Euro. Jnl of Applied Mathematics (2019), *vol.* 30, *pp.* 1153–1186 © Cambridge University Press 2019. 1153
doi:10.1017/S0956792519000044

Mean-field optimal control as Gamma-limit of finite agent controls†

M. FORNASIER[1], S. LISINI[2], C. ORRIERI[3] AND G. SAVARÉ[2]

[1] *Department of Mathematics, TU München, Boltzmannstr. 3, Garching bei München D-85748, Germany*
email: massimo.fornasier@ma.tum.de
[2] *Dipartimento di Matematica "F. Casorati", Università di Pavia, Via Ferrata 5, 27100 Pavia, Italy*
email: stefano.lisini@unipv.it; giuseppe.savare@unipv.it
[3] *Dipartimento di Matematica "G. Castelnuovo", Sapienza Università di Roma,*
Piazzale Aldo Moro 5, 00185 Roma, Italy
email: orrieri@mat.uniroma1.it

(*Received 11 March 2018*; *revised 16 December 2018*; *accepted 1 February 2019*; *first published online 8 March 2019*)

This paper focuses on the role of a government of a large population of interacting agents as a mean-field optimal control problem derived from deterministic finite agent dynamics. The control problems are constrained by a Partial Differential Equation of continuity-type without diffusion, governing the dynamics of the probability distribution of the agent population. We derive existence of optimal controls in a measure-theoretical setting as natural limits of finite agent optimal controls without any assumption on the regularity of control competitors. In particular, we prove the consistency of mean-field optimal controls with corresponding underlying finite agent ones. The results follow from a Γ-convergence argument constructed over the mean-field limit, which stems from leveraging the superposition principle.

Key words: Finite agent optimal control, mean-field optimal control, Γ-convergence, superposition principle

2010 Mathematics Subject Classification: Primary: 35Q93; 49K20. Secondary: 35Q83; 35Q70; 35Q91.

1 Introduction

In the mathematical modelling of biological, social and economical phenomena, self-organisation of multi-agent interaction systems has become a focus of applied mathematics and physics and mechanisms are studied towards the formation of global patterns. In the last years, there has been a vigorous development of literature describing collective behaviour of interacting agents [29–31, 40–42, 61], towards modelling phenomena in biology, such as motility and cell aggregation [15, 43, 44, 54], coordinated animal motion [7, 20, 23, 25–27, 31, 49, 51, 52, 56, 60, 64], coordinated human [28, 33, 58] and synthetic agent interactions and behaviour, as in the case of

† Massimo Fornasier acknowledges the financial support provided by the ERC-Starting Grant 'High-Dimensional Sparse Optimal Control' (HDSPCONTR) and the DFG-Project FO 767/7-1 'Identification of Energies from the Observation of Evolutions'. Giuseppe Savaré acknowledges the financial support provided by Cariplo foundation and Regione Lombardia via project 'Variational evolution problems and optimal transport'. Carlo Orrieri acknowledges the financial support provided by PRIN 20155PAWZB 'Large Scale Random Structures'.

cooperative robots [24, 48, 53, 59]. Part of the literature is particularly focused on studying the corresponding mean-field equations in order to simplify models for large populations of interacting agents: the effect of all the other individuals on any given individual is described by a single averaged effect. As it is very hard to be exhaustive in accounting all the developments of this very fast growing field, we refer to Refs. [18, 19, 21, 22, 62] for recent surveys.

Self-organisation is an incomplete concept, see, e.g., Ref. [12], as it is not always occurring when needed. In fact, local interactions between agents can be interpreted as distributed controls, which, however, are not always able to lead to global coordination or pattern formation. This motivated the research also of centralised optimal controls for multi-agent systems, modelling the intervention of an external government to induce desired dynamics or pattern formation. In this paper, we are concerned with the control of deterministic multi-agent systems of the type

$$\dot{x}_i(t) = \boldsymbol{F}^N(x_i(t), \boldsymbol{x}(t)) + u_i(t), \quad i = 1, \ldots, N. \tag{1.1}$$

The map $\boldsymbol{F}^N : \mathbb{R}^d \times (\mathbb{R}^d)^N \to \mathbb{R}^d$ models the interaction between the agents and $\boldsymbol{u}$ represents the action of an external controller on the system. The control is optimised by minimisation of a cost functional

$$\mathcal{E}^N(\boldsymbol{x}, \boldsymbol{u}) := \int_0^T \frac{1}{N} \sum_{i=1}^N L^N(x_i(t), \boldsymbol{x}(t))\, \mathrm{d}t + \int_0^T \frac{1}{N} \sum_{i=1}^N \psi(u_i(t))\, \mathrm{d}t, \tag{1.2}$$

where L^N is a suitable cost function used for modelling the goal of the control and capturing the work done to achieve it, and ψ is an appropriate positive convex function, which is superlinear at infinity and models the effective cost of employing the control. When the number N of agents is very large, dynamical programming for solving the optimal control problem defined by minimisation of (1.2) under the constraints (1.1) becomes computationally intractable. In fact, Richard Bellman coined the term 'curse of dimensionality' precisely to describe this phenomenon.

For situations where agents are indistinguishable, e.g., drawn independently at random from an initial probability distribution μ_0, and the dynamics $\boldsymbol{F}^N(x_i(t), \boldsymbol{x}(t)) = \boldsymbol{F}^N\big(x_i(t), \mu_t^N\big)$ depends in fact from the empirical distribution $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)}$, one may hope to invoke again the use of mean-field approximations for a tractable (approximate) solution of the control problem. By formally considering the mean-field limit of the system (1.1) for $N \to \infty$, one obtains the continuity equation of Vlasov-type:

$$\partial_t \mu_t + \nabla \cdot \big((\boldsymbol{F}(x, \mu_t) + \boldsymbol{v}_t)\mu_t\big) = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \tag{1.3}$$

where μ is the weak limit of μ^N and represents the (time-dependent) probability distribution of agents and $\boldsymbol{\nu} = \boldsymbol{v}\mu$ is a suitable vector control measure absolutely continuous w.r.t. μ and subjected to a cost functional

$$\mathcal{E}(\mu, \boldsymbol{\nu}) := \int_0^T \int_{\mathbb{R}^d} L(x, \mu_t)\, \mathrm{d}\mu_t(x)\, \mathrm{d}t + \int_0^T \int_{\mathbb{R}^d} \psi(\boldsymbol{v}(t, x))\, \mathrm{d}\mu_t(x)\, \mathrm{d}t. \tag{1.4}$$

The vector measure $\boldsymbol{\nu} = \boldsymbol{v}\mu$ can in fact be obtained as the weak limit of the sequence of finite dimensional control measures:

$$\boldsymbol{\nu}^N = \int_0^T \delta_t \otimes \boldsymbol{\nu}_t^N\, \mathrm{d}t, \quad \boldsymbol{\nu}_t^N = \frac{1}{N}\sum_{i=1}^N u_i(t)\delta_{x_i(t)}, \quad t \in [0, T].$$

Under suitable assumptions on ψ, on the convergence of $\boldsymbol{F}^N$ to $\boldsymbol{F}$ and of L^N to L and assuming for simplicity that the initial data are confined in a compact subset of $\mathbb{R}^d$, one of the main results of this paper can be summarised as follows.

Theorem 1.1 *If the initial measures $\mu_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i^N(0)}$ weakly converge to a limit probability measure μ_0, then the minimum $E^N\left(\mu_0^N\right)$ of (1.2) among all the solution of the controlled system (1.1) converges to the minimum $E(\mu_0)$ of the functional (1.4) among all the solutions of (1.3) with initial datum μ_0. Moreover, all the accumulation points (in the topology of weak convergence of measures) of the measures associated with minimisers $\boldsymbol{x}^N$, $\boldsymbol{u}^N$ of (1.2) are minima of (1.4).*

The idea of solving finite agent optimal control problems by considering a mean-field approximation has been considered since the 1960s [36,37,45] with the introduction of stochastic optimal control. The optimal control of stochastic differential equations

$$\mathrm{d}X_t^i = \boldsymbol{F}\left(X_t^i, \mu_t^N\right) + \boldsymbol{v}\left(t, X_t^i\right) + \sigma \mathrm{d}W_t^i, \quad i = 1, \ldots, N,$$

with non-degenerate diffusion and independent Brownian motions W^i has been studied for a long while via the optimal control of the law $\mu_t = \mathrm{Law}(X_t)$ constrained by a McKean–Vlasov equation:

$$\partial_t \mu_t + \nabla \cdot \big((\boldsymbol{F}(x, \mu_t) + \boldsymbol{v}(t, x))\mu_t\big) = \sigma \Delta \mu_t,$$

under a suitable control cost

$$\mathcal{E}(\mu, \boldsymbol{v}) := \int_0^T \int_{\mathbb{R}^d} L(x, \mu_t)\, \mathrm{d}\mu_t(x)\, \mathrm{d}t + \int_0^T \int_{\mathbb{R}^d} \psi(\boldsymbol{v}(t, x))\mathrm{d}\mu_t(x)\, \mathrm{d}t.$$

Most of the literature on stochastic control is focused primarily on the solution of McKean–Vlasov optimal control problems. The most popular methods are based on extending Pontryagin's maximum principle [2, 6, 9, 14, 17] or deriving a dynamic programming principle, and with it a form of a Hamilton–Jacobi–Bellman equation on a space of probability measures [8, 47, 55]. However, the rigorous justification that the McKean–Vlasov optimal control problem is consistent with the limit of optimal controls for stochastic finite agent models has been proved surprisingly just very recently [46]. The techniques used in the latter paper are largely based on martingale problems, combining ideas from the McKean–Vlasov limit theory with a well-established compactification method for stochastic control [36].

The first work addressing the consistency of mean-field optimal control for deterministic finite agent systems is [39]. In the latter paper, an analogous result as Theorem 1.1 is derived for general penalty functions ψ with polynomial growth, including the interesting case of linear growth at 0 and infinity, motivated by results of sparse controllability for finite-agent models [10, 11, 16]. Other models of sparse mean-field optimal control have been considered in Refs. [1, 13, 38]. The generality of the penalty function ψ in Ref. [39] has required to restrict the class of controls: they have been assumed to be locally Lipschitz continuous in space feedback control functions $u_i(t) = \boldsymbol{v}(t, x_i)$ with controlled time-dependent Lipschitz constants.

In this paper and in our main result Theorem 1.1, we remove this restriction, but we still impose suitable coercivity on the admissible controls, by requiring the function ψ to have superlinear growth to infinity. As sparsity of controls, i.e., the localisation of controls in space, is mainly due to the linear behaviour of the penalty function at 0, the superlinear growth at infinity does not exclude the possibility of using this model for sparse control. Moreover, in this framework, there is no need for enforcing a priori that controls are smooth feedback functions of the state variables and the limit process comes very natural in a measure-theoretical sense. In view of the minimal smoothness required to the governing interaction functions $\boldsymbol{F}^N, \boldsymbol{F}$ (they are assumed to be just continuous), there is no uniqueness of solutions in general of (1.1) and (1.3). Hence, the main results of mean-field limit are derived by leveraging the powerful machinery of the superposition principle [3, Theorem 3.4].

Finally, a comparison between deterministic and stochastic cases is in order. The consistency result obtained in [46] allows for degenerate diffusion $\sigma = 0$ (and deterministic initial condition) but does not subsume the results obtained in Ref. [39]. In the present paper, the assumptions on the system can be further weakened by employing the superposition principle, for which a stochastic counterpart is missing. It seems that sharp results for purely deterministic dynamics require in fact measure-theoretical methods, which are difficult to be directly applied to the stochastic setting. Recall also that stochastic control problems are linked to their deterministic counterpart up to non-anticipativity restrictions on the control policy. Nonetheless, deterministic consistency results in control theory could be fruitfully applied to the stochastic setting in the study of large-deviation asymptotics. We refer to Ref. [50] for an application of the present result to the analysis of fluctuations of stochastic interacting particle systems in the mean-field and small-noise regime.

The paper is organised as follows: after recalling in detail the notation and a few preliminary results on optimal transport, doubling functions and convex functionals on measures in Section 2, we describe our setting of optimal control problems in Section 3, together with the precise statements of our main results. We address the existence of solutions of the finite agent optimal control problem in Section 4. Crucial moment estimates are derived in Section 5 for feasible competitors for the mean-field control problem, which are useful for deriving compactness arguments further below. Section 6 is dedicated to the proofs of our main theorems. A relevant part is devoted to Theorem 1.1 by developing a Γ-convergence argument. While the Γ–lim inf inequality follows by relatively standard lower semicontinuity arguments, the derivation of the Γ– lim sup inequality requires a technical application of the superposition principle. Equi-coercivity and convergence of minimisers follow from compactness arguments based on moment estimates from Section 5.

2 Notation and preliminary results

Throughout the paper, we work with $\mathbb{R}^d$ as a state space and we fix a time horizon $T > 0$. We will denote by λ the normalised restriction of the Lebesgue measure to $[0, T]$, $\lambda := \frac{1}{T}\mathscr{L}^1 \llcorner [0, T]$.

Given $(\mathcal{S}, d)$ a metric space, we use the classical notation $AC([0, T]; \mathcal{S})$ for the classes of $\mathcal{S}$-valued absolute continuous curves. We indicate with $\mathcal{M}(\mathbb{R}^d), \mathcal{M}(\mathbb{R}^d; \mathbb{R}^d)$ the space of Borel (vector-valued) measures.

2.1 Probability measures and optimal transport costs

We call $\mathcal{P}(\mathbb{R}^d)$ the space of Borel probability measures. If $f : \Omega \to \mathbb{R}^h$ is a Borel map defined in a Borel subset Ω of $\mathbb{R}^d$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$ is concentrated on Ω, we will denote by $f_\sharp\mu$ the Borel measure in $\mathbb{R}^h$ defined by $f_\sharp\mu(B) := \mu(f^{-1}(B))$, for every Borel subset $B \subset \mathbb{R}^h$.

Whenever $\psi : \mathbb{R}^d \to [0, +\infty]$ is a lower semicontinuous function, we set

$$\mathcal{C}_\psi(\mu_0, \mu_1) := \inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(y - x)\, \mathrm{d}\gamma(x, y) : \gamma \in \Pi(\mu_0, \mu_1) \right\}, \tag{2.1}$$

where $\Pi(\mu_0, \mu_1)$ is the set of the optimal transport plans:

$$\Pi(\mu_0, \mu_1) := \left\{ \gamma \in \mathcal{P}\left(\mathbb{R}^d \times \mathbb{R}^d\right) : \gamma\left(B \times \mathbb{R}^d\right) = \mu_0(B), \gamma\left(\mathbb{R}^d \times B\right) = \mu_1(B) \right.$$
$$\left. \forall\, B \text{ Borel set in } \mathbb{R}^d \right\}.$$

In the particular case, when $\psi(z) := |z|$, $z \in \mathbb{R}^d$, (2.1) defines the L^1-Wasserstein distance

$$W_1(\mu_0, \mu_1) := \inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|\, \mathrm{d}\gamma(x, y) : \gamma \in \Pi(\mu_0, \mu_1) \right\}; \tag{2.2}$$

the infimum in (2.2) is always finite and attained if μ_0, μ_1 belong to the space $\mathcal{P}_1(\mathbb{R}^d)$ of Borel probability measure with finite first-order moment:

$$\mathcal{P}_1(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|\, \mathrm{d}\mu(x) < +\infty \right\}.$$

$\mathcal{P}_1(\mathbb{R}^d)$ endowed with $W_1(\mu_0, \mu_1)$ is a complete and separable metric space. In particular, we will consider absolutely continuous curves $t \mapsto \mu_t$ in $AC([0, T]; \mathcal{P}_1(\mathbb{R}^d))$. They will canonically induce a parameterised measure $\tilde{\mu} := \int \delta_t \otimes \mu_t\, \mathrm{d}\lambda(t)$ in $\mathcal{P}_1([0, T] \times \mathbb{R}^d)$ satisfying

$$\int f(t, x)\, \mathrm{d}\tilde{\mu}(t, x) = \int_0^T \int_{\mathbb{R}^d} f(t, x)\, \mathrm{d}\mu_t(x)\, \mathrm{d}\lambda(t) = \fint_0^T \int_{\mathbb{R}^d} f(t, x)\, \mathrm{d}\mu_t(x)\, \mathrm{d}t. \tag{2.3}$$

Convergence with respect to W_1 is equivalent to weak convergence (in duality with continuous and bounded functions) supplemented with convergence of first moment; equivalently, for every sequence $(\mu_n)_{n\in\mathbb{N}} \subset \mathcal{P}_1(\mathbb{R}^d)$ and candidate limit $\mu \in \mathcal{P}_1(\mathbb{R}^d)$

$$\lim_{n\to\infty} W_1(\mu_n, \mu) = 0 \quad \Leftrightarrow \quad \lim_{n\to\infty} \int \zeta\, \mathrm{d}\mu_n = \int \zeta\, \mathrm{d}\mu \quad \text{for every } \zeta \in C(\mathbb{R}^d),\ \sup_{x\in\mathbb{R}^d} \frac{\zeta(x)}{1 + |x|} < \infty.$$

In $\mathcal{P}_1(\mathbb{R}^d)$, we consider the subset $\mathcal{P}^N(\mathbb{R}^d)$ of discrete measures

$$\mathcal{P}^N(\mathbb{R}^d) := \left\{ \mu = \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \text{ for some } x_i \in \mathbb{R}^d \right\}.$$

A measure μ belongs to $\mathcal{P}^N(\mathbb{R}^d)$ if and only if $\#\,\mathrm{supp}(\mu) \le N$ and $N\mu(B) \in \mathbb{N}$ for every Borel set B of $\mathbb{R}^d$. Let us now fix an integer $N \in \mathbb{N}$ and consider vectors $\boldsymbol{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$; we will use the notation $\sigma : (\mathbb{R}^d)^N \to (\mathbb{R}^d)^N$ to denote a permutation of the coordinates of vectors in $(\mathbb{R}^d)^N$ and we set

$$\mathrm{d}_N(\boldsymbol{x}, \boldsymbol{y}) := \min_\sigma \frac{1}{N} \sum_{i=1}^N |x_i - \sigma(\boldsymbol{y})_i|, \qquad |\boldsymbol{x}|_N := \mathrm{d}_N(\boldsymbol{x}, \boldsymbol{o}) = \frac{1}{N} \sum_{i=1}^N |x_i|.$$

To every vector $\boldsymbol{x} \in (\mathbb{R}^d)^N$, we can associate the measure $\mu[\boldsymbol{x}] := \frac{1}{N}\sum_{i=1}^N \delta_{x_i} \in \mathcal{P}^N(\mathbb{R}^d)$ and we notice that by (6.37) [5, Theorem 6.0.1]

$$\mathrm{d}_N(\boldsymbol{x},\boldsymbol{y}) = W_1(\mu[\boldsymbol{x}],\mu[\boldsymbol{y}]), \quad |\boldsymbol{x}|_N = \int_{\mathbb{R}^d} |x|\,\mathrm{d}\mu[\boldsymbol{x}](x) = W_1(\mu[\boldsymbol{x}],\delta_0).$$

From now on, we say that a map $\boldsymbol{G}^N : \mathbb{R}^d \times (\mathbb{R}^d)^N \to \mathbb{R}^k$ is symmetric if

$$\boldsymbol{G}^N(x,\boldsymbol{y}) = \boldsymbol{G}^N(x,\sigma(\boldsymbol{y})) \quad \text{for every permutation } \sigma : (\mathbb{R}^d)^N \to (\mathbb{R}^d)^N.$$

Given a symmetric and continuous map $\boldsymbol{G}^N$, we can associate a function defined on measures $G^N : \mathbb{R}^d \times \mathcal{P}^N(\mathbb{R}^d) \to \mathbb{R}^k$ by setting

$$G^N(x,\mu[\boldsymbol{y}]) := \boldsymbol{G}^N(x,\boldsymbol{y}).$$

Throughout the paper, we use the following notion of convergence for symmetric maps:

Definition 2.1 We say that a sequence of symmetric maps G^N, $N \in \mathbb{N}$, $\mathcal{P}_1$-converges to $G : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^k$ uniformly on compact sets as $N \to +\infty$ if for every sequence of measure $\mu_k \in \mathcal{P}^{N_k}(\mathbb{R}^d)$ converging to μ in $\mathcal{P}_1(\mathbb{R}^d)$ as $N_k \to \infty$, we have

$$\lim_{k\to+\infty} \sup_{x\in C} \left|G^{N_k}(x,\mu_k) - G(x,\mu)\right| = 0, \quad \text{for every compact } C \subset \mathbb{R}^d.$$

2.2 Doubling and moderated convex functions

Definition 2.2 We say that $\phi : [0,+\infty) \to [0,+\infty)$ is an *admissible* function if $\phi(0) = 0$, ϕ is strictly convex and of class C^1 with $\phi'(0) = 0$, superlinear at $+\infty$ and doubling, i.e., there exists $K > 0$ such that

$$\phi(2r) \le K\big(1 + \phi(r)\big) \quad \text{for any } r \in [0,+\infty). \tag{2.4}$$

Let U be a subspace of $\mathbb{R}^d$. We say that a convex function $\psi : U \to [0,+\infty)$ is *moderated* if there exists an admissible function $\phi : [0,+\infty) \to [0,+\infty)$ and a constant $C > 0$ such that

$$\phi(|x|) - 1 \le \psi(x) \le C(1 + \phi(|x|)) \quad \text{for every } x \in U. \tag{2.5}$$

By convexity, an admissible function ϕ satisfies $\phi(r) + \phi'(r)(s-r) \le \phi(s)$ for every $r,s \in [0,+\infty)$; in particular, choosing $s = 0$ and $s = 2r$, one obtains

$$0 \le \phi(r) \le r\phi'(r) \le (\phi(2r) - \phi(r)) \le K\big(1 + \phi(r)\big) \quad \text{for every } r \in [0,+\infty). \tag{2.6}$$

It is not difficult to see that if a differentiable convex function ϕ satisfies

$$r\phi'(r) \le A(1 + \phi(r)) \quad \text{for every } r \ge R,$$

for some constants $A, R > 0$, then ϕ satisfies (2.4) with $K = \max(\mathrm{e}^A, \max_{[0,2R]}\phi)$. In fact, differentiating the function $z \mapsto (\phi(zr) + 1)$ for $z \in [1,D]$ and $r \ge R$, we get $\frac{\partial}{\partial\theta}\big(\phi(\theta r) + 1\big) = r\phi'(\theta r) \le A(1 + \phi(\theta r))$ so that

$$\phi(Dr) \le (\phi(r) + 1)\mathrm{e}^{(D-1)A} \quad D > 1,\ r > R.$$

In particular, (2.4) yields

$$\phi(Dr) \le (\phi(r)+1)\mathrm{e}^{(D-1)K} \quad D>1,\ r>0.$$

We also recall that ϕ' is monotone, i.e.,

$$(\phi'(r)-\phi'(s))(r-s) \ge 0 \quad \text{for every } s, r \ge 0. \tag{2.7}$$

The next lemma shows that it is always possible to approximate a convex superlinear function by a monotonically increasing sequence of moderated ones.

Lemma 2.3 *Let U be a subspace of $\mathbb{R}^d$ and $\psi : U \to [0, +\infty]$ be a superlinear function with $\psi(0)=0$.*

(1) There exists an admissible function $\theta : [0, +\infty) \to [0, +\infty)$ such that

$$\psi(x) \ge \theta(|x|) - \frac{1}{2} \quad \textit{for every } x \in U. \tag{2.8}$$

(2) If ψ is also convex, then there exists a sequence $\psi^N : U \to [0, +\infty)$, $N \in \mathbb{N}$, of moderated convex functions *such that*

$$\psi^N(x) \le \psi^{N+1}(x), \qquad \psi^N(x) \uparrow \psi(x) \quad \textit{as } N \to +\infty \quad \textit{for every } x \in U.$$

Proof It is not restrictive to assume $U = \mathbb{R}^d$.

Claim 1. Let us set $h(r) := \min_{|x|\ge r} \psi(x)$ and $\bar n := \min\left\{n \ge 0 : h(2^n) \ge 1\right\}$, $\bar r := 2^{\bar n}$. The map $h : [0, +\infty) \to [0, +\infty]$ is increasing, lower semicontinuous and satisfies $\lim_{r\to\infty} h(r)/r = +\infty$. By a standard result of convex analysis (see, e.g., [57, Lemma 3.7]), there exists a convex superlinear function $k : [0, +\infty) \to [0, +\infty)$ such that $h(r) \ge k(r)$ for every $r \in [0, +\infty)$, so that $\psi(x) \ge k(|x|)$ for every $x \in \mathbb{R}^d$.

Let us define the sequence $(a_n)_{n\in\mathbb{N}}$ by induction:

$$\begin{aligned}
a_n &:= 0 \quad \text{for every } n \in \mathbb{N},\ n < \bar n; \quad a_{\bar n} := 2^{-\bar n},\\
a_{n+1} &:= \min\left(2a_n, 2^{-(n-1)}\left(k(2^n) - k\left(2^{n-1}\right)\right)\right) \quad \text{for every } n \ge \bar n.
\end{aligned}$$

Since k is convex and increasing, the sequence $n \mapsto a_n$ is positive and increasing; since k is superlinear, it is also easy to check that $\lim_{n\to\infty} a_n = +\infty$.

We now consider the piecewise linear continuous function $\theta_1 : [0, +\infty) \to [0, +\infty)$ on the dyadic partition $\{0, 2^0, 2^1, 2^2, \dots, 2^n, \dots\}$, $n \in \mathbb{N}$, satisfying

$$\theta_1(r) \equiv 0 \quad \text{if } 0 \le r \le \bar r = 2^{\bar n}, \quad \theta_1'(r) = a_n \quad \text{if } 2^n < r < 2^{n+1}, \quad n \in \mathbb{N},\ n \ge \bar n.$$

Since $\theta_1' \le k'$ a.e. in $[0, +\infty)$, we have $\theta_1 \le k$. Moreover, by construction, for $0 \le r \le 2\bar r$, we have $\theta_1(r) \le \theta_1(2\bar r) = 1$ and $\theta_1'(2r) \le 2\theta_1'(r)$ if $r \ge \bar r$, so that θ_1 is also doubling since

$$\theta_1(2r) = \theta_1(2\bar r) + \int_{\bar r}^{r} 2\theta_1'(2s)\,\mathrm{d}s \le 1 + 4\int_{\bar r}^{r} \theta_1'(s)\,\mathrm{d}s = 1 + 4\theta_1(r) \quad \text{for every } r \ge \bar r.$$

Replacing now θ_1 by the convex combination $\theta_2(r) := \frac{1}{2}\theta_1(r) + \frac{1}{2}r/\bar r$, we get a strictly increasing function, still satisfying (2.8).

By possibly replacing θ_2 with $\theta_3(r) := \int_{r-1}^{r} \theta(s)\,\mathrm{d}s$ (where we set $\theta_2(s) \equiv \theta_2(0) = 0$ whenever $s < 0$), we obtain a C^1 function. Strict convexity can be eventually obtained by taking the convex combination $\theta(r) := (1-\varepsilon)\theta_3(r) + \varepsilon(\sqrt{1+r^2} - 1)$ for a sufficiently small $\varepsilon > 0$.

Claim 2. Notice that the function $x \mapsto \theta_2(|x|)$ is convex. We can define ψ^N by inf-convolution:

$$\psi^N(x) := \inf_{y\in\mathbb{R}^d} \psi(y) + N\theta_2(|x-y|), \quad x \in \mathbb{R}^d. \tag{2.9}$$

It is easy to check that the infimum in (2.9) is attained, ψ^N is convex (since it is the inf-convolution of two convex functions) and satisfies the obvious bounds

$$\psi^N(x) \le N\theta_2(|x|), \quad \psi^N(x) \le \psi(x), \quad \psi^N(x) \le \psi^{N+1}(x) \quad \text{for every } x \in \mathbb{R}^d.$$

In particular, ψ^N is continuous; since $x \mapsto \theta_2(x)$ is continuous at $x = 0$ and $\theta_2(|x|) \ge \frac{1}{2\bar r}|x|$, we easily get $\lim_{N\to\infty} \psi^N(x) = \psi(x)$ for every $x \in \mathbb{R}^d$.

It remains to show that ψ^N is moderated. Since $\psi(x) \ge \theta_2(|x|) - 1/2$ and for every $y \in \mathbb{R}^d$ the triangle inequality yields $\min(|x-y|, |y|) \ge |x|/2$, we get

$$\psi^N(x) + 1/2 \ge \inf_{y\in\mathbb{R}^d} \theta_2(|y|) + N\theta_2(|x-y|) \ge \theta_2(|x|/2) \ge \frac{1}{4}\theta_2(|x|) - \frac{1}{4}$$

and the bounds

$$\frac{1}{4}\theta(|x|) - \frac{3}{4} \le \psi^N(x) \le 4N\frac{1}{4}\theta_2(|x|).$$

By possibly replacing θ_2 with θ, we conclude. □

Let us make explicit two simple applications of the properties of Definition 2.2.

Remark 2.4 *If $\mathcal{K} \subset \mathcal{P}_1(\mathbb{R}^d)$ is a relatively compact set and $\psi : U \to [0, +\infty]$ is a superlinear function defined in a subspace U of $\mathbb{R}^d$ with $\psi(0) = 0$, then there exists an admissible function $\theta : [0, +\infty) \to [0, +\infty)$ such that*

$$\sup_{\mu\in\mathcal{K}} \int_{\mathbb{R}^d} \theta(|x|)\,\mathrm{d}\mu(x) < \infty, \quad \theta(|x|) \le 1 + \psi(x) \quad \text{for every } x \in U. \tag{2.10}$$

In fact, Prokhorov theorem yields the tightness of the set $\tilde{\mathcal{K}} := \{|x|\mu : \mu \in \mathcal{K}\}$ of finite measures, so that we can find a superlinear function $\alpha : \mathbb{R}^d \to [0, \infty)$ such that

$$\sup_{\mu\in\mathcal{K}} \int_{\mathbb{R}^d} \alpha(x)\,\mathrm{d}\mu(x) < \infty.$$

We can then apply the first statement of Lemma 2.3 with superlinear function $\alpha \wedge \psi$.

Lemma 2.5 *Let $\zeta : \mathbb{R}^d \to [0, +\infty)$ be a moderated convex function with $\zeta(0) = 0$ and let $\mu_n^i \in \mathcal{P}_1(\mathbb{R}^d)$, $i = 0, 1$, be two sequences converging to μ in $\mathcal{P}_1(\mathbb{R}^d)$ and let γ_n be the optimal plan attaining the minimum in (2.2) for $W_1(\mu_n^0, \mu_n^1)$. If*

$$\limsup_{n\to\infty} \int \zeta\,\mathrm{d}\mu_n^i \le \int \zeta\,\mathrm{d}\mu, \tag{2.11}$$

then

$$\lim_{n\to\infty}\int \zeta(y-x)\,\mathrm{d}\gamma_n(x,y)=0,\qquad \lim_{n\to\infty}\mathcal{C}_\zeta\left(\mu_n^0,\mu_n^1\right)=0. \tag{2.12}$$

Proof Let ϕ be an admissible function satisfying (2.5) for $\psi:=\zeta$. We observe that for every $x,y\in\mathbb{R}^d$

$$\phi(|y-x|)\le\phi(|x|+|y|)\le K\Big(1+\phi\left(\tfrac{1}{2}|x|+\tfrac{1}{2}|y|\right)\Big)\le K\Big(1+\phi(|x|)+\phi(|y|)\Big). \tag{2.13}$$

Inequality (2.11) shows that ζ is uniformly integrable w.r.t. μ_n (see [5, Lemma 5.1.7]) so that

$$\lim_{n\to\infty}\int\phi\big(|x|\big)\,\mathrm{d}\mu_n^i(x)=\int\phi\big(|x|\big)\,\mathrm{d}\mu(x),\qquad i=1,2,$$

whence

$$\lim_{n\to\infty}\int\Big(\phi(|x|)+\phi(|y|)\Big)\,\mathrm{d}\gamma_n(x,y)=2\int\phi(|x|)\,\mathrm{d}\mu(x)=\int\Big(\phi(|x|)+\phi(|y|)\Big)\,\mathrm{d}\gamma(x,y)$$

where $\gamma:=(x,x)_\sharp\mu$ is the weak limit of γ_n. It follows that the function $(x,y)\mapsto\phi(|x|)+\phi(|y|)$ is uniformly integrable with respect to γ_n so that, by (2.13) and [5, Lemma 5.1.7]

$$\lim_{n\to\infty}\int\phi(|y-x|)\,\mathrm{d}\gamma_n(x,y)=\int\phi(|y-x|)\,\mathrm{d}\gamma(x,y)=0.$$

Since $\zeta(y-x)\le C(1+\phi(|y-x|))$ by (2.5), we get (2.12). □

2.3 Convex functionals on measures

We are concerned with the main properties of functionals defined on measures, for a detailed treatment of this subject, we refer to [4]. Let $\psi:\mathbb{R}^h\to[0,+\infty]$ be a proper, l.s.c., convex and superlinear function, so that its recession function $\sup_{r>0}\frac{\psi(rx)}{r}=\infty$ for all $x\neq0$; we will also assume $\psi(0)=0$.

Let now Ω be an open subset of some Euclidean space, $\mu\in\mathcal{M}^+(\Omega)$ be a reference measure and $\boldsymbol{\nu}\in\mathcal{M}(\Omega;\mathbb{R}^h)$ a vector measure; we define the following functional:

$$\Psi(\boldsymbol{\nu}|\mu):=\int_\Omega\psi(\boldsymbol{v}(x))\,\mathrm{d}\mu(x)\quad\text{if }\boldsymbol{\nu}=\boldsymbol{v}\mu\ll\mu,\qquad\Psi(\boldsymbol{\nu}|\mu):=+\infty\quad\text{if }\boldsymbol{\nu}\not\ll\mu. \tag{2.14}$$

We state the main lower semicontinuity result for the functional Ψ.

Theorem 2.6 *Suppose that we have two sequences $\mu_n\in\mathcal{M}^+(\Omega)$, $\boldsymbol{\nu}_n\in\mathcal{M}(\Omega;\mathbb{R}^h)$ weakly converging to $\mu\in\mathcal{M}^+(\Omega)$ and $\boldsymbol{\nu}\in\mathcal{M}(\Omega,\mathbb{R}^h)$, respectively. Then*

$$\liminf_{n\to+\infty}\Psi(\boldsymbol{\nu}_n|\mu_n)\ge\Psi(\boldsymbol{\nu}|\mu).$$

In particular, if $\liminf_{n\to+\infty}\Psi(\boldsymbol{\nu}_n|\mu_n)<+\infty$*, we have* $\boldsymbol{\nu}\ll\mu$.

The proof can be found in Ref. [5, Lemma 9.4.3].

3 The optimal control problem and main results

Cost functional

Assume that we are given a sequence of functions $L^N : \mathbb{R}^d \times (\mathbb{R}^d)^N \to [0, +\infty)$, $N \in \mathbb{N}$, and a function $L : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to [0, +\infty)$ such that L^N is continuous and symmetric for every $N \in \mathbb{N}$ and L is continuous. We assume that

$$L^N \ \mathcal{P}_1\text{-converges to } L \text{ uniformly on compact sets, as } N \to \infty, \tag{3.1}$$

in the sense of Definition 2.1.

Assume that we are given

a subspace $U \subset \mathbb{R}^d$ and a moderated convex function $\psi : U \to [0, +\infty)$ with $\psi(0) = 0$.

We will also fix an auxiliary function ϕ satisfying (2.5).

Typical examples we consider for ψ include

- $\psi(x) = \frac{1}{p}|x|^p$, $p > 1$;
- $\psi(x) = \frac{1}{p}|x|$ for $|x| \le 1$ and $\psi(x) = \frac{1}{p}|x|^p$ for $|x| > 1$, $p > 1$.

Denoting by U^N the Cartesian product, we define a cost functional $\mathcal{E}^N : AC([0, T]; (\mathbb{R}^d)^N) \times L^1([0, T]; U^N) \to [0, +\infty)$ by

$$\mathcal{E}^N(\boldsymbol{x}, \boldsymbol{u}) := \int_0^T \frac{1}{N} \sum_{i=1}^N L^N(x_i(t), \boldsymbol{x}(t))\, dt + \int_0^T \frac{1}{N} \sum_{i=1}^N \psi(u_i(t))\, dt. \tag{3.2}$$

We consider also another cost functional $\mathcal{E} : AC([0, T]; \mathcal{P}_1(\mathbb{R}^d)) \times \mathcal{M}([0, T] \times \mathbb{R}^d; U) \to [0, +\infty)$ defined by (recall (2.3))

$$\mathcal{E}(\mu, \boldsymbol{\nu}) := \int_0^T \int_{\mathbb{R}^d} L(x, \mu_t)\, d\mu_t(x)\, dt + \Psi(\boldsymbol{\nu}|\tilde{\mu}), \tag{3.3}$$

where Ψ is defined as in (2.14). Notice that if $\Psi(\boldsymbol{\nu}|\tilde{\mu}) < \infty$, then $\boldsymbol{\nu} = \boldsymbol{v}\mu$ for a Borel vector field $\boldsymbol{v} \in L^1_{\tilde{\mu}}([0, T] \times \mathbb{R}^d; U)$ so that for λ-a.e. $t \in [0, T]$, the measure $\boldsymbol{\nu}_t := \boldsymbol{v}(t, \cdot)\mu_t$ belongs to $\mathcal{M}(\mathbb{R}^d; U)$ and we can write

$$\Psi(\boldsymbol{\nu}|\tilde{\mu}) = \int_{[0,T]\times\mathbb{R}^d} \psi(\boldsymbol{v}(t, x))\, d\tilde{\mu}(t, x) = \int_0^T \int_{\mathbb{R}^d} \psi(\boldsymbol{v}(t, x))\, d\mu_t\, dt = \int_0^T \Psi(\boldsymbol{\nu}_t|\mu_t)\, dt. \tag{3.4}$$

We shall prove below that the functional $\mathcal{E}$ is the Γ-limit of $\mathcal{E}^N$ in suitable sense [34].

The constraints (state equations)

Assume that we are given a sequence of functions $\boldsymbol{F}^N : \mathbb{R}^d \times (\mathbb{R}^d)^N \to \mathbb{R}^d$, $N \in \mathbb{N}$, symmetric and continuous and a continuous function $\boldsymbol{F} : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$. We assume that there exist constants $A, B \ge 0$ such that

$$|\boldsymbol{F}^N(x, \boldsymbol{y})| \le A + B(|x| + |\boldsymbol{y}|_N), \qquad |\boldsymbol{F}(x, \mu)| \le A + B\left(|x| + \int_{\mathbb{R}^d} |y|\, d\mu(y)\right), \tag{3.5}$$

and $\boldsymbol{F}^N$, $\boldsymbol{F}$ and U satisfy the *compatibility* condition

$$\boldsymbol{F}^N(x,\boldsymbol{y}) - \boldsymbol{F}(x,\mu) \in U \quad \text{for every } x \in \mathbb{R}^d,\ \boldsymbol{y} \in (\mathbb{R}^d)^N,\ \mu \in \mathcal{P}_1(\mathbb{R}^d). \tag{3.6}$$

Moreover, we assume that

$$\boldsymbol{F}^N\ \mathcal{P}_1\text{-converges to } \boldsymbol{F} \text{ uniformly on compact sets, as } N \to \infty, \tag{3.7}$$

in the sense of Definition 2.1.

Given $\boldsymbol{u} = (u_1, \dots, u_N) \in L^1([0,T];U^N)$, a control map, we consider the system of differential equations:

$$\dot{x}_i(t) = \boldsymbol{F}^N(x_i(t),\boldsymbol{x}(t)) + u_i(t), \qquad i = 1, \dots, N. \tag{3.8}$$

The map $\boldsymbol{F}^N : \mathbb{R}^d \times (\mathbb{R}^d)^N \to \mathbb{R}^d$ models the interaction between the agents and $\boldsymbol{u}$ represents the action of an external controller on the system. For every $\boldsymbol{u} \in L^1([0,T];U^N)$ and $\boldsymbol{x}_0 \in (\mathbb{R}^d)^N$, thanks to (3.5) and the continuity of $\boldsymbol{F}^N$, there exists a global solution in the Carathéodory sense, $\boldsymbol{x} \in AC([0,T];(\mathbb{R}^d)^N)$ of (3.8) such that $\boldsymbol{x}(0) = \boldsymbol{x}_0$. Since we have assumed only the continuity of the velocity field $\boldsymbol{F}^N$, uniqueness of solutions is not guaranteed in general. We then define the non-empty set:

$$\mathscr{A}^N := \left\{(\boldsymbol{x},\boldsymbol{u}) \in AC([0,T];\left(\mathbb{R}^d\right)^N \times L^1\left([0,T];U^N\right) :\ \boldsymbol{x} \text{ and } \boldsymbol{u} \text{ satisfy (3.8)}, \mathcal{E}^N(\boldsymbol{x},\boldsymbol{u}) < \infty\right\}.$$

Moreover, we also define for every $\boldsymbol{x}_0 \in (\mathbb{R}^d)^N$ the non-empty set

$$\mathscr{A}^N(\boldsymbol{x}_0) := \left\{(\boldsymbol{x},\boldsymbol{u}) \in \mathscr{A}^N : \boldsymbol{x}(0) = \boldsymbol{x}_0\right\}.$$

Every initial vector $\boldsymbol{x}_0 = (x_{0,1}, \dots, x_{0,N}) \in (\mathbb{R}^d)^N$ gives rise to the empirical distribution:

$$\mu_0 = \mu[\boldsymbol{x}_0] := \frac{1}{N}\sum_{i=1}^N \delta_{x_{0,i}}.$$

Similarly, every curve $\boldsymbol{x} \in AC([0,T];(\mathbb{R}^d)^N)$ is associated with the curve of probability measures:

$$\mu = \mu[\boldsymbol{x}] \in AC\left([0,T];\mathcal{P}_1\left(\mathbb{R}^d\right)\right): \qquad \mu_t := \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)}, \quad t \in [0,T],$$

and every pair $(\boldsymbol{x},\boldsymbol{u}) \in AC([0,T];(\mathbb{R}^d)^N) \times L^1([0,T];U^N)$ is linked to the *control* vector measure:

$$\boldsymbol{\nu} = \boldsymbol{\nu}[\boldsymbol{x},\boldsymbol{u}] \in \mathcal{M}([0,T]\times\mathbb{R}^d;U): \quad \boldsymbol{\nu} := \int_0^T \delta_t \otimes \boldsymbol{\nu}_t \,\mathrm{d}\lambda, \quad \boldsymbol{\nu}_t := \frac{1}{N}\sum_{i=1}^N u_i(t)\delta_{x_i(t)}.$$

We will show that for every choice of solutions and controls $(\boldsymbol{x}^N,\boldsymbol{u}^N) \in \mathscr{A}^N(\boldsymbol{x}_0^N)$ such that the cost functional $\mathcal{E}^N(\boldsymbol{x}^N,\boldsymbol{u}^N)$ remains uniformly bounded and the initial empirical distributions $\mu_0^N = \mu[\boldsymbol{x}_0^N]$ are converging to a limit measure μ_0 in $\mathcal{P}_1(\mathbb{R}^d)$ a mean-field approximation holds:

Theorem 3.1 (Compactness) *Let $(\boldsymbol{x}_0^N)_{N\in\mathbb{N}}$ be a sequence of initial data in $(\mathbb{R}^d)^N$ such that the empirical measure $\mu_0^N = \mu[\boldsymbol{x}_0^N]$ converges to a probability measure μ_0 in $\mathcal{P}_1(\mathbb{R}^d)$ as $N \to \infty$, and let $(\boldsymbol{x}^N,\boldsymbol{u}^N) \in \mathscr{A}^N(\boldsymbol{x}_0^N)$ such that the cost functional $\mathcal{E}^N(\boldsymbol{x}^N,\boldsymbol{u}^N)$ remains uniformly bounded.*

Up to extraction of a suitable subsequence, the empirical measures $\mu^N = \mu[\boldsymbol{x}^N]$ *converge uniformly in* $\mathcal{P}_1(\mathbb{R}^d)$ *to a curve of probability measures* $\mu \in AC([0,T];\mathcal{P}_1(\mathbb{R}^d))$*, the control measures* $\boldsymbol{\nu}^N = \boldsymbol{\nu}[\boldsymbol{x}^N, \boldsymbol{u}^N]$ *converge to a limit control measure* $\boldsymbol{\nu}$ *weakly* in* $\mathcal{M}([0,T]\times\mathbb{R}^d; U)$ *and* $(\mu, \boldsymbol{\nu})$ *fulfils the continuity equation:*

$$\partial_t \mu_t + \nabla \cdot \Big(\boldsymbol{F}(x, \mu_t)\mu_t + \boldsymbol{\nu}_t \Big) = 0 \quad \textit{in } (0,T)\times\mathbb{R}^d \tag{3.9}$$

in the sense of distributions.

Motivated by the above result, we define the non-empty set

$$\begin{aligned}\mathscr{A} := \Big\{ &(\mu, \boldsymbol{\nu}) \in AC([0,T];\mathcal{P}_1(\mathbb{R}^d)) \times \mathcal{M}([0,T]\times\mathbb{R}^d; U) : \\ &\mu \text{ and } \boldsymbol{\nu} \text{ satisfy (3.9) in the sense of distributions, } \mathcal{E}(\mu,\boldsymbol{\nu}) < \infty \Big\},\end{aligned}$$

and its corresponding subset associated with a given initial measure $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d)$:

$$\mathscr{A}(\mu_0) := \{(\mu, \boldsymbol{\nu}) \in \mathscr{A} : \mu(0) = \mu_0\}.$$

The elements of $\mathscr{A}^N$ can be interpreted as the trajectories $(x_1, \ldots, x_N)$ of N agents along with their strategies $(u_1, \ldots, u_N)$, whose dynamics is described by the system of Ordinary Differential Equations (ODEs) (3.8). Analogously, the elements of $\mathscr{A}$ can be interpreted as the trajectories of a continuous or discrete distribution of agents whose dynamics is described by the Partial Differential Equation (3.9) under the action of an external controller described by the measure $\boldsymbol{\nu}$.

The minimum problems

The objective of the controller is to minimise the cost functional $\mathcal{E}^N$ (resp. $\mathcal{E}$). We consider the following optimum sets, defined by the corresponding optimal control problems:

$$E^N(\boldsymbol{x}_0) := \min_{(\boldsymbol{x},\boldsymbol{u})\in\mathscr{A}^N(\boldsymbol{x}_0)} \mathcal{E}^N(\boldsymbol{x}, \boldsymbol{u}), \qquad P^N(\boldsymbol{x}_0) := \operatorname{argmin}\left\{ \mathcal{E}^N(\boldsymbol{x},\boldsymbol{u}) : (\boldsymbol{x},\boldsymbol{u}) \in \mathscr{A}^N(\boldsymbol{x}_0) \right\}, \tag{3.10}$$

$$E(\mu_0) := \min_{(\mu,\boldsymbol{\nu})\in\mathscr{A}(\mu_0)} \mathcal{E}(\mu, \boldsymbol{\nu}), \qquad P(\mu_0) := \operatorname{argmin}\{\mathcal{E}(\mu,\boldsymbol{\nu}) : (\mu,\boldsymbol{\nu}) \in \mathscr{A}(\mu_0)\}, \tag{3.11}$$

where we suppose that $\mu_0 \in D(E) := \{\mu \in \mathcal{P}_1(\mathbb{R}^d) : \mathscr{A}(\mu) \text{ is not empty}\}$.

We are interested in the rigorous justification of the convergence of the control problem (3.10) towards the corresponding infinite dimensional one (3.11).

Main results

We state now more formally our main result concerning the sequence of functionals $\mathcal{E}^N$ to $\mathcal{E}$, inspired to Γ-convergence.

Theorem 3.2 (Γ-convergence) *The following properties hold:*

- $\Gamma-\liminf$ ***inequality****: for every* $(\mu, \boldsymbol{\nu}) \in AC([0,T];\mathcal{P}_1(\mathbb{R}^d)) \times \mathcal{M}([0,T]\times\mathbb{R}^d; U)$ *and every sequence* $(\boldsymbol{x}^N, \boldsymbol{u}^N) \in AC([0,T];(\mathbb{R}^d)^N) \times L^1([0,T];U^N)$ *such that* $\mu[\boldsymbol{x}^N] \to \mu$ *in* $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$*,* $\boldsymbol{\nu}[\boldsymbol{x}^N, \boldsymbol{u}^N] \rightharpoonup^* \boldsymbol{\nu}$ *in* $\mathcal{M}([0,T]\times\mathbb{R}^d; U)$*, we have*

$$\liminf_{N\to\infty} \mathcal{E}^N\left(\boldsymbol{x}^N, \boldsymbol{u}^N\right) \geq \mathcal{E}(\mu, \boldsymbol{\nu}). \tag{3.12}$$

- $\Gamma-\limsup$ ***inequality***: *for every* $(\mu, \boldsymbol{\nu}) \in \mathscr{A}$ *such that*

$$\int_{\mathbb{R}^d} \phi(|x|)\, \mathrm{d}\mu_0(x) < \infty, \tag{3.13}$$

there exists a sequence $(\boldsymbol{x}^N, \boldsymbol{u}^N) \in \mathscr{A}^N$ *with* $x_{0,i}^N \in \operatorname{supp}(\mu_0)$ *for every* $i = 1, \ldots, N$*, such that*

$$\mu\left[\boldsymbol{x}^N\right] \to \mu \text{ in } C\left([0,T]; \mathcal{P}_1\left(\mathbb{R}^d\right)\right), \quad \boldsymbol{\nu}\left[\boldsymbol{x}^N, \boldsymbol{u}^N\right] \rightharpoonup^* \boldsymbol{\nu} \text{ in } \mathcal{M}([0,T] \times \mathbb{R}^d; U), \tag{3.14}$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \phi(|x_{0,i}^N|) = \int_{\mathbb{R}^d} \phi(|x|)\, \mathrm{d}\mu_0(x), \tag{3.15}$$

and

$$\limsup_{N\to\infty} \mathcal{E}^N\left(\boldsymbol{x}^N, \boldsymbol{u}^N\right) \le \mathcal{E}(\mu, \boldsymbol{\nu}). \tag{3.16}$$

As a combination of Theorems 3.1 and 3.2, we obtain the convergence of minima.

Theorem 3.3 *Let* $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d)$ *be satisfying* (3.13).

(1) There exists a sequence $\boldsymbol{x}_0^N \in (\mathbb{R}^d)^N$, $N \in \mathbb{N}$*, satisfying*

$$\lim_{N\to\infty} W_1\left(\mu\left[\boldsymbol{x}_0^N\right], \mu_0\right) = 0, \tag{3.17}$$

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \phi\left(\left|x_{0,i}^N\right|\right) = \int \phi(|x|)\, \mathrm{d}\mu_0(x), \tag{3.18}$$

$$\lim_{N\to\infty} E^N\left(\boldsymbol{x}_0^N\right) = E(\mu_0). \tag{3.19}$$

(2) If a sequence $\boldsymbol{x}_0^N$ *satisfies* (3.17)*, then for every choice of* $(\boldsymbol{x}^N, \boldsymbol{u}^N) \in P(\boldsymbol{x}_0^N)$ *with* $\mu^N := \mu[\boldsymbol{x}^N]$ *and* $\boldsymbol{\nu}^N := \boldsymbol{\nu}[\boldsymbol{x}^N, \boldsymbol{u}^N]$*, the collection of limit points* $(\mu, \boldsymbol{\nu})$ *of* $(\mu^N, \boldsymbol{\nu}^N)$ *in* $C([0,T]; \mathcal{P}_1(\mathbb{R}^d)) \times \mathcal{M}([0,T] \times \mathbb{R}^d; U)$ *is non-empty and contained in* $P(\mu_0)$.

(3) If moreover $U = \mathbb{R}^d$ *and* μ_0 *has compact support, then every sequence* $(\boldsymbol{x}_0^N)_{N\in\mathbb{N}}$ *satisfying* (3.17) *and uniformly supported in a compact set also satisfies* (3.18) *and* (3.19).

3.1 Examples

First-order examples

Take a continuous function $H : \mathbb{R}^d \to \mathbb{R}^d$ satisfying

$$|H(x)| \le A + B|x| \qquad \forall x \in \mathbb{R}^d$$

and set

$$\boldsymbol{F}^N(x, \boldsymbol{y}) := \frac{1}{N} \sum_{j=1}^N H(x - y_j) = \int_{\mathbb{R}^d} H(x-y)\, \mathrm{d}\mu[\boldsymbol{y}](y)$$

and

$$\boldsymbol{F}(x, \mu) := \int_{\mathbb{R}^d} H(x-y)\, \mathrm{d}\mu(y).$$

 https://doi.org/10.1017/S0956792519000044

When $H = -\nabla W$ for an even function $W \in C^1(\mathbb{R}^d)$, the system (3.8) is associated with the gradient flow of the interaction energy $\mathcal{W} : (\mathbb{R}^d)^N \to \mathbb{R}$ defined by

$$\mathcal{W}(\boldsymbol{x}) := \frac{1}{2N^2} \sum_{i,j=1}^{N} W(x_i - x_j)$$

with respect to the weighted norm $\|\boldsymbol{x}\|^2 = \frac{1}{N} \sum_{i=1}^{N} |x_i|^2$.

More generally, we can consider a continuous kernel $K(x, y) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying

$$|K(x, y)| \le A + B(|x| + |y|) \qquad \forall x, y \in \mathbb{R}^d,$$

obtaining

$$\boldsymbol{F}^N(x, \boldsymbol{y}) := \frac{1}{N} \sum_{j=1}^{N} K(x, y_j) = \int_{\mathbb{R}^d} K(x, y)\, \mathrm{d}\mu[\boldsymbol{y}](y)$$

and

$$\boldsymbol{F}(x, \mu) := \int_{\mathbb{R}^d} K(x, y)\, \mathrm{d}\mu(y).$$

An example for L^N and L is the variance:

$$L^N(x, \boldsymbol{x}) := \left| x - \frac{1}{N} \sum_{j=1}^{N} x_j \right|^2,$$

and

$$L(x, \mu) := \left| x - \int_{\mathbb{R}^d} y\, \mathrm{d}\mu(y) \right|^2.$$

A second-order example

Second-order systems can be easily reduced to first-order models if we admit controls on positions and velocities. Let us see an example where controls act only on the velocities. Assume $d = 2m$ and write the vector $x = (q, p)$, where $q \in \mathbb{R}^m$ denotes the position and $p \in \mathbb{R}^m$ the velocity.

We consider the vector field $\boldsymbol{F}^N(x, \boldsymbol{x}) = (\boldsymbol{F}_1^N(x), \boldsymbol{f}_2^N(x, \boldsymbol{x}))$ defined by

$$\boldsymbol{F}_1^N((q, p)) = p, \qquad \boldsymbol{F}_2^N((q, p), (\boldsymbol{q}, \boldsymbol{p})) = -\frac{1}{N} \sum_{j=1}^{N} \nabla W(p - p_j), \tag{3.20}$$

where the first component $\boldsymbol{F}_1^N$ is local and it is not influenced by the interaction with the other particles.

We are interested in the system

$$\begin{cases} \dot{q}_i = p_i, \\ \dot{p}_i = -\frac{1}{N} \sum_{j=1}^{N} \nabla W(p_i - p_j) + u_i, \end{cases}$$

which corresponds to (3.8) where the vector $\boldsymbol{u}$ has the particular form $\boldsymbol{u} = ((0, u_1), \ldots, (0, u_N))$, so that it is constrained to the subspace U^N where $U = \{(0, u) : u \in \mathbb{R}^m\} \subset \mathbb{R}^{2m}$. The limit vector field $\boldsymbol{F}(x, \mu) = (\boldsymbol{F}_1(x), \boldsymbol{f}_2(x, \mu))$ is defined by

$$\boldsymbol{F}_1((q, p)) = p, \qquad \boldsymbol{F}_2((q, p), \mu) = -\nabla_p W * \mu,$$

and the continuity equation

$$\partial_t \mu_t + \nabla \cdot \Big(\boldsymbol{F}(x, \mu_t)\mu_t + \boldsymbol{\nu}_t \Big) = 0$$

becomes a Vlasov-like equation

$$\partial_t \mu_t + p \cdot \nabla_q \mu_t + \nabla_p \cdot \big(\boldsymbol{F}_2(x, \mu_t)\mu_t + \boldsymbol{\nu}_t \big) = 0.$$

It is easy to check that this structure fits in our abstract setting, since $\boldsymbol{F}^N$, $\boldsymbol{f}$ satisfy the compatibility condition (3.6): for every $x \in \mathbb{R}^d$, $\boldsymbol{y} \in (\mathbb{R}^d)^N$ and $\mu \in \mathcal{P}_1(\mathbb{R}^d)$, we have $\boldsymbol{F}(x, \mu) - \boldsymbol{F}^N(x, \boldsymbol{y}) = \big(0, \boldsymbol{F}_2(x, \mu) - \boldsymbol{F}_2^N(x, \boldsymbol{y})\big) \in U^N$.

By choosing $\boldsymbol{F}_2^N$ in (3.20) as

$$\boldsymbol{F}_2^N((q,p), (\boldsymbol{q}, \boldsymbol{p})) = -\alpha p - \frac{1}{N}\sum_{j=1}^{N} \nabla W(p - p_j),$$

for some $\alpha > 0$, we obtain a model with friction in the velocity part. By choosing $\boldsymbol{F}_2^N$ in (3.20) as

$$\boldsymbol{F}_2^N((q,p), (\boldsymbol{q}, \boldsymbol{p})) = -\frac{1}{N}\sum_{j=1}^{N} a(|q - q_j|)(p - p_j),$$

where $a : [0, +\infty) \to \mathbb{R}_+$ is a continuous and non-increasing (thus bounded) function, we obtain a model of alignment in velocity. A particular and interesting example for a is given by the following decreasing function $a(|q|) = 1/(1 + |q|^2)^\gamma$ for some $\gamma \geq 0$, which yields the Cucker–Smale flocking model [31, 32].

An example for L^N and L in the second-order model is the variance of the velocities:

$$L^N((q,p), (\boldsymbol{q}, \boldsymbol{p})) := \left| p - \frac{1}{N}\sum_{j=1}^{N} p_j \right|^2,$$

and

$$L((q,p), \mu) := \left| p - \int_{\mathbb{R}^d} r_2 \, \mathrm{d}\mu(r_1, r_2) \right|^2.$$

4 The finite dimensional problem

Here we discuss the well-posedness of the finite dimensional control problem (3.10).

A first estimate on the solution is presented in the following Lemma, where we use the notation $|\boldsymbol{y}|_N = \frac{1}{N}\sum_{i=1}^{N} |y_i|$, with $\boldsymbol{y} = (y_1, \ldots, y_N) \in (\mathbb{R}^d)^N$.

Lemma 4.1 *Let* $(\boldsymbol{x}, \boldsymbol{u}) \in \mathscr{A}^N$. *Then*

$$\sup_{t \in [0,T]} |\boldsymbol{x}(t)|_N \leq \left(|\boldsymbol{x}(0)|_N + AT + \int_0^T |\boldsymbol{u}(s)|_N \, \mathrm{d}s \right) e^{2BT}, \tag{4.1}$$

where A and B are the constants of the assumption (3.5).

Proof From the integral formulation of equation (3.8), we get

$$\begin{aligned} |x_i(t)| &\leq |x_i(0)| + \int_0^t |\boldsymbol{F}^N(x_i(s), \boldsymbol{x}(s))| \mathrm{d}s + \int_0^t |u_i(s)| \, \mathrm{d}s \\ &\leq |x_i(0)| + \int_0^t (A + B(|x_i(s)| + |\boldsymbol{x}(s)|_N)) \, \mathrm{d}s + \int_0^t |u_i(s)| \, \mathrm{d}s. \end{aligned}$$

Averaging with respect to N, we obtain

$$|\boldsymbol{x}(t)|_N \leq |\boldsymbol{x}(0)|_N + AT + \int_0^T |\boldsymbol{u}(s)|_N \, \mathrm{d}s + 2B \int_0^t |\boldsymbol{x}(s)|_N \mathrm{d}s$$

and we conclude by Gronwall lemma. □

Proposition 4.2 *For every $N \in \mathbb{N}$ and $\boldsymbol{x}_0 \in (\mathbb{R}^d)^N$, the minimum problem (3.10) admits a solution, i.e., the set $P^N(\boldsymbol{x}_0)$ is not empty.*

Proof We fix $N \in \mathbb{N}$ and $\boldsymbol{x}_0 \in (\mathbb{R}^d)^N$. Let $\lambda := \inf\{\mathcal{E}^N(\boldsymbol{x}, \boldsymbol{u}) : (\boldsymbol{x}, \boldsymbol{u}) \in \mathscr{A}^N(\boldsymbol{x}_0)\}$. Since $\mathscr{A}^N(\boldsymbol{x}_0)$ is not empty, $\lambda < +\infty$. Let $(\boldsymbol{x}^k, \boldsymbol{u}^k) \in \mathscr{A}^N(\boldsymbol{x}_0)$ be a minimising sequence and $C := \sup_k \mathcal{E}^N (\boldsymbol{x}^k, \boldsymbol{u}^k) < +\infty$.

Since

$$\sup_k \int_0^T \psi(u_i^k(t)) \, \mathrm{d}t \leq C, \qquad \forall\, i = 1, \ldots, N, \tag{4.2}$$

and the function ψ is superlinear, then the sequence $\boldsymbol{u}^k$ is equi-integrable and hence weakly relatively compact in $L^1([0, T]; U^N)$. Hence there exists $\boldsymbol{u} \in L^1([0, T], U^N)$ and a subsequence, again denoted by $\boldsymbol{u}^k$, weakly convergent to $\boldsymbol{u}$ in $L^1([0, T], U^N)$.

Thanks to Lemma 4.1, the associated trajectories $\boldsymbol{x}^k$ are equi-bounded. Let us now show the equi-continuity of $x_i^k(t)$. For $s \leq t$, by equation (3.8), we have

$$x_i^k(t) - x_i^k(s) = \int_s^t \boldsymbol{F}^N(x_i^k(r), \boldsymbol{x}^k(r)) \, \mathrm{d}r + \int_s^t u_i^k(r) \, \mathrm{d}r. \tag{4.3}$$

Using the growth conditions (3.5) and (4.1), we get

$$\begin{aligned}
\left|\boldsymbol{x}^k(t) - \boldsymbol{x}^k(s)\right|_N &\leq \frac{1}{N} \sum_{i=1}^N \int_s^t \left|\boldsymbol{f}^N\left(x_i^k(r), \boldsymbol{x}^k(r)\right)\right| \, \mathrm{d}r + \int_s^t \left|\boldsymbol{u}^k(r)\right|_N \, \mathrm{d}r \\
&\leq A(t-s) + 2B \int_s^t \left|\boldsymbol{x}^k(r)\right|_N \, \mathrm{d}r + \int_s^t \left|\boldsymbol{u}^k(r)\right|_N \, \mathrm{d}r \\
&\leq A(t-s) + 2B \left(|\boldsymbol{x}_0|_N + AT + \int_0^T \left|\boldsymbol{u}^k(r)\right|_N \, \mathrm{d}r\right) e^{2BT}(t-s) + \int_s^t |\boldsymbol{u}^k(r)|_N \, \mathrm{d}r.
\end{aligned}$$

Since $\int_0^T |\boldsymbol{u}^k(r)|_N \, \mathrm{d}r$ is bounded, we have

$$\sup_k \left|\boldsymbol{x}^k(t) - \boldsymbol{x}^k(s)\right|_N \leq \tilde{C}|t-s| + \sup_k \left|\int_s^t \left|\boldsymbol{u}^k(r)\right|_N \, \mathrm{d}r\right|, \qquad \forall\, s, t \in [0, T], \tag{4.4}$$

where $\tilde{C} := A + 2B\left(|\boldsymbol{x}_0|_N + AT + \sup_k \int_0^T |\boldsymbol{u}^k(r)|_N \, \mathrm{d}r\right) e^{2BT}$. By the equi-integrability of $\boldsymbol{u}^k$, the inequality (4.4) shows the equi-continuity of $\boldsymbol{x}^k$. By Ascoli–Arzelà theorem, there exists a continuous curve $\boldsymbol{x}$ and a subsequence, again denoted by $\boldsymbol{x}^k$ such that $\boldsymbol{x}^k \to \boldsymbol{x}$ in $C([0, T]; (\mathbb{R}^d)^N)$. Passing to the limit in (4.3), we obtain

$$x_i(t) - x_i(s) = \int_s^t \boldsymbol{F}^N(x_i(r), \boldsymbol{x}(r)) \, \mathrm{d}r + \int_s^t u_i(r) \, \mathrm{d}r, \qquad i = 1, \ldots, N,$$

from which we deduce that $\boldsymbol{x}$ is absolutely continuous and solves equation (3.8). Hence $(\boldsymbol{x}, \boldsymbol{u}) \in \mathscr{A}^N(\boldsymbol{x}_0)$.

Finally, by the convexity of ψ and the continuity of L^N, we obtain the lower semicontinuity property:

$$\liminf_k \mathcal{E}^N\left(\boldsymbol{x}^k, \boldsymbol{u}^k\right) = \liminf_k \left[\int_0^T \frac{1}{N}\sum_{i=1}^N L^N\left(x_i^k(t), \boldsymbol{x}^k(t)\right)\, \mathrm{d}t + \frac{1}{N}\sum_{i=1}^N \int_0^T \psi\left(u_i^k(t)\right)\, \mathrm{d}t \right]$$
$$\geq \int_0^T \frac{1}{N}\sum_{i=1}^N L^N(x_i(t), \boldsymbol{x}(t))\, \mathrm{d}t + \frac{1}{N}\sum_{i=1}^N \int_0^T \psi(u_i(t))\, \mathrm{d}t,$$

whence the minimality of $(\boldsymbol{x}, \boldsymbol{u}) \in \mathscr{A}^N(\boldsymbol{x}_0)$. □

5 Momentum estimates

In this section, we study the set $\mathscr{A}$. We observe that if $(\mu, \boldsymbol{\nu}) \in \mathscr{A}$, then for any $\zeta \in C_c^1(\mathbb{R}^d)$, we have that the map $t \mapsto \int_{\mathbb{R}^d} \zeta\, \mathrm{d}\mu_t$ is absolutely continuous, a.e. differentiable, and

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} \zeta(x)\, \mathrm{d}\mu_t(x) = \int_{\mathbb{R}^d} \langle \boldsymbol{f}(t,x), \nabla\zeta(x)\rangle\, \mathrm{d}\mu_t(x) + \int_{\mathbb{R}^d} \langle \nabla\zeta(x), \mathrm{d}\boldsymbol{\nu}_t(x)\rangle \quad \text{for a.e. } t \in [0,T], \tag{5.1}$$

for the vector field $\boldsymbol{f}(t,x) := \boldsymbol{F}(x, \mu_t)$ satisfying the structural bounds

$$|\boldsymbol{f}(t,x)| \leq A + B\left(|x| + \int_{\mathbb{R}^d} |x|\, \mathrm{d}\mu_t\right). \tag{5.2}$$

In order to highlight the structural assumptions needed for the a priori estimates of this section, we introduce the set

$$\begin{aligned}\tilde{\mathscr{A}} := \Big\{(\mu, \boldsymbol{\nu}, \boldsymbol{f}) : \mu \in AC\left([0,T]; \mathcal{P}_1\left(\mathbb{R}^d\right)\right),\ \boldsymbol{\nu} \in \mathcal{M}\left([0,T] \times \mathbb{R}^d; U\right),\ \mathcal{E}\left(\boldsymbol{\nu}, \tilde{\mu}\right) < \infty,\\ \boldsymbol{f} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \text{ Borel function satisfying (5.1) and (5.2)}\Big\};\end{aligned} \tag{5.3}$$

the above discussion shows that if $(\mu, \boldsymbol{\nu}) \in \mathscr{A}$, then setting $\boldsymbol{f}(t,x) := \boldsymbol{F}(x, \mu_t)$, we have $(\mu, \boldsymbol{\nu}, \boldsymbol{f}) \in \tilde{\mathscr{A}}$.

Firstly, let us show a uniform bound in time of the first moment, which is the infinite dimensional version of Lemma 4.1.

Lemma 5.1 *If $(\mu, \boldsymbol{\nu}, \boldsymbol{f}) \in \tilde{\mathscr{A}}$, then the following estimate holds true*

$$\sup_{t \in [0,T]} \int_{\mathbb{R}^d} |x|\, \mathrm{d}\mu_t(x) \leq \left(\int_{\mathbb{R}^d} |x|\, \mathrm{d}\mu_0(x) + AT + |\boldsymbol{\nu}|\left((0,T) \times \mathbb{R}^d\right)\right) \mathrm{e}^{2BT}. \tag{5.4}$$

In particular, there exists a constant $M > 0$ only depending on $A, B, T, \mathcal{E}(\mu, \boldsymbol{\nu})$ and $\int_{\mathbb{R}^d} |x|\, \mathrm{d}\mu_0$ such that

$$|\boldsymbol{f}(t,x)| \leq M(1 + |x|) \quad \textit{for every } (t,x) \in [0,T] \times \mathbb{R}^d. \tag{5.5}$$

Proof Let $\zeta \in C_c^1(\mathbb{R}^d)$ be a cut-off function such that $0 \leq \zeta \leq 1$,

$$\zeta(x) = \begin{cases} 1 & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| \geq 2, \end{cases}$$

and $|\nabla\zeta|\le 1$. Let ζ_n be the sequence $\zeta_n(x) := \zeta(x/n)$. Consider now the product $\zeta_n(x)|x|$ and smooth it out in zero by substituting $|x|$ with $g_\varepsilon(x) := \sqrt{|x|^2+\varepsilon}$. Now $\zeta_n g_\varepsilon$ is a proper test function and the following equality holds true:

$$\begin{aligned}&\int_{\mathbb{R}^d}\zeta_n(x)g_\varepsilon(x)\mathrm{d}\mu_t(x) - \int_{\mathbb{R}^d}\zeta_n(x)g_\varepsilon(x)\mathrm{d}\mu_0(x)\\ &\quad = \int_0^t\int_{\mathbb{R}^d}\langle \boldsymbol{f}(s,x), \nabla(\zeta_n(x)g_\varepsilon(x))\rangle\mathrm{d}\mu_s(x)\mathrm{d}s + \int_0^t\int_{\mathbb{R}^d}\langle\nabla(\zeta_n(x)g_\varepsilon(x)), \mathrm{d}\boldsymbol{\nu}_s(x)\rangle\mathrm{d}s.\end{aligned}$$

Thanks to

$$|\nabla\zeta_n(x)|\le\frac{1}{n},\quad g_\varepsilon(x)\le|x|+\sqrt{\varepsilon},\quad |\nabla g_\varepsilon(x)| = \frac{|x|}{\sqrt{|x|^2+\varepsilon}}\le 1,$$

we can write

$$\begin{aligned}&\int_{\mathbb{R}^d}\zeta_n(x)g_\varepsilon(x)\mathrm{d}\mu_t(x) - \int_{\mathbb{R}^d}\zeta_n(x)g_\varepsilon(x)\mathrm{d}\mu_0(x)\\ &\quad\le\left(1+\frac{\sqrt{\varepsilon}}{n}\right)\int_0^t\int_{\mathbb{R}^d}|\boldsymbol{f}(s,x)|\mathrm{d}\mu_s(x)\mathrm{d}s + \left(1+\frac{\sqrt{\varepsilon}}{n}\right)\int_0^t\int_{\mathbb{R}^d}d|\boldsymbol{\nu}_s|(x)\mathrm{d}s.\end{aligned}$$

Apply now monotone convergence as $\varepsilon\to 0$ first, then let $n\to\infty$. Owing to $\zeta_n|x|\nearrow|x|$, we get

$$\begin{aligned}\int_{\mathbb{R}^d}|x|\mathrm{d}\mu_t(x) - \int_{\mathbb{R}^d}|x|\mathrm{d}\mu_0(x) &\le \int_0^t\int_{\mathbb{R}^d}|\boldsymbol{f}(s,x)|\mathrm{d}\mu_s(x)\mathrm{d}s + |\boldsymbol{\nu}|\left((0,T)\times\mathbb{R}^d\right)\\ &\le\int_0^t\int_{\mathbb{R}^d}\left[A+B\left(|x|+\int_{\mathbb{R}^d}|x|\,\mathrm{d}\mu_s(x)\right)\right]\mathrm{d}\mu_s(x)\mathrm{d}s + |\boldsymbol{\nu}|\left((0,T)\times\mathbb{R}^d\right)\\ &\le AT + 2B\int_0^t\int_{\mathbb{R}^d}|x|\mathrm{d}\mu_s(x)\mathrm{d}s + |\boldsymbol{\nu}|\left((0,T)\times\mathbb{R}^d\right),\end{aligned}$$

and we conclude by Gronwall inequality. □

Lemma 5.2 *If $(\mu,\boldsymbol{\nu},\boldsymbol{f})\in\tilde{\mathscr{A}}$ with $\boldsymbol{\nu}=\boldsymbol{v}\tilde{\mu}$, then for any $\vartheta\in C^1_{\mathrm{Lip}}(\mathbb{R}^d)$ the following equality holds:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\vartheta(x)\,\mathrm{d}\mu_t(x) = \int_{\mathbb{R}^d}\langle\boldsymbol{f}(t,x)+\boldsymbol{v}(t,x),\nabla\vartheta(x)\rangle\,\mathrm{d}\mu_t(x)\qquad \textit{for a.e. } t\in[0,T],$$

where $C^1_{\mathrm{Lip}}(\mathbb{R}^d)$ denotes the space of continuously differentiable functions with bounded gradient.

Proof Let $\vartheta\in C^1_{\mathrm{Lip}}(\mathbb{R}^d)$ and ζ_n the sequence of cut-off functions defined in the proof of Lemma 5.1. Then $\zeta_n\vartheta$ is a test function and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\zeta_n(x)\vartheta(x)\mathrm{d}\mu_t(x) &= \int_{\mathbb{R}^d}\langle\boldsymbol{f}(t,x)+\boldsymbol{v}(t,x),\nabla(\zeta_n(x)\vartheta(x))\rangle\mathrm{d}\mu_t(x)\\ &=\int_{\mathbb{R}^d}\langle\boldsymbol{f}(t,x)+\boldsymbol{v}(t,x),\nabla\zeta_n(x)\vartheta(x)+\zeta_n(x)\nabla\vartheta(x)\rangle\mathrm{d}\mu_t(x).\end{aligned}$$

Taking into account that $|\nabla\zeta_n|\le\frac{1}{n}\chi_{B_{2n}}$, the Lipschitz continuity of ϑ, the growth condition on $\boldsymbol{f}$ and Lemma 5.1, by dominated convergence, we obtain that

$$\int_{\mathbb{R}^d}\vartheta(x)\mathrm{d}\mu_t(x) = \int_{\mathbb{R}^d}\vartheta(x)\mathrm{d}\mu_0(x) + \int_{\mathbb{R}^d}\langle\boldsymbol{f}(t,x)+\boldsymbol{v}(t,x),\nabla\vartheta(x)\rangle\mathrm{d}\mu_t(x).$$

□

Now we are ready to prove the main result of this section. It involves an auxiliary admissible function $\theta : [0,\infty) \to [0,\infty)$ (according to Definition 2.2) dominated by ψ, i.e.,

$$\theta(|x|) \le 1 + \psi(x) \quad \text{for every } x \in U; \tag{5.6}$$

notice that, combining Lemma 2.3 and Remark 2.4, if $\mu_0 \in \mathcal{P}_1(\mathbb{R}^d)$, we can always find an admissible function θ satisfying (5.6) and

$$\int_{\mathbb{R}^d} \theta(|x|)\, \mathrm{d}\mu_0(x) < \infty. \tag{5.7}$$

Proposition 5.3 *Let $(\mu, \boldsymbol{\nu}, \boldsymbol{f}) \in \tilde{\mathscr{A}}$ and let θ be an admissible function satisfying (5.6) and (5.7). Then there exists a constant $C > 0$, depending only on A, B, T, $\int_{\mathbb{R}^d} |x|\, \mathrm{d}\mu_0(x)$, $\mathcal{E}(\mu, \boldsymbol{\nu})$, $\theta(1)$ and the doubling constant K of θ (see (2.4)), such that*

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^d} \theta(|x|)\, \mathrm{d}\mu_t(x) \le C \left(1 + \int_{\mathbb{R}^d} \theta(|x|)\, \mathrm{d}\mu_0(x)\right). \tag{5.8}$$

Proof Since $\mathcal{E}(\mu, \boldsymbol{\nu}) < +\infty$, we have that $\boldsymbol{\nu} = \boldsymbol{v}\tilde{\mu}$. We also set $\vartheta(x) := \theta(|x|)$, $x \in \mathbb{R}^d$.

Step 1: We start by approximating θ from below with a sequence of C^1 functions:

$$\vartheta^n(x) := \theta^n(|x|), \qquad \theta^n(r) := \begin{cases} \theta(r) & \text{if } |x| \le n \\ \theta'(n)(r-n) + \theta(n) & \text{if } r > n. \end{cases}$$

Observe that θ^n are Lipschitz since $(\theta^n)'(r) \le \theta'(n)$, for every $r \ge 0$.

$$\begin{aligned} \int_{\mathbb{R}^d} \vartheta^n(x)\mathrm{d}\mu_t(x) &= \int_{\mathbb{R}^d} \vartheta^n(x)\mathrm{d}\mu_0(x) + \int_0^t \int_{\mathbb{R}^d} \langle \boldsymbol{f}(s,x) + \boldsymbol{v}(s,x), \nabla\vartheta^n(x)\rangle \mathrm{d}\mu_s(x)\mathrm{d}s \\ &\le \int_{\mathbb{R}^d} \vartheta^n(x)\mathrm{d}\mu_0(x) + \int_0^t \int_{\mathbb{R}^d} |\boldsymbol{f}(s,x) + \boldsymbol{v}(s,x)||\nabla\vartheta^n(x)|\mathrm{d}\mu_s(x)\mathrm{d}s. \end{aligned}$$

By construction, $\theta^n(|x|) \nearrow \theta(|x|)$ $|\nabla\vartheta^n(x)| \nearrow |\nabla\vartheta(x)|$, for every $x \in \mathbb{R}^d$; we can thus pass to the limit in the relation above to get

$$\int_{\mathbb{R}^d} \vartheta(x)\mathrm{d}\mu_t(x) \le \int_{\mathbb{R}^d} \vartheta(x)\mathrm{d}\mu_0(x) + \int_0^t \int_{\mathbb{R}^d} |\boldsymbol{f}(s,x) + \boldsymbol{v}(s,x)||\nabla\vartheta(x)|\mathrm{d}\mu_s(x)\mathrm{d}s. \tag{5.9}$$

Step 2: We want to estimate the right-hand side of (5.9). Since $\theta'(r) \ge 0$ by (2.6) and $|\nabla\vartheta(x)| = \left|\theta'(|x|)\frac{x}{|x|}\right| = \theta'(|x|)$, (5.5) yields

$$\int_{\mathbb{R}^d} |\boldsymbol{f}(s,x)||\nabla\vartheta(x)|\mathrm{d}\mu_s(x) \le M \int_{\mathbb{R}^d} (1+|x|)\theta'(|x|)\mathrm{d}\mu_s(x).$$

By the monotonicity of θ' (2.7) in $[0,1]$ and (2.6), we have

$$(1+r)\theta'(r) \le 2K(1+\theta(1)+\theta(r)),$$

so that

$$\int_{\mathbb{R}^d} |\boldsymbol{f}(s,x)||\nabla\vartheta(x)|\mathrm{d}\mu_s(x) \le 2MK\left(1+\theta(1)+\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_s(x)\right).$$

Concerning the second term on the right-hand side of (5.9), we have

$$\begin{aligned}|\boldsymbol{v}(s,x)||\nabla\vartheta(x)| &\le \theta(|\boldsymbol{v}(s,x)|)+\theta^*(|\nabla\vartheta(x)|)\\ &= \theta(|\boldsymbol{v}(s,x)|)+\theta^*(\theta'(|x|))\\ &= \theta(|\boldsymbol{v}(s,x)|)+\theta'(|x|)|x|-\theta(|x|)\\ &\le \theta(|\boldsymbol{v}(s,x)|)+K(1+\theta(|x|)),\end{aligned}$$

where the equality $\theta(|x|)+\theta^*(\theta'(|x|))=\theta'(|x|)|x|$ comes from the definition of the Fenchel conjugate θ^*. What we end up with is the following:

$$\begin{aligned}\int_0^t\int_{\mathbb{R}^d}|\boldsymbol{v}(s,x)||\nabla\vartheta(x)|\mathrm{d}\mu_s(x)\mathrm{d}s &\le \int_0^T\int_{\mathbb{R}^d}\theta(|\boldsymbol{v}(s,x)|)\mathrm{d}\mu_s(x)\mathrm{d}s+Kt+K\int_0^t\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_s(x)\mathrm{d}s\\ &\le T\mathcal{E}(\mu,\boldsymbol{v})+(1+K)T+K\int_0^t\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_s(x)\mathrm{d}s.\end{aligned}$$

Summing up the two estimates, we obtain for every $t\in[0,T]$ and a suitable constant $C>0$:

$$\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_t(x)\le\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_0(x)+CT+C\int_0^t\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_s(x)\mathrm{d}s,$$

and thanks to the Gronwall inequality, we get

$$\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_t(x)\le \mathrm{e}^{CT}\left(\int_{\mathbb{R}^d}\theta(|x|)\mathrm{d}\mu_0(x)+CT\right).$$ □

6 Proof of the main theorems

6.1 The superposition principle

We first recall the superposition principle for solutions of the continuity equation

$$\partial_t\mu_t+\nabla\cdot(\boldsymbol{w}(t,\cdot)\mu_t)=0. \tag{6.1}$$

Let us denote with Γ_T the complete and separable metric space of continuous functions from $[0,T]$ to $\mathbb{R}^d$ endowed with the sup-distance and introduce the evaluation maps $e_t:\Gamma_T\to\mathbb{R}^d$ defined by $e_t(\gamma):=\gamma(t)$, for $t\in[0,T]$. The following result holds:

Theorem 6.1 (Superposition principle) *Let μ_t be a narrowly continuous weak solution to (6.1) with a velocity field $\boldsymbol{w}$ satisfying*

$$\int_0^T\int_{\mathbb{R}^d}|\boldsymbol{w}(t,x)|\,\mathrm{d}\mu_t(x)\mathrm{d}t<+\infty.$$

Then there exists $\pi\in\mathcal{P}(\Gamma_T)$ concentrated on the set of curves $\gamma\in AC([0,T];\mathbb{R}^d)$ such that

$$\dot\gamma(t)=\boldsymbol{w}(t,\gamma(t))\quad \textit{for a.e. } t\in[0,T].$$

Moreover, $\mu_t = (e_t)_\# \pi$ *for any* $t \in [0, T]$, *i.e.,*

$$\int_{\mathbb{R}^d} \vartheta(y)\, \mathrm{d}\mu_t(y) = \int_{\Gamma_T} \vartheta(\gamma(t))\, \mathrm{d}\pi(\gamma), \qquad \forall\, \vartheta \in C_b(\mathbb{R}^d).$$

For the proof, we refer to Ref. [3, Theorem 3.4].

6.2 Γ-convergence

Let us start with a preliminary lemma.

Lemma 6.2 *Let* $(\boldsymbol{x}, \boldsymbol{u}) \in AC([0, T]; (\mathbb{R}^d)^N) \times L^1([0, T]; U^N)$ *and* $\mu = \mu[\boldsymbol{x}]$, $\boldsymbol{\nu} = \boldsymbol{\nu}[\boldsymbol{x}, \boldsymbol{u}]$. *Then we have*

$$\frac{1}{N} \sum_{i=1}^{N} \psi(u_i(t)) \geq \Psi(\boldsymbol{\nu}_t | \mu_t) \qquad \textit{for a.e. } t \in [0, T]. \tag{6.2}$$

Moreover, if $(\boldsymbol{x}, \boldsymbol{u}) \in \mathscr{A}^N$, *then*

$$\frac{1}{N} \sum_{i=1}^{N} \psi(u_i(t)) = \Psi(\boldsymbol{\nu}_t | \mu_t) \qquad \textit{for a.e. } t \in [0, T]. \tag{6.3}$$

Proof Let us first compute the density of $\boldsymbol{\nu}$ w.r.t. $\tilde{\mu}$. We introduce the finite set $I_N := \{1, 2, \ldots, N\}$ with the discrete topology and the normalised counting measure $\sigma_N = \frac{1}{N} \sum_{i=1}^N \delta_i$. We can identify $\boldsymbol{x}$ with a continuous map from $[0, T] \times I_N$ to $\mathbb{R}^d$, $\boldsymbol{x}(t, i) := x_i(t)$, so that $\mu_t = \boldsymbol{x}(t, \cdot)_\sharp \sigma_N$. Similarly, we set $\boldsymbol{u}(t, i) := u_i(t)$, where $\boldsymbol{u} : [0, T] \to U^N$ is a Borel representative. In order to represent $\tilde{\mu}$ and $\boldsymbol{\nu}$, it is useful to deal with the map $\boldsymbol{y} : [0, T] \times I_N \to [0, T] \times \mathbb{R}^d$, $\boldsymbol{y}(t, i) := (t, \boldsymbol{x}(t, i))$, which yields $\tilde{\mu} = \boldsymbol{y}_\sharp\big(\lambda \otimes \sigma_N\big)$ and $\boldsymbol{\nu} = \boldsymbol{y}_\sharp\big(\boldsymbol{u} \cdot (\lambda \otimes \sigma_N)\big)$. We denote by $Y \subset [0, T] \times \mathbb{R}^d$ the range of $\boldsymbol{y}$ and by

$$X(t) = \{x \in \mathbb{R}^d : (t, x) \in Y\} = \{x \in \mathbb{R}^d : x_i(t) = x \text{ for some } i \in I_N\}$$

its fibres. For every $(t, x) \in [0, T] \times \mathbb{R}^d$, we will also consider the set

$$J(t, x) := \{i \in I_N : x_i(t) = x\} \text{ with its characteristic function } \quad \chi_{t,x}(i) := \begin{cases} 1 & \text{if } x_i(t) = x \\ 0 & \text{otherwise.} \end{cases} \tag{6.4}$$

For every $t \in [0, T]$, the collection $\{\chi_{t,x} : x \in X(t)\}$ provides a partition of unity of I_N and for every $i \in I_N$, the map $(t, x) \mapsto \chi_{t,x}$ is upper semicontinuous in $[0, T] \times \mathbb{R}^d$. The conditional measures $\tilde{\mu}_{t,x} \in \mathcal{P}(I_N)$ are then defined by

$$\tilde{\mu}_{t,x}(J) := \sigma_N(J \cap J(t, x))/\sigma_N(J(t, x)), \qquad (t, x) \in Y;$$

since for every $J \subset I_N$

$$\sigma_N(J \cap J(t, x)) = \int_J \chi_{t,x}\, \mathrm{d}\sigma_N = \frac{1}{N} \sum_{i \in J} \chi_{t,x}(i),$$

the map $(t, x) \mapsto \sigma_N(J \cap J(t, x))$ is also upper semicontinuous and $\tilde{\mu}_{t,x}$ is a Borel family.

One immediately checks that $\tilde{\mu}_{t,x}$ provides a disintegration (see, e.g., [5, Theorem 5.3.1]) of $\lambda \otimes \sigma_N$ w.r.t. the map $\boldsymbol{y}$, i.e.,

$$\lambda \otimes \sigma_N = \int_{[0,T]\times\mathbb{R}^d} \tilde{\mu}_{t,x} \,\mathrm{d}\tilde{\mu}(t,x).$$

Since $\boldsymbol{v} = \boldsymbol{y}_\sharp\big(\boldsymbol{u}\cdot(\lambda\otimes\sigma_N)\big)$, we eventually end up with the representation formula for the Borel vector field $\boldsymbol{v}$:

$$\begin{cases} \boldsymbol{v}(t,x) := \displaystyle\int_{I_N} \boldsymbol{u}(t,i)\,\mathrm{d}\tilde{\mu}_{t,x}(i) = \frac{1}{\sharp J(t,x)} \sum_{i\in J(t,x)} u_i(t) & \text{if } (t,x)\in Y, \\ \boldsymbol{v}(t,x) := 0 & \text{otherwise.} \end{cases}$$

In particular,

$$\boldsymbol{v}_t = \boldsymbol{v}(t,\cdot)\mu_t, \qquad \mu_t = \sum_{x\in X(t)} \frac{\sharp J(t,x)}{N}\delta_x,$$

and consequently

$$\Psi(\boldsymbol{v}_t|\mu_t) = \int_{\mathbb{R}^d} \psi(\boldsymbol{v}(t,x))\,\mathrm{d}\mu_t(x) = \sum_{x\in X(t)} \frac{\sharp J(t,x)}{N} \psi\left(\frac{1}{\sharp J(t,x)} \sum_{i\in J(t,x)} u_i(t)\right). \tag{6.5}$$

The convexity of ψ immediately yields

$$\Psi(\boldsymbol{v}_t|\mu_t) \le \frac{1}{N}\sum_{i=1}^N \psi(u_i(t)). \tag{6.6}$$

Let us show that equality holds in (6.6) if $(\boldsymbol{x},\boldsymbol{u}) \in \mathscr{A}^N$.

Let $\mathscr{P}$ be the collection of all the partitions P of I_N. It is clear that for every $t\in[0,T]$, the family $P_{\boldsymbol{x}}(t) := \{J(t,x) : x \in X(t)\}$ is an element of $\mathscr{P}$; moreover for every $P\in\mathscr{P}$, the set

$$S_P := \{t\in[0,T] : P_{\boldsymbol{x}}(t) = P\} \quad \text{is Borel.} \tag{6.7}$$

To show (6.7), we introduce an order relation on $\mathscr{P}$: we say that $P_1 \prec P_2$ if every element of P_1 is contained in some element of P_2. We denote by $\hat{P} := \{Q\in\mathscr{P} : P \prec Q\}$ the collection of all the partitions Q coarser than P.

It is easy to check that for every $P\in\mathscr{P}$, the set $P_{\boldsymbol{x}}^{-1}\big(\hat{P}\big) = \{t\in[0,T] : P_{\boldsymbol{x}}(t)\in\hat{P}\}$ is closed. In fact, if $P_{\boldsymbol{x}}(t)\notin\hat{P}$, then there is a set $I\in P$ not contained in any element of $P_{\boldsymbol{x}}(t)$, so that we can find two indices $i,j\in I$ belonging to different elements of $P_{\boldsymbol{x}}(t)$, i.e., $x_i(t)\neq x_j(t)$. By continuity, this relation holds in a neighbourhood U of t, so that $P_{\boldsymbol{x}}(s)\notin\hat{P}$ for every $s\in U$.

Since for every partition $P\in\mathscr{P}$ $\{P\} = \hat{P}\setminus\cup\big\{\hat{Q} : Q\in\hat{P}, Q\neq P\big\}$, it follows that

$$S_P = P_{\boldsymbol{x}}^{-1}\big(\hat{P}\big) \setminus \bigcup_{Q\in\hat{P}, Q\neq P} P_{\boldsymbol{x}}^{-1}\left(\hat{Q}\right),$$

so that S_P is the difference between closed sets and (6.7) holds.

We can therefore decompose the interval $[0,T]$ in the finite Borel partition $\{S_P : P\in\mathscr{P}\}$. On the other hand, for every partition $P\in\mathscr{P}$ and every pair of indices i,j in $I\in P$, we have $x_i(t) =$

$x_j(t)$ in S_P so that $\dot{x}_i(t) = \dot{x}_j(t)$ for λ-almost every $t \in S_P$ and consequently, by (3.8), we obtain that $u_i(t) = u_j(t)$ for λ-a.e. $t \in S_P$. We eventually deduce

$$\sharp I \psi\left(\frac{1}{\sharp I}\sum_{i\in I} u_i(t)\right) = \sum_{i\in I}\psi(u_i(t)) \quad \text{for every } I \in P_{\boldsymbol{x}}(t), \quad \lambda\text{-a.e. in } S_P,$$

and therefore, by (6.5),

$$\begin{aligned}\Psi(\boldsymbol{\nu}_t|\mu_t) &= \sum_{I\in P_{\boldsymbol{x}}(t)}\frac{\sharp I}{N}\psi\left(\frac{1}{\sharp I}\sum_{i\in I}u_i(t)\right) = \frac{1}{N}\sum_{I\in P_{\boldsymbol{x}}(t)}\sum_{i\in I}\psi(u_i(t))\\ &= \sum_{i=1}^{N}\psi(u_i(t)) \quad \text{for } \lambda\text{-a.e. } t\in S_P.\end{aligned}$$

Since $\{S_P : P \in \mathscr{P}\}$ is a finite Borel partition of $[0, T]$, we get (6.3). $\square$

Proof of Theorem 3.2 **The** lim inf **inequality.** Let $(\mu, \boldsymbol{\nu}) \in AC([0, T]; \mathcal{P}_1(\mathbb{R}^d)) \times \mathcal{M}([0, T] \times \mathbb{R}^d; U)$ and $(\boldsymbol{x}^N, \boldsymbol{u}^N) \in AC([0, T]; (\mathbb{R}^d)^N) \times L^1([0, T]; U^N)$, $N \in \mathbb{N}$, such that $\mu^N = \mu[\boldsymbol{x}^N] \to \mu$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$ and $\boldsymbol{\nu}^N = \boldsymbol{\nu}[\boldsymbol{x}^N, \boldsymbol{u}^N] \rightharpoonup^* \boldsymbol{\nu}$ in $\mathcal{M}([0, T] \times \mathbb{R}^d; U)$.

Since $L^N \geq 0$, $L^N(x, \boldsymbol{x}^N(t)) \to L(x, \mu_t)$ on compact sets and $\mu[\boldsymbol{x}^N] \rightharpoonup^* \mu$, then for every compact $K \subset \mathbb{R}^d$, by (3.1) we have

$$\begin{aligned}&\liminf_{N\to+\infty}\int_{\mathbb{R}^d} L^N\left(x, \boldsymbol{x}^N(t)\right)\, \mathrm{d}\mu_t^N(x)\\ &\qquad \geq \liminf_{N\to+\infty}\int_K L^N\left(x, \boldsymbol{x}^N(t)\right)\, \mathrm{d}\mu_t^n(x) = \int_K L(x, \mu_t)\, \mathrm{d}\mu_t(x).\end{aligned} \tag{6.8}$$

Since

$$\frac{1}{N}\sum_{i=1}^{N} L^N\left(x_i^N(t), \boldsymbol{x}^N(t)\right) = \int_{\mathbb{R}^d} L^N\left(x, \boldsymbol{x}^N(t)\right)\, \mathrm{d}\mu_t^N(x)$$

and $L \geq 0$, by (6.8) we obtain

$$\liminf_{N\to\infty}\int_0^T \frac{1}{N}\sum_{i=1}^{N} L^N\left(x_i^N(t), \boldsymbol{x}^N(t)\right)\, \mathrm{d}t \geq \int_0^T\int_{\mathbb{R}^d} L(x, \mu_t)\, \mathrm{d}\mu_t(x)\, \mathrm{d}t. \tag{6.9}$$

By (6.2) we have

$$\frac{1}{N}\sum_{i=1}^{N}\psi\left(u_i^N(t)\right) \geq \Psi\left(\boldsymbol{\nu}_t^N\,\middle|\,\mu_t^N\right) \qquad \text{for a.e. } t \in [0, T], \tag{6.10}$$

and Theorem 2.6 yields

$$\liminf_{N\to\infty}\int_0^T \Psi\left(\boldsymbol{\nu}_t^N\,\middle|\,\mu_t^N\right)\, \mathrm{d}t = \liminf_{N\to\infty}\Psi\left(\boldsymbol{\nu}^N\,\middle|\,\tilde{\mu}^N\right) \geq \Psi(\boldsymbol{\nu}|\tilde{\mu}) = \int_0^T \Psi(\boldsymbol{\nu}_t|\mu_t)\, \mathrm{d}t. \tag{6.11}$$

By (6.9), (6.10) and (6.11), it follows (3.12).

The lim sup **inequality.** Recall that ϕ is an admissible function satisfying (2.5). Let $(\mu, \boldsymbol{\nu}) \in \mathscr{A}$, such that $\mathcal{E}(\mu, \boldsymbol{\nu}) < +\infty$ and $\int_{\mathbb{R}^d}\phi(|x|)\, \mathrm{d}\mu_0(x) < \infty$.

Since $\Psi(\boldsymbol{\nu}|\tilde{\mu}) < +\infty$, we have $\boldsymbol{\nu} = \boldsymbol{v}\tilde{\mu}$ for a Borel vector field $\boldsymbol{v} : [0, T] \times \mathbb{R}^d \to U$. Since $(\mu, \boldsymbol{\nu}) \in \mathscr{A}$ the continuity equation

$$\partial_t \mu_t + \nabla \cdot \big(\boldsymbol{w}(t, \cdot)\mu_t\big) = 0 \tag{6.12}$$

holds with the vector field $\boldsymbol{w}(t, x) := \boldsymbol{f}(t, x) + \boldsymbol{v}(t, x), \boldsymbol{f}(t, x) := \boldsymbol{F}(x, \mu_t)$. By (3.5) and Lemma 5.1, we have that

$$\boldsymbol{f} \in C\left([0, T] \times \mathbb{R}^d\right), \quad |\boldsymbol{f}(t, x)| \leq M(1 + |x|), \quad \int_0^T \int_{\mathbb{R}^d} |\boldsymbol{w}(t, x)| \,\mathrm{d}\mu_t(x)\,\mathrm{d}t < +\infty. \tag{6.13}$$

By Theorem 6.1, there exists a probability measure $\pi \in \mathcal{P}(\Gamma_T)$ such that $(e_t)_\sharp \pi = \mu_t$ for every $t \in [0, T]$ and it is concentrated on the absolutely continuous solutions of the ODE:

$$\dot{\gamma}(t) = \boldsymbol{f}(t, \gamma(t)) + \boldsymbol{v}(t, \gamma(t)). \tag{6.14}$$

The strategy of the proof consists in finding an appropriate sequence of measures $\pi^N \in \mathcal{P}^N(\Gamma_T)$ narrowly convergent to π, defining $\mu_t^N := (e_t)_\sharp \pi^N$ and $\boldsymbol{x}^N$ a corresponding curve such that $\mu[\boldsymbol{x}^N] = \mu^N$. Then the objective is to construct a suitable sequence of controls $\boldsymbol{u}^N$ in such a way that the sequence $(\boldsymbol{x}^N, \boldsymbol{u}^N)$ belongs to $\mathscr{A}^N$, $\mu^N \to \mu$ in $C([0, T]; \mathcal{P}_1(\mathbb{R}^d))$, $\boldsymbol{\nu}^N = \boldsymbol{\nu}[\boldsymbol{x}^N, \boldsymbol{u}^N] \rightharpoonup^* \boldsymbol{\nu}$ in $\mathcal{M}([0, T] \times \mathbb{R}^d; U)$ and (3.16) holds.

Step 1: Definition of auxiliary functionals. We define the set

$$A := \left\{\gamma \in \Gamma_T : \gamma \in AC\left([0, T]; \mathbb{R}^d\right), (6.14) \text{ holds for a.e. } t \in [0, T]\right\}$$

and we observe that $\pi(A) = 1$.

Starting from μ and L, we define the functional $\mathcal{L} : A \to [0, +\infty)$ by

$$\mathcal{L}(\gamma) := \int_0^T L(\gamma(t), \mu_t)\,\mathrm{d}t.$$

Starting from ψ and $\boldsymbol{v}$, we define the functional $\mathcal{F} : A \to [0, +\infty)$ by

$$\mathcal{F}(\gamma) := \int_0^T \psi(\boldsymbol{v}(t, \gamma(t)))\,\mathrm{d}t.$$

By Fubini's theorem and the finiteness of $\mathcal{E}(\mu, \boldsymbol{\nu})$, we have

$$\int_0^T \int_{\mathbb{R}^d} L(x, \mu_t)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t = \int_0^T \int_A L(e_t(\gamma), \mu_t)\,\mathrm{d}\pi(\gamma)\,\mathrm{d}t = \int_A \mathcal{L}(\gamma)\,\mathrm{d}\pi(\gamma) \tag{6.15}$$

and

$$\int_0^T \int_{\mathbb{R}^d} \psi(\boldsymbol{v}(t, x))\,\mathrm{d}\mu_t(x)\,\mathrm{d}t = \int_0^T \int_A \psi(\boldsymbol{v}(t, e_t(\gamma))\,\mathrm{d}\pi(\gamma)\,\mathrm{d}t = \int_A \mathcal{F}(\gamma)\,\mathrm{d}\pi(\gamma). \tag{6.16}$$

We define the functional $\mathcal{H} : A \to [0, +\infty)$ by

$$\mathcal{H}(\gamma) := \int_0^T \phi(|\boldsymbol{v}(t, \gamma(t))|)\,\mathrm{d}t.$$

Starting by ϕ satisfying (2.5), we define the functional $\mathcal{G} : A \to [0, +\infty)$ by

$$\mathcal{G}(\gamma) := \phi(|\gamma(0)|) + \int_0^T \phi(|\gamma(t)|)\,\mathrm{d}t.$$

It is not difficult to show that $\mathcal{G}$ and $\mathcal{L}$ are continuous. Here we prove that $\mathcal{F}$ and $\mathcal{H}$ are lower semicontinuous. Let $\gamma \in A$ and $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence in A, such that $\lim_{k\to+\infty}\sup_{t\in[0,T]}|\gamma_k(t)-\gamma(t)|=0$. We define the sequence $f_k \in L^1([0,T];\mathbb{R}^d)$ by $f_k(t) := \boldsymbol{v}(t,\gamma_k(t))$.

If $\sup_{k\in\mathbb{N}} \fint_0^T \phi(|\boldsymbol{v}(t,\gamma_k(t))|)\,\mathrm{d}t < +\infty$, then by de la Vallée Poussin's criterion [4, Proposition 1.12] for equi-integrability and Dunford–Pettis theorem, there exist $g \in L^1([0,T];\mathbb{R}^d)$ and a subsequence (not relabeled) of f_k weakly convergent in $L^1([0,T];\mathbb{R}^d)$ to g such that

$$\liminf_{k\in\mathbb{N}} \fint_0^T \phi(|\boldsymbol{v}(t,\gamma_k(t))|)\,\mathrm{d}t \geq \fint_0^T \phi(|g(t)|)\,\mathrm{d}t.$$

Since γ_k satisfies

$$\gamma_k(t_2)-\gamma_k(t_1) = \int_{t_1}^{t_2} [\boldsymbol{v}(t,\gamma_k(t)) + \boldsymbol{f}(t,\gamma_k(t))]\,\mathrm{d}t, \qquad \forall\, t_1,t_2 \in [0,T] \tag{6.17}$$

and γ satisfies

$$\gamma(t_2)-\gamma(t_1) = \int_{t_1}^{t_2} [\boldsymbol{v}(t,\gamma(t)) + \boldsymbol{f}(t,\gamma(t))]\,\mathrm{d}t, \qquad \forall\, t_1,t_2 \in [0,T] \tag{6.18}$$

passing to the limit in (6.17) as $k\to\infty$, we obtain

$$\gamma(t_2)-\gamma(t_1) = \int_{t_1}^{t_2} [g(t) + \boldsymbol{f}(t,\gamma(t))]\,\mathrm{d}t.$$

By (6.18) it holds:

$$\int_{t_1}^{t_2} g(t)\mathrm{d}t = \int_{t_1}^{t_2} \boldsymbol{v}(t,\gamma(t))\mathrm{d}t, \qquad \forall t_1,t_2 \in [0,T],$$

and Lebesgue differentiation theorem yields $g(t) = \boldsymbol{v}(t,\gamma(t))$ for a.e. $t\in[0,T]$.

Step 2: Construction of $\boldsymbol{\pi}^N$. We define the function $\mathfrak{F} := (\mathcal{F},\mathcal{L},\mathcal{G},\mathcal{H}) : A \to \mathbb{R}^4$.

Notice that the finiteness of $\mathcal{E}(\mu,\boldsymbol{v})$, (6.15) and (6.16) implies that $\int_A \mathcal{F}(\gamma)\,\mathrm{d}\pi(\gamma) < +\infty$, $\int_A \mathcal{L}(\gamma)\,\mathrm{d}\pi(\gamma) < +\infty$ and $\int_A \mathcal{H}(\gamma)\,\mathrm{d}\pi(\gamma) < +\infty$. Since $\int_A \mathcal{G}(\gamma)\,\mathrm{d}\pi(\gamma) = \int_0^T\int_{\mathbb{R}^d}\phi(|x|)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t$, by Proposition 5.3, we also have that $\int_A \mathcal{G}(\gamma)\,\mathrm{d}\pi(\gamma) < +\infty$.

By Lusin's theorem applied to the space A with the measure π and the function $\mathfrak{F}$, there exists a sequence of compact sets A_k such that $A_k \subset A_{k+1} \subset A$, $\pi(A\setminus A_k) < \frac{1}{k}$, for all $k\geq 1$, and $\mathfrak{F}_{|A_k}$ is continuous. Moreover, we have

$$\lim_{k\to\infty}\pi(A_k) = \pi\left(\bigcup_{j=1}^{\infty} A_j\right) = 1, \quad \pi\left(A \setminus \bigcup_{j=1}^{\infty} A_j\right) = 0. \tag{6.19}$$

Then we define $\tilde{\pi}^k \in \mathcal{P}(\Gamma_T)$ by

$$\tilde{\pi}^k := \frac{1}{\pi(A_k)}\pi \llcorner A_k.$$

It is easy to check that $(\tilde{\pi}^k)_{k\in\mathbb{N}}$ weakly converges to π as $k \to \infty$; since for each component $\mathfrak{F}_j$, $j = 1, 2, 3, 4$, of $\mathfrak{F}$ is non-negative, Beppo Levi monotone convergence theorem yields

$$\lim_{k\to+\infty} \int_{A_k} \mathfrak{F}_j(\gamma)\,\mathrm{d}\pi(\gamma) = \int_{\bigcup_{k=1}^{\infty} A_k} \mathfrak{F}_j(\gamma)\,\mathrm{d}\pi(\gamma) = \int_A \mathfrak{F}_j(\gamma)\,\mathrm{d}\pi(\gamma), \tag{6.20}$$

and (6.19) easily yields

$$\lim_{k\to\infty} \left| \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\tilde{\pi}^k(\gamma) - \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\pi(\gamma) \right| = 0. \tag{6.21}$$

Since A_k is compact, we can find a sequence of atomic measures

$$m \mapsto \tilde{\pi}^k_m := \frac{1}{m} \sum_{i=1}^{m} \delta_{\gamma_{i,k,m}}, \qquad \gamma_{i,k,m} \in A_k,$$

narrowly convergent to $\tilde{\pi}^k$ as $m \to +\infty$. Since $\mathfrak{F}_{|A_k}$ is bounded and continuous, in particular, it holds that

$$\lim_{m\to\infty} \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\tilde{\pi}^k_m(\gamma) = \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\tilde{\pi}^k(\gamma).$$

Hence, for every $k \in \mathbb{N}$, there exists $\bar{m}(k)$ satisfying

$$W\left(\tilde{\pi}^k_m, \tilde{\pi}^k\right) \le \frac{1}{k} \quad \text{and} \quad \left| \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\tilde{\pi}^k_m(\gamma) - \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\tilde{\pi}^k(\gamma) \right| \le \frac{1}{k}, \qquad \forall\, m \ge \bar{m}(k), \tag{6.22}$$

where W is any distance metrising the weak convergence.

We define $\bar{\pi}^k := \tilde{\pi}^k_{\bar{m}(k)}$ and we clearly have that $\bar{\pi}^k \in \mathcal{P}^{\bar{m}(k)}(\Gamma_T)$, $W(\bar{\pi}^k, \pi) \to 0$ as $k \to \infty$ and, by (6.21) and (6.22),

$$\lim_{k\to\infty} \left| \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\bar{\pi}^k(\gamma) - \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\pi(\gamma) \right| = 0. \tag{6.23}$$

Since we can choose the sequence $k \mapsto \bar{m}(k)$ strictly increasing, we can consider the sequence $N \mapsto \pi^N$ such that $\pi^N \in \mathcal{P}^N(\Gamma_T)$, $\pi^N := \bar{\pi}^k$ when $\bar{m}(k) \le N < \bar{m}(k+1)$; π^N narrowly converges to π as $N \to +\infty$ and

$$\lim_{N\to+\infty} \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\pi^N(\gamma) = \int_{\Gamma_T} \mathfrak{F}(\gamma)\,\mathrm{d}\pi(\gamma). \tag{6.24}$$

Since all the components of $\mathfrak{F}$ are non-negative and lower semicontinuous maps, by a combination of [4, Proposition 1.62 (a)] and [4, Proposition 1.80], we have that (6.23) yields, in particular, that the measures

$$\sigma_1^N := \mathcal{F}\pi + \mathcal{F}\pi^N, \quad \sigma_2^N := \mathcal{G}\pi + \mathcal{G}\pi^N, \quad \sigma_3^N := \mathcal{H}\pi + \mathcal{H}\pi^N, \quad \sigma_4^N := \mathcal{L}\pi + \mathcal{L}\pi^N$$

weakly converge to $\sigma_1 := 2\mathcal{F}\pi$, $\sigma_2 := 2\mathcal{G}\pi$, $\sigma_3 := 2\mathcal{H}\pi$ and $\sigma_4 := 2\mathcal{L}\pi$, respectively. In particular, they are uniformly tight, so that for every $\varepsilon > 0$, there exists $\bar{N}(\varepsilon) \in \mathbb{N}$ and a compact set B_ε and such that

$$B_\varepsilon \subset A_N, \ \left(\pi + \pi^N\right)(\Gamma_T \backslash B_\varepsilon) + \int_{\Gamma_T \backslash B_\varepsilon} (\mathcal{F} + \mathcal{L} + \mathcal{H} + \mathcal{G})\ \mathrm{d}(\pi + \pi^N) \le \varepsilon \ \text{ for every } N \ge \bar{N}(\varepsilon). \tag{6.25}$$

Step 3: Definition of $(\boldsymbol{x}^N, \boldsymbol{u}^N)$ and convergence. We define $\mu_t^N := (e_t)_\sharp \pi^N \in \mathcal{P}^N(\mathbb{R}^d)$ and we denote by $\boldsymbol{x}^N$ a corresponding curve such that $\mu[\boldsymbol{x}^N] = \mu^N$. We define

$$\boldsymbol{f}^N(t,x) := \boldsymbol{F}^N\left(x, \boldsymbol{x}^N(t)\right) = \boldsymbol{F}\left(x, \mu_t^N\right), \quad \boldsymbol{v}^N(t,x) := \boldsymbol{v}(t,x) + \boldsymbol{f}(t,x) - \boldsymbol{f}^N(t,x) \tag{6.26}$$

$$u_i^N(t) := \boldsymbol{v}^N(t, x_i^N(t)) \tag{6.27}$$

and $\boldsymbol{u}^N = (u_1^N, \ldots, u_N^N)$. Notice that

$$\boldsymbol{f}(t,x) - \boldsymbol{f}^N(t,x) \in U \text{ and } \boldsymbol{u}^N \in U^N, \text{ thanks to the compatibility condition (3.6).}$$

We have that $\boldsymbol{v}^N := \boldsymbol{v}[\boldsymbol{x}^N, \boldsymbol{u}^N] = \boldsymbol{v}^N \mu^N$. Since each component x_i^N of $\boldsymbol{x}^N$ belongs to A, then the sequence $(\boldsymbol{x}^N, \boldsymbol{u}^N)$ belongs to $\mathscr{A}^N$, so that $(\mu^N, \boldsymbol{v}^N) \in \mathscr{A}$. Using the same computation of the proof of Proposition 5.3, taking into account that μ^N satisfies

$$\partial_t \mu_t^N + \nabla \cdot \left(\left(\boldsymbol{f}(t,\cdot) + \boldsymbol{v}(t,\cdot)\right)\mu_t^N\right) = 0, \tag{6.28}$$

with (recall (6.16) and (6.24))

$$\int_0^T \int_{\mathbb{R}^d} \psi(\boldsymbol{v}(t,x))\, \mathrm{d}\mu_t^N(x)\, \mathrm{d}t = \int_{\Gamma_T} \mathcal{F}(\gamma)\, \mathrm{d}\pi^N(\gamma) \le 1 + \mathcal{E}(\mu|\boldsymbol{v})$$

for N sufficiently large, we obtain by (6.13), (5.8) and (5.5) that

$$\sup_{N\in\mathbb{N}} \sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|\, \mathrm{d}\mu_t^N(x) < +\infty, \quad \sup_{N\in\mathbb{N}} \sup_{t\in[0,T]} \int_{\mathbb{R}^d} \phi(|x|)\, \mathrm{d}\mu_t^N(x) < +\infty, \tag{6.29}$$

which implies the uniform convergence $\mu^N \to \mu$ in $C([0,T]; \mathcal{P}_1(\mathbb{R}^d))$ and the uniform estimate

$$\left|\boldsymbol{f}^N(t,x)\right| \le M'(1+|x|) \quad \text{for every } t \in [0,T],\ x \in \mathbb{R}^d,\ N \in \mathbb{N}, \tag{6.30}$$

for a suitable constant $M' > 0$. By a direct computation, using the assumption (3.7), we obtain that $\boldsymbol{v}^N \rightharpoonup^* \boldsymbol{v}$ in $\mathcal{M}([0,T] \times \mathbb{R}^d; U)$.

Step 4: Definition and convergence of $\mathcal{F}^N$. We define $\mathcal{F}^N : A \to [0,+\infty)$ by

$$\mathcal{F}^N(\gamma) := \int_0^T \psi(\boldsymbol{v}^N(t,\gamma(t)))\, \mathrm{d}t. \tag{6.31}$$

Here we show that the sequence $\mathcal{F}^N$ converges to $\mathcal{F}$ uniformly on every compact set $\Lambda \subset A_h$ for some $h \in \mathbb{N}$. To do it, we fix $\Lambda \subset A_h$ and we prove that for any $\gamma \in \Lambda$ and every sequence $(\gamma_N)_{N\in\mathbb{N}} \subset \Lambda$ such that $\sup_{t\in[0,T]} |\gamma_N(t) - \gamma(t)| \to 0$, we have $\mathcal{F}^N(\gamma_N) \to \mathcal{F}(\gamma)$ as $N \to +\infty$.

By the assumption (3.7), we have that

$$\lim_{N\to+\infty} \left|\boldsymbol{f}(t,\gamma_N(t)) - \boldsymbol{f}^N(t,\gamma_N(t))\right| = 0, \qquad \forall\, t \in [0,T].$$

Since $\mathcal{H}$ is continuous in A_h, it holds

$$\lim_{N\to+\infty} \int_0^T \phi(|\boldsymbol{v}(t,\gamma_N(t))|)\, \mathrm{d}t = \int_0^T \phi(|\boldsymbol{v}(t,\gamma(t))|)\, \mathrm{d}t. \tag{6.32}$$

Since ϕ is strictly convex and superlinear, by Visintin's Theorem [63, Theorem 3], $\boldsymbol{v}(\cdot,\gamma_N(\cdot))$ strongly converges in $L^1(0,T)$ to $\boldsymbol{v}(\cdot,\gamma(\cdot))$. Then, using also the continuity of ψ, along a subsequence (still denoted by γ_N), we have

$$\lim_{N\to+\infty} \psi(\boldsymbol{v}(t,\gamma_N(t))+\boldsymbol{f}(t,\gamma_N(t))-\boldsymbol{f}^N(t,\gamma_N(t))=\psi(\boldsymbol{v}(t,\gamma(t))), \qquad \text{for a.e. } t\in[0,T].$$

Since by (5.5) and (6.30), we have

$$\left|\boldsymbol{f}(t,\gamma_N(t))-\boldsymbol{f}^N(t,\gamma_N(t))\right|\le (M+M')\big(1+|\gamma_N(t)|\big),$$

then, using the doubling property and the uniform convergence of γ_N, we can find a constant C such that

$$\psi\left(\boldsymbol{v}(t,\gamma_N(t))+\boldsymbol{f}(t,\gamma_N(t))-\boldsymbol{f}^N(t,\gamma_N(t))\right)\le C\Big(1+\phi(|\boldsymbol{v}(t,\gamma_N(t))|)\Big).$$

By (6.32), the generalised dominated convergence theorem (see for instance [35, Theorem 4, p. 21]) shows that

$$\lim_{N\to+\infty}\int_0^T \psi(\boldsymbol{v}^N(t,\gamma_N(t)))\,\mathrm{d}t=\int_0^T \psi(\boldsymbol{v}(t,\gamma(t)))\,\mathrm{d}t.$$

Step 5: Definition and convergence of $\mathcal{L}^N$. We define $\mathcal{L}^N : A\to[0,+\infty)$ by

$$\mathcal{L}^N(\gamma):=\int_0^T L^N(\gamma(t),\boldsymbol{x}^N(t))\,\mathrm{d}t. \tag{6.33}$$

Here we show that the sequence $\mathcal{L}^N$ converges to $\mathcal{L}$ uniformly on every compact set $\Lambda\subset A_h$ for some $h\in\mathbb{N}$. As in Step 4, we fix $\Lambda\subset A_h$ and we prove that for every sequence $(\gamma_N)_{N\in\mathbb{N}}\subset\Lambda$, with $\sup_{t\in[0,T]}|\gamma_N(t)-\gamma(t)|\to 0$, we have $\mathcal{L}^N(\gamma_N)\to\mathcal{L}(\gamma)$ as $N\to+\infty$. Indeed, by (3.1),

$$\lim_{N\to+\infty} L^N(\gamma_N(t),\boldsymbol{x}^N(t))=L(\gamma(t),\mu_t), \qquad \forall\, t\in[0,T].$$

Since $(\gamma_N)_N$ is bounded and $\mu^N\to\mu$ in $C([0,T];\mathcal{P}_1(\mathbb{R}^d))$, by (3.1) we obtain that

$$\sup_{N\in\mathbb{N}} L^N(\gamma_N(t),\boldsymbol{x}^N(t))<+\infty.$$

By dominated convergence we conclude.

Step 6: Conclusion. By the growth assumptions (6.13) and (6.30) on $\boldsymbol{f},\boldsymbol{f}^N$, the doubling property of ϕ and (2.5), we have

$$\mathcal{F}^N(\gamma)\le C(1+\mathcal{F}(\gamma)+\mathcal{G}(\gamma)) \qquad \forall\,\gamma\in A, \quad \forall\, N\in\mathbb{N}. \tag{6.34}$$

Moreover, by (3.1) and the uniform convergence of μ^N to μ, there exists a constant C such that

$$\mathcal{L}^N(\gamma)\le\mathcal{L}(\gamma)+C \qquad \forall\,\gamma\in A, \quad \forall\, N\in\mathbb{N}. \tag{6.35}$$

Fix $\varepsilon>0$ and let B_ε and $\bar N(\varepsilon)$ such that (6.25) holds and

$$\left|\int_{\Gamma_T}\mathcal{F}(\gamma)\mathrm{d}\pi^N(\gamma)-\int_{\Gamma_T}\mathcal{F}(\gamma)\mathrm{d}\pi(\gamma)\right|<\varepsilon.$$

By (6.34), (6.35) and (6.25), we have

$$\int_{\Gamma_T\setminus B_\varepsilon}\mathcal{F}^N(\gamma)\mathrm{d}(\pi+\pi^N)(\gamma)\le\varepsilon, \quad \int_{\Gamma_T\setminus B_\varepsilon}\mathcal{L}^N(\gamma)\mathrm{d}(\pi+\pi^N)(\gamma)\le\varepsilon \qquad \forall\, N\ge\bar N(\varepsilon).$$

Moreover, from the previous step, there exists $\tilde{N}(\varepsilon)$ such that

$$\sup_{\gamma\in B_\varepsilon}\left|\mathcal{F}^N(\gamma)-\mathcal{F}(\gamma)\right|\le\varepsilon,\quad \sup_{\gamma\in B_\varepsilon}\left|\mathcal{L}^N(\gamma)-\mathcal{L}(\gamma)\right|\le\varepsilon\qquad\forall\, N\ge\tilde{N}(\varepsilon).$$

Hence

$$\begin{aligned}\left|\int_{\Gamma_T}\mathcal{F}^N(\gamma)\mathrm{d}\pi^N(\gamma)-\int_{\Gamma_T}\mathcal{F}(\gamma)\mathrm{d}\pi(\gamma)\right|&\le\left|\int_{B_\varepsilon}\mathcal{F}^N(\gamma)\mathrm{d}\pi^N(\gamma)-\int_{B_\varepsilon}\mathcal{F}(\gamma)\mathrm{d}\pi(\gamma)\right|\\&\quad+\left|\int_{\Gamma_T\setminus B_\varepsilon}\mathcal{F}^N(\gamma)\mathrm{d}\pi^N(\gamma)-\int_{\Gamma_T\setminus B_\varepsilon}\mathcal{F}(\gamma)\mathrm{d}\pi(\gamma)\right|\\&\le\varepsilon+2\varepsilon,\qquad\forall\, N\ge\max\{\bar{N}(\varepsilon),\tilde{N}(\varepsilon)\},\end{aligned}$$

which shows that

$$\lim_{N\to\infty}\int_{\Gamma_T}\mathcal{F}^N(\gamma)\mathrm{d}\pi^N(\gamma)=\int_{\Gamma_T}\mathcal{F}(\gamma)\mathrm{d}\pi(\gamma).$$

Analogously we obtain

$$\lim_{N\to\infty}\int_{\Gamma_T}\mathcal{L}^N(\gamma)\mathrm{d}\pi^N(\gamma)=\int_{\Gamma_T}\mathcal{L}(\gamma)\mathrm{d}\pi(\gamma).$$ □

6.3 Convergence of minima

Proof of Theorem 3.1 Equi-continuity. Let N be fixed and $s\le t$. From the constraint (3.8), we get

$$\begin{aligned}W_1(\mu_s^N,\mu_t^N)&\le\frac{1}{N}\sum_{i=1}^N\left|x_i^N(s)-x_i^N(t)\right|\\&\le\frac{1}{N}\sum_{i=1}^N\int_s^t\left|\boldsymbol{F}^N(x_i(r),\boldsymbol{x}^N(r))\right|\,\mathrm{d}r+\frac{1}{N}\sum_{i=1}^N\int_s^t\left|u_i^N(r)\right|\,\mathrm{d}r\\&\le\tilde{C}(t-s)+\int_s^t\frac{1}{N}\sum_{i=1}^N\left|u_i^N(r)\right|\,\mathrm{d}r,\end{aligned}$$

where $\tilde{C}:=A+2B\left(\sup_N|\boldsymbol{x}^N(0)|_N+AT+\sup_N\int_0^T|\boldsymbol{u}^N(r)|_N\,\mathrm{d}r\right)e^{2BT}$ (see the proof of (4.4)) which is uniformly bounded, since μ_0^n is converging in $\mathcal{P}_1(\mathbb{R}^d)$ and $\mathcal{E}(\mu^N,\boldsymbol{\nu}^N)$ is uniformly bounded.

By Remark 2.4, we can select an admissible function θ satisfying (2.10) with $\mathcal{K}:=\{\mu_0\}\cup\{\mu_0^N:N\in\mathbb{N}\}$. The uniform bound on $\mathcal{E}(\mu^N,\boldsymbol{\nu}^N)$ implies that

$$\sup_N\int_0^T\frac{1}{N}\sum_{i=1}^N\theta\left(\left|u_i^N(r)\right|\right)\,\mathrm{d}r<+\infty,$$

by the convexity and superlinearity of θ, there exists a uniform modulus of continuity $\omega:[0,+\infty)\to[0,+\infty)$ such that $\sup_N\int_s^t\frac{1}{N}\sum_{i=1}^N|u_i^N(r)|\,\mathrm{d}r\le\omega(t-s)$.

Hence we have just shown the equi-continuity property

$$W_1\left(\mu_s^N,\mu_t^N\right)\le\omega(|t-s|)+\tilde{C}|t-s|\qquad\forall\, t,s\in[0,T].$$

Compactness. From Theorem 5.3, we have

$$\sup_{N\in\mathbb{N}}\sup_{t\in[0,T]}\int_{\mathbb{R}^d}\theta(|x|)\,\mathrm{d}\mu_t^n(x)<+\infty.$$

This implies that the family $(\mu^N)_{N\in\mathbb{N}}\subset\mathcal{P}_1(\mathbb{R}^d)$ is relatively compact (see, e.g., [5, Proposition 7.1.5]).

The application of Ascoli–Arzelà theorem provides a limit curve $\mu\in C([0,T];\mathcal{P}_1(\mathbb{R}^d))$ and a subsequence, still denoted by μ^N, such that

$$\sup_{t\in[0,T]} W_1\left(\mu\left[\boldsymbol{x}^N\right]_t,\mu_t\right)\to 0. \tag{6.36}$$

Concerning the control part, we write $\boldsymbol{\nu}^N=\boldsymbol{v}^N\mu^N$. Since $\mathcal{E}^N(\boldsymbol{x}^N,\boldsymbol{u}^N)$ is uniformly bounded, we have

$$\sup_{N\in\mathbb{N}}\int_0^T\int_{\mathbb{R}^d}\psi\left(\boldsymbol{v}^N(t,x)\right)\,\mathrm{d}\mu_t^N(x)\,\mathrm{d}t<+\infty.$$

By the superlinearity of ψ and the convergence (6.36), using the same argument of the proof of Ref. [5, Theorem 5.4.4], we obtain that there exist $\boldsymbol{v}:[0,T]\times\mathbb{R}^d\to U$ and a subsequence (again denoted by $\boldsymbol{v}^N$) such that

$$\int_0^T\int_{\mathbb{R}^d}\psi(\boldsymbol{v}(t,x))\,\mathrm{d}\mu_t(x)\,\mathrm{d}t<+\infty$$

and

$$\lim_{N\to\infty}\int_0^T\int_{\mathbb{R}^d}\boldsymbol{\xi}(t,x)\cdot\boldsymbol{v}^N(t,x)\mathrm{d}\mu_t^N(x)\,\mathrm{d}t=\int_0^T\int_{\mathbb{R}^d}\boldsymbol{\xi}(t,x)\cdot\boldsymbol{v}(t,x)\mathrm{d}\mu_t(x)\,\mathrm{d}t,$$
$$\forall\,\boldsymbol{\xi}\in C_c^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^d).$$

This proves the convergence of $\boldsymbol{\nu}^N\to\boldsymbol{\nu}:=\boldsymbol{v}\mu$ in $\mathcal{M}([0,T]\times\mathbb{R}^d;U)$ and the fact that $(\mu,\boldsymbol{\nu})$ satisfies (3.9). □

Proof of Theorem 3.3 The first two claims are standard consequence of the Γ-convergence result of Theorem 3.2 and the coercivity property stated in Theorem 3.1. We thus consider the third claim.

Let us fix $\mu_0\in\mathcal{P}_1(\mathbb{R}^d)$ with compact support and $(\mu,\boldsymbol{\nu})\in P(\mu_0)$. By Theorem 3.2, we can find a sequence of discrete solutions $(\hat{\boldsymbol{x}}^N,\hat{\boldsymbol{u}}^N)$ corresponding to initial data $\hat{\boldsymbol{x}}_0^N$ supported in $\mathrm{supp}(\mu_0)$ and measures $(\hat{\mu}^N,\hat{\boldsymbol{\nu}}^N)$ converging to $(\mu,\boldsymbol{\nu})$ such that (3.15) and (3.16) hold. Theorem 3.2 also yields $\lim_{N\to\infty}E^N(\hat{\boldsymbol{x}}_0^N)=E(\mu_0)$.

Let now $(\boldsymbol{x}_0^N)_{N\in\mathbb{N}}$ be any other sequence satisfying (3.17) with $(\boldsymbol{x}^N,\boldsymbol{u}^N)\in P(\boldsymbol{x}_0^N)$ and $\mu^N=\mu[\boldsymbol{x}^N]$, $\boldsymbol{\nu}^N=\boldsymbol{\nu}[\boldsymbol{x}^n,\boldsymbol{u}^N]$. Applying Lemma 2.5, we deduce that the associated measures μ_0^N satisfy

$$\lim_{N\to\infty}\mathcal{C}_\phi\left(\hat{\mu}_0^N,\mu_0^N\right)=0.$$

Up to a permutation of the initial points $(\hat{x}_{0,1}^N,\hat{x}_{0,2}^N,\ldots,\hat{x}_{0,N}^N)$ (and of the corresponding solutions $(\hat{\boldsymbol{x}}^N,\hat{\boldsymbol{u}}^N)$) which, however, leaves $\hat{\mu}_0^N,\hat{\mu}^N,\hat{\boldsymbol{\nu}}^N$ invariant, we may assume by (2.1) that

$$c_N=\mathcal{C}_\phi\left(\hat{\mu}_0^N,\mu_0^N\right)=\frac{1}{N}\sum_{i=1}^N\phi\left(\left|\hat{x}_{0,i}^N-x_{0,i}^N\right|\right). \tag{6.37}$$

For $0 < \delta < T$ and $\boldsymbol{y}^{N,\delta} := \delta^{-1}(\hat{\boldsymbol{x}}_0^N - \boldsymbol{x}_0^N)$, we can then define a new competitor by

$$\boldsymbol{x}^{N,\delta}(t) := \begin{cases} (1-t/\delta)\,\boldsymbol{x}_0^N + t/\delta\,\hat{\boldsymbol{x}}_0^N & \text{if } t \in [0,\delta), \\ \hat{\boldsymbol{x}}^N(t-\delta) & \text{if } t \in [\delta, T], \end{cases}$$

$$\boldsymbol{u}_i^{N,\delta}(t) := \begin{cases} \boldsymbol{y}^{N,\delta} - \boldsymbol{F}^N\left(x_i^{N,\delta}, \boldsymbol{x}^{N,\delta}(t)\right) & \text{if } t \in [0,\delta), \\ \hat{\boldsymbol{u}}^N(t-\delta) & \text{if } t \in [\delta, T]. \end{cases}$$

It is easy to check that $(\boldsymbol{x}^{N,\delta}, \boldsymbol{u}^{N,\delta}) \in \mathscr{A}(\boldsymbol{x}_0^N)$ so that $E^N(\boldsymbol{x}_0^N) \le \mathcal{E}^N(\boldsymbol{x}^{N,\delta}, \boldsymbol{u}^{N,\delta})$. On the other hand,

$$T\mathcal{E}^N\left(\boldsymbol{x}^{N,\delta}, \boldsymbol{u}^{N,\delta}\right) \le \frac{1}{N}\int_0^\delta \sum_{i=1}^N L^N\left(x_i^{N,\delta}(t), \boldsymbol{x}^{N,\delta}(t)\right)\,\mathrm{d}t + \frac{1}{N}\int_0^\delta \sum_{i=1}^N \psi\left(\boldsymbol{y}^{N,\delta} - \boldsymbol{F}^N(x_i^{N,\delta}, \boldsymbol{x}^{N,\delta}(t))\right)\,\mathrm{d}t + T\mathcal{E}^N\left(\hat{\boldsymbol{x}}^N, \hat{\boldsymbol{u}}^N\right).$$

From the doubling property and the compactness of supports of $(\boldsymbol{x}_0^N)$, applying the same argument as in the proof of Theorem 3.2, we get

$$\begin{aligned}\psi\left(\boldsymbol{y}^{N,\delta} - \boldsymbol{F}^N\left(x_i^{N,\delta}, \boldsymbol{x}^{N,\delta}(t)\right)\right) &\le C\left(1+\phi\left(\left|\hat{\boldsymbol{x}}_0^N - \boldsymbol{x}_0^N\right|/\delta\right)\right) \\ &\le C\mathrm{e}^{K/\delta}\left(1+\phi\left(\left|\hat{\boldsymbol{x}}_0^N - \boldsymbol{x}_0^N\right|\right)\right) \qquad 0<\delta<1.\end{aligned}$$

Setting $\mu_t^{N,\delta} = \mu[\boldsymbol{x}^{N,\delta}(t)]$ we get,

$$T\left(\mathcal{E}^N\left(\boldsymbol{x}^{N,\delta}, \boldsymbol{u}^{N,\delta}\right) - \mathcal{E}^N\left(\hat{\boldsymbol{x}}^N, \hat{\boldsymbol{u}}^N\right)\right) \le C\,c_N\delta\left(1+\mathrm{e}^{K/\delta}\right) + \delta\sup_{t\in[0,1]}\int_{\mathbb{R}^d} L^N\left(x, \mu_t^{N,1}\right)\,\mathrm{d}\mu_t^{N,1}. \tag{6.38}$$

If we choose $\delta = \delta(N) := -K\big(\log(c_N)\big)^{-1}$, since $\lim_{N\to\infty}\sup_{t\in[0,1]} W_1(\mu_t^{N,1}, \mu_0) = 0$, we see that the right-hand side of (6.38) tends to 0 as $N \to \infty$, so that we eventually obtain

$$\limsup_{N\to\infty} E^N\left(\boldsymbol{x}_0^N\right) \le \limsup_{N\to\infty} \mathcal{E}^N\left(\boldsymbol{x}^{N,\delta}, \boldsymbol{u}^{N,\delta}\right) \le \limsup_{N\to\infty} \mathcal{E}^N\left(\hat{\boldsymbol{x}}^N, \hat{\boldsymbol{u}}^N\right) `= E(\mu_0).$$

Acknowledgements

We wish to thank Filippo Santambrogio for useful discussions concerning the third claim of Theorem 3.3.

Conflicts of interest

None.

References

[1] ALBI, G., BONGINI, M., CRISTIANI, E. & KALISE, D. (2016) Invisible control of self-organizing agents leaving unknown environments. *SIAM J. Appl. Math.* **76**(4), 1683–1710.

[2] ALBI, G., CHOI, Y.-P., FORNASIER, M. & KALISE, D. (2017) Mean field control hierarchy. *Appl. Math. Optim.* **76**(1), 93–135.

[3] AMBROSIO, L. & CRIPPA, G. (2014) Continuity equations and ODE flows with non-smooth velocity. *Proc. Roy. Soc. Edinburgh Sect. A* **144**(6), 1191–1244.

[4] AMBROSIO, L., FUSCO, N. & PALLARA, D. (2000) *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York.

[5] AMBROSIO, L., GIGLI, N. & SAVARÉ, G. (2008) *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel.

[6] ANDERSSON, D. & DJEHICHE, B. (2011) A maximum principle for SDEs of mean-field type. *Appl. Math. Optim.* **63**(3), 341–356.

[7] BALLERINI, M., CABIBBO, N., CANDELIER, R., CAVAGNA, A., CISBANI, E., GIARDINA, I., LECOMTE, V., ORLANDI, A., PARISI, G., PROCACCINI, A., VIALE, M., & ZDRAVKOVIC, V. (2008) Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proc. Nat. Acad. Sci.* **105**(4), 1232–1237.

[8] BAYRAKTAR, E., COSSO, A. & PHAM, H. (2018) Randomized dynamic programming principle and Feynman-Kac representation for optimal control of McKean-Vlasov dynamics. *Trans. Amer. Math. Soc.* **370**(3), 2115–2160.

[9] BENSOUSSAN, A., FREHSE, J. & YAM, P. (2013) *Mean Field Games and Mean Field Type Control Theory*. Springer Briefs in Mathematics, Springer, New York.

[10] BONGINI, M. & FORNASIER, M. (2014) Sparse stabilization of dynamical systems driven by attraction and avoidance forces. *Netw. Heterog. Media* **9**(1), 1–31.

[11] BONGINI, M. & FORNASIER, M. (2017) Sparse control of multiagent systems. In: *Active Particles, Vol. 1, Advances in Theory, Models, and Applications*. Modeling and Simulation in Science, Engineering and Technology, Birkhäuser/Springer, Cham, pp. 173–228.

[12] BONGINI, M., FORNASIER, M. & KALISE, D. (2015) (Un)conditional consensus emergence under perturbed and decentralized feedback controls. *Discrete Contin. Dyn. Syst.* **35**(9), 4071–4094.

[13] BONGINI, M., FORNASIER, M., ROSSI, F. & SOLOMBRINO, F. (2017) Mean-field Pontryagin maximum principle. *J. Optim. Theory Appl.* **175**(1), 1–38.

[14] BUCKDAHN, R., DJEHICHE, B. & LI, J. (2011) A general stochastic maximum principle for SDEs of mean-field type. *Appl. Math. Optim.* **64**(2), 197–216.

[15] CAMAZINE, S., DENEUBOURG, J.-L., FRANKS, N. R., SNEYD, J., THERAULAZ, G. & BONABEAU, E. (2003) *Self-organization in Biological Systems*. Princeton Studies in Complexity, Princeton University Press, Princeton, NJ. Reprint of the 2001 original.

[16] CAPONIGRO, M., FORNASIER, M., PICCOLI, B. & TRÉLAT, E. (2013) Sparse stabilization and optimal control of the Cucker-Smale model. *Math. Control Relat. Fields* **3**(4), 447–466.

[17] CARMONA, R., DELARUE, F. & LACHAPELLE, A. (2013) Control of McKean-Vlasov dynamics versus mean field games. *Math. Financ. Econ.* **7**(2), 131–166.

[18] CARRILLO, J. A., CHOI, Y.-P. & HAURAY, M. (2014) The derivation of swarming models: mean-field limit and Wasserstein distances. In: *Collective Dynamics from Bacteria to Crowds*. CISM Courses and Lectures, Vol. 553, Springer, Vienna, pp. 1–46.

[19] CARRILLO, J. A., CHOI, Y.-P. & PEREZ, S. P. (2017) A review on attractive-repulsive hydrodynamics for consensus in collective behavior. In: *Active Particles, Vol. 1, Advances in Theory, Models, and Applications*. Modeling and Simulation in Science, Engineering and Technology, Birkhäuser/Springer, Cham, pp. 259–298.

[20] CARRILLO, J. A., D'ORSOGNA, M. R. & PANFEROV, V. (2009) Double milling in self-propelled swarms from kinetic theory. *Kinet. Relat. Models* **2**(2), 363–378.

[21] CARRILLO, J. A., FORNASIER, M., TOSCANI, G. & VECIL, F. (2010) Particle, kinetic, and hydrodynamic models of swarming. In: *Mathematical Modeling of Collective Behavior in Socio-economic and Life Sciences*. Modeling and Simulation in Science, Engineering and Technology, Birkhäuser Boston, Inc., Boston, MA, pp. 297–336.

[22] CHOI, Y.-P., HA, S.-Y. & LI, Z. (2017) Emergent dynamics of the Cucker-Smale flocking model and its variants. In: *Active Particles, Vol. 1, Advances in Theory, Models, and Applications*. Modeling and Simulation in Science, Birkhäuser/Springer, Cham, pp. 299–331.

[23] CHUANG, Y.-L., D'ORSOGNA, M. R., MARTHALER, D., BERTOZZI, A. L. & CHAYES, L. S. (2007) State transitions and the continuum limit for a 2D interacting, self-propelled particle system. *Phys. D* **232**(1), 33–47.

[24] CHUANG, Y.-L., HUANG, Y. R., D'ORSOGNA, M. R. & BERTOZZI, A. L. (2007) Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials. In: *2007 IEEE International Conference on Robotics and Automation*, IEEE, pp. 2292–2299.

[25] COUZIN, I. D. & FRANKS, N. R. (2003) Self-organized lane formation and optimized traffic flow in army ants. *Proc. R. Soc. London, B: Biol. Sci.* **270**(1511), 139–146.

[26] COUZIN, I. D., KRAUSE, J., FRANKS, N. R. & LEVIN, S. A. (2005) Effective leadership and decision-making in animal groups on the move. *Nature* **433**(7025), 513.

[27] CRISTIANI, E., PICCOLI, B. & TOSIN, A. (2010) Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints. In: *Mathematical Modeling of Collective Behavior in Socio-economic and Life Sciences*. Modeling and Simulation in Science, Engineering and Technology, Birkhäuser Boston, Inc., Boston, MA, pp. 337–364.

[28] CRISTIANI, E., PICCOLI, B. & TOSIN, A. (2011) Multiscale modeling of granular flows with application to crowd dynamics. *Multiscale Model. Simul.* **9**(1), 155–182.

[29] CUCKER, F. & DONG, J.-G. (2011) A general collision-avoiding flocking framework. *IEEE Trans. Automat. Control* **56**(5), 1124–1129.

[30] CUCKER, F. & MORDECKI, E. (2008) Flocking in noisy environments. *J. Math. Pures Appl.* **89**(3), 278–296.

[31] CUCKER, F. & SMALE, S. (2007) Emergent behavior in flocks. *IEEE Trans. Automat. Control* **52**(5), 852–862.

[32] CUCKER, F. & SMALE, S. (2007) On the mathematics of emergence. *Jpn. J. Math.* **2**(1), 197–227.

[33] CUCKER, F., SMALE, S. & ZHOU, D.-X. (2004) Modeling language evolution. *Found. Comput. Math.* **4**(3), 315–343.

[34] DAL MASO, G. (1993) *An Introduction to* Γ*-convergence*. Progress in Nonlinear Differential Equations and their Applications, Vol. 8, Birkhäuser Boston, Inc., Boston, MA.

[35] EVANS, L. C. & GARIEPY, R. F. (2015) *Measure Theory and Fine Properties of Functions*. Textbooks in Mathematics, CRC Press, Boca Raton, FL, revised edition.

[36] FLEMING, W. H. (1977) Generalized solutions in optimal stochastic control. Differential games and control theory, II (Proc. 2nd Conf., Univ. Rhode Island, Kingston, R.I., 1976), pp. 147–165. *Lecture Notes in Pure and Appl. Math.*, 30. Dekker, New York.

[37] FLORENTIN, J. J. (1961) Optimal control of continuous time, Markov, stochastic systems. *J. Electron. Control* **10**, 473–488.

[38] FORNASIER, M., PICCOLI, B. & ROSSI, F. (2014) Mean-field sparse optimal control. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **372**(2028), 20130400.

[39] FORNASIER, M. & SOLOMBRINO, F. (2014) Mean-field optimal control. *ESAIM Control Optim. Calc. Var.* **20**(4), 1123–1152.

[40] GRÉGOIRE, G. & CHATÉ, H. (2004) Onset of collective and cohesive motion. *Phys. Rev. Lett.* **92**(2), 025702.

[41] JADBABAIE, A., LIN, J. & STEPHEN MORSE, A. (2003) Correction to: "Coordination of groups of mobile autonomous agents using nearest neighbor rules" [IEEE Trans. Automat. Control **48**(6), 988–1001; MR 1986266]. *IEEE Trans. Automat. Control* **48**(9), 1675.

[42] KE, J., MINETT, J. W., AU, C.-P. & WANG, W. S.-Y. (2002) Self-organization and selection in the emergence of vocabulary. *Complexity* **7**(3), 41–54.

[43] KELLER, E. F. & SEGEL, L. A. (1970) Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.* **26**(3), 399–415.

[44] KOCH, A. L. & WHITE, D. (1998) The social lifestyle of myxobacteria. *Bioessays* **20**(12), 1030–1038.

[45] KUSHNER, H. J. (1962) Optimal stochastic control. *IRE Trans. Autom. Control* **7**(5), 120–122.

[46] LACKER, D. (2017) Limit theory for controlled McKean-Vlasov dynamics. *SIAM J. Control Optim.* **55**(3), 1641–1672.

[47] LAURIÈRE, M. & PIRONNEAU, O. (2014) Dynamic programming for mean-field type control. *C. R. Math. Acad. Sci. Paris* **352**(9), 707–713.

[48] LEONARD, N. E. & FIORELLI, E. (2001) Virtual leaders, artificial potentials and coordinated control of groups. In: *Proceedings of the 40th IEEE Conference on Decision and Control, 2001*, Vol. 3. IEEE, pp. 2968–2973.

[49] NIWA, H.-S. (1994) Self-organizing dynamic model of fish schooling. *J. Theor. Biol.* **171**(2), 123–136.

[50] ORRIERI, C. (2018) Large deviations for interacting particle systems: joint mean-field and small-noise limit. *arXiv preprint* arXiv:1810.12636.

[51] PARRISH, J. K. & EDELSTEIN-KESHET, L. (1999) Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* **284**(5411), 99–101.

[52] PARRISH, J. K., VISCIDO, S. V. & GRUNBAUM, D. (2002) Self-organized fish schools: an examination of emergent properties. *Biol. Bull.* **202**(3), 296–305.

[53] PEREA, L., ELOSEGUI, P., & GÓMEZ, G. Extension of the Cucker-Smale control law to space flight formations. *J. Guidance Control Dyn.* **32**(2), 527–537, 2009.

[54] PERTHAME, B. (2007) *Transport Equations in Biology*. Frontiers in Mathematics, Birkhäuser Verlag, Basel.

[55] PHAM, H. & WEI, X. (2018) Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM Control Optim. Calc. Var.*, 24(1), 437–461.

[56] ROMEY, W. L. (1996) Individual differences make a difference in the trajectories of simulated schools of fish. *Ecol. Modell.* **92**(1), 65–77.

[57] ROSSI, R. & SAVARÉ, G. (2003) Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2**(2), 395–431.

[58] SHORT, M. B., D'ORSOGNA, M. R., PASOUR, V. B., TITA, G. E., BRANTINGHAM, P. J., BERTOZZI, A. L. & CHAYES, L. B. (2008) A statistical model of criminal behavior. *Math. Models Methods Appl. Sci.* **18**(suppl.), 1249–1267.

[59] SUGAWARA, K. & SANO, M. (1997) Cooperative acceleration of task performance: foraging behavior of interacting multi-robots system. *Phys. D Nonlinear Phenom.* **100**(3–4), 343–354.

[60] TONER, J. & TU, Y. (1995) Long-range order in a two-dimensional dynamical xy model: how birds fly together. *Phys. Rev. Lett.* **75**(23), 4326.

[61] VICSEK, T., CZIRÓK, A., BEN-JACOB, E., COHEN, I. & SHOCHET, O. (1995) Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**(6), 1226–1229.

[62] VICSEK, T. & ZAFEIRIS, A. (2012) Collective motion. *Phys. Rep.* **517**(3–4), 71–140.

[63] VISINTIN, A. (1984) Strong convergence results related to strict convexity. *Comm. Partial Differ. Equ.* **9**(5), 439–466.

[64] YATES, C. A., ERBAN, R., ESCUDERO, C., COUZIN, I. D., BUHL, J., KEVREKIDIS, I. G., MAINI, P. K. & SUMPTER, D. J. T. (2009) Inherent noise can facilitate coherence in collective swarm motion. *Proc. Nat. Acad. Sci.* **106**(14), 5464–5469.