

Stochastic Processes for Bounded Noise*

Giovanni Colombo,[†] Paolo Dai Pra,[†] Vlastimil Krivan,[‡] and Ivo Vrkoč[§]

Abstract. The scalar differential inclusion

$$\dot{x} \in f(x) + g(x)u, \quad u \in [-1, 1], \quad x(0) = x_0 \quad (0.1)$$

is considered as a model of the dynamical system $\dot{x} = f(x)$ perturbed by the bounded noise $g(x)u$, $u \in [-1, 1]$, and the problem of constructing a nontrivial probability measure on the set \mathcal{S} of solutions to (0.1) is studied. In particular, it is shown that:

- (i) every Markov process whose probability measure is supported on \mathcal{S} is degenerate, in a sense to be specified (see Theorem 3.1);
- (ii) given a flow of probability measures μ_t on the reachable sets R_t of (0.1), satisfying a certain compatibility condition, a Markov process X_t is constructed such that its marginals are exactly μ_t and (0.1) is satisfied “from one side” (see Theorem 4.1); its finite-dimensional distributions are computed and the regularity of its sample paths is investigated (see Section 5.2);
- (iii) given a process of a type previously considered, another process Y_t is constructed through its finite-dimensional distributions, and its distribution is shown to be supported exactly on \mathcal{S} .

Finally, a model example is considered (see Section 7).

Key words. Scalar differential inclusions, Perturbed dynamical systems, Construction of Markov and non-Markov processes, Finite-dimensional distributions, Piecewise deterministic processes.

* Date received: July 18, 2000. Date revised: July 18, 2002. The work of G.C. was supported by EC Grant ERB-CIPA-CT-93-1554; that of V.K. by EC Grant ERB-CIPA-CT-92-0370 and GA ČR 201/98/0227. The stay of V.K. and I.V. at the Faculty of Biological Sciences was supported by MŠMT Grant No. VS96086. I.V. was partly supported by GA ČR Grant No. 201/95/0629.

[†] Dipartimento di Matematica Pura e Applicata, Università di Padova, via Belzoni 7, 35131 Padova, Italy. {colombo, daipra}@math.unipd.it.

[‡] Institute of Entomology, Academy of Sciences of the Czech Republic, Žitná 25, Praha, Czech Republic, and Faculty of Biological Sciences USB, Branišovská 31, 370 05 České Budějovice, Czech Republic. krivan@entu.cas.cz.

[§] Institute of Mathematics, Academy of Sciences of the Czech Republic, Žitná 25, Praha, Czech Republic, and Faculty of Biological Sciences USB, Branišovská 31, 370 05 České Budějovice, Czech Republic. Vrkoč@matsrv.math.cas.cz.

1. Introduction

This paper deals with the modeling of a dynamical system with uncertainty. We restrict ourselves to systems described by a scalar differential equation, with the uncertainty acting as a parameter within the equation. In this section, for the sake of clarity, we assume that the noise appears linearly in the equation, namely,

$$\dot{x} = f(x) + g(x)u. \quad (1.1)$$

Throughout the paper the general case

$$\dot{x} = F(x, u)$$

will always be treated. We think of (1.1) as the ordinary differential equation $\dot{x} = f(x)$, which we call the deterministic part of (1.1), perturbed by the “noise” $g(x)u$, u being a parameter possibly belonging to an infinite-dimensional space. We require f to be at least Lipschitz continuous, in order to have global existence and uniqueness for the unperturbed equation. A difficulty of the model is how to make $g(x)u$ represent a noise. There are several different approaches to this task (see, e.g. [6], [13] and [14]), each of them being possibly preferable in some situations.

The first model consists in taking a white noise process in place of u . This idea is classical and brought to innumerable applications. In this way, (1.1) is no longer an ordinary differential equation, because u is no longer well defined as a function: (1.1) becomes indeed a stochastic differential equation. The only informations which can be obtained from it are statistical properties of the solution, which is a stochastic process. This approach has been questioned, in some situations, because the statistical properties of the white noise may not be suitable to describe the observed noise. Some discussions of this problem in the field of theoretical population biology, for example, can be found in [9], [15], [1] and [16].

The second approach consists in considering (1.1) as a random differential equation. Now one chooses $u = u_t(\omega)$ as time functions, depending on a parameter ω in a probability space (Ω, \mathcal{F}, P) ; $u_t(\omega)$ must be regular enough to let (1.1) be an ordinary differential equation, for a.e. ω , and so generate a probability on the set of solutions. This approach seems to be the most natural and general one; on the other hand, it has not reached the same range of applications as the previous approach (see, however, [14]). The problem consists first in finding a process u_t which realizes the observed noise. Secondly, the joint distributions of the solution process can very rarely be computed (see [14]), and this process is almost never Markov. Therefore, the statistical properties of the solution process are difficult to derive.

The third approach may be called the “unknown deterministic noise”. The noise $u = u_t$ is thought of as an unknown function, possibly satisfying some constraints, but no statistical properties are assumed. A special but relevant case of the deterministic approach is when the noise is constrained to take a value in given sets (see [13], [6] and [10]). In particular, we assume that the noise is norm bounded, with an a priori given bound. In this setting, this approach consists in

taking u as an arbitrary measurable function, with values in a prescribed interval (or bounded set), say $[-1, +1]$. Now, for any choice of $u(\cdot)$, (1.1) is an ordinary differential equation; (1.1) can be seen as a control problem, or, equivalently, as the differential inclusion

$$\dot{x} \in f(x) + g(x)[-1, +1]. \quad (1.2)$$

This model is deterministic, if one thinks of the set of solutions as a whole, and it provides an estimate about where the actual trajectory may be found. Interest in this type of model has been increasing in the last years: this approach seems to be conceptually easy and rather general. However, it does not allow, as it is, any probabilistic treatment.

In this paper we explore the possibility of probabilistically describing a dynamical system perturbed by bounded noise, i.e. the (scalar) differential inclusion (1.2), under minimal statistical assumptions on it.

We start with a negative result (Theorem 3.1): any Markov process X_t such that a.s. its sample paths are solutions of (1.2) must be degenerate, in the sense that \dot{X}_t is a deterministic function, possibly time-dependent, of X_t .

Then two possibilities are followed. First (Sections 4 and 5) we consider the case where just one of the differential constraints of (1.2) is active, e.g. the upper one, but keep the requirement that the process we want to construct be Markov; moreover, we want to a priori assign the marginals of X_t ; this will be the only statistical assumption on (1.2) we make in this framework. We must relax the Lipschitz continuity of the sample paths, so that the differential constraint has to be interpreted in a suitable way. More precisely, we look for a Markov process X_t such that a.s.

$$X_{t+h} \leq \varphi_h^+(X_t), \quad \forall t, h \geq 0, \quad (1.3)$$

where $\varphi^+(y)$ is the maximal solution of (1.2) such that $\varphi_0^+(y) = y$. The existence statement reads: given a flow of densities $p_t(\xi)$ on the reachable set of (1.2) satisfying a necessary and sufficient compatibility condition with f and g , we construct a strong Markov process, which satisfies (1.3) and whose marginal densities are exactly $p_t(\xi)$. The idea of the construction came from a natural definition of joint distributions (see (4.7)), taking into account reachable sets of (1.2); then, from the two-dimensional distributions one can compute the transition probabilities, and prove the existence of the process by checking the Chapman–Kolmogorov identity. Moreover, we show (under some regularity of p_t) that all trajectories are càdlàg and only finitely many downwards jumps may occur, in any interval $0 < s < t$; moreover, X_t behaves deterministically between jumps, in the sense that if T_1 and T_2 are two subsequent jump times, for $t \in [T_1, T_2)$ it holds that $X_t = \varphi_{t-T_1}^+(X_{T_1})$. In other words, X_t is a piecewise deterministic process. However, it has some properties which distinguish it among all such processes (see [5, Section 2], [8, Section III] and [7]): first, it is not necessarily homogeneous, second, all joint distributions as well as the transition probabilities have an explicit expression. Knowing all joint distributions permits the explicit computation of the probability that X_t remains below a solution of (1.2), as well as an approximation of the probability of extinction at a given time (Section 5.3, Section 7). Obviously, an

entirely symmetric construction can be performed with the minimal solution in place of the maximal.

The second possibility is giving up the Markov property, and constructing a process whose trajectories are solutions of (1.2). Of course, there are many such processes, at least for particular choices of f and g ; for example, the telegraph (or Kac) process, whose trajectories are a.s. polygonal solutions of $\dot{x} \in \{-1, 1\}$. Our contribution is the construction (Section 6), for any process X_t satisfying (1.3), of a process Y_t associated with it, whose distribution Q is supported on the whole set of solutions of (1.2), and such that all of its joint distributions are known. To the best of our knowledge, Y_t is the only nontrivial stochastic process with Lipschitz trajectories whose finite-dimensional distributions have an explicit expression. This fact permits us to compute the law of the derivative process \dot{Y}_t , as well as to compute the probability that Y_t remains below a given solution of (1.2), and to approximate the extinction probability.

Finally (Section 7), some explicit computations for the model case $\dot{x} \in [-1, 1]$, $x(0) = 0$ are presented, both for X_t and for Y_t ; Section 2 contains some basic notations and results.

In order for these processes to be effective tools for modeling dynamical systems with bounded noise (only bounded from above for the process X_t), a deeper analysis of them is desirable. In particular, laws of jump times of X_t and of Y_t , as well as the construction of X_t given jump times and of Y_t with pre-assigned marginals, are facts of some interest. Forthcoming papers will be dedicated to the above properties, together with a construction of both X_t and Y_t with given marginals without density. Moreover, an interpretation of the process Y_t as driven by X_t will be provided, so that Y_t may model, for example, the evolution of a population depending on some good X_t , subject to random catastrophies.

2. Preliminaries

First, a few symbols and notations are presented. If Γ is a multifunction from a set A into a set B , we denote its graph by $\text{graph}(\Gamma) = \{(x, y) \in A \times B : y \in \Gamma(x), x \in A\}$. We write $1_A(\cdot)$ as the characteristic function of a set A . Given $a, b \in \mathbb{R}$ we write $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. The unit mass concentrated at $z \in \mathbb{R}$ is denoted by $\delta_z(dx)$. The product σ -field generated by Lebesgue-measurable and Borel-measurable subsets of \mathbb{R} is indicated by $\mathcal{L} \otimes \mathcal{B}$. The convex hull of a set A is denoted by $\text{co } A$.

Lemma 2.1. *Let U be a compact metric space, and let $F: \mathbb{R} \times U \rightarrow \mathbb{R}$ be a continuous function, with $F(\cdot, u)$ Lipschitz uniformly in u . Set $F^+(x) = \max_{u \in U} F(x, u)$, $F^-(x) = \min_{u \in U} F(x, u)$. Then the maps F^+, F^- are Lipschitz, with the same constant as F .*

Proof. Assume F to be L -Lipschitz and fix $x, y \in \mathbb{R}$. Let u^+, v^+ be such that, respectively, $F^+(x) = F(x, u^+)$, $F^+(y) = F(y, v^+)$. Assume that $F^+(x) \geq F^+(y)$.

Then $F^+(x) - F^+(y) \leq F^+(x) - F(y, u^+) \leq L|x - y|$. If $F^+(x) < F^+(y)$, it suffices to interchange x, y, u^+, v^+ . The same argument holds for F^- . ■

Consider the (scalar) differential inclusion

$$\dot{x} = F(x, u), \quad u \in U, \tag{2.1}$$

with initial condition

$$x(0) = y. \tag{2.2}$$

Fix $T > 0$ and call $\mathcal{S}(y)$ the set of Carathéodory solutions of (2.1), (2.2) on the interval $[0, T]$, i.e. the set of all absolutely continuous functions which satisfy (2.2) and, for a.e. t , (2.1); we recall that $\mathcal{S}(y)$ is bounded in $\mathcal{C}^0(0, T)$, endowed with the sup-norm topology; the reachable set at time t is indicated by $R_t(y)$. When the dependence on y is dropped, we mean that $y = x_0$ is the initial condition, fixed once for all. The minimal and maximal solutions of (2.1), (2.2), which (for $t > 0$) are the solutions of $\dot{x} = F^-(x)$ and of $\dot{x} = F^+(x)$, $x(0) = y$, respectively, are denoted by $\varphi^-(y)$, $\varphi^+(y)$. We observe that both the maximal and the minimal solutions satisfy the semigroup property $\varphi_t^+(\varphi_s^+(y)) = \varphi_{t+s}^+(y)$. Moreover, maximal and minimal solutions can be considered also backwards in time, by interchanging F^+ with F^- ; in particular, if $s < t$, $\varphi_{t-s}^+(\varphi_{s-t}^-(x)) = x$ for all x .

The following simple lemma is the fundamental tool for our constructions.

Lemma 2.2. *Let $0 < t_1 < \dots < t_n \leq T$, $x_1, \dots, x_n \in \mathbb{R}$ be such that*

$$x_i \geq \varphi_{t_i}^- \quad \text{for all } i = 1, \dots, n. \tag{2.3}$$

Then problem (2.1) with the initial condition

$$x(0) = x_0 \tag{2.4}$$

and the constraints

$$x(t_i) \leq x_i, \quad \forall i = 1, \dots, n, \tag{2.5}$$

has a set of Carathéodory solutions with a unique maximal element, denoted by $\varphi_t^+(t_1, \dots, t_n; x_1, \dots, x_n)$, which is nondecreasing in each of the variables x_1, \dots, x_n and is Lipschitz in $(t_1, \dots, t_n; x_1, \dots, x_n)$. Symmetrically, if $x_i \leq \varphi_{t_i}^+$ the problem (2.1), (2.4) with the constraints

$$x(t_i) \geq x_i, \quad \forall i = 1, \dots, n,$$

has a set of Carathéodory solutions with a unique minimal element, denoted by $\varphi_t^-(t_1, \dots, t_n; x_1, \dots, x_n)$, which is nondecreasing in each of the variables x_1, \dots, x_n and is Lipschitz in $(t_1, \dots, t_n; x_1, \dots, x_n)$.

Proof. Set $t_0 = 0$ and define, for $i = 0, \dots, n$, $x^i(\cdot)$ to be the maximal Carathéodory solution of (2.1), with the initial condition $x(t_i) = x_i$, i.e. the solution of

$$\dot{x} = F^-(x), \quad x(t_i) = x_i, \quad \text{for } t < t_i,$$

and of

$$\dot{x} = F^+(x), \quad x(t_i) = x_i, \quad \text{for } t > t_i.$$

Set

$$\varphi_t^+(t_1, \dots, t_n; x_1, \dots, x_n) = \min\{x^i(t) : i = 0, \dots, n\};$$

then φ^+ is clearly a Carathéodory solution of (2.1), which satisfies (2.5); by (2.3), (2.2) holds too. Moreover, φ^+ is maximal, because if ψ is any solution of (2.1), (2.2), (2.5), then $\psi(t) \leq x^i(t)$ for all $t \in [0, T]$ and for all $i = 0, \dots, n$. Uniqueness and monotonicity follow directly from maximality, while Lipschitz continuity is obtained from the Lipschitz dependence of x^i on the initial conditions. The proof for the minimal solution goes along the same lines. ■

3. Markov Processes with Absolutely Continuous Trajectories

The statement “every homogeneous strong Markov process on the real line with a.s. continuous trajectories is a diffusion” can be found in several references (see, e.g. [12]). This fact suggests that a homogeneous strong Markov process with Lipschitz, or absolutely continuous, trajectories must be in some sense degenerate. Since we deal with not necessarily homogeneous processes, we find it easier to prove directly a simple degeneracy property, which forbids Markovianity to processes in \mathbb{R}^n which may be reasonable candidates to model bounded noise.

Theorem 3.1. *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, X_t, P)$ be an \mathbb{R}^n -valued Markov process such that, a.s., $t \mapsto X_t$ is absolutely continuous. Let $T > 0$ be fixed and let*

$$C = \left\{ (\omega, t) \in \Omega \times [0, T] : \lim_{h \rightarrow 0^+} \frac{X_{t-h}(\omega) - X_t(\omega)}{-h} \text{ exists} \right\}.$$

Then $(P \otimes \lambda)(C) = T$ and there exists an \mathcal{F}_t -adapted stochastic process V_t such that $\dot{X}_t(\omega) = V_t(\omega)$ for all $(\omega, t) \in C$. Moreover, there exists $\Gamma: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, with $\Gamma(t, \cdot)$ Borel-measurable for all t , such that a.s. X_t is a Carathéodory solution of the ordinary differential equation

$$\dot{u} = \Gamma(t, u). \tag{3.1}$$

Corollary 3.1. *Under the same assumptions of Theorem 3.1, for all t, x it holds that*

$$P(V_t \in y + dy \mid X_t = x) = \delta_{\Gamma(t, x)}(dy).$$

Therefore, there cannot exist any Markov process associated with a differential inclusion, in the sense that the support of $P(\dot{X}_t \in y + dy \mid X_t)$ can be multivalued only on a (ω, t) set of $(P \otimes \lambda)$ -measure zero.

Proof of the Theorem. By the continuity of the trajectories, the process X_t is jointly measurable. Let $C_t = \{\omega : (\omega, t) \in C\}$, $C_\omega = \{t : (\omega, t) \in C\}$ and $d\hat{P} = dP \otimes d\lambda$. Observe that C_t is \mathcal{F}_t -measurable for all t . By the absolute continuity

assumption, $\lambda(C_\omega) = T$ for P -almost all ω . Therefore, by Fubini's theorem $\hat{P}(C) = \int \lambda(C_\omega) dP = \int_0^T P(C_t) dt$, whence $\hat{P}(C) = T$ and $P(C_t) = 1$ for a.e. $t \in [0, T]$. Set

$$V_t(\omega) = \begin{cases} \lim_{n \rightarrow \infty} \frac{X_{t-1/n}(\omega) - X_t(\omega)}{-1/n} & \text{if } \omega \in C_t, \\ 0 & \text{if } \omega \notin C_t; \end{cases} \quad (3.2)$$

by the above remark, V_t is \mathcal{F}_t -adapted. Now, since X_t is Markov, for all $t, h > 0$,

$$\frac{X_{t+h} - X_t}{h} \perp_{X_t} \mathcal{F}_t,$$

where \perp_{X_t} means conditional independence given X_t ; by passing to the limit as $h \rightarrow 0$, we obtain for all t ,

$$V_t \perp_{X_t} \mathcal{F}_t.$$

Since V_t is \mathcal{F}_t -measurable, this implies that V_t must be $\sigma(X_t)$ -measurable. Therefore, there exists $\Gamma: [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, Borel-measurable in the second variable, such that, for all t , $V_t = \Gamma(t, X_t)$. Fix now ω such that $X(\omega)$ is absolutely continuous. Since $\dot{X}_t(\omega) = V_t(\omega)$ for a.e. $t \in C_\omega$, it follows that $X(\omega)$ is a solution of (3.1). ■

4. Construction of the Markov Process

We consider the initial value problem

$$\begin{cases} \dot{x}(t) = F(x, u), & u \in U, \\ x(0) = x_0, \end{cases} \quad (4.1)$$

together with a flow $\{\mu_t : t > 0\}$ of locally finite measures on \mathbb{R} . We recall that R_t denotes the reachable set of (4.1) in time t . The following assumptions hold:

- (H1) F is continuous, $F(\cdot, u)$ is Lipschitz, uniformly with respect to u , and $F^+ - F^- > 0$ on $\bigcup_{t>0} R_t$;
- (H2) the measures μ_t admit a density $p_t(\cdot)$, positive on $(\varphi_t^-, \varphi_t^+)$; the functions $p_t(x)$ are $(\mathcal{L} \otimes \mathcal{B})$ -measurable in (t, x) .

Remark. Assumption (H1) is satisfied if $F(x, u) = f(x) + g(x)u$, $u \in [-1, +1]$, f, g are Lipschitz, and $g(x) > 0$ on $\bigcup_{t>0} R_t$.

We set, for given times $0 = t_0 < t_1 < \dots < t_n \leq T$ and points x_1, \dots, x_n ,

$$\chi_0 = x_0, \quad \chi_{i+1} = \min\{x_{i+1}, \varphi_{t_{i+1}-t_i}^+(\chi_i)\}, \quad i = 0, \dots, n-1.$$

We begin by defining a candidate to be a probability kernel.

Definition 4.1. We define, for all $0 < s \leq t$, a.e. $y \in (\varphi_s^-, \varphi_s^+)$,

$$A(t, y, s) = \frac{p_t(\varphi_{t-s}^+(y))}{p_s(y)} \frac{\mu_s(\varphi_s^-, y]}{\mu_t(\varphi_t^-, \varphi_{t-s}^+(y))]} \frac{d}{dy} \varphi_{t-s}^+(y), \quad (4.2)$$

and assume

$$A(t, y, s) \leq 1, \quad \forall t, s, \quad \text{a.e. } y. \quad (4.3)$$

We set also $A(t, x_0, 0) = 0$ for all $t > 0$. We define also, for $0 < s < t$,

$P(dx, t; y, s)$

$$= A(t, y, s) \delta_{\varphi_{t-s}^+(y)}(dx) + [1 - A(t, y, s)] 1_{(\varphi_t^-, \varphi_{t-s}^+(y))}(x) \frac{\mu_t(dx)}{\mu_t(\varphi_t^-, \varphi_{t-s}^+(y))}, \quad (4.4)$$

for a.e. $y \in (\varphi_s^-, \varphi_s^+)$,

$$P(dx, t; x_0, 0) = 1_{(\varphi_t^-, \varphi_t^+)}(x) \frac{\mu_t(dx)}{\mu_t(\varphi_t^-, \varphi_t^+)}, \quad (4.5)$$

and

$$P(dx, t; y, s) = \begin{cases} \delta_{\varphi_{t-s}^+(y)}(dx) & \text{for } y \geq \varphi_s^+, \\ \delta_{\varphi_{t-s}^-(y)}(dx) & \text{for } y \leq \varphi_s^-. \end{cases} \quad (4.6)$$

Remark. The requirement (4.3), which is necessary and sufficient for $P(dx, t; y, s)$ to be a probability measure, since $A \geq 0$ for all (t, y, s) , is a kind of compatibility between the dynamics and the flow of measures; sufficient conditions for it are given in Section 5.1. Observe that the definition of $P(dx, t; y, s)$ makes sense because $\varphi_t^+(y)$ is a.e. differentiable with respect to y (see Lemma 2.1) and because the reachable set never collapses to a singleton.

The construction is carried out by showing that the above defined probability kernels $P(dx, t; y, s)$ satisfy the Chapman–Kolmogorov identity.

Theorem 4.1. *Let $x_0 \in \mathbb{R}$, F and μ_t be given satisfying (H1), (H2) and (4.3). Then there exists a Markov process X_t , which satisfies the following requirements:*

- (M1) $X_0 = x_0$ a.s.;
- (M2) for all $t > 0$, $B \subset \mathbb{R}$ measurable,

$$P\{X_t \in B\} = \frac{\mu_t(B \cap (\varphi_t^-, \varphi_t^+))}{\mu_t(\varphi_t^-, \varphi_t^+)};$$

- (M3) for all $t, h \geq 0$

$$P(X_{t+h} \leq \varphi_h^+(X_t)) = 1.$$

Moreover, the joint distributions of the process are given by

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = \prod_{i=1}^n \frac{\mu_{t_i}(\varphi_{t_i}^-, \chi_i]}{\mu_{t_i}(\varphi_{t_i}^-, \varphi_{t_i-t_{i-1}}^+(\chi_{i-1}))}. \quad (4.7)$$

Before proving the above statement we need a lemma, whose proof is a simple computation, and is omitted.

Lemma 4.1. *Let $G: \mathbb{R} \rightarrow \mathbb{R}$ be bounded and measurable. Then, for a.e. $y \in (\varphi_s^-, \varphi_s^+)$,*

$$\int G(x)P(dx, t; y, s) = \frac{1}{p_s(y)} \frac{d}{dy} \left[\frac{\mu_s(\varphi_s^-, y]}{\mu_t(\varphi_t^-, \varphi_{t-s}^+(y))} \int_{\varphi_t^-}^{\varphi_{t-s}^+(y)} G(x)\mu_t(dx) \right].$$

Proof. *Step 1.* We show first that the probability kernels $P(dx, t; y, s)$ satisfy the Chapman–Kolmogorov identity, i.e. for every G bounded and Borel measurable $0 \leq \tau < s < t, z \in \mathbb{R}$,

$$\int \left[\int G(x)P(dx, t; y, s) \right] P(dy, s; z, \tau) = \int G(x)P(dx, t; z, \tau). \quad (4.8)$$

By Lemma 4.1, recalling the semigroup property of the maximal solutions φ^+ , it holds that

$$\begin{aligned} & \int \left[\int G(x)P(dx, t; y, s) \right] P(dy, s; z, \tau) \\ &= \frac{1}{p_\tau(z)} \frac{d}{dz} \left\{ \frac{\mu_\tau(\varphi_\tau^-, z]}{\mu_s(\varphi_s^-, \varphi_{s-\tau}^+(z))} \int_{\varphi_s^-}^{\varphi_{s-\tau}^+(z)} \left[\int G(x)P(dx, t; y, s) \right] p_s(y) dy \right\} \\ &= \frac{1}{p_\tau(z)} \frac{d}{dz} \left\{ \frac{\mu_\tau(\varphi_\tau^-, z]}{\mu_s(\varphi_s^-, \varphi_{s-\tau}^+(z))} \right. \\ & \quad \left. \times \int_{\varphi_s^-}^{\varphi_{s-\tau}^+(z)} \frac{d}{dy} \left[\frac{\mu_s(\varphi_s^-, y]}{\mu_t(\varphi_t^-, \varphi_{t-s}^+(y))} \int_{\varphi_t^-}^{\varphi_{t-s}^+(y)} G(x)\mu_t(dx) \right] dy \right\} \\ &= \frac{1}{p_\tau(z)} \frac{d}{dz} \left\{ \frac{\mu_\tau(\varphi_\tau^-, z]}{\mu_s(\varphi_s^-, \varphi_{s-\tau}^+(z))} \frac{\mu_s(\varphi_s^-, \varphi_{s-\tau}^+(z))}{\mu_t(\varphi_t^-, \varphi_{t-\tau}^+(z))} \int_{\varphi_t^-}^{\varphi_{t-\tau}^+(z)} G(x)\mu_t(dx) \right\}. \end{aligned}$$

The proof is concluded.

Step 2. We call X_t the Markov process whose transition probability is given by (4.4). Then (M1)–(M3) are immediately implied by the expression of the transition probability.

Step 3. Property (4.7) holds.

Proof. We proceed by induction on n . For $n = 1$ the result follows from the fact that, for $t > 0$,

$$P(dx, t; x_0, 0) = \frac{\mu_t(dx)}{\mu_t(\varphi_t^-, \varphi_t^+)}.$$

For general n , notice that

$$\begin{aligned}
& F_{t_1 \dots t_n}(x_1, \dots, x_n) \\
&= \frac{1}{\mu_{t_1}(\varphi_{t_1}^-, \varphi_{t_1}^+)} \int_{\varphi_{t_1}^-}^{\chi_1} P(X_{t_2} \leq x_2, \dots, X_{t_n} \leq x_n \mid X_{t_1} = \xi_1) \mu_{t_1}(d\xi_1) \\
&= (\text{by the Markov property}) \\
&= \frac{1}{\mu_{t_1}(\varphi_{t_1}^-, \varphi_{t_1}^+)} \\
&\quad \cdot \int_{\varphi_{t_1}^-}^{\chi_1} \left[\int_{\varphi_{t_2}^-}^{\chi_2} P(X_{t_3} \leq x_3, \dots, X_{t_n} \leq x_n \mid X_{t_2} = \xi_2) P(d\xi_2, t_2; \xi_1, t_1) \right] \mu_{t_1}(d\xi_1) \\
&= (\text{by Lemma 4.1}) \\
&= \frac{1}{\mu_{t_1}(\varphi_{t_1}^-, \varphi_{t_1}^+)} \int_{\varphi_{t_1}^-}^{\chi_1} \frac{1}{p_{t_1}(\xi_1)} \frac{d}{d\xi_1} \left[\frac{\mu_{t_1}(\varphi_{t_1}^-, \xi_1]}{\mu_{t_2}(\varphi_{t_2}^-, \varphi_{t_2-t_1}^+(\xi_1))} \right. \\
&\quad \cdot \left. \int_{\varphi_{t_2}^-}^{x_2 \wedge \varphi_{t_2-t_1}^+(\xi_1)} P(X_{t_3} \leq x_3, \dots, X_{t_n} \leq x_n \mid X_{t_2} = \xi_2) \mu_{t_2}(d\xi_2) \right] p_{t_1}(\xi_1) d\xi_1 \\
&= \frac{1}{\mu_{t_1}(\varphi_{t_1}^-, \varphi_{t_1}^+)} \frac{\mu_{t_1}(\varphi_{t_1}^-, \chi_1]}{\mu_{t_2}(\varphi_{t_2}^-, \varphi_{t_2-t_1}^+(\chi_1))} \\
&\quad \cdot \int_{\varphi_{t_2}^-}^{\chi_2} P(X_{t_3} \leq x_3, \dots, X_{t_n} \leq x_n \mid X_{t_2} = \xi_2) \mu_{t_2}(d\xi_2) \\
&= \frac{\mu_{t_1}(\varphi_{t_1}^-, \chi_1]}{\mu_{t_1}(\varphi_{t_1}^-, \varphi_{t_1}^+)} \frac{\mu_{t_2}(\varphi_{t_2}^-, \varphi_{t_2}^+]}{\mu_{t_2}(\varphi_{t_2}^-, \varphi_{t_2-t_1}^+(\chi_1))} F_{t_2 \dots t_n}(\chi_2, x_3, \dots, x_n) \\
&= (\text{by inductive assumption and noticing that } \chi_2 \leq \varphi_{t_2}^+) \\
&= \frac{\mu_{t_1}(\varphi_{t_1}^-, \chi_1]}{\mu_{t_1}(\varphi_{t_1}^-, \varphi_{t_1}^+)} \frac{\mu_{t_2}(\varphi_{t_2}^-, \varphi_{t_2}^+]}{\mu_{t_2}(\varphi_{t_2}^-, \varphi_{t_2-t_1}^+(\chi_1))} \frac{\mu_{t_2}(\varphi_{t_2}^-, \chi_2]}{\mu_{t_2}(\varphi_{t_2}^-, \varphi_{t_2}^+)} \prod_{i=3}^n \frac{\mu_{t_i}(\varphi_{t_i}^-, \chi_i]}{\mu_{t_i}(\varphi_{t_i}^-, \varphi_{t_i-t_{i-1}}^+(\chi_{i-1}))} \\
&= \prod_{i=1}^n \frac{\mu_{t_i}(\varphi_{t_i}^-, \chi_i]}{\mu_{t_i}(\varphi_{t_i}^-, \varphi_{t_i-t_{i-1}}^+(\chi_i))}.
\end{aligned}$$

The proof is complete. ■

Remark. For $p_t \equiv 1$, the one-dimensional distribution becomes (obviously), for $x_1 \in R_{t_1}$,

$$F_{t_1}(x_1) = \frac{x_1 - \varphi_{t_1}^-}{\varphi_{t_1}^+ - \varphi_{t_1}^-},$$

i.e. if $x_1 \in R_{t_1}$, $P\{X_{t_1} \leq x_1\}$ is the ratio between the length of the interval $[\varphi_{t_1}^-, x_1]$ and the length of the reachable set at time t_1 , otherwise it is 0 or 1; the two-

dimensional distribution has a similar form:

$$F_{t_1 t_2}(x_1, x_2) = F_{t_1}(x_1) \frac{x_2 \wedge \varphi_{t_2-t_1}^+(x_1) - \varphi_{t_2}^-}{\varphi_{t_2-t_1}^+(x_1) - \varphi_{t_2}^-},$$

i.e. if x_2 is reachable from $[\varphi_{t_1}^-, x_1]$ in time $t_2 - t_1$ (we assume $x_1 \in \text{int } R_{t_1}$), $P\{X_{t_2} \leq x_2 \mid X_{t_1} \leq x_1\}$ equals the ratio between the length of the interval $[\varphi_{t_2}^-, x_2]$, and the length of the reachable set at time $t_2 - t_1$ from $[\varphi_{t_1}^-, x_1]$, otherwise it is 1. Higher-dimensional distributions behave analogously.

5. Study of the Markov Process X_t

5.1. Condition (4.3)

We begin by observing a necessary condition that any stochastic process satisfying (1.3) must enjoy.

Proposition 5.1. *Let $(\Omega, \mathcal{F}, P, X_t)$ be a stochastic process satisfying a.s. (1.3) for all $h, t \geq 0$. Then, for all $y \in (\varphi_t^-, \varphi_t^+)$,*

$$P(X_{t+h} \leq \varphi_h^+(y)) \geq P(X_t \leq y). \quad (5.1)$$

Proof. By Lemma 2.2, the event $\{X_t \leq y\}$ is contained in $\{X_{t+h} \leq \varphi_h^+(y)\}$. ■

It is easy to see that condition (4.3) implies (5.1).

Remark 5.1 (Sufficient Conditions for (4.3)). Condition (4.3) is equivalent to

$$k: y \mapsto \frac{\mu_t(\varphi_t^-, \varphi_{t-s}^+(y))}{\mu_s(\varphi_s^-, y]} \quad \text{nonincreasing.}$$

Observe that k is the ratio between the measure of the reachable set in time $t - s$ from the interval $(\varphi_s^-, y]$, and the measure of that interval.

Assume now that F^+ is of class \mathcal{C}^1 (e.g. $F = f + g$, f, g of class \mathcal{C}^1), and let $p_t \equiv 1$. Since, for all $y \in (\varphi_s^-, \varphi_s^+)$, $A(s, y, s) = 1$, a sufficient condition for (4.3) is

$$\frac{\partial}{\partial t} A(t, y, s) \leq 0, \quad \forall t \geq s, \quad \forall y \in (\varphi_s^-, \varphi_s^+). \quad (5.2)$$

It holds that

$$\begin{aligned} \frac{\partial}{\partial t} A(t, y, s) &= \frac{\partial}{\partial t} \left(\frac{y - \varphi_s^-}{\varphi_{t-s}^+(y) - \varphi_t^-} \exp \left(\int_0^{t-s} F^{+'}(\varphi_\tau^+(y)) d\tau \right) \right) \\ &= (y - \varphi_s^-) \exp \left(\int_0^{t-s} F^{+'}(\varphi_\tau^+(y)) d\tau \right) \\ &\quad \cdot \frac{F^{+'}(\varphi_{t-s}^+(y))(\varphi_{t-s}^+(y) - \varphi_t^-) - F^+(\varphi_{t-s}^+(y)) + F^-(\varphi_t^-)}{(\varphi_{t-s}^+(y) - \varphi_t^-)^2}. \end{aligned}$$

Thus, recalling (H1), a sufficient condition for (5.2) is $y \mapsto F^+(y)$ concave down, in particular, $F(\cdot, u)$ is linear.

We observe also that (4.3) is violated, for t close to the explosion time, in the case $F(x, u) = x^2 + u$, $u \in [-1, 1]$, $p_t \equiv 1$, $x_0 = 0$; this example does not satisfy the assumption of the global existence of solutions to (4.1), but can be easily localized.

Remark 5.2 (Existence for Small Time). Condition (4.3) is satisfied by any F, p_t regular enough, provided $t > 0$ is sufficiently small. Indeed, (4.3) is equivalent to

$$\frac{\mu_s(\varphi_s^-, y]}{p_s(y)} \frac{d}{dy} \varphi_{t-s}^+(y) \leq \frac{\mu_t(\varphi_t^-, \varphi_{t-s}^+(y))]}{p_t(\varphi_{t-s}^+(y))}. \quad (5.3)$$

Assume F^+ and p_t are of class \mathcal{C}^1 on $\bigcup_{t>0} \{t\} \times (\varphi_t^-, \varphi_t^+)$ (the assumption on F^+ holds if, e.g. $F(x, u) = f(x) + g(x)u$ and both f and g are \mathcal{C}^1), and that $\liminf_{t \rightarrow 0^+, x \rightarrow 0, x \in R_t} p_t(x) > 0$, $(d/dx)F^+$ is bounded around x_0 , and $F^+(x_0) > F^-(x_0)$. Set $m(t) = \max\{(d/dy)F^+(y) : y \in R_t\}$. By developing around $t = s$, for s close to 0, (5.3) is equivalent to

$$\begin{aligned} & (y - \varphi_s^- + o(y - \varphi_s^-))(1 + m(t-s)(t-s) + o(t-s)) \\ & \leq y - \varphi_s^- + (t-s)F^+(y) - (t-s)F^-(\varphi_s^-) + o(t-s), \end{aligned}$$

which is implied by $(y - \varphi_s^-)m(t-s) < F^+(y) - F^-(\varphi_s^-)$. Since $y \in (\varphi_s^-, \varphi_s^+)$, the above inequality is satisfied for all $s > 0$ sufficiently small.

5.2. Regularity of Sample Paths

The trajectories of the process are a.s. continuous at $t = 0$ by construction. We state also the following:

Theorem 5.1. *Assume that $(t, x) \mapsto p_t(x)$ is continuous on $\bigcup_{t \in (0, T]} \{t\} \times R_t$, and that F^+ is of class \mathcal{C}^1 on the same set. Then there exists a version \hat{X}_t of X_t such that, with probability 1:*

- (1) For all $t \in [0, T], h > 0$, $\hat{X}_{t+h} \leq \varphi_h^+(\hat{X}_t)$.
- (2) \hat{X}_t has only jump discontinuities, and its sample paths are right continuous; moreover, \hat{X}_t is a Feller process, so that it has the strong Markov property; furthermore, for all $t \in [0, T]$, $\hat{X}_{t^+} = \hat{X}_t \leq \hat{X}_{t^-}$, so that jumps may occur only downwards.
- (3) Let $0 < S < T$, and define $\tau = \inf\{t \in (S, T] : X_t < \varphi_{t-S}^+(X_S)\}$. Then

$$P(X_\tau < X_{\tau^-} \mid \tau < T) = 1 \quad \text{a.s.},$$

i.e. X_t follows the maximal solution between jumps.

- (4) The number of times X_t jumps in $[S, T]$, with $0 < S < T$, is finite with probability 1.

Proof. Properties (1) and (2) are proved together. Suppose X_t is a Markov process defined on some probability space (Ω, \mathcal{F}, P) with transition probability (4.4). It follows from (4.4) that there is $N \in \mathcal{F}$ such that $P(N) = 0$ and

$$X_{t+h} \leq \varphi_h^+(X_t) \quad \text{for all } \omega \notin N, \quad (5.4)$$

for all $t, h \in \mathbb{Q}$, $t \in [0, T]$, $h > 0$. Note that, for any $t \in [0, T]$,

$$\lim_{s \rightarrow t^+, s \in \mathbb{Q}} X_s$$

exists for all $\omega \notin N$. Indeed, fix $t \in [0, T]$ and assume, by contradiction, that $\lim_{s \rightarrow t^+, s \in \mathbb{Q}} X_s$ does not exist. Then there exist $\eta > 0$ and sequences of rationals $t_n \downarrow t$, $\tau_n \downarrow t$ such that $X_{t_n} > X_{\tau_n} + \eta$ for all n , contradicting (5.4). Thus define, for $t \in [0, T]$, $\omega \notin N$,

$$\hat{X}_t(\omega) = \lim_{s \rightarrow t^+, s \in \mathbb{Q}} X_s;$$

note that $\hat{X}_t(\omega)$ is automatically right continuous for $\omega \notin N$, and (5.4) holds for \hat{X}_t for any $t, h > 0$. We now show that \hat{X}_t and X_t have the same finite distributions; so, in particular, \hat{X}_t is Markov with transition probability (4.4). Let $\hat{F}_{t_1, \dots, t_n}(x_1, \dots, x_n) = P(\hat{X}_{t_1} \leq x_1, \dots, \hat{X}_{t_n} \leq x_n)$. Note that, by (4.7) and continuity of p_t , $F_{t_1, \dots, t_n}(x_1, \dots, x_n)$ is continuous in (t_1, \dots, t_n) , and so

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = \lim_{s_1 \downarrow t_1, \dots, s_n \downarrow t_n} F_{s_1, \dots, s_n}(x_1, \dots, x_n), \quad (5.5)$$

where s_1, \dots, s_n vary in \mathbb{Q} . Moreover, by definition of \hat{X}_t and dominated convergence

$$\begin{aligned} \hat{F}_{t_1, \dots, t_n}(x_1, \dots, x_n) &= \lim_{s_1 \downarrow t_1, \dots, s_n \downarrow t_n} P(X_{s_1} \leq x_1, \dots, X_{s_n} \leq x_n) \\ &= \lim_{s_1 \downarrow t_1, \dots, s_n \downarrow t_n} F_{s_1, \dots, s_n}(x_1, \dots, x_n). \end{aligned} \quad (5.6)$$

The conclusion follows from (5.5) and (5.6). We now show that, for $\omega \notin N$, the path $\hat{X}_t(\omega)$ has the left limit in any point. This is proved as for the right limit: nonexistence of the limit would violate (5.4). To complete the proof of (2) we still have to show the Feller and the strong Markov property. First we prove that $P\{\hat{X}_t > \varphi_t^-, \forall t > 0\} = 1$. Indeed, $\{\hat{X}_t > \varphi_t^-, \forall t > 0\} = \bigcap_{T>0} \bigcup_{\gamma>0} \{\hat{X}_t \geq \varphi_t^- + \gamma, \forall t \in (0, T]\} := \bigcap_{T>0} \bigcup_{\gamma>0} \Omega_{\gamma, T}$. It is easy to see, by (4.7) and the regularity of sample paths of \hat{X}_t , that $\sup_{\gamma>0} P(\Omega_{\gamma, T}) = 1$ for all T . The Feller property now follows from Lemma 4.1; this, together with right continuity, implies the strong Markov property (see pp. 56–57 of [3]).

For the proof of (3) and (4), refer to [4]. ■

Remark 5.3 (The Infinitesimal Generator). By assuming, essentially, that the densities are of class \mathcal{C}^1 , a formal computation of the infinitesimal generator can

be performed. We omit the details. We consider the derivative

$$(L_s G)(y) = \frac{d}{dt} \left[\int G(x) P(dx, t; y, s) \right]_{|t=s}$$

for $s \geq 0$, $y \in \mathbb{R}$ and assume that G is of class \mathcal{C}^1 . For $s > 0$ and $y \in (\varphi_s^-, \varphi_s^+)$ it holds that

$$(L_s G)(y) = G'(y) \dot{\varphi}_0^+(y) - \frac{\partial}{\partial t} \Big|_{t=s} A(t, s, y) \int_{\varphi_s^-}^y [G(x) - G(y)] \frac{p_s(x)}{\mu_s(\varphi_s^-, y)} dx.$$

The expression of the generator is compatible with the properties of X_t which have already been proved, and suggests further ones. However, observe that

$$-\frac{\partial}{\partial t} \Big|_{t=s} A(t, y, s)$$

diverges at φ_s^- and is nonsmooth, so that the classical construction of the process through its generator cannot be performed. A more detailed analysis is contained in [4].

5.3. Probability that X_t Remains Below a Solution of (4.1)

The following computation will be used in Section 6. In this subsection the assumptions of Theorem 5.1 are supposed to hold, and X_t is identified with its regular version \hat{X}_t . Let φ be a solution of (4.1) such that $\varphi_t > \varphi_t^-$ for all $t > 0$. We want to compute

$$P_\varphi^T := P\{X_t \leq \varphi_t, \forall t \in [0, T]\}.$$

By the regularity of the sample paths of X_t , it holds that

$$P_\varphi^T = \lim_{n \rightarrow \infty} P\{X_{t_k^n} \leq \varphi_{t_k^n}, \forall k = 1, \dots, n\},$$

where $t_k^n = (k/n)T$. Now, since $\varphi_{t_{i+1}^n} \leq \varphi_{T/n}^+(\varphi_{t_i^n})$,

$$\begin{aligned} & P\{X_{t_k^n} \leq \varphi_{t_k^n}, \forall k = 1, \dots, n\} \\ &= F_{t_1^n \dots t_n^n}(\varphi_{t_1^n}, \dots, \varphi_{t_n^n}) \\ &= \prod_{i=0}^{n-1} \frac{\mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{t_{i+1}^n}^+]}{\mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{T/n}^+(\varphi_{t_i^n}^+)]} \\ &= \prod_{i=0}^{n-1} \frac{\mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{t_{i+1}^n}^+]}{\mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{t_{i+1}^n}^+) + \mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{T/n}^+(\varphi_{t_i^n}^+)]} \\ &= \prod_{i=0}^{n-1} \left(1 - \frac{\mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{T/n}^+(\varphi_{t_i^n}^+)]}{\mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{t_{i+1}^n}^+) + \mu_{t_{i+1}^n}(\varphi_{t_{i+1}^n}^-, \varphi_{T/n}^+(\varphi_{t_i^n}^+)]} \right). \end{aligned}$$

Note that $\mu_{i+1}^n(A) = \mu_i^n(A) + O(T/n)$, for all measurable A ; therefore,

$$\begin{aligned}
& \log P\{X_k^n \leq \varphi_k^n, \forall k = 1, \dots, n\} \\
&= \sum_{i=0}^{n-1} \log \left(1 - \frac{\mu_{i+1}^n(\varphi_{i+1}^n, \varphi_{T/n}^+(\varphi_i^n))}{\mu_{i+1}^n(\varphi_{i+1}^-, \varphi_{i+1}^n] + \mu_{i+1}^n(\varphi_{i+1}^n, \varphi_{T/n}^+(\varphi_i^n))} \right) \\
&= - \sum_{i=0}^{n-1} \left\{ \frac{\mu_{i+1}^n(\varphi_{i+1}^n, \varphi_{T/n}^+(\varphi_i^n))}{\mu_{i+1}^n(\varphi_{i+1}^-, \varphi_{i+1}^n] + \mu_{i+1}^n(\varphi_{i+1}^n, \varphi_{T/n}^+(\varphi_i^n))} \right. \\
&\quad \left. + o \left(\frac{\mu_{i+1}^n(\varphi_{i+1}^n, \varphi_{T/n}^+(\varphi_i^n))}{\mu_{i+1}^n(\varphi_{i+1}^-, \varphi_{i+1}^n] + \mu_{i+1}^n(\varphi_{i+1}^n, \varphi_{T/n}^+(\varphi_i^n))} \right) \right\} \\
&= - \sum_{i=0}^{n-1} \left\{ \frac{p_i^n(\varphi_i^n)(F^+(\varphi_i^n) - \dot{\varphi}_i^n)(T/n) + o(T/n)}{\mu_i^n(\varphi_i^-, \varphi_i^n] + O(T/n)} + o\left(\frac{T}{n}\right) \right\} \\
&= - \sum_{i=0}^{n-1} \left\{ \frac{p_i^n(\varphi_i^n)(F^+(\varphi_i^n) - \dot{\varphi}_i^n)}{\mu_i^n(\varphi_i^-, \varphi_i^n]} \frac{T}{n} + o\left(\frac{T}{n}\right) \right\}.
\end{aligned}$$

By taking the limit for $n \rightarrow \infty$ we obtain, formally,

$$P_\varphi^T = \exp \left(- \int_0^T \frac{p_t(\varphi_t)(F^+(\varphi_t) - \dot{\varphi}_t)}{\mu_t(\varphi_t^-, \varphi_t]} dt \right);$$

the above integral diverges to $+\infty$ unless φ_t coincides with φ_t^+ in a right neighborhood of $t = 0$. Thus

$$P_\varphi^T = \begin{cases} 0 & \text{if } \varphi_t < \varphi_t^+, \quad \forall t > 0, \\ \exp \left(- \int_0^T \frac{p_t(\varphi_t)(F^+(\varphi_t) - \dot{\varphi}_t)}{\mu_t(\varphi_t^-, \varphi_t]} dt \right) & \text{if } \varphi_t = \varphi_t^+, \quad \forall t \in [0, \tau_0], \quad \tau_0 > 0. \end{cases} \quad (5.7)$$

6. The Solution Set as a Stochastic Process

In Section 4 the Markov process X_t was constructed in such a way that its density p_t equals a given flow, supported on the reachable set of a differential inclusion. The construction is based on ratios of measures of reachable sets, but it does not fully take into account the structure of the differential inclusion: in fact, the family of joint distributions (4.7) uses the values of the solutions $\varphi^+(t_1, \dots, t_n; x_1, \dots, x_n)$, defined in Lemma 2.2, only at certain points. The construction of the present section is still based on maximal solutions which are given constraints below, but uses “all of them”, instead of only their values at nodal points. This permits us to define a process which is fully compatible with a differential inclusion.

6.1. Construction of Process Y_t

We recall that in Section 5.3 the probability that the process X_t remains below a solution of (4.1) was computed. We use it to define a family of joint distributions.

Definition 6.1. Let $0 < t_1 < \dots < t_n \leq T$ and let x_1, \dots, x_n be points satisfying

$$x_i \in [\varphi_{t_i}^-, \varphi_{t_i}^+], \quad i = 1, \dots, n.$$

We define

$$G_{t_1 \dots t_n}(x_1, \dots, x_n) := P\{X_t \leq \varphi_t^+(t_1, \dots, t_n; x_1, \dots, x_n), \forall t \in [0, T]\}; \quad (6.1)$$

$G_{t_1 \dots t_n}(x_1, \dots, x_n)$ can be extended to the whole \mathbb{R}^n by setting

$$G_{t_1, \dots, t_n}(x_1, \dots, x_i, \dots, x_n) = G_{t_1, \dots, t_n}(x_1, \dots, \varphi_{t_i}^+, \dots, x_n)$$

for $x_i > \varphi_{t_i}^+$, and

$$G_{t_1, \dots, t_n}(x_1, \dots, x_i, \dots, x_n) = 0$$

for $x_i < \varphi_{t_i}^-$. Now set $\varphi_t = \varphi_t^+(t_1, \dots, t_n; x_1, \dots, x_n)$. Recalling (5.7) we have

$$G_{t_1 \dots t_n}(x_1, \dots, x_n) = \exp\left(-\int_{\tau_0}^T \frac{p_t(\varphi_t)(F^+(\varphi_t) - \dot{\varphi}_t)}{\mu_t(\varphi_t^-, \varphi_t]} dt\right), \quad (6.2)$$

where $\tau_0 = \tau_0(t_1, \dots, t_n; x_1, \dots, x_n) = \max\{t : \varphi_t = \varphi_t^+\} > 0$.

We now show that the family of joint distributions given in Definition 6.1 satisfies a set of consistency conditions. The proof actually uses only one property of X_t , so we state the theorem in the most general case. We will need the following notation: given a function $G(x_1, \dots, x_n)$ and an interval $I = [a, b]$, we define for $k = 1, \dots, n$,

$$\begin{aligned} \Delta_I^k G(x_1, \dots, x_n) \\ = G(x_1, \dots, x_{k-1}, b, x_{k+1}, \dots, x_n) - G(x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n). \end{aligned}$$

Given intervals $I_1 = [a_1, b_1], \dots, I_n = [a_n, b_n]$, by applying recursively the above operation, we have

$$\Delta_{I_1}^1 \cdots \Delta_{I_n}^n G(x_1, \dots, x_n) = \sum_{j=1}^{2^n} (-1)^{n(v_j)} G(v_j^1, \dots, v_j^n), \quad (6.3)$$

where $v_j = (v_j^1, \dots, v_j^n)$ —with $\{v_j^\ell\}_j$ running over the 2^n possible choices between a_ℓ and b_ℓ , $\ell = 1, \dots, n$ —and $n(v_j)$ equals the number of lower extrema a_ℓ among the coordinates of the vector v_j (see pp. 242–243 of [11]).

Theorem 6.1. Let X_t be a stochastic process such that, for all $t, h > 0$,

$$P\{X_{t+h} \leq \varphi_h^+(X_t)\} = 1. \quad (6.4)$$

Then the family of functions

$$G_{t_1 \dots t_n}(x_1 \cdots x_n), \quad 0 < t_1 < \dots < t_n \leq T, \quad x_i \in [\varphi_{t_i}^-, \varphi_{t_i}^+],$$

defined as in (6.1) satisfies the following conditions:

- (i) $G_{t_1 \dots t_n}(x_1, \dots, x_n)$ is right-continuous, i.e. if $x_i^{(j)} \downarrow x_i$, $x_i^{(j)} \in [\varphi_{t_i}^-, \varphi_{t_i}^+]$ for all $i = 1, \dots, n$, then

$$G_{t_1 \dots t_n}(x_1^{(j)}, \dots, x_n^{(j)}) \rightarrow G_{t_1 \dots t_n}(x_1, \dots, x_n);$$

- (ii) if $x_i^{(j)} \downarrow -\infty$ for some i , then $G_{t_1 \dots t_n}(x_1, \dots, x_i^{(j)}, \dots, x_n) \rightarrow 0$; if $x_i^{(j)} \uparrow +\infty$ for all $i = 1, \dots, n$, then $G_{t_1 \dots t_n}(x_1^{(j)}, \dots, x_n^{(j)}) \rightarrow 1$;
 (iii) if $x_i \geq \varphi_{t_i}^+(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n; x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ for some i , then

$$G_{t_1 \dots t_n}(x_1, \dots, x_n) = G_{t_1 \dots t_{i-1} t_{i+1} \dots t_n}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n);$$

- (iv) let $I_1 \dots I_n$ be intervals; then

$$\Delta_{I_1}^1 \dots \Delta_{I_n}^n G_{t_1 \dots t_n}(x_1, \dots, x_n) \geq 0.$$

Remark. If X_t satisfies the condition $P\{X_{t+h} \in (\varphi_t^-, \varphi_t^+(X_t))\} = 1$ (as is the case of the process constructed in Section 4), then the first part of (ii) can be strengthened to

- if $x_i^{(j)} \downarrow \varphi_{t_i}^-$ for some i , then $G_{t_1 \dots t_n}(x_1, \dots, x_i^{(j)}, \dots, x_n) \rightarrow 0$.

Proof. (i) The continuity of $G_{t_1 \dots t_n}(x_1, \dots, x_n)$ with respect to (x_1, \dots, x_n) is a straightforward consequence of the continuity of φ_t^+ in Lemma 2.2 and of the right-continuity of the measure induced by X_t .

- (ii) It is a simple consequence of (6.4) and of the definition of $G_{t_1 \dots t_n}(x_1, \dots, x_n)$.

- (iii) If $x_i \geq \varphi_{t_i}^+(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n; x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$, then, by (6.4),

$$\varphi_t^+(t_1, \dots, t_n; x_1, \dots, x_n) = \varphi_t^+(t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n; x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n).$$

(iv) Fix x_1, \dots, x_n and let $\varepsilon_1, \dots, \varepsilon_n > 0$. Set also $I_i = [x_i, x_i + \varepsilon_i]$. Thanks to (iii), it suffices to prove (iv) only in the case where, for all $k = 1, \dots, n$,

$$x_k \leq \varphi_{t_k - t_{k-1}}^+(x_{k-1}), \quad x_k + \varepsilon_k \leq \varphi_{t_k - t_{k-1}}^+(x_{k-1} + \varepsilon_{k-1}). \quad (6.5)$$

Let $I_i = [x_i, x_i + \varepsilon_i]$. The following statement is obviously sufficient to prove (iv). For the sake of simplicity of notation, in the lemma below we omit writing times t_1, \dots, t_n in the functions φ^+ .

Lemma 6.1. For $k = 1, \dots, n$ define

$$\begin{aligned} A_k &= \{(x, t) : \varphi_t^+(x_1 + \varepsilon_1, \dots, x_{k-1} + \varepsilon_{k-1}, x_k, x_{k+1}, \dots, x_n) < x \\ &\leq \varphi_t^+(x_1 + \varepsilon_1, \dots, x_{k-1} + \varepsilon_{k-1}, x_k + \varepsilon_k, x_{k+1}, \dots, x_n)\} \end{aligned}$$

(we mean $\varepsilon_0 = 0$) and

$$B_k = \{(x, t) : x \leq \varphi_t^+(x_1 + \varepsilon_1, \dots, x_k + \varepsilon_k, x_{k+1}, \dots, x_n)\},$$

where both A_k and B_k are thought of as subsets of $\mathbb{R} \times [0, t_n]$. Then

$$\begin{aligned} &\Delta_{I_1}^1 \dots \Delta_{I_n}^n G_{t_1 \dots t_n}(x_1, \dots, x_n) \\ &= P\{\text{graph } X \cap A_1 \cap \dots \cap A_k \neq \emptyset, \text{graph}(X) \subset B_k\}. \end{aligned}$$

Proof. By induction on k . For $k = 1$ the result is trivial, since by definition

$$\Delta_{I_1}^1 G(x_1, \dots, x_n) = P\{\text{graph}(X.) \cap A_1 \neq \emptyset, \text{graph}(X.) \subset B_1\}.$$

We now show the inductive step. First, it is easy to see, recalling (6.5), that

$$\begin{aligned} A_k &= \{(x, t) : t \in (t_{k-1}, t_{k+1}), \varphi_t^+(x_{k-1} + \varepsilon_{k-1}, x_k, x_{k+1}) < x \\ &\leq \varphi_t^+(x_{k-1} + \varepsilon_{k-1}, x_k + \varepsilon_k, x_{k+1})\}. \end{aligned}$$

Now fix $k < n$. For $j = 1, \dots, k$ let \tilde{A}_j, \tilde{B}_j be defined as A_j, B_j but with $x_{k+1} + \varepsilon_{k+1}$ in place of x_{k+1} . One easily has

$$\tilde{B}_k = B_{k+1}, \quad \tilde{A}_j = A_j \quad \text{for } j < k, \quad \tilde{A}_k \supset A_k.$$

Moreover, because of the fact that the paths of X_t cannot cross upwards a maximal solution implies that

$$\text{graph}(X.) \cap \tilde{A}_k \neq \emptyset, \text{graph}(X.) \subset \tilde{B}_k \quad \Rightarrow \quad \text{graph}(X.) \cap A_k \neq \emptyset.$$

Thus, by inductive assumption

$$\begin{aligned} &\Delta_{I_1}^1 \cdots \Delta_{I_k}^k G_{I_1 \dots I_n}(x_1, \dots, x_k, x_{k+1} + \varepsilon_{k+1}, x_{k+2}, \dots, x_n) \\ &= P\{\text{graph}(X.) \cap A_1 \cap \cdots \cap A_k \neq \emptyset, \text{graph}(X.) \subset B_{k+1}\}, \end{aligned}$$

from which

$$\begin{aligned} &\Delta_{I_1}^1 \cdots \Delta_{I_k}^k \Delta_{I_{k+1}}^{k+1} G(x_1, \dots, x_n) \\ &= P\{\text{graph}(X.) \cap A_1 \cap \cdots \cap A_k \neq \emptyset, \\ &\quad \text{graph}(X.) \subset B_{k+1}, \text{graph}(X.) \cap (B_{k+1} \setminus B_k) \neq \emptyset\}, \end{aligned}$$

and the conclusion follows, since $A_{k+1} = B_{k+1} \setminus B_k$. ■

The proof of Theorem 6.1 is concluded. ■

Kolmogorov's extension theorem (see p. 253 of [2]) yields the following:

Theorem 6.2. *There exists a stochastic process $(Y_t, t \in [0, T], \mathcal{Q})$ such that*

$$\mathcal{Q}\{Y_{t_1} \leq x_1, \dots, Y_{t_n} \leq x_n\} = G_{I_1 \dots I_n}(x_1, \dots, x_n).$$

Remark. If X_t is the process constructed in Section 4, stronger consistency conditions can be proved, which permit us to perform a construction of the probability measure \mathcal{Q} directly in the set of solutions of (4.1), rather than in $\mathbb{R}^{[0, T]}$ as in Kolmogorov's theorem. This would provide regularity of sample paths together with existence; however, we think it is simpler to study them separately. Regularity, together with other properties, is the topic of the next section.

6.2. Regularity of Sample Paths of Process Y_t

We first prove a result concerning a probability at times $t, t + h$. We recall that R_t denotes the reachable set of (4.1) at time t .

Proposition 6.1. *Let $h > 0$, and assume that, for all $x \in \bigcup_{t>0} R_t$, $F(x, U)$ is convex. Then*

$$Q\{Y_{t+h} \in R_h(Y_t)\} = 1. \tag{6.6}$$

Proof. It suffices to show that, for all x ,

$$Q(Y_{t+h} > \varphi_h^+(x) \mid Y_t = x) = 0, \tag{6.7}$$

$$Q(Y_{t+h} < \varphi_h^-(x) \mid Y_t = x) = 0. \tag{6.8}$$

Property (6.7) is obvious by (6.4). To prove (6.8), fix x and a closed interval $[a, b] \subset (\varphi_{t+h}^-, \varphi_{t+h}^-(x))$. If $\varepsilon > 0$ is sufficiently small,

$$Q(Y_t \in [x - \varepsilon, x + \varepsilon], Y_{t+h} \in [a, b]) = 0.$$

Indeed,

$$\begin{aligned} & Q(Y_t \in [x - \varepsilon, x + \varepsilon], Y_{t+h} \in [a, b]) \\ &= G_{t,t+h}(x + \varepsilon, b) - G_{t,t+h}(x - \varepsilon, b) - G_{t,t+h}(x + \varepsilon, a) + G_{t,t+h}(x - \varepsilon, a) \\ &= 0, \end{aligned}$$

since $G_{t,t+h}(x + \varepsilon, b) = G_{t,t+h}(x - \varepsilon, b)$, and $G_{t,t+h}(x + \varepsilon, a) = G_{t,t+h}(x - \varepsilon, a)$, being $\varphi^+(t, t + h; x + \varepsilon, b) = \varphi^+(t, t + h; x - \varepsilon, b)$ and $\varphi^+(t, t + h; x + \varepsilon, a) = \varphi^+(t, t + h; x - \varepsilon, a)$. ■

The regularity follows easily:

Theorem 6.3. *Let $T > 0$. Then there exists a version \hat{Y}_t of Y_t such that the function $t \mapsto \hat{Y}_t$, $t \in [0, T]$, is Q -a.s. a solution of $\dot{x} \in \text{co } F(x, U)$, $x(0) = x_0$.*

Proof. By Proposition 6.1 there exists a full measure set such that, for all Y_t in it and positive $t, h \in \mathbb{Q}$,

$$\text{dist}(Y_{t+h} - Y_t, h[F^-(Y_t), F^+(Y_t)]) = o(h).$$

By uniform continuity, set, for $t \in [0, T]$, $\hat{Y}_t = \lim_{s \rightarrow t, s \in \mathbb{Q}} Y_s$. Then \hat{Y}_t is Q -a.s. Lipschitz, and thus Q -a.s. differentiable for a.e. t ; moreover, for a.e. t ,

$$\lim_{h \rightarrow 0^+} \frac{\hat{Y}_{t+h} - \hat{Y}_t}{h} \in \text{co } F(\hat{Y}_t, U) \quad Q\text{-a.s.}$$

Since a symmetrical argument holds also for $h < 0$, the sample paths of the process are a.s. solutions of $\dot{x} \in \text{co } F(x, U)$, $x(0) = x_0$. ■

Assuming $F(x, U)$ is convex, Theorem 6.3 and the following proposition imply $Q(\mathcal{S}) = 1$.

Proposition 6.2. *The solution set \mathcal{S} of (2.1), (2.2) is measurable with respect to the σ -field generated by the cylinder sets in $\mathbb{R}^{[0,1]}$.*

Proof. Let $L = \sup\{F^+(x) : x \in \bigcup_{t \in [0, T]} R_t\}$. For $n \geq 1$, $i = 1, \dots, 2^n$, set $t_i^n = i2^{-n}T$, and consider the nested sequence of measurable sets

$$S_n = \{x \in \mathbb{R}^{[0, T]} : x_{t_{i+1}^n} \in x_{t_i^n} + [\varphi_{2^{-n}T}^-(x_{t_i^n}), \varphi_{2^{-n}T}^+(x_{t_i^n})], x \text{ is } L\text{-Lipschitz}\}.$$

It is easy to show that $\mathcal{S} = \bigcap_n S_n$, and therefore \mathcal{S} is measurable. \blacksquare

The following result shows that the support of Q is all of \mathcal{S} ; thus Q is a reasonable measure on \mathcal{S} .

Proposition 6.3. *Let \mathcal{S} be endowed with the uniform convergence topology. Then the support of the probability measure Q is the whole of \mathcal{S} .*

Proof. Observe that every open set in \mathcal{S} contains a set from the family

$$\{\varphi_t \in \mathcal{S} : \varphi_t^-(t_1, \dots, t_n; x_1, \dots, x_n) < \varphi_t < \varphi_t^+(t_1, \dots, t_n; y_1, \dots, y_n), \forall t \in [0, T]\},$$

with $0 < t_1 < \dots < t_n \leq T$, $\varphi_t^- < x_i < y_i < \varphi_t^+$, $n \in \mathbb{N}$. Each of the above sets clearly has positive probability. \blacksquare

6.3. The Law of \dot{Y}_t

We consider here the process Y_t associated with the X_t constructed in Section 4. Moreover, we require that both $p_t(\xi)$ and $F^+(x)$ be of class \mathcal{C}^1 ; the regularity of F^+ ensures that the map $(t_1, \dots, t_n; x_1, \dots, x_n) \mapsto \varphi^+(t_1, \dots, t_n; x_1, \dots, x_n)$ is of class \mathcal{C}^1 . Under the above assumption, we can compute the joint densities of the process Y_t . In particular, we write $\varphi_t = \varphi_t^+(t_1, \dots, t_n; x_1, \dots, x_n)$ and set

$$H_t(t_1, \dots, t_n; x_1, \dots, x_n) = \frac{p_t(\varphi_t)[F^+(\varphi_t) - \dot{\varphi}_t]}{\mu_t(\varphi_t^-, \varphi_t]},$$

i.e. H_t is the integrand of the exponent in $G_{t_1, \dots, t_n}(x_1, \dots, x_n)$. Then it holds, for x_1 in the interior of R_{t_1} , that

$$\frac{\partial G_{t_1}(x_1)}{\partial x_1} = -G_{t_1}(x_1) \int_0^T \frac{\partial H_t(t_1; x_1)}{\partial x_1} dt,$$

and

$$\begin{aligned} & \frac{\partial G_{t_1 t_2}(x_1, x_2)}{\partial x_1} \\ &= \begin{cases} 0 & \text{if } x_2 < \varphi_{t_2-t_1}^-(x_1), \\ -G_{t_1 t_2}(x_1, x_2) \int_0^T \frac{\partial H_t(t_1, t_2; x_1, x_2)}{\partial x_1} dt & \text{if } x_2 \in (\varphi_{t_2-t_1}^-(x_1), \varphi_{t_2-t_1}^+(x_1)), \\ -G_{t_1}(x_1) \int_0^T \frac{\partial H_t(t_1; x_1)}{\partial x_1} dt & \text{if } x_2 > \varphi_{t_2-t_1}^+(x_1). \end{cases} \end{aligned} \tag{6.9}$$

Theorem 6.4. *Let $T > 0$ be fixed. There exists $D \subset \Omega \times [0, T]$ such that $(Q \otimes \lambda)(D) = T$ and for all $(\omega, t) \in D$,*

$$\dot{Y}_t \in \{F^-(Y_t), F^+(Y_t)\}. \tag{6.10}$$

Moreover,

$$Q(\dot{Y}_t = F^-(Y_t) \mid Y_t = y) = \frac{-G_t(y)}{\partial G_t(y)/\partial y} \int_0^T \frac{\partial H_t(t, y)}{\partial y} dt, \tag{6.11}$$

so that the law of \dot{Y}_t is nontrivial.

Proof. The set D can be constructed with the same argument as in Theorem 3.1. Now fix t such that \dot{Y}_t exists a.s., and let $0 < \lambda < 1$. Set, for $y \in [\varphi_t^-, \varphi_t^+]$, $F_\lambda(y) = \lambda F^-(y) + (1 - \lambda)F^+(y)$, and observe that for all $h > 0$ sufficiently small we have that

$$y + hF_\lambda(y) \in R_h(y).$$

Therefore, by (6.9),

$$\begin{aligned} & Q(Y_{t+h} - Y_t \leq hF_\lambda(Y_t)) \\ &= \int_{\varphi_t^-}^{\varphi_t^+} Q(Y_{t+h} \leq y + hF_\lambda(y) \mid Y_t = y) \frac{\partial G_t}{\partial y}(y) dy \\ &= \int_{\varphi_t^-}^{\varphi_t^+} \left. \frac{\partial G_{t,t+h}(y, z)}{\partial y} \right|_{z=y+hF_\lambda(y)} dy \\ &= - \int_{\varphi_t^-}^{\varphi_t^+} G_{t,t+h}(y, y + hF_\lambda(y)) \int_0^T \left. \frac{\partial H_t(t, t+h; y, z)}{\partial y} \right|_{z=y+hF_\lambda(y)} dt dy. \end{aligned}$$

The last expression, for $h \rightarrow 0^+$, tends to

$$- \int_{\varphi_t^-}^{\varphi_t^+} G_t(y) \int_0^T \frac{\partial H_t}{\partial y}(t, y) dt dy,$$

which does not depend on λ . Therefore, the only possibility is (6.10), and the last expression equals $Q(\dot{Y}_t = F^-(Y_t))$. The very same argument as before yields (6.11). ■

7. Computations for the Processes Associated with $\dot{x} \in [-1, 1]$

We consider the differential inclusion

$$\begin{cases} \dot{x}(t) \in [-1, +1], \\ x(0) = 0, \end{cases} \tag{7.1}$$

with the uniform distribution, $p_t \equiv 1$.

The above setting may be considered as a model of “pure bounded noise” (i.e. a perturbation of the zero dynamics) with complete uncertainty on its statistical properties. Let X_t and Y_t be the processes constructed in Sections 4 and 6 for the case described above.

7.1. Extinction Probability for X_t

Consider the process Z_t defined for $t \in \mathbb{R}$ by

$$Z_t = \frac{X_{e^t}}{e^t}.$$

Clearly, $Z_t \in [-1, 1]$ a.s. Moreover, via a change of variables, we obtain $\tilde{P}(dz, t; w, s) = P(e^t dx, e^t; e^s w, e^s)$; hence by (4.4) one obtains

$$\begin{aligned} \tilde{P}(dz, t; w, s) &= \frac{e^{-(t-s)}(w+1)}{e^{-(t-s)}(w-1)+2} \delta_{e^{-(t-s)}(w-1)+1}(dz) \\ &\quad + \frac{2(1-e^{-(t-s)})}{e^{-(t-s)}(w-1)+2} \frac{1_{(-1, e^{-(t-s)}(w-1)+1]}(z)}{e^{-(t-s)}(w-1)+2} dz. \end{aligned} \quad (7.2)$$

Note that $\tilde{P}(dz, t; w, s)$ depends on t, s only through $t-s$. So Z_t is a time-homogeneous Markov process.

Now we use Z_t to show that

$$P\{X_t \geq 0, \forall t \in [0, a]\} = 0 \quad (7.3)$$

for all $a > 0$, so that the extinction probability $P_0^a = 1$ for all $a > 0$. It is easy to see that the joint distributions of X_t are positively homogeneous with degree 0, i.e.

$$F_{\rho t_1 \dots \rho t_n}(\rho x_1, \dots, \rho x_n) = F_{t_1 \dots t_n}(x_1, \dots, x_n)$$

for all $\rho > 0$, $0 \leq t_1 < \dots < t_n$, x_1, \dots, x_n . Thus the probability in (7.3) is independent of a , and we compute it for $a = 1$. Note that

$$X_t \geq 0, \quad \forall t \in [0, 1] \quad \Rightarrow \quad Z_n \geq 0, \quad \forall n \in \mathbb{Z}, \quad n \leq 0.$$

So we show

$$P\{Z_n \geq 0, \forall n \leq 0\} = 0.$$

By the Markov property and the time homogeneity of Z_t it is enough to show that

$$\sup_{w \in [-1, 1]} P(Z_1 \geq 0 | Z_0 = w) < 1. \quad (7.4)$$

Indeed, $P\{Z_n \geq 0, \forall n \leq 0\}$ is bounded above by an infinite product of factors, all equal to the left-hand side of (7.4). By direct computation, using (7.2) we get

$$P(Z_1 \geq 0 | Z_0 = w) = \begin{cases} 1 - \frac{2(1-e^{-1})}{[e^{-1}(w-1)+2]^2} & \text{if } e^{-1}(w-1)+1 \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

So $P(Z_1 \geq 0 | Z_0 = w) \leq (1+e^{-1})/2 < 1$, concluding the proof of (7.4).

Further properties can be observed. In particular, from (7.3) and $P\{X_t \leq 0, \forall t \in [S, T]\} = 0$, which follows from (5.7), we have that, for all $\varepsilon > 0$, X_t hits a.s. 0 for infinitely many times $t \in [0, \varepsilon]$.

Finally, observe that the positive homogeneity of $F_{t_1 \dots t_n}(x_1, \dots, x_n)$ together with (7.3) imply that, for all η ,

$$\lim_{T \rightarrow +\infty} P\{X_t \geq \eta, \forall t \in [0, T]\} = 0 \quad \text{a.s.},$$

so that a.s. the process leads to extinction.

7.2. Joint Distributions for Y_t

A direct computation provides, for $x_1 \in [-t_1, t_1]$:

$$G_{t_1}(x_1) = \exp\left(\frac{x_1 - t_1}{x_1 + t_1}\right), \tag{7.5}$$

$$G_{t_1 t_2}(x_1, x_2) = \begin{cases} \exp\left(\frac{x_2 - t_2}{x_2 + t_2}\right) & \text{for } x_2 \in [-t_2, x_1 - t_2 + t_1], \\ \exp\left(\frac{x_1 - t_1}{x_1 + t_1}\right) \exp\left(\frac{x_2 - x_1 - t_2 + t_1}{x_2 + t_2}\right) & \text{for } x_2 \in [x_1 - t_2 + t_1, x_1 + t_2 - t_1], \\ \exp\left(\frac{x_1 - t_1}{x_1 + t_1}\right) & \text{for } x_2 \in (x_1 + t_2 - t_1, t_2]. \end{cases} \tag{7.6}$$

In the above formulas, we mean $e^{-a/0} = 0$ for $a > 0$. In the general case, $G_{t_1 \dots t_n}(x_1, \dots, x_n)$ appears as a product of as many factors of the type $e^{(x_i - (x_{i-1} + t_i - t_{i-1})) / (x_i + t_i)}$ as is the number of active constraints. If all points are reachable from the preceding one, i.e. $x_{i+1} \in R_{t_{i+1} - t_i}(x_i)$ for all i , we obtain that

$$G_{t_1 \dots t_n}(x_1, \dots, x_n) = \exp\left(\sum_{i=1}^n \frac{x_i - x_{i-1} - t_i + t_{i-1}}{x_i + t_i}\right), \tag{7.7}$$

where we assume $x_0 = t_0 = 0$.

7.3. The Law of \dot{Y}_t

The same computation of Theorem 6.4 yields that

$$Q(\dot{Y}_t = -1) = \int_{-t}^t e^{(s-t)/(s+t)} \frac{t-s}{(t+s)^2} ds = \int_0^{+\infty} e^{-\xi} \frac{\xi}{1+\xi} d\xi;$$

observe that this value is independent of t . It also follows that

$$Q(\dot{Y}_t = -1 \mid Y_t = a) = \frac{t-a}{2t},$$

$$Q(\dot{Y}_t = 1 \mid Y_t = a) = \frac{t+a}{2t}.$$

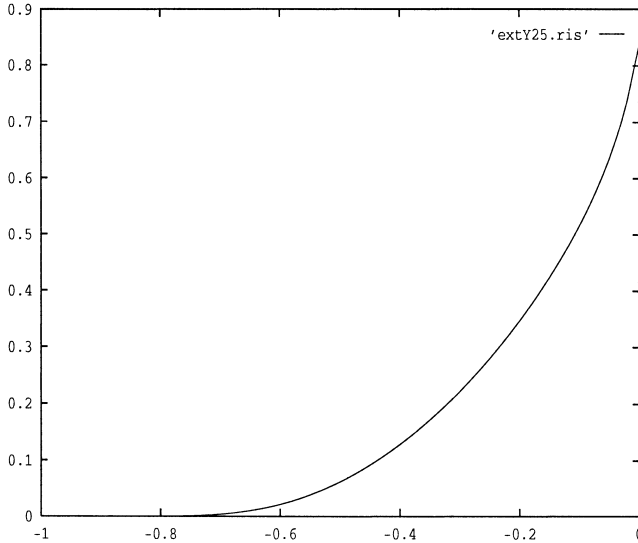


Fig. 1. $1 - Q_n(\eta)$ is plotted against η (parameters: $n = 25$, $T = 1$).

7.4. The Extinction Probability for Y_t

Let $T > 0$ be fixed and let $\eta \in (-T, T)$. We consider the probability

$$\begin{aligned} Q\{\exists t \in [0, T] : Y_t < \eta\} &= 1 - Q\{Y_t \geq \eta, \forall t \in [0, T]\} \\ &= 1 - \lim_{n \rightarrow \infty} Q\{Y_{iT/n} \geq \eta : \forall i = 1, \dots, n\}. \end{aligned}$$

Set $Q_n(\eta) = Q\{Y_{iT/n} \geq \eta : \forall i = 1, \dots, n\}$. Thanks to the regularity of trajectories and the knowledge of joint distributions, all Q_n are known. Figure 1 shows $1 - Q_n(\eta)$ plotted against η .

8. Conclusions

In this paper we give results on probabilistic modeling for bounded noise. The aim is to construct stochastic processes satisfying the following two requirements:

- (1) The support of the distribution of the process on its path space coincides with the set of solutions of a differential inclusion;
- (2) many features of the process can be analyzed by explicit computations, via its finite-dimensional distributions.

We have shown that Markov processes are not suitable models for these purposes. In fact, Markov processes with absolutely continuous trajectories are solutions, in the Carathéodory sense, of a deterministic differential equation: “randomness” may only arise from the nonuniqueness of the solution.

We have constructed non-Markovian processes that satisfy requirements (1) and

(2) for a scalar differential inclusion. Having fixed the differential inclusion, there are still several degrees of freedom in the choice of the related stochastic process, making this class of process quite flexible for identification and estimation.

The construction presented in this paper relies on the total order of the real line. Extension to higher dimensions or to partially ordered sets is currently under investigation.

Acknowledgement. We thank an anonymous referee for sharp and constructive criticisms.

References

- [1] Braumann CA (1983) Population growth in random environments, *Bull. Math. Biol.*, 45:635–641
- [2] Breiman L (1992) *Probability*, SIAM, Philadelphia, PA
- [3] Chung KL (1982) *Lectures from Markov Processes to Brownian Motion*, Springer-Verlag, Berlin
- [4] Colombo G, Dai Pra P (2001) A class of piecewise deterministic Markov processes, *Markov Process. Rel. Fields*, 7:251–287
- [5] Davis MH (1993) *Markov Models and Optimization*, Chapman & Hall, London
- [6] Di Masi GB, Gombani A, Kurzhanski AB (eds.) (1991) *Modeling, Estimation and Control of Systems with Uncertainty*, Birkhäuser, Boston, MA
- [7] Ethier SN, Kurtz Th (1986) *Markov Processes*, Wiley, New York
- [8] Gihman II, Skorohod AV (1975) *The Theory of Stochastic Processes, II*, Springer-Verlag, Berlin
- [9] Halley JM (1996) Ecology, evolution and $1/f$ -noise, *TREE*, 11:33–37
- [10] Kurzhanski AB, Vályi I (1996) *Ellipsoidal Calculus for Estimation and Control*, Birkhäuser, Boston, MA
- [11] McShane EJ (1944) *Integration*, Princeton University Press, Princeton, NJ
- [12] Ray D (1956) Stationary Markov processes with continuous paths, *Trans. Amer. Math. Soc.*, 82:52–493
- [13] Smith RS, Dahleh M (eds.) (1994) *The Modeling of Uncertainty in Control Systems: Proceedings of the 1992 Santa Barbara Workshop*, Lecture Notes in Control and Information Sciences, Volume 92, Springer-Verlag, Berlin
- [14] Soong TT (1973) *Random Differential Equations in Science and Engineering*, Academic Press, New York
- [15] Steele JH (1985) A comparison of terrestrial and marine ecological systems, *Nature*, 313:355–358
- [16] Turelli M (1977) Random environments and stochastic calculus, *Theoret. Pop. Biol.*, 12:140–178