Probabilistic Bisimulations of Switching and Resetting Diffusions

Alessandro Abate

Delft Center for Systems and Control, TU Delft, The Netherlands

a.abate@tudelft.nl

***Abstract*— This contribution presents sufficient conditions for the existence of probabilistic bisimulations between two diffusion processes that are additionally endowed with switching and resetting behaviors. A probabilistic bisimulation between two stochastic processes is defined by means of a bisimulation function, which induces an approximation metric over the distance between the two processes. The validity of the proposed sufficient conditions results in the explicit characterization of one such bisimulation function. The conditions depend on contractivity properties of the two stochastic processes.**

I. INTRODUCTION

Much research nowadays is targeted to the general problem of model simplification, particularly when the system under study is dimensionally large and complex. In many instances – both theoretical and practical – it is critical to obtain a simplified model that is, in a certain sense, equivalent to the original one. Alternatively, other approaches target qualitative simplifications and are mostly concerned with computationally scalable procedures [3]. The notion of equivalence between models can be expressed by the correspondence between the trajectories of the original (concrete) model with those of the simplified (abstract) one, and can be formalized with the concept of bisimulation [16]. Bisimulation relations are widely used in the computer sciences, particularly with regards to finite, discrete-space models [14], [5]. However, in the instance of dynamical models evolving on continuous spaces or according to probabilistic laws, the search for a bisimulation relation can lead to rather conservative results [13], [18]. In this case it is possible to rely on approximate versions of the notion of bisimulation [9], or probabilistic variants of it [5], [14] – both relaxations rely on metrics over the distance between trajectories or realizations of the models. They both work for models in continuous- and discrete-time. Whereas the notion in [9] works for continuous-space models, that in [5], [14] holds exclusively for probabilistic models over discrete spaces.

The notion of probabilistic bisimulation for models evolving on uncountable state spaces has been introduced only recently [12]. This work leverages a Lyapunov-based approach to derive sufficient algebraic conditions for the existence of probabilistic bisimulation functions between processes as general as jump linear stochastic sytems. As an alternative to [12], the work in [1] puts forward sufficient conditions for the existence and explicit characterization of a probabilistic bisimulation functions between two diffusion processes, which are based on their contractivity properties. The conditions in [1] are shown to be related to a probabilistic version of the concept of incremental stability [2]. Interestingly, both the notion of Lyapunov stability and the concept of incremental stability have been recently exploited in the study of (bi-)simulation relations for deterministic models on continuous spaces [10], [11].

This work extends the results in [1] in a number of directions. It is firstly concerned with finding conditions for the existence of probabilistic bisimulations (Section IV) of stochastic models with switches and resets, namely models that can discretely change the structure of their dynamics, as well as perform finite jumps over their continuous domains (Section II) – these models are thus probabilistic and hybrid in nature and are named “switching and resetting diffusions.” Secondly, the present work looks at probabilistic contractivity over non-identity metrics (Sections III and V-D). This allows to consider processes that present non-trivial limit sets. The study develops a number of numerical studies in Section V. Proofs are omitted due to space constraints.

II. MODEL

We consider a probabilistic and hybrid framework in the autonomous case, inspired by [7], [8]. It is characterized by continuous-time processes that evolve probabilistically over an uncountable state space, jump randomly within a finite set of modes, and are reset upon changing mode. We denote processes with bold typeset, and with normal fonts sample values, points, or functions over the state space.

The state of the system at time $t \in \mathbb{R}^+$ (non-negative reals) is given by a vector $(\mathbf{q}(t), \mathbf{x}(t)) \in \mathcal{S}$, where $\mathcal{S} \doteq \mathcal{Q} \times \mathbb{R}^n, n < \infty$, is the hybrid state space and $\mathcal{Q} = \{1, 2, \ldots, Q\}, Q \in \mathbb{N}, Q < \infty$. The discrete component $\mathbf{q}$ belongs to a finite set $\mathcal{Q}$ of modes, whereas the continuous component $\mathbf{x}$ to the Euclidean space $\mathbb{R}^n$. Within each mode, the continuous component is characterized by the following dynamics:

$$d\mathbf{x} = f(\mathbf{q}, \mathbf{x})dt + \sigma(\mathbf{q}, \mathbf{x})d\mathbf{W} + r(\mathbf{q}, \mathbf{x})d\mathbf{p}, \qquad (1)$$

where $f : \mathcal{S} \to \mathbb{R}^n$ is a vector field (which characterizes the deterministic drift), $\sigma : \mathcal{S} \to \mathbb{R}^{n \times m}$ is the diffusion matrix, $\mathbf{W}$ a standard m-dimensional Wiener process [17], $r : \mathcal{S} \to \mathbb{R}^n$ is a deterministic reset function, and $\mathbf{p}(\cdot, \cdot)$

Work supported by the European Commission under Marie Curie Grant MANTRAS 249295 and under STREP project MoVeS FP7-ICT-257005, and by NWO under VENI grant 016.103.020.

is a Poisson measure with intensity $dt \times \texttt{Leb}(dz)$, where $\texttt{Leb}$ is the Lebesgue measure over the real line. The Poisson process $\mathbf{p}$ associated with the random measure induces a set of random time instants $\tau_i, i \geq 0, \tau_0 = 0, \tau_{i+1} > \tau_i$ (called *event times*), at which the discrete component $\mathbf{q}$ of the process changes location. The initial condition for the model is sampled from a probability distribution with support over the whole hybrid state space, $\pi : \mathcal{S} \to [0,1]$. Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space over which $\mathbf{W}, \mathbf{p}, \pi$, and thus $(\mathbf{q}, \mathbf{x})$, are defined [8].

The Poisson process $\mathbf{p}$ is characterized by state-dependent rate functions $\lambda_{ij}(\cdot), i, j \in \mathcal{Q}$, describing the dynamics of the discrete component of the process, which commutes location according to the following probability:

$$P\left(\mathbf{q}(t+\delta t) = j | (\mathbf{q}(t), \mathbf{x}(t)) = (i, x)\right) = \lambda_{ij}(x)\delta t + o(dt), \tag{2}$$

where $i \neq j$, and $\forall (i,x) \in \mathcal{S}, \lambda_{ij} : \mathbb{R}^n \to \mathbb{R}^+$, and $\sum_{j \in \mathcal{Q}} \lambda_{ij}(x) = 0$. For any $(i,x) \in \mathcal{S}, j \in \mathcal{Q}$, define a function $l : \mathcal{Q} \times \mathcal{S} \to \mathbb{R}^+$ so that $l(0,(i,x)) = 0, l(j,(i,x)) = \sum_{k=1}^{j} \lambda_{ik}(x)$. Introduce a function $\tilde{h} : \mathbb{R} \times \mathcal{S} \to \mathcal{Q}$, so that

$$\tilde{h}(z,(i,x)) = \begin{cases} j - i, & \text{if } l(j-1,(i,x)) \leq z \leq l(j,(i,x)), \\ 0, & \text{else.} \end{cases}$$

Similarly, let us introduce a function $\tilde{r} : \mathbb{R} \times \mathcal{S} \to \mathbb{R}^n$ as

$$\tilde{r}(z,(i,x)) = \begin{cases} r(i,x), & \text{if } z \leq |l(i,(i,x))|, \\ 0, & \text{else,} \end{cases}$$

where 0 denotes an n-dimensional vector of zero elements.

The process described by (1)-(2) can then be equivalently expressed as

$$\begin{cases} d\mathbf{x} = f(\mathbf{q},\mathbf{x})dt + \sigma(\mathbf{q},\mathbf{x})d\mathbf{W} + \int_{\mathbb{R}} \tilde{r}(z,(\mathbf{q},\mathbf{x}))\mathbf{p}(dt,dz), \\ d\mathbf{q} = \int_{\mathbb{R}} \tilde{h}(z,(\mathbf{q},\mathbf{x}))\mathbf{p}(dt,dz). \end{cases} \tag{3}$$

A solution of (3) is a cadlag [4] stochastic process $(\mathbf{q}(t,s_0), \mathbf{x}(t,s_0))$ taking values in $\mathcal{S}$, with initial condition $(\mathbf{q}(0), \mathbf{x}(0)) = s_0$ sampled from π, which evolves within the event times $[\tau_i, \tau_{i+1}), i \geq 0$, according to the SDE

$$d\mathbf{x} = f(\mathbf{q},\mathbf{x})dt + \sigma(\mathbf{q},\mathbf{x})d\mathbf{W},$$

and which at times $\tau_i, i \geq 1$, undergoes a change of discrete location and a reset of the continuous component as follows:

$$\mathbf{x}(\tau_i) = \lim_{t \uparrow \tau_i} \mathbf{x}(t) + r\left(\lim_{t \uparrow \tau_i} (\mathbf{q}(t), \mathbf{x}(t))\right),$$
$$\mathbf{q}(\tau_i) = \tilde{h}\left(\lim_{t \uparrow \tau_i} \mathbf{p}(t), \lim_{t \uparrow \tau_i} (\mathbf{q}(t), \mathbf{x}(t))\right),$$

where $\mathbf{p}$ is the Poisson process associated with the random measure.

Let us introduce the following set of structural assumptions on the model in (3).

Assumption 1: For any pair $x, y \in \mathbb{R}^n$, $q, q' \in \mathcal{Q}$, there exists finite and positive constants K_q^1, K_q^2, K_q^3, such that:

1) *Lipschitz continuity:* $\|f(q,x) - f(q,y)\| + \|\sigma(q,x) - \sigma(q,y)\| \leq K_q^1 \|x - y\|$
2) *Bound on growth:* $\|f(q,x)\|^2 + \|\sigma(q,x)\|^2 \leq K_q^2(1 + \|x\|^2)$
3) *Bounded intensities:* $|\lambda_{qq'}(x)| < K_q^3$. Furthermore,
4) *Finite reset magnitude:* the function $r(q,x)$ takes values in a bounded domain of $\mathbb{R}^n$. □

It can be shown [8] that by upholding Assumption 1 on the components of the model (3), the solution process $(\mathbf{q}(t,s_0), \mathbf{x}(t,s_0))$ exists and is unique, for any $t \geq t_0 = 0$ and any finite initial condition $s_0 = (q_0, x_0) \in \mathcal{S}$.

It is important to have a precise characterization of the infinitesimal generator of (3) [4]. Consider a function $\psi : \mathcal{S} \to \mathbb{R}^+$, assumed to be bounded and twice continuously differentiable over the continuous domain $\mathbb{R}^n$. The infinitesimal generator of (3) is an operator $\mathcal{L}$, acting on $\psi(q,x)$, that is defined for any $q \in \mathcal{Q}$ as follows:

$$\begin{aligned} \mathcal{L}\psi(q,x) = & \frac{\partial \psi}{\partial x}(q,x) f(q,x) \\ & + \frac{1}{2}\mathrm{Tr}\left(\sigma(q,x)\sigma(q,x)^T \frac{\partial^2 \psi}{\partial x^2}(q,x)\right) \\ & + \sum_{q' \neq q} \lambda_{qq'}(x)\left(\psi(q', x + r(q,x)) - \psi(q,x)\right). \end{aligned} \tag{4}$$

The Dynkin equation [4] allows computing expectations of functions of the process, and states that, for any $s \in \mathcal{S}$, $\mathbb{E}_s \psi(\mathbf{q}(t,s), \mathbf{x}(t,s)) = \psi(s) + \mathbb{E}_s \int_0^t \mathcal{L}\psi(\mathbf{q}(l,s), \mathbf{x}(l,s))dl$.

A. Special Case: Resetting Diffusion (RD)

As a special case of (3), consider a model with probabilistic continuous dynamics and no discrete mode commutations – it resets onto its unique domain according to an arrival process $\mathbf{p}$. More precisely, consider $\mathcal{Q} = \{q\}$, the state-dependent rate function $\lambda(x) \doteq \lambda_{qq}(x)$, and the functions $f(x) \doteq f(q,x), \sigma(x) \doteq \sigma(q,x), r(x) \doteq r(q,x)$. The state space is simply $\mathbb{R}^n$. A resetting diffusion is thus described by the following dynamical relation:

$$d\mathbf{x} = f(\mathbf{x})dt + \sigma(\mathbf{x})d\mathbf{W} + r(\mathbf{x})d\mathbf{p}. \tag{5}$$

Consider a bounded, twice continuously differentiable function $\psi : \mathbb{R}^n \to \mathbb{R}^+$. The infinitesimal generator $\mathcal{L}$ is defined over points $x \in \mathbb{R}^n$ as follows:

$$\begin{aligned} \mathcal{L}\psi(x) = & \frac{\partial \psi}{\partial x}(x) f(x) + \frac{1}{2}\mathrm{Tr}\left(\sigma(x)\sigma(x)^T \frac{\partial^2 \psi}{\partial x^2}(x)\right) \\ & + \lambda(x)\left(\psi(x + r(x)) - \psi(x)\right). \end{aligned}$$

B. Special Case: Switching Diffusions (SD)

As a second special case of (3), consider the framework known as switching diffusions [7], characterized by probabilistic continuous dynamics and discrete mode changes with no associated continuous resets:

$$\begin{cases} d\mathbf{x} = f(\mathbf{q},\mathbf{x})dt + \sigma(\mathbf{q},\mathbf{x})d\mathbf{W}, \\ d\mathbf{q} = \int_{\mathbb{R}} \tilde{h}(z,(\mathbf{q},\mathbf{x}))\mathbf{p}(dt,dz). \end{cases} \tag{6}$$

Select a function $\psi : \mathcal{S} \to \mathbb{R}^+$, assumed to be bounded and twice continuously differentiable within each of the continuous domains $\mathbb{R}^n$. The infinitesimal generator $\mathcal{L}$ at point $s = (q, x)$ is defined as follows:

$$\mathcal{L}\psi(s) = \frac{\partial \psi}{\partial x}(s) f(s) + \frac{1}{2}\mathrm{Tr}\left(\sigma(s)\sigma(s)^T \frac{\partial^2 \psi}{\partial x^2}(s)\right) + \sum_{q' \neq q} \lambda_{qq'}(x)\left(\psi(q', x) - \psi(q, x)\right).$$

III. Stochastic Contractivity

The following definition extends [1], [19], and is inspired by studies of contractivity analysis for deterministic models in [15]. For any $(q, x) \in \mathcal{S}$, introduce the function $s(q, x) = x + r(q, x)$. Define $\Lambda_q = \sup_{x \in \mathbb{R}^n} |\lambda_{qq}(x)|$, which is finite as per Assumption 1.3. Denote with $\mathfrak{m}_A : A \to \mathbb{R}$ a map that acts on a finite set $A = \{a_1, \ldots, a_n\}$ and yields the maximum among the elements of the set A, $\mathfrak{m}_A(a_i) = \max_{i=1,\ldots,n; a_i \in A} |a_i|$.

***Definition** 1 ((General) Stochastic Contractivity):* Consider the process in (3). Assume that the following conditions are valid for all $q \in \mathcal{Q}$:

1) $f(q, \cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists F_q < \infty$: $\lambda_{\max}\left(\frac{\partial f}{\partial x}(\tilde{x})\right) \leq F_q$, where $\partial f / \partial x(\tilde{x})$ is the symmetric part of the Jacobian of $f(q, \cdot)$ evaluated at $\tilde{x}$, and $\lambda_{\max}(\cdot)$ is a function computing the maximum value among the real parts of the eigenvalues of a matrix
2) $\sigma(q, \cdot)$ is Lipschitz continuous, as per Assumption 1.1, with finite and positive constant $K_q^1 : (K_q^1)^2 \doteq S_q$
3) $s(q, \cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists R_q < \infty$: $s(q, \tilde{x})^T s(q, \tilde{x}) \leq R_q \, \tilde{x}^T \tilde{x}$

The system in (3) is said to be stochastically contractive (in the identity metric) if the following conditions hold:

$$\mathfrak{m}_{\mathcal{Q}}(F_q) \leq 0, \qquad \mathfrak{m}_{\mathcal{Q}}(\Lambda_q(R_q - 1)) \leq 0,$$
$$2\mathfrak{m}_{\mathcal{Q}}(F_q) + \mathfrak{m}_{\mathcal{Q}}(S_q) + \mathfrak{m}_{\mathcal{Q}}(\Lambda_q(R_q - 1)) < 0. \qquad \square$$

Let us focus on special cases of Definition 1. The first deals with resetting diffusions.

***Definition** 2 (Stochastic Contractivity for RD):* Consider the process in (5). Assume that the following conditions are valid:

1) $f(\cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists F < \infty$: $\lambda_{\max}\left(\frac{\partial f}{\partial x}(\tilde{x})\right) \leq F$
2) $\sigma(\cdot)$ is Lipschitz continuous with finite and positive constant $K^1 : (K^1)^2 \doteq S$
3) $s(\cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists R < \infty$: $s(\tilde{x})^T s(\tilde{x}) \leq R \, \tilde{x}^T \tilde{x}$

The system in (5) is said to be stochastically contractive (in the identity metric) if the following holds:

$$2F + S + \Lambda(R - 1) < 0. \qquad \square$$

The second special case deals with switching diffusions.

***Definition** 3 (Stochastic Contractivity for SD):* Consider the process in (6). Assume that the following conditions are valid for all $q \in \mathcal{Q}$:

1) $f(q, \cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists F_q < \infty$: $\lambda_{\max}\left(\frac{\partial f}{\partial x}(\tilde{x})\right) \leq F_q$
2) $\sigma(q, \cdot)$ is Lipschitz continuous, as per Assumption 1.1, with finite and positive constant $K_q^1 : (K_q^1)^2 \doteq S_q$

The system in (6) is said to be stochastically contractive (in the identity metric) if the following holds:

$$2\mathfrak{m}_{\mathcal{Q}}(F_q) + \mathfrak{m}_{\mathcal{Q}}(S_q) < 0. \qquad \square$$

***Remark** 1:* Definition 1 poses restrictions on the dynamics. Within each domain it is critical that the accrued contractivity effect of drift and reset offsets the disruptive presence of the diffusion term. Furthermore, to account for jumps between different domains, a contractivity condition has to hold for any possible pair of domains, which leads to the use of the maximization function. An numerical example in Section V-A illustrates contractive dynamics. $\square$

IV. Probabilistic Bisimulations

In this Section we recall the definition of probabilistic bisimulation of two stochastic processes by associating its existence to that of a function that relates the (expected value of the squared) distance between them. We claim that a sufficient condition for the existence of a probabilistic bisimulation function is the stochastic contractivity of a new system obtained by composing the two processes.

Consider two processes $S_i, i = 1, 2$, with solutions $(\mathbf{q}_i(t), \mathbf{x}_i(t)) \in \mathcal{S}_i$ of (3) and equipped with observations:

$$\mathbf{y}_i(t) = g_i(\mathbf{x}_i(t)), \quad i = 1, 2. \tag{7}$$

Let us assume that both observation functions take values in $\mathbb{R}^o, o \in \mathbb{N}, o < \infty$. The use of an observation map allows to compute the distance between two processes $S_i, i = 1, 2$ resorting to the standard Euclidean norm, rather than a (more complicated) distance defined over their hybrid state spaces [4]. Notice that the processes $S_i, i = 1, 2$ do not necessarily have the same state dimension (i.e., $Q_1 \neq Q_2, n_1 \neq n_2$), and are not driven by the same noise $\mathbf{W}$ nor by the same jump process $\mathbf{p}$.

***Assumption** 2:* Assume that each $g_i, i = 1, 2$, vanishes at the origin and is Lipschitz continuous with finite positive constant K_i^4. (Define $\nu = \max\{K_1^4, K_2^4\}$.) $\square$

Consider now the process made up by composing in parallel S_1 and S_2 and by subtracting their observations. The parallel composition is performed by taking the cross product of the hybrid state spaces for the two processes ($\bar{\mathcal{S}} = \mathcal{S}_1 \times \mathcal{S}_2$). The output is defined as follows:

$$\bar{\mathbf{y}}(t) = g_1(\mathbf{x}_1(t)) - g_2(\mathbf{x}_2(t)) \doteq g(\bar{\mathbf{x}}(t)), \tag{8}$$

where we have denoted with $\bar{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ (henceforth barred variables denote the composed model). Let us recall the following classical notion:

Definition 4 ((Super-)Martingale, [6]): A function $\chi : \mathbb{R}^n \to \mathbb{R}$ is called a *martingale* for a stochastic process $\mathbf{s}(t, s_0), t \geq 0$, taking values in $\mathbb{R}^n$, if for any $s_0 \in \mathbb{R}^n, t \geq 0, \mathbb{E}_{s_0}[\chi(\mathbf{s}(t, x_0))] = \chi(s_0)$. The function χ is called a *supermartingale* if $\mathbb{E}_{s_0}[\chi(\mathbf{s}(t, x_0))] \leq \chi(s_0)$. □

The following definition relates the behavior of the two processes S_1 and S_2 by upper-bounding the distance between their observations with a non-increasing function of time:

Definition 5 (Probabilistic Bisimulation Function, [12]): A continuous function $\psi : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}_0^+$ is called a probabilistic bisimulation function for the processes S_1 and S_2, solution of (3)-(7), if, considering the composed process in (8), the following holds:

1) $\forall (s_1, s_2) \in \bar{\mathcal{S}}, s_i = (q_i, x_i),\ \psi(\bar{x}) \geq \|g(\bar{x})\|^2$;
2) $\forall \bar{s}_0 = (\bar{q}_0, \bar{x}_0) \in \bar{\mathcal{S}}, \psi(\bar{\mathbf{x}}(t, \bar{x}_0))$ is a supermartingale started at $\bar{x}_0$.

If two processes S_1, S_2 (started at $\bar{s}_0$) admit a probabilistic bisimulation function, then they are said to be probabilistically bisimilar with precision $\psi(\bar{x}_0)$. □

If such a function exists, then we are allowed to state that $P_{\bar{x}_o}\left(\sup_{0 \leq s \leq t} \|g(\bar{\mathbf{x}}(s, \bar{x}_o))\|^2 \geq \delta\right) \leq \frac{\psi(\bar{x}_o)}{\delta}$: this defines a (probabilistic) bound on the distance between outputs of the two processes.

***Theorem** 1:* Consider two processes, solutions of (3) and with output (7) under Assumption 2. If the composition of S_1, S_2 (as in (8)) is stochastically contractive (as in Definition 1), then S_1 and S_2 are probabilistically bisimilar. When existing, a probabilistic bisimulation function for the two processes started at $\bar{x}$ has the form $\psi(\bar{x}) = 2\nu\|\bar{x}\|^2$. □

Introduce the function $\gamma : \mathbb{R}^o \times \mathbb{R}^o \times \mathbb{R}^+ \to \mathbb{R}^+$ as $\gamma((y_1, y_2), t) = \|g_1(x_1) - g_2(x_2)\|^2 e^{\Gamma t}$, where

$$\Gamma = 2\mathfrak{m}_{\mathcal{Q}}(F_q) + \mathfrak{m}_{\mathcal{Q}}(S_q) + \mathfrak{m}_{\mathcal{Q}}(\Lambda_q(R_q - 1))$$

is the global contractivity coefficient. It characterizes a time-dependent upper bound on the (expected value of the squared) distance between S_1, S_2 started at $(q_1, x_1), (q_2, x_2), q_1, q_2 \in \mathcal{Q}$. Similarly for the special cases:

***Corollary** 1:* If the two processes are solutions of (5) (resp. (6)) with output (7) under Assumption 2, the contractivity condifion in Definition 2 (resp. 3) on their composition (as in (8)) ensures their probabilistic bisimilarity. □

V. Case Studies and Extensions

A. Dynamical Properties of a single Contractive Switching and Resetting Diffusion

Let us consider a system made up of $\mathcal{Q} = \{1, 2\}, \mathcal{S} = \mathcal{Q} \times \mathbb{R}^2$. The drift is linear, $f(i, x) = A_i x, x = [x_1, x_2]^T$, and characterized by:

$$A_1 = \begin{bmatrix} -0.6 & 0.3 \\ -0.6 & 0.15 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -0.35 & 0 \\ 0.1 & -0.25 \end{bmatrix}.$$

The continuous dynamics are driven by a 1-dimensional Wiener process, scaled by matrices $\sigma(1, x) = 0.2[x_1\, x_2]^T, \sigma(2, x) = 0.3[x_1\, x_2]^T$. The Poisson measures are characterized by rates $\Lambda_1 = 0.41, \Lambda_2 = 0.38$, which are independent of the continuous component. The reset maps induce a rotation and an expansion of the continuous state, as follows:

$$r(i, x) = \alpha_i \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix} x,$$

where $\alpha_1 = 1.1, \alpha_2 = 4/3$, and $\theta_1 = -1, \theta_2 = -0.1$. The global coefficient of contractivity equals $-0.0778 < 0$. Let us denote the fully observed process started at $s_0 \in \mathcal{S}$ with $(\mathbf{q}(t, s_0), \mathbf{x}(t, s_0))$. Figure 1 displays the dynamical properties of the process, along with its distance from the origin (which is the limit set for the dynamics).

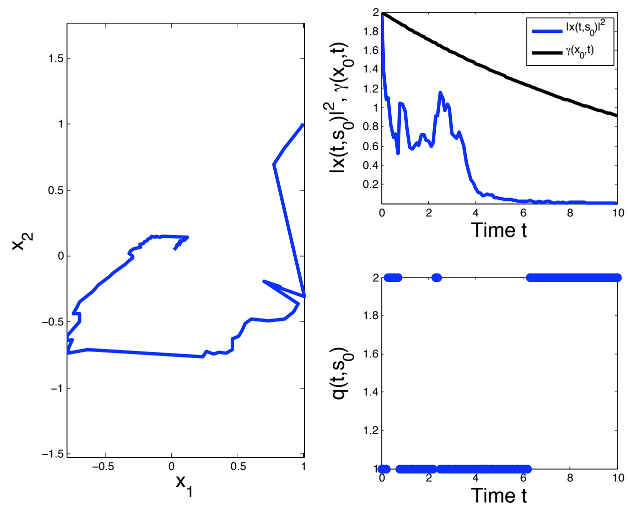

Fig. 1. A single realization of the continuous component $\mathbf{x}(t, s_0)$ of the process (in blue color, left plot). The plots on the right describe its (squared) norm, upper bounded by the function $\gamma(x_0, t) = \gamma((x_0, 0), t) = e^{-0.0778\,t}\|x_0\|^2$, and the evolution of its discrete component $\mathbf{q}(t, s_0)$. The time horizon is $T = 10$ sec, and is divided in $N = 10^2$ time intervals. The initial condition is $s_0 = (1, [1, 1]^T)$.

B. Probabilistic Bisimilarity of two Processes

We now compare the evolution of the process described in Section V-A with that of the following fully observed deterministic ODE, which takes values $x_r \in \mathbb{R}^2$: $\dot{\mathbf{x}}_r(t) = A_1 \mathbf{x}_r$.

We assume that process $(\mathbf{q}(t, s_0), \mathbf{x}(t, s_0))$ is initialized according to the uniform distribution $s_0 \sim \mathcal{U}(\mathcal{A})$, where $\mathcal{A} = \mathcal{Q} \times [-1, 1]^2$. Similarly, process $\mathbf{x}_r(t, (x_r)_0)$ is initialized according to $(x_r)_0 \sim \mathcal{U}([-1, 1]^2)$. Clearly, the contractivity coefficient of their composition again equals -0.0778. The function $\mathbb{E}_{(s_0, (x_r)_0)}[\|\mathbf{x}(t, s_0) - \mathbf{x}_r(t, (x_r)_0)\|^2], t \in [0, 20]$ is computed over 10^3 experiments. The function $\gamma(s_0, (x_r)_0, t) = e^{-0.0778\,t}\|x_0 - (x_r)_0\|^2$ and the bisimulation function $\psi(s_0, (x_r)_0) = \|x_0 - (x_r)_0\|^2$ are both averaged over the different initial conditions. Figure 2 plots the outcomes.

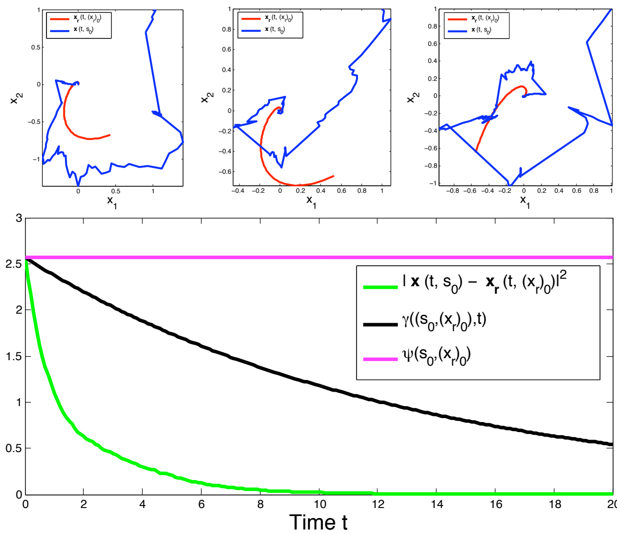

Fig. 2. Top row: three realizations of $\mathbf{x}(t, s_0)$ and of $\mathbf{x}_r(t,(x_r)_0)$. To ease the interpretation, we have kept $s_0 = (1,[1\,1]^T)$ fixed. Bottom row: average over 10^3 experiments for $\mathbb{E}_{(s_0,(x_r)_0)}[\|\mathbf{x}(t,s_0) - \mathbf{x}_r(t,(x_r)_0)\|^2], \gamma((s_0,(x_r)_0),t)$, and $\psi(s_0,(x_r)_0)$. The time horizon is $T = 20$ sec, and is divided in $N = 200$ time intervals.

C. Probabilistic Bisimilarity of Models with Outputs

Let us compare the following two linear, single-mode SDE with observation maps [12]. The second model is a "reduced" version of the first, extended model.

$$d\mathbf{x} = A\mathbf{x}\,dt + \Sigma\mathbf{x}\,d\mathbf{W}, \qquad \mathbf{y} = G\mathbf{x},$$
$$d\mathbf{x}_r = A_r\mathbf{x}_r\,dt + \Sigma_r\mathbf{x}_r\,d\mathbf{W}, \quad \mathbf{y}_r = G_r\mathbf{x}_r,$$

where (in "Matlab" notations)

$$A = \mathtt{blkdiag}(a_1,a_2,a_3), \quad A_r = \mathtt{blkdiag}(a_1,a_2);$$
$$a_1 = \begin{bmatrix} -1 & -10 \\ 10 & -1 \end{bmatrix}, a_2 = \begin{bmatrix} -2 & -20 \\ 20 & -1 \end{bmatrix}, a_3 = \begin{bmatrix} -2 & 0 \\ 0 & -2.5 \end{bmatrix};$$
$$\Sigma = 0.15 \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \Sigma_r = \Sigma(1:4;1:4);$$
$$G = \begin{bmatrix} 0.84 & -1.03 & 1.07 & -0.88 & 0.5 & 0 \\ -0.60 & -1.35 & -0.26 & -0.27 & 0 & -0.5 \end{bmatrix},$$

and $G_r = G(1:2;1:4)$. The processes are initialized deterministically in $x_0 = [111111]^T, (x_r)_0 = [1111]^T$. Based on the extended mode, we obtain a contractivity coefficient equal to $2F + S \approx -1.6 < 0$. Figure 3 displays the output of a single realization of the two processes, along with the average over 10^2 experiments of their distance and with function γ (here $\nu \approx 2.08$).

D. Non-Identity Metrics

An advantage that comes along with the use of (stochastic) contractivity analysis in the search for probabilistic bisimulation functions is the extension of the former to non-identity metrics [15], which entails the possibility of considering

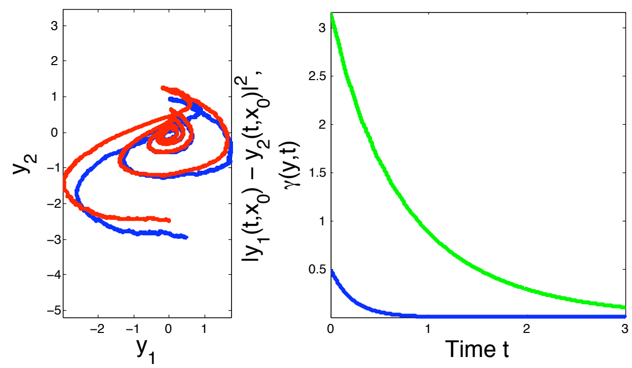

Fig. 3. Left plot: single realizations of $\mathbf{y}(t,x_0)$ (red color, full model) and of $\mathbf{y}_r(t,(x_r)_0)$ (blue color, reduced-order model), where the initial conditions have been set to $x_0 = [111111]^T, (x_r)_0 = [1111]^T$ respectively. Right plot: computation, over 10^2 experiments, of $\mathbb{E}_{(x_0,(x_r)_0)}[\|\mathbf{y}(t,x_0) - \mathbf{y}_r(t,(x_r)_0)\|^2]$ (blue plot), and plot of $\gamma((x_0,(x_r)_0),t)$ (green plot). The time horizon is $T = 3$ sec, and is divided in $N = 10^3$ time intervals.

bisimulation functions for models with non-trivial limit sets, such as closed orbits and limit cycles. For the sake of explanation, we shall focus on the RD model in (5) (the extension to models as in (3) follows intuitively).

***Definition** 6 (Stochastic Contractivity, Non-Identity Metrics):* Consider the process in (5), and a square matrix $\Theta(x), x \in \mathbb{R}^n$, such that $M(x) = \Theta^T(x)\Theta(x) \succ 0$, uniformly in x. Assume that the following holds:

1) $f(\cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists F < \infty$: $\lambda_{\max}\left(\frac{d\Theta}{dt}(x) + \Theta(x)\frac{\partial f}{\partial x}(\tilde{x})\right)\Theta^{-1}(x) \leq F$
2) $\sigma(\cdot)$ is Lipschitz continuous in the metric $M(\cdot)$ with finite and positive constant $K^1 : (K^1)^2 \doteq S$
3) $s(\cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists R < \infty$: $\lambda_{\max}\left(\Theta^T(\tilde{x})s^T(\tilde{x})M^T(\tilde{x})s(\tilde{x})\Theta(\tilde{x})\right) \leq R$

The system in (5) is said to be stochastically contractive in the metric $M(\cdot)$ if

$$2F + S + \Lambda(R-1) < 0. \qquad \square$$

In general, selecting a proper metric $M(\cdot)$ to show the contractivity property of a process can be non trivial. Ideally, a coordinate transformation $\Theta(\cdot)$ allows to define a (non unique) metric and to enforce the conditions in Definition 6 directly, as the following example shows.

a) Example of Contractivity over a Non-Identity Metric: Consider the following two-dimensional model in polar coordinates $(\rho,\theta) \in \mathbb{R}^+ \times [0, 2\pi]$:

$$\dot{\rho} = \rho(l_1 - \rho^2), \quad \dot{\theta} = l_2, \tag{9}$$

where $l_1 > 0, l_2 \in \mathbb{R}$. The model has an unstable equilibrium set at $(0,\theta)$, and a limit cycle at $(1,\theta)$. The parameter l_2 denotes the direction and the angular speed of rotation. Introduce the coordinate change $z = V_{l_1}(\rho) = \log(\rho/\sqrt{l_1})$, which translates into the dynamics $\dot{z} = l_1(1 - e^{2z})$ that are stable at the origin $z = 0$ and have a contractivity index $F = -2l_1 < 0$. It can be shown that a metric for the original model can be induced by selecting $\Theta(\rho) = \partial V_{l_1}/\partial\rho = 1/\rho$, which yields $M(\rho) = 1/\rho^2$. Let us remark that the chosen metric is not unique. $\square$

The example above leads to a general trick for selecting a possible metric: to look for a function that characterizes the limit region as its level set. Clearly this presupposes the explicit knowledge of the limit set. (Notice that the technique does not always lead to a legitimate coordinate transformation, as in Example a) – in general the matrix Θ has to be square and has to lead to a uniformly positive definite matrix M.) This can also lead to a generalization of the result in Theorem 1 for finding a bisimulation function for two processes, under proper contractivity assumptions. Consider two systems S_1, S_2, as in (5), and a compact set $\mathcal{C} \subset \mathbb{R}^n$ described as a level set of a differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ as $\mathcal{C} = \{x \in \mathbb{R}^n : V(x) = 0\}$. Select output processes $\mathbf{y}_i(t) = g_i(\mathbf{x}_i(t)) = V(\mathbf{x}_i(t))$. The following holds:

***Theorem** 2:* Consider two processes S_1, S_2 as in (5), with initial conditions respectively $x_1, x_2 \in \mathbb{R}^n$. If both systems are stochastically contractive over the set $\mathcal{C}$ (i.e., there exists a legitimate metric M as in Definition 6) then a probabilistic bisimulation function for them is $\psi(x_1, x_1) = |V(x_1) - V(x_2)|^2$, whenever V is Lipschitz. $\square$

b) Example of Probabilistic Bisimulation for Stochastically Contractive Processes over a Limit Set: Let us consider a perturbed version of (9):

$$d\rho = \rho(l_1 - \rho^2)dt + \sigma(\rho)dW, \quad \dot{\theta} = l_2, \tag{10}$$

where we assume $\sigma(\rho) = \alpha|V_{l_1}(\rho)|, \alpha > 0$. By expressing $x_1 = \rho\cos\theta, x_2 = \rho\sin\theta$, we obtain in Cartesian coordinates the following model:

$$d\mathbf{x}_1 = ((l_1 - \mathbf{x}_1^2 - \mathbf{x}_2^2)\mathbf{x}_1 - l_2\mathbf{x}_2)dt + \alpha_a|V_{l_1}(\rho)|dW,$$
$$d\mathbf{x}_2 = ((l_1 - \mathbf{x}_1^2 - \mathbf{x}_2^2)\mathbf{x}_2 + l_2\mathbf{x}_1)dt + \alpha_a|V_{l_1}(\rho)|dW.$$

The model has an attractive limit cycle at $\mathcal{C}_{l_1} = \{(x_1, x_2) : x_1^2 + x_2^2 = l_1\}$. We denote its solution from $(x_a)_0$ with $\mathbf{x}_a(t, (x_a)_0) \in \mathbb{R}^2, t \geq 0$.

We compare the above model with a second system that is endowed with a limit cycle, namely the van der Pol oscillator [20]. This can be expressed as:

$$\ddot{\mathbf{x}}_v + \beta(t)(\mathbf{x}_v^2 - 1)\dot{\mathbf{x}}_v + \mathbf{x}_v = 0. \tag{11}$$

For this second order model, if $\beta(\cdot) > 1$ the system displays relaxation oscillations, if $\beta(\cdot) \ll 1$ the system presents a circular limit cycle at $\mathcal{C} = \{((x_v)_1, (x_v)_2) : (x_v)_1^2 + (x_v)_2^2 = 4\}$, whereas as $\beta(\cdot) \approx 0$ the system behaves like a simple, undamped harmonic oscillator. Let us further perturb the model in (11) with a Wiener process with diffusion term $\alpha_v|V_4(x_v)|$. Over a simulation horizon $[0, T]$, we select a β that is a monotonically decreasing function of time, so that $\beta(0) = 1$ and $\beta(T) = 0$. Let us denote the solution of (11) from $(x_v)_0$ with $\mathbf{x}_v(t, (x_v)_0) \in \mathbb{R}^2, t \geq 0$.

Select $l_1 = 4$ and $l_2 = 1$, so that the two limit sets coincide, and $\alpha_a = \alpha_v = 1$. Let us focus on the observed process $\bar{\mathbf{y}}(t) = |g_a(\mathbf{x}_a(t)) - g_v(\mathbf{x}_v(t))|^2 = |V_4(\mathbf{x}_a(t)) - V_4(\mathbf{x}_v(t))|^2$. Figure 4 displays the outcomes. $\square$

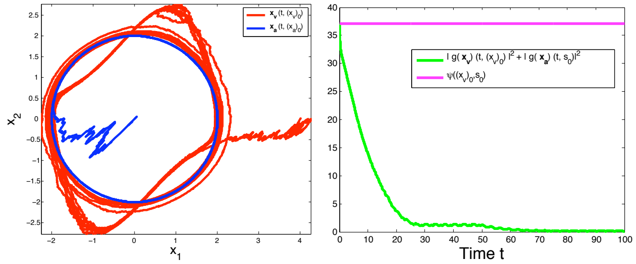

Fig. 4. Left plot: single realizations of $\mathbf{x}_a(t, (x_a)_0)$ (noisy attractor) and of $\mathbf{x}_v(t, (x_v)_0)$ (perturbed van der Pol oscillator), where the initial conditions have been set to $(x_a)_0 = (0, 0)^T, (x_v)_0 = (4, 0)^T$. To ease the interpretation, we have kept $s_0 = (1, [1\,1]^T)$ fixed. Right plot: average, over 10^2 experiments, for $\mathbb{E}_{((x_a)_0,(x_v)_0)}[|V_4(x_v(t, (x_v)_0)) - V_4(x_a(t, (x_a)_0))|^2]$, and $\psi((x_a)_0, (x_v)_0)$. The time horizon is $T = 10^2$ sec, and is divided in $N = 10^4$ time intervals.

References

[1] A. Abate. A contractivity approach for probabilistic bisimulations of diffusion processes. In *Proceedings of the 48th IEEE Conference of Decision and Control*, pages 2230–2235, Shanghai, China, December 2009.

[2] David Angeli. A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, 47:410–421, 2002.

[3] A.C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM, 2005.

[4] M. H. A. Davis. *Markov Models and Optimization*. Chapman & Hall/CRC Press, London, 1993.

[5] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labeled Markov processes. *Information and Computation*, 179(2):163–193, 2002.

[6] Richard Durrett. *Probability: Theory and Examples - Third Edition*. Duxbury Press, 2004.

[7] M.K. Ghosh, A. Arapostathis, and S. I. Marcus. Ergodic control of switching diffusions. *SIAM Journal of Control and Optimization*, 35(6):1952–1988, Nov. 1997.

[8] M.K. Ghosh and A. Bagchi. Modeling stochastic hybrid systems. In *System Modeling and Optimization*, IFIP International Federation for Information Processing 166, pages 269–280. Springer Verlag, 2005.

[9] A. Girard and G.J. Pappas. Approximation metrics for discrete and continuous systems. *IEEE Trans. on Automatic Control*, 52(5):782–798, 2007.

[10] Antoine Girard. Approximately bisimilar finite abstractions of stable linear systems. In A. Bemporad, A. Bicchi, and G. Buttazzo, editors, *Hybrid Systems: Computation and Control*, Lecture Notes in Computer Science 4416, pages 231–244. Springer Verlag, 2007.

[11] Antoine Girard, Giordano Pola, and Paulo Tabuada. Approximately bisimilar symbolic models for incrementally stable switched systems. *IEEE Transactions on Automatic Control*, 55(1):116–126, Jan 2010.

[12] A.A. Julius and G.J. Pappas. Approximate abstraction of stochastic hybrid systems. *IEEE Trans. Automatic Control*, 54(6):1193–1203, 2009.

[13] G. Lafferriere, G. J. Pappas, , and S. Sastry. O-minimal hybrid systems. *Mathematics of Control, Signals, and Systems*, 13(1):1–21, March 2000.

[14] K.G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94:1–28, 1991.

[15] W. Lohmiller and J.-J. Slotine. On contraction analysis for nonlinear systems. *Automatica*, 34:671–682, 1998.

[16] R. Milner. *Communication and concurrency*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.

[17] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, 6th edition, 2003.

[18] G.J. Pappas. Bisimilar linear systems. *Automatica*, 39(12):2035–2047, December 2003.

[19] Q.-C. Pham, N. Tabareau, and J.-J. Slotine. A contraction theory approach to stochastic incremental stability. *IEEE Transactions on Automatic Control*, 54(4):816–820, 2009.

[20] Shankar Sastry. *Nonlinear Systems: Analysis, Stability and Control*. Springer Verlag, 1999.