SIAM J. NUMER. ANAL.
Vol. 57, No. 4, pp. 1842–1874

THE ESCALATOR BOXCAR TRAIN METHOD FOR A SYSTEM OF AGE-STRUCTURED EQUATIONS IN THE SPACE OF MEASURES*

JOSÉ ANTONIO CARRILLO†, PIOTR GWIAZDA‡, KAROLINA KROPIELNICKA§, AND ANNA K. MARCINIAK-CZOCHRA¶

Abstract. The Escalator Boxcar Train (EBT) numerical method has been designed and widely used by theoretical biologists to compute solutions of one-dimensional structured population models of McKendrick–von Foerster type. Recently the method has been derived for an age-structured two-sex population model (Fredrickson–Hoppenstaedt model), which consists of three coupled hyperbolic partial differential equations with nonlocal boundary conditions. The convergence of the EBT method for the Fredrickson–Hoppenstaedt model has not been analyzed, and relevant numerical examples are still missing. In this paper, we derive a simplified EBT method for the "two-sex model" and prove its convergence. However, due to the interest in tracking specified cohorts of individuals, the analytical results cannot be analyzed in the $L1$ norm. Instead, we embedded the problem in the space of nonnegative Radon measures equipped with the bounded Lipschitz distance (the flat metric). We also present numerical examples to illustrate the results, compute the error in bounded Lipschitz distance, and compare it against the total variation (TV) distance.

Key words. particle method, Escalator Boxcar Train, structured population models, two-sex population models, system of equations, measure-valued solutions

AMS subject classifications. 65M75, 45K05, 92D25

DOI. 10.1137/18M1189427

1. Introduction.

The Escalator Boxcar Train (EBT) method is a numerical integrator that was introduced in [18] for structured population models of McKendrick–von Foerster type [38] and described by a scalar hyperbolic partial differential equation with nonlocal boundary conditions. It has been widely used by theoretical biologists, because it approximates the density function of the distribution of individuals in a way that has a clear biological interpretation. The method is based on partitioning the population into a number of cohorts, i.e., groups of individuals endowed with a similar property described by the state variable. Cohorts are described by masses localized in discrete points, and their dynamics is traced along growth trajectories given by characteristic lines of the model. A similar concept lies behind the particle

*Received by the editors May 23, 2018; accepted for publication (in revised form) April 9, 2019; published electronically July 30, 2019.
https://doi.org/10.1137/18M1189427
Funding: This work was partially supported by grant 346300 for IMPAN from the Simons Foundation, and by the matching 2015-2019 Polish MNiSW fund. The first author was supported by EPSRC grant EP/P031587/1. The second author was supported by National Center for Science grant UMO-2015/18/M/ST1/00075. The third author was supported by National Center for Science grant DEC-2012/05/E/ST1/02218. The fourth author was supported by the Emmy Noether Programme of the German Research Council (DFG).
†Department of Mathematics, Imperial College London, London, SW7 2AZ, UK (carrillo@imperial.ac.uk).
‡Institute of Mathematics, Polish Academy of Sciences, Warsaw, 00-656, Poland (pgwiazda@mimuw.edu.pl).
§Institute of Mathematics, University of Gdańsk, Gdańsk, 80-308, Poland, and Institute of Mathematics, Polish Academy of Sciences, Warsaw, 00-656, Poland (karolina.kropielnicka@mat.ug.edu.pl).
¶Institute of Applied Mathematics, Interdisciplinary Center of Scientific Computing and BioQuant, University of Heidelberg, Heidelberg, 69120, Germany (anna.marciniak@iwr.uni-heidelberg.de).

methods that have been widely applied in computational physics [5, 6, 16, 21, 40] and in fluid mechanics [17, 22].

The EBT method has been a common tool for many years, but its convergence was proved only recently for a scalar equation in [8] and in [23], where the rate of convergence was also shown. The latter results were obtained using an approach to stability, in which the underlying model was embedded in a space of nonnegative Radon measures ($\mathcal{M}^+(\mathbb{R}_+)$) with a bounded Lipschitz distance (the flat metric). This approach was proposed in [25, 26]. Such an external approximation, i.e., approximation of functions from $\mathbf{L^1}(\mathbb{R}_+)$ by objects from a space of nonnegative Radon measures, is a natural extension of the way in which the initial condition of the problem is approximated. The choice of the suitable metric is a delicate issue, since the Cauchy problem for the transport equation in the space of nonnegative Radon measures with the total variation (TV) norm is not well posed; i.e., its solutions are not strongly continuous in the TV norm with respect to initial data, time, or model parameters; see e.g., [26, Examples 1.1 and 1.2]. Following previous works, we apply the flat metric, which has proven to be useful in the analysis of a variety of transport equations models. The flat metric has been equivalently defined in the literature as a (dual) bounded Lipschitz distance [39], Kantorovich–Rubinstein distance [7], (dual) Fortet–Mourier distance [35], and Dudley distance [19]; see [27] for more details.

Other numerical methods used for the structured population models differ in discretization of the population distribution and in incorporation of nonlocal feedback. The classical upwind finite-difference method is based on mesh-dependent discretization (with a fixed or moving mesh) of the individuals distribution in the state variable [1, 3]. In contrast, the EBT method is a mesh-free method, similar to the method of characteristics that works by tracking population density along characteristic curves [11, 12, 15]. Numerical accuracy and computational efficiency of the different numerical schemes have been recently compared for a choice of structured population models [46]. It has been found that the computational performance of the method depends on the character of nonlocal and nonlinear couplings, and the EBT method performs best. An exception is a reference problem with asymmetric competition affecting individual growth rates that leads to large distances between two successive cohorts. Moreover, the EBT method allows us to avoid the problem of numerical diffusion that would lead to loss of the cohort structure.

Taking into account these numerical advantages along with the fact that the EBT method allows a clear biological interpretation in terms of the number of individuals in cohorts, we investigate the method's application to a system of structured population equations using the example of the Fredrickson–Hoppenstaedt model, a two-sex population model describing evolution of males and females and the process of heterogeneous couple formation.

The Fredrickson–Hoppenstaedt model was originally formulated in [20] and later developed in [?]. It consists of three population equations with structure, which are coupled via nonlocal boundary terms and a nonlocal and nonlinear source. Dynamics of males and females is given by McKendrick-type equations; that is, they consist of a transport equation with a growth term and boundary terms determining the influx of newborn individuals. Evolution of couples is modeled by a similar equation; however, it is equipped with the so-called marriage function. The marriage function describes the influx of new couples between males and females of particular ages at a certain time. This is a simplification, since, in reality, formation of new couples depends on many social and economic factors, such as religion, culture, education, and health, and therefore it is a much more complicated process. In the references mentioned above, the authors assume that the distribution of population is given by an unnormed density. Therefore, in the system (1.1) below, functions $u^m(t,x)$ and $u^f(t,y)$ describe the distribution density of males and females at time t and ages x and y, respectively, while $u^c(t,x,y)$ is the number of couples at time t between males at age x and females at age y. The following system of nonlinear equations describes dynamics of the population of males, females, and couples:

$$
\begin{aligned}
&\partial_t u^m(t,x) + \partial_x u^m(t,x) + c^m(t,u^m(t,\cdot),u^f(t,\cdot),x)u^m(t,x) = 0,\\
&u^m(t,0) = \int_{\mathbb{R}_+^2} b^m(t,u^m(t,x),u^f(t,y),x,y)u^c(t,x,y)dxdy,\\
&u^m(0,x) = u_0^m(x),\\
&\partial_t u^f(t,y) + \partial_y u^f(t,y) + c^f(t,u^m(t,\cdot),u^f(t,\cdot),y)u^f(t,y) = 0,\\
&u^f(t,0) = \int_{\mathbb{R}_+^2} b^f(t,u^m(t,x),u^f(t,y),x,y)u^c(t,x,y)dxdy,\\
&u^f(0,x) = u_0^f(x),\\
&\partial_t u^c(t,x,y) + \partial_x u^c(t,x,y) + \partial_y u^c(t,x,y)\\
&\qquad + c^c(t,u^m(t,\cdot),u^f(t,\cdot),u^c(t,\cdot),x,y)u^c(t,x,y) = T(t,x,y),\\
&u^c(t,x,0) = u^c(t,0,y) = 0,\\
&u^c(0,x,y) = u_0^c(x,y).
\end{aligned} \tag{1.1}
$$

Functions c^m, c^f, and c^c describe the rates of disappearance of individuals, where disappearance of males or females is caused by death, while couples disappearance is caused by divorce or death of one of the spouses. Functions b^m and b^f are birth rates of males and females. Observe that these coefficients depend on ecological pressure in a nonlinear manner; that is, they are nonlocal operators depending on the distribution of males, females, and couples.

The marriage function T models the number of new marriages of males and females of ages x and y, respectively, at time t. It also depends nonlinearly on the distribution of individuals. The choice of this function is a subject of ongoing discussion [29, 30, 37, 41], due to properties such as heterosexuality, homogeneity, consistency, and competition. In this paper we follow the formulation proposed in [33], namely

$$
\begin{aligned}
T(t,x,y) &= F(t,u^m(t,x),u^f(t,y),u^c(t,x,y),x,y)\\
&= \frac{\Theta(x,y)h(x)g(y)\left[u^m(t,x)-\int_0^\infty u^c(t,x,y)dy\right]\left[u^f(t,y)-\int_0^\infty u^c(t,x,y)dx\right]}{\gamma+\int_0^\infty h(x)\left[u^m(t,x)-\int_0^\infty u^c(t,x,y)dy\right]dx+\int_0^\infty g(y)\left[u^f(t,y)-\int_0^\infty u^c(t,x,y)dx\right]dy}.
\end{aligned} \tag{1.2}
$$

The function $\Theta(x,y) \in \mathbf{L^1}(\mathbb{R}_+^2) \cap \mathbf{L}^\infty(\mathbb{R}_+^2)$ describes the marriage rate of males of age x and females of age y. Notice that $\left[u^m(t,x)-\int_0^\infty u^c(t,x,y)dy\right]$ is the number of unmarried males, and $\left[u^f(t,y)-\int_0^\infty u^c(t,x,y)dx\right]$ is the number of unmarried females. The functions $h,g \in \mathbf{L^1}(\mathbb{R}_+) \cap \mathbf{L}^\infty(\mathbb{R}_+)$ describe the distribution of eligible males/females on the marriage market. We further assume that youngsters do not marry below a certain age a, i.e.,

$$
h(x) = g(y) = 0 \text{ for } x,y \in [0,a_0)\,. \tag{1.3}
$$

The regularity of the remaining coefficients and their nonlinear dependencies are presented in detail in section 3.2.

While [44] has shown the well-posedness of the embedding of the underlying problem (1.1) in the space of nonnegative Radon measures with a flat metric, the EBT algorithm has been so far proposed only for a partially linearized system [24], where the only nonlinearity appeared in nonlocal and highly nonlinear function T defined by (1.2), that is for a system of the following equations:

$$
\begin{aligned}
&\partial_t u^m(t,x) + \partial_x u^m(t,x) + c^m(t,x)u^m(t,x) = 0,\\
&\qquad\qquad u^m(t,0) = \int_{\mathbb{R}^2_+} b^m(t,x,y)u^c(t,x,y)dxdy,\\
&\qquad\qquad u^m(0,x) = u^m_0(x),\\[1em]
&\partial_t u^f(t,y) + \partial_y u^f(t,y) + c^f(t,y)u^f(t,y) = 0,\\
&\qquad\qquad u^f(t,0) = \int_{\mathbb{R}^2_+} b^f(t,x,y)u^c(t,x,y)dxdy,\\
&\qquad\qquad u^f(0,x) = u^f_0(x),\\[1em]
&\partial_t u^c(t,x,y) + \partial_x u^c(t,x,y) + \partial_y u^c(t,x,y) + c^c(t,x,y)u^c(t,x,y) = T(t,x,y),\\
&\qquad\qquad u^c(t,x,0) = u^c(t,0,y) = 0,\\
&\qquad\qquad u^c(0,x,y) = u^c_0(x,y).
\end{aligned}
\tag{1.4}
$$

Moreover, the method was presented in [24] only heuristically, but convergence results were still missing. In the present paper we prove the convergence of the EBT algorithm for a system of coupled equations and provide error estimates, closing the gap between practical use of the method and its mathematical foundations. We focus on a simplified EBT scheme which is a modification of the original EBT method in [24]. This simplification uses a numerical approximation of the boundary cohort that yields easier and cheaper numerics and preserves convergence. Nonlinear coupling in boundary terms requires a different and more involved treatment than in the previously considered scalar model [8, 23]. We present numerical examples and express the error as a nontrivial, two-dimensional Lipschitz-bounded distance, and compare it to the TV distance in which the method does not converge. Contrary to the one-dimensional problems, in two dimensions even computing the Wasserstein distance between sums of Dirac deltas is an *np*-complex problem. We approach the problem by approximating the Wasserstein distance using the Kullback–Leibler divergence.

The remainder of this paper is organized as follows. Section 2 is devoted to the EBT schemes. In subsection 2.1 we present a simplified EBT method for the Fredrickson–Hoppenstaedt model. In subsection 2.2 we summarize the original EBT scheme for (1.4) recently derived in [24], while in subsection 2.3 we show how to obtain a simplified EBT method from the original EBT approach [24]. Simplification of the method consists of unifying the rules (ODEs) in all cohorts, without distinction between internal and boundary rules. In section 3, the underlying problem (1.1) is reformulated as an evolution in a space of nonnegative Radon measures. We approximate solutions of this problem with a linear combination of Dirac Deltas, where the masses and corresponding localizations are obtained from the simplified EBT scheme embedded in the space of nonnegative Radon measures. Subsection 3.2 is devoted to the analytical framework in the space of nonnegative Radon measures, where we introduce the necessary notation, definitions, lemmas, and assumptions. The choice of state space allows us to prove the rate of convergence of the simplified EBT scheme in section 4. Section 5 is devoted to numerical illustrations. As the computational error measurement is not trivial in the flat metric (especially in $\mathbb{R}^2$), some necessary details are provided in subsection 5.1, while the rate of convergence obtained is illustrated in subsections 5.2 and 5.3.

2. Numerical methods based on the EBT approach. In this paper we introduce and analyze a simplified EBT method for (1.4). The current section starts with the presentation of the simplified method. In further subsections we present the original method derived in [24] and explain derivation of the simplified method from the original.

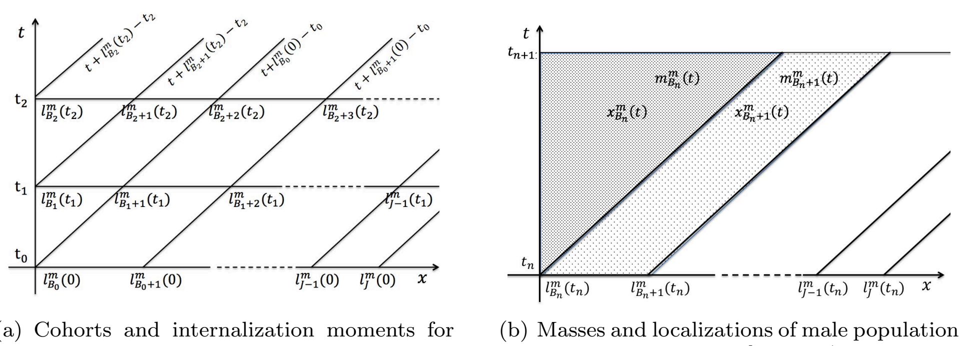
2.1. The simplified EBT method. The concept of particle methods describes grouping individuals into so-called cohorts and tracing their dynamics in time. Accordingly, the first step of the EBT algorithm consists of imposing, at time $t = 0$, initial $J - B_0$ cohorts for males and females, and $(J - B_0)^2$ for couples, as follows (see Figure 1(a)):

$$[l_i^m(0), l_{i+1}^m(0)),\ [l_j^f(0), l_{j+1}^f(0)),\ \text{and}\ [l_i^m(0), l_{i+1}^m(0)) \times [l_j^f(0), l_{j+1}^f(0)),\ i,j = B_0, \ldots, J-1, \tag{2.1}$$

respectively, in such a way that $\mathrm{supp}(u_0^m) \subset [l_{B_0}^m(0), l_J^m(0))$ and $\mathrm{supp}(u_0^f) \subset [l_{B_0}^f(0), l_J^f(0))$. Cohorts evolve in time along the characteristic lines of relevant transport operators in (1.4). As is always the case in age-structured problems, the characteristics are straight lines; see Figure 1(a).

In the next step, we impose a mesh on the time variable $t \in [0, T)$ in such a way that $t_0 = 0$ and $\bigcup_{n=0}^{N_T} [t_n, t_{n+1}) = [0, T)$; see Figure 1(a).

Remark 2.1. In each time step t_n, new boundary cohorts are created. Boundary cohorts account for the influx of new males and females. In the case of couples, new boundary cohorts also appear, but they are empty, as we do not expect newborns to form couples due to (1.3). In Figure 1(b) we illustrate cohorts and internalization moments for the male population during one time step. Notice that neither time points nor cohort boundaries are equidistant.



(a) Cohorts and internalization moments for male population.

(b) Masses and localizations of male population evolving in time interval $[t_n, t_{n+1})$.

FIG. 1.

Remark 2.2. Depending on time t, the number of cohorts changes. Let us note that in the internalization moment t_n, male and female individuals are grouped into

$J - B_n$ cohorts, while couples are grouped into $(J - B_n)^2$ cohorts, that is,

$$[l_i^m(t_n), l_{i+1}^m(t_n)),\ i = B_n, \dots, J, \qquad [l_j^f(t_n), l_{j+1}^f(t_n)),\ j = B_n, \dots, J,$$

$$[l_i^m(t_n), l_{i+1}^m(t_n)) \times [l_j^f(t_n), l_{j+1}^f(t_n)),\ i, j = B_n, \dots, J.$$

If $t \in (t_n, t_{n+1})$, then we deal with $(J - B_n + 1)$ cohorts for males and females and with $(J - B_n + 1)^2$ cohorts for couples. Obviously individuals from time t_n stay in their corresponding cohorts for time $t \in (t_n, t_{n+1})$:

$$[t + l_i^m(t_n) - t_n, t + l_{i+1}^m(t_n) - t_n),\ i = B_n, \dots, J,$$

$$[t + l_j^f(t_n) - t_n, t + l_{j+1}^f(t_n) - t_n),\ j = B_n, \dots, J,$$

$$[t + l_i^m(t_n) - t_n, t + l_{i+1}^m(t_n) - t_n) \times [t + l_j^f(t_n) - t_n, t + l_{j+1}^f(t_n) - t_n),\ i, j = B_n, \dots, J,$$

but also new boundary cohorts appear (see the triangles in Figure 1(b)), that is,

$$[0, t + l_{B_n}^m(t_n) - t_n), \qquad [0, t + l_{B_n}^f(t_n) - t_n), \qquad [0, t + l_{B_n}^m(t_n) - t_n) \times [0, t + l_{B_n}^f(t_n) - t_n).$$

Observe that the indexes of the cohorts at each internalization point move backward, i.e., $B_{n+1} = B_n - 1$, with $B_0 = 0$ typically.

The idea of the EBT algorithm is to trace the numbers of individuals in cohorts as the time changes from moment 0 to moment T, $t \in [0, T)$. Thus each cohort is characterized by its mass and location (see Figure 1(b)) as follows:

$$(m_i^m(t), x_i^m(t)),\ (m_j^f(t), x_j^f(t)),\ \text{and}\ \left(m_{ij}^c(t), (x_{ij}^c(t), y_{ij}^c(t))\right).$$

These quantities evolve in time intervals $[t_n, t_{n+1})$, $n = 0, \dots, N_T - 1$, and are governed by the following system of differential equations, which constitutes the simplified EBT scheme:

$$\tag{2.2}
\begin{aligned}
&\left\{
\begin{array}{rcl}
\frac{d}{dt} m_i^m(t) &=& -c^m(t, x_i^m(t)) m_i^m(t),\ i = B_n + 1, \dots, J,\\
\frac{d}{dt} x_i^m(t) &=& 1,\ i = B_n, \dots, J,\\
\frac{d}{dt} m_{B_n}^m(t) &=& -c^m(t, x_{B_n}^m(t)) m_{B_n}^m(t) + \sum_{i,j=B_n}^J b^m(t, x_{ij}^c(t), y_{ij}^c(t)) m_{ij}^c(t),
\end{array}
\right.\\
&\left\{
\begin{array}{rcl}
\frac{d}{dt} m_j^f(t) &=& -c^f(t, y_j^f(t)) m_j^f(t),\ j = B_n + 1, \dots, J,\\
\frac{d}{dt} y_j^f(t) &=& 1,\ j = B_n, \dots, J,\\
\frac{d}{dt} m_{B_n}^f(t) &=& -c^f(t, x_{B_n}^f(t)) m_{B_n}^f(t) + \sum_{i,j=B_n}^J b^f(t, x_{ij}^c(t), y_{ij}^c(t)) m_{ij}^c(t),
\end{array}
\right.\\
&\left\{
\begin{array}{rcl}
\frac{d}{dt} m_{ij}^c(t) &=& -c^c(t, x_{ij}^c(t), y_{ij}^c(t)) m_{ij}^c(t) + \frac{N_{ij}(t)}{D_{ij}(t)},\ i, j = B_n, \dots, J,\\
\frac{d}{dt}(x_{ij}^c(t), y_{ij}^c(t)) &=& (1, 1),\ i, j = B_n, \dots, J,
\end{array}
\right.
\end{aligned}$$

with

$$\tag{2.3}
\frac{N_{ij}(t)}{D_{ij}(t)} = \frac{\Theta(x_{ij}^c(t), y_{ij}^c(t)) h(x_{ij}^c(t)) g(y_{ij}^c(t)) \left(m_i^m(t) - \sum_{w=B_n}^J m_{iw}^c(t)\right)\left(m_j^f(t) - \sum_{v=B_n}^J m_{vj}^c(t)\right)}{\gamma + \sum_{v=B_n}^J h(x_{vj}^c(t)) \left(m_v^m(t) - \sum_{w=B_n}^J m_{vw}^c(t)\right) + \sum_{w=B_n}^J g(y_{iw}^c(t)) \left(m_w^f(t) - \sum_{v=B_n}^J m_{vw}^c(t)\right)}.$$

The definition of initial conditions for these equations depends on n as follows:

$(n=0)$ First, let us observe that $t_0 = 0$, and the initial conditions should be consistent with the initial conditions of (1.4). The boundary cohorts are defined with zero masses as follows:

$$(m^m_{B_0}(0), x^m_{B_0}(0)) = (0,0), \ (m^f_{B_0}(0), y^f_{B_0}(0)) = (0,0), \tag{2.4}$$

$$\begin{gathered}\left(m^c_{ij}(0), (x^c_{ij}(0), y^c_{ij}(0))\right) = (0, (x^m_i(0), y^f_j(0))),\\ \text{where } (i = B_0 \vee j = B_0) \wedge i, j \in \{B_0, \dots, J\}.\end{gathered} \tag{2.5}$$

Initial conditions for the internal cohorts are derived from the biological model, with mass defined as the number of individuals within a cohort as follows:

$$\begin{aligned} m^m_i(0) &= \int_{[l^m_{i-1}(0), l^m_i(0))} u^m_0(x)dx,\\ m^f_j(0) &= \int_{[l^f_{j-1}(0), l^f_j(0))} u^f_0(y)dy,\\ m^c_{ij}(0) &= \int_{[l^m_{i-1}(0), l^m_i(0)) \times [l^f_{j-1}(0), l^f_j(0))} u^c_0(x,y)dxdy, \quad i, j = B_0+1, \dots, J, \end{aligned} \tag{2.6}$$

and the location is the average value of the structural variable of the underlying cohort as follows:

$$\begin{aligned} x^m_i(0) &= \begin{cases} 0 \text{ if } m^m_i(0) = 0,\\ \frac{1}{m^m_i(0)} \int_{[l^m_{i-1}(0), l^m_i(0))} x u^m_0(x)dx \quad \text{otherwise}, \end{cases}\\ y^f_j(0) &= \begin{cases} 0 \text{ if } m^m_i(0) = 0,\\ \frac{1}{m^f_j(0)} \int_{[l^f_{j-1}(0), l^f_j(0))} x u^f_0(y)dy \quad \text{otherwise}, \end{cases}\\ (x^c_{ij}(0), y^c_{ij}(0)) &= \quad (x^m_i(0), y^f_j(0)), \qquad\qquad i, j = B_0+1, \dots, J. \end{aligned} \tag{2.7}$$

$(n>0)$ Similarly as in (2.4), (2.5), the boundary cohorts are defined with zero masses as follows:

$$(m^m_{B_n}(t_n), x^m_{B_n}(t_n)) = (0,0), \ (m^f_{B_n}(t_n), y^f_{B_n}(t_n)) = (0,0), \tag{2.8}$$

$$\begin{gathered}\left(m^c_{ij}(t_n), (x^c_{ij}(t_n), y^c_{ij}(t_n))\right) = (0, (x^m_i(t_n), y^f_j(t_n))),\\ \text{where } (i = B_n \vee j = B_n) \wedge i, j \in \{B_n, \dots, J\},\end{gathered} \tag{2.9}$$

while initial conditions for internal cohorts are obtained as an output of the $(n-1)$th step of the algorithm (in the sense of limit $t \to t_n^-$) as follows:

$$\begin{gathered} (x^m_i(t_n), m^m_i(t_n)) = \lim_{t \to t_n^-} (x^m_i(t), m^m_i(t)),\\ (y^f_j(t_n), m^f_j(t_n)) = \lim_{t \to t_n^-} (y^f_j(t), m^f_j(t)),\\ \left(m^c_{ij}(t_n), (x^c_{ij}(t_n), y^c_{ij}(t_n))\right) = \lim_{t \to t_n^-} \left(m^c_{ij}(t), (x^c_{ij}(t), y^c_{ij}(t))\right). \end{gathered} \tag{2.10}$$

2.2. The original EBT method. The original EBT method for (1.4) was derived in [24] and has the following significantly more complicated form:

$$\begin{cases} \frac{d}{dt}m_i^m(t) &= -c^m(t,x_i^m(t))m_i^m(t),\ i=B_n+1,\dots,J,\\ \frac{d}{dt}x_i^m(t) &= 1,\ i=B_n+1,\dots,J,\\ \frac{d}{dt}m_{B_n}^m(t) &= -c^m(t,0)m_{B_n}^m(t)-\partial_x c^m(t,0)\Pi_{B_n}^m(t)\\ & \quad +\sum_{i,j=B_n}^J b^m(t,x_{ij}^c(t),y_{ij}^c(t))m_{ij}^c(t),\\ \frac{d}{dt}\Pi_{B_n}^m(t) &= m_{B_n}^m(t)-c^m(t,0)\Pi_{B_n}^m(t),\\ x_{B_n}^m(t) &= \begin{cases} 0 \text{ if } m_{B_n}^m(t)=0,\\ \frac{\Pi_{B_n}^m(t)}{m_{B_n}^m(t)} \quad \text{otherwise,}\end{cases}\end{cases}$$

$$\begin{cases} \frac{d}{dt}m_j^f(t) &= -c^f(t,y_j^f(t))m_j^f(t),\ j=B_n+1,\dots,J,\\ \frac{d}{dt}y_j^f(t) &= 1,\ j=B_n+1,\dots,J,\\ \frac{d}{dt}m_{B_n}^f(t) &= -c^f(t,0)m_{B_n}^f(t)-\partial_x c^f(t,0)\Pi_{B_n}^m(t)\\ & \quad +\sum_{i,j=B_n}^J b^f(t,x_{ij}^c(t),y_{ij}^c(t))m_{ij}^c(t),\\ \frac{d}{dt}\Pi_{B_n}^f(t) &= m_{B_n}^f(t)-c^f(t,0)\Pi_{B_n}^f(t),\\ y_{B_n}^f(t) &= \begin{cases} 0 \text{ if } m_{B_n}^f(t)=0,\\ \frac{\Pi_{B_n}^m(t)}{m_{B_n}^f(t)} \quad \text{otherwise,}\end{cases}\end{cases} \tag{2.11}$$

$$\begin{cases} \frac{d}{dt}m_{ij}^c(t) &= -c^c(t,x_{ij}^c(t),y_{ij}^c(t))m_{ij}^c(t)+\frac{N_{ij}(t)}{D_{ij}(t)},\ i,j=B_n,\dots,J,\\ \frac{d}{dt}(\tilde{x}_{ij}^c(t),\tilde{y}_{ij}^c(t)) &= \left[(1,1)-(x_{ij}^c(t),y_{ij}^c(y))c^c(t,x_{ij}^c(t),y_{ij}^c(t))\right]m_{ij}^c(t)+\frac{\bar{N}_{ij}(t)}{D_{ij}(t)},\\ (x_{ij}^c(t),y_{ij}^c(t)) &= \begin{cases} 0 \text{ if } m_{ij}^c(t)=0,\\ \frac{(\tilde{x}_{ij}^c(t),\tilde{y}_{ij}^c(t))}{m_{ij}^c(t)} \quad \text{otherwise,}\end{cases}\end{cases}$$

where $\Pi_{B_n}^m(t)=\int_{[l_{i-1}^m(t_n),l_i^m(t_n))} xu^m(t,x)dx$ and $\Pi_{B_n}^f(t)=\int_{[l_{j-1}^f(t_n),l_j^f(t_n))} yu^f(t,y)dy$,

$$(\tilde{x}_{ij}^c,\tilde{y}_{ij}^c)(t)=\int_{[l_{i-1}^m(t_n),l_i^m(t_n))\times[l_{j-1}^f(t_n),l_j^f(t_n))}(x,y)u^c(t,x,y)dxdy,$$

$$\begin{aligned} N_{ij}(t) &= \Theta(x_i^m(t),y_j^f(t))h(x_i^m(t))g(y_j^f(t))m_i^m(t)m_j^f(t)\\ &\quad -\sum_{v=B_n}^J \Theta(x_i^m(t),y_{vj}^c(t))h(x_i^m(t))g(y_{vj}^c(t))m_i^m(t)m_{vj}^c(t)\\ &\quad -\sum_{w=B_n}^J \Theta(x_{iw}^c(t),y_j^f(t))h(x_{iw}^c(t))g(y_j^f(t))m_j^f(t)m_{iw}^c(t)\\ &\quad +\sum_{v,w=B_n}^J \Theta(x_{iw}^c(t),y_{vj}^c(t))h(x_{iw}^c(t))g(y_{vj}^c(t))m_{vj}^c(t)m_{iw}^c(t), \end{aligned} \tag{2.12}$$

(2.13)

$$\begin{aligned}\bar{N}_{ij}(t) = {} & (x_i^m(t), y_j^f(t))\Theta(x_i^m(t), y_j^f(t))h(x_i^m(t))g(y_j^f(t))m_i^m(t)m_j^f(t) \\ & - \sum_{v=B_n}^{J} (x_i^m(t), y_{vj}^c(t))\Theta(x_i^m(t), y_{vj}^c(t))h(x_i^m(t))g(y_{vj}^c(t))m_i^m(t)m_{vj}^c(t) \\ & - \sum_{w=B_n}^{J} (x_{iw}^c(t), y_j^f(t))\Theta(x_{iw}^c(t), y_j^f(t))h(x_{iw}^c(t))g(y_j^f(t))m_j^f(t)m_{iw}^c(t) \\ & + \sum_{v,w=B_n}^{J} (x_{iw}^c(t), y_{vj}^c(t))\Theta(x_{iw}^c(t), y_{vj}^c(t))h(x_{iw}^c(t))g(y_{vj}^c(t))m_{vj}^c(t)m_{iw}^c(t),\end{aligned}$$

and

$$D_{ij}(t) = \gamma + \sum_{i=B_n}^{J} h(x_i^m(t))m_i^m(t) - \sum_{i,j=B_n}^{J} h(x_{ij}^c(t))m_{ij}^c(t) + \sum_{j=B_n}^{J} g(y_j^f(t))m_j^f(t) - \sum_{i,j=B_n}^{J} g(y_{ij}^c(t))m_{ij}^c(t). \tag{2.14}$$

The functions $N_{ij}(t)$, $\bar{N}_{ij}(y)$, and $D_{ij}(t)$ appear in the EBT scheme via the generic marriage function (1.2). The integral of F over a cohort is approximated as

$$\int_{[l_{i-1}^m(t_k), l_i^m(t_k)) \times [l_{j-1}^f(t_k), l_j^f(t_k))} F(t, u^m(t,x), u^f(t,x), u^c(t,x,y), x, y)dxdy \simeq \frac{N_{ij}(t)}{D_{ij}(t)},$$

while its first moment is approximated as

$$\int_{[l_{i-1}^m(t_k), l_i^m(t_k)) \times [l_{j-1}^f(t_k), l_j^f(t_k))} (x,y)F(t, u^m(t,x), u^f(t,x), u^c(t,x,y), x, y)dxdy \simeq \frac{\bar{N}_{ij}(t)}{D_{ij}(t)}.$$

For further details, see [24]. Differential equations constituting the EBT method (2.11) are equipped with the initial boundary conditions (2.4)–(2.10), where, instead of (2.5) and (2.9), the following constraints are applied:

(2.15)

$$\left(m_{ij}^c(t_k), (x_{ij}^c(t_k), y_{ij}^c(t_k))\right) = (0, (0,0)), \text{ where } (i = B_n \vee j = B_n) \wedge i, j \in \{B_n, \dots, J\},$$

for all n.

2.3. Relation between the methods. The simplified method has a more transparent representation than the original EBT model for couples in [24]. The simplified method requires fewer computations while formally retaining the same order of convergence, as we will show in the next section. In the case of male and female populations, it only differs with respect to the evolution of localizations in the boundary cohorts. This simplification was already proposed in [8] for the single species case, whose convergence was proved in [23]. Here, in contrast to [24], we simplified the EBT scheme further by choosing the localization of the couples in the boundary cohorts consistently in terms of the localizations of the male and female populations. More precisely, (2.15) is imposed in contrast to (2.5)–(2.9). This results in a simpler approximation of the marriage function and implies that the characteristics of the age variable for couples and male and female populations remain the same. Moreover, we can check that under this consistent choice of the initial data for couples and female and male populations, the simplified EBT scheme is a particular case of the original EBT scheme. Notice that the formula for $(\tilde{x}^c_{ij}(t), \tilde{y}^c_{ij}(t))$ in (2.11), together with $(x^c_{ij}(t), y^c_{ij}(t)) = \frac{1}{m^c_{ij}(t)}(\tilde{x}^c_{ij}(t), \tilde{y}^c_{ij}(t))$, yields

$$\frac{d}{dt}(x^c_{ij}, y^c_{ij})(t) = \frac{d}{dt}\frac{(\tilde{x}^c_{ij}, \tilde{y}^c_{ij})(t)}{m^c_{ij}(t)} = (1,1) - \frac{(x^c_{ij}(t), y^c_{ij}(t))}{m^c_{ij}(t)}\frac{N_{ij}(t)}{D_{ij}(t)} + \frac{\bar{N}_{ij}(t)}{m^c_{ij}(t)D_{ij}(t)} \tag{2.16}$$

for all $t \in [t_n, t_{n+1})$. Note that $\frac{d}{dt}x^m_i(t) = 1$, and analyze the first component of the equality (2.16), as the same reasoning holds for $y^f_{ij}(t)$. We can rewrite this ODE as

$$\frac{d}{dt}\left(x^c_{ij}(t) - x^m_i(t)\right) = \frac{1}{m^c_{ij}(t)D_{ij}(t)}\left[\bar{N}_{ij}(t) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - x^c_{ij}(t)N_{ij}(t)\right]$$

$$\begin{aligned}
= \frac{1}{m^c_{ij}(t)D_{ij}(t)}\Bigg[&(x^m_i(t) - x^c_{ij}(t))\Theta(x^m_i(t), y^f_j(t))h(x^m_i(t))g(y^f_j(t))m^m_i(t)m^f_j(t) \\
&-(x^m_i(t) - x^c_{ij}(t))\sum_{v=B_n}^{J}\Theta(x^m_i(t), y^c_{vj}(t))h(x^m_i(t))g(y^c_{vj}(t))m^m_i(t)m^c_{vj}(t) \\
&-\sum_{w=B}^{J}(x^c_{iw}(t) - x^c_{ij}(t))\Theta(x^c_{iw}(t), y^f_j(t))h(x^c_{iw}(t))g(y^f_j(t))m^f_j(t)m^c_{iw}(t) \\
&+\sum_{v,w=B}^{J}(x^c_{iw}(t) - x^c_{ij}(t))\Theta(x^c_{iw}(t), y^c_{vj}(t))h(x^c_{iw}(t))g(y^c_{vj}(t))m^c_{vj}(t)m^c_{iw}(t)\Bigg] \\
= \frac{1}{m^c_{ij}(t)D_{ij}(t)}\Bigg[&(x^m_i(t) - x^c_{ij}(t))\Theta(x^m_i(t), y^f_j(t))h(x^m_i(t))g(y^f_j(t))m^m_i(t)m^f_j(t) \\
&-(x^m_i(t) - x^c_{ij}(t))\sum_{v=B_n}^{J}\Theta(x^m_i(t), y^c_{vj}(t))h(x^m_i(t))g(y^c_{vj}(t))m^m_i(t)m^c_{vj}(t) \\
&-\sum_{w=B}^{J}(x^c_{iw}(t) \mp x^m_i(t) - x^c_{ij}(t))\Theta(x^c_{iw}(t), y^f_j(t))h(x^c_{iw}(t))g(y^f_j(t))m^f_j(t)m^c_{iw}(t) \\
&+\sum_{v,w=B}^{J}(x^c_{iw}(t) \mp x^m_i(t) - x^c_{ij}(t))\Theta(x^c_{iw}(t), y^c_{vj}(t))h(x^c_{iw}(t))g(y^c_{vj}(t))m^c_{vj}(t)m^c_{iw}(t)\Bigg].
\end{aligned}$$

One can rewrite the previous system in terms of the auxiliary variables $z_{ij}(t) = x^c_{ij}(t) - x^m_i(t)$. Let us assume that $x^m_i(t_n) = x^c_{ij}(t_n)$ or, equivalently, $z_{ij}(t_n) = 0$ for $j = B_n, \ldots, J$. It is easy to observe from the previous expression that $z_{ij}(t) = 0$ is a solution of the system of ODEs consistent with the initial conditions at t_n, and thus by the uniqueness of the ODE system, we deduce that $z_{ij}(t) = x^c_{ij}(t) - x^m_i(t) = 0$ for all $t \in [t_n, t_{n+1})$. As a consequence, we infer that

$$\frac{d}{dt}\left(x^c_{ij}(t) - t\right) = 0,$$

which explains the last formula in the simplified EBT ODE system (2.2). Notice that in this case, when $x^m_i(t) = x^c_{ij}(t)$ and $y^f_j(t) = y^c_{ij}(t)$ for $i, j = B(t), \ldots, J$, expression $\frac{N_{ij}(t)}{D_{ij}(t)}$ simplifies significantly to (2.3).

3. Embedding in a space of measures. As stated in the introduction, we will describe and analyze the underlying problems and their solutions in a space of nonnegative Radon measures equipped with the flat metric. Setting some models of population dynamics in this space was suggested for the first time in [25]. The age-structured two-sex population model with age as a structure variable was, actually, embedded in a suitable space in [44], where the approach followed after [13, 25, 26].

Alternatively one can investigate the underlying problem not in the positive cone in the space of measures $\mathcal{M}_+(\mathbb{R}_+^N)$ but in the full Banach space, which is a closure of space of bounded Radon measures $\mathcal{M}(\mathbb{R}_+^N)$ with respect to bounded Lipschitz distance [31]. See also [45] for similar *Lipschitz-free space*. It is important to point out that the space defined in [31] is predual to $W^{1,\infty}(\mathbb{R}_+^N)$ and is essentially smaller than $\left(W^{1,\infty}(\mathbb{R}_+^N)\right)^*$.

The crucial reason for considering the predual space instead of $\left(W^{1,\infty}(\mathbb{R}_+^N)\right)^*$ is the lack of continuity of the semigroup generated by the transport operator; see Lemma 2 in [28].

Throughout this paper, $\mathcal{M}^+(\mathbb{R}_+^i)$ denotes the space of nonnegative Radon measures with bounded TV, where $\mathbb{R}_+^i = [0,\infty)^i$, $i = 1, 2$, and $B \in \mathcal{B}(\mathbb{R}_+)$ is a Borel set. We will investigate the following generalization of system (1.1) to spaces of nonnegative Radon measures:

$$
\begin{aligned}
\partial_t \mu_t^m + \partial_x \mu_t^m + \xi^m(t, \mu_t^m, \mu_t^f)\mu_t^m &= 0, && (t,x) \in [0,T] \times \mathbb{R}_+,\\
D_\lambda \mu_t^m(0^+) &= \textstyle\int_{\mathbb{R}_+^2} \beta^m(t, \mu_t^m, \mu_t^f)(z) d\mu_t^c(z), &&\\
\mu_0^m &\in \mathcal{M}^+(\mathbb{R}_+), &&\\[1ex]
\partial_t \mu_t^f + \partial_x \mu_t^f + \xi^f(t, \mu_t^m, \mu_t^f)\mu_t^f &= 0, && (t,x) \in [0,T] \times \mathbb{R}_+,\\
D_\lambda \mu_t^f(0^+) &= \textstyle\int_{\mathbb{R}_+^2} \beta^f(t, \mu_t^m, \mu_t^f)(z) d\mu_t^c(z), &&\\
\mu_0^f &\in \mathcal{M}^+(\mathbb{R}_+), &&\\[1ex]
\partial_t \mu_t^c + \partial_{z_1} \mu_t^c + \partial_{z_2} \mu_t^c + \xi^c(t, \mu_t^m, \mu_t^f, \mu_t^c)\mu_t^c &= \mathcal{T}(t, \mu_t^m, \mu_t^f, \mu_t^c), && (t,z) \in [0,T] \times \mathbb{R}_+^2,\\
\mu_t^c(\{0\} \times B) = \mu_t^c(B \times \{0\}) &= 0, &&\\
\mu_0^m &\in \mathcal{M}^+(\mathbb{R}_+^2). &&
\end{aligned}
\tag{3.1}
$$

Measures μ_t^m, $\mu_t^f \in \mathcal{M}_+(\mathbb{R}_+^1)$, and $\mu_t^c \in \mathcal{M}_+(\mathbb{R}_+^2)$ describe the distribution of males, females, and couples, respectively, at time t. Functions ξ^m, ξ^f, and ξ^c are equivalents of functions c^m, c^f, and c^c and describe the disappearance of individuals, while functions β^m and β^f constitute equivalents for b^m and b^f, birth rates of males and females, respectively. Symbols $D_\lambda \mu_t^m(0^+)$ and $D_\lambda \mu_t^f(0^+)$ denote Radon–Nikodym derivatives of μ_t^m and μ_t^f, respectively, with respect to the one-dimensional Lebesgue measure λ at point 0, as we assume that the support of the singular part of measures μ_t^m and μ_t^f does not contain 0.

Before we comment on the marriage function, let us define the distribution of single males and females, s_t^m and s_t^f, respectively. By *single* s_t^m or s_t^f we mean not only those who have never been married by time t, but also those who are divorced or widowed at time t. Let measures σ_t^m and σ_t^f be projections of μ_t^c on $\mathbb{R}_+$ and describe a distribution of males and females, respectively, who are married at time t:

$$
\sigma_t^m(B) = \mu_t^c(B \times \mathbb{R}_+) \quad \text{and} \quad \sigma_t^f(B) = \mu_t^c(\mathbb{R}_+ \times B). \tag{3.2}
$$

Now s_t^m and s_t^f can be easily defined as

$$s_t^m(B) = (\mu_t^m - \sigma_t^m)(B \times \mathbb{R}_+) \quad \text{and} \quad s_t^f(B) = (\mu_t^f - \sigma_t^f)(B \times \mathbb{R}_+) \tag{3.3}$$

Following [44] and [24], we adopt the following definition of generic marriage function:

$$\begin{aligned} \mathcal{T}(t, \mu_t^m, \mu_t^f, \mu_t^c) &= \mathcal{F}(t, \mu_t^m - \sigma_t^m, \mu_t^f - \sigma_t^f) \\ = \mathcal{F}(t, s_t^m, s_t^f) &= \frac{\Theta(x,y)h(x)g(y)}{\gamma + \int_0^\infty h(z)ds_t^m(z) + \int_0^\infty g(w)ds_t^f(w)}(s_t^m \otimes s_t^f), \end{aligned} \tag{3.4}$$

where $(s_t^m \otimes s_t^f)$ is a product measure on $\mathbb{R}_+^2$.

3.1. EBT schemes for fully nonlinear model in space of measures. The output of the numerical method should evolve in the same space as the solution of (3.1). This requirement can be easily satisfied by defining measures $\nu_{k,t}^m$, $\nu_{k,t}^f$, and $\nu_{k,t}^c$ as a linear combination of Dirac measures:

$$\nu_{k,t}^m := \sum_{i=B_n}^J m_i^m(t)\delta_{\{x_i^m(t)\}},\ \nu_{k,t}^f := \sum_{j=B_n}^J m_j^f(t)\delta_{\{y_j^f(t)\}},\ \nu_{k,t}^c := \sum_{i,j=B_n}^J m_{ij}^c(t)\delta_{\{x_{ij}^c(t), y_{ij}^c(t)\}}, \tag{3.5}$$

where $t \in [t_n, t_{n+1})$ (such that $t_{n+1} - t_n \le a_0$), and $(x_i^m(t), m_i^m(t))$, $(y_j^f(t), m_j^f(t))$, $((x_{ij}^c(t), y_{ij}^c(t)), m_{ij}^c(t))$ is the output of the EBT algorithm (2.11). The subscript k ($k = J - B_0$) is related to the approximation of the initial condition, as it is equal to the amount of initial cohorts for males and females,

$$\nu_{k,0}^m := \sum_{i=B_0}^J m_i^m(0)\delta_{\{x_i^m(0)\}},\ \nu_{k,0}^f := \sum_{j=B_0}^J m_j^f(0)\delta_{\{y_j^f(0)\}},\ \nu_{k,0}^c := \sum_{i,j=B_0}^J m_{ij}^c(0)\delta_{\{x_{ij}^c(0), y_{ij}^c(0)\}}. \tag{3.6}$$

For transparency of notation we will omit subscript k in the remainder of the paper.

For $t \in [t_n, t_{n+1})$ the system (3.1) is approximated with the following EBT scheme:

$$\left\{\begin{array}{c} \frac{d}{dt}m_i^m(t) = -\xi^m(t, \nu_t^m, \nu_t^f)(x_i^m(t))m_i^m(t),\ i = B_n+1, \dots, J, \\ \frac{d}{dt}x_i^m(t) = 1,\ i = B_n, \dots, J, \\ \frac{d}{dt}m_{B_n}^m(t) = -\xi^m(t, \nu_t^m, \nu_t^f)(x_{B_n}^m(t))m_{B_n}^m(t) \\ + \sum_{i,j=B_n}^J \beta^m(t, \nu_t^m, \nu_t^f)(x_{ij}^c(t), y_{ij}^c(t))m_{ij}^c(t), \end{array}\right.$$

$$\left\{\begin{array}{c} \frac{d}{dt}m_j^f(t) = -\xi^f(t, \nu_t^m, \nu_t^f)(y_j^f(t))m_j^f(t),\ j = B_n+1, \dots, J, \\ \frac{d}{dt}y_j^f(t) = 1,\ j = B_n, \dots, J, \\ \frac{d}{dt}m_{B_n}^f(t) = -\xi^f(t, \nu_t^m, \nu_t^f)(y_{B_n}^f(t))m_{B_n}^f(t) \\ + \sum_{i,j=B_n}^J \beta^f(t, \nu_t^m, \nu_t^f)(x_{ij}^c(t), y_{ij}^c(t))m_{ij}^c(t), \end{array}\right. \tag{3.7}$$

$$\left\{\begin{array}{c} \frac{d}{dt}m_{ij}^c(t) = -\xi^c(t, \nu_t^m, \nu_t^f, \nu_t^c)(x_{ij}^c(t), y_{ij}^c(t))m_{ij}^c(t) \\ + \frac{N_{ij}(t)}{D_{ij}(t)},\ i,j = B_n, \dots, J, \\ \frac{d}{dt}(x_{ij}^c(t), y_{ij}^c(t)) = (1,1),\ i,j = B_n, \dots, J, \end{array}\right.$$

where $\frac{N_{ij}(t)}{D_{ij}(t)}$ is defined by (2.3).

Remark 3.1. Note that application of measures (3.5) to function $\mathcal{T}$ (defined in (3.4)) results in the formula

$$\mathcal{T}(t,\nu_t^m,\nu_t^f,\nu_t^c)(x_{ij}^c(t),y_{ij}^c(t)) = \sum_{i,j=B_n}^{J} \frac{N_{ij}(t)}{D_{ij}(t)}\delta_{\{x_{ij}^c(t),y_{ij}^c(t)\}}$$

for $t \in [t_n, t_{n+1})$.

3.2. Notation, definitions, and important facts. Throughout this paper we understand that $\mathbf{C}^1(\mathbb{R}_+^N;\mathbb{R})$ is a class of differentiable functions from $\mathbb{R}_+^N$ to $\mathbb{R}$. Also in this paper, the Lipschitz bounded distance (the flat metric) d_N (defined below) is associated with space $\mathcal{M}_+(\mathbb{R}_+^N)$, where $N \in \mathbb{N}$.

Definition 3.2. *Let $\mu,\nu \in \mathcal{M}_+(\mathbb{R}_+^N)$, where $N \in \mathbb{N}$. The distance function $d_N : \mathcal{M}_+(\mathbb{R}_+^N) \times \mathcal{M}_+(\mathbb{R}_+^N) \to [0,\infty)$ is defined by*

$$d_N(\mu_1,\mu_2) = \sup\left\{\int_{\mathbb{R}_+^N} \varphi\, d(\mu_1-\mu_2):\ \varphi \in \mathbf{C}^1(\mathbb{R}_+^N;\mathbb{R}) \text{ and } \|\varphi\|_{\mathbf{W^{1,\infty}}(\mathbb{R}_+^N,\mathbb{R})} \le 1\right\},$$

where $\|\varphi\|_{\mathbf{W^{1,\infty}}(\mathbb{R}_+^N,\mathbb{R})} = \max\{\|\varphi\|_{\mathbf{L}^\infty}, \|\partial_{x_1}\varphi\|_{\mathbf{L}^\infty}, \|\partial_{x_2}\varphi\|_{\mathbf{L}^\infty},\dots,\|\partial_{x_N}\varphi\|_{\mathbf{L}^\infty}\}$, $N \in \mathbb{N}$.

The metric d_N is a distance derived from the dual norm of $\mathbf{W^{1,\infty}}(\mathbb{R}_+^N,\mathbb{R})$. For the sake of simplicity we will write $\|\cdot\|_{\mathbf{W^{1,\infty}}}$ instead of $\|\cdot\|_{\mathbf{W^{1,\infty}}(\mathbb{R}_+^N,\mathbb{R})}$ when there is no misunderstanding. Space $\left(\mathcal{M}_+(\mathbb{R}_+^N), d_n\right)$ possesses two favorable features which are necessary for investigating the convergence of the numerical scheme. It is well known [43, 10, 26] that the metric space $\left(\mathcal{M}_+(\mathbb{R}_+^N), d_N\right)$ is complete, separable and that the convergence in metric d_N is equivalent to narrow convergence of sequences in $\mathcal{M}_+(\mathbb{R}_+^N)$.

As we are concerned with linear combinations of Dirac measures, it is worth making the following observation.

Lemma 3.3. *Let $\mu = \sum_{i=1}^J m_i\delta_{x_i}$ and $\tilde\mu = \sum_{i=1}^J \tilde m_i\delta_{\tilde x_i}$, where $J \in \mathbb{N}$, x_i, $\tilde x_i \in \mathbb{R}_+^N$, and m_i, $\tilde m_i \in \mathbb{R}_+$. Then,*

$$d_N(\mu,\tilde\mu) \le \sum_{i=1}^J \left(\|x_i - \tilde x_i\| m_i + |m_i - \tilde m_i|\right).$$

Proof. The proof consists of applying the triangle inequality and Definition 3.2

to get

$$
\begin{aligned}
d_N(\mu,\tilde{\mu}) &\leq d_N\left(\sum_{i=1}^{J} m_i\delta_{x_i}, \sum_{i=1}^{J} m_i\delta_{\tilde{x}_i}\right) + d_N\left(\sum_{i=1}^{J} m_i\delta_{\tilde{x}_i}, \sum_{i=1}^{J} \tilde{m}_i\delta_{\tilde{x}_i}\right)\\
&\leq \sup\left\{\int_{\mathbb{R}_+^N} \varphi\, d\left(\sum_{i=1}^{J} m_i\delta_{x_i} - \sum_{i=1}^{J} m_i\delta_{\tilde{x}_i}\right):\ \varphi\in\mathbf{C}^1(\mathbb{R}_+^N;\mathbb{R}) \text{ and } \|\varphi\|_{\mathbf{W}^{1,\infty}(\mathbb{R}_+^N,\mathbb{R})}\leq 1\right\}\\
&\quad + \sup\left\{\int_{\mathbb{R}_+^N} \varphi\, d\left(\sum_{i=1}^{J} m_i\delta_{\tilde{x}_i} - \sum_{i=1}^{J} \tilde{m}_i\delta_{\tilde{x}_i}\right):\ \varphi\in\mathbf{C}^1(\mathbb{R}_+^N;\mathbb{R}) \text{ and } \|\varphi\|_{\mathbf{W}^{1,\infty}(\mathbb{R}_+^N,\mathbb{R})}\leq 1\right\}\\
&\leq \sup\left\{\sum_{i=1}^{J} |\varphi(x_i)-\varphi(\tilde{x}_i)|m_i:\ \varphi\in\mathbf{C}^1(\mathbb{R}_+^N;\mathbb{R}) \text{ and } \|\varphi\|_{\mathbf{W}^{1,\infty}(\mathbb{R}_+^N,\mathbb{R})}\leq 1\right\}\\
&\quad + \sup\left\{\sum_{i=1}^{J} \varphi(\tilde{x}_i)|m_i-\tilde{m}_i|:\ \varphi\in\mathbf{C}^1(\mathbb{R}_+^N;\mathbb{R}) \text{ and } \|\varphi\|_{\mathbf{W}^{1,\infty}(\mathbb{R}_+^N,\mathbb{R})}\leq 1\right\}\\
&\leq \sum_{i=1}^{J} \|x_i-\tilde{x}_i\|m_i + \sum_{i=1}^{J} |m_i-\tilde{m}_i|.
\end{aligned}
$$

DEFINITION 3.4. *Let (E,d) be a metric space. A family of bounded operators $S:[0,T]\times[0,T]\times E\to E$ is called a Lipschitz semiflow if for steps $s,t\in[0,T]$ and times τ such that $\tau,\tau+s,\tau+t,\tau+s+t\in[0,T]$ the following conditions are satisfied:*

1. $S(0;\tau)=I$,
2. $S(t+s;\tau)=S(t;\tau+s)S(s;\tau)$,
3. $d(S(t;\tau)\mu, S(s;\tau)\nu)\leq L(d(\mu,\nu)+|t-s|)$.

Let us define product spaces

$$\mathcal{U}=\mathcal{M}_+(\mathbb{R}_+)\times\mathcal{M}_+(\mathbb{R}_+)\times\mathcal{M}_+(\mathbb{R}_+^2) \text{ and } \mathcal{V}=\mathcal{M}_+(\mathbb{R}_+)\times\mathcal{M}_+(\mathbb{R}_+).$$

We will investigate the problem of approximation of weak solutions of model (3.1) in the metric space $(\mathcal{U},\mathbf{d})$, where $\mathbf{d}=d_1+d_1+d_2$.

The following definition of weak solution to (3.1) was proposed in [44] (see also [13]).

DEFINITION 3.5. *A triple $\mathbf{u}=(\mu^m,\mu^f,\mu^c):[a,b]\to\mathcal{U}$ is a weak solution to the system (3.1) on the interval $[a,b]$ if μ^m,μ^f,μ^c are narrowly continuous with respect to time and for all $(\varphi^m,\varphi^f,\varphi^c)$ such that $\varphi^m,\varphi^f\in(\mathbf{C^1}\cap\mathbf{W^{1,\infty}})([a,b]\times\mathbb{R}_+;\mathbb{R})$ and $\varphi^c\in(\mathbf{C^1}\cap\mathbf{W^{1,\infty}})([a,b]\times\mathbb{R}_+^2;\mathbb{R})$, the following equalities hold:*

$$
\begin{aligned}
&\int_a^b\int_{\mathbb{R}_+}\Big(\partial_t\varphi^i(t,x)+\partial_x\varphi^i(t,x)-\xi^i(t,\mu_t^m,\mu_t^f)\varphi^i(t,x)\Big)\,d\mu_t^i(x)dt\\
&\qquad + \int_a^b \varphi^i(t,0)\int_{\mathbb{R}_+^2}\beta^i(t,\mu_t^m,\mu_t^f)(z)d\mu_t^c(z)dt\\
&= \int_{\mathbb{R}_+}\varphi^i(b,x)d\mu_b^i(x) - \int_{\mathbb{R}_+}\varphi^i(a,x)d\mu_a^i(x) \text{ for } i=f,m
\end{aligned}
$$

and

$$\int_a^b \int_{\mathbb{R}_+} \Big(\partial_t \varphi^c(t,z) + \partial_x \varphi^c(t,z) + \partial_y \varphi^c(t,z) - \xi^c(t, \mu_t^m, \mu_t^f, \mu_t^c) \varphi^c(t,z) \Big) \, d\mu_t^c(z) dt$$

$$+ \int_a^b \int_{\mathbb{R}_+^2} \varphi^c(t,z) d\mathcal{T}(t, \mu_t^m, \mu_t^f, \mu_t^c)(z) dt$$

$$= \int_{\mathbb{R}_+} \varphi^c(b,z) d\mu_b^c(z) - \int_{\mathbb{R}_+} \varphi^c(a,z) d\mu_a^c(z).$$

Here, narrowly continuous functions are understood in the sense of narrow convergence introduced in [2].

Assumption 3.6. *We make the following assumptions on the model functions:*

$$\begin{array}{rcl} \xi^m, \xi^f & \in & \mathbf{BC}^{0,1}([0,T] \times \mathcal{V}, \mathbf{W}^{\mathbf{1},\infty}(\mathbb{R}_+;\mathbb{R})), \\ \beta^m, \beta^f & \in & \mathbf{BC}^{0,1}([0,T] \times \mathcal{V}, \mathbf{W}^{\mathbf{1},\infty}(\mathbb{R}_+^2;\mathbb{R})), \\ \xi^c & \in & \mathbf{BC}^{0,1}([0,T] \times \mathcal{U}, \mathbf{W}^{\mathbf{1},\infty}(\mathbb{R}_+^2;\mathbb{R})), \\ \mathcal{T} & \in & \mathbf{BC}^{0,1}([0,T] \times \mathcal{U}, \mathcal{M}_+(\mathbb{R}_+^2)). \end{array}$$

We understand that spaces $\mathbf{BC}^{0,1}([0,T] \times \mathcal{V}, X)$ and $\mathbf{BC}^{0,1}([0,T] \times \mathcal{U}, X)$ are spaces of X valued functions, bounded with respect to the $\|\cdot\|_X$ norm, continuous with respect to time, and Lipschitz continuous with respect to measure variables. We understand that function $\mathcal{T}$ is bounded with respect to the $\|\cdot\|_{\mathbf{W}^{\mathbf{1},\infty}}$ norm, as its values are in the space of nonnegative measures. The norm $\|\cdot\|_{\mathbf{BC}^{0,1}}$ in the $\mathbf{BC}^{0,1}$ space is defined in the following way:

$$\|f\|_{\mathbf{BC}^{0,1}} = \sup_{t \in [0,T], \mathbf{v} \in Y} \big(\|f(t,\mathbf{v})\|_X + \mathbf{Lip}_{\mathbf{v}}(f(t,\cdot)) \big),$$

where $Y = \mathcal{V}$ or $Y = \mathcal{U}$, and $\mathbf{Lip}_{\mathbf{v}}(f(t,\cdot))$ is a Lipschitz constant of $f(t,\cdot)$.

We will write $\|\cdot\|_{\mathbf{BC}}$ instead of $\|\cdot\|_{\mathbf{BC}^{0,1}}$ for the sake of clarity.

Remark 3.7. It was proved in [44], under Assumption 3.6 and Theorem 2.9 in [9], that solutions to (3.1) form Lipschitz semiflows $S : [0,T] \times [0,T] \times \mathcal{U} \to \mathcal{U}$.

Remark 3.8. Let ν_t^m, ν_t^f, ν_t^c be defined by (3.5), where $(x_i^m(t), m_i^m(t))$, $(y_j^f(t), m_j^f(t))$, $((x_{ij}^c(t), y_{ij}^c(t)), m_{ij}^c(t))$ is the output of the EBT algorithm (3.7). Then it is easy to show, using Lemma 3.3, that the solution of numerical method $\mathbf{v_t} = (\nu_t^m, \nu_t^f, \nu_t^c) : [0,T] \to (\mathcal{U}, \mathbf{d})$ is a Lipschitz continuous map. The rigorous proof can be seen in [28, Lemma 4.2].

Remark 3.9. It was shown in [34] that for a given Radon measure, the initial conditions $u_0^m(x)$ and $u_0^f(y)$ can be approximated in the flat metric with an arbitrarily good precision by a combination of Dirac measures defined by (3.6). The same reasoning can be applied to the initial condition $u_0^c(\cdot, x, y)$.

Proposition 3.10. *Let $S : [0,T] \times [0,T] \times E \to E$ be a Lipschitz semiflow. For every Lipschitz continuous map $[0,T] \ni t \mapsto \nu_t \in E$ the following estimate holds:*

$$d(\nu_t, S(t;0)\nu_0) \le L \int_{[0,t]} \liminf_{h \to 0} \frac{d(\nu_{\tau+h}, S(h;\tau)\nu_\tau)}{h} d\tau.$$

The proof of the proposition is similar to the proof of Theorem 2.9 in [9].

LEMMA 3.11. *Let us consider the equation*

$$
\begin{array}{rcll}
\partial_t \mu_t + \partial_x \mu_t + \xi(t,\mu)\mu & = & g(t)\delta_{\{0\}}, & (t,x) \in [0,T] \times \mathbb{R}_+, \\
\mu_0^m & \in & \mathcal{M}^+(\mathbb{R}_+), &
\end{array}
\tag{3.8}
$$

where $g(t) := \int_{\mathbb{R}_+^2} \beta(t,\mu_t)(z) d\mu_t^*(z)$ *with the coefficients* ξ *and* β *satisfying the conditions in Assumptions* 3.6 *for the respective coefficients. Here,* $\mu_t^*(z)$ *is a given curve of measures with finite total mass for all* $t \geq 0$ *narrowly continuous with respect to time. The weak solution is given by*

$$
\mu_t = (\Phi(t;\cdot)\#\mu_0) \exp\left(-\int_0^t \xi(\tau, \cdot - t + \tau) \mathrm{d}\tau\right) + \tilde{\mu}_t
$$

with $\Phi(t;x) = x + t$*,* $\Phi(t;\cdot)\#\mu_0$ *refers to the push forward of a measure through the map* $\Phi(t;\cdot)$*, and*

$$
\tilde{\mu}_t = \int_0^t \exp\left(-\int_\tau^t \xi(s, \cdot + s - t) \mathrm{d}s\right) g(\tau)\delta_{\{t-\tau\}} \mathrm{d}\tau.
$$

More precisely, $\tilde{\mu}_t$ *is defined by*

$$
\begin{aligned}
&\left\langle \varphi, \int_0^t \exp\left(-\int_\tau^t \xi(s, \cdot + s - t) \mathrm{d}s\right) g(\tau)\delta_{\{t-\tau\}} \mathrm{d}\tau \right\rangle \\
&\quad = \int_0^t \varphi(t-\tau) \exp\left(-\int_\tau^t \xi(s, s-\tau) \mathrm{d}s\right) g(\tau) \mathrm{d}\tau
\end{aligned}
$$

for all test functions φ*, and its Radon–Nikodym derivative with respect to the Lebesgue measure* λ *on the line is given by the bounded function*

$$
f(t,x) = \begin{cases} h(t, t-x) - \displaystyle\int_x^t \xi(t, t-\tau+x) h(t, \tau - x) \mathrm{d}\tau, & 0 \leq x \leq t, \\ 0 & x > t, \end{cases}
$$

that is, $\tilde{\mu}_t = f(t,\cdot)\lambda$*, with* $h(t,\tau) = \exp(-\int_\tau^t \xi(s, s-\tau) \mathrm{d}s) g(\tau)$*.*

Proof. Let us denote the right-hand side of (3.8) as ω_t. It is known that the semigroup of this McKendrick–von Foerster equation can be written in terms of the characteristics of the flow; see [36]. More precisely, now defining $\Phi(\tau, t; x) = x + t - \tau$ for any $t, \tau \geq 0$, we see that the unique solution reads

$$
\begin{aligned}
\mu_t = {} & \Phi(0,t;\cdot)\# \left(\mu_0 \exp\left(-\int_0^t \xi(\tau, \Phi(0,\tau;\cdot)) \mathrm{d}\tau\right)\right) \\
& + \int_0^t \Phi(\tau, t, \cdot)\# \left(\omega_\tau \exp\left(-\int_\tau^t \xi(s, \Phi(\tau, s; \cdot)) \mathrm{d}s\right)\right) \mathrm{d}\tau.
\end{aligned}
$$

By substituting $\omega_t = g(t)\delta_{\{0\}}$ and performing simple algebraic manipulations, we obtain

$$
\begin{aligned}
\mu_t = {} & (\Phi(0,t;\cdot)\#\mu_0) \exp\left(-\int_0^t \xi(\tau, \cdot - t + \tau) \mathrm{d}\tau\right) \\
& + \int_0^t \exp\left(-\int_\tau^t \xi(s, \cdot - t + s) \mathrm{d}s\right) g(\tau)\delta_{\{t-\tau\}} \mathrm{d}\tau,
\end{aligned}
$$

as stated above. The final part of the lemma is obtain by computing directly the Radon–Nikodym derivative using the formula for $\tilde{\mu}_t$ acting on test functions. Choosing for sufficiently small ε the test function defined as

$$\varphi_\varepsilon = \frac{1}{2\varepsilon}\chi_{(x-\varepsilon,x+\varepsilon)}\,, \quad \text{where } \chi_I \text{ stands for the characteristic function of the interval } I\,,$$

we can check that for $0 \le x \le t$,

$$\begin{aligned}\lim_{\varepsilon\to 0}\langle\tilde{\mu}_t,\varphi_\varepsilon\rangle &= \lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\int_{\max(0,t-x-\varepsilon)}^{\min(0,t-x+\varepsilon)} h(t,\tau)\mathrm{d}\tau \\ &= \frac{d}{dt}\int_x^t f(t,\tau-x)\mathrm{d}\tau = h(t,t-x) + \int_x^t \frac{\partial h}{\partial t}(t,\tau-x)\mathrm{d}\tau\,.\end{aligned}$$

It is straightforward to check that for $x > t$, the previous computation gives 0, finishing the proof of the result. ∎

The following lemma will be used extensively in the proof of the main theorem on the convergence of the EBT scheme

LEMMA 3.12. *Let* $t \in [\tau,\tau+h) \subset [t_n,t_{n+1})$ *and* $t \mapsto (\nu_t^m,\nu_t^f,\nu_t^c)$, $t \mapsto (\bar{\nu}_t^m,\bar{\nu}_t^f,\bar{\nu}_t^c) \in \mathcal{U}$*, where*

$$\nu_t^m := \sum_{i=B_n}^J m_i^m(t)\delta_{\{x_i^m(t)\}},\ \nu_t^f := \sum_{j=B_n}^J m_j^f(t)\delta_{\{y_j^f(t)\}},\ \nu_t^c := \sum_{i,j=B_n}^J m_{ij}^c(t)\delta_{\{x_{ij}^c(t),y_{ij}^c(t)\}}$$

and

$$\bar{\nu}_t^m := \sum_{i=B_n}^J \bar{m}_i^m(t)\delta_{\{x_i^m(t)\}},\ \bar{\nu}_t^f := \sum_{j=B_n}^J \bar{m}_j^f(t)\delta_{\{y_j^f(t)\}},\ \bar{\nu}_t^c := \sum_{i,j=B_n}^J \bar{m}_{ij}^c(t)\delta_{\{x_{ij}^c(t),y_{ij}^c(t)\}}$$

be certain Lipschitz continuous maps in d_N *such that* $(\nu_\tau^m,\nu_\tau^f,\nu_\tau^c) = (\bar{\nu}_\tau^m,\bar{\nu}_\tau^f,\bar{\nu}_\tau^c)$ *and*

$$\begin{aligned}C_1 &= \mathrm{Lip}_t(\nu^m) + \mathrm{Lip}_t(\bar{\nu}^m) + \mathrm{Lip}_t(\nu^f) + \mathrm{Lip}_t(\bar{\nu}^f),\\ C_2 &= \mathrm{Lip}_t(\nu^m) + \mathrm{Lip}_t(\bar{\nu}^m) + \mathrm{Lip}_t(\nu^f) + \mathrm{Lip}_t(\bar{\nu}^f) + \mathrm{Lip}_t(\nu^c) + \mathrm{Lip}_t(\bar{\nu}^c),\end{aligned}$$

where $\mathrm{Lip}_t(\cdot)$ *stands for the Lipschitz constant with respect to* t*. Then the following estimates hold for* $i,j = B_n,\ldots,J$*:*

$$\begin{aligned}\left|\xi^k(t,\nu_t^m,\nu_t^f)(x_i(t)) - \xi^k(t,\bar{\nu}_t^m,\bar{\nu}_t^f)(x_i(t))\right| &\le C_1 h\|\xi^k\|_{\mathbf{BC}},\ k=m,f,\\ \left|\beta^k(t,\nu_t^m,\nu_t^f)(x_{ij}(t),y_{ij}(t)) - \beta^k(t,\bar{\nu}_t^m,\bar{\nu}_t^f)(x_{ij}(t),y_{ij}(t))\right| &\le C_1 h\|\beta^k\|_{\mathbf{BC}},\ k=m,f,\\ \left|\xi^c(t,\nu_t^m,\nu_t^f,\nu_t^c)(x_{ij}(t),y_{ij}(t)) - \xi^c(t,\bar{\nu}_t^m,\bar{\nu}_t^f,\bar{\nu}_t^c)(x_{ij}(t),y_{ij}(t))\right| &\le 2C_2 h\|\xi^c\|_{\mathbf{BC}}.\end{aligned} \tag{3.9}$$

Moreover, given functions $m_i(t), \bar{m}_i(t) \in \mathrm{W}^{1,\infty}([\tau,\tau+h),\mathbb{R}_+)$*, such that* $m_i(\tau) = \bar{m}_i(\tau)$ $(i = B_n,\ldots,J)$*, and constants*

$$C_3 = \sup_{t\in[0,T]}\left\{\|\nu_t^m\|_{\mathrm{TV}} + \|\nu_t^f\|_{\mathrm{TV}} + \|\nu_t^c\|_{\mathrm{TV}} + \|\bar{\nu}_t^m\|_{\mathrm{TV}} + \|\bar{\nu}_t^f\|_{\mathrm{TV}} + \|\bar{\nu}_t^c\|_{\mathrm{TV}}\right\},$$

$$C_4 = \mathrm{Lip}_t(m_{B_n}) + \mathrm{Lip}_t(\bar{m}_{B_n}),\ C_5 = C_1C_3 + C_1,\ C_6 = 2C_2C_3 + C_2,$$

we can conclude that, for $k = m, f$ and for any $\varphi \in W^{1,\infty}$ with norm less than 1,

(3.10)

$$\left| \sum_{i=B_n}^{J} \varphi(x_i(t))[\xi^k(t, \nu_t^m, \nu_t^f)(x_i(t))m_i(t) - \xi^k(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(x_i(t))\bar{m}_i(t)] \right| \leq C_5 h \|\xi^k\|_{\mathbf{BC}},$$

$$\left| \sum_{i=B_n}^{J} \varphi(x_i(t))[\beta^k(t, \nu_t^m, \nu_t^f)(x_{ij}(t), y_{ij}(t))m_{ij}(t) \right.$$
$$\left. - \beta^k(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(x_{ij}(t), y_{ij}(t))\bar{m}_{ij}(t)] \right| \leq C_5 h \|\beta^k\|_{\mathbf{BC}},$$

$$\left| \sum_{i=B_n}^{J} \varphi(x_{ij}(t), y_{ij}(t))[\xi^c(t, \nu_t^m, \nu_t^f, \nu_t^c)(x_{ij}(t), y_{ij}(t))m_{ij}(t) \right.$$
$$\left. - \xi^c(t, \bar{\nu}_t^m, \bar{\nu}_t^f, \bar{\nu}_t^c)(x_{ij}(t), y_{ij}(t))\bar{m}_{ij}(t)] \right| \leq C_6 h \|\xi^c\|_{\mathbf{BC}},$$

and for $k = m, f$,

$$\left| \int_\tau^{\tau+h} \xi^k(t, \nu_t^m, \nu_t^f)(x_{B_n}(t))m_{B_n}(t)dt \right.$$
$$\left. - \int_\tau^{\tau+h} \xi^k(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(x_{B_n}(t))\bar{m}_{B_n}(t)dt \right| \leq (C_4 + C_1 C_3)h^2 \|\xi^k\|_{\mathbf{BC}}. \tag{3.11}$$

Proof. We will prove the above estimates for the male function ξ^m only, since the same strategy can be used for the remaining functions. To simplify the notation, we avoid the explicit t dependence of ξ^m. Let us start with (3.9). Using the triangular inequality and the assumption $(\nu_\tau^m, \nu_\tau^f, \nu_\tau^c) = (\bar{\nu}_\tau^m, \bar{\nu}_\tau^f, \bar{\nu}_\tau^c)$, we have

$$\begin{aligned}
\left|\xi^m(\nu_t^m, \nu_t^f)(x) - \xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x)\right| &\leq \left|\xi^m(\nu_t^m, \nu_t^f)(x) - \xi^m(\bar{\nu}_t^m, \nu_t^f)(x)\right| \\
&\quad + \left|\xi^m(\bar{\nu}_t^m, \nu_t^f)(x) - \xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x)\right| \\
&\leq \left|\xi^m(\nu_t^m, \nu_t^f)(x) - \xi^m(\nu_\tau^m, \nu_t^f)(x)\right| \\
&\quad + \left|\xi^m(\bar{\nu}_\tau^m, \nu_t^f)(x) - \xi^m(\bar{\nu}_t^m, \nu_t^f)(x)\right| \\
&\quad + \left|\xi^m(\bar{\nu}_t^m, \nu_t^f)(x) - \xi^m(\bar{\nu}_t^m, \nu_\tau^f)(x)\right| \\
&\quad + \left|\xi^m(\bar{\nu}_t^m, \bar{\nu}_\tau^f)(x) - \xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x)\right|
\end{aligned}$$

for all $x \geq 0$ from which

$$\begin{aligned}
&\left|\xi^m(\nu_t^m, \nu_t^f)(x) - \xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x)\right| \\
&\qquad \leq \|\xi^m\|_{\mathbf{BC}} \sup_{t \in [\tau, \tau+h)} \Big(d_1(\nu_t^m, \nu_\tau^m) + d_1(\bar{\nu}_\tau^m, \bar{\nu}_t^m) + d_1(\nu_t^f, \nu_\tau^f) + d_1(\bar{\nu}_\tau^f, \bar{\nu}_t^f) \Big) \\
&\qquad \leq C_1 h \|\xi^m\|_{\mathbf{BC}}
\end{aligned}$$

for all $x \geq 0$.

Let us move to the proof of (3.10). For any $\varphi \in W^{1,\infty}$ with norm less than 1 and for all $x \geq 0$, we have

$$\left| \sum_{i=B_n}^{J} \varphi(x)[\xi^m(\nu_t^m, \nu_t^f)(x) m_i(t) - \xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x)\bar{m}_i(t)] \right|$$

$$\leq \left| \sum_{i=B_n}^{J} \varphi(x)[\xi^m(\nu_t^m, \nu_t^f)(x) m_i(t) - \xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x) m_i(t)] \right| + \left| \sum_{i=B_n}^{J} \varphi(x)\xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x)[m_i(t) - \bar{m}_i(t)] \right|.$$

Using inequalities (3.9) and the fact that $m_i(t) \geq 0$ for $i = B_n, \ldots, J$, the first term of the right-hand side of the above inequality can be estimated as

$$\left| \sum_{i=B_n}^{J} \varphi(x)[\xi^m(\nu_t^m, \nu_t^f)(x) m_i(t) - \xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x) m_i(t)] \right|$$

$$\leq C_1 h \|\xi^m\|_{\mathbf{BC}} \sum_{i=B_n}^{J} |\varphi(x)|\, m_i(t) \leq C_1 C_3 h \|\xi^m\|_{\mathbf{BC}}\,.$$

To estimate the second term of the right-hand side, we proceed as follows:

$$\left| \sum_{i=B_n}^{J} \varphi(x)\xi^m(\bar{\nu}_t^m, \bar{\nu}_t^f)(x)[m_i(t) - \bar{m}_i(t)] \right| \leq \|\xi^m(\nu_t^m, \nu_t^f)\|_{L^\infty} d_1(\nu_t^m, \bar{\nu}_t^m) \leq C_1 h \|\xi^m\|_{\mathbf{BC}}$$

for any $\varphi \in W^{1,\infty}$ with norm less than 1 and for all $x \geq 0$. Here, we used that

$$d_1(\nu_t^m, \bar{\nu}_t^m) = \sum_{i=B_n}^{J} |m_i(t) - \bar{m}_i(t)| \leq d_1(\nu_t^m, \nu_\tau^m) + d_1(\bar{\nu}_\tau^m, \bar{\nu}_t^m),$$

since both measures are combinations of Dirac Deltas at the same locations and assumption $(\nu_\tau^m, \nu_\tau^f, \nu_\tau^c) = (\bar{\nu}_\tau^m, \bar{\nu}_\tau^f, \bar{\nu}_\tau^c)$. The proof for (3.11) is easier, as we have

$$\begin{aligned}
&\left| \int_\tau^{\tau+h} \Big(\xi^m(t, \nu_t^m, \nu_t^f)(x_{B_n}(t)) m_{B_n}^m(t) - \xi^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(x_{B_n}(t)) \bar{m}_{B_n}^m(t) \Big)\, dt \right| \\
&\leq \int_\tau^{\tau+h} \Big| \xi^m(t, \nu_t^m, \nu_t^f)(x_{B_n}(t)) - \xi^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(x_{B_n}(t)) \Big|\, \bar{m}_{B_n}^m(t) dt \\
&\quad + \int_\tau^{\tau+h} \xi^m(t, \nu_t^m, \nu_t^f)(x_{B_n}(t)) \left| m_{B_n}^m(t) - \bar{m}_{B_n}^m(t) \right| dt \\
&\leq C_1 h \|\xi^m\|_{\mathbf{BC}} \int_\tau^{\tau+h} \bar{m}_{B_n}^m(t) dt + \|\xi^m\|_{\mathbf{BC}} \int_\tau^{\tau+h} h(\mathrm{Lip}_t(m_{B_n}) + \mathrm{Lip}_t(\bar{m}_{B_n})) dt \\
&\leq (C_1 C_3 + C_4)\, h^2 \|\xi^m\|_{\mathbf{BC}}.
\end{aligned}$$

∎

4. Convergence.

THEOREM 4.1. *Let* $\mathbf{u}_t = (\mu_t^m, \mu_t^f, \mu_t^c) : [0,T] \to \mathcal{U}$ *be a solution of* (3.1), *and let* $\mathbf{v_0} = (\nu_0^m, \nu_0^f, \nu_0^c) \in \mathcal{U}$ *be an approximation of* $\mathbf{u_0} = (\mu_0^m, \mu_0^f, \mu_0^c)$ *given by formulas* (3.6), (2.6), *and* (2.7) *with an error*

$$\varepsilon_0 = \mathbf{d}(\mathbf{u_0}, \mathbf{v_0}).$$

Let $\mathbf{v_t} = (\nu_t^m, \nu_t^f, \nu_t^c) : [0,T] \to \mathcal{U}$ *be the output of the numerical method defined in section* 3.1 *with the initial condition* $\mathbf{v_0}$. *Then, there exists a constant* C *such that*

$$\mathbf{d}(\mathbf{u_t}, \mathbf{v_t}) \le \varepsilon_0 + CT\Delta t, \ t \le T, \tag{4.1}$$

where $\Delta t = \max_{n=0,\dots,N_T-1} |t_{n+1} - t_n|$, $t_n < T$ *for* $n = 0, \dots, N_T$, *and* $\Delta t \le a_0$, *where* a_0 *satisfies* (1.3).

Proof. As mentioned in Remarks 3.7 and 3.8, the solution $\mathbf{v}_t$ of the numerical method defined in section 3.1 is a Lipschitz continues map, and the problem (3.1) generates a Lipschitz semiflow S in metric space $(\mathcal{U}, \mathbf{d})$. Thus, to estimate the difference between $\mathbf{v}_t$ and $\mathbf{u}_t = S(t;0)\mathbf{u}_0$, we are allowed to use Proposition 3.10 and to consider the problem locally in time.

Without loss of generality, we can assume that interval $(\tau, \tau+h]$ does not contain internalization point time; that is, there exists n such that $(\tau, \tau+h] \subset [t_n, t_{n+1})$. A remark on this assumption will be made just after the proof.

According to Proposition 3.10, we wish to estimate the distance

$$\mathbf{d}(\mathbf{v}_{\tau+h}, S(h;\tau)\mathbf{v}_\tau) = d_1(\nu_{\tau+h}^m, S^m(h;\tau)\nu_\tau^m) + d_1(\nu_{\tau+h}^f, S^f(h;\tau)\nu_\tau^f) + d_2(\nu_{\tau+h}^c, S^c(h;\tau)\nu_\tau^c), \tag{4.2}$$

showing that it is bounded by $C(t_{n+1} - t_n)\Delta t$, with C depending only on T, and the hypotheses on the functions involved in (3.1) given by Assumption 3.6. It is straightforward from this local estimate, using the properties of the d_N-semiflow in Definition 3.4 and the triangular inequality, to deduce the global estimate (4.1) by induction.

Let us recall that components of the EBT ODE output $\mathbf{v}_t$, $t \in [t_n, t_{n+1})$, are defined as linear combinations of $(J - B_n + 1)$ Dirac measures in the case of males/females, and $(J - B_n + 1)^2$ in the case of couples,

$$\nu_t^m := \sum_{i=B_n}^{J} m_i^m(t)\delta_{\{x_i^m(t)\}}, \ \nu_t^f := \sum_{j=B_n}^{J} m_j^f(t)\delta_{\{y_j^f(t)\}}, \ \nu_t^c := \sum_{i,j=B_n}^{J} m_{ij}^c(t)\delta_{\{x_{ij}^c(t), y_{ij}^c(t)\}}.$$

Let $\bar{\mathbf{v}}_t = (\bar\nu_t^m, \bar\nu_t^f, \bar\nu_t^c)$ be a solution to (3.1) on the interval $[\tau, \tau+h]$ with initial condition $\mathbf{v}_\tau$,

$$\bar\nu_t^m := S^m(t-\tau;\tau)\nu_\tau^m, \qquad \bar\nu_t^f := S^f(t-\tau;\tau)\nu_\tau^f, \qquad \text{and} \qquad \bar\nu_t^c := S^c(t-\tau;\tau)\nu_\tau^c,$$

The semiflow $S = (S^m, S^f, S^c)$ generates the solution

$$\bar\nu_t^m = f^m(t,\cdot)\lambda^1 + \sum_{i=B_n}^{J} \bar m_i^m(t)\delta_{\{\bar x_i^m(t)\}}, \quad \bar\nu_t^f = f^f(t,\cdot)\lambda^1 + \sum_{j=B_n}^{J} \bar m_j^f(t)\delta_{\{\bar y_j^f(t)\}},$$

$$\bar\nu_t^c = \sum_{i,j=B_n}^{J} \bar m_{ij}^c(t)\delta_{\{\bar x_{ij}^c(t), \bar y_{ij}^c(t)\}},$$

where λ^1 is a Lebesgue measure, and f^m and f^f are densities arising in the boundary cohorts such that $\mathrm{supp} f^k(t,\cdot) \subset (0, t-\tau),\ k = f, m$. Observe that $\bar{\nu}^c_t$ does not have an absolutely continuous measure $f^c(t,)\lambda^2$ because newborns do not form couples. This is guaranteed by marriage function (1.2) together with assumptions (1.3) and $\Delta t \leq a_0$.

Let us write $s = \tau + h$ and note that the first component of (4.2) can be initially estimated in the following way:

$$
\begin{aligned}
d_1(\bar{\nu}^m_s, \nu^m_s) &= d_1\left(f^m(s,\cdot)\lambda^1 + \sum_{i=B_n}^{J} \bar{m}^m_i(s)\delta_{\{\bar{x}^m_i(s)\}}, \sum_{i=B_n}^{J} m^m_i(s)\delta_{\{x^m_i(s)\}} \right) \\
&\leq d_1\left(f^m(s,\cdot)\lambda^1, \bar{p}^m_{B_n}(s)\delta_{\{\bar{x}^m_{B_n}(s)\}} \right) \\
&\quad + d_1\left(\bar{p}^m_{B_n}(s)\delta_{\{\bar{x}^m_{B_n}(s)\}} + \sum_{i=B_n}^{J} \bar{m}^m_i \delta_{\{\bar{x}^m_i(s)\}}, \sum_{i=B_n}^{J} m^m_i \delta_{\{x^m_i(s)\}} \right) \\
&\leq d_1\left(f^m(s,\cdot)\lambda^1, \bar{p}^m_{B_n}(s)\delta_{\{\bar{x}^m_{B_n}(s)\}} \right) \\
&\quad + d_1\left(\left(\bar{p}^m_{B_n}(s) + \bar{m}^m_{B_n}(s)\right)\delta_{\{\bar{x}^m_{B_n}(s)\}}, m^m_{B_n}(s)\delta_{\{x^m_{B_n}(s)\}} \right) \\
&\quad + \sum_{i=B_n+1}^{J} d_1\left(\bar{m}^m_i(s)\delta_{\{\bar{x}^m_i(s)\}}, m^m_i(s)\delta_{\{x^m_i(s)\}} \right),
\end{aligned}
\tag{4.3}
$$

where $\bar{p}^m_{B_n}(s)$ is a mass generated by the absolutely continuous measures $f^m(t,\cdot)\lambda^1$, i.e., $\bar{p}^m_{B_n}(s) = \int_{\mathbb{R}_+} f^m(s,x)dx$. The previous estimate uses the fact that the flat metric is actually a norm. The second component of (4.2) concerning the female population is treated analogously, while the third component will follow easily from the expression

$$
d_2(\bar{\nu}^c_s, \nu^c_s) = d_2\left(\sum_{i,j=B_n}^{J} \bar{m}^c_{ij}(s)\delta_{\{\bar{x}^c_{ij}(s),\bar{y}^c_{ij}(s)\}}, \sum_{i,j=B_n}^{J} m^c_{ij}(s)\delta_{\{x^c_{ij}(s),y^c_{ij}(s)\}} \right). \tag{4.4}
$$

To handle the estimates (4.3) and (4.4) it is necessary to find the locations $\bar{x}^m_i(s)$, $\bar{y}^f_j(s)$, $(\bar{x}^c_{ij}(s), \bar{y}^c_{ij}(s))$ and masses $\bar{m}^m_i(s)$, $\bar{m}^f_j(s)$, $\bar{m}^c_{ij}(s)$, $\bar{p}^m_{B_n}(s)$, $\bar{p}^f_{B_n}(s)$ for $i, j = B_n, \ldots, J$, which are generated by the semiflow S.

I. Locations. Let us start with male population and observe that locations $x^m_i(s)$ and $\bar{x}^m_i(s)$, $i = B_n, \ldots, J$, are equal in $[\tau, \tau+h]$, as they are governed by the same rule ($t \in [\tau, \tau+h]$):

$$
\begin{aligned}
\tfrac{d}{dt} x^m_i(t) &= 1, \\
\tfrac{d}{dt} \bar{x}^m_i(t) &= 1, \\
\bar{x}^m_i(t) &= x^m_i(t),\ i = B_n, \ldots, J.
\end{aligned}
$$

Using a similar argument for female and couple populations, we end up with the following equalities:

$$
\begin{aligned}
x^m_i(t) &= \bar{x}^m_i(t), \\
y^f_i(t) &= \bar{y}^f_i(t), \\
(x^c_{ij}(t), y^c_{ij}(t)) &= (\bar{x}^c_{ij}(t), \bar{y}^c_{ij}(t)) \text{ for } i, j = B_n, \ldots, J.
\end{aligned}
$$

II. Masses. The following formulas for masses at time $s = \tau + h$:

$$\begin{aligned}
\bar{m}_i^m(s) &= m_i^m(\tau) - \int_\tau^{\tau+h} \xi^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(x_i^m(t))\bar{m}_i^m(t)dt,\\
\bar{p}_{B_n}^m(s) &= \int_\tau^{\tau+h} \int_{\mathbb{R}_+^2} \beta^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(z)d\bar{\nu}_t^c(z)dt\\
&\quad - \int_\tau^{\tau+h} \int_{\mathbb{R}_+} \xi^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f) f^m(t,x)dxdt,\\
\bar{m}_j^f(s) &= m_j^f(\tau) - \int_\tau^{\tau+h} \xi^f(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(y_j^f(t))\bar{m}_j^f(t)dt,\\
\bar{p}_{B_n}^f(s) &= \int_\tau^{\tau+h} \int_{\mathbb{R}_+^2} \beta^f(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(z)d\bar{\nu}_t^c(z)dt\\
&\quad - \int_\tau^{\tau+h} \int_{\mathbb{R}_+} \xi^f(t, \bar{\nu}_t^m, \bar{\nu}_t^f) f^f(t,y)dydt,\\
\bar{m}_{ij}^c(s) &= m_{ij}^c(\tau) - \int_\tau^{\tau+h} \xi^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f, \bar{\nu}_t^c)(x_{ij}^c(t), y_{ij}^c(t))\bar{m}_{ij}^c(t)dt\\
&\quad + \int_\tau^{\tau+h} \frac{\tilde{N}_{ij}(t)}{\tilde{D}_{ij}(t)}dt,
\end{aligned}$$

will be derived using three different types of test functions: $\varphi_\varepsilon(t,x)$, $\varphi_\varepsilon^i(t,x)$, and $\varphi_\varepsilon^{ij}(t,x,y)$, $i,j = B_n, \ldots, J$. In all of these test functions, the parameter should be chosen in such a way that the support of each test function intersects with the domain of only one, particular location function $x_i^m(t)$, $(x_{ij}^c, y_{ij}^c)(t)$, $t \in [\tau, \tau+h)$, and $i,j = B_n, \ldots, J$. Such a choice is possible due to the regularity of cohort boundaries, which are straight parallel lines.

(a) $\bar{m}_i^m(s)$, $\bar{m}_j^f(s)$, $i,j = B_n, \ldots, J$.

To derive the evolution of the male population mass, which is generated by the semiflow in the ith internal cohorts, we use test functions $\varphi_\varepsilon^i \in (\mathbf{C^1} \cap \mathbf{W^{1,\infty}})([\tau, \tau+h] \times \mathbb{R}_+; \mathbb{R})$ as follows:

$$\varphi_\varepsilon^i(t,x) = \begin{cases} 1 \text{ if } x \in [x_i^m(\tau) - \varepsilon, x_i^m(\tau+h) + \varepsilon],\\ \frac{x - (x_i^m(\tau) - 2\varepsilon)}{\varepsilon} \text{ if } x \in [x_i^m(\tau) - 2\varepsilon, x_i^m(\tau) - \varepsilon],\\ \frac{-x + (x_i^m(\tau+h) + 2\varepsilon)}{\varepsilon} \text{ if } x \in [x_i^m(\tau+h) + \varepsilon, x_i^m(\tau+h) + 2\varepsilon],\\ 0 \text{ otherwise.} \end{cases}$$

According to Definition 3.5, if the measure $(\bar{\nu}_t^m, \bar{\nu}_t^f, \bar{\nu}_t^c)$ is a weak solution to (3.1) on time interval $[\tau, \tau+h]$ (not $[0,T]$, as we investigate the equation locally in time), then the following equality holds:

$$\int_{\tau}^{\tau+h}\int_{\mathbb{R}_+}\Big(\partial_t\varphi_\varepsilon^i(t,x)+\partial_x\varphi_\varepsilon^i(t,x)-\xi^m(t,\bar\nu_t^m,\bar\nu_t^f)\varphi_\varepsilon^i(t,x)\Big)\,d\bar\nu_t^m(x)dt$$

$$+\int_{\tau}^{\tau+h}\varphi_\varepsilon^i(t,0)\int_{\mathbb{R}_+^2}\beta^m(t,\bar\nu_t^m,\bar\nu_t^f)(z)d\bar\nu_t^c(z)dt$$

$$=\int_{\mathbb{R}_+}\varphi_\varepsilon^i(\tau+h,x)d\bar\nu_{\tau+h}^m(x)-\int_{\mathbb{R}_+}\varphi_\varepsilon^i(\tau,x)d\bar\nu_\tau^m(x).$$

The following integrals vanish: $\int_\tau^{\tau+h}\int_{\mathbb{R}_+}\partial_t\varphi_\varepsilon^i(t,x)d\bar\nu_t^m(x)dt$, because the test functions do not depend on t; $\int_\tau^{\tau+h}\int_{\mathbb{R}_+}\partial_x\varphi_\varepsilon^i(t,x)d\bar\nu_t^m(x)dt$, because of measure $\bar\nu_t^m(x)$; and $\int_\tau^{\tau+h}\varphi_\varepsilon^i(t,0)\int_{\mathbb{R}_+^2}\beta^m(t,\bar\nu_t^m,\bar\nu_t^f)(z)d\bar\nu_t^c(z)dt$, because of the support of the test functions. Passing with ε to 0 and using the dominated convergence theorem, we finally obtain the formula for the desired coefficient $\bar m_i(s)$, $i=B_n,\dots,J$:

$$\bar m_i^m(s)=m_i^m(\tau)-\int_\tau^{\tau+h}\xi^m(t,\bar\nu_t^m,\bar\nu_t^f)(x_i^m(t))\bar m_i^m(t)dt.$$

Formulas for female masses $m_j^f(s)$, $j=1,\dots,J$, are derived using the same type of test function.

(b) $\bar p_{B_n}^m(s)$, $\bar p_{B_n}^f(s)$.
In order to find $\bar p_{B_n}^m(s)=\int_{\mathbb{R}_+}f^m(s,x)dx$, we define the test function $\varphi_\varepsilon\in(\mathbf{C^1}\cap\mathbf{W^{1,\infty}})([\tau,\tau+h]\times\mathbb{R}_+;\mathbb{R})$ in the following way:

$$\varphi_\varepsilon(t,x)=\begin{cases}1 \text{ if } x\in[0,h+\varepsilon],\\ \frac{-x+(h+2\varepsilon)}{\varepsilon} \text{ if } x\in[h+\varepsilon,h+2\varepsilon],\\ 0 \text{ otherwise.}\end{cases}$$

As previously in case (a), we apply the above test function to Definition 3.5 and observe that $\partial_t\varphi_\varepsilon=0$ and that the supp $\partial_x\varphi_\varepsilon(t,\cdot)\cap$ supp $f^m(t,\cdot)=\emptyset$, so the first two integrals vanish. Passing with ε to 0, we obtain

$$-\int_\tau^{\tau+h}\int_{\mathbb{R}_+}\xi^m(t,\bar\nu_t^m,\bar\nu_t^f)f^m(t,x)dxdt+\int_\tau^{\tau+h}\int_{\mathbb{R}_+^2}\beta^m(t,\bar\nu_t^m,\bar\nu_t^f)(z)d\bar\nu_t^c(z)dt$$
$$=\int_{\mathbb{R}_+}f^m(\tau+h,x)dx$$

and conclude that

$$\bar p_{B_n}^m(\tau+h)=-\int_\tau^{\tau+h}\int_{\mathbb{R}_+}\xi^m(t,\bar\nu_t^m,\bar\nu_t^f)f^m(t,x)dxdt$$
$$+\int_\tau^{\tau+h}\int_{\mathbb{R}_+^2}\beta^m(t,\bar\nu_t^m,\bar\nu_t^f)(z)d\bar\nu_t^c(z)dt.$$

The same reasoning should be applied to the female case.

(c) $\bar{m}^c_{ij}(s)$, $i,j = B_n,\ldots,J$.

We will find the evolution of the masses for couples $\bar{m}^c_{ij}(c)$, $i,j = B_n,\ldots,J$, using test function $\varphi^{ij}_\varepsilon \in (\mathbf{C^1}\cap\mathbf{W^{1,\infty}})([\tau,\tau+h]\times\mathbb{R}_+\times\mathbb{R}_+;\mathbb{R})$ as follows:

$$\varphi^{ij}_\varepsilon(t,x,y) = \begin{cases} 1 & \text{if } x\in[a,b], & y\in[c,d],\\ \frac{x-(a-\varepsilon)}{\varepsilon} & \text{if } x\in[a-\varepsilon,a], & y\in[c,d],\\ \frac{-x+(b+\varepsilon)}{\varepsilon} & \text{if } x\in[b,b+\varepsilon], & y\in[c,d],\\ \frac{y-(c-\varepsilon)}{\varepsilon} & \text{if } x\in[a,b], & y\in[c-\varepsilon,c],\\ \frac{-y+(d+\varepsilon)}{\varepsilon} & \text{if } x\in[a,b], & y\in[d,d+\varepsilon],\\ \frac{x+y-(a+c-\varepsilon)}{\varepsilon} & \text{if } x\in[a-\varepsilon,a], & y\in[-x+a+c-\varepsilon,c],\\ \frac{x-y-a+\varepsilon+d}{\varepsilon} & \text{if } x\in[a-\varepsilon,a], & y\in[d,x-a+\varepsilon+d],\\ \frac{-x+y+b-c+\varepsilon}{\varepsilon} & \text{if } x\in[b,b+\varepsilon], & y\in[x-b+c-\varepsilon,c],\\ \frac{-x-y+b+d+\varepsilon}{\varepsilon} & \text{if } x\in[b,b+\varepsilon], & y\in[d,-x+b+d+\varepsilon],\\ 0 & \text{otherwise}, \end{cases}$$

where

$$a = x^c_{ij}(\tau)-\varepsilon, \qquad b = x^c_{ij}(\tau+h)+\varepsilon,$$
$$c = y^c_{ij}(\tau)-\varepsilon, \qquad d = y^c_{ij}(\tau+h)+\varepsilon.$$

Again, according to Definition 3.5, if the measure $(\bar{\nu}^m_t,\bar{\nu}^f_t,\bar{\nu}^c_t)$ is a weak solution to (3.1) on time interval $[\tau,\tau+h]$, then the following equality holds:

$$\begin{aligned}\int_\tau^{\tau+h}\int_{\mathbb{R}_+}\big(&\partial_t\varphi^{ij}_\varepsilon(t,z)+\partial_x\varphi^{ij}_\varepsilon(t,z)+\partial_y\varphi^{ij}_\varepsilon(t,z)\\ &-\xi^c(t,\bar{\nu}^m_t,\bar{\nu}^f_t,\bar{\nu}^c_t)\varphi^{ij}_\varepsilon(t,z)\big)d\bar{\nu}^c_t(z)dt\\ &+\int_\tau^{\tau+h}\int_{\mathbb{R}^2_+}\varphi^{ij}_\varepsilon(t,z)d\mathcal{T}(t,\bar{\nu}^m_t,\bar{\nu}^f_t,\bar{\nu}^c_t)(z)dt\\ =\int_{\mathbb{R}_+}&\varphi^{ij}_\varepsilon(\tau+h,z)d\bar{\nu}^c_{\tau+h}(z)-\int_{\mathbb{R}_+}\varphi^{ij}_\varepsilon(\tau,z)d\bar{\nu}^c_\tau(z),\end{aligned}$$

which, after following similar arguments for the male population, leads to

$$\begin{aligned}\bar{m}^c_{ij}(s) = m^c_{ij}(\tau) &- \int_\tau^{\tau+h}\xi^m(t,\bar{\nu}^m_t,\bar{\nu}^f_t,\bar{\nu}^c_t)(x^c_{ij}(t),y^c_{ij}(t))\bar{m}^c_{ij}(t)dt\\ &+\lim_{\varepsilon\to0}\int_\tau^{\tau+h}\int_{\mathbb{R}^2_+}\varphi^{ij}_\varepsilon(t,z)d\mathcal{T}(t,\bar{\nu}^m_t,\bar{\nu}^f_t,\bar{\nu}^c_t)(z)dt.\end{aligned}$$

We now observe that the measure $\mathcal{T}$, analogously to definitions (3.2), (3.3), and (3.4), is given by

$$\mathcal{T}(t,\bar{\nu}^m_t,\bar{\nu}^f_t,\bar{\nu}^c_t)) = \frac{\Theta(x,y)h(x)g(y)}{\gamma+\int_0^\infty h(z)d\bar{s}^m_t(z)+\int_0^\infty g(w)d\bar{s}^f_t(w)}(\bar{s}^m_t\otimes\bar{s}^f_t),$$

where

$$\bar{s}_t^m(B) = (\bar{\mu}_t^m - \bar{\sigma}_t^m)(B \times \mathbb{R}_+),\ \bar{s}_t^f(B) = (\bar{\mu}_t^f - \bar{\sigma}_t^f)(B \times \mathbb{R}_+)$$

and

$$\bar{\sigma}_t^m(B) = (\bar{\mu}_t^c)(B \times \mathbb{R}_+),\ \bar{\sigma}_t^f(B) = \bar{\mu}_t^c(\mathbb{R}_+ \times B).$$

The fact that for every fixed i, j, supp $\varphi_\varepsilon^{ij}(t,\cdot) \cap$ supp $f^m(t,\cdot) = \emptyset$ and supp $\varphi_\varepsilon^{ij}(t,\cdot) \cap$ supp $f^f(t,\cdot) = \emptyset$, yields

$$\lim_{\varepsilon\to 0} \int_\tau^{\tau+h} \int_{\mathbb{R}_+^2} \varphi_\varepsilon^{ij}(t,z) d\mathcal{T}(t, \bar{\nu}_t^m, \bar{\nu}_t^f, \bar{\nu}_t^c)(z) dt = \int_\tau^{\tau+h} \frac{\tilde{N}_{ij}(t)}{\tilde{D}_{ij}(t)} dt,$$

where

$$\frac{\tilde{N}_{ij}(t)}{\tilde{D}_{ij}(t)} = \frac{\Theta(x_{ij}^c(t), y_{ij}^c(t)) h(x_{ij}^c(t)) g(y_{ij}^c(t)) \left(\bar{m}_i^m(t) - \sum_{w=B_n}^J \bar{m}_{iw}^c(t)\right)\left(\bar{m}_j^f(t) - \sum_{v=B_n}^J \bar{m}_{vj}^c(t)\right)}{\gamma + \sum_{v=B_n}^J h(x_{vj}^c(t))\left(\bar{m}_v^m(t) - \sum_{w=B_n}^J \bar{m}_{vw}^c(t)\right) + \sum_{w=B_n}^J g(y_{iw}^c(t))\left(\bar{m}_w^f(t) - \sum_{v=B_n}^J \bar{m}_{vw}^c(t)\right)}.$$

Having desired formulas for locations and masses derived, we can proceed with inequalities (4.3) and (4.4). Constants C_1, C_2, and C_3 are defined exactly like in Lemma 3.12, while C_4 is defined similarly—but with respect to the underlying mass m. Using Lemmas 3.3 and 3.12, we show the following estimates for each term:

1. $d_1\left(f^m(s,\cdot)\lambda^1, \bar{p}_{B_n}^m(s)\delta_{\{\bar{x}_{B_n}^m(s)\}}\right) = \mathcal{O}(h)\Delta t.$

   Let us observe that

$$\begin{aligned}
d_1\left(f^m(s,\cdot)\lambda^1, \bar{p}_{B_n}^m(s)\delta_{\{\bar{x}_{B_n}^m(s)\}}\right) &\le \bar{p}_{B_n}^m(s)\bar{x}_{B_n}^m(s) \\
\le \Bigg| - \int_\tau^{\tau+h} \int_0^{t-\tau} &\xi^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f) f^m(t,x) dx dt \\
+ \int_\tau^{\tau+h} \int_{\mathbb{R}_+^2} &\beta^m(t, \bar{\nu}_t^m, \bar{\nu}_t^f)(z) d\bar{\nu}_t^c(z) dt \Bigg| \bar{x}_{B_n}^m(s) \\
\le (C_7 h^2 \|\xi^m\|_{\mathbf{BC}} &+ C_8 h \|\beta^m\|_{\mathbf{BC}})\Delta t,
\end{aligned}$$

   because $\bar{x}_{B_n}^m(t) = t - \tau \le \Delta t$ and because f^m is bounded due to Lemma 3.11 since ξ^m and β^m are absolutely continuous and $\bar{\nu}^c$ is finite on $[0,T]$. Here, we defined $C_7 = \max_{t\in[\tau,\tau+h],\ x\in[0,t-\tau]} f^m(t,x)$ and $C_8 = \int_{\mathbb{R}_+^2} d\bar{\nu}_t^c(z)$.

2. $d_1\left(\left(\bar{p}^m_{B_n}(s)+\bar{m}^m_{B_n}(s)\right)\delta_{\{\bar{x}^m_{B_n}(s)\}}, m^m_{B_n}(s)\delta_{\{x^m_{B_n}(s)\}}\right)=\mathcal{O}(h^2)$. Let us observe that

$$\begin{aligned}
&d_1\left(\left(\bar{p}^m_{B_n}(s)+\bar{m}^m_{B_n}(s)\right)\delta_{\{\bar{x}^m_{B_n}(s)\}}, m^m_{B_n}(s)\delta_{\{x^m_{B_n}(s)\}}\right)\\
&\quad\le |\bar{p}^m_{B_n}(s)+\bar{m}^m_{B_n}(s)-m^m_{B_n}(s)|\\
&\quad\le \left|\int_\tau^{\tau+h}\int_0^{t-\tau}\xi^m(t,\bar{\nu}^m_t,\bar{\nu}^f_t)f^m(t,x)dxdt\right|\\
&\qquad+\left|\int_\tau^{\tau+h}\xi^m(t,\nu^m_t,\nu^f_t)(x^m_{B_n}(t))m^m_{B_n}(t)dt\right.\\
&\qquad\left.-\int_\tau^{\tau+h}\xi^m(t,\bar{\nu}^m_t,\bar{\nu}^f_t)(x^m_{B_n}(t))\bar{m}^m_{B_n}(t)dt\right|\\
&\qquad+\left|\int_\tau^{\tau+h}\sum_{i,j=B_n}^{J}\beta^m(t,\bar{\nu}^m_t,\bar{\nu}^f_t)(x^c_{ij}(t),y^c_{ij}(t))\bar{m}^c_{ij}(t)dt\right.\\
&\qquad\left.-\int_\tau^{\tau+h}\sum_{i,j=B_n}^{J}\beta^m(t,\nu^m_t,\nu^f_t)(x^c_{ij}(t),y^c_{ij}(t))m^c_{ij}(t)dt\right|\\
&\quad\le C_5h^2\|\xi^m\|_{\mathbf{BC}}+\left(C_1C_3+C_4\right)h^2\|\xi^k\|_{\mathbf{BC}}+\left(C_1C_3+C_1\right)h^2\|\beta^k\|_{\mathbf{BC}},
\end{aligned}$$

where the definitions of the masses $\bar{p}^m_{B_n}$, $\bar{m}^m_{B_n}$, $m^m_{B_n}$, and $\bar{\nu}^c_t$ were used, together with estimates similar to those in the first item.

3. $d_1\left(\sum_{i=B_n+1}^J\bar{m}^m_i(s)\delta_{\{\bar{x}^m_i(s)\}},\sum_{i=B_n+1}^J m^m_i(s)\delta_{\{x^m_i(s)\}}\right)=\mathcal{O}(h^2)$.
Using the fact that the characteristics verify $x^m_i(s)=\bar{x}^m_i(s)$ and (3.10), we get

$$\begin{aligned}
&d_1\left(\sum_{i=B_n+1}^{J}\bar{m}^m_i(s)\delta_{\{\bar{x}^m_i(s)\}},\sum_{i=B_n+1}^{J}m^m_i(s)\delta_{\{x^m_i(s)\}}\right)\\
&\quad\le\sup_{\|\varphi\|_{W^{1,\infty}}\le 1}\left|\int_\tau^{\tau+h}\sum_{i=1}^{N}\varphi(x_i(t))[\xi^m(t,\nu^m_t,\nu^f_t)(x_i(t))m_i(t)\right.\\
&\qquad\left.-\xi^m(t,\bar{\nu}^m_t,\bar{\nu}^f_t)(x_i(t))\bar{m}_i(t)]dt\right|\\
&\quad\le C_5h^2\|\xi^m\|_{\mathbf{BC}}.
\end{aligned}$$

4. $d_2\left(\sum_{i,j=B_n}^J\bar{m}^c_{ij}(s)\delta_{\{\bar{x}^c_{ij}(s),\bar{y}^c_{ij}(s)\}},\sum_{i,j=B_n}^J m^c_{ij}(s)\delta_{\{x^c_{ij}(s),y^c_{ij}(s)\}}\right)=\mathcal{O}(h^2)$.
Using the fact that the characteristics of couples satisfy $(x^c_{ij}(s),y^c_{ij}(s))=(\bar{x}^c_{ij}(s),\bar{y}^c_{ij}(s))$ and the evolution of the masses $m^c_{ij}(s)$ and $\bar{m}^c_{ij}(s)$, we get

that

$$\begin{aligned}\sum_{i,j=B_n}^{J} \left|m_{ij}^c(s) - \bar{m}_{ij}^c(s)\right| \leq &\int_\tau^{\tau+h} \sum_{i,j=1}^{N} \left|\frac{\tilde{N}_{ij}(t)}{\tilde{D}_{ij}(t)} - \frac{N_{ij}(t)}{D_{ij}(t)}\right| dt \\ &+ \int_\tau^{\tau+h} \sum_{i,j=1}^{N} |\xi^c(t, \nu_t^m, \nu_t^f, \nu_t^c)(x_{ij}(t), y_{ij}(t)) m_{ij}(t) \\ &- \xi^c(t, \bar{\nu}_t^m, \bar{\nu}_t^f, \bar{\nu}_t^c)(x_{ij}(t), y_{ij}(t)) \bar{m}_{ij}(t)| dt \\ \leq &h^2 C_9 + h^2 \|\xi^c\|_{\mathbf{BC}} (2C_2C_3 + C_2),\end{aligned}$$

where $C_9 = (8(\|h\|_{\infty,\mathbf{Lip}} + \|g\|_{\infty,\mathbf{Lip}})\|\Theta\|_{\infty,\mathbf{Lip}})\, C_2$. The estimate on C_9 was derived due to the observation that

$$\sum_{i,j=1}^{N} \left|\frac{\tilde{N}_{ij}(t)}{\tilde{D}_{ij}(t)} - \frac{N_{ij}(t)}{D_{ij}(t)}\right| \leq d_2 \left(\mathcal{T}(\nu_t^f, \nu_t^m, \nu_t^c), \mathcal{T}(\bar{\nu}_t^f, \bar{\nu}_t^m, \bar{\nu}_t^c)\right),$$

taking into account (3.4). Moreover, using [44, Lemma 2.4], we obtain

$$\begin{aligned}&d_2 \left(\mathcal{T}(\nu_t^f, \nu_t^m, \nu_t^c), \mathcal{T}(\bar{\nu}_t^f, \bar{\nu}_t^m, \bar{\nu}_t^c)\right) \\ &\qquad \leq \|\mathcal{T}\|_{\mathbf{BC}^{0,1}} \left(d_1(\nu_t^f, \bar{\nu}_t^f) + d_1(\nu_t^m, \bar{\nu}_t^m) + d_2(\nu_t^c, \bar{\nu}_t^c)\right) \\ &\qquad \leq C_2 h \left(8(\|h\|_{\infty,\mathbf{Lip}} + \|g\|_{\infty,\mathbf{Lip}})\|\Theta\|_{\infty,\mathbf{Lip}}\right).\end{aligned}$$

Obviously distance $d_1(\bar{\nu}_s^f, \nu_s^f)$ can be estimated analogously to $d_1(\bar{\nu}_s^m, \nu_s^m)$. It just has been shown that

$$\mathbf{d}(\mathbf{v}_{\tau+h}, S(h;\tau)\mathbf{v}_\tau) = (\mathcal{O}(h^2) + \mathcal{O}(h))\Delta t + \mathcal{O}(h^2).$$

Applying this local estimate to Proposition 3.10 we obtain the claim of the underlying theorem, namely

$$\mathbf{d}(\mathbf{u_t}, \mathbf{v_t}) \leq \varepsilon_0 + L \int_{[0,t]} \liminf_{h\to 0} \frac{\mathbf{d}(\mathbf{v}_{\tau+h}, S(h;\tau)\mathbf{v}_\tau)}{h} d\tau \leq \varepsilon_0 + L \int_{[0,t]} C\Delta t\, d\tau \leq \varepsilon_0 + LCt\Delta t.$$

The constant C estimates a certain combination of constants $C_1, \ldots, C_7$. ∎

Remark 4.2. At the beginning of the proof, we assumed that interval $(\tau, \tau + h]$ does not contain the internalization point t_n; then we consider $(\tau, \tau + h] = (t_n, \tau + h]$, where $\tau + h < t_{n+1}$, and observe that $\bar{m}_{B_n}^m = \bar{m}_{B_n}^f = 0$, that $\bar{m}_{ij}^c(t) = m_{ij}^c(t) = 0$, $j = B_n \vee\ j = B_n$, and that the whole argumentation of the proof does not change.

5. Numerical examples. In this section we present two numerical examples, illustrating the theoretical results. In both cases we present tables of errors indicating the rate of the convergence. In the second example we calculate the errors not only in bounded Lipschitz distance, but also in TV. As it turns out, the convergence cannot be obtained in TV, which confirms the need for the presented theory. The measurement of the error in bounded Lipschits distance is truly necessary in those calculations but is far from trivial and requires additional explanation. For this reason, in subsection 5.1 we start with the details concerning the error measurement. Subsections 5.2 and 5.3 deal with the numerical examples. In subsection 5.2 we consider an equation with the simplest possible coefficients satisfying the theoretical assumptions, whose solution is not known. For this reason Table 1 presents the errors between the numerical solutions and a reference solution. It is also worth noting here that due to the choice of trivial/constant mortality rates ($c^m = c^f = 0.1$), the number of males and females within any cohort never reaches 0, which is a nonrealistic phenomenon, because usually it is assumed that every single individual is eventually dying. To present a more probable model, we introduce the second example in subsection 5.3, where the coefficients are somewhat complicated and time-dependent. They were chosen not only to satisfy the theoretical assumptions, but also in such a way that we know the exact solution of the system. More precisely we first imposed the equations describing the evolution of males, females, and couples and derived the coefficients which meet the assumptions and for which the imposed equations satisfy the system. Given the exact solution, we can observe that all of the individuals, as well as all of the couples in any cohort, will eventually become extinct. This is in accordance to life observations. Moreover, now that the exact solution of the system is known, the errors presented in Table 2 measure the distance between the numerical and the analytical solutions.

5.1. Measurement of the error. Due to the definition of bounded Lipschitz distance (which is a supremum over bounded Lipschitz functions), the calculation of an error in the flat metric is not straightforward. First, we need to consider approximation of initial data and its error. Second, we wish to reduce the problem of computational bounded Lipschitz distance between two atomic measures to the problem of computational 1-Wasserstein distance. Third, the computational cost of Wasserstein distance in a higher dimension (2D) is troublesome itself and deserves special attention.

Every numerical computation starts with establishing the time and space steps, here Δt, Δx, and Δy. Given that in the case of age-structured population models straight lines are characteristic, it is natural to assume that $\Delta t = \Delta x = \Delta y$ and to divide the domain of solution into $T/\Delta t$ time steps, $M/\Delta t$ space cells in the case of male and female populations, and $(M/\Delta t)^2$ in the case of couples population, with M being the largest age of the population. According to Remark 3.9, the initial condition can be approximated arbitrarily accurate by a linear combination of a certain amount of Dirac measures. Special techniques of measure reconstruction described in [14] allow us to present this desired approximation with a linear combination of only $M/\Delta x$ Dirac Deltas (or $(M/\Delta x)^2$ in the case of couples) with error Δx^2. It is easy to see that the same reasoning can be applied in the 2D case. Given an exact solution at time T in subsection 5.3, we can use the same methods. Analyzing the procedures of measure reconstruction proposed in [14], it is sufficient to approximate initial conditions with formulas (2.4) and (2.5), to attain a satisfactory approximation of $\varepsilon_0 = \mathcal{O}(\Delta t^2)$. Obviously, if we present a method of order one only, this inaccuracy can be neglected.

Lemma 2.1 in [14] shows how to reduce the problem of bounded Lipschitz distance (in 1D) to some other measure expressed in terms of 1-Wasserstein distance. The same reasoning can be easily adapted to the 2D case.

LEMMA 5.1. *Let $\mu_1, \mu_2 \in \mathcal{M}_+(\mathbb{R}_+)$ be such that $M_{\mu_i} = \int_{\mathbb{R}_+} d\mu_i \neq 0$, and $\tilde{\mu}_i = \mu_i/M_{\mu_i}$ for $i = 1, 2$. Define $\rho : \mathcal{M}_+(\mathbb{R}_+) \times \mathcal{M}_+(\mathbb{R}_+) \to \mathbb{R}_+$ as*

$$\rho(\mu_1, \mu_2) = \min\{M_{\mu_1}, M_{\mu_2}\} W_1(\tilde{\mu}_1, \tilde{\mu}_2) + |M_{\mu_1} - M_{\mu_2}|,$$

where W_1 is the 1-Wasserstein distance. Then, there exists a constant

$$C_K = \frac{1}{3}\min\left\{1, \frac{2}{|K|}\right\},$$

such that

$$C_K\rho(\mu_1,\mu_2) \le d_1(\mu_1,\mu_2) \le \rho(\mu_1,\mu_2),$$

where K *is the smallest interval such that* $\mathrm{supp}(\mu_1)$, $\mathrm{supp}(\mu_2) \subseteq K$ *is the length of the interval* K*. If* K *is unbounded, we set* $C_K = 0$*.*

In all presented numerical experiments, the effective error of the method, $\mathrm{Err}(\Delta t)$, will be estimated in terms of metric ρ. To effectively compute the Wasserstein distance $W_1(\tilde\mu_1,\tilde\mu_2)$ in any dimension, we resort to the results presented in [4, 42], where the considerations start from approximation of $\tilde\mu_1$, $\tilde\mu_2$ by some atomic measures $\sum_i^{N_a} a_i\delta_{\{x_i^a\}}$ and $\sum_j^{N_b} b_j\delta_{\{x_j^b\}}$, respectively (for the sake of simplicity we assume that $N_a = N_b$). Instead of computing

$$W_1\left(\sum_i a_i\delta_{\{x_i^a\}},\sum_j b_j\delta_{\{x_j^b\}}\right):=\min\left\{\sum_{i,j}(c_{ij}\gamma_{ij}):\quad \gamma_{ij}\ge 0,\ \sum_i\gamma_{ij}=b_j,\ \sum_j\gamma_{ij}=a_i\right\} \tag{5.1}$$

we fix small $\varepsilon > 0$ and focus on

$$W_1^\varepsilon := \min\left\{\sum_{i,j}\left(c_{ij}\gamma_{ij}+\varepsilon\gamma_{ij}\log(\gamma_{ij})\right):\quad \gamma_{ij}\ge 0,\ \sum_i\gamma_{ij}=b_j,\ \sum_j\gamma_{ij}=a_i\right\},$$

which for $\varepsilon \to 0$ tends to the minimization problem (5.1) in the sense of Γ-convergence. Taking $\eta_{ij} := \mathrm{e}^{-c_{ij}/\varepsilon}$, we observe that

$$c_{ij}\gamma_{ij}+\varepsilon\gamma_{ij}\log(\gamma_{ij}) = \varepsilon\gamma_{ij}\log\left(\frac{\gamma_{ij}}{\eta_{ij}}\right) = \varepsilon\mathrm{KL}(\gamma|\eta),$$

where KL is the Kullback–Leiber divergence, that is, a sort of a distance based on a relative entropy:

$$\mathrm{KL}(\gamma|\eta) := \begin{cases} \sum_{ij}\gamma_{ij}\log\left(\frac{\gamma_{ij}}{\eta_{ij}}\right) & \text{if } \frac{\gamma_{ij}}{\eta_{ij}} > 0,\\ 0 & \text{if } \frac{\gamma_{ij}}{\eta_{ij}} = 0,\\ +\infty, & \text{if } \frac{\gamma_{ij}}{\eta_{ij}} < 0. \end{cases}$$

Given a convex set $\mathcal{C} \in R^{N_a\times N_a}$, the projection according to the KL divergence is defined as

$$P_{\mathcal{C}}^{\mathrm{KL}}(\eta) := \mathrm{argmin}_{\gamma\in\mathcal{C}}\mathrm{KL}(\gamma|\eta).$$

This means that $W_1^\varepsilon = \mathrm{KL}(P_{\mathcal{C}}^{\mathrm{KL}}(\eta)|\eta)$, where $P_{\mathcal{C}}^{\mathrm{KL}}(\eta)$ can be computed using *iterative Bergman projections*,

$$\gamma^{(0)} := \mathrm{e}^{-C/\varepsilon},\ \gamma^{(n)} := P_{\mathcal{C}}^{\mathrm{KL}}(\gamma^{(n-1)}),$$

with the entries of C defined as $c_{ij} = \|x_i^a - x_j^b\|$. It can be shown that

$$\gamma^{(n)} \to P_{\mathcal{C}}^{\mathrm{KL}}(\eta) \quad \text{as } n\to\infty.$$

For more details, see to [42]. The rate of convergence q presented in the tables of errors is given by

$$q := \lim_{\Delta t\to 0}\frac{\log[\mathrm{Err}(2\Delta t)/\mathrm{Err}(\Delta t)]}{\log 2}.$$

5.2. Example 1. In this first example we approximate system (1.1) for $t \in [0,1)$ and $(x,y) \in [0,1) \times [0,1)$, where mortality and birth rates are constant,

$$c^m(t,x) = c^f(t,y) = c^c(t,x,y) = 0.1,$$

$$b^m(t,x,y) = b^f(t,x,y) = 10,$$

and marriage function coefficients h, g, and Θ do not depend on time:

$$h(t,x) = \begin{cases} \left(\frac{1}{10} - x\right)(x-1), & \frac{1}{10} \le x \le 1, \\ 0 & \text{otherwise}, \end{cases} \qquad g(t,y) = \begin{cases} \left(\frac{1}{10} - y\right)(y-1), & \frac{1}{10} \le y \le 1, \\ 0 & \text{otherwise}, \end{cases}$$

$$\Theta(t,x,y) = \begin{cases} 10\left(\frac{1}{10} - x\right)(x-1)\left(\frac{1}{10} - y\right)(y-1), & \frac{1}{10} \le x \le 1, \frac{1}{10} \le y \le 1, \\ 0 & \text{otherwise}. \end{cases}$$

The initial conditions for the system are also trivial and given by

$$u^m(0,x) = u^f(0,y) = u^c(0,x,y) = 1.$$

Table 1 shows that the scheme behaves as an scheme of order 1 as expected from our Theorem 4.1.

Table 1

Error computed in the flat metric $\mathrm{Err}(\Delta t)$ *and its order of convergence* q.

$\Delta t = \Delta x$	$\mathrm{Err}(\Delta t)$	q
10^{-1}	$5.89 \cdot 10^{-2}$	–
$5 * 10^{-2}$	$2.57 \cdot 10^{-2}$	1.196499275
$2.5 * 10^{-2}$	$1.2 \cdot 10^{-2}$	1.09873395
$1.25 * 10^{-2}$	$5.9 \cdot 10^{-3}$	1.02424755
$6.25 * 10^{-3}$	$2.93 \cdot 10^{-3}$	1.00981429
$3.125 * 10^{-3}$	$1.46 \cdot 10^{-3}$	1.0049323
$1.5625 * 10^{-3}$	$7.39 \cdot 10^{-4}$	0.9823221
$7.8125 * 10^{-4}$	$3.7 \cdot 10^{-4}$	0.99804909

5.3. Example 2. We now present a numerical example for the system (1.1), whose exact solution is known, evolves in $[0,\infty)^3$, and is given by the following formulas:

$$u^m(t,x) = \begin{cases} \left(1 - \frac{t}{10}\right)(t-x-1)(-t+x-1), & 0 \le x \le t+1, \\ 0 & \text{otherwise}, \end{cases}$$

$$u^f(t,y) = \begin{cases} \left(1 - \frac{t}{10}\right)(t-y-1)(-t+y-1), & 0 \le y \le t+1, \\ 0 & \text{otherwise}, \end{cases}$$

$$u^c(t,x,y) = \begin{cases} \left(1 - \frac{t}{10}\right)\left(x - \frac{1}{10}\right)^2, \left(y - \frac{1}{10}\right)^2(-t+x-1)^2(-t+y-1)^2, & \frac{1}{10} \le x \le t+1 \wedge \frac{1}{10} \le y \le t+1, \\ 0 & \text{otherwise}. \end{cases}$$

The advantage of having the exact solution goes along with the disadvantage of lengthy and complicated coefficients given by

$$c^m(t,x) = \begin{cases} \frac{1}{10-t}, & 0 \le x \le t+1, \\ 0 & \text{otherwise}, \end{cases} \qquad c^f(t,y) = \begin{cases} \frac{1}{10-t}, & 0 \le y \le t+1, \\ 0 & \text{otherwise}, \end{cases} \qquad c^c(t,x,y) = 0.1,$$

$$b^m(t,x,y)=b^f(t,x,y)=\begin{cases} -\frac{9000000000000(t-1)(t+1)}{(10t+9)^{10}}, & 10x\geq 1\wedge t+1\geq x\wedge 10y\geq 1\wedge t+1\geq y,\\ 0 & \text{otherwise,}\end{cases}$$

$$h(t,x)=\begin{cases} \left(\frac{1}{10}-x\right)(-t+x-1) & \frac{1}{10}\leq x\leq t+1,\\ 0 & \text{otherwise,}\end{cases}\qquad g(t,y)=\begin{cases} \left(\frac{1}{10}-y\right)(-t+y-1) & \frac{1}{10}\leq y\leq t+1,\\ 0 & \text{otherwise,}\end{cases}$$

and

$$\Theta(t,x,y)=\frac{\mathrm{num}(t,x,y)}{\mathrm{den(t)}},$$

where

$$\mathrm{num}(t,x,y)=\frac{-\frac{t(10x(10y+199)+1990y-399)}{10000}+2x+2y-\frac{2}{5}}{\left(\frac{(t-10)(10t+9)^5(1-10x)^2(-t+x-1)}{3000000000}+\left(1-\frac{t}{10}\right)(t-x-1)\right)\left(\frac{(t-10)(10t+9)^5(1-10y)^2(-t+y-1)}{3000000000}+\left(1-\frac{t}{10}\right)(t-y-1)\right)}$$

and

$$\mathrm{den}(t)=\frac{(t-10)(10t+9)^4(10^8t^8+72\cdot 10^7t^7+2268\cdot 10^6t^6+40824\cdot 10^5t^5}{21\cdot 10^{15}}$$
$$\frac{+\,45927\cdot 10^5t^4+3306744\cdot 10^3t^3+1488034800t^2+21382637520t-51056953279)}{21\cdot 10^{15}}+1\,.$$

Given the above coefficients h, g, and Θ one can check that the marriage function is given with

$$\begin{aligned}T(t,x,y)=&\frac{1}{(-t+x-1)(-t+y-1)}\frac{(t-10)(10t+9)^5(1-10x)^2(t-x+1)^2}{3000000000}\\&+\left(1-\frac{t}{10}\right)(t-x-1)(-t+x-1)\frac{(t-10)(10t+9)^5(1-10y)^2(t-y+1)^2}{3000000000}\\&+\left(1-\frac{t}{10}\right)(t-y-1)(-t+y-1)\,.\end{aligned}$$

We first measured the error in TV and later in the flat metric. According to our expectations, TV does not show any convergence (see Table 2), while the flat metric significantly decreases the error. The rate of the error in Table 2 is one, as expected from the theoretical result in Theorem 4.1.

TABLE 2
Error computed in TV, $\mathrm{Err}_{\mathrm{TV}}(\Delta t)$*, error computed in the flat metric,* $\mathrm{Err}(\Delta t)$*, and order of convergence obtained in the flat metric* q*.*

$\mathrm{Err}_{\mathrm{TV}}(\Delta t)$	$\Delta t=\Delta x$	$\mathrm{Err}(\Delta t)$	q
$8.12*10^{-2}$	10^{-1}	$7.16\cdot 10^{-2}$	–
$7.04*10^{-2}$	$5*10^{-2}$	$4.0\cdot 10^{-2}$	0.839959587
$6.89*10^{-2}$	$2.5*10^{-2}$	$2.13\cdot 10^{-2}$	0,90914657
$6.72*10^{-2}$	$1.25*10^{-2}$	$1.14\cdot 10^{-2}$	0,901819606
$6.72*10^{-2}$	$6.25*10^{-3}$	$6.3\cdot 10^{-3}$	0,855610091
$6.71*10^{-2}$	$3.125*10^{-3}$	$3.2\cdot 10^{-3}$	0,977279923
$6.70*10^{-2}$	$1.5625*10^{-3}$	$1.6\cdot 10^{-3}$	1
$6.70*10^{-2}$	$7.8125*10^{-4}$	$8.0\cdot 10^{-4}$	1

REFERENCES

[1] L. M. Abia, O. Angulo, J. C. López-Marcos, and M. A. López-Marcos, *Numerical schemes for a size-structured cell population model with equal fission*, Math. Comput. Model., 50 (2009), pp. 653–664, https://doi.org/10.1016/j.mcm.2009.05.023.

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures Math. ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[3] O. Angulo and J. C. López-Marcos, *Numerical schemes for size-structured population equations*, Math. Biosci., 157 (1999), pp. 169–188, https://doi.org/10.1016/S0025-5564(98)10081-0.

[4] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, *Iterative Bregman projections for regularized transportation problems*, SIAM J. Sci. Comput., 37 (2015), pp. A1111–A1138, https://doi.org/10.1137/141000439.

[5] C. K. Birdsal and A. B. Langdon, *Plasma Physics Via Computer Simulation*, McGraw-Hill, New York, 1985.

[6] C. K. Birdsal and A. B. Langdon, *Plasma Physics Via Computer Simulation*, Ser. Plasma Phys., Taylor & Francis, New York, 2004.

[7] V. I. Bogachev, *Measure Theory*. Vol. I, II, Springer-Verlag, Berlin, 2007, https://doi.org/10.1007/978-3-540-34514-5.

[8] Å. Brännström, L. Carlsson, and D. Simpson, *On the convergence of the Escalator Boxcar Train*, SIAM J. Numer. Anal., 51 (2013), pp. 3213–3231, https://doi.org/10.1137/120893215.

[9] A. Bressan, *Hyperbolic Systems of Conservation Laws. The One-Dimensional Cauchy Problem*, Oxford Lecture Ser. Math. Appl. 20, Oxford University Press, Oxford, 2000,

[10] J. A. Cañizo, J. A. Carrillo, and S. Cuadrado, *Measure solutions for some models in population dynamics*, Acta Appl. Math., 123 (2013), pp. 141–156, https://doi.org/10.1007/s10440-012-9758-3.

[11] M. Campos Pinto, *Towards smooth particle methods without smoothing*, J. Sci. Comput., 65 (2015), pp. 54–82, https://doi.org/10.1007/s10915-014-9953-7.

[12] M. Campos Pinto, J. A. Carrillo, F. Charles, and Y.-P. Choi, *Convergence of a linearly transformed particle method for aggregation equations*, Numer. Math., 139 (2018), pp. 743–793, https://doi.org/10.1007/s00211-018-0958-2.

[13] 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, https://doi.org/10.1016/j.jde.2011.11.003.

[14] 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), pp. 2171–2197, https://doi.org/10.1142/S0218202514500183.

[15] A. Cohen and B. Perthame, *Optimal approximations of transport equations by particle and pseudoparticle methods*, SIAM J. Math. Anal., 32 (2000), pp. 616–636, https://doi.org/10.1137/S0036141099350353.

[16] G.-H. Cotted and P.-A. Raviart, *Particle methods for the one-dimensional Vlasov–Poisson equations*, SIAM J. Numer. Anal., 21 (1984), pp. 52–76, https://doi.org/10.1137/0721003.

[17] K. Craig and A. L. Bertozzi, *A blob method for the aggregation equation*, Math. Comp., 85 (2016), pp. 1681–1717, https://doi.org/10.1090/mcom3033.

[18] A. M. de Roos, *Numerical methods for structured population models: The Escalator Boxcar Train*, Numer. Methods Partial Differential Equations, 4 (1988), pp. 173–195, https://doi.org/10.1002/num.1690040303.

[19] J. H. M. Evers, S. C. Hille, and A. Muntean, *Mild solutions to a measure-valued mass evolution problem with flux boundary conditions*, J. Differential Equations, 259 (2015), pp. 1068–1097, https://doi.org/10.1016/j.jde.2015.02.037.

[20] A. G. Fredrickson, *A mathematical theory of age structure in sexual populations: Random mating and monogamous marriage models*, Math. Biosci., 10 (1971), pp. 117–143.

[21] K. Ganguly and H. D. Victory, Jr., *On the convergence of particle methods for multidimensional Vlasov–Poisson systems*, SIAM J. Numer. Anal., 26 (1989), pp. 249–288, https://doi.org/10.1137/0726015.

[22] J. Goodman, T. Y. Hou, and J. Lowengrub, *Convergence of the point vortex method for the 2-D Euler equations*, Comm. Pure Appl. Math., 43 (1990), pp. 415–430.

[23] 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 Differential Equations, 30 (2014), pp. 1797–1820, https://doi.org/10.1002/num.21879.

[24] P. Gwiazda, K. Kropielnicka, and A. Marciniak-Czochra, *The Escalator Boxcar Train method for a system of age-structured equations*, Netw. Heterog. Media, 11 (2016), pp. 123–143, https://doi.org/10.3934/nhm.2016.11.123.

[25] 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. Differential Equations, 248 (2010), pp. 2703–2735, https://doi.org/10.1016/j.jde.2010.02.010.

[26] P. Gwiazda and A. Marciniak-Czochra, *Structured population equations in metric spaces*, J. Hyperbolic Differ. Equ., 7 (2010), pp. 733–773, https://doi.org/10.1142/S021989161000227X.

[27] P. Gwiazda, A. Marciniak-Czochra, and H. R. Thieme, *Correction to: Measures under the flat norm as ordered normed vector space*, Positivity, 22 (2018), pp. 139–140, https://doi.org/10.1007/s11117-017-0535-4.

[28] P. Gwiazda, P. Orlinski, and A. Ulikowska, *Finite range method of approximation for balance laws in measure spaces*, Kinet. Relat. Models, 10 (2017), pp. 669–688, https://doi.org/10.3934/krm.2017027.

[29] K. P. Hadeler, *Pair formation in age-structured populations*, Acta Appl. Math., 14 (1989), pp. 91–102, https://doi.org/10.1007/BF00046676.

[30] K. P. Hadeler, R. Waldstätter, and A. Wörz-Busekros, *Models for pair formation in bisexual populations*, J. Math. Biol., 26 (1988), pp. 635–649, https://doi.org/10.1007/BF00276145.

[31] 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 Equations Operator Theory, 63 (2009), pp. 351–371, https://doi.org/10.1007/s00020-008-1652-z.

[32] F. Hoppenstaedt, *Mathematical Theories of Populations: Demographics, Genetics and Epidemics*, CBMS-NSF Reg. Conf. Ser. Appl. Math. 20, SIAM, Philadelphia, 1975, https://doi.org/10.1137/1.9781611970487.

[33] H. Inaba, *An Age-Structured Two-Sex Model for Human Population Reproduction by First Marriage*, Working Paper Ser. 15, Institute of Population Problems, Ministry of Health and Welfare, Tokyo, 1993.

[34] J. Jablonski and A. Marciniak-Czochra, *Efficient Algorithms Computing Distances between Radon Measures on R*, preprint, https://arxiv.org/abs/1304.3501, 2013.

[35] A. Lasota, J. Myjak, and T. Szarek, *Markov operators with a unique invariant measure*, J. Math. Anal. Appl., 276 (2002), pp. 343–356, https://doi.org/10.1016/S0022-247X(02)00457-2.

[36] S. Maniglia, *Probabilistic representation and uniqueness results for measure-valued solutions of transport equations*, J. Math. Pures Appl. (9), 87 (2007), pp. 601–626, https://doi.org/10.1016/j.matpur.2007.04.001.

[37] M. Martcheva and F. A. Milner, *A two-sex age-structured population model: Well posedness*, Math. Population Stud., 7 (1999), pp. 111–129, https://doi.org/10.1080/08898489909525450.

[38] A. McKendrick, *Applications of mathematics to medical problems*, Proc. Edinburgh Math. Soc., 44 (1925), pp. 98–130, https://doi.org/10.1017/S0013091500034428.

[39] H. Neunzert, *An introduction to the nonlinear Boltzmann-Vlasov equation*, in Kinetic Theories and the Boltzmann Equation (Montecatini, 1981), Lecture Notes in Math. 1048, Springer, Berlin, 1984, pp. 60–110, https://doi.org/10.1007/BFb0071878.

[40] M. C. Pinto, E. Sonnendrücker, A. Friedman, D. P. Grote, and S. M. Lund, *Noiseless Vlasov–Poisson simulations with linearly transformed particles*, J. Comput. Phys., 275 (2014), pp. 236–256, https://doi.org/10.1016/j.jcp.2014.06.032.

[41] J. Prüss and W. Schappacher, *Persistent age-distributions for a pair-formation model*, J. Math. Biol., 33 (1994), pp. 17–33, https://doi.org/10.1007/BF00160172.

[42] F. Santambrogio, *Optimal Transport for Applied Mathematicians. Calculus of Variations, PDEs, and Modeling*, Progr. Nonlinear Differential Equations Appl. 87, Birkhäuser, Cham, 2015, https://doi.org/10.1007/978-3-319-20828-2,

[43] H. Spohn, *Large Scale Dynamics of Interacting Particles*, Theoret. Math. Phys., Springer-Verlag, Berlin, 1991.

[44] A. Ulikowska, *An age-structured, two-sex model in the space of radon measures: Well posedness*, Kinet. Relat. Models, 5 (2012), pp. 873–900, https://doi.org/10.3934/krm.2012.5.873.

[45] N. Weaver, *Lipschitz Algebras*, World Scientific Publishing Co., Inc., River Edge, NJ, 1999, https://doi.org/10.1142/4100.

[46] L. Zhang, K. H. Andersen, U. Dieckmann, and Å. Brännström, *Four types of interference competition and their impacts on the ecology and evolution of size-structured populations and communities*, J. Theoret. Biol., 380 (2015), pp. 280–290, https://doi.org/10.1016/j.jtbi.2015.05.023.