Random Lifting of Set-Valued Maps

Rossana Capuani(✉), Antonio Marigonda, and Marta Mogentale

Department of Computer Sciences, University of Verona,
Strada Le Grazie 15, 37134 Verona, Italy
{rossana.capuani,antonio.marigonda}@univr.it

Abstract. In this paper we discuss the properties of particular set-valued maps in the space of probability measures on a finite-dimensional space that are constructed by mean of a suitable *lift* of set-valued map in the underlying space. In particular, we are interested to establish under which conditions some good regularity properties of the original set-valued map are inherited by the lifted one. The main motivation for the study is represented by multi-agent systems, i.e., finite-dimensional systems where the number of (microscopic) agents is so large that only macroscopical description are actually available. The macroscopical behaviour is thus expressed by the superposition of the behaviours of the microscopic agents. Using the common description of the state of a multi-agent system by mean of a time-dependent probability measure, expressing the fraction of agents contained in a region at a given time moment, the results of this paper yield regularity results for the macroscopical behaviour of the system.

Keywords: Set-valued map · Multi-agent systems · Statistical description

1 Introduction

In the last decade, the mathematical analysis of *complex systems* attracted a renewed interest from the applied mathematics community in view of its capability to model many real-life phenomena with a good degree of accuracy. In particular, the field of application of such models ranges from social dynamics (e.g., pedestrian dynamics, social network models, opinion formation, infrastructure planning) to financial markets, from big data analysis to life sciences (e.g. flocking).

All those systems are characterized by the presence of a *large* number of individuals, called *agents*, usually moving in a finite-dimensional space $\mathbb{R}^d$ under the effect of a *global field* which is possibly affected also by the current agent configuration. In its simplest setting, each agent moves along the steepest descent direction of a functional (which is the *same* for all the agents) which can also take into account interaction effects between them. The interaction between the agents may range from the simplest, e.g., avoiding collision, or attraction/repulsion effects, to more complex ones, involving also penalization of overcrowding/dispersion, or further state constraints on the density of the agents.

I. Lirkov and S. Margenov (Eds.): LSSC 2021, LNCS 13127, pp. 297–305, 2022.
https://doi.org/10.1007/978-3-030-97549-4_34

Due to the huge number of agents, a description of the motion of each agent becomes impossible. Therefore, using the simplifying assumptions that all the agents of the collective are *indistinguishable*, only a *macroscopical* (statistical) description of the system is feasible. In this sense, at each instant t of time the state of the system is described by a time-depending positive Borel measure μ_t whose meaning is the following: given a region $A \subseteq \mathbb{R}^d$, the quotient $\frac{\mu_t(A)}{\mu_t(\mathbb{R}^d)}$ represents the fraction of agents that at time t are present in the region A. If we suppose that the total amount of the agents does not change in time, we can normalize the quotient by taking $\mu_t(\mathbb{R}^d) \equiv 1$, i.e., the evolution of the system can be represented by a family of probability measures indexed by the time parameter.

Under reasonable assumptions on the agents' trajectories, the family of measures describing the evolution of the system obeys to the *continuity equation*

$$\partial_t \mu_t + \operatorname{div}(v_t \mu_t) = 0, \tag{1}$$

coupled with an initial data μ_0, representing the initial state of the system. The equation must be understood in the sense of distribution. The time-depending vector field $(t, x) \mapsto v_t(x)$ represents the macroscopical vector field along which the mass flows, and $v_t \mu_t$ is the flux. This leads to an absolutely continuous curve in the space of probability measures, endowed with the Wasserstein distance.

The link between the trajectories of the microscopic agents and the macroscopical evolution of the system is given by the *superposition principle* (see [1] and [7] in the context of differential inclusions): namely, every solution $t \mapsto \mu_t$ of (1) can be represented by the pushforward $e_t \sharp \boldsymbol{\eta}$ of a probability measure $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times C^0([0,T]))$ concentrated on pairs (x, γ), where γ is any integral solution of $\dot{\gamma}(t) = v_t(\gamma(t))$ satisfying $\gamma(0) = x$, and $e_t(x, \gamma) = \gamma(t)$ is the evaluation operator. Conversely, given a Borel family of absolutely continuous curves in $[0, T]$, any probability measure $\boldsymbol{\eta}$ concentrated on the pairs $(\gamma(0), \gamma)$ defines by the pushforward $\mu_t = e_t \sharp \boldsymbol{\eta}$ an absolutely continuous curve $t \mapsto \mu_t$, which solves (1) for a vector field $v_t(\cdot)$ representing the weighted average of the speeds of the trajectories of the agents concurring at point x at time t. We recall that in this case $v_t(\cdot)$ is an *average* and therefore it may happens that *no agent* is following the integral curves of the vector field, even if the macroscopical effects will be a displacement along it.

The problem discussed in this paper is the following. We suppose to have a set-valued map S associating to every point $x \in \mathbb{R}^d$ a set of curves in $\mathbb{R}^d$ representing the allowed trajectories of the agent which at initial time $t = 0$ is at x. Our goal is to study the corresponding properties of the set-valued map describing the family of macroscopical trajectories in the space of probability measures. It turns out that this amounts to study the properties of a new set-valued map associating to each probability measure μ the set of probability measure concentrated on the graph of $S(\cdot)$ and whose first marginal is equal to μ.

The main example is when the set $S(x)$ describes the trajectories of a differential inclusion with initial data x, and we want to derive regularity properties of the macroscopical evolutions from the properties of the set-valued map defining the differential inclusion. The regularity properties of the solution map is crucial in the study of many optimization problems and to generalize mean field models, like e.g. in [4–6].

The paper is structured as follows: in Sect. 2 we give some preliminaries and basic definitions, in Sect. 3 we prove the main results, while in Sect. 4 we explore possible extensions and further developments.

2 Preliminaries

Let (X, d_X) be a separable metric space. We denote by $\mathscr{P}(X)$ the set of Borel probability measures on X endowed with the weak* topology induced by the duality with the Banach space $C_b^0(X)$ of the real-valued continuous bounded functions on X with the uniform convergence norm. For any $p \geq 1$, we set the space of Borel probability measures with finite *p-moment* as

$$\mathscr{P}_p(X) = \left\{ \mu \in \mathscr{P}(X) : \int_X d_X^p(x, \bar{x})\, d\mu(x) < +\infty \text{ for some } \bar{x} \in X \right\}.$$

Given complete separable metric spaces (X, d_X), (Y, d_Y), for any Borel map $r : X \to Y$ and $\mu \in \mathscr{P}(X)$, we define the *push forward measure* $r\sharp\mu \in \mathscr{P}(Y)$ by setting $r\sharp\mu(B) = \mu(r^{-1}(B))$ for any Borel set B of Y.

Definition 1 (Transport plans and Wasserstein distance). *Let X be a complete separable metric space, $\mu_1, \mu_2 \in \mathscr{P}(X)$. We define the set of* admissible transport plans *between μ_1 and μ_2 by setting $\Pi(\mu_1, \mu_2) = \{\boldsymbol{\pi} \in \mathscr{P}(X \times X) : \mathrm{pr}_i\sharp\boldsymbol{\pi} = \mu_i,\ i = 1, 2\}$, where for $i = 1, 2$, we defined $\mathrm{pr}_i : X \times X \to X$ by $\mathrm{pr}_i(x_1, x_2) = x_i$. The* inverse *$\boldsymbol{\pi}^{-1}$ of a transport plan $\boldsymbol{\pi} \in \Pi(\mu, \nu)$ is defind by $\boldsymbol{\pi}^{-1} = i\sharp\boldsymbol{\pi} \in \Pi(\nu, \mu)$, where $i(x, y) = (y, x)$ for all $x, y \in X$. The p-*Wasserstein distance *between μ_1 and μ_2 is*

$$W_p^p(\mu_1, \mu_2) = \inf_{\boldsymbol{\pi} \in \Pi(\mu_1, \mu_2)} \int_{X \times X} d_X^p(x_1, x_2)\, d\boldsymbol{\pi}(x_1, x_2).$$

If $\mu_1, \mu_2 \in \mathscr{P}_p(X)$ then the above infimum is actually a minimum, and we define

$$\Pi_o^p(\mu_1, \mu_2) = \left\{ \boldsymbol{\pi} \in \Pi(\mu_1, \mu_2) : W_p^p(\mu_1, \mu_2) = \int_{X \times X} d_X^p(x_1, x_2)\, d\boldsymbol{\pi}(x_1, x_2) \right\}.$$

The space $\mathscr{P}_p(X)$ endowed with the W_p-Wasserstein distance is a complete separable metric space, moreover for all $\mu \in \mathscr{P}_p(X)$ there exists a sequence $\{\mu^N\}_{N \in \mathbb{N}} \subseteq \mathrm{co}\{\delta_x : x \in \operatorname{supp}\mu\}$ such that $W_p(\mu^N, \mu) \to 0$ as $N \to +\infty$.

Remark 1. Recalling formula (5.2.12) in [1], when X is a separable Banach space we have $W_p(\delta_0, \mu) = \mathrm{m}_p^{1/p}(\mu) = \left(\int_{\mathbb{R}^d} \|x\|_X^p\, d\mu(x) \right)^{1/p}$, for all $\mu \in \mathscr{P}_p(X), p \geq 1$. In particular, if $t \mapsto \mu_t$ is W_p-continuous, then $t \mapsto \mathrm{m}_p^{1/p}(\mu_t)$ is continuous.

Definition 2 (Set-valued maps). *Let X, Y be sets. A* set-valued *map F from X to Y is a map associating to each $x \in X$ a (possible empty) subset $F(x)$ of Y. We will write $F : X \rightrightarrows Y$ to denote a set-valued map from X to Y. The* graph *of a set-valued map F is* $\operatorname{graph} F := \{(x,y) \in X \times Y : y \in F(x)\} \subseteq X \times Y$*, while the* domain *of F is* $\operatorname{dom} F := \{x \in X : F(x) \neq \emptyset\} \subseteq X$*. Given $A \subseteq X$, we set* $\operatorname{graph}(F_{|A}) := \operatorname{graph} F \cap (A \times Y) = \{(x,y) \in A \times Y : y \in F(x)\}$*. A* selection *of F is a map $f : \operatorname{dom} F \to Y$ such that $f(x) \in F(x)$ for all $x \in \operatorname{dom} F$.*

The following Lemma is a direct consequence of [8, Theorem 7.1].

Lemma 1 (Borel selection of the metric projection). *Let X, Y be complete separable metric spaces. Assume that $S : X \rightrightarrows Y$ is a continuous set-valued map with compact nonempty images. Then there exists a Borel map $g : X \times Y \to Y$ such that $g(x,y) \in S(x)$ and $d_Y(y, g(x,y)) = d_{S(x)}(y)$ for all $(x,y) \in X \times Y$, i.e., for every $x \in X$ and $y \in Y$ we have that $g(x,y)$ is a metric projection of y on $S(x)$.*

3 Results

In this section, we will introduce the notion of random lift of set-valued maps and its main properties.

Definition 3 (Random lift of set-valued maps). *Let X, Y be complete separable metric spaces, $S : X \rightrightarrows Y$ be a set valued map. Define the set-valued map $\mathscr{P}(S) : \mathscr{P}(X) \rightrightarrows \mathscr{P}(X \times Y)$ as follows*

$$\mathscr{P}(S)(\mu) := \left\{ \boldsymbol{\eta} \in \mathscr{P}(X \times Y) : \operatorname{graph} S \supseteq \operatorname{supp} \boldsymbol{\eta}, \ \textit{and} \ \mathrm{pr}^{(1)} \sharp \boldsymbol{\eta} = \mu \right\},$$

where $\mu \in \mathcal{P}(X)$ and $\mathrm{pr}^{(1)}(x,y) = x$ for all $(x,y) \in X \times Y$. The set-valued map $\mathscr{P}(S)(\cdot)$ will be called the random lift *of $S(\cdot)$.*

Remark 2. Directly from the definition, we have that

1. $\mathscr{P}(S)$ has convex images (even if S has not convex images): indeed for all $\boldsymbol{\eta_i} \in \mathscr{P}(S)(\mu)$, $i = 1, 2$, and $\lambda \in [0,1]$, set $\boldsymbol{\eta}_\lambda = \lambda \boldsymbol{\eta_1} + (1-\lambda)\boldsymbol{\eta_2}$, and notice that $\operatorname{supp} \boldsymbol{\eta}_\lambda \subseteq \operatorname{supp} \boldsymbol{\eta_1} \cup \operatorname{supp} \boldsymbol{\eta_2} \subseteq \operatorname{graph} S$, and for all Borel set $A \subseteq X$
$$\begin{aligned} \boldsymbol{\eta}_\lambda\big((\mathrm{pr}^{(1)})^{-1}(A)\big) &= \lambda \boldsymbol{\eta_1}\big((\mathrm{pr}^{(1)})^{-1}(A)\big) + (1-\lambda)\boldsymbol{\eta_2}\big((\mathrm{pr}^{(1)})^{-1}(A)\big) \\ &= \lambda \mu(A) + (1-\lambda)\mu(A) = \mu(A), \end{aligned}$$
and so $\mathrm{pr}^{(1)} \sharp \boldsymbol{\eta}_\lambda = \mu$.
2. Given a Borel set $A \subseteq X$, $\mu \in \mathscr{P}(X)$, $\boldsymbol{\eta} \in \mathscr{P}(S)(\mu)$, we have
$$\begin{aligned} \boldsymbol{\eta}\Big((X \times Y) \backslash (\operatorname{graph}(S_{|A})\Big) &= \boldsymbol{\eta}\Big[\Big((X \times Y) \setminus \operatorname{graph} S\Big) \cup \Big((X \times Y) \setminus (A \times Y)\Big)\Big] \\ &= \boldsymbol{\eta}\Big((X \times Y) \setminus (A \times Y))\Big) = \boldsymbol{\eta}\Big((X \setminus A) \times Y\Big) = \mu(X \setminus A), \end{aligned}$$
recalling that $\operatorname{supp} \boldsymbol{\eta} \subseteq \operatorname{graph} S$ and that $\mathrm{pr}^{(1)} \sharp \boldsymbol{\eta} = \mu$.

Lemma 2 (Closure of the graph of the lift). *S has closed graph if and only if $\mathscr{P}(S)$ has closed graph.*

Proof. Suppose that S has closed graph. Indeed, let $\{(\mu_n, \boldsymbol{\eta}_n)\}_{n\in\mathbb{N}} \subseteq \operatorname{graph} \mathscr{P}(S)$ be a sequence converging to $(\mu, \boldsymbol{\eta}) \in \mathscr{P}(X) \times \mathscr{P}(X \times Y)$. Since $\mathrm{pr}^{(1)}$ is continuous, we have that $\{\mathrm{pr}^{(1)}\sharp\boldsymbol{\eta}_n\}_{n\in\mathbb{N}}$ narrowly converges to $\mathrm{pr}^{(1)}\sharp\boldsymbol{\eta}$, and therefore, since $\mathrm{pr}^{(1)}\sharp\boldsymbol{\eta}_n = \mu_n$, by passing to the limit we get $\mathrm{pr}^{(1)}\sharp\boldsymbol{\eta} = \mu$.

On the other hand, let $(x, y) \in \operatorname{supp} \boldsymbol{\eta}$. Then by [1, Proposition 5.1.8], there is a sequence $\{(x_n, y_n)\}_{n\in\mathbb{N}}$ such that $(x_n, y_n) \in \operatorname{supp} \boldsymbol{\eta}_n$ and $(x_n, y_n) \to (x, y)$ in Y. By assumption, $(x_n, y_n) \in \operatorname{graph} S$, and, since $\operatorname{graph} S$ is closed, we have $(x, y) \in \operatorname{graph} S$. Thus $\operatorname{supp} \boldsymbol{\eta} \subseteq \operatorname{graph} S$. We conclude that $\boldsymbol{\eta} \in \mathscr{P}(S)(\mu)$. Suppose now that $\mathscr{P}(S)$ has closed graph. Let $\{(x_n, y_n)\}_{n\in\mathbb{N}}$ be a sequence in $\operatorname{graph}(S)$ converging to (x, y). In particular, we have that $\{(\delta_{x_n}, \delta_{x_n} \otimes \delta_{y_n})\}_{n\in\mathbb{N}}$ is a sequence in $\operatorname{graph}(\mathscr{P}(S))$ converging to $(\delta_x, \delta_x \otimes \delta_y)$, which therefore belongs to $\operatorname{graph}(\mathscr{P}(S))$ by assumption. Thus $(x, y) \in \operatorname{graph}(S)$.

Lemma 3 (Narrow compactness). *Suppose that for every compact $K \subseteq X$ the set* $\operatorname{graph}\left(S_{|K}\right) := \operatorname{graph} S \cap (K \times Y)$ *is compact in $X \times Y$. Then, endowed $\mathscr{P}(X)$ and $\mathscr{P}(X \times Y)$ with the narrow topology, we have that for every relative compact $\mathscr{K} \subseteq \mathscr{P}(X)$, the set $\mathscr{P}(S)(\mathscr{K}) := \bigcup_{\mu\in\mathscr{K}} \mathscr{P}(S)(\mu)$ is relatively compact. Furthermore, if S has closed graph, then $\mathscr{P}(S)$ has compact images.*

Proof. The first part of the statement is a direct consequence of our assumption, Theorem 5.1.3 in [1] and Remark 2. To prove the second part, we take $\mathscr{K} = \{\mu\}$ obtaining that $\mathscr{P}(S)(\mu)$ is relatively compact. Since S has closed graph, by Lemma 2 we have that $\mathscr{P}(S)$ has closed graph, in particular it has closed images. Therefore $\mathscr{P}(S)(\mu)$ is compact.

Lemma 4 (Uniform integrability). *Suppose that S has at most* linear growth, *i.e., there exists $C, D > 0$ and $(\bar{x}, \bar{y}) \in X \times Y$ such that $d_Y(y, \bar{y}) \leq C d_X(x, \bar{x}) + D$ for all $(x, y) \in \operatorname{graph} S$. Then if $\mathscr{K} \subseteq \mathscr{P}(X \times Y)$ is a set with uniformly integrable 2-moments, we have that $\mathscr{P}(S)(\mathscr{K})$ has uniformly integrable 2-moments.*

Proof. By the linear growth, we have in particular that for every bounded $K \subseteq X$ the set $\operatorname{graph}\left(S_{|K}\right) := \operatorname{graph} S \cap (K \times Y)$ is bounded in $X \times Y$: indeed, let $R > 0$ such that $K \subseteq B(\bar{x}, R)$. Then by the linear growth we have

$$d^2_{X\times Y}((x, y), (\bar{x}, \bar{y})) = d^2_X(x, \bar{x}) + d^2_Y(y, \bar{y}) \leq R^2 + (CR + D)^2,$$

for all $(x, y) \in \operatorname{graph} S \cap (K \times Y)$. Thus $\operatorname{graph}(S_{|K}) \subseteq B((\bar{x}, \bar{y}), \sqrt{R^2 + (CR + D)^2})$.

Fix $\varepsilon > 0$. Since $\mathscr{K}$ has uniformly integrable 2-moments, we have that there exists $r_\varepsilon > 1$ such that for all $\mu \in \mathscr{K}$ it holds $\displaystyle\int_{X\setminus B(\bar{x}, r_\varepsilon)} d^2_X(x, \bar{x})\, d\mu(x) \leq \varepsilon$. In particular, for all $\mu \in \mathscr{K}$ it holds $\displaystyle\int_{X\setminus B(\bar{x}, r_\varepsilon)} \left[d^2_X(x, \bar{x}) + (C d_X(x, \bar{x}) + D)^2\right] d\mu(x)$ $\leq (1 + C + D)\varepsilon$, since on $X \setminus B(\bar{x}, r_\varepsilon)$ we have $d_X(x, \bar{x}) \geq 1$ by the choice of r_ε.

By assumption, the set graph $\left(S_{|B(\bar{x},r_\varepsilon)}\right)$ is bounded, and so there exists $k_\varepsilon > 0$ such that $B_{X\times Y}((\bar{x},\bar{y}),k) \supseteq \operatorname{graph}\left(S_{|B(\bar{x},r_\varepsilon)}\right)$ for all $k \geq k_\varepsilon$. Recalling that all $\boldsymbol{\eta} \in \mathscr{P}(S)(\mathscr{K})$ are supported on graph S, we have

$$\begin{aligned}
\int_{(X\times Y)\setminus B_{X\times Y}((\bar{x},\bar{y}),k)} & d^2_{X\times Y}((x,y),(\bar{x},\bar{y}))\, d\boldsymbol{\eta}(x,y) = \\
&= \int_{\operatorname{graph} S\setminus \operatorname{graph}\left(S_{|B(\bar{x},r_\varepsilon)}\right)} d^2_X(x,\bar{x}) + d^2_Y(y,\bar{y})\, d\boldsymbol{\eta}(x,y) \\
&\leq \int_{[X\setminus B(\bar{x},r_\varepsilon)]\times Y} \left[d^2_X(x,\bar{x}) + (C d_X(x,\bar{x}) + D)^2\right] d\boldsymbol{\eta}(x,y) \\
&= \int_{X\setminus B(\bar{x},r_\varepsilon)} \left[d^2_X(x,\bar{x}) + (C d_X(x,\bar{x}) + D)^2\right] d\mu(x) \leq (1 + C + D)\varepsilon.
\end{aligned}$$

Therefore $\mathscr{P}(S)(\mathscr{K})$ has uniformly integrable 2-moments.

Corollary 1. *Suppose that for every compact $K \subseteq X$ the set* $\operatorname{graph}\left(S_{|K}\right) := \operatorname{graph} S \cap (K \times Y)$ *is compact in $X \times Y$ and that S has at most* linear growth, *i.e., there exists $C, D > 0$ and $(\bar{x},\bar{y}) \in X\times Y$ such that $d_Y(y,\bar{y}) \leq C d_X(x,\bar{x}) + D$ for all $(x,y) \in \operatorname{graph} S$. Then the following holds true.*

1. *For every relatively compact $\mathscr{K} \subseteq \mathscr{P}_2(X)$, the set $\mathscr{P}(S)(\mathscr{K})$ is relatively compact in $\mathscr{P}_2(X\times Y)$. In particular, the restriction $\mathscr{P}(S)_{|\mathscr{P}_2(X)}$ of $\mathscr{P}(S)$ to $\mathscr{P}_2(X)$ takes values in $\mathscr{P}_2(X\times Y)$.*
2. *If furthermore S has closed graph, then $\mathscr{P}(S)_{|\mathscr{P}_2(X)}$ has closed graph in $\mathscr{P}_2(X\times Y)$, and so compact images in $\mathscr{P}_2(X\times Y)$.*

Proof. We prove (1). By Proposition 7.1.5 in [1], $\mathscr{K}$ has uniformly integrable 2-moments and it is tight. According to Lemma 3 and Lemma 4, we have that $\mathscr{P}(S)(\mathscr{K})$ has uniformly integrable 2-moments and it is tight. Again by Proposition 7.1.5 in [1], we conclude that $\mathscr{P}(S)(\mathscr{K})$ is relatively compact in $\mathscr{P}_2(X\times Y)$.

To prove (2), notice that W_2-convergence implies narrow convergence. Thus every W_2-converging sequence $\{(\mu_n,\boldsymbol{\eta}_n)\}_{n\in\mathbb{N}} \subseteq \operatorname{graph}\mathscr{P}(S) \cap \mathscr{P}_2(X\times Y)$ is narrowly converging to the same limit, say $(\mu,\boldsymbol{\eta})$. Since $\mathscr{P}(S)$ has closed graph in $\mathscr{P}(X\times Y)$, we have that the limit $(\mu,\boldsymbol{\eta})$ belongs to graph $\mathscr{P}(S)$, and since $\mu \in \mathscr{P}_2(X)$, by item (1) we have that $(\mu,\boldsymbol{\eta}) \in \operatorname{graph}\mathscr{P}(S)_{|\mathscr{P}_2(X\times Y)}$.

Theorem 1 (Lipschitz continuity). *Suppose that S is Lipschitz continuous with compact images and for every compact $K \subseteq X$ the set* $\operatorname{graph}\left(S_{|K}\right) := \operatorname{graph} S \cap (K\times Y)$ *is compact in $X\times Y$. Then*

1. *$\mathscr{P}(S)(\mathscr{P}_2(X)) \subseteq \mathscr{P}_2(X\times Y)$;*
2. *for any $\mathscr{K}$ relatively compact in $\mathscr{P}_2(X)$ we have that $\mathscr{P}(S)(\mathscr{K})$ is relatively compact in $\mathscr{P}_2(X\times Y)$. In particular, the restriction of $\mathscr{P}(S)$ to $\mathscr{P}_2(X)$ has compact images in $\mathscr{P}_2(X\times Y)$;*
3. *$\mathscr{P}(S)_{|\mathscr{P}_2(X)} : \mathscr{P}_2(X) \rightrightarrows \mathscr{P}_2(X\times Y)$ is Lipschitz continuous with* $\operatorname{Lip}\mathscr{P}(S) \leq \sqrt{1 + (\operatorname{Lip} S)^2}$.

Proof. According to the previous results, to prove (1–2) it is sufficient to show that S has at most linear growth. Fix $\bar{x} \in X$ and $\bar{y} \in S(\bar{x})$. Let $y \in S(x)$, and $y' \in S(\bar{x})$ be such that $d_{S(\bar{x})}(y) = d_Y(y, y')$. The existence of such y' follows from the compactness of $S(\bar{x})$. By Lipschitz continuity, we have

$$\begin{aligned} d_Y(y, \bar{y}) \leq & d_Y(y, y') + d_Y(y', \bar{y}) = d_{S(\bar{x})}(y) + \sup_{y_1, y_2 \in S(\bar{x})} d_Y(y_1, y_2) \\ \leq & \operatorname{Lip} S \cdot d_X(x, \bar{x}) + \sup_{y_1, y_2 \in S(\bar{x})} d_Y(y_1, y_2). \end{aligned}$$

The compactness of $S(\bar{x})$ yields the boundedness of $D := \sup_{y_1, y_2 \in S(\bar{x})} d_Y(y_1, y_2)$, the linear growth follows by taking $C = \operatorname{Lip} S$.

We prove now (3). According to Lemma 1, there exists a Borel map $g : X \times Y \to Y$ such that $d_Y(y, g(x, y)) = d_{S(x)}(y)$ and $g(x, y) \in S(x)$. Let $\mu^{(1)}, \mu^{(2)} \in \mathscr{P}_2(X)$, $\boldsymbol{\pi} \in \Pi_o^2(\mu^{(1)}, \mu^{(2)})$. Take $\boldsymbol{\eta}^{\mathbf{(1)}} \in \mathscr{P}(S)(\mu^{(1)})$ and disintegrate it w.r.t. $\mathrm{pr}^{(1)}$, i.e., $\boldsymbol{\eta}^{\mathbf{(1)}} = \mu^{(1)} \otimes \eta_x$, where $\{\eta_x\}_{x \in X}$ is family of Borel probability measures on Y, uniquely defined for $\mu^{(1)}$-a.e. $x \in X$. Define the Borel map $T : X \times Y \times X \to (X \times Y) \times (X \times Y)$ by $T(x_1, x_2, y_1) = ((x_1, y_1), (x_2, g(x_2, y_1)))$, and the measure $\theta \in \mathscr{P}(X \times X \times Y)$ by

$$\int_{X \times X \times Y} \varphi(x_1, x_2, y_1)\, d\theta(x_1, x_2, y_1) = \int_{X \times X} \int_Y \varphi(x_1, x_2, y_1)\, d\eta_{x_1}(y_1)\, d\boldsymbol{\pi}(x_1, x_2).$$

The measure $T \sharp \theta$ belongs to $\mathscr{P}((X \times Y) \times (X \times Y))$. Defined $\mathrm{pr}^{(i)}_{X \times Y} : (X \times Y) \times (X \times Y) \to X \times Y$ as $\mathrm{pr}^{(i)}_{X \times Y}((x_1, y_2), (x_2, y_2)) = (x_i, y_i)$ for $i = 1, 2$, we obtain $\mathrm{pr}^{(1)}_{X \times Y} \sharp (T \sharp \theta) = \boldsymbol{\eta}^{\mathbf{(1)}}$.

Define $\boldsymbol{\eta}^{\mathbf{(2)}} := \mathrm{pr}^{(2)}_{X \times Y} \sharp (T \sharp \theta)$. Notice that $\mathrm{pr}^{(1)} \sharp \boldsymbol{\eta}^{\mathbf{(2)}} = \mathrm{pr}^{(2)} \sharp \boldsymbol{\pi} = \mu^{(2)}$ by construction. We prove that $\operatorname{supp} \boldsymbol{\eta}^{\mathbf{(2)}} \subseteq \operatorname{graph} S$. Indeed, let A be an open set disjoint from $\operatorname{graph} S$. Then

$$\begin{aligned} \boldsymbol{\eta}^{\mathbf{(2)}}(A) = \int_{X \times Y} \chi_A(x, y)\, d\boldsymbol{\eta}^{(2)}(x, y) = \\ \iint_{(X \times Y)^2} \chi_A(x_2, g(x_2, y_1))\, d\theta(x_1, x_2, y_1) = 0, \end{aligned}$$

since $(x_2, g(x_2, y_1)) \in \operatorname{graph} S$ for all $y_1 \in Y$. We obtain that $\boldsymbol{\eta}^{\mathbf{(2)}} \in \mathscr{P}(S)(\mu^{(2)})$ and $T \sharp \theta \in \Pi(\boldsymbol{\eta}^{\mathbf{(1)}}, \boldsymbol{\eta}^{\mathbf{(2)}})$. Thus

$$\begin{aligned} W_2^2(\boldsymbol{\eta}^{\mathbf{(1)}}, \boldsymbol{\eta}^{\mathbf{(2)}}) &\leq \iiint_{X \times X \times Y} \left[d_X^2(x_1, x_2) + d_Y^2(y_1, g(x_2, y_1)) \right] d\eta_{x_1}(y_1)\, d\boldsymbol{\pi}(x_1, x_2) \\ &= W_2^2(\mu^{(1)}, \mu^{(2)}) + \iiint_{X \times X \times Y} d^2_{S(x_2)}(y_1)\, d\eta_{x_1}(y_1)\, d\boldsymbol{\pi}(x_1, x_2) \\ &\leq W_2^2(\mu^{(1)}, \mu^{(2)}) + \iint_{X \times X} (\operatorname{Lip} S)^2 \cdot d_X^2(x_1, x_2)\, d\boldsymbol{\pi}(x_1, x_2) \\ &= \left[1 + (\operatorname{Lip} S)^2 \right] \cdot W_2^2(\mu^{(1)}, \mu^{(2)}), \end{aligned}$$

the Lipschitz continuity estimate follows.

4 Applications and Extensions

The main application of the above result is the following one.

Definition 4 (Solution set-valued map). *Let X be a finite-dimensional real space, $F : [a,b] \times X \rightrightarrows X$ be a Lipschitz set-valued map with compact convex nonempty values, $I = [a,b] \subseteq \mathbb{R}$ be a compact interval, $\boldsymbol{\theta} = \{\theta_t\}_{t\in I} \in C^0(I;\mathscr{P}_2(X))$, and $\mu \in \mathscr{P}_2(X)$. We define the set-valued maps $S_I^F : X \rightrightarrows C^0(I;X)$, $\Xi_I^F : \mathscr{P}_2(\mu) \rightrightarrows \mathscr{P}_2(X \times C^0(I;X))$ and $\Upsilon_I^F : \mathscr{P}_2(\mu) \rightrightarrows C^0(I;\mathscr{P}_2(X))$ by*

$$
\begin{aligned}
S_I^F(x) :=& \{\zeta(\cdot) \in C^0(I;X) : \dot{\zeta}(t) \in F(t,\zeta(t)) \text{ for a.e. } t \in I,\, \zeta(a) = x\},\\
\Xi_I^F(\mu) :=& \{\boldsymbol{\eta} \in \mathscr{P}(X \times C^0(I;X)) : \operatorname{supp}\boldsymbol{\eta} \subseteq \operatorname{graph} S_I^F,\, e_a\sharp\boldsymbol{\eta} = \mu\},\\
\Upsilon_I^F(\mu) :=& \{e_I\sharp\boldsymbol{\eta} : \boldsymbol{\eta} \in \Xi_I^F(\mu)\}.
\end{aligned}
$$

Proposition 1. *Let X be a finite-dimensional space, $I = [a,b]$ a compact interval of $\mathbb{R}$, $F : I \times \mathscr{P}_2(X) \times X \rightrightarrows X$ be a Lipschitz set-valued map with nonempty compact and convex values. Given a Lipschitz continuous curve $\boldsymbol{\theta} = \{\theta_t\}_{t\in I} \subseteq \mathscr{P}_2(X)$, we have that the set-valued map $\Xi_I^F : \mathscr{P}_2(X) \rightrightarrows \mathscr{P}_2(X \times C^0(I;X))$ (defined as in Definition 4) enjoys the following properties:*

1. *Ξ_I^F has nonempty compact convex images;*
2. *$\operatorname{Lip}\Xi_I^F \le \sqrt{1 + e^{2(b-a)\operatorname{Lip} F\cdot(1+b-a)}}$;*
3. *for every relatively compact $\mathscr{K} \subseteq \mathscr{P}_2(X)$ we have that $\Xi_I^F(\mathscr{K})$ is relatively compact in $\mathscr{P}_2(X \times Y)$.*

Proof. Standard result in differential inclusion theory (see e.g. from Theorem 1 and Corollary 1 in Sect. 2, Chap. 2 of [2], and Filippov's Theorem, see e.g. Theorem 10.4.1 in [3]) yields all the properties needed on $S_I^F(\cdot)$ to have that its random lift Ξ_I^F enjoys the requested properties.

The notion introduced in the previous section allows to transfer informations from a Lipschitz set-valued map between complete metric separable spaces to its natural lift in the space of probability measures. It is possible, in this setting, also to add a Lipschitz dependence of F on the current state of the system. In this case the existence of trajectories in the space of probability measures follows from a straightforward application of Banach contraction principle to the set-valued map Υ_I^F.

References

1. Ambrosio, L., Gigli, N., Savare, G.: Gradient flows in metric spaces and in the space of probability measures, 2nd edn. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2008)
2. Aubin, J.-P., Cellina, A.: Differential inclusions, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 264. Springer, Cham (1984)

3. Aubin, J.-P., Frankowska, H.: Set-Valued Analysis, Modern Birkhäuser Classics, Reprint of the 1990 edition [MR1048347]. Birkhäuser Boston Inc., Boston (2009)
4. Cannarsa, P., Capuani, R.: Existence and uniqueness for mean field games with state constraints. In: Cardaliaguet, P., Porretta, A., Salvarani, F. (eds.) PDE Models for Multi-Agent Phenomena. SIS, vol. 28, pp. 49–71. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01947-1_3
5. Cannarsa, P., Capuani, R., Cardaliaguet, P.: $\mathbf{C^{1,1}}$-smoothness of constrained solutions in the calculus of variations with application to mean field games. Math. Eng. **1**(1), 174–203 (2018). https://doi.org/10.3934/Mine.2018.1.174
6. Cannarsa, P., Capuani, R., Cardaliaguet, P.: Mean field games with state constraints: from mild to pointwise solutions of the PDE system. Calculus Variat. Partial Different. Equat. **60**(3), 1–33 (2021). https://doi.org/10.1007/s00526-021-01936-4
7. Lirkov, I., Margenov, S. (eds.): LSSC 2017. LNCS, vol. 10665. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-73441-5
8. Himmelberg, C.J.: Measurable relations. Fund. Math. **87**(1), 53–72 (1975)