A multi-step Lagrangian scheme for spatially inhomogeneous evolutionary games

J. Evol. Equ.

https://doi.org/10.1007/s00028-021-00702-5

Journal of Evolution Equations

A multi-step Lagrangian scheme for spatially inhomogeneous evolutionary games

Stefano Almi, Marco Morandotti and Francesco Solombrino

Abstract. A multi-step Lagrangian scheme at discrete times is proposed for the approximation of a nonlinear continuity equation arising as a mean-field limit of spatially inhomogeneous evolutionary games, describing the evolution of a system of spatially distributed agents with strategies, or labels, whose payoff depends also on the current position of the agents. The scheme is Lagrangian, as it traces the evolution of position and labels along characteristics, and is a multi-step scheme, as it develops on the following two stages: First, the distribution of strategies or labels is updated according to a best performance criterion, and then, this is used by the agents to evolve their position. A general convergence result is provided in the space of probability measures. In the special cases of replicator-type systems and reversible Markov chains, variants of the scheme, where the explicit step in the evolution of the labels is replaced by an implicit one, are also considered and convergence results are provided.

1. Introduction

The capability of changing strategy as an adaptive response to the modification of the surrounding environment in order to maximize a certain payoff is of paramount importance in decision-making processes. Replicator-type models [21] are a particular class of dynamical models that feature this adaptivity and are well suited for studying the evolution of strategies according to their success: Given a pool of strategies, the occurrence of each of them evolves according to their performance with respect to all the others; in this way, if a strategy gives a payoff which is higher compared to the average of all strategies, it is enhanced; otherwise, it is suppressed. This criterion, in the basic replicator model, is the only one that determines the evolution of the occurrence of the strategies, which in fact is independent from all other factors, in particular from the position of the agents that play those strategies. This is a reasonable assumption, not even a restrictive one, in many cases. For example, in a financial scenario, the set of (pure) strategies U contains the financial products available to an investor.

Mathematics Subject Classification: 35Q91, 60J75, 37C10, 47J35, 58D25

Keywords: Multi-step Lagrangian scheme, Inhomogeneous evolutionary games, Replicator equation, Reversible Markov chains, Minimizing movements scheme, Implicit–explicit scheme.

The work of MM was partially supported by the *Starting grant per giovani ricercatori* of Politecnico di Torino (Grant no. 53_RIL18MORMAR). The work of FS was supported by the project *Variational methods for stationary and evolution problems with singularities and interfaces* PRIN 2017 financed by the Italian Ministry of Education, University, and Research.

Published online: 24 April 2021

Any combination of them, that is a portfolio, is called a mixed strategy: In a discrete setting such as this one, it corresponds to the fraction of the capital invested in each of the different financial products. Adapting the strategy means to allocate resources differently according to the evolution of the market, and the location the investor is at when making this decision is likely to not affect the reward of the portfolio. On the contrary, when the position influences the outcome, the system is more involved, as more feedback is available, and the adaptive optimization process relies on the mutual influence of position and strategy performance. We call such a system *spatially inhomogeneous* and make them the focus of this paper.

1.1. Overview of the problem and state of the art

The basic, spatially homogeneous, replicator equation of [21] can be enriched to include spatial dependence of the payoff function: The idea is that the same strategy adopted in two different places might originate different rewards, precisely depending on the environment. Therefore, in order to maximize the payoff players can not only adapt their strategies, but also change their position seeking for the highest possible payoff. Spatially inhomogeneous evolutionary games, introduced in [5], provide a general mathematical framework for the evolution of a distribution of players with their (distributions of) strategies: A space-dependent replicator equation governs the evolution of the distribution $\lambda \in \mathcal{P}(U)$ of the strategies $u \in U$ while the evolution of the spatial variable $x \in \mathbb{R}^d$ is determined by λ.

In the subsequent contribution [31], this approach has been suitably extended as an abstract toolbox which is capable of rigorously describing the mean-field limit of a larger class of models which share the following features:

- a multi-agent dynamics in which every agent is characterized by a label $u \in U$ (accounting for different strategies or different populations to which each individual belongs);
- exchange rates among the labels which are stochastic in nature and, therefore, are described by the evolution of a probability measure $\lambda \in \mathcal{P}(U)$.

Several other models, besides the replicator dynamics mentioned above, are included in this class. The multi-label setting can be effectively used to describe situations in which the action of every individual is weighted differently according to the species it belongs to [3,4,16,17,19]. In the theory of mean-field games or in optimal control theory, labelling is used to distinguish informed agents in the evacuations of unknown environments, to highlight the influence of key investors in the stock market or of strong leaders in opinion formation [11,13,18,39]. The addition of source and sink terms in the spirit of [34] and of label switching [38] can be successfully dealt with in this class of models. Relevant applications where label switching may occur come, for instance, from chemical reaction networks, where a particle may change its type as a result of the interaction with the others [27,32,33]; also in social dynamics, loss or gain of opinion leadership over time is a natural postulate, as it happens in [18, Section 3.b].

The framework proposed in [31] couples a nonlinear transport dynamics for the positions $x \in \mathbb{R}^d$ of the agents with a Markov-type jump process for the labels $\lambda \in \mathcal{P}(U)$ (see Sect. 2). The mean-field limit of the model was proved to be a nonlinear continuity equation of the form

$$\partial_t \Psi_t + \operatorname{div}(b_{\Psi_t} \Psi_t) = 0 \tag{1}$$

in the space of probability measures over the pairs $(x, \lambda) \in \mathbb{R}^d \times \mathcal{P}(U)$ driven by a velocity field $b_\Psi(x, \lambda)$ depending on the global state of the system $\Psi \in \mathcal{P}(\mathbb{R}^d \times \mathcal{P}(U))$. These equations are part of a general class which is of great interest in the mathematical community [6, Chapter 8] and can be studied both with a Lagrangian or a Eulerian approach. On the one hand, the nonlinear continuity equation expresses the Eulerian point of view tracing the evolution of the global state Ψ. On the other hand, a notion of solution can also be provided by the Lagrangian point of view tracing the characteristics, which are, in our case, solutions to an ODE in a suitably constructed Banach space.

Given an initial datum $\widehat{\Psi}$, a solution $t \mapsto \Psi_t$ of the initial value problem for the nonlinear continuity equation is called a Eulerian solution, whereas a curve $t \mapsto \Psi_t$ obtained via the push-forward of $\widehat{\Psi}$ through the flow map associated with the ODE

$$(\dot{x}, \dot{\lambda}) = b_{\Psi_t}(x, \lambda) \tag{2}$$

is called a Lagrangian solution. Since Lagrangian solutions are also Eulerian solutions, the equivalence of the two notions follows if one is able to prove that Eulerian solutions are also Lagrangian. For the model studied in [31], and also for other relevant ones [14], these two notions of solution are equivalent. This has been achieved by means of the *superposition principle* (see [36], and also [6, Theorem 8.2.1], [8, Theorem 7.1], and [5, Theorem 5.2]), which provides the uniqueness of Eulerian solutions [5, Theorem 5.3]. Furthermore, the Lagrangian formulation has been used to propose discretization schemes to solve the nonlinear PDE numerically [15,25,26,29,35].

Moreover, the Lagrangian point of view has been used in [5] to provide a heuristic derivation of the nonlinear continuity equation arising as the mean-field limit of the spatially inhomogeneous replicator dynamics. Let us briefly discuss this derivation. Denoting by $h = T/N$ the time step, if an agent at time $t = ih$, for $i \in \{0, \dots, N-1\}$, is in the position x with mixed strategy λ, first they optimize the strategy distribution following a homogeneous replicator dynamics of the form

$$\lambda' := \lambda + h\mathcal{T}_{\Psi_t}(x, \lambda)\,. \tag{3}$$

Here, $\mathcal{T}_{\Psi_t}(x, \lambda)$ is the payoff operator determining the enhancement or suppression of the strategies; it depends on the random state (x, λ) and also on the current distribution Ψ_t. In the setting of [5], the operator $\mathcal{T}$ is quadratic in λ. After updating the strategy portfolio, the agent updates its position x to

$$x' := x + hv(x, u), \tag{4}$$

choosing u with probability λ'. The above two equations completely determine the conditional probability of having an agent in a state (x', λ') at time $t + h$ given the distribution Ψ_t. Equivalently, the new distribution Ψ_{t+h} can be defined via duality by

$$\int_{\mathbb{R}^d \times \mathcal{P}(U)} \phi(x', \lambda') \, \mathrm{d}\Psi_{t+h}(x', \lambda') = \int_{\mathbb{R}^d \times \mathcal{P}(U)} \left(\int_U \phi(x + hv(x, u), \lambda + h\mathcal{T}_{\Psi_t}(x, \lambda)) \, \mathrm{d}\lambda'(u) \right) \mathrm{d}\Psi_t(x, \lambda)$$

where $\phi \colon \mathbb{R}^d \times \mathcal{P}(U) \to \mathbb{R}$ is of class C^1. By a formal first-order Taylor expansion, we have

$$\int_{\mathbb{R}^d \times \mathcal{P}(U)} \phi(x', \lambda') \, \mathrm{d}\Psi_{t+h}(x', \lambda') = \int_{\mathbb{R}^d \times \mathcal{P}(U)} \left[\phi(x, \lambda) + h \nabla \phi(x, \lambda) \cdot b_{\Psi_t}(x, \lambda) \right] \mathrm{d}\Psi_t(x, \lambda) + o(h),$$

where

$$b_{\Psi_t}(x, \lambda) = \begin{pmatrix} \displaystyle\int_U v(x, u) \, \mathrm{d}\lambda(u) \\ \mathcal{T}_\Psi(x, \lambda) \end{pmatrix}.$$

In the formal limit for $h \to 0$, we obtain the weak formulation of the nonlinear continuity equation (1). A related heuristic derivation has been outlined also in [2, Remark 4.1], in the context of a leader–follower dynamics which also fits in the setting of [31]. In this case, the $\mathbb{R}^d$-component of b_Ψ also depends on Ψ, whereas the λ-component acts linearly on λ, modelling a Markov chain on U.

We point out that in this paper we neglect diffusive terms as a general existence theory is still missing. Even in the literature of mean-field games, well-posedness results have been shown only for the so-called potential games [24]. For these, numerical solutions have been proposed, see, e.g., [1], which are based on the iterative solution of a backward–forward system. In our setting, a first step towards including diffusion has been made in [10], where an entropic regularization for the spatially inhomogeneous replicator dynamics (see also Sect. 4) has been considered.

1.2. Results of this paper

The main objective of this paper is to present a rigorous proof of the formal derivation described above, by means of a multi-step Lagrangian scheme. The scheme we propose is suitable for approximating all equations in the class considered in [31] (we refer to Sect. 2 for the precise details). The method is a *Lagrangian* one as it is based on the approximation of the ODE (2), and it is *multi-step* because the updates of x and λ do not happen simultaneously, but follow the heuristics described above. Indeed, first we make an incremental step in λ and then use the updated λ' to make the incremental step in x.

Since the velocity field b depends explicitly on Ψ, at each incremental step the updates of x, λ, and of the distribution Ψ involve three substeps, which are the rigorous formalization of the heuristics discussed above. To be precise,

- first we update λ to λ' in the spirit of (3) (see (8));
- then we transport λ' to the state of the system $\widetilde{\Psi}$ (see (11)). This amounts to assuming that all the agents know the optimal label distribution λ' of the other agents;
- then we update the positions x to x' in the spirit of (4) where the velocity field depends on $\widetilde{\Psi}$ (see (12)). Notice that, in our general framework, the velocity field depends on Ψ and this makes the previous step necessary;
- finally, we update the global distribution to Ψ' keeping both x' and λ' into account (see (15)).

Our first main result is Theorem 1 in Sect. 3 on the convergence of the scheme presented above.

In Sects. 4 and 5, we turn our attention to the case of the inhomogeneous replicator dynamics considered in [5] and to the leader–follower-type dynamics of [31, Section 5.1], respectively. More in general, for the second case, we assume that $\mathcal{T}_\Psi(x, \lambda)$ is a Markov chain on a finite space of an arbitrary number n of labels.

In the spatially homogeneous case, that is, when there is no x dependence in the vector field b, in both situations the evolutions of the λ-components are gradient flows of suitable energies with respect to certain metric structures, and the solution can be approximated via a minimizing movement scheme [6,22]. The spatially homogeneous replicator equation is a gradient flow with respect to the spherical Hellinger distance (36) of probability measures. (This could be obtained, for instance, for a proper choice of f in [23, formula (1.8)].) The spatially homogeneous Markov-type jump processes are the gradient flow of an entropy-like energy penalized by a distance induced by the transition matrix [28,30].

We investigate the compliance of these structures with our algorithm. More precisely, we elaborate an implicit–explicit scheme where the explicit step (3) is replaced by a minimizing movement step suggested by the aforementioned gradient flow structure (see (39) and (82), respectively). A relevant difficulty in the spatially inhomogeneous setting is that the energy and the dissipation distances that we consider may as well depend on the state Ψ, which changes from step to step. This extension is far from trivial and requires a careful analysis of the related Euler conditions, which is partially inspired by [20, Section 4.2] for the case of the replicator dynamics. This is done is Propositions 3 and 5, respectively, where we show that the deviation from the explicit scheme is uniformly controlled by the vanishing time step.

The two main results of Sects. 4 and 5 are given by Theorems 2 and 3, proving the convergence of our multi-step Lagrangian scheme to the unique solution to (1. In particular, Theorem 2 is a global-in-time convergence result for the spatially inhomogeneous replicator dynamics, whereas Theorem 3 provides a short-time existence

result for a well-prepared initial datum for spatially inhomogeneous Markov-type jump processes.

The paper is structured as follows: In Sect. 2, we introduce the structural assumptions on the systems that we consider. In Sect. 3, we describe the multi-step Lagrangian scheme, which we apply to the inhomogeneous replicator dynamics in Sect. 4 and to the inhomogeneous Markov-type jump processes in Sect. 5.

2. The mathematical setting

2.1. Basic notation.

Given a metric space (X, d_X), we denote by $\mathcal{M}(X)$ the space of signed Borel measures μ in X with finite total variation $\|\mu\|_{\mathrm{TV}}$, by $\mathcal{M}_+(X)$ and $\mathcal{P}(X)$ the convex subsets of nonnegative measures and probability measures, respectively. We say that $\mu \in \mathcal{P}_c(X)$ if $\mu \in \mathcal{P}(X)$ and the support $\operatorname{spt}\mu$ of μ is a compact subset of X. Moreover, for $K \subseteq X$ we will use the notation $\mathcal{P}(K)$ to indicate the set of measures $\mu \in \mathcal{P}(X)$ such that $\operatorname{spt}\mu \subseteq K$.

As usual, if (Z, d_Z) is another metric space, for every $\mu \in \mathcal{M}_+(X)$ and every μ-measurable function $f : X \to Z$, we define the push-forward measure $f_\#\mu \in \mathcal{M}_+(Z)$ by $(f_\#\mu)(B) := \mu(f^{-1}(B))$ for any Borel set $B \subset Z$. The push-forward measures has the same total mass as μ, namely $\mu(X) = (f_\#)\mu(Z)$.

For a Lipschitz function $f : X \to \mathbb{R}$ we set

$$\mathrm{Lip}(f) := \sup_{\substack{x,y\in X\\ x\neq y}} \frac{|f(x) - f(y)|}{\mathsf{d}_X(x, y)}$$

its Lipschitz constant. We denote by $\mathrm{Lip}(X)$ and $\mathrm{Lip}_b(X)$ the spaces of Lipschitz and bounded Lipschitz functions on X, respectively. Both are normed spaces with the norm $\|f\|_{\mathrm{Lip}} := \|f\|_\infty + \mathrm{Lip}(f)$, where $\|\cdot\|_\infty$ is the supremum norm. Furthermore, we use the notation $\mathrm{Lip}_1(X)$ for the set of functions $f \in \mathrm{Lip}_b(X)$ such that $\mathrm{Lip}(f) \leq 1$.

In a complete and separable metric space (X, d_X), we shall use the Kantorovich–Rubinstein distance W_1 in the class $\mathcal{P}(X)$, defined as

$$W_1(\mu, \nu) := \sup\left\{ \int_X \varphi \,\mathrm{d}\mu - \int_X \varphi \,\mathrm{d}\nu : \varphi \in \mathrm{Lip}_1(X) \right\}.$$

Notice that $W_1(\mu, \nu)$ is finite if μ and ν belong to the space

$$\mathcal{P}_1(X) := \left\{ \mu \in \mathcal{P}(X) : \int_X \mathsf{d}_X(x, \bar{x}) \,\mathrm{d}\mu(x) < +\infty \text{ for some } \bar{x} \in X \right\}$$

and that $(\mathcal{P}_1(X), W_1)$ is complete if (X, d_X) is complete.

If $(E, \|\cdot\|_E)$ is a Banach space and $\mu \in \mathcal{M}_+(E)$, we define the first moment $m_1(\mu)$ as

$$m_1(\mu) := \int_E \|x\|_E \,\mathrm{d}\mu\,.$$

Notice that, for a probability measure μ, finiteness of the above integral is equivalent to $\mu \in \mathcal{P}_1(E)$, whenever E is endowed with the distance induced by the norm $\|\cdot\|_E$.

For a Banach space E, the notation $C_b^1(E)$ will be used to denote the subspace of $C_b(E)$ of functions having bounded continuous Fréchet differential at each point. The notation $\nabla\phi(\cdot)$ will be used to denote the Fréchet differential. In the case of a function $\phi\colon [0,T]\times E \to \mathbb{R}$, the symbol ∂_t will be used to denote partial differentiation with respect to t. The symbol $\langle\cdot,\cdot\rangle$ will be used to denote duality products, with no further specification if the meaning is clear from the context.

2.2. Functional setting

We consider a set of pure strategies U, where U is a compact metric space, and we denote by $Y := \mathbb{R}^d \times \mathcal{P}(U)$ the state space of the system. Precisely, for every $y = (x,\lambda) \in Y$, the component $x \in \mathbb{R}^d$ describes the location of an agent in space, whereas the component $\lambda \in \mathcal{P}(U)$ describes the distribution of labels of the agent.

The correct functional space for the dynamics (see also [5,31]) is the space $\overline{Y} := \mathbb{R}^d \times \mathcal{F}(U)$, where we have set (see, e.g., [7,9] and [40, Chapter 3])

$$\mathcal{F}(U) := \overline{\operatorname{span}(\mathcal{P}(U))}^{\|\cdot\|_{\mathrm{BL}}} \subseteq (\mathrm{Lip}(U))'. \tag{5}$$

The closure in (5) is taken with respect to the *bounded Lipschitz* norm $\|\cdot\|_{\mathrm{BL}}$, defined as

$$\|\mu\|_{\mathrm{BL}} := \sup\left\{\langle\mu,\varphi\rangle : \varphi \in \mathrm{Lip}(U),\ \|\varphi\|_{\mathrm{Lip}} \le 1\right\} \qquad \text{for every } \mu \in (\mathrm{Lip}(U))'.$$

We notice that, by definition of $\|\cdot\|_{\mathrm{BL}}$, we always have

$$\|\mu\|_{\mathrm{BL}} \le \|\mu\|_{\mathrm{TV}} \qquad \text{for every } \mu \in \mathcal{M}(U).$$

In particular, $\|\lambda\|_{\mathrm{BL}} \le 1$ for every $\lambda \in \mathcal{P}(U)$.

We endow $\overline{Y}$ with the norm

$$\|y\|_{\overline{Y}} = \|(x,\lambda)\|_{\overline{Y}} := |x| + \|\lambda\|_{\mathrm{BL}}.$$

For every $R > 0$, we denote by B_R the closed ball of radius R in $\mathbb{R}^d$ and by B_R^Y the ball of radius R in Y, namely $\mathrm{B}_R^Y = \{y \in Y : \|y\|_{\overline{Y}} \le R\}$. We notice that B_R^Y is a compact set, as Y is locally compact by our assumptions on U.

As in [31], we consider, for every $\Psi \in \mathcal{P}_1(Y)$, the velocity field $v_\Psi : Y \to \mathbb{R}^d$ such that

(v_1) for every $R > 0$, $v_\Psi \in \mathrm{Lip}(\mathrm{B}_R^Y;\mathbb{R}^d)$ uniformly with respect to $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$, i.e., there exists $L_{v,R} > 0$ such that

$$|v_\Psi(y_1) - v_\Psi(y_2)| \le L_{v,R}\|y_1 - y_2\|_{\overline{Y}} \qquad \text{for every } y_1, y_2 \in Y;$$

(v_2) for every $R > 0$ there exists $L_{v,R} > 0$ such that for every $\Psi_1, \Psi_2 \in \mathcal{P}(\mathrm{B}_R^Y)$ and every $y \in \mathrm{B}_R^Y$

$$|v_{\Psi_1}(y) - v_{\Psi_2}(y)| \le L_{v,R} W_1(\Psi_1,\Psi_2);$$

(v_3) there exists $M_v > 0$ such that for every $y \in Y$ and every $\Psi \in \mathcal{P}_1(Y)$

$$|v_\Psi(y)| \le M_v\big(1 + \|y\|_{\overline{Y}} + m_1(\Psi)\big).$$

As for $\mathcal{T}$, for every $\Psi \in \mathcal{P}_1(Y)$ we assume that the operator $\mathcal{T}_\Psi : Y \to \mathcal{F}(U)$ is such that

($\mathcal{T}_0$) for every $(y, \Psi) \in Y \times \mathcal{P}_1(Y)$, the constants belong to the kernel of $\mathcal{T}_\Psi(y)$, i.e.,

$$\langle \mathcal{T}_\Psi(y), 1\rangle_{\mathcal{F}(U),\mathrm{Lip}(U)} = 0;$$

($\mathcal{T}_1$) there exists $M_{\mathcal{T}} > 0$ such that for every $y \in Y$ and every $\Psi \in \mathcal{P}_1(Y)$

$$\|\mathcal{T}_\Psi(y)\|_{\mathrm{BL}} \le M_{\mathcal{T}}\big(1 + \|y\|_{\overline{Y}} + m_1(\Psi)\big);$$

($\mathcal{T}_2$) for every $R > 0$, there exists $L_{\mathcal{T},R} > 0$ such that for every $(y_1, \Psi_1), (y_2, \Psi_2) \in \mathrm{B}_R^Y \times \mathcal{P}(\mathrm{B}_R^Y)$

$$\|\mathcal{T}_{\Psi_1}(y_1) - \mathcal{T}_{\Psi_2}(y_2)\|_{\overline{Y}} \le L_{\mathcal{T},R}\big(\|y_1 - y_2\|_{\mathrm{BL}} + W_1(\Psi_1, \Psi_2)\big);$$

($\mathcal{T}_3$) for every $R > 0$ there exists $\delta_R > 0$ such that for every $(y, \Psi) \in \mathrm{B}_R^Y \times \mathcal{P}_1(Y)$ we have

$$\mathcal{T}_\Psi(y) + \delta_R \lambda \ge 0.$$

Finally, for every $y \in Y$ and every $\Psi \in \mathcal{P}_1(Y)$ we set

$$b_\Psi(y) := \begin{pmatrix} v_\Psi(y) \\ \mathcal{T}_\Psi(y) \end{pmatrix}, \tag{6}$$

which is the velocity field driving the evolution (see (7)).

3. The multi-step Lagrangian scheme

Let $\widehat{\Psi} \in \mathcal{P}_c(Y)$ be a probability measure on Y with compact support in Y. Given $T > 0$, for every $k \in \mathbb{N} \setminus \{0\}$ we set $\tau_k : -T/k$ and, for $i \in \{0, \dots, k\}$, $t_i^k := i\tau_k$.

We now show how to construct a curve $\Psi^k : [0, T] \to \mathcal{P}_1(Y)$, defined piecewise on each time interval $[t_i^k, t_{i+1}k]$, which approximates a solution $\Psi \in C([0, 1]; \mathcal{P}_1(Y))$ of the initial value problem for the nonlinear continuity equation

$$\partial_t \Psi_t + \mathrm{div}(b_{\Psi_t}\Psi_t) = 0, \qquad \Psi_0 = \widehat{\Psi}. \tag{7}$$

Let $\Psi_0^k := \widehat{\Psi}$. In each interval $[t_i^k, t_{i+1}^k)$, assume the measure $\Psi_i^k \in \mathcal{P}_1(Y)$ to be known. With this knowledge, we update the state of the system with the following procedure, consisting of two steps.

Step 1. We update the label $\lambda_{(\hat{x},\hat{\lambda})}(t) \in \mathcal{P}(U)$ of a player that at time t_i^k sits in $\hat{x} \in \mathbb{R}^d$ with label $\hat{\lambda} \in \mathcal{P}(U)$ by setting

$$\lambda_{(\hat{x},\hat{\lambda})}(t_{i+1}^k) := \lambda_{(\hat{x},\hat{\lambda})} + \tau_k \mathcal{T}_{\Psi_i^k}\big(\hat{x}, \lambda_{(\hat{x},\hat{\lambda})}(t_i^k)\big). \tag{8}$$

At this stage, we assume that $\lambda_{(\hat{x},\hat{\lambda})}(t_{i+1}^k) \in \mathcal{P}(U)$ and we continue with the construction of the piecewise affine interpolant between $\lambda_{(\hat{x},\hat{\lambda})}(t_i^k)$ and $\lambda_{(\hat{x},\hat{\lambda})}(t_{i+1}^k)$, defined as the function $\lambda^k_{(\hat{x},\hat{\lambda}),i+1} : [t_i^k, t_{i+1}^k] \to \mathcal{P}(U)$ such that

$$\lambda^k_{(\hat{x},\hat{\lambda}),i+1}(t) := \frac{t - t_i^k}{\tau_k}\lambda_{(\hat{x},\hat{\lambda})}(t_{i+1}^k) + \left(1 - \frac{t - t_i^k}{\tau_k}\right)\lambda_{(\hat{x},\hat{\lambda})}(t_i^k)\,. \tag{9}$$

In Lemma 1, we show that the assumption $\lambda_{(\hat{x},\hat{\lambda})}(t_{i+1}^k) \in \mathcal{P}(U)$ is actually satisfied for k large enough (and therefore τ_k small enough), independently of $i = 0, \dots, k-1$. Giving Lemma 1 for granted for the time being, we define the map $\Lambda_{i+1}^k : [t_i^k, t_{i+1}^k] \times \mathbb{R}^d \times \mathcal{P}(U) \to \mathcal{P}(U)$ as

$$\Lambda_{i+1}^k(t, \hat{x}, \hat{\lambda}) := \lambda^k_{(\hat{x},\hat{\lambda}),i+1}(t) \qquad \text{for every } (t, \hat{x}, \hat{\lambda}) \in [t_i^k, t_{i+1}^k] \times \mathbb{R}^d \times \mathcal{P}(U), \tag{10}$$

and transport it to the state of the system by defining

$$\widetilde{\Psi}_{i+1}^k := (id; \Lambda_{i+1}^k(t_{i+1}^k, \cdot, \cdot))_{\#}\Psi_i^k \in \mathcal{P}_1(Y)\,. \tag{11}$$

Step 2. In the second step, we update the positions of the players. Precisely, a player that at time t_i^k sits in the position $\hat{x}$ with label $\hat{\lambda}$ will now move following the velocity field given by $v_{\widetilde{\Psi}_{i+1}^k}\left(x_{(\hat{x},\hat{\lambda})}(t_i^k), \lambda^k_{(\hat{x},\hat{\lambda}),i+1}(t_{i+1}^k)\right)$, which is determined by the updated label $\lambda^k_{(\hat{x},\hat{\lambda}),i+1}(t_{i+1}^k)$ just obtained in (8). Hence, we set

$$x_{(\hat{x},\hat{\lambda})}(t_{i+1}^k) := x_{(\hat{x},\hat{\lambda})}(t_i^k) + \tau_k v_{\widetilde{\Psi}_{i+1}^k}\left(x_{(\hat{x},\hat{\lambda})}(t_i^k), \lambda^k_{(\hat{x},\hat{\lambda}),i+1}(t_{i+1}^k)\right). \tag{12}$$

Also in this case, we can define the affine interpolant between $x_{(\hat{x},\hat{\lambda})}(t_i^k)$ and $x_{(\hat{x},\hat{\lambda})}(t_{i+1}^k)$, as a function $x^k_{(\hat{x},\hat{\lambda}),i+1} : [t_i^k, t_{i+1}^k] \to \mathbb{R}^d$, by

$$x^k_{(\hat{x},\hat{\lambda}),i+1}(t) := \frac{t - t_i^k}{\tau_k}x_{(\hat{x},\hat{\lambda})}(t_{i+1}^k) + \left(1 - \frac{t - t_i^k}{\tau_k}\right)x_{(\hat{x},\hat{\lambda})}(t_i^k). \tag{13}$$

We notice that (13), in contrast to (9), is always well defined, since $\mathbb{R}^d$ is a convex space and the velocity field is an element of $\mathbb{R}^d$.

Eventually, we define the map $X_{i+1}^k : [t_i^k, t_{i+1}^k] \times \mathbb{R}^d \times \mathcal{P}(U) \to \mathbb{R}^d$ as

$$X_{i+1}^k(t, \hat{x}, \hat{\lambda}) := x^k_{(\hat{x},\hat{\lambda}),i+1}(t) \quad \text{for every } (t, \hat{x}, \hat{\lambda}) \in [t_i^k, t_{i+1}^k] \times \mathbb{R}^d \times \mathcal{P}(U) \tag{14}$$

and we set, for $t \in [t_i^k, t_{i+1}^k]$,

$$\Psi^k(t) := \left(X_{i+1}^k(t, \cdot, \cdot); \Lambda_{i+1}^k(t, \cdot, \cdot)\right)_{\#}\Psi_i^k\,, \qquad \Psi_{i+1}^k := \Psi^k(t_{i+1}^k)\,. \tag{15}$$

For later use, we also define

$$\widetilde{\Psi}^k(t) := \widetilde{\Psi}_{i+1}^k \qquad \text{for every } t \in (t_i^k, t_{i+1}^k]\,, \tag{16}$$

$$\underline{\Psi}^k(t) := \Psi_i^k \qquad \text{for every } t \in [t_i^k, t_{i+1}^k)\,. \tag{17}$$

By an application of Gronwall inequality, in the following lemma we give an estimate of $\left|x^k_{(\hat{x},\hat{\lambda}),i+1}(t)\right|$ and $\left\|\lambda^k_{(\hat{x},\hat{\lambda}),i+1}(t)\right\|_{\mathrm{BL}}$ in terms of $|\hat{x}|$ and $\|\hat{\lambda}\|_{\mathrm{BL}}$. As a consequence, we deduce that the above construction is well defined for τ_k sufficiently small and can be iterated over $i = 0, \dots, k-1$, since the initial condition $\widehat{\Psi}$ has a compact support in Y. This indeed implies that each Ψ^k_i belongs to $\mathcal{P}_c(Y) \subseteq \mathcal{P}_1(Y)$.

Lemma 1. *Let $\widehat{\Psi} \in \mathcal{P}_c(Y)$. Then, for k large enough the curves $\Psi^k(\cdot)$, $\underline{\Psi}^k(\cdot)$, and $\widetilde{\Psi}^k(\cdot)$ are well defined from $[0,T]$ with values in $\mathcal{P}_1(Y)$. Furthermore, there exists $R > 0$ independent of k and t such that $\Psi^k(t), \underline{\Psi}^k(t), \widetilde{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_R)$.*

Proof. Along the proof of the lemma, we denote with $\lambda^k(t, x_0, \lambda_0)$ and $x^k(t, x_0, \lambda_0)$, for $(x_0, \lambda_0) \in \mathrm{spt}\widehat{\Psi} =: \mathcal{S}$, the curves obtained by iteratively solving the difference equations (8) and (12) in each interval $[t^k_i, t^k_{i+1}]$ starting from (x_0, λ_0) at time $t_0 = 0$ and using, at each node $t^k_i, i = 1, \dots, k-1, \hat{\lambda} = \lambda^k(t^k_i, x_0, \lambda_0)$ and $\hat{x} = x^k(t^k_i, x_0, \lambda_0)$ as new initial conditions.

As we have already noticed above, the curve $x^k(t, x_0, \lambda_0)$ is well defined as long as $\lambda^k(t, x_0, \lambda_0)$ and the measures $\widetilde{\Psi}^k_i$ are. Therefore, in order to prove the lemma it is enough to show that, for τ_k small enough, for every $(x_0, \lambda_0) \in \mathrm{spt}\,\widehat{\Psi}$ the piecewise linear interpolant $\lambda^k(t, x_0, \lambda_0)$ always belongs to $\mathcal{P}(U)$. This can be done recursively by arguing on each interval $[t^k_i, t^k_{i+1}], i = 0, \dots, k-1$.

To simplify our estimates, we define the piecewise constant interpolation functions

$$\begin{aligned} &\underline{x}^k(t, x_0, \lambda_0) := x^k(t^k_j, x_0, \lambda_0)\,, \quad \underline{\lambda}^k(t, x_0, \lambda_0) := \lambda^k(t^k_j, x_0, \lambda_0) && \text{for } t \in [t^k_j, t^k_{j+1})\,, \\ &\overline{\lambda}^k(t, x_0, \lambda_0) := \lambda^k(t^k_{j+1}, x_0, \lambda_0) && \text{for } t \in (t^k_j, t^k_{j+1}]\,. \end{aligned} \tag{18}$$

For $i = 0$, we have that the initial condition $\lambda_0 \in \mathcal{P}(U)$; hence, there is nothing to show. Assuming that $\lambda^k(t^k_j, x_0, \lambda_0) \in \mathcal{P}(U)$ for every $j = 0, \dots, i$ and every $(x_0, \lambda_0) \in \mathrm{spt}\,\widehat{\Psi}$, we show that $\lambda^k(t^k_{i+1}, x_0, \lambda_0) \in \mathcal{P}(U)$ for k large enough, independently of i and of the initial condition (x_0, λ_0). Since, recalling (8) and (9), we define

$$\lambda^k(t, x_0, \lambda_0) := \underline{\lambda}^k(t, x_0, \lambda_0) + (t - t^k_i)\mathcal{T}_{\underline{\Psi}^k(t)}(\underline{x}^k(t, x_0, \lambda_0), \underline{\lambda}^k(t, x_0, \lambda_0))$$

for $t \in [t^k_i, t^k_{i+1}]$; by assumptions $(\mathcal{T}_0)$ and $(\mathcal{T}_3)$, we are led to showing that the piecewise constant interpolation functions $\underline{x}^k(t, x_0, \lambda_0)$ and $\underline{\lambda}^k(t, x_0, \lambda_0)$ are bounded in $\mathbb{R}^d$ and $\mathcal{F}(U)$, respectively, uniformly with respect to $(x_0, \lambda_0) \in \mathcal{S}$ and $t \in [0, t^k_{i+1}]$, and that the bound does not depend on i. Indeed, if this is the case, let $R' > 0$ be such that $(\underline{x}^k(t, x_0, \lambda_0), \underline{\lambda}^k(t, x_0, \lambda_0)) \in \mathrm{B}^Y_{R'}$ for every $t \in [0, t^k_{i+1}]$ and every $(x_0, \lambda_0) \in \mathcal{S}$. In particular, by construction (17) of $\underline{\Psi}^k(t)$ it holds $\underline{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_{R'})$. By $(\mathcal{T}_3)$, there exists $\delta_{R'} > 0$, independent of k, i, and $(x_0, \lambda_0) \in \mathcal{S}$, such that for $t \in [t^k_i, t^k_{i+1}]$

$$\lambda_{R'} := \frac{1}{\delta_{R'}}\mathcal{T}_{\underline{\Psi}^k(t)}\big(\underline{x}^k(t, x_0, \lambda_0), \underline{\lambda}^k(t, x_0, \lambda_0)\big) + \underline{\lambda}^k(t, x_0, \lambda_0) \geq 0\,.$$

In particular, assumption $(\mathcal{T}_1)$ implies that $\lambda_{R'} \in \mathcal{F}(U)$ and satisfies

$$\left| \langle \lambda_{R'}, \eta \rangle_{\mathcal{F}(U),\mathrm{Lip}(U)} \right| \leq \|\eta\|_\infty \|\lambda_{R'}\|_{\mathrm{BL}} ,$$

so that $\lambda_{R'}$ can be extended in a unique way to a linear and continuous operator on $C(U)$. The Riesz representation theorem yields that $\lambda_{R'} \in \mathcal{M}_+(U)$. Moreover, by $(\mathcal{T}_0)$ we get

$$\langle \lambda_{R'}, 1 \rangle_{\mathcal{F}(U),\mathrm{Lip}(U)} = \left\langle \underline{\lambda}^k(t, x_0, \lambda_0), 1 \right\rangle_{\mathcal{F}(U),\mathrm{Lip}(U)} = 1,$$

which implies $\lambda_{R'} \in \mathcal{P}(U)$. By the convexity of $\mathcal{P}(U)$ we deduce that whenever $\tau_k \leq 1/\delta_{R'}$

$$\lambda^k(t, x_0, \lambda_0) = \underline{\lambda}^k(t, x_0, \lambda_0) + (t - t_i^k)\mathcal{T}_{\underline{\Psi}(t_i^k)}(\underline{x}^k(t, x_0, \lambda_0), \underline{\lambda}^k(t, x_0, \lambda_0)) \in \mathcal{P}(U)$$

for every $t \in [t_i^k, t_{i+1}^k]$. Being the upper bound R' independent of i and of $(x_0, \lambda_0) \in \mathcal{S}$, also $\delta_{R'}$ is. Hence, the trajectories $x^k(\cdot, x_0, \lambda_0)$ and $\lambda^k(\cdot, x_0, \lambda_0)$ are well defined from $[0, T]$ with values in $\mathbb{R}^d$ and $\mathcal{P}(U)$, respectively.

In order to conclude that the interpolation curves $x^k(t, x_0, \lambda_0)$ and $\lambda^k(t, x_0, \lambda_0)$ are well defined, we have to estimate $|\underline{x}^k(t, x_0, \lambda_0)|$ and $\|\underline{\lambda}^k(t, x_0, \lambda_0)\|_{\mathrm{BL}}$ for $(x_0, \lambda_0) \in \mathcal{S}$. Since we are assuming that $\lambda^k(t_j^k, x_0, \lambda_0) \in \mathcal{P}(U)$ for $j \in 0, \ldots, i$, we have that $\|\lambda^k(t_j^k, x_0, \lambda_0)\|_{\mathrm{BL}} \leq 1$, and the same holds for $\|\underline{\lambda}^k(t, x_0, \lambda_0)\|_{\mathrm{BL}}$. As for $\underline{x}^k(t, x_0, \lambda_0)$, using (12) and (v_3) we get

$$\begin{aligned} |\underline{x}^k(t, x_0, \lambda_0)| &\leq |x_0| + \int_0^{t_i^k} \left| v_{\widetilde{\Psi}^k(\tau)}(\underline{x}^k(\tau, x_0, \lambda_0), \overline{\lambda}^k(\tau, x_0, \lambda_0) \right| \mathrm{d}\tau \qquad (19) \\ &\leq |x_0| + \int_0^t M_v \Big(3 + 2 \sup_{(\hat{x},\hat{\lambda})\in\mathcal{S}} |\underline{x}_k(\tau, \hat{x}, \hat{\lambda})| \Big) \mathrm{d}\tau . \end{aligned}$$

Let us now fix $r > 0$ such that $\mathcal{S} \subseteq \mathrm{B}_r^Y$ and let

$$f_k(t) := \sup_{(\hat{x},\hat{\lambda})\in\mathcal{S}} |\underline{x}^k(t, \hat{x}, \hat{\lambda})| .$$

By taking the supremum over $\mathcal{S}$ in (19), we deduce that

$$f_k(t) \leq r + \int_0^t 3M_v(1 + f_k(\tau)) \, \mathrm{d}\tau . \qquad (20)$$

Applying the Gronwall inequality to (20), we infer that

$$f_k(t) \leq (r + 3M_v T)e^{3M_v T} . \qquad (21)$$

Setting $R' := 1 + (r + 3M_v T)e^{3M_v T}$, we have proved that the piecewise constant interpolation function $t \mapsto (\underline{x}^k(t, x_0, \lambda_0), \underline{\lambda}^k(t, x_0, \lambda_0))$ belongs to $\mathrm{B}_{R'}^Y$ for every $t \in [t_i^k, t_{i+1}^k)$ and every $(x_0, \lambda_0) \in \mathcal{S}$. In particular, we notice that the above computations

are independent of the choice of i, as long as we know that $\lambda^k(t_j^k, x_0, \lambda_0) \in \mathcal{P}(U)$ for every $j = 0, \dots, i$ and every $(x_0, \lambda_0) \in \mathcal{S}$. With this control at hand, we conclude, as explained above, that (8) and (12) are well posed.

Finally, we estimate $x^k(t, x_0, \lambda_0)$. For $(x_0, \lambda_0) \in \mathrm{spt}\widehat{\Psi}$ and $t \in [0, T]$, by (v_3) we have

$$\begin{aligned} |x^k(t, x_0, \lambda_0)| &\le |x_0| + \int_0^t \left| v_{\widetilde{\Psi}^k(\tau)}\big(\underline{x}^k(\tau, x_0, \lambda_0), \overline{\lambda}^k(\tau, x_0, \lambda_0)\big)\right| \mathrm{d}\tau \\ &\le r + 2M_v(1 + R')T\,. \end{aligned} \tag{22}$$

Setting $R := \max\{R', r + 2M_v(1 + R')T + 1\}$, we obtain that $\Psi^k(t), \widetilde{\Psi}^k(t), \underline{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}_R^Y)$ for every $t \in [0, T]$ and every $k \in \mathbb{N}$ large enough. □

In the next proposition, we show that the curve $\Psi^k(\cdot)$ solves the continuity equation (7) up to an error of order τ_k.

Proposition 1. *Let $\widehat{\Psi} \in \mathcal{P}_c(Y)$, let $\Psi^k\colon [0, T] \to \mathcal{P}_1(Y)$ be the curve defined in (15) starting from $\widehat{\Psi}$, and let $\widetilde{\Psi}^k$ be as in (16). Then, the following holds: There exists a positive constant C such that for every $\varphi \in C_b^1(\mathbb{R}^d \times \mathcal{F}(U))$, every $k \in \mathbb{N}$, every $i \in \{0, \dots, k-1\}$, and every $t \in (t_i^k, t_{i+1}^k)$,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_Y \varphi(x, \lambda)\,\mathrm{d}\Psi^k(t)(x, \lambda) = \int_Y \nabla\varphi(x, \lambda) \cdot b_{\Psi^k(t)}(x, \lambda)\,\mathrm{d}\Psi^k(t)(x, \lambda) + \vartheta_k(\varphi), \tag{23}$$

where $|\vartheta_k(\varphi)| \le C\|\varphi\|_{C_b^1}\tau_k$.

Proof. Let us fix $\varphi \in C_b^1(\mathbb{R}^d \times \mathcal{F}(U))$ and $t \in (t_i^k, t_{i+1}^k)$. By definition of $\Psi^k(t)$, we have that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int_Y \varphi(x, \lambda)\,\mathrm{d}\Psi^k(t)(x, \lambda) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_Y \varphi\big(X_{i+1}^k(t, x, \lambda), \Lambda_{i+1}^k(t, x, \lambda)\big)\,\mathrm{d}\Psi_i^k(x, \lambda) \\ &= \int_Y \nabla_x\varphi\big(X_{i+1}^k(t, x, \lambda), \Lambda_{i+1}^k(t, x, \lambda)\big) \cdot v_{\widetilde{\Psi}^k(t)}\big(x, \Lambda_{i+1}^k(t_{i+1}^k, x, \lambda)\big)\,\mathrm{d}\Psi_i^k(x, \lambda) \\ &\quad + \int_Y \nabla_\lambda\varphi\big(X_{i+1}^k(t, x, \lambda), \Lambda_{i+1}^k(t, x, \lambda)\big) \cdot \mathcal{T}_{\underline{\Psi}^k(t)}(x, \lambda)\,\mathrm{d}\Psi_i^k(x, \lambda), \end{aligned} \tag{24}$$

where $\widetilde{\Psi}^k(t)$ and $\underline{\Psi}^k(t)$ are defined in (16) and (17), respectively. In order to obtain (23) from (24), we have to estimate the following quantities:

$$\begin{aligned} I_1(x, \lambda) &:= \left| v_{\widetilde{\Psi}^k(t)}\big(x, \Lambda_{i+1}^k(t_{i+1}^k, x, \lambda)\big) - v_{\Psi^k(t)}\big(X_{i+1}^k(t, x, \lambda), \Lambda_{i+1}^k(t, x, \lambda)\big)\right|, \\ I_2(x, \lambda) &:= \left\| \mathcal{T}_{\underline{\Psi}^k(t)}(x, \lambda) - \mathcal{T}_{\Psi^k(t)}(X_{i+1}^k(t, x, \lambda), \Lambda_{i+1}^k(t, x, \lambda))\right\|_{\mathrm{BL}} \end{aligned}$$

for $(x, \lambda) \in \mathrm{spt}\Psi_i^k \subseteq \mathrm{B}_R^Y$, where R has been determined in Lemma 1.

Let us start with I_1. By triangle inequality, we have

$$\begin{aligned} I_1(x,\lambda) &\le \Big| v_{\widetilde{\Psi}^k(t)}\big(x,\, \Lambda^k_{i+1}(t^k_{i+1},x,\lambda)\big) - v_{\widetilde{\Psi}^k(t)}\big(X^k_{i+1}(t,x,\lambda),\, \Lambda^k_{i+1}(t,x,\lambda)\big)\Big| \\ &\quad + \Big| v_{\widetilde{\Psi}^k(t)}\big(X^k_{i+1}(t,x,\lambda),\, \Lambda^k_{i+1}(t,x,\lambda)\big) - v_{\Psi^k(t)}\big(X^k_{i+1}(t,x,\lambda),\, \Lambda^k_{i+1}(t,x,\lambda)\big)\Big| \\ &=: I_{1,1}(x,\lambda) + I_{1,2}(x,\lambda)\,. \end{aligned} \tag{25}$$

Since $\widetilde{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_R)$, hypothesis (v_1) implies that

$$\begin{aligned} I_{1,1}(x,\lambda) &\le L_{v,R}\big(|X^k_{i+1}(t,x,\lambda) - x| + \|\Lambda^k_{i+1}(t^k_{i+1},x,\lambda) - \Lambda^k_{i+1}(t,x,\lambda)\|_{\mathrm{BL}}\big) \\ &\le L_{v,R}\bigg(\int_{t^k_i}^{t} \big|v_{\widetilde{\Psi}^k(\tau)}\big(x,\, \Lambda^k_{i+1}(t^k_{i+1},x,\lambda)\big)\big|\,\mathrm{d}\tau + \int_t^{t^k_{i+1}} \big\|\mathcal{T}_{\Psi^k_i}(x,\lambda)\big\|_{\mathrm{BL}}\,\mathrm{d}\tau\bigg), \end{aligned}$$

where, in the second inequality, we have used the systems (9) and (13). By (v_3) and $(\mathcal{T}_3)$, we can continue with

$$\begin{aligned} I_{1,1}(x,\lambda) &\le L_{v,R}\bigg(M_v \int_{t^k_i}^{t} \big(1 + |x| + \|\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)\|_{\mathrm{BL}} + m_1(\widetilde{\Psi}^k(\tau))\big)\,\mathrm{d}\tau \\ &\qquad + M_{\mathcal{T}} \int_t^{t^k_{i+1}} \big(1 + |x| + \|\lambda\|_{\mathrm{BL}} + m_1(\Psi^k_i)\big)\,\mathrm{d}\tau\bigg) \\ &\le L_{v,R}(M_v + M_{\mathcal{T}})(1+2R)\tau_k\,. \end{aligned} \tag{26}$$

As for $I_{1,2}$, thanks to assumption (v_2) and to Lemma 1 we get

$$\begin{aligned} I_{1,2}(x,\lambda) &\le L_{v,R} W_1(\widetilde{\Psi}^k(t), \Psi^k(t)) \\ &= L_{v,R} \sup_{\eta\in \mathrm{Lip}_1(Y)} \bigg\{ \int_Y \eta(x',\lambda')\,\mathrm{d}(\widetilde{\Psi}^k(t) - \Psi^k(t))(x',\lambda') \bigg\} \\ &= L_{v,R} \sup_{\eta\in \mathrm{Lip}_1(Y)} \bigg\{ \int_Y \big[\eta(x, \Lambda^k_{i+1}(t^k_{i+1},x',\lambda')) \\ &\quad - \eta(X^k_{i+1}(t,x',\lambda'), \Lambda^k_{i+1}(t,x',\lambda'))\big]\,\mathrm{d}\Psi^k_i(x',\lambda') \bigg\} \\ &\le L_{v,R} \int_Y \big[|x - X^k_{i+1}(t,x',\lambda')| \\ &\quad + \|\Lambda^k_{i+1}(t^k_{i+1},x',\lambda')) - \Lambda^k_{i+1}(t,x',\lambda')\|_{\mathrm{BL}}\big]\,\mathrm{d}\Psi^k_i(x',\lambda') \\ &\le L_{v,R} \int_Y \bigg(\int_{t^k_i}^{t} \big|v_{\widetilde{\Psi}^k(\tau)}(x, \Lambda^k_{i+1}(t^k_{i+1},x',\lambda'))\big|\,\mathrm{d}\tau \\ &\quad + \int_t^{t^k_{i+1}} \big\|\mathcal{T}_{\Psi^k_i}(x',\lambda')\big\|_{\mathrm{BL}}\,\mathrm{d}\tau\bigg)\,\mathrm{d}\Psi^k_i(x',\lambda') \\ &\le L_{v,R}\,\tau_k \int_Y \Big(\big|v_{\widetilde{\Psi}^k(\tau)}(x, \Lambda^k_{i+1}(t^k_{i+1},x',\lambda'))\big| + \big\|\mathcal{T}_{\Psi^k_i}(x',\lambda')\big\|_{\mathrm{BL}}\Big)\,\mathrm{d}\Psi^k_i(x',\lambda')\,. \end{aligned}$$

Making use of (v_3) and $(\mathcal{T}_3)$ and recalling Lemma 1, we can continue with

$$\begin{aligned} I_{1,2}(x,\lambda) &\le L_{v,R}(M_v + M_{\mathcal{T}})\,\tau_k \int_Y \Big(1 + |x'| + \|\Lambda_{i+1}^k(t_{i+1}^k, x', \lambda')\|_{\mathrm{BL}} + \|\lambda'\|_{\mathrm{BL}} \\ &\quad + m_1(\widetilde{\Psi}^k(t)) + m_1(\Psi_i^k)\Big)\, \mathrm{d}\Psi_i^k(x', \lambda') \\ &\le 3L_{v,R}(M_v + M_{\mathcal{T}})(1+R)\tau_k\,. \end{aligned} \tag{27}$$

Combining (25)–(27), we get

$$I_1(x,\lambda) \le C_1\tau_k \tag{28}$$

for some positive constant C_1 independent of k, t, φ, and $(x,\lambda) \in \mathrm{spt}\Psi_i^k$.

Let us now estimate I_2. By Lemma 1 and by assumption $(\mathcal{T}_2)$, we get

$$\begin{aligned} I_2(x,\lambda) \le L_{\mathcal{T},R}\big(&|x - X_{i+1}^k(t,x,\lambda)| + \|\lambda - \Lambda_{i+1}^k(t,x,\lambda)\|_{\mathrm{BL}} \\ &+ W_1(\underline{\Psi}^k(t), \Psi^k(t))\big)\,. \end{aligned} \tag{29}$$

Arguing as in (25)–(28), we deduce from (29) and from the hypotheses (v_1), (v_3), and $(\mathcal{T}_2)$ that

$$I_2(x,\lambda) \le C_2\tau_k \tag{30}$$

for some positive constant C_2 independent of k, t, φ, and $(x,\lambda) \in \mathrm{spt}\Psi_i^k$.

We are now in a position to conclude the proof of the proposition. We rewrite (24) as

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\int_Y \varphi(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda) \\ &\quad= \int_Y \nabla_x\varphi\big(X_{i+1}^k(t,x,\lambda), \Lambda_{i+1}^k(t,x,\lambda)\big)\cdot \\ &\qquad\cdot v_{\Psi^k(t)}\big(X_{i+1}^k(t,x,\lambda), \Lambda_{i+1}^k(t,x,\lambda)\big)\,\mathrm{d}\Psi_i^k(x,\lambda) \\ &\qquad+ \int_Y \nabla_\lambda\varphi\big(X_{i+1}^k(t,x,\lambda), \Lambda_{i+1}^k(t,x,\lambda)\big) \\ &\qquad\cdot \mathcal{T}_{\Psi^k(t)}(X_{i+1}^k(t,x,\lambda), \Lambda^k(t,x,\lambda))\,\mathrm{d}\Psi_i^k(x,\lambda) \\ &\qquad+ \int_Y \nabla_x\varphi\big(X_{i+1}^k(t,x,\lambda), \Lambda_{i+1}^k(t,x,\lambda)\big)\cdot \big(v_{\widetilde{\Psi}^k(t)}\big(x, \Lambda_{i+1}^k(t_{i+1}^k, x, \lambda)\big) \\ &\qquad\quad - v_{\Psi^k(t)}\big(X_{i+1}^k(t,x,\lambda), \Lambda_{i+1}^k(t,x,\lambda)\big)\,\mathrm{d}\Psi_i^k(x,\lambda) \\ &\qquad+ \int_Y \nabla_\lambda\varphi\big(X_{i+1}^k(t,x,\lambda), \Lambda_{i+1}^k(t,x,\lambda)\big)\cdot \big(\mathcal{T}_{\underline{\Psi}^k(t)}(x,\lambda) \\ &\qquad- \mathcal{T}_{\Psi^k(t)}(X_{i+1}^k(t,x,\lambda), \Lambda^k(t,x,\lambda))\big)\,\mathrm{d}\Psi_i^k(x,\lambda) \\ &\qquad= \int_Y \nabla\varphi(x,\lambda)\cdot b_{\Psi^k(t)}(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda) \\ &\qquad+ \int_Y \nabla_x\varphi\big(X_{i+1}^k(t,x,\lambda), \Lambda_{i+1}^k(t,x,\lambda)\big)\cdot \big(v_{\widetilde{\Psi}^k(t)}\big(x, \Lambda_{i+1}^k(t_{i+1}^k, x, \lambda)\big) \end{aligned}$$

$$
\begin{aligned}
&- v_{\Psi^k(t)}\big(X^k_{i+1}(t,x,\lambda),\Lambda^k_{i+1}(t,x,\lambda)\big)\,\mathrm{d}\Psi^k_i(x,\lambda)\\
&+\int_Y \nabla_\lambda\varphi\big(X^k_{i+1}(t,x,\lambda),\Lambda^k_{i+1}(t,x,\lambda)\big)\cdot\big(\mathcal{T}_{\underline{\Psi}^k(t)}(x,\lambda)\\
&-\mathcal{T}_{\Psi^k(t)}(X^k_{i+1}(t,x,\lambda),\Lambda^k(t,x,\lambda))\big)\,\mathrm{d}\Psi^k_i(x,\lambda)\\
&=:\int_Y \nabla\varphi(x,\lambda)\cdot b_{\Psi^k(t)}(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)+\vartheta_k(\varphi)\,.
\end{aligned}
$$

We conclude by noticing that, thanks to (28) and (30), the term $\vartheta_k(\varphi)$ above can be estimated by

$$
|\vartheta_k(\varphi)|\le\|\varphi\|_{C^1_b}\int_Y (I_1(x,\lambda)+I_2(x,\lambda))\,\mathrm{d}\Psi^k_i(x,\lambda)\le C\|\varphi\|_{C^1_b}\tau_k\,,
$$

for a positive constant C independent of k, t, and φ. □

Theorem 1. *Let $\widehat{\Psi}\in\mathcal{P}_c(Y)$ and let $\Psi^k(\cdot)$ be defined as in (15). Then,*

$$
\lim_{k\to\infty}\sup_{t\in[0,T]} W_1(\Psi^k(t),\Psi(t))=0,
$$

where the curve $\Psi\in C([0,T];(\mathcal{P}_1(Y),W_1))$ is the unique solution of (7) with initial condition $\Psi(0)=\widehat{\Psi}$.

Proof. The existence and uniqueness of the solution to equation (7) follow from [31, Theorem 3.5], so that $\Psi\in C([0,T];(\mathcal{P}_1(Y),W_1))$ is well defined.

Let $\phi\in C^1_b([0,T]\times\overline{Y})$. In view of Proposition 1, for every $k\in\mathbb{N}, i\in\{0,\dots,k-1\}$, and every $t\in(t^k_i,t^k_{i+1})$, we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_Y\phi(t,x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)=&\int_Y\partial_t\phi(t,x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)\\
&+\int_Y\nabla\phi(t,x,\lambda)\cdot b_{\Psi^k(t)}(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)\\
&+\theta_k(\phi(t,\cdot,\cdot)),
\end{aligned}
$$

where $|\theta_k(\phi(t,\cdot,\cdot))|\le C\tau_k\|\phi\|_{C^1_b([0,T]\times\overline{Y})}$ uniformly in $[0,T]$. By integrating the previous equality over time, we deduce that

$$
\begin{aligned}
&\int_Y\phi(t,x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)-\int_Y\phi(0,x,\lambda)\,\mathrm{d}\widehat{\Psi}(x,\lambda)\\
&\quad=\int_0^t\int_Y\partial_t\phi(\tau,x,\lambda)\,\mathrm{d}\Psi^k(\tau)(x,\lambda)\,\mathrm{d}\tau\\
&\qquad+\int_0^t\int_Y\nabla\phi(\tau,x,\lambda)\cdot b_{\Psi^k(\tau)}(x,\lambda)\,\mathrm{d}\Psi^k(\tau)(x,\lambda)\,\mathrm{d}\tau\\
&\qquad+\int_0^t\theta_k(\phi(\tau,\cdot,\cdot))\,\mathrm{d}\tau\,.
\end{aligned}
\tag{31}
$$

In order to pass to the limit in (31), we have to determine a candidate limit for $\Psi^k(t)$. In Lemma 1, we have already shown that the supports of $\Psi^k(t)$ are contained in a

compact subset of $\overline{Y}$. We now show the equicontinuity of the sequence Ψ^k with respect to W_1. Given $s, t \in [0, T]$, we show that $W_1(\Psi^k(s), \Psi^k(t)) \leq L|s-t|$ for some $L > 0$ independent of k. By triangle inequality, it is enough to show it for $s, t \in [t_i^k, t_{i+1}^k]$. Arguing as in the proof of Proposition 1, we obtain

$$W_1(\Psi^k(s), \Psi^k(t)) \leq 3(M_v + M_{\mathcal{T}})(1 + R)|s - t|\,, \tag{32}$$

where R has been defined in Lemma 1. Hence, Ascoli–Arzelà theorem yields that there exists $\overline{\Psi} \in C([0, T]; (\mathcal{P}_1(Y), W_1))$ such that, up to a subsequence, $W_1(\Psi^k(t), \overline{\Psi}(t)) \to 0$ uniformly with respect to $t \in [0, T]$. In particular, $\overline{\Psi}(0) = \widehat{\Psi}$ and $\overline{\Psi}(t) \in \mathcal{P}(\mathrm{B}_R^Y)$, since $\mathrm{spt}\Psi^k(t) \subseteq \mathrm{B}_R^Y$ for every k and every t.

It remains to show that $\overline{\Psi}$ is a solution to (7), from which we would deduce that $\overline{\Psi} = \Psi$ and that the whole sequence Ψ^k converges to Ψ. The first line of (31) passes to limit as $k \to \infty$, since the test function ϕ belongs to $C_b^1([0, T] \times \overline{Y})$ and the convergence of Ψ^k in W_1 is uniform in time and implies the narrow convergence. The last term on the right-hand side of (31) tends to 0, since it holds

$$\int_0^t |\theta_k(\phi(\tau, \cdot, \cdot))|\,\mathrm{d}\tau \leq CT\tau_k \|\phi\|_{C_b^1([0,T]\times\overline{Y})}\,.$$

We conclude by estimating

$$\begin{aligned}
&\left| \int_0^t \int_{\overline{Y}} \nabla\phi(\tau, x, \lambda) \cdot b_{\Psi^k(\tau)}(x, \lambda) \mathrm{d}\Psi^k(\tau)(x, \lambda)\mathrm{d}\tau \right. \\
&\qquad \left. - \int_0^t \int_{\overline{Y}} \nabla\phi(\tau, x, \lambda) \cdot b_{\overline{\Psi}(\tau)}(x, \lambda) \mathrm{d}\overline{\Psi}(\tau)(x, \lambda)\mathrm{d}\tau \right| \\
&\leq \|\phi\|_{C_b^1} \int_0^t \int_{\overline{Y}} \|b_{\Psi^k(\tau)}(x, \lambda) - b_{\overline{\Psi}(\tau)}(x, \lambda)\|_{\overline{Y}}\, \mathrm{d}\Psi^k(\tau)(x, \lambda)\,\mathrm{d}\tau \\
&\qquad + \int_0^t \left| \int_{\overline{Y}} \nabla\phi(\tau, x, \lambda) \cdot b_{\overline{\Psi}(\tau)}(x, \lambda)\,\mathrm{d}(\Psi^k(\tau) - \overline{\Psi}(\tau))(x, \lambda) \right| \mathrm{d}\tau\,.
\end{aligned} \tag{33}$$

By [31, Proposition 3.2], Lemma 1, and Assumptions (v_2) and $(\mathcal{T}_2)$, the first term on the right-hand side of (33) can be estimated by

$$\|\phi\|_{C_b^1}(L_{R,v} + L_{R,\mathcal{T}}) \int_0^t W_1(\Psi^k(\tau), \overline{\Psi}(\tau))\,\mathrm{d}\tau \to 0$$

as $k \to \infty$ uniformly with respect to $t \in [0, T]$.

As for the second term, we first notice that, by [31, Proposition 3.2] and Lemma 1, the function $(x, \lambda) \mapsto b_{\overline{\Psi}(\tau)}(x, \lambda)$ is continuous from $\overline{Y}$ to $\overline{Y}$ and is bounded on B_R^Y. Since $\Psi^k(\tau)$ converges narrowly to $\overline{\Psi}(\tau)$, for $\tau \in [0, t]$ we have

$$\lim_{k\to\infty} \left| \int_{\overline{Y}} \nabla\phi(\tau, x, \lambda) \cdot b_{\overline{\Psi}(\tau)}(x, \lambda)\,\mathrm{d}(\Psi^k(\tau) - \overline{\Psi}(\tau))(x, \lambda) \right| = 0\,.$$

Furthermore, by (v_3) and $(\mathcal{T}_1)$ we have the uniform bound

$$\begin{aligned}\left|\int_{\overline{Y}} \nabla\phi(\tau,x,\lambda)\cdot b_{\overline{\Psi}(\tau)}(x,\lambda)\,\mathrm{d}(\Psi^k(\tau)-\overline{\Psi}(\tau))(x,\lambda)\right| \\ \leq 4\|\phi\|_{C_b^1}(M_v+M_{\mathcal{T}})(1+R)\|\Psi^k(\tau)-\overline{\Psi}(\tau)\|_{\mathrm{TV}} \\ \leq 3\|\phi\|_{C_b^1}(M_v+M_{\mathcal{T}})(1+R)\end{aligned}$$

for $\tau\in[0,t]$. Thus, by dominated convergence also the second term on the right-hand side of (33) tends to zero as $k\to\infty$.

Eventually, we infer that passing to the limit $k\to\infty$ in (31) we get the equality

$$\begin{aligned}\int_Y \phi(t,x,\lambda)\,\mathrm{d}\overline{\Psi}(t)(x,\lambda)-\int_Y \phi(0,x,\lambda)\,\mathrm{d}\widehat{\Psi}(x,\lambda) \\ =\int_0^t\int_Y \partial_t\phi(\tau,x,\lambda)\,\mathrm{d}\overline{\Psi}(\tau)(x,\lambda) \\ +\int_0^t\int_Y \nabla\phi(\tau,x,\lambda)\cdot b_{\overline{\Psi}(\tau)}(x,\lambda)\,\mathrm{d}\overline{\Psi}(\tau)(x,\lambda)\end{aligned}$$

for every $\phi\in C_b^1([0,T]\times\overline{Y})$ and every $t\in[0,T]$. This concludes the proof of the theorem. □

4. An implicit–explicit scheme for the inhomogeneous replicator dynamics

We discuss in this section a different discrete-time approximation of the continuity equation (7) for the operator $\mathcal{T}_\Psi: Y\to\mathcal{F}(U)$ corresponding to the transition operators considered in [5] (see also [31, Section 5]) for the replicator equation (first introduced in [37]; see also [21,41]), namely

$$\begin{aligned}\mathcal{T}_\Psi(x,\lambda) := \Bigg(\int_{\overline{Y}}\int_U J(x,u,x',u')\,\mathrm{d}\lambda'(u')\,\mathrm{d}\Psi(x',\lambda') \\ -\int_U\int_{\overline{Y}}\int_U J(x,u,x',u')\,\mathrm{d}\lambda'(u')\,\mathrm{d}\Psi(x',\lambda')\,\mathrm{d}\lambda(u)\Bigg)\lambda\end{aligned} \tag{34}$$

defined for every $\Psi\in\mathcal{P}_1(Y)$ and every $y=(x,\lambda)\in Y$. In (34) we consider a function $J:(\mathbb{R}^d\times U)^2\to\mathbb{R}$ such that

(J_1) J is locally Lipschitz continuous with respect to all of its variables;

(J_2) there exists $M_J>0$ such that for every $(x,u,x',u')\in(\mathbb{R}^d\times U)^2$

$$|J(x,u,x',u')|\leq M_J(1+|x|+|x'|)\,.$$

For simplicity of notation, from now on we will write

$$\begin{aligned}(J*\Psi)(x,u) &:= \int_{\overline{Y}}\int_U J(x,u,x',u')\,\mathrm{d}\lambda'(u')\,\mathrm{d}\Psi(x',\lambda')\,, \\ \langle J*\Psi,\lambda\rangle(x) &:= \int_U (J*\Psi)(x,u)\,\mathrm{d}\lambda(u)\,,\end{aligned}$$

so that (34) can be written as

$$\mathcal{T}_\Psi(x,\lambda) = \big((J*\Psi)(x,\cdot) - \langle J*\Psi,\lambda\rangle(x)\big)\lambda\,. \tag{35}$$

The following proposition holds.

Proposition 2. [31, Proposition 5.8] *Under the assumptions* (J_1)–(J_2)*, the operator* $\mathcal{T}_\Psi$ *defined in (34) satisfies the conditions* $(\mathcal{T}_0)$–$(\mathcal{T}_3)$.

We now introduce the *spherical Hellinger distance* between probability measures

$$\begin{aligned}\mathrm{HS}^2(\lambda_1,\lambda_2) := \inf\Big\{ &\frac{1}{4}\int_0^1\int_U |w_t(u)|^2\,\mathrm{d}\rho_t(u)\,\mathrm{d}t \,:\, \rho\in C([0,1];\mathcal{P}(U)),\\ &w\in L^2([0,1]\times U;\mathrm{d}\rho),\ \dot\rho_t = \Big(w_t - \int_U w_t\,\mathrm{d}\rho_t\Big)\rho_t,\\ &\rho_0=\lambda_1,\ \rho_1=\lambda_2\Big\},\end{aligned}$$

defined for every $\lambda_1,\lambda_2\in\mathcal{P}(U)$, where $L^2([0,1]\times U;\mathrm{d}\rho)$ denotes the space of functions which are square-integrable with respect to the measure ρ. For later use, we also define the *Hellinger distance* between nonnegative measures: For every $\mu_1,\mu_2\in\mathcal{M}_+(U)$, we set

$$\begin{aligned}\mathrm{H}^2(\mu_1,\mu_2) &:= \inf\Big\{\frac{1}{4}\int_0^1\int_U |w_t(u)|^2\,\mathrm{d}\rho_t(u)\,\mathrm{d}t : \rho\in C([0,1];\mathcal{M}_+(U)),\\ &\qquad w\in L^2([0,1]\times U;\mathrm{d}\rho),\ \dot\rho_t = w_t\,\rho_t,\ \rho_0=\mu_1,\ \rho_1=\mu_2\Big\}\\ &= \int_U\left[\left(\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu^*}\right)^{\frac12} - \left(\frac{\mathrm{d}\mu_2}{\mathrm{d}\mu^*}\right)^{\frac12}\right]^2\mathrm{d}\mu^*,\end{aligned}$$

where $\mu^*\in\mathcal{M}_+(U)$ is such that $\mu_1,\ \mu_2\ll\mu^*$. We notice that HS^2 can be expressed in terms of H^2 through

$$\mathrm{HS}^2(\lambda_1,\lambda_2) = \arccos\left(1-\frac{\mathrm{H}^2(\lambda_1,\lambda_2)^2}{2}\right)\qquad \text{for every } \lambda_1,\lambda_2\in\mathcal{P}(U)\,, \tag{36}$$

and that the following chain of inequalities holds:

$$\begin{aligned}&\|\lambda_1-\lambda_2\|_{\mathrm{BL}} \le \|\lambda_1-\lambda_2\|_{\mathrm{TV}}\\ &\quad\le 2\,\mathrm{H}(\lambda_1,\lambda_2)\le 2\,\mathrm{HS}(\lambda_1,\lambda_2)\end{aligned}\qquad \text{for every } \lambda_1,\lambda_2\in\mathcal{P}(U)\,. \tag{37}$$

In the spatially homogeneous case, the replicator equation is a *generalized minimizing movement* [6] for the functional

$$\mathcal{J}_{\mathrm{hom}}(\lambda) := -\frac18\int_U\int_U J(u,u')\,\mathrm{d}\lambda(u)\,\mathrm{d}\lambda(u')$$

with respect to the spherical Hellinger distance. In the spatially inhomogeneous setting, the payoff functional has a bilinear dependence on Ψ and λ: For every $\Psi \in \mathcal{P}_1(Y)$ and every $(x, \lambda) \in Y$, we set

$$\mathcal{J}_\Psi(x, \lambda) := -\frac{1}{4} \langle (J * \Psi), \lambda \rangle , \tag{38}$$

(the factor $\frac{1}{4}$ instead of $\frac{1}{8}$ is due to the dependence on λ which is now linear). We modify the scheme in Sect. 3 by replacing the finite difference (8) with a minimizing movement. Namely, in the interval $[t_i^k, t_{i+1}^k)$ let $\Psi_i^k \in \mathcal{P}(U)$ be given and define, for every $(\hat{x}, \hat{\lambda}) \in Y$,

$$\lambda_{(\hat{x},\hat{\lambda}),i+1} := \operatorname{argmin} \left\{ \mathcal{J}_{\Psi_i^k}(\hat{x}, \lambda) + \frac{1}{2\tau_k} \mathrm{HS}^2(\lambda, \hat{\lambda}) : \lambda \in \mathcal{P}(U) \right\}. \tag{39}$$

Notice that the measure $\lambda_{(\hat{x},\hat{\lambda}),i+1} \in \mathcal{P}(U)$ is well defined, as $\mathcal{P}(U)$ is compact and the functional in (39) is strictly convex. Therefore, we can define $\lambda^k_{(\hat{x},\hat{\lambda}),i+1}$, Λ^k_{i+1}, and $\tilde{\Psi}^k_{i+1}$ exactly as in (9), (10), and (11), respectively. The second step (12) in the space variable remains instead the same, so that $x^k_{(\hat{x},\hat{\lambda}),i+1}$, X^k_{i+1}, Ψ^k_{i+1} are as in (13), (14), and (15), respectively. We further refer to (15), (16), and (17) for the definition of the interpolation curves Ψ^k, $\widetilde{\Psi}^k$, and $\underline{\Psi}^k$.

The next lemma gives an estimate on the size of the support of Ψ^k_{i+1} and $\widetilde{\Psi}^k_{i+1}$, showing that they belong to $\mathcal{P}_c(Y) \subseteq \mathcal{P}_1(Y)$ for every $k \in \mathbb{N}$ and every $i \in \{0, \dots, k-1\}$.

Lemma 2. *Let $\widehat{\Psi} \in \mathcal{P}_c(Y)$. Then, there exists $R > 0$ such that, for every $k \in \mathbb{N}$ and every $t \in [0, T]$, $\Psi^k(t)$, $\underline{\Psi}^k(t)$, $\widetilde{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_R)$.*

Proof. Let us define the piecewise constant interpolation functions $\underline{\lambda}^k$, $\overline{\lambda}^k$, and $\underline{x}^k$ as in (18), and let x^k and λ^k be the corresponding piecewise affine interpolations. Then, by (39) we have that $\underline{\lambda}^k(t, x_0, \lambda_0)$, $\overline{\lambda}^k(t, x_0, \lambda_0) \in \mathcal{P}(U)$ for every $t \in [0, T]$ and every $(x_0, \lambda_0) \in \mathrm{spt}\widehat{\Psi}$, so that

$$\|\underline{\lambda}^k(t, x_0, \lambda_0)\|_{\mathrm{BL}}, \|\overline{\lambda}^k(t, x_0, \lambda_0)\|_{\mathrm{BL}} \leq 1 .$$

Following step by step the proof of (19) and (22), we also deduce that there exists $R > 0$ such that

$$|x_k(t, x_0, \lambda_0)| \leq R \qquad \begin{array}{l} \text{for every } t \in [0, T], \text{ every } (x_0, \lambda_0) \in \mathrm{spt}\,\widehat{\Psi}, \\ \text{and every } k \in \mathbb{N} . \end{array} \tag{40}$$

We notice that, being the step (39) defined through a minimization in $\mathcal{P}(U)$ and not through a finite difference, the estimate (40) holds for every k, and not only for k large. Moreover, (40) yields that $\Psi^k(t)$, $\underline{\Psi}^k(t)$, $\widetilde{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_R)$. □

In order to write the equivalent of Proposition 1, we have to determine an approximate Euler–Lagrange equation for the minimization problem (39). This is the content of the following proposition, written here for generic Ψ, x, and λ.

Proposition 3. *Let $R > 0$. Assume that $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$, $(x, \lambda) \in \mathrm{B}_R^Y$, and let $\tilde{\lambda} \in \mathcal{P}(U)$ be the solution to*

$$\min \left\{ \mathcal{J}_\Psi(x, \rho) + \frac{1}{2\tau_k} \mathrm{HS}^2(\rho, \lambda) : \rho \in \mathcal{P}(U) \right\}. \tag{41}$$

Then, there exists a constant $C = C(R) > 0$ such that

$$\mathrm{HS}(\tilde{\lambda}, \lambda) \leq C\tau_k, \tag{42}$$

$$\left\| \frac{\tilde{\lambda} - \lambda}{\tau_k} - \mathcal{T}_\Psi(x, \tilde{\lambda}) \right\|_{\mathrm{BL}} \leq C\tau_k(1 + \tau_k). \tag{43}$$

Proof. Inequality (42) follows from the minimality of $\tilde{\lambda}$. Indeed, we have that

$$\frac{1}{2\tau_k} \mathrm{HS}^2(\tilde{\lambda}, \lambda) \leq \left| \mathcal{J}_\Psi(x, \lambda) - \mathcal{J}_\Psi(x, \tilde{\lambda}) \right|. \tag{44}$$

By definition (38) of $\mathcal{J}_\Psi$, by (J_1), by the assumptions $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$, $(x, \lambda) \in \mathrm{B}_R^Y$, and by (37), we continue in (44) with

$$\frac{1}{2\tau_k} \mathrm{HS}^2(\tilde{\lambda}, \lambda) \leq \frac{M_J}{2}(1 + R)\|\tilde{\lambda} - \lambda\|_{\mathrm{TV}} \leq M_J(1 + R)\mathrm{HS}(\tilde{\lambda}, \lambda). \tag{45}$$

From (45), we deduce (42).

In order to prove (43), we write explicitly the Euler–Lagrange equation of (41). Here, we follow the lines of [20, Section 4]. For every $\varphi \in \mathrm{Lip}(U)$ with $\|\varphi\|_{\mathrm{Lip}} \leq 1$, we consider the auxiliary system

$$\begin{cases} \partial_\varepsilon \lambda_\varepsilon = (\varphi - \langle \varphi, \lambda_\varepsilon \rangle)\lambda_\varepsilon, \\ \lambda_0 = \tilde{\lambda}. \end{cases} \tag{46}$$

In view of [12, Section I.3, Theorem 1.4, Corollary 1.1], the ODE system (46) admits a unique solution $\lambda_\varepsilon^\varphi \in \mathcal{P}(U)$ for $\varepsilon > 0$. Moreover, if $\tilde{\lambda} \ll \mu$, it is easy to check that $\lambda_\varepsilon^\varphi \ll \mu$ for $\varepsilon > 0$. In the sequel, we fix $\mu^* \in \mathcal{P}(U)$ such that $\lambda, \tilde{\lambda} \ll \mu^*$.

Given $\lambda_\varepsilon^\varphi$, the Euler–Lagrange equation of (41) reads

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \mathcal{J}_\Psi(x, \lambda_\varepsilon^\varphi) + \frac{1}{2\tau_k} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \mathrm{HS}^2(\lambda_\varepsilon^\varphi, \lambda) = 0. \tag{47}$$

We compute the two derivatives appearing in (47) separately. In view of (46), we have that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \mathcal{J}_\Psi(x, \lambda_\varepsilon^\varphi) &= -\frac{1}{4} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \langle (J * \Psi), \lambda_\varepsilon^\varphi \rangle = -\frac{1}{4} \big\langle (J * \Psi), \big(\varphi - \langle \varphi, \tilde{\lambda} \rangle\big)\tilde{\lambda} \big\rangle \\ &= -\frac{1}{4} \big\langle \big((J * \Psi) - \langle (J * \Psi), \tilde{\lambda} \rangle\big)\tilde{\lambda}, \varphi \big\rangle = -\frac{1}{4} \big\langle \mathcal{T}_\Psi(x, \tilde{\lambda}), \varphi \big\rangle, \end{aligned} \tag{48}$$

where, in the last equality, we have used (35).

To compute the second term on the left-hand side of (47), we first notice that, since $\tilde{\lambda}, \lambda, \lambda_\varepsilon^\varphi \ll \mu^*$, we can write

$$\mathrm{HS}^2(\lambda_\varepsilon^\varphi, \lambda) = \arccos\left(1 - \frac{\mathrm{H}^2(\lambda_\varepsilon^\varphi, \lambda)^2}{2}\right),$$
$$\mathrm{H}^2(\lambda_\varepsilon^\varphi, \lambda) = \int_U \left[\left(\frac{\mathrm{d}\lambda_\varepsilon^\varphi}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} - \left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}}\right]^2 \mathrm{d}\mu^*.$$

Defining $\delta_k(\tilde{\lambda}, \lambda) \in [0, 1]$ such that $1 - \delta_k(\tilde{\lambda}, \lambda) = \frac{1}{\sqrt{1 - \frac{\mathrm{H}^2(\tilde{\lambda}, \lambda)^2}{4}}}$, we have that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \mathrm{HS}^2(\lambda_\varepsilon^\varphi, \lambda) = \big(1 - \delta_k(\tilde{\lambda}, \lambda)\big) \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \mathrm{H}^2(\lambda_\varepsilon^\varphi, \lambda) \\
&= 2\big(1 - \delta_k(\tilde{\lambda}, \lambda)\big) \int_U \left[\left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} - \left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}}\right] \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \left(\frac{\mathrm{d}\lambda_\varepsilon^\varphi}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} \mathrm{d}\mu^* \\
&= \big(1 - \delta_k(\tilde{\lambda}, \lambda)\big) \int_U \left[\left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} - \left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}}\right] \left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{-\frac{1}{2}} \big(\varphi - \langle \varphi, \tilde{\lambda}\rangle\big)\, \mathrm{d}\tilde{\lambda} \\
&= \big(1 - \delta_k(\tilde{\lambda}, \lambda)\big) \left\langle \big(\varphi - \langle \varphi, \tilde{\lambda}\rangle\big)\mu^*, \left[\left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} - \left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}}\right] \left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} \right\rangle.
\end{aligned} \tag{49}$$

Using the algebraic equality $2(a - b)a = a^2 - b^2 + (a - b)^2$, we continue in (49) with

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} \mathrm{HS}^2(\lambda_\varepsilon^\varphi, \lambda) &= \frac{\big(1 - \delta_k(\tilde{\lambda}, \lambda)\big)}{2} \left\langle \big(\varphi - \langle \varphi, \tilde{\lambda}\rangle\big)\mu^*, \left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*} - \frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right) \right\rangle \\
&\quad + \frac{\big(1 - \delta_k(\tilde{\lambda}, \lambda)\big)}{2} \left\langle \big(\varphi - \langle \varphi, \tilde{\lambda}\rangle\big)\mu^*, \left[\left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} - \left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}}\right]^2 \right\rangle \\
&= \frac{\big(1 - \delta_k(\tilde{\lambda}, \lambda)\big)}{2} \left\langle \tilde{\lambda} - \lambda, \big(\varphi - \langle \varphi, \tilde{\lambda}\rangle\big) \right\rangle \\
&\quad + \frac{\big(1 - \delta_k(\tilde{\lambda}, \lambda)\big)}{2} \left\langle \big(\varphi - \langle \varphi, \tilde{\lambda}\rangle\big)\mu^*, \left[\left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} - \left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}}\right]^2 \right\rangle \\
&= \frac{\big(1 - \delta_k(\tilde{\lambda}, \lambda)\big)}{2} \langle \tilde{\lambda} - \lambda, \varphi \rangle \\
&\quad + \frac{\big(1 - \delta_k(\tilde{\lambda}, \lambda)\big)}{2} \left\langle \big(\varphi - \langle \varphi, \tilde{\lambda}\rangle\big)\mu^*, \left[\left(\frac{\mathrm{d}\tilde{\lambda}}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}} - \left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac{1}{2}}\right]^2 \right\rangle,
\end{aligned} \tag{50}$$

where, in the last equality, we have used the fact that $\tilde{\lambda}, \lambda \in \mathcal{P}(U)$.

In order to conclude with (43), we estimate $\delta_k(\tilde{\lambda}, \lambda)$ and the last term on the right-hand side of (50). In view of (36), (37), and (42), it is easy to check that

$$\delta_k(\tilde{\lambda}, \lambda) \le c\tau_k^2 \tag{51}$$

for some positive constant $c = c(R) > 0$. Since $\|\varphi\|_{\mathrm{Lip}} \le 1$ and (45) holds, we have that

$$\begin{aligned}&\left|\frac{(1-\delta_k(\tilde\lambda,\lambda))}{2}\left\langle\big(\varphi-\langle\varphi,\tilde\lambda\rangle\big)\mu^*,\left[\left(\frac{\mathrm{d}\tilde\lambda}{\mathrm{d}\mu^*}\right)^{\frac12}-\left(\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^*}\right)^{\frac12}\right]^2\right\rangle\right| \\ &\le \mathrm{H}^2(\tilde\lambda,\lambda)\le 4M_J^2(1+R)^2\tau_k^2\,.\end{aligned} \tag{52}$$

Combining (37), (42), and (47)–(52), we deduce that

$$\left\|\frac{\tilde\lambda-\lambda}{\tau_k}-\mathcal{T}_\Psi(x,\tilde\lambda)\right\|_{\mathrm{BL}}\le\delta_k(\tilde\lambda,\lambda)\left\|\frac{\tilde\lambda-\lambda}{\tau_k}\right\|_{\mathrm{BL}}+8M_J^2(1+R)^2\tau_k\le C\tau_k(1+\tau_k)\,,$$

for some positive constant $C = C(R)$. This concludes the proof of the proposition. □

We are now in a position to state the equivalent of Proposition 1.

Proposition 4. *There exists $C > 0$ such that for every $\varphi \in C_b^1(\mathbb{R}^d \times \mathcal{F}(U))$, every $k \in \mathbb{N}$, every $i \in \{0,\dots,k-1\}$, and every $t \in (t_i^k, t_{i+1}^k)$,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_Y\varphi(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)=\int_Y\nabla\varphi(x,\lambda)\cdot b_{\Psi^k(t)}(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)+\vartheta_k(\varphi), \tag{53}$$

where $|\vartheta_k(\varphi)| \le C\|\varphi\|_{C_b^1}\tau_k$.

Proof. Along the proof, we denote by C a generic positive constant independent of i, k, t, and φ, that may vary from line to line.

We follow step by step the proof of Proposition 1. For every test function $\varphi \in C_b^1(\mathbb{R}^d \times \mathcal{F}(U))$ and every $t \in (t_i^k, t_{i+1}^k)$, by definition of $\Psi^k(t)$ we have that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_Y\varphi(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda)&=\frac{\mathrm{d}}{\mathrm{d}t}\int_Y\varphi\big(X_{i+1}^k(t,x,\lambda),\Lambda_{i+1}^k(t,x,\lambda)\big)\,\mathrm{d}\Psi_i^k(x,\lambda)\\&=\int_Y\nabla_x\varphi\big(X_{i+1}^k(t,x,\lambda),\Lambda_{i+1}^k(t,x,\lambda)\big)\cdot v_{\widetilde\Psi^k(t)}\big(x,\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)\big)\,\mathrm{d}\Psi_i^k(x,\lambda)\\&\quad+\int_Y\nabla_\lambda\varphi\big(X_{i+1}^k(t,x,\lambda),\Lambda_{i+1}^k(t,x,\lambda)\big)\cdot\dot\Lambda_{i+1}^k(t,x,\lambda)\,\mathrm{d}\Psi_i^k(x,\lambda)\\&=\int_Y\nabla_x\varphi\big(X_{i+1}^k(t,x,\lambda),\Lambda_{i+1}^k(t,x,\lambda)\big)\cdot v_{\widetilde\Psi^k(t)}\big(x,\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)\big)\,\mathrm{d}\Psi_i^k(x,\lambda)\\&\quad+\int_Y\nabla_\lambda\varphi\big(X_{i+1}^k(t,x,\lambda),\Lambda_{i+1}^k(t,x,\lambda)\big)\cdot\frac{\big(\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)-\lambda\big)}{\tau_k}\,\mathrm{d}\Psi_i^k(x,\lambda)\,.\end{aligned} \tag{54}$$

In order to deduce (53) from (54), we need to estimate

$$\begin{aligned}I_1(x,\lambda)&:=\left|v_{\widetilde\Psi^k(t)}\big(x,\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)\big)-v_{\Psi^k(t)}\big(X_{i+1}^k(t,x,\lambda),\Lambda_{i+1}^k(t,x,\lambda)\big)\right|,\\I_2(x,\lambda)&:=\left\|\frac{\big(\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)-\lambda\big)}{\tau_k}-\mathcal{T}_{\Psi^k(t)}\big(X_{i+1}^k(t,x,\lambda),\Lambda_{i+1}^k(t,x,\lambda)\big)\right\|_{\mathrm{BL}}\end{aligned}$$

for $(x,\lambda)\in \operatorname{spt}\Psi_i^k \subseteq \mathrm{B}_R^Y$, where R has been determined in Lemma 2.

Let us start with I_1. By triangle inequality we have

$$\begin{aligned} I_1(x,\lambda) \;\le\; & \Big| v_{\widetilde{\Psi}^k(t)}\big(x,\, \Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)\big) - v_{\widetilde{\Psi}^k(t)}\big(X_{i+1}^k(t,x,\lambda),\, \Lambda_{i+1}^k(t,x,\lambda)\big)\Big| \\ & + \Big| v_{\widetilde{\Psi}^k(t)}\big(X_{i+1}^k(t,x,\lambda),\, \Lambda_{i+1}^k(t,x,\lambda)\big) \\ & - v_{\Psi^k(t)}\big(X_{i+1}^k(t,x,\lambda),\, \Lambda_{i+1}^k(t,x,\lambda)\big)\Big| \\ =:\; & I_{1,1}(x,\lambda) + I_{1,2}(x,\lambda). \end{aligned} \tag{55}$$

Since $\widetilde{\Psi}^k(t)\in\mathcal{P}(\mathrm{B}_R^Y)$, hypothesis (v_1) implies that

$$\begin{aligned} I_{1,1}(x,\lambda) &\le L_{v,R}\big(|X_{i+1}^k(t,x,\lambda)-x| + \|\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda) - \Lambda_{i+1}^k(t,x,\lambda)\|_{\mathrm{BL}}\big) \\ &\le L_{v,R}\bigg(\int_{t_i^k}^t \big| v_{\widetilde{\Psi}^k(\tau)}\big(x,\, \Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)\big)\big|\,\mathrm{d}\tau \\ &\quad + \int_t^{t_{i+1}^k} \left\| \frac{\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)-\lambda}{\tau_k}\right\|_{\mathrm{BL}} \mathrm{d}\tau\bigg). \end{aligned}$$

By (v_3), Lemma 2, and Proposition 3, we can continue with

$$\begin{aligned} I_{1,1}(x,\lambda) \le L_{v,R}\bigg(& M_v \int_{t_i^k}^t \big(1+|x|+\|\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)\|_{\mathrm{BL}} + m_1(\widetilde{\Psi}^k(\tau))\big)\,\mathrm{d}\tau \\ & + \frac{2}{\tau_k}\int_t^{t_{i+1}^k} \mathrm{HS}\big(\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda),\lambda\big)\,\mathrm{d}\tau\bigg) \le C\tau_k\,. \end{aligned} \tag{56}$$

As for $I_{1,2}$, thanks to assumption (v_2) and to Lemma 2 we get

$$\begin{aligned} I_{1,2}(x,\lambda) &\le L_{v,R} W_1(\widetilde{\Psi}^k(t),\Psi^k(t)) \\ &= L_{v,R} \sup_{\eta\in \mathrm{Lip}_1(Y)} \bigg\{ \int_Y \eta(x',\lambda')\,\mathrm{d}(\widetilde{\Psi}^k(t)-\Psi^k(t))(x',\lambda')\bigg\} \\ &= L_{v,R} \sup_{\eta\in \mathrm{Lip}_1(Y)} \bigg\{ \int_Y \big[\eta(x,\Lambda_{i+1}^k(t_{i+1}^k,x',\lambda')) \\ &\quad - \eta(X_{i+1}^k(t,x',\lambda'),\Lambda_{i+1}^k(t,x',\lambda'))\big]\,\mathrm{d}\Psi_i^k(x',\lambda')\bigg\} \\ &\le L_{v,R}\int_Y \Big(|x-X_{i+1}^k(t,x',\lambda')| \\ &\quad + \|\Lambda_{i+1}^k(t_{i+1}^k,x',\lambda')) - \Lambda_{i+1}^k(t,x',\lambda')\|_{\mathrm{BL}}\Big)\,\mathrm{d}\Psi_i^k(x',\lambda') \\ &\le L_{v,R}\int_Y \bigg(\int_{t_i^k}^t \big| v_{\widetilde{\Psi}^k(\tau)}(x,\Lambda_{i+1}^k(t_{i+1}^k,x',\lambda'))\big|\,\mathrm{d}\tau \\ &\quad + \int_t^{t_{i+1}^k} \left\| \frac{\Lambda_{i+1}^k(t_{i+1}^k,x',\lambda')-\lambda'}{\tau_k}\right\|_{\mathrm{BL}}\mathrm{d}\tau\bigg)\,\mathrm{d}\Psi_i^k(x',\lambda') \end{aligned}$$

$$\leq L_{v,R}\,\tau_k \int_Y \left(\left| v_{\widetilde{\Psi}^k(\tau)}(x, \Lambda^k_{i+1}(t^k_{i+1}, x', \lambda')) \right| + \left\| \frac{\Lambda^k_{i+1}(t^k_{i+1}, x', \lambda') - \lambda}{\tau_k} \right\|_{\mathrm{BL}} \right) \mathrm{d}\Psi^k_i(x', \lambda')\,.$$

Arguing as in (56), we infer that

$$I_{1,2}(x,\lambda) \leq C\,\tau_k \qquad \text{for every } (x,\lambda) \in \operatorname{spt}\Psi^k_i\,. \tag{57}$$

Combining (55)–(57), we get

$$I_1(x,\lambda) \leq C\,\tau_k \qquad \text{for every } (x,\lambda) \in \operatorname{spt}\Psi^k_i\,. \tag{58}$$

Let us now estimate I_2. By triangle inequality, we have

$$\begin{aligned} I_2(x,\lambda) &\leq \left\| \frac{\big(\Lambda^k_{i+1}(t^k_{i+1}, x, \lambda) - \lambda\big)}{\tau_k} - \mathcal{T}_{\Psi^k_i}(x, \Lambda^k_{i+1}(t^k_{i+1}, x, \lambda)) \right\|_{\mathrm{BL}} \\ &\quad + \left\| \mathcal{T}_{\Psi^k(t)}\big(X^k_{i+1}(t, x, \lambda), \Lambda^k_{i+1}(t, x, \lambda)\big) - \mathcal{T}_{\Psi^k_i}\big(x, \Lambda^k_{i+1}(t^k_{i+1}, x, \lambda)\big) \right\|_{\mathrm{BL}} \\ &=: I_{2,1}(x,\lambda) + I_{2,2}(x,\lambda)\,. \end{aligned} \tag{59}$$

By Proposition 3, we have that

$$I_{2,1}(x,\lambda) \leq C\,\tau_k \qquad \text{for every } (x,\lambda) \in \operatorname{spt}\Psi^k_i\,. \tag{60}$$

By $(\mathcal{T}_2)$, (v_3), Lemma 2, and Proposition 3, and repeating the arguments of (57), we get

$$\begin{aligned} I_{2,2}(x,\lambda) &\leq L_{\mathcal{T},R}\bigg(\int_{t^k_i}^{t} \left| v_{\widetilde{\Psi}^k(\tau)}\big(x, \Lambda^k_{i+1}(t^k_{i+1}, x, \lambda)\big) \right| \mathrm{d}\tau \\ &\quad + \int_t^{t^k_{i+1}} \left\| \frac{\Lambda^k_{i+1}(t^k_{i+1}, x, \lambda) - \lambda}{\tau_k} \right\|_{\mathrm{BL}} \mathrm{d}\tau + W_1(\Psi^k(t), \Psi^k_i) \bigg) \\ &\leq L_{\mathcal{T},R}\bigg(\int_{t^k_i}^{t} \left| v_{\widetilde{\Psi}^k(\tau)}\big(x, \Lambda^k_{i+1}(t^k_{i+1}, x, \lambda)\big) \right| \mathrm{d}\tau \\ &\quad + \int_t^{t^k_{i+1}} \left\| \frac{\Lambda^k_{i+1}(t^k_{i+1}, x, \lambda) - \lambda}{\tau_k} \right\|_{\mathrm{BL}} \mathrm{d}\tau \\ &\quad + \int_Y \int_{t^k_i}^{t} \bigg(\left| v_{\widetilde{\Psi}^k(\tau)}\big(x', \Lambda^k_{i+1}(t^k_{i+1}, x', \lambda')\big) \right| \\ &\quad + \left\| \frac{\Lambda^k_{i+1}(t^k_{i+1}, x', \lambda') - \lambda'}{\tau_k} \right\|_{\mathrm{BL}} \bigg) \mathrm{d}\tau\, \mathrm{d}\Psi^k_i(x', \lambda') \bigg) \\ &\leq 4 L_{\mathcal{T},R}\, M_v (1+R) \tau_k + 4\,\mathrm{HS}\big(\Lambda^k_{i+1}(t^k_{i+1}, x, \lambda), \lambda\big) \leq C\,\tau_k\,. \end{aligned} \tag{61}$$

Combining (59)–(61), we obtain that

$$I_2(x,\lambda) \leq C\,\tau_k \qquad \text{for every } (x,\lambda) \in \operatorname{spt}\Psi^k_i\,. \tag{62}$$

Equality (53) follows from (58) and (62) as in the proof of Proposition 1. □

Finally, we prove the convergence of the sequence Ψ^k to the solution $\Psi \in C([0,T];\mathcal{P}_1(Y))$ of the continuity equation (7).

Theorem 2. *Let $\widehat{\Psi} \in \mathcal{P}_c(Y)$. Then, the sequence of curves $\Psi^k\colon [0,T] \to \mathcal{P}_1(Y)$ converges to the unique solution $\Psi \in C([0,T];\mathcal{P}_1(Y))$ of (7) in W_1, uniformly with respect to $t \in [0,T]$.*

Proof. Since the operator $\mathcal{T}_\Psi$ defined in (34) satisfies the property $(\mathcal{T}_0)$–$(\mathcal{T}_3)$, we only have to check that the sequence Ψ^k is compact in $C([0,T];\mathcal{P}_1(Y))$. The rest of the proof works as for Theorem 1 using Proposition 4 instead of Proposition 1.

In view of Lemma 2, it is enough to show that Ψ^k is equi-Lipschitz with respect to W_1. Let us fix $k \in \mathbb{N}$, $i \in \{0,\dots,k-1\}$, and $s \le t \in [t_i^k, t_{i+1}^k]$. Then,

$$\begin{aligned}
W_1(\Psi^k(t),\Psi^k(s)) &= \sup\left\{\int_Y \eta(x,\lambda)\,\mathrm{d}(\Psi^k(t)-\Psi^k(s))(x,\lambda) :\ \eta\in \mathrm{Lip}_1(Y)\right\}\\
&\le \int_Y \Big(\big|X_{i+1}^k(t,x,\lambda)-X_{i+1}^k(s,x,\lambda)\big|\\
&\quad + \big\|\Lambda_{i+1}^k(t,x,\lambda)-\Lambda_{i+1}^k(s,x,\lambda)\big\|_{\mathrm{BL}}\Big)\,\mathrm{d}\Psi_i^k(x,\lambda)\\
&\le \int_Y \left(\int_s^t \big|v_{\widetilde{\Psi}^k(\tau)}\big(x,\Lambda_{i+1}^k(\tau,x,\lambda)\big)\big|\,\mathrm{d}\tau\right.\\
&\quad \left. + \int_s^t \left\|\frac{\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda)-\lambda}{\tau_k}\right\|_{\mathrm{BL}}\mathrm{d}\tau\right)\mathrm{d}\Psi_i^k(x,\lambda)\,.
\end{aligned}$$

Therefore, by (v_2), Lemma 2, and Proposition 3 we get

$$\begin{aligned}
W_1(\Psi^k(t),\Psi^k(s)) &\le 2M_v(1+R)|t-s|\\
&\quad + 2|t-s|\int_Y \frac{\mathrm{HS}\big((\Lambda_{i+1}^k(t_{i+1}^k,x,\lambda),\lambda\big)}{\tau_k}\,\mathrm{d}\Psi_i^k(x,\lambda)\\
&\le C|t-s|
\end{aligned}$$

for some positive constant C independent of k and t. □

5. An implicit–explicit scheme for reversible Markov chains

In this section, we show how to adapt the scheme developed in Sect. 4 to a reversible Markov chain on n states. In particular, we will prove the convergence of such scheme for short time.

For fixed $n \in \mathbb{N}$, we consider the set of strategies

$$\Lambda_n := \left\{\lambda=(\lambda_1,\dots,\lambda_n)\in\mathbb{R}^n :\ \lambda_h>0,\ \sum_{h=1}^n \lambda_h = 1\right\}.$$

In the notation of Sects. 3 and 4, the closure $\overline{\Lambda}_n$ can be identified with the set of probability measures $\mathcal{P}(U)$ for $U := \{e_h : h=1,\dots,n\}$, e_h being the elements

of the canonical basis of $\mathbb{R}^n$. Keeping the notation of the previous sections, we set $Y := \mathbb{R}^d \times \overline{\Lambda}_n$. Furthermore, we define

$$\Lambda_n^\delta := \{\lambda \in \Lambda_n : \lambda_h \geq \delta\} \quad \text{for every } \delta > 0\,, \qquad \mathbb{R}^n_0 := \left\{\xi \in \mathbb{R}^n : \sum_{h=1}^n \xi_h = 0\right\},$$

$$\mathrm{B}^Y_{R,\delta} := \mathrm{B}^Y_R \cap \left(\mathbb{R}^d \times \Lambda_n^\delta\right) \quad \text{for } \delta,\ R > 0\,.$$

A Markov chain is characterized by a matrix $\mathrm{Q} \in \mathbb{M}^n$, whose element $\mathrm{Q}_{h\ell} \geq 0$, $h \neq \ell$, indicates the rate of moving from the state ℓ to the state h. In our setting, we consider a more general map $\mathcal{Q}\colon \mathbb{R}^d \times \mathcal{P}_1(Y) \to \mathbb{M}^n$ satisfying the following properties:

($\mathcal{Q}_0$) for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$ and every $h, \ell = 1, \dots, n$, $\mathcal{Q}_{h\ell}(x, \Psi) \geq 0$ for $h \neq \ell$, and $\mathcal{Q}_{hh}(x, \Psi) = -\sum_{\ell \neq h} \mathcal{Q}_{\ell h}(x, \Psi)$;

($\mathcal{Q}_1$) for every $(x, \Psi) \in \mathbb{R}^d \times \mathcal{P}_1(Y)$, $\mathcal{Q}(x, \Psi)$ is *reversible*, that is, there exists a unique $\sigma = \sigma(x, \Psi) \in \Lambda_n$ such that

$$\mathcal{Q}_{h\ell}(x, \Psi)\sigma_\ell = \mathcal{Q}_{\ell h}(x, \Psi)\sigma_h \qquad \text{for every } h, \ell = 1, \dots, n\,;$$

($\mathcal{Q}_2$) $\mathcal{Q}$ is locally Lipschitz, that is, for every $R > 0$ there exists $L_{\mathcal{Q},R} > 0$ such that for every $x_1, x_2 \in \mathrm{B}_R$ and every $\Psi_1, \Psi_2 \in \mathcal{P}(\mathrm{B}^Y_R)$

$$|\mathcal{Q}(x_1, \Psi_1) - \mathcal{Q}(x_2, \Psi_2)| \leq L_{\mathcal{Q},R}\big(|x_1 - x_2| + W_1(\Psi_1, \Psi_2)\big)\,;$$

($\mathcal{Q}_3$) there exists $M_\mathcal{Q} > 0$ such that for every $x \in \mathbb{R}^d$ and every $\Psi \in \mathcal{P}_1(Y)$

$$|\mathcal{Q}(x, \Psi)| \leq M_\mathcal{Q}\big(1 + |x| + m_1(\Psi)\big)\,.$$

Remark 1. We remark that ($\mathcal{Q}_1$) is always satisfied, for instance, when $\mathcal{Q}(x, \Psi)$ is a tridiagonal matrix for every $x \in \mathbb{R}^d$ and $\Psi \in \mathcal{P}_1(Y)$, see, e.g., [30, Section 5.1].

Remark 2. We notice that if for every $y = (x, \lambda) \in Y$ and every $\Psi \in \mathcal{P}_1(Y)$ we set $\mathcal{T}_\Psi(y) := \mathcal{Q}(x, \Psi)\lambda$, then the operator $\mathcal{T}\colon Y \times \mathcal{P}_1(Y) \to \Lambda_n$ satisfies properties ($\mathcal{T}_0$)–($\mathcal{T}_3$) of Sect. 2.

Following [28,30], for every $y = (x, \lambda) \in \mathbb{R}^d \times \Lambda_n$ and every $\Psi \in \mathcal{P}_1(Y)$ we consider the *entropy* E and the *Onsager matrix* K

$$E(x, \lambda, \Psi) := \sum_{h=1}^n \lambda_h \ln\left(\frac{\lambda_h}{\sigma_h(x, \Psi)}\right), \tag{63}$$

$$K(x, \lambda, \Psi) := \sum_{\ell=2}^n \sum_{h=1}^{\ell-1} \mathcal{Q}_{h\ell}(x, \Psi)\sigma_\ell(x, \Psi)\, \Phi\left(\frac{\lambda_h}{\sigma_h(x, \Psi)}, \frac{\lambda_\ell}{\sigma_\ell(x, \Psi)}\right)(e_h - e_\ell) \otimes (e_h - e_\ell), \tag{64}$$

where $\Phi\colon [0, +\infty) \times [0, +\infty) \to [0, +\infty)$ is defined as

$$\Phi(a, b) := \frac{a - b}{\ln a - \ln b} \quad \text{for } a \neq b, \qquad \Phi(a, a) = a\,,$$

so that Φ is analytic. Clearly, $E(x,\cdot,\Psi)$ and $K(x,\cdot,\Psi)$ can be extended to $\overline{\Lambda}_n$ by continuity. Moreover, we notice that for every $(x,\lambda)\in\mathbb{R}^d\times\Lambda_n$ and every $\Psi\in\mathcal{P}_1(Y)$, the matrix $K(x,\lambda,\Psi)$ is symmetric and positive definite when acting on $\mathbb{R}^n_0$. We denote by $G(x,\lambda,\Psi)$ its inverse on $\mathbb{R}^n_0$. The matrix G is a Riemannian tensor on $\mathbb{R}^n_0$. For every $x\in\mathbb{R}^d$ and $\Psi\in\mathcal{P}_1(Y)$, we define the Riemannian metric $\mathsf{d}_{(x,\Psi)}\colon \Lambda_n\times\Lambda_n\to[0,+\infty)$ as

$$\mathsf{d}_{(x,\Psi)}(\lambda_1,\lambda_2) := \inf\left\{\int_0^1 \left\langle G(x,\rho(s),\Psi)\rho'(s),\rho'(s)\right\rangle^{\frac{1}{2}} \mathrm{d}s : \rho\in C^1([0,1];\Lambda_n),\ \rho(0)=\lambda_1,\ \rho(1)=\lambda_2\right\}, \tag{65}$$

for every $\lambda_1,\lambda_2\in\Lambda_n$. The metric $\mathsf{d}_{(x,\Psi)}$ can be extended to $\overline{\Lambda}_n\times\overline{\Lambda}_n$ in a continuous way.

In the next two lemmas, we collect some properties of E, K, G, and $\mathsf{d}_{(x,\Psi)}$.

Lemma 3. *Let $\delta, R>0$. Then, the following facts hold:*

(i) there exists a positive constant $\eta=\eta(R)$ such that $\sigma_h(x,\Psi)\geq\eta$ for every $x\in \mathrm{B}_R$, every $\Psi\in\mathcal{P}(\mathrm{B}_R^Y)$, and every $h=1,\dots,n$;

(ii) there exist two positive constants $c_1=c_1(R)$ and $c_2=c_2(R)$ such that for every $x\in\mathrm{B}_R$, every $\lambda\in\Lambda_n$, every $\Psi\in\mathcal{P}(\mathrm{B}_R^Y)$, and every $\mu\in\mathbb{R}^n_0$,

$$c_1|\mu|^2\leq\langle G(x,\lambda,\Psi)\mu,\mu\rangle\,, \tag{66}$$

$$\langle K(x,\lambda,\Psi)\mu,\mu\rangle\leq c_2|\mu|^2\,; \tag{67}$$

(iii) there exist two positive constants $c_3=c_3(\delta,R)$ and $c_4=c_4(\delta,R)$ such that for every $(x,\lambda)\in\mathrm{B}^Y_{R,\delta}$, every $\Psi\in\mathcal{P}(\mathrm{B}_R^Y)$, and every $\mu\in\mathbb{R}^n_0$,

$$\langle G(x,\lambda,\Psi)\mu,\mu\rangle\leq c_3|\mu|^2\,, \tag{68}$$

$$c_4|\mu|^2\leq\langle K(x,\lambda,\Psi)\mu,\mu\rangle\,; \tag{69}$$

(iv) $G(x,\cdot,\Psi)$ is Lipschitz continuous in Λ_n^δ, uniformly with respect to $x\in\mathrm{B}_R$ and $\Psi\in\mathcal{P}(\mathrm{B}_R^Y)$, that is, there exists $L_{G,\delta,R}>0$ such that for every $\lambda_1,\lambda_2\in\Lambda_n^\delta$

$$|G(x,\lambda_1,\Psi)-G(x,\lambda_2,\Psi)|\leq L_{G,\delta,R}|\lambda_1-\lambda_2|; \tag{70}$$

(v) $E(x,\cdot,\Psi)$ is Lipschitz continuous in Λ_n^δ, uniformly with respect to $x\in\mathrm{B}_R$ and $\Psi\in\mathcal{P}(\mathrm{B}_R^Y)$, namely there exists $L_{E,\delta,R}>0$ such that for every $\lambda_1,\lambda_2\in\Lambda_n^\delta$

$$|E(x,\lambda_1,\Psi)-E(x,\lambda_2,\Psi)|\leq L_{E,\delta,R}|\lambda_1-\lambda_2|; \tag{71}$$

(vi) for every $\alpha\in(0,1)$ the energy $E(x,\cdot,\Psi)$ is α-Hölder continuous in $\overline{\Lambda}_n$, uniformly with respect to $x\in\mathrm{B}_R$ and $\Psi\in\mathcal{P}(\mathrm{B}_R^Y)$, that is, for every $\alpha\in(0,1)$ there exists $C_{E,\alpha,R}>0$ such that for every $\lambda_1,\lambda_2\in\overline{\Lambda}_n$

$$|E(x,\lambda_1,\Psi)-E(x,\lambda_2,\Psi)|\leq C_{E,\alpha,R}|\lambda_1-\lambda_2|^\alpha\,. \tag{72}$$

Remark 3. The constants $c_1(R)$ and $c_4(\delta, R)$ can be assumed to be decreasing with respect to R, while $c_2(R)$, $c_3(\delta, R)$, $L_{G,\delta,R}$, $L_{E,\delta,R}$, and $C_{E,\alpha,R}$ can be assumed to be increasing with respect to R.

Proof. (*Proof of Lemma 3*) In view of $(\mathcal{Q}_1)$ and $(\mathcal{Q}_2)$, we have that the function $(x, \Psi) \mapsto \sigma(x, \Psi)$ is continuous from $\mathbb{R}^d \times \mathcal{P}_1(Y) \to \Lambda_n$. Hence, there exists $\eta = \eta(R) > 0$ such that for every $x \in \mathrm{B}_R$, every $\Psi \in \mathcal{P}_1(\mathrm{B}_R^Y)$, and every $h \in \{1, \dots, n\}$, $\sigma_h(x, \Psi) \geq \eta > 0$, so that (i) holds.

From (i), (64), the regularity of Φ, and $(\mathcal{Q}_3)$, we further deduce that (67) holds for a suitable constant $c_2 = c_2(R)$.

For every $(x, \lambda) \in \mathbb{R}^d \times \Lambda_n$ and every $\Psi \in \mathcal{P}_1(Y)$, we have that $K(x, \lambda, \Psi)$ is symmetric, positive semi-definite on $\mathbb{R}^n$, and positive definite on $\mathbb{R}^n_0$. Since K is continuous with respect to (x, λ, Ψ), we deduce that there exists a positive constant $c_4 = c_4(\delta, R) \leq c_2$ such that inequality (69) holds for every $(x, \lambda) \in \mathrm{B}^Y_{R,\delta}$ and every $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$. Since G is the inverse of K on $\mathbb{R}^n_0$, (67) and (69) imply (66) and (68) with $c_1(R) := c_2(R)^{-1}$ and $c_3(\delta, R) := c_4(\delta, R)^{-1}$. This concludes the proof of (ii) and (iii).

The Lipschitz continuity (iv) of $G(x, \cdot, \Psi)$ in Λ_n^δ follows from the regularity of $K(x, \cdot, \Psi)$, from (i)–(iii), and from the identity, on $\mathbb{R}^n_0$,

$$G(x, \lambda_1, \Psi) - G(x, \lambda_2, \Psi) = G(x, \lambda_1, \Psi)\big(K(x, \lambda_2, \Psi) - K(x, \lambda_1, \Psi)\big)G(x, \lambda_2, \Psi).$$

As for (v), we notice that for $x \in \mathrm{B}_R$, $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$, and $\lambda \in \Lambda_n^\delta$, the ratio $\lambda_h/\sigma_h(x, \Psi)$ is bounded from below and from above by δ and by $1/\eta$, respectively. Since the function $a \mapsto a \ln a$ is locally Lipschitz continuous in $(0, +\infty)$, we have that there exists $L = L(\delta, R) > 0$ such that (71) holds.

Finally, we note that the function $a \mapsto a \ln a$ belongs to $W^{1,p}([0, A])$ for every $p \in [1, +\infty)$ and every $A < +\infty$. In view of (i), for every $x \in \mathrm{B}_R$, every $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$, and every $\lambda \in \overline{\Lambda}_n$, the ratio $\lambda_h/\sigma_h(x, \Psi)$ is bounded above by $1/\eta$. Hence, by Sobolev embedding in dimension one we infer that for every $\alpha \in (0, 1)$ there exists $C = C(\alpha, R) > 0$ such that (72) holds. □

Before stating the main properties of the distance $\mathsf{d}_{(x,\Psi)}$, we define, for every $x \in \mathbb{R}^d$, every $\Psi \in \mathcal{P}_1(Y)$, and every $\lambda, \lambda_1, \lambda_2 \in \Lambda_n$, the norm

$$\|\lambda_1 - \lambda_2\|_{G(x,\lambda,\Psi)} := \langle G(x, \lambda, \Psi)(\lambda_1 - \lambda_2), \lambda_1 - \lambda_2 \rangle^{\frac{1}{2}},$$

which is well defined in view of (66) and (68).

Lemma 4. *Let $\delta, R > 0$ and let $c_1, c_3 > 0$ be the constants determined in (66) and (68). Then, the following facts hold:*

(i) there exists a positive constant $m_1 = m_1(R)$ such that for every $x \in \mathrm{B}_R$ and every $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$

$$m_1|\lambda_1 - \lambda_2| \leq \mathsf{d}_{(x,\Psi)}(\lambda_1, \lambda_2) \quad \text{for every } \lambda_1, \lambda_2 \in \Lambda_n; \tag{73}$$

(ii) there exist two positive constants $m_2 = m_2(\delta, R)$ *and* $m_3 = m_3(\delta, R)$ *such that for every* $x \in \mathrm{B}_R$, *every* $\Psi \in \mathcal{P}(B_R^Y)$, *and every* $\lambda_1, \lambda_2 \in \Lambda_n^\delta$

$$\mathsf{d}_{(x,\Psi)}(\lambda_1, \lambda_2) \le m_2|\lambda_1 - \lambda_2|, \tag{74}$$

$$\mathsf{d}_{(x,\Psi)}(\lambda_1, \lambda_2) \le \|\lambda_1 - \lambda_2\|_{G(x,\lambda_1,\Psi)} + m_3|\lambda_1 - \lambda_2|^{\frac{3}{2}}; \tag{75}$$

(iii) there exists a positive constant $m_4 = m_4(\delta, R)$ *such that for every* $x \in \mathrm{B}_R$, *every* $\Psi \in \mathcal{P}(B_R^Y)$, *and every* $\lambda_1, \lambda_2 \in \Lambda_n^\delta$ *satisfying*

$$\sqrt{\frac{c_3}{c_1}}\,|\lambda_1 - \lambda_2| < \min\left\{\mathrm{dist}(\lambda_1, \partial\Lambda_n^\delta), \mathrm{dist}(\lambda_2, \partial\Lambda_n^\delta)\right\} \tag{76}$$

we have

$$\|\lambda_1 - \lambda_2\|_{G(x,\lambda_1,\Psi)} \le \mathsf{d}_{(x,\Psi)}(\lambda_1, \lambda_2) + m_4|\lambda_1 - \lambda_2|^{\frac{3}{2}}. \tag{77}$$

Remark 4. The constant $m_1(R)$ can be assumed to be decreasing with respect to R, while $m_2(\delta, R)$, $m_3(\delta, R)$, and $m_4(\delta, R)$, can be assumed to be increasing with respect to R.

Proof. (*Proof of Lemma 4*) Point (i) is a consequence of (66). We now prove (ii). Given $x \in \mathrm{B}_R$, $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$, and $\lambda_1, \lambda_2 \in \Lambda_n^\delta$, we have that the curve

$$\rho(s) := (1 - s)\lambda_1 + s\lambda_2 \qquad s \in [0, 1]$$

is a competitor for the infimum in the definition of $\mathsf{d}_{(x,\Psi)}(\lambda_1, \lambda_2)$ in (65). Moreover, by convexity, $\rho(s) \in \Lambda_n^\delta$ for every $s \in [0, 1]$. Therefore, applying (iii) of Lemma 3 we get

$$\mathsf{d}_{(x,\Psi)}(\lambda_1, \lambda_2) \le \int_0^1 \langle G(x, \rho(s), \Psi)(\lambda_2 - \lambda_1), \lambda_2 - \lambda_1\rangle^{\frac{1}{2}}\,\mathrm{d}s \le \sqrt{c_3}|\lambda_1 - \lambda_2|,$$

which is (74) with $m_2 = \sqrt{c_3}$.

Combining, instead, (iii) and (iv) of Lemma 3, we can further estimate

$$\begin{aligned}
\mathsf{d}_{(x,\Psi)}(\lambda_1, \lambda_2) &\le \int_0^1 \langle G(x, \rho(s), \Psi)(\lambda_2 - \lambda_1), \lambda_2 - \lambda_1\rangle^{\frac{1}{2}}\,\mathrm{d}s \\
&\le \langle G(x, \lambda_1, \Psi)(\lambda_2 - \lambda_1), \lambda_2 - \lambda_1\rangle^{\frac{1}{2}} \\
&\quad + \int_0^1 \left|\left\langle \big(G(x, \rho(s), \Psi) - G(x, \lambda_1, \Psi)\big)(\lambda_2 - \lambda_1), \lambda_2 - \lambda_1\right\rangle\right|^{\frac{1}{2}}\mathrm{d}s \\
&\le \|\lambda_1 - \lambda_2\|_{G(x,\lambda_1,\Psi)} + \int_0^1 \left(L_{G,\delta,R}|\lambda_1 - \rho(s)|\right)^{\frac{1}{2}}|\lambda_1 - \lambda_2|\,\mathrm{d}s \\
&\le \|\lambda_1 - \lambda_2\|_{G(x,\lambda_1,\Psi)} + \sqrt{L_{G,\delta,R}}\,|\lambda_1 - \lambda_2|^{\frac{3}{2}},
\end{aligned} \tag{78}$$

from which we conclude (75) with $m_3 = \sqrt{L_{G,\delta,R}}$.

Finally, let x, Ψ, λ_1, and λ_2 be as in point (iv). For every $\varepsilon > 0$ let $\rho_\varepsilon \in C^1([0,1];\Lambda_n)$ with $\rho_\varepsilon(0)=\lambda_1$ and $\rho_\varepsilon(1)=\lambda_2$ be such that

$$\int_0^1 \left\langle G(x,\rho_\varepsilon(s),\Psi)\rho_\varepsilon'(s),\rho_\varepsilon'(s)\right\rangle^{\frac{1}{2}} \mathrm{d}s \le \mathsf{d}_{(x,\Psi)}(\lambda_1,\lambda_2)+\varepsilon\,. \tag{79}$$

In view of (74) and of (66), we deduce from (79) that

$$\sqrt{c_1}\int_0^1 |\rho_\varepsilon'(s)|\,\mathrm{d}s \le m_2|\lambda_1-\lambda_2|+\varepsilon = \sqrt{c_3}|\lambda_1-\lambda_2|+\varepsilon\,. \tag{80}$$

Hence, (76) and (80) imply that

$$\int_0^1 |\rho_\varepsilon'(s)|\,\mathrm{d}s < \min\left\{\mathrm{dist}(\lambda_1,\partial\Lambda_n^\delta),\mathrm{dist}(\lambda_2,\partial\Lambda_n^\delta)\right\}+\frac{\varepsilon}{\sqrt{c_1}}\,.$$

Therefore, for ε small enough we may assume that $\rho_\varepsilon(s)\in\Lambda_n^\delta$ for every $s\in[0,1]$. For such ε, we estimate

$$\begin{aligned}
\|\lambda_1-\lambda_2\|_{G(x,\lambda_1,\Psi)} &\le \int_0^1 \left\langle G(x,\lambda_1,\Psi)\rho_\varepsilon'(s),\rho_\varepsilon'(s)\right\rangle^{\frac{1}{2}}\mathrm{d}s\\
&\le \int_0^1 \left\langle G(x,\rho_\varepsilon(s),\Psi)\rho_\varepsilon'(s),\rho_\varepsilon'(s)\right\rangle^{\frac{1}{2}}\mathrm{d}s\\
&\quad+\int_0^1 \left|\left\langle \big(G(x,\lambda_1,\Psi)-G(x,\rho_\varepsilon(s),\Psi)\big)\rho_\varepsilon'(s),\rho_\varepsilon'(s)\right\rangle\right|^{\frac{1}{2}}\mathrm{d}s\\
&\le \mathsf{d}_{(x,\Psi)}(\lambda_1,\lambda_2)+\int_0^1 \left|\left\langle \big(G(x,\lambda_1,\Psi)-G(x,\rho_\varepsilon(s),\Psi)\big)\rho_\varepsilon'(s),\rho_\varepsilon'(s)\right\rangle\right|^{\frac{1}{2}}\mathrm{d}s+\varepsilon\,.
\end{aligned}$$

Since $\rho_\varepsilon(s)\in\Lambda_n^\delta$ for every $s\in[0,1]$, by (iv) of Lemma 3 and by (80), we have that

$$\begin{aligned}
&\|\lambda_1-\lambda_2\|_{G(x,\lambda_1,\Psi)}\\
&\quad\le \mathsf{d}_{(x,\Psi)}(\lambda_1,\lambda_2)+\sqrt{L_{G,\delta,R}}\int_0^1 |\lambda_1-\rho_\varepsilon(s)|^{\frac{1}{2}}|\rho_\varepsilon'(s)|\,\mathrm{d}s+\varepsilon\\
&\quad\le \mathsf{d}_{(x,\Psi)}(\lambda_1,\lambda_2)+\sqrt{L_{G,\delta,R}}\left(\int_0^1 |\rho_\varepsilon'(s)|\,\mathrm{d}s\right)^{\frac{3}{2}}+\varepsilon\\
&\quad\le \mathsf{d}_{(x,\Psi)}(\lambda_1,\lambda_2)+\sqrt{L_{G,\delta,R}}\left(\frac{c_3}{c_1}\right)^{\frac{3}{4}}|\lambda_1-\lambda_2|^{\frac{3}{2}}+\varepsilon\left(1+\sqrt{\frac{L_{G,\delta,R}}{c_1^{3/2}}}\right).
\end{aligned}\tag{81}$$

Thus, we conclude (77) by passing to the limit in (81) as $\varepsilon\to 0$. In particular, $m_4=\sqrt{L_{G,\delta,R}}\left(\frac{c_3}{c_1}\right)^{\frac{3}{4}}$. □

We now rewrite the multi-step scheme presented in Sect. 4 in the language of Markov chains and show its short-time convergence to a solution of the continuity equation (7), where for $\Psi\in\mathcal{P}_1(Y)$ the field $b_\Psi: Y\to Y$ is now defined as

$$b_\Psi(x,\lambda) := \begin{pmatrix} v_\Psi(x,\lambda)\\ \mathcal{Q}(x,\Psi)\lambda\end{pmatrix}$$

for a velocity field $v_\Psi : Y \to \mathbb{R}^d$ satisfying properties (v_1)–(v_3) of Sect. 2.

Let us fix a time step $\tau_k > 0$, $k \in \mathbb{N}$, such that $\tau_k \to 0$ as $k \to \infty$, and let $t_i^k := i\tau_k$ for $i \in \mathbb{N}$. For $i = 0$ we set $\Psi_0^k := \widehat{\Psi} \in \mathcal{P}_1(Y)$. For $i > 0$, assume we are given $\Psi_i^k \in \mathcal{P}_1(Y)$. Then, similarly to (39), the label of an agent sitting in position $\hat{x} \in \mathbb{R}^d$ with label $\hat{\lambda} \in \overline{\Lambda}_n$ is updated by solving the minimizing movement

$$\min \left\{ E(\hat{x}, \lambda, \Psi_i^k) + \frac{1}{2\tau_k} \mathsf{d}^2_{(\hat{x},\Psi_i^k)}(\lambda, \hat{\lambda}) : \lambda \in \overline{\Lambda}_n \right\}. \tag{82}$$

Since $\overline{\Lambda}_n$ is compact, (82) admits at least one solution $\lambda_{(\hat{x},\hat{\lambda}),i+1}$.[1] Therefore, we can define $\lambda^k_{(\hat{x},\hat{\lambda}),i+1}$, Λ^k_{i+1}, and $\tilde{\Psi}^k_{i+1}$ exactly as in (9), (10), and (11), respectively. The step (12) in the space variable remains the same, and $x^k_{(\hat{x},\hat{\lambda}),i+1}$, X^k_{i+1}, Ψ^k_{i+1} are as in (13), (14), and (15). Furthermore, we refer to (15), (16), and (17) for the definition of the interpolation curves Ψ^k, $\widetilde{\Psi}^k$, and $\underline{\Psi}^k$.

Repeating step by step the proofs of Lemmas 1 and 2, we obtain the following uniform estimate on $\Psi^k(t)$, $\widetilde{\Psi}^k(t)$, and $\underline{\Psi}^k(t)$.

Lemma 5. *Let $\widehat{\Psi} \in \mathcal{P}_c(Y)$. Then there exists an increasing continuous function $R\colon [0,+\infty) \to [0,+\infty)$ such that for every $T \in [0,+\infty)$, every $k \in \mathbb{N}$, and every $t \in [0,T]$, $\Psi^k(t), \underline{\Psi}^k(t), \widetilde{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_{R(T)})$.*

Proof. The statement follows by the arguments of Lemmas 1 and 2 . In particular, we gave there an explicit formula for $R(T)$ as a function of $T \in [0,+\infty)$, which turns out to be continuous and increasing. □

Also in the current setting, we need to write an approximate Euler–Lagrange equation associated with (82). This is done in Proposition 5, for proving which we need the following lemma.

Lemma 6. *Let $f : \mathbb{R}^N \to \mathbb{R}\cup\{+\infty\}$ be a convex function, let $A \in \mathbb{M}^N$ be a symmetric and positive definite matrix, and let $\|\cdot\|_A : \mathbb{R}^N \to [0,+\infty)$ be the norm associated with A, namely $\|\xi\|_A^2 := \langle A\xi, \xi\rangle$, for all $\xi \in \mathbb{R}^N$. For a fixed $\zeta \in \mathbb{R}^N$ and $c > 0$, assume that ξ_0 solves*

$$\min \left\{ f(\xi) + c\|\xi - \zeta\|_A^2 \right\}. \tag{83}$$

Then ξ_0 also solves

$$\min \left\{ f(\xi) + c\|\xi - \zeta\|_A^2 - c\|\xi - \xi_0\|_A^2 \right\}. \tag{84}$$

Proof. It is enough to observe that the problem (84) can be equivalently rewritten as

$$\min \left\{ f(\xi) + 2c\langle \xi, A(\xi_0 - \zeta)\rangle \right\};$$

hence, it is a convex minimization problem. Since ξ_0 solves (83), we have $-2cA(\xi_0 - \zeta) \in \partial f(\xi_0)$, which is exactly the Euler condition for the above problem.

□

[1] The arguments in [30, Section 2.3] can also be used to show that (82) admits indeed a unique solution for τ_k sufficiently small. However, uniqueness is not needed in our framework.

Proposition 5. *Let $\delta, R > 0$ and let $m_1(R)$, $C_{E,\alpha,R}$, and $L_{E,\delta,R}$ be the constants determined in Lemmas 3 and 4. Assume that $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$ and $(x, \lambda) \in \mathrm{B}_{R,\delta}^Y$, and let $\tilde{\lambda}$ be a solution to*

$$\min \left\{ E(x, \rho, \Psi) + \frac{1}{2\tau_k} \mathsf{d}^2_{(x,\Psi)}(\rho, \lambda) : \rho \in \overline{\Lambda}_n \right\}. \tag{85}$$

Then, the following facts hold:

(i) for every $\alpha \in (0, 1)$

$$|\tilde{\lambda} - \lambda| \le \left(\frac{2\, C_{E,\alpha,R}}{m_1^2} \right)^{1/(2-\alpha)} \tau_k^{1/(2-\alpha)}; \tag{86}$$

(ii) if $\tilde{\lambda} \in \Lambda_n^\delta$, then

$$|\tilde{\lambda} - \lambda| \le \frac{2\, L_{E,\delta,R}}{m_1^2}\, \tau_k\,; \tag{87}$$

(iii) if $\lambda, \tilde{\lambda} \in \Lambda_n^\delta$ and μ is the unique solution to

$$\min \left\{ E(x, \rho, \Psi) + \frac{1}{2\tau_k} \|\rho - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)} : \rho \in \overline{\Lambda}_n \right\}, \tag{88}$$

then, for every $\alpha \in (0, 1)$ we have

$$|\mu - \lambda| \le \left(\frac{2\, C_{E,\alpha,R}}{m_1^2} \right)^{1/(2-\alpha)} \tau_k^{1/(2-\alpha)}. \tag{89}$$

If, in addition, $\mu \in \Lambda_n^\delta$, then

$$|\mu - \lambda| \le \frac{2\, L_{E,\delta,R}}{m_1^2}\, \tau_k\,. \tag{90}$$

Finally, if $\lambda, \tilde{\lambda} \in \Lambda_n^\delta$ satisfy (76), there exists a positive constant $C = C(\delta, R)$ such that

$$\left| \frac{\tilde{\lambda} - \lambda}{\tau_k} - \mathcal{Q}(x, \Psi)\tilde{\lambda} \right| \le C \tau_k^{1/4}. \tag{91}$$

Proof. By the minimality of $\tilde{\lambda}$, by (vi) of Lemma 3, and by (i) of Lemma 4 we have that for every $\alpha \in (0, 1)$

$$\frac{m_1^2}{2\tau_k} |\tilde{\lambda} - \lambda|^2 \le \left| E(x, \lambda, \Psi) - E(x, \tilde{\lambda}, \Psi) \right| \le C_{E,\alpha,R} |\tilde{\lambda} - \lambda|^\alpha, \tag{92}$$

where $m_1 = m_1(R)$ and $C_{E,\alpha,R}$ are defined in Lemmas 3 and 4, respectively. From (92), we deduce (86). In a similar way, we deduce (89), recalling that $m_1 = \sqrt{c_1}$, where c_1 has been determined in (66).

If we further assume that $\tilde{\lambda} \in \Lambda_n^\delta$, by minimality of $\tilde{\lambda}$, by (v) of Lemma 3, and by (i) of Lemma 4, we have that

$$\frac{m_1^2}{2\tau_k}|\lambda - \tilde{\lambda}|^2 \le \frac{1}{2\tau_k}\mathsf{d}^2_{(x,\Psi)}(\lambda, \tilde{\lambda}) \le |E(x, \lambda, \Psi) - E(x, \tilde{\lambda}, \Psi)| \le L_{E,\delta,R}|\lambda - \tilde{\lambda}|\,.$$

Hence, we deduce (87). Moreover, if $\mu \in \Lambda_n^\delta$, in the very same way we get (90).

In order to prove (91), we first estimate the Euclidean norm $|\mu - \tilde{\lambda}|$. Denote by $\chi_{\overline{\Lambda}_n}$ the characteristic function of the convex set $\overline{\Lambda}_n$ in the sense of convex analysis. Since $E(x, \cdot, \Psi)$ is convex in $\overline{\Lambda}_n$, we can apply Lemma 6 with $f(\cdot) = E(x, \cdot, \Psi) + \chi_{\overline{\Lambda}_n}(\cdot)$, $\xi_0 = \mu$, $c = \frac{1}{2\tau_k}$, $\zeta = \lambda$, and $A = G(x, \tilde{\lambda}, \Psi)$ obtaining

$$\begin{aligned} E(x, \mu, \Psi) + \frac{1}{2\tau_k}\|\mu - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)} + \frac{1}{2\tau_k}\|\mu - \tilde{\lambda}\|^2_{G(x,\tilde{\lambda},\Psi)} \\ \le E(x, \tilde{\lambda}, \Psi) + \frac{1}{2\tau_k}\|\tilde{\lambda} - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)}\,. \end{aligned}$$

Reordering the terms in the previous inequality and adding and subtracting on the right-hand side the terms $\frac{1}{2\tau_k}\mathsf{d}^2_{(x,\Psi)}(\tilde{\lambda}, \lambda)$ and $\frac{1}{2\tau_k}\mathsf{d}^2_{(x,\Psi)}(\mu, \lambda)$, we obtain

$$\begin{aligned} \frac{1}{2\tau_k}\|\mu - \tilde{\lambda}\|^2_{G(x,\tilde{\lambda},\Psi)} \le\ & E(x, \tilde{\lambda}, \Psi) + \frac{1}{2\tau_k}\mathsf{d}^2_{(x,\Psi)}(\tilde{\lambda}, \lambda) \\ & - E(x, \mu, \Psi) - \frac{1}{2\tau_k}\mathsf{d}^2_{(x,\Psi)}(\mu, \lambda) \\ & - \frac{1}{2\tau_k}\|\mu - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)} + \frac{1}{2\tau_k}\|\tilde{\lambda} - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)} \\ & + \frac{1}{2\tau_k}\mathsf{d}^2_{(x,\Psi)}(\mu, \lambda) - \frac{1}{2\tau_k}\mathsf{d}^2_{(x,\Psi)}(\tilde{\lambda}, \lambda)\,. \end{aligned} \tag{93}$$

By the minimality of $\tilde{\lambda}$, inequality (93) simplifies to

$$\|\mu - \tilde{\lambda}\|^2_{G(x,\tilde{\lambda},\Psi)} \le \|\tilde{\lambda} - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)} - \|\mu - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)} + \mathsf{d}^2_{(x,\Psi)}(\mu, \lambda) - \mathsf{d}^2_{(x,\Psi)}(\tilde{\lambda}, \lambda)\,. \tag{94}$$

Since $x \in \mathrm{B}_R$, $\Psi \in \mathcal{P}(\mathrm{B}_R^Y)$, $\lambda, \tilde{\lambda}, \mu \in \Lambda_n^\delta$, and $\lambda, \tilde{\lambda}$ satisfy (76), we deduce from (94), from (ii) of Lemma 3, and from (ii)–(iii) of Lemma 4 that

$$\begin{aligned} c_1^2|\mu - \tilde{\lambda}|^2 \le\ & \left(d_{(x,\Psi)}(\tilde{\lambda}, \lambda) + m_4|\tilde{\lambda} - \lambda|^{\frac{3}{2}}\right)^2 \\ & + \left(\|\mu - \lambda\|_{G(x,\tilde{\lambda},\Psi)} + m_3|\mu - \lambda|^{\frac{3}{2}}\right)^2 \\ & - \|\mu - \lambda\|^2_{G(x,\tilde{\lambda},\Psi)} - \mathsf{d}^2_{(x,\Psi)}(\tilde{\lambda}, \lambda)\,. \end{aligned} \tag{95}$$

Developing the squares and using (iii) of Lemma 3 and (ii) of Lemma 4, we continue in (95) with

$$c_1^2|\mu - \tilde{\lambda}|^2 \le m_4^2|\tilde{\lambda} - \lambda|^3 + m_3^2|\mu - \lambda|^3 + 2\,m_2\,m_4|\tilde{\lambda} - \lambda|^{\frac{5}{2}} + 2\sqrt{c_3}\,m_3|\mu - \lambda|^{\frac{5}{2}}. \tag{96}$$

Combining (96) with (87) and (90), we deduce

$$|\mu - \tilde{\lambda}| \leq \widetilde{C}\tau_k^{5/4}\,. \tag{97}$$

for some positive constant $\widetilde{C} = \widetilde{C}(\delta, R)$ independent of k.

We are now in a position to conclude (91). The minimality of μ, indeed, implies that for every $\xi \in \mathbb{R}^n_0$

$$\langle D_\lambda E(x,\mu,\Psi),\xi\rangle + \frac{1}{\tau_k}\big\langle G(x,\tilde{\lambda},\Psi)(\mu-\lambda),\xi\big\rangle = 0\,.$$

By a simple algebraic manipulation, we rewrite the previous equality as

$$\begin{aligned}\langle D_\lambda E(x,\mu,\Psi),\xi\rangle + \frac{1}{\tau_k}\big\langle G(x,\mu,\Psi)(\tilde{\lambda}-\lambda),\xi\big\rangle &\\ = \frac{1}{\tau_k}\big\langle G(x,\tilde{\lambda},\Psi)(\tilde{\lambda}-\mu),\xi\big\rangle &\\ + \frac{1}{\tau_k}\big\langle \big(G(x,\mu,\Psi)-G(x,\tilde{\lambda},\Psi)\big)(\tilde{\lambda}-\lambda),\xi\big\rangle.&\end{aligned} \tag{98}$$

Taking $\xi = K^\top(x,\mu,\Psi)\omega$ for $\omega \in \mathbb{R}^n$ in (98), we get that

$$\begin{aligned}K(x,\mu,\Psi)\,D_\lambda E(x,\mu,\Psi) + \frac{1}{\tau_k}(\tilde{\lambda}-\lambda) = \frac{1}{\tau_k}K(x,\mu,\Psi)\,G(x,\tilde{\lambda},\Psi)(\tilde{\lambda}-\mu)&\\ + \frac{1}{\tau_k}K(x,\mu,\Psi)\,\big(G(x,\mu,\Psi)-G(x,\tilde{\lambda},\Psi)\big)(\tilde{\lambda}-\lambda).&\end{aligned}$$

Since $K(x,\mu,\Psi)D_\lambda E(x,\mu,\Psi) = -\mathcal{Q}(x,\Psi)\mu$ (see [30, Theorem 3.1]), we actually have

$$\begin{aligned}-\mathcal{Q}(x,\Psi)\tilde{\lambda} + \frac{1}{\tau_k}\big(\tilde{\lambda}-\lambda\big) = \frac{1}{\tau_k}K(x,\mu,\Psi)\,G(x,\tilde{\lambda},\Psi)\big(\tilde{\lambda}-\mu\big)&\\ + \frac{1}{\tau_k}K(x,\mu,\Psi)\,\big(G(x,\mu,\Psi)-G(x,\tilde{\lambda},\Psi)\big)(\tilde{\lambda}-\lambda)&\\ + \mathcal{Q}(x,\Psi)(\tilde{\lambda}-\mu)\,.&\end{aligned} \tag{99}$$

Combining (ii)–(iv) of Lemma 3 with the inequalities (87), (90), (97), and (99), and with the assumptions $x \in \mathrm{B}_R$, $\Psi \in \mathcal{P}(\mathrm{B}^Y_R)$, and $\tilde{\lambda}, \mu \in \Lambda^\delta_n$, we get (91), and therefore, the proof is concluded. □

Lemma 7. *Let $r > 0$, $\eta > \delta > 0$, and $\widehat{\Psi} \in \mathcal{P}(\mathrm{B}^Y_{r,\eta})$. Then, there exists $T_f > 0$ such that for every $k \in \mathbb{N}$ large enough and every $t < T_f$ the following hold:*

(i) $\Psi^k(t), \underline{\Psi}^k(t), \widetilde{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_{R(T_f),\delta})$, where $R\colon [0,+\infty) \to [0,+\infty)$ is the function determined in Lemma 5;

(ii) if $t \in [t^k_i, t^k_{i+1})$ for some $i \in \mathbb{N}$, for every $(x,\lambda) \in \operatorname{spt}\underline{\Psi}^k(t)$

$$\sqrt{\frac{c_3(\delta,R(T_f))}{c_1(R(T_f))}}\,|\Lambda^k_{i+1}(t,x,\lambda)-\lambda| < \min\Big\{\operatorname{dist}(\lambda,\partial\Lambda^\delta_n), \operatorname{dist}(\Lambda^k_{i+1}(t,x,\lambda),\partial\Lambda^\delta_n)\Big\}\,.$$

Proof. Since $\widehat{\Psi} \in \mathcal{P}(\mathrm{B}^Y_{r,\eta})$, we deduce from Lemma 5 that for every $T > 0$, every $k \in \mathbb{N}$, and every i such that $i\tau_k \le T$ we have $\Psi^k_i, \widetilde{\Psi}^k_i \in \mathcal{P}(\mathrm{B}^Y_{R(T)})$. Hence, in order to conclude the lemma we have to study the behaviour of the labels $\lambda \in \overline{\Lambda}_n$ along the multi-step scheme.

Along the proof of the lemma we denote by $\lambda^k(t, x_0, \lambda_0)$ and $x^k(t, x_0, \lambda_0)$, for $(x_0, \lambda_0) \in \operatorname{spt} \widehat{\Psi}$, the curves obtained by iteratively solving (82) and the difference equation (12) in each interval $[t^k_i, t^k_{i+1}]$ starting from (x_0, λ_0) at time $t_0 = 0$ and using, at each node t^k_i, $\hat{\lambda} = \lambda^k(t^k_i, x_0, \lambda_0)$ and $\hat{x} = x^k(t^k_i, x_0, \lambda_0)$ as new initial conditions. As in (18), we define the piecewise constant interpolations $\overline{x}^k(t, x_0, \lambda_0)$, $\underline{x}^k(t, x_0, \lambda_0)$ and $\overline{\lambda}^k(t, x_0, \lambda_0)$, $\underline{\lambda}^k(t, x_0, \lambda_0)$.

The assumption $\widehat{\Psi} \in \mathcal{P}(\mathrm{B}^Y_{r,\eta})$ means that for every $(x_0, \lambda_0) \in \operatorname{spt}\widehat{\Psi}$ we have $\lambda_0 \in \Lambda^\eta_n$. Since the measures $\Psi^k(t)$ and $\widetilde{\Psi}^k(t)$ are supported on pairs of the form $(x^k(t, x_0, \lambda_0), \lambda^k(t, x_0, \lambda_0))$ and $(\underline{x}^k(t, x_0, \lambda_0), \overline{\lambda}^k(t, x_0, \lambda_0))$, respectively, we are led to estimate the number of steps needed by (82) to exit Λ^δ_n, knowing that the initial label $\lambda_0 \in \Lambda^\eta_n$.

Let us fix $\overline{\alpha} \in (0, 1)$. For $k \in \mathbb{N}$ such that $\tau_k \le 1$, we claim that the properties (i) and (ii) hold with $R = R(t^k_i)$ for every $t \in [0, t^k_i]$ until the following conditions are fulfilled:

$$\sum_{j=i}^{i-1} \frac{2\,L(j-1,k)}{m_1^2(j-1,k)}\,\tau_k + \left(\frac{2\,C(i-1,k)}{m_1^2(i-1,k)}\right)^{1/(2-\overline{\alpha})} \sqrt{\tau_k} < \eta - \delta, \tag{100}$$

$$\sum_{j=i}^{i-1} \frac{2\,L(j-1,k)}{m_1^2(j-1,k)}\,\tau_k + \frac{2\,C(i-1,k)}{m_1^2(i-1,k)} \left(\frac{c_3(i-1,k)}{c_1(i-1,k)}\right)^{1/2} \tau_k < \eta - \delta\,, \tag{101}$$

where we have set $L(j,k) := L_{E,\delta,R(j\tau_k)}$, $C(j,k) := C_{E,\overline{\alpha},R(j\tau_k)}$, $m_1(j,k) := m_1(R(j\tau_k))$, $c_1(j,k) := c_1(R(j\tau_k))$, and $c_3(j,k) := c_3(\delta, R(j\tau_k))$.

Given the claim for granted, for every $k \in \mathbb{N}$ let us denote with $i_k \in \mathbb{N}$ the first index for which at least one of the two inequalities (100) or (101) is violated. For simplicity, let us assume that it is always (100) to be violated in i_k. Hence,

$$\sum_{j=1}^{i_k-1} \frac{2\,L(j-1,k)}{m_1^2(j-1,k)}\,\tau_k \ge \eta - \delta - \left(\frac{2\,C(i_k-1,k)}{m_1^2(i_k-1,k)}\right)^{1/(2-\overline{\alpha})} \sqrt{\tau_k}\,.$$

Since $L_{E,\delta,R}$ is increasing with respect to R, $m_1(R)$ is decreasing with respect to R, and $R(t)$ determined in Lemma 5 is increasing with respect to t, we also have that

$$\frac{2\,L(i_k-1,k)}{m_1^2(i_k-1,k)}\,(i_k-1)\tau_k \ge \eta - \delta - \left(\frac{2\,C(i_k-1,k)}{m_1^2(i_k-1,k)}\right)^{1/(2-\overline{\alpha})} \sqrt{\tau_k}\,,$$

from which we deduce that $i_k\tau_k$ is bounded away from 0. Therefore, there exists $T_f > 0$ such that $T_f < (i_k - 1)\tau_k$ for every k large enough. A similar estimate can be

obtained if (101) is violated, and we conclude that there exists $T_f > 0$ such that (i) and (ii) hold.

Let us prove the claim. For fixed $i \in \mathbb{N}$, assume that (100)–(101) hold and that $\lambda^k(t_j^k, x_0, \lambda_0) \in \Lambda_n^\delta$ for $0 \le j < i$. Then, by (ii) of Proposition 5 we have that for every $1 \le j < i$

$$|\lambda^k(t_j^k, x_0, \lambda_0) - \lambda^k(t_{j-1}^k, x_0, \lambda_0)| \le \frac{2\,L(j-1,k)}{m_1^2(j-1,k)}\,\tau_k\,. \tag{102}$$

By (i) of Proposition 5, we have, since $\tau_k \le 1$,

$$|\lambda^k(t_i^k, x_0, \lambda_0) - \lambda^k(t_{i-1}^k, x_0, \lambda_0)| \le \left(\frac{2\,C(i-1,k)}{m_1^2(i-1,k)}\right)^{1/(2-\overline{\alpha})} \sqrt{\tau_k}\,. \tag{103}$$

Hence, by (100), (102), (103), and by triangle inequality, we deduce that

$$\begin{aligned}|\lambda^k(t_i^k, x_0, \lambda_0) - \lambda_0| &\le \sum_{j=1}^{i} |\lambda^k(t_j^k, x_0, \lambda_0) - \lambda^k(t_{j-1}^k, x_0, \lambda_0)| \\ &\le \sum_{j=i}^{i-1} \frac{2\,L(j-1,k)}{m_1^2(j-1,k)}\,\tau_k + \left(\frac{2\,C(i-1,k)}{m_1^2(i-1,k)}\right)^{1/(2-\overline{\alpha})} \sqrt{\tau_k} < \eta - \delta,\end{aligned}$$

which implies that $\lambda^k(t_i^k, x_0, \lambda_0) \in \Lambda_n^\delta$ as $\lambda_0 \in \Lambda_n^\eta$. Since (100) is independent of the particular choice of the initial condition $(x_0, \lambda_0) \in \operatorname{spt}\widehat{\Psi} \subseteq \mathrm{B}^Y_{r,\eta}$, we infer that $\Psi^k(t), \underline{\Psi}^k(t), \widetilde{\Psi}^k(t) \in \mathcal{P}(\mathrm{B}^Y_{R(t_i^k),\delta})$ for every $t \in [0, t_i^k]$.

Let us now denote by $\mu_i^k \in \overline{\Lambda}_n$ the solution to the minimum problem

$$\min_{\rho \in \overline{\Lambda}_n} \left\{ E\left(x^k(t_{i-1}^k, x_0, \lambda_0), \rho, \Psi_{i-1}^k\right) + \frac{1}{2\tau_k} \|\rho - \lambda^k(t_{i-1}^k, x_0, \lambda_0)\|^2_{G(x^k(t_{i-1}^k, x_0, \lambda_0), \lambda^k(t_i^k, x_0, \lambda_0), \Psi_{i-1}^k)} \right\}.$$

Then, by (iii) of Proposition 5 we get that

$$|\mu_i^k - \lambda^k(t_{i-1}^k, x_0, \lambda_0)| \le \left(\frac{2\,C(i-1,k)}{m_1^2(i-1,k)}\right)^{1/(2-\overline{\alpha})} \sqrt{\tau_k}\,.$$

Therefore, by triangle inequality and by (100) we obtain

$$\begin{aligned}|\mu_i^k - \lambda_0| &\le \sum_{j=1}^{i-1} |\lambda^k(t_j^k, x_0, \lambda_0) - \lambda^k(t_{j-1}^k, x_0, \lambda_0)| + |\mu_i^k - \lambda^k(t_{i-1}^k, x_0, \lambda_0)| \\ &\le \sum_{j=i}^{i-1} \frac{2\,L(j-1,k)}{m_1^2(j-1,k)}\,\tau_k + \left(\frac{2\,C(i-1,k)}{m_1^2(i-1,k)}\right)^{1/(2-\overline{\alpha})} \sqrt{\tau_k} < \eta - \delta,\end{aligned}$$

which yields $\mu_i^k \in \Lambda_n^\delta$.

Since $\lambda^k(t_j^k, x_0, \lambda_0) \in \Lambda_n^\delta$ for $0 \le j \le i$, by (ii) of Proposition 5, by (101), and by (102), we have that

$$\begin{aligned}
&\sum_{j=1}^{i-1} |\lambda^k(t_j^k, x_0, \lambda_0) - \lambda^k(t_{j-1}^k, x_0, \lambda_0)| \\
&\qquad + \left(\frac{c_3(i-1,k)}{c_1(i-1,k)}\right)^{1/2} |\lambda^k(t_i^k, x_0, \lambda_0) - \lambda^k(t_{i-1}^k, x_0, \lambda_0)| \\
&\qquad \le \sum_{j=i}^{i-1} \frac{2\,L(j-1,k)}{m_1^2(j-1,k)}\,\tau_k + \frac{2\,C(i-1,k)}{m_1^2(i-1,k)} \left(\frac{c_3(i-1,k)}{c_1(i-1,k)}\right)^{1/2} \tau_k < \eta - \delta,
\end{aligned}$$

which in turn implies

$$\begin{aligned}
&\left(\frac{c_3(i-1,k)}{c_1(i-1,k)}\right)^{1/2} |\lambda^k(t_i^k, x_0, \lambda_0) - \lambda^k(t_{i-1}^k, x_0, \lambda_0)| \\
&< \min\left\{ \mathrm{dist}\big(\lambda^k(t_i^k, x_0, \lambda_0), \partial\Lambda_n^\delta\big), \mathrm{dist}\big(\lambda^k(t_{i-1}^k, x_0, \lambda_0), \partial\Lambda_n^\delta\big) \right\}.
\end{aligned} \tag{104}$$

Since all the above estimates are independent of the particular choice of $(x_0, \lambda_0) \in \mathrm{spt}\,\widehat{\Psi}$ and since, for $t \in [t_{i-1}^k, t_i^k)$, the measure $\underline{\Psi}^k(t)$ has support

$$\mathrm{spt}\,\underline{\Psi}^k(t) \subseteq \left\{ \big(x^k(t_{i-1}^k, x_0, \lambda_0), \lambda^k(t_{i-1}^k, x_0, \lambda_0)\big) : (x_0, \lambda_0) \in \mathrm{spt}\,\widehat{\Psi} \right\} \subseteq \mathrm{B}^Y_{R(t_{i-1}^k),\delta},$$

we deduce that (ii) holds. □

We are now in a position to prove the short-time convergence of the multi-step Lagrangian scheme for reversible Markov chains. We start by showing the equivalent of Propositions 1 and 4 .

Proposition 6. *Let $r > 0$, $\eta > \delta > 0$, $\widehat{\Psi} \in \mathcal{P}(\mathrm{B}^Y_{r,\eta})$, and let $T_f > 0$ be as in Lemma 7. Then, there exists $C > 0$ such that for every $\varphi \in C_b^1(\mathbb{R}^d \times \overline{\Lambda}_n)$, every $k \in \mathbb{N}$ large enough, every $i \in \mathbb{N}$ such that $(i+1)\tau_k < T_f$, and every $t \in (t_i^k, t_{i+1}^k)$,*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_Y \varphi(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda) = \int_Y \nabla\varphi(x,\lambda) \cdot b_{\Psi^k(t)}(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda) + \vartheta_k(\varphi), \tag{105}$$

where $|\vartheta_k(\varphi)| \le C\|\varphi\|_{C_b^1}\tau_k^{1/4}$.

Proof. Along the proof, we denote by C a generic positive constant independent of i, k, t, and φ, that may vary from line to line.

We follow step by step the proof of Propositions 1 and 4. Let i and k be as in the statement of the proposition, and let us set $R := R(T_f)$. For every test function

$\varphi \in C^1_b(\mathbb{R}^d \times \Lambda_n)$ and every $t \in (t^k_i, t^k_{i+1})$, by definition of $\Psi^k(t)$ we have that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_Y \varphi(x,\lambda)\,\mathrm{d}\Psi^k(t)(x,\lambda) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_Y \varphi\big(X^k_{i+1}(t,x,\lambda),\,\Lambda^k_{i+1}(t,x,\lambda)\big)\,\mathrm{d}\Psi^k_i(x,\lambda)\\ &= \int_Y \nabla_x\varphi\big(X^k_{i+1}(t,x,\lambda),\,\Lambda^k_{i+1}(t,x,\lambda)\big)\cdot v_{\widetilde{\Psi}^k(t)}\big(x,\,\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)\big)\,\mathrm{d}\Psi^k_i(x,\lambda)\\ &\quad+\int_Y \nabla_\lambda\varphi\big(X^k_{i+1}(t,x,\lambda),\,\Lambda^k_{i+1}(t,x,\lambda)\big)\cdot\frac{\big(\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)-\lambda\big)}{\tau_k}\,\mathrm{d}\Psi^k_i(x,\lambda)\,.\end{aligned} \tag{106}$$

In order to deduce (105) from (106), we estimate

$$\begin{aligned}I_1(x,\lambda) &:= \Big|v_{\widetilde{\Psi}^k(t)}\big(x,\,\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)\big)-v_{\Psi^k(t)}\big(X^k_{i+1}(t,x,\lambda),\,\Lambda^k_{i+1}(t,x,\lambda)\big)\Big|\,,\\ I_2(x,\lambda) &:= \left|\frac{\big(\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)-\lambda\big)}{\tau_k}-\mathcal{Q}\big(X^k_{i+1}(t,x,\lambda),\,\Psi^k(t)\big)\Lambda^k_{i+1}(t,x,\lambda)\right|\end{aligned}$$

for $(x,\lambda)\in \mathrm{spt}\Psi^k_i \subseteq \mathrm{B}^Y_{R,\delta}$, the last inclusion being a consequence of Lemma 7.

Arguing as in the proof of (55)–(57) and using (87), we get that

$$\begin{aligned}I_1 &\le L_{v,R}\int_{t^k_i}^{t^k_{i+1}}\left(M_v\Big(1+|x|+|\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)|+m_1(\widetilde{\Psi}^k(\tau))\Big)\right.\\ &\qquad\left.+\left|\frac{\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)-\lambda}{\tau_k}\right|\right)\mathrm{d}\tau\\ &\quad+L_{v,R}\,\tau_k\int_Y\left(\big|v_{\widetilde{\Psi}^k(\tau)}(x',\Lambda^k_{i+1}(t^k_{i+1},x',\lambda'))\big|\right.\\ &\qquad\left.+\left|\frac{\Lambda^k_{i+1}(t^k_{i+1},x',\lambda')-\lambda'}{\tau_k}\right|\right)\mathrm{d}\Psi^k_i(x',\lambda')\\ &\le 4\,L_{v,R}\,M_v(1+R)\tau_k+\frac{4\,L_{E,\delta,R}}{m_1^2}\tau_k = C\,\tau_k\,.\end{aligned} \tag{107}$$

Let us now estimate I_2. By triangle inequality, we have

$$\begin{aligned}I_2 &\le \left|\frac{\big(\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)-\lambda\big)}{\tau_k}-\mathcal{Q}\big(x,\Psi^k_i\big)\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)\right|\\ &\quad+\big|\mathcal{Q}\big(X^k_{i+1}(t,x,\lambda),\Psi^k(t)\big)\Lambda^k_{i+1}(t^k_{i+1},x,\lambda)-\mathcal{Q}\big(x,\Psi^k_i\big)\Lambda^k_{i+1}(t,x,\lambda)\big|\\ &=: I_{2,1}+I_{2,2}\,.\end{aligned} \tag{108}$$

By (iii) of Proposition 5 and by Lemma 7, we have that

$$I_{2,1}\le C\,\tau_k^{1/4}\,. \tag{109}$$

By $(\mathcal{Q}_2)$, (v_3), Lemmas 5 and 7 , and (ii) of Proposition 5, we get

$$\begin{aligned} I_{2,2} &\le L_{\mathcal{Q},R}\bigg(\int_{t_i^k}^{t} \left| v_{\widetilde{\Psi}^k(\tau)}\big(x, \Lambda_{i+1}^k(t_{i+1}^k, x, \lambda)\big)\right| \mathrm{d}\tau \\ &\quad + \int_t^{t_{i+1}^k} \left| \frac{\Lambda_{i+1}^k(t_{i+1}^k, x, \lambda) - \lambda}{\tau_k}\right| \mathrm{d}\tau\bigg) \\ &\le 2L_{\mathcal{Q},R}\, M_v(1+R)\,\tau_k + \frac{2\,L_{E,\delta,R}}{m_1^2}\,\tau_k = C\,\tau_k\,. \end{aligned} \tag{110}$$

Combining (109) and (110), we obtain that

$$I_2 \le C\,\tau_k\,. \tag{111}$$

Finally, equality (105) follows from (107) and (111) as in the proof of Propositions 1 and 4. □

We finally conclude with the main result of this section.

Theorem 3. *Let $r > 0$, $\eta > \delta > 0$, and $\widehat{\Psi} \in \mathcal{P}(\mathrm{B}_{r,\eta}^Y)$. Then, there exists $T_f > 0$ such that the sequence of curves $\Psi^k\colon [0, T_f] \to \mathcal{P}_1(Y)$ converges to the unique solution $\Psi \in C([0, T_f]; \mathcal{P}_1(Y))$ of (7) in W_1, uniformly with respect to $t \in [0, T_f]$*

Proof. Let $T_f > 0$ be as in Lemma 7, so that the curves Ψ^k, $\widetilde{\Psi}^k$, and $\underline{\Psi}^k$ are well defined in the interval $[0, T_f]$ and (105) holds. Since the operator $\mathcal{T}_\Psi(x, \lambda) := \mathcal{Q}(x, \Psi)\lambda$ satisfies the property $(\mathcal{T}_0)$–$(\mathcal{T}_3)$, we only have to check that the sequence Ψ^k is compact in $C([0, T_f]; \mathcal{P}_1(Y))$. The rest of the proof works as for Theorem 1, with the obvious modifications due to the fact that the rest θ_k in Proposition 6 is now controlled by $\tau_k^{1/4}$ and not by τ_k.

In view of Lemma 7, it is enough to show that Ψ^k is equi-Lipschitz with respect to W_1. Let us fix $k \in \mathbb{N}$, $i \in \mathbb{N}$ such that $i\tau_k \le T_f$, and $s \le t \in [t_i^k, t_{i+1}^k]$, and let $R := R(T_f)$. Then,

$$\begin{aligned} W_1(\Psi^k(t), \Psi^k(s)) &= \sup_{\eta \in \mathrm{Lip}_1(Y)} \left\{ \int_Y \eta(x,\lambda)\,\mathrm{d}(\Psi^k(t) - \Psi^k(s))(x,\lambda)\right\} \\ &\le \int_Y \Big(\big|X_{i+1}^k(t,x,\lambda) - X_{i+1}^k(s,x,\lambda)\big| \\ &\quad + \big|\Lambda_{i+1}^k(t,x,\lambda) - \Lambda_{i+1}^k(s,x,\lambda)\big|\Big)\,\mathrm{d}\Psi_i^k(x,\lambda) \\ &\le \int_Y \bigg(\int_s^t \left| v_{\widetilde{\Psi}^k(\tau)}\big(x, \Lambda_{i+1}^k(\tau, x, \lambda)\big)\right| \mathrm{d}\tau \\ &\quad + \int_s^t \left| \frac{\Lambda_{i+1}^k(t_{i+1}^k, x, \lambda) - \lambda}{\tau_k}\right| \mathrm{d}\tau\bigg)\,\mathrm{d}\Psi_i^k(x,\lambda)\,. \end{aligned}$$

Therefore, by (v_2), Lemma 7, and Proposition 5 we get

$$W_1(\Psi^k(t), \Psi^k(s)) \le 2M_v(1+R)|t-s| + \frac{2\,L_{E,\delta,R}}{m_1^2}\,|t-s|,$$

where the constants $L_{E,\delta,R}$ and $m_1 = m_1(R)$ have been determined in Lemmas 3 and 4 , respectively, and are independent of k, i, and t. □

6. Concluding remarks

In this paper, we have proposed a multi-step Lagrangian scheme for spatially inhomogeneous evolutionary games. The scheme is fully explicit, traces the evolutions of positions and labels along the characteristics, and consists of two approximation steps: in the first step the agents update their beliefs on the strategy; in the second step, they update their position accordingly. Theorem 1 in Sect. 3 provides a general convergence result for the proposed scheme. Theorem 2 in Sect. 4 and Theorem 3 in Sect. 5 deal with the special cases of inhomogeneous replicator dynamics and reversible Markov chains, respectively. Differently from Theorem 1, they contain a variant in that the explicit step for the update of the strategies is replaced by an implicit one, based on a minimizing movement justified by the gradient flow structure of the problems at hand.

We notice that we presented the approximation steps in the natural order: Agents tend to update their information on the environment before changing their position. Undertaking the reverse order in the updates would bring to the same convergence result. The partial update in (11) would have to be modified accordingly, namely by placing the *id* map in the second component and using the updated lifted map X^k_{i+1} in the first component (where now x evolves with the "old" label/strategy distribution).

Using an adaptive time step, chosen on the basis of the rate of change of positions and labels/strategies, deserves further analysis. We foresee that the same convergence result given by Theorem 1 holds, provided that the choice of the time step is such that

$$\lim_{k\to\infty} \max_{i\in\{0,\dots,k-1\}} (t^k_{i+1} - t^k_i) = 0,$$

where we notice that $t^k_{i+1} - t^k_i = \tau_k$ is constant in our work.

Acknowledgements

The authors wish to thank Giuseppe Savaré for a useful discussion regarding the content of Sect. 4. MM and FS are members of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Instituto Nazionale di Alta Matematica "Francesco Severi" (INdAM).

Funding Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement.

REFERENCES

[1] Achdou, Y., Capuzzo-Dolcetta, I.: Mean field games: numerical methods. SIAM J. Numer. Anal. **48**(3), 1136–1162 (2010). https://doi.org/10.1137/090758477.

[2] Albi, G., Bongini, M., Rossi, F., Solombrino, F.: Leader formation with mean-field birth and death models. Math. Models Methods Appl. Sci. **29**(4), 633–679 (2019). https://doi.org/10.1142/S0218202519400025.

[3] Albi, G., Pareschi, L., Zanella, M.: Boltzmann-type control of opinion consensus through leaders. Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **372**(2028), 20140138, 18 (2014). https://doi.org/10.1098/rsta.2014.0138.

[4] Albi, G., Pareschi, L., Zanella, M.: Opinion dynamics over complex networks: kinetic modelling and numerical methods. Kinet. Relat. Models **10**(1), 1–32 (2017). https://doi.org/10.3934/krm.2017001.

[5] Ambrosio, L., Fornasier, M., Morandotti, M., Savaré, G.: Spatially inhomogeneous evolutionary games. Comm. Pure Appl. Math. (2021)

[6] Ambrosio, L., Gigli, N., Savaré, G.: Gradient flows in metric spaces and in the space of probability measures, second edn. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2008)

[7] Ambrosio, L., Puglisi, D.: Linear extension operators between spaces of lipschitz maps and optimal transport. J. Reine Angew. Math. (2019). https://doi.org/10.1515/crelle-2018-0037. (published online)

[8] Ambrosio, L., Trevisan, D.: Well-posedness of Lagrangian flows and continuity equations in metric measure spaces. Anal. PDE **7**(5), 1179–1234 (2014). https://doi.org/10.2140/apde.2014.7.1179.

[9] Arens, R.F., Eells Jr., J.: On embedding uniform and topological spaces. Pacific J. Math. **6**, 397–403 (1956).

[10] Bonafini, M., Fornasier, M., Schmitzer, B.: Data-driven entropic spatially inhomogeneous evolutionary games. Preprint (2021)

[11] Bongini, M., Buttazzo, G.: Optimal control problems in transport dynamics. Math. Models Methods Appl. Sci. **27**(3), 427–451 (2017). https://doi.org/10.1142/S0218202517500063.

[12] Brézis, H.: Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York (1973)

[13] Burger, M., Pinnau, R., Totzeck, C., Tse, O., Roth, A.: Instantaneous control of interacting particle systems in the mean-field limit. J. Comput. Phys. **405**, 109181, 20 (2020). https://doi.org/10.1016/j.jcp.2019.109181.

[14] Cañizo, J.A., Carrillo, J.A., Rosado, J.: A well-posedness theory in measures for some kinetic models of collective motion. Math. Models Methods Appl. Sci. **21**(3), 515–539 (2011). https://doi.org/10.1142/S0218202511005131.

[15] Carrillo, J.A., Matthes, D., Wolfram, M.T.: Lagrangian schemes for wasserstein gradient flows. Preprint (2020). arXiv:2003.03803

[16] Cirant, M.: Multi-population mean field games systems with Neumann boundary conditions. J. Math. Pures Appl. (9) **103**(5), 1294–1315 (2015). https://doi.org/10.1016/j.matpur.2014.10.013.

[17] Di Francesco, M., Fagioli, S.: Measure solutions for non-local interaction PDEs with two species. Nonlinearity **26**(10), 2777–2808 (2013). https://doi.org/10.1088/0951-7715/26/10/2777.

[18] Düring, B., Markowich, P., Pietschmann, J.F., Wolfram, M.T.: Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **465**(2112), 3687–3708 (2009). https://doi.org/10.1098/rspa.2009.0239.

[19] Fornasier, M., Piccoli, B., Rossi, F.: Mean-field sparse optimal control. Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **372**(2028), 20130400, 21 (2014). https://doi.org/10.1098/rsta.2013.0400.

[20] Gallouët, T.O., Monsaingeon, L.: A JKO splitting scheme for Kantorovich-Fisher-Rao gradient flows. SIAM J. Math. Anal. **49**(2), 1100–1130 (2017). 10.1137/16M106666X.

[21] Hofbauer, J., Sigmund, K.: Evolutionary games and population dynamics. Cambridge University Press, Cambridge (1998). 10.1017/CBO9781139173179.

[22] Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998). 10.1137/S0036141096303359.

[23] Kondratyev, S., Vorotnikov, D.: Spherical Hellinger-Kantorovich gradient flows. SIAM J. Math. Anal. **51**(3), 2053–2084 (2019). 10.1137/18M1213063.

[24] Lasry, J.M., Lions, P.L.: Mean field games. Jpn. J. Math. **2**(1), 229–260 (2007). 10.1007/s11537-007-0657-8.

[25] Leclerc, H., Mérigot, Q., Santambrogio, F., Stra, F.: Lagrangian discretization of crowd motion and linear diffusion. Preprint (2020). arXiv:1905.08507

[26] Legendre, G., Turinici, G.: Second-order in time schemes for gradient flows in Wasserstein and geodesic metric spaces. C. R. Math. Acad. Sci. Paris **355**(3), 345–353 (2017). 10.1016/j.crma.2017.02.001.

[27] Lim, T.S., Lu, Y., Nolen, J.H.: Quantitative Propagation of Chaos in a Bimolecular Chemical Reaction-Diffusion Model. SIAM J. Math. Anal. **52**(2), 2098–2133 (2020). 10.1137/19M1287687.

[28] Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261**(8), 2250–2292 (2011). 10.1016/j.jfa.2011.06.009.

[29] Matthes, D., Plazotta, S.: A variational formulation of the BDF2 method for metric gradient flows. ESAIM Math. Model. Numer. Anal. **53**(1), 145–172 (2019). 10.1051/m2an/2018045.

[30] Mielke, A.: Geodesic convexity of the relative entropy in reversible Markov chains. Calc. Var. Partial Differential Equations **48**(1-2), 1–31 (2013). 10.1007/s00526-012-0538-8.

[31] Morandotti, M., Solombrino, F.: Mean-field Analysis of Multipopulation Dynamics with Label Switching. SIAM J. Math. Anal. **52**(2), 1427–1462 (2020). 10.1137/19M1273426.

[32] Mozgunov, P., Beccuti, M., Horvath, A., Jaki, T., Sirovich, R., Bibbona, E.: A review of the deterministic and diffusion approximations for stochastic chemical reaction networks. Reac. Kinet. Mech. Cat. (123), 289–312 (2018)

[33] Oelschläger, K.: On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes. Probab. Theory Related Fields **82**(4), 565–586 (1989). 10.1007/BF00341284.

[34] Piccoli, B., Rossi, F.: Measure-theoretic models for crowd dynamics. In: Crowd dynamics. Vol. 1, Model. Simul. Sci. Eng. Technol., pp. 137–165. Birkhäuser/Springer, Cham (2018)

[35] Plazotta, S.: A BDF2-approach for the non-linear Fokker-Planck equation. Discrete Contin. Dyn. Syst. **39**(5), 2893–2913 (2019). 10.3934/dcds.2019120.

[36] Smirnov, S.K.: Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows. Algebra i Analiz **5**(4), 206–238 (1993)

[37] Taylor, P.D., Jonker, L.B.: Evolutionarily stable strategies and game dynamics. Math. Biosci. **40**(1-2), 145–156 (1978). 10.1016/0025-5564(78)90077-9.

[38] Thai, M.N.: Birth and death process in mean field type interaction. Bernoulli (to appear) (2015)

[39] Toscani, G.: Kinetic models of opinion formation. Commun. Math. Sci. **4**(3), 481–496 (2006). http://projecteuclid.org/euclid.cms/1175797553

[40] Weaver, N.: Lipschitz algebras. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ (2018). Second edition of [MR1832645]

[41] Weibull, J.W.: Evolutionary game theory. MIT Press, Cambridge, MA (1995). With a foreword by Ken Binmore

A multi-step Lagrangian scheme for spatially

Stefano Almi
Faculty of Mathematics, University of Vienna
Oskar-Morgenstern-Platz 1
1090 Wien
Austria
E-mail: stefano.almi@univie.ac.at

Marco Morandotti
Dipartimento di Scienze Matematiche
"G. L. Lagrange"
Politecnico di Torino
Corso Duca degli Abruzzi, 24
10129 Torino
Italy
E-mail: marco.morandotti@polito.it

Francesco Solombrino
Dipartimento di Matematica e Applicazioni
"R. Caccioppoli"
Università di Napoli Federico II, via Cintia
80126 Napoli
Italy
E-mail: francesco.solombrino@unina.it

Accepted: 5 April 2021