ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea Magistrale in Matematica

A mathematical macroscopic model for the onset and progression of Alzheimer's disease

Tesi di Laurea Magistrale in Analisi Matematica

Relatore:
Chiar.mo Prof.
Maria Carla Tesi

Presentata da:
Angela Fino

Sessione Unica
Anno Accademico 2017-2018

Contents

Prefazione

In questa tesi ci occupiamo dello studio di un modello matematico macroscopico che descrive l'insorgenza e l'evoluzione della malattia di Alzheimer (AD) nel cervello umano. Abbiamo iniziato questo lavoro di tesi riesaminando l'articolo "Well-posedness of a mathematical model for Alzheimer's disease" di Michiel Bertsch, Bruno Franchi, Maria Carla Tesi ed Andrea Tosin [6].
Il modello è basato sulla cosiddetta ipotesi di "*amyloid cascade*" (si veda [17], [23], [31] per una completa bibliografia), insieme alla "*prionoid hypothesis*" (si veda [9] e [33]). Da un punto di vista matematico, il modello consiste in un'equazione di trasporto accoppiata con un sistema non lineare di equazioni di tipo-Smoluchowski, con l'aggiunta di un termine diffusivo. In particolare, l'equazione di trasporto in un intervallo limitato coinvolge una misura di probabilità, introdotta nel modello per descrivere il *grado di* $malfunzionamento$ dei neuroni.
In vista delle caratteristiche dei fenomeni biologici di cui ci stiamo occupando, il principale segno distintivo di tale sistema è che la velocità di trasporto dipende dalla soluzione dell'equazione di Smoluchowski, che, a sua volta, contiene un termine sorgente che dipende dalla soluzione dell'equazione di trasporto: in questo modo i due gruppi di equazioni non possono essere disaccoppiati.

Lo scopo principale di questa tesi è quello di studiare la buona-posizione del modello matematico, analizzando in dettaglio un risultato di esistenza ed unicità della soluzione del problema introdotto sopra, accoppiato con opportune condizioni iniziali e al bordo [6]. La strategia utilizzata e descritta approfonditamente nel corso di questo lavoro di tesi, può essere sintetizzata come segue:

1. per trattare le misure di probabilità legate allo stato di salute del neurone, viene introdotto uno spazio metrico definito considerando lo spazio delle misure di probabilità su $[0, 1]$ dotato della distanza di Wasserstein. Si riscrive poi il sistema in termini delle caratteristiche dell'equazione di trasporto, mostrando che il nuovo problema così ottenuto è equivalente a quello originale. In particolare, si lavora tramite il trasporto di misure lungo le caratteristiche, via push-forward. Ci preme evidenziare come, sotto opportune ipotesi, le caratteristiche esistono in senso classico. La maggiore difficoltà incontrata è dovuta alla forte nonlinearità del sistema.

2. Si risolve il sistema che coinvolge le caratteristiche tramite l'uso di un argomento di contrazione che garantisce l'esistenza locale e l'unicità della soluzione. Il fatto che la distanza di Wasserstein dipenda dall'azione delle misure sulle funzioni *Lipschitz* induce una difficoltà tecnica quando si prova ad applicare un argomento iterativo per ottenere l'esistenza locale della soluzione. Tuttavia, questa difficoltà può essere superata mediante una scelta accurata degli spazi in cui utilizzare il teorema del punto fisso. In particolare, la mappa alla quale applicare tale teorema è una contrazione non sul suo dominio, ma sul suo insieme immagine.

3. La soluzione locale (rispetto al tempo t) del problema in termini delle caratteristiche viene estesa globalmente: mediante delle stime a priori, l'esistenza globale è provata per il sistema con le caratteristiche, e quindi anche per il problema originario equivalente.

Rimandiamo alla lettura dell'introduzione per una descrizione più dettagliata del modello macroscopico qui studiato, del suo background biologico e dell' organizzazione di questo lavoro di tesi.

Introduction

In this work we deal with the study of a macroscopic mathematical model which describes the onset and progression of Alzheimer's disease (AD) in the human brain. We begin this thesis by reviewing the paper by Michiel Bertsch, Bruno Franchi, Maria Carla Tesi and Andrea Tosin "Well-posedness of a mathematical model for Alzheimer's disease", see [6].
In particular, the mathematical model here studied is the result of a research carried on in the last few years by several authors on both microscopic and macroscopic mathematical models for AD, and presented in a series of papers [1], [4], [5], [13] and [15]. The macroscopic model was proposed in [4] and in [5] and it is based on the so-called *amyloid cascade hypothesis* (see [17], [23], [31] for a complete bibliography), together with the *prionoid hypothesis* (see [9] and [33]) which is meant to mirror the spreading of the disease in the neuronal net through neuron-to-neuron transmission.

From a biomedical point of view, AD is a neurodegenerative disease caused by gradual brain cell death, whose symptoms develop slowly. It is the prevalent form of late life dementia in the world today. Among the major ten causes of death, AD is the only one that cannot be prevented and cured. In 2015 it was estimate that about 44 million people suffered of AD, mostly older than 65 years. Until 2040, this number of patients is expected to double every 20 years [29]. Moreover, AD has a huge social and economic impact both on the society and on the family of the patients [7], [21], [25]. The definition of a mathematical model which tries to describe such type of disease provides another prospective apart from the classical *in vivo* and *in vitro* approaches. On the other hand, in order to clarify the mathematical choices made during modeling tries, it is necessary to take into consideration their biological background.

Below we briefly describe the main factors involved in this type of disease. According to the most recent biomedical literature on AD, focusing on neurons and their interconnections it is known that beta-amyloid (Aβ) and its toxic oligomeric isoforms Aβ_{40} and Aβ_{42} are the main causes of the cerebral damage. Amyloid is a general term for protein fragments that the body produces normally. In particular, monomeric Aβ is regularly produced by neurons at the level of the neuronal membrane by intramembranous proteolysis of APP (amyloid precursor protein) and naturally eliminated -among others- by the microglia. However in Alzheimer's disease, by unknown reasons perhaps due to external or partially genetic factors, an imbalance between produced and cleared Aβ begins to occur: less and less Aβ is taken away and in the meantime other Aβ is produced. So fragments of soluble Aβ diffuse through the brain tissue and eventually accumulate to form long, hard and insoluble fibrils of larger aggregates. These spherical deposits are known as senile plaques and they can be observed through medical imaging by a PET (positron emission tomography) scan. However, the presence of plaques does not seem related only to AD, due to the fact they have been observed in patients without any symptoms of dementia. Recently, some authors have suggested that they can be even protective to healthy neurons (see [17]). In fact, soluble Aβ seems to be the principal cause of neuronal death and eventually of dementia, since high levels of soluble Aβ_{40} and Aβ_{42} correlate much better with the presence and degree of cognitive deficits than plaque statistics. We can consider this statement as a sort of up-dated version of the amyloid cascade hypothesis. Moreover, the recent biomedical literature has proposed a *neuron-to-neuron prion-like transmission* of the disease as propagation mechanism of the neural damage in the neuronal net (see [9],[28]).

Another aspect that we would like to underline is the role of the microglia in the phagocytic process of soluble Aβ: the presence of soluble Aβ induces a general inflammation that activates the microglia which in turn secretes proinflammatory innate cytokines (see [16]). So, this neurotoxic effect is once again linked to high levels of soluble Aβ.

In order to model the different aspects of AD, as observed in [5], it is necessary to take into considerations different spatial and temporal scales. Indeed, the description of the role of a single neuron needs microscopic spatial scales, while the description of the diffusion process in the brain needs macroscopic spatial and short temporal scales. Finally, the global evolution of the disease occurs in large temporal scales. In particular, it has been observed that plaques of Aβ form extraordinarily quickly (minutes, hours), while the first signs of dementia can occur long after (years, decades) the first microscopic changes in the brain. For this reason, the model has been developed in two steps. First, in [1] and [15] these considerations have led the authors to treat the problem with a microscopic model, characterized by the use of a microscopic scale, say a multiple of the size of the soma of a single neuron (from 4 to 100 μm). Then, the interest in modeling the *degree of malfunctioning* of demaged neurons has led the authors to formulate a macroscopic model for AD. Here the brain is modelled with a macroscopic scale, considering it as a continuous medium in which neurons are seen as points of this set.

The main purpose of this thesis is to study the mathematical well-posedness of such macroscopic model, by analyzing in detail the paper [6] in all its mathematical aspects. From this point of view, the model consists of a transport equation coupled with a system of nonlinear diffusion equations (a Smoluchowski-type system with diffusion). In view of the features of the biological phenomena we are dealing with, the main hallmark of such system is that the transport velocity depends on the solution of the Smoluchowski equation, which, in turn, contains a source term that depends on the solution of the transport equation, so that the two groups of equations cannot be separated and they must be considered together. We refer the reader to [27] for a complete bibliography about the use of transport equations in mathematical models in life sciences.

Let us briefly introduce the model, referring to [4] and [5] for its complete description. At the basis of the model there is the idea to set up a flexible model which can be further developed in different directions. The main features that the model aims to describe are the following:

- agglomeration phenomena: by means of the so-called Smoluchowski equation;
- uniform diffusion: by the usual Fourier diffusion equation;
- production of Aβ, in the monomeric form: by a source term $\mathcal{F}$;
- outward flow of the cerebrospinal fluid through two inner disjoint 'holes' representing the sections of the cerebral ventricles;
- isolate our portion of brain from the environment: by a homogeneous Neumann condition on the outer boundary;
- neuron-to-neuron prion-like transmission.

At the same time, the model discards some phenomena involved in the development of AD such as:

- the tortuosity of the brain tissue;
- the role of twisted fibers of the τ-protein;
- the degradation of plaques into smaller agglomerations.

As anticipated above, different mathematical choices of spatial and temporal scales can be made when we have to describe biological phenomema. In our case the model considers a large portion of cerebral tissue, represented by an open and bounded set $\Omega \subseteq \mathbb{R}^n$ (for instance a 3-dimensional region with diameter of the order of 10cm, as suggested in [5]). As for the time scale, in order to describe in a single model both the production, diffusion, agglomeration and plaque-formation phenomena of Aβ (which occur in hours) and the evolution of the disease (which occurs in years), we have to compare two different time scales. For this reason it will be present in our equations a small parameter $\varepsilon > 0$ which links a rapid τ-scale for diffusion and agglomeration dynamics with a slow t-scale for the progression of AD. This relation can be expressed as $\Delta t = \varepsilon \Delta \tau$.

Let $x \in \Omega$ be the space variable and let $t \geq 0$ be the time variable. Moreover, let us fix $N \in \mathbb{N}$ which will represent the size of a senile plaque. To describe the behaviour of β-amyloid in Ω we consider a vector-valued function

$$u = (u_1, \dots, u_N), \quad \text{where } u_m = u_m(x,t) \quad (1 \leq m \leq N-1)$$

represents the molar concentration at the point $x \in \Omega$ and at time $t \geq 0$ of soluble Aβ-assemblies of m monomers. Therefore, we consider u_m as the concentration of Aβ assemblies of polymers of length m.
We denote by

$$u_N = u_N(x,t)$$

the molar concentration of insoluble assemblies of at least N monomers, i.e. u_N represents a cluster of oligomers of length greater or equal to N. For its own nature, u_N is slightly different from others u_m's.
As we have anticipated above, one of the main goals of the model is to describe the infection from one neuron to another inside the neuronal net. Indeed, a prion-like transmission seems to occur. In particular, the spread of AD is only in one direction, i.e. if a healthy neuron A is close to a damaged neuron B, it is possible that neuron A is infected by neuron B. On the contrary, the opposite "infection" is not possible, i.e. the damaged neuron B cannot be "recovered" by neuron A.
So, in order to describe this process we introduce another parameter $a \in [0,1]$ which models the degree of malfunctioning of a neuron. In particular $a = 0$ has the meaning "the neuron is healty", while $a = 1$ stands for "the neuron

is dead". Even if this parameter could seem artificial, the same authors underline how, instead, it has biological counterpart, since it can be compared with medical images from Fluorodeoxyglucose PET, which detect the cerebral glucose metabolism.
To measure the health state of neurons, a probability measure is introduced in the model. Given $x \in \Omega$ and $t \geq 0$, $f = f_{x,t}$ is a probability measure supported in $[0,1]$ and $df_{x,t}(a)$ denotes the fraction of neurons at point x and time t with degree of malfunctioning between a and $a + da$. At this point, the health state becomes our variable of interest. Indeed, we are interested in describing how it evolves during the progression of AD. Roughly speaking, now the idea is to fix ourselves at a point x, i.e. at a certain neuron, and try to understand how its degree of malfunctioning varies when time evolves. For this reason, the equation which involves the health state of neurons involves a divergence operator with respect to a and not to the space variable.

The progression of AD is determined by the *deterioration rate* of the health state of the neurons

$$v = v_x(a,t) \geq 0$$

through the continuity equation

$$\partial_t f + \partial_a (f v[f]) = 0.$$

Here the notation $v[f]$ is used to stress that the deterioration rate depends on f itself. We would like to underline how the deterioration rate must be nonnegative since the infection can take place in only one direction, since a demaged neuron cannot be cured by a healthier one.
Now we need a "costitutive law" for the rate v:

$$\begin{aligned}(v[f])_x(a,t) :=& \int_\Omega \left(\int_{[0,1]} \mathcal{K}(x,a,y,b) df_{y,t}(b) \right) dy \\ &+ \mathcal{S}(x,a,u_1(x,t),\ldots,u_{N-1}(x,t)).\end{aligned} \tag{1}$$

The first integral term describes the possible propagation of AD through the neural network: the degree of malfunctioning of a neuron deteriorates if its neighbours have a worse health state. For this reason, $\mathcal{K}$ is nonnegative and healthier neighboring neurons cannot improve the state of a neuron: $\mathcal{K}(x,a,y,b) \geq 0$ for all $x,y \in \Omega$ and $a,b \in [0,1]$ and $\mathcal{K}(x,a,y,b) = 0$ if $a > b$. Tipically $\mathcal{K}(x,a,y,b)$ is defined by the following expression

$$\mathcal{K}(x,a,y,b) = \mathcal{G}_x(a,b)H(x,y)$$

where $\mathcal{G}_x$ compares the healthy states, while H takes into considerations the neighboring neurons. For example, in [4], $\mathcal{G}_x$ does not depend explicity on x and has the form

$$\mathcal{G}_x(a,b) = C_{\mathcal{G}}(b-a)^+. \tag{2}$$

Here p^+ denotes the positive part of a real number, i.e. $p^+ := \max\{p, 0\}$. Concerning H, a typical choice is $H(x,y) = h(|x-y|)$, where $h(r)$ is a nonnegative and decreasing function which vanishes at some $r = r_0$ and satisfies $\int_{|y|<r_0} h(|y|)dy = 1$. In particular, in the limit $r_0 \to 0$, (1) becomes of the form

$$(v[f])_x(a,t) := \int_{[0,1]} \mathcal{G}_x(a,b)df_{x,t}(b) + \mathcal{S}(x,a,u_1(x,t),\ldots,u_{N-1}(x,t)). \tag{3}$$

From now on we consider the deterioration rate $v[f]$ as in (3).

The second term $\mathcal{S}$ in (1) is nonnegative and takes into account the dependence of the deterioration rate of the health state on high levels of toxic soluble Aβ oligomers. Notice that $\mathcal{S}$ does not depend on the plaque u_N, since we are not considering admissible a production of insoluble Aβ from the degradation of plaques. For instance, an elementary definition for $\mathcal{S}$ is the following:

$$\mathcal{S} = C_{\mathcal{S}}(1-a)\left(\sum_{m=1}^{N-1} m u_m(x,t) - \overline{U}\right)^+. \tag{4}$$

The constant $\overline{U} > 0$ can be seen as the minimal threshold value above which the amount of toxic Aβ becomes dangerous. Indeed, we recall that there is a physiologic Aβ production even in a healthy person. In particular, we are assuming that the toxicity of soluble Aβ polymers does depend on m, even if, to our best knowledge, quantitative data are only available for very short molecules. For long molecules any analytic expression different from the one given here, would be arbitrary. We would like to stress that if $\mathcal{K}$ and $\mathcal{S}$ are chosen as above, they vanish for $a = 1$ since the damaged neuron is considered dead.

Up to now, the choices made in the model describe the evolution and the propagation of AD. However, we have not considered its onset yet. In fact, the continuity equation $\partial_t f + \partial_a(fv[f]) = 0$ is conservative by its own nature. This means that a healthy person will continue to be healthy. Unfortunately, this does not necessarily happen. So, to describe the onset of AD, we modify the continuity equation by adding a *jump operator* to the right-hand side:

$$\partial_t f + \partial_a(fv[f]) = J[f], \tag{5}$$

where $J[f]$ takes into account the probability that, at a certain point of life, some neurons become ill because of still unknown reasons, maybe due to external agents or genetic factors. So, we assume that in small, randomly chosen parts of the cerebral tissue, concentrated for instance in the hippocampus, the degree of malfunctioning of neurons randomly jumps to higher values. Mathematically, $(J[f])_{x,t}$ is a signed measure defined by

$$d(J[f])_{x,t}(a) := \eta(t)\chi(x,t)\left[\left(\int_{[0,1]} P(t,b,a)df_{x,t}(b)\right)da - df_{x,t}(a)\right], \tag{6}$$

where da is the usual Lebesque measure in $[0,1]$. In addition, $P(t,b,a)$ is the probability to jump from state b to a (worst) state a. Obviously $P(t,b,a) = 0$ if $a < b$, since a damaged neuron cannot recover. For instance, we can choose $P(t,b,a)$ as:

$$P(t,b,a) \equiv P(b,a) = \begin{cases} \frac{2}{1-b} & \text{if } b \leq a \leq \frac{1}{2}(1+b), \\ 0 & \text{otherwise.} \end{cases}$$

Finally, $\eta > 0$ is the jump frequency and $\chi(x,t)$ is a characteristic function which localizes the region of the brain infected by AD.

Before writing the final system for the model, we have to define the source term which describes the production of Aβ monomers by neurons. This source term depends on the degree of malfunctioning of neurons and it increases if neurons are damaged. It can be chosen as follows:

$$\mathcal{F}[f](x,t) = C_{\mathcal{F}} \int_0^1 (\mu_0 + a)(1-a)\, df_{x,t}(a). \tag{7}$$

Notice that $(\mu_0 + a)(1-a) = \mu_0$ if $a = 0$. Indeed, the small constant $\mu_0 > 0$ is meant to mirror the physiologic Aβ production that is always present even for a healthy brain. On the other hand, $(\mu_0 + a)(1-a) = 0$ if $a = 1$. Thus, $\mathcal{F}[f]$ vanishes since there is no production if the neuron is dead. So, the final system of equations for $f, u_1, \ldots, u_N$ is:

$$\begin{cases} \partial_t f + \partial_a (f v[f]) = J[f] & \text{in } \Omega \times [0,1] \times (0,T], \\ \varepsilon \partial_t u_1 - d_1 \Delta u_1 = R_1 := & \\ \qquad - u_1 \displaystyle\sum_{j=1}^{N} a_{1,j} u_j + \mathcal{F}[f] - \sigma_1 u_1 & \text{in } Q_T = \Omega \times (0,T], \\ \varepsilon \partial_t u_m - d_m \Delta u_m = R_m := -u_m \displaystyle\sum_{j=1}^{N} a_{m,j} u_j & \\ \qquad + \dfrac{1}{2} \displaystyle\sum_{j=1}^{m-1} a_{j,m-j} u_j u_{m-j} - \sigma_m u_m & \text{in } Q_T \quad (2 \le m < N), \\ \varepsilon \partial_t u_N = \frac{1}{2} \sum_{\substack{j+k \ge N \\ k,j < N}} a_{j,k} u_j u_k & \text{in } Q_T. \end{cases} \tag{8}$$

We recall that $\varepsilon > 0$ is a small parameter which compares the two different time scales: the dynamics of Aβ occur in hours, while the evolution of AD occurs in years. The diffusion coefficients $d_m > 0$ are small when m is large, since big assemblies diffuse less. Concerning the equation for u_N, there is not the diffusion term since plaques do not move. Moreover, $a_{N,N} = 0$ since experimental data show that large oligomers do not aggregate with each other. The phagocytic activity of the microglia and other bulk clearance mechanisms are modeled by the linear term $-\sigma_m u_m$. It is reasonable that σ_m is inversely proportional to m. For this reason, we assume that it does not take place in the equation for u_N, since plaques are too big to be eaten. The quadratic term in u_m is given by the Smoluchowski equation and models the aggregation of Aβ polymers. For a detailed study of the aggregation process and the choice of coagulation rates $a_{i,j}$ we refer the reader to [1] and [15].

We assume that $\partial\Omega$ consists of smooth disjointed boundaries, $\partial\Omega_0$ and $\partial\Omega_1$, where $\partial\Omega_1$ represents the disjoint union of the boundaries of the cerebral ventricles, through which Aβ is removed by the cerebrospinal fluid by an outward flow through the choroid plexus (cf. [22] and [32]). In order to solve the system (8), we have to couple it with initial-boundary conditions:

$$\begin{cases} f_{x,0} = (f_0)_x & \text{if } x \in \Omega, \\ u_i(x,0) = u_{0i}(x) & \text{if } x \in \Omega, \quad 1 \le i \le N, \\ \partial_n u_i(x,t) = 0 & \text{if } x \in \partial\Omega_0, \quad t > 0, \quad 1 \le i < N, \\ \partial_n u_i(x,t) = -\gamma_i u_i(x,t) & \text{if } x \in \partial\Omega_1, \quad t > 0, \quad 1 \le i < N, \end{cases} \tag{9}$$

where n is the outward pointing normal on $\partial\Omega$.
We would like to stress some important aspects of this model:

- the choice of looking for a measure $f_{x,t}$, to study the degree of malfunctioning of a neuron during AD evolution, comes from the model itsel. In fact, a "healthy brain" ($a = 0$) would correspond to $f_{x,t}(a) = \delta_0(a)$, where $\delta_0(a)$ is the Dirac measure centered at the origin. Indeed, as we show in Remark 3.5, $\delta_0(a)$ is a weak solution of the continuity equation. So, if we add the jump term $J[f]$ to the equation and we consider a damaged brain, we are induced to search how the measure $\delta_0(a)$ changes.

- The system (8) is deeply coupled. Indeed, the transport equation for f contains a dependence on $u_1, \dots, u_{N-1}$ in the deterioration rate $v[f]$ through $\mathcal{S}$, whereas the diffusion-agglomeration system depends on f through the source term $\mathcal{F}$. Observe, that if $\mathcal{S} \equiv 0$ in (3), then the equation for f can be separated from the rest of the system and, in this case, may be possibly studied alone by relying on the results reported in [10] and [12]. However, the assumption $\mathcal{S} \equiv 0$ would imply a propagation of the disease due only to prionic diffusion, which is a controversial topic in the medical literature. For this reason, the authors prefer to also take into account the toxic contribution of Aβ oligomers, i.e., $\mathcal{S} \not\equiv 0$ in (3). This assumption requires studying the system (8) in its coupled form.

- All the functions present in the model have a specific qualitative clinical counterpart in routinely observable phenomena: the health state of the different brain regions (by means of a PET measuring the cerebral glucose metabolism), the amount of Aβ in the cerebral spinal fluid, and the Aβ plaques (by means of amyloid-PET scans).

Concerning the global well-posedness of the model, we deal with weak solutions of problem (8)-(9). See Chapter 3 for this definition and for the hypotheses on the data.
The main result in [6] is the following one:

Theorem 0.1. *Let $\Omega \subset \mathbb{R}^n$ be an open and bounded set with a smooth boundary $\partial\Omega$, which is the disjoint union of $\partial\Omega_0$ and $\partial\Omega_1$. Let $T > 0$ and $N \in \mathbb{N}$, and let hypotheses (H_1-H_6) be satisfied. Then problem* $(8) - (9)$ *has a unique solution in $\Omega \times [0, T]$ in the sense of Definition* 3.3.

The strategy used in [6] for proving Theorem 0.1 can be summarized as follows:

(i) to treat the measures $f_{x,t}$ we introduce a metric space $X_{[0,1]}$ by endowing the space of Borel probability measures on $[0, 1]$ with the Wasserstein distance. See Chapter 2 for its definition. We rewrite the system in terms of the characteristics of the transport equation for f, showing that the new system is equivalent to the original one. In particular, we deal with the transport of measures along the characteristics via push-forward. We stress that under the assumptions stated for the problem the characteristics exist in the classical sense. The major difficulty arises from the strong nonlinearity of the system: the transport equation depends nonlinearly on both its solution (through an integral operator) and the solution of the Smoluchowski system, which in turn depends on the solution of the transport equation.

(ii) We solve the equivalent system by the use of a contraction argument which guarantees uniqueness and local existence of the solution. In particular, the map to which to apply the fixed point theorem is a contraction not on its domain, but on its image set. As observed in [6], the fact that the Wasserstein distance can be characterized by the action of the measures on *Lipschitz* functions yields a technical difficulty when we try to apply an iteration argument in order to obtain the local existence of a solution. This difficulty can be circumvented thanks to an ad hoc formulation of the standard fixed point theorem.

(iii) The local (w.r.t. t) solution can be extended: through a priori bounds, the global existence is proved.

In the long run, we would like to extend our investigation to a model which incorporates those important aspects that the macroscopic model deliberately neglects, as the role of tau protein, the multifaceted mechanism of the microglia and disintegration processes of the plaques. This for several reasons: to make the model more realistic, to obtain optimal quantitative agreement with clinical data and also to investigate the possible formation of patterns in the distribution of the level of malfunctioning of the brain.
Specifically, this thesis is organized as follows. In Chapter 1, we introduce notations, definitions and results of measure theory, which we will use throughout this work. Next, in Chapter 2 we deal with the family of Borel probability measures on a locally compact separable metric space, introducing the notions of Narrow, Weak* and Wasserstein convergence. We prove some technical results in order to have, under certain assumptions, equivalent convergences with respect to the previous topologies. In Chapter 3, we describe the hypoteses on the data and formulate the main result of global well-posedness. Then, in Chapter 4 we reformulate the problem (8)-(9) in

terms of the characteristics for the transport equation for f and we prove the equivalence between the new problem and the original one. In Chapter 5, we exhibit the proof of Theorem 5.1, that guaratees the local existence of the solution and its uniqueness. This is proved by the following steps: by Lemma 5.3 and Lemma 5.5 we construct the map to which to apply a contraction argument. Then, Proposition 5.8 reformulates the standard point fixed theorem. Finally, in the second section of this Chapter, we prove a priori bounds which imply global existence. From Chapter 3 to Chapter 5 we strictly follow the paper [6]. In particular, we review the proofs showing all the details. We conclude the work with an Appendix. In Appendix A we present some useful tools concerning the Narrow Convergence and the Wasserstein Distance, referring to [3]. In Appendix B we list some auxiliary theorems by showing in many cases a detailed proof. The main goal of this Appendix, hopefully, is to make this thesis as self-consistent as possible.

Chapter 1

Prerequisites of Measure Theory

In this chapter we introduce some basic notions of measure theory that can be given in the abstract setting. The purpose is to recall those concepts that will be useful throughout this work. To do this, we strictly follow the organization and the notations of the the first chapter of [2]. Let us start with the notions of positive and real measures in sets equipped with a σ-algebra of subsets of a set given.

Definition 1.1. Let X be a nonempty set and let $\mathcal{E}$ be a collection of subsets of X.

(i) $\mathcal{E}$ is called an *algebra* if $\emptyset \in \mathcal{E}, E_1 \cup E_2 \in \mathcal{E}$ and $X \setminus E_1 \in \mathcal{E}$ whenever $E_1, E_2 \in \mathcal{E}$.

(ii) $\mathcal{E}$ is called a $\sigma - algebra$ if it is an algebra and for any $(E_h)_{h\in\mathbb{N}} \subset \mathcal{E}$ its union $\bigcup_{h\in\mathbb{N}} E_h$ belongs to $\mathcal{E}$.

(iii) For any collection $\mathcal{G}$ of subsets of X, the σ-algebra generated by $\mathcal{G}$ is the smallest σ-algebra containing $\mathcal{G}$. If (X, τ) is a topological space, we denote by $\mathcal{B}(X)$ the σ-algebra of Borel subsets of X, i.e., the σ-algebra generated by the open subsets of X.

(iv) If $\mathcal{E}$ is a σ-algebra in X, we call the pair $(X, \mathcal{E})$ a *measure space*.

Note that an algebra is closed not only under finite unions, but also under finite intersections by the De Morgan laws. A σ-algebra, instead, is closed under countable intersections and unions. Moreover, it is easy to check that the intersection of any family of σ-algebras is a σ-algebra, therefore the definition of generated σ-algebra is well posed. In addition, we underline that if (X, τ) is a topological space, it is closed under finite intersections of closed subsets, while the measure space $(X, \mathcal{B}(X))$ is closed under countable intersections of closed subsets. So, the concepts of *topology* and $\sigma - algebra$ are different.
Now we are ready to define the notion of positive measure on a measure space.

Definition 1.2 (**positive measure**). Let $(X, \mathcal{E})$ be a measure space and $\mu : \mathcal{E} \longrightarrow [0, +\infty]$.

(i) We say that μ is *additive* if, for $E_1, E_2 \in \mathcal{E}$, $E_1 \cap E_2 = \emptyset$, we have

$$\mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2).$$

(ii) We say that μ is $\sigma - subadditive$ if, for $E \in \mathcal{E}, (E_h)_{h \in \mathbb{N}} \subset \mathcal{E}$,

$$E \subset \bigcup_{h=0}^{+\infty} E_h \Rightarrow \mu(E) \leq \sum_{h=0}^{+\infty} \mu(E_h).$$

(iii) We say that μ is a *positive measure* if $\mu(\emptyset) = 0$ and μ is $\sigma - additive$ on $\mathcal{E}$, i.e. for any sequence $(E_h)_{h \in \mathbb{N}}$ of pairwise disjoint elements of $\mathcal{E}$

$$\mu\left(\bigcup_{h=0}^{+\infty} E_h\right) = \sum_{h=0}^{+\infty} \mu(E_h).$$

(iv) We say that μ is $finite$ if $\mu(X) < +\infty$.

(v) We say that $E \subset X$ is $\sigma - finite$ w.r.t a positive measure μ if it is the union of an increasing sequence of sets with finite measure. If X is σ-finite, we say that μ is σ-finite.

Since μ is a positive measure, $\sum_{h=0}^{+\infty} \mu(E_h)$ represents the sum of a series with positive terms.
A probability measure is a particular case of positive measure. As we will see throughout this thesis, we will work with spaces of probability measures, since in the mathematical model for Alzheimer's disease here studied, a probability measure takes into account of the degree of malfunction of a neuron.

Definition 1.3. A positive measure μ such that $\mu(X) = 1$ is called a *probability measure*.

Let us now see some simple examples of σ-algebras and positive measures.

Example 1. (i) The family of all subsets of X is the largest σ-algebra of subsets of X.

(ii) $\{\emptyset, X\}$ is the smallest σ-algebra of subsets of X.

(iii) $\mathcal{B}(\mathbb{R})$ is generated by $\mathcal{G} = \{]a, b[; a, b \in \mathbb{Q}\}$ or by $\mathcal{G}' = \{(-\infty, a[; a \in \mathbb{Q}\}$.

(iv) (**Dirac measure**) Let $(X, \mathcal{E})$ be a measure space, $x_0 \in X$, $A \in \mathcal{E}$

$$\delta_{x_0}(A) := \begin{cases} 1 & \text{if } \; x_0 \in A \\ 0 & \text{if } \; x_0 \notin A. \end{cases}$$

(v) Let $(X, \mathcal{E})$ be a measure space, $A \in \mathcal{E}$

$$\mu(A) := \begin{cases} 0 & \text{if } \; A \; \text{ is finite or countable} \\ +\infty & \text{otherwise.} \end{cases}$$

(vi) The so-called **Counting measure** on the measure space of all subsets of X is defined by

$$\mu(A) := \begin{cases} |A| & \text{if } \; A \; \text{ is finite} \\ +\infty & \text{otherwise.} \end{cases}$$

(vii) (**Discrete probability measure**) Let $(X, \mathcal{E})$ be a measure space with X at most countable. We assume that $\forall x_i \in X$, $\{x_i\} \in \mathcal{E}, i \in \mathbb{N}$ and let $p_i \in \mathbb{R}$ be the probability assigned to the point x_i such that $0 \leq p_i \leq 1$ and $\sum_{i\in\mathbb{N}} p_i = 1$. We define the probability measure μ by posing $\mu(A) := \sum_{x_i \in A} p_i$, for $A \in \mathcal{E}$.

Remark 1.4. It follows from definition that any positive measure is monotone, i.e., if $A, B \in \mathcal{E}, A \subseteq B$, then $\mu(A) \leq \mu(B)$. Moreover, if $(E_h)_{h\in\mathbb{N}}$ is an increasing sequence of sets (respectively a decreasing sequence of sets with $\mu(E_0) < +\infty$), then

$$\mu\left(\bigcup_{h=0}^{\infty} E_h\right) = \lim_{h\to+\infty} \mu(E_h), \quad \text{resp. } \mu\left(\bigcap_{h=0}^{\infty} E_h\right) = \lim_{h\to+\infty} \mu(E_h).$$

Before introducing the notion of real or signed measure, we would like to point out how, by definition, a positive measure is not necessarily finite. For this reason, according to the following definition, a positive measure, in general, is not a particular case of real measure.

Definition 1.5 (**Signed measure**)**.** Let $(X, \mathcal{E})$ be a measure space.

(i) We say that $\mu : \mathcal{E} \longrightarrow \mathbb{R}$ is a *signed measure* if $\mu(\emptyset) = 0$ and $\forall (E_h)_{h\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{E}$

$$\mu\left(\bigcup_{h=0}^{+\infty} E_h\right) = \sum_{h=0}^{+\infty} \mu(E_h).$$

(ii) If μ is a signed measure, we define its *total variation* $|\mu|$ for every $E \in \mathcal{E}$ as follows:

$$|\mu|(E) := \sup\left\{\sum_{h=0}^{+\infty} |\mu(E_h)|; E_h \in \mathcal{E} \text{ pairwise disjoint}, E = \bigcup_{h=0}^{+\infty} E_h\right\}$$

Remark 1.6. 1. Notice that in (i) it is necessary to request the absolute convergence of the series: in fact, the sum of the series could depend on the order of its terms, while the union does not.

2. A signed measure takes only finite values.

3. Some author call a measure as in Definition 1.5 a “real” measure.

Remark 1.7. A probability measure is a positive measure that is also a real measure, since it is finite. Moreover, real measures form a real vector space, since it is immediate to check from Definition 1.5 that they can be added and multiplied by real numbers.

Remark 1.8. If μ is a probability measure then $|\mu|$ coincides with μ.

Theorem 1.9 (See [2], Theorem 1.6)**.** *Let μ be a real measure on $(X, \mathcal{E})$; then $|\mu|$ is a positive finite measure.*

For the proof of this Theorem we refer the reader to [2], Theorem 1.6, pp. 4-5.
Let us continue this recall to general measure theory, by introducing further definitions of $\mu -$ *negligible* sets and *measurable* functions.

Definition 1.10 (**μ-negligible set**)**.** Let μ be a positive measure on $(X, \mathcal{E})$.

(i) A subset $N \subset X$ is called $\mu -$ *negligible* if $\exists E \in \mathcal{E}$ such that $N \subset E$ and $\mu(E) = 0$.

(ii) We say that a property $P(x)$ depending on the point $x \in X$ *holds* $\mu -$ *a.e. in* X if the set where P fails is a μ-negligible set.

(iii) Let $\mathcal{E}_\mu$ be the collection of all the subsets of X of the form $F = E \cup N$, with $E \in \mathcal{E}$ and N μ-negligible; then $\mathcal{E}_\mu$ is a σ-algebra which is called $\mu -$ *completion* of $\mathcal{E}$, and we say that $E \subset X$ is $\mu -$ *measurable* if $E \in \mathcal{E}_\mu$. The measure μ extends to $\mathcal{E}_\mu$ by setting $\mu(F) = \mu(E)$, for F as above.

Definition 1.11 (**measurable function**)**.** Let $f : X \longrightarrow Y$ be a function, with $(X, \mathcal{E})$ a measure space and (Y, d) a metric space.

(i) The function f is said to be $\mathcal{E} -$ *measurable* if $f^{-1}(A) \in \mathcal{E}$ for every open set $A \subset Y$.

(ii) If μ is a positive measure on $(X, \mathcal{E})$, the function f is said to be μ − *measurable* if it is $\mathcal{E}_\mu$−measurable.

In particular, if f is $\mathcal{E}$−measurable, then $f^{-1}(B) \in \mathcal{E}\ \forall B \in \mathcal{B}(Y)$. In addition, if (X, d') is a metric space too, and we consider $(X, \mathcal{E})$ with $\mathcal{E} = \mathcal{B}(X)$, then we say that f is a *Borel measurable* function if it is $\mathcal{B}(X)$−measurable.

Notice that if X, Y are metric spaces, any continuous function $f : X \to Y$ is Borel measurable. So, we can now restrict our discussion to measures defined on the σ-algebra of a metric space. From now on, in this chapter, (X, d) denotes a locally compact and separable metric space, l.c.s. metric space for short. In particular, we recall that a topological space is separable if it contains a countable, dense subset, while most commonly a topological space is called locally compact, if every point $x \in X$ has a compact neighbourhood, i.e., there exists an open set U and a compact set K, such that $x \in U \subseteq K$. There are other common definitions of locally compact space, but they are all equivalent if X is a Hausdorff space, but this is always true if X is a metric space, as in our case. Indeed if $x, y \in X, x \neq y$, then their distance $d(x, y) > 0$. Thus we can find two disjoint open balls centered at x and at y respectively, by taking a radius $r < \frac{d(x,y)}{2}$.
Obviously, a compact space is locally compact, but the converse is not true: for example $\mathbb{R}^n$ is locally compact but it is not compact.

Definition 1.12 (Borel and Radon measures). Let X be a l.c.s. metric space and consider the measure space $(X, \mathcal{B}(X))$ with its Borel σ-algebra.

(i) A positive measure on $(X, \mathcal{B}(X))$ is called a *Borel measure*. If μ is a Borel measure such that $\mu(K) < +\infty$, for every compact $K \subset X$, μ is called a *positive Radon measure*.

(ii) A (real) set function defined on the relatively compact (i.e. their closure is compact) Borel subsets of X which is a real measure on $(K, \mathcal{B}(K))$ for every compact $K \subset X$ is called a *(real) Radon measure* on X. If

$\mu : \mathcal{B}(X) \longrightarrow \mathbb{R}$ is a real measure, according to Definition 1.5, then we say that it is a *finite Radon measure*. We denote by $\mathcal{M}_{loc}(X)$ the space of the real Radon measures on X and by $\mathcal{M}(X)$ the space of the finite Radon measures on X.

Remark 1.13. Observe that if X is a l.c.s. metric space, then every Borel probability measure on X is a finite Radon measure. Moreover, if μ is a Radon measure and $\sup\{|\mu|(K) : K \subset X, K \text{ compact}\} < +\infty$, then it can be extended to the whole of $\mathcal{B}(X)$ obtaining a finite Radon measure, which we still denote with μ. We can extend it in this way: let us assume μ is positive, as for example a probability measure, and set $\mu(E) = \sup\{\mu(E \cap K) : K \subset X, K \text{ compact}\}, \forall E \in \mathcal{B}(X)$.

The following result allows us to approximate, under certain hypotheses, measurable sets through compact or open sets as a sort of inner or outer regularity of measures. Once again, for the proof of the following Proposition we refer the reader to the Proposition 1.43 of [2].

Proposition 1.14 (**Inner and outer regularity of measures**, see [2], Proposition 1.43)**.** *Let X be a l.c.s. metric space and μ a Borel measure on X; let E be a $\mu-$measurable set.*

(i) If μ is σ-finite, then

$$\mu(E) = \sup\{\mu(K) : K \subset E, K \text{compact}\}.$$

In this case we say that μ is inner regular.

(ii) Assume that a sequence $(X_h)_{h\in\mathbb{N}}$ of open sets in X exists such that $\mu(X_h) < +\infty \ \ \forall h$ and $X = \bigcup_{h\in\mathbb{N}} X_h$, then

$$\mu(E) = \inf\{\mu(A) : E \subset A, \quad A \ \text{ open}\}.$$

In this case we say that μ is outer regular.

If μ is inner and outer regular, we simply say that μ is regular.

Remark 1.15. We would like to point out how this result always applies to a Borel probability measure on any given l.c.s. metric space X. In fact, X is σ-finite since it is the union of an increasing sequence of concentric balls of the form $B_n(x)$ with $n \in \mathbb{N}$, $x \in X$ fixed and $\mu(B_n(x)) \leq \mu(X) = 1 < +\infty$, where $B_n(x)$ denotes the open ball of radius n centered at $x \in X$. Therefore μ is σ-finite on X, according to Definition 1.2. Hence (i) and (ii) hold.

Remark 1.16. Observe that according to Definition 1.12, a finite positive Borel measure is also a positive Radon measure, since $\forall K \subset X$ compact $\mu(K) \leq \mu(X) < +\infty$. Thus, a finite positive Borel measure is a regular finite positive Radon measure, by arguing as above. Note that this always holds for a Borel probability measure.

Finally, before moving on to the next chapter, in which we are going to introduce a distance between two probability measures, the Wasserstein distance, it will be useful to recall some operations on measures, such as push forward and restriction and the notion of the support of a measure.
Let us start with the definition of support.

Definition 1.17 (**Support of a measure**). Let μ be a positive Borel measure on the l.c.s. metric space X. We call the *support* of μ and we write suppμ, the closed set

$$\mathrm{supp}\mu := \overline{\{x \in X; \mu(U) > 0 \ \forall \ \text{neighbourhood} \ U \ \text{of} \ x\}}.$$

The operation of restriction, instead, allows us to define a new measure by restricting μ to a subset of the σ-algebra.

Definition 1.18 (**Restriction**). Let μ be a positive or real measure on the measure space $(X, \mathcal{E})$. If $E \in \mathcal{E}$, we set $\mu \llcorner E(F) = \mu(E \cap F) \ \forall F \in \mathcal{E}$.

It is clear that $\mu \llcorner E$ is a measure. In addition, we can note that if μ is a Borel (Radon) measure and E is a Borel set, then $\mu \llcorner E$ is a Borel (Radon) measure, too.

Finally, we would like to recall another important operation which will be essential for the study of the well-posedness of the mathematical model for Alzheimer's disease, presented in the Introduction of this work. We are talking about the push-forward of a measure, which allows us to map a measure on a metric space X to another one, Y. In this thesis, the push-forward's operation will be used in order to rewrite and resolve the system of the model in terms of the characteristics. Before defining it, we refer to the first chapter of [2] for the definition of the integral respect to a positive or real measure for measurable positive functions and its extension to real summable and integrable functions.

Definition 1.19 (Push-forward). Let $(X, \mathcal{E})$ and $(Y, \mathcal{F})$ be measurable spaces and let $\Phi : X \longrightarrow Y$ be such that $\Phi^{-1}(F) \in \mathcal{E}$ whenever $F \in \mathcal{F}$. For any positive or real measure μ on $(X, \mathcal{E})$ we define a measure $\Phi_\sharp\mu$ in $(Y, \mathcal{F})$ by

$$\Phi_\sharp\mu(F) := \mu(\Phi^{-1}(F)) \quad \forall F \in \mathcal{F}.$$

Equivalently, the measure $\Phi_\sharp\mu$ can be characterized by

$$\int_Y f d(\Phi_\sharp\mu) = \int_X f \circ \Phi d\mu$$

with f a real function on Y summable w.r.t $\Phi_\sharp\mu$ (then f $\circ\Phi$ is summable w.r.t μ).

Remark 1.20. Notice that the push-forward of a probability measure is a probability measure. From now on, in this work, we will take in the previous definition the measure space with the Borel σ-algebra $(X, \mathcal{B}(X))$, $\Phi : X \longrightarrow X$, such that $\Phi^{-1}(B) \in \mathcal{B}(X)$ $\forall B \in \mathcal{B}(X)$ and μ a Borel probability measure on X. So, $\Phi_\sharp\mu$ is a Borel probability measure too and we have

$$\int_X f(x) d(\Phi_\sharp\mu)(x) = \int_X f(\Phi(x)) d\mu(x)$$

for every bounded Borel measurable function f defined on X.

Chapter 2

Narrow, Weak* and Wasserstein convergence

In this chapter we introduce some tools which will be used throughout this work. In particular, we present some results from [3], referring mainly to the Chapter 5 and 7. These results, together with the definition of the Wasserstein distance between two probability measures, will allow us to define a solution for the mathematical model for Alzheimer's disease. As we have seen in the previous chapter, X denotes a locally compact separable metric space, with metric d. We denote by $\mathcal{B}(X)$ the Borel σ-algebra of the Borel subsets of X, by $\mathcal{P}(X)$ the family of all Borel probability measures on X. By the Remark 1.13 every $\mu \in \mathcal{P}(X)$ is a positive finite Radon measure. Let us start by giving the definition of Narrow and Weak* convergence, according to the probabilistic terminology.

Definition 2.1. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}(X)$ and $\mu \in \mathcal{P}(X)$. We say that

(i) $\mu_n \to \mu$ *narrowly* as $n \to +\infty$ if

$$\lim_{n\to+\infty} \int_X f(x)d\mu_n(x) = \int_X f(x)d\mu(x) \tag{2.1}$$

for every function $f \in C_b^0(X)$, the space of continuous and bounded real functions defined on X.

(ii) $\mu_n \to \mu$ *weakly** as $n \to +\infty$ if

$$\lim_{n\to+\infty} \int_X f(x)d\mu_n(x) = \int_X f(x)d\mu(x) \tag{2.2}$$

for every function $f \in C^0_c(X)$, the space of compactly supported continuous real functions defined on X.

Remark 2.2. Note that narrow convergence implies weak* convergence, since a compactly supported continuous real function is bounded on X by the generalization of Weierstrass Theorem to metric spaces (see Theorem B.1). Moreover, narrow and weak* convergences are equivalent if X is compact, since every continuous and bounded real function defined on X is compactly supported if X is compact.

Let us recall the notion of relatively compact sets: a relatively compact subset K of a topological space Y is a subset whose closure is compact. More generally when sequences may be used to test for compactness, the criterion for relative compactness becomes that any sequence in K has a subsequence convergent in Y. Now, if we consider $\mathcal{P}(X)$ with the Narrow topology, relatively compact sets can be characterized by the following useful result, which requires the definition of *tight* set of $\mathcal{P}(X)$.

Definition 2.3. We say that a set $\mathcal{K} \subset \mathcal{P}(X)$ is *tight* if $\forall \varepsilon > 0\ \exists K_\varepsilon$ compact in X such that $\mu(X \setminus K_\varepsilon) \leq \varepsilon\ \forall \mu \in \mathcal{K}$.

Remark 2.4. Notice that if the metric space (X, d) is compact, then every set of Borel probability measures on X is tight. Indeed, we can choose $K_\varepsilon = X$ in the definition above. Alternatively, this property can be obtained by Proposition A.5, as shown in Appendix.

Theorem 2.5 (**Prokhorov**, see [3], Theorem 5.1.3)**.** *If a set $\mathcal{K} \subset \mathcal{P}(X)$ is tight, then $\mathcal{K}$ is relatively compact in $\mathcal{P}(X)$. Conversely, if X is complete, then every relatively compact subset of $\mathcal{P}(X)$ is tight.*

In order to introduce the Wasserstein distance and present a very useful result which connects the Wasserstein convergence and the Narrow convergence, we have to give the definition of uniform integrability w.r.t. a given set $\mathcal{K} \subset \mathcal{P}(X)$.

Definition 2.6. (i) We say that a Borel measurable function $g : X \to [0, +\infty]$ is *uniformly integrable* w.r.t. a given set $\mathcal{K} \subset \mathcal{P}(X)$ if

$$\lim_{k \to +\infty} \int_{\{x : g(x) \geq k\}} g(x) d\mu(x) = 0$$

uniformly w.r.t. $\mu \in \mathcal{K}$, i.e.

$$\lim_{k \to +\infty} \sup_{\mu \in \mathcal{K}} \int_{\{x : g(x) \geq k\}} g(x) d\mu(x) = 0.$$

(ii) If d is a given metric for X, in the particular case of $g(x) := d(x, \bar{x})^p$, for some (and thus any) $\bar{x} \in X$ and a given $p \geq 1$, we say that the set $\mathcal{K} \subset \mathcal{P}(X)$ has *uniformly integrable* $p-$ *moments* if

$$\lim_{k \to +\infty} \int_{X \setminus B_k(\bar{x})} d(x, \bar{x})^p d\mu(x) = 0$$

uniformly w.r.t. $\mu \in \mathcal{K}$.

Since in the above definition $g : X \to [0, +\infty]$, we recall that the family of rays $]a, +\infty]$ is a neighbourhood basis for $+\infty$. Moreover, $B_k(\bar{x})$ denotes the open ball of radius k centered at $\bar{x}$.
Now, the main goal of this thesis is to prove the well-posedness of the mathematical model for Alzheimer's disease, presented in the introduction of this work. In order to do this, we need some tools which we recall from [3], without proofs. In particular, we will apply a useful lemma which provides a characterization of a sequence of measures with uniformly integrable p-moments in terms of (2.1), but extending its validity to functions unbounded but "p-growth". We say that a function $f : X \longrightarrow \mathbb{R}$ is p-growth if

$$|f(x)| \leq A + Bd(x, \bar{x})^p$$

$\forall x \in X$, for some $A, B \geq 0$ and $\bar{x} \in X$.
Notice that a bounded function is always a p-growth function.

We denote by $\mathcal{P}_p(X)$ the following subset of Borel probability measures on X:

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) : \int_X d(x, \bar{x})^p d\mu(x) < +\infty \text{ for some } \bar{x} \in X \right\}.$$

Lemma 2.7 (See [3], Lemma 5.1.7.). *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}(X)$ narrowly convergent to $\mu \in \mathcal{P}(X)$. The family $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathcal{P}(X)$ has uniformly integrable p-moments if and only if* (2.1) *holds for every continuous p-growth function $f : X \longrightarrow \mathbb{R}$.*

In the previous chapter we have defined the notion of push-forward of a measure. The tightness condition permits, under certain assumptions, to ensure the narrow convergence of the push-forward of a sequence of probability measures.

Lemma 2.8 (See [3], Lemma 5.2.1.). *Let X_1, X_2 be separable metric spaces, X_2 locally compact and let $\Phi_n : X_1 \longrightarrow X_2$ be Borel measurable functions uniformly converging to Φ on compact subsets of X_1. Let $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{P}(X_1)$ be a tight sequence narrowly converging to $\mu \in \mathcal{P}(X_1)$. If Φ is continuous, then $(\Phi_n)_\sharp\mu_n \longrightarrow \Phi_\sharp\mu$ narrowly as $n \to +\infty$.*

Proof. Let $f : X_2 \longrightarrow \mathbb{R}$ be a bounded continuous function. Let us start to prove

$$\liminf_{n\to+\infty} \int_{X_2} f d(\Phi_n)_\sharp\mu_n \geq \int_{X_2} f d(\Phi)_\sharp\mu. \tag{2.3}$$

We can assume that $f \geq 0$, eventually adding a constant to f. By hypothesis, Φ_n converges uniformly to Φ on every compact $K \subset X_1$. Thus, by Proposition B.3 , proved in Appendix, $f \circ \Phi_n$ converges uniformly to $f \circ \Phi$ on K. Therefore

$$\begin{aligned}
\liminf_{n\to+\infty} \int_{X_2} f d(\Phi_n)_\sharp\mu_n &= \liminf_{n\to+\infty} \int_{X_1} f \circ \Phi_n d\mu_n && (2.4)\\
&\geq \liminf_{n\to+\infty} \int_K f \circ \Phi_n d\mu_n && (2.5)\\
&\geq \liminf_{n\to+\infty} \int_K f \circ \Phi d\mu_n && (2.6)\\
&= \liminf_{n\to+\infty} \left(\int_{X_1} f \circ \Phi d\mu_n - \int_{X_1\setminus K} f \circ \Phi d\mu_n \right) && (2.7)
\end{aligned}$$

$$\geq \liminf_{n\to+\infty} \int_{X_1} f \circ \Phi d\mu_n + (-\sup f) \sup_n \mu_n(X_1 \setminus K) \tag{2.8}$$

$$= \int_{X_1} f \circ \Phi d\mu + (-\sup f) \sup_n \mu_n(X_1 \setminus K) \tag{2.9}$$

In particular (2.5) holds because $f \geq 0$ and $I_{X_1} \geq I_K$, where I_A with $A \in \mathcal{B}(X_1)$ denotes the indicator function such that $I_A(x) = 0$ when $x \notin A$ and $I_A(x) = 1$ when $x \in A$; (2.6) holds since $f \circ \Phi_n$ converges uniformly to $f \circ \Phi$ on K. Indeed

$$\left| \int_K f \circ \Phi_n d\mu_n - \int_K f \circ \Phi d\mu_n \right| \leq \int_K |f \circ \Phi_n - f \circ \Phi| d\mu_n \leq$$

$$\leq \sup_K |f \circ \Phi_n - f \circ \Phi| \mu_n(K) \leq \sup_K |f \circ \Phi_n - f \circ \Phi|$$

since μ_n are probability measures. Thus, letting $n \to +\infty$, the sup goes to zero for the uniform convergence and we obtain

$$\liminf_{n\to+\infty} \left(\int_K f \circ \Phi_n d\mu_n - \int_K f \circ \Phi d\mu_n \right) = 0.$$

Now, we can apply Proposition B.2

$$0 = \liminf_{n\to+\infty} \left(\int_K f \circ \Phi_n d\mu_n - \int_K f \circ \Phi d\mu_n \right) \tag{2.10}$$

$$\leq \liminf_{n\to+\infty} \left(\int_K f \circ \Phi_n d\mu_n \right) + \limsup_{n\to+\infty} \left(- \int_K f \circ \Phi d\mu_n \right) \tag{2.11}$$

$$= \liminf_{n\to+\infty} \left(\int_K f \circ \Phi_n d\mu_n \right) - \liminf_{n\to+\infty} \left(\int_K f \circ \Phi d\mu_n \right) \tag{2.12}$$

Hence we obtain (2.6)

$$\liminf_{n\to+\infty} \left(\int_K f \circ \Phi d\mu_n \right) \leq \liminf_{n\to+\infty} \left(\int_K f \circ \Phi_n d\mu_n \right).$$

Moreover, (2.7) holds because $I_K = I_{X_1} - I_{X_1 \setminus K}$; for (2.8) we have used

$$\int_{X_1 \setminus K} f \circ \Phi d\mu_n \leq \sup f \mu_n(X_1 \setminus K) \leq \sup f \sup_n \mu_n(X_1 \setminus K),$$

where $\sup f$ is a constant since f is bounded. Thus, by applying (B.3) in Proposition B.2 we obtain (2.8); finally, (2.9) holds since $f \circ \Phi$ is continuous

and bounded and μ_n converges narrowly to μ.
Now, since $\{\mu_n\}$ is tight we can find an increasing sequence of compact set K_m such that

$$\lim_m \sup_n \mu_n(X_1 \setminus K_m) = 0.$$

Putting $K = K_m$ in the inequality above and letting $m \uparrow +\infty$ we obtain (2.3). To conclude we obtain the inequality for the lim sup by replacing f to $-f$. □

To introduce the Wasserstein distance, we present first the concept of marginals of a Borel probability measure on a product space and the notion of transport plane. Let X_1 and X_2 be complete separable metric spaces and we assume for our purpose that X_1, X_2 are also locally compact. We consider X_1, X_2 as measure spaces with their respective Borel σ-algebra. Let $\pi^1 : X_1 \times X_2 \longrightarrow X_1$ and $\pi^2 : X_1 \times X_2 \longrightarrow X_2$ be the projection operators such that $\pi^i((x_1, x_2)) = x_i$ with $i \in \{1, 2\}$.
From now on, when $\boldsymbol{X} = X_1 \times ... \times X_k$ is a product space we will often use **bold** letters to indicate Borel measures $\boldsymbol{\mu} \in \mathcal{P}(\boldsymbol{X})$. First, we recall how a product space is defined; for our purposes we limits to consider a product of measure spaces with the Borel σ-algebra, but it can be defined in the same way starting from any σ-algebra.

Definition 2.9. We define the *product space* the pair

$$(X_1 \times X_2, \mathcal{B}(X_1) \times \mathcal{B}(X_2))$$

where $\mathcal{B}(X_1) \times \mathcal{B}(X_2)$ is the product σ-algebra defined as the smallest σ-algebra which makes the projection operators measurable. Equivalently, we can define the product σ-algebra as the σ-algebra generated by the sets $(\pi^i)^{-1}(A)$ with $A \in \mathcal{B}(X_i)$ and $i \in \{1, 2\}$.

In particular, if we consider $X_1 \times X_2$ as topological space with the product topology and we take now its Borel σ-algebra $\mathcal{B}(X_1 \times X_2)$, then by Theorem B.4 we have that

$$\mathcal{B}(X_1 \times X_2) = \mathcal{B}(X_1) \times \mathcal{B}(X_2),$$

since X_1, X_2 are separable metric spaces and thus they have a countable basis of open sets. In the same way, if $N \geq 2$ is an integer and $X_1, ..., X_N$ are separable metric spaces we have $\mathcal{B}(X_1 \times ... \times X_N) = \mathcal{B}(X_1) \times ... \times \mathcal{B}(X_N)$.

We introduce now the notion of transport plane between two probability measures.

Definition 2.10. Let $N \geq 2$ be an integer and $i, j = 1, ..., N$, we denote by π^i, $\pi^{i,j}$ the projection operators defined on the product space $X_1 \times ... \times X_N$, respectively defined by

$$\pi^i : (x_1, ..., x_N) \mapsto x_i \in X_i, \qquad \pi^{i,j} : (x_1, ..., x_N) \mapsto (x_i, x_j) \in X_i \times X_j.$$

(i) If $\boldsymbol{\mu} \in \mathcal{P}(X_1 \times ... \times X_N)$, the *marginals* of $\boldsymbol{\mu}$ are the probability measures

$$\mu^i := \pi^i_\sharp \boldsymbol{\mu} \in \mathcal{P}(X_i) \qquad \boldsymbol{\mu}^{ij} := \pi^{i,j}_\sharp \boldsymbol{\mu} \in \mathcal{P}(X_i \times X_j).$$

(ii) If $\mu^i \in \mathcal{P}(X_i)$, $i = 1, ..., N$, the class of *multiple plans* with marginals μ^i is defined by

$$\Gamma(\mu^1, ..., \mu^N) := \left\{ \boldsymbol{\mu} \in \mathcal{P}(X_1 \times ... \times X_N) : \pi^i_\sharp \boldsymbol{\mu} = \mu^i, \ \ i = 1, ..., N \right\}.$$

(iii) In the case $N = 2$ a measure

$$\boldsymbol{\mu} \in \Gamma(\mu^1, \mu^2) := \left\{ \boldsymbol{\mu} \in \mathcal{P}(X_1 \times X_2) : \pi^1_\sharp \boldsymbol{\mu} = \mu^1, \ \ \pi^2_\sharp \boldsymbol{\mu} = \mu^2 \right\}$$

is also called *transport plane* between μ^1 and μ^2.

Remark 2.11. The product measure $\mu^1 \times \mu^2$ (see Theorem B.5 in Appendix) is a transport plane which belongs to $\Gamma(\mu^1, \mu^2)$, since $\pi^1_\sharp(\mu^1 \times \mu^2) = \mu^1$ and $\pi^2_\sharp(\mu^1 \times \mu^2) = \mu^2$. Indeed, if $A \in \mathcal{B}(X_1)$

$$\begin{aligned} \pi^1_\sharp(\mu^1 \times \mu^2)(A) &= (\mu^1 \times \mu^2)((\pi^1)^{-1}(A)) \\ &= \mu^1 \times \mu^2(A \times X_2) \\ &= \mu^1(A)\mu^2(X_2) \\ &= \mu^1(A). \end{aligned}$$

Similarly for $\pi^2_\sharp(\mu^1 \times \mu^2) = \mu^2$.

We would like to present the Wasserstein distance by briefly contextualizing it within the so-called optimal transport problem. However, in this work, we will assume the general results on optimal transportation problems between probability measures, referring to the chapters 5 and 6 of [3] for a detailed discussion. In this way we can have the right tools to build the Alzheimer's disease model setting discussed in this thesis.
In particular, when we refer to the optimal transportation problem, we can consider a strong formulation of the problem with transport maps due to Monge and a weak formulation with transport plans due to Kantorovich. As regards the first one, let X, Y be complete and separable metric spaces, let $c : X \times Y \longrightarrow [0, +\infty]$ be a Borel cost function and let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ be given. The optimal transport problem, in Monge's formulation, is given by

$$\inf \left\{ \int_X c(x, t(x)) d\mu(x) : t_\sharp \mu = \nu \right\}$$

where $t : X \longrightarrow Y$ is a Borel map, called transport map since it carries a measure on X to a measure on Y via push-forward. However, sometimes there may not be a transport map t such that $t_\sharp \mu = \nu$. So, Monge's formulation can be ill posed. Kantorovich's formulation, instead, eludes this problem by considering

$$\min \left\{ \int_{X\times Y} c(x, y) d\gamma(x, y) : \gamma \in \Gamma(\mu, \nu) \right\}$$

since, by Remark 2.11, $\mu \times \nu \in \Gamma(\mu, \nu)$. As for the existence of an optimal transport plane, this is guaranteed when c is a lower semicontinuous function. We recall that if (X, d) is a metric space, we say that a function $c : X \to [-\infty, +\infty]$ is *lower semicontinuous* if

$$\lim_{r\to 0} \left(\inf_{y\in B_r(x)\setminus\{x\}} c(y) \right) \geq c(x), \qquad \forall x \in X.$$

This fact is proved in Chapter 5 and 6 of [3]: the idea is to use a lower semicontinuity property for semicontinuous functions bounded from below and the fact that $\Gamma(\mu,\nu)$ is tight (this property is equivalent to the tightness of μ,ν, a property always guaranteed in a complete separable metric space). In particular, if we consider as cost function $c(x,y)=d(x,y)^p$, with $p\geq 1$, there is an optimal transport plane, since the metric d is Lipschitz-continuous, thus continuous, which implies that d^p is also continuous, and therefore it is lower semicontinuous. Finally, this carries us to the Wasserstein distance.

Definition 2.12. Let X be a complete separable locally compact metric space and $p\geq 1$. The p-th *Wasserstein distance* between $\mu^1,\mu^2\in\mathcal{P}_p(X)$ is the number denoted by $\mathcal{W}_p(\mu^1,\mu^2)$ and defined by

$$\mathcal{W}_p^p(\mu^1,\mu^2):=\inf\left\{\int_{X\times X} d(x_1,x_2)^p d\boldsymbol{\mu}(x_1,x_2):\boldsymbol{\mu}\in\Gamma(\mu^1,\mu^2)\right\},$$

where $\Gamma(\mu^1,\mu^2)\subset\mathcal{P}(X^2)$ is the set of all transport planes between μ^1 and μ^2.

The function defined above is indeed a distance. We can prove the triangle inequality using Remark A.2. Indeed, if $\mu^1,\mu^2,\mu^3\in\mathcal{P}_p(X)$ and γ^{12} is optimal between μ^1 and μ^2, γ^{23} is optimal between μ^2 and μ^3, then by Remark A.2 we can find $\gamma\in\mathcal{P}(X\times X\times X)$ such that $\pi^{1,2}_\sharp\gamma=\gamma^{12}$, $\pi^{2,3}_\sharp\gamma=\gamma^{23}$ and such that if we take $\gamma^{13}:=\pi^{13}_\sharp\gamma$, γ^{13} belongs to $\Gamma(\mu^1,\mu^3)$. Now, since γ^{12} is optimal between μ^1 and μ^2 and $\pi^{1,2}_\sharp\gamma=\gamma^{12}$, we have

$$\begin{aligned}\mathcal{W}_p^p(\mu^1,\mu^2)&=\int_{X^2} d(x_1,x_2)^p d\gamma^{12}(x_1,x_2)\\&=\int_{X^2} d(x_1,x_2)^p d(\pi^{1,2}_\sharp\gamma)(x_1,x_2)\\&=\int_{X^3} d(\pi^{1,2}(x_1,x_2,x_3))^p d\gamma(x_1,x_2,x_3)\\&=\int_{X^3} d(x_1,x_2)^p d\gamma(x_1,x_2,x_3).\end{aligned}$$

If now we start from γ^{23}, that is optimal between μ^2 and μ^3, and we use $\pi^{2,3}_\sharp\gamma=\gamma^{23}$, then in similar way we can get

$$\mathcal{W}_p^p(\mu^2,\mu^3)=\int_{X^3} d(x_2,x_3)^p d\gamma(x_1,x_2,x_3).$$

Now, since $\gamma^{13} := \pi^{13}_{\sharp}\gamma \in \Gamma(\mu^1, \mu^3)$, we have

$$\begin{aligned}
\mathcal{W}_p^p(\mu^1, \mu^3) &= \inf \left\{ \int_{X^2} d(x_1, x_3)^p d\boldsymbol{\mu}(x_1, x_3) : \boldsymbol{\mu} \in \Gamma(\mu^1, \mu^3) \right\} \\
&\leq \int_{X^2} d(x_1, x_3)^p d\gamma^{13}(x_1, x_3) \\
&= \int_{X^2} d(x_1, x_3)^p d(\pi^{13}_{\sharp}\gamma)(x_1, x_3) \\
&= \int_{X^3} d(\pi^{13}(x_1, x_2, x_3))^p d\gamma(x_1, x_2, x_3) \\
&= \int_{X^3} d(x_1, x_3)^p d\gamma(x_1, x_2, x_3).
\end{aligned}$$

The standard triangle inequality of the L^p distance allows us to conclude the proof, where the L^p distance is defined as follows: we consider
$L^p(\mu; Y) := \big\{ r : X \to Y \mu - \text{measurable} : \int_X d_Y(r(x), \bar{y})^p d\mu(x) < +\infty$
for some (and thus any) $\bar{y} \in Y \big\}$,
with the distance

$$\boldsymbol{d}(r, s)_{L^p(\mu;Y)} := \left(\int_X d_Y(r(x), s(x))^p d\mu(x) \right)^{\frac{1}{p}}.$$

Thus, for the previous computations

$$\begin{aligned}
\mathcal{W}_p(\mu^1, \mu^3) &\leq \left(\int_{X^3} d(x_1, x_3)^p d\gamma(x_1, x_2, x_3) \right)^{\frac{1}{p}} \\
&= \boldsymbol{d}(x_1, x_3)_{L^p(\gamma;X)} \\
&\leq \boldsymbol{d}(x_1, x_2)_{L^p(\gamma;X)} + \boldsymbol{d}(x_2, x_3)_{L^p(\gamma;X)} \\
&= \mathcal{W}_p(\mu^1, \mu^2) + \mathcal{W}_p(\mu^2, \mu^3).
\end{aligned}$$

The following result guarantees the completeness of the Wasserstein metric so that we can work with a complete metric space of Borel probability measures with finite p-moments.

Proposition 2.13 (See Proposition A.4)**.** *Let X be a complete separable locally compact metric space. If $\mu \in \mathcal{P}(X)$ has compact support then for any $\bar{x} \in X$ and $p \geq 1$*

$$\int_X d(x, \bar{x})^p d\mu(x) < +\infty.$$

In particular, μ has finite p-moment, i.e. $\mu \in \mathcal{P}_p(X)$. Endowed with the p-th Wasserstein distance $\mathcal{W}_p$, $\mathcal{P}_p(X)$ is a complete metric space.

Now, as we have anticipated before, the main goal of this chapter is obtaining some result that connects the Wasserstein convergence and the narrow or equivalently the weak* convergence. To do this, we need to know that the limit of a weakly* converging sequence, all supported within a fixed compact, which does not depend by the terms of the sequence, is itself still supported within that compact.

Proposition 2.14. *Let X be a locally compact separable metric space. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_p(X)$ such that $\mu_n \longrightarrow \mu$ as $n \to +\infty$ weakly*. Suppose there exist a compact set K such that* $\mathrm{supp}\mu_n \subset K$ $\forall n \in \mathbb{N}$ *and an open set O satisfying*

$$K \subset O \quad \textit{and} \quad X \setminus O \neq \emptyset.$$

Then $\mathrm{supp}\mu \subset K$.

Before giving the proof, let us introduce some notations: in a metric space (X, d), the distance of a point $x \in X$ from a set $K \subseteq X$ is defined by

$$d(x, K) := \inf_{y \in K} d(x, y).$$

Moreover, the distance between two sets $A, B \subseteq X$ is defined by

$$d(A, B) := \inf_{x \in A, y \in B} d(x, y).$$

Proof. Let H be any compact set such that $H \subset X \setminus K$, hence $d(H, K) > 0$. If we consider now

$$H_n := \overline{\left\{ x \in X; d(x, H) < \frac{1}{2n} \right\}}$$

with $n \in \mathbb{N}$, then for n sufficiently large, i.e. $n \geq n_1$, we have

$$H_n \cap K = \emptyset. \tag{2.13}$$

Moreover, $H \subset X \setminus K$ where $X \setminus K$ is open, since K is compact and thus closed, and $X \setminus (X \setminus K) = K \neq \emptyset$. So, by Proposition B.9 for $n \geq n_2$, H_n is compact. Let's set

$$n \geq \max(n_1, n_2) \tag{2.14}$$

and let $\phi_n : X \longrightarrow \mathbb{R}$ be a continuous function,

$$\phi_n(x) := \left(1 - \frac{d(x, H)}{\frac{1}{2n}}\right)^+,$$

where $p^+ := \max(p, 0)$. By definition, $\phi_n \equiv 1$ on H and $\mathrm{supp}\phi_n \subset H_n$. Thus $\mathrm{supp}\phi_n$ is compact because it is closed and contained in a compact set. Now, by (2.14) and (2.13) $\int_X \phi_n d\mu_j = 0$ (remember that $\mathrm{supp}\mu_j \subset K \ \ \forall j \in \mathbb{N}$). Hence, since $\mu_j \longrightarrow \mu$ weakly*, i.e.

$$0 = \int_X \phi_n d\mu_j \longrightarrow \int_X \phi_n d\mu,$$

as $j \to +\infty$, we have

$$\int_X \phi_n d\mu = 0.$$

Moreover, for any point x fixed we have $\phi_n(x) \longrightarrow I_H(x)$ as $n \to +\infty$, where I_H denotes the indicator function on H. Indeed, if $x \in H$, $\phi_n(x) = 1 \longrightarrow 1 = I_H(x)$ as $n \to +\infty$, while if $x \in X \setminus H$, $\phi_n(x) \longrightarrow 0$ as $n \to +\infty$, by definition of ϕ_n. In addition, $|\phi_n| \leq 1 \ \forall n \in \mathbb{N}$ and, since μ is a probability measure,

$$1 \in L^1(X; \mu) := \left\{ f : X \longrightarrow \overline{\mathbb{R}} \ \ \mu - \text{measurable}; \int_X |f| d\mu < +\infty \right\}.$$

Thus, by the dominated convergence Theorem B.10

$$0 = \int_X \phi_n d\mu \longrightarrow \int_X I_H d\mu = \mu(H) \text{ as } n \to +\infty.$$

Hence $\mu(H) = 0 \ \forall H$ compact, $H \subset X \setminus K$.

Now, $\forall x \in X \setminus K$ there exists an open ball $B(x, r) \subset X \setminus K$, since $X \setminus K$ is open. So, if we show that $\mu(B(x, r)) = 0$, then by definition of support we can conclude that $\mathrm{supp}\mu \subset K$. Indeed, by Remark 1.15 μ is inner regular.

$$\begin{aligned} \mu(B(x, r)) &= \sup\{\mu(H), \ \ H \ \text{ compact}, \ \ H \subset B(x, r)\} \\ &= 0 \end{aligned}$$

since $H \subset B(x,r) \subset X \setminus K$, and we have already shown that $\mu(H) = 0$, $\forall H$ compact, $H \subset X \setminus K$. $\square$

Another useful result, which we will use later, connects the support of a measure with that of its push-forward.

Proposition 2.15. *Let X, Y be locally compact separable metric spaces. Let $\mu \in \mathcal{P}(X)$ and $r : X \longrightarrow Y$ be an open continuous injective function, then*

$$r(\text{supp}\mu) \subseteq \text{supp}(r_\sharp\mu) = \overline{r(\text{supp}\mu)}.$$

In particular, if r is also a closed function

$$\text{supp}(r_\sharp\mu) = r(\text{supp}\mu).$$

Proof. Let us start to prove that $\text{supp}(r_\sharp\mu) \subseteq \overline{r(\text{supp}\mu)}$. We recall that, by definition,

$$\text{supp}(r_\sharp\mu) := \overline{\{y \in Y; (r_\sharp\mu)(U) > 0, \ \ \forall\text{neighborhood } \ U \ \text{ of } \ y\}}.$$

Since $\overline{r(\text{supp}\mu)}$ is closed and contains $r(\text{supp}\mu)$, we have but to prove that

$$\{y \in Y; (r_\sharp\mu)(U) > 0, \ \ \forall\text{neighborhood } \ U \ \text{ of } \ y\} \subseteq r(\text{supp}\mu).$$

Hence we could conclude that $\overline{r(\text{supp}\mu)}$ contains the closure of the previous set, i.e. $\text{supp}(r_\sharp\mu) \subseteq \overline{r(\text{supp}\mu)}$.
Now, let $\bar{y} \in \{y \in Y; (r_\sharp\mu)(U) > 0, \ \forall\text{neighborhood } \ U \ \text{ of } \ y\}$. We can say that any neighborhood V of $r^{-1}(\bar{y})$ can be represented as $V = r^{-1}(W)$ with W a neighborhood of $\bar{y}$. Indeed, since r is injective, $V = r^{-1}(r(V))$. Hence we can set $W := r(V)$ that is a neighborhood of $\bar{y}$ since r is open. Therefore, for any neighborhood V of $r^{-1}(\bar{y})$, we have

$$\begin{aligned} \mu(V) &= \mu(r^{-1}(W)) \\ &= (r_\sharp\mu)(W) > 0 \end{aligned}$$

since W is a neighborhood of $\bar{y}$ and

$$\bar{y} \in \{y \in Y; (r_\sharp\mu)(U) > 0, \ \ \forall\text{neighborhood } \ U \ \text{ of } \ y\}.$$

This means that $r^{-1}(\bar{y}) \in \text{supp}\mu$, i.e. $\bar{y} \in r(\text{supp}\mu)$.
Let's show now the opposite inclusion $\overline{r(\text{supp}\mu)} \subseteq \text{supp}(r_\sharp\mu)$. For definition of closure, since the support of a measure is a closed set, it is sufficient to show that $r(\text{supp}\mu) \subseteq \text{supp}(r_\sharp\mu)$. Let $y \in r(\text{supp}\mu)$ and U be any neighborhood of y. Since $y \in r(\text{supp}\mu)$, there exists $x \in \text{supp}\mu$ such that $r(x) = y$. Moreover, r is continuous, thus $r^{-1}(U)$ is a neighborhood of x. Hence

$$(r_\sharp\mu)(U) = \mu(r^{-1}(U)) > 0$$

since $x \in \text{supp}\mu$. This holds for any U neighborhood of y, i.e. $y \in \text{supp}(r_\sharp\mu)$. □

Remark 2.16. Notice that it is sufficient the continuity of r to have

$$r(\text{supp}\mu) \subseteq \text{supp}(r_\sharp\mu).$$

We have now all the necessary tools to investigate the link between the the Wasserstein convergence of a sequence of Borel probability measures and the weak* convergence of the same sequence.

Proposition 2.17. *Let X be a complete locally compact separable metric space. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_p(X)$. We have*

*(i) if $\mathcal{W}_p(\mu_n, \mu) \to 0$ as $n \to +\infty$, then $\mu_n \to \mu$ as $n \to +\infty$ weakly**;*

(ii) suppose there exists a compact set K such that $\text{supp}\mu_n \subset K$ *for all $n \in \mathbb{N}$ and an open set $\mathcal{O}$ satisfying*

$$K \subset \mathcal{O} \ \ \text{and} \ \ X \setminus \mathcal{O} \neq \emptyset.$$

Then

$$\mathcal{W}_p(\mu_n, \mu) \to 0 \ \ \text{as} \ \ n \to +\infty$$

if and only if $\mu_n \to \mu$ as $n \to +\infty$ weakly or, equivalently, narrowly.*

Proof. As for (i), we apply Proposition A.4 and we obtain $\mu_n \to \mu$ narrowly and by Remark 2.2 also weakly*. As for (ii), let us start to prove that if the support of every term of the sequence is contained in a fixed compact K satisfying our assumptions, then, the weak* convergence of the sequence implies the narrow one. Let $f : X \longrightarrow \mathbb{R}$ be any bounded continuous real function. We can easily construct a continuous function $\tilde{f}$ such that

$$\text{supp}\tilde{f} \subset \mathcal{O} \quad \text{and} \quad \tilde{f} \equiv f \quad \text{in} \quad K. \tag{2.15}$$

For example, we can take $\tilde{f}$ as follows: first, let $\varepsilon > 0$ and K_ε compact be given by Proposition B.9; we take now

$$\tilde{f}(x) = f(x)\left(1 - \frac{d(x,K)}{\varepsilon}\right)^+$$

which is a continuous function satisfying (2.15) with compact support, since $\text{supp}\tilde{f} \subset K_\varepsilon$ and $K_\varepsilon \subset \mathcal{O}$.
Thus

$$\begin{aligned}\lim_{n\to+\infty}\int_X f(x)d\mu_n(x) &= \lim_{n\to+\infty}\int_K f(x)d\mu_n(x) = \lim_{n\to+\infty}\int_K \tilde{f}(x)d\mu_n(x)\\ &= \lim_{n\to+\infty}\int_X \tilde{f}(x)d\mu_n(x) = \int_X \tilde{f}(x)d\mu(x) = \int_K \tilde{f}(x)d\mu(x) \qquad (2.16)\\ &= \int_K f(x)d\mu(x) = \int_X f(x)d\mu(x)\end{aligned}$$

where we have used repeatedly $\text{supp}\mu_n$, $\text{supp}\mu \subset K$ (by Proposition 2.14) and weak* convergence with $\tilde{f}$. So, we have proved weak* convergence $\Rightarrow$ narrow convergence. Now, in order to have $\lim\limits_{n\to+\infty}\mathcal{W}_p(\mu_n,\mu) = 0$, we apply again Proposition A.4. To do this, we have but to prove that μ_n's have uniformly integrable p-moments. By Lemma 2.7 the assertion will follow by showing that

$$\lim_{n\to+\infty}\int_X f(x)d\mu_n(x) = \int_X f(x)d\mu(x) \tag{2.17}$$

for any continuous real p-growth function f.
Take now such p-growth function f and we construct $\tilde{f}$ as above. Now, by arguing exactly as in (2.16), we obtain (2.17) for any p-growth function f. Hence (μ_n) has uniformly integrable p-moments. Thus by Proposition A.4 we have

$$\lim_{n\to+\infty}\mathcal{W}_p(\mu_n,\mu) = 0.$$

□

Remark 2.18. If X is compact, then the assertion (ii) is trivial. Indeed, we have already pointed out that narrow convergence and weak* convergence are equivalent on compact metric spaces. Thus we can apply Proposition A.4. Indeed the μ_n's have uniformly integrable p-moments, by Lemma 2.7. However, since X is compact we don't need to construct $\tilde{f}$ because any continuous p-growth function on X is bounded. Therefore (2.17) holds due to the narrow convergence hypothesis.

Remark 2.19. Throughout this thesis, we will work with Borel probability measures μ on $[0,1]$. So, in order to apply the previous proposition and thus to find a compact K satisfying our assumptions, we will often consider μ as the restriction on $[0,1]$ of a Borel probability measure defined on $\mathbb{R}$ and supported in $[0,1]$, which we will still denote with μ to avoid cumbersome notations. In this way we can always find an open set $\mathcal{O}$ in $\mathbb{R}$, such that $[0,1] \subset \mathcal{O}$ and $\mathbb{R} \setminus \mathcal{O} \neq \emptyset$.

Proposition 2.20. *Let X, Y be complete separable metric spaces. In addition, let X be compact and Y be locally compact. Let assume that for any compact set $K \subset Y$ there exists an open set $\mathcal{O}$ such that*

$$K \subset \mathcal{O} \qquad \textit{and} \qquad Y \setminus \mathcal{O} \neq \emptyset.$$

Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{P}_p(X)$. Let $\phi_n : X \to Y$ be a sequence of continuous injective (and hence open) maps that converges uniformly to a continuous injective (and hence open) map $\phi : X \to Y$. Then, if $p \geq 1$

$$\lim_{n\to+\infty} \mathcal{W}_p(\mu_n, \mu) = 0 \qquad \Longleftrightarrow \qquad \lim_{n\to+\infty} \mathcal{W}_p((\phi_n)_\sharp\mu_n, \phi_\sharp\mu) = 0.$$

Before giving the proof, we would like to point out how the assertion: "continuous injective (and hence open) map $\phi : X \to Y$ " holds by Theorem B.13 since X is compact.

Proof. Let us start to prove $\Rightarrow$).
By Remark 2.4, the sequence $(\mu_n)_{n\in\mathbb{N}}$ is tight. Thus by Proposition 2.17,

point (i), by Remark 2.2 and by Lemma 2.8

$$\lim_{n\to+\infty} \mathcal{W}_p(\mu_n, \mu) = 0 \Rightarrow (\phi_n)_\sharp \mu_n \to \phi_\sharp \mu$$

narrowly as $n \to +\infty$. Now, if $\phi(X) = Y$, then Y is compact, since X is compact and ϕ is continuous. Therefore, by Remark 2.18, we obtain $\lim_{n\to+\infty} \mathcal{W}_p((\phi_n)_\sharp \mu_n, \phi_\sharp \mu) = 0$. Otherwise, set

$$K_0 := \overline{\{y \in Y; d(y, \phi(X)) < \varepsilon\}}.$$

with $\varepsilon > 0$ given by Proposition B.9 and thus K_0 is compact. Moreover, by Theorem B.13, ϕ_n is a closed function $\forall n \in \mathbb{N}$. Thus, by Proposition 2.15

$$\text{supp}((\phi_n)_\sharp \mu_n) = \phi_n(\text{supp}\mu_n) \subset \phi_n(X).$$

Now, since ϕ_n converges uniformly to ϕ, if $n > \bar{n}$, then $\phi_n(X) \subset K_0$. Hence

$$\text{supp}((\phi_n)_\sharp \mu_n) \subset K_0$$

that is compact. By assumption, there is an open set $\mathcal{O}_0$ such that

$$K_0 \subset \mathcal{O}_0 \quad \text{and} \quad Y \setminus \mathcal{O}_0 \neq \emptyset.$$

Thus by Proposition 2.17, part (ii), $\lim_{n\to+\infty} \mathcal{W}_p((\phi_n)_\sharp \mu_n, \phi_\sharp \mu) = 0$. This proves the first part of the statement.
Let's prove $\Leftarrow$) and suppose now $\lim_{n\to+\infty} \mathcal{W}_p((\phi_n)_\sharp \mu_n, \phi_\sharp \mu) = 0$.
We notice now that the sequence $(\mu_n)_{n\in\mathbb{N}}$ is tight (again by Remark 2.4), and hence by Theorem 2.5, it is relatively compact w.r.t. the narrow convergence. Therefore, there exists a subsequence $(\mu_{n_j})_{j\in\mathbb{N}}$ converging narrowly to $\nu \in \mathcal{P}(X)$. By Proposition 2.17 (or better by Remark 2.18), $\mathcal{W}_p(\mu_{n_j}, \nu) \to 0$ as $j \to +\infty$, and then $\mathcal{W}_p((\phi_{n_j})_\sharp \mu_{n_j}, \phi_\sharp \nu) \to 0$ as $j \to +\infty$ (by the first part of the present proposition). Thus the uniqueness of the Wasserstein limit yields $\phi_\sharp \nu = \phi_\sharp \mu$ and eventually $\nu = \mu$. Indeed, let $B \subset X$ be an open set (and hence $\phi(B)$ is open in Y), since ϕ is injective, we have

$$\nu(B) = \nu(\phi^{-1}(\phi(B))) = (\phi_\sharp \nu)(\phi(B)) = (\phi_\sharp \mu)(\phi(B)) = \mu(\phi^{-1}(\phi(B))) = \mu(B).$$

By Coincidence criterion (see Proposition B.6), if ν and μ coincide on open sets, then $\nu = \mu$ on $\mathcal{B}(X)$.
Therefore $\lim_{j\to+\infty} \mathcal{W}_p(\mu_{n_j}, \mu) = 0$.
A standard argument in metric spaces makes it possible to recover the limit for the full sequence $(\mu_n)_{n\in\mathbb{N}}$. Indeed, let suppose by contradiction that $\mathcal{W}_p(\mu_n, \mu) \nrightarrow 0$ as $n \to +\infty$. Then, $\exists \varepsilon > 0$ and $\exists$ a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ such that

$$\mathcal{W}_p(\mu_{n_k}, \mu) \geq \varepsilon \quad \forall k \in \mathbb{N}. \tag{2.18}$$

By arguing as above with $(\mu_n)_{n\in\mathbb{N}}$ replaced by $(\mu_{n_k})_{k\in\mathbb{N}}$, we find a subsequence $(\mu_{n_{k_j}})_{j\in\mathbb{N}}$ of $(\mu_{n_k})_{k\in\mathbb{N}}$ such that $\mathcal{W}_p(\mu_{n_{k_j}}, \mu) \to 0$ as $j \to +\infty$, i.e.

$$\forall \tilde{\varepsilon} > 0 \;\; \exists j_{\tilde{\varepsilon}} \in \mathbb{N} \;\text{ such that }\; \forall j > j_{\tilde{\varepsilon}}, \;\; \mathcal{W}_p(\mu_{n_{k_j}}, \mu) < \tilde{\varepsilon}.$$

But now, if we take $\tilde{\varepsilon} = \varepsilon$ we have $\mathcal{W}_p(\mu_{n_{k_j}}, \mu) < \varepsilon \; \forall j > j_\varepsilon$ and, at the same time, $\mathcal{W}_p(\mu_{n_{k_j}}, \mu) \geq \varepsilon$ (because this holds for all terms of $(\mu_{n_k})_{k\in\mathbb{N}}$).
So we got a contradiction. This would conclude the proof.
However for greater completeness, we show how to construct the subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$. We are assuming that $\mathcal{W}_p(\mu_n, \mu) \nrightarrow 0$. This means that

$$\exists \varepsilon > 0 \;\text{ such that }\; \forall k \in \mathbb{N} \;\; \exists p_k > k \;\text{ such that }\; \mathcal{W}_p(\mu_{p_k}, \mu) \geq \varepsilon.$$

We choose $k = 1 \Rightarrow \mathcal{W}_p(\mu_{p_1}, \mu) \geq \varepsilon$;
we choose $k = p_1 \Rightarrow \mathcal{W}_p(\mu_{p_2}, \mu) \geq \varepsilon$, $p_2 > p_1$;
we choose $k = p_2 \Rightarrow \mathcal{W}_p(\mu_{p_3}, \mu) \geq \varepsilon$, $p_3 > p_2$;
and so on. Thus $(p_k)_{k\in\mathbb{N}}$ is an increasing sequence of indeces, therefore it gives a subsequence $(\mu_{p_k})_{k\in\mathbb{N}}$ such that $\mathcal{W}_p(\mu_{p_k}, \mu) \geq \varepsilon \; \forall k \in \mathbb{N}$. □

Corollary 2.21. *Let $(X, d_1), (Y, d_2)$ be complete separable metric spaces satisfying the assumption of Proposition* 2.20. *If $I \subset \mathbb{R}$ is a compact interval, let $\phi : X \times I \to Y$ be a continuous map such that for any $t \in I$ the map $x \to \phi(x, t)$ is injective and open.*
If $t \in I$, let $\mu_t \in \mathcal{P}(X)$.
Then $t \mapsto \mu_t$ is continuous (w.r.t. the Wasserstein topology) if and only if $t \mapsto \phi(\cdot, t)_\sharp \mu_t$ is continuous (w.r.t. the Wasserstein topology).

Proof. We show the continuity of the function $t \mapsto \mu_t$ (w.r.t. the Wasserstein topology) using the well know approach by sequences.
Let $t_0 \in I$, $(t_n)_{n\in\mathbb{N}} \subset I$, $t_n \neq t_0$ $\forall n \in \mathbb{N}$ be any sequence such that $t_n \to t_0$ as $n \to +\infty$. Since X and I are compact and ϕ is continuous, then by Heine Cantor Theorem ϕ is uniformly continuous, i.e.

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \text{such that} \;\; \forall (x,t),(y,s) \in X \times I, \;\; d_1(x,y) + |t-s| < \delta$$

$$d_2(\phi(x,t), \phi(y,s)) < \varepsilon.$$

Let consider now $d_2(\phi(x,t_n), \phi(x,t_0))$. By uniform continuity if $d_1(x,x) + |t_n - t_0| = |t_n - t_0| < \delta$, then $d_2(\phi(x,t_n), \phi(x,t_0)) < \varepsilon$. For sure $|t_n - t_0| < \delta$ since $t_n \to t_0$ as $n \to +\infty$.
Thus $\forall \tilde{\varepsilon} > 0$ $\exists n_{\tilde{\varepsilon}} \in \mathbb{N}$ such that $\forall n \geq n_{\tilde{\varepsilon}}$, $|t_n - t_0| < \tilde{\varepsilon}$. Therefore, if we take $\tilde{\varepsilon} = \delta$, there exists $n_\delta \in \mathbb{N}$ such that $|t_n - t_0| < \delta$ $\forall n \geq n_\delta$. Hence by uniform continuity of ϕ, we get

$$d_2(\phi(x,t_n), \phi(x,t_0)) < \varepsilon$$

$\forall n \geq n_\delta, \forall x \in X$, that is the definition of uniform convergence of the function $\phi(\cdot, t_n)$ to the function $\phi(\cdot, t_0)$. Thus by Proposition 2.20

$$\lim_{n\to+\infty} \mathcal{W}_p(\mu_{t_n}, \mu_{t_0}) = 0 \;\; \text{iff} \; \lim_{n\to+\infty} \mathcal{W}_p(\phi(\cdot, t_n)_\sharp \mu_{t_n}, \phi(\cdot, t_0)_\sharp \mu_{t_0}) = 0.$$

□

Chapter 3

Problem Statement and Main Result

In the previous chapters we have introduced some useful tools that we will use to investigate the mathematical well-posedness of the macroscopic model for AD presented in the Introduction of this work. In this chapter we state the problem studied in this thesis by describing the hypotheses on the data and giving the definition of an its solution. Finally, we introduce the main result of existence and uniqueness of the solution, whose proof will be provided in the following chapters.

Throughout this thesis we set $T > 0$, $N \in \mathbb{N}$, while $\Omega \subset \mathbb{R}^n$ is an open and bounded set with a smooth boundary $\partial\Omega$, which is the disjoint union of $\partial\Omega_0$ and $\partial\Omega_1$.
To treat the measures $f_{x,t}$ that take into account the degree of malfunctioning of neurons, we introduce a metric space $X_{[0,1]}$.

Definition 3.1. The space $\mathcal{P}([0,1])$ of Borel probability measures on $[0,1]$ endowed with the Wasserstein distance $\mathcal{W}_1$ is denoted by $X_{[0,1]}$, i.e.

$$X_{[0,1]} := (\mathcal{P}([0,1]), \mathcal{W}_1).$$

We refer to Chapter 2 for the definition of the Wasserstein distances $\mathcal{W}_p$. By Proposition A.4, $X_{[0,1]}$ is a complete separable metric space since $[0,1]$ is complete respect to the induced Euclidean metric because it is a closed set in $\mathbb{R}$, which is complete. Moreover, by Proposition 2.17 and by Remark 2.19, a sequence $(\mu_n)_{n\in\mathbb{N}}$ converges in $X_{[0,1]}$ if and only if it converges narrowly or weakly*.

We denote by $C([0,T];X_{[0,1]})$ the space of continuous functions from the interval $[0,T]$ to $X_{[0,1]}$, i.e.

$$C([0,T];X_{[0,1]}) = \left\{t \mapsto \mu_t \in X_{[0,1]}; \mathcal{W}_1(\mu_t,\mu_{t_0}) \to 0 \text{ as } t \to t_0, t_0 \in [0,T]\right\}.$$

It is easy to see that, endowed with the distance

$$\max_{0\leq t\leq T} \mathcal{W}_1((\mu_1)_t,(\mu_2)_t),$$

$C([0,T];X_{[0,1]})$ is a complete metric space.

3.1 Hypotheses on the data

Let us start to recall some notations. Below we denote by ∂_a, ∇_u, etc., distributional derivatives, while C denotes a generic constant. Moreover, we recall the following spaces

(i) $L^\infty(\Omega; C([0,1]\times[0,+\infty)^{N-1})) = \Big\{$measurable functions on Ω such that $x \mapsto \mathcal{S}(x,a,u_1,\ldots,u_{N-1}) = \mathcal{S}_x(a,u_1,\ldots,u_{N-1}) \in C([0,1]\times[0,+\infty)^{N-1})$ for a.e. $x \in \Omega$ and $\sup_{x\in\Omega}\big(\sup_{[0,1]\times[0,+\infty)^{N-1}} |\mathcal{S}(x,a,u_1,\ldots,u_{N-1})|\big) < +\infty\Big\}$;

(ii) $H^1(\Omega) = \{f \in L^2(\Omega); \forall i = 1,\ldots,n \;\; \exists\partial_{x_i} f \text{ and } \partial_{x_i} f \in L^2(\Omega)\}$.
If $f \in H^1(\Omega)$ we define

$$\|f\|_{H^1(\Omega)} := \left(\int_\Omega |f|^2 + |\nabla f|^2 dx\right)^{\frac{1}{2}};$$

(iii) $L^2([0,T];H^1(\Omega)) = \Big\{t \mapsto f(t,\cdot) \in H^1(\Omega); \|f(t,\cdot)\|_{H^1(\Omega)}$ is a measurable function on $[0,T]$ and $\int_0^T \|f(t,\cdot)\|^2_{H^1(\Omega)} dt < +\infty\Big\}$;

(iv) $H^1([0,T];H^1(\Omega)) = \Big\{f(t,\cdot) \in H^1(\Omega), \partial_t f(t,\cdot) \in H^1(\Omega) \forall t \in [0,T]$; $\|f(t,\cdot)\|_{H^1(\Omega)}$, $\|\partial_t f(t,\cdot)\|_{H^1(\Omega)}$ are measurable functions on $[0,T]$ and $\int_0^T \left(\|f(t,\cdot)\|^2_{H^1(\Omega)} + \|\partial_t f(t,\cdot)\|^2_{H^1(\Omega)}\right) dt < +\infty\Big\}$.

Throughout the work we shall use the following assumptions on the data.

(H_1) $\varepsilon, C_{\mathcal{F}}, \mu_0, d_i, \sigma_i, \gamma_i, a_{i,j}$ are positive constants $(1 \le i < N, 1 \le j \le N)$;

(H_2) $u_{0i} \in C(\overline{\Omega})$ is nonnegative $(i = 1, \dots, N)$ and $(f_0)_x \in X_{[0,1]}$ for a.e. $x \in \Omega$;

(H_3) χ is the characteristic function of a measurable set $Q_0 \subseteq Q_T = \Omega \times [0,T]$; the function $\eta \in C([0,T])$ is nonnegative;

(H_4) for a.e. $x \in \Omega$, $\mathcal{G}_x \in C([0,1]^2)$, $\mathcal{G}_x(1,b) = 0$ for $b \in [0,1]$, and

$$-C \le \partial_a \mathcal{G}_x \le 0, \quad |\partial_b \mathcal{G}_x| \le C \text{ in } [0,1]^2; \tag{3.1}$$

(H_5) $\mathcal{S} \in L^\infty(\Omega; C([0,1] \times [0,+\infty)^{N-1}))$, $\mathcal{S}(x,1,u_1,\dots,u_{N-1}) = 0$ for $u_i \ge 0$ and a.e. $x \in \Omega$, and for all compact sets $K \subset [0,+\infty)^{N-1}$ there exists a constant $C(K)$ such that for a.e. $x \in \Omega$

$$-C(K) \le \partial_a \mathcal{S}(x,a,u) \le 0, |\nabla_u \mathcal{S}(x,a,u)| \le C(K) \text{ for } a \in [0,1], u \in K; \tag{3.2}$$

(H_6) $P \in C([0,T] \times [0,1]^2)$, P is nonnegative for all $t \in [0,T]$

$$\int_0^1 P(t,b,a)da = 1 \text{ for } b \in [0,1], \quad P(t,b,a) = 0 \text{ if } a < b, \tag{3.3}$$

and there exists $L > 0$ such that, for all $a', a'', b', b'' \in [0,1]$ and $t \in [0,T]$

$$|P(t,b',a') - P(t,b'',a'')| \le L(|b' - b''| + |a' - a''|). \tag{3.4}$$

Remark 3.2. Notice that it is easy to check that all functions presented in the Introduction of this thesis satisfy all the above assumptions.

3.2 Main result

We introduce some additional notation. Let $\mathcal{M}([0,1])$ be the space of signed Radon measures on the interval $[0,1]$. Then $\mathcal{M}([0,1])$ is the dual of

$C([0,1])$ (see [2], Remark 1.57). We say that $\mu : \Omega \times (0,T) \to \mathcal{M}([0,1])$ is weakly* measurable if for any $\rho \in C([0,1])$ the map

$$(x,t) \mapsto \int \rho(a) d\mu_{x,t}(a) \tag{3.5}$$

is measurable in $\Omega \times (0,T)$. We say that

$$f \in \mathcal{L}(\Omega; C([0,T]; X_{[0,1]}))$$

if $f \in C([0,T]; X_{[0,1]})$ for a.e. $x \in \Omega$ and f is weakly* measurable as a function from $\Omega \times (0,T)$ in $\mathcal{M}([0,1])$. In particular, if $f \in \mathcal{L}(\Omega; C([0,T]; X_{[0,1]}))$, then, by the Fubini theorem (see in Appendix), for all $\psi \in C([0,1] \times \overline{\Omega} \times [0,T])$

$$x \mapsto \int_0^T \left(\int \psi(a,x,t) df_{x,t}(a) \right) dt$$

is measurable and belongs to $L^\infty(\Omega)$. Indeed

$$\left| \int_0^T \left(\int \psi(a,x,t) df_{x,t}(a) \right) dt \right| \leq C \int_0^T \int df_{x,t}(a) dt = CT < +\infty$$

since $f_{x,t}$ is a probability measure and ψ is bounded.

We can now define what we mean by "solution" of problem (8)-(9). In order to define a solution in weak sense, we use a typical technique that it is often used to rewrite a differential equation in a form where less smoothness is required to define a "solution". The basic idea is to multiply the PDE by a smooth "test function", integrate one or more times over some domain, and finally use integration by parts to move derivatives onto the smooth test function. In this way the resulting equation involves fewer derivatives on the "solution", and hence requiring less regularity.

Definition 3.3. An $(N+1)$-ple $(f, u_1, \ldots, u_N)$ is called a solution of problem (8)-(9) in $\Omega \times [0,T]$ if

(i) $f \in \mathcal{L}(\Omega; C([0,T]; X_{[0,1]}))$;

(ii) $u_i \in C(\overline{Q}_T)$ and $u_i \geq 0$ in Q_T for $1 \leq i \leq N$;

(iii) the first equation in (8) is satisfied in a weak sense: for a.e. $x \in \Omega$

$$\int_0^\tau \Big(\int (\partial_t \phi + v_x \partial_a \phi) df_{x,t} + \int \phi dJ_{x,t} \Big) dt = \int \phi(\cdot, \tau) df_{x,\tau} - \int \phi(\cdot, 0) d(f_0)_x$$

for all $\tau \in [0, T]$ and $\phi \in C^1([0,1] \times [0,T])$, where the function v is defined by (3) and the signed measure J by (6);

(iv) if $1 \leq i < N$, $u_i \in L^2([0,T]; H^1(\Omega))$ and

$$d_i \int_0^T \left[\int_\Omega \nabla u_i(x,t) \cdot \nabla \psi(x,t) dx + \gamma_i \int_{\partial\Omega_1} u_i(x,t)\psi(x,t) d\sigma(x) \right] dt$$
$$= \varepsilon \int\int_{Q_T} u_i \partial_t \psi dx dt + \varepsilon \int_\Omega u_{0i} \psi(x,0) dx + \int\int_{Q_T} R_i \psi dx dt \qquad (3.6)$$

for all $\psi \in H^1([0,T]; H^1(\Omega))$, $\psi(x,T) = 0$, where R_i is defined as in (8) and $\mathcal{F}$ (which is part of R_1) by (7);

(v) $\partial_t u_N \in C(\overline{Q}_T)$, $u_N(\cdot, 0) = u_{0N}$ in Ω, and the equation for u_N in (8) is satisfied in Q_T.

Remark 3.4. (a) It follows from (6) and (H_6) that, for a.e. $x \in \Omega$, $\int dJ_{x,t}(a) = 0$ for $t \in [0,T]$. Indeed, by Fubini theorem,

$$\begin{aligned} \int dJ_{x,t}(a) &= \eta\chi \left[\int_0^1 \int_0^1 P(t,b,a) df_{x,t}(b) da - \int_0^1 df_{x,t}(a) \right] \\ &= \eta\chi \left[\int_0^1 \int_0^1 P(t,b,a) da\, df_{x,t}(b) - 1 \right] \\ &= \eta\chi \left[\int_0^1 1 df_{x,t}(b) - 1 \right] \\ &= \eta\chi \,[1-1] = 0 \end{aligned}$$

since $f_{x,t}$ is a probability measure on $[0,1]$.

(b) It follows from (3.1) - (3.2) that, for a.e. $x \in \Omega$, v_x is Lipschitz continuous w.r.t. a, uniformly in $t \in [0,T]$, i.e. there exists a constant L, which does not depend on t, such that

$$|v_x(a^{'}, t) - v_x(a^{''}, t)| \leq L |a^{'} - a^{''}| \quad \forall a^{'}, a^{''} \in [0,1], \forall t \in [0,T].$$

Indeed, if $(x,t) \mapsto (u_1,\dots,u_{N-1})$ is continuous on $\overline{Q}_T$, $(u_1,\dots,u_{N-1})$ belongs to a compact set K of $\mathbb{R}^{N-1}$. Therefore, by (3.1) - (3.2)

$$-C(K) \le \partial_a \mathcal{S} \le 0 \text{ and } -C \le \partial_a \mathcal{G}_x \le 0,$$

hence

$$\partial_a v_x(a,t) = \int_{[0,1]} \partial_a \mathcal{G}(x,a,b) df_{x,t}(b) + \partial_a \mathcal{S}(x,a,u_1(x,t),\dots,u_{N-1}(x,t)) \tag{3.7}$$

and

$$-C - C(K) \le \partial_a v_x(a,t) \le 0. \tag{3.8}$$

(c) As observed in [6], the concept of weak solution of the first order transport equation, defined in Definition 3.3(iii), needs some explanation. In particular, it follows from $(H_4 - H_5)$ that, for a.e. $x \in \Omega$, $v_x(1,t) = 0$ for $t \in [0,T]$ and $v_x(a,t) \ge 0$ for $a \in [0,1]$ and $t \in [0,T]$. This implies that formally the "flux" fv vanishes at $a = 1$, a condition which allows the choice of continuous test functions $\phi(x,a,t)$ without any restriction at $a = 1$. Since $v \ge 0$ at $a = 0$, characteristics (see the next chapter) "enter the domain $[0,1]$" at $a = 0$; so we need a boundary condition at $a = 0$ which, according to Definition 3.3(iii), is again the no flux condition. Actually this is imposed by the condition that $f_{x,\tau}$ is a probability measure in $[0,1]$: choosing $\phi \equiv 1$ it follows from Definition 3.3(iii) and from point (a) that for a.e. $x \in \Omega$

$$\int df_{x,\tau} = \int d(f_0)_x = 1 \text{ for } \tau \in (0,T].$$

Remark 3.5. We would like to stress that if we consider the continuity equation $\partial_t f + \partial_a(fv[f]) = 0$ without the jump operator, then the Dirac measure centered at the origin $\delta_0(a)$, which would correspond to a "healthy brain", is a weak solution of the continuity equation in the sense of Definition 3.3 (iii), without the integral term for J. Indeed, for $f_{x,t} = \delta_0$, the deterioration rate in (3) becomes

$$\begin{aligned}(v[\delta_0])_x(a,t) &= \int_0^1 \mathcal{G}_x(a,b) d\delta_0(b) + \mathcal{S}(x,a,u_1,\dots,u_{N-1}) \\ &= \mathcal{G}_x(a,0) + \mathcal{S}(x,a,u_1,\dots,u_{N-1}).\end{aligned}$$

Let $\tau \in [0,T]$, $\phi \in C^1([0,1]\times[0,T])$ and $u := (u_1, \ldots, u_{N-1})$ then, $\delta_0(a)$ is a solution if it satisfies

$$\int_0^\tau \Big(\int \partial_t \phi(a,t) d\delta_0(a)\Big) dt + \int_0^\tau \Big(\int (\mathcal{G}_x(a,0) + \mathcal{S}(x,a,u))\partial_a \phi(a,t) d\delta_0(a)\Big) dt$$
$$= \int \phi(a,\tau) d\delta_0(a) - \int \phi(a,0) d\delta_0(a)$$

$$\Longleftrightarrow$$

$$\int_0^\tau \partial_t \phi(0,t) dt + \int_0^\tau (\mathcal{G}_x(0,0) + \mathcal{S}(x,0,u_1,\ldots,u_{N-1}))\partial_a \phi(0,t) dt$$
$$= \phi(0,\tau) - \phi(0,0)$$

But now, if we consider $\mathcal{G}_x$ and $\mathcal{S}$ defined as in (2) and (4), then in a "healthy brain" the following term

$$\int_0^\tau (\mathcal{G}_x(0,0) + \mathcal{S}(x,0,u_1,\ldots,u_{N-1}))\partial_a \phi(0,t) dt = 0,$$

since both $\mathcal{G}_x(0,0)$ and $\mathcal{S}(x,0,u_1,\ldots,u_{N-1})) = C_{\mathcal{S}} \left(\sum_{m=1}^{N-1} m u_m(x,t) - \bar{U}\right)^+$ are equal to zero. Indeed, if the neuron is not damaged the amount of toxic Aβ oligomers does not exceed the threshold value $\bar{U}$. Therefore, the previous equation becomes

$$\int_0^\tau \partial_t \phi(0,t) dt = \phi(0,\tau) - \phi(0,0)$$

that is obviously satisfied for all $\tau \in [0,T]$ and $\phi \in C^1([0,1]\times[0,T])$.

We are now ready to present the main result studied in this thesis. Recalling [6], our purpose is to show that the problem (8) – (9) is well-posed.

Theorem 3.6. *Let $\Omega \subset \mathbb{R}^n$ be an open and bounded set with a smooth boundary $\partial\Omega$, which is the disjoint union of $\partial\Omega_0$ and $\partial\Omega_1$. Let $T > 0$ and $N \in \mathbb{N}$, and let hypotheses $(H_1$-$H_6)$ be satisfied. Then problem* (8) – (9) *has a unique solution in $\Omega \times [0,T]$ in the sense of Definition* 3.3.

We will prove this theorem by proceeding by steps. In particular, the purpose of the next chapter is to introduce the characteristics for the first order transport equation for f. Then we will rewrite the original problem (8) in terms of the characteristics and we will prove that the resulting system is equivalent to the original one.

Chapter 4

Statement of the problem in terms of characteristics

In this chapter we are going to reformulate problem (8)-(9) in terms of the so-called characteristics for the first order transport equation for f. As we will see, this provides us a new system that is equivalent to the original one. Once this equivalence is proved, in the following chapter we will prove the existence and uniqueness of the solution for the reformulated system. We stress that the proof of the equivalence between the new problem involving the characteristics and the original one will request some efforts. Indeed, the major difficulty arises from the strong non linearity of the system: the transport equation depends nonlinearly on both its solution (f), through an integral operator ($v[f]$), and the solution of the Smoluchowski system, which in turn depends on the solution of the transport equation ($\mathcal{F}[f]$).

Let us start to introduce the characteristics.
Let $f \in \mathcal{L}(\Omega; C([0,T]; X_{[0,1]}))$ and $u_i \in C(\overline{Q}_T)$, and let $v[f]$ be defined by (3). As we have seen in Remark 3.4 (b), for a.e. $x \in \Omega$ the map $a \mapsto v_x(a,t)$ is Lipschitz continuous, uniformly respect to $t \in [0,T]$. Thus, the following Cauchy problem issued from $y \in [0,1]$,

$$\begin{cases} \partial_t A_x(y,t) = v_x(A_x(y,t),t) \quad \text{for } 0 < t \leq T, \\ A_x(y,0) = y, \end{cases} \tag{4.1}$$

has a unique solution, by Cauchy-Lipischitz Theorem. In particular, the solution of (4.1) is called *characteristic* of the transport equation $\partial_t f + \partial_a(fv[f]) = 0$ which starts from y at the time $t = 0$.
We would like to point out that under our hypotheses the characteristics exist in the classical sense.
In particular, the solution of (4.1) satisfies

$$\begin{cases} 0 \leq A_x(y_1,t) < A_x(y_2,t) \leq A_x(1,t) = 1 & \text{if } 0 \leq y_1 < y_2 \leq 1, 0 \leq t \leq T, \\ A_x(y,t_1) \leq A_x(y,t_2) & \text{if } y \in [0,1], 0 \leq t_1 \leq t_2 \leq T. \end{cases} \tag{4.2}$$

Indeed, as for the first chain of inequalities, $A_x(1,t) = 1$ since for a.e. $x \in \Omega$, $v_x(1,t) = 0$ $\forall t \in [0,T]$, by our assumptions $(H_4 - H_5)$. Thus, $a = 1$ is a so-called steady state and the corresponding constant solution $A_x(1,t) \equiv 1$ satisfies (4.1) with initial data $y = 1$. Moreover, it follows from $(H_4 - H_5)$ that, for a.e. $x \in \Omega$, $v_x(a,t) \geq 0$ for $a \in [0,1]$ and $t \in [0,T]$. Hence $\partial_t A_x(y,t) \geq 0$, i.e. the function $t \mapsto A_x(y,t)$ is increasing. Therefore, if the initial data $y_1 \in [0,1]$, i.e. $A_x(y_1,0) = y_1 \geq 0$, then $A_x(y_1,t) \geq 0$ $\forall t \in [0,T]$. Finally, for uniqueness of the solution of a Cauchy problem, we have $A_x(y_1,t) < A_x(y_2,t)$ if $0 \leq y_1 < y_2 \leq 1, 0 \leq t \leq T$. As for the second inequality in (4.2), it follows again by monotony of the map $t \mapsto A_x(y,t)$.
In particular, for a.e. $x \in \Omega$, the function $y \mapsto A_x(y,t)$ is injective for all $t \in [0,T]$. Indeed, let suppose that $A_x(y_1,t) = A_x(y_2,t)$. If $y_1 < y_2 \Rightarrow A_x(y_1,t) < A_x(y_2,t)$, if $y_2 < y_1 \Rightarrow A_x(y_2,t) < A_x(y_1,t)$. It follows that $y_1 = y_2$. In addition, observe that, for a.e. $x \in \Omega$, $y \mapsto A_x(y,t)$ is continuous by a classical result on continuous dependence on initial data [19] and

$$\partial_y A_x(y,t) = \exp\left(\int_0^t \partial_a v_x(A_x(y,s),s)ds\right) > 0 \text{ for all } t \in [0,T]. \tag{4.3}$$

Now, before proceeding to reformulate the original problem in terms of the characteristics, we prove the following result.

Proposition 4.1. *Let* $f \in \mathcal{L}(\Omega; C([0,T]; X_{[0,1]}))$ *and* $u_i \in C(\overline{Q}_T)$. *Let* $v[f]$ *and* $J[f]$ *be defined by* (3) *and* (6). *Let, for a.e.* $x \in \Omega$, $A_x(y,t)$ *be the* *solution of* (4.1) *for any* $y \in [0,1]$. *If* f *satisfies* (5) *in the sense of Definition* 3.3(iii), *then, for a.e.* $x \in \Omega$,

$$\mathrm{supp} f_{x,t}, \ \ \mathrm{supp} J_{x,t} \subseteq [A_x(0,t), 1] \ \textit{for} \ t \in [0,T]. \tag{4.4}$$

Proof. A_x is well defined for a.e. $x \in \Omega$. We fix such x and also $\tau \in [0,T]$. Let $h \in C^1(\mathbb{R})$ be nondecreasing and satisfy $h \equiv 0$ in $(-\infty, 0]$ and $h \equiv 1$ in $[1, +\infty)$. Let $\delta > 0$ and set for a.e. $x \in \Omega$

$$h_\delta(s) = h\left(\frac{s}{\delta}\right) \text{ for } s \in \mathbb{R}, \qquad \psi_\delta(a,t) = h_\delta(A_x(0,t) - a) \text{ for } a \in [0,1], t \in [0,T].$$

Then, by construction, ψ_δ is of class C^1 and

$$\partial_a \psi_\delta(a,t) = \partial_a h_\delta(A_x(0,t) - a) = \partial_a h\left(\frac{A_x(0,t) - a}{\delta}\right) = -\frac{1}{\delta} h'\left(\frac{A_x(0,t) - a}{\delta}\right),$$

$$\begin{aligned}\partial_t \psi_\delta(a,t) &= \partial_t h_\delta(A_x(0,t) - a) = \partial_t h\left(\frac{A_x(0,t) - a}{\delta}\right) \\ &= \frac{1}{\delta}\partial_t A_x(0,t) h'\left(\frac{A_x(0,t) - a}{\delta}\right) = \frac{1}{\delta} v_x(A_x(0,t),t) h'\left(\frac{A_x(0,t) - a}{\delta}\right).\end{aligned}$$

Now, by assumption, f satisfies (5) in the sense of Definition 3.3(iii). We can use ψ_δ as a test function in Definition 3.3(iii). Since $A_x(0,0) = 0$, $\psi_\delta(a,0) = h_\delta(-a) = 0$ if $a \geq 0$. Hence $\int \psi_\delta(\cdot,0) d(f_0)_x = 0$. Therefore, the test function relation

$$\int_0^\tau \left(\int (\partial_t \psi_\delta + v_x \partial_a \psi_\delta) df_{x,t} \right) dt = \int \psi_\delta(\cdot,\tau) df_{x,\tau} - \int_0^\tau \left(\int \psi_\delta dJ_{x,t} \right) dt$$

implies that

$$\int \psi_\delta(\cdot,\tau) df_{x,\tau} - \int_0^\tau \left(\int \psi_\delta dJ_{x,t} \right) dt \to 0 \text{ as } \delta \to 0 \tag{4.5}$$

if we prove that

$$\int_0^\tau \left(\int (\partial_t \psi_\delta + v_x \partial_a \psi_\delta) df_{x,t} \right) dt \to 0 \text{ as } \delta \to 0. \tag{4.6}$$

To prove (4.6) we observe that

$$
\begin{aligned}
|\partial_t\psi_\delta + v_x\partial_a\psi_\delta| &= \left|\frac{v_x(A_x(0,t),t)-v_x(a,t)}{\delta}h^{'}\left(\frac{A_x(0,t)-a}{\delta}\right)\right| \\
&\le C\left|\frac{A_x(0,t)-a}{\delta}h^{'}\left(\frac{A_x(0,t)-a}{\delta}\right)\right| \\
&\le C\sup_{s\in\mathbb{R}}|sh^{'}(s)|
\end{aligned}
$$

by the Lipschitz continuity of $a\mapsto v_x(a,t)$, uniformly w.r.t. t, by Remark 3.4(b). In particular C is a some constant which does not depend on δ. Notice that $\sup_{s\in\mathbb{R}}|sh^{'}(s)|<+\infty$. Indeed, $sh^{'}(s)$ is continuous on $\mathbb{R}$ because h is C^1, $h^{'}(s)\equiv 0$ for $s\le 0$ and $s\ge 1$, while on the compact $0\le s\le 1$, $sh^{'}(s)$ is bounded.

Hence

$$
\Big|\int_0^\tau\left(\int(\partial_t\psi_\delta+v_x\partial_a\psi_\delta)df_{x,t}\right)dt\Big| \tag{4.7}
$$

$$
\le C\int_0^\tau\left(\int\left|\frac{A_x(0,t)-a}{\delta}h^{'}\left(\frac{A_x(0,t)-a}{\delta}\right)\right|df_{x,t}\right)dt \tag{4.8}
$$

Now, if $\frac{A_x(0,t)-a}{\delta}\le 0$ or $\frac{A_x(0,t)-a}{\delta}\ge 1$, then $h^{'}\left(\frac{A_x(0,t)-a}{\delta}\right)=0$, by construction of h. Thus we consider a such that $0<\frac{A_x(0,t)-a}{\delta}<1$ and $a\ge 0$, i.e.

$$
\begin{cases} a > A_x(0,t)-\delta \\ a < A_x(0,t) \\ a\ge 0\end{cases} \iff a\in(A_x(0,t)-\delta, A_x(0,t))\cap[0,A_x(0,t)).
$$

Hence

$$
\begin{aligned}
(4.8) &= C\int_0^\tau\left(\int\Big|\frac{A_x(0,t)-a}{\delta}h^{'}\left(\frac{A_x(0,t)-a}{\delta}\right)\Big|\right. \\
&\left. df_{x,t}\llcorner(A_x(0,t)-\delta,A_x(0,t))\cap[0,A_x(0,t))\right)dt \\
&\le C\sup_{s\in\mathbb{R}}|sh^{'}(s)|\int_0^\tau\left(\int df_{x,t}\llcorner(A_x(0,t)-\delta,A_x(0,t))\cap[0,A_x(0,t))\right)dt \\
&= C\tilde{C}\int_0^\tau\left(\int df_{x,t}\llcorner(A_x(0,t)-\delta,A_x(0,t))\cap[0,A_x(0,t))\right)dt.
\end{aligned}
$$

So, we have obtained

$$\Big|\int_0^\tau \left(\int (\partial_t\psi_\delta + v_x\partial_a\psi_\delta)df_{x,t}\right)dt\Big| \\ \leq C\tilde{C}\int_0^\tau \left(\int df_{x,t}\llcorner(A_x(0,t)-\delta, A_x(0,t))\cap[0,A_x(0,t))\right)dt \quad (4.9)$$

where we recall that, here and in the following, the symbol $\llcorner$ denotes the restriction of a measure to a measurable subset, as we have seen in Definition 1.18. Now, since $\bigcap_{\delta>0}(A_x(0,t)-\delta, A_x(0,t))\cap[0,A_x(0,t)) = \emptyset$ and

$$\left|\int df_{x,t}\llcorner(A_x(0,t)-\delta, A_x(0,t))\cap[0,A_x(0,t))\right| \leq 1 \text{ for } t\in[0,\tau],$$

(4.6) follows from (4.9) and the dominated convergence theorem. Therefore (4.5) follows.

Moreover by the dominated convergence theorem

$$\int \psi_\delta(\cdot,\tau)df_{x,\tau} \to \int df_{x,\tau}\llcorner[0,A_x(0,\tau)) \text{ as } \delta\to 0 \quad (4.10)$$

and

$$\int_0^\tau \left(\int \psi_\delta dJ_{x,t}\right)dt \to \int_0^\tau \left(\int dJ_{x,t}\llcorner[0,A_x(0,t))\right)dt \text{ as } \delta\to 0. \quad (4.11)$$

In particular, as for (4.10), $|\psi_\delta(\cdot,\tau)| \leq 1 \in L^1_{df_{x,\tau}}$ and $\psi_\delta(\cdot,\tau)\to I_{[0,A_x(0,\tau))}$ as $\delta\to 0$. Indeed,

$$\text{if } 0\leq a < A_x(0,\tau) \Rightarrow \psi_\delta(a,\tau) = h\left(\frac{A_x(0,\tau)-a}{\delta}\right) \to 1 \text{ as } \delta\to 0,$$
$$\text{if } a\geq A_x(0,\tau) \Rightarrow \psi_\delta(a,\tau) = h\left(\frac{A_x(0,\tau)-a}{\delta}\right) \to 0 \text{ as } \delta\to 0.$$

Hence we obtain (4.10) by the dominated convergence theorem.

As for, (4.11):

$$\begin{aligned}\left|\int \psi_\delta dJ_{x,t}\right| &= \left|\int \psi_\delta(a,t)\eta\chi\left(\int P(t,b,a)df_{x,t}(b)\right)da - \int \psi_\delta(a,t)\eta\chi df_{x,t}(a)\right| \\ &\leq \eta\chi\left|\int \psi_\delta(a,t)\left(\int P(t,b,a)df_{x,t}(b)\right)da\right| \\ &+ \eta\chi\int |\psi_\delta(a,t)|df_{x,t}(a) \\ &\leq C_1\eta\left|\int_0^1 \psi_\delta(a,t)da\right| + \eta\int df_{x,t}(a) \\ &\leq (C_1+1)\eta \leq (C_1+1)\sup_{0\leq t\leq T}\eta(t) = C_2 < \infty \text{ and } \int_0^\tau C_2 < +\infty\end{aligned}$$

where we have used $0 \leq \psi_\delta \leq 1$, (H_3), $\sup_{[0,T]\times[0,1]^2} P(t,b,a) < +\infty$ by (H_6) and the fact that $f_{x,t}$ is a probability measure.
Moreover, $\int \psi_\delta(a,t)dJ_{x,t}(a) \to \int dJ_{x,t} \llcorner [0, A_x(0,t))$ as $\delta \to 0$. Indeed, as for the first term in $\int \psi_\delta(a,t)dJ_{x,t}(a)$,

$$\begin{aligned}\eta\chi\int \psi_\delta(a,t)\Big(\int_0^1 P(t,b,a)df_{x,t}(b)\Big)da \to \\ \eta\chi\int I_{[0,A_x(0,t))}(a)\left(\int_0^1 P(t,b,a)df_{x,t}(b)\right)da \\ = \eta\chi\int_0^{A_x(0,t)}\left(\int_0^1 P(t,b,a)df_{x,t}(b)\right)da\end{aligned} \tag{4.12}$$

as $\delta \to 0$ by the dominated convergence theorem since

$$\left|\psi_\delta(a,t)\int_0^1 P(t,b,a)df_{x,t}(b)\right| \leq C \in L^1_{da} \ \ (0 \leq a \leq 1)$$

and

$$\psi_\delta(a,t)\int_0^1 P(t,b,a)df_{x,t}(b) \to I_{[0,A_x(0,t))}(a)\left(\int_0^1 P(t,b,a)df_{x,t}(b)\right) \text{ as } \delta \to 0.$$

As for the second term, we have already shown by (4.10) that

$$\eta\chi\int \psi_\delta(a,t)df_{x,t}(a) \to \eta\chi\int df_{x,t} \llcorner [0, A_x(0,t)) \text{ as } \delta \to 0. \tag{4.13}$$

Thus by (4.12) and (4.13) we have

$$\int \psi_\delta(a,t)dJ_{x,t}(a) \to \int dJ_{x,t} \llcorner [0, A_x(0,t)) \text{ as } \delta \to 0,$$

hence we can conclude that (4.11) holds by the dominated convergence theorem.
Finally, since both (4.5) and (4.10), (4.11) hold, then by uniqueness of the limit we have

$$\int df_{x,\tau} \llcorner [0, A_x(0,\tau)) = \int_0^\tau \left(\int dJ_{x,t} \llcorner [0, A_x(0,t)) \right) dt. \tag{4.14}$$

Now, it follows from (6), Fubini's theorem, and (3.3) that

$$\begin{aligned}
&\int dJ_{x,t}(a) \llcorner [0, A_x(0,t)) \\
&= \eta(t)\chi(x,t) \left[\int \left(\int_0^{A_x(0,t)} P(t,b,a)da \right) df_{x,t}(b) - \int df_{x,t}(a) \llcorner [0, A_x(0,t)) \right] \\
&= \eta\chi \Big[\int \left(\int_b^{A_x(0,t)} P(t,b,a)da \right) df_{x,t}(b) \llcorner [0, A_x(0,t)) \\
&- \int df_{x,t} \llcorner [0, A_x(0,t)) \Big] \qquad (4.15) \\
&\leq \eta\chi \left[\int df_{x,t}(b) \llcorner [0, A_x(0,t)) - \int df_{x,t}(a) \llcorner [0, A_x(0,t)) \right] = 0.
\end{aligned}$$

Combined with (4.14), it gives

$$0 \leq \int df_{x,\tau} \llcorner [0, A_x(0,\tau)) = \int_0^\tau \left(\int dJ_{x,t} \llcorner [0, A_x(0,\tau)) \right) dt \leq 0$$

hence $\int df_{x,\tau} \llcorner [0, A_x(0,\tau)) = 0$ for $\tau \in [0,T]$ and thus $\text{supp} f_{x,t} \subseteq [A_x(0,t), 1]$ for $t \in [0,T]$. In addition, if we come back to (4.15) with $\int df_{x,t} \llcorner [0, A_x(0,t)) = 0$, we have

$$\begin{aligned}
\int dJ_{x,t}(a) \llcorner [0, A_x(0,t)) &= \eta\chi \int \left(\int_b^{A_x(0,t)} P(t,b,a)da \right) df_{x,t}(b) \llcorner [0, A_x(0,t)) \\
&\geq 0.
\end{aligned}$$

Hence

$$0 \leq \int dJ_{x,t}(a) \llcorner [0, A_x(0,t)) \leq 0 \Rightarrow \int dJ_{x,t}(a) \llcorner [0, A_x(0,t)) = 0,$$

which implies $\text{supp} J_{x,t} \subseteq [A_x(0,t), 1]$ for $t \in [0,T]$. □

We are now ready to reformulate the original problem in terms of the characteristics. Specifically, we shall see below that the measure f can be obtained by transporting along the characteristics a suitable measure g. In particular f is the push forward of g through A (cf. Definition 1.19), where g satisfies

$$\begin{cases} \partial_t A_x(y,t) = \int \mathcal{G}_x(A_x(y,t), A_x(\xi,t))dg_{x,t}(\xi) + \mathcal{S}(x, A_x(y,t), u_1, \dots, u_{N-1}), \\ \partial_t g_{x,t}(y) = \eta\chi\Big[\partial_y A_x(y,t) \int P(t, A_x(\xi,t), A_x(y,t))dg_{x,t}(\xi) - g_{x,t}(y)\Big], \\ \varepsilon\partial_t u_1 - d_1\Delta u_1 = \tilde{R}_1 \\ \qquad := -u_1 \displaystyle\sum_{j=1}^{N} a_{1,j}u_j - \sigma_1 u_1 + C_{\mathcal{F}} \int (\mu_0 + A_x(\xi,t))(1 - A_x(\xi,t))dg_{x,t}(\xi), \\ \varepsilon\partial_t u_m - d_m\Delta u_m = \tilde{R}_m := -u_m \sum_{j=1}^{N} a_{m,j}u_j + \frac{1}{2}\sum_{j=1}^{m-1} a_{j,m-j}u_j u_{m-j} - \sigma_m u_m, \\ \varepsilon\partial_t u_N = \frac{1}{2} \sum_{\substack{j+k\geq N \\ k,j<N}} a_{j,k}u_j u_k, \end{cases} \tag{4.16}$$

where $x \in \Omega$, $y \in [0,1]$, $t \in (0,T]$, and $2 \leq m < N$, with initial and boundary conditions

$$\begin{cases} g_{x,0}(y) = (f_0)_x(y), \quad A_x(y,0) = y & \text{if } x \in \Omega, \quad 0 \leq y \leq 1, \\ u_i(x,0) = u_{0i}(x) & \text{if } x \in \Omega, \quad 1 \leq i \leq N, \\ \partial_n u_i(x,t) = 0 & \text{if } x \in \partial\Omega_0, \quad t \in (0,T], \quad 1 \leq i < N, \\ \partial_n u_i(x,t) = -\gamma_i u_i(x,t) & \text{if } x \in \partial\Omega_1, \quad t \in (0,T], \quad 1 \leq i < N. \end{cases} \tag{4.17}$$

Notice that this new system has $N+2$ equations, while the original one has $N+1$. Indeed, as we have already anticipated, the idea is to find f as the push forward of g through A, i.e. $f := A_\sharp g$, so we have an extra unknown variable, that is the characteristic. In particular, $(4.16)_1$ is the same equation of $(4.1)_1$ by writing $v_x(A_x(y,t),t)$ with the push forward, according to Definition 1.19, i.e.

$$
\begin{aligned}
\partial_t A_x(y,t) &= v_x(A_x(y,t),t) \\
&= \int \mathcal{G}_x(A_x(y,t),b)df_{x,t}(b) + \mathcal{S}(x,A_x(y,t),u_1,\ldots,u_{N-1}) \\
&= \int \mathcal{G}_x(A_x(y,t),A_x(\xi,t))dg_{x,t}(\xi) + \mathcal{S}(x,A_x(y,t),u_1,\ldots,u_{N-1}).
\end{aligned}
$$

Similarly for $(4.16)_3$, while in $(4.16)_2$ the right-hand side reformulates the jump term $J[f]$ by using again the push-forward through A. In particular, the measure da which appears in (6) becomes formally $\partial_y A_x(y,t)dy$, by the relation $a = A_x(y,t)$.

Definition 4.2. The $(N+2)$-ple $(A, g, u_1, \ldots, u_N)$ is called a solution of problem (4.16)-(4.17) in $\Omega \times [0,T]$ if

(i) $g \in \mathcal{L}(\Omega; C([0,T]; X_{[0,1]}))$;

(ii) $A, \partial_t A \in L^\infty(\Omega; C([0,1] \times [0,T]; [0,1]))$;

(iii) $u_i \in C(\overline{Q}_T)$ and $u_i \geq 0$ in Q_T for $1 \leq i \leq N$;

(iv) for a.e. $x \in \Omega$, A_x satisfies $(4.16)_1$ and $A_x(y,0) = y$ for $y \in [0,1]$;

(v) $(4.16)_2$ for g is satisfied in a weak sense for a.e. $x \in \Omega$: for all $\tau \in (0,T]$ and $\phi \in C([0,1] \times [0,T])$ with $\partial_t \phi \in C([0,1] \times [0,T])$

$$
\begin{aligned}
&\int \phi(y,\tau)dg_{x,\tau}(y) - \int \phi(y,0)d(f_0)_x(y) - \int_0^\tau \left(\int \partial_t \phi(y,t) dg_{x,t}(y) \right) dt \\
&= \int_0^\tau \eta\chi \Bigg[\int_0^1 \phi(y,t)\partial_y A_x(y,t) \left(\int P(t, A_x(\xi,t), A_x(y,t)) dg_{x,t}(\xi) \right) dy \\
&\quad - \int \phi(y,t) dg_{x,t}(y) \Bigg] dt; \qquad (4.18)
\end{aligned}
$$

(vi) if $1 \leq i < N$, $u_i \in L^2([0,T]; H^1(\Omega))$, and

$$
\begin{aligned}
d_i \int_0^T &\left[\int_\Omega \nabla u_i(x,t) \cdot \nabla \psi(x,t) dx + \gamma_i \int_{\partial\Omega_1} u_i(x,t)\psi(x,t) d\sigma(x) \right] dt \\
&= \varepsilon \int\int_{Q_T} u_i \partial_t \psi dx dt + \varepsilon \int_\Omega u_{0i}\psi(x,0)dx + \int\int_{Q_T} \tilde{R}_i \psi dx dt \quad (4.19)
\end{aligned}
$$

for all $\psi \in H^1([0,T]; H^1(\Omega))$, $\psi(x,T) = 0$, where $\tilde{R}_i$ is defined as in (4.16);

(vii) $\partial_t u_N \in C(\overline{Q}_T)$, $u_N(\cdot,0) = u_{0N}$ in Ω, and the equation for u_N in (4.16) is satisfied in Q_T.

In the remainder of this chapter we prove the equivalence of problems (8)-(9) and (4.16)-(4.17), proceeding by steps. Let us start with the first step by proving the following result which ensures that if we have a solution of (4.16)-(4.17), then we can find, by push-forward, a solution of (8)-(9).

Theorem 4.3. *Let hypotheses* $(H_1 - H_6)$ *be satisfied. Let* $(A, g, u_1, \dots, u_N)$ *be a solution of* $(4.16) - (4.17)$ *in* $\Omega \times [0,T]$ *and set, for a.e.* $x \in \Omega$,

$$f_{x,t} := A_x(\cdot,t)_\sharp g_{x,t} \text{ for all } t \in [0,T].$$

Then $(f, u_1, \dots, u_N)$ *is a solution of problem* $(8) - (9)$ *in* $\Omega \times [0,T]$.

Notice that if $(A, g, u_1, \dots, u_N)$ is a solution of (4.16)-(4.17), then, in order to prove that $(f, u_1, \dots, u_N)$, defined as above, is a solution of (8)-(9), it is sufficient to verify Definition 3.3(i),(iii) and (iv) (only for $i = 1$) since the other requests are the same presented in Definition 4.2 and thus satisfied by assumptions.

Proof. Since, for a.e. $x \in \Omega$, $g_{x,t}$ is a Borel regular probability measure in $[0,1]$ for $t \in [0,T]$, so is $f_{x,t}$ by Remark 1.20, Remark 1.15 and Remark B.15. As we have said above, for a.e. $x \in \Omega$ the function $y \mapsto A_x(y,t)$ with $y \in [0,1]$ is continuous and by (4.3) it is also injective for $t \in [0,T]$, so that by Theorem B.13 it is open and closed and thus by Proposition 2.15, recalling that $A_x(1,t) = 1$

$$\mathrm{supp} f_{x,t} = A_x(\mathrm{supp} g_{x,t}, t) \subseteq A_x([0,1], t) = [A_x(0,t), 1]. \tag{4.20}$$

In particular, $f_{x,t} \in X_{[0,1]}$ for a.e. $x \in \Omega$, since $[A_x(0,t), 1] \subseteq [0,1]$. In addition, by Theorem B.13 and by Corollary 2.21, the map $t \mapsto f_{x,t}$ belongs to $C([0,T];X_{[0,1]})$ for a.e. $x \in \Omega$. Finally, $f \in \mathcal{L}(\Omega; C([0,T];X_{[0,1]}))$ since f is weakly* measurable. Indeed, $\forall \rho \in C([0,1])$ the map

$$(x,t) \mapsto \int \rho(a) df_{x,t}(a) = \int \rho(A_x(y,t)) dg_{x,t}(y)$$

is measurable in $\Omega \times (0,T)$ since $(x,t) \mapsto \int \rho(A_x(y,t)) dg_{x,t}(y)$ is measurable, being $y \mapsto \rho(A_x(y,t))$ continuous and $g \in \mathcal{L}(\Omega; C([0,T];X_{[0,1]}))$.
Now let, for a.e. $x \in \Omega$, v be defined by (3) and J by (6). By (4.20) and (3.3), $\int P(t,b,a) df_{x,t}(b) = 0$ if $a < A_x(0,t)$, $\text{supp} J_{x,t} \subseteq \text{supp} f_{x,t}$, whence

$$\text{supp} J_{x,t} \subseteq [A_x(0,t), 1]. \tag{4.21}$$

To avoid cumbersome notations, we set $B_x(\cdot,t) := A_x^{-1}(\cdot,t)$. Since

$$A_x(\cdot,t) : [0,1] \to [A_x(0,t), A_x(1,t)] = [A_x(0,t), 1],$$

is injective and surjective, $B_x(\cdot,t)$ is well-defined in $[A_x(0,t),1]$,

$$B_x(A_x(y,t),t) \equiv y \text{ for } y \in [0,1], \text{ and } A_x(B_x(a,t),t) \equiv a \text{ for } a \in [A_x(0,t),1].$$

Since $\text{supp} f_{x,t} \subseteq [A_x(0,t),1]$, integrals of functions of $B_x(\cdot,t)$ w.r.t. $f_{x,t}$ are well-defined.

By Definition 4.2(iv), $\partial_t A_x(y,t) = v_x(A_x(y,t),t)$ for a.e. $x \in \Omega$. By (4.3), B_x is Lipschitz continuous w.r.t. y for a.e. $x \in \Omega$. Differentiating the identity $A_x(B_x(y,t),t) = y$ with respect to t and y, we obtain that

$$\begin{cases} \partial_y A_x(B_x(a,t),t)\partial_t B_x(a,t) + \partial_t A_x(B_x(a,t),t) = 0, \\ \partial_y A_x(B_x(y,t),t)\partial_y B_x(y,t) = 1 \end{cases} \iff$$

$$\begin{cases} \partial_y A_x(B_x(a,t),t)\partial_t B_x(a,t) = -\partial_t A_x(B_x(a,t),t) = \\ \qquad\qquad - v_x(A_x(B_x(a,t),t),t) = -v_x(a,t), \\ \partial_y A_x(B_x(y,t),t)\partial_y B_x(y,t) = 1, \end{cases} \tag{4.22}$$

so that $\partial_t B_x(y,t)\partial_y A_x(B_x(y,t),t)\partial_y B_x(y,t) = \partial_t B_x(y,t) \cdot 1 = \partial_t B_x(y,t)$ and, by (4.22),

$$\partial_t B_x(y,t) = -v_x(y,t)\partial_y B_x(y,t). \tag{4.23}$$

Let us prove Definition 3.3(iii). Let $\psi \in C^1([0,1]\times[0,T])$. Let $\tau \in (0,T]$. Let x be fixed such that $A_x, \partial_t A_x \in C([0,1]\times[0,T];[0,1])$ (by assumption A is solution, thus it satisfies Definition 4.2(ii)) and set

$$\phi(y,t) = \psi(A_x(y,t),t) \text{ for } y \in [0,1]$$

and

$$\begin{aligned} C_\phi &= -\int \phi(y,\tau)dg_{x,\tau}(y) + \int \phi(y,0)d(f_0)_x(y) \\ &= -\int \phi(B_x(A_x(y,\tau),\tau),\tau)dg_{x,\tau}(y) + \int \phi(y,0)d(f_0)_x(y) \\ &= -\int \phi(B_x(a,\tau),\tau)df_{x,\tau}(a) + \int \phi(a,0)d(f_0)_x(a), \end{aligned}$$

where the last equality holds by definition of push-forward with $f_{x,\tau} = A_x(\cdot,\tau)_\sharp g_{x,\tau}$. Since ϕ satisfies the conditions in Definition 4.2(v), it follows that

$$\begin{aligned} &-\int_0^T \left(\int \partial_t\phi(y,t)dg_{x,t}(y)\right)dt \\ &= \int_0^T \eta\chi\Bigg[\int_0^1 \phi(y,t)\partial_y A_x(y,t)\left(\int P(t,A_x(\xi,t),A_x(y,t))dg_{x,t}(\xi)\right)dy \\ &- \int \phi(y,t)dg_{x,t}(y)\Bigg]dt + C_\phi \\ &= \int_0^T \eta\chi\Bigg[\int_0^1 \phi(B_x(A_x(y,t),t),t)\partial_y A_x(y,t)\left(\int P(t,b,A_x(y,t))df_{x,t}(b)\right)dy \\ &- \int \phi(B_x(A_x(y,t),t),t)dg_{x,t}(y)\Bigg]dt + C_\phi \\ &= \int_0^T \eta\chi\Bigg[\int_{A_x(0,t)}^{A_x(1,t)} \phi(B_x(a,t),t)\left(\int P(t,b,a)df_{x,t}(b)\right)da \\ &- \int \phi(B_x(A_x(y,t),t),t)dg_{x,t}(y)\Bigg]dt + C_\phi \end{aligned} \tag{4.24}$$

where we have used the Definition 1.19 of push-forward and the relation $da = \partial_y A(x,y,t)dy$.

On the other hand, the first left-hand side of (4.24) can be written as

$$
\begin{aligned}
-\int_0^T \left(\int \partial_t \phi(y,t) dg_{x,t}(y) \right) dt &= -\int_0^T \left(\int \partial_t \phi(B_x(A_x(y,t),t),t) dg_{x,t}(y) \right) dt \\
&= -\int_0^T \left(\int \partial_t \phi(B_x(a,t),t) df_{x,t}(a) \right) dt.
\end{aligned}
\tag{4.25}
$$

Let $a \in [A_x(0,t),1]$. Then $\psi(a,t) = \psi(A_x(B_x(a,t),t),t) = \phi(B_x(a,t),t)$ and

$$
\partial_t \psi(a,t) = \partial_y \phi(B_x(a,t),t) \partial_t B_x(a,t) + \partial_t \phi(B_x(a,t),t). \tag{4.26}
$$

Since we know, by (4.20), that $\text{supp} f_{x,t} \subseteq [A_x(0,t),1]$, it follows from (4.24)-(4.25)-(4.26) that

$$
\begin{aligned}
&-\int_0^T \left(\int \partial_t \psi(a,t) df_{x,t}(a) \right) dt = \\
&-\int_0^T \left(\int \partial_y \phi(B_x(a,t),t) \partial_t B_x(a,t) df_{x,t}(a) \right) dt \\
&-\int_0^T \left(\int \partial_t \phi(B_x(a,t),t) df_{x,t}(a) \right) dt \\
&= -\int_0^T \left(\int \partial_y \phi(B_x(a,t),t) \partial_t B_x(a,t) df_{x,t}(a) \right) dt - \int_0^T \int \partial_t \phi(y,t) dg_{x,t}(y) dt
\end{aligned}
$$

$$
\begin{aligned}
&= -\int_0^T \left(\int \partial_y \phi(B_x(a,t),t) \partial_t B_x(a,t) df_{x,t}(a) \right) dt \\
&+ \int_0^T \eta\chi \left[\int_{A_x(0,t)}^{A_x(1,t)} \phi(B_x(a,t),t) \left(\int P(t,b,a) df_{x,t}(b) \right) da \right. \\
&\left. - \int \phi(B_x(A_x(y,t),t),t) dg_{x,t}(y) \right] dt + C_\phi \\
&= -\int_0^T \left(\int \partial_y \phi(B_x(a,t),t) \partial_t B_x(a,t) df_{x,t}(a) \right) dt \\
&+ \int_0^T \eta\chi \left[\int_{A_x(0,t)}^{A_x(1,t)} \psi(a,t) \left(\int P(t,b,a) df_{x,t}(b) \right) da \right. \\
&\left. - \int \psi(a,t) df_{x,t}(a) \right] dt + C_\phi \\
&= -\int_0^T \left(\int \partial_y \phi(B_x(a,t),t) \partial_t B_x(a,t) df_{x,t}(a) \right) dt \\
&+ \int_0^T \left(\int \psi(a,t) dJ_{x,t}(a) \right) dt + C_\phi. \qquad (4.27)
\end{aligned}
$$

By (4.23),

$$
\begin{aligned}
-\int \partial_y \phi(B_x(a,t),t) \partial_t B_x(a,t) df_{x,t}(a) \qquad (4.28) \\
= \int \partial_y \phi(B_x(a,t),t) \partial_a B_x(a,t) v_x(a,t) df_{x,t}(a) \\
= \int \partial_a \psi(a,t) v_x(a,t) df_{x,t}(a), \qquad (4.29)
\end{aligned}
$$

whence, by (4.27), the first equation in (8) is satisfied in the weak sense:

$$
\begin{aligned}
-\int_0^T \left(\int \partial_t \psi(a,t) df_{x,t}(a) \right) dt = \int_0^T \left(\int \partial_a \psi(a,t) v_x(a,t) df_{x,t}(a) \right) dt \\
+ \int_0^T \left(\int \psi(a,t) dJ_{x,t}(a) \right) dt - \int \psi(a,\tau) df_{x,\tau}(a) + \int \psi(a,0) d(f_0)_x(a).
\end{aligned}
\qquad (4.30)
$$

So we have proved that Definition 3.3(iii) holds.

Concerning the Smoluchowski system in (8), it is sufficient to notice that the third equation in (4.16) and the second equation in (8) coincide via the push-forward, since

$$\int (\mu_0 + a)(1-a)df_{x,t}(a) = \int (\mu_0 + A_x(\xi,t))(1 - A_x(\xi,t))dg_{x,t}(\xi).$$

□

The proof of the equivalence between problems (8)-(9) and (4.16)-(4.17) is completed by the following result.

Theorem 4.4. *Let* $(f, u_1, \dots, u_N)$ *be a solution of* (8) – (9) *in* $\Omega \times [0,T]$ *and let* $A_x(y,t)$ *be defined by* (4.1). *Then there exists a probability measure* $g_{x,t}$ *such that*

$$f_{x,t} := A_x(\cdot,t)_\sharp g_{x,t},$$

and $(A, g, u_1, \dots, u_N)$ *is a solution of problem* (4.16) – (4.17) *in* $\Omega \times [0,T]$.

Proof. As before, we fix $x \in \Omega$ for a.e. $x \in \Omega$, and $t \in [0,T]$. We consider the map

$$A_x(\cdot,t) : [0,1] \mapsto [A_x(0,t), 1].$$

By Proposition 4.1, $f_{x,t} = f_{x,t} \llcorner [A_x(0,t),1]$. Hence, by Theorem 1.20 in [24], there exists a positive Radon measure $g_{x,t}$ on $[0,1]$ such that

$$f_{x,t} = f_{x,t} \llcorner [A_x(0,t),1] = A_x(\cdot,t)_\sharp g_{x,t}.$$

Obviously, since $f_{x,t}$ is a probability measure, $g_{x,t}$ is a probability measure too and belongs to $X_{[0,1]}$. By Theorem B.13 and by Corollary 2.21, the map $t \mapsto g_{x,t}$ is continuous w.r.t. the Wasserstein metric. Moreover, $g_{x,t} \to (f_0)_x$ as $t \to 0$ since $A_x(y,0) = y$. Indeed, $t \mapsto g_{x,t}$ is continuous w.r.t. the Wasserstein metric, therefore $g_{x,t} \to g_{x,0}$ as $t \to 0$. On the other hand, $f_{x,t} = A_x(\cdot,t)_\sharp g_{x,t}$ and $t \mapsto f_{x,t}$ is continuous w.r.t. the Wasserstein metric since f is solution of (8)-(9) by assumption. Thus $f_{x,t} \to f_{x,0}$ as $t \to 0$ and $A_x(\cdot,t)_\sharp g_{x,t} \to A_x(\cdot,0)_\sharp g_{x,0}$ as $t \to 0$. Hence, by uniqueness of the limit (in the Wasserstein metric) $f_{x,0} = A_x(\cdot,0)_\sharp g_{x,0}$, i.e. $f_{x,0} = g_{x,0}$ since $A_x(\cdot,0)$ is the identity map. But now by hypotheses $f_{x,0} = (f_0)_x$, whence $g_{x,0} = (f_0)_x$. Therefore g satisfies the qualitative assumptions in order to be a solution of (4.16) and (4.17). To complete the proof of the theorem, it is enough to check the identities in the proof of Theorem 4.3 in the opposite direction. □

Chapter 5

Existence and Uniqueness Result

As we have seen in Chapter 4 the problem (8)-(9) can be reformulated in terms of the characteristics providing a new system (4.16)-(4.17) which we have proved by Theorems 4.3 and 4.4 to be equivalent to the original one. Now, the goal of this chapter is to prove the main result of this thesis, Theorem 3.6. In order to do this, we have organized this chapter in two sections: in the first one we prove local (with respect to t) existence and uniqueness of a solution of problem (4.16)-(4.17), while in the second section we show that this solution can be continued in $[0, T]$. In this way Theorem 3.6 will be proved.

5.1 Local existence and Uniqueness

The purpose of this section is to prove the following result.

Theorem 5.1. *Let $\Omega \subset \mathbb{R}^n$ be an open and bounded set with a smooth boundary $\partial\Omega$, which is the disjoint union of smooth manifolds $\partial\Omega_0$ and $\partial\Omega_1$. Let $T > 0$ and $N \in \mathbb{N}$, and let hypotheses $(H_1 - H_6)$ be satisfied. Then there exists $\tau \in (0, T]$ such that problem $(4.16) - (4.17)$ has a unique solution in $\Omega \times [0, \tau]$.*

The proof is based on a contraction argument. In particular, we recall that the Wasserstein distance $\mathcal{W}_1$ is the metric involved by the probability measures f. By Kantorovich-Rubinstein duality (see Proposition A.3), the Wasserstein distance depends on the action of the measures on Lipschitz functions. This fact yields a technical difficulty when we try to apply an iteration argument in order to obtain the local existence of a solution. In order to circumvent this difficulty we will introduce an ad hoc formulation of the standard fixed point theorem.

To this purpose, we introduce a suitable metric space. Let us consider the following normed spaces

$$L^\infty(\Omega; C([0,1]\times[0,\tau];[0,1])), \text{ and } C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N)$$

endowed with the following standard norms:
if $A \in L^\infty(\Omega; C([0,1]\times[0,\tau];[0,1]))$, define

$$\|A\| := \sup_{x\in\Omega} \max_{[0,1]\times[0,\tau]} |A_x(y,t)|;$$

if $u \in C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N)$, define

$$\|u\| := \max_{\overline{\Omega}\times[0,\tau]} \|u(x,t)\|_{\mathbb{R}^N}.$$

Definition 5.2. Let $\tau \in (0,T]$ be given. We denote by $(\mathcal{X}_\tau, d)$ the complete metric space

$$\mathcal{X}_\tau := L^\infty(\Omega; C([0,1]\times[0,\tau];[0,1]))\times\mathcal{L}(\Omega; C([0,\tau]; X_{[0,1]}))\times C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N),$$

where $L^\infty(\Omega; C([0,1]\times[0,\tau];[0,1]))$ and $C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N)$ are endowed with their natural metrics as normed spaces, and $\mathcal{L}(\Omega; C([0,\tau]; X_{[0,1]}))$ is endowed with the metric

$$\sup_{x\in\Omega} \max_{t\in[0,\tau]} \mathcal{W}_1(f_{x,t}, g_{x,t})$$

(Notice that condition (3.5) passes to the limit w.r.t. the $\mathcal{W}_1$-convergence, by Proposition 2.17).

We denote by $\mathcal{X}_{\tau,\rho}$ the closed ball in $\mathcal{X}_\tau$ of radius $\rho > 0$ centered at (y, f_0, u_0).

Notice that, for the moment, we have given up the nonnegativity of u_i, since $u = (u_1, \dots, u_N) \in C(\overline{\Omega} \times [0, \tau]; \mathbb{R}^N)$. However it will be recovered during the proof of Theorem 5.1. For this reason we define $\mathcal{S}$ also for negative values of u_i, by requiring that $\mathcal{S}$ is even with respect to u_i for each $i = 1, \dots, N-1$.

We must construct the map to which we can apply the contraction argument. We will do this proceeding by steps.
Let us start to prove the following Lemma whose meaning is that to reformulate the Cauchy problem (4.1) presented in Chapter 4 by replacing the rate $v_x(a, t)$ with an analogous rate $\hat{v}_x(a, t)$ defined by fixing a point in this new metric space $\mathcal{X}_\tau$.

Lemma 5.3. *Let* $(\hat{A}, g, u) \in \mathcal{X}_T$ *and set, for a.e.* $x \in \Omega$,

$$\hat{v}_x(a,t) := \int \mathcal{G}_x(a, \hat{A}_x(\xi, t)) dg_{x,t}(\xi) + \mathcal{S}(x, a, u_1, \dots, u_{N-1}) \geq 0. \tag{5.1}$$

Then, for a.e. $x \in \Omega$, *the Cauchy problem*

$$\begin{cases} \partial_t \underline{A}_x(y,t) = \hat{v}_x(\underline{A}_x(y,t), t) & \text{for } t > 0, \\ \underline{A}_x(y, 0) = y \in [0,1], \end{cases} \tag{5.2}$$

has a unique solution defined for all $t \in (0, T]$, *and the function* $y \mapsto \underline{A}_x(y,t)$ *is continuous, strictly increasing (and thus open) on* $[0,1]$, *and maps* $[0,1]$ *onto* $[\underline{A}_x(0,t), 1]$ *for all* $t \in [0, T]$. *Finally, the map* $(x, y, t) \mapsto \underline{A}_x(y,t)$ *belongs to* $L^\infty(\Omega; C([0,1] \times [0,T]; [0,1]))$.

Notice that the function $y \mapsto \underline{A}_x(y,t)$ has features similar to $y \mapsto A_x(y,t)$ defined by (4.1).

Proof. Let us start to specify that a continuous, strictly increasing function on $[0,1]$ is open by Theorem B.13 since it is injective.
Now, we claim that, for a.e. $x \in \Omega$, the map $(a,t) \mapsto \hat{v}_x(a,t)$ is continuous and Lipschitz continuous w.r.t. $a \in [0,1]$, uniformly in $t \in [0,T]$.

By (3.2) this is trivial for the map $(a,t) \mapsto \mathcal{S}(x, a, u_1(x,t), \dots, u_{N-1}(x,t))$, since $(x,t) \mapsto (u_1, \dots, u_{N-1})$ is continuous on $\overline{\Omega} \times [0,T]$ and $(u_1, \dots, u_{N-1})$

belongs to a compact set of $\mathbb{R}^{N-1}$. It remains to show that

$$(a,t) \mapsto \int \mathcal{G}_x(a, \hat{A}_x(\xi,t))dg_{x,t}(\xi)$$

is continuous and uniformly Lipschitz continuous with respect to $a \in [0,1]$.

Let $a, a_0 \in [0,1]$ and $t, t_0 \in (0,T]$ be given. Then

$$\begin{aligned}
&\Big|\int \mathcal{G}_x(a, \hat{A}_x(\xi,t))dg_{x,t}(\xi) - \int \mathcal{G}_x(a_0, \hat{A}_x(\xi,t_0))dg_{x,t_0}(\xi)\Big| \\
&\quad\le \Big|\int \mathcal{G}_x(a, \hat{A}_x(\xi,t))dg_{x,t}(\xi) - \int \mathcal{G}_x(a_0, \hat{A}_x(\xi,t_0))dg_{x,t}(\xi)\Big| \\
&\quad+ \Big|\int \mathcal{G}_x(a_0, \hat{A}_x(\xi,t_0))dg_{x,t}(\xi) - \int \mathcal{G}_x(a_0, \hat{A}_x(\xi,t_0))dg_{x,t_0}(\xi)\Big| := I_1 + I_2,
\end{aligned} \tag{5.3}$$

by adding and subtracting the same term $\int \mathcal{G}_x(a_0, \hat{A}_x(\xi,t_0))dg_{x,t}(\xi)$ and using the triangle inequality. Now, $(a,\xi,t) \mapsto \mathcal{G}_x(a, \hat{A}_x(\xi,t))$ is uniformly continuous in $[0,1]^2 \times [0,T]$, since, by our assumptions, it is continuous in a compact set. Therefore $I_1 \to 0$ as $(a,t) \to (a_0,t_0)$. Indeed,

$$\begin{aligned}
&\Big|\int \mathcal{G}_x(a, \hat{A}_x(\xi,t))dg_{x,t}(\xi) - \int \mathcal{G}_x(a_0, \hat{A}_x(\xi,t_0))dg_{x,t}(\xi)\Big| \\
&\le \int \big|\mathcal{G}_x(a, \hat{A}_x(\xi,t)) - \mathcal{G}_x(a_0, \hat{A}_x(\xi,t_0))\big| dg_{x,t}(\xi)
\end{aligned} \tag{5.4}$$

Now, we know by uniform continuity that

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \text{ such that } \big|\mathcal{G}_x(a, \hat{A}_x(\xi,t)) - \mathcal{G}_x(a_0, \hat{A}_x(\xi_0,t_0))\big| < \varepsilon,$$

$$\forall (a,\xi,t);\, |(a,\xi,t) - (a_0,\xi_0,t_0)| < \delta.$$

Hence, if we take (a,ξ,t), (a_0,ξ,t_0), such that $|(a,\xi,t)-(a_0,\xi,t_0)| = |(a,t)-(a_0,t_0)| < \delta$ (we have chosen $\xi_0 = \xi$) we obtain

$$\begin{aligned}
(5.4) &\le \varepsilon \int dg_{x,t}(\xi) \\
&= \varepsilon.
\end{aligned}$$

Thus, we have shown that $\forall \varepsilon$ there exists $\delta > 0$ such that $|I_1| < \varepsilon$ for all (a,t) such that $0 < |(a,t) - (a_0,t_0)| < \delta$, i.e. $I_1 \to 0$ as $(a,t) \to (a_0,t_0)$.

As for I_2, since $\xi \mapsto \mathcal{G}_x(a_0, \hat{A}_x(\xi, t_0))$ is continuous in $[0,1]$ and for a.e. $x \in \Omega$ $t \mapsto g_{x,t} \in C([0,T]; X_{[0,1]})$, then $g_{x,t} \to g_{x,t_0}$ w.r.t. the Wasserstein distance as $t \to t_0$, i.e. $\mathcal{W}_1(g_{x,t}, g_{x,t_0}) \to 0$ as $t \to t_0$. Therefore, by Proposition 2.17 and by Remark 2.2, $t \mapsto g_{x,t}$ is narrowly continuous, then $I_2 \to 0$ as $t \to t_0$.

Similarly, by (3.1), for a.e. $x \in \Omega$ and all $\xi \in [0,1]$ and $t \in [0,T]$,

$$|\mathcal{G}_x(a, \hat{A}_x(\xi,t)) - \mathcal{G}_x(a', \hat{A}_x(\xi,t))| \le C|a - a'| \text{ for } a, a' \in [0,1].$$

This completes the proof of the claim, which implies, for a.e. $x \in \Omega$, the existence and uniqueness of the solution of problem (5.2) for all $y \in [0,1]$. By a standard argument,

$$\partial_y \underline{A}_x(y,t) = \exp\left[\int_0^t \partial_a \hat{v}_x(\underline{A}_x(y,s), s)ds\right] > 0. \tag{5.5}$$

So, by (5.5), (3.1) and (3.2), we have $0 < C_1 \le \partial_y \underline{A}_x(y,t) \le C_2$ for some constants C_1, C_2 which depend on the compact set $K \subset \mathbb{R}^{N-1}$ which contains $(u_1(x,t), \dots, u_{N-1}(x,t))$. Finally, by a classical result on continuous dependence on initial data [19], we have that $(x,y,t) \mapsto \underline{A}_x(y,t)$ belongs to $L^\infty(\Omega; C([0,1] \times [0,T]; [0,1]))$. □

Remark 5.4. It follows from the proof of Lemma 5.3 (in particular from (5.5)) that $\underline{A}_x(\xi, s)$ is Lipschitz continuous in ξ, uniformly w.r.t. x and s.

Lemma 5.5. *Let $(\hat{A}, g, u) \in \mathcal{X}_T$. Let for a.e. $x \in \Omega$, $\underline{A}$ be defined as in Lemma* 5.3 *and $(F[g])_{x,t}$ be the signed measure on $[0,1]$ defined by*

$$d(F[g])_{x,t} := \eta\chi\left[\partial_y \underline{A}_x(y,t) \int P(t, \underline{A}_x(\xi,t), \underline{A}_x(y,t)) dg_{x,t}(\xi) dy - dg_{x,t}(y)\right]$$

for $0 < t \le T$. Then, for a.e. $x \in \Omega$,

(i) the integral equation

$$\underline{g}_{x,t} = (f_0)_x + \int_0^t (F[\underline{g}])_{x,s} ds \tag{5.6}$$

has a unique solution $t \mapsto \underline{g}_{x,t}$ which belongs to $C([0,T]; X_{[0,1]})$;[1]

[1]In the integral equation (5.6), the integral term has the following meaning: if $t \to \mu_t$ is a continuous map from $[0,T]$ to $X_{[0,1]}$, for any Borel set $B \subset [0,1]$, we set $\left(\int_0^t \mu_s ds\right)(B) := \int_0^t \mu_s(B) ds$.

(ii) the measure $\underline{g}_{x,t}$ *is a weak solution of the system*

$$\begin{cases} \partial_t \underline{g}_{x,t}(y) = \eta\chi\Big[\partial_y \underline{A}_x(y,t) \int P(t, \underline{A}_x(\xi,t), \underline{A}_x(y,t)) d\underline{g}_{x,t}(\xi) - \underline{g}_{x,t}(y)\Big], \\ \underline{g}_{x,0} = (f_0)_x \end{cases} \tag{5.7}$$

in the sense of (4.18).

Proof. First of all, we observe that for a.e. $x \in \Omega$ and $s \in [0,T]$ and for all $g \in X_{[0,1]}$,

$$\int d(F[g])_{x,s} = 0. \tag{5.8}$$

This assertion is obvious if $\chi(x,s) = 0$. If $\chi(x,s) = 1$, by Fubini's theorem, by (3.3) and by posing $b = \underline{A}_x(y,s)$ (implying the relation $db = \partial_y \underline{A}_x(y,s)dy$), we get

$$\begin{aligned} \frac{1}{\eta}\int d(F[g])_{x,s} &= \int \Big(\int P(s, \underline{A}_x(\xi,s), \underline{A}_x(y,s)) \partial_y \underline{A}_x(y,s) dy\Big) dg_{x,s}(\xi) \\ &\quad - \int dg_{x,s}(y) \\ &= \int \Big(\int_{\underline{A}_x(0,s)}^{\underline{A}_x(1,s)=1} P(s, \underline{A}_x(\xi,s), b) db\Big) dg_{x,s}(\xi) - \int dg_{x,s}(y) \\ &= \int \Big(\int_0^1 P(s, \underline{A}_x(\xi,s), b) db\Big) dg_{x,s}(\xi) - \int dg_{x,s}(y) \\ &= \int dg_{x,s}(\xi) - \int dg_{x,s}(y) = 0, \end{aligned}$$

where $\int_{\underline{A}_x(0,s)}^1 P(s, \underline{A}_x(\xi,s), b) db = \int_0^1 P(s, \underline{A}_x(\xi,s), b) db$ recalling that, by (3.3)
$P(s, \underline{A}_x(\xi,s), b) = 0$ if $b < \underline{A}_x(\xi,s)$, and $\underline{A}_x(0,s) \le \underline{A}_x(\xi,s)$ because $\xi \mapsto \underline{A}_x(\xi,s)$ is increasing by Lemma 5.3. Therefore for $0 \le b < \underline{A}_x(0,s)$ we have $P(s, \underline{A}_x(\xi,s), b) = 0$, so we can extend the integral in $[0,1]$.

We set for a.e. $x \in \Omega$ (from now on we fix such x),

$$q_t := e^{\int_0^t \eta(s)\chi(x,s)ds} g_{x,t} \quad \text{for } t \in [0,T].$$

Let Y be the set of such q, i.e. $q \in Y$ if the map $t \mapsto e^{-\int_0^t \eta(s)\chi(x,s)ds} q_t$ belongs to $C([0,T]; X_{[0,1]})$. Then Y naturally inherits a metric from $C([0,T]; X_{[0,1]})$,

$$d_Y(q_1, q_2) := \max_{t\in[0,T]} \mathcal{W}_1 \left(e^{-\int_0^t \eta(s)\chi(x,s)ds} (q_1)_t, e^{-\int_0^t \eta(s)\chi(x,s)ds} (q_2)_t \right),$$

so Y is a complete metric space.

The equation for g translates into

$$\partial_t q_t(y) = L q_t(y) := \eta\chi(x,t)\partial_y \underline{A}_x(y,t) \int P(t, \underline{A}_x(\xi,t), \underline{A}_x(y,t)) dq_t(\xi) \geq 0,$$

and the corresponding integral equation is

$$q_t = (f_0)_x + \int_0^t L q_s ds \quad \text{for } t \in [0,T]. \tag{5.9}$$

Indeed, formally

$$\begin{aligned}
\partial_t q_t(y) &= \partial_t (e^{\int_0^t \eta(s)\chi(x,s)ds} g_{x,t}) \\
&= \eta(t)\chi(x,t) e^{\int_0^t \eta(s)\chi(x,s)ds} g_{x,t} + e^{\int_0^t \eta(s)\chi(x,s)ds} \partial_t g_{x,t} \\
&= \eta\chi q_t + \\
&+ e^{\int_0^t \eta(s)\chi(x,s)ds} \eta\chi \left[\partial_y \underline{A}_x(y,t) \int P(t, \underline{A}_x(\xi,t), \underline{A}_x(y,t)) dg_{x,t}(\xi) - g_{x,t}(y) \right] \\
&= \eta\chi q_t - \eta\chi e^{\int_0^t \eta(s)\chi(x,s)ds} g_{x,t} + \\
&+ \eta\chi \partial_y \underline{A}_x(y,t) \int P(t, \underline{A}_x(\xi,t), \underline{A}_x(y,t)) e^{\int_0^t \eta(s)\chi(x,s)ds} dg_{x,t}(\xi) \\
&= \eta\chi q_t - \eta\chi q_t + \eta\chi \partial_y \underline{A}_x(y,t) \int P(t, \underline{A}_x(\xi,t), \underline{A}_x(y,t)) dq_t(\xi) \\
&= \eta\chi(x,t) \partial_y \underline{A}_x(y,t) \int P(t, \underline{A}_x(\xi,t), \underline{A}_x(y,t)) dq_t(\xi).
\end{aligned}$$

Now, we consider the map

$$q \mapsto (f_0)_x + \int_0^t L q_s ds. \tag{5.10}$$

To this map we will apply a contraction argument. In particular, one easily checks that, by (5.8), for all $q \in Y$

$$\int dL q_t = \eta(t)\chi(x,t) e^{\int_0^t \eta(s)\chi(x,s)ds} \text{ for } t \in [0,T], \quad (f_0)_x + \int_0^t L q_s ds \in Y. \tag{5.11}$$

We prove that $(f_0)_x + \int_0^t Lq_s ds \in Y$. By definition of the space Y we have to show that the map $t \mapsto e^{-\int_0^t \eta(s)\chi(x,s)}((f_0)_x + \int_0^t Lq_s ds) \in C([0,T], X_{[0,1]})$. Let us start to prove that $\forall t \in [0,T]$, $e^{-\int_0^t \eta(s)\chi(x,s)}((f_0)_x + \int_0^t Lq_s ds) \in X_{[0,1]}$. Recalling that $(f_0)_x \in X_{[0,1]}$ and $q_t \in Y$, by Fubini's Theorem we have

$$
\begin{aligned}
&\int e^{-\int_0^t \eta(s)\chi(x,s)ds} d\Big((f_0)_x + \int_0^t Lq_s ds\Big)(y) \\
&= \int e^{-\int_0^t \eta(s)\chi(x,s)ds} d(f_0)_x(y) + \int e^{-\int_0^t \eta(s)\chi(x,s)} d\Big(\int_0^t Lq_s ds\Big)(y) \\
&= e^{-\int_0^t \eta(s)\chi(x,s)ds} + \\
&+ \int_0^t e^{-\int_0^t \eta\chi}\Big(\int \eta(s)\chi(x,s)\partial_y \underline{A}_x(y,s) \\
&\Big(\int P(t, \underline{A}_x(\xi,s), \underline{A}_x(y,s)) dq_s(\xi)\Big) dy\Big) ds \\
&= e^{-\int_0^t \eta(s)\chi(x,s)ds} + \\
&+ \int_0^t e^{-\int_0^s \eta\chi} e^{-\int_s^t \eta\chi}\Big(\int \eta(s)\chi(x,s)\Big(\int P(t, \underline{A}_x(\xi,s), b) db\Big) dq_s(\xi)\Big) ds \\
&= e^{-\int_0^t \eta(s)\chi(x,s)ds} + \int_0^t e^{-\int_s^t \eta\chi} \eta(s)\chi(x,s) \int e^{-\int_0^s \eta\chi} dq_s(\xi) ds \\
&= e^{-\int_0^t \eta(s)\chi(x,s)ds} + \int_0^t e^{\int_t^s \eta(u)\chi(x,u)du} \eta(s)\chi(x,s) ds \\
&= e^{-\int_0^t \eta(s)\chi(x,s)ds} + \Big[e^{\int_t^s \eta\chi}\Big]_{s=0}^{s=t} \\
&= e^{-\int_0^t \eta(s)\chi(x,s)ds} + 1 - e^{\int_t^0 \eta(u)\chi(x,u)du} \\
&= e^{-\int_0^t \eta(s)\chi(x,s)ds} + 1 - e^{-\int_0^t \eta(u)\chi(x,u)du} = 1.
\end{aligned}
$$

Concerning the continuity of the map $t \mapsto e^{-\int_0^t \eta(s)\chi(x,s)}((f_0)_x + \int_0^t Lq_s ds) \in C([0,T]; X_{[0,1]})$, by Proposition 2.17 it is enough to show that it is narrowly continuous. Let $\rho \in C([0,1])$.

$$\int \rho(y)e^{-\int_0^t \eta\chi} d\Big((f_0)_x + \int_0^t Lq_s ds\Big)(y) = \int \rho(y)e^{-\int_0^t \eta\chi} d(f_0)_x(y) + \int \rho(y)e^{-\int_0^t \eta\chi} d\Big(\int_0^t Lq_s ds\Big)(y) =: I_1 + I_2.$$

As for I_1, by the dominated convergence theorem $I_1 \to \int \rho(y)e^{-\int_0^{t_0} \eta\chi} d(f_0)_x(y)$ as $t \to t_0$.
As for I_2:

$$I_2 = \int \rho(y)e^{-\int_0^t \eta\chi}\Big(\int_0^t \eta\chi\partial_y \underline{A}_x(y,s)\Big(\int P(s, \underline{A}_x(\xi,s), \underline{A}_x(y,s))dq_s(\xi)\Big)ds\Big)dy,$$

let us show that we can apply the dominated convergence theorem:

$$\begin{aligned}
&\Big|\rho(y)e^{-\int_0^t \eta\chi}\Big(\int_0^t \eta\chi\partial_y \underline{A}_x(y,s)\Big(\int P(s, \underline{A}_x(\xi,s), \underline{A}_x(y,s))dq_s(\xi)\Big)ds\Big)\Big| \\
&\quad \le \max_{y\in[0,1]} |\rho(y)| \max_{[0,T]} \eta \ \ C_2 \int_0^T \int P(s, \underline{A}_x(\xi,s), \underline{A}_x(y,s))dq_s(\xi)ds \\
&\quad \le \max_{y\in[0,1]} |\rho(y)| \max_{[0,T]} \eta \ \ C_2 C \int_0^T \int e^{\int_0^s \eta\chi} e^{-\int_0^s \eta\chi} dq_s(\xi)ds \\
&\quad \le \tilde{C} e^{\int_0^T \eta(u)\chi(x,u)du} \int_0^T \int e^{-\int_0^s \eta(u)\chi(x,u)du} dq_s(\xi)ds \\
&\quad = \bar{C} \int_0^T 1ds \\
&\quad \le \bar{C}T \in L^1_{dy} \ \ (y \in [0,1])
\end{aligned}$$

where we have used (5.5), (H_3), (H_6), and the assumption $q_t \in Y$, i.e. $e^{-\int_0^t \eta\chi} q_t \in C([0,T]; X_{[0,1]})$. Moreover,

$$\rho(y)e^{-\int_0^t \eta\chi}\Big(\int_0^t \eta\chi\partial_y \underline{A}_x(y,s)\Big(\int P(s, \underline{A}_x(\xi,s), \underline{A}_x(y,s))dq_s(\xi)\Big)ds\Big)$$

converges to

$$\rho(y)e^{-\int_0^{t_0} \eta\chi}\Big(\int_0^{t_0} \eta\chi\partial_y \underline{A}_x(y,s)\Big(\int P(s, \underline{A}_x(\xi,s), \underline{A}_x(y,s))dq_s(\xi)\Big)ds\Big)$$

as $t \to t_0$. Hence by the dominated convergence theorem

$$\int \rho(y) e^{-\int_0^t \eta(s)\chi(x,s)ds} d\Big(\int_0^t Lq_s ds\Big)(y) \longrightarrow \int \rho(y) e^{-\int_0^{t_0} \eta(s)\chi(x,s)ds} d\Big(\int_0^{t_0} Lq_s ds\Big)(y)$$

as $t \to t_0$. Therefore, $(f_0)_x + \int_0^t Lq_s ds \in Y$.
Let us come back to consider the map (5.10). If we show that it is a contraction, i.e., it is such that for all $q_1, q_2 \in Y$

$$J_0(q_1,q_2) := d_Y\left((f_0)_x + \int_0^t L(q_1)_s ds, (f_0)_x + \int_0^t L(q_2)_s ds\right) \le C d_Y(q_1,q_2), \tag{5.12}$$

it follows from a standard contraction argument that the map (5.10) has a unique fixed point in a sufficiently small interval $[0,\tau]$ and that (5.9) has a unique local solution $\underline{q}$ which can be continued in $[0,T]$.

To prove (5.12) we use the characterization of the $\mathcal{W}_1$-distance given in Proposition A.3:

$$d_Y(q_1,q_2) = \max_{t\in[0,T]} \left[e^{-\int_0^t \eta(s)\chi(x,s)ds} \sup\left\{ \int \phi d(q_1-q_2)_t ; \phi \in \mathrm{Lip}_1([0,1],\mathbb{R}) \right\} \right], \tag{5.13}$$

where $\mathrm{Lip}_1([0,1],\mathbb{R})$ is the space of Lipschitz continuous functions $\Phi : [0,1] \to \mathbb{R}$ with Lipschitz constant not greater than 1.
Hence

$$J_0(q_1,q_2) = \max_{t\in[0,T]} \left[e^{-\int_0^t \eta(s)\chi(x,s)ds} \sup\{ I_\phi(t); \phi \in \mathrm{Lip}_1([0,1],\mathbb{R})\} \right],$$

where

$$I_\phi(t) := \int \phi d \int_0^t (L(q_1)_s - L(q_2)_s) ds$$

and $L(q_1)_s - L(q_2)_s$ is given by

$$\eta(s)\chi(x,s)\left(\int P(s, \underline{A}_x(\xi,s), \underline{A}_x(y,s)) \partial_y \underline{A}_x(y,s) d(q_1-q_2)_s(\xi) \right).$$

By Fubini's theorem, $I_\phi(t)$ is equal to

$$\int \phi(y)\int_0^t \Big(\eta\chi\Big(\int P(s,\underline{A}_x(\xi,s),\underline{A}_x(y,s))\partial_y\underline{A}_x(y,s)d(q_1-q_2)_s(\xi)\Big)dy\Big)ds$$
$$=\int_0^t \eta\chi\Big(\int_{\underline{A}_x(0,s)}^1 \phi(B_x(b,s))\Big(\int P(s,\underline{A}_x(\xi,s),b)d(q_1-q_2)_s(\xi)\Big)db\Big)ds$$
$$=\int_0^t \eta\chi\Big(\int\Big(\int_{\underline{A}_x(0,s)}^1 \phi(B_x(b,s))P(s,\underline{A}_x(\xi,s),b)db\Big)d(q_1-q_2)_s(\xi)\Big)ds.$$

By (5.11), $I_\phi(t)=0$ if ϕ is constant, so we may assume that $\phi(0)=0$. Indeed, we could always take $\tilde\phi(y)=\phi(y)-\phi(0)$, whence $\tilde\phi(0)=0$ and $|\tilde\phi(y^{'})-\tilde\phi(y^{''})|=|\phi(y^{'})-\phi(y^{''})|\leq|y^{'}-y^{''}|$, i.e. $\tilde\phi\in \mathrm{Lip}_1([0,1],\mathbb{R})$.
Hence, we can assume that $\phi(0)=0$. So, $|\phi(y)|=|\phi(y)-\phi(0)|\leq C|y-0|\leq |y|\leq 1$, i.e., $|\phi|\leq 1$ and, by (3.4),

$$\Big|\int_{\underline{A}_x(0,s)}^1 \phi(B_x(b,s))(P(s,\underline{A}_x(\xi^{'},s),b)-P(s,\underline{A}_x(\xi^{''},s),b))db\Big|$$
$$\leq \int_{\underline{A}_x(0,s)}^1 |P(s,\underline{A}_x(\xi^{'},s),b)-P(s,\underline{A}_x(\xi^{''},s),b)|db\leq L|\xi^{'}-\xi^{''}|.$$

Therefore the function

$$\xi\mapsto \frac{1}{L}\int_{\underline{A}_x(0,s)}^1 \phi(B_x(b,s))P(s,\underline{A}_x(\xi,s),b)db\in \mathrm{Lip}_1([0,1],\mathbb{R}),$$

whence

$$I_\phi(t)=\int_0^t \eta\chi\Big(\int\Big(\int_{\underline{A}_x(0,s)}^1 \frac{L}{L}\phi(B_x(b,s))P(s,\underline{A}_x(\xi,s),b)db\Big)d(q_1-q_2)_s(\xi)\Big)ds$$
$$\leq L\int_0^t \eta\chi\sup\Big\{\int\phi d(q_1-q_2)_s;\phi\in \mathrm{Lip}_1([0,1],\mathbb{R})\Big\}ds$$
$$\leq L\max_{[0,T]}\eta\int_0^t \sup\Big\{\int\phi d(q_1-q_2)_s;\phi\in \mathrm{Lip}_1([0,1],\mathbb{R})\Big\}ds.$$

By passing to the sup:

$$\sup\{I_\phi(t);\phi\in \mathrm{Lip}_1([0,1],\mathbb{R})\}$$
$$\leq L\max_{[0,T]}\eta\int_0^t \sup\Big\{\int\phi d(q_1-q_2)_s;\phi\in \mathrm{Lip}_1([0,1],\mathbb{R})\Big\}ds$$

and

$$
\begin{aligned}
& e^{-\int_0^t \eta(s)\chi(x,s)ds} \sup\{I_\phi(t); \phi \in \mathrm{Lip}_1([0,1],\mathbb{R})\} \\
& \leq L \max_{[0,T]} \eta \int_0^t \Big[e^{-\int_0^t \eta(u)\chi(x,u)du} \cdot \\
& \cdot \sup \Big\{ \int \phi d(q_1 - q_2)_s ; \phi \in \mathrm{Lip}_1([0,1],\mathbb{R}) \Big\} \Big] ds \\
& \leq L \max_{[0,T]} \eta \int_0^t \max_{s\in[0,T]} \Big[e^{-\int_0^s \eta(u)\chi(x,u)du} \cdot \\
& \cdot \sup \Big\{ \int \phi d(q_1 - q_2)_s ; \phi \in \mathrm{Lip}_1([0,1],\mathbb{R}) \Big\} \Big] ds \\
& \leq LT \max_{[0,T]} \eta \max_{s\in[0,T]} \Big[e^{-\int_0^s \eta(u)\chi(x,u)du} \sup \Big\{ \int \phi d(q_1 - q_2)_s ; \phi \in \mathrm{Lip}_1([0,1],\mathbb{R}) \Big\} \Big] \\
& = LT \max_{[0,T]} \eta \; d_Y(q_1, q_2).
\end{aligned}
$$

Thus (5.12) follows by taking the sup of the left-hand side term:

$$J_0(q_1, q_2) \leq LT \max_{t\in[0,T]} \eta(t) \; d_Y(q_1, q_2).$$

Therefore, the map (5.10) has a unique fixed point in a sufficiently small interval $[0, \tau]$, which can be continued in $[0, T]$.
Setting

$$\underline{g}_{x,t} = e^{-\int_0^t \eta(s)\chi(x,s)ds} \underline{q}_t \text{ for } t \in [0, T],$$

we have completed the proof of part (i) of the Lemma.

Fix an $x \in \Omega$ for which (5.6) and (5.9) (for $\underline{q}$) are valid. Since the map $t \mapsto \underline{g}_{x,t}$ is continuous in the weak* topology (and so is $t \mapsto \underline{q}_t$) by Proposition 2.17, and P and $\underline{A}_x$ are continuous functions, then we can say that the map

$$(y, t) \mapsto \int P(t, \underline{A}_x(\xi, t), \underline{A}_x(y, t)) d\underline{q}_t(\xi)$$

is continuous in $[0, 1] \times [0, T]$ by arguing in the same way as in (5.3). In particular, it is also bounded. Hence $L(\underline{q}, \cdot) \in L^\infty((0, 1) \times (0, T))$, where we recall that $L\underline{q}_t := \eta\chi(x, t)\partial_y \underline{A}_x(y, t) \int P(t, \underline{A}_x(\xi, t), \underline{A}_x(y, t)) dq_t(\xi) \geq 0$.

We set $\tilde{q} = \underline{q} - (f_0)_x$. By (5.9)

$$\tilde{q}_t = \underline{q}_t - (f_0)_x = (f_0)_x - (f_0)_x + \int_0^t L\underline{q}_s ds = \int_0^t L(\tilde{q}_s + (f_0)_x) ds$$

i.e.,

$$\tilde{q}_t = \int_0^t L(\tilde{q}_s + (f_0)_x) ds \text{ for } t \in [0, T].$$

Since, by boundedness of $L(\tilde{q}_s + (f_0)_x)(y)$, $t \mapsto \tilde{q}_t(y)$ is absolutely continuous in $[0, T]$ for a.e. $y \in (0, 1)$. This means that for all $\tau \in (0, T]$ and $\psi \in L^\infty([0, 1] \times [0, T])$ with $\partial_t \psi \in L^\infty([0, 1] \times [0, T])$, we have

$$\begin{aligned}\int_0^1 \psi(y, \tau)\tilde{q}_\tau(y) dy &= \int_0^1 \Big(\int_0^\tau (\psi(y, t)\tilde{q}_t(y))' dt \Big) dy \\ &= \int\int_{(0,1)\times(0,\tau)} [\partial_t \psi(y, t)\tilde{q}_t(y) + \psi(y, t)L((\tilde{q}_t + (f_0)_x)(y))] dy dt.\end{aligned} \tag{5.14}$$

Finally let $\phi(y, t) \in C([0, 1] \times [0, T])$ with $\partial_t \phi \in C([0, 1] \times [0, T])$ (we recall that x is fixed). We substitute the function $\psi(y, t) = e^{-\int_0^t \eta(s)\chi(x,s) ds} \phi(y, t)$ into (5.14). Since

$$\partial_t \psi(y, t) = e^{-\int_0^t \eta(s)\chi(x,s) ds} (-\eta\chi\phi(y, t) + \partial_t \phi(y, t)),$$

ψ and $\partial_t \psi$ are continuous with respect to y and, by a straightforward calculation, (5.14) transforms into

$$\begin{aligned}&\int \phi(y, \tau) d\underline{g}_{x,\tau}(y) - \int \psi(y, 0) d(f_0)_x(y) \\ &= \int e^{\int_0^\tau \eta(s)\chi(x,s)} \psi(y, \tau) d\underline{g}_{x,\tau}(y) - \int \psi(y, 0) d(f_0)_x(y) \\ &= \int \psi(y, \tau) d\underline{q}_\tau(y) - \int \psi(y, 0) d(f_0)_x(y) \\ &= \int \psi(y, \tau) d\tilde{q}_\tau(y) + \int \psi(y, \tau) d(f_0)_x(y) - \int \psi(y, 0) d(f_0)_x(y) \\ &= \int\int_{(0,1)\times(0,\tau)} [\partial_t \psi(y, t)\tilde{q}_t(y) + \psi(y, t)L((\tilde{q}_t + (f_0)_x)(y))] dy dt + \\ &+ \int \psi(y, \tau) d(f_0)_x(y) - \int \psi(y, 0) d(f_0)_x(y)\end{aligned}$$

$$
\begin{aligned}
&= \int\int_{(0,1)\times(0,\tau)} e^{-\int_0^t \eta(s)\chi(x,s)ds}(-\eta\chi\phi(y,t) + \partial_t\phi(y,t))\tilde{q}_t(y)dydt+ \\
&+ \int\int_{(0,1)\times(0,\tau)} \psi(y,t)L((\tilde{q}_t + (f_0)_x)(y))dydt+ \\
&+ \int(\psi(y,\tau) - \psi(y,0))d(f_0)_x(y) \\
&= \int\int_{(0,1)\times(0,\tau)} \partial_t\phi(y,t)e^{-\int_0^t \eta(s)\chi(x,s)ds}\underline{q}_t(y)dydt \\
&- \int\int_{(0,1)\times(0,\tau)} \partial_t\phi(y,t)e^{-\int_0^t \eta(s)\chi(x,s)ds}d(f_0)_x(y)dt \\
&- \int\int_{(0,1)\times(0,\tau)} \eta(t)\chi(x,t)\phi(y,t)e^{-\int_0^t \eta(s)\chi(x,s)ds}(\underline{q}_t - (f_0)_x)(y)dydt+ \\
&+ \int\int_{(0,1)\times(0,\tau)} \psi(y,t)L\underline{q}_t(y)dydt + \int_0^\tau \Big(\int \partial_t\psi(y,t)d(f_0)_x(y)\Big)dt \\
&= \int\int_{(0,1)\times(0,\tau)} \partial_t\phi(y,t)d\underline{g}_{x,t}(y)dt \\
&- \int\int_{(0,1)\times(0,\tau)} \partial_t\phi(y,t)e^{-\int_0^t \eta(s)\chi(x,s)ds}d(f_0)_x(y)dt \\
&- \int\int_{(0,1)\times(0,\tau)} \eta(t)\chi(x,t)\phi(y,t)d\underline{g}_{x,t}(y)dt \\
&- \int\int_{(0,1)\times(0,\tau)} -\eta(t)\chi(x,t)\phi(y,t)e^{-\int_0^t \eta(s)\chi(x,s)ds}d(f_0)_x(y)dt+ \\
&+ \int\int_{(0,1)\times(0,\tau)} \psi(y,t)L\underline{q}_t(y)dydt + \int_0^\tau \Big(\int \partial_t\psi(y,t)d(f_0)_x(y)\Big)dt \\
&= \int\int_{(0,1)\times(0,\tau)} \partial_t\phi(y,t)d\underline{g}_{x,t}(y)dt \\
&- \int\int_{(0,1)\times(0,\tau)} \partial_t\psi(y,t)d(f_0)_x(y)dt \\
&- \int\int_{(0,1)\times(0,\tau)} \eta(t)\chi(x,t)\phi(y,t)d\underline{g}_{x,t}(y)dt+ \\
&+ \int\int_{(0,1)\times(0,\tau)} \Big[\phi(y,t)\eta\chi\partial_y\underline{A}_x(y,t) \\
&\int P(t,\underline{A}_x(\xi,t),\underline{A}_x(y,t))e^{-\int_0^t \eta\chi ds}d\underline{q}_t(\xi)\Big]dydt \\
&+ \int_0^\tau \Big(\int \partial_t\psi(y,t)d(f_0)_x(y)\Big)dt \\
&= \int_0^\tau \Big(\int \phi(y,t)d(F[\underline{g}])_{x,t}(y) + \int \partial_t\phi(y,t)d\underline{g}_{x,t}(y)\Big)dt.
\end{aligned}
$$

Therefore, we have obtained

$$\int \phi(y,\tau)d\underline{g}_{x,\tau}(y) - \int \psi(y,0)d(f_0)_x(y) \\ = \int_0^\tau \Big(\int \phi(y,t)d(F[\underline{g}])_{x,t}(y) + \int \partial_t\phi(y,t)d\underline{g}_{x,t}(y)\Big)dt$$

for all $\tau \in [0,T]$ and for all $\phi(y,t) \in C([0,1]\times[0,T])$ with $\partial_t\phi \in C([0,1]\times[0,T])$. Since $\psi(y,0) = \phi(y,0)$, this implies that $\underline{g}_{x,t}(y)$ satisfies the equation of the system in sense of (4.18):

$$\int \phi(y,\tau)d\underline{g}_{x,\tau}(y) - \int \phi(y,0)d(f_0)_x(y) \\ = \int_0^\tau \Big(\int \phi(y,t)d(F[\underline{g}])_{x,t}(y) + \int \partial_t\phi(y,t)d\underline{g}_{x,t}(y)\Big)dt.$$

□

Let $(\hat{A}, g, u) := (\hat{A}, g, u_1, \dots, u_N) \in \mathcal{X}_{\tau,\rho}$. By Lemma 5.3, $(\hat{A}, g, u)$ uniquely defines a function $\underline{A} \in L^\infty(\Omega; C([0,1]\times[0,\tau];[0,1]))$, and, by Lemma 5.5, $\underline{A}$ uniquely defines a measure $\underline{g} \in \mathcal{L}(\Omega; C([0,T]; X_{[0,1]}))$. Let $\underline{u} := (\underline{u}_1, \dots, \underline{u}_N)$ be a weak solution of the problem

$$\begin{cases} \varepsilon\partial_t\underline{u}_m - d_m\Delta\underline{u}_m = F_m(\underline{A}, \underline{g}, u) \ \ (1 \le m < N), & \text{in } Q_\tau = \Omega\times(0,\tau] \\ \varepsilon\partial_t\underline{u}_N = F_N(\underline{A}, \underline{g}, u) & \text{in } Q_\tau = \Omega\times(0,\tau] \end{cases} \tag{5.15}$$

with initial and boundary conditions

$$\begin{cases} \underline{u}_i(x,0) = u_{0i}(x) & \text{if } x \in \Omega, \\ \partial_n\underline{u}_i(x,t) = 0 & \text{if } x \in \partial\Omega_0, \ \ t>0, \quad (1 \le i \le N). \\ \partial_n\underline{u}_i(x,t) = -\gamma_i\underline{u}_i(x,t) & \text{if } x \in \partial\Omega_1\times(0,\tau]. \end{cases} \tag{5.16}$$

Here we have set

$$\begin{cases} F_1(\underline{A},\underline{g},u) := -u_1 \displaystyle\sum_{j=1}^{N} a_{1,j}u_j - \sigma_1 u_1 + \\ \qquad\qquad + C_{\mathcal{F}} \displaystyle\int_0^1 (\mu_0 + \underline{A}_x(\xi,t))(1-\underline{A}_x(\xi,t))d\underline{g}_{x,t}(\xi), \\ F_m(\underline{A},\underline{g},u) := -u_m \sum_{j=1}^{N} a_{m,j}u_j + \frac{1}{2}\sum_{j=1}^{m-1} a_{j,m-j}u_j u_{m-j} - \sigma_m u_m, \\ F_N(\underline{A},\underline{g},u) := \frac{1}{2} \sum_{\substack{j+k\geq N \\ k,j<N}} a_{j,k}u_j u_k. \end{cases}$$

Notice that, since $(\hat{A},g,u) \in \mathcal{X}_{\tau,\rho}$, $F_i \in L^\infty(\Omega \times [0,\tau])$ $(i=1,\ldots,N)$ and its norm only depends on the compact set $K \subset \mathbb{R}^N$ containing $(u_1,\ldots,u_N)$. We also observe that system (5.15)-(5.16) consists of $N-1$ (uncoupled) scalar linear heat equations with linear boundary conditions and an ordinary differential equation. Therefore it has a unique weak solution $\underline{u}$. More precisely, following [26] and recalling that $\mathcal{X}_{\tau,\rho}$ denotes the closed ball of radius $\rho > 0$ centered at (y, f_0, u_0) in $\mathcal{X}_\tau$, we have the following result.

Proposition 5.6 (see Theorems 2.11, 3.2, and 3.3 in [26])**.** *Let* $(\hat{A},g,u) \in \mathcal{X}_{\tau,\rho}$*. For all* $1 \leq i < N$ *there exists a unique*

$$\underline{u}_i \in C([0,\tau];L^2(\Omega)) \cap L^2([0,\tau];H^1(\Omega))$$

such that

$$d_i \int_0^\tau \left[\int_\Omega \nabla\underline{u}_i(x,t)\cdot\nabla\psi(x,t)dx + \gamma_i \int_{\partial\Omega_1} \underline{u}_i(x,t)\psi(x,t)d\sigma(x)\right] dt$$

$$= \varepsilon \int\int_{Q_\tau} \underline{u}_i \partial_t \psi dxdt + \varepsilon \int_\Omega u_{0i}\psi(x,0)dx + \int\int_{Q_\tau} F_i(\underline{A},\underline{g},u)\psi dxdt$$

for all $\psi \in H^1([0,\tau];H^1(\Omega))$, $\psi(x,\tau) = 0$.
Let $\underline{u}_N(x,t) = u_{0N}(x) + \int_0^\tau F_N(\underline{A},\underline{g},u)ds$ *and* $\underline{u} = (\underline{u}_1,\ldots,\underline{u}_N)$*. Then* $\underline{u} \in C(\overline{Q}_\tau;\mathbb{R}^N)$, $\underline{u}(\cdot,0) = \underline{u}_0$ *and, for* $1 \leq i \leq N$,

$$\|\underline{u}_i\|_{C(\overline{Q}_\tau;\mathbb{R})} \leq C\{\|u_{0i}\|_{L^\infty(\Omega)} + \|F_i\|_{L^r(Q_\tau;\mathbb{R})}\} \text{ if } r > n,\ \ \frac{1}{r} + \frac{n}{2r} < 1.$$

In particular $\|\underline{u}_i\|_{C(\overline{Q}_\tau;\mathbb{R})} \leq C\{\|u_{0i}\|_{C(\overline{\Omega})} + \tau^{\frac{1}{r}}\|F_i\|_{C(\overline{Q}_\tau;\mathbb{R})}\}$.

So, by involving the main results seen so far in this chapter, we are now ready to define the map to which we shall apply a contraction argument. Let $\rho > 0$ be fixed. We know that if $(\hat{A}, g, u) \in \mathcal{X}_{\tau,\rho}$, then we can find uniquely $\underline{A}$ as in Lemma 5.3, $\underline{g}$ as in Lemma 5.5 and $\underline{u}$ as in Proposition 5.6 introduced above. Using this notation, we set

$$\mathcal{H}(\hat{A}, g, u) := (\underline{A}, \underline{g}, \underline{u}) \text{ for } (\hat{A}, g, u) \in \mathcal{X}_{\tau,\rho}. \tag{5.17}$$

Let $\mathcal{T}_d$ denote the metric topology of $\mathcal{X}_{\tau,\rho}$ and $\mathcal{T}$ the weaker topology on $\mathcal{X}_{\tau,\rho}$ which is obtained by endowing $L^\infty(\Omega; C([0,1] \times [0,\tau]; [0,1]))$ with the L^1-topology on $\Omega \times [0,1] \times [0,\tau]$.
In particular, the following proposition ensures that if $\tau > 0$ is sufficiently small, $\mathcal{H}$ is well defined as a map from $\mathcal{X}_{\tau,\rho}$ to $\mathcal{X}_{\tau,\rho}$; moreover, it is a contraction not at the first step on $\mathcal{X}_{\tau,\rho}$ but when we re-apply it to its image $\mathcal{H}(\mathcal{X}_{\tau,\rho})$, where $\mathcal{H}(\mathcal{X}_{\tau,\rho}) \subset \mathcal{X}_{\tau,\rho}$. Finally, $\mathcal{H} : (\mathcal{X}_{\tau,\rho}, \mathcal{T}_d) \to (\mathcal{X}_{\tau,\rho}, \mathcal{T})$ is continuous.

Proposition 5.7. *Let $\rho > 0$ be fixed and let $\mathcal{H}(\hat{A}, g, u)$ be defined by* (5.17). *If $\tau > 0$ is sufficiently small, then $\mathcal{H} : \mathcal{X}_{\tau,\rho} \to \mathcal{X}_{\tau,\rho}$, $(\underline{A}_n, \underline{g}_n, \underline{u}_n) \to (\underline{A}, \underline{g}, \underline{u})$ in $\mathcal{T}$, if $(\hat{A}_n, g_n, u_n) \to (\hat{A}, g, u)$ in $\mathcal{T}_d$, and $\mathcal{H}$ is a contraction on $\mathcal{H}(\mathcal{X}_{\tau,\rho})$.*

Proof. Let us start to prove that $\mathcal{H}(\mathcal{X}_{\tau,\rho}) \subset \mathcal{X}_{\tau,\rho}$ if τ is sufficiently small. We recall that $\mathcal{X}_{\tau,\rho}$ denotes the closed ball in $\mathcal{X}_\tau$ of radius ρ centered at (y, f_0, u_0). By Proposition 5.6, $\|\underline{u}(\cdot,t) - u_0\|_{C(\overline{\Omega};\mathbb{R}^N)} \to 0$ as $t \to 0^+$. This means that $\forall \rho > 0$ there exists $\delta > 0$, such that $\forall t$, $0 < |t| < \delta$ we have $\max_{\overline{\Omega}} \|\underline{u}(\cdot,t) - u_0\|_{\mathbb{R}^N} < \varepsilon$.
Thus, since $\max_{\overline{\Omega}\times[0,\tau]} \|\underline{u}(x,t) - u_0\|_{\mathbb{R}^N} \leq \max_{[0,\tau]}(\max_{\overline{\Omega}} \|\underline{u}(\cdot,t) - u_0\|_{\mathbb{R}^N}) \leq \varepsilon$ if $\tau < \delta$ (for example we can choose $\varepsilon < \frac{\rho}{\sqrt{3}}$, if we want that $(\underline{A}, \underline{g}, \underline{u})$ belongs to the closed ball of radius ρ). So, to prove that $(\underline{A}, \underline{g}, \underline{u})$ belongs to the closed ball $\mathcal{X}_{\tau,\rho}$ for $\tau > 0$ sufficiently small, by arguing as above, it remains to show that, as $t \to 0^+$,

$$\sup_{x\in\Omega, 0\leq y\leq 1} |\underline{A}_x(y,t) - y| \to 0, \quad \sup_{x\in\Omega} \mathcal{W}_1(\underline{g}_{x,t}, (f_0)_x) \to 0. \tag{5.18}$$

By (5.2), recalling that $y \in [0,1]$, $\hat{v}_x(1,t) = 0$ and $(a,t) \mapsto \hat{v}_x(a,t)$ is continuous and Lipschitz continuous w.r.t. $a \in [0,1]$, uniformly in $t \in [0,T]$, as we have seen in the proof of Lemma 5.3 by using assumptions (3.1) and (3.2), we obtain

$$\begin{aligned}
|\underline{A}_x(y,t)-y| &= |\underline{A}_x(y,t)-\underline{A}_x(y,0)| \\
&= \Big|\int_0^t \partial_t \underline{A}_x(y,s)ds\Big| \\
&= \Big|\int_0^t \hat{v}_x(\underline{A}_x(y,s),s)ds\Big| \\
&= \Big|\int_0^t \big(\hat{v}_x(\underline{A}_x(y,s),s)-\hat{v}_x(1,s)\big)ds\Big| \\
&\le C_1\int_0^t |\underline{A}_x(y,s)-1|ds \le C_1\int_0^t \big(|\underline{A}_x(y,s)-y|+|y-1|\big)ds \\
&\le C_1\int_0^t \big(|\underline{A}_x(y,s)-y|+1\big)ds \\
&\le C_1\int_0^t |\underline{A}_x(y,s)-y|ds + C_2\tau.
\end{aligned}$$

Therefore, by Gronwall's Lemma B.17

$$|\underline{A}_x(y,t)-y| \le C_2\tau e^{C_1 t} \to 0$$

as $\tau \to 0^+$ (which also implies $t \to 0^+$). This gives $(5.18)_1$.
As for $(5.18)_2$, it easily follows from Lemma 5.3(i) and its proof: indeed, for a.e. $x \in \Omega$, $t \mapsto \underline{g}_{x,t}$ is the unique solution of the integral equation

$$\underline{g}_{x,t} = (f_0)_x + \int_0^t (F[\underline{g}])_{x,s}ds$$

and it belongs to $C([0,T];X_{[0,1]})$. Thus, by continuity $\underline{g}_{x,t} \to \underline{g}_{x,0} = (f_0)_x$ as $t \to 0^+$ in the Wasserstein topology, i.e. $\mathcal{W}_1(\underline{g}_{x,t},(f_0)_x) \to 0$ as $t \to 0^+$. This concludes the proof and therefore $\mathcal{H}$ maps $\mathcal{X}_{\tau,\rho}$ in $\mathcal{X}_{\tau,\rho}$ since $\mathcal{H}(\mathcal{X}_{\tau,\rho}) \subset \mathcal{X}_{\tau,\rho}$ if τ is sufficiently small, i.e. $\mathcal{H}:\mathcal{X}_{\tau,\rho} \to \mathcal{X}_{\tau,\rho}$.
Now, let us prove the $(\mathcal{T}_d,\mathcal{T})$-continuity of $\mathcal{H}$. Let $\hat{A}_n, \hat{A} \in L^\infty(\Omega; C([0,1]\times[0,\tau];[0,1]))$ be such that $(\hat{A}_n, g_n, u_n) \to (\hat{A}, g, u)$ in $\mathcal{X}_{\tau,\rho}$ as $n \to +\infty$. We have to show that $\underline{A}_n \to \underline{A}$ in $L^1(\Omega\times[0,1]\times[0,\tau])$ as $n \to +\infty$.
Since $|\underline{A}_n| \le 1$ $\forall n$ and $1 \in L^1(\Omega\times[0,1]\times[0,\tau])$, then by the dominated

convergence theorem, $\underline{A}_n \to \underline{A}$ in $L^1(\Omega \times [0,1] \times [0,\tau])$ if we prove that

$$\underline{A}_n \to \underline{A} \text{ a.e. in } \Omega \times [0,1] \times [0,\tau] \text{ as } n \to +\infty. \tag{5.19}$$

To prove (5.19) we observe that by (5.2)

$$\begin{aligned}
&\big|(\underline{A}_n)_x(y,t) - \underline{A}_x(y,t)\big| = \big|(\underline{A}_n)_x(y,t) - y + y - \underline{A}_x(y,t)\big| \\
&= \big|(\underline{A}_n)_x(y,t) - (\underline{A}_n)_x(y,0) + \underline{A}_x(y,0) - \underline{A}_x(y,t)\big| \\
&= \Big| \int_0^t \int \mathcal{G}_x((\underline{A}_n)_x(y,s), (\hat{A}_n)_x(\xi,s)) d(g_n)_{x,s}(\xi) ds \\
&+ \int_0^t \mathcal{S}(x, (\underline{A}_n)_x(y,s), u_n(x,s)) ds \\
&- \int_0^t \int \mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s)) dg_{x,s}(\xi) ds \\
&- \int_0^t \mathcal{S}(x, \underline{A}_x(y,s), u(x,s)) ds \Big| \\
&\leq \int_0^t \Big| \int [\mathcal{G}_x((\underline{A}_n)_x(y,s), (\hat{A}_n)_x(\xi,s)) - \mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s))] d(g_n)_{x,s}(\xi) \Big| ds \\
&+ \int_0^t \Big| \int \mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s)) d(g_n - g)_{x,s}(\xi) \Big| ds \\
&+ \int_0^t \big|\mathcal{S}(x, (\underline{A}_n)_x(y,s), u_n(x,s)) - \mathcal{S}(x, (\underline{A}_n)_x(y,s), u(x,s))\big| ds \\
&+ \int_0^t \big|\mathcal{S}(x, (\underline{A}_n)_x(y,s), u(x,s)) - \mathcal{S}(x, \underline{A}_x(y,s), u(x,s))\big| ds \\
&=: I_1 + I_2 + I_3 + I_4,
\end{aligned} \tag{5.20}$$

where $I_j = I_j(x,y,t)$ for $j = 1,2,3,4$.
Now, since $(\hat{A}_n, g_n, u_n), (\hat{A}, g, u) \in \mathcal{X}_{\tau,\rho}$, we have in particular

$$\max_{\overline{\Omega} \times [0,\tau]} \|u_n - u\|_{\mathbb{R}^N} \leq 2\rho.$$

Therefore, the constant $C(K)$ given by (3.2) depends on ρ, if we consider as compact $K \subset \mathbb{R}^{N-1}$ a closed ball in $\mathbb{R}^{N-1}$ of radius depending on ρ in order to contain u_n, u. So, it follows easily from (3.2) that

$$I_3 \leq C_\rho t \sup_{x \in \Omega, 0 \leq s \leq \tau} |u_n(x,s) - u(x,s)| \leq C_\rho t d((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)).$$

Concerning I_4, always by (3.2) we have

$$I_4 \leq C_\rho \int_0^t |(\underline{A}_n)_x(y,s) - \underline{A}_x(y,s)| ds.$$

By (3.1)

$$\begin{aligned}
&I_1 \leq \\
&\int_0^t \int \Big| \mathcal{G}_x((\underline{A}_n)_x(y,s), (\hat{A}_n)_x(\xi,s)) - \mathcal{G}_x(\underline{A}_x(y,s), (\hat{A}_n)_x(\xi,s)) \Big| d(g_n)_{x,s}(\xi) ds \\
&+ \int_0^t \int \Big| \mathcal{G}_x(\underline{A}_x(y,s), (\hat{A}_n)_x(\xi,s)) - \mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s)) \Big| d(g_n)_{x,s}(\xi) ds \\
&\leq C \int_0^t \int \Big| (\underline{A}_n)_x(y,s) - \underline{A}_x(y,s) \Big| d(g_n)_{x,s}(\xi) ds \\
&+ C \int_0^t \int \Big| (\hat{A}_n)_x(\xi,s) - \hat{A}_x(\xi,s) \Big| d(g_n)_{x,s}(\xi) ds \\
&= C \int_0^t \Big| (\underline{A}_n)_x(y,s) - \underline{A}_x(y,s) \Big| \int d(g_n)_{x,s}(\xi) ds \\
&+ C \int_0^t \int \sup_{x \in \Omega, 0 \leq y \leq 1, 0 \leq s \leq \tau} \Big| (\hat{A}_n)_x(\xi,s) - \hat{A}_x(\xi,s) \Big| d(g_n)_{x,s}(\xi) ds \\
&= C \int_0^t \Big| (\underline{A}_n)_x(y,s) - \underline{A}_x(y,s) \Big| ds + Ct \sup_{x \in \Omega, 0 \leq y \leq 1, 0 \leq s \leq \tau} \Big| (\hat{A}_n)_x(\xi,s) - \hat{A}_x(\xi,s) \Big| \\
&\leq C \int_0^t \Big| (\underline{A}_n)_x(y,s) - \underline{A}_x(y,s) \Big| ds + Ctd((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\Big| (\underline{A}_n)_x(y,t) - \underline{A}_x(y,t) \Big| \leq I_1 + I_2 + I_3 + I_4 \\
&\leq (C + C_\rho) \int_0^t \Big| (\underline{A}_n)_x(y,s) - \underline{A}_x(y,s) \Big| ds + (C + C_\rho) td((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)) \\
&+ I_2(x,y,t) \\
&\leq C'_\rho \int_0^t \Big| (\underline{A}_n)_x(y,s) - \underline{A}_x(y,s) \Big| ds + C'_\rho \tau d((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)) + I_2(x,y,\tau).
\end{aligned}$$

By Gronwall's Lemma B.17

$$\begin{aligned}
&\left|(\underline{A}_n)_x(y,t) - \underline{A}_x(y,t)\right| \\
&\leq \Big(C'_\rho \tau d((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)) + I_2(x,y,\tau)\Big) e^{C'_\rho t} \\
&\leq \Big(C'_\rho \tau d((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)) + I_2(x,y,\tau)\Big) e^{C'_\rho \tau} \to 0 \qquad (5.21)
\end{aligned}$$

as $n \to +\infty$ since $d((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)) \to 0$ and $I_2(x,y,\tau) \to 0$ as $n \to +\infty$ by Proposition 2.17, as we will show below. In particular this proves (5.19) and thus $\underline{A}_n \to \underline{A}$ in $L^1(\Omega \times [0,1] \times [0,\tau])$.
Let us now prove what we have stated above, i.e. that $I_2(x,y,\tau) \to 0$ as $n \to +\infty$. We recall that

$$I_2(x,y,\tau) = \int_0^\tau \Big| \int \mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s)) d(g_n - g)_{x,s}(\xi) \Big| ds.$$

Since

$$\mathcal{W}_1((g_n)_{x,s}, g_{x,s}) \leq \sup_{x\in\Omega} \max_{0\leq t\leq\tau} \mathcal{W}_1((g_n)_{x,t}, g_{x,t}) \leq d((\hat{A}_n, g_n, u_n), (\hat{A}, g, u)),$$

we have $\mathcal{W}_1((g_n)_{x,s}, g_{x,s}) \to 0$ as $n \to +\infty$. Therefore, by Proposition 2.17 $(g_n)_{x,s} \to g_{x,s}$ weakly* as $n \to +\infty$ and thus

$$\int \mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s)) d(g_n - g)_{x,s}(\xi) \to 0$$

as $n \to +\infty$.
Moreover,

$$\begin{aligned}
&\Big| \int \mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s)) d(g_n - g)_{x,s}(\xi) \Big| \\
&\leq \int \Big|\mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s))\Big| d(g_n)_{x,s}(\xi) + \int \Big|\mathcal{G}_x(\underline{A}_x(y,s), \hat{A}_x(\xi,s))\Big| dg_{x,s}(\xi) \\
&\leq 2C < +\infty.
\end{aligned}$$

Hence by the dominated convergence theorem $I_2(x,y,\tau) \to 0$ as $n \to +\infty$.
Notice that to complete the proof of the $(\mathcal{T}_d, \mathcal{T})$-continuity of the map $\mathcal{H}$ it remains to prove that $\underline{g}_n \to \underline{g}$ in $\mathcal{L}(\Omega; C([0,\tau]; X_{[0,1]}))$ and $\underline{u}_n \to \underline{u}$ in

$C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N)$. This will be done in the sequel, when we prove that $\mathcal{H}$ is a contraction.

In the remainder of the proof, we show that $\mathcal{H}$ is a contraction on $\mathcal{H}(\mathcal{X}_{\tau,\rho})$ if τ is small enough. Let $(\hat{A}^1,g^1,u^1),(\hat{A}^2,g^2,u^2)\in\mathcal{H}(\mathcal{X}_{\tau,\rho})$. Repeating the same arguments leading to (5.21), we attain that

$$
\begin{aligned}
&|\underline{A}^1_x(y,t)-\underline{A}^2_x(y,t)|\\
&\quad\le\Big(C'_\rho\tau d((\hat{A}^1,g^1,u^1),(\hat{A}^2,g^2,u^2))+\\
&\qquad+\int_0^\tau\Big|\int\mathcal{G}_x(\underline{A}^2_x(y,s),\hat{A}^2_x(\xi,s))d(g^1-g^2)_{x,s}(\xi)\Big|ds\Big)e^{C'_\rho\tau}.
\end{aligned}
$$

Since $(\hat{A}^2,g^2,u^2)\in\mathcal{H}(\mathcal{X}_{\tau,\rho})$, it follows from Remark 5.4 and (3.1) that $\hat{A}^2_x(\xi,s)$ and $\mathcal{G}_x(\underline{A}^2_x(y,s),\hat{A}^2_x(\xi,s))$ are Lipschitz continuous in ξ, uniformly w.r.t. x and s. Therefore, by Proposition A.3,

$$
\begin{aligned}
&|\underline{A}^1_x(y,t)-\underline{A}^2_x(y,t)|\\
&\le\Big(C'_\rho\tau d((\hat{A}^1,g^1,u^1),(\hat{A}^2,g^2,u^2))+C_\rho\int_0^\tau\mathcal{W}_1(g^1_{x,s},g^2_{x,s})ds\Big)e^{C'_\rho\tau}\\
&\le e^{C'_\rho\tau}(C'_\rho+C_\rho)\tau d((\hat{A}^1,g^1,u^1),(\hat{A}^2,g^2,u^2)),
\end{aligned}
$$

whence

$$\sup_{x,y,t}|\underline{A}^1_x(y,t)-\underline{A}^2_x(y,t)|\le e^{C'_\rho\tau}(C'_\rho+C_\rho)\tau d((\hat{A}^1,g^1,u^1),(\hat{A}^2,g^2,u^2)).\quad(5.22)$$

Let us pass to consider now $\mathcal{W}_1(\underline{g}^1_{x,t},\underline{g}^2_{x,t})$. In view of the definition of $\underline{g}^1,\underline{g}^2$, we may reproduce the arguments in the proof of Lemma 5.5 and obtain that

$$
\begin{aligned}
\mathcal{W}_1(\underline{g}^1_{x,t},\underline{g}^2_{x,t})&\le C\max_{[0,T]}\eta\ t\sup_{0\le s\le t}\mathcal{W}_1(g^1_{x,s},g^2_{x,s}) &(5.23)\\
&\le C\tau d((\hat{A}^1,g^1,u^1),(\hat{A}^2,g^2,u^2)). &(5.24)
\end{aligned}
$$

Finally we have to estimate the third component $\max_{\overline{\Omega}\times[0,\tau]}|\underline{u}^1-\underline{u}^2|$. Set $\underline{U}=\underline{u}^1-\underline{u}^2$ and $\underline{U}=(\underline{U}_1,\dots,\underline{U}_N)$. Then, $\underline{U}$ is a weak solution in the sense of Proposition 5.6 of a system similar to (5.15)-(5.16) with F_j replaced by $\tilde{F}_j:=$

$F_j(\underline{A}^1, \underline{g}^1, u^1) - F_j(\underline{A}^2, \underline{g}^2, u^2)$, $j = 1, \ldots, N$, and u_0 replaced by $\underline{U}(x, 0) = 0$. By Proposition 5.6,

$$\|\underline{u}^1 - \underline{u}^2\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N)} = \|\underline{U}\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N)} \leq \sum_{i=1}^{N} \|\underline{U}_i\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R})}$$
$$\leq C\tau^{\frac{1}{r}} \sum_{i=1}^{N} \|\tilde{F}_i\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R})}. \quad (5.25)$$

If $k > 1$, F_k is a polynomial in the components of u and, since u_1, u_2 are uniformly bounded by a constant which depends on ρ in $\overline{\Omega} \times [0, \tau]$,

$$\|\tilde{F}_k\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R})} \leq C_\rho \sum_i \|u_i^1 - u_i^2\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R})} \text{ if } k > 1,$$

and hence

$$\|\tilde{F}_k\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R})} \leq C_\rho d((\hat{A}^1, g^1, u^1), (\hat{A}^2, g^2, u^2)) \text{ if } k > 1. \quad (5.26)$$

The same argument applies to the polynomial terms of $\tilde{F}_1$, so if we call

$$I := \int_0^1 (\mu_0 + \underline{A}_x^1(\xi, t))(1 - \underline{A}_x^1(\xi, t)) d\underline{g}_{x,t}^1(\xi)$$
$$- \int_0^1 (\mu_0 + \underline{A}_x^2(\xi, t))(1 - \underline{A}_x^2(\xi, t)) d\underline{g}_{x,t}^2(\xi),$$

we have to estimate $\|I\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R})}$. Arguing as above,

$$|I| \leq J_1 + J_2$$
$$:= \int_0^1 |(\mu_0 + \underline{A}_x^1(\xi, t))(1 - \underline{A}_x^1(\xi, t)) - (\mu_0 + \underline{A}_x^2(\xi, t))(1 - \underline{A}_x^2(\xi, t)) d\underline{g}_{x,t}^1(\xi)$$
$$+ \Big| \int_0^1 (\mu_0 + \underline{A}_x^2(\xi, t))(1 - \underline{A}_x^2(\xi, t)) d(\underline{g}_{x,t}^1 - \underline{g}_{x,t}^2)(\xi) \Big|.$$

By a straightward calculation and by (5.22)

$$J_1 = \int_0^1 \left| \left(\underline{A}_x^2(\xi, t) - \underline{A}_x^1(\xi, t)\right)\left((\mu_0 - 1) + \underline{A}_x^2(\xi, t) + \underline{A}_x^1(\xi, t)\right) \right| d\underline{g}_{x,t}^1(\xi)$$
$$\leq \int_0^1 \left|\underline{A}_x^2(\xi, t) - \underline{A}_x^1(\xi, t)\right| \left(|\mu_0 - 1| + \left|\underline{A}_x^2(\xi, t) + \underline{A}_x^1(\xi, t)\right|\right) d\underline{g}_{x,t}^1(\xi)$$
$$\leq (2 + |\mu_0 - 1|) \sup_{x,\xi,t} \left|\underline{A}_x^2(\xi, t) - \underline{A}_x^1(\xi, t)\right| \int_0^1 d\underline{g}_{x,t}^1(\xi)$$
$$\leq C e^{C'_\rho \tau} (C'_\rho + C_\rho) \tau d((\hat{A}^1, g^1, u^1), (\hat{A}^2, g^2, u^2)).$$

Concerning J_2, by Remark 5.4, the map $\xi \mapsto (\mu_0 + \underline{A}_x^2(\xi,t))(1-\underline{A}_x^2(\xi,t))$ is Lipschitz continuous w.r.t. ξ, uniformly w.r.t. x and t. Thus, by Proposition A.3 and by (5.24),

$$J_2 \leq C_\rho \mathcal{W}_1(\underline{g}^1_{x,t}, \underline{g}^2_{x,t}) \leq C_\rho \tau d((\hat{A}^1, g^1, u^1), (\hat{A}^2, g^2, u^2)),$$

so that

$$\|\tilde{F}_1\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R})} \leq C_\rho \tau d((\hat{A}^1, g^1, u^1), (\hat{A}^2, g^2, u^2)) + \\ + C_\rho d((\hat{A}^1, g^1, u^1), (\hat{A}^2, g^2, u^2)).$$

Combining the last estimate with (5.26) and finally with (5.25), we obtain that

$$\|\underline{u}^1 - \underline{u}^2\|_{C(\overline{\Omega}\times[0,\tau];\mathbb{R}^N)} \leq C_\rho \tau^{\frac{1}{r}} d((\hat{A}^1, g^1, u^1), (\hat{A}^2, g^2, u^2)). \tag{5.27}$$

It follows from (5.22), (5.24) and (5.27) that $\mathcal{H}$ is a contraction on $\mathcal{H}(\mathcal{X}_{\tau,\rho})$ if τ is small enough.
This also concludes the proof of the $(\mathcal{T}_d, \mathcal{T})$-continuity of $\mathcal{H}$, since it follows that $\underline{g}_n \to \underline{g}$ and $\underline{u}_n \to \underline{u}$. □

At this point we have the right tools to complete the proof of Theorem 5.1. In order to do this, we need a minor modification of the classical Banach-Caccioppoli fixed point theorem.

Proposition 5.8. *Let (X,d) be a complete metric space and let $\mathcal{T}_d$ be the topology induced by d. Let $\mathcal{T}$ be a Hausdorff topology on X which is weaker than $\mathcal{T}_d$. If $\mathcal{H} : X \to X$ is a contraction on $\mathcal{H}(X)$ which is $(\mathcal{T}_d, \mathcal{T})$-continuous, then $\mathcal{H}$ has a unique fixed point.*

Proof. We start carrying out the standard iteration procedure, by defining the iterative sequence

$$x_{n+1} = \mathcal{H}(x_n), \tag{5.28}$$

starting from a point $x_0 \in \mathcal{H}(X)$, so that $x_n \in \mathcal{H}(X)$ for all $n \geq 0$. By a standard argument, it is a Cauchy sequence. Thus, by the completeness of

(X, d), we may assume that $x_n \to \bar{x} \in X$ as $n \to +\infty$. When $\mathcal{H}$ is a contraction on all points of X, and, hence, in particular is Lipschitz continuous from X to X, we can conclude the proof taking the limit as $n \to +\infty$ in (5.28). In our case this argument has to be slightly adapted since $\mathcal{H}$ is a contraction only on $\mathcal{H}(X)$. But since $\mathcal{H}$ is $(\mathcal{T}_d, \mathcal{T})$-continuous in X, the following holds: on one side $x_{n+1} \to \bar{x}$ as $n \to +\infty$ with respect to the topology $\mathcal{T}_d$ and thus also with respect to the topology $\mathcal{T}$, since it is weaker than $\mathcal{T}_d$; on the other hand $\mathcal{H}(x_n) \to \mathcal{H}(\bar{x})$ as $n \to +\infty$ w.r.t. the topology $\mathcal{T}$. Therefore, by (5.28) and by the uniqueness of the limit in $\mathcal{T}$, we can conclude that $\bar{x} = \mathcal{H}(\bar{x})$. □

Proof of Theorem 5.1. By Proposition 5.7 and the fixed point theorem (Proposition 5.8), the problem (4.16)-(4.17) has a unique solution in the sense of Definition 4.2 in $\Omega \times [0, \tau]$, for τ small enough, if we show the nonnegativity of u_i:

$$u_i \geq 0 \text{ in } \Omega \times [0, \tau] \quad (i = 1, \dots, N). \tag{5.29}$$

If we suppose to have already shown that $u_i \geq 0$ for $i = 1, \dots, N-1$, then for $i = N$, (5.29) is trivially satisfied.
For $i = 1, \dots, N-1$, (5.29) formally follows from the maximum principle. Let we make this precise if $i = 1$. If $1 < i < N-1$ the proof is even easier, since the $i - th$ equation of the system (4.16) has not the integral term.

Since $f = C_{\mathcal{F}} \int_0^1 (\mu_0 + A_x(\xi, t))(1 - A_x(\xi, t)) dg_{x,t}(\xi)$ is nonnegative and belongs to $L^\infty(Q_\tau)$, there exists a sequence of smooth nonnegative functions $(f_k)_{k\in\mathbb{N}}$ such that $f_k \to f$ as $k \to +\infty$ in $L^r(Q_\tau)$, where $r > n$ and $\frac{1}{r} + \frac{n}{2r} < 1$. We can also approximate $h = \sum_{j=1}^{N} a_{1,j} u_j + \sigma_1 \in C(\overline{Q}_\tau)$ uniformly by smooth functions h_k.
Now, let v_k be the unique smooth solution of

$$\begin{cases} \varepsilon \partial_t v_k = d_1 \Delta v_k - v_k h_k + f_k & \text{in } Q_\tau, \\ v_k(x, 0) = u_{01}(x) & \text{if } x \in \Omega, \\ \partial_n v_k(x, t) = 0 & \text{if } x \in \partial\Omega_0, t > 0, \\ \partial_n v_k(x, t) = -\gamma_1 v_k(x, t) & \text{if } x \in \partial\Omega_1, t > 0. \end{cases} \tag{5.30}$$

Since $\gamma_1 > 0$, $f_k \geq 0$ in Q_τ, and $u_{01} \geq 0$ in Ω, it follows from the maximum principle that $v_k \geq 0$ in Q_τ (see [1]).

On the other hand, $w_k := u_1 - v_k$ is a weak solution of

$$\begin{cases} \varepsilon\partial_t w_k = d_1\Delta w_k - w_k h_k - u_1(h - h_k) + f - f_k & \text{in } Q_\tau, \\ w_k(x,0) = 0 & \text{if } x \in \Omega, \\ \partial_n w_k(x,t) = 0 & \text{if } x \in \partial\Omega_0, t > 0, \\ \partial_n w_k(x,t) = -\gamma_1 w_k(x,t) & \text{if } x \in \partial\Omega_1, t > 0. \end{cases} \tag{5.31}$$

So it follows from Theorem 3.2 in [26] that $w_k \to 0$ as $k \to +\infty$ uniformly on $\overline{Q}_\tau$ and therefore $v_k \to u_1$ as $k \to +\infty$ uniformly on $\overline{Q}_\tau$. Thus also $u_1 \geq 0$ in Q_τ. □

5.2 Global existence

In this section we complete the proof of Theorem 3.6. We recall that problems (4.16)-(4.17) and (8)-(9) are equivalent, as we have proved in Chapter 4. So, if we show that the local solution of problem (4.16)-(4.17), whose uniqueness and local existence has been proved in the previous section, can be continued to the whole interval $[0, T]$, then we will complete the proof of Theorem 3.6.

In order to show the global existence (with respect to t) of the solution, we argue by contradiction. Let (A, g, u) be the local solution of (4.16)-(4.17) and let us suppose that the maximal interval of existence is $[0, \tau^*)$ for some $\tau^* < T$.

First, we provide an a priori estimate for $u(x, t)$.

Since

$$C_{\mathcal{F}} \int_0^1 (\mu_0 + A_x(\xi, t))(1 - A_x(\xi, t)) dg_{x,t}(\xi) \leq C_1 \text{ in } \Omega \times [0, \tau^*) \tag{5.32}$$

for some constant C_1, it follows formally from the maximum principle that

$$u_1(x,t) \leq \sup_\Omega u_{01} + C_1 t \text{ for } x \in \Omega, 0 \leq t < \tau^*.$$

Similarly, if $u_1, \dots, u_{m-1}$ are bounded in $L^\infty(\Omega\times[0,\tau^*))$ for some $1 < m < N$ then,

$$\frac{1}{2}\sum_{j=1}^{m-1} a_{j,m-j}u_j u_{m-j} \leq C_m \text{ in } \Omega\times[0,\tau^*)$$

for some constant C_m, and it follows formally from the maximum principle that

$$u_m(x,t) \leq \sup_\Omega u_{0m} + C_m t \text{ for } x\in\Omega, 0\leq t<\tau^*.$$

In both cases the use of the maximum principle is justified as in the proof of (5.29). Concerning the boundedness of u_N in $\Omega\times[0,\tau^*)$, it follows from that of $u_1, \dots, u_{N-1}$, so we have obtained that, for some $C_u > 0$,

$$|u| \leq C_u \text{ in } \Omega\times[0,\tau^*). \tag{5.33}$$

This a priori estimate for $u(x,t)$ allows us to show the existence of

$$\lim_{t\to\tau^*} u(x,t) =: u(x,\tau^*).$$

Indeed, in view of (5.32) and (5.33), it follows from standard regularity theory for weak solutions of parabolic equations that u is uniformly (Hölder) continuous in $\Omega\times[0,\tau^*)$ (see, for example, Theorem 1, p.111 in [30]). Hence u can be extended to $\Omega\times[0,\tau^*]$ as a continuous function.
Let us pass to show the existence of

$$\lim_{t\to\tau^*} A_x(y,t) =: A_x(y,\tau^*).$$

Arguing as in the proof of Lemma 5.3 we obtain that $A_x(y,t)$ and $\hat{v}_x(A_x(y,t),t)$ are Lipshitz continuous with respect to y, uniformly with respect to $x\in\Omega$ and $t\in[0,\tau^*)$. By the boundedness of $\hat{v}_x(A_x(y,t),t)$, the map $t\mapsto A_x(y,t)$ is Lipschitz continuous on $[0,\tau^*)$. Hence $A_x(y,\tau^*) := \lim_{t\to\tau^*} A_x(y,t)$ exists and is Lipschitz continuous with respect to y, uniformly with respect to $x\in\Omega$.

Finally, we can affirm that there exists

$$\lim_{t\to\tau^*} g_{x,t} =: g_{x,\tau^*}.$$

Indeed, we can reproduce the arguments of the proof of Lemma 5.5 and we obtain that the map $t \mapsto g_{x,t}$ is Lipschitz continuous from $[0, \tau^*)$ to $X_{[0,1]}$, which, we recall, is endowed with the Wasserstein distance $\mathcal{W}_1$.

Hence, if we consider the functions g_{x,τ^*} and $u(x, \tau^*)$ as the given initial data, then the local existence theorem provides a solution in $[\tau^*, \tau_1]$, for some $\tau_1 \in [\tau^*, T]$. Therefore $[0, \tau^*)$ is not the maximal interval of existence.
So, we have found a contradiction.

This proves the global existence of the solution and completes the proof of Theorem 3.6.

Appendix A

Narrow Convergence and the Wasserstein Distance: some useful tools

The main purpose of this Appendix is helping the reader in the comprehension of all the steps of this thesis. Indeed we cite some useful results which we have used in this work, in order to adapt more or less well-known results in literature to our modeling needs for Alzheimer's disease. For proofs and further details we refer to [3].
Throughout this Appendix, X denotes a complete separable and locally compact metric space with metric d, unless other specifications. We remind that a positive Borel measure μ on X such that $\mu(X) = 1$ is called probability measure and we write $\mu \in \mathcal{P}(X)$. Moreover, we denote by $\mathcal{P}_p(X)$ the following subset of Borel probability measures on X

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) : \int_X d(x, \bar{x})^p d\mu(x) < +\infty \text{ for some } \bar{x} \in X \right\}.$$

Let us start introducing a useful Lemma which allows to prove the triangle inequality for the Wasserstein distance.

Lemma A.1 (See [3], Lemma 5.3.2)**.** *Let X_1, X_2, X_3 be complete separable and locally compact metric spaces and let $\gamma^{12} \in \mathcal{P}(X_1 \times X_2)$, $\gamma^{13} \in \mathcal{P}(X_1 \times X_3)$*

such that $\pi^1_\sharp\gamma^{12} = \pi^1_\sharp\gamma^{13} = \mu^1$*. Then there exists* $\boldsymbol{\mu} \in \mathcal{P}(X_1 \times X_2 \times X_3)$ *such that*

$$\pi^{1,2}_\sharp\boldsymbol{\mu} = \gamma^{12}, \pi^{1,3}_\sharp\boldsymbol{\mu} = \gamma^{13}. \tag{A.1}$$

We denote by $\Gamma^1(\gamma^{12}, \gamma^{13})$ *the subset of plans* $\boldsymbol{\mu} \in \mathcal{P}(X_1 \times X_2 \times X_3)$ *satisfying* (A.1).

Remark A.2. (See [3], Remark 5.3.3) A similar situation occurs when $\gamma^{12} \in \mathcal{P}(X_1 \times X_2)$ and $\gamma^{23} \in \mathcal{P}(X_2 \times X_3)$. Now we say that $\mu \in \Gamma^2(\gamma^{12}, \gamma^{23})$ if $\pi^{1,2}_\sharp\boldsymbol{\mu} = \gamma^{12}, \pi^{2,3}_\sharp\boldsymbol{\mu} = \gamma^{23}$. Notice that $\Gamma^2(\gamma^{12}, \gamma^{23})$ is not empty if and only if $\pi^2_\sharp\gamma^{12} = \pi^2_\sharp\gamma^{23}$. Indeed, this follows by using $\pi^i_\sharp(\pi^{i,j}_\sharp\boldsymbol{\mu}) = (\pi^i \circ \pi^{i,j})_\sharp\boldsymbol{\mu}$ and the definition of $\Gamma^2(\gamma^{12}, \gamma^{23})$. In this case there exists a measure $\boldsymbol{\mu} \in \Gamma^2(\gamma^{12}, \gamma^{23})$ constructed as in the proof of the previous lemma (see [3]) such that $\pi^{1,3}_\sharp\boldsymbol{\mu}$ belongs by construction to $\Gamma(\mu^1, \mu^3)$.

In Chapter 2 we have given the definition of the p-th Wasserstein distance between two probability measures μ and ν involving all transport plans between them. However, when $p = 1$ there is a particular case which allows us to have a characterization of the Wasserstein distance with a sort of duality form. This is expressed by the following Proposition.

Proposition A.3 (**Kantorovich-Rubinstein duality**; cf. equation (7.1.2) of [3])**.** *If* $\mu, \nu \in \mathcal{P}_1(X)$ *have compact support, then*

$$\mathcal{W}_1(\mu, \nu) = \sup\left\{\int_X \Phi d(\mu - \nu) : \Phi \in \mathrm{Lip}_1(X, \mathbb{R})\right\},$$

where $\mathrm{Lip}_1(X, \mathbb{R})$ *is the space of Lipschitz continuous functions* $\Phi : X \longrightarrow \mathbb{R}$ *with Lipschitz constant not greater than 1.*

We consider now the space $\mathcal{P}_p(X)$ endowed with the p-th Wasserstein distance: we are interested in investigating its compactness and completeness and the relation between the Wasserstein convergence and the narrow convergence. About this we cite the following result.

Proposition A.4 (See [3], Proposition 7.1.5)**.** $\mathcal{P}_p(X)$ *endowed with the p-Wasserstein distance is a separable metric space which is complete if X is* *complete. A set $\mathcal{K} \subset \mathcal{P}_p(X)$ is relatively compact if and only if it is p-uniformly integrable and tight. In particular, for a given sequence $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{P}_p(X)$ we have*

$$\lim_{n\to+\infty} \mathcal{W}_p(\mu_n,\mu) = 0 \iff \begin{cases} \mu_n \longrightarrow \mu \ \text{ narrowly as } \ n \to +\infty, \\ (\mu_n)_{n\in\mathbb{N}} \ \text{has uniformly integrable } p\text{−moments.} \end{cases}$$

The property of tightness of a set of probability measures is often requested. So, in order to check if a set $\mathcal{K} \subset \mathcal{P}(X)$ is tight, we can use the following result which provides an integral condition for this property.

Proposition A.5 (See [3], Remark 5.1.5)**.** *Let X be a locally compact separable metric space and $\mathcal{K} \subset \mathcal{P}(X)$. If there exists a function $\phi : X \longrightarrow [0,+\infty]$, whose sublevels $\{x \in X : \phi(x) \le c\}$ are compact in X, such that*

$$\sup_{\mu\in\mathcal{K}} \int_X \phi(x) d\mu(x) < +\infty$$

then $\mathcal{K}$ is tight.

Proof. It is sufficient to apply Chebichev inequality (see Proposition A.6) which shows that the tightness condition in Definition 2.3 is satisfied by the family of sublevels of ϕ. Indeed, we have

$$\frac{1}{t}\int_X \phi d\mu \le \frac{1}{t}\sup_{\mu\in\mathcal{K}}\int_X \phi d\mu \longrightarrow 0$$

as $t \to +\infty$, since $\sup_{\mu\in\mathcal{K}} \int_X \phi(x)d\mu(x) < +\infty$. This means that $\forall\varepsilon > 0 \ \exists M_\varepsilon > 0$ such that $\forall t \ge M_\varepsilon$

$$\frac{1}{t}\sup_{\mu\in\mathcal{K}}\int_X \phi d\mu \le \varepsilon.$$

Thus, by posing $t = t_\varepsilon = M_\varepsilon$ and by using Chebichev inequality,

$$\begin{aligned} \mu(X \setminus \{x \in X; \phi(x) \le t_\varepsilon\}) &= \mu(\{x \in X; \phi(x) > t_\varepsilon\}) \\ &\le \frac{1}{t_\varepsilon}\int_X \phi d\mu \\ &\le \varepsilon. \end{aligned}$$

Thus, if we take $K_\varepsilon := \{x \in X; \phi(x) \le t_\varepsilon\}$, it is compact by hypothesis and it satisfies the tightness condition. ∎

Obviously, if X is compact, every $\mathcal{K} \subset \mathcal{P}(X)$ is tight, by choosing $\phi \equiv 1$ on X.

Proposition A.6 (**Chebichev inequality**, see [2], Remark 1.18)**.** *If $f : X \longrightarrow [0, +\infty]$, f μ-measurable is such that $\int_X f d\mu < +\infty$, then for any $t > 0$*

$$\mu(\{x \in X : f(x) > t\}) \leq \frac{1}{t} \int_X f d\mu.$$

Appendix B

Some auxiliary theorems

Let us recall some notations: in a metric space (X, d), the distance of a point $x \in X$ from a set K is defined by

$$d(x, K) := \inf_{y \in K} d(x, y).$$

Moreover, the distance between two sets $A, B \subset X$ is defined by

$$d(A, B) := \inf_{x \in A, y \in B} d(x, y).$$

Theorem B.1. *Let (X, d) be a metric space and let $K \subseteq X$ be compact. If $f : K \longrightarrow \mathbb{R}$ is continuous, then f must attain a maximum and a minimum.*

Proposition B.2. *Let a_n and b_n be real bounded sequences. Then*

$$\limsup_{n \to +\infty}(a_n + b_n) \geq \limsup_{n \to +\infty} a_n + \liminf_{n \to +\infty} b_n \tag{B.1}$$

$$\liminf_{n \to +\infty}(a_n + b_n) \leq \liminf_{n \to +\infty} a_n + \limsup_{n \to +\infty} b_n \tag{B.2}$$

$$\liminf_{n \to +\infty}(a_n + b_n) \geq \liminf_{n \to +\infty} a_n + \liminf_{n \to +\infty} b_n \tag{B.3}$$

In particular, if one of the sequences converges, then the equality holds in (B.3)*.*

Proof. Let us start to prove the first inequality. We have, by definition of sup and inf:

$$\sup_{k \geq n}(a_k + b_k) \geq \sup_{k \geq n} a_k + \inf_{k \geq n} b_k$$

$$\lim_{n\to+\infty}\sup_{k\geq n}(a_k+b_k)\geq\lim_{n\to+\infty}\sup_{k\geq n}a_k+\lim_{n\to+\infty}\inf_{k\geq n}b_k$$

Hence by definition of lim sup and lim inf (i.e.$\limsup_{n\to+\infty}a_n := \lim_{n\to+\infty}\sup_{k\geq n}a_n$ and $\liminf_{n\to+\infty}a_n := \lim_{n\to+\infty}\inf_{k\geq n}a_n$):

$$\limsup_{n\to+\infty}(a_n+b_n)\geq\limsup_{n\to+\infty}a_n+\liminf_{n\to+\infty}b_n.$$

By replacing in the previous inequality a_n and b_n with $-a_n$ and $-b_n$ and using $\limsup_{n\to+\infty}(-a_n) = -\liminf_{n\to+\infty}a_n$ and $\liminf_{n\to+\infty}(-a_n) = -\limsup_{n\to+\infty}a_n$, we obtain the second inequality.

Finally, the last inequality follows from the definition of lim inf. □

Proposition B.3. *Let (X_1,d_1), (X_2,d_2) be separable metric spaces, X_2 locally compact and let $\phi_n : X_1 \longrightarrow X_2$ be Borel measurable functions uniformly converging to ϕ on compact subsets of X_1. If $\phi : X_1 \longrightarrow X_2$ is continuous, then for any compact set $K\subset X_1$ the uniform convergence of ϕ_n to ϕ on K implies the uniform convergence of $f\circ\phi_n$ to $f\circ\phi$ on K, where $f : X_2 \longrightarrow \mathbb{R}$ is a continuous function.*

Proof. Let $K\subset X_1$ be compact. Since ϕ is continuous, $\phi(K)\subset X_2$ is compact. Let U be a neighborhood of $\phi(K)$, for example $U = \bigcup_{y\in\phi(K)} B_{d_2}(y,r)$ with $r>0$ and B_{d_2} an open ball respect to the metric d_2 of radius r and centered at y. Since X_2 is locally compact, by Proposition B.8 there exists W compact such that $\phi(K)\subset W\subset U\subset X_2$.

We consider now $0<\varepsilon<d_2(\phi(K),\partial W)$, where ∂W denotes the boundary of the set. Moreover, we know that ϕ_n converges uniformly to ϕ on K, i.e. $\forall\varepsilon>0$ (and thus also for $0<\varepsilon<d_2(\phi(K),\partial W)$ fixed) $\exists n_\varepsilon\in\mathbb{N}$ such that

$$\sup_{x\in K} d_2(\phi_n(x),\phi(x))\leq\varepsilon \qquad \forall n\geq n_\varepsilon$$

Therefore,

$$d_2(\phi_n(x),\phi(K)) = \inf_{y\in\phi(K)} d_2(\phi_n(x),y)\leq d_2(\phi_n(x),\phi(x))\leq\varepsilon$$

$\forall n \geq n_\varepsilon, \forall x \in K$. This means that

$$\phi_n(K) \subset \{y \in X_2; d_2(y, \phi(K)) \leq \varepsilon\} \subset W \qquad \forall n \geq n_\varepsilon.$$

Now, we remember that f is continuous, thus by Heine Cantor it is uniformly continuous on W, since W is compact, i.e. $\forall \bar{\varepsilon} > 0\ \exists \delta > 0$ such that $\forall x, y \in W$, if $d_2(x, y) < \delta$

$$|f(x) - f(y)| < \bar{\varepsilon}.$$

Now, if we take $\varepsilon < \min(d_2(\phi(K), \partial W), \delta)$, by the uniform convergence

$$d_2(\phi_n(x), \phi(x)) < \varepsilon < \delta \qquad \forall n \geq n_\varepsilon \qquad \forall x \in K.$$

Hence by uniform continuity of f we have proved that $\forall \bar{\varepsilon} > 0\ \exists \bar{n} = n_\varepsilon \in \mathbb{N}$ such that

$$|f(\phi_n(x)) - f(\phi(x))| < \bar{\varepsilon} \qquad \forall n \geq \bar{n} \qquad \forall x \in K$$

that is the definition of uniform convergence of $f \circ \phi_n$ to $f \circ \phi$. □

Theorem B.4. *Let X, Y be topological spaces with a countable basis of open sets. Then the Borel σ-algebra of the topological product space $X \times Y$ coincides with the product σ-algebra*

$$\mathcal{B}(X \times Y) = \mathcal{B}(X) \times \mathcal{B}(Y).$$

Proof. The product σ-algebra contains the rectangles $U \times V$ whenever U and V vary in the basis of open sets of the first and the second space respectively. Indeed, $U \times V = (\pi^1)^{-1}(U) \cap (\pi^2)^{-1}(V) \in \mathcal{B}(X) \times \mathcal{B}(Y)$. In addition, they form a countable basis for the open sets of the topological product $X \times Y$, thus they generate $\mathcal{B}(X \times Y)$. Hence $\mathcal{B}(X) \times \mathcal{B}(Y) \supseteq \mathcal{B}(X \times Y)$. The opposite inclusion follows from definition of the product topology and the product σ-algebra. □

Theorem B.5 (Fubini). *Let $(X_1, \mathcal{E}_1)$, $(X_2, \mathcal{E}_2)$ be measurable spaces and μ_1, μ_2 be positive σ-finite measures in X_1, X_2 respectively. Then, there is a unique positive σ-finite measure μ on $(X_1 \times X_2, \mathcal{E}_1 \times \mathcal{E}_2)$ such that*

$$\mu(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2) \ \ \forall E_1 \in \mathcal{E}_1 \ \ \forall E_2 \in \mathcal{E}_2.$$

Furthermore, for any μ-measurable function $u : X_1 \times X_2 \to [0, +\infty]$ we have that

$$x \mapsto \int_{X_2} u(x,y)d\mu_2(y) \text{ and } y \mapsto \int_{X_1} u(x,y)d\mu_1(x)$$

are respectively μ_1-measurable and μ_2-measurable and

$$\begin{aligned}\int_{X_1\times X_2} ud\mu &= \int_{X_1}\left(\int_{X_2} u(x,y)d\mu_2(y)\right)d\mu_1(x)\\ &= \int_{X_2}\left(\int_{X_1} u(x,y)d\mu_1(x)\right)d\mu_2(y).\end{aligned}$$

Proposition B.6 (**Coincidence criterion**, see [2], Proposition 1.8)**.** *Let μ, ν be positive measures on the measure space $(X, \mathcal{E})$, and let $\mathcal{G} \subset \mathcal{E}$ be a family closed under finite intersection; assume that $\mu(E) = \nu(E)$ for every $E \in \mathcal{G}$, and that there exists a sequence (X_h) in $\mathcal{G}$ such that $X = \bigcup_h X_h$ and $\mu(X_h) = \nu(X_h) < +\infty$ for any h. Then μ and ν coincide on the σ-algebra generated by $\mathcal{G}$.*

We prove now some very useful results for a locally compact topological space that we have often used during this thesis in order to have compact sets. We recall that a Hausdorff topological space X is said locally compact if every point $x \in X$ has a compact neighborhood.

Proposition B.7 (See Corollaire Proposition 9 in [8])**.** *Let X be a locally compact topological space, every point of X has a fundamental system of compact neighborhoods.*

This property can be generalized as follows.

Proposition B.8 (See [8], Proposition 10.)**.** *Let X be a locally compact topological space, every compact set K has a fundamental system of compact neighborhoods.*

Proof. Let U be any neighborhood of K: $\forall x \in K$ there exists a compact neighborhood $W(x)$ of x contained in U. The interiors of the sets $W(x)$ form an open cover of K when x varies in K, thus , since K is compact, there is a finite number of points $x_i \in K$ $(1 \leq i \leq n)$ such that the interiors of $W(x_i)$ form a cover of K. Then $V := \bigcup_{i=1}^{n} W(x_i)$ is a compact neighborhood of K contained in U, since a finite union of compact sets is compact. $\square$

Proposition B.9. *Let (X, d) be a locally compact metric space. Let K be a compact set and O an open set such that*

$$K \subset O \quad \text{and} \quad X \setminus O \neq \emptyset.$$

Then there exists $\varepsilon > 0$ such that the set

$$K_\varepsilon := \overline{\{x \in X : d(x, K) < \varepsilon\}}$$

is compact and $K \subset K_\varepsilon \subset O$.

Proof. Let us consider a compact neighborhood V of K contained in O, which exists by Proposition B.8. Now, we take $\varepsilon > 0$ such that $\varepsilon < d(\partial K, \partial V)$. Then $K \subset K_\varepsilon \subset V$, and K_ε is compact because it is a closed set contained in a compact set. $\square$

Theorem B.10 (Dominated Convergence Theorem). *Let $u, u_h : X \longrightarrow \overline{\mathbb{R}}$ be μ-measurable functions and assume that $u_h(x) \longrightarrow u(x)$ for μ-a.e. $x \in X$ as $h \to +\infty$. If*

$$\int_X \sup_h |u_h| d\mu < +\infty$$

then

$$\lim_{h \to +\infty} \int_X u_h d\mu = \int_X u d\mu.$$

Definition B.11. Let X and Y be topological spaces. A function $f : X \longrightarrow Y$ is a local homeomorphism if for every point $x \in X$ there exists an open set U that contains x such that the image $f(U)$ is open in Y and the restriction $f_{|_U} : U \longrightarrow f(U)$ is a homeomorphism (with the respective induced topologies on U and on $f(U)$).

Theorem B.12. *Every local homeomorphism is an open map.*

Proof. Let $f : X \longrightarrow Y$ be a local homeomorphism and $V \subseteq X$ be open in X; $f(V)$ is open iff $\forall y \in f(V)\ \exists U \subseteq Y$ open in Y such that $y \in U \subseteq f(V)$. Let $y \in f(V)$ and $x \in V$ such that $f(x) = y$. Then, since f is a local homeomorphism, $\exists A \subseteq X$ open in X; $B \subseteq Y$ open in Y such that

$$x \in A; \quad f(A) = B \quad \text{and} \quad f_{|_A} : A \to B \quad \text{is a homeomorphism.}$$

Then $y \in f(V \cap A)$ and $U = f(V \cap A)$ is open in B since $V \cap A$ is open in A and $f_{|_A}$ is open because it is a homeomorphism. Moreover, U is also open in Y since B is open in Y. □

Theorem B.13. *Let (X, d_1), (Y, d_2) be metric spaces with X compact and let $f : X \longrightarrow Y$ be a continuous injective function. Then f is open and closed.*

Proof. Let us start to prove that f is closed. Let C be a closed set in X, then it is compact, since X is compact. Hence $f(C)$ is compact in Y because f is continuous and thus $f(C)$ is also closed in Y.
To prove that f is open, by Theorem B.12 it is sufficient to show that it is a local homeomorphism.
Let x be any point in X and let $B_2(f(x), r)$ be an open ball in Y centered at $f(x)$. Since f is continuous $U := f^{-1}(B_2(f(x), r))$ is an open neighborhood of x. Now, we want to show that

$$f_{|_U} : U \longrightarrow f(U)$$

is a homeomorphism. First, $f_{|_U}$ is bijective, since f is injective. Second, it is continuous because f is continuous and $U, f(U)$ are open respectively in X, Y. Finally, we have to show that $f_{|_U}^{-1} : f(U) \to U$ is continuous.
Let C be a closed set in U, i.e. $C = S \cap U$ with S a closed set in X. Then

$$(f_{|_U}^{-1})^{-1}(C) = f_{|_U}(C) = f(S \cap U) = f(S) \cap f(U)$$

where the last equality holds since f is injective. Now $f(S) \cap f(U)$ is closed in $f(U)$ because $f(S)$ is closed in Y since we have already proved that f is closed. Therefore $f_{|_U}^{-1}$ is continuous and this concludes the proof. □

Proposition B.14 (See Corollario 6.8. in [11])**.** *Let X be a separable metric space, then every subspace Y of X is separable.*

Remark B.15. Notice that $[0,1]$ with the induced Euclidean metric is a separable metric space. Indeed, $\mathbb{R}$ is separable since $\mathbb{Q}$ is dense in $\mathbb{R}$.

Lemma B.16 (Gronwall's Lemma (strong formulation))**.** *Let $u \in C^1([0,T];\mathbb{R})$ such that*

$$\frac{du(t)}{dt} \leq Bu(t), \quad \forall t \in [0,T],$$

then

$$u(t) \leq u(0)e^{Bt} \quad \forall t \in [0,T].$$

Proof. Indeed, we have that

$$\frac{d}{dt}[u(t)e^{-Bt}] = e^{-Bt}[\frac{du(t)}{dt} - Bu(t)] \leq 0.$$

By integration, it follows that

$$u(t)e^{-Bt} - u(0) \leq 0, \text{ hence } u(t)e^{-Bt} \leq u(0).$$

So, the result follows immediately. □

Lemma B.17 (Gronwall's Lemma (weak formulation))**.** *Let $u \in C([0,T];\mathbb{R})$ such that*

$$u(t) \leq A + B\int_0^t u(s)ds \quad \forall t \in [0,T],$$

then

$$u(t) \leq Ae^{Bt}, \quad \forall t \in [0,T].$$

Proof. Let us setting $w(t) = A + B\int_0^t u(s)ds \in C^1$. In particular,

$$\frac{dw(t)}{dt} = Bu(t) \leq B\left(A + B\int_0^t u(s)ds\right) = Bw(t).$$

Therefore, by the strong formulation of Gronwall's Lemma we have $w(t) \leq w(0)e^{Bt}$, and since $w(0) = A$ we obtain

$$u(t) \leq w(t) \leq Ae^{Bt}.$$

□

Bibliography

[1] Y. Achdou, B. Franchi, N. Marcello, and M. C. Tesi, *A qualitative model for aggregation and diffusion of β-amyloid in Alzheimer's disease*, J. Math. Biol., 67 (2013), pp. 1369-1392.

[2] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Clarendon Press, Oxford 2000.

[3] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed., Lectures in Math. ETH Zürich, Birkhäuser, Basel, 2008.

[4] M. Bertsch, B. Franchi, N. Marcello, M. C. Tesi, and A. Tosin, *Alzheimer's disease: A mathematical model for onset and progression*, Math. Med. Biol., 34 (2017), pp. 193-214.

[5] M. Bertsch, B. Franchi, M. C. Tesi, and A. Tosin, *Microscopic and macroscopic models for the onset and progression of Alzheimer's disease*, J. Phys. A, 50 (2017), 414003.

[6] M. Bertsch, B. Franchi, M. C. Tesi, and A. Tosin, *Well-posedness of a mathematichal model for Alzheimer's disease*, SIAM J. Math. Anal., 50 (2018), pp. 2362-2388.

[7] K. Blennow, M. J. de Leon, H. Zetterberg, *Alzheimer's disease*, Lancet Neurol., 368 (2006), pp. 387-403.

[8] N. Bourbaki, *Topologie Générale*, Chp 9, Springer 2007.

[9] H. Braak and K. Del Tredici, *Alzheimer's pathogenesis: Is there neuron-to-neuron propagation?*, Acta Neuropathol., 121 (2011), pp. 589-595.

[10] J. A. Carrillo, R. M. Colombo, P. Gwiazda, and A. Ulikowska, *Structured populations, cell growth and measure valued balance laws*, J. Differential Equations, 252 (2012), pp. 3245-3277.

[11] V. Checcucci, A. Tognoli, E. Vesentini, *Lezioni di topologia generale*, Feltrinelli Editore Milano, 1967.

[12] J. H. M. Evers, S. C. Hille, and A. Muntean, *Measure-valued mass evolution problems with flux boundary conditions and solution-dependent velocities*, SIAM J. Math. Anal., 48 (2016), pp. 1929-1953.

[13] B. Franchi and S. Lorenzani, *From a microscopic to a macroscopic model for Alzheimer disease: Two-scale homogenization of the Smoluchowski equation in perforated domains*, J. Nonlinear Sci., 26 (2016), pp. 717-753.

[14] B. Franchi and S. Lorenzani, *Smoluchowski equation with variable coefficients in perforated domains: Homogenization and applications to mathematical models in medicine*, in Harmonic Analysis, Partial Differential Equations and Applications, Appl. Numer. Harmon. Anal., Springer, Cham, Switzerland, 2017, pp. 49-67.

[15] B. Franchi and M. C. Tesi, *A qualitative model for aggregation-fragmenattion and diffusion of β-amyloid in Alzheimer's disease*, Rend. Semin. Mat. Univ. Politec. Torino, 7 (2012), pp. 75-84.

[16] W. S. T. Griffin, J. G. Sheng, M. C. Royston, S. M. Gentleman, J. E. McKenzie, D. I. Graham, G. W. Roberts and R. E. Mrak, *Glial-neuronal interactions in Alzheimer's disease: the potential role of a cytokine cycle in disease progression*, Brain Pathol. 8 (1998), pp. 65-72.

[17] C. Haass and D. J. Selkoe, *Soluble protein oligomers in neurodegeneration: Lessons from the Alzheimer's amyloid beta-peptide*, Nat. Rev. Mol. Cell. Biol., 108 (2007), pp. 101-112.

[18] W. Hao and A. Friedman, *Mathematical model on Alzheimer's disease*, BMC Systems Biol., 108 (2016), 108.

[19] P. Hartman, *Ordinary Differential Equations*, 2nd ed., SIAM, 2002.

[20] M. Helal, E. Hingant, L. Pujo-Menjouet, and G. F. Webb, *Alzheimer's disease: Analysis of a mathematical model incorporating the role of prions*, J. Math. Biol., 69 (2013), pp. 1- 29.

[21] M. D. Hurd, P. Martorell, A. Delavande, K. J. Mullen and K. M. Langa, *Monetary costs of dementia in the United States*, New Engl. J. Med., 368 (2013), pp. 1326-34.

[22] J. J. Iliff, M. Wang, Y. Liao, B. A. Plogg, W. Peng, G. A. Gundersen, H. Benveniste, G. E. Vates, R. Deane, S. A. Goldman, E. A. Nagelhus, and M. Nedergaard, *A paravascular pathway facilitates CSF flow through the brain parenchyma and the clearance of interstitial solutes, including amyloid* β, Sci. Transl. Med., 4 (2012), 147ra111.

[23] E. Karran, M. Mercken, and B. De Strooper, *The amyloid cascade hypothesis for Alzheimer's disease: An appraisal for the development of therapeutics*, Nat. Rev. Drug Discov., 10 (2011), pp. 698-712.

[24] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge Stud. Adv. Math. 44, Cambridge University Press, Cambridge, 1995.

[25] M. P. Mattson, *Pathways towards and away from Alzheimer's disease*, Nature, 430 (2004), pp. 631-9.

[26] R. Nittka, *Inhomogeneous parabolic Neumann problems*, Czechoslovak Math. J., 64 (2014), pp. 703-742.

[27] B. Perthame, *Transport Equations in Biology*, Front. Math., Birkhäuser, Basel, 2007.

[28] A. Raj, A. Kuceyeski and M. Weiner, *A network diffusion model of disease progression in dementia*, Neuron, 73 (2012), pp. 1204-15.

[29] C. Reitz, C. Brayne, and R. Mayeux, *Epidemiology of Alzheimer disease*, Nat. Rev. Neurol., 7 (2011), pp. 137-152.

[30] F. Rothe, *Global solutions of reaction-diffusion systems*, Lecture Notes in Math. 1072, Springer, Berlin, 1984.

[31] D. J. Selkoe and J. Hardy, *The amyloid hypothesis of alzheimer's disease at 25 years*, EMBO Molec. Med., 8 (2016), pp. 595-608.

[32] J-M. Serot, J. Zmudka, and P. Jouanny, *A possible role for CSF turnover and choroid plexus in the pathogenesis of late onset Alzheimer's disease*, J. Alzheimer's Dis., 30 (2012), pp. 17-26.

[33] O. G. Tatarnikova, M. A. Orlov, and B. N.V, *Beta-amyloid and tau protein: Structure, interaction and prion-like properties*, Biochem. (Moscow), 80 (2015), pp. 1800-1819.