ESAIM: COCV 24 (2018) 377–399
https://doi.org/10.1051/cocv/2017003

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

OPTIMIZATION IN STRUCTURE POPULATION MODELS THROUGH THE ESCALATOR BOXCAR TRAIN *, **, ***

Rinaldo M. Colombo[1,a], Piotr Gwiazda[1,2] and Magdalena Rosińska[1,2]

Abstract. The Escalator Boxcar Train (EBT) is a tool widely used in the study of balance laws motivated by structure population dynamics. This paper proves that the approximate solutions defined through the EBT converge to exact solutions. Moreover, this method is rigorously shown to be effective also in computing optimal controls. As preliminary results, the well posedness of classes of PDEs and of ODEs comprising various biological models is also obtained. A specific application to welfare policies illustrates the whole procedure.

Mathematics Subject Classification. 65M75, 35R06, 92D25.

Received January 8, 2016. Revised November 7, 2016. Accepted January, 12, 2017.

1. Introduction

This paper is devoted to the well posedness, to the numerical approximation and to the optimal control of renewal equations motivated by physiologically structured population models and whose solutions attain values in spaces of measures.

The dynamics of populations which are heterogeneous with respect to some individual property can be described through initial – boundary value problems for a class of nonlinear first order partial differential equations (PDE), called renewal equations. Within this class, one of the first PDE models devoted to population biology is the renewal equation introduced by Kermack and McKendrick with reference to epidemiology, see [23, 24]. There, the time since infection, *i.e.*, the age, plays the role of a structure parameter, due to its essential role in the spreading of the epidemic. Equations of the same class are later proposed by von Förster in [33] to describe the process of cell division. The recent monograph [10] provides an extensive theoretical and empirical treatment of the ecology of ontogenetic growth and development of organisms, emphasizing the importance of an individual–based perspective in understanding the dynamics of populations and communities. Classical

Keywords and phrases. Escalator boxcar train, structure population model.

* *R.M.C. was partially supported by the 2015 GNAMPA project* Balance Laws in the Modeling of Physical, Biological and Industrial Processes *and by the CaRiPLo project 2013-0893.*

** *The research of M.R. received funding from the National Science Centre, DEC-2012/05/E/ST1/02218.*

*** *The research of P.G. received funding from the National Science Centre, Poland, 2014/13/B/ST1/03094.*

[1] University of Brescia, 25121 Brescia BS, Italie.
[a] Corresponding author: rinaldo.colombo@unibs.it
[2] University of Warsaw, 00-927 Warszawa, Pologne.
pgwiazda@mimuw.edu.pl; mrosinska@pzh.gov.pl

Article published by EDP Sciences

analytic studies on these equations are settled in $\mathbf{L^1}$ and go back, for instance, to the monographs of Webb [34], Iannelli [22] or Thieme [31].

The space of positive Radon measures is introduced in biological applications in [27]. Indeed, whenever the distribution of individuals is concentrated on discrete values of structure parameters, for instance at the initial time, the resulting population density may well lack absolute continuity with respect to the Lebesgue measure. One is thus lead to consider the problem

$$\begin{cases} \partial_t \mu + \partial_x \left(b(t)(\mu, x)\ \mu\right) + c(t)(\mu, x)\ \mu = 0 \\ b(t)(\mu(t), 0)\ D_\lambda \mu(t)(0+) = \int_0^{+\infty} \beta(t)(\mu(t), x)\ \mathrm{d}\mu(t)(x) \\ \mu(0) = \mu_o \end{cases} \tag{1.1}$$

where $t \in \mathbb{R}_+$ is time and $x \in \mathbb{R}_+$ is a biological parameter, typically age or size. The unknown μ is a time dependent, non-negative and finite Radon measure. The growth function b and the mortality rate c are strictly positive, while the birth function β is non-negative. By $D_\lambda \mu(0+)$ we denote the Radon–Nikodym derivative of μ with respect to the Lebesgue measure λ computed at 0. The initial datum μ_o is a non-negative Radon measure.

Clearly, as soon as the initial datum in (1.1) is regular, solutions to (1.1) can be found in spaces of regular functions, such as $\mathbf{C^1}$. Nevertheless, we choose a measure theoretic setting for the reasons stated here below and in Remark 1.2.

Remark 1.1. *A priori*, the present approach overcomes a key inconsistency between the $\mathbf{L^1}$ norm and empirical data. Even if we assume that the distribution of a real population is absolutely continuous with respect to the Lebesgue measure, so that an $\mathbf{L^1}$ distribution density exists, typical experimental data provide information on percentiles, *i.e.*, on the number of individuals in some range of the structural variable (age, size, *etc.*). In the case of epidemiology or demography, for instance, birth cohorts are typically used (*e.g.*, individuals born in a given year). A detailed discussion of this topic can be found for instance in ([20], Sect. 5).

The analysis of solutions to (1.1) in spaces of positive Radon measures was initiated in [11], where the authors show the weak$*$ continuity of solutions with respect to time and initial data. They also point out the key relevance of the dependence of solutions on the various model parameters, which was obtained in [6, 18, 20, 32].

The Lipschitz continuous dependence of solutions in measure spaces on time and initial datum is a preliminary step towards the convergence of the so called *particle methods*. These are numerical algorithms whose starting idea is the representation of a heterogeneous population as a sum of Dirac masses evolving in time. This representation is consistent with the usual experimental attitude of concentrating real data in discrete cohorts evolving in time. On these grounds, the numerical algorithm usually referred to as the *Escalator Boxcar Train* (EBT) is introduced back in [9]. Remarkably, in spite of the wide success of this method, a convergence proof of the EBT appears only rather recently in [3]. A key role in this result is played by the bounded Lipschitz distance introduced in [18]. Detailed estimates on the order of convergence are then provided in [17].

Another numerical method effective in the computation of solutions to structured population models is proposed in [7]. Here, a key role is played by the *operator splitting* method. According to it, the measure valued semigroup generated by renewal equations can be approximated through the iterated application of simpler semigroups. More precisely, a problem involving both transport terms and nonlocal growth terms is approximated through two problems, each involving only one of the two processes. The analytic framework established in [6] allows a detailed control of the convergence rate of the algorithm.

Remark 1.2. *A posteriori*, even if the solution to (3.1) can be regular, the EBT approximation necessarily requires to be set in the space $\mathcal{M}(\mathbb{R}^+)$ of Radon measures. The main result of the present paper is the convergence of the optimal solution to the approximated problem (defined in the space of measures) to the optimal solution to the exact problem. Therefore, the use of measure theoretic tools is unavoidable and the choice of which metric to adopt on $\mathcal{M}(\mathbb{R}^+)$ is crucial.

From the measure theoretic point of view, the above mentioned results rely on the use of Wasserstein (or Monge−Kantorovich) type metrics, adapted to the nonconservative character of (1.1). This methodology was proposed in [18] for a flat metric (bounded Lipschitz distance) and in [20] for a Wasserstein metric, suitably modified to deal with nonnegative Radon measures with integrable first moment. A relevant advantage of this approach is in providing a structure of a space appropriate both to compare solutions and to study their stability. Remark that precise estimates on the continuous dependence of solutions on the modeling parameters plays a key role in the numerical approximations and in calibrating the model on the basis of experimental data. We refer to [29] for the definition and properties of a similar metric structure.

Similar techniques based on particle methods are usual tools in simulating kinetic models for more than three decades in physics, see for instance [30] and the references therein. Recent applications include for instance the porous medium equation [28, 35] and the isentropic Euler equations in fluid mechanics [16, 35]. Other particle methods are found also in the study of problems related to crowd dynamics and pedestrians flow, see [13, 14, 29], as well as in the description of the collective motion of large groups of agents, see [5]. Differently from the case of structured population models, the original particle methods are mainly designed for problems where the total mass, or number of individuals, is conserved.

Aiming at the optimal control of the solution to (1.1), we introduce therein a control parameter u, possibly time and/or state dependent, attaining values in a given set $\mathcal{U}$. Therefore, we obtain:

$$\begin{cases} \partial_t \mu + \partial_x \left(b(t,u)(\mu, x)\ \mu \right) + c(t,u)(\mu, x)\ \mu = 0 \\ b(t,u)(\mu(t), 0)\ D_\lambda \mu(t)(0+) = \displaystyle\int_0^{+\infty} \beta(t,u)(\mu(t), x)\ \mathrm{d}\mu(t)(x) \\ \mu(0) = \mu_o \end{cases} \tag{1.2}$$

Together with (1.2), we are given a cost functional

$$\mathcal{J}(u) = \int_0^{+\infty} j\left(t, u(t), \mu(t)\right) \mathrm{d}t$$

and we provide below a *constructive* algorithm to find, within a suitable function space, a control function u_* optimal in the sense that

$$\mathcal{J}(u_*) = \min_{u(t) \in \mathcal{U}} \mathcal{J}(u).$$

As is well known, solutions to conservation or balance laws typically depend in a Lipschitz continuous way on the initial datum as well as from the functions defining the equation. This does not allow the use of differential tools in the search for the optimal control.

Here, *constructive* should be understood in the following sense: on the basis of the control problem for (1.2), we define a sequence of control problems for a system of ordinary differential equations and prove that the corresponding sequence of optimal controls converges to an optimal control for the original problem. More precisely, we approximate the solution to (1.2) by means of the EBT algorithm as defined in ([17], Sect. III). The functional $\mathcal{J}$ computed along approximate solutions is proved to be a smooth, namely $\mathbf{C^1}$, function of the control parameter u and this allows to exhibit the existence of an optimal control for each approximate problem. A limiting procedure constructively ensures the existence of the optimal control for the original problem (1.2).

The next section presents results on the well posedness of (1.1) and the results on the escalator boxcar train algorithm that allow to obtain our main result, namely the construction of a sequence of controls that converge to an optimal control for (1.2). Section 3 is devoted to a possible application of the theory here developed. The technical proofs are deferred to Section 4, with a final Appendix that gathers necessary results concerning ordinary differential equations.

2. Main results

Throughout, we denote $\mathbb{R}_+ = [0, +\infty[$. Let (M, d_M) be a metric space and $(V, \| \ \|_V)$ be a normed space. Then, $\mathbf{C^0}(M; V)$, respectively $\mathbf{C^{0,1}}(M; V)$ is the space of continuous, respectively Lipschitz continuous, functions defined on M and attaining values in V, equipped with the norm

$$\|\varphi\|_{\mathbf{C^0}(M,V)} = \sup_{x\in M} \left\|\varphi(x)\right\|_V, \text{ respectively} \tag{2.1}$$

$$\|\varphi\|_{\mathbf{C^{0,1}}(M,V)} = \max\left\{\sup_{x\in M}\left\|\varphi(x)\right\|_V, \sup_{x_1,x_2\in M,\ x_1\neq x_2} \frac{\left\|\varphi(x_2)-\varphi(x_1)\right\|_V}{d_M(x_1,x_2)}\right\}. \tag{2.2}$$

Given $T>0$ and a function $f\colon [0,T]\to V$, we set

$$\mathrm{TV}_V(f) = \sup\left\{\sum_{i=1}^n \left\|f(t_i)-f(t_{i-1})\right\|_V : n\in\mathbb{N} \text{ and } \begin{array}{ll} t_i\in[0,T] & \text{for } i=0,\ldots,n \\ t_{i-1}<t_i & \text{for } i=1,\ldots,n\end{array}\right\}. \tag{2.3}$$

The space $\mathcal{M}^+(\mathbb{R}_+)$ of positive Radon measure on $\mathbb{R}_+$ is equipped with the flat distance

$$d(\mu',\mu'') = \sup\left\{\int_{\mathbb{R}_+}\varphi\ \mathrm{d}(\mu'-\mu'') : \varphi\in\mathbf{C^1}(\mathbb{R}_+;[-1,1]) \text{ with } \mathbf{Lip}(\varphi)\leq 1\right\}, \tag{2.4}$$

see ([6], Sect. 2). Below, for positive $T,\mathcal{L}$ and $\mathcal{C}$, we use the space $\mathcal{F}$ of functions

$$f\colon [0,T]\to\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R}_+) \tag{$\boldsymbol{\mathcal{F}}$}$$

with the properties:

($\mathcal{F}_1$) f is bounded: $\left\|f(t)\right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}\leq\mathcal{L}$ for all $t\in[0,T]$.

($\mathcal{F}_2$) f has bounded total variation in time: $\mathrm{TV}_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R}_+)}(f)\leq\mathcal{C}$.

Throughout, the constants $T,\mathcal{L}$ and $\mathcal{C}$ are kept fixed and the dependence of $\mathcal{F}$ on them is omitted. In $\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+$ we use the distance

$$d_{\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+}\left((\mu_1,x_1),(\mu_2,x_2)\right) = d(\mu_1,\mu_2)+|x_2-x_1|,$$

where d is as in (2.4). Therefore, ($\mathcal{F}_1$) also implies that f is Lipschitz continuous in μ and a uniformly in t, in the sense that for all $t\in[0,T]$, $\mu_1,\mu_2\in\mathcal{M}^+(\mathbb{R}_+)$ and $x_1,x_2\in\mathbb{R}_+$,

$$\left|f(t)(\mu_1,x_1)-f(t)(\mu_2,x_2)\right|\leq\mathcal{L}\left(d(\mu_1,\mu_2)+|x_1-x_2|\right).$$

2.1. PDE – Well posedness

As a first step, we need to extend the well posedness of (1.1) obtained in ([6], Thm. 2.11) to the case of functions b and c being only of bounded variation in time. First, recall the definition of solution to (1.1) attaining as values Radon measures.

Definition 2.1 ([19], Def. 3.1)**.** Fix $T>0$ and let $\mu_o\in\mathcal{M}^+(\mathbb{R}_+)$ and $b,c,\beta\in\mathcal{F}$. By *solution* to (1.1) we mean a function $\mu\colon[0,T]\to\mathcal{M}^+(\mathbb{R}_+)$ with the following properties:

(1) μ is Lipschitz continuous with respect to the flat distance (2.4);

(2) for all $\varphi\in(\mathbf{C^1}\cap\mathbf{C^{0,1}})([0,T]\times\mathbb{R}_+;\mathbb{R})$

$$\begin{aligned}\int_{\mathbb{R}_+}\varphi(T,x)\ \mathrm{d}\mu(T)(x)-\int_{\mathbb{R}_+}\varphi(0,x)\,\mathrm{d}\mu_o(x) = {}& \int_0^T\int_{\mathbb{R}_+}\partial_t\varphi(t,x)\,\mathrm{d}\mu(t)(x)\,\mathrm{d}t \\ &+\int_0^T\int_{\mathbb{R}_+}\big(\partial_x\varphi(t,x)\ b(t)(\mu(t),x) \\ &-\varphi(t,x)\ c(t)(\mu(t),x)\big)\,\mathrm{d}\mu(t)(x)\,\mathrm{d}t \\ &+\int_0^T\int_{\mathbb{R}_+}\varphi(t,0)\ \beta(t)(\mu(t),x)\ \mathrm{d}\mu(t)(x)\,\mathrm{d}t\,.\end{aligned}$$

We now weaken the assumptions on the regularity in time used in ([6], Thm. 2.11).

Theorem 2.2. *Fix $T > 0$. Let $b, c, \beta \in \mathcal{F}$. Then, for any $\mu_o \in \mathcal{M}^+(\mathbb{R}_+)$, problem (1.1) admits a unique solution in the sense of Definition 2.1. Moreover, there exists a constant C dependent only on $\mathcal{C}, \mathcal{L}$ and T such that if for $i = 1, 2$, μ^i is the solution to (1.1) with initial data μ_o^i and b, c, β replaced by b_i, c_i, β_i, then,*

$$\begin{aligned} d\left(\mu^1(t), \mu^2(t)\right) \le d(\mu_o^1, \mu_o^2)\, e^{C\,t} + C\,t\, e^{C\,t} \Bigg(& \sup_{t\in[0,T]} \left\|b_1(t) - b_2(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \\ & + \sup_{t\in[0,T]} \left\|c_1(t) - c_2(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \\ & + \sup_{t\in[0,T]} \left\|\beta_1(t) - \beta_2(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \Bigg). \end{aligned} \tag{2.5}$$

The proof is deferred to Section 4.

Aiming at the study of (1.2), we extend the definition of $\mathcal{F}$ as follows. Fix $T > 0$ and a compact subset $\mathcal{U}$ of $\mathbb{R}^N$, for fixed positive $T, \mathcal{L}, \mathcal{C}$ and a positive integer N, we introduce the space $\mathcal{F}^u$ of functions

$$f : [0, T] \times \mathcal{U} \to \mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R}) \tag{$\boldsymbol{\mathcal{F}^u}$}$$

with the properties:

($\mathcal{F}_1^u$) f is bounded: $\left\|f(t, u)\right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \le \mathcal{L}$ for all $t \in [0, T]$ and all $u \in \mathcal{U}$.

($\mathcal{F}_2^u$) f has bounded total variation in t uniformly in u: $\mathrm{TV}_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R}_+)}\left(f(\cdot, u)\right) \le \mathcal{C}$ for all $u \in \mathcal{U}$.

($\mathcal{F}_3^u$) f is Lipschitz continuous in the control uniformly in time: for all $t \in [0, T]$ and for all $u_1, u_2 \in \mathcal{U}$, $\left\|f(t, u_1) - f(t, u_2)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \le \mathcal{L}\left\|u_1 - u_2\right\|$.

As above, we remark that ($\mathcal{F}_1^u$) ensures that $f(t; u)$ is Lipschitz continuous in μ and a uniformly in t and u: for all $t \in [0, T]$, $u \in \mathcal{U}$, $\mu_1, \mu_2 \in \mathcal{M}^+(\mathbb{R}_+)$ and $x_1, x_2 \in \mathbb{R}_+$,

$$\left|f(t, u)(\mu_1, x_1) - f(t, u)(\mu_2, x_2)\right| \le \mathcal{L}\left(d(\mu_1, \mu_2) + |x_1 - x_2|\right).$$

In ($\mathcal{F}_2^u$), the total variation is computed as in (2.3), keeping u fixed. Throughout, the constants $T, \mathcal{L}$ and $\mathcal{C}$ are kept fixed and the dependence of $\mathcal{F}^u$ on them is omitted.

The extension of Definition 2.1 from the case of (1.1) to that of (1.2) is immediate.

Corollary 2.3. *Fix $T > 0$ and a compact subset $\mathcal{U}$ of $\mathbb{R}^N$. Let $b, c, \beta \in \mathcal{F}^u$. Then, for any $\mu_o \in \mathcal{M}^+(\mathbb{R}_+)$ and any $u \in \mathbf{BV}([0, T]; \mathcal{U})$, problem (1.2) admits a unique solution. Moreover, there exists a constant C dependent only on $\mathcal{L}, \mathcal{C}$ and T such that if for $i = 1, 2$, μ^i is the solution to (1.1) with initial data μ_o^i, and b, c, β, u replaced by b_i, c_i, β_i, u_i, then,*

$$\begin{aligned} d\left(\mu^1(t), \mu^2(t)\right) \le d(\mu_o^1, \mu_o^2)\, e^{C\,t} + C\,t\, e^{C\,t} \Bigg(& \sup_{t\in[0,T]} \left\|b_1(t) - b_2(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathcal{U}\times\mathbb{R}_+;\mathbb{R})} \\ & + \sup_{t\in[0,T]} \left\|c_1(t) - c_2(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathcal{U}\times\mathbb{R}_+;\mathbb{R})} \\ & + \sup_{t\in[0,T]} \left\|\beta_1(t) - \beta_2(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathcal{U}\times\mathbb{R}_+;\mathbb{R})} \\ & + \sup_{t\in[0,T]} \left\|u_1(t) - u_2(t)\right\| \Bigg). \end{aligned} \tag{2.6}$$

The proof is in Section 4.

2.2. ODE – Well posedness

We first present the approximation algorithm introduced in [9], see [3, 17] for the present simplified version. Fix a positive time T. For any $n \in \mathbb{N} \setminus \{0\}$ and for the time step Δt, approximate the initial datum μ_o in (1.2) by means of a linear combination μ_0^n of Dirac deltas centered at $x_0^0, x_0^1, \ldots, x_0^n$ with masses $m_0^0, \ldots, m_0^n$ and approximate the initial datum with the measure [3]

$$\mu_0^n = \sum_{i=0}^{n} m_0^i \, \delta_{x_0^i}.$$

On the time interval $[0, \Delta t[$, we approximate the solution to (1.2) with the measure

$$\mu^n(t) = \sum_{i=0}^{n} m^i(t) \, \delta_{x^i(t)}$$

where

$$\begin{cases} \dot{x}^i = b\left(t, u(t)\right)\left(\mu^n(t), x^i(t)\right) & i = 0, \ldots, n \\ \dot{m}^0 = -c\left(t, u(t)\right)\Big(\mu^n(t), x^0(t)\Big)\; m^0 + \displaystyle\sum_{i=1}^{n} \beta\left(t, u(t)\right)\left(\mu^n(t), x^i(t)\right)\; m^i & \\ \dot{m}^i = -c\left(t, u(t)\right)\left(\mu^n(t), x^i(t)\right)\; m^i & i = 1, \ldots, n \\ x^i(0) = x_0^i & i = 0, \ldots, n \\ m^0(0) = 0 & \\ m^i(0) = m_0^i & i = 1, \ldots, n \end{cases} \tag{2.7}$$

Define $x_1^i = \lim_{t \to \Delta t-} x^i(t)$ and $m_1^i = \lim_{t \to \Delta t-} m^i(t)$ for $i = 0, \ldots, n$. Iteratively, for $k \geq 1$, we prolong μ^n, $x^{-k+1}, \ldots, x^n$ and $m^{-k+1}, \ldots, m^n$ on the interval $\left[k\,\Delta t, (k+1)\,\Delta t\right[$ solving

$$\begin{cases} \dot{x}^i = b\left(t, u(t)\right)\left(\mu^n(t), x^i(t)\right) & i = -k, \ldots, n \\ \dot{m}^{-k} = -c\left(t, u(t)\right)\Big(\mu^n(t), x^{-k}(t)\Big)\; m^{-k} + \displaystyle\sum_{i=-k+1}^{n} b\left(t, u(t)\right)\left(\mu^n(t), x^i(t)\right)\; m^i & \\ \dot{m}^i = -c\left(t, u(t)\right)\left(\mu^n(t), x^i(t)\right)\; m^i & i = -k+1, \ldots, n \\ x^i(k\Delta t) = x_k^i & i = -k+1, \ldots, n \\ m^i(k\Delta t) = m_k^i & i = -k+1, \ldots, n \\ x^{-k}(k\Delta t) = 0 & \\ m^{-k}(k\Delta t) = 0 & \end{cases} \tag{2.8}$$

where $x_k^i = \lim_{t \to k\Delta t-} x^i(t)$, $m_k^i = \lim_{t \to k\Delta t-} m^i(t)$ for $i = 0, \ldots, n$ and

$$\mu^n(t) = \sum_{i=-k+1}^{n} m^i(t) \, \delta_{x^i(t)}.$$

To describe the hypotheses on b, c, β ensuring the well posedness of (2.7)–(2.8) it is of use to introduce, for positive T and L, the set $\widetilde{\mathcal{F}}^u$ of functions

$$f \colon [0, T] \times \mathcal{U} \to \mathbf{C}^{\mathbf{0,1}}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R}_+) \tag{$\widetilde{\boldsymbol{\mathcal{F}}}^{\boldsymbol{u}}$}$$

such that

$$f(t; u)(\mu, a) = \tilde{f}\left(t, \int_{\mathbb{R}_+} \bar{f}(\alpha)\, \mathrm{d}\mu(\alpha)\,, a; u\right)$$

[3]We note that since linear combinations of Dirac deltas are dense in the space of bounded Radon measures equipped with bounded Lipschitz distance [21], we can approximate every initial datum with a linear combination of Dirac deltas. However, constructing an optimal approximation is not easy, some suboptimal algorithms are given in ([7], Sect. 2.3).

where:

($\widetilde{\mathcal{F}}_1^u$) The map $\bar f \in \mathbf{C^1}(\mathbb{R}_+;\mathbb{R}_+)$ is bounded.
($\widetilde{\mathcal{F}}_2^u$) The map $(A,x;u) \to \tilde f(t,A,x;u)$ is in $\mathbf{C^1}(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{U};\mathbb{R}_+)$ for a.e. $t \in [0,T]$.
($\widetilde{\mathcal{F}}_3^u$) The map $t \to \tilde f(t,A,x;u)$ is in $\mathbf{L}^\infty([0,T];\mathbb{R}_+)$ for all $A \in \mathbb{R}_+$, $x \in \mathbb{R}_+$ and $u \in \mathcal{U}$.
($\widetilde{\mathcal{F}}_4^u$) $\tilde f$ is Lipschitz continuous in A, x, u uniformly in t:

$$\left|\tilde f(t,A_1,x_1;u_1) - \tilde f(t,A_2,x_2;u_2)\right| \le L\left(|A_1 - A_2| + |x_1 - x_2| + \|u_1 - u_2\|\right).$$

The next result ensures the well posedness of the Cauchy Problem for the system of ordinary differential equations (4.2)–(4.3).

Theorem 2.4. *Fix $n, N \in \mathbb{N}\setminus\{0\}$, $T, L > 0$ and a compact subset $\mathcal{U}$ of $\mathbb{R}^N$. Let $b, c, \beta \in \tilde{\mathcal{F}}^u$. Then, for any control $u \in \mathbf{BV}([0,T];\mathcal{U})$ and any initial datum $(x_0^1,\dots,x_0^n) \in \mathbb{R}_+^{n+1}$, $(m_0^1,\dots,m_0^n) \in \mathbb{R}^n$, problem (2.7)–(2.8) admits a unique solution $t \to (x,m)_u(t)$ defined for all $t \in [0,T]$. Moreover, the map $u \to (x,m)_u$ is in $\mathbf{C^1}(\mathbf{BV}([0,T];\mathcal{U});\mathbf{C^0}([0,T];\mathbb{R}_+^{n+1} \times \mathbb{R}^n))$.*

The proof directly follows from Lemma 4.2, which shows that Lemma A.1 can be applied, and from the usual properties of the Nemitsky operator.

Theorem 2.5. *Consider an arbitrary $n \in \mathbb{N}\setminus\{0\}$ and fix $T > 0$ and a compact subset $\mathcal{U}$ of $\mathbb{R}^N$. Let $b, c, \beta \in \mathcal{F}^u$ and $u \in \mathbf{BV}([0,T];\mathcal{U})$. Fix $\mu_o \in \mathcal{M}^+(\mathbb{R}_+)$, $(x_o^1,\dots,x_o^n) \in \mathbb{R}_+^n$, $(m_o^1,\dots,m_o^n) \in \mathbb{R}_+^n$. Let μ solve problem (1.2) in the sense of Definition 2.1 and (m,x) solve problem (2.7)–(2.8) with time step Δt. Then, there exists a positive C independent of n, u, Δt such that for all $t \in [0,T]$,*

$$d\left(\mu_t, \sum_{i=-n}^{n} m^i(t)\,\delta_{x^i(t)}\right) \le C \cdot \left[\Delta t + d\left(\mu_o, \sum_{i=0}^{n} m_0^i\,\delta_{x_0^i}\right)\right].$$

In specific numerical implementations of the present method, the quantity $d(\mu_o, \sum_{i=0}^n m_0^i\,\delta_{x_0^i})$ is typically of the same order of the size of the space mesh Δx.

2.3. Optimal control

A general cost functional defined on the controls in $\mathbf{BV}([0,T];\mathcal{U})$ is

$$\begin{aligned} \widetilde{\mathcal{J}} :\ & \mathbf{BV}([0,T];\mathcal{U}) \to \mathbb{R} \\ & u \quad \to \int_0^T j\left(t, u(t), \int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_u(t)(\xi)\right)\mathrm{d}t \end{aligned} \tag{2.9}$$

where $\gamma \in \mathbf{C^{0,1}}(\mathbb{R}_+;\mathbb{R}_+)$, μ_u is the solution to (1.2) corresponding to the control u with b, β, c and μ_o satisfying the assumptions of Theorem 2.2, and $j\colon [0,T] \times \mathcal{U} \times \mathbb{R}_+ \to \mathbb{R}_+$ being such that:

($\mathbf{J_1}$) $j \ge 0$;
($\mathbf{J_2}$) the map $t \to j(t,x;u)$ is measurable for all $x \in \mathbb{R}_+$, $u \in \mathcal{U}$ and there exists a $\hat u \in \mathbf{BV}([0,T];\mathcal{U})$ such that $\mathcal{J}(\hat u) < +\infty$;
($\mathbf{J_3}$) there exist $L \in \mathbf{L^1}([0,T];\mathbb{R}_+)$ and a nondecreasing $\omega \in \mathbf{C^0}(\mathbb{R}_+;\mathbb{R}_+)$, with $\omega(0) = 0$, such that

$$\left|j(t,x_1,u_1) - j(t,x_2,u_2)\right| \le L(t)\ \omega\big(|x_1 - x_2| + |u_1 - u_2|\big)$$

for a.e. $t \in [0,T]$, for all $x_1, x_2 \in \mathbb{R}_+$ and all $u_1, u_2 \in \mathcal{U}$.

Having to consider also costs related to the adjustments in the values of the control, it is natural to seek the minimization of

$$\begin{aligned} \mathcal{J} \,:\, \mathbf{BV}([0,T];\mathcal{U}) &\to \mathbb{R} \\ u &\to \widetilde{\mathcal{J}}(u) + \mathrm{TV}_{\mathbb{R}^N}(u). \end{aligned} \tag{2.10}$$

As a first result, we prove the existence of an optimal control.

Theorem 2.6. *Fix $T > 0$ and a compact subset $\mathcal{U}$ of $\mathbb{R}^N$. For all $b, c, \beta \in \mathcal{F}^u$, $u \in \mathbf{BV}([0,T];\mathcal{U})$ and $\mu_o \in \mathcal{M}^+(\mathbb{R}_+)$, let μ_u be the solution to problem (1.2). With reference to the cost functional (2.9), $\gamma \in \mathbf{C^{0,1}}(\mathbb{R}_+;\mathbb{R}_+)$ and j satisfies $(\mathbf{J_1})$, $(\mathbf{J_2})$, $(\mathbf{J_3})$. Then, there exists a control minimizing $\mathcal{J}$ as defined in (2.10):*

$$\exists\, u^* \in \mathbf{BV}([0,T];\mathcal{U}) : \quad \mathcal{J}(u^*) = \inf_{u \in \mathbf{BV}([0,T];\mathcal{U})} \mathcal{J}(u).$$

We now pass to the discrete counterpart of Theorem 2.6, substituting the evolution described by (1.2) with the approximation provided by the Escalator Boxcar Train (2.7)–(2.8). At the same time, also the functionals (2.7)–(2.8) have to be computed on linear combination of Dirac deltas.

Theorem 2.7. *Fix $T > 0$ and a compact subset $\mathcal{U}$ of $\mathbb{R}^N$. Let $b, c, \beta \in \mathcal{F}^u$ and $u \in \mathbf{BV}([0,T];\mathcal{U})$. For any $n \in \mathbb{N} \setminus \{0\}$ and $\Delta t_n > 0$, fix an initial datum $(x_o^1, \ldots, x_o^n) \in \mathbb{R}_+^{n+1}, (m_o^1, \ldots, m_o^n) \in \mathbb{R}^n$ in (2.7)–(2.8) and call $(x^{-n}, \ldots, x^n)$, $(m^{-n}, \ldots, m^n)$ the corresponding solution. Further, define the cost functionals*

$$\begin{aligned} \widetilde{\mathcal{J}}_n \,:\, \mathbf{BV}([0,T];\mathcal{U}) &\to \mathbb{R} \\ u &\to \int_0^T j\left(t, u(t), \int_{\mathbb{R}_+} \gamma(\xi)\, \mathrm{d}\mu_u^n(t)(\xi)\right) \mathrm{d}t \end{aligned} \tag{2.11}$$

$$\begin{aligned} \mathcal{J}_n \,:\, \mathbf{BV}([0,T];\mathcal{U}) &\to \mathbb{R} \\ u &\to \widetilde{\mathcal{J}}_n(u) + \mathrm{TV}_{\mathbb{R}^N}(u). \end{aligned} \tag{2.12}$$

where $\mu_u^n(t) = \sum_{i=-n}^n m^i(t)\, \delta_{x^i(t)}$, $\gamma \in (\mathbf{C^1} \cap \mathbf{C^{0,1}})(\mathbb{R}_+;\mathbb{R}_+)$, j satisfies $(\mathbf{J_1})$, $(\mathbf{J_2})$, $(\mathbf{J_3})$ and there exists a $\hat{u} \in \mathbf{BV}([0,T];\mathcal{U})$ such that $\mathcal{J}(\hat{u}) < +\infty$. Then, there exists a control minimizing $\mathcal{J}_n$:

$$\exists u_n^* \in \mathbf{BV}([0,T];\mathcal{U}) : \quad \mathcal{J}_n(u_n^*) = \inf_{u \in \mathbf{BV}([0,T];\mathcal{U})} \mathcal{J}_n(u).$$

The above theorems yield the following corollary, which is the main result of the present work. It ensures that the Escalator Boxcar Train algorithm can also be used to solve optimal control problems.

Corollary 2.8. *With the same assumptions and notation as in Theorem 2.6 and in Theorem 2.7, if*

$$\lim_{n \to +\infty} \Delta t_n = 0 \qquad \textit{and} \qquad \lim_{n \to +\infty} d\left(\mu_o, \sum_{i=-n}^{n} m_o^i(t) \delta_{x_o^i(t)}\right) = 0$$

then,

$$\lim_{n \to +\infty} \mathcal{J}_n(u_n^*) = \inf_{u \in \mathbf{BV}([0,T];\mathcal{U})} \mathcal{J}(u) \tag{2.13}$$

and, up to a subsequence,

$$\lim_{n \to +\infty} \left\| u_n^* - u^* \right\|_{\mathbf{L}^\infty([0,T];\mathbb{R})} = 0 \quad \textit{where} \quad \mathcal{J}(u^*) = \inf_{u \in \mathbf{BV}([0,T];\mathcal{U})} \mathcal{J}(u). \tag{2.14}$$

3. The McKendrick – Von Förster model in welfare policies

The McKendric – Von Förster model for population growth, equipped with an integral functional to be maximized, provides a first example of a system fitting within (1.2), where the results in the Sections 2.1 and 2.2 can be applied.

Consider a population described by the amount $n = n(t,x)$ of people that at time t have the age x. Call $-d$, with $d = d(x)$, the population mortality rate. We thus obtain:

$$\begin{cases} \partial_t n + \partial_x n = -d(x)\, n \\ n(t,0) = \int_0^{+\infty} \tilde\beta(x)\, n(t,x)\,\mathrm{d}x \\ n(0,x) = n_o(x). \end{cases}$$

Here, $\tilde\beta$ describes the natality rate of the population of age a at time t.

Introduce a policy to sustain birth rate. It is then natural to assume that a control parameter, say u, enters the birth functions. The parameter u, possibly vector valued, reflects a government policy to foster natality, helping through *ad hoc* acts the families with children.

$$\begin{cases} \partial_t n_u + \partial_x n_u = -d(x)\, n_u \\ n_u(t,0) = \int_0^{+\infty} \tilde\beta(x,u)\, n_u(t,x)\,\mathrm{d}x \\ n_u(0,x) = n_o(x). \end{cases} \tag{3.1}$$

From the governmental point of view, the income of the state welfare can be described by the functional

$$\mathcal{J}(u) = \int_0^{+\infty} \mathrm{e}^{-\lambda\, t} \left(\int_0^{+\infty} w(x)\, n_u(t,x)\,\mathrm{d}x - u(t)\, n_u(t,0) \right) \mathrm{d}t\,. \tag{3.2}$$

The weight $w = w(x)$ is positive all through the active age interval, *i.e.*, all during the period where individuals, paying taxes, sustain the state. On the contrary, w is negative when individuals receive services from the state, *e.g.*, during childhood and retirement. The term $\mathrm{e}^{\lambda t}$ is motivated by the need of describing an interest rate, as typically used in economical models, see for instance [2].

Lemma 3.1. *Fix a compact $\mathcal{U}$ in $\mathbb{R}^N$ and $\alpha \in\,]0,1[$. System (3.1) fits into (1.2) setting*

$$b(t,u)(\mu,x) = 1, \qquad c(t,u)(\mu,x) = d(x), \qquad \beta(t,u)(\mu,x) = \tilde\beta(x,u).$$

Moreover, if

$$d \in \mathbf{C^{0,1}}(\mathbb{R}_+, \mathbb{R}), \qquad \tilde\beta \in \mathbf{C^{0,1}}(\mathcal{U} \times \mathbb{R}_+; \mathbb{R})$$

then, for all $u \in \mathbf{C^{1,\alpha}_b}([0,T];\mathcal{U})$, Theorem 2.4 applies.

In the present case, equations (2.7)–(2.8) take the form, for $t \in [0, \Delta t]$

$$\begin{cases} \dot x^i = 1 & i = 0,\ldots,n \\ \dot m^0 = -d(x^0)\, m^0 + \sum_{i=0}^{n} \tilde\beta\left(x^i, u(t)\right) m^i & \\ \dot m^i = -d(x^i)\, m^i & i = 1,\ldots,n \\ x^i(0) = x^i_o & i = 0,\ldots,n \\ m^i(0) = m^i_o & i = 0,\ldots,n \end{cases}$$

while for $t \in [k\,\Delta t, (k+1)\Delta t]$ the solution to the above system is extended as follows

$$\begin{cases} \dot{x}^i = 1 & i = -k, \ldots, n \\ \dot{m}^{-k} = -d(x^{-k})\, m^{-k} + \displaystyle\sum_{i=-k}^{n} \tilde{\beta}\left(x^i, u(t)\right) m^i & \\ \dot{m}^i = -d(x^i)\, m^i & i = -k+1, \ldots, n \\ x^{-k}(k\,\Delta t) = 0 & \\ m^{-k}(k\,\Delta t) = 0. & \end{cases}$$

Note that the variables x^i decouple and it is immediate to obtain

$$x^i(t) = t - i\,\Delta t \qquad \text{for} \qquad t \geq \min\{-i\,\Delta t, 0\} \qquad \text{and} \qquad i = -k, \ldots, n.$$

The discretized version of the cost functional (3.2) is

$$\mathcal{J}^n(u) = \sum_{k=0}^{+\infty} \int_{k\,\Delta t}^{(k+1)\Delta t} \mathrm{e}^{-\lambda\, t} \left(\sum_{i=-k+1}^{n} w(t - i\,\Delta t)\, m_u^i(t) - u(t)\, m_u^{-k}(t) \right) \mathrm{d}t$$

4. Technical details

4.1. Proofs related to Section 2.1

Lemma 4.1. *Fix $T > 0$ and a normed space X. Let $x \in \mathbf{BV}([0,T]; X)$. Then, for any $\varepsilon > 0$, there exists $n \in \mathbb{N}$, $\{t_1, t_2, \ldots, t_n\} \subset [0,T]$ and $\{x_1, x_2, \ldots, x_n\} \in X$ such that, setting $x_\varepsilon(t) = \sum_{i=1}^n x_i\, \chi_{[t_{i-1}, t_i[}(t)$,*

$$\sup_{t\in[0,T]} \left\| x(t) - x_\varepsilon(t) \right\|_X \leq \varepsilon, \quad x_\varepsilon([0,T]) \subseteq x([0,T]) \quad \text{ and } \quad \mathrm{TV}_X(x_\varepsilon) \leq \mathrm{TV}_X(x).$$

Proof. The construction of the function x_ε follows, for instance, from ([1], Thm. 1.2, Chap. 1). The inclusion and the bound on the total variation are immediate, since the x_i are chosen among the values attained by x. □

Proof of Theorem 2.2. On the space $X = \mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R})$, define the norm $\|\cdot\|_X$ as in (2.1) and apply Lemma 4.1 to the maps $b, c, \beta \colon [0,T] \to X$. For every $\varepsilon > 0$, there exists $n \in \mathbb{N}$, $\{t_1, t_2, \ldots, t_n\} \subset [0,T]$ and piecewise constant functions $b_\varepsilon, c_\varepsilon, \beta_\varepsilon \colon [0,T] \to X$ such that

$$\begin{aligned} \sup\nolimits_{t\in[0,T]} \left\| b(t) - b_\varepsilon(t) \right\|_X &\leq \varepsilon, & \mathrm{TV}_X(b_\varepsilon) &\leq \mathrm{TV}_X(b), \\ \sup\nolimits_{t\in[0,T]} \left\| c(t) - c_\varepsilon(t) \right\|_X &\leq \varepsilon, & \mathrm{TV}_X(c_\varepsilon) &\leq \mathrm{TV}_X(c), \\ \sup\nolimits_{t\in[0,T]} \left\| \beta(t) - \beta_\varepsilon(t) \right\|_X &\leq \varepsilon, & \mathrm{TV}_X(\beta_\varepsilon) &\leq \mathrm{TV}_X(\beta). \end{aligned}$$

Moreover, the inclusion proved in Lemma 4.1 ensures that

$$\begin{aligned} \sup_{t\in[0,T]} \left\| b_\varepsilon(t) \right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} &\leq \mathcal{L} + \varepsilon, \\ \sup_{t\in[0,T]} \left\| c_\varepsilon(t) \right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} &\leq \mathcal{L} + \varepsilon, \\ \sup_{t\in[0,T]} \left\| \beta_\varepsilon(t) \right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} &\leq \mathcal{L} + \varepsilon. \end{aligned}$$

By construction, the sequences $b_\varepsilon, c_\varepsilon$ and β_ε converge to b, c and β uniformly on $[0,T]$. Hence, they are all Cauchy sequences.

Fix $\varepsilon > 0$. For all $i = 1, \ldots, n$, ([6], Thm. 2.11), or ([19], Thm. 4.6), ([20], Thm. 1.3), can be recursively applied on the interval $[t_{i-1}, t_i]$ to the problem

$$\begin{cases} \partial_t \mu_i + \partial_x \left(b_\varepsilon(t)(\mu_i, x)\, \mu_i \right) + c_\varepsilon(t)(\mu_i, x)\, \mu_i = 0 \\ b_\varepsilon(t)(\mu_i(t), 0)\, D_\lambda \mu_i(t)(0+) = \int_0^{+\infty} \beta_\varepsilon(t)(\mu_i, \alpha)\, \mathrm{d}\mu_i(t)(\alpha) \\ \mu_i(t_{i-1}) = \mu_{i-1}^o \end{cases}$$

where $\mu_0^o = \mu_o$ and $\mu_i^o = \lim_{t \to t_i -} \mu_{i-1}(t)$ for $i = 1, \ldots, n-1$. Define $\mu^\varepsilon(t)$ by $\mu^\varepsilon(t) = \mu_i(t)$ whenever $t \in [t_{i-1}, t_i[$.

By ([6], (iv) in Thm. 2.8), for any $\varepsilon, \varepsilon' > 0$ sufficiently small,

$$\begin{aligned} d\left(\mu^\varepsilon(t), \mu^{\varepsilon'}(t)\right) \le C\, t\, \mathrm{e}^{C\,t} \Bigg(& \sup_{t \in [0,T]} \left\| b_\varepsilon(t) - b_{\varepsilon'}(t) \right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R})} \\ & + \sup_{t \in [0,T]} \left\| c_\varepsilon(t) - c_{\varepsilon'}(t) \right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R})} \\ & + \sup_{t \in [0,T]} \left\| \beta_\varepsilon(t) - \beta_{\varepsilon'}(t) \right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R})} \Bigg). \end{aligned}$$

Therefore, by the completeness of the space $\mathbf{C^0}\left([0,T]; \mathcal{M}^+(\mathbb{R}_+)\right)$, there exists a measure valued map $\mu \in \mathbf{C^0}\left([0,T]; \mathcal{M}^+(\mathbb{R}_+)\right)$ such that $\lim_{\varepsilon \to 0} \sup_{t \in [0,T]} d\left(\mu^\varepsilon(t), \mu_t\right) = 0$.

To prove that μ solves (1.1) in the sense of Definition 2.1, observe that by construction

$$\begin{aligned} & \int_{\mathbb{R}_+} \varphi(T, x)\, \mathrm{d}\mu^\varepsilon(T)(x) - \int_{\mathbb{R}_+} \varphi(0, x)\, \mathrm{d}\mu_o^\varepsilon(x) \\ = & \int_0^T \int_{\mathbb{R}_+} \partial_t \varphi(t, x)\; \mathrm{d}\mu^\varepsilon(t)(x)\, \mathrm{d}t \\ & + \int_0^T \int_{\mathbb{R}_+} \left(\partial_x \varphi(t, x)\; b_\varepsilon(t)(\mu^\varepsilon(t), x) - \varphi(t, x)\; c_\varepsilon(t)(\mu^\varepsilon(t), x) \right)\; \mathrm{d}\mu^\varepsilon(t)(x)\, \mathrm{d}t \\ & + \int_0^T \int_{\mathbb{R}_+} \varphi(t, 0)\; \beta_\varepsilon(t)(\mu^\varepsilon(t), x)\; \mathrm{d}\mu^\varepsilon(t)(x)\, \mathrm{d}t\,. \end{aligned}$$

and the limit $\varepsilon \to 0$ can pass inside the integral sign thanks to the uniform convergences $\mu^\varepsilon \to \mu$, $b_\varepsilon \to b$, $c_\varepsilon \to c$ and $\beta_\varepsilon \to \beta$ on the time interval $[0, T]$.

A further application of ([6], (iv) in Thm. 2.1) proves the stability estimate (2.5). □

Proof of Corollary 2.3. Note first that if $b, c, \beta \in \mathcal{F}^u$ and $u \in \mathbf{BV}([0,T]; \mathcal{U})$, then the maps b^u, c^u, β^u defined by $b^u(t) = b\left(t, u(t)\right)$, $c^u(t) = c\left(t, u(t)\right)$ and $\beta^u(t) = \beta\left(t, u(t)\right)$ all satisfy $b^u, c^u, \beta^u \in \mathcal{F}$. Therefore, Theorem 2.2 applies, ensuring the existence of a solution to (1.2).

Concerning the stability estimates, with obvious notations, by (2.5) we have:

$$\begin{aligned} d\left(\mu^1(t), \mu^2(t)\right) \le d(\mu_o^1, \mu_o^2)\, \mathrm{e}^{C\,t} + C\, t\, \mathrm{e}^{C\,t} \Bigg(& \sup_{t \in [0,T]} \left\| b_1^{u_1}(t) - b_2^{u_2}(t) \right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R})} \\ & + \sup_{t \in [0,T]} \left\| c_1^{u_1}(t) - c_2^{u_2}(t) \right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R})} \\ & + \sup_{t \in [0,T]} \left\| \beta_1^{u_1}(t) - \beta_2^{u_2}(t) \right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+) \times \mathbb{R}_+; \mathbb{R})} \Bigg). \end{aligned}$$

Observe now that

$$\begin{aligned}\left\|b_1^{u_1}(t)-b_2^{u_2}(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} &\le \left\|b_1^{u_1}(t)-b_1^{u_2}(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}+\left\|b_1^{u_2}(t)-b_2^{u_2}(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}\\ &\le \mathcal{L}_u\left\|u_1(t)-u_2(t)\right\|+\left\|b_1(t)-b_2(t)\right\|_{\mathbf{C^0}(\mathcal{U}\times\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}.\end{aligned}$$

Entirely analogous estimates can be proved for the term $\left\|c_1^{u_1}(t)-c_2^{u_2}(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}$ as well as for the term $\left\|\beta_1^{u_1}(t)-\beta_2^{u_2}(t)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}$, allowing to obtain (2.6). □

4.2. Proofs related to Section 2.2

Aiming at the well posedness of (2.7)–(2.8) we rewrite it as

$$\begin{cases}\dot{x}=f(t,x,m,u)\\ \dot{m}=g(t,x,m,u)\\ x(0)=x_o\\ m(0)=m_o\end{cases}\tag{4.1}$$

where

$$x=(x^{-n},\ldots,x^n) \qquad\qquad m=(m^{-n},\ldots,m^n)$$

$$x_o^i=\begin{cases} i\,\Delta t & i=0,\ldots,n\\ 0 & i=-n,\ldots,-1\end{cases} \qquad\qquad m_o^i=\begin{cases}\mu_o\left(\left[i\,\Delta t,(i+1)\Delta t\right[\right) & i=0,\ldots,n\\ 0 & i=-n,\ldots,-1\end{cases}$$

$$f\,:\,[0,T]\times\mathbb{R}_+^{2n+1}\times\mathbb{R}_+^{2n+1}\times\mathcal{U}\to\mathbb{R}^{2n+1} \qquad\qquad g\,:\,[0,T]\times\mathbb{R}_+^{2n+1}\times\mathbb{R}_+^{2n+1}\times\mathcal{U}\to\mathbb{R}^{2n+1}$$

the functions f_i, g_i being defined, for $i=-n,\ldots,n$, by

$$f_i(t,x,m;u)=\begin{cases} b(t,u)\left(\sum_{j=-n}^{n} m^j\delta_{x^j},x^i\right) & t\ge\max\{-i\,\Delta t,\,0\}\\ 0 & t<\max\{-i\,\Delta t,\,0\}\end{cases}\tag{4.2}$$

and

$$g_i(t,x,m;u)=\begin{cases} -c(t,u)\left(\sum_{j=-n}^{n} m^j\delta_{x^j},x^i\right)\, m^i & t>\max\{(1-i)\,\Delta t,\,0\}\\ -c(t,u)\left(\sum_{j=-n}^{n} m^j\delta_{x^j},x^i\right)\, m^i & \\ \quad+\sum_{\ell=-n}^{n}\beta(t,u)\left(\sum_{j=-n}^{n} m^j\delta_{x^j},x^\ell\right)\, m^\ell & t\in[\max\{-i\Delta t,0\},\max\{(1-i)\Delta t,0\}]\\ 0 & t<\max\{-i\,\Delta t,\,0\}\end{cases}\tag{4.3}$$

Lemma 4.2. *Fix positive T, L and let $b,c,\beta\in\tilde{\mathcal{F}}^u$. Then, the map f and g defined in (4.2) and (4.3) satisfy the following conditions*:

($\mathbf{f_1}$) $t\to(f,g)(t,x,m;u)$ *is measurable for all* $x\in\mathbb{R}_+$, $m\in\mathbb{R}_+$ *and* $u\in\mathcal{U}$;
($\mathbf{f_2}$) $(x,m;u)\to(f,g)(t,x,m;u)$ *is in* $\mathbf{C^1}$ *for a.e.* $t\in[0,T]$;
($\mathbf{f_3}$) $(t,x,m)\to(f,g)(t,x,m;u)$ *is sublinear in* (x,m), *uniformly in* t *and for all* u.

Proof. We detail the proof that f satisfies the above properties, the case of g being entirely similar.

The measurability of $t \to f(t,x,y;u)$ is immediate. To verify the differentiability, introduce the standard base $(e_{-n}, e_{-n+1} \dots, e_{n-1}, e_n)$ of $\mathbb{R}^{2n+1}$ and compute for $i = -n, \dots, n$, $t > \max\{-i\Delta t; 0\}$ and for a (small) $h \in \mathbb{R}$

$$\begin{aligned}
&f_i(t, x + he_i, m; u) - f_i(t, x, m; u)\\
&= b(t,u)\left(\sum_{j=-n}^{n} m^j \delta_{x^j + \delta_{ij} he_j}, x^i + he^i\right) - b(t,u)\left(\sum_{j=-n}^{n} m^j \delta_{x^j}, x^i\right)\\
&= \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j + \delta_{ij} he_i), x^i + he_i; u\right) - \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j), x^i; u\right)\\
&= \partial_y \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j), x^i; u\right) m^i \bar{b}'(x^i) h + \partial_x \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j), x^i; u\right) h + o(h)
\end{aligned}$$

as $h \to 0$, while for $\ell \neq i$ and for $t > \max\{-i\Delta t; -\ell\Delta t, 0\}$

$$\begin{aligned}
&f_i(t, x + he_\ell, m; u) - f_i(t, x, m; u)\\
&= b(t,u)\left(\sum_{j=-n}^{n} m^j \delta_{x^j + \delta_{\ell j} he_\ell}, x^i\right) - b(t,u)\left(\sum_{j=-n}^{n} m^j \delta_{x^j}, x^i\right)\\
&= \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j + \delta_{\ell j} he_\ell), x^i; u\right) - \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j), x^i; u\right)\\
&= \partial_y \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j), x^i; u\right) m^\ell \bar{b}'(x^\ell) h + o(h) \text{ as } h \to 0,
\end{aligned}$$

proving the differentiability of f_i with respect to x. Let now $i, \ell = -n, \dots, n$:

$$\begin{aligned}
&f_i(t, x, m + he_\ell; u) - f_i(t, x, m; u)\\
&= b(t,u)\left(\sum_{j=-n}^{n} (m^j + \delta_{\ell j} he_\ell)\delta_{x^j}, x^i\right) - b(t,u)\left(\sum_{j=-n}^{n} m^j \delta_{x^j}, x^i\right)\\
&= \tilde{b}\left(t, \sum_{j=-n}^{n} (m^j + \delta_{\ell j} he_\ell)\, \bar{b}(x^j), x^i; u\right) - \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^j), x^i; u\right)\\
&= \partial_y \tilde{b}\left(t, \sum_{j=-n}^{n} m^j\, \bar{b}(x^i), x^i; u\right) \bar{b}(x^\ell) h + o(h) \text{ as } h \to 0,
\end{aligned}$$

so that f_i is differentiable also with respect to m. The differentiability with respect to u is immediate.

Finally, we prove that $(t, x, m) \to (f, g)(t, x, m; u)$ is sublinear in (x, m), uniformly in t and for all u:

$$\begin{aligned}
\left|f_i(t,x,m;u)\right| &\le \left|f_i(t,0,0;u)\right| + \left|f_i(t,x,m;u) - f_i(t,0,0;u)\right| \\
&= \left|\left(b(t,u)(0,0)\right)\right| + \left|b(t,u)\left(\sum_{j=-n}^{n} m^j \delta_{x^j}, x^i\right) - b(t,u)(0,0)\right| \\
&= \left|\tilde b(t,0,0;u)\right| + \left|\tilde b\left(t, \sum_{j=-n}^{n} m^j \bar b(x^j), x^i; u\right) - \tilde b(t,0,0;u)\right| \\
&\le \left|\tilde b\,(t,0,0;u)\right| + L\left(\left|\sum_{j=-n}^{n} m^j \bar b(x^j)\right| + x^i\right) \\
&\le \left\|\tilde b\,(\cdot,0,0;u)\right\|_{\mathbf{L}^\infty([0,T];\mathbb{R}_+)} + L\sqrt{2n+1}\left(\|m\|_{\mathbb{R}^{2n+1}} + \|x\|_{\mathbb{R}^{2n+1}}\right)
\end{aligned}$$

and the first summand above is bounded by $(\tilde{\mathcal{F}}_3^u)$, completing the proof. □

Below, we call *semiflow* (or process) on the set M a map $S\colon M \times [0,\delta] \times [0,T] \to M$ such that $S(0,t) = \mathbf{Id}$ for all $t \in [0,T]$, and $S(t_3, t_1 + t_2) \circ S(t_2, t_1) = S(t_2 + t_3, t_1)$ for all t_1, t_2, t_3 such that $t_1, t_1 + t_2 \in [0,T]$, $t_2, t_3, t_2 + t_3 \in [0,\delta]$. The semiflow is Lipschitz continuous if the map $\mu \to S(t, t_o)\mu$ is Lipschitz continuous, uniformly in $t_o \in [0,T]$ and in $t \in [0,\delta]$.

Lemma 4.3. *Let (M, d_M) be a metric space and $S\colon M \times [0,\delta] \times [0,T] \to M$ a Lipschitz semiflow with Lipschitz constant L. For every Lipschitz continuous map $\mu\colon [0,T] \to M$, the following estimate holds:*

$$d_M\left(\mu(t), S(t,0,\mu_0)\right) \le L \int_0^t \liminf_{h\to 0+} \frac{d_M\left(\mu(\tau+h), S(h,\tau,\mu(\tau))\right)}{h}\, \mathrm{d}\tau \tag{4.4}$$

For a proof, see ([4], Thm. 2.9) or, in the present non autonomous case, ([17], Prop. 4.1 or [8], Proof of Thm. 3.15).

Lemma 4.4 ([17], Lem. 7.3)**.**
Let $n \in \mathbb{N}$, $m, m' \in \mathbb{R}^n$ and $x, x' \in \mathbb{R}^n$. Then, with reference to the distance d defined in (2.4),

$$d\left(\sum_{i=1}^{n} m_i \delta_{x_i}, \sum_{i=1}^{n} m'_i \delta_{x'_i}\right) \le \max\left\{1, \sum_{i=1}^{n} |m_i|\right\} \sum_{i=1}^{n} \left(\left|m_i - m'_i\right| + \left|x_i - x'_i\right|\right).$$

Proof of Theorem 2.5*.* The proof relies on Lemma 4.4. First, we prove that the map

$$\begin{aligned}
\mu^n \;:\; [0,T] &\to \mathcal{M}^+(\mathbb{R}_+) \\
t &\to \sum_{i=-n}^{n} m^i(t)\, \delta_{x_i(t)},
\end{aligned} \tag{4.5}$$

where $t \to (x^i, m^i)(t)$ solves (4.1)–(4.2)–(4.3), is Lipschitz continuous with respect to the metric d defined in (2.4). Indeed, by Lemma 4.4

$$\begin{aligned} d\left(\mu^n(t), \mu^n(s)\right) &\le \max\left\{1, \sum_{i=-n}^{n} \left|m^i(t)\right|\right\} \sum_{i=-n}^{n} \left(\left|m^i(t) - m^i(s)\right| + \left|x^i(t) - x^i(s)\right|\right) \\ &\le \max\left\{1, \sum_{i=-n}^{n} \left|m^i(t)\right|\right\} \sum_{i=-n}^{n} \left(\mathbf{Lip}(x_i) + \mathbf{Lip}(m_i)\right)\ (t-s). \end{aligned} \tag{4.6}$$

Moreover,

$$\begin{aligned} \mathbf{Lip}(x^i) &\le \|f_i\|_{\mathbf{C^0}([0,T]\times\mathbb{R}_+\times\mathbb{R}_+\times\mathcal{U};\mathbb{R})} && \text{[by (4.1)]} \\ &\le \sup_{(t,u)\in[0,T]\times\mathcal{U}} \left\|b(t,u)\right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} && \text{[by (4.2)]} \\ &\le \mathcal{L} && \text{[by } (\mathcal{F}_1^u)\text{]} \\ \left|m^i(t)\right| &\le |m_o| \exp\left(T \sup_{(t,u)\in[0,T]\times\mathcal{U}} \left\|c(t,u)\right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}\right) \\ &\qquad \times \exp\left((2n+1)T \sup_{(t,u)\in[0,T]\times\mathcal{U}} \left\|\beta(t,u)\right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}\right) && \text{[by (4.3)]} \\ &\le |m_o| \mathrm{e}^{\left(2(n+1)\mathcal{L}\right)T} && \text{[by } (\mathcal{F}_1^u)\text{]} \\ \mathbf{Lip}(m^i) &\le \|g_i\|_{\mathbf{C^0}([0,T]\times\mathbb{R}_+\times\mathbb{R}_+\times\mathcal{U};\mathbb{R})} && \text{[by (4.1)]} \\ &\le \sup_{(t,u)\in[0,T]\times\mathcal{U}} \left\|c(t,u)\right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \sup_{t\in[0,T]} \left|m^i(t)\right| \\ &\qquad + (2n+1)\left\|\beta(t,u)\right\|_{\mathbf{C^{0,1}}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \sup_{t\in[0,T]} \left|m^i(t)\right| && \text{[by (4.3)]} \\ &\le 2(n+1)\mathcal{L}|m_o| \mathrm{e}^{\left(2(n+1)\mathcal{L}\right)T} && \text{[by } (\mathcal{F}_1^u)\text{]} \end{aligned}$$

These estimates, inserted in (4.6), complete the proof of the Lipschitz continuity of μ^n with respect to the metric d defined in (2.4).

By the above computations and Corollary 2.3, we can thus use Lemma 4.4, where S is the semiflow generated by (1.2) and μ is replaced by μ^n as defined in (4.5), obtaining

$$\begin{aligned} d\left(\sum_{i=-n}^{n} m^i(t)\,\delta_{x^i(t)}, \mu_u(t)\right) &= d\left(\mu^n(t), S(t,0)\mu_o\right) \\ &\le d\left(\mu^n(t), S(t,0)\mu^n(0)\right) + d\left(S(t,0)\mu^n(0), S(t,0)\mu_o\right) \\ &\le \max\left\{\mathbf{Lip}(S), \mathrm{e}^{Ct}\right\} \left[\int_0^t \liminf_{h\to 0+} \frac{d(\mu^n_{\tau+h}, S(h,\tau)\mu^n_\tau)}{h}\,\mathrm{d}\tau + d(\mu^n(0), \mu_o)\right]. \end{aligned} \tag{4.7}$$

The rest of the proof is devoted to estimate the integrand in the latter term above.

Without loss of generality, we may assume that $\tau \in [0, \Delta t[$ and that h is so small that $[\tau, \tau+h] \subset [0, \Delta t[$. Define

$$\mu_\tau(t) := S(t,\tau)\,\mu^n(\tau).$$

Then, for $t \in [\tau, \Delta t[$, the map $t \to \mu_\tau(t)$ solves problem (1.2) with initial datum $\mu^n(\tau) = \sum_{i=-n}^{n} m^i(\tau)\,\delta_{x^i(\tau)} = \sum_{i=0}^{n} m^i(\tau)\,\delta_{x^i(\tau)}$ assigned at time τ.

As in ([17], Proof of Thm. 4.3), $\mu_\tau(t)$ can be represented as

$$\mu_\tau(t) = \sum_{i=0}^{n} M^i(\tau+h)\,\delta_{y^i(\tau+h)} + \pi(t,\cdot)\,\mathrm{d}x$$

for suitable maps $M^0, \ldots, M^n$, $y^0, \ldots, y^n$. Here, the density $\pi(t,\cdot)$ is the absolutely continuous part $\mu_\tau(t)$ which arises from the boundary terms due to the non local part in (1.2) and supported inside $[x_o^0, y^0(t)]$. Denote the total mass of $\pi(t,\cdot)$ by

$$M^\pi(t) = \int_{x_o^0}^{y^0(t)} \pi(t,x)\,\mathrm{d}x\,.$$

Using the implicit representation formula for the solution to the transport equation in spaces of measures [26], we obtain:

$$\begin{aligned}
y^i(\tau+h) &= x^i(\tau) + \int_\tau^{\tau+h} b(t,u)\left(\mu_\tau(t-\tau), y^i(t)\right)\mathrm{d}t\\
M^i(\tau+h) &= m^i(\tau) + \int_\tau^{\tau+h} c(t,u)\left(\mu_\tau(t-\tau), y^i(t)\right) M^i(t)\,\mathrm{d}t\\
M^\pi(\tau+h) &= M^\pi(\tau) + \int_\tau^{\tau+h}\Bigg(\int_{x_o^0}^{y^0(t)} -c(t,u)(\mu_\tau(t-\tau),x)\;\mathrm{d}\mu_\tau(t-\tau)(x)\\
&\qquad + \int_{x_o^0}^{+\infty}\beta(t,u)(\mu_\tau(t-\tau),x)\;\mathrm{d}\mu_\tau(t-\tau)(x)\Bigg)\,\mathrm{d}t\\
&= M^\pi(\tau) + \int_\tau^{\tau+h}\int_{x_o^0}^{y^0(t)}\Big[\Big(-c(t,u)(\mu_\tau(t-\tau),x) + \beta(t,u)(\mu_\tau(t-\tau),x)\Big)\,\mathrm{d}\mu_\tau(t-\tau)(x)\Big]\,\mathrm{d}t\\
&\qquad + \sum_{i=0}^n\int_\tau^{\tau+h}\beta(t,u)\left(\mu_\tau(t-\tau), y^i(t)\right)\;M^i(t)\,\mathrm{d}t\\
&\le M^\pi(\tau) + 2\mathcal{L}\int_\tau^{\tau+h}\int_{x_o^0}^{y^0(t)}\pi(t,x)\,\mathrm{d}x\,\mathrm{d}t + \sum_{i=0}^n\int_\tau^{\tau+h}\beta(t,u)\left(\mu_\tau(t-\tau), y^i(t)\right)M^i(t)\,\mathrm{d}t\\
&= M^\pi(\tau) + \mathcal{O}(h^2) + \sum_{i=0}^n\int_\tau^{\tau+h}\beta(t,u)\left(\mu_\tau(t-\tau), y^i(t)\right)M^i(t)\,\mathrm{d}t\;,
\end{aligned}$$

where with $\mathcal{O}(h^k)$ we denote a quantity that can be bounded by the product of h^k with a constant dependent only on $T, \mathcal{L}$ and $\mathcal{C}$.

Above, $(\mathcal{F}_1^u)$ ensures a bound on c and β. We also used the uniform boundedness of $\pi(t,\cdot)$ on $[0,T]$ and the estimate

$$\left|y^0(t) - x_o^0\right| \le \sup_{t\in[0,T]}\sup_{u\in\mathcal{U}}\left\|b(t,u)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})}\,h \le \mathcal{L}\,h$$

for $t\in[\tau,\tau+h)$. For $t\in[0,\Delta t[$ define the time dependent measure

$$\xi(t) = \sum_{i=0}^n p^i(t)\,\delta_{y^i(t)} \quad\text{where}\quad \begin{cases} p^o(t) = M^o(t) + M^\pi(t),\\ p^i(t) = M^i(t), \quad \text{for } i = 1,\ldots,n\end{cases} \tag{4.8}$$

in other words, in the measure $\xi(t)$ the mass created due to the boundary condition, described by the density $\pi(t,\cdot)$, is shifted to the closest Dirac delta. We note that:

$$d\left(\mu_\tau(t),\mu^n(t+\tau)\right) \le d\left(\mu_\tau(t),\xi(t+\tau)\right) + d\left(\xi(t+\tau),\mu^n(t+\tau)\right). \tag{4.9}$$

Recalling that $t\to y^i(t)$ is Lipschitz continuous with Lipschitz constant

$$\mathbf{Lip}(y^i) \le \sup_{t\in[0,T]}\sup_{u\in\mathcal{U}}\left\|b(t,u)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \le \mathcal{L}$$

and that the total mass is uniformly bounded on $[0,T]$, the first term in the right hand side of (4.9) is estimated as follows:

$$\begin{aligned}
d\left(\mu_\tau(h), \xi(\tau+h)\right) &= d\left(\pi(\tau+h,\cdot), M^\pi(\tau+h)\,\delta_{y^0(\tau+h)}\right)\\
&\le \left|y^0(\tau+h)\right| M^\pi(\tau+h)\\
&\le \mathcal{L}\,\Delta t\left[\sup_{t\in[0,T],u\in\mathcal{U}} \left\|\beta(t,u)\right\|_{\mathbf{C^0}(\mathcal{M}^+(\mathbb{R}_+)\times\mathbb{R}_+;\mathbb{R})} \int_\tau^{\tau+h} \sum_{i=0}^{n} M^i(t)\,\mathrm{d}t + \mathcal{O}(h^2)\right]\\
&\le \mathcal{L}\,\Delta t\left(\mathcal{L}\,C(T)\,h + \mathcal{O}(h^2)\right)\\
&= \Delta t\left(\mathcal{O}(h) + \mathcal{O}(h^2)\right).
\end{aligned}$$

To bound the second term in (4.9), we want to use Lemma 4.4. Hence, we preliminary obtain the following estimates on $\left|x^i(\tau+h)-y^i(\tau+h)\right|$ and $\left|m^i(\tau+h)-p^i(\tau+h)\right|$:

$$\begin{aligned}
\left|x^i(\tau+h)-y^i(\tau+h)\right| &\le \int_\tau^{\tau+h} \left|b(t,u)\left(\mu^n(t),x^i(t)\right) - b(t,u)\left(\mu_\tau(t-\tau),y^i(t)\right)\right|\mathrm{d}t\\
&\le \int_\tau^{\tau+h} \left|b(t,u)\left(\mu^n(t),x^i(t)\right) - b(t,u)\left(\mu_\tau(t-\tau),x^i(t)\right)\right|\mathrm{d}t\\
&\quad + \int_\tau^{\tau+h} \left|b(t,u)\left(\mu_\tau(t-\tau),x^i(t)\right) - b(t,u)\left(\mu_\tau(t-\tau),y^i(t)\right)\right|\mathrm{d}t\\
&\le \mathcal{L}\int_\tau^{\tau+h} d\left(\mu^n(t),\mu_\tau(t-\tau)\right)\mathrm{d}t + \mathcal{L}\int_\tau^{\tau+h}\left|x^i(t)-y^i(t)\right|\mathrm{d}t\\
&\le \mathcal{L}\int_\tau^{\tau+h}\left(\underset{\tau}{\mathbf{Lip}}(\mu^n)\,h + d\left(\mu^n(\tau),\mu_\tau(0)\right) + \underset{\tau}{\mathbf{Lip}}\left(\mu_\tau(t)\right)\,h\right)\mathrm{d}t\\
&\quad + \int_\tau^{\tau+h}\left(\underset{\tau}{\mathbf{Lip}}(x^i)\,h + \left|x^i(\tau)-y^i(\tau)\right| + \underset{\tau}{\mathbf{Lip}}(y^i)\,h\right)\mathrm{d}t\\
&\le \mathcal{O}(h^2),
\end{aligned}$$

since $\mu^n(\tau) = \mu_\tau(0)$ and $x^i(\tau) = y^i(\tau)$. Entirely analogous estimates can be used to bound the term $\sum_{i=0}^n \left|m^i(\tau+h)-p^i(\tau+h)\right|$, taking into account (4.8) and the estimate for $M^\pi(t)$:

$$\begin{aligned}
&\sum_{i=0}^{n}\left|m^i(\tau+h)-p^i(\tau+h)\right|\\
&\le \sum_{i=0}^{n}\int_\tau^{\tau+h}\left|c(t,u)\left(\mu^n(t),x^i(t)\right)\,m^i(t) - c(t,u)\left(\mu_\tau(t-\tau),y^i(t)\right)\,M^i(t)\right|\mathrm{d}t\\
&\quad + \sum_{i=0}^{n}\int_\tau^{\tau+h}\left|\beta(t,u)\left(\mu^n(t),x^i(t)\right)\,m^i(t) - \beta(t,u)\left(\mu_\tau(t-\tau),y^i(t)\right)\,M^i(t)\right|\mathrm{d}t + \mathcal{O}(h^2)\\
&\le \sum_{i=0}^{n}\int_\tau^{\tau+h}\left(\left|c(t,u)\left(\mu^n(t),x^i(t)\right)\right| + \left|\beta(t,u)\left(\mu^n(t),x^i(t)\right)\right|\right)\left|m^i(t)-M^i(t)\right|\mathrm{d}t\\
&\quad + \sum_{i=0}^{n}\int_\tau^{\tau+h} M^i(t)\left|c(t,u)\left(\mu^n(t),x^i(t)\right) - c(t,u)\left(\mu^n(t),y^i(t)\right)\right|\mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=0}^{n}\int_{\tau}^{\tau+h} M^i(t)\left|c(t,u)\Big(\mu^n(t),y^i(t)\Big)-c(t,u)\Big(\mu_\tau(t-\tau),y^i(t)\Big)\right|\mathrm{d}t\\
&+\sum_{i=0}^{n}\int_{\tau}^{\tau+h} M^i(t)\left|\beta(t,u)\Big(\mu^n(t),x^i(t)\Big)-\beta(t,u)\Big(\mu^n(t),y^i(t)\Big)\right|\mathrm{d}t\\
&+\sum_{i=0}^{n}\int_{\tau}^{\tau+h} M^i(t)\left|\beta(t,u)\Big(\mu^n(t),y^i(t)\Big)-\beta(t,u)\Big(\mu_\tau(t-\tau),y^i(t)\Big)\right|\mathrm{d}t+\mathcal{O}(h^2)\\
\le{}& \left\|(c,\beta)\right\|_{\mathbf{C^0}}\sum_{i=0}^{n}\int_{\tau}^{\tau+h}\left|m^i(t)-M^i(t)\right|\mathrm{d}t+\left\|(c,\beta)\right\|_{\mathbf{C^{0,1}}}\sum_{i=0}^{n}\int_{\tau}^{\tau+h}M^i(t)\left|x^i(t)-y^i(t)\right|\mathrm{d}t\\
&+\left\|(c,\beta)\right\|_{\mathbf{C^{0,1}}}\sum_{i=0}^{n}\int_{\tau}^{\tau+h}M^i(t)\;d\left(\mu^n(t),\mu_\tau(t-\tau)\right)\mathrm{d}t+\mathcal{O}(h^2)\\
\le{}& \left\|(c,\beta)\right\|_{\mathbf{C^0}}\sum_{i=0}^{n}\int_{\tau}^{\tau+h}\Big(\mathbf{Lip}(m^i)h+\left|m^i(\tau)-M^i(\tau)\right|+\mathbf{Lip}(M^i)h\Big)\mathrm{d}t\\
&+\left\|(c,\beta)\right\|_{\mathbf{C^{0,1}}}\sum_{i=0}^{n}\int_{\tau}^{\tau+h}M^i(t)\Big(\mathbf{Lip}(x^i)h+\left|x^i(\tau)-y^i(\tau)\right|+\mathbf{Lip}(y^i)h\Big)\mathrm{d}t\\
&+\left\|(c,\beta)\right\|_{\mathbf{C^{0,1}}}\sum_{i=0}^{n}\int_{\tau}^{\tau+h}M^i(t)\Big(\mathbf{Lip}(\mu^n)h+d\left(\mu^n(\tau),\mu_\tau(0)\right)+\mathbf{Lip}(\mu_\tau)h\Big)\mathrm{d}t+\mathcal{O}(h^2)\\
\le{}& \left\|(c,\beta)\right\|_{\mathbf{C^0}}h^2\sum_{i=0}^{n}\Big(\mathbf{Lip}(m^i)+\mathbf{Lip}(M^i)\Big)\\
&+\left\|(c,\beta)\right\|_{\mathbf{C^{0,1}}}h^2\left(2\|b\|_{\mathbf{C^0}}+\mathbf{Lip}(\mu^n)+\mathbf{Lip}(\mu_\tau)\right)\sum_{i=0}^{n}M^i(t)+\mathcal{O}(h^2)\\
={}&\mathcal{O}(h^2).
\end{aligned}
$$

Inserting the obtained estimates in the integrand in (4.7), we get:

$$
\liminf_{h\to0^+}\frac{1}{h}\,d\left(\mu_\tau(h),\mu^n(\tau+h)\right)\le\liminf_{h\to0^+}\frac{1}{h}\left(\Delta t\left(\mathcal{O}(h)+\mathcal{O}(h^2)\right)+\mathcal{O}(h^2)\right)=\mathcal{O}(1)\;\Delta t,
$$

completing the proof. □

4.3. Proofs related to Section 2.3

Lemma 4.5. *Fix $T>0$ and for all $u\in\mathbf{BV}([0,T];\mathcal{U})$ call μ_u the corresponding solution to (1.2). Assume there exists a constant $\mathcal{L}$ such that for all u_1, $u_2\in\mathbf{BV}([0,T];\mathcal{U})$*

$$
d(\mu_{u_1},\mu_{u_2})\le\mathcal{L}\;\|u_1-u_2\|_{\mathbf{L}^\infty([0,T];\mathbb{R})}.
$$

Let $\gamma\in\mathbf{C^{0,1}}(\mathbb{R}_+;\mathbb{R}_+)$ and j satisfy $(\mathbf{J_1})$–$(\mathbf{J_3})$. Then, the functional $\mathcal{J}$ defined in (2.9)–(2.10) is lower semi-continuous with respect to the $\mathbf{L}^\infty$–norm.

Proof. Let $u_n\in\mathbf{BV}([0,T];\mathcal{U})$ be a sequence converging to $u\in\mathbf{BV}([0,T];\mathcal{U})$ in the $\mathbf{L}^\infty$-norm. First we recall that by ([12], Thm. 1, Sect. 5.2.1) $\mathrm{TV}_{\mathbb{R}^N}(u)\le\liminf_{n\to\infty}\mathrm{TV}_{\mathbb{R}^N}(u_n)$.

Next we show the sequential continuity of the map $\widetilde{\mathcal{J}}$ defined in (2.9), using (**J3**) and the fact that ω is a nondecreasing function by ($\mathbf{J_3}$).

$$\begin{aligned}
\left|\widetilde{\mathcal{J}}(u_n) - \widetilde{\mathcal{J}}(u)\right| &\le \int_0^T \left| j\left(t, u_n(t), \int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_{u_n}(t)(\xi)\right) - j\left(t, u(t), \int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_u(t)(\xi)\right)\right| \mathrm{d}t \\
&\le \int_0^T L(t)\,\omega\left(\left|\int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_{u_n}(t)(\xi) - \int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_u(t)(\xi)\right| + \left|u_n(t) - u(t)\right|\right) \mathrm{d}t \\
&\le \omega\left(\|\gamma\|_{\mathbf{L}^\infty(\mathbb{R}_+;\mathbb{R}_+)} \sup_{t\in[0,T]} d\left(\mu_{u_n}(t), \mu_u(t)\right) + \|u_n - u\|_{\mathbf{L}^\infty([0,T];\mathbb{R}^N)}\right) \int_0^T L(t)\,\mathrm{d}t \\
&\le \omega\left((1 + \mathcal{L}\|\gamma\|_{\mathbf{L}^\infty(\mathbb{R}_+;\mathbb{R}_+)})\|u_n - u\|_{\mathbf{L}^\infty([0,T];\mathbb{R}^N)}\right) \int_0^T L(t)\,\mathrm{d}t \\
&\to 0 \quad \text{in } \mathbf{L}^\infty([0,T];\mathbb{R}^N) \text{ as } n \to +\infty,
\end{aligned}$$

completing the proof. □

Proof of Theorem 2.6. Note that if $u \in \mathbf{BV}([0,T];\mathcal{U})$ then $\mathcal{J}(u) < \infty$, and therefore $\inf_u \mathcal{J}(u) < \infty$. Let ε_n be a strictly decreasing sequence converging to 0. Correspondingly, there exists a sequence $u_{\varepsilon_n} \in \mathbf{BV}([0,T];\mathcal{U})$ such that

$$\mathcal{J}(u_{\varepsilon_n}) \le \inf_u \mathcal{J}(u) + \varepsilon_n$$

and, without loss of generality, we may also assume that $\mathcal{J}(u_{\varepsilon_n}) \le \inf_u \mathcal{J}(u) + 1$ for all n. Moreover, by ($\mathbf{J_1}$) and (2.10)

$$\mathrm{TV}(u_{\varepsilon_n}) \le \mathcal{J}(u_{\varepsilon_n}) \le \inf_u \mathcal{J}(u) + 1$$

So that Helly Theorem *e.g.*, ([4], Chap. 2, Thm. 2.3) can be applied, showing that, up to a subsequence, u_{ε_n} converges pointwise and in $\mathbf{L}^p$, for every $p < \infty$, to a function $u^* \in \mathbf{BV}([0,T];\mathcal{U})$. Note that Lemma 4.5 can be applied, since the Lipschitz continuity of $u \to \mu_u$ is proved in ([20], Thm. 3.1). Therefore,

$$\begin{aligned}
\mathcal{J}(u^*) &= \mathcal{J}(\lim_{n\to+\infty} u_{\varepsilon_n}) && [\text{by the definition of } u^*] \\
&\le \liminf_{n\to\infty} \mathcal{J}(u_{\varepsilon_n}) && [\text{by Lemma 4.5}] \\
&= \inf_{u\in\mathbf{BV}([0,T];\mathcal{U})} \mathcal{J}(u) && [\text{by the definition of } u_{\varepsilon_n}],
\end{aligned}$$

completing the proof. □

Proof of Theorem 2.7. The proof follows the same lines as that of Theorem 2.6. □

Proof of Corollary 2.8. We first prove the uniform convergence $\mathcal{J}_n \to \mathcal{J}$ of the costs on $\mathbf{BV}([0,T];\mathcal{U})$, using ($\mathbf{J_3}$), (2.9), (2.10), (2.11), (2.12) and Theorem 2.5, for all n, we have:

$$\begin{aligned}
&\sup_{u\in\mathbf{BV}([0,T];\mathcal{U})} \left|\mathcal{J}_n(u) - \mathcal{J}(u)\right| \\
&= \sup_{u\in\mathbf{BV}([0,T];\mathcal{U})} \left|\widetilde{\mathcal{J}}_n(u) - \widetilde{\mathcal{J}}(u)\right| \\
&\le \sup_{u\in\mathbf{BV}([0,T];\mathcal{U})} \int_0^T \left| j\left(t, u(t), \int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_u^n(t)(\xi)\right) - j\left(t, u(t), \int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_u(t)(\xi)\right)\right| \mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
&\le \sup_{u\in \mathbf{BV}([0,T];\mathcal{U})} \int_0^T L(t)\, \omega\left(\left|\int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_u^n(t)(\xi) - \int_{\mathbb{R}_+} \gamma(\xi)\,\mathrm{d}\mu_u(t)(\xi)\right|\right) \mathrm{d}t \\
&\le \int_0^T L(t)\,\mathrm{d}t \sup_{u\in \mathbf{BV}([0,T];\mathcal{U})} \omega\left(\|\gamma\|_{L^\infty} \sup_{t\in[0,T]} d\left(\mu_u^n(t), \mu_u(t)\right)\right) \\
&\le \int_0^T L(t)\,\mathrm{d}t\ \omega\left(C\,\|\gamma\|_{L^\infty}\left[\Delta t_n + d\left(\mu_o, \sum_{i=0}^n m_o^i\,\delta_{x_o^i}\right)\right]\right) \\
&\to 0 \quad \text{as } n \to +\infty,
\end{aligned}
$$

which immediately implies (2.13).

Using ($\mathbf{J_2}$), the same procedure used in the proof of Theorem 2.6 ensures that $\mathrm{TV}(u_n)$ is bounded uniformly in n. By Helly Theorem (*e.g.*, [4], Chap. 2, Thm. 2.3), up to a subsequence, $u_n \to \bar{u}$ a.e. on $[0,T]$, proving (2.14). Using Lemma 4.5, which can be applied thanks to ([20], Thm. 3.1), and the uniform convergence of $\mathcal{J}_n$ to $\mathcal{J}$ proved above, we have:

$$
\begin{aligned}
\mathcal{J}(\bar{u}) &\le \liminf_{n\to\infty} \mathcal{J}(u_n) \\
&= \liminf_{n\to\infty} \left(\mathcal{J}(u_n) + \mathcal{J}_n(u_n) - \mathcal{J}_n(u_n)\right) \\
&\le \liminf_{n\to\infty} \mathcal{J}_n(u_n) + \lim_{n\to\infty}\left(\sup_{u\in \mathbf{BV}([0,T];\mathcal{U})} \left|\mathcal{J}_n(u) - \mathcal{J}(u)\right|\right) \\
&\le \lim_{n\to\infty} \mathcal{J}_n(u_n) \\
&= \inf_{u\in \mathbf{BV}([0,T];\mathcal{U})} \mathcal{J}(u),
\end{aligned}
$$

where (2.13) was used to obtain the last equality. □

Appendix A. ODE Results

For completeness, we collect here a few basic ODE results using exactly the spaces and norms of use above.

Lemma A.1. *Fix $T > 0$ and a compact $\mathcal{U} \subset \mathbb{R}^M$. Let $f\colon [0,T] \times \mathbb{R}^N \times \mathcal{U} \to \mathbb{R}^N$ be such that*

($\mathbf{f_1}$) *$t \to f(t,x;u)$ is measurable for all $x \in \mathbb{R}^N$ and $u \in \mathcal{U}$;*

($\mathbf{f_2}$) *$(x,u) \to f(t,x;u)$ is in $\mathbf{C^1}$ for a.e. $t \in [0,T]$ and for every compact set $\mathcal{K} \subset \mathbb{R}^N$ there exists a constant $L_{\mathcal{K}} > 0$ such that for a.e. $t \in [0,T]$, for all $x_1, x_2 \in \mathcal{K}$ and for all $u_1, u_2 \in \mathcal{U}$,*

$$
\left\|f(t,x_1;u_1) - f(t,x_2;u_2)\right\| \le L_{\mathcal{K}}\left(\|x_1 - x_2\| + \|u_1 - u_2\|\right).
$$

($\mathbf{f_3}$) *$(t,x) \to f(t,x;u)$ is sublinear in $x \in \mathbb{R}^N$, uniformly in $t \in [0,T]$ and for all $u \in \mathcal{U}$.*

Then, for all $x_o \in \mathbb{R}^N$ and all $u \in \mathbf{L}^\infty([0,T];\mathcal{U})$, the problem

$$
\begin{cases} \dot{x} = f(t,x;u) \\ x(0) = x_o \end{cases} \tag{A.1}
$$

admits a unique solution $X(u)\colon [0,T] \to \mathbb{R}^N$. The map $X\colon \mathbf{L}^\infty([0,T];\mathcal{U}) \to \mathbf{C^1}([0,T];\mathbb{R}^N)$ is Gateaux differentiable in any direction $v \in \mathbf{L}^\infty([0,T];\mathcal{U})$ and the directional derivative $D_v X(u)$ solves the Cauchy problem

$$
\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} D_v X(u) = \partial_x f\left(t, X(u); u\right) D_v X(u) + \partial_v f\left(t, X(u); u\right) v \\ \left(D_v X(u)\right)(0) = 0. \end{cases} \tag{A.2}
$$

Proof. The map X is well defined by the standard theory of Caratheodory ODEs, see for instance [15]. Moreover, there exists a compact $\Omega \subset \mathbb{R}^N$ such that for all $u \in \mathbf{L}^\infty([0,T];\mathcal{U})$, $\big(X(u)\big)([0,T]) \subset \Omega$.

To prove the directional differentiability, call g the solution to the linear problem (A.2) and use the integral form of both Cauchy problems (A.1) and (A.2) to obtain

$$
\begin{aligned}
&\frac{1}{h}\big(X(u+hv)-X(u)\big)(t)-g(t)\\
&=\int_0^t \frac{f\Big(\tau,\big(X(u+hv)\big)(\tau),(u+hv)(\tau)\Big)-f\Big(\tau,\big(X(u)\big)(\tau),u(\tau)\Big)}{h}\,\mathrm{d}\tau-g(t)\\
&=\int_0^t\int_0^1 \partial_x f\Big(\tau,\big(X(u+\vartheta hv)\big)(\tau),(u+\vartheta hv)(\tau)\Big)\,\mathrm{d}\vartheta\,\frac{\big(X(u+hv)-X(u)\big)(\tau)}{h}\,\mathrm{d}\tau\\
&\quad+\int_0^t\int_0^1 \partial_v f\Big(\tau,\big(X(u+\vartheta hv)\big)(\tau),(u+\vartheta hv)(\tau)\Big)\,\mathrm{d}\vartheta\, v(\tau)\,\mathrm{d}\tau\\
&\quad-\int_0^t\left(\partial_x f\Big(\tau,\big(X(u)\big)(\tau);u\Big)\,g(\tau)+\partial_v f\Big(\tau,\big(X(u)\big)(\tau);u\Big)\;v\right)\mathrm{d}\tau\\
&=\int_0^t\int_0^1\left[\partial_x f\Big(\tau,\big(X(u+\vartheta hv)\big)(\tau),(u+\vartheta hv)(\tau)\Big)-\partial_x f\left(\tau,X(u);u\right)\right]\mathrm{d}\vartheta\,\frac{\big(X(u+hv)-X(u)\big)(\tau)}{h}\,\mathrm{d}\tau\\
&\quad+\int_0^t \partial_x f\left(\tau,X(u);u\right)\left(\frac{\big(X(u+hv)-X(u)\big)(\tau)}{h}-g(\tau)\right)\mathrm{d}\tau\\
&\quad+\int_0^t\int_0^1\left[\partial_v f\Big(\tau,\big(X(u+\vartheta hv)\big)(\tau),(u+\vartheta hv)(\tau)\Big)-\partial_v f\left(\tau,X(u);u\right)\right]\mathrm{d}\vartheta\, v(\tau)\,\mathrm{d}\tau
\end{aligned}
$$

By the sublinearity of f, for every $u, v \in \mathcal{U}$, we can find a compact set $\mathcal{K}$ such that $\big(X(u+hv)\big)(t)$, $\big(X(u)\big)(t)$ and $\big(X(u+\vartheta hv)\big)(t) \in \mathcal{K}$.

Lusin's Theorem ([25], Thm. 1), applied to $(\partial_x f, \partial_v f) \in \mathbf{L}^\infty([0,T];\mathbf{C^{0,1}}(\mathcal{K}\times\mathcal{U};\mathbb{R}^{2N^2}))$, ensures that for any $\varepsilon > 0$, there exists a compact set $K \subset [0,T]$ such that the Lebesgue measure of $[0,T]\setminus K$ is smaller than ε and both $\partial_x f$ and $\partial_v f$ are continuous on $K \times \Omega \times \mathcal{U}$, hence also uniformly continuous. Therefore, if h is sufficiently small,

$$
\begin{aligned}
\sup_{K\times\Omega\times\mathcal{U}}\ \sup_{\vartheta\in[0,1]}\left\|\partial_x f\Big(\tau,\big(X(u+\vartheta hv)\big)(\tau),(u+\vartheta hv)(\tau)\Big)-\partial_x f\left(\tau,X(u);u\right)\right\| &\le \varepsilon,\\
\sup_{K\times\Omega\times\mathcal{U}}\ \sup_{\vartheta\in[0,1]}\left\|\partial_v f\Big(\tau,\big(X(u+\vartheta hv)\big)(\tau),(u+\vartheta hv)(\tau)\Big)-\partial_v f\left(\tau,X(u);u\right)\right\| &\le \varepsilon.
\end{aligned}
$$

Introduce now the quantity

$$
\delta_h(t)=\left\|\frac{1}{h}\big(X(u+hv)-X(u)\big)(t)-g(t)\right\|.
$$

Since $\|\partial_x f\| \le L$, $\|\partial_v f\| \le L$, $\|f\|_{\mathbf{L}^\infty([0,T]\times\Omega\times\mathcal{U};\mathbb{R}^N)} < +\infty$, the above estimates lead to

$$
\begin{aligned}
\delta_h(t) &\le (2\,L\,\varepsilon+\varepsilon\,t)\|f\|_{\mathbf{L}^\infty([0,T]\times\Omega\times\mathcal{U};\mathbb{R}^N)}+\int_0^t L\,\delta_h(\tau)\,\mathrm{d}\tau+(2\,L\,\varepsilon+\varepsilon\,t)\|v\|_{\mathbf{L}^\infty([0,T];\mathbb{R}^M)}\\
&=(2\,L+t)\left(\|f\|_{\mathbf{L}^\infty([0,T]\times\Omega\times\mathcal{U};\mathbb{R}^N)}+\|v\|_{\mathbf{L}^\infty([0,T];\mathbb{R}^M)}\right)\varepsilon+\int_0^t L\,\delta_h(\tau)\,\mathrm{d}\tau\,.
\end{aligned}
$$

An application of Gronwall Lemma yields that for all $\varepsilon > 0$, if h is sufficiently small

$$\delta_h(t) \leq (2\,L + t)\left(\|f\|_{\mathbf{L}^\infty([0,T]\times\Omega\times\mathcal{U};\mathbb{R}^N)} + \|v\|_{\mathbf{L}^\infty([0,T];\mathbb{R}^M)}\right)\varepsilon\, \mathrm{e}^{L\,t}$$

completing the proof. □

References

[1] H. Amann and J. Escher, Analysis. II. Translated from the 1999 German original by Silvio Levy and Matthew Cargo. Birkhäuser Verlag, Basel (2008).

[2] O. Blanchard and S. Fischer, Lectures on Macroeconomics. MIT Press (1989).

[3] Å. Brännström, L. Carlsson and D. Simpson, On the convergence of the escalator boxcar train. *SIAM J. Numer. Anal.* **51** (2013) 3213–3231.

[4] A. Bressan, Hyperbolic systems of conservation laws the one-dimensional Cauchy problem. Vol. 20 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford (2000).

[5] J.A. Cañizo, J.A. Carrillo and J. Rosado, A well-posedness theory in measures for some kinetic models of collective motion. *Math. Models Methods Appl. Sci.* **21** (2011) 515–539.

[6] J.A. Carrillo, R.M. Colombo, P. Gwiazda and A. Ulikowska, Structured populations, cell growth and measure valued balance laws. *J. Differ. Equ.* **252** (2012) 3245–3277.

[7] J.A. Carrillo, P. Gwiazda and A. Ulikowska, Splitting-particle methods for structured population models: convergence and applications. *Math. Models Methods Appl. Sci.* **24** (2014) 2171–2197.

[8] R.M. Colombo and G. Guerra, Hyperbolic balance laws with a non local source. *Comm. Partial Differ. Equ.* **32** (2007) 1917–1939.

[9] A.M. de Roos, Numerical methods for structured population models: the escalator boxcar train. *Numer. Methods Partial Differ. Equ.* **4** (1988) 173–195.

[10] A.M. de Roos and L. Persson, Population and community ecology of ontogenetic development. *Monographs in population biology*. Princeton University Press, Princeton (2013).

[11] O. Diekmann and P. Getto, Boundedness, global existence and continuous dependence for nonlinear dynamical systems describing physiologically structured populations. *J. Differ. Equ.* **215** (2005) 268–319.

[12] L.C. Evans and R.F. Gariepy, Measure theory and fine properties of functions. *Studies in Advanced Mathematics*. CRC Press, Boca Raton, FL (1992).

[13] J.H.M. Evers, S.C. Hille and A. Muntean, Well-posedness and approximation of a measure-valued mass evolution problem with flux boundary conditions. *C.R. Math. Acad. Sci. Paris* **352** (2014) 51–54.

[14] J.H.M. Evers, S.C. Hille and A. Muntean. Mild solutions to a measure-valued mass evolution problem with flux boundary conditions. *J. Differ. Equ.* **259** (2015) 1068–1097.

[15] A.F. Filippov, *Differential equations with discontinuous righthand sides*. Kluwer Academic Publishers Group, Dordrecht (1988). Translated from the Russian.

[16] W. Gangbo and M. Westdickenberg, Optimal transport for the system of isentropic Euler equations. *Comm. Partial Differ. Equ.* **34** (2009) 1041–1073.

[17] P. Gwiazda, J. Jabłoński, A. Marciniak−Czochra and A. Ulikowska, Analysis of particle methods for structured population models with nonlocal boundary term in the framework of bounded Lipschitz distance. *Numer. Methods Partial Differ. Equ.* **30** (2014) 1797–1820.

[18] P. Gwiazda, T. Lorenz and A. Marciniak−Czochra, A nonlinear structured population model: Lipschitz continuity of measure-valued solutions with respect to model ingredients. *J. Differ. Equ.* **248** (2010) 2703–2735.

[19] P. Gwiazda, T. Lorenz and A. Marciniak−Czochra, A nonlinear structured population model: Lipschitz continuity of measure-valued solutions with respect to model ingredients. *J. Differ. Equ.* **248** (2010) 2703–2735.

[20] P. Gwiazda and A. Marciniak−Czochra, Structured population equations in metric spaces. *J. Hyperbolic Differ. Equ.* **7** (2010) 733–773.

[21] S.C. Hille and D.T.H. Worm, Embedding of semigroups of Lipschitz maps into positive linear semigroups on ordered Banach spaces generated by measures. *Integral Equ. Oper. Theory* **63** (2009) 351–371.

[22] M. Iannelli, Mathematical theory of age-structured population dynamics, Vol. 7 of *Applied Mathematics Monographs*. Giardini editori e stampatori in Pisa (1995).

[23] W. Kermack and A. McKendrick, Contributions to the mathematical theory of epidemics–I. *Bull. Math. Biol.* **53** (1991) 33–55.

[24] W. Kermack and A. McKendrick, Contributions to the mathematical theory of epidemics–III. *Bull. Math. Biol.* **53** (1991) 57–87.

[25] P.A. Loeb and E. Talvila, Lusin's theorem and Bochner integration. *Sci. Math. Jpn* **60** (2004) 113–120.

[26] S. Maniglia, Probabilistic representation and uniqueness results for measure-valued solutions of transport equations. *J. Math. Pures Appl.* **87** (2007) 601–626.

[27] J.A.J. Metz and O. Diekmann, The dynamics of physiologically structured populations, Papers from the colloquium held in Amsterdam (1983). Vol. 68 of *Lect. Notes in Biomathematics*. Springer-Verlag, Berlin (1986).

[28] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Differ. Equ.* **26** (2001) 101–174.

[29] B. Piccoli and F. Rossi, Generalized Wasserstein distance and its application to transport equations with source. *Arch. Ration. Mech. Anal.* **211** (2014) 335–358.

[30] P.-A. Raviart, An analysis of particle methods, vol. 1127 of *Lect. Notes Math.* Springer, Berlin (1985).

[31] H.R. Thieme, Mathematics in population biology. *Princeton Series in Theoretical and Computational Biology*. Princeton University Press, Princeton, NJ (2003).

[32] A. Ulikowska, An age-structured two-sex model in the space of Radon measures: well posedness. *Kinet. Relat. Models* **5** (2012) 873–900.

[33] H. Von Förster, Some remarks on changing populations. The Kinetics of Cellular Proliferation. Shalton Press, New York (1959).

[34] G.F. Webb, Theory of nonlinear age-dependent population dynamics, vol. 89 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York (1985).

[35] M. Westdickenberg and J. Wilkening, Variational particle schemes for the porous medium equation and for the system of isentropic Euler equations. *ESAIM: M2AN* **44** (2010) 133–166.