A Contractivity Approach
for Probabilistic Bisimulations of Diffusion Processes

Alessandro Abate

Abstract— This work is concerned with the problem of characterizing and computing probabilistic bisimulations of diffusion processes. A probabilistic bisimulation relation between two such processes is defined through a bisimulation function, which induces an approximation metric on the expectation of the (squared norm of the) distance between the two processes. We introduce sufficient conditions for the existence of a bisimulation function, based on the use of contractivity analysis for probabilistic systems. Furthermore, we show that the notion of stochastic contractivity is related to a probabilistic version of the concept of incremental stability. This relationship leads to a procedure that constructs a discrete approximation of a diffusion process. The procedure is based on the discretization of space and time. Given a diffusion process, we raise sufficient conditions for the existence of such an approximation, and show that it is probabilistically bisimilar to the original process, up to a certain approximation precision.

I. INTRODUCTION

In order to cope with the increasing complexity and scaling of real-world engineering systems and with the intractability of their corresponding mathematical models, a number of studies have explored the development of techniques aimed at relating or simplifying a model into a tractable one that is in some sense equivalent to the first. Equivalence is usually expressed with the notion of language correspondence, or with that of bisimulation. Less stringent versions of the notion of bisimulation (namely, the concept of simulation) are also used to express the idea of "inclusion" between the model dynamics. Furthermore, since frequently the *exact* notion of bisimulation translates into rather conservative requirements on the models under study (this is in particular true for "dynamically rich" models, such as hybrid and probabilistic ones), the concept of *approximate* bisimulation has been introduced as a relaxed version of that of strict bisimulation. This approximate concept requires the use of proper metrics (or pseudo-metrics) on the system trajectories.

While the notion of bisimulation is quite well known in computer science (in particular for the analysis of discrete models such as automata or transition systems), the study of exact or approximate bisimulation relations for deterministic, continuous dynamical systems has a more recent history. A few early results [1], [2] aimed at introducing exact notions, and have been more recently generalized to deal with approximate versions [7]. This has allowed the investigation of continuous- or discrete-time, nonlinear, controlled dynamical [8], [17] and even switched models [9].

Department of Aeronautics and Astronautics, Stanford University, CA, USA and Delft Center for Systems and Control, TU Delft, The Netherlands. Email: a.abate@tudelft.nl

With regards to probabilistic models, notions of similarity or bisimilarity have been in use for discrete-space models within the formal verification and the model checking community [4], [12]. It is of interest to extend the applicability of these notions to continuous models – as argued above, perhaps approximate notions would better fit the purpose. Recently, the contribution in [10] has introduced a notion of probabilistic bisimulation to set up an approximate correspondence between pairs of continuous stochastic processes. This result leverages the use of proper metrics on the system realizations, as well as Lyapunov techniques. From a different perspective but for similar discrete-time systems, [5] defines an approximate relationship between two processes by computing the distance between their distributions.

In this work, we set base from the concept of probabilistic bisimulation introduced in [10]. We also draw inspiration from the contribution in [8] and from that in [17], where for deterministic systems the existence of approximate bisimulations is shown and put in relationship with the concept of stability (classical [19], and incremental [3]). We extend the notion of incremental stability to a concept that holds on probabilistic models, and formally relate this concept to the notion of stochastic contractivity [15]. (This notion is the probabilistic extension of similar studies for deterministic models [13].) Under proper assumptions on the model, we propose a constructive procedure to formally approximate a diffusion process. The procedure performs a discretization of space and time, and is formally underpinned by the theory of weak approximations [11]. We show that this procedure induces a probabilistic bimulation of the original diffusion, and compute the relationship between the bisimulation precision and an approximation parameter.

The contribution develops as follows. Section II deals with the concept of stochastic contractivity. In Section III we show that the notion of stochastic contractivity is related to a probabilistic version of the concept of incremental stability. By leveraging both concepts, in Section IV we raise conditions on two diffusion processes to be probabilistically bisimilar, and show how to find a bisimulation function. Furthermore, this relationship leads to a procedure that attains a discrete approximation of a diffusion process, as discussed in Section V. We raise sufficient conditions for the existence of such an approximation and show that it is probabilistically bisimilar to the original system, up to a certain precision. Section VI recapitulates the results and lists future extensions.

II. CONTRACTION THEORY FOR DIFFUSION PROCESSES

Let us consider a diffusion process, taking values in the n-dimensional Euclidean space $\mathbb{R}^n$ and being characterized by the following stochastic differential equation (SDE):

$$d\mathbf{x} = f(\mathbf{x})dt + \sigma(\mathbf{x})d\mathbf{W}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a vector field (which characterizes the deterministic drift), $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ is the diffusion matrix, and $\mathbf{W}(t)$ a standard m-dimensional Wiener process [14].

Throughout this work, we assume that the following structural properties are in order:

Assumption 1: For any pair $x, y \in \mathbb{R}^n$, there exists finite and positive constants K_1, K_2, such that:
1) Lipschitz continuity: $\|f(x)-f(y)\|+\|\sigma(x)-\sigma(y)\| \leq K_1\|x-y\|$;
2) Growth bound: $\|f(x)\|^2+\|\sigma(x)\|^2 \leq K_2(1+\|x\|^2)$. $\square$

In Assumption 1, we have implicitly relied on the Euclidean norm for vectors and matrices. It can be shown [14] that by upholding Assumption 1 on the components of the SDE (1), its solution process exists and is unique, for any $t \geq t_0$ and any finite initial condition $x_0 \in \mathbb{R}^n$ at initial time $t_0 = 0$. We shall denote such a solution process $\mathbf{x}(t, x_0)$.[1]

The following definition is inspired by [16], which extends earlier studies for deterministic models [13].

Definition 1 (Stochastic Contractivity): Consider the SDE (1). Assume that the following conditions are valid:
1) $f(\cdot)$ is such that, for all $\tilde{x} \in \mathbb{R}^n, \exists \Lambda < \infty :$ $\lambda_{\max}\left(\frac{\partial f}{\partial x}(\tilde{x})\right) \leq \Lambda$, where $\partial f/\partial x(\tilde{x})$ is the symmetric part of the Jacobian of f 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) $\sigma(\cdot)$ is Lipschitz continuous, as per Assumption 1.1, with finite and positive constant $K_1 : (K_1)^2 \doteq K$.

Then the system in (1) is said to be stochastically contractive (in the identity metric) if $2\Lambda + K < 0$. $\square$

Example 1 (Special Case: deterministic ODE): Let us simplify the SDE (1) by considering $\sigma(x) \equiv 0, \forall x \in \mathbb{R}^n$. This reduces (1) to a (deterministic) ordinary differential equation (ODE), for which we raise Assumption 1.1. Furthermore, let us assume that the first assertion in Definition 1 is valid. Then, it can be shown [13] that if $\Lambda < 0$ then any solution path of the ODE, which is finitely bounded away from a second trajectory of the system, converges exponentially to such a trajectory. In particular, this convergence holds if the chosen trajectory is taken to coincide with an equilibrium point of the vector field $f(\cdot)$. This behavior should be evident in the simpler case of a linear ODE, where $f(x) = Ax, x \in \mathbb{R}^n$. $\square$

Remark 1: Definition 1 can be extended to non-identity and state-dependent metrics on the Euclidean space. Such

metrics scale non-uniformly the distance between points on the space under study and can be particularly useful for the study of nonlinear systems. In the following, we shall consider only the identity metric.

It is also possible to introduce a definition of contractivity over a strict subset $\mathcal{C}$ of the Euclidean space: in particular, it can be of interest to determine, based on the structure and properties of $f(\cdot)$ and $\sigma(\cdot)$, what is the *domain of contractivity* $\mathcal{C} \subseteq \mathbb{R}^n$, that is what is the set of points for which the conditions in Definition 1 are verified. $\square$

III. PROBABILISTIC INCREMENTAL STABILITY

Contractivity is a property that can be used to study certain stability properties of a system. In this Section we plan to employ the definition of stochastic contractivity, introduced in the previous Section, to show a probabilistic version of the notion of *incremental stability*. This probabilistic extension is inspired by the work in [3], where the concept of incremental stability is introduced and its properties analyzed for deterministic models. To the best of the author's knowledge, the concept of probabilistic incremental stability has been first mentioned in [16]. In Example 2, we export the probabilistic definition back to the deterministic models discussed in [3].

Let us first introduce some nomenclature: let $\mathbb{R}_0^+$ denote the non-negative reals. Consider a continuous function $\alpha : \mathbb{R}_0^+ \to \mathbb{R}_0^+$; α is said to belong to class $\mathcal{K}$ if it is strictly increasing and if $\alpha(0) = 0$. A continuous function $\beta : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is said to be in class $\mathcal{K}_\infty$ if it belongs to class $\mathcal{K}$ and moreover if $\beta(x) \to \infty$ as $x \to \infty$. A continuous function $\gamma : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is said to belong to class $\mathcal{KL}$ if, for any fixed $x \in \mathbb{R}_0^+$, the function $\gamma(\cdot, x)$ is in class $\mathcal{K}_\infty$ and if, for any fixed $y \in \mathbb{R}_0^+$, the function $\gamma(y, \cdot)$ is decreasing and such that $\gamma(y, t) \to 0$, as $t \to \infty$.

Definition 2 (Probabilistic Incremental Stability): The SDE (1) is said to be probabilistically incrementally stable (in p^{th} mean, $p \in \mathbb{N}$) if there exists a $\mathcal{KL}$ function γ such that for any $t \in \mathbb{R}_0^+$, any $x_1, x_2 \in \mathbb{R}^n$, the following condition is satisfied:

$$\mathbb{E}_{x_1,x_2}\left[\|\mathbf{x}(t, x_1) - \mathbf{x}(t, x_2)\|^p\right] \leq \gamma(\|x_1 - x_2\|^p, t). \qquad \square$$

Equivalently, we will also state that a solution process of the SDE in (1) is probabilistically incrementally stable (in p^{th} mean). Notice that the subscript in the expectation sign above, as intuitive, highlights a conditional operation:

$$\mathbb{E}_{x_1,x_2}\left[\|\mathbf{x}(t, x_1) - \mathbf{x}(t, x_2)\|^p\right] =$$
$$\mathbb{E}\left[\|\mathbf{x}(t, x_1) - \mathbf{x}(t, x_2)\|^p | \mathbf{x}(0, x_1) = x_1, \mathbf{x}(0, x_2) = x_2\right].$$

Remark 2: Definition 2 characterizes a notion which is related to the presence of convergent dynamics. As customary in the study of convergence of stochastic processes, various characterizations of this notion can be given, depending on the particular kind of convergence (for instance, in *probability*, in p^{th} *mean*, or *almost surely*). Similarly, Definition 2 can be modified accordingly, for instance: the SDE (1) is said to be probabilistically incrementally stable in *probability* if

---

[1]Throughout the paper, stochastic processes will be denoted with bold font, while points on the state space or sample values with regular typeset.

there exists a function $\gamma : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow [0,1]$ such that, for any fixed $x \in \mathbb{R}_0^+$, $\gamma(\cdot, x)$ is in class $\mathcal{K}$ and if, for any fixed $y \in \mathbb{R}_0^+$, the function $\gamma(y, \cdot)$ is decreasing and such that $\gamma(y, t) \rightarrow 0$, as $t \rightarrow \infty$, and such that for any $\epsilon > 0$, any $t \in \mathbb{R}_0^+$, any $x_1, x_2 \in \mathbb{R}^n$, the following holds: $P(\|\mathbf{x}(t, x_1) - \mathbf{x}(t, x_2)\| > \epsilon) \leq \gamma(\|x_1 - x_2\|, t)$.

In this work the selection of a notion based on the mean value comes from its adaptability to the known deterministic case (see the following Example 2), as well as to the concepts that will be discussed later in this this work. □

***Example** 2 (Deterministic incremental stability, [3]):* Consider the deterministic ODE in (1), obtained by setting $\sigma(x) \equiv 0, \forall x \in \mathbb{R}^n$. Let us uphold Assumption 1.1. System (1) is said to be *(globally) incrementally asymptotically stable* if there exists a $\mathcal{KL}$ function γ such that for any $t \in \mathbb{R}_0^+, x_1, x_2 \in \mathbb{R}^n$, the following is satisfied:

$$\|\mathbf{x}(t, x_1) - \mathbf{x}(t, x_2)\| \leq \gamma(\|x_1 - x_2\|, t). \qquad \square$$

As anticipated, the following result shows that stochastic contractivity of a diffusion process is sufficient for its probabilistic incremental stability (in mean square) to hold.

***Theorem** 1:* If system (1) is stochastically contractive, then its solution process is probabilistically incrementally stable (in mean square). □

Let us conclude the Section by remarking that it is possible to extend the notions developed here to the case of processes initialized according to a general probabilistic law.

IV. PROBABILISTIC BISIMULATIONS OF DIFFUSION PROCESSES

In this Section we recall the definition of probabilistic bisimulation function for certain classes of stochastic processes. We show that, given two diffusions, there exists a probabilistic bisimulation function that applies to them if a new system, composed by the two original diffusions, is probabilistically incrementally stable. This allows one to argue that a sufficient condition on the composed system for the existence of a probabilistic bisimulation function (relating its two components) is its stochastic contractivity.

Let us consider two diffusions S_1, S_2, along with their observations, which are described by the following systems:

$$\begin{array}{rcl} d\mathbf{x}_i & = & f_i(\mathbf{x}_i)dt + \sigma_i(\mathbf{x}_i)d\mathbf{W}_i; \\ \mathbf{y}_i(t) & = & g_i(\mathbf{x}_i(t)), \qquad i = 1, 2. \end{array} \quad (2)$$

We assume that the systems have heterogeneous state dimensions, but equal output size and noise dimension: namely, $x_i \in \mathbb{R}^{n_i}, \sigma_i(\cdot) \in \mathbb{R}^{n_i \times d_i}, y_i \in \mathbb{R}^{m_i}$, where it is *not* necessary that n_1 is equal to n_2, but where $d_1 = d_2 = d, m_1 = m_2 = m$. Moreover, we assume that the observation functions g_i vanish at the origin, $g_i(0) = 0$, and that they are Lipschitz continuous. This allows to state that, for any $x_i \in \mathbb{R}^{n_i}, \exists 0 \leq \nu < \infty : \|g_i(x_i)\| \leq \nu\|x_i\|$.

Consider now the following process on the Euclidean space $\mathbb{R}^{n_1 + n_2}$, which is composed by S_1 and S_2 (in parallel) and observed on $\mathbb{R}^m$ as follows:

$$\bar{x} = \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right], \quad \bar{W} = \left[\begin{array}{c} W_1 \\ W_2 \end{array} \right], \quad \bar{y} = [I - I] \left[\begin{array}{c} y_1 \\ y_2 \end{array} \right], \quad (3)$$

where we have denoted with I the m-dimensional unity matrix. Its dynamics are generated by:

$$d\bar{\mathbf{x}} = f(\bar{x})dt + \sigma(\bar{x})d\bar{\mathbf{W}}, \qquad \bar{\mathbf{y}}(t) = g(\bar{\mathbf{x}}(t)), \quad (4)$$

where

$$f(\bar{x}) \doteq \left[\begin{array}{c} f_1(x_1) \\ f_2(x_2) \end{array} \right], \sigma(\bar{x}) \doteq \left(\begin{array}{cc} \sigma_1(x_1) & 0 \\ 0 & \sigma_2(x_2) \end{array} \right),$$

$$g(\bar{x}) \doteq [I - I] \left[\begin{array}{c} g_1(x_1) \\ g_2(x_2) \end{array} \right] = [g_1(x_1) - g_2(x_2)].$$

Let us recall the following classical notion:

***Definition** 3 ((Super-) Martingale, [6]):* A function $\chi : \mathbb{R}^n \rightarrow \mathbb{R}$ is called a *martingale* for the process $\mathbf{x}(t, x_0), t \geq 0$, solution of (1), if for any $x(0) = x_0 \in \mathbb{R}^n, t \geq 0, \mathbb{E}_{x_0}[\chi(\mathbf{x}(t, x_0))] = \chi(x_0)$. The function χ is called a *supermartingale* if $\mathbb{E}_{x_0}[\chi(\mathbf{x}(t, x_0))] \leq \chi(x_0)$. □

Let us now introduce the following definition, which relates the behavior of systems S_1 and S_2 by bounding their distance with a non-increasing function of time:

***Definition** 4 (Probabilistic Bisimulation Function, [10]):* A continuous function $\psi : \mathbb{R}^{n_1 + n_2} \rightarrow \mathbb{R}_0^+$ is called a probabilistic bisimulation function for the diffusion processes S_1 and S_2 in (2) if, considering the composed process in (3)-(4), the following holds:

1) $\forall \bar{x} \in \mathbb{R}^{n_1 + n_2}, \psi(\bar{x}) \geq \|g(\bar{x})\|^2$;
2) $\forall \bar{x}_0 \in \mathbb{R}^{n_1 + n_2}, \psi(\bar{\mathbf{x}}(t, \bar{x}_0))$ is a supermartingale started at $\bar{x}_0$.

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

The next contribution of this work formally relates the property of stochastic contractivity of each of the two diffusions with the shared condition of probabilistic bisimilarity between them. The contractivity is intended to hold for the same metric (e.g., the identity one) for both processes.

***Theorem** 2:* Consider two diffusion processes, solutions of systems S_1, S_2 as in (2). If the composition of S_1, S_2, as defined in (3)-(4), is stochastically contractive, then S_1, S_2 are probabilistically bisimilar.
When existing, a probabilistic bisimulation function has the form $\psi(\bar{x}) = 2\nu\|\bar{x}\|^2$. □

Theorem 2 can be exploited to compute an upper bound on the (mean square) distance between pairs of realizations of S_1 and of S_2. More precisely, selecting a parameter $\delta > 0$, any two points $x_i^0 \in \mathbb{R}^{n_i}$, and by resorting to the properties

of the bisimulation function (as described in Definition 4) and to the Markov inequality [6], the following holds:

$$P_{x_1^0, x_2^0} \left(\sup_{0 \leq t < \infty} \| \bar{\mathbf{y}} \left(t, (x_1^0, x_2^0)^T \right) \|^2 \geq \delta \right)$$

$$= P_{x_1^0, x_2^0} \left(\sup_{0 \leq t < \infty} \| \mathbf{y}(t, x_1^0) - \mathbf{y}(t, x_1^0) \|^2 \geq \delta \right)$$

$$\leq P_{x_1^0, x_2^0} \left(\sup_{0 \leq t < \infty} \psi \left(\bar{\mathbf{x}} \left(t, (x_1^0, x_2^0)^T \right) \right) \geq \delta \right)$$

$$\leq \psi \left(\begin{array}{c} x_1^0 \\ -x_2^0 \end{array} \right) \Big/ \delta = \frac{2\nu}{\delta} \| \bar{x}^0 \|^2. \quad (5)$$

A closer look to Theorem 1 and additional observations on the notion of probabilistic incremental stability suggest that the bound in (5) can be further elaborated, exploiting the monotonically decreasing property in time of function ψ. Given any $x_i^0 \in \mathbb{R}^{n_i}, i = 1, 2, 0 \leq T < \infty$, and fixing a parameter $\delta > 0$, the following holds:

$$P_{x_1^0, x_2^0} \left(\sup_{T \leq t < \infty} \| \bar{\mathbf{y}} \left(t, (x_1^0, x_2^0)^T \right) \|^2 \geq \delta \right)$$
$$\leq \psi \left(\bar{\mathbf{x}} \left(T, (x_1^0, x_2^0)^T \right) \right) / \delta \leq \frac{2\nu}{\delta} \| \bar{x}^0 \|^2 e^{(2\Lambda + K)T}.$$

As argued in [10], the bound in (5) suggests a number of potential applications in formal verification problems.

Remark 3 (Connections with the Literature): The condition described in Theorem 2 for the existence of a probabilistic bisimulation of two diffusion processes can be put in relationship with the results in [10, Lemma 9]. There, a class of stochastic hybrid systems, encompassing random jumps between heterogeneous domains, is studied. The contractivity condition on the systems dynamics in Theorem 2 subsumes the negative-definiteness condition in [10, (40)] (which is indeed related to the classical notion of stability). Furthermore, the scalar quantity 2ν in (5) establishes the validity of the more complicated condition [10, (39)] on the output maps.

The presented conditions have a few potential advantages. First, the contractivity conditions are directly computable on the system dynamics (in fact, as we argued above, it is possible to characterize the portion of the state space where such conditions are valid, and this region is–in a certain sense–invariant). Second, the probabilistic bisimulation function is directly obtained. Third, the conditions are applicable to nonlinear dynamics. (We further elaborate on potential extensions to hybrid models in Section VI.) Finally, this work allows composition of models with separate Wiener processes, whereas [10] drives the system composition by a unique noise process.

Notice also the relationship between the presented results and the work in [8], which established the existence of approximate bisimulation relations for stable linear deterministic models. Finally, let us stress the connections with the results in [17], where the notion of incremental stability [3] is used to compute approximate bisimulation relations for nonlinear systems. We shall draw further comparisons with the work in [17] in the next Section. □

V. COMPUTABLE PROBABILISTIC BISIMULATIONS OF A DIFFUSION PROCESS

In this Section we propose a procedure to approximate an n-dimensional diffusion process $\mathbf{x}(t, x_0)$, solution of the SDE (1) and initialized at x_0, by a discrete-time, discrete-state Markov chain (MC) $\{\mathbf{v}_k, k \geq 0\}$, properly initialized over the discretized state. The procedure follows [11], which introduces general approximation techniques for stochastic processes and discusses their convergence properties, though not from the perspective of the theory of probabilistic bisimulations.

We raise assumptions on a diffusion process and conditions on a discretization procedure, which ensure that a new continuous-time process, obtained by a piecewise-constant interpolation of the MC approximation obtained from the procedure, is a probabilistic bisimulation of the original diffusion, with a given precision. We argue that if the domain of definition of the diffusion process is bounded, then the procedure is decidable, and it yields a MC approximation with finite cardinality. The MC (or its continuous-time interpolation) can be employed in verification problems, according to the relationship established in (5).

A. Discretization Procedure

The discretization procedure that approximates the original diffusion with a MC is characterized by a gridding parameter $\delta > 0$. Let us introduce a homogeneous grid over $\mathbb{R}^n$ of side $\delta > 0$, denoted with $\mathbb{Z}_\delta^n$ and defined as

$$\mathbb{Z}_\delta^n = \{(m_1 \delta, m_2 \delta, \ldots, m_n \delta) | (m_1, m_2, \ldots, m_n) \in \mathbb{Z}^n\}.$$

For any $\delta > 0$, the time step for the MC is denoted with Δt_δ, and is taken so that $\Delta t_\delta = o(\delta)$. The MC evolves on $\mathbb{Z}_\delta^n$ and is referred to as $\{\mathbf{v}_k, k \geq 0\}$. It is characterized by a set of transition probabilities

$$p_\delta(z \rightarrow z') = P_\delta(\mathbf{v}_{k+1} = z' | \mathbf{v}_k = z),$$

which are selected to locally approximate the behavior of the process $\mathbf{x}(\cdot)$. For any point $z \in \mathbb{Z}_\delta^n$, define its neighbors to be the set of points

$$\mathcal{N}_\delta(z) = \{z + (i_1 \delta, i_2 \delta, \ldots, i_n \delta) | (i_1, i_2, \ldots, i_n) \in \mathcal{I}\},$$

where $\mathcal{I} \subseteq \{-1, 0, 1\}^n \setminus \{(0, 0, \ldots, 0)\}$. A trajectory of the MC, located at point z, is allowed to evolve to any of the points in $\mathcal{N}_\delta(z)$ according to the following transition probability:

$$p_\delta(z \rightarrow z') = \pi_\delta(z'|z), z' \in \mathcal{N}_\delta(z) \cup \{z\},$$

where $\pi_\delta(z'|z)$ is an appropriate function of the drift f and the diffusion term σ in (1), both evaluated at z. Assume now that the MC is located at point z at some time k. Introduce

$$m_\delta(z) = \frac{1}{\Delta t_\delta} \mathbb{E}_\delta[\mathbf{v}_{k+1} - \mathbf{v}_k | \mathbf{v}_k = z];$$

$$V_\delta(z) = \frac{1}{\Delta t_\delta} \mathbb{E}_\delta[(\mathbf{v}_{k+1} - \mathbf{v}_k)(\mathbf{v}_{k+1} - \mathbf{v}_k)^T | \mathbf{v}_k = z].$$

For any z, the set of neighbors $\mathcal{N}_\delta(z)$ and the distributions $\pi_\delta(\cdot|z)$ ought to be selected so that, as $\delta \to 0$,

$$m_\delta(z) \to f(x), \qquad V_\delta(z) \to \sigma(x)\sigma^T(x),$$

for all $x \in \mathbb{R}^n$ and where, for $\delta > 0$, z is the point in $\mathbb{Z}^n_\delta$ that is closest to x. See Example V-C or [18] for an instance of such a selection.

Given these choices, it can be shown that the process $\{\mathbf{v}_k, k \geq 0\}$ is Markov. Now let $\{\tau_k, k \geq 0\}$ be an i.i.d. sequence of random variables that are independent of $\{\mathbf{v}_k, k \geq 0\}$, and which are exponentially distributed with mean Δt_δ. Denote with $\mathbf{v}(t, z_0), t \geq 0$, the continuous-time process which is obtained by piecewise constant interpolation of $\{\mathbf{v}_k, k \geq 0\}$, with initial condition in $z_0 \in \mathbb{Z}^n$. The following result is proven in [11, Theorem 4.10]:

Proposition 1: Let us select any point $x_0 \in \mathbb{R}^n$. As $\delta \to 0$, the process $\mathbf{v}(t, z_0)$, which is the continuous-time process interpolated from the discrete-time MC $\{\mathbf{v}_k, k \geq 0\}$ and started from the point $z_0 \in \mathbb{Z}^n_\delta$ closest to x_0, converges weakly to the process $\mathbf{x}(t, x_0)$, solution of (1). □

B. Existence of a Probabilistic Bisimulation

The weak convergence in Proposition 1 is expressed in terms of point-wise convergence of probability distributions. This is an interesting result, however it is not enough to establish a relationship of probabilistic bisimulation between the original stochastic process $\mathbf{x}(t, x_0)$ and the approximated one $\mathbf{v}(t, z_0)$. Next we show that, under proper conditions on the diffusion process $\mathbf{x}(t, \cdot)$, for certain choices of bisimulation precision $\epsilon > 0$ it is possible to select a parameter $\delta < \epsilon$ that induces a MC approximation $\mathbf{v}(t, \cdot)$ that is probabilistically bisimilar to the original diffusion, with precision ϵ.

Theorem 3: Consider the diffusion process $\mathbf{x}(t, x_0)$, solution of (1) and started from $x_0 \in \mathcal{C}$, where $\mathcal{C} \subset \mathbb{R}^n$, and assume it is stochastically contractive (i.e., assume that it verifies the statements in Definition 1, in particular that $2\Lambda + K < 0$). Then, given a value $\epsilon > 0$ there exists a parameter $\delta > 0$ that induces a locally consistent MC $\{\mathbf{v}_k, k \geq 0\}$ (starting from a point $z_0 \in \mathbb{Z}^n \cap \mathcal{C}$ closest to x_0), which approximates $\mathbf{x}(t, x_0)$ on the uniform grid $\mathbb{Z}^n \cap \mathcal{C}$, and such that $\mathbf{v}(t, z_0)$ is probabilistically bisimilar to $\mathbf{x}(t, x_0)$ with precision ϵ, if the following holds:

$$n\delta^2 - \frac{K\eta + 1}{K + 2\Lambda} < \epsilon,$$

where $\eta = \sup_{z_o \in \mathbb{Z}^n \cap \mathcal{C}} \|z_o\|^2$. The precision of the probabilistic bisimulation can be lower bounded by $-\frac{K\eta+1}{K+2\Lambda}$. □

Corollary 1: Assume that the domain $\mathcal{C}$ where (1) is considered is a bounded set. Then the space of the MC abstraction obtained with the proposed procedure has finite cardinality, and the approximation procedure is decidable. □

Remark 4: The result in Theorem 3 resembles a similar solution proposed in [17] for the case of deterministic systems. There, incremental stability is used as the underlying assumption on the system dynamics (in continuous time), which allows the introduction of a procedure that obtains a discrete-space, discrete-time MC. Notice that [17] also allows the presence of a control input, which represents a potential extension of the present work. □

C. Example: Finite Bisimulation of a 2D Diffusion

Let us consider a two-dimensional SDE, with linear drift and diffusion terms, and one-dimensional noise: $d\mathbf{x} = A\mathbf{x}dt + \sigma\mathbf{x}d\mathbf{W}$. Let us select the following parameters:

$$A = \begin{bmatrix} -2 & -1 \\ 1 & -0.9 \end{bmatrix}, \qquad \sigma = 0.1.$$

The dynamics are analyzed on a compact set $\mathcal{C} = [-2, 2]^2$. On $\mathcal{C}$, let us introduce grids of side $\delta \in D = \{0.020, 0.015, 0.010, 0.0075, 0.005\}$. The time discretization for the approximate bisimulations is assumed to be $\Delta t_\delta = \rho\delta^2, \delta \in D$, whereas that for the diffusion is taken to be equal to $\rho(0.005)^2$. Let us consider a time horizon of $T = 4.5$ sec. The parameters of the diffusion yield a contractivity index of $2\Lambda + K = -1.7$, which means that the precision of the bisimulations will be, at best, approximately $\epsilon \approx 0.64$, as per Theorem 3.

The choice of the neighboring points, as well as that of the transition probabilities for the approximations, follows [18]. We have run $m = 100$ tests, in each of which we integrate the dynamics of the diffusion and of the approximate bisimulations over the horizon $[0, T]$. Figure 1 plots a single realization of these processes, of which the diffusion starts from $x_0 = [\sqrt{3}/2, \sqrt{2}]^T$. Furthermore, we have computed for each of the m tests the squared distance, in time, between the approximate bisimulation and the diffusion. Figure 2 plots the average in time, over the m experiments, of this distance. Notice that the average distance is sensibly lower than the computed theoretical bound in Theorem 3, which draws to the search of refinements for it. □

VI. CONCLUSIONS

This work has drawn inspiration from recent results on computable (approximate) bisimulations for deterministic systems, and puts forward three main statements:

1) that, for certain classes of diffusion processes, the concept of (stochastic) contractivity is related to a notion of probabilistic incremental stability (in mean square);

2) that both concepts can be useful in studying the existence of probabilistic bisimulation relations for diffusion processes;

3) and finally that under certain conditions it is possible to construct a (finite) probabilistic bisimulation of a diffusion process, given a certain precision, by selecting a proper parameter for a spatial discretization, and that the construction induces a locally consistent approximation (with the structure of a Markov chain) of the original process.
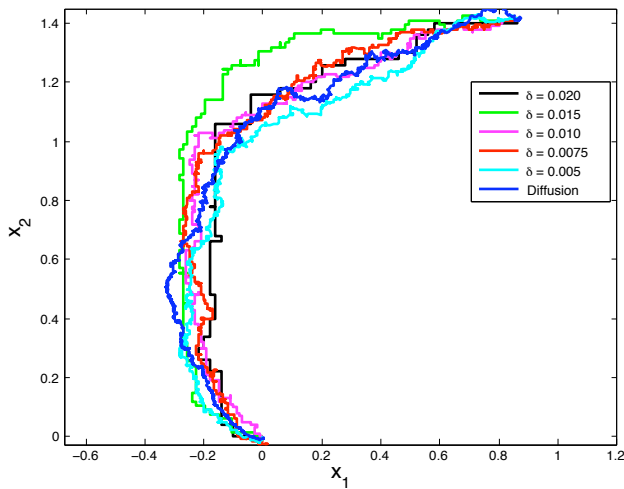
Fig. 1. A single realization of the diffusion process (plotted in blue color), and of the approximated processes. The diffusion starts from $x_0 = [\sqrt{3}/2, \sqrt{2}]^T$. Different colors correspond to different discretization steps (as in the legend).
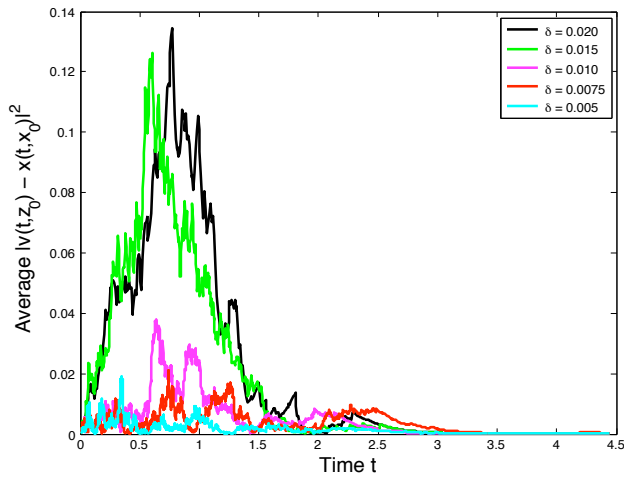

Fig. 2. Average values of the squared distance between the approximated process and the original diffusion, in time. The average is taken over $m = 100$ experiments. Different colors correspond to different discretization steps (as in the legend).

The results are prone to be extended to classes of continuous-time Stochastic Hybrid Systems, along directions that have already been explored in [10]. In particular, the discretization procedure described in Section V can be extended to a general class of switched and hybrid models [18]. A few early results on stochastic contractivity for hybrid models have been derived in [15]. The investigation of formal verification procedures aided by the use of probabilistic bisimulations is also an enticing next step. Furthermore, the study of controlled models certainly represents an important goal, for its relevance in practical applications. Finally, the study of discrete-time, stochastic models appears to require a qualitatively different approach, which may leverage adjacent studies such as [5].

REFERENCES

[1] R. Alur and D.L. Dill. A theory of timed automata. *Theoretical Computer Science*, 126:183–235, 1994.
[2] R. Alur, T. Henzinger, G. Lafferriere, and G. Pappas. Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, 88(2):971–984, July 2000.
[3] David Angeli. A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, 47:410–421, 2002.
[4] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labeled Markov processes. *Information and Computation*, 179(2):163–193, 2002.
[5] A. D'Innocenzo, A. Abate, M.D. Di Benedetto, and S. Sastry. Approximate abstractions of discrete-time controlled stochastic hybrid systems. In *Proceedings of the 47th IEEE Conference of Decision and Control*, pages 221–226, Cancun, MX, December 2008.
[6] Richard Durrett. *Probability: Theory and Examples - Third Edition*. Duxbury Press, 2004.
[7] A. Girard and G.J. Pappas. Approximation metrics for discrete and continuous systems. *IEEE Trans. on Automatic Control*, 52(5):782–798, 2007.
[8] 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.
[9] Antoine Girard, Giordano Pola, and Paulo Tabuada. Approximately bisimilar symbolic models for incrementally stable switched systems. *Under review*, 2008.
[10] A.A. Julius and G.J. Pappas. Approximate abstraction of stochastic hybrid systems. *IEEE Trans. Automatic Control*, 54(6):1193–1203, 2009.
[11] H. J. Kushner and P. G. Dupuis. *Numerical methods for stochastic control problems in continuous time*. Springer Verlag, 1992.
[12] K.G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and Computation*, 94:1–28, 1991.
[13] W. Lohmiller and J.-J. Slotine. On contraction analysis for nonlinear systems. *Automatica*, 34:671–682, 1998.
[14] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, 6th edition, 2003.
[15] Q.-C. Pham. Analysis of discrete and hybrid stochastic systems by nonlinear contraction theory. In *Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Hanoi, Vietnam, 2008.
[16] 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.
[17] Giordano Pola, Antoine Girard, and Paulo Tabuada. Approximately bisimilar symbolic models for nonlinear control systems. *Automatica*, 44(10):2508–2516, October 2008.
[18] M. Prandini and J. Hu. A stochastic approximation method for reachability computations. In H.A.P. Blom and J. Lygeros, editors, *Stochastic hybrid systems: theory and safety applications*, Lecture Notes in Control and Informations Sciences 337, pages 107–139. Springer Verlag, 2006.
[19] Shankar Sastry. *Nonlinear Systems: Analysis, Stability and Control*. Springer Verlag, 1999.