Stochastic Relative Degree and Path-Wise Control of Nonlinear Stochastic Systems

Alberto Mellone ⊙, *Student Member, IEEE*, and Giordano Scarciotti ⊙, *Senior Member, IEEE*

Abstract—In this article, we address the path-wise control of systems described by a set of nonlinear stochastic differential equations. For this class of systems, we introduce a notion of stochastic relative degree and a change of coordinates, which transforms the dynamics to a stochastic normal form. The normal form is instrumental for the design of a state-feedback control, which linearizes and makes the dynamics deterministic. We observe that this control is *idealistic*, i.e., it is not practically implementable because it employs a feedback of the Brownian motion (which is never available) to cancel the noise. Using the idealistic control as a starting point, we introduce a hybrid control architecture, which achieves *practical* path-wise control. This hybrid controller uses measurements of the state to perform periodic compensations for the noise contribution to the dynamics. We prove that the hybrid controller retrieves the idealistic performances in the limit as the compensating period approaches zero. We address the problem of asymptotic output tracking, solving it in the idealistic and in the practical framework. We finally validate the theory by means of a numerical example.

Index Terms—Feedback linearization, normal form, output tracking, relative degree, stochastic systems.

I. INTRODUCTION

A POINT of departure in the study of nonlinear deterministic systems is the definition of the relative degree of the system and, consequently, of a change of coordinates that is able to transform the differential equations in a so-called normal form that makes analysis and control easier. These ideas were first introduced in the seminal work [1], where the authors solved the problem of static state-feedback noninteracting control. The theory of normal forms was later addressed in [2] and [3] for the control and observation of time-varying nonlinear systems, and a systematic overview of normal forms was given in [4]. The problem of feedback linearization of single-input single-output and multi-input systems was introduced in [5] and [6], respectively, and a systematic procedure to find the feedback-linearizing control was provided in [7] and [8]. In [9], the notion

of zero dynamics was introduced and later employed in [10] to tackle the problem of asymptotic stabilization of nonlinear systems.

In this article, we introduce the notion of stochastic relative degree to develop a theory of path-wise feedback control for a general class of systems described by nonlinear stochastic differential equations. The advantage of using this stochastic framework lies in the fact that it allows us to address uncertainties characterized by probabilistic properties that cannot be captured by the classical deterministic robust designs. Modeling in the stochastic framework is flexible and lends itself to mechanical systems (e.g., the quarter-car model), electro-mechanical systems (e.g., the suspended gyro), aerospace systems (e.g., the satellite dynamics), and mathematical finance. A survey of these applications can be found in [11, Sec. 1.9] and references therein. Stochastic differential equations are also at the basis of methodological applications, such as stochastic H_∞ control (see [12]–[14]) and filtering and optimal control (see [15] and [16]).

Some notions of normal forms for stochastic systems have been introduced in the literature. For example, in [17] and [18], Stratonovich calculus was used to obtain a normal form for purely diffusive processes, while [19] employed coordinate changes to introduce symmetries for stochastic differential equations. One of the first works suggesting the convenience of a normal form for control of stochastic systems is [20], where the change of coordinates makes the dynamics quasi-linear. In [21] and [22], similar coordinate transformations are proposed in order to reduce the system dynamics to canonical forms that are amenable to specific control strategies. Specifically, strict-feedback noise-prone dynamics are obtained, which allow achieving optimal globally stabilizing back-stepping controllers as proposed in [23].

The differences between the normal forms mentioned above and the one we introduce in this article are both technical and in scope. From a technical viewpoint, we propose a coordinate projection and a feedback control that *annihilate* the noise in the linearized coordinates. From an objective viewpoint, the stochastic relative degree and the normal form are mere instrumental means for the design of a practical hybrid controller, which achieves output tracking in a path-wise fashion. In other words, the controller that we develop compensates for each *specific* realization of the stochastic disturbance.

More generally, by defining a new stochastic normal form, the goal of this article is to address the path-wise control of stochastic systems described by a general class of nonlinear stochastic differential equations. We show that the implementation of feedback laws that perfectly linearize the system dynamics in a new set of coordinates comes with insurmountable causality issues, hence, the attribute *idealistic* with which we refer to these control

laws. In fact, they employ a feedback of the noise, which is not practically available. However, with these idealistic controls at hand, we introduce hybrid nonlinear controllers that incorporate a causal estimator of the Brownian motion. This controller is practically realisable and, although achieving only approximate feedback linearization and tracking, its performance can be arbitrarily improved by tuning an underlying parameter, retrieving the idealistic case as a limit behavior.

While preliminary work has been published in [24]–[26], in this article, we present several additional contributions. The main novelties introduced are as follows.

1) All the results are now proved, which provides a substantial theoretical contribution.
2) By leveraging the Itô-Stratonovich equivalence, we are able to obtain sharper results in the theory of the stochastic normal form.
3) We provide a characterization of the solvability of the feedback linearization problem.
4) The theory of the practical control has been revised and made sharper.
5) An analysis of the control challenges arising when the input appears in the diffusion term of the equation has been added.
6) Practical asymptotic output tracking has been addressed and solved for the first time.
7) A number of technical results regarding uniform asymptotic stability of nonlinear time-varying stochastic systems have been added and proved. For the sake of readability, they have been gathered in the Appendix.
8) A comprehensive example illustrates the path-wise output tracking of a nonlinear stochastic system both in the idealistic and practical scenarios.

The rest of this article is organized as follows. In Section II, we recall some preliminary notions on stochastic systems. Section III introduces the stochastic relative degree and normal form. In Section IV, we address the problem of feedback linearization in the idealistic framework. In Section V, we propose a hybrid controller, which practically approximates the idealistic linearizing control. Section VI addresses the problem of asymptotic output tracking. In Section VII, we validate the theory by means of a numerical example. Section VIII concludes this article. Finally, technical lemmas, which are instrumental to prove some of the results of this article have been collected and proved in the Appendix.

Notation: The symbol $\mathbb{Z}$ denotes the set of integer numbers, while $\mathbb{R}$ and $\mathbb{C}$ denote the fields of real and complex numbers, respectively; by adding the subscript "<0" ("≥ 0," "0") to any symbol indicating a set of numbers, we denote that subset of numbers with negative (nonnegative, zero) real part. Where convenient, the symbol ∂_x^n is used as a shorthand for the operator $\partial^n/\partial x^n$, while $\alpha^{(n)}$ indicates the nth time derivative of α. The Lie derivative of the smooth scalar function $h(x)$ along the vector field $f(x)$ is denoted by $\mathcal{L}_f h(x)$. We use the recursive notation $\mathcal{L}_f^k h(x) = \mathcal{L}_f \mathcal{L}_f^{k-1} h(x)$, with $\mathcal{L}_f^0 h(x) = h(x)$. Given two smooth vector fields $f(x)$ and $g(x)$, we define the operator $\mathrm{ad}_f g(x) = (\partial_x g(x)) f(x) - (\partial_x f(x)) g(x)$, and, recursively, $\mathrm{ad}_f^k g(x) = \mathrm{ad}_f \mathrm{ad}_f^{k-1} g(x)$ with $\mathrm{ad}_f^0 g(x) = g(x)$. $(\nabla, \mathfrak{A}, \mathfrak{P})$ is a probability space given by the set ∇, the σ-algebra $\mathfrak{A}$ defined on ∇ and the probability measure $\mathfrak{P}$ on the measurable space

$(\nabla, \mathfrak{A})$. A *stochastic process* with state space $\mathbb{R}^n$ is a family $\{x_t, \ t \in \mathbb{R}\}$ of $\mathbb{R}^n$-valued random variables, i.e., for every fixed $t \in \mathbb{R}, x_t(\cdot)$ is an $\mathbb{R}^n$-valued random variable and, for every fixed $w \in \nabla, x.(w)$ is an $\mathbb{R}^n$-valued function of time [27, Sec. 1.8]. For ease of notation, we often indicate a stochastic process $\{x_t, \ t \in \mathbb{R}\}$ simply with x_t (this is common in the literature, see, e.g., [27]). With a slight abuse of notation, any subscript different from the symbol "t" indicates the corresponding component of the vector x_t, e.g., x_i is the i-th component of the vector x_t. All mappings appearing as integrands in stochastic integrals are assumed to be integrable in the corresponding sense, namely Itô's (see, e.g., [16, Def. 3.1.4]) or Stratonovich's (equivalent to deterministic integrability).

II. PRELIMINARIES

In this section, we shortly recall the theory of generalized stochastic processes and define differential operators that will be used in the remainder of this article.

Let $C_0^\infty(\mathbb{R})$ be the space of all infinitely differentiable functions on $\mathbb{R}$ with compact support [28, Def. 1.2.1]. The following definitions characterize the notions of distribution (also known as generalized function), distributional derivative, and generalized stochastic process.

Definition 1 (see [29, Def. 3.1]): Let X be an open subset of $\mathbb{R}$. A *distribution on X* is a linear form ψ on $C_0^\infty(\mathbb{R})$ that is also continuous in the sense that

$$\lim_{j \to \infty} \psi(\varphi_j) = \psi(\varphi) \qquad \text{as} \qquad \lim_{j \to \infty} \varphi_j = \varphi \ \ in \ \ C_0^\infty(\mathbb{R}).$$

Definition 2 (see [28, Def. 3.1.1]): For any distribution ψ, its *distributional derivative* $\dot{\psi}$ is defined as the distribution that satisfies

$$\dot{\psi}(\varphi) = -\psi(\dot{\varphi}) \qquad \forall \varphi \in C_0^\infty(\mathbb{R}).$$

Note that generalized functions have derivatives of all order, which are generalized functions as well.

Definition 3 (see [27, Sec. 3.2]): A *generalized stochastic process* is a random generalized function in the sense that a random variable $\psi(\varphi)$ is assigned to every $\varphi \in C_0^\infty$, where ψ is, with probability 1, a generalized function.

We now look at the Brownian motion as a generalized stochastic process. Therefore, its distributional derivative is always defined [27, Sec. 3.2]. In particular, the generalized stochastic process given by such a derivative has zero mean value and covariance function given by the generalized function $\delta(t - s)$, $t, s \in \mathbb{R}$, i.e., the Dirac delta. Consequently, the derivative of the generalized Brownian motion is the generalized *white noise* [27, Sec. 3.2]. In the remainder, with a slight abuse of notation, we refer to generalized Brownian motion and generalized white noise omitting the attribute "generalized" and we denote them by simply $\mathcal{W}_t$ and ξ_t, respectively, with $\xi_t = \dot{\mathcal{W}}_t$. It should be emphasized that the just mentioned time derivative is meant in the sense of distributions, and not as the limit of the difference quotient as the increment tends to zero, which instead applies to differentiable functions in the classical sense.

Consider the nonlinear single-input, single-output stochastic system expressed in the shorthand integral notation by

$$dx_t = (f(x_t) + g(x_t)u)dt + (l(x_t) + m(x_t)u)d\mathcal{W}_t$$
$$y_t = h(x_t) \tag{1}$$

with $x_t \in \mathbb{R}^n, u \in \mathbb{R}, y_t \in \mathbb{R}$ and $f : \mathbb{R}^n \to \mathbb{R}^n, g : \mathbb{R}^n \to \mathbb{R}^n$, $l : \mathbb{R}^n \to \mathbb{R}^n$, $m : \mathbb{R}^n \to \mathbb{R}^n$, $h : \mathbb{R}^n \to \mathbb{R}$ *smooth* functions, i.e., they admit continuous partial derivatives of any order. We assume that, for a fixed initial condition $x_{t=0}$, the solution of (1) is unique. Note that, in the light of the previous discussion, system (1) can be rewritten in the following differential notation:

$$\dot{x}_t = f(x_t) + g(x_t)u + (l(x_t) + m(x_t)u)\xi_t, \quad y_t = h(x_t). \tag{2}$$

Note that when ξ_t is (generalized) white noise, as in this case, then the differential equation (2) is equivalent to the integral equation (1) if the latter is interpreted in Itô's sense [27, Sec. 10.3]. Given the equivalence of the two representations in the framework of generalized stochastic processes, in the remainder of this article, (1) and (2) are used interchangeably, as convenient, to refer to the same nonlinear stochastic system. We refer the reader to [27, Ch. 3] for a detailed discussion on the relation between Brownian motion and white noise, and the representations (1) and (2).

III. STOCHASTIC RELATIVE DEGREE AND NORMAL FORM

In this section, we introduce the concept of stochastic relative degree and show that a suitable coordinate transformation brings the system into a simpler form, which is convenient for analysis and control.

We first introduce three new operators, which are fundamental to systematically define repeated time derivatives of stochastic processes. The first one, which indicates the second derivative of h along the vector fields f and g, is defined as

$$^g\mathcal{G}_f h(x) = g(x)^\top \partial_x^2[h] \, f(x) = \sum_{j=1}^n g_j(x) \sum_{i=1}^n \frac{\partial^2 h}{\partial x_j \partial x_i} f_i(x).$$

Similarly to the Lie derivative, we use the notation $^b\mathcal{G}_a{}^g\mathcal{G}_f h(x) = b(x)^\top \partial_x^2[^g\mathcal{G}_f h] \, a(x)$, and $^g\mathcal{G}_f^k h(x) = g(x)^\top \partial_x^2[^g\mathcal{G}_f^{k-1}h] \, f(x)$, to indicate the reiterated operations. The second operator $^l\mathcal{S}_f h$, which we call the *stochastic Lie derivative*, indicates the derivative of h along the drift vector field f and diffusion vector field l, namely

$$^l\mathcal{S}_f h(\xi_t, x) = \mathcal{L}_f h(x) + \mathcal{L}_l h(x)\xi_t + \frac{1}{2}{}^l\mathcal{G}_l h(x).$$

The reiterated application of this operator can be defined if the white noise does not appear explicitly. That is, if $^l\mathcal{S}_f h(\xi_t, x) = {}^l\mathcal{S}_f h(x)$ is a deterministic expression, we use the notation $^l\mathcal{S}_f^2 h(\xi_t, x) = {}^l\mathcal{S}_f{}^l\mathcal{S}_f h(\xi_t, x)$ and, iteratively, if $^l\mathcal{S}_f^{k-1} h(\xi_t, x) = {}^l\mathcal{S}_f^{k-1} h(x)$ is deterministic, $^l\mathcal{S}_f^k h(\xi_t, x) = {}^l\mathcal{S}_f{}^l\mathcal{S}_f^{k-1} h(\xi_t, x)$, with $^l\mathcal{S}_f^0 h(x) = h(x)$ by definition. Finally, we define the third operator

$$^m_g\mathcal{A}_l h(\xi_t, x) = \mathcal{L}_g h(x) + \mathcal{L}_m h(x)\xi_t + {}^m\mathcal{G}_l h(x). \tag{3}$$

Having defined the operators $\mathcal{G}$, $\mathcal{S}$, and $\mathcal{A}$, it is easy to see that, by using Itô's formula, the first derivative of the output of system (2) is given by

$$y_t^{(1)} = {}^l\mathcal{S}_f h(\xi_t, x_t) + {}^m_g\mathcal{A}_l h(\xi_t, x_t)u + \frac{1}{2}{}^m\mathcal{G}_m h(x_t)u^2. \tag{4}$$

We now define the concept of stochastic relative degree and then point out the rationale of such a definition.

Definition 4 (Stochastic Relative Degree): System (2) is said to have *stochastic relative degree* r at a point $\bar{x}$ if

(ND) $\mathcal{L}_l{}^l\mathcal{S}_f^k h(x) = 0$ and $\mathcal{L}_m{}^l\mathcal{S}_f^k h(x) = 0$ for all x in a neighborhood of $\bar{x}$ and for all $k \in \{0, \ldots, r-2\}$.

(CD) $\mathcal{L}_g{}^l\mathcal{S}_f^k h(x) + {}^m\mathcal{G}_l{}^l\mathcal{S}_f^k h(x) = 0$, $\mathcal{L}_m{}^l\mathcal{S}_f^k h(x) = 0$ and $^m\mathcal{G}_m{}^l\mathcal{S}_f^k h(x) = 0$ for all x in a neighborhood of $\bar{x}$ and all $k \in \{0, \ldots, r-2\}$.

(RD) $\mathcal{L}_g{}^l\mathcal{S}_f^{r-1} h(\bar{x}) + {}^m\mathcal{G}_l{}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq 0$ or $\mathcal{L}_m{}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq 0$ or $^m\mathcal{G}_m{}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq 0$.

As it is formally proved in the following proposition, in Definition 4, condition (ND) (which stands for noise decoupling) ensures that the noise ξ_t does not appear in y_t and its first $r-1$ derivatives. In the remainder, we will omit the dependency of the operators $\mathcal{S}$ and $\mathcal{A}$ on the white noise ξ_t whenever this does not appear explicitly [for instance, because of condition (ND)]. Condition (CD) (for control decoupling) ensures that the control input u does not appear in y_t and its first $r-1$ derivatives. Finally, condition (RD) (for relative degree) ensures that the control u appears in the r-th derivative of y_t, thus defining the relative degree of the system. These observations are formally gathered in the following result, the proof of which can be found in Appendix A.

Proposition 1: Suppose that system (2) has stochastic relative degree $r > 0$ at $\bar{x}$. Then

$$y_t^{(k)} = {}^l\mathcal{S}_f^k h(x_t) \qquad \forall k \in \{0, \ldots, r-1\} \tag{5}$$

i.e., the first $r-1$ derivatives of y_t do not depend explicitly on ξ_t nor u. Moreover, if, at time $\bar{t}$, $x_{\bar{t}} = \bar{x}$, then

$$y_{t=\bar{t}}^{(r)} = {}^l\mathcal{S}_f^r h(\xi_{\bar{t}}, \bar{x}) + {}^m_g\mathcal{A}_l{}^l\mathcal{S}_f^{r-1} h(\xi_{\bar{t}}, \bar{x})u(\bar{t})$$
$$+ \frac{1}{2}{}^m\mathcal{G}_m{}^l\mathcal{S}_f^{r-1} h(\bar{x})u(\bar{t})^2 \tag{6}$$

where either $^m_g\mathcal{A}_l{}^l\mathcal{S}_f^{r-1} h(\xi_{\bar{t}}, \bar{x})$ or $^m\mathcal{G}_{ml}{}^l\mathcal{S}_f^{r-1} h(\bar{x})$ are nonzero.

The previous result shows that, analogously to the deterministic case, the stochastic relative degree is equal to the order of the derivative of the output at time $\bar{t}$ in which the input $u(\bar{t})$ explicitly appears. Two observations are in order: first, while the white noise does not appear in all the derivatives up to order $r-1$ because of condition (ND), it may or may not appear in the rth derivative; second, differently from the deterministic case, the control u appears linearly and/or quadratically in (6).

Remark 1: Condition (ND) is a type of *disturbance decoupling* condition, as we suppose that the noise does not appear in y_t and its successive $r-1$ derivatives. For a deterministic analogous, see, e.g., [30, Sec. 4.6]. If $\mathcal{L}_l{}^l\mathcal{S}_f^k h(x) \neq 0$ for a $k < r-1$, the differentiation of y_t up to the rth time would require us to introduce successive derivatives of the white noise, which is theoretically and practically challenging. We exclude this possibility with (ND).

Remark 2: The equality $\mathcal{L}_m{}^l\mathcal{S}_f^k h(x) = 0$ appears in both conditions (ND) and (CD). This repetition is unnecessary. However, since $\mathcal{L}_m{}^l\mathcal{S}_f^k h(x)$ multiplies both ξ_t and u in the expression of the derivative $y_t^{(k+1)}$, we believe that for the sake of clarity it is beneficial to require it to be zero in both the noise and control decoupling conditions.

Remark 3: Consider a linear stochastic system, i.e., system (2) with $f(x_t) = Ax_t$, $g(x_t) \equiv B$, $l(x_t) = Fx_t$ and $m(x_t) \equiv G$, where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $F \in \mathbb{R}^{n \times n}$ and $G \in \mathbb{R}^{n \times 1}$. By applying Definition 4, system (2) in the linear case has stochastic relative degree r if

1) $CA^k B = 0$ and $CA^k G = 0$ for all $k \in \{0, \ldots, r-2\}$.
2) $CA^{r-1}B \neq 0$ or $CA^{r-1}G \neq 0$.

Note that the conditions are remarkably simple and reminiscent of the deterministic case. In fact, $^l\mathcal{S}_f^k h(x) = CA^k x$, i.e., it is linear in x, for all $k \in \{0, \ldots, r-1\}$, hence $^l\mathcal{G}_l\,^l\mathcal{S}_f^k h(x) \equiv$ $^m\mathcal{G}_l\,^l\mathcal{S}_f^k h(x) \equiv\,^m\mathcal{G}_m\,^l\mathcal{S}_f^k h(x) \equiv 0$ because the Hessian of a linear function with respect to x is identically zero. Assuming that the relative degree of this system is, for instance, $r > 2$, it follows that

$$y_t^{(k)} = CA^k x_t \qquad \forall k \in \{0, \ldots, r-1\}$$
$$y_t^{(r)} = CA^{r-1}(A + F\xi_t)x_t + CA^{r-1}(B + G\xi_t)u.$$

Having defined the notion of stochastic relative degree and having discussed its interpretation, we are now interested in finding a diffeomorphism $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ that locally (i.e., in a neighborhood $\bar{U}$ of $\bar{x} \in U \subset \mathbb{R}^n$) transforms system (2) in such a way that its dynamics is somewhat "simpler." The diffeomorphism we are looking for is a direct consequence of the definition of the stochastic relative degree previously given. To simplify the exposition, we make the following assumption on the stochastic Lie derivatives of $y_t = h(x_t)$ along the drift vector f and the diffusion vector l.

Assumption 1: Let r be the stochastic relative degree of system (2) at $\bar{x}$. Then, the row vectors

$$\partial_x[h]_{x=\bar{x}}, \ \partial_x\left[^l\mathcal{S}_f h\right]_{x=\bar{x}}, \ \ldots, \ \partial_x\left[^l\mathcal{S}_f^{r-1} h\right]_{x=\bar{x}}$$

are linearly independent.

Observe that if Assumption 1 holds, then necessarily $r \leq n$.

Remark 4: For deterministic nonlinear systems, the linear independence of the gradients of the first $r-1$ successive derivatives of the output at $\bar{x}$ is a consequence of the relative degree being defined, see, e.g., [30, Lemma 4.1.1]. In Section III-A, we prove that this property holds in the present setting for the case in which $m \equiv 0$ near $\bar{x}$. However, a proof that a similar result holds in general ($m \neq 0$) is missing. For simplicity, we use Assumption 1 at this stage to develop the theory of normal form for the most general class of systems, and we note that no counter-example has been found for which this property is not satisfied when $m \neq 0$ near $\bar{x}$.

Proposition 2: Suppose that system (2) has stochastic relative degree r at $\bar{x}$ and let Assumption 1 hold. Set

$$\phi_1(x) = h(x), \ \phi_2(x) = {}^l\mathcal{S}_f h(x), \ \ldots, \ \phi_r(x) = {}^l\mathcal{S}_f^{r-1}h(x).$$

If $r < n$, then there exist smooth functions $\phi_{r+1}(x), \ldots, \phi_n(x)$, with $\phi_j \in \mathbb{R}$ for all $j \in \{r+1, \ldots, n\}$, such that the Jacobian of the mapping

$$\Phi(x) = \begin{bmatrix} \phi_1(x) & \phi_2(x) & \cdots & \phi_n(x) \end{bmatrix}^\top \tag{7}$$

is invertible at $\bar{x}$ almost surely, thus defining a coordinate transformation in a neighborhood of $\bar{x}$. Then, the state-space representation of system (2) in the transformed state

$z_t = \Phi(x_t)$ is

$$\dot{z}_i = z_{i+1}, \qquad\qquad i = 1, \ldots, r-1$$
$$\dot{z}_r = c(\xi_t, z_t) + b(\xi_t, z_t)u + a(z_t)u^2$$
$$\dot{z}_j = p_j(\xi_t, z_t) + q_j(\xi_t, z_t)u + s_j(z_t)u^2, \quad j = r+1, \ldots, n$$
$$y_t = z_1 \tag{8}$$

where the mappings[1] c, b, p_j, and q_j are affine in ξ_t.

Proof: By Assumption 1, the matrix

$$\left[\partial_x[h(x)]^\top \quad \cdots \quad \partial_x\left[{}_l\mathcal{S}_f^{r-1}h(x)\right]^\top \right]^\top$$

has rank r at $\bar{x}$. If $r < n$, let $\gamma_{r+1}(x), \ldots, \gamma_n(x)$, with $\gamma_j \in \mathbb{R}^n$ for all $j \in \{r+1, \ldots, n\}$, be any set of $n-r$ vectors such that

$$\left[\partial_x[h(\bar{x})]^\top \cdots \partial_x\left[{}_l\mathcal{S}_f^{r-1}h(\bar{x})\right]^\top \ \gamma_{r+1}(\bar{x}) \ \cdots \ \gamma_n(\bar{x}) \right]^\top$$

has rank n. Note that this is possible because there always exist $n-r$ linearly independent vectors $\gamma_{r+1}(\bar{x}), \ldots, \gamma_n(\bar{x})$ that complete the first r linearly independent vectors $\partial_x[h(\bar{x})]^\top, \ldots, \partial_x\left[{}_l\mathcal{S}_f^{r-1}h(x)\right]^\top$ to a basis of $\mathbb{R}^n$. Let $\phi_j(x)$ be any smooth function such that $\partial_x[\phi_j(x)] = \gamma_j^\top(x)$ for $j = r+1, \ldots, n$. Then, $\Phi(x)$ as defined in (7) is a local diffeomorphism in a neighborhood of $\bar{x}$ and, therefore, it defines a local change of coordinates $z_t = \Phi(x_t)$ for the stochastic system (2). Applying Itô's lemma and since the system has relative degree r, the following holds

$$\dot{z}_i = {}^l\mathcal{S}_f^i h(x_t) = \phi_{i+1}(x_t) = z_{i+1}, \qquad i = 1, \ldots, r-1.$$

Moreover

$$\dot{z}_r = {}^l\mathcal{S}_f^r h(x_t) + {}_g^m\mathcal{A}_l\,{}^l\mathcal{S}_f^{r-1}h(x_t)u + \frac{1}{2}{}^m\mathcal{G}_m\,{}^l\mathcal{S}_f^{r-1}h(x_t)u^2.$$

We now set

$$c(\xi_t, z_t) = {}^l\mathcal{S}_f^r h\left(\xi_t, \Phi^{-1}(z_t)\right)$$
$$b(\xi_t, z_t) = {}_g^m\mathcal{A}_l\,{}^l\mathcal{S}_f^{r-1}h\left(\xi_t, \Phi^{-1}(z_t)\right)$$
$$a(z_t) = \frac{1}{2}{}^m\mathcal{G}_m\,{}^l\mathcal{S}_f^{r-1}h\left(\Phi^{-1}(z_t)\right)$$

thus obtaining

$$\dot{z}_r = c(\xi_t, z_t) + b(\xi_t, z_t)u + a(z_t)u^2.$$

As for the remaining $n-r$ components of z_t, by applying Itô's lemma to the functions $\phi_j(x_t)$ and setting

$$p_j(\xi_t, z_t) = {}^l\mathcal{S}_f\phi_j\left(\xi_t, \Phi^{-1}(z_t)\right)$$
$$q_j(\xi_t, z_t) = {}_g^m\mathcal{A}_l\phi_j\left(\xi_t, \Phi^{-1}(z_t)\right)$$
$$s_j(z_t) = \frac{1}{2}{}^m\mathcal{G}_m\phi_j\left(\Phi^{-1}(z_t)\right)$$

yields

$$\dot{z}_j = p_j(\xi_t, z_t) + q_j(\xi_t, z_t)u + s_j(z_t)u^2, \qquad j = r+1, \ldots, n.$$

---

[1]To keep the statement of the proposition concise, the mappings a, b, c, p_j, q_j, s_j are defined in the proof.

The proof is completed by observing that $y_t = h(x_t) = z_1$ and that, by the definitions of the operators $\mathcal{S}$ and $\mathcal{A}$, the coefficients c, b, p_j, and q_j are affine in ξ_t. ∎

Note that it might be possible to find smooth functions $\phi_{r+1}, \ldots, \phi_n$ such that the dynamics of the last $n - r$ transformed coordinates is independent of the input u, i.e., $q_j(\cdot, z_t) \equiv 0$, $s_j(\cdot, z_t) \equiv 0$, for all $j \in \{r+1, \ldots, n\}$, in a neighborhood of $\Phi(\bar{x})$. This observation motivates the following definition.

Definition 5 (Stochastic Normal Form): Let x_t be the unique solution of (2) and $z_t = \Phi(x_t)$ be a local diffeomorphism in a subset U of $\mathbb{R}^n$ such that

$$\dot{z}_i = z_{i+1}, \qquad\qquad\qquad i = 1, \ldots, r-1$$
$$\dot{z}_r = c(\xi_t, z_t) + b(\xi_t, z_t)u + a(z_t)u^2$$
$$\dot{z}_j = p_j(\xi_t, z_t), \qquad\qquad j = r+1, \ldots, n$$
$$y_t = z_1. \tag{9}$$

System (9) is said to be the *stochastic normal form* of system (2).

Remark 5: The fact that the coefficients c, b, p_j, and q_j are affine in ξ_t guarantees that system (8), which is written in the differential notation [as in (2)], can always be equivalently written in the integral notation [as in (1)].

For compactness, in the remainder, we use the definitions $p = \begin{bmatrix} p_{r+1} & \cdots & p_n \end{bmatrix}^\top$, $q = \begin{bmatrix} q_{r+1} & \cdots & q_n \end{bmatrix}^\top$, and $s = \begin{bmatrix} s_{r+1} & \cdots & s_n \end{bmatrix}^\top$. Obviously, if the stochastic relative degree at $\bar{x}$ is equal to the order of the system, then the system admits a stochastic normal form in a neighborhood U of $\bar{x}$ (because p, q, and s have dimension zero).

Remark 6: For deterministic systems, it can be proved (see, e.g., [30, Prop. 4.1.3]) that functions $\phi_{r+1}, \ldots, \phi_n$ always exist such that a normal form exists when $r < n$. While a proof that this property holds in general in the present setting is missing, in Section III-A, we prove that this property holds for the case in which $m \equiv 0$ near $\bar{x}$.

Remark 7: The notion of stochastic relative degree and normal form presented are consistent with Itô's interpretation. This is without loss of generality, as all the results of this section can be obtained also in other formalisms, *e.g.,* Stratonovich's formalism [31].

A. Sharper Results for $m \equiv 0$

In the previous section, we introduced the concept of relative degree and of normal form for a class of nonlinear stochastic systems in which the control input appears both in the drift and in the diffusion terms of the stochastic differential equation. This allowed us to provide as general as possible definitions. However, for the remainder of this article, it is beneficial to make the standing assumption that the control u does not enter the diffusion term of the stochastic differential equation (i.e., $m \equiv 0$) in a neighborhood of $\bar{x}$. On the one hand, this subclass of systems is more common in the literature (see, e.g., [21], [22] and references therein). On the other hand, this assumption allows us to achieve sharper results in terms of nonlinear control of stochastic systems, both in the idealistic case, i.e., when the noise process is assumed available) and, in the practically implementable controller that we develop in Section V-B. A discussion of the general case (i.e., $m \neq 0$) is given in Section V-C. Therefore, for the time being, we assume $m(x_t) \equiv 0$ near $\bar{x}$ and

we consider systems of the form

$$dx_t = (f(x_t) + g(x_t)u)dt + l(x_t)d\mathcal{W}_t, \qquad y_t = h(x_t) \tag{10}$$

or, equivalently

$$\dot{x}_t = f(x_t) + g(x_t)u + l(x_t)\xi_t, \qquad y_t = h(x_t). \tag{11}$$

In order to prove results in this section, it is useful to introduce the *Stratonovich equivalent* of system (10), given by

$$dx_t = (f_S(x_t) + g(x_t)u)dt + l(x_t) \circ d\mathcal{W}_t, \quad y_t = h(x_t) \tag{12}$$

where $f_S(x) = f(x) - \frac{1}{2}\frac{\partial l(x)}{\partial x}l(x)$, and the symbol $\circ$ denotes the fact that the stochastic integral is meant in Stratonovich's sense. The Itô solution x_t of (10) is identical to the Stratonovich solution of (12) (see, e.g., [32]).

We now discuss what implications follow from restricting ourselves to the class of systems (11). First, the operator $_g^m\mathcal{A}_l$, reduces to $\mathcal{L}_g$, the Lie derivative along the only control vector field g. Second, as the control input does not appear in the diffusion term, the derivative of any function of the state x_t is never quadratic in the control input u. In fact one can observe that u^2 multiplies the operator $^m\mathcal{G}_m$, e.g., in (4), and that this operator is identically zero near $\bar{x}$ because so is the vector field m. Having said this, the definition of stochastic relative degree becomes simpler and for the sake of clarity, it is useful to rewrite it.

Definition 6 (Stochastic Relative Degree $-m \equiv 0$): System (11) is said to have *stochastic relative degree* r at a point $\bar{x}$ if

(ND) $\mathcal{L}_l{}^l\mathcal{S}_f^k h(x) = 0$ for all x in a neighborhood of $\bar{x}$ and for all $k \in \{0, \ldots, r-2\}$.

(CD) $\mathcal{L}_g{}^l\mathcal{S}_f^k h(x) = 0$ for all x in a neighborhood of $\bar{x}$ and all $k \in \{0, \ldots, r-2\}$.

(RD) $\mathcal{L}_g{}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq 0$.

By considering a class of systems of the form (11), we can achieve sharper results in terms of definition of a stochastic normal form. Indeed, we can now prove the claim previously assumed in Assumption 1. To this end, we first prove a technical result.

Lemma 1: Let r be the stochastic relative degree of system (11) at $\bar{x}$. Then

$$^l\mathcal{S}_f^k h = \mathcal{L}_{f_S}^k h \qquad \forall k \in \{0, \ldots, r-1\} \quad \text{near } \bar{x}. \tag{13}$$

Proof: First observe that by (ND)

$$\partial_x \left(\mathcal{L}_l{}^l\mathcal{S}_f^k h \right) = l^\top \frac{\partial^2{}^l\mathcal{S}_f^k h}{\partial x^2} + \frac{\partial^l\mathcal{S}_f^k h}{\partial x}\frac{\partial l}{\partial x} = 0$$

hence

$$l^\top \frac{\partial^2{}^l\mathcal{S}_f^k h}{\partial x^2} = -\frac{\partial^l\mathcal{S}_f^k h}{\partial x}\frac{\partial l}{\partial x} \tag{14}$$

for all k in $\{0, \ldots, r-2\}$ near $\bar{x}$. Then, the claim follows by induction. In fact, (13) trivially holds for $k = 0$. Note that for some $0 < k < r - 1$ we have, in view of (14), that

$$^l\mathcal{S}_f^{k+1} h = \mathcal{L}_f{}^l\mathcal{S}_f^k h + \frac{1}{2}l^\top\frac{\partial^2{}^l\mathcal{S}_f^k h}{\partial x^2}l$$

$$= \mathcal{L}_f{}^l\mathcal{S}_f^k h - \frac{1}{2}\frac{\partial^l\mathcal{S}_f^k h}{\partial x}\frac{\partial l}{\partial x}l = \mathcal{L}_{f-\frac{1}{2}\frac{\partial l}{\partial x}l}{}^l\mathcal{S}_f^k h.$$

Rewriting this equation using in the Lie derivative notation and assuming by the inductive hypothesis that (13) holds for $0 < k < r - 1$ yields

$$^l\mathcal{S}_f^{k+1} h = \mathcal{L}_{f - \frac{1}{2}\frac{\partial l}{\partial x} l} \mathcal{L}_{f_S}^k h = \mathcal{L}_{f_S} \mathcal{L}_{f_S}^k h = \mathcal{L}_{f_S}^{k+1} h.$$

∎

The fact that the stochastic Lie derivatives of h along f and l are identical to the Lie derivatives of h along f_S up to order $r - 1$ is crucial to prove the rest of the results in this and the following section. In fact, this equivalence allows us to leverage the deterministic techniques, because the Stratonovich differentiation rule is formally identical to the deterministic one. We can then use the identities in Lemma 1 to translate the results to the Itô system (11).

We are now ready to prove the following.

Lemma 2: Let r be the stochastic relative degree of system (11) at $\bar{x}$. Then, the row vectors

$$\partial_x[h]_{x=\bar{x}}, \ \partial_x \left[{}^l\mathcal{S}_f h \right]_{x=\bar{x}}, \ \ldots, \ \partial_x \left[{}^l\mathcal{S}_f^{r-1} h \right]_{x=\bar{x}}$$

are linearly independent.

Proof: Lemma 1 and the definition of stochastic relative degree of (11) imply that the set of conditions $\mathcal{L}_g \mathcal{L}_{f_S}^k h = 0$ near $\bar{x}$ for all k in $\{0, \ldots, r - 2\}$ and $\mathcal{L}_g \mathcal{L}_{f_S}^{r-1} h \neq 0$ at $\bar{x}$ hold. By [30, Lemma 4.1.1], this set of conditions in turn implies the linear independence of

$$\partial_x[h]_{x=\bar{x}}, \ \partial_x[\mathcal{L}_{f_S} h]_{x=\bar{x}}, \ \ldots, \ \partial_x \left[\mathcal{L}_{f_S}^{r-1} h \right]_{x=\bar{x}}.$$

Using again Lemma 1, the claim follows. ∎

For systems of the form (11), the result in Proposition 2 can be specialized to the existence of a coordinate transformation $z_t = \Phi(x_t)$ such that the transformed dynamics is given by

$$\dot{z}_i = z_{i+1}, \qquad\qquad i = 1, \ldots, r-1$$
$$\dot{z}_r = c(\xi_t, z_t) + b(z_t)u$$
$$\dot{z}_j = p_j(\xi_t, z_t) + q_j(z_t)u, \qquad j = r+1, \ldots, n$$

where the coefficients c and p_j are as before and

$$b(z_t) = \mathcal{L}_g {}^l\mathcal{S}_f^{r-1} h \left(\Phi^{-1}(z_t) \right), \quad q_j(z_t) = \mathcal{L}_g \phi_j \left(\Phi^{-1}(z_t) \right).$$

Observe that, unlike the general case in which $m \neq 0$, since now the coefficients q_j are simply the Lie derivatives of ϕ_j along g for $j = r+1, \ldots, n$, it is easy to show that it is always possible to find $\phi_{r+1}, \ldots, \phi_n$ such that said coefficients are identically zero for x in a neighborhood of $\bar{x}$. Exploiting Lemma 1, the proof is analogous to the one reported in [30, Prop. 4.1.3]. Thus, we can conclude that it is always possible to find a coordinate transformation $\Phi(x)$ in a subset U of $\mathbb{R}^n$ such that the dynamics in the transformed state $z_t = \Phi(x_t)$ is

$$\dot{z}_i = z_{i+1}, \qquad\qquad i = 1, \ldots, r-1$$
$$\dot{z}_r = c(\xi_t, z_t) + b(z_t)u \qquad\qquad (15)$$
$$\dot{z}_j = p_j(\xi_t, z_t), \qquad\qquad j = r+1, \ldots, n$$
$$y_t = z_1$$

which is the normal form of system (11).

Remark 8: The notion of normal form which we propose introduces elements of significant novelty, compared to works on analogous topics like [20]–[22]. By condition (ND) in Definition 6, the state is projected onto new coordinates, of which the first $r - 1$ have a noise-free dynamics. Not only is this a fundamental difference with respect to past works, but it also allows us to introduce, as shown in Sections IV and V, control laws performing path-wise control of nonlinear stochastic systems. To the best of the authors' knowledge, this problem has not been systematically addressed in the literature.

IV. IDEALISTIC LINEARIZATION VIA STATE FEEDBACK

We now turn our attention to the problem of exact feedback linearization. Specifically, in this section, we formulate and solve an idealistic version of the problem, by assuming that the control input is a function of the noise ξ_t. Although unrealistic because it implies that the noise is available for feedback, this assumption is necessary to develop the theory of *exact* linearization via feedback. This theory is instrumental for the introduction of a practically implementable controller in Section V which, by approximating the noise contribution to the dynamics via measurements of the state, achieves a *practical* result. The feedback linearization problem is formally defined as follows.

Problem 1 (Exact Feedback Linearization): Consider the nonlinear stochastic system

$$\dot{x}_t = f(x_t) + g(x_t)u + l(x_t)\xi_t. \qquad (16)$$

Given a point $\bar{x}$, the problem of *exact feedback linearization* consists in finding a neighborhood U of $\bar{x}$, a feedback law $u_t = k(\xi_t, x_t, v)$ affine in ξ_t, with $v \in \mathbb{R}$, defined on U and a stochastic coordinate transformation $z_t = \Phi(x_t)$ defined on U such that the closed-loop system

$$\dot{x}_t = f(x_t) + g(x_t)k(\xi_t, x_t, v) + l(x_t)\xi_t$$

in the coordinates $z_t = \Phi(x_t)$, is linear, deterministic, and controllable.

Remark 9: By requiring that the control $u_t = k(\xi_t, x_t, v)$ is an explicit function of the white noise, as in the statement of Problem 1, we may be enlarging the class of systems (11) beyond what could be allowed. In the general case where k is any nonlinear function of ξ_t, the resulting closed-loop system in the derivative notation, i.e., (10), is not anymore equivalent to the one in the differential notation, i.e., (11). In fact, this equivalence is preserved only when the resulting closed-loop dynamics is affine in the white noise. Therefore, we hereby define an idealistic control law to be *admissible* if the dynamics of the closed-loop system is affine in ξ_t. This is achieved easily by requiring that k is affine in ξ_t as done in the statement of Problem 1. Note that the idealistic controls that we introduce in this article are all admissible by construction.

Proposition 3: Problem 1 is solvable if and only if there exists a real valued function $h(x_t)$ such that system (16) with the output $y_t = h(x_t)$ has stochastic relative degree n at $\bar{x}$.

Proof: Sufficiency. Suppose $y_t = h(x_t)$ is such that the stochastic relative degree of system (16) is n at $\bar{x}$. Then, by Proposition 2, there exists a change of coordinates

$$z_t = \Phi(x_t) = \begin{bmatrix} h(x_t) & {}^l\mathcal{S}_f h(x_t) & \cdots & {}^l\mathcal{S}_f^{n-1} h(x_t) \end{bmatrix}^\top$$

for all x in a neighborhood U of $\bar{x}$, such that the normal form of system (16) is

$$\dot{z}_i = z_{i+1}, \quad i = 1, \ldots, n-1$$

$$\dot{z}_n = c(\xi_t, z_t) + b(z_t)u.$$

By the definition of stochastic relative degree, there exists a neighborhood U of $\bar{x}$ such that $b(z_t) \neq 0$ in $\Phi(U)$. Therefore, the control law

$$u_t = \tilde{k}(\xi_t, z_t, v) = \frac{1}{b(z_t)}(-c(\xi_t, z_t) + v) \qquad (17)$$

is well-defined in $\Phi(U)$. Therefore, in the neighborhood U of $\bar{x}$ the feedback law $u_t = \tilde{k}(\xi_t, \Phi(x_t), v) = k(\xi_t, x_t, v)$ defined on U brings the transformed system with state $z_t = \Phi(x_t)$ into the form

$$\dot{z}_t = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} z_t + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} v = Az_t + Bv$$

where A and B are such that $\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$ has rank n. The transformed system is linear, deterministic and controllable, hence, Problem 1 is solved.

Necessity: This proof is formally identical to the one of [30, Lemma 4.2.1]. In fact, it is possible to show that the stochastic relative degree is invariant under coordinate change and feedback by using the arguments in [30] on the Stratonovich equivalent system (12) and then use Lemma 1 to show that these arguments hold for the Itô system (11) as well. ∎

Remark 10: As the coefficient c is in general a function of ξ_t, the feedback linearizing control (17) could require the knowledge of the exact value of the white noise for all t, because this is the only way that the noisy dynamics of the open loop system (16) can be rendered deterministic via feedback. Of course, the assumption that the noise can be used in the feedback loop is unrealistic. However, building on the results of this section, in Section V, we design a practical hybrid control scheme. This hybrid scheme implements a deterministic state feedback law which periodically compensates, in an approximate way, the noise.

Remark 11: The control $u_t = k(\xi_t, x_t, v)$ designed in the proof of Proposition 3 is affine in the variable ξ_t, because in (17) the coefficient b is not an explicit function of ξ_t, whilst the coefficient c is affine in ξ_t. Therefore, the feedback linearizing control is always admissible because replacing its expression in (11) leaves the closed-loop dynamics affine in ξ_t.

To conclude this section, we provide necessary and sufficient conditions for the existence of an output $y_t = h(x_t)$ which makes the stochastic relative degree of system (16) equal to n at a point $\bar{x}$.

Theorem 1: Problem 1 is solvable if and only if
1) the matrix $\begin{bmatrix} g(\bar{x}) & \mathrm{ad}_{f_S}g(\bar{x}) & \cdots & \mathrm{ad}_{f_S}^{n-1}g(\bar{x}) \end{bmatrix}$ is invertible and
2) the distribution $\mathrm{span}\{g, \mathrm{ad}_{f_S}g, \ldots, \mathrm{ad}_{f_S}^{n-2}g\}$ is involutive near $\bar{x}$.

Proof: By Proposition 3, Problem 1 is solvable if and only if it is possible to find an output function $h(x_t)$ satisfying the conditions in Definition 6. By Lemma 1, this amounts to the set of conditions
1) $\mathcal{L}_g \mathcal{L}_{f_S}^k h(x) = 0$ for all x in a neighborhood of $\bar{x}$ and all $k \in \{0, \ldots, r - 2\}$;

2) $\mathcal{L}_g \mathcal{L}_{f_S}^{r-1} h(\bar{x}) \neq 0$,

which is a set of partial differential equations and a nontriviality condition. By [30, Lemma 4.1.2] these are equivalent to
1) $\mathcal{L}_g \mathcal{L}_{\mathrm{ad}_{f_S}g}^k h(x) = 0$ for all x in a neighborhood of $\bar{x}$ and all $k \in \{0, \ldots, r - 2\}$.
2) $\mathcal{L}_g \mathcal{L}_{\mathrm{ad}_{f_S}g}^{r-1} h(\bar{x}) \neq 0$.

Following the proof of [30, Lemma 4.2.2], it is possible to show that the previous set of conditions is equivalent to the set of conditions in the statement of this theorem. ∎

If the stochastic relative degree r of system (11) is strictly less than n at $\bar{x}$, we indicate by $\zeta_t = \begin{bmatrix} z_1 & \cdots & z_r \end{bmatrix}^\top$ the first r components of the transformed state $z_t = \Phi(x_t)$ and by $\eta_t = \begin{bmatrix} z_{r+1} & \cdots & z_n \end{bmatrix}^\top$ the remaining $n - r$. Note that if $r < n$ the control law (17) linearizes only the dynamics of ζ_t, thus, we say that it partially feedback linearizes system (15).

The zero dynamics of a nonlinear stochastic system is, analogously to the deterministic case, the dynamics of the internal variable η_t when the input and the initial conditions are chosen in such a way that the output is constrained to be identically zero. This is achieved by setting $\zeta_0 = 0$ and $u_{z,t} = -c(\xi_t, 0, \eta_t)/b(0, \eta_t)$. Note that the control $u_{z,t}$ is affine in ξ_t, hence admissible. We now provide a definition that extends the concept of zero dynamics to nonlinear stochastic systems of the form (11).

Definition 7 (Zero Dynamics): The stochastic differential equation

$$\dot{\eta}_t = p(\xi_t, 0, \eta_t)$$

is called the *zero dynamics* of system (11).

The properties of the zero dynamics are fundamental in studying the problem of asymptotic output tracking, which is the topic of Section VI.

V. PRACTICAL LINEARIZATION VIA STATE FEEDBACK

In this section, we introduce a causal method to estimate the sequence of variations of the Brownian motion between successive time instants. We then show that the use of these estimates is beneficial in the design of practical feedback linearizing controls.

A. Estimation of the Brownian Motion

The estimation of the Brownian variations is performed by periodically sampling the state. This method was first introduced in [33] in order to practically solve the problem of output regulation of *linear* stochastic systems. We now extend it to the present context of *nonlinear* stochastic systems.

Let $\{t_k\}_{k \in \mathbb{Z}_{\geq 0}}$ be a sequence of equally spaced sampling times, with $t_k - t_{k-1} = \varepsilon$ for all $k \in \mathbb{Z}_{>0}$. Define the differences $\Delta W_\varepsilon(k) = \mathcal{W}_{t_k} - \mathcal{W}_{t_{k-1}}$ and $\Delta x(k) = x_{t_k} - x_{t_{k-1}}$. Our aim is to show that it is possible to compute a causal estimate $\widehat{\Delta W}_\varepsilon(k)$ of the quantity $\Delta W_\varepsilon(k)$ by comparing the samples of the state of the system at times t_{k-1} and t_k. In particular, we want this estimate to "converge," in a sense to be defined, to the stochastic differential $d\mathcal{W}_t$ as the sampling period ε converges to zero. Let $\mathfrak{L}_I$ be the space of functions that are integrable in Itô's sense. Then with the notation $\widehat{\Delta W}_\varepsilon \xrightarrow{\varepsilon} d\mathcal{W}_t$ we mean that for all $\alpha \in \mathfrak{L}_I \lim_{\varepsilon \to 0} \sum_k \alpha(t_{k-1}, w) \widehat{\Delta W}_\varepsilon(k) = \int_0^t \alpha(\tau, w) d\mathcal{W}_\tau$ for almost all $w \in \nabla$. For ease of notation,

define $F_{t_k} = f(x_{t_k}) + g(x_{t_k})u_{t_k}$ and $L_{t_k} = l(x_{t_k})$, which are the drift and diffusion coefficients, respectively, of system (10) evaluated at time t_k. We make the following assumption.

Assumption 2: There exists $\delta > 0$ such that $|L_{t_k}| > \delta$ almost surely for all $k \in \mathbb{Z}_{\geq 0}$.

The rationale of this assumption is explained later, in Remark 12 at the end of Section V. Under Assumption 2, the Moore–Penrose left pseudoinverse of L_{t_k}, i.e., $L_{t_k}^\dagger = (L_{t_k}^\top L_{t_k})^{-1} L_{t_k}^\top$, is well-defined almost surely. The following Lemma extends the results in [33] to systems with nonlinear drift and diffusion terms.

Lemma 3: Consider system (11) and let Assumption 2 hold. Let $\left\{ \Delta \widehat{W}_\varepsilon(k) \right\}_{k>0}$ be a sequence of scalars defined as

$$\Delta \widehat{W}_\varepsilon(k) = L_{t_{k-1}}^\dagger \left[\Delta x(k) - F_{t_{k-1}}\varepsilon \right]. \tag{18}$$

Then, $\Delta \widehat{W}_\varepsilon(k) \xrightarrow{\varepsilon} d\mathcal{W}_t$ almost surely.

Proof: Let $k \in \mathbb{Z}_{>0}$. By [34, Th. 7.1]

$$\Delta x(k) = F_{t_{k-1}}\varepsilon + L_{t_{k-1}}\Delta W_\varepsilon(k) + o\left(\varepsilon^2\right)$$

holds, where $o\left(\varepsilon^2\right)$, which is the *one-step truncation error* of the forward-Euler scheme, is an infinitesimal of the same order of ε^2. The previous expression can be rewritten as

$$L_{t_{k-1}}\Delta W_\varepsilon(k) = \Delta x(k) - F_{t_{k-1}}\varepsilon + o\left(\varepsilon^2\right).$$

Since $L_{t_{k-1}}$ has full column rank almost surely, the expression

$$\Delta W_\varepsilon(k) = L_{t_{k-1}}^\dagger \left[\Delta x(k) - F_{t_{k-1}}\varepsilon + o\left(\varepsilon^2\right) \right]$$

holds almost surely. Defining $\Delta \widehat{W}_\varepsilon(k)$ as in (18) yields

$$\Delta \widehat{W}_\varepsilon(k) = \Delta W_\varepsilon(k) + L_{t_{k-1}}^\dagger o\left(\varepsilon^2\right).$$

almost surely. Let $\alpha_t \in \mathfrak{L}_I$. Then

$$\sum_k \alpha_{t_{k-1}} \Delta \widehat{W}_\varepsilon(k) = \sum_k \alpha_{t_{k-1}} \left(\Delta W_\varepsilon(k) + L_{t_{k-1}}^\dagger o\left(\varepsilon^2\right) \right).$$

Taking the limit of both sides as ε tends to zero yields $\Delta \widehat{W}_\varepsilon \xrightarrow{\varepsilon} d\mathcal{W}_t$, since for all $\alpha \in \mathfrak{L}_I$

$$\lim_{\varepsilon \to 0} \sum_k \alpha_{t_{k-1}} L_{t_{k-1}}^\dagger o\left(\varepsilon^2\right) = 0 \quad \text{almost surely.} \qquad \blacksquare$$

B. Hybrid Control Law

In this section, we discuss how the sequence of estimates $\{\Delta \widehat{W}_\varepsilon(k)\}_k$ can be used to design a practically implementable hybrid feedback control law that approximates the idealistic input (17), namely $u_t^{\text{lin}} = (-c(\xi_t, z_t) + v)/b(z_t)$, which linearizes the dynamics of the first r components of system (15). We show that the accuracy can be improved by reducing the sampling time ε. We formally state the problem as follows.

Problem 2 (Practical Feedback Linearization): Consider system (15) and let z_t^{lin} be its solution when $u_t = u_t^{\text{lin}}$. The problem of *practical feedback linearization* consists in finding a control law $\hat{u}_t^{\text{lin}}(\varepsilon)$, depending on the sampling rate ε, such that, if z_t is the solution of (15) when $u_t = \hat{u}_t^{\text{lin}}$, then for every $\sigma > 0$

$$\lim_{\varepsilon \to 0} \mathfrak{P}\left(|z_t - z_t^{lin}| \geq \sigma \right) = 0. \tag{19}$$

The meaning of Problem 2 is to find a causal controller for which the closed-loop system dynamics converges in probability

to the idealistic closed-loop dynamics as the sampling time is made smaller. To begin with, since the coefficient c is affine in ξ_t, we can express it as $c(\xi_t, z_t) = c_d(z_t) + c_s(z_t)\xi_t$ for some mappings c_d and c_s. Note that since in practice ξ_t is not known, a naive approximation of ξ_t boils down to replacing it with its expectation, i.e., zero. Therefore, a first causal approximation of the coefficient c is given by only the term $c_d(z_t)$, in turn implying that the control u_t^{lin} can be practically approximated by the naive law

$$u_t^{zn,\text{lin}} = \frac{-c_d(z_t) + v}{b(z_t)}.$$

We call this basic feedback law the *zero-noise control*. By replacing $c_s(z_t)\xi_t$ with zero in the expression of $c(\xi_t, z_t)$, the zero-noise control does not perform any form of stochastic compensation when performing feedback linearization. Consequently, there is no guarantee that the closed-loop behavior of system (15) with $u_t = u_t^{zn,\text{lin}}$ is any close to its idealistic behavior, i.e., when $u_t = u_t^{\text{lin}}$.

The goal of this section is to improve the performance of the zero-noise control $u_t^{zn,\text{lin}}$ by leveraging the estimated sequence $\left\{ \Delta \widehat{W}_\varepsilon(k) \right\}_k$ and show that we can recover the idealistic behavior in probability. Therefore, we define $\hat{u}_t^{\text{lin}} = u_t^{zn,\text{lin}} + u_t^s$, with u_t^s to be specified in such a way that $\hat{u}_t^{\text{lin}}$ is a "better" approximation of u_t^{lin} than $u_t^{zn,\text{lin}}$. By replacing the control $\hat{u}_t^{\text{lin}}$ in (15), the dynamics of the transformed system becomes

$$\begin{aligned} \dot{z}_i &= z_{i+1}, && i = 1, \dots, r-1 \\ \dot{z}_r &= v + c_s(z_t)\xi_t + b(z_t)u_t^s \\ \dot{\eta}_t &= p(\xi_t, \zeta_t, \eta_t) \\ y_t &= z_1 \end{aligned} \tag{20}$$

where the term $c_s(z_t)\xi_t$ in the dynamics of the r-th component is due to the fact that the approximating control $u_t^{zn,\text{lin}}$ cannot cancel the noisy dynamics as the idealistic control u_t^{lin} does.

We now want to design the control u_t^s employing the estimates $\left\{ \Delta \widehat{W}_\varepsilon(k) \right\}_k$ introduced in Section V-A, to reduce the contribution of the term $c_s(z_t)\xi_t$ onto the dynamics of the system. Since the quantity $\Delta \widehat{W}_\varepsilon(k)$ carries information on the evolution of the noise between t_{k-1} and t_k, we look at the evolution of z_r between these two consecutive sampling times. The value of z_r at time t_k is given by

$$z_{r,t_k} = z_{r,t_{k-1}} + \int_{t_{k-1}}^{t_k} v d\tau + \beta_d(k) + \int_{t_{k-1}}^{t_k} b(z_\tau)u_\tau^s d\tau$$

where

$$\beta_d(k) = \int_{t_{k-1}}^{t_k} c_s(z_\tau) d\mathcal{W}_\tau$$

is the contribution of the noise on the dynamics of z_r between the two sampling times. Our goal is to minimise this contribution using u_t^s and the estimate $\Delta \widehat{W}_\varepsilon(k)$ obtained at time t_k. The fact that $\Delta \widehat{W}_\varepsilon(k)$ is only available *a posteriori*, at the end of the sampling period, suggests that u_t^s should induce a jump variation at time t_k in the state z_r in order to compensate for the quantity $\beta_d(k)$. In other words, the dynamics of the closed-loop system should be hybrid. It is then necessary to introduce a

simplified jump notation. At time t_k, we denote by z_{t_k} the state before the jump, and by $z_{t_k^+}$ the state after the jump. The flow dynamics of the closed-loop hybrid system we seek is, therefore, given by

$$
\begin{aligned}
\dot{z}_i &= z_{i+1}, & i = 1, \dots, r-1 \\
\dot{z}_r &= v + c_s(z_t)\xi_t \\
\dot{\eta}_t &= p(\xi_t, \zeta_t, \eta_t)
\end{aligned}
\tag{21}
$$

for all $t \in \mathbb{R}_{\geq 0}$, while the jump dynamics is given by

$$
\begin{aligned}
z_{i,t_k^+} &= z_{i,t_k}, & i = 1, \dots, r-1 \\
z_{r,t_k^+} &= z_{r,t_k} + b(z_{t_k})u^*(k) \\
\eta_{t_k^+} &= \eta_{t_k}
\end{aligned}
\tag{22}
$$

for all $k \in \mathbb{Z}_{>0}$, where $\{u^*(k)\}_k$ is a yet to be defined sequence of scalars depending on $\left\{\Delta\widehat{W}_\varepsilon(k)\right\}_k$. Alternatively, the aforementioned hybrid dynamics (21) and (22) can be equivalently produced by using in (20) an impulsive control u_t^s given by

$$
u_t^s = \sum_{i=1}^{k} u^*(i)\delta(t - t_i), \qquad t \leq t_k
$$

where $\delta(t)$ is a Dirac delta. As a result of the hybrid dynamics induced by the control u_t^s, the expression of z_r before the jump at time t_k is given by

$$
z_{r,t_k} = z_{r,t_{k-1}^+} + \int_{t_{k-1}}^{t_k} v\,d\tau + \beta_d(k)
$$

while after the jump by

$$
\begin{aligned}
z_{r,t_k^+} &= z_{r,t_k} + b(z_{t_k})u^*(k) \\
&= z_{r,t_{k-1}^+} + \int_{t_{k-1}}^{t_k} v\,d\tau + \beta_d(k) + b(z_{t_k})u^*(k).
\end{aligned}
$$

Thus, we have reduced the problem of approximate partial feedback linearization to the problem of finding the sequence $\{u^*(k)\}_k$ such that the contribution of the term $\beta_d(k) + b(z_{t_k})u^*(k)$ is minimized. In particular, we look for a sequence $\{u^*(k)\}_k$, which retrieves the exact linearization of the dynamics of ζ_t as ε tends to zero. We are now ready to solve Problem 2.

Theorem 2: Consider system (15) and let Assumption 2 hold. Let the control $\hat{u}_t^{\text{lin}}$ be given by

$$
\hat{u}_t^{\text{lin}} = \frac{-c_d(z_t) + v}{b(z_t)} - \sum_{i=1}^{k} \frac{c_s(z_{t_{i-1}^+})\Delta\widehat{W}_\varepsilon(i)}{b(z_{t_i})}\delta(t - t_i), t \leq t_k
$$

with $\Delta\widehat{W}_\varepsilon(k)$ given by (18). Then $\hat{u}_t^{\text{lin}}$ solves Problem 2.

Proof: It is trivial to observe that, since system (15) is in normal form, (19) holds if and only if it holds for the r-th component of the state z_t. Thus, we now focus on the r-th component. Under the control $\hat{u}_t^{\text{lin}}$, the jump of z_r at t_k is given by

$$
z_{r,t_k^+} = z_{r,t_k} - c_s(z_{t_{k-1}^+})\Delta\widehat{W}_\varepsilon(k)
$$

thus, in light of the previous discussion

$$
z_{r,t_k^+} = z_{r,t_{k-1}^+} + \int_{t_{k-1}}^{t_k} v\,d\tau + \int_{t_{k-1}}^{t_k} c_s(z_\tau)d\mathcal{W}_\tau
$$

$$
- c_s(z_{t_{k-1}^+})\Delta\widehat{W}_\varepsilon(k).
$$

This can be rewritten as

$$
z_{r,t_k^+} = z_{r,t_0} + \int_{t_0}^{t_k} v\,d\tau + \int_{t_0}^{t_k} c_s(z_\tau)d\mathcal{W}_\tau
$$

$$
- \sum_{i=1}^{k} c_s(z_{t_{i-1}^+})\Delta\widehat{W}_\varepsilon(i).
$$

By Lemma 3

$$
\lim_{\varepsilon\to 0} \sum_{i=1}^{k} c_s(z_{t_{i-1}^+})\Delta\widehat{W}_\varepsilon(i) = \int_{t_0}^{t_k} c_s(z_\tau)d\mathcal{W}_\tau \quad \text{almost surely}
$$

hence

$$
\lim_{\varepsilon\to 0} z_{r,t_k^+} = z_{r,t_0} + \int_{t_0}^{t_k} v\,d\tau = z_{r,t_k}^{\text{lin}} \quad \text{almost surely.} \tag{23}
$$

Now, consider any $t \in (t_{k-1}, t_k)$, for some k, and let $0 < \bar{\varepsilon} < \varepsilon$ be given by $\bar{\varepsilon} = t - t_{k-1}$. Then

$$
z_{r,t} = z_{r,t_{k-1}^+} + \int_{t_{k-1}}^{t} v\,d\tau + \int_{t_{k-1}}^{t} c_s(z_\tau)d\mathcal{W}_\tau
$$

which can be discretized with the Euler–Maruyama method as

$$
z_{r,t} = z_{r,t_{k-1}^+} + v(z_{t_{k-1}^+})\bar{\varepsilon} + c_s(z_{t_{k-1}^+})\Delta W_{\bar{\varepsilon}}(k) + o\left(\bar{\varepsilon}^2\right)
$$

with $\Delta W_{\bar{\varepsilon}}(k) = \mathcal{W}_t - \mathcal{W}_{t_{k-1}}$, and where $o\left(\bar{\varepsilon}^2\right)$ is the one-step truncation error, which is of order $\bar{\varepsilon}^2$. Similarly, the discretized dynamics of $z_{r,t}^{\text{lin}}$ is

$$
z_{r,t}^{\text{lin}} = z_{r,t_{k-1}}^{\text{lin}} + v\left(z_{t_{k-1}}^{\text{lin}}\right)\bar{\varepsilon} + o\left(\bar{\varepsilon}^2\right).
$$

The difference $z_{r,t} - z_{r,t}^{\text{lin}}$ is, therefore,

$$
z_{r,t} - z_{r,t}^{\text{lin}} = z_{r,t_{k-1}^+} - z_{r,t_{k-1}}^{\text{lin}} + \left(v(z_{t_{k-1}^+}) - v\left(z_{t_{k-1}}^{\text{lin}}\right)\right)\bar{\varepsilon}
$$

$$
+ c_s(z_{t_{k-1}^+})\Delta W_{\bar{\varepsilon}}(k) + o\left(\bar{\varepsilon}^2\right)
\tag{24}
$$

Observe that, by the properties of the Brownian motion, $\Delta W_{\bar{\varepsilon}}(k)$ is a normally distributed random variable with zero expectation and variance $\bar{\varepsilon}$. Moreover, c_s is bounded near zero, because it is the Lie derivative of a smooth function. By taking the limit of (24) as ε (hence $\bar{\varepsilon}$) goes to zero and using (23), we have that for every σ

$$
\lim_{\varepsilon\to 0} \mathfrak{P}(\left|z_{r,t} - z_{r,t}^{\text{lin}}\right| \geq \sigma) = 0
$$

hence, the claim follows. ∎

The previous proposition states that, when u_t^s is selected as

$$
u_t^s = -\sum_{i=1}^{k} \frac{c_s(z_{t_{i-1}^+})\Delta\widehat{W}_\varepsilon(i)}{b(z_{t_i})}\delta(t - t_i), \quad t \leq t_k \tag{25}
$$

the control $\hat{u}_t^{\text{lin}}$ approximates the idealistic control u_t^{lin} as ε tends to zero, in the sense that the dynamics of the variable ζ_t can be made approximately linear with an accuracy increasing as ε decreases.

Remark 12: The rationale of Assumption 2 is that the noise is persistently exciting, so it can be estimated by measuring the state of the system. From a technical viewpoint, this assumption

is necessary for the development of the theory because it ensures the boundedness of $L_{t_k}^\dagger$. However, Assumption 2 should not be considered practically restrictive. In fact, if $|L_{t_k}| < \bar{\delta}$, for an arbitrarily small $\bar{\delta}$, this would imply that the noise contribution to the dynamics of the system is sufficiently small at time t_k to be considered negligible. In that case, it is simply possible to avoid performing the stochastic compensation at time t_k and still obtain good control performance.

C. Remarks on the Control Input in the Diffusion: $m \neq 0$

At the beginning of Section IV, we made the standing assumption that $m(x) \equiv 0$ near $\bar{x}$, thus restricting the class of systems considered to those with the control input appearing only in the drift term of the stochastic differential equations. This assumption allowed us to obtain sharper analytical results. In this section, we discuss the rationale of such assumption from a control viewpoint, which has implications both in the idealistic and the practical frameworks.

As shown in (9), for general m, the dynamics of the r-th component of the transformed state $z_t = \Phi(x_t)$ is a quadratic function of the control input. This implies that, in the case that $a(z_t) = (1/2)^m \mathcal{G}_m^l \mathcal{S}_f^{r-1} h\left(\Phi^{-1}(z_t)\right)$ is not zero for x near $\bar{x}$, the idealistic feedback linearizing control would have the form

$$u_t^{\text{lin}} = \frac{-b(\xi_t, z_t) \pm \sqrt{b(\xi_t, z_t)^2 - 4a(z_t)(c(\xi_t, z_t) - v)}}{2a(z_t)}$$

which has real values if and only if the input v is such that $b(\xi_t, z_t)^2 - 4a(z_t)(c(\xi_t, z_t) - v) \geq 0$. The corresponding zero-noise control would have the form

$$u_t^{zn,\text{lin}} = \frac{-b_d(z_t) \pm \sqrt{b_d(z_t)^2 - 4a(z_t)(c_d(z_t) - v)}}{2a(z_t)}$$

where the input v should enforce $b_d(z_t)^2 - 4a(z_t)(c_d(z_t) - v) \geq 0$. It is evident that, although in both cases the control input can be forced to be real by an appropriate choice of v, from a practical viewpoint such a choice may not leave any space for a control law achieving objectives, such as, for instance, output tracking. Moreover, even though we may be able to define a zero-noise control, the construction of a hybrid controller presents even more challenges, as the quadratic u_t^{lin} is not affine in ξ_t. Of course, one might make the standing assumption that $a(z_t) \equiv 0$ for x near $\bar{x}$. This class of systems would include the systems for which $m \equiv 0$, which we address in detail in this article, but also those systems with $m \neq 0$ with the additional assumption that $^m \mathcal{G}_m^l \mathcal{S}_f^{r-1} h\left(\Phi^{-1}(z_t)\right) = 0$. However, the design of controllers poses substantial technical challenges even in this case. For the sake of completeness, we briefly give an overview of these issues. Suppose $m \neq 0$ and $a(z_t) \equiv 0$ for x near $\bar{x}$. The coefficient b in (9) is, therefore, affine in ξ_t, i.e., $b(\xi_t, z_t) = b_d(z_t) + b_s(z_t)\xi_t$. In the idealistic framework, a feedback linearizing control would require a division by $b(\xi_t, z_t)$. Unless further, possibly restrictive, assumptions are made, such a division may result in a nonadmissible control. Additionally, the fact that the coefficient b depends on the noise affects impulsive compensations as well. In fact, the dynamics of z_r when the control $\hat{u}_t^{\text{lin}}$ is applied is

$$\dot{z}_r = v + c_s(z_t)\xi_t + b_d(z_t)u_t^s + b_s(z_t)u_t^s\xi_t.$$

Thus, any compensating control u_t^s inevitably introduces noise at time t_k^+, which cannot be compensated for because ξ_{t_k} is a

random variable that is independent of the process ξ_t, $t < t_k$. This makes it impossible to conclude on the convergence in probability of z_{r,t_k^+} to z_{r,t_k}^{lin} as ε approaches zero, because the random variable ξ_{t_k} can take arbitrarily large values in $\mathbb{R}$ with nonzero probability.

VI. ASYMPTOTIC OUTPUT TRACKING

In this section, we first design an idealistic controller, i.e., using the white noise in the feedback loop, to make the output asymptotically track a reference signal. Then, we show that a practical feedback control law leveraging the causal stochastic compensations introduced in Section V is able to retrieve the idealistic result, in the limit of the sampling time ε going to zero.

We start by introducing the idealistic control and we show that, under suitable stability hypotheses on the zero dynamics, it is possible to control the system so that its output tracks reference trajectories while its internal variables remain bounded almost surely. First observe that, as long as $z_t = \Phi(x_t)$ is chosen such that $q \equiv 0$ near zero (which is always possible, see Section III-A), the zero dynamics of system (15) is affine in ξ_t. Let it be expressed as

$$\dot{\eta}_t = p(\xi_t, 0, \eta_t) = p_d(0, \eta_t) + p_s(0, \eta_t)\xi_t.$$

Consider a reference signal y_R, which is continuously differentiable r times with values in a neighborhood of zero. We assume that the initial state of the transformed system (15) is arbitrary while in a neighborhood of zero and we seek a feedback control u_t that makes the output y_t of the system asymptotically converge to y_R. Let

$$v(\zeta_t, y_R(t)) = y_R^{(r)} - \sum_{i=1}^r d_{i-1}\left(\zeta_i - y_R^{(i-1)}\right)$$

with $d_i \in \mathbb{R}$ for $i = 0, \ldots, r-1$ to be determined, and the admissible idealistic feedback control law be given by

$$u_t^{\text{track}} = -\frac{c(\xi_t, \zeta_t, \eta_t) - v(\zeta_t, y^R(t))}{b(\zeta_t, \eta_t)}.$$

Define the tracking error $e_t := y_t - y_R(t)$. Then, the control u_t^{track} forces the dynamics of the tracking error to be $e_t^{(r)} + d_{r-1}e_t^{(r-1)} + \cdots + d_1 e_t^{(1)} + d_0 e_t$, which can be made exponentially stable by selecting the coefficients d_i such that

$$\Lambda(s) = s^r + d_{r-1}s^{r-1} + \cdots + d_1 s + d_0, \qquad (26)$$

which is the characteristic polynomial of the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -d_0 & -d_1 & -d_2 & \cdots & -d_{r-1} \end{bmatrix}, \qquad (27)$$

has roots with negative real parts. We also study the boundedness of ζ_t and of the internal variable η_t under the control u_t^{track}, when y_R and its first $r-1$ derivatives are bounded. Define $\zeta_R(t) = \begin{bmatrix} y_R(t) & \cdots & y_R^{(r-1)}(t) \end{bmatrix}^\top$ and $\theta_t = \begin{bmatrix} e_t & \cdots & e_t^{(r-1)} \end{bmatrix}^\top$. Then, the following result, the proof of

which relies on the definition of strict Lyapunov function and some technical lemmas presented in Appendix B, holds.

Theorem 3: Consider system (15). Suppose $y_R(t)$, $y_R^{(1)}(t), \ldots, y_R^{(r-1)}(t)$ are bounded. Let $\eta_{R,t}$ be the solution of

$$\dot{\eta}_{R,t} = p(\xi_t, \zeta_R(t), \eta_{R,t}), \quad \eta_{R,0} = 0 \qquad (28)$$

and let p_d and p_s be Lipschitz continuous. Moreover, assume that there exists a strict Lyapunov function $V(\eta_R, t)$ for (28) such that $\frac{\partial V}{\partial \eta_{R,i}}(x, t)$ and $\frac{\partial^2 V}{\partial \eta_{R,i} \partial \eta_{R,j}}(x, t)$ are bounded for all x in a neighborhood of the origin and $t \geq 0$. Suppose that the roots of the polynomial $\Lambda(s)$ in (26) have negative real part. Then, for sufficiently small $\epsilon_R > 0$, if

$$\left| z_i(\bar{t}) - y_R^{(i-1)}(\bar{t}) \right| < \epsilon_R, \quad 1 \leq i \leq r, \qquad \|\eta_{\bar{t}} - \eta_{R,\bar{t}}\| < \epsilon_R,$$

then for all $\epsilon > 0$ there exists $\delta > 0$ such that

$$\left| z_i(\bar{t}) - y_R^{(i-1)}(\bar{t}) \right| < \delta \to \left| z_i(t) - y_R^{(i-1)}(t) \right| < \epsilon$$

$$1 \leq i \leq r, \text{ for all } t \geq \bar{t} \geq 0$$

$$\|\eta_{\bar{t}} - \eta_{R,\bar{t}}\| < \delta \to \|\eta_t - \eta_{R,t}\| < \epsilon \text{ for all} t \geq \bar{t} \geq 0$$

almost surely, i.e., the response z_i and η_t, $t \geq \bar{t} \geq 0$, of system (15) under the control law u_t^{track} is bounded almost surely.

Proof: System (15) under the control law u_t^{track} can be rewritten in the form

$$\dot{\theta}_t = A\theta_t, \quad \dot{\eta}_t = p(\xi_t, \zeta_R(t) + \theta_t, \eta_t) \qquad (29)$$

with A given by (27), which has characteristic polynomial $\Lambda(s)$. Therefore, θ_t has asymptotically stable dynamics. Let $\nu_t = \eta_t - \eta_{R,t}$ and $P(\xi_t, \nu_t, \theta_t, t) = p(\xi_t, \zeta_R(t) + \theta_t, \eta_{R,t} + \nu_t) - p(\xi_t, \zeta_R(t), \eta_{R,t})$. Note that the system

$$\dot{\nu}_t = P(\xi_t, \nu_t, \theta_t, t), \quad \dot{\theta}_t = A\theta_t \qquad (30)$$

is in the form (36) (see Appendix B). Moreover, system (30) additionally satisfies the hypotheses of Lemma 8 in Appendix B because of the assumption of Lipschitz continuity of p_d and p_s and of the existence of V as in the statement. Then, by Lemma 8, $(0, \eta_{R,t})$ is an almost surely uniformly stable solution of (29) and the claimed estimates follow.

The previous theorem solves the idealistic local asymptotic output tracking problem, i.e., the output $y_t = z_1$ asymptotically converges to y_R whilst the state z_t remains bounded almost surely.

We now focus on practical output tracking. While it can be proved (see Proposition 4 in Appendix C) that under some technical assumptions a zero-noise control is sufficient to asymptotically stabilize the equilibrium at the origin, a controller not performing any sort of compensation for the stochastic disturbances does not guarantee asymptotic tracking. This is because if the system tracks a nonzero reference, then the states will not converge to zero and so the noise will enter the dynamics in a persistent fashion. Hence, the noise might drive the states away from the desired trajectory, possibly inducing instability. On the contrary, the control law $\hat{u}_t^{\text{track}} = u_t^{zn,\text{track}} + u_t^s$, with

$$u_t^{zn,\text{track}} = -\frac{c_d(\zeta_t, \eta_t) - v(\zeta_t, y^R(t))}{b(\zeta_t, \eta_t)}$$

and u_t^s given by (25) is able to prevent this when ε tends to zero by approximately compensating for the Brownian-induced disturbances, as shown in the following result.

Corollary 1: Consider system (15) and suppose that Assumption 2 and the assumptions in Theorem 3 hold. Then, the control law $\hat{u}_t^{\text{track}}$ is such that y_t converges to y_R in probability and the state z_t, $t \geq \bar{t} \geq 0$, of system (15) is bounded in probability in the limit as ε approaches zero.

Proof: By Theorem 3 the control u_t^{track} makes the output y_t of (15) asymptotically converge to y_R while keeping the internal states bounded. Let z_t^{track} and z_t be the state of system (15) when $u_t = u_t^{\text{track}}$ and $u_t = \hat{u}_t^{\text{track}}$ are applied, respectively. Then, by Theorem 2, for every $\sigma > 0$

$$\lim_{\varepsilon \to 0} \mathfrak{P} \left(\left| z_t - z_t^{\text{track}} \right| \geq \sigma \right) = 0$$

and the claim follows.   ∎

VII. ILLUSTRATIVE EXAMPLE

In this section, we illustrate the validity of the theory by means of a numerical example. Consider the following nonlinear stochastic system in the form (11) with

$$f(x_t) = \begin{bmatrix} s_2(1 + x_1) \\ -2 \tan x_2 \\ f_3(x_t) \end{bmatrix}, \ g(x_t) = \begin{bmatrix} e^{x_3} \\ 0 \\ e^{x_3} \end{bmatrix}, \ l(x_t) = \begin{bmatrix} x_1 \\ \frac{-2x_1}{c_2} \\ x_1^2 \end{bmatrix}$$

with $f_3(x_t) = 2x_3 + x_1 s_2 - \frac{2 s_2 x_1^2}{c_2^2}$ and s_i and c_i denoting $\sin x_i$ and $\cos x_i$, respectively. Let the output of the system be $y_t = x_1 + s_2 - x_3$. We are interested in analyzing the system around the origin, i.e., we set $\bar{x} = 0$. The goal is to bring the system to its normal form and to perform asymptotic output tracking. The first step is determining the stochastic relative degree of the system at zero. We set $z_1 = y_t = x_1 + s_2 - x_3$ and we compute its derivative applying Itô's formula (we omit the procedure for brevity), thus obtaining $dz_1 = (-2x_3 - s_2)dt$. As neither the input u nor the noise appear in this expression, the stochastic relative degree, if defined, is higher than one at the origin. We set $z_2 = -2x_3 - s_2$ and, by computing its derivative, we obtain

$$dz_2 = \left(2s_2 - 4x_3 - 2x_1 s_2 + \frac{6x_1^2 s_2}{c_2^2} - 2e^{x_3} u \right) dt + 4x_1 d\mathcal{W}_t.$$

The system has, therefore, stochastic relative degree $r = 2$ at the origin. Moreover

$$\tilde{c}_d(x_t) = c_d(\Phi(x_t)) = 2s_2 - 4x_3 - 2x_1 s_2 + \frac{6x_1^2 s_2}{c_2^2}$$

$$\tilde{c}_s(x_t) = c_s(\Phi(x_t)) = -2e^{x_3}$$

$$\tilde{b}(x_t) = b(\Phi(x_t)) = 4x_1.$$

Setting $z_3 = x_1 - x_3$ makes the coordinate change $z = \Phi(x)$ a diffeomorphism in a neighborhood of the origin, with

$$\Phi(x) = \begin{bmatrix} x_1 + s_2 - x_3 \\ -s_2 - 2x_3 \\ x_1 - x_3 \end{bmatrix}.$$

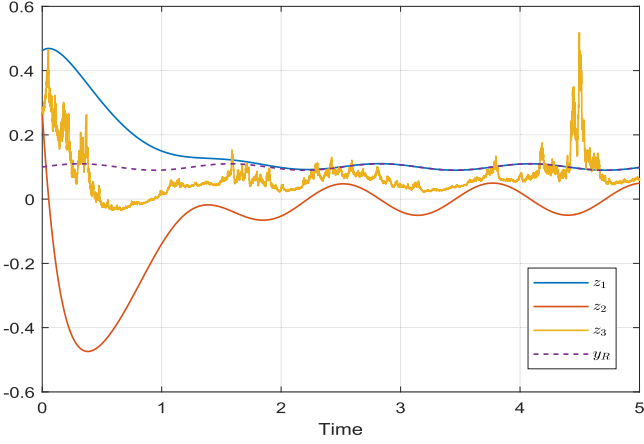In this new set of coordinates the dynamics of the system is given by

$$\dot{z}_1 = z_2$$

Fig. 1. Time history of the state z_t under the control u_t^{track}, achieving idealistic output tracking.

$$\dot{z}_2 = c_d(z_t) + c_s(z_t)\xi_t + b(z_t)u$$

$$\dot{z}_3 = s_2 - 2x_3 + \frac{2x_1^2 s_2}{c_2^2} + 2x_1\xi_t = \tilde{p}(\xi_t, \Phi(x_t)) \qquad (31)$$

which is in the stochastic normal form. The zero dynamics of the system is obtained by equating $z_1 = 0$ and $z_2 = 0$, $z_3 = \eta_t$, which yields $x_1 = (3/2)\eta_t$, $s_2 = -\eta_t$. Replacing these in the third equation in (31) we get the zero dynamics as follows:

$$\dot{\eta}_t = p(\xi_t, 0, \eta_t) = -2\eta_t + \frac{9\eta_t^3}{2(\eta_t^2 - 1)} + 3\eta_t\xi_t.$$

Its first approximation around the origin is $\dot{\eta}_t = -2\eta_t + 3\eta_t\xi_t = A\eta_t + F\eta_t\xi_t$, which is asymptotically stable almost surely because $A - F^2/2 < 0$. Therefore, the zero dynamics of the system is almost surely asymptotically stable. We now choose a reference signal of the form $y_R(t) = \beta + \alpha\cos(\omega t)$ and illustrate that u_t^{track} achieves asymptotic output tracking. To this end, we first performed a simulation in the idealistic scenario in which the noise is used in the feedback law. We selected the coefficients $d_0 = 12$ and $d_1 = 7$ in the input v so the characteristic polynomial $\Lambda(s)$ in (26) has roots in -3 and -4, thus guaranteeing asymptotic stability of the linearized subsystem in the coordinates ζ_t. The reference signal y_R is characterised by $\beta = 0.1$, $\alpha = 0.01$ and $\omega = 5$. The nonlinear stochastic differential equations were integrated using the Euler–Maruyama scheme with period $\Delta t = 10^{-6}$. In Fig. 1, we show the time history of the state in the coordinates z_t when u_t^{track} is applied. Observe that the first two components (blue and red lines) display, as expected, linear and deterministic behaviors, while the internal variable z_3 has noisy dynamics. Because of the properties of the zero dynamics, z_3 stays bounded under u_t^{track}. Moreover, the component z_1 (blue line), which by the definition of normal form is the output y_t of the system, asymptotically converges to the reference signal y_R (purple/dashed line).

We now turn our attention to practically realisable controllers. To this end, we first discuss the numerical implementation of the simulations in the cases where stochastic compensations are performed with a period ε. The continuous-time dynamics was integrated, as usual, using the Euler–Maruyama numerical
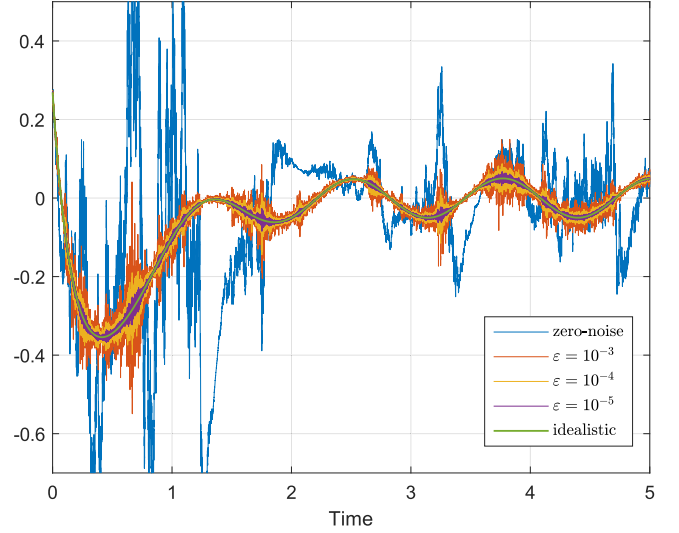

Fig. 2. Time history of the component z_2 under the controls $u_t^{zn,\text{track}}$ (blue line), $\hat{u}_t^{\text{track}}$ with $\varepsilon = 10^{-3}$ (red line), $\varepsilon = 10^{-4}$ (yellow line), $\varepsilon = 10^{-5}$ (purple line), and u_t^{track}.
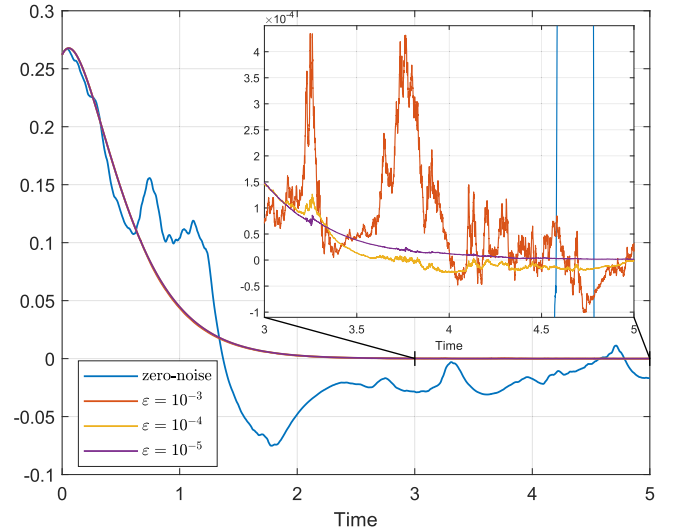

Fig. 3. Time history of the tracking error $y_t - y_R$ under the controls $u_t^{zn,\text{track}}$ (blue line) and $\hat{u}_t^{\text{track}}$ with $\varepsilon = 10^{-3}$ (red line), $\varepsilon = 10^{-4}$ (yellow line), and $\varepsilon = 10^{-5}$ (purple line).

scheme with fixed period $\Delta t = 10^{-6}$. The stochastic compensations were performed with a period ε, which must necessarily satisfy $\varepsilon > \Delta t$. We performed simulations with values of ε of 10^{-3}, 10^{-4}, and 10^{-5}, in order to illustrate the limit behavior of the solutions as ε decreases. Note that it could be possible to select smaller ε as long as Δt is decreased accordingly.

We illustrate in Fig. 2 that when the control law with compensations is employed and ε is decreased, the trajectory of the state under the idealistic control is retrieved. In Fig. 2, we consider again the output tracking setting and display the time history of the component $z_r = z_2$ under the controls $u_t^{zn,\text{track}}$ (blue line), $\hat{u}_t^{\text{track}}$, with varying ε (10^{-3} (red line), 10^{-4} (yellow line), 10^{-5} (purple line)), and u_t^{track} (green line). Observe that, since the

control $\hat{u}_t^{\text{track}}$ improves the noise compensations as ε decreases, the trajectories of z_2 under the compensated control tend to the idealistic trajectory.

To conclude, Fig. 3 shows a comparison between the time histories of the tracking error $z_1 - y_R = y_t - y_R$ when the controls applied are, respectively, $u_t^{zn,\text{track}}$ (blue line) or $\hat{u}_t^{\text{track}}$ with decreasing values of ε, namely 10^{-3} (red line), 10^{-4} (yellow line), 10^{-5} (purple line). While the hybrid controller produces, for any of the values of ε, strikingly better asymptotic performances than the zero-noise control, observe that the tracking error is made smaller and smaller as ε is decreased. This is in line with the fact that the component $z_1 = y_t$ under $\hat{u}_t^{\text{track}}$ tends, as ε approaches zero, to z_1 under the control u_t^{track}, which in turn asymptotically approaches y_R.

VIII. CONCLUSION

In this article, we have addressed the path-wise control of nonlinear stochastic systems. First, we have introduced a notion of stochastic relative degree and normal form. Then, leveraging these, we have presented a feedback linearizing controller. We have observed that this controller is not causal, hence, referred to as idealistic, because it requires a feedback of the noise. To overcome this limitation, we have introduced a hybrid control architecture that estimates the Brownian motion from measurements of the state and uses these estimates to periodically compensate, in approximate way, for the noise. We have proved that the performances of the idealistic control law are retrieved when compensations at a frequency tending to infinity are performed. Finally, we have solved the problem of asymptotic output tracking, both in the idealistic and practical framework, and we have provided an illustrative example.

APPENDIX

A. Technical Lemmas and Proof of Proposition 1

Lemma 4: Let $x \in U \subset \mathbb{R}^n$ and $k \in \{0, \ldots, r-2\}$. Then, ${}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^k h(\xi_t, x) = 0$ for all $\xi_t \in \mathbb{R}$ if and only if $\mathcal{L}_g \, {}^l\mathcal{S}_f^k h(x) + {}^m\mathcal{G}_l \, {}^l\mathcal{S}_f^k h(x) = 0$ and $\mathcal{L}_m \, {}^l\mathcal{S}_f^k h(x) = 0$.

Proof: Without loss of generality assume $k = 0$. Then, the expression of ${}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^k h(\xi_t, x)$ for $k = 0$ is given in (3). The sufficiency is trivial. As for the necessity, observe that if $\mathcal{L}_m h(x) \neq 0$ then the randomness induced by the white noise implies that ${}^m_g \mathcal{A}_l h(\xi_t, x) \neq 0$ almost surely. Therefore, $\mathcal{L}_m h(x) = 0$ is a necessary condition for ${}^m_g \mathcal{A}_l h(\xi_t, x)$ to be zero for all $\xi_t \in \mathbb{R}$. As a consequence, also $\mathcal{L}_g h(x) + {}^m\mathcal{G}_l h(x) = 0$ is a necessary condition for ${}^m_g \mathcal{A}_l h(\xi_t, x)$ to be zero for all $\xi_t \in \mathbb{R}$. ∎

Lemma 5: Let $\bar{x} \in \mathbb{R}^n$. Then, ${}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^{r-1} h(\xi_t, \bar{x}) \neq 0$ almost surely if and only if $\mathcal{L}_g \, {}^l\mathcal{S}_f^{r-1} h(\xi_t, \bar{x}) + {}^m\mathcal{G}_l \, {}^l\mathcal{S}_f^{r-1} h(\xi_t, \bar{x}) \neq 0$ or $\mathcal{L}_m \, {}^l\mathcal{S}_f^{r-1} h(\xi_t, \bar{x}) \neq 0$.

Proof: Observe that ${}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^{r-1} h(\xi_t, \bar{x}) = \mathcal{L}_g \, {}^l\mathcal{S}_f^{r-1} h(\bar{x}) + \mathcal{L}_m \, {}^l\mathcal{S}_f^{r-1} h(\bar{x})\xi_t + {}^m\mathcal{G}_l \, {}^l\mathcal{S}_f^{r-1} h(\bar{x})$. Then, the necessity is trivial. As for the sufficiency, observe that if $\mathcal{L}_g \, {}^l\mathcal{S}_f^{r-1} h(\bar{x}) + {}^m\mathcal{G}_l \, {}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq 0$ or $\mathcal{L}_m \, {}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq 0$ then the randomness induced by the white noise implies that $\mathcal{L}_g \, {}^l\mathcal{S}_f^{r-1} h(\bar{x}) +$

${}^m\mathcal{G}_l \, {}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq -\mathcal{L}_m \, {}^l\mathcal{S}_f^{r-1} h(\bar{x})\xi_t$ almost surely, hence the claim follows. ∎

Proof of Proposition 1: We prove the first part of the proposition by induction. Equation (5) trivially holds for $k = 0$. Now, suppose that it holds for any $k \in \{1, \ldots, r-2\}$. Then

$$y_t^{(k+1)} = {}^l\mathcal{S}_f^{k+1} h(\xi_t, x_t) + {}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^k h(\xi_t, x_t)u$$
$$+ \frac{1}{2} {}^m\mathcal{G}_m \, {}^l\mathcal{S}_f^k h(\xi_t, x_t)u^2. \tag{32}$$

The first term on the right-hand side expands to

$${}^l\mathcal{S}_f^{k+1} h(\xi_t, x_t) = \mathcal{L}_f \, {}^l\mathcal{S}_f^k h(x_t) + \mathcal{L}_l \, {}^l\mathcal{S}_f^k h(x_t)\xi_t$$
$$+ \frac{1}{2} {}^l\mathcal{G}_l \, {}^l\mathcal{S}_f^k h(x_t)$$

which, by the first equality in condition (ND), reduces to

$${}^l\mathcal{S}_f^{k+1} h(x_t) = \mathcal{L}_f \, {}^l\mathcal{S}_f^k h(x_t) + \frac{1}{2} {}^l\mathcal{G}_l \, {}^l\mathcal{S}_f^k h(x_t)$$

i.e., ${}^l\mathcal{S}_f^{k+1} h(x_t)$ is independent of ξ_t for all $k \in \{0, \ldots, r-2\}$. Going back to (32), by Lemma 4, the first two equalities in condition (CD) are equivalent to ${}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^k h(\xi_t, x_t) = 0$. Moreover, by the last equality in condition (CD), ${}^m\mathcal{G}_m \, {}^l\mathcal{S}_f^k h(\xi_t, x_t) = 0$. In conclusion, $y_t^{(k+1)} = {}^l\mathcal{S}_f^{k+1} h(x_t)$ for all $k \in \{0, \ldots, r-2\}$, which proves that conditions (ND) and (RD) imply (5).

To prove (6), observe that by Lemma 5, the first two inequalities in condition (RD) are equivalent to ${}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^{r-1} h(\xi_t, \bar{x}) \neq 0$. Therefore, condition (RD) implies that either ${}^m_g \mathcal{A}_l \, {}^l\mathcal{S}_f^{r-1} h(\xi_t, \bar{x}) \neq 0$ or ${}^m\mathcal{G}_m \, {}^l\mathcal{S}_f^{r-1} h(\bar{x}) \neq 0$, thus making the control $u(\bar{t})$ appear in the expression of the r-th derivative of y_t. ∎

B. Almost Sure Stability of Perturbed Stochastic Systems

Consider the stochastic time-varying system

$$\dot{x}_t = f_d(x_t, t) + f_s(x_t, t)\xi_t = \tilde{f}(\xi_t, x_t, t) \tag{33}$$

for which we introduce the following Lipschitz assumption [35].

Assumption 3: f_d and f_s are locally Lipschitz continuous for all $t \geq 0$.

We now introduce a concept of almost sure total stability for nonlinear time-varying stochastic system. Consider the following extension of the definition of strict Lyapunov function for stationary stochastic systems given in [36] to the case of time-varying stochastic systems.

Definition 8: Consider the autonomous system (33). A smooth function $V : U \times \mathbb{R}_{\geq 0} \to \mathbb{R}$, where U is a bounded domain, is said to be a *strict Lyapunov function* for system (33) if

1) $W_1(x) \leq V(x, t) \leq W_2(x)$ for all $x \in U$, $t \geq 0$, $V(0, \cdot) = 0$ and W_1 and W_2 are continuous positive definite functions on U;
2) there exists a positive definite function $W_3(x)$ such that $\frac{\partial V}{\partial t}(x, t) + \mathcal{L}_{f_d} V(x, t) + \frac{1}{2} {}^{f_s}\mathcal{G}_{f_s} V(x, t) < -W_3(x)$ and $\mathcal{L}_{f_s} V(x, t) = 0$ for all $x \in U$, $t \geq 0$.

It is possible to show that the existence of a strict Lyapunov function for system (2) is sufficient to conclude the almost sure (uniform) asymptotic stability of its equilibrium point (as defined, e.g., in [37] and [38]). Namely, the following result is an extension of [36, Th. 2.5].

Lemma 6: Under Assumption 3, if there exists a strict Lyapunov function for system (33), then the equilibrium point at the origin of system (33) is almost surely uniformly asymptotically stable.

The proof follows by replacing in the proof of [36, Th. 2.6] the time-invariant infinitesimal generator of V with its time-varying version.

Remark 13: By dropping the condition that $\mathcal{L}_{f_s}V(x,t) = 0$ in Definition 8, the results can be reformulated in the context of the weaker stability in probability. See [27, Th. 11.2.8] for an equivalent of Lemma 6 in this context.

Consider now a persistently perturbed version of system (33), namely

$$dx_t = (f_d(x_t,t) + p_d(x_t,t))dt + (f_s(x_t,t) + p_s(x_t,t))d\mathcal{W}_t \tag{34}$$

where p_d and p_s are Lipschitz. In the following, we formalize the idea that if p_d and p_s are small enough perturbations, then the stability properties of (33) are analogous to those of (34). In other words, we provide an extension to stochastic systems of the concept of total stability presented, e.g., in [39, Def. 56.1]. To this end, let $\tilde{p}(\xi_t,x_t,t) = p_d(x_t,t) + p_s(x_t,t)\xi_t$.

Definition 9: Consider the stochastic systems (33) and the perturbed system

$$\dot{x}_t = \tilde{f}(\xi_t,x_t,t) + \tilde{p}(\xi_t,x_t,t)$$
$$= f_d(x_t,t) + p_d(x_t,t) + (f_s(x_t,t) + p_s(x_t,t))\xi_t. \tag{35}$$

Suppose that there exists a solution $x_{p,t}(\bar{t},\bar{x})$ of (35). The equilibrium $x = 0$ of (33) is said to be *totally stable almost surely* if for each $\epsilon > 0$ there exist positive $\delta_1(\epsilon), \delta_2(\epsilon)$ and $\delta_3(\epsilon)$ such that $\|x_{p,t}(\bar{t},\bar{x})\| < \epsilon$ for all $t \geq \bar{t}$ almost surely provided that $\|\bar{x}\| < \delta_1, \|p_d(x,t)\| < \delta_2$ and $\|p_s(x,t)\| < \delta_3$ for all (x,t) satisfying $t \geq \bar{t}$ and $\|x\| \leq \epsilon$.

The following lemma provides some properties that a strict Lyapunov function must satisfy in order for system (33) to have an almost surely totally stable equilibrium at zero.

Lemma 7: Consider systems (33) and (35). Suppose that Assumption 3 holds. If there exists a strict Lyapunov function V for system (33) such that
1) $\frac{\partial V}{\partial x_i}(x,t)$ and $\frac{\partial^2 V}{\partial x_i \partial x_j}(x,t)$ are bounded for all $x \in U$ and $t \geq 0$,
2) $\mathcal{L}_{p_s}V(x,t) = 0$ for all $x \in U$ and $t \geq 0$,

then system (33) has an almost surely totally stable equilibrium at $x = 0$.

Proof: The proof follows from [39, Th. 56.3] by noticing that the time derivatives of V for (33) and (35) differ by the term

$$\frac{\partial V}{\partial x}p_d + (f_s + p_s)^\top \frac{\partial^2 V}{\partial x^2}(f_s + p_s)$$

which, by Assumption 3 and by the hypotheses of this lemma, can be made so small that the time derivative of V for (35) is

negative. The rest of the proof is analogous to that of [39, Th. 56.3]. ∎

Finally, the following lemma gives important stability properties of nonlinear time-varying stochastic systems driven by almost surely stable systems.

Lemma 8: Consider the stochastic system

$$\dot{y}_t = \tilde{a}(\xi_t,y_t,z_t,t) = \tilde{a}_d(y_t,z_t,t) + \tilde{a}_s(y_t,z_t,t)\xi_t$$
$$\dot{z}_t = \tilde{b}(\xi_t,z_t) = \tilde{b}_d(z_t) + \tilde{b}_s(z_t)\xi_t \tag{36}$$

and assume the following.
A1) $(y,z) = (0,0)$ is an equilibrium of (36) and $\tilde{a}_d$ and $\tilde{a}_s$ are Lipschitz continuous for all $t \geq 0$ in a neighborhood of zero.
A2) There exists a strict Lyapunov function $V(x,t)$ for the first equation of system (36) such that $\frac{\partial V}{\partial x_i}(x,t)$ and $\frac{\partial^2 V}{\partial x_i \partial x_j}(x,t)$ are bounded for all $x \in U$ and $t \geq 0$.
A3) The equilibrium $z = 0$ of $\dot{z}_t = \tilde{b}(\xi_t,z_t)$ is stable almost surely.

Then the equilibrium $(y,z) = (0,0)$ of (36) is (uniformly) stable almost surely.

Proof: Set $\tilde{f}(\xi_t,y_t,t) = \tilde{a}(\xi_t,y_t,0,t)$ and $\tilde{p}(\xi_t,y_t,t) = \tilde{a}(\xi_t,y_t,z_t(\bar{t},\bar{z}),t) - \tilde{a}(\xi_t,y_t,0,t)$, with $z_t(\bar{t},\bar{z})$ solution of $\dot{z}_t = \tilde{b}(\xi_t,z_t)$ satisfying $z_{\bar{t}} = \bar{z}$. Then, the first equation in (36) has the form (35). If $\|z_t\| < \epsilon_z$ almost surely for all $t > \bar{t}$, then, by Assumption A1 there exist $\mu_1 > 0$ and $\mu_2 > 0$ such that

$$\|p_d(y,t)\| = \|a_d(y,z_t(\bar{t},\bar{z}),t) - a_d(y,0,t)\| < \mu_1 \epsilon_z$$
$$\|p_s(y,t)\| = \|a_s(y,z_t(\bar{t},\bar{z}),t) - a_s(y,0,t)\| < \mu_2 \epsilon_z$$

almost surely for all y in a neighborhood of zero and all $t \geq \bar{t}$. Observe that by Assumption A2, the first and second conditions of Lemma 7 are satisfied for the first equation in (36). Therefore, the first equation in (36) is totally stable at $y = 0$. By Definition 9, for all ϵ_y, there exist $\delta_1 > 0$, $\delta_2 > 0$, and $\delta_3 > 0$ such that $\|\bar{y}\| < \delta_1$, $\|p_d(y,t)\| < \delta_2$, and $\|p_s(y,t)\| < \delta_3$, for all (y,t) such that $\|y\| < \epsilon_y$ and $t \geq \bar{t} \geq 0$, imply

$$\|y_t(\bar{t},\bar{y})\| < \epsilon_y \quad \text{almost surely for all } t \geq \bar{t}.$$

Finally, by Assumption A3, it is possible to find a δ_z such that $\|\bar{z}\| < \delta_z$ yields $\|z_t\| < \bar{\delta} = \min\{\delta_2,\delta_3\}/\max\{\mu_1,\mu_2\}$ for all $t \geq \bar{t}$, which, for what stated earlier, makes $\|p_d(y,t)\| < \mu_1\bar{\delta} \leq \delta_2$ and $\|p_s(y,t)\| < \mu_2\bar{\delta} \leq \delta_3$ be satisfied almost surely. Therefore, for a bounded initial condition $(\bar{y},\bar{z})$ in a neighborhood of the equilibrium $(0,0)$, the trajectory (y_t,z_t) is bounded almost surely in a neighborhood of $(0,0)$, which concludes the proof. ∎

C. Almost Sure Stability Under Zero-Noise Control

Lemma 9: Consider the system

$$\dot{y}_t = \hat{A}y_t + \gamma(\xi_t,y_t,z_t) \quad \dot{z}_t = \hat{f}(\xi_t,y_t,z_t) \tag{37}$$

with γ and $\hat{f}$ affine in ξ_t and suppose that $\gamma(\xi_t,0,z) = 0$ for z in a neighborhood of zero and that $\frac{\partial \gamma}{\partial y}(\xi_t,0,0) = 0$. If $\dot{z}_t = f(\xi_t,0,z_t)$ has an almost surely asymptotically stable equilibrium at $z = 0$ and the eigenvalues of A all have negative real part, then system (37) has an almost surely asymptotically stable equilibrium at $(y,z) = (0,0)$.

Proof: Consider the stochastic subsystem $\dot{z}_t = f(\xi_t, 0, z_t)$. Since this has an almost surely asymptotically stable equilibrium at 0 by assumption, by [40], it is possible to decompose the system in the Oseledec subspaces E_c and E_s relative to the zero and negative Lyapunov exponents of the linearized dynamics, respectively. Let $z_t^c \in E_c$ be the projection of z_t onto E_c and $z_t^s \in E_s$ the projection of z_t onto E_s; then for some F_c, F_s, g_c, and g_s

$$\dot{z}_t^c = F_c(\xi_t)z_t^c + g_c(\xi_t, 0, z_t^c, z_t^s)$$
$$\dot{z}_t^s = F_s(\xi_t)z_t^s + g_s(\xi_t, 0, z_t^c, z_t^s) \tag{38}$$

where the Lyapunov exponents associated to F_c are zero and those associated to F_s are negative. Let $z_t^s = \pi_1(\xi_t, z_t^c)$ be a stochastic center manifold (see [40]) at zero for (38). Then, by the reduction principle (see [40, Th. 7.3]), since the subsystem $\dot{z}_t = f(\xi_t, 0, z_t)$ has an almost surely asymptotically stable equilibrium at $z = 0$ by assumption, then the equilibrium at the origin of

$$\dot{z}_t^c = F_c(\xi_t)z_t^c + g_c(\xi_t, 0, z_t^c, \pi_1(\xi_t, z_t^c)) \tag{39}$$

is asymptotically stable almost surely as well. Consider now the full system (37), for which a center manifold at zero is described by the pair

$$z_t^s = \chi_1(\xi_t, z_t^c), \quad y_t = \chi_2(\xi_t, z_t^c).$$

By the reduction principle, the dynamics (37) has an almost surely asymptotically stable equilibrium at (0,0) if the reduced system

$$\dot{z}_t^c = F_c(\xi_t)z_t^c + g_c(\xi_t, \chi_2(\xi_t, z_t^c), z_t^c, \chi_1(\xi_t, z_t^c)) \tag{40}$$

has an almost surely asymptotically stable equilibrium at 0. It is easy to see that $\chi_2(\xi_t, z_t^c) = 0$, because, by the properties of γ, the first approximation of the dynamics of y_t reduces to $\dot{y}_t = Ay_t$, which is asymptotically stable by assumption. Then, since $\chi_1(\xi_t, z_t^c) = \pi_1(\xi_t, z_t^c)$, (40) reduces to (39), which has been proved to have an almost surely asymptotically stable equilibrium at zero. Hence, system (37) has an almost surely asymptotically stable equilibrium at $(y, z) = (0, 0)$. ∎

Consider system (15) and the control

$$u_t^{zn,\text{stab}} = -\frac{c_d(\zeta_t, \eta_t) - v(\zeta_t)}{b(\zeta_t, \eta_t)}$$

with c_d (and c_s) as defined in Section V-B and $v(\zeta_t) = -d_{r-1}\zeta_r - d_{r-2}\zeta_{r-1} - \cdots - d_0\zeta_1$, with $d_i \in \mathbb{R}$ for $i = 1, \ldots, r - 1$ such that the polynomial (26) has roots with negative real part. Then, the following holds.

Proposition 4: Consider system (15) and suppose that the equilibrium at $\eta = 0$ of the zero dynamics of system (15) is locally asymptotically stable almost surely. Additionally, suppose that $c_s(0, \eta) \equiv 0$ for η in a neighborhood of zero and that $\frac{\partial c_s}{\partial \zeta}(0, 0) = 0$. Then, the control law $u_t^{zn,\text{stab}}$ renders the equilibrium $(\zeta, \eta) = (0, 0)$ asymptotically stable almost surely.

Proof: Observe that, if $u_t = u_t^{zn,\text{stab}}$ in (15) then, the closed-loop dynamics is given by

$$\dot{\zeta}_t = A\zeta_t + Bc_s(\zeta_t, \eta_t)\xi_t, \quad \dot{\eta}_t = p(\xi_t, \zeta_t, \eta_t)$$

where A is given by (27), which has characteristic polynomial (26). This system is in the form (37) with $\gamma(\xi_t, \zeta_t, \eta_t) = Bc_s(\zeta_t, \eta_t)\xi_t$ and, by the assumptions on c_s, γ satisfies the hypotheses of Lemma 9. Therefore, the equilibrium at the origin is asymptotically stable almost surely. ∎

Remark 14: When studying asymptotic stabilization, the convergence of the states to the equilibrium at the origin contradicts the persistence of excitation condition of Assumption 2. However, following from the discussion in Remark 12, if the hypotheses of Proposition 4 are satisfied, one might perform practical asymptotic stabilization using a control $\hat{u}_t^{stab} = u_t^{zn,\text{stab}} + u_t^s$ until the states are in an arbitrarily small neighborhood of zero, and then switch to just the zero-noise control $u_t^{zn,\text{stab}}$. Doing this has the advantage of obtaining state trajectories that are closer, as ε goes to zero, to the idealistic ones when the system is away from the equilibrium, thus ensuring a more predictable behavior of the system.

REFERENCES

[1] A. Isidori, A. Krener, C. Gori-Giorgi, and S. Monaco, "Nonlinear decoupling via feedback: A differential geometric approach," *IEEE Trans. Autom. Control*, vol. AC-26, no. 2, pp. 331–345, Apr. 1981.

[2] M. Zeitz, "Controllability canonical (phase-variable) form for non-linear time-variable systems," *Int. J. Control*, vol. 37, no. 6, pp. 1449–1457, 1983.

[3] D. Bestle and M. Zeitz, "Canonical form observer design for non-linear time-variable systems," *Int. J. Control*, vol. 38, no. 2, pp. 419–431, 1983.

[4] A. J. Krener, "Normal forms for linear and nonlinear systems," *Contemporary Math.*, vol. 68, pp. 157–189, 1987.

[5] R. Brockett, "Feedback invariants for nonlinear systems," *IFAC Proc. Volumes*, vol. 11, no. 1, pp. 1115–1120, 1978.

[6] B. Jakubczyk and W. Respondek, "On linearization of control systems," *Bull. de l'Académie Polonaise des Sci.. Série des Sci. Mathématiques*, vol. 28, pp. 517–522, 1980.

[7] R. Su, "On the linear equivalents of nonlinear systems," *Syst. Control Lett.*, vol. 2, no. 1, pp. 48–52, 1982.

[8] L. Hunt and G. Meyer, "Global transformations of nonlinear systems," *IEEE Trans. Autom. Control*, vol. AC-28, no. 1, pp. 24–31, Jan. 1983.

[9] C. I. Byrnes and A. Isidori, "A frequency domain philosophy for nonlinear systems, with applications to stabilization and to adaptive control," in *Proc. 23rd IEEE Conf. Decis. Control*, 1984, pp. 1569–1573.

[10] C. I. Byrnes and A. Isidori, "Local stabilization of minimum-phase nonlinear systems," *Syst. Control Lett.*, vol. 11, no. 1, pp. 9–17, 1988.

[11] T. Damm, *Rational Matrix Equations in Stochastic Control (Lecture Notes in Control and Information Sciences, no. 297)*. Berlin, Germany: Springer, 2004.

[12] D. Hinrichsen and A. J. Pritchard, "Stochastic H^∞," *SIAM J. Control Optim.*, vol. 36, no. 5, pp. 1504–1538, 1998.

[13] E. Gershon, "Robust reduced-order H_∞ output-feedback control of retarded stochastic linear systems," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2898–2904, Nov. 2013.

[14] H. Hua, J. Cao, G. Yang, and G. Ren, "Voltage control for uncertain stochastic nonlinear system with application to energy internet: Non-fragile robust H_∞ approach," *J. Math. Anal. Appl.*, vol. 463, no. 1, pp. 93–110, 2018.

[15] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations (Stochastic Modelling and Applied Probability)*. New York, NY, USA: Springer, 1999.

[16] B. Øksendal, *Stochastic Differential Equations*, 6th ed. Berlin, Germany: Springer-Verlag, 2003.

[17] L. Arnold and P. Imkeller, "Normal forms for stochastic differential equations," *Probability Theory Related Fields*, vol. 110, no. 4, pp. 559–588, May 1998.

[18] L. Arnold, *Random Dynamical Systems (Springer Monographs in Mathematics)*. Berlin, Germany: Springer-Verlag, 2003.

[19] G. Gaeta and N. Rodríguez Quintero, "Lie-point symmetries and stochastic differential equations," *J. Phys. A: Math. Gen.*, vol. 32, no. 48, pp. 8485–8505, 1999.

[20] T. Lahdhiri and A. Alouani, "An introduction to the theory of exact stochastic feedback linearization for nonlinear stochastic systems," *IFAC Proc. Volumes*, vol. 29, no. 1, pp. 3904–3909, 1996.

[21] Z. Pan, "Differential geometric condition for feedback complete linearization of stochastic nonlinear system," *Automatica*, vol. 37, no. 1, pp. 145–149, 2001.

[22] Z. Pan, "Canonical forms for stochastic nonlinear systems," *Automatica*, vol. 38, no. 7, pp. 1163–1170, 2002.

[23] Z. Pan and T. Basar, "Backstepping controller design for nonlinear stochastic systems under a risk-sensitive cost criterion," *SIAM J. Control Optim.*, vol. 37, no. 3, pp. 957–995, 1999.

[24] A. Mellone and G. Scarciotti, "Normal form and exact feedback linearisation of nonlinear stochastic systems: The ideal case," in *Proc. IEEE 58th Conf. Decis. Control*, 2019, pp. 3503–3508.

[25] A. Mellone and G. Scarciotti, "The zero dynamics of nonlinear stochastic systems: Stabilization and output tracking in the ideal case," in *Proc. 21st IFAC World Congr.*, 2020, pp. 5053–5058.

[26] A. Mellone and G. Scarciotti, "Approximate feedback linearisation and stabilisation of nonlinear stochastic systems," in *Proc. 21st IFAC World Congr.*, 2020, pp. 5059–5064.

[27] L. Arnold, *Stochastic Differential Equations*. Hoboken, NJ, USA: Wiley, 1974.

[28] L. Hörmander, *The Analysis of Linear Partial Differential Operators I (Classics in Mathematics)*. Berlin, Germany: Springer-Verlag, 1983.

[29] J. J. Duistermaat and J. A. C. Kolk, *Distributions (Cornerstones)*. Cambridge, MA, USA: Birkäuser, 1995.

[30] A. Isidori, *Nonlinear Control Systems (Communications and Control Engineering)*. London, U.K.: Springer-Verlag, 1995.

[31] A. Mellone and G. Scarciotti, "A note on the Itô and Stratonovich stochastic relative degree and normal form," in *Proc. 59th IEEE Conf. Decis. Control*, 2020, pp. 4306–4311.

[32] E. Wong and M. Zakai, "On the convergence of ordinary integrals to stochastic integrals," *Ann. Math. Statist.*, vol. 36, no. 5, pp. 1560–1564, Oct. 1965.

[33] A. Mellone and G. Scarciotti, "Output regulation of linear stochastic systems," *IEEE Trans. Autom. Control*, vol. 67, no. 4, pp. 1728–1743, Apr. 2022.

[34] T. Gard, *Introduction to Stochastic Differential Equations (Monographs and Textbooks in Pure and Applied Mathematics)*. New York, NY, USA: Marcel Dekker, 1988.

[35] H. K. Khalil, *Nonlinear Systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.

[36] M. Bardi and A. Cesaroni, "Almost sure stabilizability of controlled degenerate diffusions," *SIAM J. Control Optim.*, vol. 44, no. 1, pp. 75–98, 2005.

[37] F. Kozin, "On almost sure stability of linear systems with random coefficients," *J. Math. Phys.*, vol. 42, no. 1–4, pp. 59–67, 1963.

[38] R. Z. Khas'minskii, "Necessary and sufficient conditions for the asymptotic stability of linear stochastic systems," *Theory Probability Appl.*, vol. 12, no. 1, pp. 144–147, 1967.

[39] W. Hahn and A. P. Baartz, *Stability of Motion (Grundlehren Der Mathematischen Wissenschaften)*. Berlin, Germany: Springer-Verlag, 1967.

[40] P. Boxler, "A stochastic version of center manifold theory," *Probability Theory Related Fields*, vol. 83, no. 4, pp. 509–545, Dec. 1989.

Alberto Mellone (Student Member, IEEE) was born in Galatina (Lecce), Italy, in 1993. He received the B.S. degree in information engineering from the University of Salento, Lecce, Italy in 2015 and the M.S. degree in robotics and automation engineering from the University of Pisa, Pisa, Italy in 2018.

In 2018 he joined the Control and Power Group, Imperial College London, London, U.K., where he pursued a Ph.D. looking into analysis and control of stochastic systems. He graduated in 2022 with a thesis on path-wise output regulation and nonlinear control of stochastic systems.

Giordano Scarciotti (Senior Member, IEEE) was born in Frascati (Rome), Italy, in 1988. He received his B.Sc. and M.Sc. degrees in automation engineering from the University of Rome "Tor Vergata," Italy, in 2010 and 2012, respectively.

In 2012 he joined the Control and Power Group, Imperial College London, U.K., where he obtained a Ph.D. degree in 2016 with a thesis on approximation, analysis and control of large-scale systems. He also received an M.Sc. degree in Applied Mathematics from Imperial College in 2020. He is currently a Senior Lecturer in the Control and Power Group. His current research interests are focused on analysis and control of uncertain systems, model reduction and optimal control. He was a visiting scholar at New York University in 2015 and at University of California Santa Barbara in 2016 and he is currently a Visiting Fellow of Shanghai University. He is the recipient of an Imperial College Junior Research Fellowship (2016), of the IET Control and Automation Ph.D. Award (2016), the Eryl Cadwaladr Davies Prize (2017) and an ItalyMadeMe award (2017). He is a member of the IEEE CSS Conference Editorial Board, of the IFAC and IEEE CSS Technical Committees on Nonlinear Control Systems and has served in the International Programme Committee of multiple conferences. He is the National Organising Committee Chair for the EUCA European Control Conference 2022 and the Invited Session Chair for IFAC Symposium on Nonlinear Control Systems 2022.