Trudy Instituta Matematiki i Mekhaniki UrO RAN

Online First 2026

UDK 517.977.56, 517.977.54

VALUE FUNCTION OF THE OPTIMAL CONTROL PROBLEM FOR NONLOCAL BALANCE EQUATION[1]

Yu. Averboukh

This paper studies an optimal control problem for a system governed by a nonlocal balance equation, which models the evolution of a particle distribution. In the examined model, particles move according to a vector field and may disappear. The phase space for this problem is the space of non-negative measures. We prove the existence of an optimal relaxed control, establish a dynamic programming principle, and demonstrate that the value function is a viscosity solution of the corresponding Hamilton–Jacobi equation on the space of non-negative measures.

Keywords: controlled balance equation, optimal control problem, viscosity solution, space of non-negative measures.

Ю. В. Авербух. Функция цены в задаче оптимального управления уравнением баланса.

Данная статья посвящена задаче оптимального управления для системы, описываемой нелокальным уравнением баланса, которое моделирует эволюцию распределения частиц. В рассматриваемой модели частицы движутся в соответствии с векторным полем и могут исчезать. Фазовым пространством для данной задачи является пространство неотрицательных мер. Мы доказываем существование оптимального обобщенного управления, устанавливаем принцип динамического программирования и показываем, что функция значения является вязкостным решением соответствующего уравнения Гамильтона — Якоби в пространстве неотрицательных мер.

Ключевые слова: управляемое уравнение баланса, задача оптимального управления, вязкостное решение, пространство неотрицательных мер.

MSC: 49L12, 35F20, 49L20, 49L25, 82C21

DOI: 10.21538/0134-4889-2026-32-1-fon-01

1. Introduction

This paper considers the following control problem on the space of non-negative measures:

$$\text{minimize } G(m_T) + \int_s^T L(t, m_t, u(t))dt \tag{1.1}$$

subject to the nonlocal balance equation

$$\partial_t m_t + \operatorname{div}(f(t, x, m_t, u(t))m_t) = g(t, x, m(t), u(t))m_t, \quad m_s = \mu, \quad u(t) \in U. \tag{1.2}$$

Here, $t \in [0, T]$ is the time variable, m_t is a measure on $\mathbb{R}^d$, and $u(t)$ is an external control. We will consider only the case of non-positive function g.

Nonlocal balance equation (1.2) has a natural interpretation. The measure m_t describes the distribution of particles in an infinite system, where each particle moves according to the vector field f, which depends on time, the current distribution, and the external control $u(t)$. Additionally, each particle can disappear with a probability rate $-g$, which also depends on time, the particle's

[1]The research was funded by the Russian Science Foundation (project no. 24-21-00373, https://rscf.ru/project/24-21-00373/).

state, the current distribution, and the control $u(t)$. Thus, the problem involves optimizing the distribution of particles.

Note that, if $g \equiv 0$, the nonlocal balance equation reduces to the nonlocal continuity equation. A key feature in this case is the conservation of particles. The control problem for the nonlocal continuity equation is well studied; we refer to [5;8–11] and references therein for recent developments.

We now provide a brief survey of the literature on the nonlocal balance equation, where usually the case of a general function g is examined. The positive values of the function g refer to the appearance of particles. First, we mention papers [1;4;7;12;13;15;19;23;28], which discuss various models of crowd and opinion dynamics that lead to balance equations. Since the phase space for the nonlocal balance equation is the space of non-negative measures, its mathematical analysis requires an appropriate metric. A common approach is to generalize the celebrated Wasserstein distance [29]. A concept based on the Kantorovich-Rubinstein duality was used in [25]. An extension of the Benamou-Brenier formula was considered in [17]. Our approach relies on one proposed by Piccoli and Rossi [20–24] where the metric on the space of measures involves the idea of alignment. In [2], this metric was interpreted using the Wasserstein metric on the space of probability measures over an augmented space. The properties of the nonlocal balance equation have been studied in [2; 20–25], where existence and uniqueness results were proved. Additionally, the solution can be represented via a distribution of weighted curves; this superposition principle is derived in [2;18].

The control problem for the balance equation was recently studied in [16;26], where the authors derived the Pontryagin maximum principle for the linear case (i.e., when f and g do not depend on the measure variable).

As mentioned above, this paper explores the control problem for the nonlinear nonlocal balance equation. The assumption that g is non-positive means that the term only acts as a sink for particles. The main results are as follows:

- the existence of an optimal control;
- the dynamic programming principle for the value function of the examined problem;
- an infinitesimal necessary condition on the value function in the form of viscosity inequalities, which can be regarded as a viscosity solution to a Hamilton–Jacobi equation in the space of non-negative measures.

As a byproduct of the last result, we introduce some concepts of nonsmooth analysis in the space of non-negative measures.

The rest of the paper is organized as follows. Section 2 contains general notation and various concepts for extending the Wasserstein metric to the space of non-negative measures. In Section 3, we discuss the properties of solutions to the nonlocal balance equation (1.2) and prove the existence of an optimal control for problem (1.1), (1.2). The dynamic programming principle is derived in Section 4. Finally, Section 5 provides the fact that the value function is a viscosity solution of a Hamilton–Jacobi equation corresponding to control problem (1.1), (1.2).

2. Preliminaries

2.1. General notation

- If $X_1, \ldots, X_n$ are some sets, $i_1, \ldots, i_k$ are indices from $\{1, \ldots, n\}$, then $\mathrm{p}^{i_1,\ldots,i_k} : X_1 \times \ldots \times X_n \to X_{i_1} \times \ldots \times X_{i_k}$ is the operator assigning to $(x_1, \ldots, x_n)$ the tuple $(x_{i_1}, \ldots, x_{i_k})$. Furthermore, Id denotes the identity mapping.
- If X is a set, $Y \subset X$, then $\mathbb{1}_Y$ is a characteristic function for the set Y.
- If $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ are measurable spaces, m is a measure on $\mathcal{F}$, whereas $h : \Omega \to \Omega'$ is a $\mathcal{F}/\mathcal{F}'$ mapping, then $h\sharp m$ stands for the push-forward measure defined by the rule: for each $\Upsilon \in \mathcal{F}'$,
$$(h\sharp m)(\Upsilon) \triangleq m(h^{-1}(\Upsilon)).$$

- If (X, ρ_X) is a Polish space, then $C_b(X)$ denotes the space of all bounded continuous functions. We consider on $C_b(X)$ the standard sup-norm.
- The Polish space (X, ρ_X) is always endowed with the Borel σ-algebra which is denoted by $\mathcal{B}(X)$. The space of non-negative finite measures on X is denoted by $\mathcal{M}(X)$. Recall that $m \in \mathcal{M}(X)$ is a probability provided $m(X) = 1$. The space of probabilities on X is denoted by $\mathcal{P}(X)$. If $m \in \mathcal{M}(X)$, then $\operatorname{supp}(m)$ denotes its support, i.e., the minimal closed set Υ s.t. $m(X \setminus \Upsilon) = 0$. We consider on $\mathcal{M}(X)$ the topology of narrow convergence, i.e., a sequence $\{m_n\}_{n=1}^\infty \subset \mathcal{M}(X)$ converges narrowly to $m \in \mathcal{M}(X)$ provided that, for each $\phi \in C_b(X)$,

$$\int_X \phi(x) m_n(dx) \to \int_X \phi(x) m(dx) \text{ as } n \to \infty.$$

  Notice that $\mathcal{P}(X)$ is closed w.r.t. the topology of narrow convergence. In Section 2.2, we introduce a different approach to the topology on $\mathcal{M}(X)$ based on an extension of the Wasserstein metric. Notice that, if X is compact, then, due to the Prokhorov theorem, $\mathcal{P}(X)$ is also compact.
- In what follows, we consider on on each time interval $[s, r]$ the Lebesgue σ-algebra, i.e., we enlarge the Borel σ-algebra on $[s, r]$ by adding all null sets. Each function measurable h defined on a time interval $[s, r]$ is assumed to be Lebesgue measurable.
- $\mathbb{R}^d$ denotes the space of column-vectors, whilst $\mathbb{R}^{d,*}$ is the space of row-vectors. We denote the standard Euclidean norm on $\mathbb{R}^d$ by $|\cdot|$. The same symbol is used for the norm on $\mathbb{R}^{d,*}$. If $R > 0$, then $\mathbb{B}_R$ denotes the closed ball in $\mathbb{R}^d$ of the radius R centered at the origin.
- If $\phi : \mathbb{R}^d \to \mathbb{R}$, $x \in \mathbb{R}^d$, then $\nabla\phi(x)$ denotes the derivative of ϕ at x. We assume that $\nabla\phi(x)$ is a row-vector.
- The space of all continuously differentiable functions on $\mathbb{R}^d$ which are bounded with its derivative is denoted by $C_b^1(\mathbb{R}^d)$. Furthermore, $C_c^1(\mathbb{R}^d)$ stands for the set of all functions from $C_b^1(\mathbb{R}^d)$ with a compact support.

2.2. Wasserstein distance and its extension

The aim of this section is to recall the celebrated Wasserstein distance on the space of probability measures and its extensions to the space of non-negative measures. The standard Wasserstein metric is defined on the space of probabilities measures with the finite p-th moment. In what follows, if (X, ρ_X) is a Polish space, then we denote by $\mathscr{P}_p(X)$ the set of probabilities $m \in \mathcal{P}(X)$ s.t., for some (equivalently, every) $x_* \in X$,

$$\int_X \rho_X^p(x, x_*) m(dx) < \infty.$$

The Wasserstein metric on X is defined by the rule: if $m_1, m_2 \in \mathscr{P}_p(X)$,

$$W_p(m_1, m_2) \triangleq \left[\inf_{\pi \in \Pi(m_1, m_2)} \int_{X \times X} \rho_X^p(x_1, x_2) \pi(d(x_1, x_2)) \right]^{1/p}.$$

Here $\Pi(m_1, m_2)$ denotes the set of measures π on $X \times X$ such that $\mathrm{p}^i \sharp \pi = m_i$. It is well known that the space $\mathscr{P}_p(X)$ endowed with the metric W_p is Polish.

Now, let $m_1, m_2 \in \mathcal{M}(X)$. We write $m_1 \leqq m_2$ provided, for each Borel set $\Upsilon \subset X$, $m_1(\Upsilon) \leq m_2(\Upsilon)$. Furthermore, $\|m\|$ denotes the total mass of a measure m. Notice that, if $m_1, m_2 \in \mathcal{M}(X)$ are such that $\|m_1\| = \|m_2\|$, then the set $\Pi(m_1, m_2)$ is well-defined. Now, we introduce the PRW-metric (Piccoli–Rossi–Wasserstein metric also known as generalized Wassesrstein metric) $\mathcal{W}_{p,b}$ for the parameters $p \geq 1$ and $b > 0$ letting

$$\big(\mathcal{W}_{p,b}(m_1, m_2)\big)^p \triangleq \inf\bigg\{b^p\|m_1 - \hat{m}_1\| + b^p\|m_2 - \hat{m}_2\| + \int\limits_{X\times X} \rho_X^p(x_1, x_2)\pi(d(x_1, x_2)) : \hat{m}_1 \leqq m_1,\, \hat{m}_2 \leqq m_2,\, \|\hat{m}_1\| = \|\hat{m}_2\|,\, \pi \in \Pi(\hat{m}_1, \hat{m}_2)\bigg\}.$$

There is an equivalent description of this metric [2]. It relies on the augmented space $X \cup \{\star\}$, where $\star$ is an external point. We assume that the distance on $X \cup \{\star\}$ is defined by the rule:

$$\rho_\star(x_1, x_2) \triangleq \begin{cases} \rho_X(x_1, x_2), & x_1, x_2 \in X, \\ b, & x_1 \in X, x_2 = \star \text{ or } x_1 = \star, x_2 \in X, \\ 0, & x_1 = x_2 = \star. \end{cases}$$

Furthermore, if $R \geq \|m\|$, then $(m \triangleright_R)$ is a probability on $X \cup \{\star\}$ defined by the rule: for each Borel $\Upsilon \in X \cup \{\star\}$,

$$(m \triangleright_R)(\Upsilon) \triangleq R^{-1}\big(m(\Upsilon \cap X) + (R - \|m\|)\mathbb{1}_\Upsilon(\star)\big).$$

It is proved [2, Proposition 1] that, for $m_1, m_2 \in \mathcal{M}(X)$, $R \geq \|m_1\| \vee \|m_2\|$,

$$\mathcal{W}_{p,b}(m_1, m_2) \triangleq R^{1/p}W_p((m_1 \triangleright_R), (m_2 \triangleright_R)).$$

Below we consider only the case when $p = 1$. Additionally, let $\mathscr{M}_1(\mathbb{R}^d)$ denote the space of all non-negative measures $m \in \mathcal{M}(\mathbb{R}^d)$ s.t. $\int_{\mathbb{R}^d} |x| m(dx) < \infty$. Furthermore,

$$\mathscr{M}_1^\rho(\mathbb{R}^d) \triangleq \{m \in \mathscr{M}_1(\mathbb{R}^d) : \|m\| \leq \rho\}.$$

Certainly, $\mathscr{M}_1^\rho(\mathbb{R}^d)$ inherits the PRW-metric. However, it is convenient to consider a metric relying on projections of probability measures on $\mathbb{R}^d \times [0, \rho]$. In what follows, an element of $\mathbb{R}^d \times [0, \rho]$ is regarded as a pair (x, w), where $x \in \mathbb{R}^d$, $w \in \mathbb{R}$. We interpret a pair (x, w) as a 'massed particle'; where w represents a mass. If $\nu \in \mathscr{P}_1(\mathbb{R}^d \times [0, \rho])$, then we let $\lfloor \nu \rfloor$ to be a measure on $\mathbb{R}^d$ defined by the rule: for each $\phi \in C_b(\mathbb{R}^d)$,

$$\int\limits_{\mathbb{R}^d} \phi(x)\lfloor \nu \rfloor(dx) \triangleq \int\limits_{\mathbb{R}^{d+1}} w\phi(y)\nu(d(y, w)).$$

Notice that, given $m \in \mathcal{M}_1^\rho(\mathbb{R}^d)$, the probability $\nu \triangleq (\mathrm{Id}, \|m\|)\sharp(\|m\|^{-1}m)$ is s.t. $\lfloor \nu \rfloor = m$ and lies in $\mathscr{P}_1(\mathbb{R}^d \times [0, \rho])$ for sufficiently large ρ.

Definition 1. The first induced Wasserstein metric on $\mathscr{M}_1^\rho(\mathbb{R}^d)$ is defined by the rule: for $m_1, m_2 \in \mathscr{M}_1^\rho(\mathbb{R}^d)$,

$$\mathbb{W}_{1,\rho}(m_1, m_2) \triangleq \inf\big\{W_1(\nu_1, \nu_2) : \nu_i \in \mathscr{P}_1(\mathbb{R}^d \times [0, \rho]),\, \lfloor \nu_i \rfloor = m_i,\, i = 1, 2\big\}.$$

The following statement can be proved in the same way as [2, Proposition 4].

Proposition 1. *If $\nu_1, \nu_2 \in \mathscr{P}_1(\mathbb{R}^d \times [0, \rho])$, then*

$$\mathcal{W}_{1,b}(\lfloor \nu_1 \rfloor, \lfloor \nu_2 \rfloor) \leq C_0 W_1(\nu_1, \nu_2),$$

where $C_0 \triangleq \rho \vee b$. In particular, if $m_1, m_2 \in \mathscr{M}_1^\rho(\mathbb{R}^d)$, then

$$\mathcal{W}_{1,b}(m_1, m_2) \leq C_0 \mathbb{W}_{1,\rho}(m_1, m_2). \tag{2.3}$$

Finally, given $R, \rho > 0$, we denote by $\mathcal{M}^\rho(\mathbb{B}_R)$ the set of all measures $m \in \mathcal{M}(\mathbb{R}^d)$ s.t. m is concentrated on $\mathbb{B}_R$ and $\|m\| \leq \rho$. Notice that $\mathcal{M}^\rho(\mathbb{B}_R) \subset \mathscr{M}_1^\rho(\mathbb{R}^d)$. Thus, $\mathcal{M}^\rho(\mathbb{R}^d)$ inherits the metric $\mathbb{W}_{1,\rho}$. Moreover, if $m \in \mathcal{M}^\rho(\mathbb{B}_R)$, whereas $\nu \in \mathscr{P}_1(\mathbb{R}^d \times [0, \rho])$ is s.t. $\lfloor \nu \rfloor = m$, then ν is supported on $\mathbb{B}_R \times [0, \rho]$.

3. Balance equation

In what follows, we fix $\rho > 0$ and consider the functions

- $f : [0,T] \times \mathbb{R}^{d+1} \times \mathscr{M}_1^\rho(\mathbb{R}^d) \times U \to \mathbb{R}^d$,
- $g : [0,T] \times \mathbb{R}^{d+1} \times \mathscr{M}_1^\rho(\mathbb{R}^d) \times U \to \mathbb{R}$,
- $L : [0,T] \times \mathscr{M}_1^\rho(\mathbb{R}^d) \times U \to \mathbb{R}$,
- $G : \mathscr{M}_1^\rho \to \mathbb{R}$.

We assume the following:

(H1) the set U is a metric compact;

(H2) the function f is bounded by a constant C_f^0;

(H3) there exists $R > 0$ s.t., for each $t \in [0,T]$, $|x| \geq R$, $m \in \mathscr{M}_1^\rho(\mathbb{R}^d)$, $u \in U$, $f(t,x,m,u) = 0$;

(H4) the function g takes values in $[-C_g^0, 0]$;

(H5) the functions f and g are uniformly continuous w.r.t. t and u;

(H6) the functions f and g are Lipschitz continuous w.r.t. x and m, i.e., there exist constants C_f^1 and C_g^1 s.t., for every $x_1, x_2 \in \mathbb{R}^d$, $m_1, m_2 \in \mathscr{M}_1^\rho(\mathbb{R}^d)$,

$$|f(t,x_1,m_1,u) - f(t,x_2,m_2,u)| \leq C_f^1(|x_1 - x_2| + \mathbb{W}_{1,\rho}(m_1,m_2)),$$

$$|g(t,x_1,m_1,u) - g(t,x_2,m_2,u)| \leq C_g^1(|x_1 - x_2| + \mathbb{W}_{1,\rho}(m_1,m_2));$$

(H7) the function G is continuous.

Remark 1. Certainly, one can consider that the functions f, g, L and G are defined on the whole space $\mathcal{M}(\mathbb{R}^d)$ and the function f, g, L are considered w.r.t. the measure variable in the metric $\mathcal{W}_{1,b}$, while the function G is continuous in this metric. From Proposition 1, the hypotheses of the paper will follow for each $\rho > 0$ and revised Lipschitz constants.

Remark 2. The assumption that the functions f, g, L are Lipschitz continuous w.r.t. the measure variable when the corresponding space is endowed with the metric $\mathbb{W}_{1,\rho}$ looks quite natural. In particular, the functions $\mathscr{M}_1^\rho(\mathbb{R}^d) \ni m \mapsto F\Big(\int_{\mathbb{R}^d} \psi(x)m(dx)\Big)$ are Lipschitz continuous, whenever F and ψ are Lipschitz continuous. This directly follows from the definitions of the operation $\lfloor \cdot \rfloor$ and the metric $\mathbb{W}_{1,\rho}$.

We follow the standard approach of control theory and consider a relaxation of problem (1.1), (1.2). To this end, for $s, r \in [0,T]$, $s < r$, we denote by $\mathcal{U}_{s,r}$ the set of all finite measures ξ on $[s,r] \times U$ s.t. $\mathrm{p}^1 \sharp \xi$ is equal to the Lebesgue measure. Each element of $\mathcal{U}_{s,r}$ is called a relaxed control. On $\mathcal{U}_{s,r}$, we consider the topology of the narrow convergence. The disintegration theorem [14, III-70] states that, given a relaxed control ξ there exists, a weakly measurable function $[s,r] \ni t \mapsto \xi(\cdot|t) \in \mathcal{P}(U)$ satisfying the following condition: for each $\phi \in C([s,r] \times U)$,

$$\int_{[s,r]\times U} \phi(t,u)\xi(d(t,u)) = \int_s^r \phi(t,u)\xi(du|t)dt. \tag{3.1}$$

Recall that a function $\beta : [s,r] \to \mathcal{P}(U)$ is called weakly measurable provided that the mapping $\int_U \phi(u)\beta(t,du)$ is measurable for every $\phi \in C(U)$. Conversely, given a weakly measurable function

$[s,r]\ni t\mapsto \xi(\cdot|t)\in\mathcal{P}(U)$, one can define the measure ξ on $[s,r]\times U$ by rule (3.1). Thus, we identify a relaxed control ξ with its disintegration.

Moreover, if $u(\cdot):[s,r]\to U$ is measurable, then one can define the relaxed control $\xi_{u(\cdot)}$ by the rule $\xi_{u(\cdot)}(\cdot|t)\triangleq\delta_{u(t)}(\cdot)$. Hereinafter, δ_z stands for the Dirac measure concentrated at z. Notice that, for each $\phi\in C([s,r]\times U)$,

$$\int_s^r\int_U \phi(t,u)\xi_{u(\cdot)}(du|t)dt=\int_s^r\phi(t,u(t))dt.$$

Using the relaxation, one can rewrite nonlocal balance equation (1.2) as follows

$$\partial_t m_t+\operatorname{div}\left(\int_U f(t,x,m_t,u)\xi(du|t)\cdot m_t\right)=\int_U g(t,x,m_t,u)\xi(du|t)\cdot m_t. \tag{3.2}$$

Definition 2. Let $s,r\in[0,T]$, $s<r$. We say that a function $[s,r]\ni t\mapsto m_t\in\mathcal{M}(\mathbb{R}^d)$ is a solution of (3.2) provided, for each $\phi\in C^1_c((s,r)\times\mathbb{R}^d)$,

$$\int_s^r\int_{\mathbb{R}^d}\left(\partial_t\phi(t,x)+\nabla\phi(t,x)\int_U f(t,x,m_t,u)\xi(du|t)+\phi(t,x)\int_U g(t,x,m_t,u)\xi(du|t)\right)m_t(dx)dt=0.$$

From [2, Theorem 7], it follows that, given $\mu\in\mathscr{M}_1^\rho(\mathbb{R}^d)$ and $\xi\in\mathcal{U}_{s,r}$, there exists a unique solution to (3.2) satisfying the initial condition

$$m_s=\mu. \tag{3.3}$$

We will show later (see Remark 3) that, in this case, $m_t\in\mathscr{M}_1^\rho(\mathbb{R}^d)$.

Now, let us recall an equivalent description of the solution to the nonlocal balance equation. To this end, we introduce the space of $\Gamma^\rho_{s,r}$ that consists of all continuous curves $\gamma(\cdot)=(x(\cdot),w(\cdot))$ from $[s,r]$ to $\mathbb{R}^d\times[0,\rho]$. If $\gamma(\cdot)=(x(\cdot),w(\cdot))\in\Gamma^\rho_{s,r}$, then e_t is an evaluation operator assigning to a curve $\gamma(\cdot)\in\Gamma^\rho_{s,r}$ the vector $\gamma(t)\in\mathbb{R}^{d+1}$. Furthermore, if $\eta\in\mathscr{P}_1(\Gamma^\rho_{s,r})$ we let $\lfloor\eta\rfloor_t\triangleq\lfloor e_t\sharp\eta\rfloor$. This means that $\lfloor\eta\rfloor_t$ is a measure on $\mathbb{R}^d$ defined by the rule: for every $\phi\in C_b(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d}\phi(x)\lfloor\eta\rfloor_t\triangleq\int_{\Gamma^\rho_{s,r}}w(t)\phi(x(t))\eta(d(x(\cdot),w(\cdot))).$$

In [2, Theorem 7] the following variant of the superposition principle was derived.

Proposition 2. *A flow of probabilities $t\mapsto m_t\in\mathcal{M}_1^\rho(\mathbb{R}^d)$ is a solution of (3.2) for some relaxed control $\xi\in\mathcal{U}_{s,r}$ if and only if, for some $C>0$, there exists a probability $\eta\in\mathscr{P}_1(\Gamma^\rho_{s,r})$ satisfying the following conditions*:

(M1) $m_t=\lfloor\eta\rfloor_t$, $t\in[s,r]$;

(M2) *η-a.e. curve $(x(\cdot),w(\cdot))$ solves*

$$\frac{d}{dt}x(t)=\int_U f(t,x(t),\lfloor\eta\rfloor_t,u)\xi(du|t),$$
$$\frac{d}{dt}w(t)=\int_U g(t,x(t),\lfloor\eta\rfloor_t,u)\xi(du|t)\cdot w(t).$$

Moreover, [2, Theorem 6] states that given $\nu\in\mathscr{P}_1(\mathbb{R}^d\times[0,\rho])$ there exists a unique measure $\eta\in\mathscr{P}_1(\Gamma^{\rho_1}_{s,r})$ satisfying the initial condition $e_s\sharp\eta=\nu$ whenever ρ_1 is sufficiently large.

Remark 3. Due to Hypothesis (H4), if $w(s) \in [0,\rho]$, then, for each $t \in [s,r]$, one has that

$$w(t) \in \big[0,\rho\big].$$

Thus, we always can assume that η satisfying conditions (M1), (M2) is defined on $\mathscr{P}_1(\Gamma^{\rho}_{s,r})$. This gives that, if $m_\cdot$ is a solution of (3.2) for some relaxed control $\xi \in \mathcal{U}_{s,r}$ and the initial measure $\mu \in \mathscr{M}_1^{\rho}(\mathbb{R}^d)$, then $m_t \in \mathscr{M}_1^{\rho}(\mathbb{R}^d)$ for each $t \in [s,r]$.

Notice also that, thanks to Hypothesis (H3), if $x(s) \in \mathbb{B}_R$, then, for each $t \in [s,r]$, $x(t) \in \mathbb{B}_R$. Therefore, $m_t \in \mathcal{M}^{\rho}(\mathbb{B}_R)$ provided that $\mathrm{supp}(\mu) \subset \mathbb{B}_R$.

In what follows, we denote the solution of (3.2) satisfying (3.3) for the relaxed control ξ by $m_\cdot[s,\mu,\xi]$.

Proposition 3. *If $\eta \in \mathscr{P}_1(\Gamma^{\rho}_{s,T})$ satisfies condition (M2), then, for each $r \in [s,T]$,*

$$W_1(e_s\sharp\eta, e_r\sharp\eta) \leq C_1(r-s).$$

If, additionally, $m_t \triangleq \lfloor \eta \rfloor_t$, then

$$\mathbb{W}_{1,\rho}(m_s, m_r) \leq C_1(r-s).$$

Here C_1 is a constant.

The first part of the proposition follows directly from condition (M2) and the fact that the functions f and g are bounded. The second part is a consequence of condition (M1) and Proposition 1.

From now we restrict our attention to the space $\mathcal{M}^{\rho}(\mathbb{B}_R)$. By Remark 3 it is invariant under controlled dynamics (3.2).

The main result of this section is the following.

Theorem 1. *Let*

- *$s, r \in [0,T]$, $s < r$;*
- *for each $n \in \mathbb{N} \cup \{\infty\}$, $\mu^n \in \mathcal{M}^{\rho}(\mathbb{B}_R)$, $\xi^n \in \mathcal{U}_{s,r}$, $m^n_\cdot \triangleq m_\cdot[s,\mu^n,\xi^n]$.*

Assume that $\mathbb{W}_{1,\rho}(\mu^n,\mu^\infty) \to 0$ and $\xi^n \to \xi^\infty$ as $n \to \infty$. Then,

$$\sup_{t\in[s,r]} \mathbb{W}_{1,\rho}(m^n_t, m^\infty_t) \to 0 \text{ as } n \to \infty.$$

Proof. For each $n \in \mathbb{N} \cup \{+\infty\}$, we consider the following system of ODE

$$\begin{aligned}
\frac{d}{dt}x(t) &= \int_U f(t,x(t),m^n_t,u)\xi^n(du|t),\\
\frac{d}{dt}w(t) &= \int_U g(t,x(t),m^n_t,u)\xi^n(du|t)\cdot w(t).
\end{aligned} \tag{3.4}$$

We denote the solution of the first equation in (3.4) with the initial condition $x(s) = y$ by $x^n(\cdot,y)$. Additionally, let $w^n(\cdot,y,z)$ denote the solution of the second equation in (3.4) for $x(\cdot) = x^n(\cdot,y)$ and the initial condition $w(s) = z$. Furthermore, let $\mathcal{X}^n(y,z)$ assign to $(y,z) \in \mathbb{R}^{d+1}$ the whole trajectory $(x^n(\cdot,y), w^n(\cdot,y,z))$. If $\nu^n \in \mathscr{P}_1(\mathbb{B}_R \times [0,\rho])$ is s.t. $\lfloor \nu^n \rfloor = \mu^n$, $\eta^n \triangleq \mathcal{X}^n\sharp\nu^n$, then $\lfloor \eta^n \rfloor_t = m^n_t$. This follows from [2, Lemma 8].

For each $(y,z) \in \mathbb{B}_R \times [0,\rho]$, we denote by $F_i(t,y,z,u)$ the value $f_i(t,x^\infty(t,y),m^\infty_t,u)$. Here, f_i is the i-th coordinate of the function f. Furthermore,

$$F_{d+1}(t,y,z,u) \triangleq g(t,x^\infty(t,y),m^\infty_t,u)w^\infty(t,y,z).$$

Notice that the function F_i, $i = 1,\ldots,d+1$ are bounded and Lipschitz continuous on $\mathbb{B}_R \times [0,\rho]$.

Now, let us choose $\varepsilon > 0$. Due to the Lipschitz continuity of the functions F_i, there exists a finite number of points $\{(y_k^\varepsilon, z_k^\varepsilon)\}_{k=1}^{K^\varepsilon} \subset \mathbb{B}_R \times [0,\rho]$ satisfying the following condition: for each $(y,z) \in \mathbb{B}_R \times [0,\rho]$ there exists a number $k \in \{1,\ldots,K^\varepsilon\}$ s.t. given $i = 1,\ldots,d+1$, $t \in [0,T]$, $u \in U$,

$$|F_i(t,y,z,u) - F_i(t,y_k^\varepsilon,z_k^\varepsilon,u)| \leq \varepsilon. \tag{3.5}$$

Furthermore, the boundedness of the functions F_i gives that there exists a partition of the time interval $[s,r]$ $\{t_i^\varepsilon\}_{i=0}^{I^\varepsilon}$ s.t., for every $t \in [t_i^\varepsilon, t_{i+1}^\varepsilon]$, $i = 0,\ldots,I^\varepsilon - 1$,

$$|F_i(t,y,z,u) - F_i(t_i^\varepsilon,y,z,u))| \leq \varepsilon, \tag{3.6}$$

whenever $(y,z) \in \mathbb{B}_R \times [0,\rho]$, $u \in U$.

For each n, let ν^n and $\nu^{\infty,n}$ be probabilities from $\mathscr{P}_1(\mathbb{B}_R \times [0,\rho])$ s.t.

- $\lfloor \nu^n \rfloor = \mu^n$;
- $\lfloor \nu^{\infty,n} \rfloor = \mu^\infty$;
- $\mathbb{W}_{1,\rho}(\mu^n, \mu^\infty) + \varepsilon \geq W_1(\nu^n, \nu^{\infty,n})$.

Finally, we denote by π^n an optimal plan between ν^n and $\nu^{\infty,n}$.

As we mentioned above, there exist distributions of curves $\eta^n, \eta^{\infty,n} \in \mathscr{P}_1(\Gamma_{s,r}^\rho)$ s.t. $\eta^n = \mathcal{X}^n \sharp \nu^n$, $\eta^{\infty,n} = \mathcal{X}^\infty \sharp \nu^{\infty,n}$. Moreover, $m_t^n = \lfloor \eta^n \rfloor_t$, $m_t^\infty = \lfloor \eta^{\infty,n} \rfloor_t$. Thus,

$$\mathbb{W}_{1,\rho}(m_t^n, m_t^\infty) \leq W_1(e_t \sharp \eta^n, e_t \sharp \eta^{\infty,n}). \tag{3.7}$$

Let N be s.t. for each $n \geq N$ one has that

$$\mathbb{W}_{1,\rho}(\mu^n, \mu^\infty) \leq \varepsilon. \tag{3.8}$$

$$\left| \int_s^{t_i^\varepsilon} \int_U F_i(\tau, y_k^\varepsilon, z_k^\varepsilon, u) \xi^n(du|\tau) d\tau - \int_s^{t_i^\varepsilon} \int_U F_i(\tau, y_k^\varepsilon, z_k^\varepsilon, u) \xi^\infty(du|\tau) d\tau \right| \leq \varepsilon \tag{3.9}$$

whenever $i \in 0,\ldots,I^\varepsilon - 1$, $k \in 1,\ldots K^\varepsilon$. Such an N exists because ξ^n converges narrowly to ξ^∞.

From Hypothesis (H6), we have that, for each $(y^n,z^n),(y^\infty,z^\infty) \in \mathbb{B}_R \times [0,\rho]$,

$$\begin{aligned} |x^n(t,y^n) - x^\infty(t,y^\infty)| &\leq |y^n - y^\infty| \\ &+ C_f^1 \int_s^t \big(|x^n(\tau,y^n) - x^\infty(\tau,y^\infty)| + \mathbb{W}_{1,\rho}(m_\tau^n, m_\tau^\infty) \big) d\tau \\ &+ \sum_{i=1}^d \left| \int_s^t \int_U F_i^0(\tau, y^\infty, z^\infty, u)(\xi^n(du|t) - \xi^\infty(du|t)) dt \right|. \end{aligned}$$

Similarly, thanks to Hypotheses (H4), (H6), we obtain the estimate:

$$\begin{aligned} |w^n(t,y^n,z^n) - w^\infty(t,y^\infty,z^\infty)| &\leq |z^n - z^\infty| \\ &+ \int_s^t \big(C_g^0 |w^n(\tau,y^n,z^n) - w^\infty(\tau,y^\infty,z^\infty)| + C_g^1 |x^n(\tau,y^n) - x^\infty(\tau,y^\infty)| \big) d\tau \\ &+ \int_s^t \mathbb{W}_{1,\rho}(m_\tau^n, m_\tau^\infty) d\tau + \left| \int_s^t \int_U F_{d+1}^0(\tau, y^\infty, z^\infty, u)(\xi^n(du|t) - \xi^\infty(du|t)) dt \right|. \end{aligned}$$

Hence, from Gronwall's inequality,

$$|x^n(t,y^n)-x^\infty(t,y^\infty)|+|w^n(t,y^n,z^n)-w^\infty(t,y^\infty,z^\infty)|\leq C_0'\big(|y^n-y^\infty|+|z^n-z^\infty|\big)$$

$$+C_1'\int_s^t \mathbb{W}_{1,\rho}(m_\tau^n,m_\tau^\infty)d\tau+C_1'\sum_{i=1}^{d+1}\left|\int_s^t\int_U F_i^0(\tau,y^\infty,z^\infty,u)(\xi^n(du|t)-\xi^\infty(du|t))dt\right|.$$

Here C_0', C_1' are some constants.

From now on, we assume that $n\geq N$, where N is chosen according to (3.9). Take additionally into account the choice of points $(y_k^\varepsilon,z_k^\varepsilon)$ (see (3.5)) and times t_a^ε (see (3.6)), we conclude that

$$\left|\int_s^t\int_U F_i^0(\tau,y^\infty,z^\infty,u)(\xi^n(du|t)-\xi^\infty(du|t))dt\right|\leq 5\varepsilon.$$

Thus,

$$|x^n(t,y^n)-x^\infty(t,y^\infty)|+|w^n(t,y^n,z^n)-w^\infty(t,y^\infty,z^\infty)|$$

$$\leq C_0'\big(|y^n-y^\infty|+|z^n-z^\infty|\big)+C_1'\int_s^t \mathbb{W}_{1,\rho}(m_\tau^n,m_\tau^\infty)d\tau+C_2'\varepsilon.$$

Here, $C_2'=C_1'5(d+1)$.

Integrating the latter inequality w.r.t. π^n and using the definition of η^n and $\eta^{\infty,n}$, we conclude that

$$W_1(e_t\sharp\eta^n,e_t\sharp\eta^{\infty,n})\leq C_0'W_1(\nu^n,\nu^{\infty,n})+C_1'\int_s^t \mathbb{W}_{1,\rho}(m_\tau^n,m_\tau^\infty)d\tau+C_2'\varepsilon.$$

Now, let us use (3.7), (3.8) and the fact that $W_1(\nu^n,\nu^{\infty,n})\leq \mathbb{W}_{1,\rho}(\mu^n,\mu^\infty)+\varepsilon$. Thus,

$$\mathbb{W}_{1,\rho}(m_t^n,m_t^\infty)\leq C_1'\int_s^t \mathbb{W}_{1,\rho}(m_\tau^n,m_\tau^\infty)d\tau+\big(C_2'+2C_0')\varepsilon.$$

Using Gronwall's inequality once again, we conclude that, for every $n\geq N$, where N satisfies (3.8), (3.9),

$$\mathbb{W}_{1,\rho}(m_t^n,m_t^\infty)\leq C_3'\varepsilon,$$

where C_3' is a constant that does not depend on n. This gives the desired convergence. □

Corollary 1. *If $\mu\in\mathcal{M}^\rho(\mathbb{B}_R)$, then there exists an optimal relaxed control for optimal control problem (1.1), (1.2), i.e., a relaxed control ξ^* s.t., for every $\xi\in\mathcal{U}_{s,r}$,*

$$G(m_T[s,\mu,\xi^*])+\int_s^T\int_U L(t,m_t[s,\mu,\xi^*])\xi^*(du|t)dt\leq G(m_T[s,\mu,\xi])+\int_s^T\int_U L(t,m_t[s,\mu,\xi])\xi(du|t)dt.$$

This statement directly follows from the compactness of $\mathcal{U}_{s,T}$, the continuous dependence of $m_\cdot[s,\mu,\xi]$ on the relaxed control ξ (see Theorem 1) and the continuity of the functions L and G in the metric $\mathbb{W}_{1,\rho}$. The latter is a consequence of Hypotheses (H6), (H7) and inequality (2.3).

4. Value function and dynamic programming principle

We introduce the value function as follows. First, if $s \in [0, T]$, $\mu \in \mathcal{M}^\rho(\mathbb{B}_R)$, $\xi \in \mathcal{U}_{s,T}$, we put

$$J[s, \mu, \xi] \triangleq G(m_T[s, \mu, \xi]) + \int_s^T \int_U L(t, m_t[s, \mu, \xi])\xi(du|t)dt.$$

Furthermore,

$$\mathrm{Val}(s, \mu) \triangleq \max\left\{J[s, \mu, \xi] : \xi \in \mathcal{U}_{s,T}\right\}$$

Due to Corollary 1, this definition is correct. Below we consider on $\mathcal{M}^\rho(\mathbb{B}_R)$ the metric $\mathbb{W}_{1,\rho}$.

Proposition 4. *The function* Val *is continuous on* $[0, T] \times \mathcal{M}^\rho(\mathbb{B}_R)$.

Proof. First, notice that inequality (2.3) and the assumptions imply that the functions L and G are continuous w.r.t. the metric $\mathbb{W}_{1,\rho}$. The fact that $\mu \mapsto \mathrm{Val}(s, \mu)$ is continuous w.r.t. the metric $\mathbb{W}_{1,\rho}$ follows from this and Theorem 1. Proposition 3 give the continuity w.r.t. the time variable. □

Now let us establish the dynamic programming principle.

Theorem 2. *Let* $s \in [0, T]$, $\mu \in \mathcal{M}^\rho(\mathbb{B}_R)$, $r \in (s, T]$. *Then,*

$$\mathrm{Val}(s, \mu) = \min\left\{\mathrm{Val}(r, m_r[s, \mu, \xi]) + \int_s^r \int_U L(t, m_t[s, \mu, \xi], u)\xi(du|t)dt : \xi \in \mathcal{U}_{s,r}\right\}.$$

Proof. Notice that given $\xi_1 \in \mathcal{U}_{s,r}$, $\xi_2 \in \mathcal{U}_{r,T}$, one can define the concatenation of these relaxed controls using their disintegration

$$(\xi_1 \diamond_r \xi_2)(\cdot|t) \triangleq \begin{cases} \xi_1(\cdot|t), & t \in [s, r), \\ \xi_2(\cdot|t), & t \in [r, T]. \end{cases}$$

Moreover,

$$\mathrm{Val}(s, \mu) = \min_{\xi \in \mathcal{U}_{r,T}} \Bigg[G(m_T[s, \mu, \xi]) + \int_r^T \int_U L(t, m_t[s, \mu, \xi], u)\xi(t|du)dt$$
$$+ \int_s^r \int_U L(t, m_t[s, \mu, \xi], u)\xi(du|t)dt\Bigg].$$

Notice that each relaxed control $\xi \in \mathcal{U}_{s,T}$ can be expressed in the form $\xi = \xi' \diamond_r \xi''$, where ξ' (respectively, ξ'') is the restriction of the relaxed control ξ on $[s, r]$ (respectively, $[r, T]$). Thus,

$$\mathrm{Val}(s, \mu) \geq \min\left\{\mathrm{Val}(r, m_r[s, \mu, \xi]) + \int_s^r \int_U L(t, m_t[s, \mu, \xi], u)\xi(du|t)dt : \xi \in \mathcal{U}_{s,r}\right\}.$$

To derive the converse inequality, we first choose a control $\xi_1 \in \mathcal{U}_{s,r}$ s.t.

$$\mathrm{Val}(r, m_r[s, \mu, \xi_1]) + \int_s^r \int_U L(t, m_t[s, \mu, \xi_1], u)\xi_1(du|t)dt$$

$$= \min\left\{\mathrm{Val}(r, m_r[s, \mu, \xi]) + \int_s^r \int_U L(t, m_t[s, \mu, \xi], u)\xi(du|t)dt : \xi \in \mathcal{U}_{s,r}\right\}.$$

The existence of such a control follows from Corollary 1. Furthermore, we choose $\xi_2 \in \mathcal{U}_{r,T}$ satisfying

$$\mathrm{Val}(r,\mu') = G(m_T[r,\mu',\xi_2]) + \int_r^T \int_U L(t, m_t[r,\mu',\xi_2], u)\xi_2(t|du)dt$$

where $\mu' \triangleq m_r[s,\mu,\xi_1]$. Letting $\hat{\xi} \triangleq \xi_1 \diamond_r \xi_2$, and taking into account the fact that

$$m_t[s,\mu,\hat{\xi}] = \begin{cases} m_t[s,\mu,\xi_1], & t \in [s,r], \\ m_t[r,\mu',\xi_2], & t \in [r,T], \end{cases}$$

we conclude that

$$G(m_T[s,\mu,\hat{\xi}]) + \int_r^T \int_U L(t, m_t[s,\mu,\hat{\xi}], u)\hat{\xi}(t|du)dt$$

$$= G(m_T[r,\mu',\xi_2]) + \int_r^T \int_U L(t, m_t[r,\mu',\xi_2], u)\xi_2(t|du)dt + \int_s^r \int_U L(t, m_t[s,\mu,\xi_1], u)\xi_1(t|du)dt$$

$$= \min\left\{ \mathrm{Val}(r, m_r[s,\mu,\xi]) + \int_s^r \int_U L(t, m_t[s,\mu,\xi], u)\xi(du|t)dt \,:\, \xi \in \mathcal{U}_{s,r} \right\}.$$

Therefore,

$$\mathrm{Val}(s,\mu) \le \min\left\{ \mathrm{Val}(r, m_r[s,\mu,\xi]) + \int_s^r \int_U L(t, m_t[s,\mu,\xi], u)\xi(du|t)dt \,:\, \xi \in \mathcal{U}_{s,r} \right\}. \qquad \square$$

5. Viscosity solution of the Hamilton–Jacobi equations

In this section, we give the infinitesimal characterization of the value function. It states that a value function is a viscosity solution of a corresponding Hamilton–Jacobi equation in the space of non-negative measures. To this end, we first introduce some concept of nonsmooth analysis on $\mathscr{M}_1(\mathbb{R}^d)$. First, we recall the following operation proposed in [3]. For a measure $\mu \in \mathscr{M}_1(\mathbb{R}^d)$, a function $(\xi,\zeta) \in L^1(\mathbb{R}^d,\mu;\mathbb{R}^d\times\mathbb{R})$ s.t. ζ is μ-essentially bounded and a number $\tau > 0$, $\Theta^\tau_\mu[\xi,\zeta] \in \mathscr{M}_1(\mu)$ is defined by the rule: if $\phi \in C_b(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \phi(x)\Theta^\tau_\mu[\xi,\zeta](dx) \triangleq \int_{\mathbb{R}^d} \phi(z+\tau\xi(z)) \exp\big[\tau\zeta(z)\big]\mu(dz).$$

Below, we will consider only $\mu \in \mathcal{M}^\rho(\mathbb{B}_R)$. Additionally, for $\mu \in \mathcal{M}^\rho(\mathbb{B}_R)$, we put

$$\Xi(\mu) \triangleq \big\{(\xi,\zeta) \in L^1(\mathbb{R}^d,\mu;\mathbb{R}^d\times\mathbb{R}) : \xi(x) = 0 \text{ for } x \notin \mathbb{B}_R,\ \zeta(x) \in [-C^0_g, 0] \text{ for } x \in \mathbb{R}^d\big\}.$$

We will consider on $\Xi(\mu)$ the topology inherited from the standard L^1-space. Notice that, for each $\mu \in \mathcal{M}^\rho(\mathbb{B}_R)$, $(\xi,\zeta) \in \Xi(\mu)$ and $\tau > 0$, $\Theta^\tau_\mu(\xi,\zeta) \in \mathcal{M}^\rho(\mathbb{B}_R)$.

The constructions of the nonsmooth analysis presented below may depend on parameters R and ρ. However, we do not indicate this dependence.

In what follows, we fix a function $\phi : \mathcal{M}^\rho(\mathbb{B}_R) \to \mathbb{R}$, $s \in [0,T]$ and a measure $\mu \in \mathscr{M}_1(\mathbb{R}^d)$.

Definition 3. Let $\alpha \in \mathbb{R}$, $(\xi,\zeta) \in \Xi(\mu)$, then the lower directional derivative is defined by the rule:

$$d_D^-\phi(s,\mu;\alpha,\xi,\zeta) \triangleq \liminf_{\substack{h\downarrow 0,\alpha'\to\alpha,\\ (\xi',\zeta')\in\Xi(\mu),\,(\xi',\zeta')\to(\xi,\zeta)}} \frac{\phi(s+h\alpha',\Theta_\mu^h[\xi,\zeta])-\phi(s,\mu)}{h}.$$

The upper directional derivative is defined similarly:

$$d_D^+\phi(s,\mu;\alpha,\xi,\zeta) \triangleq \limsup_{\substack{h\downarrow 0,\alpha'\to\alpha,\\ (\xi',\zeta')\in\Xi(\mu),\,(\xi',\zeta')\to(\xi,\zeta)}} \frac{\phi(s+h\alpha',\Theta_\mu^h[\xi,\zeta])-\phi(s,\mu)}{h}.$$

Definition 4. The directional subdifferential of the function ϕ at (s,μ) contains all triples (a,p,q), where $a\in\mathbb{R}$, $(p,q)\in L^\infty(\mathbb{R}^d,\mu;\mathbb{R}^{d+1,*})$ s.t., for each $\alpha\in\mathbb{R}$ and $(\xi,\zeta)\in\Xi(\mu)$,

$$a\alpha + \int_{\mathbb{R}^d}\big[p(x)\xi(x)+q(x)\zeta(x)\big]\mu(dx) \leq d_D^-\phi(s,\mu;\alpha,\xi,\zeta).$$

The directional subdifferential of the function ϕ at μ is denoted by $\partial_D^-\phi(s,\mu)$.

Similarly, the directional superdifferential of the function ϕ at (s,μ) is denoted by $\partial_D^+\phi(s,\mu)$ and consists of all triples (a,p,q), where $a\in\mathbb{R}$, $(p,q)\in L^\infty(\mathbb{R}^d,\mu;\mathbb{R}^{d+1,*})$ satisfying the following condition: if $\alpha\in\mathbb{R}$, $(\xi,\zeta)\in\Xi(\mu)$,

$$a\alpha + \int_{\mathbb{R}^d}\big[p(x)\xi(x)+q(x)\zeta(x)\big]\mu(dx) \geq d_D^+\phi(s,\mu;\alpha,\xi,\zeta).$$

The Hamiltonian for the examined control problem is defined by the rule: for $s\in[0,T]$, $\mu\in\mathscr{M}_1(\mathbb{R}^d)$, $(p,q)\in L^\infty(\mathbb{R}^d,\mu;\mathbb{R}^{d+1,*})$,

$$H(s,\mu,p,q) \triangleq \min_{u\in U}\Bigg[\int_{\mathbb{R}^d}\big(p(x)f(s,x,\mu,u)+q(x)g(s,x,\mu,u)\big)\mu(dx)+L(s,\mu,u)\Bigg].$$

We consider the following Hamilton–Jacobi equation.

$$\partial_t\varphi + H(t,\mu,\nabla\varphi) = 0. \tag{5.1}$$

This is a formal equation, where we assume that $t\in[0,T]$, $\mu\in\mathcal{M}^\rho(\mathbb{B}_R)$, $\nabla\phi$ is a function from $L^\infty(\mathbb{R}^d,\mu;\mathbb{R}^{d+1,*})$. We will interpret this equation in the viscosity sense following approach of [6;27].

Definition 5. We say that a function $\varphi:[0,T]\times\mathcal{M}^\rho(\mathbb{B}_R)$ is a viscosity supersolution to (5.1) provided that

$$a + H(s,\mu,a,p,q) \leq 0$$

whenever $(a,p,q)\in\partial_D^-\varphi(s,\mu)$.

A function $\varphi:[0,T]\times\mathcal{M}^\rho(\mathbb{B}_R)$ is a viscosity subsolution to (5.1) if

$$a + H(s,\mu,a,p,q) \geq 0$$

for each $(a,p,q)\in\partial_D^+\varphi(s,\mu)$.

We say that a function $\varphi:[0,T]\times\mathcal{M}^\rho(\mathbb{B}_R)$ is a viscosity solution of (5.1) if it is a super- and subsolution simultaneously.

The main result of this section is as follows.

Theorem 3. *The value function of optimal control problem (1.1), (1.2) on $[0,T]\times\mathcal{M}^\rho(\mathbb{B}_R)$ is a viscosity solution of (5.1).*

First, we need the following auxiliary statement.

Lemma 1. *Let*

- $s \in [0,T]$, $\mu \in \mathcal{M}^\rho(\mathbb{B}_R)$;
- $\{h_n\}_{n=1}^\infty$ *be a sequence of positive numbers converging to zero*;
- *for each* n, $\xi_{(n)} \in \mathcal{U}_{s,s+h_n}$; $\zeta_{(n)}$ *be a time-averaging of* $\xi_{(n)}$ *over the interval* $[s, s+h_n]$ *defined for each* $\phi \in C(U)$ *by the rule:*

$$\int_U \phi(u)\zeta_{(n)}(du) \triangleq \frac{1}{h_n} \int_{[s,r]\times U} \phi(u)\xi_{(n)}(d(t,u));$$

- *for each* n,

$$v_{(n)}(y) \triangleq \frac{1}{h_n} \int_{[s,s+h_n]\times U} f(t, X^{s,t}_{(n)}(y), m[t,s,\mu], u)\xi^n(d(t,u)), \tag{5.2}$$

$$b_{(n)}(y) \triangleq \frac{1}{h_n} \int_{[s,s+h_n]\times U} g(t, X^{s,t}_{(n)}(y), m[t,s,\mu], u)\xi^n(d(t,u)), \tag{5.3}$$

where $t \mapsto X^{s,t}_{(n)}(y)$ *stands for the solution map for the initial value problem*

$$\frac{d}{dt}x(t) = \int_U f(t, x(t), m.[s,\mu,\xi_n], u)\xi_{(n)}(du|t), \quad x(t) = y.$$

Assume that $\zeta_{(n)} \to \zeta$ *as* $n \to \infty$ *and put*

$$v(y) = \int_U f(t,y,\mu,u)\zeta(du), \tag{5.4}$$

$$b(y) = \int_U g(t,y,\mu,u)\zeta(du). \tag{5.5}$$

Then, $(v_{(n)}, b_{(n)}), (v,b) \in \Xi(\mu)$ *and* $(v_{(n)}, b_{(n)}) \to (v,b)$ *as* $n \to \infty$.

Proof. First, notice that, due to Hypothesis (H2), the functions $v_{(n)}, v$ are bounded; moreover, $b_{(n)}, b$ take values in $[-C_g^0, 0]$ (this is due to Hypothesis (H4)). Hypotheses (H3) implies that $v_{(n)}$ and v are equal to zero outside the ball $\mathbb{B}_R$. Thus, $(v_{(n)}, b_{(n)}), (v,b) \in \Xi(\mu)$.

Now let us prove that $v_{(n)} \to v$ pointwise. Indeed, recall that $\mathcal{W}_{1,b}(m_t[s,\mu,\xi_{(n)}],\mu) \le C_2|t-s|$ (see Proposition 3). Similarly, notice that $|X^{s,t}_{(n)}(y) - y| \le C_1''|t-s|$, where C_1'' is a constant. Thus, for each $y \in \mathbb{R}^d$, we have that

$$\|v_{(n)}(y) - v(y)\| \le \left| \int_U f(s,y,\mu,u)\zeta_{(n)}(du) - \int_U f(s,y,\mu,u)\zeta(du) \right|$$

$$+ \frac{1}{h_n}\left| \int_{[s,s+h_n]\times U} f(t,y,\mu,u)\xi_{(n)}(d(t,u)) - \int_{[s,s+h_n]\times U} f(s,y,\mu,u)\xi_{(n)}(d(t,u)) \right| + C_f^1(C_2 + C_1'')h_n.$$

Using the uniform continuity of the function f w.r.t. the time variable and the fact that $\zeta_{(n)} \to \zeta$ narrowly, we prove that $v_{(n)} \to v$ pointwise.

The fact that $b_{(n)} \to b$ pointwise is proved analogously.

Since the functions $(v_{(n)}, b_{(n)})$ are uniformly bounded, we have that $(v_{(n)}, b_{(n)}) \to (v,b)$ in the L^1-sense. □

Proof of Theorem 3. First, we prove that Val is a viscosity supersolution to (5.1). To this end, we fix $s \in [0, T]$, $\mu \in \mathcal{M}^\rho(\mathbb{B}_R)$ and $(a, p, q) \in \partial_D^- \operatorname{Val}(s, \mu)$. Now, let $\tau > 0$. Due to the dynamic programming principle, there exists a relaxed control $\xi_h \in \mathcal{U}_{s,s+h}$ s.t.

$$\operatorname{Val}(s, \mu) = \operatorname{Val}(s + h, m_{s+h}[s, \mu, \xi_h]) + \int_s^{s+h}\int_U L(t, m_t[s, \mu, \xi_h], u)\xi_h(du|t). \tag{5.6}$$

As above, let $t \mapsto X_h^{s,t}(y)$ denote a solution map for the initial value problem

$$\frac{d}{dt}x(t) = \int_U f(t, x(t), m_t[s, \mu, \xi_h], u)\xi_h(du|t), \quad x(t) = y.$$

We put

$$v_h(y) \triangleq \frac{1}{h}\int_{[s,s+h_n]\times U} f(t, X_h^{s,t}(y), m_t[s, \mu, \hat{\xi}_h], t)\xi_h(d(u, t)),$$

$$b_h(y) \triangleq \frac{1}{h}\int_{[s,s+h_n]\times U} g(t, X_h^{s,t}(y), m_t[s, \mu, \xi_h], t)\xi_h(d(u, t)).$$

Now we define the averaging of the measure ξ_h over $[s, s + h]$ by the rule: for $\phi \in C(U)$,

$$\int_U \phi(u)\zeta_h(du) \triangleq \frac{1}{h}\int_{[s,s+h]\times U} \phi(u)\xi_h(d(t, u)).$$

Notice that, due to the compactness of $\mathcal{P}(U)$ (this is a consequence of the celebrated Prokhorov theorem) there exists a sequence $\{h_n\}_{n=1}^\infty \subset (0, +\infty)$ s.t.

- $h_n \to 0$ as $n \to \infty$;
- $\zeta_{(n)} \triangleq \zeta_{h_n}$ converges narrowly to some $\zeta \in \mathcal{P}(U)$.

In what follows, we put $\xi_{(n)} \triangleq \xi_{h_n}$, $v_{(n)} \triangleq v_{h_n}$, $b_{(n)} \triangleq v_{h_n}$. By Proposition 1, $(v_{(n)}, b_{(n)})$ converge to (v, b), where

$$v(y) \triangleq \int_U f(s, y, \mu, u)\zeta(du), \ b(y) \triangleq \int_U g(s, y, \mu, u)\zeta(du).$$

Furthermore,

$$m_{s+h}[s, \mu, \xi_h] = \Theta_\mu^h(v_h, b_h). \tag{5.7}$$

Finally, from the very definition of the measures $\zeta_{(n)}$, the fact that the sequence $\{\zeta_{(n)}\}_{n=1}^\infty$ converges narrowly to ζ, Proposition 3 and Hypothesis (H5), (H6), we obtain that

$$\left|\int_s^{s+h_n}\int_U L(t, m_t[s, \mu, \xi_{(n)}], u)\xi_{(n)}(du|t)dt - h_n\int_U L(s, \mu, u)\zeta(du|t)\right| \leq h_n\varsigma_n,$$

where $\varsigma_n \to 0$ as $n \to \infty$. Thus, from (5.6), (5.7), it follows that, for each n,

$$h_n\sigma_n \geq \operatorname{Val}(s + h_n, \Theta_\mu^h(v_{(n)}, b_{(n)})) - \operatorname{Val}(s, \mu) + h_n\int_U L(s, \mu, u)\zeta(du).$$

Recall that, by the definition of the lower directional derivative,

$$\liminf_{n\to\infty} \frac{\operatorname{Val}(s + h_n, \Theta_\mu^h(v_{(n)}, b_{(n)})) - \operatorname{Val}(s, \mu)}{h_n} \geq d_D^- \operatorname{Val}(s, \mu; 1, v, b).$$

Since $(a,p,q)\in\partial_D^-\operatorname{Val}(s,\mu)$,

$$0\geq\liminf_{n\to\infty}\frac{\operatorname{Val}(s+h_n,\Theta_\mu^h(v_{(n)},b_{(n)}))-\operatorname{Val}(s,\mu)}{h_n}+\int\limits_U L(s,\mu,u)\zeta(du)$$

$$\geq a+\int\limits_{\mathbb{R}^d}\big[p(y)v(y)+q(y)b(y)\big]\mu(dy)+\int\limits_U L(s,\mu,u)\zeta(du)$$

$$=a+\int\limits_U\Bigg[\int\limits_{\mathbb{R}^d}\big(p(y)f(s,y,\mu,u)+q(y)g(s,y,\mu,u)\big)\mu(dy)+L(s,\mu,u)\Bigg]\zeta(du).$$

In the last step, we used the definition of the functions v, b and applied the Tonelli theorem.

Since, for each $u\in U$,

$$\int\limits_{\mathbb{R}^d}\big(p(y)f(s,y,\mu,u)+q(y)g(s,y,\mu,u)\big)\mu(dy)+L(s,\mu,u)\geq H(s,\mu,p,q),$$

we conclude that Val is a viscosity supersolution to (5.1).

To prove the fact that Val is a viscosity subsolution to (5.1), we fix $s\in[0,T]$, $\mu\in\mathcal{M}^\rho(\mathbb{B}_R)$ and $(a,p,q)\in\partial_D^+\operatorname{Val}(s,\mu)$. There exists $\bar{u}$ s.t.

$$\int\limits_{\mathbb{R}^d}\big(p(y)f(s,y,\mu,\bar{u})+q(y)g(s,y,\mu,\bar{u})\big)\mu(dy)+L(s,\mu,\bar{u})=H(s,\mu,p,q). \tag{5.8}$$

Let $\bar{\zeta}\triangleq\delta_{\bar{u}}$ and let $\bar{\xi}$ be a constant relaxed control s.t. $\bar{\xi}(\cdot|t)\triangleq\bar{\zeta}(\cdot)=\delta_{\bar{u}}(\cdot)$. Furthermore, let $\{h_n\}_{n=1}^\infty\subset(0,+\infty)$ converge to 0. Now we use notation (5.2)–(5.5) for $\xi_{(n)}\triangleq\bar{\xi}$, $\zeta_{(n)}=\bar{\zeta}$, $\zeta=\bar{\zeta}$. In this case, Proposition 1 gives that $(v_{(n)},b_{(n)})$ converge to (v,b). Furthermore, notice that

$$m_{s+h_n}[s,\mu,\bar{\xi}]=\Theta_\mu^{h_n}(v_{(n)},b_{(n)}).$$

Hence, the dynamic programming principle (see Theorem 2) gives that

$$\operatorname{Val}(s,\mu)\leq\operatorname{Val}(s+h_n,\Theta_\mu^{h_n}(v_{(n)},b_{(n)}))+\int\limits_s^{s+h_n}L(t,m_t[s,\mu,\bar{\xi}],\bar{u})dt. \tag{5.9}$$

As above, from Proposition 3 and Hypothesis (H5), (H6), it follows that, if $t\in[s,s+h_n]$,

$$\big|L(t,m_t[s,\mu,\bar{\xi}],\bar{u})dt-L(s,\mu,\bar{u})\big|\leq\varsigma_n',$$

where $\varsigma_n'\to 0$ as $n\to\infty$. Therefore, thanks to (5.9), we conclude that

$$-h_n\varsigma_n'\leq\operatorname{Val}(s+h_n,\Theta_\mu^{h_n}(v_{(n)},b_{(n)}))-\operatorname{Val}(s,\mu)+h_nL(s,\mu,\bar{u}). \tag{5.10}$$

Recall that (see Definitions 3, 4) that

$$\limsup_{n\to\infty}\frac{\operatorname{Val}(s+h_n,\Theta_\mu^{h_n}(v_{(n)},b_{(n)}))-\operatorname{Val}(s,\mu)}{h_n}$$

$$\leq d_D^+\operatorname{Val}(s,\mu;1,v,b)\leq a+\int\limits_{\mathbb{R}^d}\big[p(y)v(y)+q(q)b(y)\big]\mu(dy).$$

This and (5.10) imply that

$$0\leq\int\limits_{\mathbb{R}^d}\big(p(y)f(s,y,\mu,\bar{u})+q(y)g(s,y,\mu,\bar{u})\big)\mu(dy)+L(s,\mu,\bar{u}).$$

Taking into account the choice of the control $\bar{u}$ (see (5.8)), we derive the fact that Val is a viscosity subsolution to (5.1). □

REFERENCES

1. Albi G., Bongini M., Rossi F., Solombrino F. Leader formation with mean-field birth and death models. *Math. Models Methods Appl. Sci.*, 2019, vol. 29, no. 4, pp. 633–679. https://doi.org/10.1142/S0218202519400025
2. Averboukh Yu. Nonlocal balance equation: representation and approximation of solution. *J. Dyn. Diff. Equat.*, 2025, vol. 37, pp. 2461–2495. https://doi.org/10.1007/s10884-024-10373-8
3. Averboukh Yu.V. Nagumo-type viability theorem for nonlocal balance equation. *Izv. IMI UdGU*, 2024, vol. 64, pp. 3–16. https://doi.org/10.35634/2226-3594-2024-64-01
4. Ayi N., Pouradier Duteil N. Mean-field and graph limits for collective dynamics models with time-varying weights. *J. Differ. Equat.*, 2021, vol. 299, pp. 65–110. https://doi.org/10.1016/j.jde.2021.07.010
5. Badreddine Z., Frankowska H. Hamilton–Jacobi inequalities on a metric space. *J. Differ. Equat.*, 2021, vol. 271, pp. 1058–1091. https://doi.org/10.1016/j.jde.2020.09.026
6. Bardi M., Capuzzo-Dolcetta I. Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations. Boston, Birkhäuser Boston, 1997, 574 p. https://doi.org/10.1007/978-0-8176-4755-1
7. Bayen A., Friedrich J., Keimer A., Pflug L., Veeravalli T. Modeling multilane traffic with moving obstacles by nonlocal balance laws. *SIAM J. Appl. Dyn. Syst.*, 2022, vol. 21, no. 2, pp. 1495–1538. https://doi.org/10.1137/20M1366654
8. Bonnet B.A. Pontryagin maximum principle in Wasserstein spaces for constrained optimal control problems. *ESAIM: Control Optim. Calc. Var.*, 2019, vol. 25, art. no. 52, 38 p. https://doi.org/10.1051/cocv/2019044
9. Bonnet B., Frankowska H. Mean-field optimal control of continuity equations and differential inclusions. In: *Proc. 2020 59th IEEE Conference on Decision and Control* (*CDC*), Jeju, Korea (South), 2020, pp. 470–475. https://doi.org/10.1109/CDC42340.2020.9303835
10. Bonnet B., Frankowska H. Semiconcavity and sensitivity analysis in mean-field optimal control and applications. *J. Math. Pures Appl.*, 2022, vol. 157, pp. 282–345. https://doi.org/10.1016/j.matpur.2021.11.001
11. Bonnet B., Rossi F. The Pontryagin maximum principle in the Wasserstein space. *Calc. Var. Partial Differ. Equat.*, 2019, vol. 58, no. 1, art. no. 11, 36 p. https://doi.org/10.1007/s00526-018-1447-2
12. Carlini E., Festa A., Silva F.J., Wolfram M.-T. A semi-lagrangian scheme for a modified version of the hughes' model for pedestrian flow. *Dyn. Games Appl.*, 2017, vol. 7, no. 4, pp. 683–705. https://doi.org/10.1007/s13235-016-0202-6
13. Colombo R.M., Marcellini F. Nonlocal systems of balance laws in several space dimensions with applications to laser technology. *J. Differ. Equat.*, 2015, vol. 259, no. 11, pp. 6749–6773. https://doi.org/10.1016/j.jde.2015.08.005
14. Dellacherie C., Meyer P.-A. *Probabilities and potential*. Ser. North-Holland Math. Studies, Hermann, North-Holland Pub. Co., 1978, vol. 29, 463 p. ISBN-10: 072040701X.
15. Duteil N.P. Mean-field limit of collective dynamics with time-varying weights. *Netw. Heterog. Media.*, 2022, vol. 17, no. 2, pp. 129–161. https://doi.org/10.3934/nhm.2022001
16. Goncharova E.V., Pogodaev N.I., Staritsyn M.V. Exact formulas for the increment of the cost functional in optimal control of linear balance equation. *Bull. Irkutsk State Univ. Ser. Math.*, 2025, vol. 51, pp. 3–20. https://doi.org/10.26516/1997-7670.2025.51.3
17. Kondratyev S., Monsaingeon L., Vorotnikov D. A new optimal transport distance on the space of finite Radon measures. *Adv. Differential Equations*, 2016, vol. 21, pp. 1117–1164. https://doi.org/10.57262/ade/1476369298
18. Maniglia S. Probabilistic representation and uniqueness results for measure-valued solutions of transport equations. *J. Math. Pures Appl.*, 2007, vol. 87, no. 6, pp. 601–626. https://doi.org/10.1016/j.matpur.2007.04.001
19. McQuade S., Piccoli B., Duteil N.P. Social dynamics models with time-varying influence. *Math. Models Methods Appl. Sci.*, 2019, vol. 29, no. 4, pp. 681–716. https://doi.org/10.1142/S0218202519400037
20. Piccoli B., Duteil N.P. Control of collective dynamics with time-varying weights. In: Salvarani F. (eds.) *Conf. paper "Recent advances in kinetic equations and applications"*, Springer INdAM Ser., vol. 48, Cham, Springer, 2021, pp. 289–308. https://doi.org/10.1007/978-3-030-82946-9_12
21. Piccoli B., Rossi F. Generalized Wasserstein distance and its application to transport equations with source. *Arch. Ration. Mech. Anal.*, 2014, vol. 211, no. 1, pp. 335–358. https://doi.org/10.1007/s00205-013-0669-x

22. Piccoli B., Rossi F. On properties of the generalized Wasserstein distance. *Arch. Ration. Mech. Anal.*, 2016, vol. 222, pp. 1339–1365. https://doi.org/10.1007/s00205-016-1026-7
23. Piccoli B., Rossi F. Measure-theoretic models for crowd dynamics. In: Gibelli L., Bellomo N. (eds.) *Crowd Dynamics, vol. 1. Modeling and simulation in science, engineering and technology*. Cham, Birkhäuser, 2018, pp. 137–165. https://doi.org/10.1007/978-3-030-05129-7_6
24. Piccoli B., Rossi F., Tournus M. *A Wasserstein norm for signed measures, with application to nonlocal transport equation with source term*, 2019, 26 p. https://doi.org/10.48550/arXiv.1910.05105
25. Pogodaev N.I., Staritsyn M.V. Nonlocal balance equations with parameters in the space of signed measures. *Mat. Sb.*, 2022, vol. 213, no. 1, pp. 63–87. https://doi.org/10.1070/SM9516
26. Pogodaev N.I., Staritsyn M.V. optimal control of nonlocal balance equations in the space of nonnegative measures. *Sib. Math. J.*, 2025, vol. 66, pp. 576–593. https://doi.org/10.1134/S0037446625020223
27. Subbotin A.I. *Generalized solutions of first-order PDEs: the dynamical optimization perspective*, Ser. Systems & Control: Foundations & Applications (SCFA), Boston, Birkhäuser, 1994, 314 p. https://doi.org/10.1007/978-1-4612-0847-1
28. Thai M.-N. *Birth and death process in mean field type interaction*, 2015, 19 p. https://doi.org/10.48550/arXiv.1510.03238
29. Villani C. *Optimal transport: old and new*, Ser. Grundlehren der mathematischen Wissenschaften; vol. 338. Berlin, Heidelberg, Springer, 2009, 976 p. https://doi.org/10.1007/978-3-540-71050-9

Received October 10, 2025
Revised October 31, 2025
Accepted November 3, 2025
Published online November 19, 2025

Funding Agency: The research was funded by the Russian Science Foundation (project no. 24-21-00373, https://rscf.ru/project/24-21-00373/).

Yurii Vladimirovich Averboukh, Dr. Sci. in Math., Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences, Yekaterinburg, 620077 Russia, e-mail: ayv@imm.uran.ru .

Cite this article as: Yu. Averboukh. Value function of the optimal control problem for nonlocal balance equation. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2026.
https://doi.org/10.21538/0134-4889-2026-32-1-fon-01