Consistency of vanishingly smooth fictitious play

Michel Benaïm
Institut de Mathématiques, Université de Neuchâtel, Rue Emile-Argand 11. Neuchâtel. Switzerland
email: michel.benaim@unine.ch

Mathieu Faure
Aix-Marseille School of Economics (GREQAM, AMU), Centre de la Vieille Charité. 2, rue de la vieille charité, Marseille
email: mathieu.faure@univ-amu.fr

We discuss consistency of *Vanishingly Smooth Fictitious Play*, a strategy in the context of game theory, which can be regarded as a *smooth fictitious play procedure*, where the smoothing parameter is time-dependent and asymptotically vanishes. This answers a question initially raised by Drew Fudenberg and Satoru Takahashi.

---

1. Introduction and background   A recurring question in the theory of repeated games is to define properly a notion of *good strategy* for a player facing an unknown environment. Consequently, in this paper, we are not concerned with the formalisation of strategic interactions between rational players, but rather between a *decision maker* and *nature*. Not much is known about the latter, no assumption is made on its payoff function, its thinking process or its rationality. We take the point of view of the former, whose objective is to maximize his/her average payoff in the long run. A naive approach in this direction is to assume that the game is zero-sum and to look for optimal strategies. However, the fact that his/her opponent might not try to maximize his/her payoff could lead to bad outcomes. A possible definition of *good strategy* for the decision maker has been proposed by Hannan (see [16]). It is closely related to the concept of *regret*. After n stages, the regret of the decision maker is the difference between the payoff that he could have obtained if he knew in advance the empirical moves of nature and the average payoff he actually got. A good strategy for the decision maker may then be defined as a strategy which ensures that, regardless of the behaviour of nature, the regret asymptotically goes to zero. Such a strategy is called *consistent*. Consistent strategies are known to exist for a long time and can be constructed, for instance, using so-called *block-annealing* procedures (see e.g. [5], [10], [11] and [18]). For a complete bibliography on the topic, see the last quoted paper. Also, for a recent comprehensive overview about consistency in games, see [22] (in french). However *fictitious play* strategies are known to be non-consistent (see [14]) while *smooth fictitious play* strategies have been shown to be "almost" consistent by Fudenberg and Levine [13] (see section 1.2 for a rigorous exposition). In this paper, we consider a time-varying smooth fictitious play with a smoothing parameter decreasing to zero, that we call *vanishingly smooth fictitious play* (VSFP). VSFP strategies initially behave like smooth fictitious play and asymptotically like fictitious play. The main objective of this work is to answer the following question raised to us by Drew Fudenberg and Satoru Takahashi: "are VSFP strategies consistent?"

1.1 Notation   We consider a two-player finite game in normal form. I and L are the (finite) set of moves of respectively player 1 (the decision maker) and player 2 (the nature). The map $\pi : I \times L \to \mathbb{R}$ denotes the payoff function of player 1. The sets of mixed strategies available to players are denoted $X = \Delta(I)$ and $Y = \Delta(L)$, where

$$\Delta(I) := \left\{ x \in \mathbb{R}_+^I \ | \ \sum_{i \in I} x_i = 1 \right\},$$

and analogously for $\Delta(L)$. As usual π is extended to $X \times Y$ by multilinearity:

$$\forall x \in X, \, y \in Y, \, \pi(x,y) = \sum_{i \in I} \sum_{l \in L} \pi(i,l) x_i y_l.$$

In the following, $(i_1, ..., i_n, ...)$ (respectively $(l_1, ..., l_n, ...)$) will denote the sequence of actions picked by player 1 (resp. his/her opponent). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, endowed with a filtration $(\mathcal{F}_n)_n$.

Formally, a *strategy* for player 1 is a choice of an adapted process $(i_n)_n$ on $(\Omega, (\mathcal{F}_n)_n, \mathbb{P})$, Throughout the paper, we assume that the agents play independently: specifically, for $(i, l) \in I \times L$, we have

$$\mathbb{P}\left(i_{n+1} = i, \ l_{n+1} = l \mid \mathcal{F}_n\right) = \mathbb{P}\left(i_{n+1} = i \mid \mathcal{F}_n\right)\mathbb{P}\left(l_{n+1} = l \mid \mathcal{F}_n\right).$$

Finally, we call

$$\overline{x}_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{i_k}$$

the average moves of player 1 at time n, $\overline{y}_n$ the average moves of player 2 and

$$\overline{\pi}_n = \frac{1}{n}\sum_{k=1}^{n}\pi(i_k, l_k)$$

the average payoff to player 1.

1.2 Consistency, definition and comments   We now introduce $\Pi : Y \to \mathbb{R}$, defined by

$$\Pi(y) := \max_{i \in I}\pi(i, y).$$

A strategy of the decision maker is *consistent* if, against any strategy of nature, the average payoff obtained by player 1 is at least as much as if the sequence of empirical moves of nature was known in advance, and the decision maker had played a best response against it. More precisely, let us define the *average regret evaluation* along a sequence of moves $h_n = (i_1, l_1, ...i_n, l_n)$:

$$e_n := \max_{i \in I}\pi\left(i, \frac{1}{n}\sum_{m=1}^{n}l_m\right) - \frac{1}{n}\sum_{m=1}^{n}\pi(i_m, l_m) = \Pi(\overline{y}_n) - \overline{\pi}_n.$$

DEFINITION 1.1 *A strategy for player 1 is said to be consistent if, for any strategy of nature,*

$$\limsup_n e_n \leq 0, \ \ \mathbb{P} - almost \ surely.$$

It is η-consistent if

$$\limsup_n e_n \leq \eta, \ \ \mathbb{P} - almost \ surely.$$

Given $y \in Y$, we call $br(y)$ the set of best responses of player 1 to y, namely,

$$br(y) = \mathsf{Argmax}_{x \in X}\pi(x, y).$$

The discrete-time *fictitious play* (FP) process has been introduced in [6]. We say that player 1 uses a FP strategy, with prior $\overline{y}_0$ if, for $n \geq 1$,

$$\mathbb{P}(i_{n+1} = \cdot \mid \mathcal{F}_n) \in br(\gamma_n),$$

where $\gamma_n = \frac{1}{n+1}\overline{y}_0 + \frac{n}{n+1}\overline{y}_n$. It is well known that this strategy is not consistent. A simple example is given by the following (see e.g. [14]).

EXAMPLE 1.1 Assume that the game is matching pennies, i.e. the payoff matrix of player 1 is given by

$$\begin{array}{cc} & \begin{array}{cc} H & T \end{array} \\ \begin{array}{c} H \\ T \end{array} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

and the prior is $\overline{y}_0 = (1/3, 2/3)$. If player two acts accordingly to the deterministic rule *heads* (H) on odd stages and *tails* (T) on even stages, then player 1 and 2 always play the opposite and the average regret satisfies $\lim_{n\to\infty} e_n = 1/2$.

However, η-consistency can be achieved by small modifications of fictitious play, which are usually called *stochastic fictitious play strategies*. Originally, stochastic fictitious play was introduced by Fudenberg and Kreps in [12] and the concept behind this is that players use fictitious play in a game where payoff functions are perturbed by some random variables in the spirit of Harsanyi [17]. On the subject, see also [13], [14] or [2]. In this paper, we adopt another point of view and assume that player 1 chooses to randomize his/her moves by adding a small perturbation function to his/her initial payoff map π.

The class of perturbation functions considered here is the following. It is directly inspired by [14] or [20]. Let $\rho : Int(X) \to \mathbb{R}$ be a C^2 map such that:

(A1) There exists $\lambda > 0$ such that, for all $x \in Int(X)$ and $h \in TX$,

$$\langle D^2 \rho(x) h, h \rangle \geq \lambda \|h\|^2.$$

(A2) the first derivative of ρ verifies

$$\lim_{x \to \partial X} \|\nabla \rho(x)\| = +\infty.$$

We introduce the perturbed payoff function $\tilde{\pi}$ defined, for $x \in X$, $y \in Y$ and $\beta > 0$ by

$$\tilde{\pi}(x, y, \beta) = \pi(x, y) + \frac{1}{\beta} \rho(x).$$

Under $(A1)$ and $(A2)$, the function $\tilde{\pi}$ enjoys the following property:

   (i) For all $y \in Y$, $\beta > 0$, $\mathsf{Argmax}_{x \in X} \tilde{\pi}(x, y, \beta)$ reduces to one point and defines a continuous map **br** from $Y \times \mathbb{R}^*_+$ to $Int(X)$.

   (ii) There exists $L > 0$ such that the map $y \mapsto \mathbf{br}(\beta, y)$ is Lipschitz continuous, with Lipschitz constant $L\beta$.

The map $(y, \beta) \in Y \times \mathbb{R}^*_+ \mapsto \mathbf{br}(y, \beta)$ is usually called a *smooth best response* map.

Throughout the remainder of the paper, we assume that ρ verifies $(A1)$ and $(A2)$. We call such a map is a *good* perturbation function.

REMARK 1.1 *Let $\rho : x \in X \mapsto \rho(x) = -\sum_{i \in I} x_i \log x_i$ be the entropy function. It is a particular case of a good perturbation function, and the resulting smooth best response is the so-called* logit map, *given by*

$$\mathbf{L}(\beta, y)_i = \frac{\exp(\beta \pi(i, y))}{\sum_{k \in I} \exp(\beta \pi(k, y))}$$

DEFINITION 1.2 *Player 1 plays accordingly to a smooth fictitious play strategy, with the parameter $\beta > 0$ (SFP(β)) if*

$$\mathbb{P}\left(i_{n+1} = i \mid \mathcal{F}_n\right) = \mathbf{br}(\overline{y}_n, \beta)_i, \ \ \forall n \geq 1.$$

THEOREM 1.1 (FUDENBERG AND LEVINE, 1995) *For any $\eta > 0$, there exists $\beta_0 > 0$ such that a SFP(β) strategy is η-consistent for any $\beta > \beta_0$.*

Smooth fictitious play is closely related to the so-called *exponential weight algorithm* and also to the *follow the perturbed leader algorithm* (see [7], chapters 4.2 and 4.3), even if the link with the latter is less obvious. In [23], the authors discuss the consistency of continuous-time versions of FP and SFP.

1.3 Vanishingly smooth fictitious play A related natural strategy is given by the following. Recall that **br** is a smooth best response function, induced by a good perturbation function.

DEFINITION 1.3 *Let $(\beta_n)_n$ be a sequence going to infinity. The vanishingly smooth fictitious play strategy induced by β_n (and **br**) for player 1 is defined by*

$$\mathbb{P}\left(i_{n+1} = i \mid \mathcal{F}_n\right) = \mathbf{br}\left(\overline{y}_n, \beta_n\right)_i \ \ \forall n \geq 1.$$

We use the notation VSFP(β_n) in the sequel. Consistency is not verified for any choice of $(\beta_n)_n$. If this sequence increases too fast, then consistency might fail to hold, as shown by the following example.

EXAMPLE 1.2 Assume that, once again the game is 2-player matching pennies and that nature uses the deterministic strategy described in example 1.1. Then, if player one plays accordingly to a VSFP strategy induced by the logit map, $\beta_n = n$ and prior $\overline{y}_0 = (1/3, 2/3)$, we have

$$\gamma_{2n} = \left(\frac{1}{2} - \frac{1}{6(2n+1)}, \frac{1}{2} + \frac{1}{6(2n+1)}\right) \ \text{ and } \ \gamma_{2n+1} = \left(\frac{1}{2} + \frac{1}{6(n+1)}, \frac{1}{2} - \frac{1}{6(n+1)}\right).$$

After a few lines of calculus (left to the reader) one gets:

$$\mathbb{E}\left(\delta_{l_{2n+1}} \mid \mathcal{F}_n\right) = \mathbf{L}(\gamma_{2n}, \beta_{2n}) = \left(\frac{1}{1 + \exp\left(\frac{2n}{3(2n+1)}\right)}, \frac{1}{1 + \exp\left(-\frac{2n}{3(2n+1)}\right)}\right).$$

Hence $(\pi(i_{2n+1}, l_{2n+1}))_n$ is a sequence of independent random variables taking values in $\{0, 1\}$, such that

$$\lim_n \mathbb{P}\left(\pi(i_{2n+1}, l_{2n+1}) = 1\right) = \lim_n \frac{1}{1 + e^{2n/3(2n+1)}} = \frac{1}{1 + e^{1/3}} = 1/2 - a$$

with $a > 0$. Similarly, $(\pi(i_{2n}, l_{2n}))_n$ is a sequence of independent random variables taking values in $\{0, 1\}$ and

$$\lim_n \mathbb{P}\left(\pi(i_{2n}, l_{2n}) = 1\right) = \frac{1}{1 + e^{2/3}} = 1/2 - b$$

with $b > 0$.

Therefore, consistency is not satisfied for VSFP strategies with $\beta_n = n$ since $e_n \to (a+b)/2 > 0$.

We now can state our main result.

THEOREM 1.2 *Any VSFP(β_n) strategy, with $\beta_n \leq n^\nu$ for some $\nu < 1$, is consistent.*

In [4], the authors prove the same result as Theorem 1.1 using stochastic approximation methods. Specifically, they consider the state variable $(\overline{x}_n, \overline{y}_n, \overline{\pi}_n)_n$, write it as a stochastic approximation process relative to some differential inclusion, and prove that it almost surely converges to the *consistency set*:

$$\{(x, y, \pi) : \ \Pi(y) - \pi \leq \eta\} .$$

This is the approach taken in this paper. In section 2 we show how our state variable can be written as a stochastic approximation algorithm, relative to some non-autonomous differential inclusion. A concept of Lyapunov function with respect to a set A for non-autonomous systems is introduced in section 3 and, in Proposition 3.2, we establish that A attracts the so-called *perturbed solutions*, under the right conditions. In our specific case, we also prove that there exists a Lyapunov function relative to the consistency set. The proof of our main result, Theorem 1.2, is given in Section 4. It consists in showing that $(\overline{x}_n, \overline{y}_n, \overline{\pi}_n)_n$ is almost surely a perturbed solution with good properties and applying the results of Section 3. In the appendix, we provide some general stability results for non-autonomous differential inclusions, namely we estimate the deviation of so-called *perturbed solutions* from the set of solutions curves.

2. Stochastic approximations

2.1 A stochastic difference inclusion   As it was previously mentioned, we are interested in the asymptotic behavior of the state variable $v_n := (\overline{x}_n, \overline{y}_n, \overline{\pi}_n) \in M := X \times Y \times [-\|\pi\|_\infty, \|\pi\|_\infty]$, where $\|\pi\|_\infty := \max_{i,l} |\pi(i, l)|$. We have

$$\overline{x}_{n+1} - \overline{x}_n - \frac{1}{n+1} \left(\delta_{i_{n+1}} - \mathbb{E}_\sigma(\delta_{i_{n+1}} \mid \mathcal{F}_n)\right) = \frac{1}{n+1}\left(-\overline{x}_n + \mathbf{br}(\overline{y}_n, \beta_n)\right) .$$

Writing the analogous recursive formulas for $\overline{y}_n$ and $\overline{\pi}_n$, we obtain that

$$v_{n+1} - v_n - \frac{1}{n+1} U_{n+1} \in \frac{1}{n+1} F_n(v_n),$$

where

- the noise sequence
$$U_{n+1} = (v_{n+1} - v_n) - \mathbb{E}(v_{n+1} - v_n \mid \mathcal{F}_n)$$
  is a bounded martingale difference,
- the set valued map F_n is given by

$$F_n(x, y, \pi) := \{(\mathbf{br}(y, \beta_n) - x, \tau - y, \pi(\mathbf{br}(y, \beta_n), \tau) - \pi, \ \tau \in Y\} . \tag{1}$$

2.2 Stochastic approximations relative to non-autonomous differential inclusions   On a more general level, let $M \subset \mathbb{R}^d$ and $F : \mathbb{R}_+ \times M \rightrightarrows \mathbb{R}^d$ be a set-valued map taking values in the set of non-empty, compact, convex subsets of $\mathbb{R}^d$. We say that F is *regular* if :

(R1) $s \mapsto F(t, w)$ is measurable, for each $w \in M$;

$(R2)$ for any $t \in \mathbb{R}^+$, the map $w \mapsto F(t, w)$ has a closed graph, i.e.

$$\{(w, w') \in M \times M : \ w' \in F(t, w)\}$$

is closed;

$(R3)$ The map F is uniformly bounded, i.e., $\sup_{t,w} \sup_{w' \in F(t,w)} \|w'\| \leq \|F\|_\infty < +\infty$.

Consider a discrete time stochastic process $(v_n)_n$ in M, defined by the recursive formula

$$v_{n+1} - v_n - \gamma_{n+1} U_{n+1} \in \gamma_{n+1} F_n(v_n), \tag{2}$$

where $F_n : M \rightrightarrows \mathbb{R}^d$ is a set-valued map, $(\gamma_n)_n$ is a positive sequence, decreasing to 0 and $(U_n)_n$ a sequence of M-valued random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. Set $\tau_n := \sum_{i=1}^n \gamma_i$ and $m(s) := \sup\{j \mid \tau_j \leq s\}$. We make the following additional assumptions:

$(SA1)$ For all $c > 0$,

$$\sum_n e^{-c/\gamma_n} < \infty,$$

$(SA2)$ $(U_n)_n$ is uniformly bounded (by $\|U\|_\infty$) and

$$\mathbb{E}\left(U_{n+1} \mid \mathcal{F}_n\right) = 0,$$

$(SA3)$ The map $F : \mathbb{R}_+ \times M \rightrightarrows M$, given by

$$F(t, w) := F_{m(t)}(w)$$

is regular.

DEFINITION 2.1 *If the conditions $(SA1)$, $(SA2)$ and $(SA3)$ are met, we say that (v_n) is a good stochastic approximation algorithm relative to F.*

Call $v(\cdot)$ the continuous time affine interpolated process induced by $(v_n)_n$ and $\overline{\gamma}(\cdot)$ (resp. $\overline{U}(\cdot)$) the piecewise constant deterministic processes induced by $(\gamma_n)_n$ (resp. $(U_n)_n$):

$$v(\tau_i + s) = v_i + s \frac{v_{i+1} - v_i}{\gamma_{i+1}} \text{ for } s \in [0, \gamma_{i+1}], \ \overline{\gamma}(\tau_i + s) := \gamma_{i+1} \text{ for } s \in [0, \gamma_{i+1}[,$$

and analogously for $\overline{U}$.

LEMMA 2.1 *For almost every $s \in \mathbb{R}_+$, $v(\cdot)$ is differentiable and we have*

$$\dot{v}(s) - \overline{U}(s) \in F(s, v_{m(s)}).$$

PROOF. We have

$$v(s) = v_{m(s)} + \frac{v_{m(s)+1} - v_{m(s)}}{\gamma_{m(s)+1}}(s - \tau_{m(s)})$$

Hence, if $s \notin \{\tau_n, \ n \in \mathbb{N}^*\}$, $v(\cdot)$ is differentiable and

$$\dot{v}(s) = \frac{v_{m(s)+1} - v_{m(s)}}{\gamma_{m(s)+1}}.$$

Consequently

$$\dot{v}(s) - \overline{U}(s) \in F_{m(s)}(v_{m(s)}) = F(s, v_{m(s)}).$$

$\square$

In the sequel, we use the notation $\overline{v}(s) := v_{m(s)}$. Notice that $\overline{v}$ is a piecewise constant map on $\mathbb{R}_+$.

Let us come back to the particular case of section 2.1, where $v_n = (\overline{x}_n, \overline{y}_n, \overline{\pi}_n)$ and F_n is given by (1).

LEMMA 2.2 *$(v_n)_n$ is a good stochastic approximation algorithm with step size $\gamma_n = 1/n$, relative to the map F given by*

$$F(t, w) := F_{m(t)}(w) = \left\{ (\mathbf{br}(y, \beta_{m(t)}) - x, \tau - y, \pi(\mathbf{br}(y, \beta_{m(t)}), \tau) - \pi, \ \tau \in Y \right\}. \tag{3}$$

PROOF. We only need to prove that F is a regular set-valued map. The fact that F has non-empty compact convex values is straightforward, as well as measurability. Also, the map F takes values in a compact set. Thus F is uniformly bounded. Given $s \in \mathbb{R}_+$, we now need to check upper semi-continuity of $v \mapsto F(s, v)$, which is equivalent to $\{(v, w), \ w \in F(s, v)\}$ being closed. Let (x_n, y_n, π_n) converge to (x, y, π) and $\tau_n \to_n \tau \in Y$. We then have $\lim_n \mathbf{br}(y_n, \beta_{m(s)}) = \mathbf{br}(y, \beta_{m(s)})$. Hence,

$$
\begin{aligned}
\lim_n \quad & \left(\mathbf{br}(y_n, \beta_{m(s)}) - x_n, \tau_n - y_n, \pi\left(\mathbf{br}(y_n, \beta_{m(s)}), \tau_n\right) - \pi_n\right) \\
= \quad & \left(\mathbf{br}(y, \beta_{m(s)}) - x, \tau - y, \pi\left(\mathbf{br}(y, \beta_{m(s)}), \tau\right) - \pi\right) \in F(s, x, y, \pi).
\end{aligned}
$$

$\square$

In the particular case where F is an autonomous set-valued map (i.e. F does not depend on $t \in \mathbb{R}_+$), stochastic approximation algorithms described above have been studied in [3] where it is proved that there is a deep relationship between the asymptotic behavior of (v_n) and the solutions of the autonomous differential inclusion

$$
\dot{\mathbf{w}} \in F(\mathbf{w}).
$$

In particular, they show that, if there exists a global attractor A for the deterministic dynamics, then the limit set of $(v_n)_n$ is contained in A.

Unfortunately, in our case, the mean deterministic system associated to our random process $(v_n)_n$ is a non-autonomous differential inclusion, as we will see later on.

3. Lyapunov functions relative to nonautonomous differential inclusions

3.1 Perturbed solutions and uniform Lyapunov functions Let us consider the *non-autonomous differential inclusion*

$$
\dot{\mathbf{w}}(s) \in F(s, \mathbf{w}(s)), \ \ s \in [a, b] \tag{4}
$$

A map $\mathbf{w} : [a, b] \to M$ is a solution of (4) if it is absolutely continuous and, for almost every $s \in [0, T]$, $\dot{\mathbf{w}}(s) \in F(s, \mathbf{w}(s))$. The existence of solutions from any initial condition is guaranteed under various sets of assumptions, in particular for regular F (see Section 5.2 for more details)

DEFINITION 3.1 *A map $v : \mathbb{R}_+ \to M$ is a* perturbed solution *of the non-autonomous differential inclusion $\dot{\mathbf{w}}(s) \in F(s, \mathbf{w}(s))$ if there is a locally integrable function $\overline{U} : \mathbb{R}_+ \to \mathbb{R}^d$ such that*

$(PS1)$ *v is absolutely continuous,*

$(PS2)$ *$\lim_{t \to +\infty} \Delta(t, t + T) = 0$ where*

$$
\Delta(t, t + T) := \sup_{h \in [0, T]} \int_t^{t+h} \overline{U}(s) ds
$$

$(PS3)$ *$\dot{v}(s) - \overline{U}(s) \in F(s, \overline{v}(s))$ for some measurable map $\overline{v} : \mathbb{R}_+ \to M$ such that*

$$
\|v(s) - \overline{v}(s)\| \leq \delta(s),
$$

with $\delta(s) \downarrow_s 0$.

REMARK 3.1 *Notice that, in the autonomous case, this is Definition (II) in [3]*

PROPOSITION 3.1 *Let $v(\cdot)$ be the continuous time affine interpolated process associated to a good stochastic approximation. Then v is almost surely a perturbed solution, with $\overline{v}(s) = v_{m(s)}$ and $\delta(s) = c\overline{\gamma}(s)$ (where c is some positive constant).*

PROOF. This is a direct consequence of Lemma 2.1 and Proposition 4.4 in [1]. We will provide more details in the particular case we are interested in, in Section 4. $\square$

We now define a concept of Lyapunov function adapted to non-autonomous differential inclusions.

DEFINITION 3.2 *Let A be a compact set in M and V be an open neighbourhood of A. A smooth map $\Phi : \mathbb{R}_+ \times V \to \mathbb{R}_+$ is called a* uniform Lyapunov function *for the non-autonomous differential inclusion (4) with respect to A if the following hold:*

a) *we have*

$$A = \left\{ w \in V : \ \exists s_n \uparrow +\infty, \ \lim_{n \to +\infty} \Phi(s_n, w) = 0 \right\},$$

b) *There exist two maps* $\lambda : \mathbb{R}_+^* \to]0,1[$ *and* $\varepsilon : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ *with the property that*

$$\lim_{T \to +\infty} \lambda(T) = 0 \quad and \quad \lim_{t \to +\infty} \varepsilon(t, T) = 0, \ \forall T > 0;$$

and, for any $t > 0, T > 0$ *and any solution* $\mathbf{w}$ *on* $[t, t+T]$, *we have*

$$\Phi(t+s, \mathbf{w}(t+s)) \leq \lambda(s)\Phi(t, \mathbf{w}(t)) + \varepsilon(t, T), \ \forall s \in [0, T].$$

If $V = M$ *then* Φ *is called a* global uniform Lyapunov function.

REMARK 3.2 *Assumption a) is checked in particular if the somewhat more explicit condition is verified:*

a′) *there exists a continuous map* $g : V \to \mathbb{R}+$ *such that*

$$A = \{ w \in M : g(w) = 0 \}, \ \|g(w) - \Phi(s, w)\| \to_{s \to +\infty} 0,$$

uniformly in $w \in V$.

The following lemma will be useful to prove the main result of this section, namely Proposition 3.2.

LEMMA 3.1 *Let* $(\Phi_k)_{k \geq k_0}$, $(\lambda_k)_{k \geq k_0}$ *and* $(\eta_k)_{k \geq k_0}$ *be positive sequences of real numbers such that* $0 < \lambda_k < 1$ *and*

(i) *for any* $k \geq k_0$

$$\Phi_{k+1} \leq \lambda_k \Phi_k + \eta_{k+1};$$

(ii) *for* $k \geq k_0 + 1$, *denoting* $H_k := \Pi_{i=k_0}^{k-1} \lambda_i$ *and* $\tilde{H}_k = H_k \sum_{i=0}^{k-1} H_i^{-1} \eta_i$, *we have* $\lim_{k \to \infty} H_k = \lim_{k \to \infty} \tilde{H}_k = 0$.

Then $\lim_{k \to \infty} \Phi_k = 0$.

PROOF. Without loss of generality, we assume that $k_0 = 0$. A simple recursive argument yields

$$\Phi_k \leq H_k \left(\Phi_0 + \sum_{i=1}^{k} H_i^{-1} \eta_i \right)$$

and the proof is complete. □

We say that Φ is *uniformly Lipschitz* if there exists $L_\Phi > 0$ such that, for any $s \geq 0$ and $w, w' \in M$,

$$|\Phi(s, w) - \Phi(s, w')| \leq L_\Phi \|w - w'\|.$$

We now need to define Lipschitz continuity for non-autonomous set-valued maps: call d_H the Hausdorff distance, given by

$$d_H(E, F) = \max \left\{ \sup_{x \in E} d(x, F), \ \sup_{y \in F} d(y, E) \right\}.$$

Recall that d_H is a pseudo-metric on the set of non-empty subsets of M and a metric if we restrict to the non-empty compact sets of M. We say that F is *Hausdorff continuous* if it is continuous with respect of the Hausdorff metric:

$$\lim_{t \to t_0, w \to w_0} d_H(F(t, w), F(t_0, w_0)) = 0.$$

If F is Hausdorff continuous, we call it L-Lipschitz, for an integrable function $L : [a, b] \to \mathbb{R}_+$ if

$$d_H(F(t, w), F(t, w')) \leq L(t) \|w - w'\|, \ \text{for a.e. } t \in [a, b], \ \forall \ w, w'$$

We now state the main result of this section. Corollary 5.1 (stated in the appendix) plays an important role here, as it gives upper bound for the deviation of perturbed solutions from actual solutions of the deterministic system. For convenience of the reader, we chose to postpone this technical result to Section 5.2.

PROPOSITION 3.2 *Assume that v is a perturbed solution relative to a regular Lipschitz map F (with $L : \mathbb{R}_+ \to \mathbb{R}_+$) and that Φ is a global uniform Lyapunov function with respect to a compact set A and the differential inclusion (4). Assume also that there exists a sequence of positive real numbers $(T_k)_k$ such that*

(i) $S_k := \sum_{i=1}^{k} T_i \to +\infty$,

(ii) *there exists $k_0 \in \mathbb{N}$ and a sequence $(r_k)_k$ such that, for any $k \geq k_0$*

$$R(S_k, S_{k+1}) \leq r_k,$$

with R defined by (7) in Corollary 5.1,

(iii) Φ *is uniformly Lipschitz, with constant L_Φ,*

(iv) *denoting $H_k := \Pi_{i=k_0+1}^{k} \lambda(T_i)$ and $\eta_k := \varepsilon(S_{k-1}, T_k) + L_\Phi r_{k-1}$, we have*

$$\lim_{k \to +\infty} H_k \sum_{i=1}^{k} H_i^{-1} \eta_i = 0.$$

Then the limit set of v, $\mathcal{L}((v(s))_{s>0}) := \{v^* : \exists s_n \uparrow +\infty, \lim_n v(s_n) = v^*\}$ is contained in A.

PROOF. First, by Corollary 5.1, for any $k \in \mathbb{N}$, there exists a solution $\mathbf{w}^\mathbf{k}$ on $[S_k, S_{k+1}]$ such that $\mathbf{w}^\mathbf{k}(S_k) = v(S_k)$ and

$$\sup_{s \in [S_k, S_{k+1}]} \|v(s) - \mathbf{w}^\mathbf{k}(s)\| \leq R(S_k, S_{k+1}).$$

By (ii) the sequence of solutions curves $(\mathbf{w}^\mathbf{k})_{k \geq k_0}$ is such that

$$\sup_{s \in [S_k, S_{k+1}]} \|v(s) - \mathbf{w}^\mathbf{k}(s)\| \leq r_k.$$

On the other hand, by definition of Φ and $\mathbf{w}^\mathbf{k}$, we have

$$\Phi(S_{k+1}, \mathbf{w}^\mathbf{k}(S_{k+1})) \leq \lambda(T_{k+1})\Phi(S_k, \mathbf{w}^\mathbf{k}(S_k)) + \varepsilon(S_k, T_{k+1}).$$

Hence, by (iii) and (iv), for any $k \geq k_0$,

$$
\begin{aligned}
\Phi(S_{k+1}, v(S_{k+1})) &\leq \Phi(S_{k+1}, \mathbf{w}^\mathbf{k}(S_{k+1})) + L_\Phi \left\| v(S_{k+1}) - \mathbf{w}^\mathbf{k}(S_{k+1}) \right\| \\
&\leq \lambda(T_{k+1})\Phi(S_k, v(S_k)) + L_\Phi r_k + \varepsilon(S_k, T_{k+1}) \\
&= \lambda(T_{k+1})\Phi(S_k, v(S_k)) + \eta_{k+1}
\end{aligned}
$$

Clearly, $H_k \to 0$, by definition on λ. Calling $\Phi_k := \Phi(S_k, v(S_k))$ and $\lambda_k := \lambda(T_{k+1})$ we have $\Phi_k \to 0$ by Lemma 3.1. Now let v_* be a limit point of $v(s)$: $v_* = \lim_n v(s_n)$, for some sequence $s_n \uparrow_n +\infty$. Call $k(n) := \sup\{k \in \mathbb{N} : S_k \leq s_n\}$. For n large enough, $k(n) \geq k_0$ and

$$\Phi(s_n, v(s_n)) \leq \lambda(s_n - S_{k(n)})\Phi(S_{k(n)}, v(S_{k(n)}) + L_\Phi r_{k(n)} + \varepsilon(S_{k(n)}, s_n - S_{k(n)}) \to_{n \to +\infty} 0.$$

We therefore have

$$\Phi(s_n, v_*) \leq \Phi(s_n, v(s_n)) + L_\Phi \|v_* - v_n\| \to_{n \to +\infty} 0.$$

Consequently $v_* \in A$ and the proof is complete. □

3.2 A Lyapunov function for the differential inclusion induced by (1) We now focus on the particular case of Section 2.1 and prove that there exists a global Lyapunov function with respect to the so-called *consistency set*.

THEOREM 3.1 *Let $A = \{(x, y, \pi) \in M \mid \Pi(y) - \pi \leq 0\}$. There exists a global uniform Lyapunov function Φ relative to the compact set A and the non-autonomous differential inclusion*

$$\dot{\mathbf{w}}(s) \in F(s, \mathbf{w}(s)). \tag{5}$$

PROOF.   Let $\Phi : \mathbb{R}_+ \times M \to \mathbb{R}_+$ be defined by

$$\Phi(s, x, y, \pi) = \begin{cases} \tilde{\Pi}(y, \beta_{m(s)}) - \pi & \text{if } \tilde{\Pi}(y, \beta_{m(s)}) \geq \pi \\ 0 & \text{if } \tilde{\Pi}(y, \beta_{m(s)}) < \pi. \end{cases}$$

where

$$\tilde{\Pi} : Y \times \mathbb{R}_+^* \to \mathbb{R}, \ (y, \beta) \mapsto \max_{x \in X} \tilde{\pi}(x, y, \beta) = \tilde{\pi}(\mathbf{br}(y, \beta), y, \beta).$$

We prove that properties a') and b) (of respectively Remark 3.2 and Definition 3.2) hold. Notice that

$$A = \{(x, y, \pi) : g(x, y, \pi) = 0\} \quad \text{and} \quad \|g(x, y, \pi) - \Phi(s, x, y, \pi)\| \to_{s \to +\infty} 0$$

uniformly, where $g(x, y, \pi) := \max\{0, \Pi(y) - \pi\}$. Let t and T be positive real numbers and $\mathbf{w}(s) := (x(s), y(s), \pi(s))$ be a solution of the non-autonomous differential inclusion (5) on $[t, t + T]$, such that $\pi(s) \leq \tilde{\Pi}(y(s), \beta_{m(s)})$. Thus

$$\dot{y}(s) = \tau(s) - y(s), \ \ \dot{\pi}(s) = \pi(\mathbf{br}(y(s), \beta_{m(s)}), \tau(s)) - \pi(s),$$

where $\tau(s) \in Y, \ \forall s$. Let

$$\Psi(s) := \Phi(s, x(s), y(s), \pi(s)) = \tilde{\pi}\left(\mathbf{br}(y(s), \beta_{m(s)}), y(s), \beta_{m(s)}\right) - \pi(s).$$

Recall that $\beta_{m(s)}$ is piecewise constant on $[t, t + T]$. Hence, for almost every $s \in [t, t + T]$, we have

$$\begin{aligned}
\dot{\Psi}(s) &= \tilde{\pi}\left(\mathbf{br}(y(s), \beta_{m(s)}), \dot{y}(s), \beta_{m(s)}\right) - \dot{\pi}(s) \\
&= \tilde{\pi}\left(\mathbf{br}(y(s), \beta_{m(s)}), \tau(s), \beta_{m(s)}\right) - \tilde{\pi}\left(\mathbf{br}(y(s), \beta_{m(s)}), y(s), \beta_{m(s)}\right) \\
&\quad -\pi(\mathbf{br}(y(s), \beta_{m(s)}), \tau(s)) + \pi(s) \\
&\leq -\Psi(s) + \frac{1}{\beta_{m(s)}} \rho(\mathbf{br}(y(s), \beta_{m(s)})) \leq -\Psi(s) + \frac{1}{\beta_{m(s)}},
\end{aligned}$$

where we recall that ρ denotes the perturbation function. The first equality is obtained using the enveloppe theorem and the fact that $\tilde{\pi}$ is linear in its second argument. Since β_n is increasing, we therefore have, for almost every $s \in [t, t + T]$,

$$\dot{\Psi}(s) \leq -\left(\Psi(s) - \frac{1}{\beta_{m(t)}}\right).$$

Thus, applying Gronwall's lemma to the function $s \mapsto \Psi(s) - \frac{1}{\beta_{m(t)}}$, we obtain

$$\Psi(t + T) \leq e^{-T} \Psi(t) + \frac{1 - e^{-T}}{\beta_{m(t)}} \leq e^{-T} \Psi(t) + \frac{T}{\beta_{m(t)}}.$$

Consequently, Φ is a global uniform Lyapunov function with respect to A, which proves the result.   □

4. Proof of Theorem 1.2   We are now ready to prove our main result. We already proved that the interpolated random process induced by $(v_n)_n$ is almost surely a perturbed solution of the differential inclusion (5) with F given by (3) and $\delta(s) = c\overline{\gamma}(s)$, and that there exists a global uniform Lyapunov function Φ with respect to

$$A = \{(x, y, \pi) \in M \mid \Pi(y) - \pi \leq 0\},$$

see respectively Proposition 3.1 and Theorem 3.1. Therefore we now check that the assumptions of Proposition 3.2 hold. Be aware that we have not used the particular form of the parameter sequence $(\beta_n)_n$ so far. Recall that $\beta_n = n^\nu$, for some $\nu \in (0, 1)$.

Notice that $\gamma_n = 1/n$. Therefore we have $\tau_n \sim \log n$ and $m(s) = \mathcal{O}(e^s)$[1]. Recall that, given positive real numbers t and T, $\Delta(t, t + T)$ denotes the random variable

$$\sup_{h \in [0, T]} \int_t^{t+h} \overline{U}(s) ds.$$

The next lemma (proved in [8] or [1] for instance) gives an upper bound of the quantity $\mathbb{P}\left(\Delta(t, t + T) \geq \alpha\right)$.

---

[1] more precisely, $\frac{e-1}{e} e^s \leq m(s) \leq e^s - 1$

LEMMA 4.1 *There exists positive constants C and C' (depending on $\|U\|_\infty$) such that, for any $\alpha > 0$,*

$$\mathbb{P}\left(\Delta(t, t+T) \geq \alpha\right) \leq C \exp\left(\frac{-\alpha^2 e^t}{C'T}\right).$$

The set-valued map F is regular and $L(\cdot)$-Lipschitz, with the same Lipschitz constant as the map $(s, y) \mapsto \mathbf{br}(y, \beta_{m(s)})$. Hence $L(s) = L\beta_{m(s)}$, for some constant L (by property (ii) of $\mathbf{br}$). Hence, (by replacing ν by $\nu' > \nu$) we can assume without loss of generality that $L(s) \leq e^{\nu s}$. In the next proposition, we see that assumptions (i) and (ii) of Proposition 3.2 hold, if we choose $T_k = (\nu k)^{-1}$.

PROPOSITION 4.1 *If we choose $T_k := (\nu k)^{-1}$ there exist some constant $r > 1$ with the property that, with probability one, there exists $k_0 \in \mathbb{N}$ such that points (i) and (ii) of Proposition 3.2 are verified for v, with $r_k = k^{-r}$*

PROOF. Point (i) clearly holds. We now need to prove (ii). In this particular case, the quantity $R(S_k, S_{k+1})$ satisfies

$$R(S_k, S_{k+1}) \leq (\Delta(S_k, S_{k+1}) + c\overline{\gamma}(S_k)) \exp\left(\int_{S_k}^{S_{k+1}} L(\tau)d\tau\right).$$

By our choice of the sequence $(T_k)_k$, $\exp(\nu S_k) \leq \exp(1 + \log k) \leq 3k$. Hence

$$\exp\left(\int_{S_k}^{S_{k+1}} L(\tau)d\tau\right) \leq \exp(T_{k+1}e^{\nu S_{k+1}}) \leq C_0,$$

for some constant C_0 which depends on ν. Choose $r \in (1, \frac{\nu+1}{2\nu})$. By Lemma 4.1,

$$\mathbb{P}\left(\Delta(S_k, S_{k+1}) \exp\left(\int_{S_k}^{S_{k+1}} L(\tau)d\tau\right) \geq \frac{1}{2k^r}\right) \leq C \exp\left(\frac{-k^{-2r}e^{S_k}}{4C'C_0T_{k+1}}\right)$$

$$\leq C \exp\left(\frac{-k^{-2r+1/\nu}}{C'C_0\nu^{-1}(k+1)^{-1}}\right)$$

$$\leq C \exp\left(\frac{-k^{-2r+1+1/\nu}}{C_1'}\right)$$

for some positive constant C_1'. Now, since $\overline{\gamma}(S_k) \leq 2e^{-S_k} \leq 2k^{-1/\nu}$ and $r < 1/\nu$, we have for k large enough

$$c\overline{\gamma}(S_k) \exp\left(\int_{S_k}^{S_{k+1}} L(\tau)d\tau\right) \leq \frac{3c}{k^{1/\nu}} \leq \frac{1}{2k^r}.$$

Consequently, if we call A_k the event

$$\left\{(\Delta(S_k, S_{k+1}) + c\overline{\gamma}(S_k)) \exp\left(\int_{S_k}^{S_{k+1}} L(\tau)d\tau\right) \geq \frac{1}{k^r}\right\},$$

then

$$\mathbb{P}(A_k) \leq C \exp\left(\frac{-k^{-2r+1+1/\nu}}{C_1'}\right).$$

By an application of the Borel Cantelli lemma, with probability one, there exists $k_0 \in \mathbb{N}$ such that, for any $k \geq \mathbb{N}$,

$$R(S_k, S_{k+1}) \leq (\Delta(S_k, S_{k+1}) + c\overline{\gamma}(S_k)) \exp\left(\int_{S_k}^{S_{k+1}} L(\tau)d\tau\right) \leq \frac{1}{k^r},$$

which yields the result. $\qquad\square$

Consequently, points (i) and (ii) of Proposition 3.2 are almost surely satisfied for $k \geq k_0$, with $T_k = (\nu k)^{-1}$ and $r_k = k^{-r}$, $r > 1$. We now need to check points (iii) and (iv).

Let b be a positive constant and consider the map $\phi : Y \times [-\|\pi\|_\infty, \|\pi\|_\infty] \to \mathbb{R}_+$, given by

$$\phi(y, \pi) = \begin{cases} \tilde{\Pi}(y, b) - \pi & \text{if } \tilde{\Pi}(y, b) \geq \pi \\ 0 & \text{if } \tilde{\Pi}(y, b) < \pi. \end{cases}$$

Let (y, π) be such that $\tilde{\Pi}(y, b) > \pi$. Then, by Lemma 6.2 in [4] (see also [15]), we have

$$
\begin{aligned}
\frac{\partial}{\partial y}\phi(y, \pi)(h) &= \frac{\partial}{\partial y}\tilde{\pi}(\mathbf{br}(y, b), y, b)(h) \\
&= \pi(\mathbf{br}(y, b), h)
\end{aligned}
$$

and

$$
\frac{\partial}{\partial \pi}\phi(y, \pi) = -1.
$$

Thus

$$
|\phi(y, \pi) - \phi(y', \pi')| \le \|\pi\|_\infty \|y - y'\| + |\pi - \pi'|
$$

and ϕ is Lipschitz with Lipschitz constant that does not depend on b, which means that the map $v \mapsto \Phi(s, v)$ is uniformly Lipschitz.

We now prove point (iv). By Theorem 3.1, Φ is a global uniform Lyapunov function relative to

$$
A = \{(x, y, \pi) \in M \mid \Pi(y) - \pi \le 0\},
$$

with $\lambda(T) = e^{-T}$ and $\varepsilon(t, T) = \frac{T}{\beta_{m(t)}}$. Hence

$$
\eta_{k+1} = L_\Phi k^{-r} + \frac{T_{k+1}}{\beta_{m(S_k)}} \le L_\Phi k^{-r} + c\frac{T_{k+1}}{k},
$$

for some positive constant c. We have $\sum_i \eta_i < \infty$ and $H_k = e^{-\sum_{i=k_0}^k T_i}$, which converges to zero. Thus point (iv) is checked (see point b) of Lemma 5.1 for a proof).

As a consequence, Proposition 3.2 applies and

$$
\mathcal{L}((v(s))_{s>0}) \subset A
$$

almost surely. In particular

$$
\limsup_n e_n \le 0, \quad \text{almost surely}
$$

and Theorem 1.2 is proved.

5. Appendix

5.1 Sufficient conditions for Lemma 3.1, (ii) to hold

LEMMA 5.1 *Point (ii) of Lemma 3.1 is verified in the following cases:*

    a) $\lambda_k = \lambda < 1$ *and* $\lim_{k\to\infty} \eta_k = 0,$

    b) $\lim_{k\to\infty} H_k = 0$ *and* $\sum_i \eta_i < +\infty.$

PROOF.   For point a), $H_k = \lambda^k$ and we have

$$
\begin{aligned}
\tilde{H}_{k+k'} &= \lambda^{k+k'}\left(\sum_{i=1}^{k} H_i^{-1}\eta_i + \sum_{i=k+1}^{k+k'} H_i^{-1}\eta_i\right) \\
&\le \lambda^{k'}\max_{i=1,\dots,k}\eta_i + \eta_{k+1}\sum_{i=0}^{k'-1}\lambda^i \\
&\le \lambda^{k'}\max_{i=0,\dots,k}\eta_i + \eta_{k+1}\frac{1}{1-\lambda},
\end{aligned}
$$

which gives the result.

For the second point, remember that $(H_k)_k$ is a decreasing sequence. Hence

$$
\begin{aligned}
\tilde{H}_{k+k'} &\le H_{k+k'}\left(\sum_{i=1}^{k} H_i^{-1}\eta_i + H_{k+k'}^{-1}\sum_{i=k+1}^{k+k'} \eta_i\right) \\
&\le H_{k+k'}\left(\sum_{i=1}^{k} H_i^{-1}\eta_i\right) + \sum_{i=k+1}^{+\infty} \eta_i.
\end{aligned}
$$

Given $\varepsilon > 0$, by choosing k large enough, the second term is smaller than ε. Then we can pick k' large enough so that the first term is also smaller than ε and the proof is complete.   □

5.2 Stability of one-sided Lipschitz differential inclusions   Let $M \subset \mathbb{R}^d$. Consider a set-valued map $F : \mathbb{R}_+ \times M \rightrightarrows M$ taking values in the set of non-empty, compact, convex subsets of M. Given $a < b$, let us consider the non-autonomous differential inclusion (4):

$$\dot{\mathbf{w}}(s) \in F(s, \mathbf{w}(s)), \ \ s \in [a,b]. \tag{6}$$

For $A \subset M$ we let $F^{-1}(A) = \{(s, w) \in [a, b] \times M : \ F(s, w) \cap A \neq \emptyset\}$. We say that F is measurable if $F^{-1}(A)$ is measurable, for any closed set $A \subset M$. It is *upper semi-continuous* (USC) (resp. *lower semi-continuous* (LSC)) if, for any closed (resp. open) set $A \subset M$, $F^{-1}(A)$ is closed (resp. open) in $[a, b] \times M$. If M is compact, F is upper semi-continuous if and only if its graph is closed.

We now introduce a regularity condition:

DEFINITION 5.1 (RELAXED ONE-SIDED LIPSCHITZ) *we say that the set-valued map F is* Relaxed One-sided Lipschitz *(ROSL) on $[a, b] \times M$ if there exists an integrable map $L : [a, b] \to M$ such that, for any t, t' in $[a, b]$ $w, w' \in M$ and any $y \in F(t, w)$ there exists $y' \in F(t', w')$ with*

$$< w' - w \mid y' - y > \leq L(t) \|x' - x\|^2, \ \ \forall t \in [a, b].$$

REMARK 5.1 *If F is $L(\cdot)$-Lipschitz then it is $L(\cdot)$-ROSL.*

The question of existence of solutions to (4) has been studied extensively. One of the first result on the topic was proved by Filippov (see [9]) and says that if $F(\cdot, \cdot)$ is Hausdorff continuous on any closed set of $[a, b] \times M$ then, for any $w_0 \in M$, there exists a solution $\mathbf{w}(\cdot)$ of (4), with $\mathbf{w}(a) = x_0$. Under less restrictive assumptions, the same result still holds (see [21]; on the topic, see also [19]).

THEOREM 5.1 (OLECH, 1975) *Assume that F is regular. Then there exists a solution $\mathbf{w}(\cdot)$ of (4), with $\mathbf{w}(a) = w_0$.*

The following result will prove useful to establish Theorem 5.2.

LEMMA 5.2 *Let y be a continuously differentiable function on $[a, b]$ and f, g be non-negative, continuous maps. If, for every $s \in [a, b]$, $\|\dot{y}(s)\| \leq f(s)\|y(s)\| + g(s)$ then*

$$\|y(s)\| \leq \|y(a)\| \exp\left(\int_a^s f(\tau)d\tau\right) + \int_a^s g(u)\exp\left(\int_u^s f(\tau)d\tau\right)ds$$

PROOF.   Notice that

$$\|y(s)\| \leq \|y(a)\| + \int_a^s \|\dot{y}(u)\|du \leq \|y(a)\| + \int_a^s g(u)du + \int_a^s f(u)\|y(u)\|du$$

and apply the integral form of Gronwall's lemma.   □

In the remaining of this section, we assume that F is regular. The set of solution trajectories on $[a, b]$ (resp. starting in w_0) will be labelled $\mathcal{S}(a, b)$ (resp. $\mathcal{S}(w_0, a, b)$).

THEOREM 5.2 *Let $W : [a, b] \to M$ be an absolutely continuous function such that there exists a measurable map $\overline{v} : [a, b] \to M$ and a bounded measurable map $r : [a, b] \to \mathbb{R}_+$ which satisfy, for almost every $s \in [a, b]$,*

$$d(W(s), \overline{v}(s)) \leq r(s), \ \ \dot{W}(s) \in F(s, \overline{v}(s)).$$

Then

    a) *if F is ROSL with respect to the integrable function L, then there exists a solution $\mathbf{w} : [a, b] \to M$ of (4) such that $\mathbf{w}(a) = W(a)$ and*

$$\sup_{s \in [a,b]} \|\mathbf{w}(s) - W(s)\|^2 \leq \int_a^b \alpha(s) \exp\left(4\int_s^b L(\tau)d\tau\right)ds,$$

    where $\alpha(s) = 4L(s)r^2(s) + 4r(s)\|F\|_\infty$.

b) *if we now assume that F is Lipschitz continuous, with respect to L then the conclusions of a) trivially still hold and* **w** *can also be chosen such that*

$$\sup_{s\in[a,b]} \|\mathbf{w}(s) - W(s)\| \le \int_a^b r(s)L(s) \exp\left(\int_s^b L(\tau)d\tau\right) ds.$$

PROOF. We prove the first point. Consider the set-valued map $G : [a,b] \times M \rightrightarrows M$ given by

$$G(s,x) := \left\{ v \in F(s,w) : \ (w - W(s) \mid v - \dot{W}(s)) \le 2L(s)\|w - W(s)\|^2 + \frac{1}{2}\alpha(s) \right\}.$$

For any (s,w), the set $G(s,w)$ is non-empty. Indeed, by the ROSL condition, since $\dot{W}(s) \in F(s, \overline{v}(s))$, there exists $v \in F(s,w)$ such that

$$(w - \overline{v}(s) \mid v - \dot{W}(s)) \le L(s)\|w - \overline{v}(s)\|^2.$$

Hence we have

$$
\begin{aligned}
(w - W(s) \mid v - \dot{W}(s)) &\le L(s)\|w - \overline{v}(s)\|^2 + \|\overline{v}(s) - W(s)\|(\|v\| + \|\dot{W}(s)\|) \\
&\le 2L(s)\|W(s) - w\|^2 + 2L(s)r(s)^2 + 2r(s)\|F\|_\infty \\
&= 2L(s)\|W(s) - w\|^2 + \frac{1}{2}\alpha(s).
\end{aligned}
$$

Now clearly, the set $G(s,w)$ is compact and convex. The map $w \mapsto G(s,w)$ has a closed graph, for any $s \in [a,b]$. Finally It is measurable in s since every map involved is measurable. Consequently, there exists a solution to the non-autonomous differential inclusion

$$\dot{\mathbf{w}}(s) \in G(s, \mathbf{w}(s)),$$

with initial condition $\mathbf{w}(a) = W(a)$. In particular, **w** is a solution of (4) and we also have, for almost every s

$$(\mathbf{w}(s) - W(s) \mid \dot{\mathbf{w}}(s) - \dot{W}(s)) \le 2L(s)\|W(s) - \mathbf{w}(s)\|^2 + \frac{1}{2}\alpha(s).$$

Hence , for almost every s, we have

$$
\begin{aligned}
\frac{d}{ds}\|\mathbf{w}(s) - W(s)\|^2 &= 2(\mathbf{w}(s) - W(s) \mid \dot{\mathbf{w}}(s) - \dot{W}(s)) \\
&\le 4L(s)\|W(s) - \mathbf{w}(s)\|^2 + \alpha(s)
\end{aligned}
$$

and point a) follows from the differential form of Gronwall's lemma.

When the Lipschitz continuity holds, let us consider the set-valued map $H : [a,b] \times M \rightrightarrows M$ given by

$$H(s,w) := \left\{ v \in F(s,w) : \ \|v - \dot{W}(s)\| \le L(s)\|w - W(s)\| + L(s)r(s) \right\}.$$

The fact that H has non-empty values follows from Lipschitz continuity: given s and w, since $\dot{W}(s) \in F(s, \overline{v}(s))$, there exists $v \in F(s,w))$ such that

$$\|v - \dot{W}(s)\| \le L(s)\|w - \overline{v}(s)\| \le L(s)\left(\|w - W(s)\| + \|W(s) - \overline{v}(s)\|\right).$$

Hence $v \in H(s,w) \ne \emptyset$. Also $H(s,w)$ is convex and compact, the map $w \mapsto H(s,w)$ has a closed graph and $s \mapsto H(s,w)$ is measurable. Thus, there exists a solution **x** to the non-autonomous differential inclusion

$$\dot{\mathbf{w}}(s) \in H(s, \mathbf{w}(s)),$$

with initial condition $\mathbf{w}(a) = W(a)$. In particular, **w** is a solution of (4) and we also have, for almost every s

$$\|\dot{\mathbf{w}}(s) - \dot{W}(s)\| \le L(s)\|\mathbf{w}(s) - W(s)\| + L(s)r(s)$$

By Gronwall's lemma (see Lemma 5.2), we then have

$$\sup_{s\in[a,b]} \|\mathbf{w}(s) - W(s)\| \le \int_a^b L(s)r(s) \exp\left(\int_s^b L(\tau)d\tau\right) ds$$

and point b) is proved. $\square$

COROLLARY 5.1 *Let $v : [a, b] \to M$ be an absolutely continuous map. Assume that there exist measurable maps $\overline{v} : [a, b] \to M$, $\delta : [a, b] \to \mathbb{R}_+$ bounded and $\overline{U} : [a, b] \to M$ integrable such that, for almost every $s \in [a, b]$,*

$$\dot{v}(s) - \overline{U}(s) \in F(s, \overline{v}(s)), \quad \|v(s) - \overline{v}(s)\| \leq \delta(s).$$

Then if F is $L(\cdot)$-Lipschitz, there exists a solution $\mathbf{w}$ on $[a, b]$ such that $\mathbf{w}(a) = v(a)$ and

$$\sup_{s \in [a,b]} \|v(s) - \mathbf{w}(s)\| \leq R(a, b),$$

where

$$R(a, b) = \Delta(a, b) \exp\left(\int_a^b L(\tau)d\tau\right) + \sup_{s \in [a,b]} \delta(s) \left(\exp\left(\int_a^b L(\tau)d\tau\right) - 1\right) \tag{7}$$

and $\Delta(a, b) = \sup_{s \in [a,b]} \|\int_a^s \overline{U}(\tau)d\tau\|$.

PROOF. Define $W : [a, b] \to M$ by

$$W(s) := v(s) - \int_a^s \overline{U}(\tau)d\tau.$$

Clearly, W is absolutely continuous and, for any s for which v is differentiable, we have $\dot{W}(s) = \dot{v}(s) - \overline{U}(s) \in F(s, \overline{v}(s))$. Additionally,

$$\|W(s) - \overline{v}(s)\| \leq \|v(s) - \overline{v}(s)\| + \|\int_a^s \overline{U}(\tau)d\tau\| \leq \delta(s) + \left\|\int_a^s \overline{U}(\tau)d\tau\right\|.$$

By a direct application of Theorem 5.2 with $r(s) = \delta(s) + \left\|\int_a^s \overline{U}(\tau)d\tau\right\|$, there exists a solution $\mathbf{w}$ such that $\mathbf{w}(a) = W(a) = v(a)$ and

$$\sup_{s \in [a,b]} \|v(s) - \mathbf{w}(s)\| \leq \Delta(a, b) + \int_a^b L(s) \left(\delta(s) + \|\int_a^s \overline{U}(\tau)d\tau\|\right) \exp\left(\int_s^b L(\tau)d\tau\right) ds$$

$$\leq \Delta(a, b) + (\sup_{s \in [a,b]} \delta(s) + \Delta(a, b)) \int_a^b L(s) \exp\left(\int_s^b L(\tau)d\tau\right) ds \leq R(a, b).$$

$\square$

References

[1] M. Benaïm, *Dynamics of stochastic approximation algorithms*, Séminaire de probabilités de Strasbourg **33** (1999), 1–68.

[2] M. Benaïm and M.W. Hirsch, *Mixed Equilibria and Dynamical Systems Arising from Fictitious Play in Perturbed Games*, Games and Economic Behavior **29** (1999), no. 1-2, 36–72.

[3] M. Benaïm, J. Hofbauer, and S. Sorin, *Stochastic approximations and differential inclusions*, I. SIAM Journal on Optimization and Control **44** (2005), 328–348.

[4] _____, *Stochastic Approximations and Differential Inclusions. Part II: Applications*, Mathematics of Operations Research **31** (2006), 673–695.

[5] D. Blackwell, *Controlled random walks*, Proceedings of the International Congress of Mathematicians, vol. 3, 1954, pp. 336–338.

[6] G.W. Brown, *Iterative solution of games by fictitious play*, Activity analysis of production and allocation **13** (1951), no. 1, 374–376.

[7] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*, Cambridge Univ Pr, 2006.

[8] M. Duflo, *Random iterative models*, Springer Verlag, 1997.

[9] AF Filippov, *The existence of solutions of generalized differential equations*, Mathematical Notes **10** (1971), no. 3, 608–611.

[10] D.P. Foster and R.V. Vohra, *A randomization rule for selecting forecasts*, Operations Research (1993), 704–709.

[11] _____, *Asymptotic calibration*, Biometrika **85** (1998), no. 2, 379–390.

[12] D. Fudenberg and D. Kreps, *Learning mixed equilibria*, Games and Economic Behavior **5** (1993), no. 3, 320–367.

[13] D. Fudenberg and D.K. Levine, *Consistency and cautious fictitious play*, Journal of Economic Dynamics and Control **19** (1995), no. 5-7, 1065–1089.

[14] _____, *The Theory of Learning in Games*, MIT Press, 1998.

[15] _____, *Conditional Universal Consistency* 1*, Games and Economic Behavior **29** (1999), no. 1-2, 104–130.

[16] J. Hannan, *Approximation to Bayes risk in repeated play*, Contributions to the Theory of Games **3** (1957), 97–139.

[17] J.C. Harsanyi, *Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points*, International Journal of Game Theory **2** (1973), no. 1, 1–23.

[18] S. Hart and A. Mas-Colell, *A general class of adaptive strategies*, Journal of Economic Theory **98** (2001), no. 1, 26–54.

[19] CJ Himmelberg and FS Van Vleck, *Existence of solutions for generalized differential equations with unbounded right-hand side* 1*, Journal of Differential Equations **61** (1986), no. 3, 295–320.

[20] J. Hofbauer and W.H. Sandholm, *On the Global Convergence of Stochastic Fictitious Play*, Econometrica **70** (2002), no. 6, 2265–2294.

[21] C. Olech, *Existence of solutions of non-convex orientor fields*, Boll. Unione Mat. Ital. **11** (1975), 189–197.

[22] V. Perchet, *Approchabilité, Calibration et Regret dans les Jeux à Informations Partielles*, Ph.D. thesis, UPMC, 2010.

[23] Y. Viossat, S. Sorin, and J. Hofbauer, *Time average replicator and best reply dynamics*, Mathematics of Operations Research **34** (2009), no. 2, 263.