



Filippov's Theorem for Mutational Inclusions in a Metric Space *

Hélène Frankowska, Thomas Lorenz

► To cite this version:

Hélène Frankowska, Thomas Lorenz. Filippov's Theorem for Mutational Inclusions in a Metric Space
*. 2022. hal-03680046

HAL Id: hal-03680046

<https://hal.science/hal-03680046>

Preprint submitted on 27 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Filippov's Theorem for Mutational Inclusions in a Metric Space*

Hélène Frankowska and Thomas Lorenz

Abstract

This article is devoted to an extension of the celebrated Filippov theorem to the metric space setting. We deal with fairly general metric spaces, where derivatives of time-dependent functions are replaced by mutations and solutions of differential equations/inclusions are mutational primitives of (time-dependent) maps of transitions. As an example of application we discuss measure-valued solutions to a controlled transport equation and state the Filippov theorem in this context. We also show that whenever a transport equation is generated by Lipschitz vector fields its classical weak solutions coincide with its mutational solutions. Our abstract setting applies as well to systems on the space of nonempty compact subsets of \mathbb{R}^n endowed with the Pompeiu-Hausdorff distance and to continuity equations/inclusions on Wasserstein spaces of Borel probability measures.

Keywords. Filippov theorem, metric space, transport equation, mutational analysis.

AMS subject classification. 54E40, 58D25, 35R06, 34K09.

1 Introduction

The use of (topological) vector spaces has been and still is predominant in models arising in applied sciences. Modeling in such mathematical setting, however, usually requires that all states are endowed with units of measures and thus quantified. This has proved to be well suited for physical environments, of course, and has been extended to economics by introducing quantified goods and commodities.

However, not all relevant entities can be quantified straightforwardly. In social dynamics, for instance, the interactions between agents are strongly influenced by mutual feelings, individual opinions and the past experiences. This renders them particularly imprecise and calls for taking means of states of all the agents, like in the mean-field games. In the recent years, social sciences related models are often formulated on the so-called Wasserstein (metric) spaces of probability measures. States in such models are the occupation probability of agents, that is positive reals from the interval $[0, 1]$ associated to subsets of a given space and obeying measure theoretic laws.

On the other hand, deterministic approach to uncertainties uses sets instead of singletons to describe possible outcomes of a system (with uncertain initial conditions or uncertain parameters). Nevertheless, nonempty subsets (of a fixed basic set/space) do not have an obvious linear structure – unless considerations are restricted to some very special classes, like convex subsets of a vector space. The need to deal with dynamic systems on metric spaces, so that closed or even compact subsets can be handled, resulted in developments of the so called morphological analysis on the space of nonempty compact subsets of \mathbb{R}^n endowed with the Pompeiu-Hausdorff distance, see [8, 12, 41].

Our goal is to develop analysis of dynamical systems with states in a metric space (E, d) . For this aim we apply the mutational concept introduced by Aubin [6–8] as a counterpart to differential equations. It is based on specifying a class Θ of some semi-dynamical systems $\vartheta : [0, 1] \times E \longrightarrow E$ (called *transitions*) which are used to define first-order approximations of a curve $x : [0, T] \longrightarrow E$ at time $t \in [0, T)$ just in terms of the metric d on E :

$$\lim_{h \downarrow 0} \frac{1}{h} \cdot d(\vartheta(h, x(t)), x(t+h)) = 0 \quad (1)$$

*This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-18-1-0254.

(more details are presented in Section 2.1 below). This mutational approach has already proved to be useful in a broad class of examples, like random closed sets, structured population models and traffic flow models (see, e.g., [18, 19, 24, 27, 33, 34, 42, 43, 45–47, 52] and references therein).

Motivated by Carathéodory solutions to ordinary differential equations, given $f : [0, T] \times E \longrightarrow \Theta$, a curve $x : [0, T] \longrightarrow E$ is considered a solution to the mutational equation with f on the right-hand side, if $x(\cdot)$ is absolutely continuous and if at a.e. $t \in [0, T]$, the transition $\vartheta := f(t, x(t)) \in \Theta$ satisfies the infinitesimal condition (1). This notion extends to differential inclusions on metric spaces.

Let us underline that the notion of metric gradient introduced in [3, Chapter 1] for absolutely continuous mappings $x : [0, T] \longrightarrow E$ differs from transitions. For instance, in the case of a reflexive Banach space $(E, \|\cdot\|)$, the metric gradient at a point t of differentiability of $x(\cdot)$ is equal to $\|x'(t)\|$, while the associated transition is $\vartheta(h, z) = z + hx'(t)$ for $(h, z) \in [0, 1] \times E$. In [3, Part I] curves of maximal slope are investigated with respect to the upper gradient of a given extended real-valued functional on (E, d) . In contrast, transitions induce curves without any optimisation features behind. In short, the objective of mutational analysis is to extend theory of ordinary differential equations to metric spaces. In particular, example provided in Section 6 implies that for the 1-Wasserstein space any weak solution to the continuity equation generated by a bounded Lipschitz vector field is a mutational solution. Hence our main result derived in a general metric setting applies to controlled continuity equations on this Wasserstein space with generating Lipschitz vector fields that are not necessarily gradients. Also any Lipschitz, locally bounded set-valued map $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ having nonempty convex compact values defines a transition on the metric space $\mathcal{K}(\mathbb{R}^n)$ of nonempty compact subsets of \mathbb{R}^n endowed with the Pompeiu-Hausdorff distance, see [42]. These transitions are described by reachable sets from initial conditions that are elements of $\mathcal{K}(\mathbb{R}^n)$. Hence in this case the flow of sets is generated by the set-valued map F , while in [3], roughly speaking, dynamics of systems on metric spaces are induced by “gradients” of some given functions.

In the theory of differential inclusions, the celebrated theorem of Filippov combines the existence of solutions with a priori estimates on the distance of a given curve from the set of all the solutions to a differential equation/inclusion. It provides a richer information than the Cauchy-Lipschitz existence theorem even in the case of ordinary differential equations and has many applications ranging from relaxation, invariance, stability, to the uniqueness of solutions to Hamilton-Jacobi equations. To the best of our knowledge, the only Filippov-like result in a mutational setting so far was presented by Doyen in [18, § 7] and concerns the special case of autonomous morphological inclusions (i.e., systems on the space of nonempty compact subsets of \mathbb{R}^n). In the present paper we extend it to a much broader framework of non-autonomous mutational inclusions in general metric spaces.

Let us first recall a finite dimensional version of Filippov’s theorem (see, e.g., [9, 21, 54]). Under suitable assumptions about the set-valued map $F : [0, T] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ which is $k(t)$ –Lipschitz in the second variable with $k \in L^1([0, T])$, the following statement holds :

Let an absolutely continuous curve $y : [0, T] \longrightarrow \mathbb{R}^n$ be given. Then for each $x_0 \in \mathbb{R}^n$, there exists a solution $x : [0, T] \longrightarrow \mathbb{R}^n$ to the differential inclusion $x' \in F(t, x)$ such that $x(0) = x_0$ and

- $|x(t) - y(t)| \leq \eta(t) := |x_0 - y(0)| e^{\|k\|_{L^1([0, t])}} + \int_0^t e^{\|k\|_{L^1([s, t])}} \cdot \text{dist}(y'(s), F(s, y(s))) ds, \forall t \in [0, T],$
- $|x'(t) - y'(t)| \leq k(t) \cdot \eta(t) + \text{dist}(y'(t), F(t, y(t)))$ for a.e. $t \in [0, T]$.

The above inequalities mean that the distance between $x(\cdot)$ and $y(\cdot)$ can be estimated from the above by the distance between their initial conditions and the violation by $y(\cdot)$ of the prescribed dynamic.

The main goal of this article is to get the counterpart of the Filippov theorem for mutational inclusions in a metric space (E, d) supplied with a class Θ of transitions. The set-valued map $F : [0, T] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ is replaced here by a set-valued map $\mathfrak{F} : [0, T] \times E \rightsquigarrow \Theta$. Then, given an absolutely continuous curve $y : [0, T] \longrightarrow E$, we look for an absolutely continuous curve $x : [0, T] \longrightarrow E$ such that firstly, at a.e. $t \in [0, T]$, a transition $\vartheta \in \mathfrak{F}(t, x(t)) \subset \Theta$ satisfies condition (1) and secondly, similar estimates w.r.t. y are fulfilled, see Theorem 3.4.

To illustrate our result, we consider a transport inclusion on the metric space of all finite real-valued Radon measures on \mathbb{R}^n , that are present, for instance, in models of traffic flow and cancer cell migration. Various hyperbolic models using time-dependent probability or Radon measures have already been investigated in the literature by Carrillo, Colombo, Goatin, Piccoli, Rossi and collaborators, see for instance [1, 14–16, 25, 26, 48–51] and the references therein. In comparison with many former results, we extend the setting in mainly two aspects:

- The transport equation is considered with a (time-dependent) control $u(t) \in U$ and so, becomes a quasilinear first-order partial differential inclusion $\partial_t \mu \in \bigcup_{u \in U} (-\operatorname{div}_{\mathbf{x}}(f(t, \mu, u) \mu) + g(t, \mu, u) \cdot \mu)$.
- The values of the solutions are (possibly signed) Radon measures on \mathbb{R}^n . This opens the door to modeling the evolution of the “total mass” quantitatively. Whenever $g \equiv 0$, the above inclusion is usually called a continuity inclusion, and probability measures (as initial states) stay probability measures for all times.

The outline of this paper is as follows. Section 2 specifies the mutational setting and discusses notions of transitions, mutational derivatives, primitives and their key properties. The corresponding proofs can be found in Section 4. In Section 3 we state the main result in Theorem 3.4, postponing its proof to Section 5. Section 6 is devoted to an example in the metric space of Radon measures on \mathbb{R}^n , where our main result is illustrated for trajectories of a transport inclusion by establishing a link between weak solutions to transport equations and the developed abstract mutational framework. Finally, the appendix consists of some new analytical tools used in this article such as variants of Scorza-Dragoni-like theorems about almost lower semicontinuous set-valued maps and about almost closed graphs.

2 Preliminaries

Below we always consider the Lebesgue measure on $[0, T]$ and the Borel measure on (pseudo-) metric spaces. Consistently, given a metric space E and a pseudo-metric space Θ , “measurability” of a function $[0, T] \times E \rightarrow \Theta$ is understood in the sense of Lebesgue-Borel measurability (unless stated otherwise). We denote by $L^1([0, T])$ the set of all Lebesgue integrable functions from $[0, T]$ into \mathbb{R}_+ .

2.1 Mutational Setting in a Metric Space

Consider a metric space (E, d) where $d : E \times E \rightarrow \mathbb{R}$ denotes a metric and a lower semicontinuous “magnitude” function $\|\cdot\| : E \rightarrow [0, \infty)$. For a nonempty subset $A \subset E$, we denote $\|A\|_\infty := \sup_{x \in A} \|x\|$ and $\|\emptyset\|_\infty := \infty$. A subset $A \subset E$ is called *bounded* w.r.t. $\|\cdot\|$ if $\|A\|_\infty < \infty$. For any $r \geq 0$ define the sublevel set $E_r := \{x \in E \mid \|x\| \leq r\}$.

In the next definition that extends an earlier one by Aubin [8], we consider a mapping $(t, x) \mapsto \vartheta(t, x)$ referring to t as the time variable and to x as the state variable.

Definition 2.1 (Transition) *A function $\vartheta : [0, 1] \times E \rightarrow E$ is called a transition on $(E, d, \|\cdot\|)$ if :*

- (i) *for every $x \in E$: $\vartheta(0, x) = x$,*
- (ii) *ϑ has the semigroup property: $\vartheta(t, \vartheta(s, x)) = \vartheta(s+t, x)$ for all $x \in E$ and $s, t \in [0, 1]$ with $s+t \leq 1$.*
- (iii) *for every $r \geq 0$, there exists $\alpha_r(\vartheta) \in [0, \infty)$ such that for any $x, y \in E_r$,*

$$\limsup_{h \downarrow 0} \frac{d(\vartheta(h, x), \vartheta(h, y)) - d(x, y)}{h} \leq \alpha_r(\vartheta) \cdot d(x, y),$$

- (iv) *for every $r \geq 0$, there is $\beta_r(\vartheta) \geq 0$ such that for any $x \in E_r$, $\vartheta(\cdot, x) : [0, 1] \rightarrow (E, d)$ is $\beta_r(\vartheta)$ -Lipschitz.*

(v) at the initial time ϑ has sublinear growth w.r.t. time and $\|\cdot\|$:

$$\liminf_{h \downarrow 0} \frac{\|\vartheta(h, x)\| - \|x\|}{h} \leq \gamma(\vartheta) \cdot (1 + \|x\|) \quad \forall x \in E.$$

Conditions (i) – (iv) of Definition 2.1 are typical properties of a quasi-contractive semigroup (often considered for operators on a Banach space). Originally, Aubin suggested an infinitesimal criterion instead of semigroup property (ii). They prove to be equivalent to each other (as verified in section 4 below).

Proposition 2.2 (Infinitesimal characterization of semigroup property) *Let $\vartheta : [0, 1] \times E \longrightarrow E$ satisfy the conditions (i), (iii), (iv) of Definition 2.1. Then ϑ has the semigroup property (ii) if and only if for every $x \in E$ and $t \in [0, 1]$:*

$$\lim_{h \downarrow 0} \frac{1}{h} \cdot d(\vartheta(t+h, x), \vartheta(h, \vartheta(t, x))) = 0.$$

The parameter $\gamma(\vartheta)$ in the above sublinear growth condition (v) lays the basis for a priori bounds of $\vartheta(\cdot, x) : [0, 1] \longrightarrow E$. They concern only the “magnitude” function $\|\cdot\|$, not the metric d on E , in general.

Proposition 2.3 (A priori bound of growth w.r.t. $\|\cdot\|$) *Let $\vartheta_1, \dots, \vartheta_K$ be finitely many transitions on $(E, d, \|\cdot\|)$ with $\Gamma := \max_{k \in \{1, \dots, K\}} \gamma(\vartheta_k) < \infty$. For any $x_0 \in E$ and partition $0 = t_0 < t_1 < \dots < t_K = T$ with $\sup_k (t_k - t_{k-1}) \leq 1$, define the curve $x(\cdot) : [0, T] \longrightarrow E$ by*

$$x(0) := x_0, \quad x(t) := \vartheta_k(t - t_{k-1}, x(t_{k-1})) \quad \text{for } t \in (t_{k-1}, t_k], \quad k \in \{1, \dots, K\}.$$

Then, $\|x(t)\| \leq (\|x_0\| + \Gamma \cdot t) \cdot e^{\Gamma \cdot t}$ for every $t \in [0, T]$.

The following proposition clarifies why each transition can be interpreted as a *quasi-contractive* semigroup. It is an immediate consequence of Proposition 2.9 below and so, we do not give a separate proof.

Proposition 2.4 (Lipschitz dependence of transitions on initial state) *Let $\vartheta : [0, 1] \times E \longrightarrow E$ be a transition and $r \geq 0$. Set $R := (r + \gamma(\vartheta)) \cdot e^{\gamma(\vartheta)}$.*

Then for all $x, y \in E$ with $\|x\| \leq r, \|y\| \leq r$ and every $h \in [0, 1]$, $d(\vartheta(h, x), \vartheta(h, y)) \leq d(x, y) \cdot e^{\alpha_R(\vartheta) h}$.

To compare the evolutions of two states along two transitions we introduce a pseudo-distance between two transitions that localizes w.r.t. $\|\cdot\|$ Aubin’s original definition given in [8].

Definition 2.5 (Pseudo-distance on transitions) *Let $\Theta(E, d, \|\cdot\|)$ or shortly Θ denote a fixed non-empty subset of transitions on $(E, d, \|\cdot\|)$. For $r \geq 0$ with $E_r \neq \emptyset$ and any $\vartheta \in \Theta$, define the restriction map $\iota_r(\vartheta) := \vartheta|_{[0, 1] \times E_r}$ and let $\Theta_r := \{\iota_r(\vartheta) \mid \vartheta \in \Theta\}$. Set for any ϑ, τ in Θ_r*

$$D_r(\vartheta, \tau) := \sup_{x \in E_r} \limsup_{h \downarrow 0} \frac{1}{h} \cdot d(\vartheta(h, x), \tau(h, x))$$

and let $D_r(\cdot, \cdot) = 0$ whenever $E_r = \emptyset$.

To simplify, we use also the notation $D_r(\vartheta, \tau) := D_r(\iota_r \vartheta, \iota_r \tau)$ for all $r \geq 0$ and $\vartheta, \tau \in \Theta$ (instead of Θ_r).

Remark 2.6 Let Θ be a subset of transitions on $(E, d, \|\cdot\|)$ and $r \geq 0$.

(a) For all $\vartheta, \tau \in \Theta_r$, the definition of Θ_r guarantees the existence of transitions $\tilde{\vartheta}, \tilde{\tau} \in \Theta$ with $\vartheta = \tilde{\vartheta}|_{[0, 1] \times E_r}$ and $\tau = \tilde{\tau}|_{[0, 1] \times E_r}$. Then the property (iv) from Definition 2.1 of $\tilde{\vartheta}, \tilde{\tau}$ implies

$$D_r(\vartheta, \tau) \leq \sup_{x \in E_r} \limsup_{h \downarrow 0} \frac{1}{h} \cdot \left(d(\vartheta(h, x), x) + d(x, \tau(h, x)) \right) \leq \beta_r(\tilde{\vartheta}) + \beta_r(\tilde{\tau}) < \infty.$$

- (b) Clearly for each $r \geq 0$, D_r is a pseudo-metric on both Θ_r and Θ . (Θ, D_r) might not be Hausdorff though because any transitions $\vartheta, \tau \in \Theta$ which coincide in $[0, 1] \times E_r$ but not in $[0, 1] \times (E \setminus E_r)$ satisfy $D_r(\vartheta, \tau) = 0$.

We consider it an advantage that our analytical tools do not require D_r to be positive definite in the general theory. In examples (as in section 6 below), the assumptions about the coefficients are often formulated for a distance which might be larger than D_r , but is simpler to work with (see also, e.g., [42, §§ 1, 2], [34, 43–45]).

- (c) In general, given $\vartheta \in \Theta$, $\iota_r(\vartheta)$ is not a transition on E_r because it may happen that $\vartheta(h, x) \notin E_r$ for some $h \in (0, 1]$ and $x \in E_r$.

From now on, we suppose that a subset Θ of transitions is fixed, i.e., de facto we deal with $(E, d, \|\cdot\|, \Theta)$.

Corollary 2.7 *Let $r \geq 0$ be such that $E_r \neq \emptyset$, $\vartheta, \tau \in \Theta$ and $\Gamma := \max\{\gamma(\vartheta), \gamma(\tau)\}$. If $D_R(\vartheta, \tau) = 0$ for some $R > (r + \Gamma)e^\Gamma$, then $\iota_r(\vartheta) = \iota_r(\tau)$. Furthermore, if $D_R(\vartheta, \tau) = 0$ for all $R > 0$, then $\vartheta = \tau$.*

Define

$$D(\vartheta, \tau) := \sum_{k=1}^{\infty} 2^{-k} \min\{1, D_k(\vartheta, \tau)\} \quad \forall \vartheta, \tau \in \Theta.$$

Clearly D metrize the topology induced by the family $(D_r(\cdot, \cdot))_{r \geq 0}$.

Proposition 2.8 *(Θ, D) is a metric space and the canonical embedding $(\Theta, D) \longrightarrow (\Theta, D_r)$, $\vartheta \longmapsto \vartheta$ is continuous for every $r \geq 0$. If (Θ, D_k) is separable for every integer $k \geq 1$, then (Θ, D) is separable. Moreover if $\sup_{\theta \in \Theta} \gamma(\theta) < \infty$ and (Θ, D_k) is complete for every integer $k \geq 1$, then (Θ, D) is complete.*

To compare the evolution of two states in E along two (possibly different) transitions we have the following result.

Proposition 2.9 *Let $\vartheta, \tau \in \Theta$ and $r \geq 0$ be given. Then for any $x, y \in E$ with $\|x\| \leq r$, $\|y\| \leq r$ and all $t_1, t_2 \in [0, 1]$, the following estimate is satisfied for every $h \in [0, 1]$ with $\max\{t_1 + h, t_2 + h\} \leq 1$*

$$d(\vartheta(t_1 + h, x), \tau(t_2 + h, y)) \leq (d(\vartheta(t_1, x), \tau(t_2, y)) + h \cdot D_R(\vartheta, \tau)) \cdot e^{\alpha_R(\vartheta) h}$$

for any $R \geq (r + \max\{\gamma(\vartheta), \gamma(\tau)\}) \cdot e^{\max\{\gamma(\vartheta), \gamma(\tau)\}}$.

2.2 Mutations of a Curve as Counterpart of Time Derivatives

The notion of first-order approximation leads to the so-called mutation of a curve – as a counterpart of its derivative w.r.t. time:

Definition 2.10 ([8]) *The mutation of a curve $x : [0, T] \longrightarrow E$ at time $t \in [0, T]$ is defined as*

$$\dot{x}(t) := \left\{ \vartheta \in \Theta \mid \lim_{h \downarrow 0} \frac{1}{h} \cdot d(\vartheta(h, x(t)), x(t+h)) = 0 \right\}.$$

The curve x is called mutable at $t \in [0, T]$ if $\dot{x}(t) \neq \emptyset$.

Remark 2.11

- (a) For every transition $\vartheta \in \Theta$ and initial state $x_0 \in E$, the curve $x(\cdot) := \vartheta(\cdot, x_0) : [0, 1] \longrightarrow E$ has ϑ in its mutation $\dot{x}(t)$ at each time $t \in [0, 1]$. This observation results from (ii) of Definition 2.1 and lays the basis for constructing solutions to mutational equations by means of Euler method.

- (b) Observe that the mutation $\dot{x}(t)$ is closed w.r.t. the pseudo-metric D_R for every $R \geq \|x(t)\|$. Indeed fix such R and consider $\vartheta_n \in \dot{x}(t)$ converging to some $\vartheta \in \Theta$. Since for any $h \in [0, 1]$, $d(\vartheta(h, x(t)), x(t+h)) \leq d(\vartheta_n(h, x(t)), x(t+h)) + d(\vartheta_n(h, x(t)), \vartheta(h, x(t)))$, dividing by $h > 0$ and taking the upper limit yields

$$\limsup_{h \downarrow 0} \frac{1}{h} \cdot d(\vartheta(h, x(t)), x(t+h)) \leq D_R(\vartheta_n, \vartheta).$$

Then taking the limit when $n \rightarrow \infty$ we deduce that $\vartheta \in \dot{x}(t)$.

Proposition 2.12 (Absolutely continuous mutable curves: Modulus) *Consider an absolutely continuous $x : [0, T] \rightarrow E$ with the following properties:*

- (i) $x(\cdot)$ is mutable almost everywhere in $[0, T]$ and $R := \sup_{t \in [0, T]} \|x(t)\| < \infty$.
- (ii) There exists $B_R \in L^1([0, T])$ such that for a.e. $t \in [0, T]$, some $\vartheta_t \in \dot{x}(t)$ satisfies $\beta_R(\vartheta_t) \leq B_R(t)$.

Then $d(x(t_1), x(t_2)) \leq \int_{t_1}^{t_2} B_R(s) ds$ for all $0 \leq t_1 < t_2 \leq T$.

The following estimate generalizes Proposition 2.9 when $t_1 = t_2$.

Proposition 2.13 *Let $x, y : [0, T] \rightarrow E$ be absolutely continuous and satisfy:*

- (i) $r := \sup_t \|y(t)\| < \infty$ and $R := \max \{ \sup_t \|x(t)\|, r \} < \infty$.
- (ii) $x(\cdot)$ and $y(\cdot)$ are mutable almost everywhere in $[0, T]$.
- (iii) For some $A_R, \Delta_r \in L^1([0, T])$ and for a.e. $t \in [0, T]$, there is a transition $\vartheta \in \dot{x}(t)$ with

$$\alpha_R(\vartheta) \leq A_R(t), \quad \inf_{\tau \in \dot{y}(t)} D_r(\vartheta, \tau) \leq \Delta_r(t).$$

Then $d(x(t), y(t)) \leq d(x(0), y(0)) \cdot e^{\int_0^t A_R(s) ds} + \int_0^t \Delta_r(s) \cdot e^{\int_s^t A_R(\sigma) d\sigma} ds$ for every $t \in [0, T]$.

Finally, to be able to use some measurable selection theorems, we need an appropriate topological structure of Θ motivated by a compromise: On the one hand, it should be convenient to verify in examples. On the other hand, it is to provide useful conclusions of convergence in (Θ, D_r) for all $r \geq 0$ – so that r can be chosen a posteriori whenever required.

From now on we *assume* the existence of a metric D_{loc} on Θ such that the convergence w.r.t. D_{loc} always implies the convergence w.r.t. D_r for every $r \geq 0$, i.e., for every sequence $(\vartheta_k)_{k \in \mathbb{N}}$ in Θ

$$\lim_{k \rightarrow \infty} D_{\text{loc}}(\vartheta_k, \vartheta) = 0 \implies \lim_{k \rightarrow \infty} D_r(\vartheta_k, \vartheta) = 0, \quad \forall r \geq 0. \quad (2)$$

It is worth mentioning that the inverse implication is not required in general, i.e., D_{loc} does not have to metrize the topology induced by the family $(D_r(\cdot, \cdot))_{r \geq 0}$. The metric D_{loc} might characterize an even finer topology on Θ instead. Proposition 2.8 provides sufficient conditions for the existence of such metric.

Remark 2.14 Aubin's original approach is motivated by the example of so-called morphological equations [6–8]: $\mathcal{K}(\mathbb{R}^n)$ denotes the set of nonempty compact subsets of \mathbb{R}^n and is supplied with the Pompeiu-Hausdorff metric \mathcal{d} . Every bounded Lipschitz set-valued map $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ with nonempty compact convex values specifies an autonomous differential inclusion $x' \in F(x)$ and, its reachable sets induce the so-called morphological transition $\vartheta_F : [0, 1] \times \mathcal{K}(\mathbb{R}^n) \rightarrow \mathcal{K}(\mathbb{R}^n)$ (e.g., [8, § 3.7.1], [42, § 1.9] for details). Supply $\mathcal{K}(\mathbb{R}^n)$ with the magnitude function $\|M\|_\infty := \sup_{x \in M} |x|$ for $M \in \mathcal{K}(\mathbb{R}^n)$. Then we can dispense with assuming the boundedness of F , i.e., every Lipschitz set-valued map $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ with nonempty compact convex values induces a transition ϑ_F on $(\mathcal{K}(\mathbb{R}^n), \mathcal{d}, \|\cdot\|_\infty)$ in the sense of Definition 2.1 (as

a special case of the example in [42, § 2.7]). In particular, Filippov's theorem for ordinary differential inclusions leads to the following parameters and estimates for any of these maps $F, G : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ and every $r \geq 0$

$$\begin{aligned} \alpha_r(\vartheta_F) &= \text{Lip } F, & \gamma(\vartheta_F) &= \max \{ \|F(0)\|_\infty, \text{Lip } F \} \\ \beta_r(\vartheta_F) &= \sup_{x \in \mathbb{B}_R} \|F(x)\|_\infty, & D_r(\vartheta_F, \vartheta_G) &\leq \lim_{\varepsilon \downarrow 0} \sup_{x \in \mathbb{B}_{r+\varepsilon}} d(F(x), G(x)) = \sup_{x \in \mathbb{B}_r} d(F(x), G(x)), \end{aligned}$$

where $\text{Lip } F$ denotes the Lipschitz constant of F and $R := (r + \gamma(\vartheta_F)) e^{\gamma(\vartheta_F)}$. Here we can choose D_{loc} as the following metric of uniform convergence on bounded sets

$$D_{\text{loc}}(\vartheta_F, \vartheta_G) := \sum_{k=1}^{\infty} 2^{-k} \cdot \min \left\{ 1, \sup_{x \in \mathbb{B}_k} d(F(x), G(x)) \right\}.$$

Then the implication (2) even holds as an equivalence.

The implication (2) can be reformulated equivalently as one of the following properties:

- the continuity of each restriction map $\iota_r : (\Theta, D_{\text{loc}}) \longrightarrow (\Theta_r, D_r)$, $\vartheta \longmapsto \vartheta|_{[0,1] \times E_r}$ ($r \geq 0$) and
- the continuity of every canonical embedding $(\Theta, D_{\text{loc}}) \longrightarrow (\Theta, D_r)$, $\vartheta \longmapsto \vartheta$ ($r \geq 0$).

In Section 6 we provide an example of (Θ, D_{loc}) as above on the space of finite Radon measures.

Let us summarize all the general hypotheses about $E, d, \|\cdot\|, \Theta$ and D_{loc} used in the sequel:

HYPOTHESIS $(\mathbb{H}\Theta)$:

- (i) (E, d) is a metric space and the magnitude $\|\cdot\| : E \longrightarrow [0, \infty)$ is given.
- (ii) For every $R \geq 0$, the set $E_R = \{x \in E \mid \|x\| \leq R\}$ is complete and separable w.r.t. d .
- (iii) Θ is a given nonempty subset of transitions on $(E, d, \|\cdot\|)$ supplied with the family of pseudo-metrics D_r ($r \geq 0$).
- (iv) There exists a function $D_{\text{loc}} : \Theta \times \Theta \longrightarrow \mathbb{R}$ such that (Θ, D_{loc}) is a complete separable metric space and for every $R \geq 0$, the restriction map $\iota_R : (\Theta, D_{\text{loc}}) \longrightarrow (\Theta_R, D_R)$ is continuous.

In particular, $(\mathbb{H}\Theta)$ (ii) implies that $\|\cdot\|$ is lower semicontinuous. As a consequence of $(\mathbb{H}\Theta)$ (iv), both (Θ, D_R) and (Θ_R, D_R) are *weakly Suslin spaces*, i.e., they are the images of complete separable metric spaces under continuous mappings, but not necessarily Hausdorff (see, e.g., [35, 55, 56]). This allows us to apply a measurable selections result by Leese [38–40], cf. Proposition B.5, and deduce its consequences B.7, C.1. They are used in subsequent proofs about mutational primitives and solutions.

Proposition 2.15 *Assume $(\mathbb{H}\Theta)$ and let $x : [0, T] \longrightarrow E$ be continuous, bounded w.r.t. $\|\cdot\|$ and mutable almost everywhere in $[0, T]$. Then, the mutation $\dot{x} : [0, T] \rightsquigarrow (\Theta, D_{\text{loc}})$ is measurable with closed values.*

Corollary 2.16 *Under the assumptions of Proposition 2.15, there exists a measurable selection θ of \dot{x} , i.e., a measurable map $\theta : [0, T] \longrightarrow (\Theta, D_{\text{loc}})$ satisfying $\theta(t) \in \dot{x}(t)$ for a.e. $t \in [0, T]$.*

Proposition 2.15 generalizes [42, Lemma 5.3] because it avoids some further assumptions. The proof is based on constructing essentially the same approximating sequence of set-valued maps, but their measurability is concluded from closedness of their graphs instead of their respective upper semicontinuity. Then Corollary 2.16 results from the measurable selection theorem, cf. [37], [10, Theorem 8.1.3]. Hence, we do not give the proof in detail here.

2.3 Mutational Primitives

This subsection concerns “integration” (in time): For a given single-valued curve $\theta : [0, T] \longrightarrow \Theta$ of transitions, we aim to find an underlying curve $x(\cdot)$ such that $\theta(\cdot)$ is a selection of its mutation $\dot{x}(\cdot)$.

Definition 2.17 ([8]) *Consider $\theta : [0, T] \longrightarrow \Theta$. A function $x : [0, T] \longrightarrow E$ is called a mutational primitive of $\theta(\cdot)$ if it is absolutely continuous, bounded w.r.t. $\|\cdot\|$ and $\theta(t) \in \dot{x}(t)$ for a.e. $t \in [0, T]$.*

The following estimate extends Proposition 2.13 to mutational primitives and is the key tool for all subsequent constructions.

Proposition 2.18 (Continuous dependence of mutational primitives) *Let $r, R > 0$ and for $k = 1, 2$ let $\theta_k : [0, T] \longrightarrow (\Theta, D_r)$ be measurable, $x_k : [0, T] \longrightarrow E$ be any mutational primitive of $\theta_k(\cdot)$ satisfying*

- (i) $\sup_t \|x_2(t)\| \leq r, \quad \max \{ \sup_t \|x_1(t)\|, r \} \leq R.$
- (ii) *There is $A_R \in L^1([0, T])$ with $\alpha_R(\theta_1(t)) \leq A_R(t)$ for a.e. $t \in [0, T]$.*
- (iii) *There is $\Psi_r \in L^1([0, T])$ such that $D_r(\theta_1(t), \theta_2(t)) \leq \Psi_r(t)$ for a.e. $t \in [0, T]$.*

Then $D_r(\theta_1(\cdot), \theta_2(\cdot)) : [0, T] \longrightarrow \mathbb{R}$ is integrable and, the following inequality holds for every $t \in [0, T]$,

$$d(x_1(t), x_2(t)) \leq d(x_1(0), x_2(0)) \cdot e^{\int_0^t A_R(\sigma) d\sigma} + \int_0^t D_r(\theta_1(s), \theta_2(s)) \cdot e^{\int_s^t A_R(\sigma) d\sigma} ds.$$

Corollary 2.19 (Convergence of mutational primitives) *Let $0 \leq r < R$, $(\theta_k)_{k \in \mathbb{N}}$ be a sequence of measurable functions from $[0, T]$ into (Θ, D_r) , $(x_k)_{k \in \mathbb{N}}$ be a sequence of functions from $[0, T]$ into E and $\theta : [0, T] \longrightarrow (\Theta, D_r)$ be measurable. Suppose:*

- (i) *Each $x_k(\cdot)$ is a mutational primitive of $\theta_k(\cdot)$ and $\sup_t \|x_k(t)\| \leq r$.*
- (ii) *There exist $A_R, B_r, \Gamma \in L^1([0, T])$ such that for a.e. $t \in [0, T]$ and all $k \in \mathbb{N}$,*

$$\alpha_R(\theta_k(t)) \leq A_R(t), \quad \beta_r(\theta_k(t)) \leq B_r(t), \quad \max \{ \gamma(\theta_k(t)), \gamma(\theta(t)) \} \leq \Gamma(t).$$

- (iii) $[0, T] \ni t \longmapsto D_r(\theta(0), \theta(t))$ *is integrable.*

- (iv) $\|D_r(\theta(\cdot), \theta_k(\cdot))\|_{L^1([0, T])} \longrightarrow 0 \quad (k \longrightarrow \infty)$ *and $(x_k(0))_{k \in \mathbb{N}}$ has a limit $x_0 \in E$.*

Then $(x_k(\cdot))_{k \in \mathbb{N}}$ converges uniformly to an absolutely continuous $x : [0, T] \longrightarrow (E, d)$ which is a mutational primitive of $\theta(\cdot)$ with $x(0) = x_0$, $\sup_t \|x(t)\| \leq r$,

$$d(x(t_1), x(t_2)) \leq \int_{t_1}^{t_2} B_r(s) ds \quad (t_1, t_2 \in [0, T], t_1 \leq t_2).$$

Corollary 2.20 (A priori growth bound for mutational primitives) *Let $\theta : [0, T] \longrightarrow \Theta$ and suppose that there exists $\Gamma \in L^1([0, T])$ such that $\gamma(\theta(t)) \leq \Gamma(t)$ for a.e. $t \in [0, T]$. Assume that $r \geq 0$ and $R > (r + \|\Gamma\|_{L^1}) \cdot e^{\|\Gamma\|_{L^1}}$ have the following properties: $\theta : [0, T] \longrightarrow (\Theta, D_R)$ is measurable and for some $B_R \in L^1([0, T])$*

- (i) $\beta_R(\theta(t)) \leq B_R(t)$ *for a.e. $t \in [0, T]$.*
- (ii) *For every $\rho \geq R$, there exists $A_\rho \in L^1([0, T])$ with $\alpha_\rho(\theta(t)) \leq A_\rho(t)$ for a.e. $t \in [0, T]$.*
- (iii) *For some (and thus, every) $\vartheta_0 \in \Theta$, there is $\Psi_R \in L^1([0, T])$ with $D_R(\theta(t), \vartheta_0) \leq \Psi_R(t)$ for a.e. $t \in [0, T]$.*

Then every mutational primitive $x : [0, T] \longrightarrow E$ of $\theta(\cdot)$ with $\|x(0)\| \leq r$ satisfies for all $t \in [0, T]$,

$$\|x(t)\| \leq \left(\|x(0)\| + \int_0^t \Gamma(s) ds \right) \cdot e^{\int_0^t \Gamma(s) ds} < R.$$

The above Corollary and Proposition 2.18 yield the following uniqueness result.

Corollary 2.21 (Uniqueness of mutational primitives) *Let $x_0 \in E$, $r := \|x_0\|$ and $\theta : [0, T] \longrightarrow \Theta$, R satisfy the assumptions of Corollary 2.20. Then there exists at most one mutational primitive $x : [0, T] \longrightarrow E$ of $\theta(\cdot)$ with $x(0) = x_0$.*

Proposition 2.22 (Existence of mutational primitives) *Let $\theta : [0, T] \longrightarrow \Theta$ and suppose that there exists $\Gamma \in L^1([0, T])$ such that $\gamma(\theta(t)) \leq \Gamma(t)$ for a.e. $t \in [0, T]$. If $x_0 \in E$ and $R > (\|x_0\| + \|\Gamma\|_{L^1}) e^{\|\Gamma\|_{L^1}}$ are such that $\theta : [0, T] \longrightarrow (\Theta, D_R)$ is measurable and*

(i) *For some (and thus, every) $\vartheta_0 \in \Theta$, there is $\Psi_R \in L^1([0, T])$ with $D_R(\theta(t), \vartheta_0) \leq \Psi_R(t)$ for a.e. $t \in [0, T]$.*

(ii) *For some $A_R, B_R \in L^1([0, T])$ it holds $\alpha_R(\theta(t)) \leq A_R(t)$, $\beta_R(\theta(t)) \leq B_R(t)$ for a.e. $t \in [0, T]$,*

then there exists at least one mutational primitive $x : [0, T] \longrightarrow E$ of $\theta(\cdot)$ with $x(0) = x_0$ and $\|x(t)\| \leq (\|x_0\| + \|\Gamma\|_{L^1([0, t])}) e^{\|\Gamma\|_{L^1([0, t])}}$ for every $t \in [0, T]$. This mutational primitive is unique if in addition for every $\rho \geq R$, there is $A_\rho \in L^1([0, T])$ satisfying $\alpha_\rho(\theta(t)) \leq A_\rho(t)$ for a.e. $t \in [0, T]$.

3 Main Results

Recall that for a set-valued map $F : X \rightsquigarrow Y$ between two nonempty sets X, Y its graph is defined by

$$\text{Graph } F(\cdot) = \{(x, y) \in X \times Y \mid y \in F(x)\} \subset X \times Y.$$

As usual, $\mathcal{B}(Z)$ denotes the family of Borel subsets of a topological space Z , $\mathcal{L}([0, T])$ the family of Lebesgue measurable subsets of $[0, T]$ and \mathcal{L}^1 the Lebesgue measure on $[0, T]$. For a set-valued map $\mathfrak{F} : [0, T] \times E \rightsquigarrow \Theta$ and $R \geq 0$, define $\mathfrak{F}_R : [0, T] \times E \rightsquigarrow \Theta_R$ by

$$\mathfrak{F}_R(t, x) := \{\vartheta_{|[0, 1] \times E_R} \mid \vartheta \in \mathfrak{F}(t, x)\} \subset \Theta_R.$$

We first recall a slightly modified notion of solution to a mutational inclusion introduced in [42].

Definition 3.1 *Let a set-valued map $\mathfrak{F} : [0, T] \times E \rightsquigarrow \Theta$ be given. A curve $x : [0, T] \longrightarrow E$ is called a solution to the mutational inclusion*

$$\dot{x}(\cdot) \cap \mathfrak{F}(\cdot, x(\cdot)) \neq \emptyset$$

in $(E, d, \|\cdot\|)$ if it satisfies the following conditions:

(a) *$x(\cdot)$ is absolutely continuous and $\sup_t \|x(t)\| < \infty$.*

(b) *$\dot{x}(t) \cap \mathfrak{F}(t, x(t)) \neq \emptyset$ for a.e. $t \in [0, T]$.*

In general, sets $\dot{x}(t)$ are not singletons. The next result is a measurable selection theorem for solutions to mutational inclusions.

Proposition 3.2 *Assume $(\mathbb{H}\Theta)$ and that $\mathfrak{F} : [0, T] \times E \rightsquigarrow \Theta$ has nonempty values. Consider a continuous curve $x : [0, T] \longrightarrow (E, d)$ and suppose that $\sup_t \|x(t)\| \leq R$ for some real R . If the graph of $(\mathfrak{F}_R)_{|[0, T] \times E_R} : [0, T] \times E_R \rightsquigarrow \Theta_R$ belongs to the σ -algebra $\mathcal{L}([0, T]) \otimes \mathcal{B}(E_R, d) \otimes \mathcal{B}(\Theta_R, D_R)$, then the following statements are equivalent:*

- (a) $\dot{x}(t) \cap \mathfrak{F}(t, x(t)) \neq \emptyset$ for almost every $t \in [0, T]$.
- (b) There is a measurable $\theta : [0, T] \longrightarrow (\Theta, D_R)$ such that $\theta(t) \in \dot{x}(t) \cap \mathfrak{F}(t, x(t))$ a.e. in $[0, T]$.

Before stating our main result we recall the following definition.

Definition 3.3 Let Y be a pseudo-metric space with the pseudo-metric $D(\cdot, \cdot)$. For nonempty subsets $M_1, M_2 \subset Y$, set

$$\text{dist}(M_1, M_2) := \inf_{y \in M_1, z \in M_2} D(y, z), \quad h^\sharp(M_1, M_2) := \sup_{y \in M_1} \text{dist}(y, M_2) \in [0, \infty].$$

The Pompeiu-Hausdorff pseudo-distance between M_1 and M_2 is defined as

$$d(M_1, M_2) := \max \{h^\sharp(M_1, M_2), h^\sharp(M_2, M_1)\} \in [0, \infty].$$

A set-valued map $F : (E_R, d) \rightsquigarrow (\Theta, D_R)$ with nonempty values is called λ -Lipschitz on E_R if it satisfies for $D = D_R$:

$$d(F(x), F(y)) \leq \lambda d(x, y) \quad \forall x, y \in E_R.$$

We are ready to state our main result whose proof is postponed to Section 5.

Theorem 3.4 Assume $(\mathbf{H}\Theta)$. Consider a set-valued map $\mathfrak{F} : [0, T] \times E \rightsquigarrow \Theta$ with nonempty values, $x_0 \in E$ and an absolutely continuous curve $y : [0, T] \longrightarrow E$ which is mutable a.e. in $[0, T]$. Further suppose that for some $R \geq 0$:

- (i) $\exists \Gamma \in L^1([0, T])$ satisfying $\sup \{\gamma(\vartheta) \mid \vartheta \in \mathfrak{F}(t, z), z \in E\} \leq \Gamma(t)$ for a.e. $t \in [0, T]$.
- (ii) $R > \max \{ \sup_t \|y(t)\|, (\|x_0\| + \|\Gamma\|_{L^1}) \cdot e^{\|\Gamma\|_{L^1}} \}$.
- (iii) There exist $A_R, B_R \in L^1([0, T])$ satisfying for a.e. $t \in [0, T]$,
$$\sup \{\alpha_R(\vartheta) \mid \vartheta \in \mathfrak{F}(t, z), z \in E_R\} \leq A_R(t), \quad \sup \{\beta_R(\vartheta) \mid \vartheta \in \mathfrak{F}(t, z), z \in E_R\} \leq B_R(t).$$
- (iv) For all $t \in [0, T]$ and $x \in E_R$, $\mathfrak{F}_R(t, x)$ is complete and closed in (Θ_R, D_R) .
- (v) For every $x \in E_R$, $\mathfrak{F}_R(\cdot, x) : [0, T] \rightsquigarrow (\Theta_R, D_R)$ is measurable.
- (vi) There is $\lambda_R \in L^1([0, T])$ such that $\mathfrak{F}_R(t, \cdot)|_{E_R} : (E_R, d) \rightsquigarrow (\Theta_R, D_R)$ is $\lambda_R(t)$ -Lipschitz for a.e. t .
- (vii) $\exists \Psi_R \in L^1([0, T])$ and $\vartheta_0 \in \Theta$ satisfying for a.e. $t \in [0, T]$, $\text{dist}_R(\dot{y}(t), \mathfrak{F}(t, y(t))) < \Psi_R(t)$ and
$$\{\vartheta \in \dot{y}(t) \mid \text{dist}_R(\vartheta, \mathfrak{F}(t, y(t))) \leq \Psi_R(t), D_R(\vartheta_0, \vartheta) \leq \Psi_R(t)\} \neq \emptyset.$$

Then for each $\kappa > 1$, there exists a solution $x : [0, T] \longrightarrow E$ of the mutational inclusion

$$\dot{x}(\cdot) \cap \mathfrak{F}(\cdot, x(\cdot)) \neq \emptyset, \quad x(0) = x_0,$$

such that:

- (a) $\|x(t)\| \leq (\|x_0\| + \|\Gamma\|_{L^1([0, t])}) \cdot e^{\|\Gamma\|_{L^1([0, t])}}$ for every $t \in [0, T]$.
- (b) $d(x(t), y(t)) \leq \eta_\kappa(t) :=$

$$d(x_0, y(0)) \cdot e^{\int_0^t (A_R(\sigma) + \kappa \lambda_R(\sigma)) d\sigma} + \int_0^t \kappa \cdot \text{dist}_R(\dot{y}(s), \mathfrak{F}(s, y(s))) e^{\int_s^t (A_R(\sigma) + \kappa \lambda_R(\sigma)) d\sigma} ds.$$

- (c) A measurable selection $\theta : [0, T] \longrightarrow (\Theta, D_R)$ of $\dot{x}(\cdot) \cap \mathfrak{F}(\cdot, x(\cdot))$ (a.e.) satisfies for a.e. $t \in [0, T]$

$$\text{dist}_R(\theta(t), \dot{y}(t)) \leq \kappa \lambda_R(t) \cdot \eta_\kappa(t) + \kappa \cdot \text{dist}_R(\dot{y}(s), \mathfrak{F}(s, y(s))).$$

4 Proofs of Results in Section 2

Proof of Proposition 2.3. It is given via induction with respect to k : The claim is obvious at time $t_0 = 0$. Now assume this estimate on the interval $[0, t_{k-1}]$. The function $\llbracket x(\cdot) \rrbracket : [t_{k-1}, t_k] \rightarrow \mathbb{R}$ is lower semicontinuous. By Definition 2.1 (ii) and (v), for each $t \in [t_{k-1}, t_k]$

$$\begin{aligned} \liminf_{h \downarrow 0} \frac{\llbracket x(t+h) \rrbracket - \llbracket x(t) \rrbracket}{h} &= \liminf_{h \downarrow 0} \frac{1}{h} \cdot (\llbracket \vartheta_k(t+h-t_{k-1}, x(t_{k-1})) \rrbracket - \llbracket x(t) \rrbracket) \\ &= \liminf_{h \downarrow 0} \frac{1}{h} \cdot (\llbracket \vartheta_k(h, x(t)) \rrbracket - \llbracket x(t) \rrbracket) \leq \gamma(\vartheta_k) \cdot (1 + \llbracket x(t) \rrbracket) < \infty. \end{aligned}$$

Gronwall's inequality A.1 and the induction hypothesis imply for $t \in [t_{k-1}, t_k]$

$$\begin{aligned} \llbracket x(t) \rrbracket &\leq (\llbracket x(t_{k-1}) \rrbracket + \gamma(\vartheta_k) \cdot (t - t_{k-1})) \cdot e^{\gamma(\vartheta_k) \cdot (t - t_{k-1})} \\ &\leq ((\llbracket x_0 \rrbracket + \Gamma \cdot t_{k-1}) \cdot e^{\Gamma t_{k-1}} + \Gamma \cdot (t - t_{k-1})) \cdot e^{\Gamma \cdot (t - t_{k-1})} \\ &\leq (\llbracket x_0 \rrbracket + \Gamma \cdot t) \cdot e^{\Gamma t}. \end{aligned}$$

□

Proof of Proposition 2.2. “ \implies ” The infinitesimal criterion of ϑ is an obvious consequence of the semigroup property. “ \impliedby ” Fix $x \in E$ and $s \in [0, 1]$ arbitrarily. By Proposition 2.3, $r := \|\vartheta([s, 1], x)\|_\infty < \infty$. Define $\psi : [0, 1-s] \rightarrow [0, \infty)$ by $\psi(t) = d(\vartheta(t, \vartheta(s, x)), \vartheta(s+t, x))$. It is Lipschitz due to the condition (iv) of Definition 2.1. Furthermore, $\psi(0) = 0$. Then (iii) of Definition 2.1 implies that for every $t \in (0, 1-s)$ such that ψ is differentiable at t and for any $R > \max\{r, \|\vartheta([0, 1-s], \vartheta(s, x))\|_\infty\}$,

$$\begin{aligned} \psi'(t) &= \lim_{h \downarrow 0} \frac{1}{h} \cdot (d(\vartheta(t+h, \vartheta(s, x)), \vartheta(s+t+h, x)) - \psi(t)) \\ &\leq \limsup_{h \downarrow 0} \frac{1}{h} \cdot (d(\vartheta(t+h, \vartheta(s, x)), \vartheta(h, \vartheta(t, \vartheta(s, x)))) + \\ &\quad d(\vartheta(h, \vartheta(t, \vartheta(s, x))), \vartheta(h, \vartheta(s+t, x))) - \psi(t) + \\ &\quad d(\vartheta(h, \vartheta(s+t, x)), \vartheta(s+t+h, x))) \leq \alpha_R(\vartheta) \cdot \psi(t) \end{aligned}$$

and we conclude from Gronwall's inequality that $\psi(\cdot) = 0$ on $[0, 1-s]$. □

Proof of Corollary 2.7. By Proposition 2.3 for $R' := (r + \Gamma)e^\Gamma < R$ and any $(h, x) \in [0, 1] \times E_r$ we have $\vartheta(h, x), \tau(h, x) \in E_{R'}$. Fix $x \in E_r$ and define $\psi(h) := d(\vartheta(h, x), \tau(h, x))$ for $h \in [0, 1]$. Then ψ is Lipschitz by (iv) of Definition 2.1. Consider $t \in (0, 1)$ such that ψ is differentiable at t . Since for all small $\delta > 0$,

$$\psi(t + \delta) = d(\vartheta(\delta, \vartheta(t, x)), \tau(\delta, \tau(t, x))) \leq d(\vartheta(\delta, \vartheta(t, x)), \vartheta(\delta, \tau(t, x))) + d(\vartheta(\delta, \tau(t, x)), \tau(\delta, \tau(t, x))),$$

using that $D_R(\vartheta, \tau) = 0$, we obtain

$$\psi'(t) \leq \limsup_{\delta \downarrow 0} \frac{d(\vartheta(\delta, \vartheta(t, x)), \vartheta(\delta, \tau(t, x))) - d(\vartheta(t, x), \tau(t, x))}{\delta} \leq \alpha_R(\vartheta)\psi(t).$$

Since $\psi(0) = 0$, the Gronwall inequality yields $\psi(\cdot) = 0$. Consequently, $\vartheta(h, x) = \tau(h, x)$ for any $h \in [0, 1]$. This proves the first statement. To prove the second one, consider any $x \in E$ and set $r := \llbracket x \rrbracket$. Then, by the first claim, $\vartheta(h, x) = \tau(h, x)$ for any $h \in [0, 1]$. The arbitrariness of $x \in E$ implies separability. □

Proof of Proposition 2.8. Corollary 2.7 implies that (Θ, D) is a metric space. Clearly the canonical embedding $(\Theta, D) \rightarrow (\Theta, D_k)$ is continuous for any integer $k \geq 1$. Pick any $r \geq 0$ and observe that for any integer $k \geq r$ we have $D_k \geq D_r$. Thus the canonical embedding $(\Theta, D) \rightarrow (\Theta, D_r)$ is continuous for every $r \geq 0$.

To prove the separability statement of Proposition, for every integer $k \geq 1$ consider a dense subset $(\theta_{kj})_{j \in \mathbb{N}}$ in (Θ, D_k) . We claim that $\bigcup_{k,j} \{\theta_{kj}\}$ is dense in (Θ, D) . Indeed fix any $\theta \in \Theta$, $\varepsilon > 0$ and consider n such that $\sum_{k \geq n+1} 2^{-k} < \frac{\varepsilon}{2}$. Let j be such that $D_n(\theta, \theta_{nj}) < \frac{\varepsilon}{2}$. Since $D_k(\theta, \theta_{nj}) \leq D_n(\theta, \theta_{nj})$ for every $1 \leq k \leq n$ we deduce that $\sum_{k=1}^n 2^{-k} D_k(\theta, \theta_{nj}) < \frac{\varepsilon}{2}$. Thus $D(\theta, \theta_{nj}) \leq \varepsilon$. Arbitrariness of $\varepsilon > 0$ yields separability of (Θ, D) .

To prove the last statement define $\Gamma := \sup_{\theta \in \Theta} \gamma(\theta)$ and consider a Cauchy sequence $(\theta_j)_{j \in \mathbb{N}}$ in (Θ, D) . By the definition of $D(\cdot, \cdot)$ we know that $(\theta_j)_{j \in \mathbb{N}}$ is Cauchy in (Θ, D_k) for any k and, by completeness, it converges to some limit in (Θ, D_k) , which may be not unique. Let k^0 be the smallest integer such that $E_{k^0} \neq \emptyset$ and consider any $R^0 > (k^0 + \Gamma)e^\Gamma$. Since for any two limits ϑ^0 and τ of $(\theta_j)_{j \in \mathbb{N}}$ in (Θ, D_{R^0}) we have $D_{R^0}(\vartheta^0, \tau) = 0$, Corollary 2.7 yields $\iota_{k^0}(\vartheta^0) = \iota_{k^0}(\tau)$. We construct $\theta \in \Theta$ using an induction argument.

Set $\theta(h, x) = \vartheta^0(h, x)$ for every $(h, x) \in [0, 1] \times E_{k^0}$ and consider any $R^1 > \max\{R^0, (k^0 + 1 + \Gamma)e^\Gamma\}$. By Corollary 2.7 if ϑ^1 and τ are limits of $(\theta_j)_{j \in \mathbb{N}}$ in (Θ, D_{R^1}) , then $\iota_{k^0+1}(\vartheta^1) = \iota_{k^0+1}(\tau)$. Since ϑ^1 is also a limit of $(\theta_j)_{j \in \mathbb{N}}$ in (Θ, D_{R^0}) we deduce that $\iota_{k^0}\vartheta^0 = \iota_{k^0}\vartheta^1$. We extend θ on $[0, 1] \times E_{k^0+1}$ by setting $\theta(h, x) = \vartheta^1(h, x)$ for any $(h, x) \in [0, 1] \times E_{k^0+1}$. Assume that for some $k \geq 1$ we already constructed reals $R^0 < R^1 < \dots < R^k$ with $R^i > (k_0 + i + \Gamma)e^\Gamma$ and $\vartheta^i \in \Theta$ for $i = 0, 1, \dots, k$ with $\lim_{j \rightarrow \infty} D_{R^i}(\vartheta^i, \theta_j) = 0$ satisfying $\iota_{k_0+i}(\vartheta^i) = \iota_{k_0+i}(\vartheta^{i+1})$ for any $1 \leq i < k$ and defined $\theta(h, x) = \vartheta^k(h, x)$ for $(h, x) \in [0, 1] \times E_{k_0+k}$. Consider next any $R^{k+1} > \max\{R^k, (k_0 + k + 1 + \Gamma)e^\Gamma\}$ and any limit ϑ^{k+1} of $(\theta_j)_{j \in \mathbb{N}}$ in $(\Theta, D_{R^{k+1}})$. By the same arguments as before $\iota_{k_0+k}(\vartheta^k) = \iota_{k_0+k}(\vartheta^{k+1})$ and for any other limit τ of $(\theta_j)_{j \in \mathbb{N}}$ in $(\Theta, D_{R^{k+1}})$ we have $\iota_{k_0+k+1}(\vartheta^{k+1}) = \iota_{k_0+k+1}(\tau)$. Set $\theta(h, x) = \vartheta^{k+1}(h, x)$ for $(h, x) \in [0, 1] \times E_{k_0+k+1}$. In this way we obtain θ defined on $[0, 1] \times E$. Since $\vartheta^k \in \Theta$ for all k , we deduce that θ is a transition.

We claim that θ_j converge to θ in (Θ, D) . Fix any $\varepsilon > 0$ and consider n such that $\sum_{k \geq k_0+n+1} 2^{-k} < \frac{\varepsilon}{2}$. By construction, $\lim_{j \rightarrow \infty} D_{R^n}(\vartheta^n, \theta_j) = 0$ and $\vartheta^n(h, x) = \theta(h, x)$ for any $(h, x) \in [0, 1] \times E_{k_0+n}$. Then $D_{k_0+n}(\vartheta^n, \theta_j) < \frac{\varepsilon}{2}$ for all large j , because $R^n > k_0 + n$. Therefore, for all $1 \leq k \leq k_0 + n$

$$D_k(\theta, \theta_j) \leq D_{k_0+n}(\theta, \theta_j) < \frac{\varepsilon}{2}$$

whenever j is sufficiently large. Consequently, $\sum_{k=1}^{k_0+n} 2^{-k} D_k(\theta, \theta_j) < \frac{\varepsilon}{2}$ implying that $\sum_{k \geq 1} 2^{-k} D_k(\theta, \theta_j) < \varepsilon$. The arbitrariness of $\varepsilon > 0$ implies our claim. \square

Proof of Proposition 2.9. Define $\psi : [0, 1 - \max\{t_1, t_2\}] \rightarrow [0, \infty)$ by $\psi(h) = d(\vartheta(t_1 + h, x), \tau(t_2 + h, y))$. Proposition 2.3 ensures that for R as in the statement of Proposition 2.9, $\|\vartheta(h, x)\| \leq R$, $\|\tau(h, y)\| \leq R$ for each h . ψ is Lipschitz as so are $\vartheta(\cdot, x)$ and $\tau(\cdot, y)$. Moreover, it satisfies for every $h \in (0, 1 - \max\{t_1, t_2\})$ such that ψ is differentiable at h

$$\begin{aligned} \psi'(h) &= \lim_{\delta \downarrow 0} \frac{1}{\delta} \left(d(\vartheta(t_1 + h + \delta, x), \tau(t_2 + h + \delta, y)) - d(\vartheta(t_1 + h, x), \tau(t_2 + h, y)) \right) \\ &\leq \limsup_{\delta \downarrow 0} \frac{1}{\delta} \left(d(\vartheta(t_1 + h + \delta, x), \vartheta(\delta, \vartheta(t_1 + h, x))) + d(\vartheta(\delta, \tau(t_2 + h, y)), \tau(\delta, \tau(t_2 + h, y))) \right. \\ &\quad \left. + d(\vartheta(\delta, \vartheta(t_1 + h, x)), \vartheta(\delta, \tau(t_2 + h, y))) - d(\vartheta(t_1 + h, x), \tau(t_2 + h, y)) \right. \\ &\quad \left. + d(\tau(\delta, \tau(t_2 + h, y)), \tau(t_2 + h + \delta, y)) \right) \leq D_R(\vartheta, \tau) + \alpha_R(\vartheta) \cdot \psi(h). \end{aligned}$$

This and Gronwall's inequality complete the proof. \square

Proof of Proposition 2.12. Let $J \subset [0, T]$ denote the subset of $t \in [0, T]$ at which $\dot{x}(t) \neq \emptyset$ and the assumption (ii) holds. J has full measure in $[0, T]$. For every $t \in J$, there is $\vartheta_t \in \dot{x}(t)$ with $\beta_R(\vartheta_t) \leq B_R(t)$

and so,

$$\limsup_{h \downarrow 0} \frac{d(x(t+h), x(t))}{h} \leq \limsup_{h \downarrow 0} \frac{1}{h} \cdot d(x(t+h), \vartheta_t(h, x(t))) + \limsup_{h \downarrow 0} \frac{1}{h} \cdot d(\vartheta_t(h, x(t)), x(t)) \leq B_R(t).$$

Fix any $t_1 \in [0, T]$ and define $\psi : [t_1, T] \rightarrow \mathbb{R}$ by $\psi(s) = d(x(t_1), x(s))$. It is absolutely continuous and satisfies for all $s \in J$ such that $\psi'(s)$ does exist

$$\psi'(s) = \lim_{h \downarrow 0} \frac{d(x(t_1), x(s+h)) - d(x(t_1), x(s))}{h} \leq \limsup_{h \downarrow 0} \frac{d(x(s+h), x(s))}{h} \leq B_R(s)$$

implying the desired inequality. \square

Proof of Proposition 2.13. Define $\psi : [0, T] \rightarrow \mathbb{R}$ by $\psi(t) = d(x(t), y(t))$. It is absolutely continuous by assumption. Now let $J \subset (0, T)$ denote the subset of full measure such that $\dot{y}(t) \neq \emptyset$, ψ is differentiable at t for all $t \in J$ and both inequalities in (iii) of our Proposition are satisfied. Fix $\varepsilon > 0$ arbitrarily and choose any $t \in J$. Let $\vartheta \in \dot{x}(t)$ be as in (iii) and pick $\tau \in \dot{y}(t)$ with $D_r(\vartheta, \tau) \leq \Delta_r(t) + \varepsilon$.

Similarly to the proof of Proposition 2.9 we get

$$\begin{aligned} \psi'(t) &\leq \limsup_{h \downarrow 0} \frac{1}{h} \cdot \left(d(x(t+h), \vartheta(h, x(t))) + d(\vartheta(h, x(t)), \vartheta(h, y(t))) - d(x(t), y(t)) + \right. \\ &\quad \left. d(\vartheta(h, y(t)), \tau(h, y(t))) + d(\tau(h, y(t)), y(t+h)) \right) \\ &\leq \alpha_R(\vartheta) \cdot d(x(t), y(t)) + D_r(\vartheta, \tau) \leq A_R(t) \cdot \psi(t) + \Delta_r(t) + \varepsilon. \end{aligned}$$

Gronwall's inequality and the arbitrariness of $\varepsilon > 0$ complete the proof. \square

Proof of Proposition 2.18. $D_r(\theta_1(\cdot), \theta_2(\cdot)) : [0, T] \rightarrow \mathbb{R}$ is measurable because both $\theta_1(\cdot)$ and $\theta_2(\cdot)$ are measurable and D_r is Lipschitz. Assumption (iii) of Proposition 2.18 guarantees its integrability. The rest of the proof is essentially as the one of Proposition 2.13. The only difference concerns the selections $\theta_1(\cdot)$, $\theta_2(\cdot)$ of $\dot{x}_1(\cdot)$, $\dot{x}_2(\cdot)$ respectively which are now given explicitly as measurable functions from $[0, T]$ into (Θ, D_r) . \square

The proof of Corollary 2.19 uses the following variant of Lebesgue points for measurable metric space-valued functions:

Lemma 4.1 *Let Y be a pseudo-metric space and $\psi : [0, T] \rightarrow Y$ and $\delta : Y \times Y \rightarrow [0, \infty)$ satisfy:*

- (i) $\psi(\cdot)$ is measurable, $\delta(\cdot)$ is continuous and satisfies the triangle inequality.
- (ii) For some (and thus, for every) $y_0 \in Y$, the function $M(\cdot) := \max\{\delta(y_0, \psi(\cdot)), \delta(\psi(\cdot), y_0)\} : [0, T] \rightarrow \mathbb{R}$ is integrable.

Then, $\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} \delta(\psi(t), \psi(s)) ds = 0$ holds for a.e. $t \in [0, T]$.

Proof of Lemma 4.1. Choose any sequence $(\varepsilon_\ell)_{\ell \in \mathbb{N}}$ in $(0, 1)$ with $\sum_{\ell=1}^{\infty} \varepsilon_\ell < \infty$. For each $\ell \in \mathbb{N}$, Lusin's Theorem B.6 provides a compact subset $I_\ell \subset [0, T]$ with $\mathcal{L}^1([0, T] \setminus I_\ell) < \varepsilon_\ell$ such that the restriction $\psi|_{I_\ell} : I_\ell \rightarrow Y$ is continuous. Set $\tilde{J}_k := \bigcap_{\ell \geq k} I_\ell \subset [0, T]$ for $k \in \mathbb{N}$.

Let $\chi_{[0, T] \setminus \tilde{J}_k} : \mathbb{R} \rightarrow \{0, 1\}$ denote the characteristic function of $[0, T] \setminus \tilde{J}_k \subset \mathbb{R}$ for each $k \in \mathbb{N}$. Due to assumption (ii), the product $\chi_{[0, T] \setminus \tilde{J}_k} \cdot M : [0, T] \rightarrow \mathbb{R}$ is integrable. Hence the set

$$\begin{aligned} J_k &:= \left\{ t \in [0, T] \mid \lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} |\chi_{[0, T] \setminus \tilde{J}_k}(t) - \chi_{[0, T] \setminus \tilde{J}_k}(s)| ds = 0 \right\} \\ &\cap \left\{ t \in [0, T] \mid \lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} |(\chi_{[0, T] \setminus \tilde{J}_k} \cdot M)(t) - (\chi_{[0, T] \setminus \tilde{J}_k} \cdot M)(s)| ds = 0 \right\} \end{aligned}$$

is of full measure. We obtain for all $k \in \mathbb{N}$ and $t \in \tilde{J}_k \cap J_k$

$$\begin{cases} \lim_{h \downarrow 0} \frac{1}{h} \cdot \mathcal{L}^1([t, t+h] \setminus \tilde{J}_k) &= \chi_{[0,T] \setminus \tilde{J}_k}(t) &= 0 \\ \lim_{h \downarrow 0} \frac{1}{h} \cdot \int_t^{t+h} (\chi_{[0,T] \setminus \tilde{J}_k} \cdot M)(s) \, ds &= \chi_{[0,T] \setminus \tilde{J}_k}(t) \cdot M(t) &= 0 \\ \lim_{h \downarrow 0} \frac{1}{h} \cdot \int_{[t, t+h] \cap \tilde{J}_k} \delta(\psi(t), \psi(s)) \, ds &= 0 \end{cases}$$

because $\tilde{J}_k \ni s \mapsto \delta(\psi(t), \psi(s))$ is a composition of continuous functions. Hence, the inequality

$$\begin{aligned} & \int_t^{t+h} \delta(\psi(t), \psi(s)) \, ds \\ &= \int_{[t, t+h] \cap \tilde{J}_k} \delta(\psi(t), \psi(s)) \, ds + \int_{[t, t+h] \setminus \tilde{J}_k} \delta(\psi(t), \psi(s)) \, ds \\ &\leq \int_{[t, t+h] \cap \tilde{J}_k} \delta(\psi(t), \psi(s)) \, ds + \int_{[t, t+h] \setminus \tilde{J}_k} \left(\delta(\psi(t), y_0) + \delta(y_0, \psi(s)) \right) \, ds \\ &\leq \int_{[t, t+h] \cap \tilde{J}_k} \delta(\psi(t), \psi(s)) \, ds + \int_{[t, t+h] \setminus \tilde{J}_k} (M(t) + M(s)) \, ds \end{aligned}$$

implies $\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} \delta(\psi(t), \psi(s)) \, ds = 0$ for all $k \in \mathbb{N}$ and $t \in \tilde{J}_k \cap J_k$.

Finally, this last equality holds for a.e. $t \in [0, T]$ since

$$\tilde{J}_k \subset \tilde{J}_{k+1} \quad \forall k \in \mathbb{N}, \quad \mathcal{L}^1([0, T] \setminus \tilde{J}_k) \leq \sum_{\ell=k}^{\infty} \mathcal{L}^1([0, T] \setminus I_\ell) \leq \sum_{\ell=k}^{\infty} \varepsilon_\ell \longrightarrow 0 \quad (k \rightarrow \infty).$$

□

Proof of Corollary 2.19. Proposition 2.18 implies that for all $k, \ell \in \mathbb{N}$ and $t \in [0, T]$,

$$\begin{aligned} d(x_k(t), x_\ell(t)) &\leq e^{\|A_R\|_{L^1}} \left(d(x_k(0), x_\ell(0)) + \int_0^t D_r(\theta_k(s), \theta_\ell(s)) \, ds \right) \\ &\leq e^{\|A_R\|_{L^1}} \left(d(x_k(0), x_\ell(0)) + \|D_r(\theta_k(\cdot), \theta_\ell(\cdot))\|_{L^1} \right). \end{aligned}$$

Due to assumption (iv), $(x_k(\cdot))_{k \in \mathbb{N}}$ is a Cauchy sequence in $C^0([0, T], (E, d))$.

(E_r, d) is complete by assumption (HΘ) (ii) and thus, $(x_k(\cdot))_{k \in \mathbb{N}}$ converges uniformly to a function $x : [0, T] \rightarrow (E_r, d)$ with $x(0) = x_0$. In particular, we have $\|x(t)\| \leq r < R$ for all $t \in [0, T]$.

$x(\cdot)$ is absolutely continuous because for all $t_1, t_2 \in [0, T]$ ($t_1 < t_2$), assumption (ii) and Proposition 2.12 lead to

$$d(x(t_1), x(t_2)) = \lim_{k \rightarrow \infty} d(x_k(t_1), x_k(t_2)) \leq \int_{t_1}^{t_2} B_r(s) \, ds.$$

It remains to prove that $\theta(t) \in \mathring{x}(t)$ for a.e. $t \in [0, T]$, i.e., $\lim_{h \downarrow 0} \frac{1}{h} \cdot d(\theta(t)(h, x(t)), x(t+h)) = 0$. The construction of $x(\cdot)$ and Proposition 2.18 guarantee for every $t \in [0, T]$ and all sufficiently small $h > 0$

$$\begin{aligned} d(\theta(t)(h, x(t)), x(t+h)) &= \lim_{k \rightarrow \infty} d(\theta(t)(h, x(t)), x_k(t+h)) \\ &\leq \limsup_{k \rightarrow \infty} e^{\|A_R\|_{L^1}} \left(d(x(t), x_k(t)) + \int_0^h D_r(\theta(t), \theta_k(t+\sigma)) \, d\sigma \right) \\ &\leq \limsup_{k \rightarrow \infty} e^{\|A_R\|_{L^1}} \int_0^h \left(D_r(\theta(t), \theta(t+\sigma)) + D_r(\theta(t+\sigma), \theta_k(t+\sigma)) \right) \, d\sigma \\ &\leq \limsup_{k \rightarrow \infty} e^{\|A_R\|_{L^1}} \left(\int_0^h D_r(\theta(t), \theta(t+\sigma)) \, d\sigma + \|D_r(\theta, \theta_k)\|_{L^1} \right) \\ &\stackrel{(iv)}{=} e^{\|A_R\|_{L^1}} \cdot \int_0^h D_r(\theta(t), \theta(t+\sigma)) \, d\sigma. \end{aligned}$$

Finally, Lemma 4.1 implies that for a.e. $t \in [0, T)$

$$\lim_{h \downarrow 0} \frac{1}{h} \cdot d(\theta(t)(h, x(t)), x(t+h)) \leq e^{\|A_R\|_{L^1}} \cdot \lim_{h \downarrow 0} \frac{1}{h} \cdot \int_0^h D_r(\theta(t), \theta(t+\sigma)) d\sigma = 0.$$

□

The next goal is to prove Corollary 2.20. On the one hand, the transition-valued curve $\theta : [0, T] \rightarrow (\Theta, D_R)$ is assumed to be (merely) measurable. On the other hand, Proposition 2.3 provides only a priori bound w.r.t. $\|\cdot\|$ so far and it is restricted to piecewise constant curves of transitions.

This gap is to be bridged by a Lusin-type argument relating measurability to almost continuity. The following lemma lays the foundations and is a key tool for several similar conclusions below.

Lemma 4.2 *Let (Y, d_Y) be a pseudo-metric space, $\tilde{J} \subset [0, T]$ be a compact subset such that $\mathcal{L}^1([0, T] \setminus \tilde{J}) < \frac{\varepsilon}{2}$ for some $\varepsilon > 0$. Then for any measurable $\psi : [0, T] \rightarrow Y$ and $\varepsilon_\psi > 0$ there exist a compact subset $J_\varepsilon \subset \tilde{J}$ and a finite partition $0 = t_0 < t_1 < \dots < t_L = T$ with the following properties:*

- (a) $\mathcal{L}^1([0, T] \setminus J_\varepsilon) < \varepsilon$, $\psi|_{J_\varepsilon} : J_\varepsilon \rightarrow Y$ is continuous.
- (b) For each $\ell \in \{1, \dots, L\}$, it holds $d_Y(\psi(s_1), \psi(s_2)) < \varepsilon_\psi$ for all $s_1, s_2 \in [t_{\ell-1}, t_\ell] \cap J_\varepsilon$.
- (c) For each $\ell \in \{1, \dots, L\}$ with $t_{\ell-1} \notin J_\varepsilon$, $[t_{\ell-1}, t_\ell] \cap J_\varepsilon = \emptyset$.

Proof. Lusin's Theorem B.6 provides a compact subset $J_\varepsilon \subset \tilde{J}$ with $\mathcal{L}^1([0, T] \setminus J_\varepsilon) < \varepsilon$ such that $\psi|_{J_\varepsilon}$ is continuous. Hence, ψ is even uniformly continuous in J_ε and so, there exists $\delta_\varepsilon \in (0, \frac{\varepsilon}{2})$ satisfying $d_Y(\psi(s_1), \psi(s_2)) < \varepsilon_\psi$ for all $s_1, s_2 \in J_\varepsilon$ with $|s_1 - s_2| < \delta_\varepsilon$.

Let $0 = \tau_0 < \tau_1 < \dots < \tau_K = T$ denote any finite partition of $[0, T]$ with $\max_k |\tau_k - \tau_{k-1}| < \delta_\varepsilon$. A further step of refinement leads the final partition $0 = t_0 < t_1 < \dots < t_L = T$ of $[0, T]$ according to the following rule for $k = 1, \dots, K$: Whenever $\tau_{k-1} \notin J_\varepsilon$ and $(\tau_{k-1}, \tau_k) \cap J_\varepsilon \neq \emptyset$, then the next node is $t_\ell := \min((\tau_{k-1}, \tau_k) \cap J_\varepsilon) > \tau_{k-1}$. Otherwise set $t_\ell := \tau_k$. □

Proof of Corollary 2.20. Let $x(\cdot)$ be a mutational primitive of θ with $\|x(0)\| \leq r$ and $x_0 := x(0)$. Set $\rho := \max\{R, \sup_t \|x(t)\|\}$. Choose any $j > \max\{\alpha_\rho(\vartheta_0), \beta_R(\vartheta_0), \gamma(\vartheta_0)\}$ sufficiently large such that the measurable set $I_j := \{t \in [0, T] \mid \Psi_R(t) \leq j, A_\rho(t) \leq j, B_R(t) \leq j \text{ and } \Gamma(t) \leq j\}$ satisfies

$$\left(\|x_0\| + \|\Gamma\|_{L^1} + \gamma(\vartheta_0) \cdot \mathcal{L}^1([0, T] \setminus I_j) \right) \cdot e^{\|\Gamma\|_{L^1} + \gamma(\vartheta_0) \cdot \mathcal{L}^1([0, T] \setminus I_j)} < R.$$

Define $\theta_j : [0, T] \rightarrow \Theta$ and $A_{\rho,j}, B_{R,j}, \Gamma_j : [0, T] \rightarrow \mathbb{R}$ as

$$\begin{aligned} \theta_j(t) &:= \begin{cases} \theta(t) & \text{if } t \in I_j \\ \vartheta_0 & \text{if } t \notin I_j \end{cases} & A_{\rho,j}(t) &:= \begin{cases} A_\rho(t) & \text{if } t \in I_j \\ \alpha_\rho(\vartheta_0) & \text{if } t \notin I_j \end{cases} \\ B_{R,j}(t) &:= \begin{cases} B_R(t) & \text{if } t \in I_j \\ \beta_R(\vartheta_0) & \text{if } t \notin I_j, \end{cases} & \Gamma_j(t) &:= \begin{cases} \Gamma(t) & \text{if } t \in I_j \\ \gamma(\vartheta_0) & \text{if } t \notin I_j. \end{cases} \end{aligned}$$

For each index $k \in \mathbb{N}$, consider a compact subset $\tilde{J}_k \subset [0, T)$ with $\mathcal{L}^1([0, T] \setminus \tilde{J}_k) < \frac{1}{2j^k}$ such that $\alpha_\rho(\theta(t)) \leq A_\rho(t)$, $\beta_R(\theta(t)) \leq B_R(t)$, $\gamma(\theta(t)) \leq \Gamma(t)$ and $D_R(\theta(t), \vartheta_0) \leq \Psi_R(t)$ hold for every $t \in \tilde{J}_k$.

Lemma 4.2 applied to \tilde{J}_k and $\theta_j : [0, T] \rightarrow (\Theta, D_R)$, $A_{\rho,j}, B_{R,j}, \Gamma_j, \Psi_R \in L^1([0, T])$ provides a compact subset $J_{j,k} \subset \tilde{J}_k$ and a finite partition $0 = t_{j,k,0} < t_{j,k,1} < \dots < t_{j,k,L(j,k)} = T$ with the following properties for each index $\ell \in \{1, \dots, L(j,k)\}$:

- $\mathcal{L}^1([0, T] \setminus J_{j,k}) < \frac{1}{j^k}$, $0 < |t_{j,k,\ell} - t_{j,k,\ell-1}| < \frac{1}{k}$
- $(\theta_j)|_{J_{j,k}} : J_{j,k} \rightarrow (\Theta, D_R)$ and $(A_{\rho,j})|_{J_{j,k}}, (B_{R,j})|_{J_{j,k}}, (\Gamma_j)|_{J_{j,k}}, (\Psi_R)|_{J_{j,k}} : J_{j,k} \rightarrow \mathbb{R}$ are continuous

- $\alpha_\rho(\theta_j(t)) \leq A_{\rho,j}(t)$, $\beta_R(\theta_j(t)) \leq B_{R,j}(t)$, $\gamma(\theta_j(t)) \leq \Gamma_j(t)$ and $D_R(\vartheta_0, \theta_j(t)) \leq \Psi_R(t)$ for all $t \in J_{j,k}$
- for all $s, t \in [t_{j,k,\ell-1}, t_{j,k,\ell}] \cap J_{j,k}$, $\max \left\{ D_R(\theta_j(s), \theta_j(t)), |A_{\rho,j}(s) - A_{\rho,j}(t)|, |B_{R,j}(s) - B_{R,j}(t)|, |\Gamma_j(s) - \Gamma_j(t)| \right\} < \frac{1}{Tk}$
- if $t_{j,k,\ell-1} \notin J_{j,k}$ then $[t_{j,k,\ell-1}, t_{j,k,\ell}) \subset [0, T] \setminus J_{j,k}$.

We define the piecewise constant functions $\theta_{j,k} : [0, T] \longrightarrow \Theta$ and $\Gamma_{j,k} : [0, T] \longrightarrow \mathbb{R}$ as follows: for $t \in [t_{j,k,\ell-1}, t_{j,k,\ell})$

$$\begin{cases} \text{if } t_{j,k,\ell-1} \in J_{j,k}, \text{ then set } \theta_{j,k}(t) := \theta_j(t_{j,k,\ell-1}), & \Gamma_{j,k}(t) := \gamma(\theta_j(t_{j,k,\ell-1})) \leq j, \\ \text{if } t_{j,k,\ell-1} \notin J_{j,k}, \text{ then set } \theta_{j,k}(t) := \vartheta_0, & \Gamma_{j,k}(t) := \gamma(\vartheta_0). \end{cases}$$

For all $t \in [0, T] \setminus J_{j,k}$, we have the obvious bound

$$\max \{ D_R(\vartheta_0, \theta_j(t)), D_R(\vartheta_0, \theta_{j,k}(t)), \alpha_\rho(\theta_{j,k}(t)), \beta_R(\theta_{j,k}(t)), \Gamma_{j,k}(t) \} \leq j.$$

Furthermore, every $t \in J_{j,k}$ is contained in a subinterval $[t_{j,k,\ell-1}, t_{j,k,\ell})$ with $t_{j,k,\ell-1} \in J_{j,k}$ and thus,

$$\begin{aligned} D_R(\theta_j(t), \theta_{j,k}(t)) &\leq \frac{1}{Tk}, & \alpha_\rho(\theta_{j,k}(t)) &\leq A_{\rho,j}(t) + \frac{1}{Tk}, \\ \beta_R(\theta_{j,k}(t)) &\leq B_{R,j}(t) + \frac{1}{Tk}, & \Gamma_{j,k}(t) &\leq \Gamma_j(t) + \frac{1}{Tk}. \end{aligned}$$

Next we construct a mutational primitive $x_{j,k} : [0, T] \longrightarrow E$ of $\theta_{j,k}(\cdot)$ with $x_{j,k}(0) = x_0$ in a piecewise way as indicated in Remark 2.11 (a), i.e., for $\ell \in \{1, \dots, L(j, k)\}$,

$$x_{j,k}(0) = x_0, \quad x_{j,k}(t) = \theta_{j,k}(t_{j,k,\ell-1}) (t - t_{j,k,\ell-1}, x_{j,k}(t_{j,k,\ell-1})) \quad \text{for } t \in (t_{j,k,\ell-1}, t_{j,k,\ell}].$$

As in the proof of Proposition 2.3, we conclude from condition (v) in Definition 2.1 (via induction w.r.t. the index ℓ of subintervals) that for every $t \in (t_{j,k,\ell-1}, t_{j,k,\ell}]$ and $k \in \mathbb{N}$,

$$\begin{aligned} \|x_{j,k}(t)\| &\leq (\|x_{j,k}(t_{j,k,\ell-1})\| + \gamma(\theta_{j,k}(t_{j,k,\ell-1})) \cdot (t - t_{j,k,\ell-1})) \cdot e^{\gamma(\theta_{j,k}(t_{j,k,\ell-1})) \cdot (t - t_{j,k,\ell-1})} \\ &\leq (\|x_0\| + \int_0^t \Gamma_{j,k}(s) ds) \cdot e^{\int_0^t \Gamma_{j,k}(s) ds}. \end{aligned}$$

The choice of $\Gamma_{j,k}$ and $J_{j,k}$ implies

$$\begin{aligned} \int_0^t \Gamma_{j,k}(s) ds &\leq \int_{[0,t] \cap J_{j,k}} (\Gamma_j(s) + \frac{1}{Tk}) ds + \int_{[0,t] \setminus J_{j,k}} \Gamma_{j,k}(s) ds \\ &\leq \int_{[0,t] \cap J_{j,k}} \Gamma_j(s) ds + \frac{1}{k} + j \cdot \mathcal{L}^1([0, t] \setminus J_{j,k}) \\ &\leq \int_0^t \Gamma_j(s) ds + \frac{2}{k} \leq \int_0^t \Gamma(s) ds + \gamma(\vartheta_0) \cdot \mathcal{L}^1([0, t] \setminus I_j) + \frac{2}{k}. \end{aligned}$$

Hence, $\limsup_{k \rightarrow \infty} \sup_{[0,t]} \|x_{j,k}(\cdot)\| \leq (\|x_0\| + \int_0^t \Gamma_j(s) ds) \cdot e^{\int_0^t \Gamma_j(s) ds} < R$.

According to Corollary 2.19 (applied with $r := R$), $(x_{j,k}(\cdot))_{k \in \mathbb{N}}$ converges uniformly (as $k \rightarrow \infty$) to a mutational primitive $x_j : [0, T] \longrightarrow E$ of $\theta_j(\cdot)$ with $x_j(0) = x_0$. The lower semicontinuity of $\|\cdot\|$ guarantees

$$\|x_j(t)\| \leq \liminf_{k \rightarrow \infty} \|x_{j,k}(t)\| \leq (\|x_0\| + \int_0^t \Gamma_j(s) ds) \cdot e^{\int_0^t \Gamma_j(s) ds} < R$$

for all $t \in [0, T]$. On the other hand, $\lim_{j \rightarrow \infty} \mathcal{L}^1([0, t] \setminus I_j) = 0$. This, the Lebesgue dominated convergence theorem and Corollary 2.19 (applied with $r := R$) imply that $(x_j(\cdot))_{j \in \mathbb{N}}$ converges uniformly (as $j \rightarrow \infty$) to a mutational primitive $\tilde{x} : [0, T] \longrightarrow E$ of $\theta(\cdot)$ with $\tilde{x}(0) = x_0$. Proposition 2.18 (applied with $r := R$) yields $x = \tilde{x}$. The lower semicontinuity of $\|\cdot\|$ and the last inequality complete the proof. \square

Proof of Proposition 2.22. We apply the same construction as in the proof of Corollary 2.20:

Choose any $j \in \mathbb{N}$ sufficiently large such that $j > \max\{\alpha_R(\vartheta_0), \beta_R(\vartheta_0), \gamma(\vartheta_0)\}$ and the measurable set $I_j := \{t \in [0, T] \mid \Psi_R(t) \leq j, A_R(t) \leq j, B_R(t) \leq j \text{ and } \Gamma(t) \leq j\}$ satisfies

$$\left(\|x_0\| + \|\Gamma\|_{L^1} + \gamma(\vartheta_0) \cdot \mathcal{L}^1([0, T] \setminus I_j) \right) \cdot e^{\|\Gamma\|_{L^1} + \gamma(\vartheta_0) \cdot \mathcal{L}^1([0, T] \setminus I_j)} < R.$$

Now define $\theta_j : [0, T] \longrightarrow \Theta$ and $A_{R,j}, B_{R,j}, \Gamma_j : [0, T] \longrightarrow \mathbb{R}$ as before. We already know that for every large j there exists a mutational primitive $x_j : [0, T] \longrightarrow E$ of $\theta_j(\cdot)$ with $x_j(0) = x_0$ and $\|x_j(t)\| < R$ for all $t \in [0, T]$.

Since $\|D_R(\theta(\cdot), \theta_j(\cdot))\|_{L^1} \longrightarrow 0$ ($j \rightarrow \infty$), Corollary 2.19 yields the uniform convergence of $(x_j(\cdot))_{j \in \mathbb{N}}$ to a mutational primitive $x : [0, T] \longrightarrow E$ of $\theta(\cdot)$ with $x(0) = x_0$ and $\|x(t)\| < R$ for all $t \in [0, T]$. \square

5 Proofs of Results in Section 3

Proof of Proposition 3.2. The implication “(b) \implies (a)” is obvious. The implication “(a) \implies (b)” results from Leese’s selection statement in Proposition B.5 in combination with Proposition 2.15. Let $x : [0, T] \longrightarrow E$ be continuous with $\sup_t \|x(t)\| \leq R$ and satisfy statement (a). Fix any $\vartheta_0 \in E$.

We claim that the set-valued map $F : [0, T] \rightsquigarrow (\Theta, D_{\text{loc}})$ defined by $F(t) = \dot{x}(t)$ if $\dot{x}(t) \neq \emptyset$ and $F(t) = \{\vartheta_0\}$ otherwise is measurable. Indeed, by Proposition 2.15, $\dot{x}(\cdot) : [0, T] \rightsquigarrow (\Theta, D_{\text{loc}})$ is measurable. In particular, the set $\{t \mid \dot{x}(t) = \emptyset\}$ is measurable. Consider any open set $O \subset (\Theta, D_{\text{loc}})$. If it does not contain ϑ_0 , then the set $\{t \mid F(t) \cap O \neq \emptyset\} = \{t \mid \dot{x}(t) \cap O \neq \emptyset\}$ is measurable. If O contains ϑ_0 then $\{t \mid \dot{x}(t) \cap O \neq \emptyset\}$ is measurable and $\{t \mid F(t) \cap O \neq \emptyset\} = \{t \mid \dot{x}(t) \cap O \neq \emptyset\} \cup \{t \mid \dot{x}(t) = \emptyset\}$. Consequently, F is measurable with nonempty images.

Due to hypothesis (HΘ) (iv) (i.e., the continuity of $\iota_R : (\Theta, D_{\text{loc}}) \longrightarrow (\Theta_R, D_R)$), the composition $F_R = \iota_R \circ F : [0, T] \rightsquigarrow (\Theta_R, D_R)$ is also measurable. Its values are nonempty and closed due to Remark 2.11 (b).

Therefore the set-valued map $G : [0, T] \rightsquigarrow E_R \times (\Theta_R, D_R)$ defined by $G(t) = \{x(t)\} \times F_R(t)$ is measurable with nonempty closed images. Proposition B.4 implies that G is graph measurable. Consequently, by assumption,

$$\text{Graph } G \cap \text{Graph } (\mathfrak{F}_R)_{|[0, T] \times E_R} \in \mathcal{L}([0, T]) \otimes \mathcal{B}(E_R, d) \otimes \mathcal{B}(\Theta_R, D_R).$$

This means

$$\{\{t\} \times \{x(t)\} \times (F_R(t) \cap \mathfrak{F}_R(t, x(t))) \mid t \in [0, T]\} \in \mathcal{L}([0, T]) \otimes \mathcal{B}(E_R, d) \otimes \mathcal{B}(\Theta_R, D_R).$$

Define $\Phi(t) = \{x(t)\} \times (F_R(t) \cap \mathfrak{F}_R(t, x(t)))$ if the intersection $F_R(t) \cap \mathfrak{F}_R(t, x(t))$ is nonempty and $\Phi(t) = \{x(t)\} \times \{\iota_R \vartheta_0\}$ otherwise. Then the set-valued map $\Phi : [0, T] \rightsquigarrow E_R \times (\Theta_R, D_R)$ is graph measurable with nonempty values. By Proposition B.5 it has a measurable selection $z(t) \in \Phi(t)$ for $t \in [0, T]$. Let $\bar{\theta}(t) \in \Theta_R$ be such that $z(t) = (x(t), \bar{\theta}(t))$. For each $t \in [0, T]$ with $\iota_R \dot{x}(t) \cap \mathfrak{F}_R(t, x(t)) \neq \emptyset$, we have $\bar{\theta}(t) \in F_R(t) \cap \mathfrak{F}_R(t, x(t)) = \iota_R \dot{x}(t) \cap \mathfrak{F}_R(t, x(t))$ by the definition of F , Φ and let $\theta(t) \in \mathfrak{F}(t, x(t)) \subset \Theta$ such that $\bar{\theta}(t) = \iota_R \theta(t)$. Then $\theta(t) \in \dot{x}(t)$ results from $\bar{\theta}(t) \in F_R(t) = \iota_R \dot{x}(t)$ and Definition 2.10 of $\dot{x}(t)$. For all remaining $t \in [0, T]$ set $\theta(t) := \vartheta_0$. Finally, $\theta : [0, T] \longrightarrow (\Theta, D_R)$ is measurable because $z(\cdot)$ is measurable and ι_R maps every open set in (Θ, D_R) to an open set in (Θ_R, D_R) . \square

Lemma 5.1 (e.g., [9, p.122]) *Consider non-negative $\varphi, \lambda \in L^1([0, T])$, $k \in \mathbb{N}$, $\kappa \geq 1$ and define $\Lambda : [0, T] \longrightarrow [0, \infty)$ by $\Lambda(t) := \kappa \int_0^t \lambda(s) ds$. Then for every $t \in [0, T]$,*

$$\int_0^t \kappa \lambda(s) \cdot \left(\int_0^s \varphi(r) \frac{(\Lambda(s) - \Lambda(r))^{k-1}}{(k-1)!} dr \right) ds = \int_0^t \varphi(s) \frac{(\Lambda(t) - \Lambda(s))^k}{k!} ds.$$

Proof of Theorem 3.4.

STEP 1 Define $\psi : [0, T] \longrightarrow \mathbb{R} \cup \{\infty\}$ by $\psi(t) = \text{dist}_R(\dot{y}(t), \mathfrak{F}(t, y(t)))$. We claim that it is measurable.

Due to HYPOTHESIS $(\mathbb{H}\Theta)$, the pseudo-metric spaces (Θ_R, D_R) and (E_R, d) are separable and weakly Suslin. Furthermore, $(\mathfrak{F}_R)_{|[0, T] \times E_R} : [0, T] \times (E_R, d) \rightsquigarrow (\Theta_R, D_R)$ is graph measurable with nonempty closed values as a consequence of Propositions B.4, B.7 and assumptions (iv) – (vi). Furthermore, assumption (vi) implies that for a.e. $t \in [0, T]$, the graph of $\mathfrak{F}_R(t, \cdot)_{|E_R} : (E_R, d) \rightsquigarrow (\Theta_R, D_R)$ is closed. By Proposition C.1 for each $\varepsilon > 0$, there is a closed subset $J_\varepsilon \subset [0, T]$ with $\mathcal{L}^1([0, T] \setminus J_\varepsilon) < \varepsilon$ such that the graph of $(\mathfrak{F}_R)_{|J_\varepsilon \times E_R} : J_\varepsilon \times (E_R, d) \rightsquigarrow (\Theta_R, D_R)$ is closed. Due to the continuity of $y(\cdot)$, the graph of the composition

$$\mathfrak{F}_R(\cdot, y(\cdot))_{|J_\varepsilon} : J_\varepsilon \rightsquigarrow (\Theta_R, D_R), \quad t \mapsto \mathfrak{F}_R(t, y(t))$$

is also closed and thus, it belongs to $\mathcal{L}([0, T]) \otimes \mathcal{B}(\Theta_R, D_R)$. Set $\Phi(t) = \mathfrak{F}_R(t, y(t))$ for $t \in J_\varepsilon$ and $\Phi(t) = \{\iota_R \vartheta_0\}$ for $t \in [0, T] \setminus J_\varepsilon$. Then the graph of Φ belongs to $\mathcal{L}([0, T]) \otimes \mathcal{B}(\Theta_R, D_R)$. Proposition B.5 guarantees a Castaing representation of Φ , i.e., there exists a sequence $(\eta_{\varepsilon, j})_{j \in \mathbb{N}}$ of measurable functions $\eta_{\varepsilon, j} : [0, T] \longrightarrow (\Theta_R, D_R)$ such that for all $t \in [0, T]$, $\{\eta_{\varepsilon, j}(t) \mid j \in \mathbb{N}\}$ is dense in $\Phi(t)$ w.r.t. D_R . From continuity of $D_R(\cdot, \cdot)$ we deduce that for any $\vartheta \in \Theta$, the function $D_R(\vartheta, \eta_{\varepsilon, j}(\cdot))$ is measurable on $[0, T]$. Define $g(t, \vartheta) = \text{dist}_R(\vartheta, \mathfrak{F}(t, y(t)))$ for $\vartheta \in \Theta$, $t \in [0, T]$. Therefore,

- For every $\vartheta \in \Theta$, the marginal function $J_\varepsilon \ni t \longmapsto g(t, \vartheta) := \inf_{j \in \mathbb{N}} D_R(\vartheta, \eta_{\varepsilon, j}(t))$ is measurable (e.g., [10, Lemma 8.2.12], [13, Theorem 2.1.5]).
- For every $t \in [0, T]$, the function $\Theta \ni \vartheta \longmapsto \inf_{\tau \in \mathfrak{F}(t, y(t))} D_R(\vartheta, \tau)$ is 1-Lipschitz w.r.t. D_R . Considering its composition with the continuous embedding $(\Theta, D_{\text{loc}}) \longrightarrow (\Theta, D_R)$ by $(\mathbb{H}\Theta)$ (iv), we obtain its continuity w.r.t. D_{loc} .

Consider any sequence $\varepsilon_n \rightarrow 0+$. Since $\bigcup_{n \geq 1} J_{\varepsilon_n}$ is of full measure in $[0, T]$, we deduce that $g(\cdot, \vartheta)$ is measurable for every $\vartheta \in \Theta$. By Proposition 2.15, $\dot{y}(\cdot) : [0, T] \rightsquigarrow (\Theta, D_{\text{loc}})$ is measurable with closed images. Set $F(t) = \dot{y}(t)$ if $\dot{y}(t) \neq \emptyset$ and $F(t) = \{\vartheta_0\}$ otherwise. Then $F : [0, T] \rightsquigarrow (\Theta, D_{\text{loc}})$ is measurable with closed nonempty images. Furthermore $\psi(t) = \inf_{\vartheta \in F(t)} g(t, \vartheta)$ a.e. in $[0, T]$. Since the metric space (Θ, D_{loc}) is complete and separable, [10, Theorem 8.2.11] implies our claim. Due to assumption (vii), ψ is even integrable.

STEP 2 In this step we construct a sequence $(\theta_j(\cdot))_{j \in \mathbb{N}}$ of measurable functions $[0, T] \longrightarrow (\Theta, D_R)$ and the corresponding sequence of their primitives $(x_j(\cdot))_{j \in \mathbb{N}}$ in a recursive way.

Now fix $\kappa > 1$ arbitrarily and let $g(\cdot, \cdot)$ be as in STEP 1. We already know that it is measurable with respect to the first variable and continuous with respect to the second variable on (Θ, D_{loc}) . By [10, Theorem 8.2.9] applied to $H(\cdot) \equiv \Theta$, the set-valued maps

$$[0, T] \ni t \longmapsto \{\vartheta \in (\Theta, D_{\text{loc}}) \mid g(t, \vartheta) \leq \sqrt{\kappa} \psi(t)\}, \quad [0, T] \ni t \longmapsto \{\vartheta \in (\Theta, D_{\text{loc}}) \mid D_R(\vartheta_0, \vartheta) \leq \Psi_R(t)\}$$

are measurable with closed values. Due to Proposition 2.15, [10, Theorem 8.2.4] (about intersections) and assumption (vii), the set-valued map $G : [0, T] \rightsquigarrow (\Theta, D_{\text{loc}})$ defined by

$$G(t) := \{\vartheta \in \dot{y}(t) \mid g(t, \vartheta) \leq \sqrt{\kappa} \psi(t), D_R(\vartheta_0, \vartheta) \leq \Psi_R(t)\}$$

is measurable with nonempty closed values a.e. in $[0, T]$. Set $\widehat{G}(t) = G(t)$ if $G(t) \neq \emptyset$ and $\widehat{G}(t) = \{\vartheta_0\}$ otherwise. Then $\widehat{G} : [0, T] \rightsquigarrow (\Theta, D_{\text{loc}})$ is measurable with closed nonempty values.

Thus, the selection theorem of Kuratowski and Ryll-Nardzewski (e.g., [37], [10, Theorem 8.1.3]) provides a measurable $\tau_0 : [0, T] \longrightarrow (\Theta, D_{\text{loc}})$ such that $\tau_0(t) \in \widehat{G}(t)$ for all $t \in [0, T]$. In particular, for a.e. $t \in [0, T]$,

$$\tau_0(t) \in \mathring{y}(t), \quad \text{dist}_R(\tau_0(t), \mathfrak{F}(t, y(t))) \leq \sqrt{\kappa} \psi(t), \quad D_R(\vartheta_0, \tau_0(t)) \leq \Psi_R(t).$$

Next we claim that there exists a measurable function $\theta_0 : [0, T] \longrightarrow (\Theta, D_R)$ satisfying for a.e. $t \in [0, T]$,

$$\theta_0(t) \in \mathfrak{F}(t, y(t)), \quad D_R(\tau_0(t), \theta_0(t)) \leq \sqrt{\kappa} \cdot \text{dist}_R(\tau_0(t), \mathfrak{F}(t, y(t))) \leq \kappa \psi(t).$$

Indeed, the function $[0, T] \ni t \longmapsto \text{dist}_R(\tau_0(t), \mathfrak{F}(t, y(t)))$ is integrable due to three arguments: First, it is bounded a.e. by $\sqrt{\kappa} \Psi_R$. Second, $\phi : [0, T] \times (\Theta, D_R) \longrightarrow [0, \infty)$ defined by $\phi(t, \vartheta) := g(t, \vartheta)$ is a Carathéodory function because by **STEP 1** it is measurable in time and it is Lipschitz with respect to ϑ in (Θ, D_R) . Third, the measurable function $\tau_0 : [0, T] \longrightarrow (\Theta, D_R)$ is the pointwise limit of a sequence of simple functions (Remark **B.3**) and so, the standard arguments (known for complete separable metric spaces) now guarantee the measurability of the function $[0, T] \ni t \longmapsto \phi(t, \tau_0(t))$ (see, e.g., [10, Lemma 8.2.3]). Hence, the set-valued map

$$[0, T] \rightsquigarrow (\Theta, D_{\text{loc}}), \quad t \mapsto \{\vartheta \in \Theta \mid D_R(\tau_0(t), \vartheta) \leq \sqrt{\kappa} \cdot \phi(t, \tau_0(t))\}$$

is measurable with nonempty closed values (due to [10, Theorem 8.2.9]) and so is its composition with the continuous restriction $\iota_R : (\Theta, D_{\text{loc}}) \longrightarrow (\Theta_R, D_R)$, i.e.,

$$[0, T] \rightsquigarrow (\Theta_R, D_R), \quad t \mapsto \{\vartheta \in \Theta_R \mid D_R(\tau_0(t), \vartheta) \leq \sqrt{\kappa} \cdot \phi(t, \tau_0(t))\}$$

is measurable with nonempty closed values. The latter is graph measurable due to Proposition **B.4**. Consider any sequence $\varepsilon_n \rightarrow 0+$. By **STEP 1** the graph of $\mathfrak{F}_R(\cdot, y(\cdot))|_{J_{\varepsilon_n}} : J_{\varepsilon_n} \rightsquigarrow (\Theta_R, D_R)$ belongs to $\mathcal{L}([0, T]) \otimes \mathcal{B}(\Theta_R, D_R)$. Since $\bigcup_{n \geq 1} J_{\varepsilon_n}$ is of full measure in $[0, T]$ we deduce that the graph of $\mathfrak{F}_R(\cdot, y(\cdot))$ also belongs to $\mathcal{L}([0, T]) \otimes \mathcal{B}(\Theta_R, D_R)$. Hence, the map

$$[0, T] \rightsquigarrow (\Theta_R, D_R), \quad t \mapsto \{\vartheta \in \Theta_R \mid \vartheta \in \mathfrak{F}_R(t, y(t)), D_R(\tau_0(t), \vartheta) \leq \sqrt{\kappa} \cdot \text{dist}_R(\tau_0(t), \mathfrak{F}(t, y(t)))\}$$

(as an intersection) is graph measurable with nonempty values. Proposition **B.5** provides a measurable selection $\bar{\theta}_0 : [0, T] \longrightarrow (\Theta_R, D_R)$. For a.e. $t \in [0, T]$, let $\theta_0(t) \in \mathfrak{F}(t, y(t))$ be such that its restriction (to $[0, 1] \times E_R$) is $\bar{\theta}_0(t)$. Then $\theta_0 : [0, T] \longrightarrow (\Theta, D_R)$ is measurable since ι_R maps every open set in (Θ, D_R) to an open set in (Θ_R, D_R) . Furthermore $\theta_0(\cdot)$ is a selection of $[0, T] \rightsquigarrow (\Theta, D_R)$, $t \mapsto \{\vartheta \in \mathfrak{F}(t, y(t)) \mid D_R(\tau_0(t), \vartheta) \leq \sqrt{\kappa} \cdot \text{dist}_R(\tau_0(t), \mathfrak{F}(t, y(t)))\}$ (as required). Moreover for a.e. $t \in [0, T]$, we have

$$\begin{aligned} D_R(\vartheta_0, \theta_0(t)) &\leq D_R(\vartheta_0, \tau_0(t)) + D_R(\tau_0(t), \theta_0(t)) \\ &\leq D_R(\vartheta_0, \tau_0(t)) + \sqrt{\kappa} \cdot \text{dist}_R(\tau_0(t), \mathfrak{F}(t, y(t))) \\ &\leq \Psi_R(t) + \kappa \psi(t) \leq \Psi_R(t) (1 + \kappa). \end{aligned}$$

This is the basis for constructing two sequences $(x_j(\cdot))_{j \in \mathbb{N}}$, $(\theta_j(\cdot))_{j \in \mathbb{N}}$ recursively:

- According to Proposition **2.22** there exists a mutational primitive $x_j : [0, T] \longrightarrow E$ of $\theta_{j-1}(\cdot)$ with $x_j(0) = x_0$ and $\sup_{t \in [0, T]} \|x_j(t)\| < R$.
- Construct $\theta_j : [0, T] \longrightarrow (\Theta, D_R)$ as a measurable function satisfying for a.e. $t \in [0, T]$,

$$\theta_j(t) \in \mathfrak{F}(t, x_j(t)), \quad D_R(\theta_{j-1}(t), \theta_j(t)) \leq \kappa \cdot \text{dist}_R(\theta_{j-1}(t), \mathfrak{F}(t, x_j(t))) \leq \kappa \lambda_R(t) \cdot d(x_{j-1}(t), x_j(t)).$$

Hence, the assumptions of Proposition **2.22** are satisfied in each step of recursion. Proposition **2.12** implies for every $t_1, t_2 \in [0, T]$ ($t_1 \leq t_2$)

$$d(x_j(t_1), x_j(t_2)) \leq \int_{t_1}^{t_2} B_R(s) \, ds.$$

STEP 3 We claim that for $\Lambda_R(t) := \kappa \cdot \int_0^t \lambda_R(s) \, ds$ and all $j \in \mathbb{N}$:

$$d(x_j(t), x_{j+1}(t)) \leq \int_0^t \frac{(\Lambda_R(t) - \Lambda_R(s))^{j-1}}{(j-1)!} \kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^t A_R \, d\sigma} \, ds.$$

This upper bound results, similarly [9, p.122], from Proposition 2.18 and assumption (vi) of our theorem by means of induction w.r.t. j and Lemma 5.1.

STEP 4 We claim that $(x_j(\cdot))_{j \in \mathbb{N}}$ is a uniform Cauchy sequence. Thus, it tends uniformly to some absolutely continuous $x : [0, T] \longrightarrow (E, d)$ and $\sup_{[0, T]} \|x(\cdot)\| \leq R$.

The a priori estimate in STEP 3 guarantees for every $k \in \mathbb{N}$

$$\sum_{j=1}^k \sup_{t \in [0, T]} d(x_j(t), x_{j+1}(t)) \leq \int_0^T \left(\kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^T (A_R + \kappa \lambda_R) \, d\sigma} \right) ds$$

and thus, $(x_j(\cdot))_{j \in \mathbb{N}}$ is a uniform Cauchy sequence in $C^0([0, T], (E_R, d))$. Its limit $x : [0, T] \longrightarrow (E_R, d)$ satisfies for all $t_1, t_2 \in [0, T]$ ($t_1 \leq t_2$),

$$d(x(t_1), x(t_2)) = \lim_{j \rightarrow \infty} d(x_j(t_1), x_j(t_2)) \leq \int_{t_1}^{t_2} B_R(s) \, ds.$$

STEP 5 We claim that $(\theta_j(\cdot))_{j \in \mathbb{N}}$ converges to some $\theta : [0, T] \longrightarrow (\Theta, D_R)$ almost everywhere in $[0, T]$ and $\theta(t) \in \mathfrak{F}(t, x(t))$ for a.e. $t \in [0, T]$.

By STEP 2 for a.e. $t \in [0, T]$ and all $j \in \mathbb{N}$, $D_R(\theta_{j-1}(t), \theta_j(t)) \leq \kappa \lambda_R(t) \cdot d(x_{j-1}(t), x_j(t))$ and so, all further conclusions in this step are drawn for a.e. $t \in [0, T]$ (without indicating it explicitly). From STEP 3 we obtain

$$\sum_{j=2}^{\infty} D_R(\theta_{j-1}(t), \theta_j(t)) \leq \kappa \lambda_R(t) \int_0^t \kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^t (A_R + \kappa \lambda_R) \, d\sigma} \, ds.$$

Hence, $(\theta_j(t))_{j \in \mathbb{N}}$ is a Cauchy sequence in Θ w.r.t. D_R for a.e. $t \in [0, T]$. Moreover, the construction of $\theta_j(t) \in \mathfrak{F}(t, x_j(t))$ and assumption (vi) provide in terms of the excess $h_R^\#$

$$\text{dist}_R(\iota_R \theta_j(t), \mathfrak{F}_R(t, x(t))) \leq h_R^\#(\mathfrak{F}_R(t, x_j(t)), \mathfrak{F}_R(t, x(t))) \leq \lambda_R(t) \cdot d(x_j(t), x(t)) \xrightarrow{j \rightarrow \infty} 0.$$

According to assumption (iv), $\mathfrak{F}_R(t, x(t)) \subset \Theta_R$ is complete w.r.t. D_R and thus, there exists $\bar{\theta}(t) \in \mathfrak{F}_R(t, x(t))$ with $D_R(\iota_R \theta_j(t), \bar{\theta}(t)) \longrightarrow 0$ ($j \longrightarrow \infty$). Let $\theta(t) \in \mathfrak{F}(t, x(t)) \subset \Theta$ (instead of Θ_R) be such that its restriction (to $[0, 1] \times E_R$) is $\bar{\theta}(t)$. Then, $D_R(\theta_j(t), \theta(t)) \longrightarrow 0$ ($j \longrightarrow \infty$).

STEP 6 We show here that $x(\cdot)$ is a mutational primitive of $\theta(\cdot)$ and thus, it is a solution of the mutational inclusion $\dot{x}(\cdot) \cap \mathfrak{F}(\cdot, x(\cdot)) \neq \emptyset$ (a.e. in $[0, T]$).

This results from Corollary 2.19. In general, every pointwise limit of measurable functions $[0, T] \longrightarrow (\Theta, D_R)$ is measurable (since the proof of [20, Proposition 4.2.2], for example, can be easily adapted to functions with values in a pseudo-metric space). Thus, $\theta : [0, T] \longrightarrow (\Theta, D_R)$ is measurable.

Finally, the inequalities in STEP 5 guarantee for every index $k \in \mathbb{N}$

$$D_R(\theta_k(t), \theta(t)) \leq \sum_{j=k+1}^{\infty} D_R(\theta_{j-1}(t), \theta_j(t)) \leq \kappa \lambda_R(t) \cdot \int_0^t \kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^t (A_R + \kappa \lambda_R) \, d\sigma} \, ds.$$

Due to assumptions about $A_R, \lambda_R \in L^1([0, T])$, the last line provides a dominating function of t in $L^1([0, T])$ as required for verifying all the assumptions of Corollary 2.19.

STEP 7 We next check that $x(\cdot)$ satisfies $d(x(t), y(t)) \leq \eta_\kappa(t)$ for all $t \in [0, T]$.

The construction of $x(\cdot)$ as the uniform limit of $(x_j(\cdot))_{j \in \mathbb{N}}$ and the a priori estimate of STEP 3 (with the same arguments as in STEP 4) lead to

$$\begin{aligned} d(x(t), y(t)) &\leq \lim_{k \rightarrow \infty} d(x_k(t), x_1(t)) + d(x_1(t), y(t)) \leq \sum_{j=1}^{\infty} d(x_j(t), x_{j+1}(t)) + d(x_1(t), y(t)) \\ &\leq \int_0^t e^{\Lambda_R(t) - \Lambda_R(s)} \kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^t A_R d\sigma} ds + d(x_1(t), y(t)) \\ &= \int_0^t \kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^t (A_R + \kappa \lambda_R) d\sigma} ds + d(x_1(t), y(t)) \end{aligned}$$

for every $t \in [0, T]$. We conclude from the choice of $\theta_0(\cdot)$, $\tau_0(\cdot)$ in STEP 2 and Proposition 2.18

$$\begin{aligned} d(y(t), x_1(t)) &\leq d(y(0), x_0) \cdot e^{\int_0^t A_R d\sigma} + \int_0^t D_R(\tau_0(s), \theta_0(s)) \cdot e^{\int_s^t A_R d\sigma} ds \\ &\leq d(y(0), x_0) \cdot e^{\int_0^t A_R d\sigma} + \int_0^t \kappa \cdot \psi(s) \cdot e^{\int_s^t A_R d\sigma} ds. \end{aligned}$$

Combining the two inequalities we get

$$\begin{aligned} &\int_0^t \kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^t (A_R + \kappa \lambda_R) d\sigma} ds \\ &\leq \int_0^t \kappa \lambda_R(s) \cdot \left(d(y(0), x_0) \cdot e^{\int_0^s A_R d\rho} + \int_0^s \kappa \cdot \psi(\sigma) \cdot e^{\int_\sigma^s A_R d\rho} d\sigma \right) \cdot e^{\int_s^t (A_R + \kappa \lambda_R) d\rho} ds \\ &\leq d(y(0), x_0) \cdot e^{\int_0^t A_R d\rho} \cdot \int_0^t \kappa \lambda_R(s) \cdot e^{\int_s^t \kappa \lambda_R d\rho} d\sigma + \int_0^t \kappa \lambda_R(s) e^{\int_s^t \kappa \lambda_R d\rho} \left(\int_0^s \kappa \cdot \psi(\sigma) \cdot e^{\int_\sigma^s A_R d\rho} d\sigma \right) ds \\ &= d(y(0), x_0) \cdot e^{\int_0^t A_R d\rho} \cdot \int_0^t \kappa \lambda_R(s) \cdot e^{\int_s^t \kappa \lambda_R d\rho} d\sigma + \int_0^t \kappa \cdot \psi(s) \cdot e^{\int_s^t A_R d\rho} \left(\int_s^t \kappa \lambda_R(\sigma) e^{\int_\sigma^t \kappa \lambda_R d\rho} d\sigma \right) ds \\ &= d(y(0), x_0) \cdot e^{\int_0^t A_R d\rho} \cdot \left(e^{\int_0^t \kappa \lambda_R d\rho} - 1 \right) + \int_0^t \kappa \cdot \psi(s) \cdot e^{\int_s^t A_R d\rho} \left(e^{\int_s^t \kappa \lambda_R d\rho} - 1 \right) ds. \end{aligned}$$

STEP 8 We next claim that for a.e. $t \in [0, T]$, $D_R(\tau_0(t), \theta(t)) \leq \kappa \psi(t) + \kappa \lambda_R(t) \cdot \eta_\kappa(t)$.

In STEP 2 and STEP 3, we have concluded that for a.e. $t \in [0, T]$ and all $j \in \mathbb{N}$,

$$D_R(\tau_0(t), \theta_0(t)) \leq \kappa \psi(t), \quad D_R(\theta_{j-1}(t), \theta_j(t)) \leq \kappa \lambda_R(t) \cdot d(x_{j-1}(t), x_j(t)).$$

Similarly to STEP 7, it implies that for a.e. $t \in [0, T]$

$$\begin{aligned} D_R(\tau_0(t), \theta(t)) &\leq D_R(\tau_0(t), \theta_0(t)) + \sum_{j=1}^{\infty} D_R(\theta_{j-1}(t), \theta_j(t)) \\ &\leq D_R(\tau_0(t), \theta_0(t)) + \kappa \lambda_R(t) \cdot \sum_{j=1}^{\infty} d(x_{j-1}(t), x_j(t)) \\ &\leq \kappa \psi(t) + \kappa \lambda_R(t) \cdot \int_0^t \kappa \lambda_R(s) \cdot d(y(s), x_1(s)) \cdot e^{\int_s^t (A_R + \kappa \lambda_R) d\sigma} ds. \end{aligned}$$

STEP 7 contains an upper bound of the last integral and so, we obtain

$$\begin{aligned} D_R(\tau_0(t), \theta(t)) &\leq \kappa \psi(t) + \kappa \lambda_R(t) \cdot d(y(0), x_0) \cdot e^{\int_0^t A_R d\rho} \left(e^{\int_0^t \kappa \lambda_R d\rho} - 1 \right) + \\ &\quad + \kappa \lambda_R(t) \cdot \int_0^t \kappa \psi(s) \cdot e^{\int_s^t A_R d\rho} \left(e^{\int_s^t \kappa \lambda_R d\rho} - 1 \right) ds \\ &\leq \kappa \psi(t) + \kappa \lambda_R(t) \cdot \eta_\kappa(t). \end{aligned}$$

□

6 Example: Measure-Valued Solutions to Transport Inclusions

Below $C_c^0(\mathbb{R}^n)$ denotes the space of continuous functions $\mathbb{R}^n \rightarrow \mathbb{R}$ with compact support and $C_0^0(\mathbb{R}^n)$ its closure with respect to the supremum norm. Let $\mathcal{M}(\mathbb{R}^n)$ be the set of all finite real-valued Radon measures on \mathbb{R}^n , i.e., it is the dual space of $(C_0^0(\mathbb{R}^n), \|\cdot\|_{L^\infty})$ [2, Remark 1.57], $\mathcal{M}_c(\mathbb{R}^n)$ denotes the subset of measures with compact support and $\mathcal{M}^+(\mathbb{R}^n)$ denotes the subset of measures $\mu \geq 0$. For each $\mu \in \mathcal{M}(\mathbb{R}^n)$, $|\mu|(\mathbb{R}^n)$ abbreviates the total variation of μ . Recall that a nonempty subset $\mathcal{S} \subset \mathcal{M}(\mathbb{R}^n)$ is called (*uniformly*) *tight* if for all $\varepsilon > 0$ there is a compact $K_\varepsilon \subset \mathbb{R}^n$ with $\sup_{\mu \in \mathcal{S}} |\mu|(\mathbb{R}^n \setminus K_\varepsilon) < \varepsilon$. A sequence $(\mu_n)_{n \in \mathbb{N}}$ in $\mathcal{M}(\mathbb{R}^n)$ is said to *converge narrowly* to $\mu \in \mathcal{M}(\mathbb{R}^n)$ if for every bounded $\varphi \in C^0(\mathbb{R}^n)$, $\lim_{n \rightarrow \infty} \int_{\mathbb{R}^n} \varphi d\mu_n = \int_{\mathbb{R}^n} \varphi d\mu$.

Definition 6.1 ([13]) *The mapping $d_{\mathcal{M}} : \mathcal{M}(\mathbb{R}^n) \times \mathcal{M}(\mathbb{R}^n) \rightarrow [0, \infty)$ defined by*

$$d_{\mathcal{M}}(\mu, \nu) := \sup \left\{ \int_{\mathbb{R}^n} \psi d(\mu - \nu) \mid \psi \in C^1(\mathbb{R}^n), \|\psi\|_{L^\infty} \leq 1, \|\nabla \psi\|_{L^\infty} \leq 1 \right\}$$

is called the Kantorovich-Rubinstein metric on $\mathcal{M}(\mathbb{R}^n)$.

Whenever μ and ν are Borel probability measures on \mathbb{R}^n (and thus, $\mu, \nu \in \mathcal{M}^+(\mathbb{R}^n)$), $d_{\mathcal{M}}(\mu, \nu)$ coincides with their so-called Wasserstein distance of order $p = 1$ (e.g., [13, Theorem 8.10.45]). Recall that every narrowly continuous function $[0, T] \rightarrow \mathcal{M}(\mathbb{R}^n)$ with tight values has a bounded total variation.

Definition 6.2 *Let $T > 0$, a metric space $U \neq \emptyset$ and functions $f : [0, T] \times \mathcal{M}(\mathbb{R}^n) \times U \rightarrow W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$ and $g : [0, T] \times \mathcal{M}(\mathbb{R}^n) \times U \rightarrow W^{1,\infty}(\mathbb{R}^n)$ be given. Consider a measurable $u : [0, T] \rightarrow U$ (called control). Then $\mu : [0, T] \rightarrow \mathcal{M}(\mathbb{R}^n)$ written as $t \mapsto \mu_t$ is called a weak solution to the transport equation*

$$\partial_t \mu = -\operatorname{div}_{\mathbf{x}}(f(t, \mu_t, u(t)) \mu) + g(t, \mu_t, u(t)) \cdot \mu \quad (3)$$

on $[0, T]$ if μ is narrowly continuous and for any $0 \leq t_1 < t_2 \leq T$ and $\varphi \in C_c^1(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \varphi d\mu_{t_2} - \int_{\mathbb{R}^n} \varphi d\mu_{t_1} = \int_{t_1}^{t_2} \int_{\mathbb{R}^n} (f(s, \mu_s, u(s), \mathbf{x}) \cdot \nabla_{\mathbf{x}} \varphi(\mathbf{x}) + g(s, \mu_s, u(s), \mathbf{x}) \varphi(\mathbf{x})) d\mu_s(\mathbf{x}) ds.$$

Let S_k be a closed separable subset of $W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^k)$ for $k \in \{1, n\}$, (U, d_U) a compact metric space of controls and

$$f : [0, T] \times \mathcal{M}(\mathbb{R}^n) \times U \rightarrow S_n, \quad g : [0, T] \times \mathcal{M}(\mathbb{R}^n) \times U \rightarrow S_1$$

have the following properties:

- $\widehat{C}_{f,\infty} := \sup \{ \|f(t, \zeta, u)\|_{L^\infty(\mathbb{R}^n, \mathbb{R}^n)} \mid t \in [0, T], \zeta \in \mathcal{M}(\mathbb{R}^n), u \in U \} < \infty$ and $\widehat{C}_{g,\infty} := \sup \{ \|g(t, \zeta, u)\|_{L^\infty(\mathbb{R}^n)} \mid t \in [0, T], \zeta \in \mathcal{M}(\mathbb{R}^n), u \in U \} < \infty$.
- For every $r > 0$, there exists $C_r \in [0, \infty)$ such that for all $t \in [0, T]$, $u \in U$ and $\zeta \in \mathcal{M}(\mathbb{R}^n)$ with $|\zeta|(\mathbb{R}^n) \leq r$, $\|\partial_{\mathbf{x}} f(t, \zeta, u)\|_{L^\infty(\mathbb{R}^n, \mathbb{R}^{n \times n})}$, $\|\nabla_{\mathbf{x}} g(t, \zeta, u)\|_{L^\infty(\mathbb{R}^n, \mathbb{R}^n)} \leq C_r$.
- For all $\zeta \in \mathcal{M}(\mathbb{R}^n)$ and $u \in U$, $f(\cdot, \zeta, u), g(\cdot, \zeta, u) : [0, T] \rightarrow (L^\infty, \|\cdot\|_{L^\infty})$ are measurable.
- For every $r > 0$, there is $\Lambda_r \in L^1([0, T])$ such that for each $u \in U$ and a.e. $t \in [0, T]$, $f(t, \cdot, u)$ and $g(t, \cdot, u)$ are $\Lambda_r(t)$ -Lipschitz on $\{\mu \in \mathcal{M}(\mathbb{R}^n) \mid |\mu|(\mathbb{R}^n) \leq r\}$ w.r.t. $d_{\mathcal{M}}, \|\cdot\|_{L^\infty}$.
- For a.e. $t \in [0, T]$ and each $\zeta \in \mathcal{M}(\mathbb{R}^n)$, $f(t, \zeta, \cdot), g(t, \zeta, \cdot) : U \rightarrow (L^\infty, \|\cdot\|_{L^\infty})$ are continuous.

Then Theorem 3.4 leads to the following statement:

Theorem 6.3 Consider an absolutely continuous curve $\eta : [0, T] \longrightarrow (\mathcal{M}(\mathbb{R}^n), d_{\mathcal{M}})$ and measurable $b_\eta : [0, T] \longrightarrow (S_n, \|\cdot\|_{L^\infty})$, $c_\eta : [0, T] \longrightarrow (S_1, \|\cdot\|_{L^\infty})$. Suppose

- (i) $\eta([0, T]) \subset \mathcal{M}(\mathbb{R}^n)$ is tight and bounded w.r.t. the total variation.
- (ii) η is a weak solution of $\partial_t \eta = -\operatorname{div}_{\mathbf{x}}(b_\eta(t) \eta) + c_\eta(t) \cdot \eta$ in $[0, T]$.
- (iii) There is $\Psi \in L^1([0, T])$ satisfying for a.e. $t \in [0, T]$,

$$\inf_{u \in U} \left(\|b_\eta(t) - f(t, \eta_t, u)\|_{L^\infty(\mathbb{R}^n, \mathbb{R}^n)} + \|c_\eta(t) - g(t, \eta_t, u)\|_{L^\infty(\mathbb{R}^n)} \right) < \Psi(t).$$

Then for every $\kappa > 1$ and $\mu_0 \in \mathcal{M}(\mathbb{R}^n)$, there exist a control $u(\cdot)$ and a weak solution μ of the transport equation (3) such that

- (1.) μ is Lipschitz w.r.t. $d_{\mathcal{M}}$ and has tight values with $|\mu_t|(\mathbb{R}^n) \leq (|\mu_0| + \widehat{C}_{g, \infty} T) \cdot e^{\widehat{C}_{g, \infty} T}$, $\mu(0) = \mu_0$.
- (2.) $d_{\mathcal{M}}(\mu_t, \eta_t) \leq \delta_R(t)$, $\forall t \in [0, T]$, where $R := \sup \{ (|\mu_0| + \widehat{C}_{g, \infty} T) \cdot e^{\widehat{C}_{g, \infty} T}, |\eta_t|(\mathbb{R}^n) \mid t \in [0, T] \} + 1$,

$$\delta_R(t) := d_{\mathcal{M}}(\mu_0, \eta_0) \cdot e^{\int_0^t (\widehat{C}_{g, \infty} + 2 C_R + 2 \kappa R \Lambda_R(\sigma)) d\sigma} + \int_0^t \kappa R \cdot \Psi(s) \cdot e^{\int_s^t (\widehat{C}_{g, \infty} + 2 C_R + 2 \kappa R \Lambda_R(\sigma)) d\sigma} ds.$$
- (3.) For a.e. $t \in [0, T]$,

$$\|f(t, \mu_t, u(t)) - b_\eta(t)\|_{L^\infty(\mathbb{R}^n, \mathbb{R}^n)} + \|g(t, \mu_t, u(t)) - c_\eta(t)\|_{L^\infty(\mathbb{R}^n)} \leq \kappa R \cdot (2 \Lambda_R(t) \delta_R(t) + \Psi(t)).$$
- (4.) Whenever $\mu_0 \in \mathcal{M}^+(\mathbb{R}^n)$, we have $\mu_t \in \mathcal{M}^+(\mathbb{R}^n)$ for every $t \in [0, T]$.
- (5.) If $\mu_0 \in \mathcal{M}(\mathbb{R}^n)$ has compact support in $\overline{\mathbb{B}}_\rho \subset \mathbb{R}^n$, then $\mu_t \in \mathcal{M}(\mathbb{R}^n)$ has its support in $\overline{\mathbb{B}}_{\rho + \widehat{C}_{f, \infty} t}$ for every $t \in [0, T]$.
- (6.) If, in addition, $g \equiv 0 \in C^0(\mathbb{R}^n)$, then $|\mu_t|(\mathbb{R}^n)$ is constant in t .

To link the above result with more general Theorem 3.4 we introduce the corresponding transitions. But first we cite several (rather technical) results about Radon measures and measure-valued solutions to linear transport equations from [42, § 2.5], [34, §§ 6,7] which serve as tools here.

Lemma 6.4 ([34, Proposition 40])

- (1.) For any tight sequence $(\mu_n)_{n \in \mathbb{N}}$ and μ in $\mathcal{M}(\mathbb{R}^n)$, the following equivalences hold

$$\lim_{n \rightarrow \infty} d_{\mathcal{M}}(\mu_n, \mu) = 0 \ \& \ \sup_{n \in \mathbb{N}} |\mu_n|(\mathbb{R}^n) < \infty \iff \mu_n \rightarrow \mu \text{ weak}^* \iff \mu_n \rightarrow \mu \text{ narrowly } (n \rightarrow \infty).$$

- (2.) The set $\{\mu \in \mathcal{M}(\mathbb{R}^n) \mid |\mu|(\mathbb{R}^n) \leq r\}$ is complete w.r.t. $d_{\mathcal{M}}$ for any $r > 0$.
- (3.) Every tight set $\mathcal{S} \subset \mathcal{M}(\mathbb{R}^n)$ with $\sup_{\mu \in \mathcal{S}} |\mu|(\mathbb{R}^n) < \infty$ is relatively compact w.r.t. $d_{\mathcal{M}}$.

Due to statement (3.) of Lemma 6.4, $\{\mu \in \mathcal{M}(\mathbb{R}^n) \mid |\mu|(\mathbb{R}^n) \leq r\}$ is separable w.r.t. $d_{\mathcal{M}}$ for each $r > 0$ because $\mathcal{M}_c(\mathbb{R}^n)$ is dense in $(\mathcal{M}(\mathbb{R}^n), d_{\mathcal{M}})$ and for each $m \in \mathbb{N}$, $\mathcal{M}(\overline{\mathbb{B}}_m)$ is tight. Hence, $(\mathcal{M}(\mathbb{R}^n), d_{\mathcal{M}}, |\cdot|(\mathbb{R}^n))$ fulfills the hypothesis (HΘ) (i), (ii).

For every Lipschitz $\mathbf{b} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, let $\mathbf{X}_{\mathbf{b}}(\cdot, \mathbf{x}_0) : [0, T] \longrightarrow \mathbb{R}^n$ denote the unique solution of $\mathbf{x}' = \mathbf{b}(\mathbf{x})$ in $[0, T]$ with the initial state $\mathbf{x}_0 \in \mathbb{R}^n$.

Lemma 6.5 ([34, Propositions 43 and 44 (2.)]) *For any $\mathbf{b} \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$, $c \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R})$ and $\nu_0 \in \mathcal{M}(\mathbb{R}^n)$, there exists a unique weak solution $\mu : [0, T] \rightarrow \mathcal{M}(\mathbb{R}^n)$ written as $t \mapsto \mu_t$ to the linear problem*

$$\partial_t \mu_t + \operatorname{div}_{\mathbf{x}}(\mathbf{b} \mu_t) = c \mu_t \quad \text{in } [0, T], \quad \mu_0 = \nu_0. \quad (4)$$

Moreover for all $\varphi \in C_c^1(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \varphi d\mu_t = \int_{\mathbb{R}^n} \varphi(\mathbf{X}_{\mathbf{b}}(t, \mathbf{x})) \cdot \exp\left(\int_0^t c(\mathbf{X}_{\mathbf{b}}(s, \mathbf{x})) ds\right) d\nu_0(\mathbf{x}).$$

From now on, we use the notation $\vartheta_{\mathbf{b},c}(t, \nu_0) := \mu_t$ for $t \in [0, 1]$ with the weak solution μ to initial value problem (4). It is worth mentioning that $\nu_0 \in \mathcal{M}^+(\mathbb{R}^n)$ implies $\vartheta_{\mathbf{b},c}(t, \nu_0) \in \mathcal{M}^+(\mathbb{R}^n)$ for every $t \in [0, 1]$.

Lemma 6.6 ([34, Proposition 46 and its remark]) *Let $\mathbf{b}, \bar{\mathbf{b}} \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$ and $c, \bar{c} \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R})$. The following statements hold for any $\mu_0, \nu_0 \in \mathcal{M}(\mathbb{R}^n)$ and $t, h \in [0, 1]$ with $t + h \leq 1$*

- (1.) $\vartheta_{\mathbf{b},c}(0, \mu_0) = \mu_0$, $\vartheta_{\mathbf{b},c}(h, \vartheta_{\mathbf{b},c}(t, \mu_0)) = \vartheta_{\mathbf{b},c}(t + h, \mu_0)$, $|\vartheta_{\mathbf{b},c}(h, \mu_0)|(\mathbb{R}^n) \leq e^{\|c\|_{L^\infty} h} \cdot |\mu_0|(\mathbb{R}^n)$
- (2.) $d_{\mathcal{M}}(\vartheta_{\mathbf{b},c}(t, \mu_0), \vartheta_{\mathbf{b},c}(t + h, \mu_0)) \leq h (\|\mathbf{b}\|_{L^\infty} + \|c\|_{L^\infty}) e^{\|c\|_{L^\infty}} \cdot |\mu_0|(\mathbb{R}^n)$
- (3.) $d_{\mathcal{M}}(\vartheta_{\mathbf{b},c}(h, \mu_0), \vartheta_{\mathbf{b},c}(h, \nu_0)) \leq d_{\mathcal{M}}(\mu_0, \nu_0) \cdot e^{(\operatorname{Lip} \mathbf{b} + \|c\|_{W^{1,\infty}}) h}$
- (4.) $d_{\mathcal{M}}(\vartheta_{\mathbf{b},c}(h, \mu_0), \vartheta_{\bar{\mathbf{b}},\bar{c}}(h, \mu_0))$
 $\leq (\|\mathbf{b} - \bar{\mathbf{b}}\|_{L^\infty} \cdot e^{h \|\nabla c\|_{L^\infty}} + \|c - \bar{c}\|_{L^\infty}) h e^{h \cdot (\operatorname{Lip} \mathbf{b} + \max\{\|c\|_{L^\infty}, \|\bar{c}\|_{L^\infty}\})} \cdot |\mu_0|(\mathbb{R}^n).$

Now Definitions 2.1 and 2.5 lead directly to the following conclusions about $\vartheta_{\mathbf{b},c}$.

Corollary 6.7 *For every $\mathbf{b} \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$ and $c \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R})$, the function $\vartheta_{\mathbf{b},c} : [0, 1] \times \mathcal{M}(\mathbb{R}^n) \rightarrow \mathcal{M}(\mathbb{R}^n)$ is a transition on $(\mathcal{M}(\mathbb{R}^n), d_{\mathcal{M}}, |\cdot|(\mathbb{R}^n))$ with the parameters for $r \geq 0$*

$$\alpha_r(\vartheta_{\mathbf{b},c}) := \operatorname{Lip} \mathbf{b} + \|c\|_{W^{1,\infty}}, \quad \beta_r(\vartheta_{\mathbf{b},c}) := r e^{\|c\|_{L^\infty}} \cdot (\|\mathbf{b}\|_{L^\infty} + \|c\|_{L^\infty}), \quad \gamma(\vartheta_{\mathbf{b},c}) := \|c\|_{L^\infty}.$$

The related pseudo-metric D_r satisfies for all $\mathbf{b}, \bar{\mathbf{b}} \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^n)$, $c, \bar{c} \in W^{1,\infty}(\mathbb{R}^n, \mathbb{R})$ and $r \geq 0$

$$D_r(\vartheta_{\mathbf{b},c}, \vartheta_{\bar{\mathbf{b}},\bar{c}}) \leq r \cdot (\|\mathbf{b} - \bar{\mathbf{b}}\|_{L^\infty(\mathbb{R}^n, \mathbb{R}^n)} + \|c - \bar{c}\|_{L^\infty(\mathbb{R}^n, \mathbb{R})}).$$

This implies that $\Theta := \{\vartheta_{\mathbf{b},c} \mid \mathbf{b} \in S_n, c \in S_1\}$ supplied with $D_{\text{loc}}(\vartheta_{\mathbf{b},c}, \vartheta_{\bar{\mathbf{b}},\bar{c}}) := \|\mathbf{b} - \bar{\mathbf{b}}\|_{W^{1,\infty}} + \|c - \bar{c}\|_{W^{1,\infty}}$ satisfies hypothesis $(\mathbb{H}\Theta)$ (iv).

The mutational Filippov Theorem 3.4 yields the following existence result.

Lemma 6.8 *Under all the assumptions of Theorem 6.3, for every $\kappa > 1$ and $\mu_0 \in \mathcal{M}(\mathbb{R}^n)$, there exist $\mu : [0, T] \rightarrow \mathcal{M}(\mathbb{R}^n)$ with $\mu(0) = \mu_0$ and measurable $(b_\mu, c_\mu) : [0, T] \rightarrow L^\infty(\mathbb{R}^n, \mathbb{R}^n) \times L^\infty(\mathbb{R}^n, \mathbb{R})$ such that for $R, \delta_R(t)$ defined as in Theorem 6.3 the inequality (2.) of Theorem 6.3 holds true and*

- (1.) μ is Lipschitz w.r.t. $d_{\mathcal{M}}$, $|\mu_t|(\mathbb{R}^n) \leq (|\mu_0| + \widehat{C}_{g,\infty} T) \cdot e^{\widehat{C}_{g,\infty} T} < R$ for all t .
- (2.) For a.e. $t \in [0, T]$, $(b_\mu(t), c_\mu(t)) \in \{(f, g)(t, \mu_t, u) \mid u \in U\} \subset S_n \times S_1$.
- (3.) For a.e. $t \in [0, T]$, $\lim_{h \downarrow 0} \frac{1}{h} \cdot d_{\mathcal{M}}(\mu_{t+h}, \vartheta_{b_\mu(t), c_\mu(t)}(h, \mu_t)) = 0$.
- (4.) For a.e. $t \in [0, T]$, $\|b_\mu(t) - b_\eta(t)\|_{L^\infty} + \|c_\mu(t) - c_\eta(t)\|_{L^\infty} \leq \kappa R \cdot (2 \Lambda_R(t) \delta_R(t) + \Psi(t)).$

A measurable selection theorem provides some measurable $u : [0, T] \longrightarrow U$ satisfying for a.e. $t \in [0, T]$ $b_\mu(t) = f(t, \mu_t, u(t))$, $c_\mu(t) = g(t, \mu_t, u(t))$ (see, e.g., [10, Theorem 8.2.10]). In regard to Theorem 6.3, it remains to prove that $\mu : [0, T] \longrightarrow \mathcal{M}(\mathbb{R}^n)$ is a weak solution of $\partial_t \mu = -\operatorname{div}_{\mathbf{x}}(f(t, \mu_t, u(t)) \mu) + g(t, \mu_t, u(t)) \cdot \mu$ in $[0, T]$ with tight values and the supplementary properties (4.) – (6.) of Theorem 6.3. Having specified measurable b_μ , c_μ , however, we can use some results about weak solutions to nonautonomous linear transport equations. Results of [34, Section 3 and 6] imply the following statement.

Lemma 6.9 *There exists a unique narrowly continuous weak solution $\tilde{\mu} : [0, T] \longrightarrow \mathcal{M}(\mathbb{R}^n)$ of the nonautonomous linear equation $\partial_t \tilde{\mu}_t + \operatorname{div}_{\mathbf{x}}(b_\mu(t) \tilde{\mu}_t) = c_\mu(t) \tilde{\mu}_t$ with $\tilde{\mu}_0 = \mu_0$ and, it satisfies :*

- (1.) *$\tilde{\mu}$ is Lipschitz w.r.t. $d_{\mathcal{M}}$ and has tight values with $|\tilde{\mu}_t|(\mathbb{R}^n) \leq (|\mu_0| + \widehat{C}_{g,\infty} T) \cdot e^{\widehat{C}_{g,\infty} T} < R$ for all t .*
- (2.) *For a.e. $t \in [0, T]$, $\lim_{h \downarrow 0} \frac{1}{h} \cdot d_{\mathcal{M}}(\tilde{\mu}_{t+h}, \vartheta_{b_\mu(t), c_\mu(t)}(h, \tilde{\mu}_t)) = 0$.*

In addition, we conclude from Lemma 6.5 (by means of Euler approximations) that $\tilde{\mu}_t$ is characterized by

$$\int_{\mathbb{R}^n} \varphi d\tilde{\mu}_t = \int_{\mathbb{R}^n} \varphi(\mathbf{X}_{b_\mu}(t, \mathbf{x})) \cdot \exp\left(\int_0^t c_\mu(\mathbf{X}_{b_\mu}(s, \mathbf{x})) ds\right) d\mu_0(\mathbf{x})$$

for all $\varphi \in C_c^1(\mathbb{R}^n)$ with $\mathbf{X}_{b_\mu}(\cdot, \mathbf{x}_0) : [0, T] \longrightarrow \mathbb{R}^n$ denoting the unique Carathéodory solution of $\mathbf{x}' = b_\mu(t, \mathbf{x})$ in $[0, T]$ with $\mathbf{X}_{b_\mu}(0, \mathbf{x}_0) = \mathbf{x}_0 \in \mathbb{R}^n$. It guarantees the properties (4.) – (6.) of Theorem 6.3 for $\tilde{\mu}_t$. Finally, the statements (3.) of Lemma 6.8 and (2.) of Lemma 6.9 lead to

$$\lim_{h \downarrow 0} \frac{1}{h} \cdot (d_{\mathcal{M}}(\mu_{t+h}, \tilde{\mu}_{t+h}) - d_{\mathcal{M}}(\mu_t, \tilde{\mu}_t)) \leq (\widehat{C}_{g,\infty} + 2C_R) \cdot d_{\mathcal{M}}(\mu_t, \tilde{\mu}_t) \quad \text{for a.e. } t \in [0, T],$$

and, since $d_{\mathcal{M}}(\mu_0, \tilde{\mu}_0) = 0$, Gronwall's inequality ensures $\mu_t = \tilde{\mu}_t$ for every $t \in [0, T]$.

Appendix. Tools in Topology and Non-smooth Analysis

A Gronwall's Inequality for Lower Semicontinuous Functions

Proposition A.1 *Let $\psi : [a, b] \longrightarrow [0, \infty)$ be lower semicontinuous and $f, g \in C^0([a, b])$ satisfy $f(\cdot) \geq 0$ and*

$$\liminf_{h \downarrow 0} \frac{\psi(t+h) - \psi(t)}{h} \leq f(t) \cdot \psi(t) + g(t) \quad \text{for every } t \in [a, b].$$

Then, for every $t \in [a, b]$ and $\ell(t) := \int_a^t f(s) ds$ we have

$$\psi(t) \leq \psi(a) \cdot e^{\ell(t)} + \int_a^t e^{\ell(t)-\ell(s)} g(s) ds.$$

Proof. Without any loss of generality we assume that $a = 0$. Consider the closed set $K := \operatorname{epi}(\psi) \cup ([b, \infty) \times \mathbb{R})$, where $\operatorname{epi}(\psi)$ denotes the epigraph of ψ and define $f(t) = f(b)$, $g(t) = g(b)$ for $t > b$. It follows from our assumption, that for any $s \in [0, b)$ the vector $(1, f(s) \cdot \psi(s) + g(s)) \in \mathbb{R}^2$ belongs to the contingent cone $T_K(s, \psi(s))$ to K at $(s, \psi(s))$ implying that for any $y \geq \psi(s)$, $(1, f(s) \cdot \psi(s) + g(s)) \in T_K(s, y)$ and, because $f \geq 0$, that $(1, f(s) \cdot y + g(s)) \in T_K(s, y)$. Furthermore, for any $(s, y) \in K$ with $s \geq b$ we have $\{1\} \times \mathbb{R} \subset T_K(s, y)$. By the viability theorem, see for instance, [9, p.180] there exists a solution to the system

$$\begin{cases} s' = 1, & s(0) = 0 \\ y' = f(s)y + g(s), & y(0) = \psi(0) \end{cases}$$

satisfying $(s(t), y(t)) \in K$ for all $t \in [0, T]$ for some $0 < T < b$. Observing that $s(t) = t$ we deduce that y is uniquely defined on $[0, \infty)$. Using that ψ is lower semicontinuous we obtain $\psi(t) \leq y(t)$ for all $t \in [0, T]$. From the viability theorem and the Zorn lemma it follows, in the usual way, that $y(\cdot)$ satisfies the inequality $\psi(\cdot) \leq y(\cdot)$ on $[0, b]$. Since $y(t) = \psi(0) \cdot e^{\ell(t)} + \int_0^t e^{\ell(t)-\ell(s)} g(s) ds$ the proof follows. \square

B Measurability of Set-Valued Maps in Topological Spaces

The example of transitions given as a function of time requests extending standard results about measurable set-valued maps (see, e.g., [10, Ch. 8], [31, Ch. 2]) to mappings with values in a *pseudo-metric* space (thus not necessarily Hausdorff).

Proposition B.1 *Let X and Y be topological spaces. If Y is second countable, i.e., Y has a countable base, then $\mathcal{B}(X \times Y) = \mathcal{B}(X) \otimes \mathcal{B}(Y)$.*

Proof. The special case of Hausdorff spaces X, Y is considered in [13, Lemma 6.4.2], but the additional assumption is not used in its proof. \square

Definition B.2 *Consider a measurable space (Ω, \mathcal{A}) , a topological space X and a set-valued map $F : \Omega \rightsquigarrow X$.*

(a) *F is called measurable if the inverse image of each open set is measurable, i.e., for every open $O \subset X$,*

$$F^{-1}(O) \stackrel{\text{Def.}}{=} \{\omega \in \Omega \mid F(\omega) \cap O \neq \emptyset\} \in \mathcal{A}.$$

(b) *F is called graph measurable if its graph belongs to $\mathcal{A} \otimes \mathcal{B}(X)$.*

We also need the following fact.

Remark B.3 Consider a measurable space (Ω, \mathcal{A}) , a separable pseudo-metric space (Y, d_Y) and a measurable function $f : \Omega \rightarrow Y$. Then there exists a sequence $(f_j)_{j \in \mathbb{N}}$ of simple functions $\Omega \rightarrow Y$ converging to f pointwise in Ω .

Indeed, the identity $Y \rightarrow Y$ is the pointwise limit of a sequence $(g_j)_{j \in \mathbb{N}}$ of Borel measurable functions $g_j : Y \rightarrow Y$ each of which has finitely many values as the proof of [20, Proposition 4.2.6] also holds for a separable pseudo-metric space (Y, d_Y) . Hence, $f_j := g_j \circ f : \Omega \rightarrow Y$ has the claimed properties.

The following two propositions follow from Leese [38–40], see also [55, Theorem 4.2 (b) and Theorem 5.10]. Though originally they were stated on an arbitrary topological space X with a measure defined on a Suslin family of subsets of X , in order to simplify the presentation, we restrict our attention to $X = [0, T]$, with $T > 0$, supplied with the Lebesgue measure, which suffices for our proofs.

Proposition B.4 *Let Y be a topological space having a countable base and $F : [0, T] \rightsquigarrow Y$ be a measurable set-valued map with nonempty closed values. Then F is graph measurable.*

Proposition B.5 *Let Y be a weakly Suslin space and $F : [0, T] \rightsquigarrow Y$ be a graph measurable set-valued map with nonempty values. Then F has a Castaing representation, i.e., there is a sequence $(f_n)_{n \in \mathbb{N}}$ of measurable selections of F such that for each $t \in [0, T]$, $\{f_n(t) \mid n \in \mathbb{N}\}$ is dense in $F(t)$.*

Recall that a topological space is called a weakly Suslin space if it is the image of a complete separable metric space under a continuous mapping [35, 55, 56] and $f : [0, T] \rightarrow Y$ is called a measurable selection of F if $f(t) \in F(t)$ for every $t \in [0, T]$.

We also state the following extension of Lusin's Theorem relating the measurability of a vector-valued function to its almost continuity. It can be proved in the same way as in [13, Theorem 7.14.25], [23], [36, Theorem 4.1], where a metric space Y was considered.

Theorem B.6 *Let Y be a pseudo-metric space. A function $f : [0, T] \rightarrow Y$ is measurable if and only if it is almost continuous, i.e., if for every $\varepsilon > 0$, there is a compact subset $K_\varepsilon \subset [0, T]$ with $\mathcal{L}^1([0, T] \setminus K_\varepsilon) < \varepsilon$ such that f is continuous on K_ε .*

Finally, we formulate a Scorza-Dragoni type result for so-called *measurable/Lipschitz* maps (essentially in the sense of [10, Definition 9.5.1]). For a complete separable metric space X , a set-valued map from $[0, T] \times X$ into a metric space Y is not necessarily Lebesgue-Borel measurable if it is Lebesgue measurable w.r.t. the first variable and upper semicontinuous (or lower semicontinuous) w.r.t. the second variable. Counterexamples are presented in [4], [31, Ch. 2, Example 7.2]. This general observation explains why we prefer the assumption of graph measurability in Proposition 3.2.

The Lipschitz continuity w.r.t. the second variable, however, is sufficient for further conclusions about joint regularity as the next proposition specifies. In comparison with Himmelberg's results about almost semicontinuity [28, 29], we dispense with the Hausdorff condition on the value space Y and so, its proof cannot rely on characterizing elements in terms of distances as usual.

Proposition B.7 *Let X be a complete separable metric space, Y be a weakly Suslin pseudo-metric space and $F : [0, T] \times X \rightsquigarrow Y$ be a set-valued map with nonempty values satisfying the following properties:*

- (i) *for each $x \in X$, the map $F(\cdot, x) : [0, T] \rightsquigarrow Y$ is graph measurable,*
- (ii) *there exists $\lambda \in L^1([0, T])$ such that for a.e. $t \in [0, T]$, $F(t, \cdot) : X \rightsquigarrow Y$ is $\lambda(t)$ -Lipschitz.*

Then for every $\varepsilon > 0$, there exists a closed subset $S_\varepsilon \subset [0, T]$ with $\mathcal{L}^1([0, T] \setminus S_\varepsilon) < \varepsilon$ such that the restriction $F|_{S_\varepsilon \times X}$ is lower semicontinuous.

Consequently, F is Lebesgue-Borel measurable. If in addition all values of F are closed then its graph is contained in $\mathcal{L}([0, T]) \otimes \mathcal{B}(X) \otimes \mathcal{B}(Y)$.

Proof. Let $\{x_1, x_2, \dots\}$ denote a dense subset of X . For each index $k \in \mathbb{N}$, Proposition B.5 guarantees a Castaing representation of $F(\cdot, x_k) : [0, T] \rightsquigarrow Y$, i.e., there exists a sequence $(f_{k,\ell})_{\ell \in \mathbb{N}}$ of Lebesgue measurable functions $[0, T] \rightarrow Y$ such that for every $t \in [0, T]$, the set $\{f_{k,\ell}(t) \mid \ell \in \mathbb{N}\}$ is dense in $F(t, x_k) \subset Y$. Fixing $\varepsilon > 0$ arbitrarily, Lusin-type Theorem B.6 provides a closed subset $S_\varepsilon \subset [0, T]$ (inductively w.r.t. $k, \ell \in \mathbb{N}$) satisfying $\mathcal{L}^1([0, T] \setminus S_\varepsilon) < \varepsilon$ and

- for every $t \in S_\varepsilon$, $F(t, \cdot) : X \rightsquigarrow Y$ is $\lambda(t)$ -Lipschitz,
- $\lambda|_{S_\varepsilon} : S_\varepsilon \rightarrow \mathbb{R}$ is continuous and, for all $k, \ell \in \mathbb{N}$, $(f_{k,\ell})|_{S_\varepsilon} : S_\varepsilon \rightarrow Y$ is continuous.

Then, the restriction $\tilde{F} := F|_{S_\varepsilon \times X} : S_\varepsilon \times X \rightsquigarrow Y$ is lower semicontinuous in the sense that for every open set $O \subset Y$, the inverse image $\tilde{F}^{-1}(O) = \{(t, x) \in S_\varepsilon \times X \mid F(t, x) \cap O \neq \emptyset\}$ is open in $S_\varepsilon \times X$.

Indeed, for each $(t_0, x_0) \in \tilde{F}^{-1}(O)$, choose $y_0 \in F(t_0, x_0) \cap O$ and $\rho > 0$ with $\mathbb{B}_\rho(y_0) \subset O \subset Y$. Set $\Lambda := 1 + \sup_{S_\varepsilon} \lambda(\cdot) < \infty$. There exists $k = k(x_0, \Lambda) \in \mathbb{N}$ with $d_X(x_0, x_k) < \frac{\rho}{8\Lambda}$. Next we select $\ell = \ell(k) \in \mathbb{N}$ with

$$d_Y(f_{k,\ell}(t_0), y_0) \leq h^\sharp(F(t_0, x_k), F(t_0, x_0)) + \frac{\rho}{8} \leq \Lambda \cdot d(x_k, x_0) + \frac{\rho}{8} < \frac{\rho}{4}.$$

The continuity of $(f_{k,\ell})|_{S_\varepsilon}$ provides some $\delta_{k,\ell} > 0$ such that for all $t \in S_\varepsilon$,

$$|t - t_0| < \delta_{k,\ell} \implies d_Y(f_{k,\ell}(t), f_{k,\ell}(t_0)) < \frac{\rho}{4}.$$

Finally, we obtain for all $(t, x) \in S_\varepsilon \times X$ with $|t - t_0| < \delta_{k,\ell}$ and $d_X(x, x_0) < \frac{\rho}{8\Lambda}$,

$$\begin{aligned} \text{dist}_Y(y_0, F(t, x)) &\leq d_Y(y_0, f_{k,\ell}(t_0)) + \text{dist}_Y(f_{k,\ell}(t_0), F(t, x)) \\ &< \frac{\rho}{4} + d_Y(f_{k,\ell}(t_0), f_{k,\ell}(t)) + \Lambda \cdot d_X(x_k, x) < \frac{\rho}{4} + \frac{\rho}{4} + \Lambda \cdot 2 \frac{\rho}{8\Lambda} < \rho \\ \implies \quad \emptyset &\neq F(t, x) \cap \mathbb{B}_\rho(y_0) \subset F(t, x) \cap O. \end{aligned}$$

The same condition on inverse images of open sets implies the measurability of $F|_{S_\varepsilon \times X}$.

By means of a sequence $\varepsilon_n \downarrow 0$, we obtain the Lebesgue-Borel measurability of $F : [0, T] \times X \rightsquigarrow Y$. Finally, whenever in addition F has closed values, its graph measurability is concluded from Proposition B.4. \square

C Scorza-Dragoni-Like Theorem about Almost Closed Graphs

There are many versions of the Scorza-Dragoni theorem available in the literature (see, e.g., [4, 5, 11, 17, 22, 30–32, 53]). We shall prove here the following one.

Proposition C.1 *Let (X, d_X) and (Y, d_Y) be separable weakly Suslin pseudo-metric spaces and $F : [0, T] \times X \rightsquigarrow Y$ be a set-valued map with nonempty closed values. Assume:*

- (i) *For a.e. $t \in [0, T]$, the graph of $F(t, \cdot) : X \rightsquigarrow Y$ is closed in $X \times Y$.*
- (ii) *The graph of F belongs to the σ -algebra $\mathcal{L}([0, T]) \otimes \mathcal{B}(X) \otimes \mathcal{B}(Y)$.*

Then, for each $\varepsilon > 0$, there exists a closed subset $J_\varepsilon \subset [0, T]$ with $\mathcal{L}^1([0, T] \setminus J_\varepsilon) < \varepsilon$ such that the graph of $F|_{J_\varepsilon \times X} : J_\varepsilon \times X \rightsquigarrow Y$ is closed.

Lemma C.2 *Under the assumptions of Proposition C.1, the function $[0, T] \ni t \mapsto \text{dist}(\xi, \text{Graph } F(t, \cdot))$ is measurable for each $\xi \in X \times Y$.*

Proof of Lemma C.2. By assumption (ii) of Proposition C.1, the graph of F belongs to the σ -algebra

$$(\mathcal{L}([0, T]) \otimes \mathcal{B}(X)) \otimes \mathcal{B}(Y) = \mathcal{L}([0, T]) \otimes (\mathcal{B}(X) \otimes \mathcal{B}(Y))$$

and, we know that $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{B}(X \times Y)$ from Proposition B.1. Hence, the graph of F belongs to $\mathcal{L}([0, T]) \otimes \mathcal{B}(X \times Y)$. Since X and Y are weakly Suslin spaces, $X \times Y$ is also a weakly Suslin space. Thus, Proposition B.5 guarantees a Castaing representation of the set-valued map $[0, T] \ni t \mapsto \text{Graph } F(t, \cdot)$, i.e., there is a sequence $(g_j)_{j \in \mathbb{N}}$ of measurable functions $g_j : [0, T] \rightarrow X \times Y$ such that $\{g_j(t) \mid j \in \mathbb{N}\}$ is dense in $\text{Graph } F(t, \cdot) \subset X \times Y$ for every $t \in [0, T]$.

Finally, for every $\xi \in X \times Y$, we conclude successively that the real-valued functions

$$[0, T] \ni t \mapsto d_{X \times Y}(\xi, g_j(t)) \quad (j \in \mathbb{N}), \quad [0, T] \ni t \mapsto \text{dist}(\xi, \text{Graph } F(t, \cdot)) = \inf_{j \in \mathbb{N}} d_{X \times Y}(\xi, g_j(t))$$

are also measurable (e.g., [10, Lemma 8.2.12], [13, Theorem 2.1.5]). □

Proof of Proposition C.1. Note that

$$d_{X \times Y} : (X \times Y) \times (X \times Y) \rightarrow \mathbb{R}, \quad ((x_1, y_1), (x_2, y_2)) \mapsto d_X(x_1, x_2) + d_Y(y_1, y_2)$$

is a pseudo-metric and, $(X \times Y, d_{X \times Y})$ is separable. Let $\{a_k \in X \times Y \mid k \in \mathbb{N}\}$ denote an arbitrary dense subset of $X \times Y$ and consider the functions

$$[0, T] \ni t \mapsto \varphi_k(t) := \text{dist}(a_k, \text{Graph } F(t, \cdot)) \quad (k \in \mathbb{N}).$$

According to Lemma C.2, each φ_k is measurable. Moreover for each $t \in [0, T]$ at which $\text{Graph } F(t, \cdot)$ is closed in $X \times Y$ (due to assumption (i) of Proposition C.1),

$$\text{Graph } F(t, \cdot) = \bigcap_{k \in \mathbb{N}} \{(x, y) \in X \times Y \mid d_{X \times Y}(a_k, (x, y)) \geq \varphi_k(t)\} \subset X \times Y.$$

Indeed, the inclusion “ \subset ” is obvious. For proving “ \supset ”, choose any $(x, y) \in (X \times Y) \setminus \text{Graph } F(t, \cdot)$. As $\text{Graph } F(t, \cdot)$ is closed, we have $\delta := \text{dist}((x, y), \text{Graph } F(t, \cdot)) > 0$. There is some $k \in \mathbb{N}$ with $d_{X \times Y}(a_k, (x, y)) < \frac{\delta}{3}$ and, the triangle inequality leads to $\varphi_k(t) \geq \frac{2}{3}\delta > d_{X \times Y}(a_k, (x, y))$.

For $\varepsilon > 0$ fixed arbitrarily and each index $k \in \mathbb{N}$, Lusin’s theorem B.6 guarantees a closed subset $A_{\varepsilon, k} \subset [0, T]$ with $\mathcal{L}^1([0, T] \setminus A_{\varepsilon, k}) < \frac{\varepsilon}{2^k}$ such that the restriction $(\varphi_k)|_{A_{\varepsilon, k}}$ is continuous and $\text{Graph } F(t, \cdot)$ closed in $X \times Y$ for every $t \in A_{\varepsilon, k}$. Then, $J_\varepsilon := \bigcap_{k \in \mathbb{N}} A_{\varepsilon, k}$ is a closed subset of $[0, T]$ with $\mathcal{L}^1([0, T] \setminus J_\varepsilon) < \varepsilon$. Finally, $\text{Graph } F|_{J_\varepsilon \times X} = \bigcap_{k \in \mathbb{N}} \{(t, x, y) \in J_\varepsilon \times X \times Y \mid d_{X \times Y}(a_k, (x, y)) \geq \varphi_k(t)\}$ is closed since $d_{X \times Y}(a_k, \cdot) : X \times Y \rightarrow \mathbb{R}$ and $(\varphi_k)|_{J_\varepsilon} : J_\varepsilon \rightarrow \mathbb{R}$ ($k \in \mathbb{N}$) are continuous. □

References

- [1] A. Aggarwal, R. M. Colombo, and P. Goatin. Nonlocal systems of conservation laws in several space dimensions. *SIAM J. Numer. Anal.*, 53(2):963–983, 2015.
- [2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2005.
- [4] J. Appell. Multifunctions of two variables: examples and counterexamples. In *Topology in Nonlinear Analysis (Warsaw, 1994)*, volume 35 of *Banach Center Publ.*, pages 119–128. Polish Acad. Sci. Inst. Math., Warsaw, 1996.
- [5] J. Appell, E. De Pascale, N. H. Thái, and P. P. Zabrejko. Multi-valued superpositions. *Dissertationes Math. (Rozprawy Mat.)*, 345:97, 1995.
- [6] J.-P. Aubin. A note on differential calculus in metric spaces and its applications to the evolution of tubes. *Bull. Polish Acad. Sci. Math.*, 40(2):151–162, 1992.
- [7] J.-P. Aubin. Mutational equations in metric spaces. *Set-Valued Anal.*, 1(1):3–46, 1993.
- [8] J.-P. Aubin. *Mutational and Morphological Analysis*. Systems & Control: Foundations & Applications. Birkhäuser, Boston, MA, 1999. Tools for shape evolution and morphogenesis.
- [9] J.-P. Aubin and A. Cellina. *Differential Inclusions*, volume 264 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1984. Set-valued maps and viability theory.
- [10] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*, volume 2 of *Systems & Control: Foundations & Applications*. Birkhäuser, Boston, MA, 1990.
- [11] D. Aversa. Lusin type theorems for multifunctions, Scorza Dragoni’s property and Carathéodory selections. *Boll. Unione Mat. Ital., VII. Ser., A*, 8(2):193–202, 1994.
- [12] Z. Badreddine and H. Frankowska. Hamilton-Jacobi inequalities on a metric space. *J. Differential Equations*, 271:1058–1091, 2021.
- [13] V. I. Bogachev. *Measure Theory. Vol. I, II*. Springer, Berlin, 2007.
- [14] R. Borsche, R. M. Colombo, M. Garavello, and A. Meurer. Differential equations modeling crowd interactions. *J. Nonlinear Sci.*, 25(4):827–859, 2015.
- [15] J. A. Cañizo, J. A. Carrillo, and S. Cuadrado. Measure solutions for some models in population dynamics. *Acta Appl. Math.*, 123:141–156, 2013.
- [16] J. A. Carrillo, R. M. Colombo, P. Gwiazda, and A. Ulikowska. Structured populations, cell growth and measure valued balance laws. *J. Differential Equations*, 252(4):3245–3277, 2012.
- [17] C. Castaing, M. Moussaoui, and A. Syam. Multivalued differential equations on closed convex sets in Banach spaces. *Set-Valued Anal.*, 1(4):329–353, 1993.
- [18] L. Doyen. Filippov and invariance theorems for mutational inclusions of tubes. *Set-Valued Anal.*, 1(3):289–303, 1993.
- [19] L. Doyen. Mutational equations for shapes and vision-based control. *J. Math. Imaging Vis.*, 5(2):99–109, 1995.
- [20] R. M. Dudley. *Real Analysis and Probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.
- [21] H. Frankowska. A priori estimates for operational differential inclusions. *J. Differential Equations*, 84(1):100–128, 1990.
- [22] H. Frankowska, S. Plaskacz, and T. Rzeżuchowski. Measurable viability theorems and the Hamilton-Jacobi-Bellman equation. *J. Differential Equations*, 116(2):265–305, 1995.
- [23] D. H. Fremlin. Measurable functions and almost continuous functions. *Manuscripta Math.*, 33(3-4):387–405, 1980/81.

- [24] S. Gautier and K. Pichard. Viability results for mutational equations with delay. *Numer. Funct. Anal. Optim.*, 24(3-4):273–284, 2003.
- [25] P. Goatin and S. Scialanga. Well-posedness and finite volume approximations of the LWR traffic flow model with non-local velocity. *Netw. Heterog. Media*, 11(1):107–121, 2016.
- [26] X. Gong and M. Kawski. Weak measure-valued solutions of a nonlinear hyperbolic conservation law. *SIAM J. Math. Anal.*, 53(4):4417–4444, 2021.
- [27] A. Gorre. Evolutions of tubes under operability constraints. *J. Math. Anal. Appl.*, 216(1):1–22, 1997.
- [28] C. J. Himmelberg. Precompact contraction of metric uniformities, and the continuity of $F(t, x)$. *Rend. Sem. Mat. Univ. Padova*, 50:185–188 (1974), 1973.
- [29] C. J. Himmelberg. Correction to: “Precompact contraction of metric uniformities, and the continuity of $F(t, x)$ ” (Rend. Sem. Mat. Univ. Padova **50** (1973), 185–188 (1974)). *Rend. Sem. Mat. Univ. Padova*, 51:361 (1975), 1974.
- [30] C. J. Himmelberg and F. S. Van Vleck. An extension of Brunovský’s Scorza Dragoni type theorem for unbounded set-valued functions. *Math. Slovaca*, 26(1):47–52, 1976.
- [31] S. Hu and N. S. Papageorgiou. *Handbook of Multivalued Analysis. Vol. I*, volume 419 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1997. Theory.
- [32] J. Jarník and J. Kurzweil. On conditions on right hand sides of differential relations. *Časopis Pěst. Mat.*, 102(4):334–349, 426, 1977.
- [33] P. E. Kloeden and T. Lorenz. Stochastic morphological evolution equations. *J. Differential Equations*, 251(10):2950–2979, 2011.
- [34] P. E. Kloeden and T. Lorenz. Nonlocal multi-scale traffic flow models: analysis beyond vector spaces. *Bull. Math. Sci.*, 6(3):453–514, 2016.
- [35] A. Kucia. Scorza Dragoni type theorems. *Fund. Math.*, 138(3):197–203, 1991.
- [36] J. Kupka and K. Prikry. The measurability of uncountable unions. *Amer. Math. Monthly*, 91(2):85–97, 1984.
- [37] K. Kuratowski and C. Ryll-Nardzewski. A general theorem on selectors. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 13:397–403, 1965.
- [38] S. J. Leese. Multifunctions of Souslin type. *Bull. Austral. Math. Soc.*, 11:395–411, 1974.
- [39] S. J. Leese. *Set-Valued Functions and Selectors*. Doctoral thesis. Keele University, 1974.
- [40] S. J. Leese. Corrigendum: “Multifunctions of Souslin type” (Bull. Austral. Math. Soc. **11** (1974), 395–411). *Bull. Austral. Math. Soc.*, 13(1):159–160, 1975.
- [41] T. Lorenz. Morphological control problems with state constraints. *SIAM J. Control Optim.*, 48(8):5510–5546, 2010.
- [42] T. Lorenz. *Mutational Analysis*, volume 1996 of *Lecture Notes in Mathematics*. Springer, Berlin, 2010. A joint framework for Cauchy problems in and beyond vector spaces.
- [43] T. Lorenz. Nonlocal hyperbolic population models structured by size and spatial position: well-posedness. *Discrete Contin. Dyn. Syst. Ser. B*, 24(8):4547–4628, 2019.
- [44] T. Lorenz. Partial differential inclusions of transport type with state constraints. *Discrete Contin. Dyn. Syst. Ser. B*, 24(3):1309–1340, 2019.
- [45] T. Lorenz. Viability in a non-local population model structured by size and spatial position. *J. Math. Anal. Appl.*, 491(1):124249, 50, 2020.
- [46] J. A. Murillo Hernández. Tangential regularity in the space of directional-morphological transitions. *J. Convex Anal.*, 13(2):423–441, 2006.
- [47] L. Najman. Euler method for mutational equations. *J. Math. Anal. Appl.*, 196(3):814–822, 1995.
- [48] B. Piccoli. Measure Differential Equations. *Arch. Ration. Mech. Anal.*, 233(3):1289–1317, 2019.

- [49] B. Piccoli and F. Rossi. Transport equation with nonlocal velocity in Wasserstein spaces: convergence of numerical schemes. *Acta Appl. Math.*, 124:73–105, 2013.
- [50] B. Piccoli and F. Rossi. Generalized Wasserstein distance and its application to transport equations with source. *Arch. Ration. Mech. Anal.*, 211(1):335–358, 2014.
- [51] B. Piccoli and A. Tosin. Time-evolving measures and macroscopic modeling of pedestrian flow. *Arch. Ration. Mech. Anal.*, 199(3):707–738, 2011.
- [52] K. Pichard and S. Gautier. Equations with delay in metric spaces: the mutational approach. *Numer. Funct. Anal. Optim.*, 21(7-8):917–932, 2000.
- [53] A. A. Tolstonogov. On the Scorza-Dragoni theorem for multivalued mappings with a variable domain. *Mat. Zametki*, 48(5):109–120, 160, 1990.
- [54] R. Vinter. *Optimal Control*. Modern Birkhäuser Classics. Birkhäuser, Boston, MA, 2010. Paperback reprint of the 2000 edition.
- [55] D. H. Wagner. Survey of measurable selection theorems. *SIAM J. Control Optimization*, 15(5):859–903, 1977.
- [56] D. H. Wagner. Survey of measurable selection theorems: an update. In *Measure Theory, Oberwolfach 1979 (Proc. Conf., Oberwolfach, 1979)*, volume 794 of *Lecture Notes in Math.*, pages 176–219, Berlin-New York, 1980. Springer.

Hélène Frankowska
 CNRS, IMJ-PRG, Sorbonne Université,
 Case 247, 4 Place Jussieu, 75252 Paris, France
 e-mail: helene.frankowska@imj-prg.fr

Thomas Lorenz
 Heidelberg University, Faculty of Mathematics and Computer Science,
 69120 Heidelberg, Germany
 e-mail: thomas.lorenz@mathinf.uni-heidelberg.de