is generally very complicated. Most inverse problems are ill-posed in the sense that one of the following properties is not satisfied (ill-posedness according to Hadamard):

1. There is a solution (existence).
2. The solution is unique (uniqueness).
3. The solution depends continuously on the data (stability).

In dealing with inverse problems, it is therefore necessary to examine whether the above conditions are satisfied and, if not, to find a way of handling the

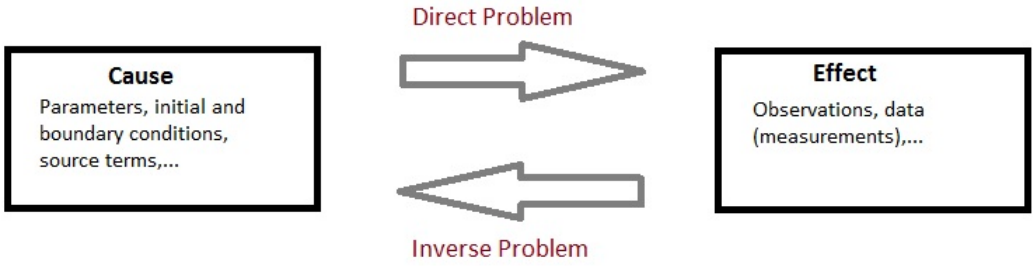

Figure 29: Schematic representation of a direct versus an inverse problem.

ill-posedness. Stability is particularly crucial in applications: If the solution does not depend continuously on the data, small data errors lead to exploding errors in the solution.

Numerical differentiation may serve as an example of the difficulties we face in ill-posed inverse problems. Differentiation and integration are inverse to one another. Numerical integration of a given function $\phi \in C([0,1])$ is a direct problem:

$$(T_D\phi)(x) = \int_0^x \phi(t)\,dt; \quad x \in [0,1]; \phi \in C([0,1]), \tag{4.30}$$

while the corresponding inverse problem is given by:

$$\text{Solve } (T_D\phi)(x) = g(x); \text{ for a given } g \in C([0,1]), \tag{4.31}$$

where T_D is the integral operator defined by the direct problem. The inverse problem is ill-posed in the sense of Hadamard, as the following example shows. Let us assume that we have noisy data $g^\delta \in C([0,1])$ satisfying an error bound

$$\|g - g^\delta\|_\infty \leq \delta$$

for a small δ. The functions

$$g_n^\delta(x) = g(x) + \delta \sin\left(\frac{nx}{\delta}\right); \; n = 2,3,4,\ldots$$

satisfy the error bound. However, for the derivatives

$$(g_n^\delta)'(x) = g'(x) + n\cos\left(\frac{nx}{\delta}\right); \; n = 2,3,4,\ldots$$

we find that

$$\|(g_n^\delta)'(x) - g'\|_\infty = n,$$

which blows up as $n \to \infty$. When we look at the approximate solution of (4.31) obtained by using the central difference quotients

$$(R_h g)(x) = \frac{g(x+h) - g(x-h)}{2h}; \quad x \in [0,1],\ h > 0, \tag{4.32}$$

we can see another typical property of ill-posed inverse problems. Assuming that g is periodic with period 1 in order to make $(R_h g)$ well defined near the boundaries, and using Taylor expansion, we obtain the approximiation error

$$\|g' - R_h g\|_\infty \leq \frac{h}{2}\|g''\|_\infty$$

if $g \in C^2([0,1])$. For noisy data g^δ, the total error can then be estimated by

$$\begin{aligned}\|g' - R_h g^\delta\|_\infty &\leq \|g' - R_h g\|_\infty + \|R_h g - R_h g^\delta\|_\infty \\ &\leq \frac{h}{2}\|g''\|_\infty + \frac{\delta}{h}.\end{aligned}$$

The first term on the right hand side is the approximation error, which tends to 0 as h goes to 0, the second term is the data noise error, which blows up as h goes to 0. In order to obtain a good approximation, we must balance the two error terms by choosing an adequate discretization parameter h. The optimal h is that which minimizes the total error, thus

$$h = \left(\frac{2}{\|g'\|_\infty}\right)^{\frac{1}{2}} \delta^{\frac{1}{2}}.$$

This simple example has illustrated several typical properties of ill-posed inverse problems:

- amplification of high frequency errors;
- dependence of ill-posedness on the choice of norms;
- trade-off between accuracy and stability in the choice of the discretization parameter;
- dependence of the optimal choice of a discretization parameter on the smoothness of the solution.

4.4.1 Algorithms for the solution of linear inverse problems

We will now introduce some common methods for the stable solution of linear ill-posed operator equations

$$T\phi = g, \tag{4.33}$$

where $T : X \to Y$ is a bounded linear injective operator, X and Y are Hilbert spaces, and $g \in \text{range}(T) =: R(T)$. The methods that will be introduced here are the Tikhonov regularization, which is used in the solution method proposed by El Badia et al. [10] (see 4.5.1), and the discrepancy principle as an example of a so-called parameter-choice rule. Parameter-choice rules are used to find an optimal discretization or regularization parameter.

Tikhonov regularization

One of the most commonly used methods for the regularization of linear ill-posed inverse problems is the Tikhonov regularization, sometimes also called the Ridge regression. One of its advantages is that there are no restrictions on the operator T. The method works for any T regardless of its symmetry, definiteness, or rank, provided that the regularization parameter α, which is defined below, is positive. The idea of the Tikhonov regularization is to rewrite problem (4.33) with noisy data g^δ as a minimization problem:

$$\begin{aligned} &\text{Find } \phi \in X \text{ such that} \\ &\qquad \|T\phi - g^\delta\|^2 \to \min. \end{aligned} \tag{4.34}$$

Again, the solution to this problem does not depend continuously on the data. One possible way of restoring stability is to add a penalty term to the functional. We then obtain the so-called Tikhonov functional

$$J_\alpha(\phi) = \|T\phi - g^\delta\|^2 + \alpha\|\phi - \phi_0\|^2 \tag{4.35}$$

for an initial guess ϕ_0. If no initial guess suggests itself, we take $\phi_0 = 0$. The number $\alpha > 0$ as in (4.35) is called the regularization parameter. It can be proved that the Tikhonov functional has a unique minimum ϕ_α^δ for all $\alpha > 0$ and for all initial guesses $\phi_0 \in X$. Moreover, ϕ_α^δ depends continuously on the perturbed data. The minimum is then given by

$$\phi_\alpha^\delta = (T^*T + \alpha I)^{-1}(T^*g^\delta + \alpha\phi_0), \tag{4.36}$$

where T^* is the conjugate transpose of T and the operator $T^*T + \alpha I$ is boundedly invertible. The proof can be found in appendix C.

Remark 4.1 (Iterated Tikhonov reularization). As mentioned above, we use $\phi_0 = 0$ as a default initial guess. Depending on the actual function, this may not be very accurate. One can try to improve the approximation by applying the Tikhonov regularization again, now using the obatined function ϕ_α^δ as a new initial guess. This leads to the so-called iterated Tikhonov regularization:

$$\begin{aligned}\phi_{\alpha,0}^\delta &= 0 \\ \phi_{\alpha,n+1} &= (T^*T + \alpha I)^{-1}(T^*g^\delta + \alpha\phi_{\alpha,n}^\delta); \quad n \geq 0.\end{aligned}$$

Note that only one operator $T^*T + \alpha I$ has to be inverted in order to compute $\phi_{\alpha,n}^\delta$ for any $n \in \mathbb{N}$. Thus, computing $\phi_{\alpha,n}^\delta$ for any $n \geq 2$ is not much more expensive than computing $\phi_{\alpha,1}^\delta$ when a method such as LU factorization is used. By induction, the following expression for $\phi_{\alpha,n}^\delta$ is found:

$$\phi_{\alpha,n}^\delta = (\alpha I + T^*T)^{-n}(T^*T)^{-1}\left((\alpha I + T^*T)^n - \alpha^n I\right)T^*g^\delta.$$

Discrepancy principle

We have seen that the Tikhonov regularization is a very useful method since it yields a unique solution depending continuously on the noisy data and since it can be applied to any operator T. A crucial problem concerning the Tikhonov regularization as well as other regularization methods is the choice of an optimal regularization parameter α. We can decompose the total error $\phi - \phi_\alpha^\delta$ for $\phi_0 = 0$ as follows:

$$\begin{aligned}\phi - \phi_\alpha^\delta &= \phi - (\alpha I + T^*T)^{-1}T^*T\phi + (T^*T + \alpha I)^{-1}T^*(g - g^\delta) \\ &= \alpha(\alpha I + T^*T)^{-1}\phi + (T^*T + \alpha I)^{-1}T^*(g - g^\delta).\end{aligned}$$

We have now split the total error into an approximation error

$$\alpha(\alpha I + T^*T)^{-1}\phi$$

and a data noise error

$$(T^*T + \alpha I)^{-1}T^*(g - g^\delta).$$

Formally, the approximation error tends to 0 as α goes to 0, while the data noise error tends to T^{-1}, which is unbounded. We therefore expect the data noise error to blow up as α tends to 0. Hence, there is a trade-off between accuracy and stability as we have already seen in the introductory example. The trade-off situation is illustrated in figure 30.

The optimal value of the regularization parameter α balances between desired accuracy and necessary stability. It depends on both the data noise

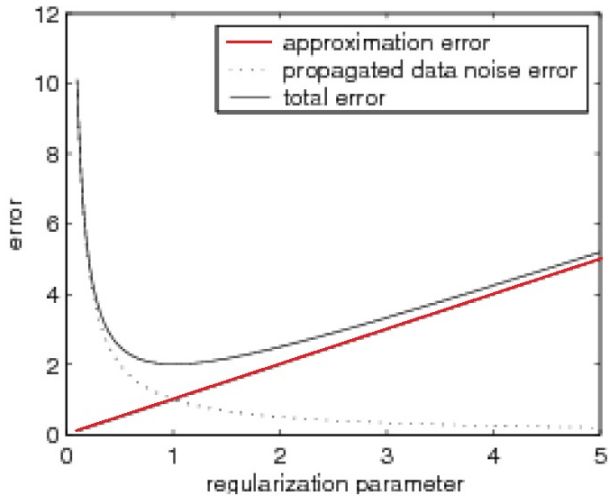

Figure 30: Dependence of the errors on α. [19]

level δ and the exact solution ϕ. Many strategies have been proposed for choosing α. The most well-known is Morozov's discrepancy principle. The idea is to not try to satisfy the operator equation more accurately than the data-noise error. More precisely, we want to take the largest $\alpha = \alpha(\delta, g^\delta)$ such that the residual $\|T\phi_\alpha^\delta - g^\delta\|$ is not larger than $\tau\delta$ with $\tau \geq 1$ being some fixed parameter:

$$\alpha(\delta, g^\delta) = \sup\{\alpha > 0 \mid \|T\phi_\alpha^\delta - g^\delta\| \leq \tau\delta\}. \tag{4.37}$$

It can be shown that the function $\alpha \mapsto \|T\phi_\alpha^\delta - g^\delta\|$ is monotonically increasing for most regularization methods including Tikhonov regularization. Usually, finding α such that

$$\tau_1\delta \leq \|T\phi_\alpha^\delta - g^\delta\| \leq \tau_2\delta$$

for given $1 \leq \tau_1 < \tau_2$ is sufficient. This can be done by a bisection algorithm or by Newtons's method applied to $\frac{1}{\alpha}$ as unknown.

4.5 The inverse problem

So far, we have worked with the direct problem assuming that some intial conditions are given. In practice, however, this approach is not adequate. The only available data in humans are concentrations measured in the lumbar region. Reconstructing brain-ECF concentrations from lumbar CSF measurements is an inverse problem. In this section, we will formulate the inverse problem. In view of the ill-posedness of inverse problems, three questions will have to be answered:

1.) Identifiability: Is the source term F, which must be recovered, uniquely determined by the measurements, that is the right hand side of equation (4.41) below?

2.) Stability: If so, do the source parameters continuously depend on the data or not?

3.) Identification: Is there a constructive algorithm for determining the source?

Unfortunately, indentifiability is often not given. In [10], however, El Badia et al. have proved for a problem similar to ours that a point source is identifiable if certain conditions are met regarding the points where the measurements are taken. El Badia et al. also proved a local stability result and proposed an algorithm for recovering the source parameters, that is location and intensity of the source F. In section 4.5.1, we will present the results given in [10] and apply them to our problem. First, let us formulate the inverse problem. The operator L defined by

$$L[c] = \partial_t c(x,t) - \kappa \partial_{xx} c(x,t) + v \partial_x c(x,t) \tag{4.38}$$

is a linear parabolic partial differential operator. It is known that the problem

$$L[c](x,t) = F(x,t) \tag{4.39}$$

has a unique solution, if certain conditions on F are met and suitable boundary and initial conditions are given. This is the direct problem that was studied in section 4.3. We will assume that c is smooth enough so we can postulate the existence of point values. Then, we can define the so-called observation operator

$$B[F] \coloneqq \{c(a,t;F), c(b,t;F); 0 < t < T\} \tag{4.40}$$

for $0 < a < b < x_L$. In accordance with the results in [10], we will assume that we have measurements in two points, even though we have only one actual point of measurement, namely the lumbar region, where and where alone CSF can be accessed. A way of circumventing this problem and adding a second point of measurement will be given in the following subsection. T is the duration of the experiment. The inverse source problem is now stated as follows:

(ISP): Given the measurements $\{d_1(t), d_2(t), 0 < t < T\}$ of the concentration in two points $a, b \in (0, x_L)$, find the source F such that

$$B[F] = \{d_1(t), d_2(t); 0 < t < T\}. \tag{4.41}$$

In accordance with the assumptions in [10], we will assume that the source F becomes inactive after time T^* with $0 < T^* < T$. We will assume F to be a point source since El Badia et al. proved identifiability for point sources only. This assumption is no unreasonable simplification, because the source in our model is located in the lateral ventricles, which are very small compared to the whole CSF system.

4.5.1 Solving the inverse problem for a point source

In this section, we will examine the solvability of the inverse problem using the results in [10]. We will assume here that we start injecting the drug into the CSF space at time $t = 0$ and stop the injection at some time T^* with $0 < T^* < T$, meaning that the source becomes inactive after time T^* but we continue collecting data till a later time T. Note that T^* and T might take high values depending on the drug in question.

With these assumptions in place, let us answer the three questions.

Identifiability

Since the source is considered to be a point source, it is of the form

$$F(x,t) = \lambda(t)\delta(x-S) = c_{in}c_{ecf}(t)\delta(x-S), \tag{4.42}$$

where $\delta(x)$ is the so-called delta function or delta distribution (see appendix B) and S is the location of the source. Since the source is assumed to become inactive after time T^*, we assume $\lambda(t) = 0$ for $t > T^*$. Also, we will presume that the unknown intensity function λ belongs to $L^2(0,T)$. In this case, it is known that the direct problem has a unique solution, which belongs to $L^2(0,T;H^1(0,x_L)) \cap C([0,T];L^2(0,x_L))$ [26]. By the Sobolev embedding theorem (A.1), we can therefore define the values of $c(x,t)$ in any point $(x,t) \in (0,x_L) \times (0,T)$. This allows us to define the observation parameter and formulate the inverse problem. In order to determine the parameters of the source F, that is the intensity λ and the location S, we will denote the solution as $c(x,t;\lambda,S)$ and the observation operator as

$$B[\lambda,S] = \{c(a,t;\lambda,S), c(b,t;\lambda,S); 0 < t < T\}.$$

For proving the identifiability of the point source, certain conditions must be imposed on the points a, b, where data is collected. We only have access to the CSF in the lumbar region, but in order to use the results in [10] we need two measurements, one of which is to be taken upstream from the source and one downstream, that is $0 < a < S < b < x_L$, where S is the location

of the source. Since the source is located in the lateral ventricles, that is $S \in (0, x_0)$, we can use a close to 0 and $b \in (x_0, x_1)$, that is b is in the lumbar region (recall figure 6). The concentration upstream from the source should be negligible, because the CSF flow is unidirectional, the effect of diffusion is assumed to be small enough so the drug does not diffuse against the flow direction, and no drug is naturally present in the body. We may therefore simply take $c(a, t; \lambda, S) = 0$ for all times t. Then, in order to guarantee the identifiability of the point source, one of the points a, b needs to be strategic in the following sense.

Definition 4.1. Let $\{\psi_n\}$ be a complete orthonormal family of continuous functions in $L^2(0, x_L)$. A point $b \in (0, x_L)$ is said to be strategic relative to the family $\{\psi_n\}$ if

$$\psi_n(b) \neq 0 \quad \forall n.$$

This condition is needed in the proofs in [10]. The orthonormal family ψ_n will be used to express the solution in terms of Fourier series. The definition of a strategic point relative to the ψ_n is then used to prove that $c(b, t) = 0$ implies that all coefficients in the Fourier series must vanish. In order to find an expression for the ψ_n, let us consider the Sturm-Liouville problem

$$-\kappa\psi''(x) + \rho\psi(x) = \mu\psi(x); \quad 0 < x < x_L \tag{4.43}$$

$$\psi(0) = \psi'(x_L) - \alpha\psi(x_L) = 0, \tag{4.44}$$

where $\alpha = \frac{-v}{2\kappa}$, $\rho = \alpha^2\kappa$, and v and κ are taken from (4.38). We are interested in the family $\{\psi_n\}$ of complete orthonormal eigenfunctions to the above problem. For this, rewrite (4.43) as

$$-\kappa\psi''(x) + (\rho - \mu)\psi = 0. \tag{4.45}$$

The general solution to this equation is

$$\psi_n = c_n \sin(\beta_n x) + \tilde{c}_n \cos(\beta_n x). \tag{4.46}$$

Using the initial condition in (4.44), we find that $\psi_n(0) = 0$, which implies $\tilde{c}_n = 0$ and hence

$$\psi_n(x) = c_n \sin(\beta_n x).$$

By the boundary condition in (4.44), $\psi_n'(x_L) = \alpha\psi_n(x_L)$. Since

$$\psi_n'(x_L) = -c_n\beta_n \cos(\beta_n x_L),$$

we obtain the equation

$$-c_n\beta_n \cos(\beta_n x_L) = c_n\alpha \sin(\beta_n x_L).$$

By dividing by $c_n \neq 0$ and multiplying by x_L, we find that the β_n, $n \geq 0$ are the positive solutions to

$$\beta x_L \cot(\beta x_L) = \alpha x_L. \tag{4.47}$$

The c_n are normalization coefficients. The associated eigenvalues μ_n to ψ_n are

$$\mu_n = \rho + \kappa \beta_n^2. \tag{4.48}$$

Hence,

$$x_L \beta_n = (2n+1)\frac{\pi}{2} - \delta_n; \quad 0 < \delta_n < \frac{\pi}{2}, \quad \lim_{n\to\infty} \delta_n = 0. \tag{4.49}$$

And thus,

$$\rho < \mu_n < \mu_{n+1}, \quad \mu_n \approx \frac{\kappa}{x_L^2}\pi^2 n^2 \text{ at } \infty. \tag{4.50}$$

In the following, we will use the ψ_n as defined by (4.43), (4.44). This is convenient since El Badia et al. [10] proved identifiability only if one of the points where measurements are taken is strategic with respect to the family $\{\psi_n\}$. For the family we have just introduced, any point $p \in (0, x_L)$ such that $\psi_n(p) = 0$ has to satisfy $p = \frac{m\pi}{\beta_n} y x_L$ for some integer m. The set of such points is countable, so an arbitrary point $b \in (0, x_L)$ is strategic with probability 1. In the following, let $L[\cdot]$ be the operator defined by (4.38). In order to prove the main theorem on the identifiability of the source, we need the following lemma (see lemma 1 in [10]).

Lemma 4.1. *Let $T^* < T$ and $b \in (0, x_L)$ be a strategic point relative to the family $\{\psi_n\}$. Suppose that $w = w(x,t)$ satisfies*

$$\begin{aligned} &L[w](x,t) = 0; \quad 0 < x < x_L, \ T^* < t < T \\ &w(0,t) = \partial_x w(x_L, t) = 0; \quad T^* < t < T \\ &w(\cdot, T^*) \in L^2(0, x_L). \end{aligned} \tag{4.51}$$

Then

$$w(b,t) = 0 \text{ for all } t \in (T^*, T) \Rightarrow w(\cdot, T^*) = 0 \text{ in } L^2(0, x_L).$$

The proof of this lemma can be found in appendix D. We will now state and prove the main theorem on the identifiability of the solution (see theorem 1 in [10]). The theorem states that if we have the same observations in points a and b, one of them being strategic, for two source terms $F_j = \lambda_j(t)\delta(x - S_j)$, $j = 1, 2$ for all $0 < t < T$, then $F_1 = F_2$ almost everywere. It is thus a uniqueness result.

Theorem 4.1. *Suppose $F_j(x,t) = \lambda_j(t)\delta(x-S_j)$, where $\lambda_j \in L^2(0,T)$ is such that for $j=1,2$, $\lambda_j(t) \geq 0$ with $\lambda_j(t) = 0$ for $t \in (T^*,T)$, and $S_j \in (a,b)$, $j=1,2$. If at least one of the points a,b is strategic with respect to the family $\{\psi_n\}$, then $B[\lambda_1,S_1] = B[\lambda_2,S_2]$ implies $\lambda_1(t) = \lambda_2(t)$ almost everywhere in $(0,T)$ and $S_1 = S_2$.*

Proof. Let $c_j,\ j=1,2$ be the solutions of

$$L[c_j](x,t) = \lambda_j(t)\delta(x-S_j); \quad 0 < x < x_L,\ 0 < z < T \tag{4.52}$$
$$c_j(0,t) = \partial_x c_j(x_L,t) = 0; \quad 0 < t < T \tag{4.53}$$
$$c_j(x,0) = 0; \quad 0 < x < x_L. \tag{4.54}$$

Again, the second condition in (4.53) applies when we neglect elimination in the arachnoid villi. Since there is no elimination and the effect of diffusion is assumed to be small, it is assumed that the change of concentration with respect to space is neglectable at x_L due to the distance to the source.

Consider the difference $u = c_2 - c_1$, which is the solution to

$$L[u](x,t) = \lambda_2(t)\delta(x-S_2) - \lambda_1\delta(x-S_1); \quad 0 < x < x_L,\ 0 < t < T \tag{4.55}$$
$$u(0,t) = \partial_x u(x_L,t) = 0; \quad 0 < t < T \tag{4.56}$$
$$u(x,0) = 0; \quad 0 < x < x_L. \tag{4.57}$$

Now, $B[\lambda_1,S_1] = B[\lambda_2,S_2]$ implies that

$$u(a,t) = u(b,t) = 0; \quad 0 < t < T. \tag{4.58}$$

First consider $u \in (0,x_L) \times (T^*,T)$. Then $\lambda_j(t) = 0$, $j=1,2$, and we obtain the problem

$$L[u](x,t) = 0; \quad 0 < x < x_L,\ T^* < t < T$$
$$u(0,t) = \partial_x u(x_L,t) = 0; \quad T^* < t < T.$$

Since $u(b,t) = 0$ in (T^*,T) by (4.58), lemma 4.1 implies that

$$u(x,T^*) = 0; \quad 0 < x < x_L. \tag{4.59}$$

Now consider $u \in (0,x_L) \times (0,T^*)$ and the corresponding problem

$$L[u](x,t) = \lambda_2(t)\delta(x-S_2) - \lambda_1(t)\delta(x-S_1);\ 0 < x < x_L,\ 0 < t < T^* \tag{4.60}$$
$$u(0,t) = \partial_x u(x_L,t) = 0; \quad 0 < t < T^* \tag{4.61}$$
$$u(x,0) = 0; \quad 0 < x < x_L. \tag{4.62}$$

Since $0 < a < S_j < b < x_L$ and $\delta(x - S_j) = 0$ for all $x \neq S_j$, $j = 1, 2$, we see that the right hand side of (4.60) vanishes identically in $(0, a) \times (0, T^*) \cup (b, x_L) \times (0, T^*)$. Therefore we deduce from (4.60)-(4.62) that

$$u = 0 \quad \text{in } (0, a) \times (0, T^*) \cup (b, x_L) \times (0, T^*). \tag{4.63}$$

And thus, by (4.58),

$$\partial_x u(a, t) = \partial_x u(b, t) = 0; \quad 0 < t < T^*. \tag{4.64}$$

Let r_i, $i = 1, 2$ be the solutions of the equation $-\kappa r^2 - vr = 0$ and let $u_i = e^{r_i x}$. Through multiplying (4.60) by u_i and integrating with respect to x and t, we obtain

$$\int_a^b \int_0^{T^*} L[u](x, t) u_i(x) dt dx = e^{r_i S_2} \int_0^{T^*} \lambda_2(t) dt - e^{r_i S_1} \int_0^{T^*} \lambda_2(t) dt, \tag{4.65}$$

where the $e^{r_i S_j}$, $i, j = 1, 2$ term results from the integral property of delta, that is from $\int_a^b \delta(x - S_j) u_i(x) dx = u_i(S_j)$, because $S \in (a, b)$. Since $u \in L^2(0, T; H^1(0, x_L)) \cap C(0, t; L^2(0, x_L))$, using integration by parts, we see that the left hand side of (4.65) vanishes according to (4.58), (4.59), and (4.64). Therefore,

$$\begin{aligned} e^{r_1 S_2} \int_0^{T^*} \lambda_2(t) dt &= e^{r_1 S_1} \int_0^{T^*} \lambda_1(t) dt \\ e^{r_2 S_2} \int_0^{T^*} \lambda_2(t) dt &= e^{r_2 S_1} \int_0^{T^*} \lambda_1(t) dt. \end{aligned}$$

Since $r_1 \neq r_2$ and $\int_0^{T^*} \lambda_j(t) dt > 0$, we obtain $S_1 = S_2$ and $\int_0^{T^*} \lambda_1(t) dt = \int_0^{T^*} \lambda_2(t) dt$. So let $S = S_1 = S_2$ and $y(x, t) = e^{\alpha x} u(x, t)$ with $\alpha = \frac{-v}{2\kappa}$ as above. Then, u is a solution to (4.55) if and only if y is a solution to the heat equation:

$$\partial_t y - \kappa \partial_{xx} y + \rho y = e^{\alpha S} (\lambda_2(t) - \lambda_1(t)) \delta(x - S); \; 0 < x < x_L, 0 < t < T \tag{4.66}$$

$$y(0, t) = \partial_x y(x_L, t) - \alpha y(x_L, t) = 0; \quad 0 < t < T \tag{4.67}$$

$$y(x, 0) = 0; \quad 0 < x < x_L, \tag{4.68}$$

where $\rho = \alpha^2 \kappa$. Then, y is given by the Fourier expansion

$$y(x, t) = \sum_n a_n(t) \psi_n(x)$$

with $a_n(t) = \int_0^{x_L} y(x, t) \psi_n(x) dx$ and $a_n(0) = 0$. Since

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle y(x, t), \psi_n(x) \rangle = \left\langle \frac{\partial}{\partial t} y(x, t), \psi_n(x) \right\rangle + \left\langle y(x, t), \frac{\partial}{\partial t} \psi_n(x) \right\rangle = \left\langle \frac{\partial}{\partial t} y(x, t), \psi_n(x) \right\rangle,$$

we obtain

$$\begin{aligned} a_n'(t) &= \int_0^{x_L} \partial_t y(x,t)\psi_n(x)dx \\ &= \int_0^{x_L} (\kappa\partial_{xx}y(x,t) - \rho y(x,t))\psi_n(x)dx + e^{\alpha S}(\lambda_2(t) - \lambda_1(t))\psi_n(S) \\ &= -\mu_n a_n(t) + e^{\alpha S}(\lambda_2(t) - \lambda_1(t))\psi_n(S), \end{aligned}$$

using (4.66) for the expression of $\partial_t y(x,t)$. The μ_n are the eigenvalues with respect to ψ_n as defined by (4.43) and (4.48). Therefore,

$$u(x,t) = e^{-\alpha x}y(x,t) = \sum_{n=0}^{\infty} \psi_n(x)\psi_n(S)e^{-\alpha(x-S)} \int_0^t (\lambda_2 - \lambda_1)(\zeta)e^{-\mu_n(t-\zeta)}d\zeta,$$

which is obtained by integrating $a_n'(t)$ with respect to t. We can write this as

$$u(x,t) = \int_0^t (\lambda_2 - \lambda_1)(\zeta)\Phi(x,t-\zeta)d\zeta$$

with

$$\Phi(x,t) = \sum_{n=0}^{\infty} \psi_n(S)\psi_n(x)e^{-\alpha(x-S)}e^{-\mu_n t}. \tag{4.69}$$

Here we have switched integration and summation, which is justified by Lebesgue's theorem of dominated convergence and the estimate

$$\begin{aligned} \sum_{n_0}^{\infty} |\psi_n(S)\psi_n(x)e^{-\alpha(x-S)}|e^{-\mu_n t} &\le M \sum_{n_0}^{\infty} e^{-Kn^2 t} && (4.70) \\ &\le M \int_0^{\infty} e^{-Kt\xi^2}d\xi && (4.71) \\ &= M_1 \frac{1}{\sqrt{t}}, && (4.72) \end{aligned}$$

where M, M_1, K are constants.

The first estimation (4.70) is obtained by using (4.50) for n_0 sufficiently large, while equation (4.72) is obtained by using $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$ with $x = \xi\sqrt{Kt},\ d\xi = \frac{1}{\sqrt{Kt}}dx$.

Since $u(b,t) = u(a,t) = 0$ for $0 < t < T$, we get

$$u(b,t) = \int_0^t (\lambda_2 - \lambda_1)(\xi)\Phi(b,t-\xi)d\xi = 0; \quad \forall t \in (0,T).$$

According to theorem D.1, the functions $(\lambda_2 - \lambda_1)$ and Φ must vanish identically at least in some intervals $(0,T')$, $(0,T'')$ respectively, where T', T'' are such that $T' + T'' \ge T$. If $\Phi = 0$ in $(0,T'')$ with $T'' > 0$, analytic continuation

also yields $\Phi = 0$ in $(0, \infty)$, which implies $\psi_n(S)\psi_n(b) = 0$ for all n. Since b is strategic, that is $\psi_n(b) \neq 0$ for all n, this implies $\psi_n(S) = 0$ for all n, which is impossible according to the representation of β_n in (4.49). Hence $(\lambda_2 - \lambda_1) = 0$, and thus

$$\lambda_2(t) = \lambda_1(t);\ 0 < t < T.$$

□

Stability

Stability is crucial in numerical applications. If the source F does not depend continuously on the data, small perturbations and rounding errors are amplified. El Badia et al. [10] proved a local Lipschitz stability result using Gâteaux differentiability. These arguments are based on those used by Bellout and Friedmann in [4].

Definition 4.2. Let X, Y be locally convex topological spaces, for example Banach spaces, $U \subset X$ open and $F : X \to Y$. The Gâteaux differential $dF(u; \psi)$ of F at $u \in U$ in direction $\psi \in X$ is defined as

$$dF(u; \psi) = \lim_{\tau \to 0} \frac{F(u + \tau\psi) - F(u)}{\tau}$$

if this limit exists. If the limit exists for all $\psi \in X$, then F is said to be Gâteaux differentiable at $u \in U$.

Now consider the set

$$G(T^*) = \{(\lambda, S) \in L^2(0, x_L) \times (a, b) \mid \lambda(t) \geq 0 \text{ with } \lambda(t) = 0 \text{ for } t \geq T^*\}.$$

Then, for $(\lambda, S), (\mu, \tau) \in G(T^*)$ and h sufficiently small, $(\lambda + h\mu, S + h\tau) \in G(T^*)$. We can therefore define the source term

$$F^h(x, t) = (\lambda(t) + h\mu(t))\delta(x - (S + h\tau)). \tag{4.73}$$

Moreover, by using Taylor expansion with respect to h, we see that there is a point S_h satisfying $|S - S_h| < |\tau h|$ such that

$$F^h(x, t) = F(x, t) + h\hat{F}(x, t) + h^2\tilde{F}(x, t),$$

where

$$\hat{F}(x, t) = \mu(t)\delta(x - S) - \lambda(t)\tau\delta'(x - S)$$

$$\tilde{F}(x,t) = -\mu(t)\delta'(x-S) + \frac{1}{2}\lambda(t)\tau^2\delta''(x-S_h).$$

Hence,

$$c(x,t;F^h) = c(x,t;F) + hc(x,t;\hat{F}) + h^2c(x,t;\tilde{F}),$$

where $c(x,t;\hat{F})$ and $c(x,t;\tilde{F})$ are solutions of (4.39) with the initital and boundary conditions

$$\begin{aligned} c(0,t) &= \partial_x c(x_L,t) = 0 \\ c(x,0) &= 0 \end{aligned}$$

and with $\hat{F}$ and $\tilde{F}$, respectively, as source terms. Therefore,

$$\lim_{h\to 0}\frac{B[F^h]-B[F]}{h} = \{c(a,t;\hat{F}), c(b,t;\hat{F})\}. \tag{4.74}$$

We can now state the local Lipschitz stability result as mentioned earlier. For the proof, see appendix D.

Theorem 4.2. *Let $\lambda, \mu \in G(T^*)$. If $(\mu,\tau) \neq (0,0)$ for $0 < t < T$, then*

$$\lim_{h\to 0}\frac{B[F^h]-B[F]}{h} \neq 0.$$

This theorem means that with $h \to 0$ we can still distinguish between the source terms F^h and F by means of the measurements taken in points a and b as long as the error in the measurements is $o(h)$.

Identification

Having answered the questions of identifiability and stability, we are left with the third question. If the source term is identifiable and the solution depends continuously on the data, how can we determine the source parameters? For the identification of the source term $F(x,t)$, we need to find the location S and the intensity λ. The method presented in [10] requires knowledge of the functions $c(x,T^*)$ for $0 < x < x_L$ and $\partial_x c(a,t)$, $\partial_x c(b,t)$ for $0 < t < T^*$. So before we can try to recover the source parameters, we need to find these functions. This is done by solving three systems.

The first step is to determine $c_{T^*} = c(x,T^*)$ from the system

$$\begin{aligned} &L[c](x,t) = 0; \quad 0 < x < x_L,\ T^* < t < T \\ &c(0,t) = \partial_x c(x_L,t) = 0; \quad T^* < t < T \\ &c(x,T^*) = c_{T^*}(x); \quad 0 < x < x_L \end{aligned}$$

by using the data $\{d_1(t), T^* < t < T\}$ and $\{d_2(t), T^* < t < T\}$. Again, $F = 0$ since $\lambda(t) = 0$ in (T^*, T). Tikhonov regularization can be used to solve the system and hence to identify c_{T^*}.

Then, $\partial_x c(a,t)$ and $\partial_x c(b,t)$ can be determined by solving two direct problems. For identifying $\partial_x c(a,t)$, we use the data $\{d_1(t), 0 < t < T^*\}$ and solve the system

$$\begin{aligned} &L[c](x,t) = 0; \quad 0 < x < a,\ 0 < t < T^* \\ &c(0,t) = 0; \quad 0 < t < T^* \\ &c(a,t) = d_1(t); \quad 0 < t < T^* \\ &c(x,0) = 0; \quad 0 < x < a. \end{aligned}$$

Here, $F = 0$ since $S \notin (0,a)$. Similarly, for $\partial_x c(b,t)$ we have to solve the system

$$\begin{aligned} &L[c](x,t) = 0; \quad b < x < x_L; \quad 0 < t < T^* \\ &c(b,t) = d_2(t); \quad 0 < t < T^* \\ &\partial_x c(x_L,t) = 0; \quad 0 < t < T^* \\ &c(x,0) = 0; \quad b < x < x_L \end{aligned}$$

with $F = 0$ since $S < b$.

If we suppose now that $c(x,T^*)$, $\partial_x c(a,t)$ and $\partial_x c(b,t)$ are completely determined in their domains, we can try to recover the location S. Therefore, consider the system

$$\begin{aligned} &L[c](x,t) = \lambda(t)\delta(x-S); \quad 0 < x < x_L,\ 0 < t < T^* \\ &c(0,t) = \partial_x c(x_L,t) = 0; \quad 0 < t < T^* \\ &c(x,0) = 0; \quad 0 < x < x_L. \end{aligned}$$

Through multiplying the first equation by $e^{r_i x}$ with r_i as before and integrating with respect to x and t over $(0,x_L) \times (0,T^*)$, in which we use integration by parts, we obtain

$$\begin{aligned} e^{r_i S} \int_0^{T^*} \lambda(t)dt = -\kappa e^{r_i b} \int_0^{T^*} \partial_x c(b,t)dt + \kappa e^{r_i a} \int_0^{T^*} \partial_x c(a,t)dt \\ + \int_a^b c(x,T^*)e^{r_i x}dx; \quad i = 1,2. \end{aligned} \tag{4.75}$$

Here, the integrals with respect to x are from a to b because $L[c] = 0$ for $x \in (0,a) \times (b,x_L)$. The formula allows us to determine S since all functions appearing on the right hand side are completely determined in their domains

(and set equal to 0 everywhere else). More precisely, by taking the logarithm on both sides of (4.75) for $i = 1, 2$, we obtain a system

$$\ln\left(\int_0^{T^*} \lambda(t)dt\right) + r_1 S = \ln(A(r_1))$$
$$\ln\left(\int_0^{T^*} \lambda(t)dt\right) + r_2 S = \ln(A(r_2)),$$

where $A(r_i)$ denotes the right hand side of (4.75) for $i = 1, 2$. Now, by substracting the second equation from the first, we obtain

$$(r_1 - r_2)S = \ln\left(\frac{A(r_1)}{A(r_2)}\right),$$

which leads to

$$S = \frac{1}{r_1 - r_2} \ln\left(\frac{A(r_1)}{A(r_2)}\right). \tag{4.76}$$

In the last step, we can now try to reconstruct the intensity function λ. This function gives information on the brain-ECF concentrations, in which we are interested. Recall that λ was of the form $\lambda(t) = c_{in}c_{ecf}(t)$, so finding λ gives indeed the concentration we have been searching for. El Badia et al. [10] present two methods for solving the following problem: Given measurements $\{d_2(t); 0 < t < T\}$, determine λ such that

$$d_2(t) = \int_0^t \lambda(\zeta)\Phi(b, t - \zeta)\,\mathrm{d}\zeta; \quad 0 < t < T.$$

Here, $\Phi(x, t)$ is the function (4.69) in the proof of theorem (4.1). Both methods transform the above equation into a linear system. For the first method, the convolution equation is replaced by its approximated version using the trapezoidal rule. To this end, set

$$h = \frac{T^*}{M}, \quad t_m = mh,\ m = 1, \ldots, M, \quad 0 < t_1 < \cdots < t_k < \cdots < t_{m-1} < t_m$$

and denote

$$y_k = d_2(t_k) \quad \text{and} \quad \lambda_k = \lambda(t_k).$$

In each interval (t_k, t_{k+1}), we can now approximate the corresponding integral

$$\int_{t_k}^{t_{k+1}} \lambda(\zeta)\Phi(b, t - \zeta)d\zeta$$

by the trapezoidal rule

$$\frac{h}{2}\left(\Phi(b, t_{m-k-1})\lambda_{k+1} + \Phi(b, t_{m-k})\lambda_k\right).$$

Then,

$$y_m = \frac{h}{2}\sum_{k=0}^{m-1}(\Phi(b, t_{m-k-1})\lambda_{k+1} + \Phi(b, t_{m-k})\lambda_k); \quad m = 1, \ldots, M,$$

which leads to the linear system

$$A\Lambda = Y \tag{4.77}$$

with $Y = (y_k)_k = (d_2(t_k))_k$ and $\Lambda = (\lambda_k)_k$. The second method consists in decomposing λ on a finite Fourier basis l_k

$$\lambda(t) = \sum_{k=1}^{m} \theta_k l_k(t),$$

which leads to the linear system

$$A\Theta = Y. \tag{4.78}$$

The coefficients of A are given by

$$\int_0^{t_m} l_k(\zeta)\Phi(b, t_m - \zeta)d\zeta. \tag{4.79}$$

4.6 Model improvements

We will now discuss some possible improvements of the model derived in this thesis. The model so far neglects the effects of drug elimination or reabsorption of CSF along the spine. A certain portion of CSF may be absorbed into the blood by the spinal venules or drain into the lymph nodes in the neck. The experiments mentioned in McComb [27], in which dyed fluid was injected into the CSF space of rabbits, suggest that about 26% of the CFS drains into the cervical lymphatics, which would imply that drug elimination also takes place in the SAS and hence affects the concentration in the lumbar region. Similar results can be found in Koh et al. [22], where human values are given as well. On the other hand, since CSF is eliminated by flowing into the blood, we would expect the CSF containing the dye to be filtered from the blood into the lymphatics. This would lead to an overestimation of the actual lymphatic contribution to CSF drainage [22]. It is thus not yet

clear to what extend CSF drainage into the lymphatic system needs to be considered. It is also conceivable that some drug compounds may react with certain contents of the CSF, creating a new substance which is no longer detectable as the drug in question. Hence, drug elimination may take place by reaction with other substances along the CSF space.

All of this might mean that neglecting elmination in the model yields insufficiently accurate results. One could therefore add to the equation a reaction term which deals with the above possibilities of absorption, drainage, or reaction, yielding the PDE

$$\partial_t c(x,t) + v\partial_x c(x,t) - \kappa\partial_{xx} c(x,t) + Rc(x,t) = F(x,t); \quad 0 < x < x_L;\ 0 < t < T.$$

Here, R is the reaction coefficient, which is the fraction of the total concentration eliminated in the ways specified in the preceding paragraph. All proofs in section 4.5.1 can be adapted to this case using $\rho = \alpha^2\kappa + R$ with $\alpha = \frac{-v}{2\kappa}$ as before. Of course, R needs to be estimated or determined experimentally. The experimental setup could be as follows: The drug in question is injected into the plasma as usual. The drug is bound to H_1 (hydrogen). Then, the same drug is injected at the same dose but bound to H_3 into the CSF space in the cervival region, that is in the neck. Here, H_3 denotes tritium, a radioactive isotope of hydrogen. It contains one proton and 2 neutrons, whereas the most common isotope of hydrogen H_1 (or simply H) contains one proton and no neutrons. Tritium is often used as a tracer in pharmaceutical or chemical studies. It reacts in the same manner as hydrogen, but is distinguishable from simple hydrogen due to its radioactivity. Tritium can be used to monitor, for example, metabolic processes because it can easily be followed when participating in biochemical reactions. In the lumbar region, CSF is extracted and the concentrations of the drug bound to H_1 and bound to H_3 are measured separately (see figure 31). Comparing these two concentrations should give not only the loss of concentration due to reaction or absorption in the thoracic and lumbar regions of the SAS (in form of the loss of drug bound to H_3), but also loss due to brain metabolism or reabsorption into the cervical lymphatics. The same fraction of the drug bound to H_1 is lost along the thoracic and lumbar regions of the SAS. The additional difference to the initial dose corresponds to the fraction lost due to brain metabolism or ealier drainage.

Moreover, we have shown solvability for a point source only. It may be desirable to have solvability conditions and a stability result also for non-point sources. The assumption we made was that exchange between brain ECF and CSF is only possible in one single point located in the lateral ventricles. Of course, in reality we expect exchange to be possible everywhere in the

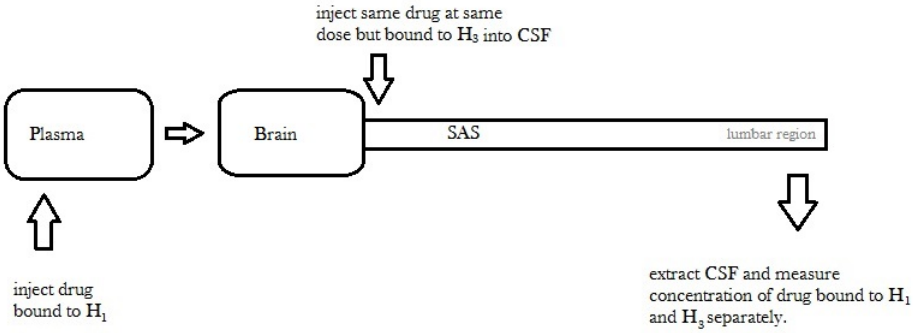

Figure 31: Schematic setup of an experiment to determine the reaction/elimination coefficient R.

lateral ventricles. Also, drugs may enter the CSF in the lateral ventricles directly from the blood via the BCSFB, which has been neglected in our model. Neglecting the input from transport across the BCSFB leads to an overestimation of brain-ECF concentrations. In order to incorporate the transport across the BCSFB, the source term must be modified. For this, it is important to have a reliable estimation of the BCSFB contribution to drug transport into the CSF in order to distinguish between brain-ECF concentration and transport across the BCSFB in the source term. So far, the coefficients are assumed to be constants. However, the coefficients may be time dependent or depend on the spatial variable. Since some transport processes are faster than others, c_{in} might be time dependent. Similarly, we may find it appropriate to use $v = v(x)$ or even $v(x,t)$ instead of assuming a constant velocity. The experiments in [29] and [17] suggest that the CSF flow velocity depends on the location within the CSF space as well as the time within the cardiac cycle.

So far, we can only model the change in the CSF concentration, whereas the compartmental models also describe the change in plasma and brain-ECF concentrations. Since the PDE describes all of the CSF compartments together, it is possible to incorporate the PDE model into the compartmental models by replacing the ODEs for the CSF compartments with the PDE while continuing to use the ODEs for plasma and brain ECF concentrations. The resulting model will describe the full process of drug distribution from injection into the blood to distribution along the CSF space.

Finally, a generalization to a 2D or even 3D model might be more accurate.

5 Result

The model developed in this thesis can be seen as a basic model calling for further research. Many simplifications were made to prove the solvability of the problem of reconstructing brain-ECF values from CSF measurements in the lumbar region. Not enough data is available to test the method for solving this inverse problem. The numerical results in [10] suggest that the method presented works quite well if the data-noise level is small enough, despite the fact that the errors are multiplied by solving the problems for $c(x, T^*)$ and $\partial_x c(a,t),\ \partial_x c(b,t)$ first and using the obtained functions for solving the last problem. For the actual validation of the model, it is necessary to have brain-ECF and lumbar-CSF data for the same species and obtained under the same conditions, that is in the same study. Only by means of such data would it be possible to see how accurately the model – along with the method introduced in section 4.5 – reconstructs brain-ECF values from lumbar-CSF measurements. Also, reliable estimates for c_{in} and κ need to be given for the drug in question.

The main problems we faced were due to insufficient data or difficulties with the solvability of the model. Many simplifications we made were necessary to guarantee the solvability of the inverse problem. The method introduced by El Badia et al. works only for point sources. Most existing methods for inverse advection-diffusion equations allowing more flexibility in the coefficients have only been proved to work for time independent source terms. For this reason, they cannot be used for reconstructing brain-ECF concentrations. Hence, further research is necessary on the cases of a non-point source and of variable velocity. Concerning the data problem, recall the following figure (figure 32 = figure 12 above), which is a reproduction of figure 7 in Westerhout et al. [39]. It must be noted that the time scales in different parts of the figure vary from 120 min up to 1680 min. Therefore, the full elimination of methotrexate is not yet visible in some plots. In parts D and E of this figure, we can clearly see the travelling-wave effect as obtained for the analytical solution from the advection term (recall figure 10). In A to C, this effect is not visible. The SAS in a rat is very short while the time scale in the figure is rather coarse, which makes it probable that the travelling-wave effect is present but too small to be visible in the figure. It is remarkable that the humanized model seems to underestimate the real brain-ECF and CSF concentrations. There are three main reasons for this observation. Firstly, it may be due to errors in extrapoliting from the rat to the human model. Secondly, the models were obtained using data from healthy rats while the human data was obtained under disease conditions. The disease state may have affected the permeability of the barriers and the CSF turnover, which then affects

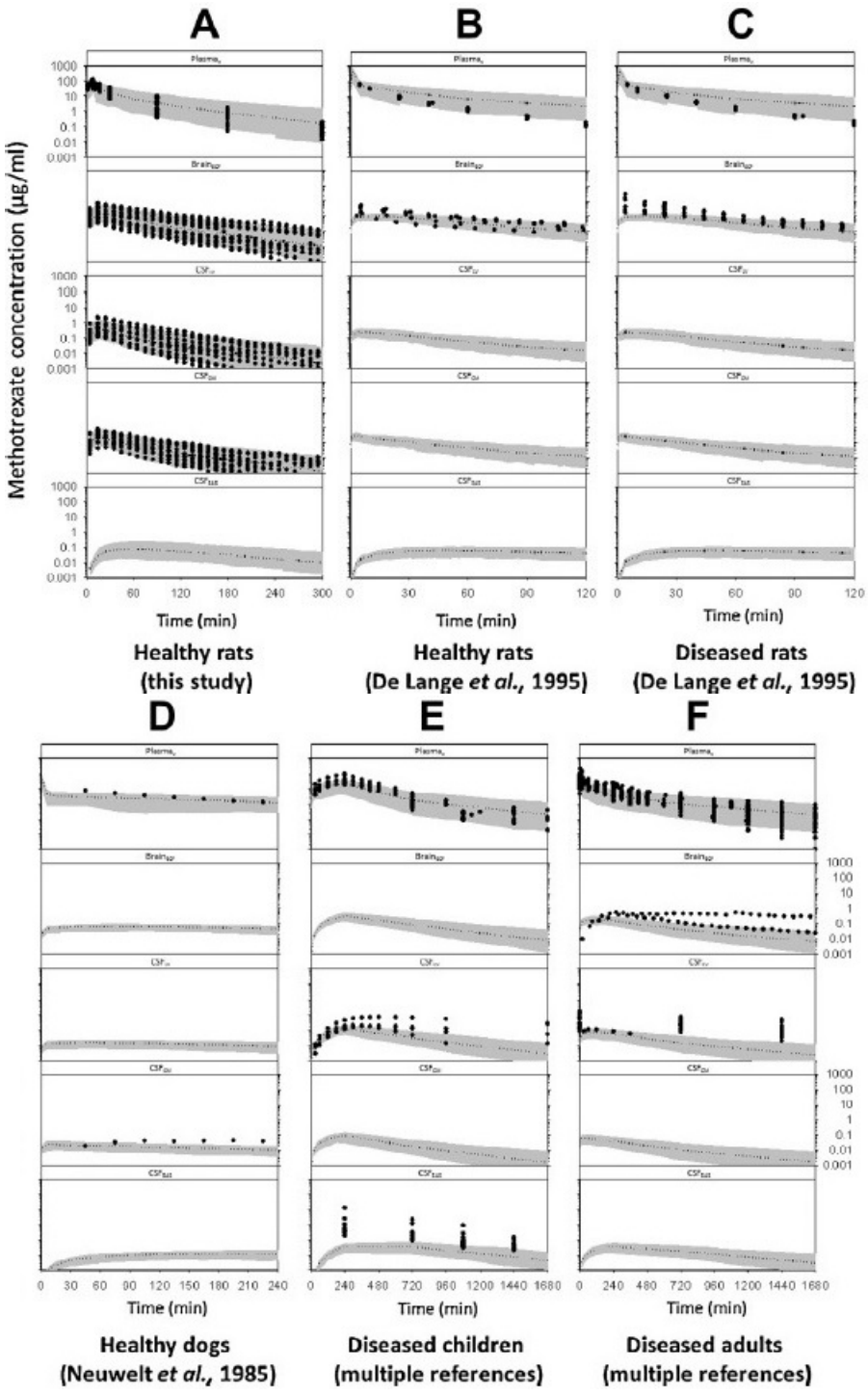

Figure 32: Concentration-time profiles for methotrexate concentrations in different species. Black dots are observed data. Grey areas show the 95% confidence interval. Reproduction of figure 7 in [39].

Reference	Subjects	Disease state	Dose (i.v.)	$Brain_{ECF}$ or CSF (sampling method)	Analysis
This manuscript	Rats (225–275 g; $n = 32$)	Healthy	40 and 80 mg/kg (10 min)	$Brain_{ECF}$, CSF_{LV} and CSF_{CM} (microdialysis)	LC/MS/MS
De Lange et al. (1995)	Rats (160–200 g; $n = 12$)	Brain tumor (implanted rhabdomyo-sarcoma)	75 mg/kg (bolus)	$Brain_{ECF}$ (microdialysis)	HPLC
Neuwelt et al. (1985)	Dogs (20–25 kg; $n = 4$)	Healthy	4 mg/kg (bolus) + Evans blue	CSF_{CM} (Cisterna magna sampling)	Radioimmuno-assay
Aumente et al. (2006)	Children (0.5–17 y; $n = 49$)	Acute lymphoblastic leukemia	3 g/m^2 (4 h) + remission-induction therapy (>24 h after methotrexate)	N.A.	Fluorescence polarization immunoassay
Blakeley et al. (2009)	Human adults (>18 y; $n = 2$)	Recurrent high grade gliomas	12 g/m^2 (4 h) + sodium bicarbonate	$Brain_{ECF}$ (intratumoral microdialysis)	LC/MS
Bore et al. (1987)	Human adults (17–67 y; $n = 8$)	Sarcoma, carcinoma, lung metastasis	50 mg/m^2 (bolus) (no co-medication)	N.A.	Radioimmuno-assay
Chatelut et al. (1991)	Children (2–17 y; $n = 4$)	Medulloblastoma or ependymo-blastoma	8 g/m^2 (4 h) + sodium bicarbonate (i.v. hydration)	CSF_{LV} (ventriculo-peritoneal derivation)	Enzymatic assay
Colom et al. (2009)	Children ($n = 14$)	Osteosarcoma	12 g/m^2 (4 h) + leucovorin (24 h after methotrexate)	N.A.	Fluorescence polarization immunoassay
Creaven et al. (1976)	Human adults ($n = 9$)	Small-cell lung carcinoma	15 mg/m^2 (bolus) (no co-medication)	N.A.	[^{3}H]Metho-trexate
Glantz et al. (1998)	Human adults (>18 y; $n = 16$)	Histologically diagnosed primary tumor	8 g/m^2 (4 h) + leucovorin (24 h after methotrexate)	CSF_{LV} (ventricular reservoirs)	Fluorescence polarization immunoassay
Herman et al. (1989)	Human adults (16–80 y; $n = 41$)	Rheumatoid arthritis	10 mg/m^2 (bolus) (no co-medication)	N.A.	Radiochemical-ligand binding assay
Shapiro et al. (1975)	Human adults (18–63 y; $n = 21$)	Leukemia and carcinoma	50 mg (bolus) (no co-medication)	CSF_{LV} (Ommaya reservoir)	Microbiologic disk assay
Stewart et al. (1990)	Human adults (30–78 y; $n = 15$)	Rheutmatoid arthritis	15 mg (5 min) (received naproxen 3 days before)	N.A.	Radioenzymatic assay
Vassal et al. (1990)	Children (3–15 y; $n = 29$)	NonHodgkin's lymphoma	3 g/m^2 (3 h) + sodium bicarbonate (i.v. hydration)	CSF_{SAS} (lumbar puncture)	Fluorescence polarization immunoassay

Figure 33: Data used for figure 32, n=number of subjects (sample size). Reproduction of table 4 in [39].

the drug distribution and hence the concentration profiles. This is supported by the fact that we see a similar underestimation in diseased rats, albeit in the first part of the time interval, while in humans it is stronger later on. However, the last reason may be the most important: The human data was obtained in multiple studies. Several studies have shown that CSF turnover and permeability of the barriers are highly sensible to a great variety of parameters, so measurements obtained under different conditions in different studies will differ severely. We can see in figure 33, which is a reproduction of table 4 in Westerhout et al. [39], that the subjects from whom the data used in figure 32 was obtained had different diseases or disease states. Moreover, some studies measured concentrations from as few as two people. The human data is therefore of very limited comparability. Yet again, we face the problem of insufficient data.

A Sobolev spaces

In sections 4.4 and 4.5, we used certain properties of Lebesgue functions and Sobolev spaces. The definitions and major theorems shall be given here. In order to define the Sobolev spaces, we need a weaker form of differentiation. The idea behind the weak derivative is related to that behind integration by parts. Before giving the definition of a weak derivative, we introduce a multiindex as follows.

Definition A.1 (Multiindex)**.** A vector

$$\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$$

is called a multiindex. The order of α is $|\alpha| = \alpha_1 + \cdots + \alpha_n$. In the symbol D^α, the α_i indicates how many times we differentiate in the ith direction.

Example A.1. Let $n = 3$ and $u \in C^3$. Then, for $\alpha = (1, 0, 2)$ we have $|\alpha| = 3$ and

$$D^\alpha u = \frac{\partial^3 u}{\partial x_1 \partial x_3^2}.$$

Definition A.2 (Weak derivative)**.** Let $\Omega \subset \mathbb{R}^n$ open and define

$$L^{1,loc}(\Omega) = \{f : \Omega \to \mathbb{R} \mid f \in L^1(K) \text{ for any compact set } K \subset \Omega\}.$$

If there exists $v \in L^{1,loc}(\Omega)$ such that

$$\int_\Omega u(x) D^\alpha \phi(x) dx = (-1)^\alpha \int_\Omega v(x)\phi(x) dx; \quad \forall \phi \in C_0^\infty(\Omega),$$

then v is called the weak derivative of y of order α. In this case, we write $v = D^\alpha u$.

In the above definition, $C_0^\infty(\Omega)$ is the set of $C^\infty(\Omega)$ having a compact support in Ω. Note that $\operatorname{supp}\phi = \{x \in \Omega \mid \phi(x) \neq 0\}$ implies that any $\phi \in C_0^\infty(\Omega)$ vanishes on the boundary $\partial\Omega$.

Definition A.3 (Sobolev spaces)**.** Let $\Omega \in \mathbb{R}^n$.

i.) The Sobolev space $W^{k,p}(\Omega)$ is the space of all functions $u \in L^p$ such that for every multiindex α with $|\alpha| \leq k$ the weak derivative $D^\alpha u$ exists and belongs to $L^p(\Omega)$, that is

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega) \text{ for all } |\alpha| \leq k\}.$$

ii.) On $W^{k,p}(\Omega)$ we introduce the Sobolev norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \leq k} \int_\Omega |D^\alpha u|^p dx \right)^{1/p} = \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{1/p}; \qquad 1 \leq p < \infty$$

$$\|u\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha| \leq k} \operatorname{ess\,sup}_{x \in \Omega} |D^\alpha u| \; ; \qquad p = \infty.$$

The Sobolev spaces are Banach spaces with respect to the Sobolev norm.

iii.) For $p = 2$, $W^{k,p}$ is a Hilbert space. We write $W^{k,2}(\Omega) = H^k(\Omega)$ to distinguish this case from the other Sobolev spaces.

iv.) $H^s_0(\Omega)$ is the closure of $C^\infty_0(\Omega)$ in $H^s(\Omega)$.

Of course, ii.) and iii.) are not merely definitions; a proof can be found in [6].

Theorem A.1 (Sobolev embedding theorem)**.** *Let $1 \leq p < \infty$ and $k \in \mathbb{N}_0$. Let further $\Omega \subset \mathbb{R}^n$ bounded open subset with C^1 boundary. Then the following continuous embeddings hold.*

i.) If $kp < n$:

$$W^{k,p}(\Omega) \hookrightarrow L^q(\Omega); \quad \textit{with } \frac{1}{q} = \frac{1}{p} - \frac{k}{n}.$$

ii.) If $kp = n$:

$$W^{k,p} \hookrightarrow L^q(\Omega); \quad \forall 1 \leq q < \infty.$$

iii.) If $kp > n$:

$$W^{k,p}(\Omega) \hookrightarrow C(\bar{\Omega}).$$

On $C(\bar{\Omega})$ the norm $\max_{x \in \bar{\Omega}} |u(x)|$ is imposed.

A proof of this theorem can be found in [6] as well as [2].

B The Dirac delta function

The delta function plays an important role in modelling point sources. Since a point source is only active in one point, it should be zero elsewhere, which yields a source term of the form

$$F(x,t) = \lambda(t)\delta(x - S), \tag{B.1}$$

where λ is the intensity of the source and S the location. There are different ways to define the delta function $\delta(x)$. One possibility is to define it by its properties

- $\delta(x) = \begin{cases} 0\,; & x \neq 0 \\ \infty\,; & x = 0 \end{cases}$
- $\int_{-\infty}^{\infty} \delta(x) f(x)\, dx = f(0)$.

The second property implies that

$$\int_a^b \delta(x) f(x - S)\, dx = f(S)$$

if $a < S < b$ or zero otherwise. It must be noted that the delta function is in fact not a function, which is why it is also called delta distribution. Another definition often found in the literature is $\delta(x) = H'(x)$, where the derivative is to be understood in the weak sense and H is the so-called Heavyside function

$$H(x) = \begin{cases} 0\,; & x < 0 \\ 1\,; & x > 0 \end{cases}.$$

We can also define δ as the limit of a sequence of functions. For example, the delta function can be defined as one of the following limits for $\epsilon \to 0$:

$$\delta(x) = \frac{1}{\pi} \lim_{\epsilon \to 0} \frac{\epsilon}{x^2 + \epsilon^2} \tag{B.2}$$

$$= \lim_{\epsilon \to 0} \frac{1}{2} \epsilon |x|^{\epsilon - 1} \tag{B.3}$$

$$= \lim_{\epsilon \to 0} \frac{1}{2\sqrt{\pi\epsilon}} e^{-\frac{x^2}{4\epsilon}} \tag{B.4}$$

$$= \lim_{\epsilon \to 0} \frac{1}{\pi x} \sin(\frac{x}{\epsilon}). \tag{B.5}$$

Moreover, it is possible to find a weak derivative of $\delta(x)$. Using definition A.2, one finds that $\delta'(x)$ is defined by the property

$$\int_a^b \delta'(x) f(x)\, dx = -\int_a^b \delta(x) f'(x)\, dx = f'(0)$$

for $a < x < b$. From (B.2) we can even derive an expression for $\delta(x)$ as the limit of a sequence of functions [33]

$$\delta'(x) = -\frac{2}{\pi} \lim_{\epsilon \to 0} \frac{x\epsilon}{(\epsilon^2 + x^2)^2}. \tag{B.6}$$

C Tikhonov regularization: proofs

We will prove that the Tikhonov functional

$$J_\alpha(\phi) \coloneqq \|T\phi - g^\delta\|^2 + \alpha\|\phi - \phi_0\|^2 \tag{C.1}$$

for perturbed data g^δ and $\alpha > 0$, has a unique minimum for all initial guesses ϕ_0. Moreover, we will show that the mimimum ϕ_α^δ depends continuously on the perturbed data. Some preparations are needed for the proof of the main theorem.

Definition C.1. Let $X,\ Y$ be normed spaces and $U \subset X$ an open subset. A mapping $F : U \to Y$ is called Fréchet differentiable at $\phi \in U$ if there exists a bounded linear operator $F'[\phi] : X \to Y$ such that

$$\lim_{h\to 0} \frac{1}{\|h\|}(F(\phi + h) - F(\phi) - F'[\phi]h) = 0.$$

$F'[\phi]$ is called the Fréchet derivative of F at ϕ. F is called Fréchet differentiable if it is differentiable at every $\phi \in U$.

For Fréchet differentiable functions F the following lemma holds.

Lemma C.1. *Let U be an open subset of a normed space X. If $F : U \to \mathbb{R}$ is Fréchet differentiable at ϕ and if $\phi \in U$ is a local minimum of F, then $F'[\phi] = 0$.*

Proof. Let ϕ be a local minimum of F and assume that $F'[\phi] \neq 0$. After changing the sign of h if necessary, assume that $F'[\phi]h < 0$. Then

$$\lim_{\epsilon \searrow 0} \frac{F(\phi + \epsilon h) - F(\phi)}{\epsilon} = F'[\phi]h < 0.$$

This contradicts the assumption that ϕ was a local minimum of F. □

Lemma C.2. *The Tikhonov functional is Fréchet differentiable for every $\alpha \geq 0$ and the Fréchet derivative is given by*

$$J'_\alpha[\phi]h = 2\,Re\langle T^*(T\phi - g^\delta) + \alpha(\phi - \phi_0), h\rangle. \tag{C.2}$$

Proof. Follows from

$$J_\alpha(\phi + h) - J_\alpha(\phi) - J'_\alpha[\phi]h = \|Th\|^2 + \alpha\|h\|^2.$$

□

Finally, recall the following theorem.

Theorem C.1 (Lax-Milgram)**.** *Let X be a Hilbert space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and $a : X \times X \to \mathbb{K}$ a sesquilinearform. Let there exist constants $0 < c \leq C < \infty$ such that for all $x, y \in X$:*

i) $|a(x, y)| \leq C\|x\|_X \|y\|_X$,

ii) $Re\, a(x, x) \geq c\|x\|_X^2$.

Then, there is a unique linear operator $A : X \to X$ such that

$$a(y, x) = \langle y, Ax\rangle_X; \quad \forall x, y \in X.$$

Furthermore, A is invertible and

$$\|A\| \leq C; \quad \|A^{-1}\| \leq \frac{1}{c}.$$

A proof can be found in [25].

Now we can state and prove the following useful theorem.

Theorem C.2. *The Tikhonov functional J_α has a unique minimum ϕ_α^δ in X for all $\alpha > 0$, $g^\delta \in Y$ and $\phi_0 \in X$. This minimum is given by*

$$\phi_\alpha^\delta = (T^*T + \alpha I)^{-1}(T^* g^\delta + \alpha\phi_0), \tag{C.3}$$

where T^* is the conjugate transpose of T. The operator $T^*T + \alpha I$ is boundedly invertible, so ϕ_α^δ depends continuously on g^δ.

Proof. Let us assume that ϕ_α^δ minimizes the Tikhonov functional J_α. Then $J'[\phi_\alpha^\delta]h = 0$ for all $h \in X$ by lemma C.1 and lemma C.2. Chosing $h = T^*(T\phi - g^\delta) + \alpha(\phi - \phi_0)$ in (C.2) implies that

$$(T^*T + \alpha I)\phi_\alpha^\delta = T^* g^\delta + \alpha\phi_0,$$

since in this case

$$J'_\alpha[\phi]h =< h, h >= \|h\|^2 = 0 \Rightarrow h = 0.$$

By theorem C.1 and the inequality

$$Re\langle (T^*T + \alpha I)\phi, \phi\rangle = \|T\phi\|^2 + \alpha\|\phi\|^2 \geq \alpha\|\phi\|^2,$$

the operator $T^*T + \alpha I$ is invertible, since condition i) of theorem C.1 follows from the Cauchy-Schwarz inequality. Note that for all $h \in X \setminus \{0\}$, the function $\psi(t) := J_\alpha(\phi_\alpha^\delta + th)$ is a polynomial of degree 2 with $\psi \geq 0$ and $\psi'(0) = 0$, which implies that $\psi(t) \geq \psi(0)$ for all $t \in \mathbb{R}$ and equality only when $t = 0$. Therefore, ϕ_α^δ defined as in (C.3) indeed minimizes the Tikhonov functional J_α. □

D Proofs of the theorems in section 4.5

In the following, we will give the theorems used in section 4.5 and the proofs by El Badia et al. in [10].

Lemma D.1. *Let $T^* < T$ and $b \in (0, x_L)$ be a strategic point relative to the family $\{\psi_n\}$. Suppose that $w = w(x,t)$ satisfies*

$$\begin{aligned} &L[w](x,t) = 0; \quad 0 < x < x_L,\ T^* < t < T \\ &w(0,t) = \partial_x w(x_L,t) = 0; \quad T^* < t < T \\ &w(\cdot, T^*) \in L^2(0, x_L). \end{aligned} \tag{D.1}$$

Then

$$w(b,t) = 0 \text{ for all } t \in (T^*, T) \Rightarrow w(\cdot, T^*) = 0 \text{ in } L^2(0, x_L)$$

Proof. Let

$$z(x,t) = e^{\alpha x} w(x,t) \tag{D.2}$$

with $\alpha = \frac{-v}{2\kappa}$. Then, w is the solution to (D.1) if and only if z is the solution to the following equation:

$$\begin{aligned} &\partial_t z - \kappa \partial_{xx} z + \rho z = 0; \quad 0 < x < x_L,\ T^* < t < T \\ &z(0,t) = \partial_x z(x_L,t) - \alpha z(x_L,t) = 0; \quad T^* < t < T \\ &z(\cdot, T^*) \in L^2(0, x_L). \end{aligned} \tag{D.3}$$

Thus, z is given by the Fourier expansion in L^2-sense

$$z(x,t) = \sum_{n=0}^{\infty} \langle z(\cdot, T^*), \psi_n \rangle_{L^2} \psi_n(x) e^{-\mu_n (t - T^*)}, \tag{D.4}$$

where the μ_n are the eigenvalues (4.48) and $\langle f, g \rangle_{L^2} = \int_0^{x_L} f(x) g(x) dx$ denotes the usual inner product in $L^2(0, x_L)$. It can be seen from (4.50) that the right hand side of (D.4) converges uniformly for all $t \geq t_0 > T^*$ by the Cauchy criterion since $e^{-\mu_n (t - T^*)}$ goes to 0 exponentially fast as $n \to \infty$. Also, the right hand side gives an analytic function of $t \in (T^*, \infty)$ for every $x \in (0, x_L)$. This gives a sense to $z(b,t)$ and $w(b,t)$ for $t > T^*$. Now, if $w(b,t) = 0$ for all $t \in (T^*, T)$, then by definition

$$z(b,t) = 0; \quad \forall t \in (T^*, T).$$

By analytic continuation, we can conclude that

$$\sum_{n=0}^{\infty} \langle z(\cdot, T^*), \psi_n \rangle_{L^2(0,x_L)} \psi_n(b) e^{-\mu_n (t - T^*)} = 0; \quad \forall t \in (T^*, \infty). \tag{D.5}$$

Therefore, we may deduce successively that all coefficients in (D.5) vanish, so

$$\langle z(\cdot, T^*), \psi_n \rangle_{L^2(0,x_L)} \psi_n(b) = 0; \quad \forall n \in \mathbb{N}.$$

Since b is strategic with respect to the ψ_n, so $\psi_n(b) \neq 0$ for all n, this means

$$\langle z(\cdot, T^*), \psi_n \rangle_{L^2(0,x_L)} = 0; \quad \forall n \in \mathbb{N}.$$

Thus $z(\cdot, T^*) = 0$ in $L^2(0, x_L)$ and therefore, by the definition of $z(x,t)$, we also have $w(b,t) = 0$ in $L^2(0, x_L)$. □

In the proof of the main theorem on the identifiability of the solution, we also used the following theorem. A proof can be found in [9].

Theorem D.1 (Titchmarsh)**.** *Let $\phi(t), \psi(t)$ be integrable functions such that*

$$\int_0^x \phi(t)\psi(x-t)\,dt = 0$$

almost everywhere in the interval $0 < x < \nu$. Then there exist $\lambda \geq 0, \mu \geq 0$ with $\lambda + \mu \geq \nu$ such that $\phi(t) = 0$ almost everywhere in $(0, \lambda)$ and $\psi(t) = 0$ almost everywhere in $(0, \mu)$.

For the proof of the stability result, consider the set

$$G(T^*) = \{(\lambda, S) \in L^2(0, x_L) \times (a, b) \mid \lambda(t) \geq 0 \text{ with } \lambda(t) = 0 \text{ for } t \geq T^*\}.$$

Then, for $(\lambda, S), (\mu, \tau) \in G(T^*)$ and h sufficiently small, $(\lambda + h\mu, S + h\tau) \in G(T^*)$. We can therefore define the source term

$$F^h(x,t) = (\lambda(t) + h\mu(t))\delta(x - (S + h\tau)). \tag{D.6}$$

Moreover, by using Taylor expansion with respect to h, we see that there is a point S_h satisfying $|S - S_h| < |\tau h|$ such that

$$F^h(x,t) = F(x,t) + h\hat{F}(x,t) + h^2\tilde{F}(x,t),$$

where

$$\begin{aligned}
\hat{F}(x,t) &= \mu(t)\delta(x - S) - \lambda(t)\tau\delta'(x - S) \\
\tilde{F}(x,t) &= -\mu(t)\delta'(x - S) + \frac{1}{2}\lambda(t)\tau^2\delta''(x - S_h).
\end{aligned}$$

Hence,

$$c(x,t;F^h) = c(x,t;F) + hc(x,t;\hat{F}) + h^2 c(x,t;\tilde{F}),$$

where $c(x,t;\hat{F})$ and $c(x,t;\tilde{F})$ are solutions of (4.39) with the initital and boundary conditions

$$\begin{aligned} c(0,t) &= \partial_x c(x_L,t) = 0 \\ c(x,0) &= 0 \end{aligned}$$

and with $\hat{F}$ and $\tilde{F}$, respectively, as source terms. Therefore,

$$\lim_{h\to 0} \frac{B[F^h] - B[F]}{h} = \{c(a,t;\hat{F}), c(b,t;\hat{F})\}. \tag{D.7}$$

Furthermore, since $\lambda, \mu \in L^2(0,x_L)$, it can be proved by means of a transformation method that the function $c(\cdot,\cdot;\hat{F})$ belongs to $L^2((0,x_L)\times(0,T))$, and hence it also belongs to $C([0,T];H^{-1}(0,x_L))$, where H^{-1} denotes the dual space to H_0^1.[4] This regularity result is needed to justify the integration by parts below.

Now we can prove the local Lipschitz stability result as given in section 4.5 (see theorem 2 in [10]).

Theorem D.2. *Let $\lambda, \mu \in G(T^*)$. If $(\mu,\tau) \neq (0,0)$ for $0 < t < T$, then*

$$\lim_{h\to 0} \frac{B[F^h] - B[F]}{h} \neq 0.$$

Proof. We will prove that

$$\{c(a,t;\hat{F}), c(b,t;\hat{F})\} \neq \{0,0\}; \quad 0 < t < T$$

by showing that $\{c(a,t;\hat{F}), c(b,t;\hat{F})\} = \{0,0\}$ implies that $(\lambda,\tau) = (0,0)$. So suppose that $c(a,t;\hat{F}) = c(b,t;\hat{F}) = 0$. First, using the same techniques as in the proof of theorem 4.1, we can show that

$$c(x,T^*;\hat{F}) = 0; \quad 0 < x < x_L \tag{D.8}$$

and

$$\partial_x c(a,t;\hat{F}) = \partial_x c(b,t;\hat{F}) = 0; \quad 0 < t < T^*. \tag{D.9}$$

Note that in (T^*,T) the problem becomes

$$\begin{aligned} &L[c](x,t) = 0; \quad 0 < x < x_L,\ T^* < t < T \\ &c(0,t) = \partial_x c(x_L,t) = 0; \quad T^* < t < T \\ &c(x,0) = 0, \end{aligned}$$

[4] El Badia et al. [10] adopt this method from Lions [26]. See this for details.

because $\hat{F}(x,t) = \mu(t)\delta(x-S) - \lambda(t)\tau\delta'(x-S)$ and $\mu, \lambda = 0$ for $t \geq T^*$. Since $c(b,t;\hat{F}) = 0$ by assumption, lemma 4.1 implies that

$$c(x,T^*;\hat{F}) = 0; \quad 0 < x < x_L.$$

In $(0,T^*)$, we now have the problem

$$L[c](x,t) = \mu(t)\delta(x-S) - \lambda(t)\tau\delta'(x-S) \tag{D.10}$$
$$c(0,t) = \partial_x c(x_L,t) = 0 \tag{D.11}$$
$$c(x,0) = 0. \tag{D.12}$$

By definfition, $\delta(x_S) = 0$ for all $x \neq 0$. Also, $\delta'(x-S)$ can be written as the limit

$$\delta'(x-S) = -\frac{2}{\pi}\lim_{\epsilon\to 0}\frac{\epsilon(x-S)}{(\epsilon^2+(x-S)^2)^2}$$

according to (B.6). Now, since $0 < a < S < b < x_L$, $\delta(x-S) = 0$ and also $\delta'(x-S) = 0$ in $(0,a)\times(0,T^*)\cup(b,x_L)\times(0,T^*)$, which means

$$c(x,t;\hat{F}) = 0 \text{ in } (0,a)\times(0,T^*)\cup(b,x_L)\times(0,T^*).$$

By (D.10)-(D.12), we conclude that

$$\partial_x c(a,t;\hat{F}) = \partial_x c(b,t;\hat{F}).$$

Now consider an infinitely differentiable function $\xi \in \mathcal{D}(0,x_L)$ such that $\xi(x) = 1$ in a neighbourhood of S. Through multiplying the equation

$$L[c](x,t) = \hat{F}(x,t)$$

with $L[c]$ as in (4.38) by $\xi(x)e^{r_i x}$, where the r_i are the solutions to the characteristic equation $-\kappa r^2 - vr = 0$, and integrating with respect to x from 0 to x_L, we obtain

$$\begin{aligned}\mu(t)e^{r_i S} + \lambda(t)\tau r_i e^{r_i S} = \\ \int_0^{x_L}(\partial_t c(x,t) - v\partial_x c(x,t) - \kappa\partial_{xx}c(x,t))\xi(x)e^{r_i x}dx.\end{aligned} \tag{D.13}$$

By using integration by parts and (D.8) and (D.9), we find that

$$\mu(t)e^{r_i S} + \lambda(t)\tau r_i e^{r_i S} = \left\langle \partial_t c(x,t;\hat{F}), \xi(x)e^{r_i x}\right\rangle_{H^{-1},H_0^1}, \quad 0 < t < T^*, \tag{D.14}$$

since all other terms vanish. Here $\langle\cdot,\cdot\rangle_{H^{-1},H_0^1}$ is the duality bracket between H^{-1} and H_0^1 (that is a generalized form of the L^2 inner product). Now, since

$$\partial_t c(x,t;\hat{F}) \in L^1(0,T;H^{-1}(0,x_L)) \text{ and } \xi(x)e^{r_i x} \in H_0^1(0,x_L)$$

$$c(x, T^*; \hat{F}) = c(x, 0; \hat{F}) = 0; \quad 0 < x < x_L$$

we have

$$\int_0^{T^*} \langle \partial_t c(x,t;\hat{F}), \xi(x) e^{r_i x} \rangle_{H^{-1}, H_0^1} dt = \left\langle \int_0^{T^*} \partial_t c(x,t;\hat{F}) dt, \xi(x) e^{r_i x} \right\rangle_{H^{-1}, H_0^1} = 0.$$

Thus,

$$\int_0^{T^*} \mu(t) dt + \tau r_1 \int_0^{T^*} \lambda(t) dt = 0$$
$$\int_0^{T^*} \mu(t) dt + \tau r_2 \int_0^{T^*} \lambda(t) dt = 0.$$

Since $r_1 \neq r_2$ and $\int_0^{T^*} \lambda(t) dt > 0$, this implies $\tau = 0$ and $\int_0^{T^*} \mu(t) dt = 0$. Using the same arguments as in the proof of theorem (4.1), we find that $\int_0^{T^*} \mu(t) dt = 0$ implies that $\mu(t) = 0$, which ends the proof of the theorem. $\square$

References

[1] N. J. Abbott et al.: *Structure and function of the blood–brain barrier*; Neurobiology of Disease, 37(1):13-25; 2010.

[2] R. A. Adams: *Sobolev Spaces*; Academic Press; 1975.

[3] H. Bauer: *Measure and Integration Theory*; De Gruyter; 2001.

[4] H. Bellout, A. Friedman: *Identification problems in potential theory*; Archieve for Rational Mechanics and Analysis, 101(2):143-160; 1988.

[5] D. Braess: *Finite Elements*; Cambridge University Press; 2007.

[6] A. Bressan: *Lecture Notes on Sobolev Spaces* (lecture notes); Pennsylvania State University; 2012.

[7] H. S. Carslaw, J. C. Jaeger: *Conduction of Heat in Solids*, second edition; Oxford University Press; 1959.

[8] F. C. Courtice, W. J. Simmonds: *The removal of protein from the subarachnoid space*; Australian Journal of Experimental Biology and Medical Sciences, 29(4):255-263; 1951.

[9] R. Doss: *An elemantary proof of Titchmarsh's convolution theorem*; Proceedings of the American Mathematical Society, 104(1):181-184; 1988.

[10] A. El Badia et al.: *Identification of a point source in a linear advection-dispersion-reaction equation: application to a pollution source problem*; Inverse Problems, 21(3):1121-1136; 2005.

[11] A. El Badia, A. Hamdi: *Inverse source problem in an advection-dispersion-reaction system: application to water pollution*; Inverse Problems, 23(5):2103-2120; 2007.

[12] A. El Jai, A. J. Pritchard: *Capteurs et actionneurs dans l'analyse des systèmes distribués*; Masson; 1986.

[13] K. Eriksson: *Computational Differential Equations*, vol. 1; Cambridge University Press; 1996.

[14] M. van Esdonk: *The impact of a change in cerebrospinal fluid turnover on the prediction of brain target site concentrations in rats and humans* (master thesis); Universiteit Leiden; 2014.

[15] O. Forster: *Analysis 3*; Vieweg Verlag; 1999.

[16] M. T. van Genuchten, W. J. Alves: *Analytical Solutions of the One-Dimensional Convective-Dispersive Solute Transport Equation*; U.S. Department of Agriculture, Technical Bulletin No. 1661; 1982.

[17] S. Gupta et al.: *Cerebrospinal fluid dynamics in the human cranial subarachnoid space: an overlooked mediator of cerebral disease. I. Computational model*; Journal of the Royal Society Interface, 7(49):1195-1204; 2010.

[18] R. Herzog: *Skript zur Vorlesung Einführung in Sobolevräume* (lecture notes); Technische Universität Chemnitz; 2012.

[19] T. Hohage: *Inverse Problems* (lecture notes); Georg-August-Universität Göttingen; 2002.

[20] A. Jüngel: *Mathematische Modellierung mit Differentialgleichungen* (lecture notes); Johannes-Gutenberg-Universität Mainz; 2003.

[21] J. B. Keller: *Inverse Problems*; The American Mathematical Monthly, 83(2):107-118; 1976.

[22] L. Koh et al.: *Integration of the subarachnoid space and lymphatics: Is it time to embrace a new concept of cerebrospinal fluid absorption?*; Cerebrospinal Fluid Research, 2:6 (published online); 2005.

[23] E. C. M. de Lange: *Utility of CSF in translational neuroscience*; Journal of Pharmacokinetics and Pharmacodynamics, 40(3):315-326; 2013.

[24] E.C.M. de Lange: *The mastermind approach to CNS drug therapy: translational prediction of human brain distribution, target site kinetics, and therapeutic effects*; Fluids and Barriers of the CNS, 10:12 (published online); 2013.

[25] C. De Lellis: *Funktionalanalysis* (lecture notes); Universität Zürich; 2008.

[26] J. L. Lions: *Pointwise control for distributed systems*; Control and Estimation in Distributed Parameters Systems, ed H. T. Banks; SIAM; 1992.

[27] J. G. McComb: *Recent research into the nature of cerebrospinal fluid formation and absorption*; Journal of Neurosurgery, 59(3):369-383; 1983.

[28] J. Molenaar, E. van Groesen: *Continuum Modeling in the Physical Sciences*; Society for Industrial and Applied Mathematics (SIAM); 2007.

[29] N. Nitz et al.: *Flow Dynamics of Cerebrospinal Fluid: Assessment with Phase-Contrast Velocity MR Imaging Performed with Retrospective Cardiac Gating*; Radiology 183(2):395-405; 1992.

[30] A. Ogata, R. B. Banks: *A Solution of the Differential Equation of Longitudinal Dispersion in Porous Media*; U.S. Geological Survey Professional Paper 411-A:A1-A7; 1961.

[31] L. Ridolfi, M. Macis: *Identification of Source Terms in Nonlinear Convection Diffusion Phenomena by Sinc Collocation-Interpolation Methods*; Mathematical and Computer Modelling, 26(2):69-79; 1997.

[32] F. Roberts, D. Freshwater-Turner: *Pharmacokinetics and anaesthesia*; Continuing Education in Anaesthesia, Critical Care & Pain, 7(1): 25-29; 2007.

[33] H. Schlitt: *Systemtheorie für regellose Vorgänge*; Springer; 1960.

[34] K. Selmaj: *Pathophysiology of the blood-brain barrier*; Immunoneurology, ed M. Chofflon, L. Steinman, 175-191; Springer; 1996.

[35] S. A. Socolofsky, G. H. Jirka: *Special Topics in Mixing and Transport Processes in the Environment*, vol.5 (lecture notes); Texas A&M University; 2005.

[36] J. G. Veening, H. P. Barendregt: *The regulation of brain states by neuroactive substances distributed via the cerebrospinal fluid; a review*; Cerebrospinal Fluid Research, 7:1 (published online); 2010.

[37] J. Westerhout et al.: *Preclinical Prediction of Human Brain Target Site Concentrations: Considerations in Extrapolating to the Clinical Setting*; Journal of Pharmaceutical Sciences, 100(9):3577-3593; 2011.

[38] J. Westerhout et al.: *Physiologically Based Pharmakokinetic Modeling to Investigate Regional Brain Distribution Kinetics in Rats*; The AAPS Journal, 14(3):543-553; 2012.

[39] J. Westerhout et al.: *Prediction of methotrexate CNS distribution in different species – Influence of disease conditions*; European Journal of Pharmaceutical Sciences, 57:11-24; 2014.