SIAM J. CONTROL OPTIM.
Vol. 45, No. 3, pp. 885–926

PRACTICAL OUTPUT-FEEDBACK RISK-SENSITIVE CONTROL FOR STOCHASTIC NONLINEAR SYSTEMS WITH STABLE ZERO-DYNAMICS*

YUN-GANG LIU† AND JI-FENG ZHANG‡

Abstract. This paper addresses the design problem of practical (or satisfaction) output-feedback controls for stochastic strict-feedback nonlinear systems in observer canonical form with stable zero-dynamics under long-term average tracking risk-sensitive cost criteria. The cost function adopted here is of the quadratic-integral type usually encountered in practice, rather than the quartic-integral one used to avoid difficulty in control design and performance analysis of the closed-loop system. A sequence of coordinate diffeomorphisms is introduced to separate the zero-dynamics from the entire system, so that the transformed system has an appropriate form suitable for integrator backstepping design. For any given risk-sensitivity parameter and desired cost value, by using the integrator backstepping methodology, an output-feedback control is constructively designed such that (a) the closed-loop system is bounded in probability and (b) the long-term average risk-sensitive cost is upper bounded by the desired value. In addition, this paper does not require the uniform boundedness of the gain functions of the system noise. Furthermore, an example is given to show the effectiveness of the theory.

Key words. nonlinear system, stochastic system, integrator backstepping methodology, risk-sensitive control, output-feedback control, zero dynamics

AMS subject classifications. 93B52, 93C10, 93E03, 93E15, 93E20

DOI. 10.1137/S0363012903439185

1. Introduction.

Research on global stabilization control design for nonlinear systems has been accelerated over the last two decades. After the celebrated characterization of the feedback linearizable systems (see [13]), a breakthrough was achieved with the introduction of the integrator backstepping design methodology (see [20]), which provides a general constructive tool for designing global stabilization controls for nonlinear systems in or feedback equivalent to strict-feedback form. Since the early 1990s, a series of research results on strict-feedback systems have been obtained by using this methodology together with other design tools, such as nonlinear damping, tuning functions, and MT filters (see, e.g., [8], [15], [18], [19], [22], [23], [32], [34], and [38]).

The research on risk-sensitive control can be traced back to the early 1970s, when Jacobson introduced the linear exponential quadratic Gaussian (LEQG) problem (see [14]). Then, Whittle put a risk-sensitivity parameter into the cost, and solved the linear discrete-time problem (see [39]). Bensoussan and van Schuppen considered the continuous-time case in their paper [4]. But the significance of the risk-sensitive control was not fully realized until the 1990s. It has been known that risk-sensitive control is more general than H_∞ control and H_2 control, and closely related to differential game problems (see, e.g., [9], [10], [17], [31], [37], and [40]). For example, when

*Received by the editors December 30, 2003; accepted for publication (in revised form) February 18, 2006; published electronically July 17, 2006. This work was supported by the National Natural Science Foundation of China under grants 60274021, 60221301, 60334040, 60304002, and the Science and Technical Development Plan of Shandong Province under grant 2004GG4204014.

http://www.siam.org/journals/sicon/45-3/43918.html

†School of Control Science and Engineering, Shandong University, Jinan 250061, China (lygfr@sdu.edu.cn).

‡Corresponding author. Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, China (jif@iss.ac.cn).

the noise vanishes, the large deviation limit of the risk-sensitive control is nothing but a deterministic differential game problem. These connections have initialized and accelerated the research on stochastic risk-sensitive controls over the last decade.

The design of controls for strict-feedback stochastic nonlinear systems has received intense investigation recently (see, e.g., [1], [5], [6], [7], [11], [26], [27], [28], [29], [33], and [35]), where [7], [11], and [33] considered full state-feedback control design, and [1], [5], [6], [26], [27], [28], [29], and [35] considered output-feedback control design. Under the assumption (A), "the disturbance vector field vanishes at the origin," [5], [7], and [11] studied the problem of designing a control to asymptotically stabilize the closed-loop systems in the large. Meanwhile, [1], [6], [26], [27], [28], [29], [33], and [35] considered the control design to achieve the boundedness in probability of the closed-loop system without using assumption (A). Specifically, [7] considered the disturbance attenuation problem; [35] considered the stabilization problem of systems with stable zero-dynamics; [33], [26], [1], and [29] considered the design of satisfaction control under a quadratic, a quartic regulation, and a quadratic tracking risk-sensitive cost criterion, respectively. [1] used the assumption (B), "the gain functions of stochastic noise are uniformly bounded," while [26], [29], and [33] did not; [27] and [28] considered the reduced-order observer-based stabilization control design of the single-input multioutput stochastic nonlinear systems.

This paper studies the problem of output-feedback control design for a class of stochastic nonlinear systems in observer canonical form with stable zero-dynamics under a quadratic tracking risk-sensitive cost criterion. In general, the design of output-feedback control is more difficult and challenging than that of full state-feedback control. Since the early 1990s, a general framework for studying output-feedback control problems has been developed. The key thought is to first introduce the so-called *information state*, which is a generalization of observer or filter, and then, by a measure transformation, to change the output-feedback control design problem into a full state-feedback problem of an augmented system (see, e.g., [2], [3], [12], [16], and [17]). However, generally speaking, the equality (or inequality) of the information state satisfied is infinite-dimensional, to which an explicit finite-dimensional solution exists only for linear or special nonlinear systems (see [2]). The method of this paper is different from the information state one and can be used to deal with more general inherently nonlinear systems. The objective of this paper is very practical: to search for a *satisfaction* control rather than an optimal one. This makes it possible to avoid the measure transformation. In order to get the explicit formula of the control, strict-feedback nonlinear systems are considered. The main results of this paper indicate that for any given risk-sensitivity parameter and desired tracking risk-sensitive cost value, a dynamic output-feedback control can always be constructively designed so that the closed-loop system is bounded in probability and the long-time average risk-sensitive cost is upper bounded by the desired value. While [1] considered assumption (B) to be essential, the current paper does not use this assumption. In addition, the value range of the characteristic parameter of the value function used for backstepping design is enlarged from $\frac{2}{3}$ (see [26]) to set $(\frac{1}{2}, 1)$. This provides control designers with a freedom in choosing the value function.

The remainder of the paper is organized as follows. Section 2 provides some notation. Section 3 describes the system model and formulates the control objective to be studied. Section 4 describes the constructive design procedure of the control by employing an integrator backstepping approach, and presents several important lemmas for the closed-loop performance analysis. Section 5 addresses the main results of this paper. Section 6 gives a simulation example to illustrate our theoretical findings. Section 7 gives some concluding remarks. The paper ends with two appendices. Appendix A introduces the definitions of stability notions *asymptotically stable in the large* and *bounded in probability*, and gives a key theorem of sufficient conditions for the solvability of the control problem. Appendix B provides some technical lemmas that play an important role in the control design and performance analysis.

2. Notation. Throughout this paper, $\mathbb{N}$ denotes the set of all natural numbers; $\mathbb{R}$ denotes the set of all real numbers, and $\mathbb{R}^n$ denotes the real n-dimensional space, $n \in \mathbb{N}$; $\mathcal{C}^i$ denotes the set of all functions with continuous partial derivative up to ith order, $i \in \mathbb{N}$, and $\mathcal{C}^\infty$ denotes the set of all smooth functions; for a given vector or matrix W, we use $W^\top$ to denote its transpose; $\mathrm{Tr}(W)$ denotes its trace when W is square, i.e., the sum of all elements on the main diagonal line; we use $|W|$ to denote the absolute value for scalar numbers, and $\|W\|$ to denote the Euclidean norm for vectors and the corresponding induced norm for matrices; we also introduce the Frobenius norm of W defined by $\|W\|_F = \sqrt{\mathrm{Tr}(W^\top W)}$ with properties: $\|W\| \leq \|W\|_F$ and $\|WV\|_F \leq \|W\|\|V\|_F$ for any matrix V with appropriate dimension; for any $x \in \mathbb{R}^n$, x_i denotes its ith element, $x_{[i]}$ denotes the column vector consisting of the first i elements of x in the original order, i.e., $x_{[i]} = [x_1, \ldots, x_i]^\top$; for any given ith continuously differentiable function $y_d(t)$, $y_d^{(i)}(t)$ denotes the ith derivative with respect to the time variable t, the first and second derivatives are denoted by $\dot{y}_d$ and $\ddot{y}_d$, respectively, and $y_d^{[i]}$ denotes the $(i+1)$-dimensional column vector consisting of $y_d,\ \dot{y}_d, \ldots, y_d^{(i)}$, i.e., $y_d^{[i]} = [y_d, \dot{y}_d, \ddot{y}_d, \ldots, y_d^{(i)}]^\top$. Obviously, $x_{[1]} = x_1$, $x_{[n]} = x$, $y_d^{[0]} = y_d$. $0_{i\times j}$ denotes the $(i \times j)$-dimensional matrix with all zero elements and will be written as 0 for brevity when there is no confusion caused. We use I_i to denote the $i \times i$ identity matrix. For a set A, I_A denotes the characteristic function of the set. For any given symmetric matrix P, $\lambda_{\max}(P)$ and $\lambda_{\min}(P)$ denote its maximum and minimum eigenvalue, respectively.

In addition, when a function shows up for the first time, we will clearly write out its arguments, and then, for simplicity of expression in later use, we sometimes drop the arguments when no confusion is caused.

For a given stochastic system

$$\mathrm{d}x = f(t,x)\,\mathrm{d}t + h(t,x)\,\mathrm{d}w, \quad x(t_0) = x_0,$$

define a differential operator $\mathcal{L}$:

$$\mathcal{L}V(t,\,x) = \frac{\partial V(t,\,x)}{\partial t} + \frac{\partial V(t,\,x)}{\partial x} f(t,\,x) + \frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2 V(t,\,x)}{\partial x^2} h(t,\,x)\,(h(t,\,x))^\top\right),$$

where x is an n-dimensional state vector, $n \in \mathbb{N}$; $f:\ [0,\ \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ and $h:\ [0,\ \infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times s}$, $s \in \mathbb{N}$, are assumed to be continuous in t and locally Lipschitz in x; w is an s-dimensional vector-valued Brownian motion defined on a probability space $(\Omega,\ \mathcal{F},\ \mathcal{P})$; and $V:\ [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^1$ in t and $\mathcal{C}^2$ in x.

3. Problem formulation.

3.1. System model. We consider the stochastic nonlinear systems in observer canonical form with zero-dynamics of the form (see [35]):

$$
\begin{aligned}
\mathrm{d}x_1 &= x_2\,\mathrm{d}t + f_1(y)\,\mathrm{d}t + h_1(y)\,\mathrm{d}w,\\
&\ \ \vdots\\
\mathrm{d}x_{\rho-1} &= x_\rho\,\mathrm{d}t + f_{\rho-1}(y)\,\mathrm{d}t + h_{\rho-1}(y)\,\mathrm{d}w,\\
\mathrm{d}x_\rho &= x_{\rho+1}\,\mathrm{d}t + f_\rho(y)\,\mathrm{d}t + b_m g(y)u\,\mathrm{d}t + h_\rho(y)\,\mathrm{d}w,\\
&\ \ \vdots\\
\mathrm{d}x_{n-1} &= x_n\,\mathrm{d}t + f_{n-1}(y)\,\mathrm{d}t + b_1 g(y)u\,\mathrm{d}t + h_{n-1}(y)\,\mathrm{d}w,\\
\mathrm{d}x_n &= f_n(y)\,\mathrm{d}t + b_0 g(y)u\,\mathrm{d}t + h_n(y)\,\mathrm{d}w,\\
y &= x_1,
\end{aligned}
\tag{3.1}
$$

where $x = [x_1, \ldots, x_n]^\top$ is the n-dimensional state vector, $n \in \mathbb{N}$, and its initial value $x(t_0) = x_0$ is fixed but unknown; u is the scalar control input; y is the scalar measurable output; $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$, are the system nonlinearities depending only on output y; $h_i : \mathbb{R} \to \mathbb{R}^{1\times s}$, $i = 1, \ldots, n$, are the gain functions of the system noise depending only on y, $s \in \mathbb{N}$; $g : \mathbb{R} \to \mathbb{R}$ is the nonlinear gain function of the control input u depending only on y; $w \in \mathbb{R}^s$ is a vector-valued standard Brownian motion defined on probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with Ω being a sample space, $\mathcal{F}$ being a filtration, and $\mathcal{P}$ being the probability measure, $s \in \mathbb{N}$; $m \in \mathbb{N}$ satisfies $0 \le m < n$; and $\rho = n - m \in \mathbb{N}$ is the relative degree of the system.

The main results of this paper are based on the following assumptions:

A1. The nonlinear functions f_i and h_i $(i = 1, \ldots, n)$ are smooth. That is, $f_i \in \mathcal{C}^\infty$ and $h_i \in \mathcal{C}^\infty$; the nonlinear function g is continuous; and, for any $y \in \mathbb{R}$, $g(y) \neq 0$.

A2. All the roots of the polynomial $b_m s^m + \cdots + b_1 s + b_0$, $b_m \neq 0$, have negative real parts.

A3. Desired system output y_d is deterministic, and it and its derivatives $\dot{y}_d, \ldots, y_d^{(\rho)}$ are known and bounded; i.e., there exist known positive constants $C_{y_d^{(i)}}$, $i = 0, \ldots, \rho$, that bound the reference trajectory y_d and its derivatives.

Assumption A1 is standard for this class of control problems, to ensure that f_i and h_i $(i = 1, \ldots, n)$ are local Lipschitz functions and, together with assumption A3, to ensure the global boundedness of $h_i(y_d)$ $(i = 1, \ldots, n)$. Assumption A2 ensures that the zero-dynamics are stable.

Unlike the problem of feedback stabilization, there is no need to require that the origin $x = \mathbf{0}_{n\times 1}$ be the equilibrium point of the open-loop system. This is because the purpose of the tracking control is to make the system output conform to the time-varying desired system output $y_d(t)$, rather than to steer the system state to the origin $x = \mathbf{0}_{n\times 1}$.

3.2. Control objective. The goal of control design is to make the solution process of the system (3.1) be bounded in probability and the following quadratic tracking risk-sensitive cost criterion achieve a predefined long-term cost value:

$$
J_\theta(y) = \limsup_{T\to\infty} \frac{1}{T}\,\frac{2}{\theta} \ln\left(E\left(\exp\left(\frac{\theta}{2}\int_0^T (y - y_d)^2\,\mathrm{d}t\right)\right)\right).
\tag{3.2}
$$

That is, for any given positive cost value R_l (arbitrarily close to zero), the risk-sensitive cost $J_\theta(y)$ is not greater than R_l, where θ is called the risk-sensitivity parameter and $y - y_d$ is called the output tracking error. When $\theta > 0$, the cost function weights heavily the large deviation of $y - y_d$ through the exponential operator, which leads to a risk-averse control design problem. The greater the value of θ, the more conservative is the controller. Actually, by the value of θ, the risk-sensitive problem can be classified (see [10] and [31]) as follows: (i) when $\theta > 0$, it is a risk-averse problem; (ii) when $\theta < 0$, it is a risk-seeking problem; (iii) when $\theta \to 0$, the cost function converges to a standard integral cost, and so it is known as a risk-neutral problem.

In this paper, we will study only the case where θ is positive.

For convenience of expression, we give the following definition.

DEFINITION 3.1. *For a given positive risk-sensitivity parameter θ, a controller u is said to achieve a guaranteed risk-sensitive cost R_l ($R_l > 0$) if the following inequality holds for the output of the closed-loop system:*

$$J_\theta(y) \le R_l.$$

In addition to the purposes of cost upper bound, we are also interested in achieving *boundedness in probability* for the closed-loop system. This notion, together with the *asymptotical stability in the large*, was introduced in the classical book [21] and has now been widely used. For the sake of the self-containedness of this paper, we will restate these two notions in Appendix A.

The system (3.1) can be rewritten into the following compact form:

$$\mathrm{d}x = f(x)\,\mathrm{d}t + Bg(x)u\,\mathrm{d}t + h(x)\,\mathrm{d}w, \tag{3.3}$$

where

$$f(x) = \begin{bmatrix} x_2 + f_1(x_1) \\ \vdots \\ x_n + f_{n-1}(x_1) \\ f_n(x_1) \end{bmatrix}, \qquad g(x) = g(x_1),$$

$$B = \begin{bmatrix} \mathbf{0}_{(n-m-1)\times 1} \\ b_m \\ \vdots \\ b_0 \end{bmatrix}, \qquad h(x) = \begin{bmatrix} h_1(x_1) \\ h_2(x_1) \\ \vdots \\ h_n(x_1) \end{bmatrix}.$$

If $\rho = 1$, then $m = n - \rho = n - 1$. For this special case, vector B defined above is simply $[b_{n-1}, \dots, b_1, b_0]^\top$.

For tracking purposes, the controller to be designed is time-varying in general, and so is the resulting closed-loop system, even though the original system is not. Thus, as in [26] and [33], with the long-term average risk-sensitive cost criterion $J_\theta(y)$, for a given desired cost value $R_l > 0$, a practical risk-sensitive output-feedback tracking control is designed as

$$\begin{cases} \dot{\xi} = \alpha\left(t,\, \xi,\, y\right), & \alpha \in \mathcal{C}^1, \\ u = \mu\left(t,\, \xi,\, y\right), & \mu \in \mathcal{C}^1, \end{cases} \tag{3.4}$$

so that there exists a nonnegative value function $V(t, x, \xi)$, which is $\mathcal{C}^1$ in t and $\mathcal{C}^2$ in $(x,\, \xi)$ and radially unbounded with respect to x and ξ, satisfying the following Hamilton–Jacobi–Bellman (HJB) inequality:

$$\begin{aligned} &\frac{\partial V}{\partial t} + \begin{bmatrix} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial \xi} \end{bmatrix} \begin{bmatrix} f + Bg\mu \\ \alpha \end{bmatrix} + \frac{\theta}{4}\frac{\partial V}{\partial x} hh^\top \left(\frac{\partial V}{\partial x}\right)^\top \\ &\qquad + \frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2 V}{\partial x^2} hh^\top\right) + (y - y_d)^2 \;\le\; R_l. \end{aligned} \tag{3.5}$$

From (3.5), it is easy to see that the essential difference between the stochastic HJB and deterministic HJB equations is that the former has the Itô term $\frac{1}{2}\mathrm{Tr}(\frac{\partial^2 V}{\partial x^2} hh^\top)$. How to deal with this term is the key to the control design and performance analysis.

4. Output-feedback risk-sensitive control design. We shall design the output-feedback tracking controller in three steps. First, we introduce an observer to rebuild the system states. With the observer dynamics in the loop, we introduce a sequence of coordinate diffeomorphisms transforming the system into a lower triangular structure which is amenable to the application of integrator backstepping methodology. Then, we describe the control design procedure and present several lemmas, which will be used for the performance analysis of the closed-loop systems in the next section.

4.1. Observer design. Since the states of (3.1), except for the state x_1 which can be obtained directly since $y = x_1$, are unknown and need an observer to rebuild,

$$\begin{aligned}
\dot{\widehat{x}}_1 &= \widehat{x}_2 + k_1(y - \widehat{x}_1) + f_1(y),\\
&\ \vdots\\
\dot{\widehat{x}}_{\rho-1} &= \widehat{x}_\rho + k_{\rho-1}(y - \widehat{x}_1) + f_{\rho-1}(y),\\
\dot{\widehat{x}}_\rho &= \widehat{x}_{\rho+1} + k_\rho(y - \widehat{x}_1) + f_\rho(y) + b_m g(y)u,\\
&\ \vdots\\
\dot{\widehat{x}}_{n-1} &= \widehat{x}_n + k_{n-1}(y - \widehat{x}_1) + f_{n-1}(y) + b_1 g(y)u,\\
\dot{\widehat{x}}_n &= k_n(y - \widehat{x}_1) + f_n(y) + b_0 g(y)u,
\end{aligned} \tag{4.1}$$

where $k_1, k_2, \ldots, k_n$ are design constants such that all the roots of polynomial $s^n + k_1 s^{n-1} + \cdots + k_n$ have negative real parts. The initial condition for observer (4.1) is set by certain value $\widehat{x}(t_0) = \widehat{x}_0$.

Let $\widehat{x} = [\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_n]^\top$. Both system output y and observer state vector $\widehat{x}$ are available for control design. Denote the state estimation error as $\widetilde{x} = x - \widehat{x}$. Then we have

$$\mathrm{d}\widetilde{x} = \begin{bmatrix} -k_1 & & & \\ \vdots & & I_{n-1} & \\ -k_n & 0 & \cdots & 0 \end{bmatrix} \widetilde{x}\,\mathrm{d}t + h(y)\,\mathrm{d}w \stackrel{\triangle}{=} A\widetilde{x}\,\mathrm{d}t + h(y)\,\mathrm{d}w. \tag{4.2}$$

Thus, with observer dynamics (4.1) in the loop, we have the following entire system:

$$\begin{aligned}
\mathrm{d}\widetilde{x} &= A\widetilde{x}\,\mathrm{d}t + h(y)\,\mathrm{d}w,\\
\mathrm{d}y &= (\widehat{x}_2 + \widetilde{x}_2)\,\mathrm{d}t + f_1(y)\,\mathrm{d}t + h_1(y)\,\mathrm{d}w,\\
\dot{\widehat{x}}_2 &= \widehat{x}_3 + k_2(y - \widehat{x}_1) + f_2(y),\\
&\ \vdots\\
\dot{\widehat{x}}_{\rho-1} &= \widehat{x}_\rho + k_{\rho-1}(y - \widehat{x}_1) + f_{\rho-1}(y),\\
\dot{\widehat{x}}_\rho &= \widehat{x}_{\rho+1} + k_\rho(y - \widehat{x}_1) + f_\rho(y) + b_m g(y)u,\\
&\ \vdots\\
\dot{\widehat{x}}_{n-1} &= \widehat{x}_n + k_{n-1}(y - \widehat{x}_1) + f_{n-1}(y) + b_1 g(y)u,\\
\dot{\widehat{x}}_n &= k_n(y - \widehat{x}_1) + f_n(y) + b_0 g(y)u.
\end{aligned} \tag{4.3}$$

System (4.3) has three parts, which are associated with the estimation error $\widetilde{x}$, system output y, and observer states $\widehat{x}_2, \ldots, \widehat{x}_\rho$, respectively. In particular, when $\rho = 1$, then $m = n - \rho = n - 1$. For this case, the second subequation of (4.3) shall be replaced by the following equation:

$$\mathrm{d}y = (\widehat{x}_2 + \widetilde{x}_2)\,\mathrm{d}t + f_1(y)\,\mathrm{d}t + b_{n-1}g(y)u\mathrm{d}t + h_1(y)\,\mathrm{d}w.$$

In the next subsections, with this entire system as starting point, we shall search for the desired controller.

4.2. Coordinate diffeomorphisms. To prepare for the backstepping design in the next subsection, we introduce a series of ρ coordinate diffeomorphisms (see [36]) so as to convert the entire system (4.3) into zero-dynamics canonical form, which is amenable to the application of integrator backstepping methodology.

The idea of such coordinate diffeomorphisms was first introduced in [30] and significantly modified in Chapter 8 of [24]. Our presentation, including the two cases of $\rho = 1$ and $\rho > 1$, is much more direct and easier to implement.

4.2.1. Case of $\boldsymbol{\rho = 1}$. When $\rho = 1$, $m = n - \rho = n - 1$. This means that control input appears in every subequation of (3.1) and (4.1). In this case, one coordinate transformation is sufficient to obtain the desired structure.

Let $\varsigma_0 = [y, \widehat{x}_2, \ldots, \widehat{x}_n]^\top$. Then, by (4.3), we have the following dynamics for ς_0:

$$\mathrm{d}\varsigma_0 = D_0\varsigma_0\,\mathrm{d}t + G_0(y, \widetilde{x}_1)\,\mathrm{d}t + [1, \mathbf{0}_{1\times(n-1)}]^\top \widetilde{x}_2\,\mathrm{d}t + g(y)B_0 u\,\mathrm{d}t + H_0(y)\,\mathrm{d}w, \tag{4.4}$$

where

$$\begin{aligned}
D_0 &= \begin{bmatrix} 0 & & & \\ \vdots & & I_{n-1} & \\ 0 & 0 & \cdots & 0 \end{bmatrix},\\
G_0 &= [f_1(y),\ f_2(y) + k_2\widetilde{x}_1, \ldots, f_n(y) + k_n\widetilde{x}_1]^\top,\\
B_0 &= [b_{n-1}, \ldots, b_1, b_0]^\top,\\
H_0 &= [(h_1(y))^\top,\ \mathbf{0}_{s\times(n-1)}]^\top.
\end{aligned}$$

By coordinate transformation we would like to transform the vector B_0 into one with all elements being zero except the first element, b_{n-1}. Let $\varsigma_1 = T_1\varsigma_0$, where T_1 is the same as I_n, except with the first column replaced by $[1, -b_{n-2}/b_{n-1}, \ldots, -b_0/b_{n-1}]^\top$. Then, T_1^{-1} is also the same as I_n, except with the first column replaced by $\left[1, \frac{b_{n-2}}{b_{n-1}}, \ldots, \frac{b_0}{b_{n-1}}\right]^\top$.

Then we have

$$\mathrm{d}\varsigma_1 = D_1\varsigma_1\,\mathrm{d}t + G_1(y, \widetilde{x}_1)\,\mathrm{d}t + [1, L_2^\top]^\top \widetilde{x}_2\,\mathrm{d}t + g(y)B_1 u\,\mathrm{d}t + H_1(y)\,\mathrm{d}w,$$

where $D_1 = T_1 D_0 T_1^{-1}$ is the same as D_0, except with the first and second columns replaced by $[d_{11}, \ldots, d_{1,n}]^\top$ and $[1, -b_{n-2}/b_{n-1}, \ldots, -b_0/b_{n-1}]^\top$, respectively:

$$\begin{aligned}
G_1 &= T_1 G_0(y, \widetilde{x}_1) \triangleq \left[\overline{g}_{11}(y) + \overline{d}_{11}\widetilde{x}_1, \ldots, \overline{g}_{1n}(y) + \overline{d}_{1n}\widetilde{x}_1\right]^\top,\\
L_2 &= [-b_{n-2}/b_{n-1}, \ldots, -b_1/b_{n-1},\ -b_0/b_{n-1}]^\top,\\
B_1 &= T_1 B_0 = [b_{n-1},\ \mathbf{0}_{1\times m}]^\top,\\
H_1 &= T_1 H_0(y) \triangleq \left[(\widehat{h}_1(y))^\top, \ldots, (\widehat{h}_n(y))^\top\right]^\top.
\end{aligned}$$

Denote $\eta = \varsigma_1 = [\eta_1, \ldots, \eta_n]^\top$ and $\zeta = [\eta_2, \eta_3, \ldots, \eta_n]^\top$. Then, the dynamics of $\widetilde{x}$, ζ, and η_1 can be expressed as follows:

$$
\begin{aligned}
\mathrm{d}\widetilde{x} &= A\widetilde{x}\,\mathrm{d}t + h(y)\,\mathrm{d}w,\\
\mathrm{d}\zeta &= \begin{bmatrix} -\frac{b_{n-2}}{b_{n-1}} & & & \\ \vdots & & I_{n-2} & \\ -\frac{b_1}{b_{n-1}} & & & \\ -\frac{b_0}{b_{n-1}} & 0 & \cdots & 0 \end{bmatrix} \zeta\,\mathrm{d}t + \begin{bmatrix} d_{12} \\ d_{13} \\ \vdots \\ d_{\rho n} \end{bmatrix} y\,\mathrm{d}t\\
&\quad + \begin{bmatrix} \overline{g}_{12}(y) + \overline{d}_{12}\widetilde{x}_1 \\ \overline{g}_{13}(y) + \overline{d}_{13}\widetilde{x}_1 \\ \vdots \\ \overline{g}_{1n}(y) + \overline{d}_{1n}\widetilde{x}_1 \end{bmatrix} \mathrm{d}t - \frac{1}{b_{n-1}} \begin{bmatrix} b_{n-2} \\ \vdots \\ b_1 \\ b_0 \end{bmatrix} \widetilde{x}_2\mathrm{d}t + \begin{bmatrix} \widehat{h}_2(y) \\ \widehat{h}_3(y) \\ \vdots \\ \widehat{h}_n(y) \end{bmatrix} \mathrm{d}w \qquad (4.5)\\
&\triangleq E\zeta\,\mathrm{d}t + L_1\widetilde{x}_1\,\mathrm{d}t + L_2\widetilde{x}_2\,\mathrm{d}t + G(y)\,\mathrm{d}t + \Psi(y)\,\mathrm{d}w,\\
\mathrm{d}\eta_1 &= \left[1,\, \mathbf{0}_{1\times(n-2)}\right]\zeta\,\mathrm{d}t + (\overline{g}_{11}(y) + d_{11}y)\,\mathrm{d}t\\
&\quad + \overline{d}_{11}\widetilde{x}_1\,\mathrm{d}t + \widetilde{x}_2\,\mathrm{d}t + b_{n-1}g(y)u\,\mathrm{d}t + \widehat{h}_1(y)\,\mathrm{d}w,
\end{aligned}
$$

where

$$
\begin{aligned}
y &= \eta_1,\\
L_1 &= [\overline{d}_{12},\, \overline{d}_{13}, \ldots, \overline{d}_{1n}]^\top,\\
G &= [\overline{g}_{12}(y) + d_{12}y,\, \overline{g}_{13}(y) + d_{13}y, \ldots, \overline{g}_{1n}(y) + d_{1n}y]^\top.
\end{aligned}
$$

This system is equivalent to the entire system (4.3) under the transformation $[\eta_1, \zeta^\top]^\top = T_1[y, \widehat{x}_2, \ldots, \widehat{x}_n]^\top$. The structure of (4.5) makes the design of an output feedback controller much easier (see the latter design procedure for details).

4.2.2. Case of $\rho > 1$. Let us now give the coordinate transformations for the case of $\rho > 1$. From the ρ transformations below one can see that there exist some essential differences between this case and the case of $\rho = 1$.

Let $\overline{g}_{01}(y) = f_1(y)$, $\overline{d}_{0i} = k_i$, and $\overline{g}_{0i}(y) = f_i(y)$ $(i = 2, \ldots, n)$. Then we have the following dynamics for $\varsigma_0 = [y, \widehat{x}_2, \ldots, \widehat{x}_n]^\top$:

$$
\mathrm{d}\varsigma_0 = D_0\varsigma_0\,\mathrm{d}t + G_0(y, \widetilde{x}_1)\,\mathrm{d}t + [1,\, \mathbf{0}_{1\times(n-1)}]^\top\widetilde{x}_2\,\mathrm{d}t + g(y)B_0u\,\mathrm{d}t + H_0(y)\,\mathrm{d}w,
$$

where matrix D_0 and function H_0 are the same as those of (4.4), and

$$
\begin{aligned}
G_0 &= [f_1(y),\, f_2(y) + k_2\widetilde{x}_1, \ldots, f_n(y) + k_n\widetilde{x}_1]^\top\\
&\triangleq [\overline{g}_{01}(y) + \overline{d}_{01}\widetilde{x}_1,\, \overline{g}_{02}(y) + \overline{d}_{02}\widetilde{x}_1, \ldots, \overline{g}_{0n}(y) + \overline{d}_{0n}\widetilde{x}_1]^\top\\
&\triangleq [g_{01}(y, \widetilde{x}_1),\, g_{02}(y, \widetilde{x}_1), \ldots, g_{0n}(y, \widetilde{x}_1)]^\top,\\
B_0 &= [\mathbf{0}_{1\times(\rho-1)}, b_m, b_{m-1}, \ldots, b_0]^\top.
\end{aligned}
$$

By the first transformation, we would like to transform the matrix B_0 into one with all elements being zero except the ρth element, b_m. Let $\varsigma_1 = T_1\varsigma_0$, where T_1 is the same as I_n, except with the ρth column replaced by $\left[\mathbf{0}_{1\times(\rho-1)}, 1, -\frac{b_{m-1}}{b_m}, \ldots, -\frac{b_0}{b_m}\right]^\top$.

Then T_1^{-1} is also the same as I_n, except with the ρth column replaced by $\left[\mathbf{0}_{1\times(\rho-1)}, 1, -\frac{b_{m-1}}{b_m}, \dots, -\frac{b_0}{b_m}\right]^\top$.

Then we have

$$\mathrm{d}\varsigma_1 = D_1\varsigma_1\,\mathrm{d}t + G_1(y,\widetilde{x}_1)\,\mathrm{d}t + \left[1, \mathbf{0}_{1\times(n-1)}\right]^\top \widetilde{x}_2\,\mathrm{d}t + g(y)B_1u\,\mathrm{d}t + H_1(y)\,\mathrm{d}w,$$

where $D_1 = T_1D_0T_1^{-1}$ is the same as D_0 except with the ρth and $(\rho+1)$st columns replaced by $[\mathbf{0}_{1\times(\rho-2)}, 1, d_{11}, \dots, d_{1,m+1}]^\top$ and $[\mathbf{0}_{1\times(\rho-1)}, 1, -b_{m-1}/b_m, \dots, -b_0/b_m]^\top$, respectively,

$$\begin{aligned}
G_1 &= T_1G_0(y,\widetilde{x}_1)\\
&\triangleq [\overline{g}_{11}(y) + \overline{d}_{11}\widetilde{x}_1, \dots, \overline{g}_{1n}(y) + \overline{d}_{1n}\widetilde{x}_1]^\top\\
&\triangleq [g_{11}(y,\widetilde{x}_1), \dots, g_{1n}(y,\widetilde{x}_1)]^\top,\\
B_1 &= T_1B_0 = [\mathbf{0}_{1\times(\rho-1)},\, b_m,\, \mathbf{0}_{1\times m}]^\top,\\
H_1 &= T_1H_0(y) = H_0(y).
\end{aligned}$$

By the ith ($i = 2, \dots, \rho-1$) transformation, we would like to transform the $(\rho-i+2)$nd column of the matrix D_{i-1} into the $(\rho-i+1)$st unit vector. Let $\varsigma_i = T_i\varsigma_{i-1}$, where T_i is the same as I_n except with the $(\rho-i+1)$st column replaced by $[\mathbf{0}_{1\times(\rho-i)}, 1, -d_{i-1,1}, \dots, -d_{i-1,m+i-1}]^\top$. Then, T_i^{-1} is also the same as I_n except with the $(\rho-i+1)$st column replaced by $[\mathbf{0}_{1\times(\rho-i)}, 1, d_{i-1,1}, \dots, d_{i-1,m+i-1}]^\top$.

This leads to

$$\mathrm{d}\varsigma_i = D_i\varsigma_i\,\mathrm{d}t + G_i(y,\widetilde{x}_1)\,\mathrm{d}t + [1, \mathbf{0}_{1\times(n-1)}]^\top\widetilde{x}_2\,\mathrm{d}t + g(y)B_iu\,\mathrm{d}t + H_i(y)\,\mathrm{d}w,$$

where $D_i = T_iD_{i-1}T_i^{-1}$ is the same as D_0 except with the $(\rho-i+1)$st and $(\rho+1)$st columns replaced by $[\mathbf{0}_{1\times(\rho-i-1)}, 1, d_{i1}, \dots, d_{i,i+m}]^\top$ and $[\mathbf{0}_{1\times(\rho-1)}, 1, -b_{m-1}/b_m, \dots, -b_0/b_m]^\top$, respectively,

$$\begin{aligned}
G_i &= T_iG_{i-1}(y,\widetilde{x}_1)\\
&\triangleq [\overline{g}_{i1}(y) + \overline{d}_{i1}\widetilde{x}_1, \dots, \overline{g}_{in}(y) + \overline{d}_{in}\widetilde{x}_1]^\top\\
&\triangleq [g_{i1}(y,\widetilde{x}_1), \dots, g_{in}(y,\widetilde{x}_1)]^\top,\\
B_i &= T_iB_{i-1} = B_1 = [\mathbf{0}_{1\times(\rho-1)},\, b_m,\, \mathbf{0}_{1\times m}]^\top,\\
H_i &= T_iH_{i-1}(y) = H_0(y).
\end{aligned}$$

Finally, by the last transformation, we would like to transform the second column of the matrix $D_{\rho-1}$ into the first unit vector. Let $\varsigma_\rho = T_\rho\varsigma_{\rho-1}$, where T_ρ is the same as I_n except with the first column replaced by $[1, -d_{\rho-1,1}, \dots, -d_{\rho-1,n-1}]^\top$. Then, T_ρ^{-1} is also the same as I_n except with the first column replaced by $[1, d_{\rho-1,1}, \dots, d_{\rho-1,n-1}]^\top$.

This leads to

$$\begin{aligned}
\mathrm{d}\varsigma_\rho = {}& D_\rho\varsigma_\rho\,\mathrm{d}t + G_\rho(y,\widetilde{x}_1)\,\mathrm{d}t + [1, -d_{\rho-1,1}, \dots, -d_{\rho-1,n-1}]^\top\,\widetilde{x}_2\,\mathrm{d}t\\
&+ g(y)B_\rho u\,\mathrm{d}t + H_\rho(y)\,\mathrm{d}w,
\end{aligned}$$

where $D_\rho = T_\rho D_{\rho-1} T_\rho^{-1}$ is the same as D_0 except with the first and $(\rho+1)$st column replaced by $[d_{\rho 1}, \ldots, d_{\rho n}]^\top$ and $[\mathbf{0}_{1\times(\rho-1)}, 1, -b_{m-1}/b_m, \ldots, -b_0/b_m]^\top$, respectively,

$$\begin{aligned}
G_\rho &= T_\rho G_{\rho-1}(y, \widetilde{x}_1) \\
&\triangleq \left[\overline{g}_{\rho 1}(y) + \overline{d}_{\rho 1}\widetilde{x}_1, \ldots, \overline{g}_{\rho n}(y) + \overline{d}_{\rho n}\widetilde{x}_1\right]^\top \\
&\triangleq \left[g_{\rho 1}(y, \widetilde{x}_1), \ldots, g_{\rho n}(y, \widetilde{x}_1)\right]^\top, \\
B_\rho &= T_\rho B_{\rho-1} = B_1, \\
H_\rho &= T_\rho H_{\rho-1} = T_\rho H_0 \triangleq \left[(\widehat{h}_1(y))^\top, \ldots, (\widehat{h}_n(y))^\top\right]^\top.
\end{aligned}$$

Denote $\eta = \varsigma_\rho = [\eta_1, \ldots, \eta_n]^\top$ and $\zeta = [\eta_{\rho+1}, \ldots, \eta_n]^\top$. Then $\eta_1 = y$, and the dynamics of estimation error, the zero-dynamics of ζ, and the lower triangular form for the dynamics of $\eta_1, \ldots, \eta_\rho$ can be expressed as follows:

$$\begin{aligned}
\mathrm{d}\widetilde{x} &= A\widetilde{x}\,\mathrm{d}t + h(y)\,\mathrm{d}w, \\
\mathrm{d}\zeta &= \begin{bmatrix} -\frac{b_{m-1}}{b_m} & & \\ \vdots & I_{m-1} & \\ -\frac{b_1}{b_m} & & \\ -\frac{b_0}{b_m} & 0 \;\cdots & 0 \end{bmatrix} \zeta\,\mathrm{d}t + \begin{bmatrix} d_{\rho,\rho+1} \\ \vdots \\ d_{\rho n} \end{bmatrix} y\,\mathrm{d}t \\
&\quad + \begin{bmatrix} -d_{\rho-1,\rho} \\ \vdots \\ -d_{\rho-1,n-1} \end{bmatrix} \widetilde{x}_2\,\mathrm{d}t + \begin{bmatrix} g_{\rho,\rho+1}(y,\widetilde{x}_1) \\ \vdots \\ g_{\rho n}(y,\widetilde{x}_1) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \widehat{h}_{\rho+1}(y) \\ \vdots \\ \widehat{h}_n(y) \end{bmatrix} \mathrm{d}w \\
&\triangleq E\zeta\,\mathrm{d}t + L_1\widetilde{x}_1\,\mathrm{d}t + L_2\widetilde{x}_2\,\mathrm{d}t + G(y)\,\mathrm{d}t + \Psi(y)\,\mathrm{d}w, \\
\mathrm{d}\eta_1 &= d_{\rho 1} y\,\mathrm{d}t + (\eta_2 + \widetilde{x}_2)\,\mathrm{d}t + g_{\rho 1}(y,\widetilde{x}_1)\,\mathrm{d}t + \widehat{h}_1(y)\,\mathrm{d}w, \\
\mathrm{d}\eta_2 &= [d_{\rho 2} y + \eta_3 + g_{\rho 2}(y,\widetilde{x}_1) - d_{\rho-1,1}\widetilde{x}_2]\,\mathrm{d}t + \widehat{h}_2(y)\,\mathrm{d}w, \\
&\;\;\vdots \\
\mathrm{d}\eta_{\rho-1} &= [d_{\rho,\rho-1} y + \eta_\rho + g_{\rho,\rho-1}(y,\widetilde{x}_1) - d_{\rho-1,\rho-2}\widetilde{x}_2]\,\mathrm{d}t + \widehat{h}_{\rho-1}(y)\,\mathrm{d}w, \\
\mathrm{d}\eta_\rho &= \left[1,\, \mathbf{0}_{1\times(m-1)}\right]\zeta\,\mathrm{d}t + d_{\rho\rho} y\,\mathrm{d}t + g_{\rho\rho}(y,\widetilde{x}_1)\,\mathrm{d}t \\
&\quad + b_m g(y) u\,\mathrm{d}t - d_{\rho-1,\rho-1}\widetilde{x}_2\,\mathrm{d}t + \widehat{h}_\rho(y)\,\mathrm{d}w,
\end{aligned} \tag{4.6}$$

where

$$\begin{aligned}
y &= \eta_1, \\
L_1 &= [\overline{d}_{\rho,\rho+1}, \ldots, \overline{d}_{\rho n}]^\top, \\
G &= [\overline{g}_{\rho,\rho+1}(y) + d_{\rho,\rho+1} y, \ldots, \overline{g}_{\rho,n}(y) + d_{\rho n} y]^\top.
\end{aligned}$$

This system is equivalent to the entire system (4.3) under the transformation $[\eta_1, \ldots, \eta_\rho, \zeta^\top]^\top = T_\rho \cdots T_1 [y,\, \widehat{x}_2, \ldots, \widehat{x}_n]^\top$. The structure of (4.6) allows the design of an output feedback controller by using integrator backstepping methodology.

4.3. Control design procedure. We now start to design the desired controller with the estimation error $\widetilde{x}$ and the zero-dynamics ζ (given by (4.5) for the case of

$\rho = 1$ and (4.6) for the case of $\rho > 1$). To do so, let $\chi = [\zeta^\top,\ \widetilde{x}^\top]^\top \in \mathbb{R}^{n+m}$. Then for both $\rho = 1$ and $\rho > 1$ we have

$$\text{(4.7)} \qquad \mathrm{d}\chi = \begin{bmatrix} E & L \\ \mathbf{0}_{n\times m} & A \end{bmatrix} \chi\,\mathrm{d}t + \begin{bmatrix} G(y) \\ \mathbf{0}_{n\times 1} \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \Psi(y) \\ h(y) \end{bmatrix} \mathrm{d}w$$
$$\triangleq W\chi\,\mathrm{d}t + F(y)\,\mathrm{d}t + \Phi(y)\,\mathrm{d}w,$$

where $L = [L_1,\ L_2,\ \mathbf{0}_{m\times(n-2)}]$, and F and Φ are $\mathcal{C}^\infty$.

For the objective of a tight controller, the dynamics χ would be partitioned as $\chi = [\chi_a^\top,\ \chi_b^\top]^\top$, where $\chi_a = [\zeta^\top,\ \widetilde{x}_1]^\top \in \mathbb{R}^{m+1}$ is available for feedback design, while $\chi_b = [\widetilde{x}_2, \ldots, \widetilde{x}_n]^\top \in \mathbb{R}^{n-1}$ is not. Furthermore, χ_a and χ_b satisfy the following stochastic differential equations, respectively:

$$\text{(4.8a)} \qquad \mathrm{d}\chi_a = \begin{bmatrix} E & L_1 \\ \mathbf{0}_{1\times m} & -k_1 \end{bmatrix} \chi_a\,\mathrm{d}t + \begin{bmatrix} L_2 & \mathbf{0}_{m\times(n-2)} \\ 1 & \mathbf{0}_{1\times(n-2)} \end{bmatrix} \chi_b\,\mathrm{d}t$$
$$+ \begin{bmatrix} G(y) \\ 0 \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \Psi(y) \\ h_1(y) \end{bmatrix} \mathrm{d}w$$
$$\triangleq W_a\chi_a\,\mathrm{d}t + L_a\chi_b\,\mathrm{d}t + F_a(y)\,\mathrm{d}t + \Phi_a(y)\,\mathrm{d}w,$$

$$\text{(4.8b)} \qquad \mathrm{d}\chi_b = \begin{bmatrix} 0 & & & \\ \vdots & & I_{n-2} & \\ 0 & 0 & \cdots & 0 \end{bmatrix} \chi_b\,\mathrm{d}t + \begin{bmatrix} 0 & \cdots & 0 & -k_2 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -k_n \end{bmatrix} \chi_a\,\mathrm{d}t$$
$$+ \begin{bmatrix} h_2(y) \\ \vdots \\ h_n(y) \end{bmatrix} \mathrm{d}w$$
$$\triangleq W_b\chi_b\,\mathrm{d}t + L_b\chi_a\,\mathrm{d}t + \Phi_b(y)\,\mathrm{d}w,$$

where F_a, Φ_a, and Φ_b are $\mathcal{C}^\infty$.

Remark 4.1. From subsection 4.2 we know that E, G, L, and Ψ in (4.7)–(4.8) are differently defined with respect to $\rho = 1$ and $\rho > 1$, respectively, and so are W, W_a, W_b, F, F_a, F_b, L_a, L_b, and Φ_a. Thus, for the sake of the unambiguousness, these two cases will be separately handled below.

We are now in a position to develop a recursive construction procedure for the desired risk-sensitive controller.

4.3.1. Initial assignment. First, we present the initial assignment for the entire design procedure.

By assumption A2, we know that matrix E is Hurwitz. This, together with the Hurwitz property of matrix A, implies that W is also Hurwitz. Therefore, there exists a symmetric and positive definite matrix P such that

$$\text{(4.9)} \qquad W^\top P + PW = -I_{n+m}.$$

We introduce a value function (or *Lyapunov* function) for the χ system:

$$\text{(4.10)} \qquad V_0(\chi) = \phi(\xi(\chi)) = \delta(c + \xi(\chi))^\gamma - \delta c^\gamma, \quad \xi = \chi^\top P\chi,$$

where $0 < \delta \le 1$, $c > 0$, $\frac{1}{2} < \gamma < 1$. Design constants c and δ will be specified later. Constant γ is pregiven and called the *characteristic parameter* of value function V_0. Clearly, $V_0(\chi)$ is positive definite and radially unbounded, and it vanishes at the origin $\chi = \mathbf{0}_{(n+m)\times 1}$.

Remark 4.2. Risk-sensitive control is much different from stochastic stabilization, and thus the methods developed by [5], [6], and [35] are not suitable for our control objective. Therefore, here we introduce a subquadratic function V_0 characterized by γ (see (4.10)), by which a method suitable for output-feedback risk-sensitive control design is developed.

Let $z_1 = y - y_d$ be the tracking error. Then, by assumptions A1 and A3, there exist a vector-valued smooth function $\overline{F}(y_d, z_1)$ and a matrix-valued smooth function $\overline{\Phi}(y_d, z_1)$ such that

$$F(y) = F(z_1 + y_d) = F(y_d) + z_1\overline{F}(y_d, z_1), \tag{4.11a}$$

$$\Phi(y) = \Phi(z_1 + y_d) = \Phi(y_d) + z_1\overline{\Phi}(y_d, z_1). \tag{4.11b}$$

LEMMA 4.1. *There exist a smooth vector-valued function $\sigma_0(\chi, y)$, a smooth function $N_0(y_d, \chi_a, z_1)$, and smooth $r_0(y_d)$, $C_0(y_d)$ such that*

$$\mathrm{d}V_0 \le \sigma_0\,\mathrm{d}w - \frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t - \frac{r_0\,\|\chi\|^2}{(c + \chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t + N_0 z_1\,\mathrm{d}t + C_0\,\mathrm{d}t. \tag{4.12}$$

Proof. By (4.7) and the Itô formula we have

$$\begin{aligned}\mathrm{d}V_0 = {} & -\frac{\partial\phi}{\partial\xi}\|\chi\|^2\,\mathrm{d}t + \sigma_0(y, \chi)\,\mathrm{d}w - \frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t + \frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t \\ & + 2\frac{\partial\phi}{\partial\xi}\chi^\top PF(y)\,\mathrm{d}t + \frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2 V_0}{\partial\chi^2}\Phi(y)(\Phi(y))^\top\right)\mathrm{d}t,\end{aligned} \tag{4.13}$$

where $\sigma_0 = \frac{\partial V_0}{\partial\chi}\Phi$ is a row vector-valued function. In the above equality, we have used the technique of subtracting from and adding term $\frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t$ to its right-hand side. Notice that $\frac{\partial V_0}{\partial\chi} = \frac{\partial\phi}{\partial\xi}\cdot\frac{\partial\xi}{\partial\chi}$ and $\frac{\partial\phi}{\partial\xi} = \frac{\delta\gamma}{(c+\xi)^{1-\gamma}}$. Then we have

$$\sigma_0 = \frac{\partial\phi}{\partial\xi}\cdot\frac{\partial\xi}{\partial\chi}\Phi = \frac{2\delta\gamma}{(c+\xi)^{1-\gamma}}\chi^\top P\Phi(y). \tag{4.14}$$

Let $\xi_a = \chi_a^\top(P_1 - P_2P_3^{-1}P_2^\top)\chi_a$, where $P = \begin{bmatrix} P_1 & P_2 \\ P_2^\top & P_3\end{bmatrix}$, $P_1 \in \mathbb{R}^{(m+1)\times(m+1)}$, $P_2 \in \mathbb{R}^{(m+1)\times(n-1)}$, $P_3 \in \mathbb{R}^{(n-1)\times(n-1)}$. Clearly, since P is positive definite, so is $P_1 - P_2P_3^{-1}P_2^\top$. Then ξ_a is available for feedback design and satisfies $0 \le \xi_a \le \xi$. For the first term of the second line on the right-hand side of (4.13), by using (4.11a), we have

$$\begin{aligned} 2\frac{\partial\phi}{\partial\xi}\chi^\top PF(y) &= 2\frac{\partial\phi}{\partial\xi}\chi^\top P(F(y_d) + \overline{F}(y_d, z_1)z_1) \\ &= \frac{2\delta\gamma\chi^\top PF(y_d)}{(c+\xi)^{1-\gamma}} + \frac{2\delta\gamma\chi^\top P\overline{F}(y_d, z_1)}{(c+\xi)^{1-\gamma}}z_1 \\ &\le \frac{\delta\gamma\varepsilon_{01}^2\|\chi\|^2}{(c+\xi)^{1-\gamma}} + \frac{\delta\gamma\|PF(y_d)\|^2}{\varepsilon_{01}^2(c+\xi)^{1-\gamma}} + \frac{\delta\gamma\|P\overline{F}(y_d, z_1)\|^2}{\varepsilon_{02}^2(c+\xi)^{1-\gamma}}z_1^2 + \frac{\delta\gamma\varepsilon_{02}^2\|\chi\|^2}{(c+\xi)^{1-\gamma}} \\ &\le \frac{\delta\gamma(\varepsilon_{01}^2 + \varepsilon_{02}^2)\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}} + \frac{\delta\gamma\|PF(y_d)\|^2}{\varepsilon_{01}^2 c^{1-\gamma}} + \frac{\delta\gamma\|P\overline{F}(y_d, z_1)\|^2}{\varepsilon_{02}^2(c+\xi_a)^{1-\gamma}}z_1^2. \end{aligned} \tag{4.15}$$

Here and hereafter, ε_{01}, ε_{02}, ε_{03}, and ε_{04} are positive design constants to be determined later.

For the term $\frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t$ on the right-hand side of (4.13), by (4.11b) we have

$$
\begin{aligned}
\frac{\theta}{4}\sigma_0\sigma_0^\top &= \frac{\theta\delta^2\gamma^2\chi^\top P\Phi(y)(\Phi(y))^\top P\chi}{(c+\chi^\top P\chi)^{2-2\gamma}} \\
&= \frac{\theta\delta^2\gamma^2\chi^\top P\left(\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)\cdot\left(\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)^\top P\chi}{(c+\chi^\top P\chi)^{2-2\gamma}} \\
&= \frac{\theta\delta^2\gamma^2\chi^\top P\Phi(y_d)(\Phi(y_d))^\top P\chi}{(c+\chi^\top P\chi)^{2-2\gamma}} \\
&\quad+\frac{\theta\delta^2\gamma^2\left[\chi_a^\top,\,\chi_b^\top\right]P\left(2\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)(\overline{\Phi}(y_d,\,z_1))^\top P[\chi_a^\top,\,\chi_b^\top]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1 \\
&= \frac{\theta\delta^2\gamma^2\chi^\top P\Phi(y_d)(\Phi(y_d))^\top P\chi}{(c+\chi^\top P\chi)^{2-2\gamma}} \\
&\quad+\frac{\theta\delta^2\gamma^2\left[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}\right]P\overline{\Phi}(y_d,\,z_1)(\overline{\Phi}(y_d,\,z_1))^\top P\left[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}\right]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1^2 \\
&\quad+\frac{\theta\delta^2\gamma^2\chi^\top P\left(2\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)(\overline{\Phi}(y_d,\,z_1))^\top P\left[\mathbf{0}_{1\times(m+1)},\,\chi_b^\top\right]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1 \\
&\quad+\frac{2\theta\delta^2\gamma^2\chi^\top P\Phi(y_d)(\overline{\Phi}(y_d,\,z_1))^\top P\left[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}\right]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1 \\
&\quad+\frac{\theta\delta^2\gamma^2\left[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}\right]P\overline{\Phi}(y_d,\,z_1)(\overline{\Phi}(y_d,\,z_1))^\top P[\mathbf{0}_{1\times(m+1)},\,\chi_b^\top]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1^2 \\
&\le \frac{\theta\delta^2\gamma^2\|P\Phi(y_d)\|_F^2}{c^{1-\gamma}}\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}} \\
&\quad+\frac{\theta\delta^2\gamma^2\left[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}\right]P\overline{\Phi}(y_d,\,z_1)(\overline{\Phi}(y_d,\,z_1))^\top P\left[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}\right]^\top}{(c+\xi_a)^{2-2\gamma}}z_1^2 \\
&\quad+\frac{\theta\delta^2\gamma^2\|P\|^2\left\|\left(2\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)(\overline{\Phi}(y_d,\,z_1))^\top\right\|_F^p}{p\varepsilon_{03}^p(c+\xi_a)^{p-\gamma(p+1)}}z_1^p \\
&\quad+\frac{(p-1)\theta\delta^2\gamma^2\|P\|^2\varepsilon_{03}^{\frac{p}{p-1}}}{p\lambda_{\min}^{\frac{1}{p-1}}(P)}\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}} \\
&\quad+\frac{\theta\delta^2\gamma^2\|P\|^4}{\varepsilon_{04}^2(c+\xi_a)^{3-3\gamma}}\|\chi_a\|^2\left(\|\Phi(y_d)\|_F^2+\frac{z_1^2}{2}\|\overline{\Phi}(y_d,\,z_1)\|_F^2\right)\|\overline{\Phi}(y_d,\,z_1)\|_F^2z_1^2 \\
&\quad+\frac{3\theta\delta^2\gamma^2\varepsilon_{04}^2\|\chi\|^2}{2(c+\chi^\top P\chi)^{1-\gamma}},
\end{aligned}
\tag{4.16}
$$

where p is a positive even integer (that is, it takes values in set $\{2,4,6,8,\ldots\}$) and satisfies the inequality $p\ge\frac{\gamma}{1-\gamma}$ (or $\gamma\le\frac{p}{p+1}$). Let $q=\frac{p}{p-1}$. Then, p and q satisfy $\frac{1}{p}+\frac{1}{q}=1$. In the inequality (4.16), we have used the Young's inequality

$$
x^\top y\le\frac{\varepsilon^pW^p(x,\,y)}{p}\|x\|^p+\frac{1}{q\varepsilon^qW^q(x,\,y)}\|y\|^q\qquad\forall\,x,y\in\mathbb{R}^n,\,\varepsilon>0,\,W(x,\,y)>0,
$$

to get

$$\begin{aligned}
&\frac{\theta\delta^2\gamma^2\chi^\top P\left(2\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)\left(\overline{\Phi}(y_d,\,z_1)\right)^\top P[\mathbf{0}_{1\times(m+1)},\,\chi_b^\top]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1\\
&\le \frac{\theta\delta^2\gamma^2\|P\|^2W^p(\chi)\left\|\left(2\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)\left(\overline{\Phi}(y_d,\,z_1)\right)^\top\right\|^p}{p\varepsilon_{03}^p(c+\xi)^{2p(1-\gamma)}}z_1^p\\
&\quad+\frac{\theta\delta^2\gamma^2\|P\|^2\varepsilon_{03}^q\|\chi\|^{2q}}{qW^q(\chi)}\\
&\le \frac{\theta\delta^2\gamma^2\|P\|^2\left\|\left(2\Phi(y_d)+z_1\overline{\Phi}(y_d,\,z_1)\right)\left(\overline{\Phi}(y_d,\,z_1)\right)^\top\right\|_F^p}{p\varepsilon_{03}^p(c+\xi_a)^{p-\gamma(p+1)}}z_1^p\\
&\quad+\frac{(p-1)\theta\delta^2\gamma^2\|P\|^2\varepsilon_{03}^{\frac{p}{p-1}}}{p\lambda_{\min}^{\frac{1}{p-1}}(P)}\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}},
\end{aligned}$$

with $W(\chi)=(c+\xi)^{1+\frac{\gamma}{p}-\gamma}$, and

$$\begin{aligned}
&\frac{2\theta\delta^2\gamma^2\chi^\top P\Phi(y_d)(\overline{\Phi}(y_d,\,z_1))^\top P\left[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}\right]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1\\
&\le \frac{\theta\delta^2\gamma^2\left(\frac{1}{\varepsilon_{04}^2(W(\chi))^2}\left\|P\Phi(y_d)(\overline{\Phi}(y_d,z_1))^\top P[\chi_a^\top,\mathbf{0}_{1\times(n-1)}]^\top\right\|^2z_1^2+\varepsilon_{04}^2(W(\chi))^2\|\chi\|^2\right)}{(c+\chi^\top P\chi)^{2-2\gamma}}\\
&\le \frac{\theta\delta^2\gamma^2\|P\|^4}{\varepsilon_{04}^2(c+\xi_a)^{3-3\gamma}}\|\chi_a\|^2\left\|\Phi(y_d)\right\|_F^2\left\|\overline{\Phi}(y_d,\,z_1)\right\|_F^2z_1^2+\frac{\theta\delta^2\gamma^2\varepsilon_{04}^2\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}},
\end{aligned}$$

$$\begin{aligned}
&\frac{\theta\delta^2\gamma^2[\chi_a^\top,\,\mathbf{0}_{1\times(n-1)}]P\overline{\Phi}(y_d,\,z_1)(\overline{\Phi}(y_d,\,z_1))^\top P[\mathbf{0}_{1\times(m+1)},\,\chi_b^\top]^\top}{(c+\chi^\top P\chi)^{2-2\gamma}}z_1^2\\
&\le \frac{\theta\delta^2\gamma^2\left(\frac{1}{\varepsilon_{04}^2W^2(\chi)}\left\|[\chi_a^\top,\mathbf{0}_{1\times(n-1)}]P\overline{\Phi}(y_d,z_1)(\overline{\Phi}(y_d,z_1))^\top P\right\|^2z_1^4+\varepsilon_{04}^2W^2(\chi)\|\chi_b\|^2\right)}{2(c+\chi^\top P\chi)^{2-2\gamma}}\\
&\le \frac{\theta\delta^2\gamma^2\|P\|^4}{2\varepsilon_{04}^2(c+\xi_a)^{3-3\gamma}}\|\chi_a\|^2\left\|\overline{\Phi}(y_d,\,z_1)\right\|_F^4z_1^4+\frac{\theta\delta^2\gamma^2\varepsilon_{04}^2\|\chi\|^2}{2(c+\chi^\top P\chi)^{1-\gamma}},
\end{aligned}$$

with $W(\chi)=(c+\xi)^{\frac{1}{2}-\frac{1}{2}\gamma}$.

For the last term on the right-hand side of (4.13), by (4.11b), we have

$$\begin{aligned}
(4.17)\qquad &\frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2V_0}{\partial\chi^2}\Phi\Phi^\top\right)\\
&=\mathrm{Tr}\left(\left(\frac{\delta\gamma P}{(c+\chi^\top P\chi)^{1-\gamma}}-\frac{2\delta\gamma(1-\gamma)P\chi\chi^\top P}{(c+\chi^\top P\chi)^{2-\gamma}}\right)\Phi\Phi^\top\right)\\
&\le \frac{2\delta\gamma}{c^{1-\gamma}}\mathrm{Tr}\left((\Phi(y_d))^\top P\Phi(y_d)\right)+\frac{2\delta\gamma\mathrm{Tr}\left((\overline{\Phi}(y_d,\,z_1))^\top P\overline{\Phi}(y_d,\,z_1)\right)}{(c+\xi_a)^{1-\gamma}}z_1^2.
\end{aligned}$$

Substituting (4.15), (4.16), and (4.17) into (4.13), we get (4.12) with

$$r_0 = \delta\gamma(1-\varepsilon_{01}^2-\varepsilon_{02}^2) - \theta\delta^2\gamma^2\left(\frac{3}{2}\varepsilon_{04}^2 + \frac{\|P\Phi(y_d)\|_F^2}{c^{1-\gamma}} + \frac{(p-1)\varepsilon_{03}^{\frac{p}{p-1}}\|P\|^2}{p\lambda_{\min}^{\frac{1}{p-1}}(P)}\right), \tag{4.18}$$

$$\begin{aligned}N_0 = {} & \frac{\delta\gamma\left\|P\overline{F}(y_d,\, z_1)\right\|^2}{\varepsilon_{02}^2(c+\xi_a)^{1-\gamma}} z_1 \\ & + \frac{\theta\delta^2\gamma^2\left[\chi_a^\top,\, \mathbf{0}_{1\times(n-1)}\right] P\overline{\Phi}(y_d,\, z_1)(\overline{\Phi}(y_d,\, z_1))^\top P\left[\chi_a^\top,\, \mathbf{0}_{1\times(n-1)}\right]^\top}{(c+\xi_a)^{2-2\gamma}} z_1 \\ & + \frac{\theta\delta^2\gamma^2\|P\|^2\left\|\left(2\Phi(y_d)+z_1\overline{\Phi}(y_d,\, z_1)\right)\left(\overline{\Phi}(y_d,\, z_1)\right)^\top\right\|_F^p}{p\varepsilon_{03}^p(c+\xi_a)^{p-\gamma(p+1)}} z_1^{p-1} \\ & + \frac{\theta\delta^2\gamma^2\|P\|^4}{\varepsilon_{04}^2(c+\xi_a)^{3-3\gamma}}\|\chi_a\|^2\left(\|\Phi(y_d)\|_F^2 + \frac{z_1^2}{2}\|\overline{\Phi}(y_d,\, z_1)\|_F^2\right)\|\overline{\Phi}(y_d,\, z_1)\|_F^2 z_1 \\ & + \frac{2\delta\gamma\mathrm{Tr}\left((\overline{\Phi}(y_d,\, z_1))^\top P\overline{\Phi}(y_d,\, z_1)\right)}{(c+\xi_a)^{1-\gamma}} z_1,\end{aligned}$$

$$C_0 = \frac{\delta\gamma}{c^{1-\gamma}}\left(\frac{\left\|PF(y_d)\right\|^2}{\varepsilon_{01}^2} + 2\mathrm{Tr}\left(P\Phi(y_d)(\Phi(y_d))^\top\right)\right). \qquad \square \tag{4.19}$$

The control design procedure will be presented for the two cases of $\rho = 1$ and $\rho > 1$ separately in subsections 4.3.2 and 4.3.3 below.

4.3.2. Control design for the case of $\rho = 1$. Let us now present the control design for the system (3.1) with $\rho = 1$. From (4.5) and (4.7) we obtain the following overall systems:

$$\mathrm{d}\chi = W\chi\,\mathrm{d}t + F(y)\,\mathrm{d}t + \Phi(y)\,\mathrm{d}w, \tag{4.20a}$$

$$\mathrm{d}\eta_1 = b_{n-1}g(y)u\,\mathrm{d}t + d_1\chi_b\,\mathrm{d}t + g_1(y,\, \chi_a)\,\mathrm{d}t + \widehat{h}_1(y)\,\mathrm{d}w, \tag{4.20b}$$

where

$$\begin{aligned} y &= \eta_1, \\ d_1 &= [1,\ \mathbf{0}_{1\times(n-2)}], \\ g_1 &= \overline{g}_{11}(y) + \overline{d}_{11}[\mathbf{0}_{1\times(n-1)},\ 1]\chi_a + d_{11}y + \left[1,\ \mathbf{0}_{1\times(n-1)}\right]\chi_a. \end{aligned}$$

It is easy to check that g_1 and $\widehat{h}_1$ are $\mathcal{C}^\infty$.

Let $\alpha_1 = b_{n-1}g(y)u$, $S_1 = d_1$, $F(y_d^{[1]},\, \chi_1,\, \eta_1) = g_1(y,\, \chi_a) - \dot{y}_d$, and $\Psi_1(y) = \widehat{h}_1(y)$. Then, by (4.20), we have the dynamics of tracking error $z_1 = \eta_1 - y_d$:

$$\mathrm{d}z_1 = (\alpha_1 + S_1\chi_b + F_1)\,\mathrm{d}t + \Psi_1\,\mathrm{d}w. \tag{4.21}$$

Let $V_1 = V_0 + \Xi_1(y_d)z_1^2$, where V_0 is defined by (4.10) and Ξ_1 is to be specified later.

Then, by (4.12) and (4.21), we have

(4.22)

$$
\begin{aligned}
\mathrm{d}V_1 &= \mathrm{d}V_0 + 2z_1\Xi_1(\alpha_1 + F_1 + S_1\chi_b)\,\mathrm{d}t + \frac{\partial \Xi_1}{\partial t} z_1^2\,\mathrm{d}t + \Xi_1\Psi_1\Psi_1^\top\,\mathrm{d}t + 2\Xi_1\Psi_1 z_1\,\mathrm{d}w \\
&\le \sigma_1(y_d,\, \chi,\, z_1)\,\mathrm{d}w - \frac{\theta}{4}\sigma_1\sigma_1^\top\,\mathrm{d}t - \beta_1\Xi_1 z_1^2\,\mathrm{d}t + \beta_1\Xi_1 z_1^2\,\mathrm{d}t - r_0(y_d)\frac{\|\chi\|^2}{(c+\xi)^{1-\gamma}}\,\mathrm{d}t \\
&\quad + 2z_1\Xi_1(y_d)(\alpha_1 + F_1)\,\mathrm{d}t + \frac{\partial \Xi_1}{\partial t} z_1^2\,\mathrm{d}t + N_0(y_d,\, \chi_a,\, z_1)z_1\,\mathrm{d}t \\
&\quad + \frac{\theta}{4}\sigma_1\sigma_1^\top\,\mathrm{d}t - \frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t + M_1\chi_b z_1\,\mathrm{d}t + \Xi_1\Psi_1\Psi_1^\top\,\mathrm{d}t + C_0(y_d)\,\mathrm{d}t,
\end{aligned}
$$

where $M_1 = 2\Xi_1 S_1$, $\sigma_1 = \sigma_0 + 2\Xi_1\Psi_1 z_1$. In the above inequality, we have used the technique of subtracting from and adding to its right-hand side the terms $\frac{\theta}{4}\sigma_1\sigma_1^\top\,\mathrm{d}t$ and $\beta_1\Xi_1 z_1^2\,\mathrm{d}t$. Here and hereafter, $\beta_1, \beta_1, \ldots, \beta_\rho$ are positive design constants to be determined.

Since $\sigma_0 = \frac{\partial V_0}{\partial \chi}\Phi(y)$ is unavailable for feedback design, so is the term $\frac{\theta}{4}\sigma_1\sigma_1^\top\,\mathrm{d}t - \frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t$ on the right-hand side of (4.22). Therefore, we give the following estimate:

(4.23)

$$
\begin{aligned}
\frac{\theta}{4}(\sigma_1\sigma_1^\top - \sigma_0\sigma_0^\top) &= \theta\Xi_1\sigma_0\Psi_1^\top z_1 + \theta\Xi_1^2\Psi_1\Psi_1^\top z_1^2 \\
&= \theta\Xi_1^2\Psi_1\Psi_1^\top z_1^2 + \theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial \chi}\right)^\top\right\|^2 \\
&\quad + \frac{\theta}{4\varepsilon_{11}^2}\Xi_1^2\Psi_1\Phi^\top\Phi\Psi_1^\top z_1^2 - \theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial \chi}\right)^\top - \frac{\Xi_1}{2\varepsilon_{11}^2}\Phi\Psi_1^\top z_1\right\|^2 \\
&\le -\theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial \chi}\right)^\top - \frac{\Xi_1}{2\varepsilon_{11}^2}\Phi\Psi_1^\top z_1\right\|^2 + \frac{4\theta\delta^2\gamma^2\varepsilon_{11}^2\|P\|^2}{c^{1-\gamma}} \\
&\quad \cdot \frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}} + \theta\Xi_1^2\left(\Psi_1\Psi_1^\top + \frac{1}{4\varepsilon_{11}^2}\Psi_1\Phi^\top\Phi\Psi_1^\top\right)z_1^2,
\end{aligned}
$$

where (and whereafter) ε_{11} and ε_1 are positive design constants to be specified.

Define

$$
\Delta_{11}(y_d, \chi, z_1) = \theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial \chi}\right)^\top - \frac{\Xi_1}{2\varepsilon_{11}^2}\Phi\Psi_1^\top z_1\right\|^2, \tag{4.24a}
$$

$$
\Delta_{12}(y_d, \chi, z_1) = \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\|\chi_b\|^{\frac{p_1}{p_1-1}} + \frac{\|M_1^\top\|^{p_1}}{p_1\varepsilon_1^{p_1}}z_1^{p_1} - M_1\chi_b z_1. \tag{4.24b}
$$

Clearly, $\Delta_{11} \ge 0$. Thus also, by Young's inequality, it is easy to see that $\Delta_{12} \ge 0$.

If p_1 takes values in set $\{4,\, 6,\, 8,\, 10, \ldots\}$ and satisfies $p_1 \ge \frac{2\gamma}{2\gamma-1}$, then we can give an upper bound for "$M_1\chi_b z_1$":

$$
\begin{aligned}
(4.25)\qquad M_1\chi_b z_1 &= -\Delta_{12} + \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\|\chi_b\|^{\frac{p_1}{p_1-1}} + \frac{\|M_1^\top\|^{p_1}}{p_1\varepsilon_1^{p_1}} z_1^{p_1}\\
&\le -\Delta_{12} + \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\lambda_{\max}^{1-\gamma}(P)\left(\mathcal{M}_\gamma(c) + \frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\right)\\
&\quad + \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\mathcal{K}\left(\frac{p_1}{p_1-1}, 2\gamma\right) + \frac{\|M_1^\top\|^{p_1}}{p_1\varepsilon_1^{p_1}} z_1^{p_1},
\end{aligned}
$$

where $\mathcal{K}(a_1, a_2)$ is defined in Lemma B.3.

Since there exist smooth functions $\overline{\Psi}_1(y_d, z_1)$ and $\overline{\Psi}_{11}(y_d, z_1)$ satisfying

$$\Psi_1 = \overline{\Psi}_1(y_d, z_1) = \overline{\Psi}_1(y_d) + z_1\overline{\Psi}_{11}(y_d, z_1),$$

for the fourth term of the last line on the right-hand side of (4.22) we have

$$
\begin{aligned}
(4.26)\qquad \Xi_1\Psi_1\Psi_1^\top &= \Xi_1\|(\overline{\Psi}_1(y_d) + z_1\overline{\Psi}_{11}(y_d, z_1))^\top\|^2\\
&\le 2\Xi_1\|(\overline{\Psi}_1(y_d))^\top\|^2 + 2\Xi_1\|(\overline{\Psi}_{11}(y_d, z_1))^\top\|^2 z_1^2.
\end{aligned}
$$

Choose

$$(4.27)\qquad \Xi_1 = \frac{\kappa_1}{1+\|(\overline{\Psi}_1(y_d))^\top\|^2},$$

where (and whereafter) $\kappa_1, \kappa_2, \ldots, \kappa_\rho$ are positive design constants to be determined.

Thus, by substituting (4.23), (4.25), (4.26), and (4.27) into (4.22), and via some straightforward calculations, we get

$$
\begin{aligned}
(4.28)\qquad \mathrm{d}V_1 &\le -z_1^2\,\mathrm{d}t + \sigma_1\,\mathrm{d}w - \frac{\theta}{4}\sigma_1\sigma_1^\top\,\mathrm{d}t - \frac{r_1(y_d)\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t\\
&\quad - \Xi_1\beta_1 z_1^2\,\mathrm{d}t + 2z_1\Xi_1\left(\alpha_1 - \overline{\alpha}_1(y_d^{[1]}, \chi_a, \eta_1)\right)\mathrm{d}t\\
&\quad - \Delta_1(y_d, \chi, z_1)\,\mathrm{d}t + 2\Xi_1 z_1 z_2\,\mathrm{d}t + C_1(y_d^{[1]})\,\mathrm{d}t,
\end{aligned}
$$

where

$$(4.29)\qquad r_1 = r_0(y_d) - \frac{4\theta\delta^2\gamma^2\varepsilon_{11}^2\|P\|^2}{c^{1-\gamma}} - \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\lambda_{\max}^{1-\gamma}(P),$$

$$
\begin{aligned}
(4.30)\qquad N_1 &= F_1 + \frac{\beta_1 z_1}{2} + \frac{z_1}{2\Xi_1}\frac{\partial\Xi_1}{\partial y_d}\dot{y}_d + \frac{\|M_1^\top\|^{p_1}}{2p_1\Xi_1\varepsilon_1^{p_1}} z_1^{p_1-1}\\
&\quad + \theta\Xi_1\left(\frac{\Psi_1\Psi_1^\top}{2} + \frac{\Psi_1\Phi^\top\Phi\Psi_1^\top}{8\varepsilon_{11}^2}\right) z_1 + \|(\overline{\Psi}_{11}(y_d, z_1))^\top\|^2 z_1,
\end{aligned}
$$

$$(4.31)\qquad \overline{\alpha}_1 = \left.\left\{-N_1 - \frac{z_1}{2\Xi_1} - \frac{N_0}{2\Xi_1}\right\}\right|_{z_1=\eta_1-y_d},$$

$$\Delta_1 = \Delta_{11} + \Delta_{12}, \qquad \text{with } \Delta_{11} \text{ and } \Delta_{12} \text{ being defined by (4.24)},$$

$$
\begin{aligned}
(4.32)\qquad C_1 &= C_0(y_d) + \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\lambda_{\max}^{1-\gamma}(P)\mathcal{M}_\gamma(c) + 2\Xi_1\|(\overline{\Psi}_1(y_d))^\top\|^2\\
&\quad + \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\mathcal{K}\left(\frac{p_1}{p_1-1}, 2\gamma\right).
\end{aligned}
$$

It is easy to check that C_1, r_1, and $\overline{\alpha}_1$ are $\mathcal{C}^\infty$.

Thus, we can choose the function $\alpha_1(y_d^{[1]}, \chi_a, \eta_1)$ in the following form:

$$\alpha_1 = \overline{\alpha}_1(y_d^{[1]}, \chi_a, \eta_1). \tag{4.33}$$

From this and the definition of α_1, i.e., $\alpha_1 = b_{n-1}g(y)u$, we immediately obtain the following risk-sensitive controller:

$$u = \frac{\alpha_1}{b_{n-1}g(y)} = \frac{1}{b_{n-1}g(y)}\,\overline{\alpha}_1(y_d^{[1]}, \chi_a, \eta_1). \tag{4.34}$$

Then, by (4.28) and (4.33), we have

$$\begin{aligned} \mathrm{d}V_1 \le{} & -z_1^2\,\mathrm{d}t + \sigma_1\,\mathrm{d}w - \frac{\theta}{4}\sigma_1\sigma_1^\top\,\mathrm{d}t - \frac{r_1(y_d)\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t \\ & - \Xi_1\beta_1 z_1^2\,\mathrm{d}t - \Delta_1(y_d, \chi, z_1)\,\mathrm{d}t + C_1(y_d^{[1]})\,\mathrm{d}t. \end{aligned} \tag{4.35}$$

4.3.3. Control design for the case of $\rho > 1$. This subsection investigates the control design for the system (3.1) with $\rho > 1$. From the procedure addressed below, we know that the control design for this case is more complicated than that for the case of $\rho = 1$ given in subsection 4.3.2.

First, from (4.6) and (4.7) we obtain the following overall systems amenable for integrator backstepping design:

$$\begin{aligned} \mathrm{d}\chi &= W\chi\,\mathrm{d}t + F(y)\,\mathrm{d}t + \Phi(y)\,\mathrm{d}w, \\ \mathrm{d}\eta_1 &= \eta_2\,\mathrm{d}t + d_1\chi_b\,\mathrm{d}t + g_1(y, \chi_a)\,\mathrm{d}t + \widehat{h}_1(y)\,\mathrm{d}w, \\ \mathrm{d}\eta_2 &= \eta_3\,\mathrm{d}t + d_2\chi_b\,\mathrm{d}t + g_2(y, \chi_a)\,\mathrm{d}t + \widehat{h}_2(y)\,\mathrm{d}w, \\ &\;\;\vdots \\ \mathrm{d}\eta_{\rho-1} &= \eta_\rho\,\mathrm{d}t + d_{\rho-1}\chi_b\,\mathrm{d}t + g_{\rho-1}(y, \chi_a)\,\mathrm{d}t + \widehat{h}_{\rho-1}(y)\,\mathrm{d}w, \\ \mathrm{d}\eta_\rho &= b_m g(y)u\,\mathrm{d}t + d_\rho\chi_b\,\mathrm{d}t + g_\rho(y, \chi_a)\,\mathrm{d}t + \widehat{h}_\rho(y)\,\mathrm{d}w, \end{aligned} \tag{4.36}$$

where

$$\begin{aligned} & y = \eta_1, \\ & d_1 = [1, \mathbf{0}_{1\times(n-2)}], \quad d_i = [-d_{\rho-1,i-1}, \mathbf{0}_{1\times(n-2)}],\ i = 2,\ldots,\rho, \\ & g_i = g_{\rho i}\left(y, [\mathbf{0}_{1\times m}, 1]\chi_a\right) + d_{\rho i}y, \quad i = 1,\ldots,\rho-1, \\ & g_\rho = g_{\rho\rho}\left(y, [\mathbf{0}_{1\times m}, 1]\chi_a\right) + d_{\rho\rho}y + [1, \mathbf{0}_{1\times m}]\,\chi_a. \end{aligned}$$

It is easy to check that g_i, $\widehat{h}_i$, $i = 1,\ldots,\rho$, are $\mathcal{C}^\infty$.

Below is the backstepping design procedure, which involves ρ steps in all.

Step 1. Define variable $z_2 = \eta_2 - \alpha_1(y_d^{[1]}, \chi_a, \eta_1)$ and value function $V_1 = V_0 + \Xi_1(y_d)z_1^2$ for this step, where α_1 is a smooth function known as a virtual control law and Ξ_1 is a positive and smooth function. Both α_1 and Ξ_1 will be specified in this step.

From (4.36) it follows that

$$\mathrm{d}z_1 = (z_2 + \alpha_1 + F_1(y_d^{[1]}, \chi_a, \eta_1))\,\mathrm{d}t + S_1\chi_b\,\mathrm{d}t + \Psi_1(\eta_1)\,\mathrm{d}w, \tag{4.37}$$

where $F_1 = g_1(y, \chi_a) - \dot{y}_d$, $S_1 = d_1$, $\Psi_1 = \widehat{h}_1(y)$.

Clearly, (4.37) has the same structure as that of (4.21). Then, as in the case of $\rho = 1$, the virtual controller α_1 can be given by (4.33), which is such that

$$\begin{aligned}\text{(4.38)}\qquad \mathrm{d}V_1 \le{}& -z_1^2\,\mathrm{d}t + \sigma_1(y_d,\,\chi,\,z_1)\,\mathrm{d}w - \frac{\theta}{4}\sigma_1\sigma_1^\top\,\mathrm{d}t - \frac{r_1(y_d)\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t\\ &- \Xi_1\beta_1 z_1^2\,\mathrm{d}t - \Delta_1(y_d,\,\chi,\,z_1)\,\mathrm{d}t + 2\Xi_1 z_1 z_2\,\mathrm{d}t + C_1(y_d^{[1]})\,\mathrm{d}t,\end{aligned}$$

where σ_1, r_1, Ξ_1, Δ_1, C_1 are defined as in the case of $\rho = 1$.

This completes *Step* 1.

Step i ($i = 2, \dots, \rho-1$). Suppose that from step 1 through to step $i-1$ we have obtained $z_j = \eta_j - \alpha_{j-1}(y_d^{[j-1]}, \chi_a, \eta_{[j-1]})$, $j = 1, \dots, i$, and value function

$$V_{i-1} = V_0 + \sum_{j=1}^{i-1} \Xi_j(y_d^{[j-1]},\,\chi_a,\,z_{[j-1]})z_j^2$$

satisfying

$$\begin{aligned}\text{(4.39)}\qquad \mathrm{d}V_{i-1} \le{}& -z_1^2\,\mathrm{d}t + \sigma_{i-1}(y_d^{[i-2]},\,\chi,\,z_{[i-1]})\,\mathrm{d}w - \frac{\theta}{4}\sigma_{i-1}\sigma_{i-1}^\top\,\mathrm{d}t\\ &- r_{i-1}(y_d)\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t - \Delta_{i-1}(y_d^{[i-2]},\,\chi,\,z_{[i-1]})\,\mathrm{d}t\\ &- \sum_{j=1}^{i-1}\Xi_j\left(\beta_j - 2\sum_{m=j+1}^{i-1}(m-1)\,\kappa_m\right)z_j^2\,\mathrm{d}t\\ &+ 2\Xi_{i-1}z_{i-1}z_i\,\mathrm{d}t + C_{i-1}(y_d^{[i-1]},\,\chi_a,\,z_{[i-1]})\,\mathrm{d}t,\end{aligned}$$

where

$$\begin{aligned}\mathrm{d}z_1 &= (z_2 + \alpha_1 + F_1(y_d^{[1]},\,\chi_a,\,\eta_1))\mathrm{d}t + S_1\chi_b\mathrm{d}t + \Psi_1(\eta_1)\mathrm{d}w,\\ \mathrm{d}z_j &= (z_{j+1} + \alpha_j + F_j(y_d^{[j]},\,\chi_a,\,\eta_{[j]}))\mathrm{d}t + S_j(y_d^{[j-1]},\,\chi_a,\,\eta_{[j-1]})\chi_b\mathrm{d}t\\ &\quad + \Psi_j(y_d^{[j-1]},\,\chi_a,\,\eta_{[j-1]})\mathrm{d}w,\quad j = 2,\,3,\dots,i-1,\\ \sigma_{i-1} &= \sigma_0 + \sum_{j=1}^{i-1}\left(2\Xi_j z_j\Psi_j + z_j^2\frac{\partial\Xi_j}{\partial\chi_a}\Phi_a + z_j^2\sum_{k=1}^{j-1}\frac{\partial\Xi_j}{\partial z_k}\Psi_k\right),\end{aligned}$$

$$\text{(4.40)}\qquad r_{i-1} = r_1(y_d),$$

$$\text{(4.41)}\qquad C_{i-1} = C_1(y_d^{[1]}) + \sum_{j=2}^{i-1} 2\Xi_j\|(\overline{\Psi}_j(y_d^{[j-1]},\,\chi_a,\,\mathbf{0}_{(j-1)\times 1}))^\top\|^2,$$

and $\Delta_{i-1} = \Delta_{i-1,1}(y_d^{[i-2]},\,\chi,\,z_{[i-1]}) + \Delta_{i-1,2}(y_d^{[i-2]},\,\chi_b,\,z_{[i-1]})$. Here $\Delta_{i-1,1}$ and $\Delta_{i-1,2}$ are given as follows:

$$\Gamma_j(y_d^{[j-1]},\,\chi_a,\,z_{[j]}) = \Psi_j + \frac{z_j}{2\Xi_j}\frac{\partial\Xi_j}{\partial\chi_a}\Phi_a + \frac{z_j}{2\Xi_j}\sum_{k=1}^{j-1}\frac{\partial\Xi_j}{\partial z_k}\Psi_k,\quad j = 1,\dots,i-1,$$

$$\Delta_{i-1,1} = \theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial\chi}\right)^\top - \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Phi\Gamma_j^\top z_j\right\|^2,$$

$$M_k(y_d^{[k-1]},\,\chi_a,\,z_{[k]}) = 2\Xi_k S_k + z_k\sum_{j=1}^{k-1}\frac{\partial\Xi_k}{\partial z_j}S_j + z_k\frac{\partial\Xi_k}{\partial\chi_a}L_a,\quad 1\le k\le i-1,$$

$$\text{(4.42)}\qquad \Delta_{i-1,2} = \frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\|\chi_b\|^{\frac{p_1}{p_1-1}} + \frac{\left\|\sum_{k=1}^{i-1}M_k^\top z_k\right\|^{p_1}}{p_1\varepsilon_1^{p_1}} - \left(\sum_{k=1}^{i-1}M_k z_k\right)\chi_b.$$

It should be noted that Ξ_i, F_j, S_j, Γ_j, Ψ_j, M_j, $j = 1, \ldots, i-1$, α_j, $\overline{\alpha}_j$, $i = 0, 1, \ldots, i-1$, σ_{i-1}, r_{i-1}, C_{i-1} are $\mathcal{C}^\infty$. By Young's inequality, it is easy to see that $\Delta_{i-1,2} \geq 0$.

Let $\varrho_{i-1}(y_d^{[i-2]}, \chi_a, z_{[i-2]}, z_{i-1}) = \frac{1}{p_1 \varepsilon_1^{p_1}} \| \sum_{k=1}^{i-2} M_k^\top z_k + M_{i-1}^\top z_{i-1} \|^{p_1}$. Then

$$\begin{aligned}\Upsilon_{i-1}(y_d^{[i-2]}, \chi_a, z_{[i-1]}) &\triangleq \frac{1}{p_1 \varepsilon_1^{p_1}} \left(\left\| \sum_{k=1}^{i-2} M_k^\top z_k + M_{i-1}^\top z_{i-1} \right\|^{p_1} - \left\| \sum_{k=1}^{i-2} M_k^\top z_k \right\|^{p_1} \right) \\ &= \varrho_{i-1}(y_d^{[i-2]}, \chi_a, z_{[i-2]}, z_{i-1}) - \varrho_{i-1}(y_d^{[i-2]}, \chi_a, z_{[i-2]}, 0).\end{aligned}$$

Thus, by using the identity (see [25])

$$f(X) - f(0) = \left(\int_0^1 \frac{\partial f(s)}{\partial s} \bigg|_{s=\beta X} d\beta \right) X,$$

we have

$$\Upsilon_{i-1} = z_{i-1} \overline{\Upsilon}_{i-1}(y_d^{[i-2]}, \chi_a, z_{[i-1]}), \tag{4.43}$$

where

$$\overline{\Upsilon}_{i-1} = \int_0^1 \frac{\partial \varrho_{i-1}(\cdot, s)}{\partial s} \bigg|_{s=\alpha z_{i-1}} d\alpha.$$

Let $z_{i+1} = \eta_{i+1} - \alpha_i(y_d^{[i]}, \chi_a, \eta_{[i]})$, where α_i is a $\mathcal{C}^\infty$ function to be defined later. Then we have

$$\begin{aligned}\mathrm{d}z_i = {} & (z_{i+1} + \alpha_i + F_i(y_d^{[i]}, \chi_a, \eta_{[i]}))\,\mathrm{d}t \\ & + S_i(y_d^{[i-1]}, \chi_a, \eta_{[i-1]})\chi_b\,\mathrm{d}t + \Psi_i(y_d^{[i-1]}, \chi_a, \eta_{[i-1]})\,\mathrm{d}w,\end{aligned} \tag{4.44}$$

where

$$\begin{aligned}F_i &= g_i(y, \chi_a) - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \eta_j}(\eta_{j+1} + g_j(y, \chi_a)) - \frac{\partial \alpha_{i-1}}{\partial \chi_a}(W_a \chi_a + F_a(y)) \\ &\quad - \sum_{j=0}^{i-1} \frac{\partial \alpha_{i-1}}{\partial y_d^{(j)}} y_d^{(j+1)} - \frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2 \alpha_{i-1}}{\partial([\chi_a^\top, \eta_{[i-1]}^\top]^\top)^2} \begin{bmatrix} \Phi_a \\ \widehat{h}_1 \\ \vdots \\ \widehat{h}_{i-1} \end{bmatrix} \begin{bmatrix} \Phi_a \\ \widehat{h}_1 \\ \vdots \\ \widehat{h}_{i-1} \end{bmatrix}^\top \right), \\ S_i &= d_i - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \eta_j} d_j - \frac{\partial \alpha_{i-1}}{\partial \chi_a} L_a, \\ \Psi_i &= \widehat{h}_i(y) - \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial \eta_j} \widehat{h}_j(y) - \frac{\partial \alpha_{i-1}}{\partial \chi_a} \Phi_a(y)\end{aligned}$$

all are smooth functions.

Now we introduce the value function for this step as follows:

$$V_i = V_{i-1} + \Xi_i(y_d^{[i-1]},\, \chi_a,\, z_{[i-1]})z_i^2, \tag{4.45}$$

where Ξ_i is a positive smooth weighting function to be determined below in this step. By (4.39) and (4.44), we have

$$\begin{aligned}
\mathrm{d}V_i &= \mathrm{d}V_{i-1} + 2z_i\Xi_i(z_{i+1} + \alpha_i + F_i + S_i\chi_b)\,\mathrm{d}t + z_i^2\sum_{j=0}^{i-1}\frac{\partial \Xi_i}{\partial y_d^{(j)}}y_d^{(j+1)}\,\mathrm{d}t \\
&\quad + z_i^2\left(\frac{\partial \Xi_i}{\partial \chi_a}(W_a\chi_a + F_a + L_a\chi_b) + \sum_{j=1}^{i-1}\frac{\partial \Xi_i}{\partial z_j}(z_{j+1} + \alpha_j + F_j + S_j\chi_b)\right)\mathrm{d}t \\
&\quad + \frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2(\Xi_i z_i^2)}{\partial([\chi_a^\top,\, z_{[i]}^\top]^\top)^2}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}^\top\right)\mathrm{d}t \\
&\quad + 2\Xi_i z_i\Psi_i\,\mathrm{d}w + z_i^2\left(\frac{\partial \Xi_i}{\partial \chi_a}\Phi_a + \sum_{j=1}^{i-1}\frac{\partial \Xi_i}{\partial z_j}\Psi_j\right)\mathrm{d}w \\
&\le -z_1^2\,\mathrm{d}t + \sigma_i(y_d^{[i-1]}, \chi, z_{[i]})\,\mathrm{d}w - \frac{\theta}{4}\sigma_i\sigma_i^\top\,\mathrm{d}t + \frac{\theta}{4}(\sigma_i\sigma_i^\top - \sigma_{i-1}\sigma_{i-1}^\top)\,\mathrm{d}t \\
&\quad - \beta_i\Xi_i z_i^2\,\mathrm{d}t + \beta_i\Xi_i z_i^2\,\mathrm{d}t - \frac{r_{i-1}(y_d)\,\|\chi\|^2}{(c + \chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t \\
&\quad - \sum_{j=1}^{i-1}\Xi_j\left(\beta_j - 2\sum_{m=j+1}^{i-1}(m-1)\,\kappa_m\right)z_j^2\,\mathrm{d}t - \Delta_{i-1}\,\mathrm{d}t \\
&\quad + 2z_i\Xi_i\big(z_{i+1} + \alpha_i + \overline{F}_i(y_d^{[i]}, \chi_a, z_{[i]})\big)\,\mathrm{d}t + z_iM_i(y_d^{[i-1]}, \chi_a, z_{[i]})\chi_b\,\mathrm{d}t \\
&\quad + C_{i-1}\,\mathrm{d}t + \frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2(\Xi_i z_i^2)}{\partial([\chi_a^\top,\, z_{[i]}^\top]^\top)^2}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}^\top\right)\mathrm{d}t,
\end{aligned} \tag{4.46}$$

where

$$\begin{aligned}
\sigma_i &= \sigma_{i-1} + 2\Xi_i z_i\Psi_i + z_i^2\frac{\partial \Xi_i}{\partial \chi_a}\Phi_a + z_i^2\sum_{j=1}^{i-1}\frac{\partial \Xi_i}{\partial z_j}\Psi_j, \\
\overline{F}_i &= F_i + \frac{\Xi_{i-1}}{\Xi_i}z_{i-1} + \frac{z_i}{2\Xi_i}\left(\sum_{j=0}^{i-1}\frac{\partial \Xi_i}{\partial y_d^{(j)}}y_d^{(j+1)}\right. \\
&\qquad\qquad \left. + \sum_{j=1}^{i-1}\frac{\partial \Xi_i}{\partial z_j}(z_{j+1} + \alpha_j + F_j) + \frac{\partial \Xi_i}{\partial \chi_a}(W_a\chi_a + F_a)\right), \\
M_i &= 2\Xi_i S_i + z_i\sum_{j=1}^{i-1}\frac{\partial \Xi_i}{\partial z_j}S_j + z_i\frac{\partial \Xi_i}{\partial \chi_a}L_a
\end{aligned}$$

all are smooth functions. Here, we have used the technique of subtracting and adding terms $\frac{\theta}{4}\sigma_i\sigma_i^\top\,dt$ and $\beta_i\Xi_i z_i^2\,dt$ to the right-hand side of inequality (4.46).

Let $\overline{\sigma}_{i-1}(y_d^{[i-2]}, \chi, z_{[i-1]}) = \sigma_{i-1} - \frac{\partial V_0}{\partial \chi}\Phi = \sum_{j=1}^{i-1}(2\Xi_j z_j \Psi_j + z_j^2 \frac{\partial \Xi_j}{\partial \chi_a}\Phi_a + z_j^2 \sum_{k=1}^{j-1}\frac{\partial \Xi_j}{\partial z_k}\Psi_k)$. Then by noticing that $\overline{\sigma}_{i-1}$ is independent of χ_b, we have

(4.47)

$$\begin{aligned}
&-\Delta_{i-1,1} + \frac{\theta}{4}(\sigma_i\sigma_i^\top - \sigma_{i-1}\sigma_{i-1}^\top)\\
&\quad= -\theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial\chi}\right)^\top - \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Phi\Gamma_j^\top z_j\right\|^2 + \theta\Xi_i\sigma_{i-1}\Gamma_i^\top z_i + \theta\Xi_i^2\Gamma_i\Gamma_i^\top z_i^2\\
&\quad= -\theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial\chi}\right)^\top - \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Phi\Gamma_j^\top z_j\right\|^2\\
&\qquad+ \theta\Xi_i\frac{\partial V_0}{\partial\chi}\Phi\Gamma_i^\top z_i + \theta\Xi_i(\overline{\sigma}_{i-1} + \Xi_i\Gamma_i z_i)\Gamma_i^\top z_i\\
&\quad= -\theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial\chi}\right)^\top - \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Phi\Gamma_j^\top z_j\right\|^2\\
&\qquad+ \left(\theta\left(\frac{\partial V_0}{\partial\chi} - \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Gamma_j\Phi^\top z_j\right)\Xi_i\Phi\Gamma_i^\top z_i - \frac{\theta}{4\varepsilon_{11}^2}\Xi_i^2\Gamma_i\Phi^\top\Phi\Gamma_i^\top z_i^2\right)\\
&\qquad+ \frac{\theta}{4\varepsilon_{11}^2}\Xi_i^2\Gamma_i\Phi^\top\Phi\Gamma_i^\top z_i^2 + \frac{\theta}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Gamma_j\Phi^\top z_j\Xi_i\Phi\Gamma_i^\top z_i\\
&\qquad+ \theta\Xi_i(\overline{\sigma}_{i-1} + \Xi_i\Gamma_i z_i)\Gamma_i^\top z_i\\
&\quad= -\Delta_{i1} + \theta\Xi_i\left(\frac{1}{4\varepsilon_{11}^2}\Xi_i\Gamma_i\Phi^\top\Phi z_i\right.\\
&\qquad\qquad\left.+ \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Gamma_j\Phi^\top\Phi z_j + \overline{\sigma}_{i-1} + \Xi_i\Gamma_i z_i\right)\Gamma_i^\top z_i,
\end{aligned}$$

where

$$\Gamma_i(y_d^{[i-1]},\, \chi_a,\, z_{[i]}) = \Psi_i + \frac{z_i}{2\Xi_i}\frac{\partial\Xi_i}{\partial\chi_a}\Phi_a + \frac{z_i}{2\Xi_i}\sum_{j=1}^{i-1}\frac{\partial\Xi_i}{\partial z_j}\Psi_j,$$

$$\Delta_{i1}(y_d^{[i-1]},\, \chi,\, z_{[i]}) = \theta\varepsilon_{11}^2\left\|\left(\frac{\partial V_0}{\partial\chi}\right)^\top - \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i}\Xi_j\Phi\Gamma_j^\top z_j\right\|^2. \tag{4.48}$$

Similar to (4.25), by using (4.42) we have

$$
\begin{aligned}
(4.49)\qquad & M_i\chi_b z_i-\Delta_{i-1,2}\\
&=\left(\sum_{k=1}^{i}M_kz_k\right)\chi_b-\frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\|\chi_b\|^{\frac{p_1}{p_1-1}}-\frac{\left\|\sum_{j=1}^{i-1}M_j^\top z_j\right\|^{p_1}}{p_1\varepsilon_1^{p_1}}z_1^{p_1}\\
&=-\left(\frac{\varepsilon_1^{\frac{p_1}{p_1-1}}(p_1-1)}{p_1}\|\chi_b\|^{\frac{p_1}{p_1-1}}+\frac{\left\|\sum_{j=1}^{i}M_j^\top z_j\right\|^{p_1}}{p_1\varepsilon_1^{p_1}}-\sum_{j=1}^{i}M_jz_j\chi_b\right)\\
&\quad+\frac{1}{p_1\varepsilon_1^{p_1}}\left(\left\|\sum_{j=1}^{i-1}M_j^\top z_j+M_i^\top z_i\right\|^{p_1}-\left\|\sum_{j=1}^{i-1}M_j^\top z_j\right\|^{p_1}\right)\\
&=-\Delta_{i2}+\Upsilon_i,
\end{aligned}
$$

where

$$
\begin{aligned}
(4.50)\qquad \Delta_{i2}(y_d^{[i-1]},\chi_b,z_{[i]})&=\frac{(p_1-1)\varepsilon_1^{\frac{p_1}{p_1-1}}}{p_1}\|\chi_b\|^{\frac{p_1}{p_1-1}}+\frac{1}{p_1\varepsilon_1^{p_1}}\left\|\sum_{j=1}^{i}M_j^\top z_j\right\|^{p_1}\\
&\quad-\sum_{j=1}^{i}M_jz_j\chi_b,\\
\Upsilon_i(y_d^{[i-1]},\chi_a,z_{[i]})&=\frac{1}{p_1\varepsilon_1^{p_1}}\left(\left\|\sum_{j=1}^{i-1}M_j^\top z_j+M_i^\top z_i\right\|^{p_1}-\left\|\sum_{j=1}^{i-1}M_j^\top z_j\right\|^{p_1}\right).
\end{aligned}
$$

By Young's inequality, it is easy to see that $\Delta_{i2}\ge 0$. And similar to (4.43), there exists a smooth function $\overline{\Upsilon}_i(y_d^{[i-1]},\,\chi_a,\,z_{[i]})$ such that

$$\Upsilon_i=z_i\overline{\Upsilon}_i(y_d^{[i-1]},\,\chi_a,\,z_{[i]}).$$

By assumptions A1 and A3, we know that there exist vector-valued smooth functions $\overline{\Psi}_i(y_d^{[i-1]},\chi_a,\eta_{[i-1]})$ and $\overline{\Psi}_{ij}(y_d^{[i-1]},\,\chi_a,\,z_{[j]})$, $j=1,\dots,i-1$, such that

$$
\begin{aligned}
\Psi_i(y_d^{[i-1]},\chi_a,\eta_{[i-1]})&=\overline{\Psi}_i(y_d^{[i-1]},\chi_a,z_{[i-1]})\\
&=\overline{\Psi}_i(y_d^{[i-1]},\chi_a,\mathbf{0}_{(i-1)\times 1})+\sum_{j=1}^{i-1}\overline{\Psi}_{ij}(y_d^{[i-1]},\chi_a,z_{[j]})z_j.
\end{aligned}
$$

Then, for the last term on the right-hand side of (4.46), we have

(4.51)

$$
\begin{aligned}
&\frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2(\Xi_iz_i^2)}{\partial([\chi_a^\top,\,z_{[i]}^\top]^\top)^2}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}^\top\right)\\
&=\frac{1}{2}\mathrm{Tr}\left(\begin{bmatrix}\dfrac{\partial^2\Xi_i}{\partial([\chi_a^\top,\,z_{[i-1]}^\top]^\top)^2}z_i^2 & 2\left(\dfrac{\partial\Xi_i}{\partial[\chi_a^\top,\,z_{[i-1]}^\top]^\top}\right)^\top z_i\\ 2\dfrac{\partial\Xi_i}{\partial[\chi_a^\top,\,z_{[i-1]}^\top]^\top}z_i & 2\Xi_i\end{bmatrix}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}^\top\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\mathrm{Tr}\left(\left[\begin{array}{cc} \dfrac{\partial^2 \Xi_i}{\partial([\chi_a^\top,\, z_{[i-1]}^\top]^\top)^2} z_i & 2\left(\dfrac{\partial \Xi_i}{\partial[\chi_a^\top,\, z_{[i-1]}^\top]^\top}\right)^\top \\ 2\dfrac{\partial \Xi_i}{\partial[\chi_a^\top,\, z_{[i-1]}^\top]^\top} & 0 \end{array}\right]\left[\begin{array}{c}\Phi_a\\ \Psi_1 \\ \vdots \\ \Psi_i\end{array}\right]\left[\begin{array}{c}\Phi_a\\ \Psi_1 \\ \vdots \\ \Psi_i\end{array}\right]^\top\right) z_i \\
&\quad + \Xi_i \left(\overline{\Psi}_i(y_d^{[i-1]},\, \chi_a,\, \mathbf{0}_{(i-1)\times 1}) + \sum_{j=1}^{i-1} \overline{\Psi}_{ij}(y_d^{[i-1]},\, \chi_a,\, z_{[j]}) z_j\right) \\
&\qquad \times \left(\overline{\Psi}_i(y_d^{[i-1]},\, \chi_a,\, \mathbf{0}_{(i-1)\times 1}) + \sum_{j=1}^{i-1} \overline{\Psi}_{ij}(y_d^{[i-1]},\, \chi_a,\, z_{[j]}) z_j\right)^\top \\
&\le \frac{1}{2}\mathrm{Tr}\left(\left[\begin{array}{cc} \dfrac{\partial^2 \Xi_i}{\partial([\chi_a^\top,\, z_{[i-1]}^\top]^\top)^2} z_i & 2\left(\dfrac{\partial \Xi_i}{\partial[\chi_a^\top,\, z_{[i-1]}^\top]^\top}\right)^\top \\ 2\dfrac{\partial \Xi_i}{\partial[\chi_a^\top,\, z_{[i-1]}^\top]^\top} & 0 \end{array}\right]\left[\begin{array}{c}\Phi_a\\ \Psi_1 \\ \vdots \\ \Psi_i\end{array}\right]\left[\begin{array}{c}\Phi_a\\ \Psi_1 \\ \vdots \\ \Psi_i\end{array}\right]^\top\right) z_i \\
&\quad + 2\Xi_i \left\|(\overline{\Psi}_i(y_d^{[i-1]},\, \chi_a,\, \mathbf{0}_{(i-1)\times 1}))^\top\right\|^2 \\
&\quad + 2(i-1)\Xi_i \sum_{j=1}^{i-1} \Xi_j^{-1} \left\|(\overline{\Psi}_{ij}(y_d^{[i-1]},\, \chi_a,\, z_{[j]}))^\top\right\|^2 \Xi_j z_j^2.
\end{aligned}
$$

Choose

(4.52)

$$
\Xi_i = \frac{\kappa_i}{1 + \left\|(\overline{\Psi}_i(y_d^{[i-1]},\, \chi_a,\, \mathbf{0}_{(i-1)\times 1}))^\top\right\|^2 + \sum_{j=1}^{i-1} \Xi_j^{-1}\left\|(\overline{\Psi}_{ij}(y_d^{[i-1]},\, \chi_a,\, z_{[j]}))^\top\right\|^2}.
$$

Then we have

$$
\begin{aligned}
&2\Xi_i\left\|(\overline{\Psi}_i(y_d^{[i-1]},\, \chi_a,\, \mathbf{0}_{(i-1)\times 1}))^\top\right\|^2 \le 2\kappa_i, \\
&2(i-1)\Xi_i \sum_{j=1}^{i-1} \Xi_j^{-1}\left\|(\overline{\Psi}_{ij}(y_d^{[i-1]},\, \chi_a,\, z_{[j]}))^\top\right\|^2 \Xi_j z_j^2 \le 2(i-1)\kappa_i \sum_{j=1}^{i-1} \Xi_j z_j^2.
\end{aligned}
$$

By substituting (4.47)–(4.52) into (4.46), we get

$$
\begin{aligned}
\text{(4.53)} \qquad \mathrm{d}V_i &\le -z_1^2\,\mathrm{d}t + \sigma_i\,\mathrm{d}w - \frac{\theta}{4}\sigma_i\sigma_i^\top\,\mathrm{d}t - r_i(y_d)\frac{\|\chi\|^2}{(c + \chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t \\
&\quad - \sum_{j=1}^{i} \Xi_j\left(\beta_j - 2\sum_{m=j+1}^{i}(m-1)\,\kappa_m\right) z_j^2\,\mathrm{d}t - \Delta_i(y_d^{[i-1]}, \chi, z_{[i]})\,\mathrm{d}t \\
&\quad + 2\Xi_i z_i z_{i+1}\,\mathrm{d}t + 2z_i\Xi_i(\alpha_i - \overline{\alpha}_i(y_d^{[i]},\, \chi_a,\, \eta_{[i]}))\,\mathrm{d}t \\
&\quad + C_i(y_d^{[i]},\, \chi_a,\, z_{[i-1]})\,\mathrm{d}t,
\end{aligned}
$$

where

(4.54)

$r_i = r_{i-1}(y_d)$, with r_{i-1} being defined by (4.40),

$\Delta_i = \Delta_{i1} + \Delta_{i2}$, with Δ_{i1} and Δ_{i2} being defined by (4.48) and (4.50), respectively,

$$N_i(y_d^{[i]},\, \chi_a,\, z_{[i]}) = \overline{F}_i + \frac{\overline{\Upsilon}_i}{2\Xi_i} + \frac{\beta_i z_i}{2}$$
$$+ \frac{\theta}{2}\left(\frac{1}{4\varepsilon_{11}^2}\Xi_i\Gamma_i\Phi^\top\Phi z_i + \frac{1}{2\varepsilon_{11}^2}\sum_{j=1}^{i-1}\Xi_j\Gamma_j\Phi^\top\Phi z_j + \overline{\sigma}_{i-1} + \Xi_i\Gamma_i z_i\right)\Gamma_i^\top$$
$$+ \frac{1}{4\Xi_i}\mathrm{Tr}\left(\begin{bmatrix} \dfrac{\partial^2\Xi_i}{\partial([\chi_a^\top,\, z_{[i-1]}^\top]^\top)^2} z_i & 2\left(\dfrac{\partial\Xi_i}{\partial[\chi_a^\top,\, z_{[i-1]}^\top]^\top}\right)^\top \\ 2\dfrac{\partial\Xi_i}{\partial[\chi_a^\top,\, z_{[i-1]}^\top]^\top} & 0 \end{bmatrix}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}\begin{bmatrix}\Phi_a\\ \Psi_1\\ \vdots\\ \Psi_i\end{bmatrix}^\top\right),$$

$$\overline{\alpha}_i = \left\{-N_i(y_d^{[i]},\, \chi_a,\, z_{[i]})\right\}\Big|_{z_j=\eta_j-\alpha_{j-1},\, j=1,\ldots,i},$$

$$C_i = C_{i-1} + 2\Xi_i\left\|(\overline{\Psi}_i(y_d^{[i-1]},\, \chi_a,\, \mathbf{0}_{(i-1)\times 1}))^\top\right\|^2, \quad C_{i-1} \text{ is given by (4.41).} \tag{4.55}$$

Now, we choose virtual controller $\alpha_i(y_d^{[i]},\, \chi_a,\, \eta_{[i]})$ as follows:

$$\alpha_i = \overline{\alpha}_i(y_d^{[i]},\, \chi_a,\, \eta_{[i]}). \tag{4.56}$$

Substituting (4.56) into (4.53), we have

$$\begin{aligned}\mathrm{d}V_i \le{}& -z_1^2\,\mathrm{d}t + \sigma_i\,\mathrm{d}w - \frac{\theta}{4}\sigma_i\sigma_i^\top\,\mathrm{d}t - r_i(y_d)\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t - \Delta_i\,\mathrm{d}t\\ &- \sum_{j=1}^{i}\Xi_j\left(\beta_j - 2\sum_{m=j+1}^{i}(m-1)\,\kappa_m\right)z_j^2\,\mathrm{d}t\\ &+ 2\Xi_i z_i z_{i+1}\,\mathrm{d}t + C_i(y_d^{[i]},\, \chi_a,\, z_{[i]})\,\mathrm{d}t.\end{aligned} \tag{4.57}$$

This completes *Step* i.

Step ρ. It is easy to see that the results of Step i hold also for $i=\rho$, where $\eta_{\rho+1} = b_m g(y)u$. Define the value function V_ρ as in (4.45) with $i=\rho$ for this step. Then, V_ρ satisfies (4.57) with $i=\rho$. Set $z_{\rho+1}=0$. Then, we arrive at the controller

$$u(y_d^{[\rho]},\, \chi_a,\, \eta_{[\rho]}) = \frac{1}{b_m g(y)}\alpha_\rho(y_d^{[\rho]},\, \chi_a,\, \eta_{[\rho]}), \tag{4.58}$$

where α_ρ is defined by letting $i=\rho$ in (4.56). Let $V_\rho = V_{\rho-1} + \Xi(y_d^{[\rho-1]},\, \chi_a,\, z_{[\rho-1]})z_\rho^2$. Then we have

$$\begin{aligned}\mathrm{d}V_\rho \le{}& -z_1^2\,\mathrm{d}t + \sigma_\rho(y_d^{[\rho-1]},\, \chi_a,\, z)\,\mathrm{d}w - \frac{\theta}{4}\sigma_\rho\sigma_\rho^\top\,\mathrm{d}t\\ &- r_\rho(y_d)\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t - \Delta_\rho(y_d^{[\rho-1]},\, \chi,\, z)\,\mathrm{d}t\\ &- \sum_{j=1}^{\rho}\Xi_j\left(\beta_j - 2\sum_{m=j+1}^{\rho}(m-1)\,\kappa_m\right)z_j^2\,\mathrm{d}t + C_\rho(y_d^{[\rho]},\, \chi_a,\, z_{[\rho]})\,\mathrm{d}t,\end{aligned} \tag{4.59}$$

where Ξ_ρ, σ_ρ, r_ρ, Δ_ρ, and C_ρ are defined in the same way as in Step i ($i = 2, \ldots, \rho-1$), with i being replaced by ρ.

So far, we have completed the entire backstepping design.

4.4. Properties of the design procedure. In this subsection, we give several properties of the design procedure above. To avoid duplication of the expression, here only the case of $\rho > 1$ is considered, since the case of $\rho = 1$ has the same properties.

By Lemmas B.1 and B.2, we have

$$
\begin{aligned}
r_\rho(y_d)\frac{\|\chi\|^2}{(c+\xi)^{1-\gamma}} &\geq \frac{r_\rho(y_d)}{(\lambda_{\max}(P))^{1-\gamma}}\|\chi\|^{2\gamma} - r_\rho(y_d)\mathcal{M}_\gamma(c) \\
&\geq \frac{r_\rho(y_d)}{\lambda_{\max}(P)}\left((c+\xi)^\gamma - c^\gamma\right) - r_\rho(y_d)\mathcal{M}_\gamma(c) \\
&= \frac{r_\rho(y_d)}{\delta\lambda_{\max}(P)}\phi(\xi) - r_\rho(y_d)\mathcal{M}_\gamma(c),
\end{aligned} \tag{4.60}
$$

where $r_\rho(y_d) = r_1(y_d)$ is defined by (4.54) and ξ and $\phi(\xi)$ are defined in (4.10).

Define

$$
\overline{r}(y_d) = \frac{r_\rho(y_d)}{\delta\lambda_{\max}(P)}, \tag{4.61}
$$

$$
\overline{\beta}_i = \beta_i - 2\sum_{j=i+1}^{\rho}(j-1)\,\kappa_j, \quad i = 1, \ldots, \rho, \tag{4.62}
$$

$$
C(y_d^{[\rho]},\, \chi_a,\, z_{[\rho]}) = C_\rho + r_\rho(y_d)\mathcal{M}_\gamma(c). \tag{4.63}
$$

Then, by (4.59) (or (4.35) for the case of $\rho = 1$) and (4.60)–(4.63), we have

$$
\mathrm{d}V_\rho \leq \sigma_\rho\,\mathrm{d}w - \frac{\theta}{4}\sigma_\rho\sigma_\rho^\top\,\mathrm{d}t - z_1^2\,\mathrm{d}t - \overline{r}\,\phi(\xi)\,\mathrm{d}t - \sum_{i=1}^{\rho}\Xi_i\overline{\beta}_i z_i^2\,\mathrm{d}t - \Delta_\rho\,\mathrm{d}t + C\,\mathrm{d}t. \tag{4.64}
$$

The following lemma presents the method specifying the design constants.

LEMMA 4.2. *For any given cost value $R_l > 0$, risk-sensitivity parameter $\theta > 0$, and characteristic parameter $\gamma \in (\frac{1}{2}, 1)$, there always exist positive design constants $\delta, c, \varepsilon_{01}, \varepsilon_{02}, \varepsilon_{03}, \varepsilon_{04}, \varepsilon_1, \varepsilon_{11}, \beta_1, \ldots, \beta_\rho, \kappa_1, \ldots, \kappa_\rho$, such that the following inequalities hold:*

$$
r_\rho \;\geq\; r \;>\; 0, \quad \overline{\beta}_1 \;>\; 0, \ldots, \overline{\beta}_\rho \;>\; 0, \quad \textit{and} \quad C_\rho \;\leq\; R_l, \tag{4.65}
$$

where r is constant.

Proof. The proof can be accomplished by properly selecting a set of design constants.

Design constants δ, ε_{01}, and ε_{02} are chosen such that

$$
0 < \delta \leq 1, \tag{4.66}
$$

$$
0 < \varepsilon_{01} < \frac{\sqrt{2}}{4}, \tag{4.67}
$$

$$
0 < \varepsilon_{02} < \frac{\sqrt{2}}{4}. \tag{4.68}
$$

For given $\gamma \in (\frac{1}{2}, 1)$, even numbers $p \in \{2, 4, 6, \ldots\}$ and $p_1 \in \{4, 6, 8, \ldots\}$ are chosen such that

$$\frac{p_1}{2(p_1-1)} \le \gamma \le \frac{p}{p+1}.$$

For example, when $\gamma = \frac{2}{3}$, even numbers $p \ge 2$ and $p_1 \ge 4$ are proper; when $\gamma = \frac{4}{5}$, then even numbers $p \ge 4$ and $p_1 \ge 4$ are proper; when $\gamma = \frac{3}{5}$, then even numbers $p \ge 2$ and $p_1 \ge 6$ are proper.

Then, for given $\gamma \in (\frac{1}{2}, 1)$, risk-sensitivity parameter θ, desired positive risk-sensitive cost value R_l, given output y_d, selected δ, ε_{01}, p, and p_1, and design constants ε_{03}, ε_{04}, c, and ε_1 are chosen such that

$$0 < \varepsilon_{03} < \left(\frac{p\lambda_{\min}^{\frac{1}{p-1}}(P)}{12\theta\delta\gamma(p-1)\|P\|^2} \right)^{\frac{p-1}{p}}, \tag{4.69}$$

$$0 < \varepsilon_{04} < \sqrt{\frac{1}{18\theta\delta\gamma}}, \tag{4.70}$$

$$c \ge \max\left\{1, \left(\max\left\{ \frac{5\delta\gamma \max_{|y_d|\le C_{y_d}} \|PF(y_d)\|^2}{R_l\varepsilon_{01}^2}, 12\theta\delta\gamma \max_{|y_d|\le C_{y_d}} \left(\|P\Phi(y_d)\|_F^2\right), \frac{10\delta\gamma \max_{|y_d|\le C_{y_d}} \mathrm{Tr}\big(P\Phi(y_d)\Phi^\top(y_d)\big)}{R_l} \right\}\right)^{\frac{1}{1-\gamma}} \right\}, \tag{4.71}$$

$$0 < \varepsilon_1 < \delta \left(\min\left\{ \frac{p_1 R_l}{5(p_1-1)(1+\mathcal{K}(\frac{p_1}{p_1-1}, 2\gamma))}, \quad \frac{p_1 R_l}{5(p_1-1)\lambda_{\max}^{1-\gamma}(P)\mathcal{M}_\gamma(c)}, \frac{\gamma p_1}{8(p_1-1)\lambda_{\max}^{1-\gamma}(P)} \right\}\right)^{\frac{p_1-1}{p_1}}. \tag{4.72}$$

For given γ and selected constants c and δ, constant ε_{11} is chosen such that

$$0 < \varepsilon_{11} < \sqrt{\frac{c^{1-\gamma}}{16\theta\delta\gamma\|P\|^2}}. \tag{4.73}$$

Design constants $\kappa_1, \ldots, \kappa_\rho$ are chosen such that

$$0 < \kappa_i = \min\left\{1, \quad \frac{R_l}{10\rho}\right\}, \quad i = 1, \ldots, \rho. \tag{4.74}$$

For given κ_i's, design constants $\beta_1, \ldots, \beta_\rho$ are chosen such that

$$\beta_i > 2 \sum_{m=i+1}^{\rho} (m-1)\,\kappa_m, \quad i = 1, \ldots, \rho. \tag{4.75}$$

Thus, by (4.54) with $i = \rho$, (4.29), (4.18), (4.67)–(4.73), we have $r_\rho(y_d) > \frac{1}{8}\delta\gamma > 0$; by (4.62) and (4.75) we have $\overline{\beta}_1 > 0, \ldots, \overline{\beta}_\rho > 0$; by (4.59) (or (4.35) for the case of $\rho = 1$), (4.32), (4.19), (4.67), (4.71), (4.72), and (4.75), we have $C_\rho \le R_l$. $\square$

For the value function V_ρ, we have the following lemma.

LEMMA 4.3. *There are positive definite, continuous, and radially unbounded functions $W_1(\chi, z)$ and $W_2(\chi, z)$ such that*

$$W_1(\chi, z) \le V_\rho(y_d^{[\rho-1]}, \chi, z) \le W_2(\chi, z). \tag{4.76}$$

Proof. Define

$$W_1(\chi, z) = V_0(\chi) + \sum_{i=1}^{\rho} \min_{|y_d|\le C_{y_d},\ldots,|y_d^{(i-1)}|\le C_{y_d^{(i-1)}}} \Xi_i(y_d^{[i-1]}, \chi_a, z_{[i-1]}) z_i^2,$$

$$W_2(\chi, z) = V_0(\chi) + \sum_{i=1}^{\rho} \max_{|y_d|\le C_{y_d},\ldots,|y_d^{(i-1)}|\le C_{y_d^{(i-1)}}} \Xi_i(y_d^{[i-1]}, \chi_a, z_{[i-1]}) z_i^2.$$

Then inequality (4.76) holds.

We now show that $W_1(\chi, z)$ and $W_2(\chi, z)$ are positive definite, continuous, and radially unbounded. Clearly, W_1 and W_2 are continuous. In fact, based on $W_1(\chi, z) \le W_2(\chi, z)$ and $W_2(\mathbf{0}_{(n+m)\times 1}, \mathbf{0}_{\rho\times 1}) = 0$, it suffices to show that so is $W_1(\chi, z)$.

Let us next prove the positive definition and radial unboundedness of W_1 by induction. It is clear that $V_0(\chi)$ is positive definite and radially unbounded by the definition (4.10).

By the definition (4.27) of Ξ_1, assumption A3, and the smoothness of $\|\overline{\Psi}_1\|^2$, we see that $\min_{|y_d|\le C_{y_d}} \Xi_1(y_d)$ is existent and positive. Thus, $V_0(\chi)+\min_{|y_d|\le C_{y_d}} \Xi_1(y_d)z_1^2$ is positive definite and radially unbounded.

Clearly, $\min_{|y_d|\le C_{y_d},|y_d^{(1)}|\le C_{y_d^{(1)}}} \Xi_2(y_d^{[1]}, \chi_a, z_1)$ is positive and continuous with respect to (χ_a, z_1). Thus, by Lemma B.4 in Appendix B we can get the positive definiteness and radial unboundedness of $V_0(\chi)+\min_{|y_d|\le C_{y_d}} \Xi_1(y_d)z_1^2+\min_{|y_d|\le C_{y_d},|y_d^{(1)}|\le C_{y_d^{(1)}}} \Xi_2(y_d^{[1]}, \chi_a, z_1)z_2^2$.

Suppose that $V_0(\chi) + \sum_{j=1}^{i-1} \min_{|y_d|\le C_{y_d},\ldots,|y_d^{(j-1)}|\le C_{y_d^{(j-1)}}} \Xi_j(y_d^{[j-1]}, \chi_a, z_{[j-1]})z_j^2$ $(i = 3, 4, \ldots, \rho)$ is positive definite and radially unbounded. Then, from the positiveness and continuity of $\min_{|y_d|\le C_{y_d},\ldots,|y_d^{(i-1)}|\le C_{y_d^{(i-1)}}} \Xi_i(y_d^{[i-1]}, \chi_a, z_{[i-1]})$ and Lemma B.4 in Appendix B, we obtain the positive definiteness and radial unboundedness of

$$V_0(\chi) + \sum_{j=1}^{i} \min_{|y_d|\le C_{y_d},\ldots,|y_d^{(j-1)}|\le C_{y_d^{(j-1)}}} \Xi_j(y_d^{[j-1]}, \chi_a, z_{[j-1]}) z_i^2.$$

Thus, by induction, $W_1(\chi, z)$ is positive definite and radially unbounded. □

The following two properties are largely straightforward.

PROPERTY 4.1. *If the design constants are chosen such that inequalities* (4.65) *hold, then we have*

$$\begin{aligned} \mathrm{d}V_\rho \le{}& -(y-y_d)^2\,\mathrm{d}t + \sigma_\rho\,\mathrm{d}w - \frac{\theta}{4}\sigma_\rho\sigma_\rho^\top\,\mathrm{d}t \\ &- l\big(y_d^{[\rho-1]}, \chi, z_{[\rho]}\big)\,\mathrm{d}t + R_l\,\mathrm{d}t + r(y_d)\mathcal{M}_\gamma(c)\mathrm{d}t, \end{aligned} \tag{4.77}$$

where $l = \overline{r}(y_d)\phi(\xi) + \sum_{i=1}^{\rho} \Xi_i\overline{\beta}_i z_i^2 + \Delta_\rho$ is nonnegative.

Proof. By (4.59) (or (4.35) for the case of $\rho = 1$) and (4.65) one can easily get (4.77). □

PROPERTY 4.2. *If design constants are chosen such that inequalities* (4.65) *hold, then*

$$\mathcal{L}V_\rho \leq -z_1^2 + R_l, \tag{4.78}$$

$$\mathcal{L}V_\rho \leq -c_1 V_\rho + c_2, \tag{4.79}$$

where constants c_1 and c_2 satisfy

$$\begin{cases} c_1 = \min\left\{\frac{r}{\delta\lambda_{\max}(P)}, \overline{\beta}_1, \overline{\beta}_2, \ldots, \overline{\beta}_\rho\right\} > 0, \\ c_2 = R_l + \max_{y_d} r_\rho(y_d)\mathcal{M}_\gamma(c). \end{cases} \tag{4.80}$$

Proof. When design constants are chosen such that inequalities (4.65) hold, then (4.78) comes from (4.77), and (4.79) comes from (4.61)–(4.64), $V_\rho = \phi(\xi) + \sum_{i=1}^{\rho} \Xi_i z_i^2$, and

$$\overline{r}(y_d) \geq c_1, \overline{\beta}_1 \geq c_1, \ldots, \overline{\beta}_\rho \geq c_1, \quad \text{and} \quad C \leq c_2. \qquad \square$$

5. Main results. In this section, we summarize the main results of this paper as a theorem.

THEOREM 5.1. *Consider the system* (3.1) *and the tracking risk-sensitive cost criterion* (3.2). *Suppose that assumptions* A1–A3 *hold. Then, for any given risk-sensitivity parameter $\theta > 0$ and desired cost value $R_l > 0$, there exists an output-feedback controller such that the closed-loop system*

1. *has a unique solution on* $[0, \infty)$ *almost surely,*
2. *admits a guaranteed cost value R_l for the risk-sensitive cost criterion* (3.2)*,*
3. *is bounded in probability.*

Proof. We prove this theorem only for the case of $\rho > 1$ by construction. The proof for the case $\rho = 1$ is similar and straightforward, and so is omitted here.

For any given risk-sensitivity parameter $\theta > 0$ and desired cost value $R_l > 0$, section 4 provides a constructive design procedure of an output-feedback risk-sensitive controller. From Lemma 4.2, it is easily known that there are design constants such that inequalities (4.65) hold. Then, Lemma 4.3, (4.59), and the first two statements of Theorem A.1 imply statements 1 and 2.

Property 4.2, together with the third statement of Theorem A.1, leads directly to the boundedness in probability of $[\chi^\top, z_{[\rho]}^\top]^\top$. To show statement 3, let us first show the boundedness in probability of $[\eta_1, \ldots, \eta_\rho]^\top$.

By $\eta_1 = y = z_1 + y_d$ and assumption A3, it is easy to see that η_1 is bounded in probability. Suppose that $[\eta_1, \ldots, \eta_{k-1}]^\top$ is bounded in probability for k ($k = 2, \ldots, \rho$). Then by $\eta_k = z_k + \alpha_{k-1}(y_d^{[k-1]}, \chi_a, \eta_{[k-1]})$, the smoothness of α_{k-1}, and assumption A3, we know that η_k, and hence $[\eta_1, \ldots, \eta_k]^\top$, is bounded in probability. Therefore, by induction, $[\eta_1, \ldots, \eta_\rho]^\top$ is bounded in probability.

Thus, by $[\eta_1, \ldots, \eta_\rho, \zeta^\top]^\top = T_\rho, \ldots, T_1[y, \widehat{x}_2, \ldots, \widehat{x}_n]^\top$ and $\chi = [\zeta^\top, \widetilde{x}^\top]^\top$, it is easy to derive that $\widetilde{x}$ and $[y, \widehat{x}_2, \ldots, \widehat{x}_n]^\top$ are bounded in probability. This, together with $y = x_1$, $\widehat{x}_1 = y - \widetilde{x}_1$, and $[x_2, \ldots, x_n]^\top = [\widetilde{x}_2 + \widehat{x}_2, \ldots, \widetilde{x}_n + \widehat{x}_n]$, leads to the boundedness in probability of $[x^\top, \widehat{x}^\top]^\top$. That is, statement 3 is true. □

Remark 5.1. As for the value range of characteristic parameter γ in value function V_ρ (or V_0 given by (4.10)), the following two points are considered. First, since χ_b is unknown, in order to guarantee stability of the closed-loop system, we use

$-\frac{r_0(y_d)\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t$ to dominate the term $M_1\chi_b z_1\,\mathrm{d}t$ on the right-hand side of (4.22). This requires that the power 2γ of χ_b (or χ) in the term $\frac{r_0(y_d)\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\,\mathrm{d}t$ on the right-hand side of (4.22) be greater than 1, the power of χ_b of $M_1\chi_b z_1\,\mathrm{d}t$, that is, $\gamma > 1/2$. Second, to deal with the term $\sigma_0\,\mathrm{d}w$ on the right-hand side of (4.13), we use the technique of subtracting and adding $\frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t$ on the right-hand side of (4.13). The negative term $-\frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t$ is used to control the term $\sigma_0\,\mathrm{d}w$, while the positive term $\frac{\theta}{4}\sigma_0\sigma_0^\top\,\mathrm{d}t$ is dominated by term $-\frac{\partial\phi}{\partial\xi}\|\chi\|^2\,\mathrm{d}t$ and the system input. Thus, it is natural to require that the power 2γ of χ in $\frac{\partial\phi}{\partial\xi}\|\chi\|^2$ be greater than $4\gamma-2$, the power of χ in $\sigma_0\sigma_0^\top$; that is, $2\gamma > 4\gamma - 2$, or equivalently, $\gamma < 1$.

6. Example. Consider the second-order system

$$\begin{aligned}
\mathrm{d}x_1 &= x_2\mathrm{d}t + u\mathrm{d}t + \frac{1}{2}y^2\mathrm{d}w,\\
\mathrm{d}x_2 &= u\,\mathrm{d}t,\\
y &= x_1.
\end{aligned}$$

The purpose is to design u based on only y such that the output y of the closed-loop system tracks the sinusoidal signal:

$$y_d(t) = a\sin(\omega t), \quad a = 2,\ \omega = 2.$$

Clearly, in this case, we have $n = 2$, $m = 1$, $\rho = 1$, and $h(y) = [\frac{1}{2}y^2,\ 0]^\top$.

Design the following state observer:

$$\begin{aligned}
\dot{\widehat{x}}_1 &= \widehat{x}_2 + k_1(y - \widehat{x}_1) + u, \quad && k_1 = 1,\\
\dot{\widehat{x}}_2 &= k_2(y - \widehat{x}_1) + u, && k_2 = 1.
\end{aligned}$$

Then, the estimation error $\widetilde{x} = [x_1 - \widehat{x}_1,\ x_2 - \widehat{x}_2]^\top$ satisfies the following equation:

$$\mathrm{d}\widetilde{x} = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}\widetilde{x}\mathrm{d}t + h(y)\mathrm{d}w \triangleq A\widetilde{x}\mathrm{d}t + h(y)\mathrm{d}w.$$

Set $\varsigma_0 = [y,\ \widehat{x}_2]^\top$. Then we have the following dynamical equation for ς_0:

$$\mathrm{d}\varsigma_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\varsigma_0\mathrm{d}t + \begin{bmatrix} 0 \\ \widetilde{x}_1 \end{bmatrix}\mathrm{d}t + \begin{bmatrix} 1 \\ 0 \end{bmatrix}\widetilde{x}_2\mathrm{d}t + \begin{bmatrix} 1 \\ 1 \end{bmatrix}u\,\mathrm{d}t + \begin{bmatrix} \frac{1}{2}y^2 \\ 0 \end{bmatrix}\mathrm{d}w.$$

Let $\varsigma_1 = T_1\varsigma_0$, $T_1 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$, $T_1^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Then we have

$$\mathrm{d}\varsigma_1 = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\varsigma_1\mathrm{d}t + \begin{bmatrix} 0 \\ \widetilde{x}_1 \end{bmatrix}\mathrm{d}t + \begin{bmatrix} 1 \\ -1 \end{bmatrix}\widetilde{x}_2\mathrm{d}t + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u\,\mathrm{d}t + \begin{bmatrix} \frac{1}{2}y^2 \\ -\frac{1}{2}y^2 \end{bmatrix}\mathrm{d}w.$$

Let $\eta = [\eta_1,\ \eta_2]^\top = \varsigma_1$ and $\zeta = \eta_2$. Then we have $\eta_1 = y$ and the following dynamics used to control design:

$$\begin{aligned}
\mathrm{d}\widetilde{x} &= A\widetilde{x}\mathrm{d}t + h(y)\mathrm{d}w,\\
\mathrm{d}\zeta &= (-\zeta - y + \widetilde{x}_1 - \widetilde{x}_2)\mathrm{d}t - \frac{1}{2}y^2\,\mathrm{d}w,\\
\mathrm{d}y &= (\zeta + y + u + \widetilde{x}_2)\mathrm{d}t + \frac{1}{2}y^2\,\mathrm{d}w.
\end{aligned}$$

In this case, we have

$$W = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

By solving Lyapunov equation $W^\top P + PW = -I_3$, we get

$$P = \begin{bmatrix} 0.5000 & 0.3333 & -0.1667 \\ 0.3333 & 1.5000 & -0.6667 \\ -0.1667 & -0.6667 & 1.6667 \end{bmatrix}.$$

Clearly, P is symmetric and positive definite. The eigenvalues of P are 0.3978, 0.9476, 2.3213, and thus, $\lambda_{\min}(P) = 0.3978$, $\lambda_{\max}(P) = 2.3213$.

Let $\gamma = \frac{2}{3}$, $p = 2$, $p_1 = 4$, $z_1 = y - y_d$, $\chi_a = [\zeta,\, \widetilde{x}_1]^\top$, $\xi_a = \chi_a^\top (P_1 - P_2 P^{-1} P_2^\top)\chi_a = 0.4833\zeta^2 + 1.2333\widetilde{x}_1^2 + 0.5332\zeta\widetilde{x}_1$. Then we design the controller $u(y_d^{[1]},\, \chi_a,\, \eta_1) = \alpha_1(y_d^{[1]},\, \chi_a,\, \eta_1)$ as follows:

$$\alpha_1 = \left. \left\{ -N_1 - \frac{z_1}{2\Xi_1} - \frac{N_0}{2\Xi_1} \right\} \right|_{z_1 = \eta_1 - y_d}$$

with $\Xi_1 = \frac{4\kappa_1}{4 + y_d^4}$ and

$$\begin{aligned} N_0 = {} & \frac{0.2593\delta}{\varepsilon_{02}^2} \cdot \frac{1}{(c + \xi_a)^{\frac{1}{3}}} + \frac{4\theta\delta^2}{9} \cdot \frac{(y + y_d)^2 (0.5834\widetilde{x}_1 - 0.0833\zeta)^2 z_1}{(c + \xi_a)^{\frac{2}{3}}} \\ & + \frac{3.2261\theta\delta^2}{\varepsilon_{04}^2} \cdot \frac{(\zeta^2 + \widetilde{x}_1^2)\left(2y_d^4 + z_1^2 (y + y_d)^2\right)(y + y_d)^2 z_1}{c + \xi_a} \\ & + \frac{0.2993\theta\delta^2}{\varepsilon_{03}^2} \cdot (y + y_d)^2 (y^2 + y_d^2)^2 z_1 + \frac{\delta}{3} \cdot \frac{(y + y_d)^2 z_1}{(c + \xi_a)^{\frac{1}{3}}}, \\ N_1 = {} & \zeta - \dot{y}_d + \frac{\beta_1}{2} z_1 + \frac{z_1}{2\Xi_1} \frac{\partial \Xi_1}{\partial y_d} \dot{y}_d + \frac{2}{\varepsilon_1^4} \Xi_1^3 z_1^3 + \frac{\theta}{8} \Xi_1 \left(1 + \frac{y^4}{8\varepsilon_{11}^2}\right) y^4 z_1 \\ & + \frac{1}{4}(y + y_d)^2 z_1. \end{aligned}$$

Here, the desired cost value R_l is set to 0.5. Accordingly, the design constants in Lemma 4.2 are chosen as $\delta = 0.9$, $\theta = 0.2$, $\varepsilon_{01} = 0.3$, $\varepsilon_{02} = 0.3$, $\varepsilon_{03} = 0.32$, $\varepsilon_{04} = 0.68$, $c = 100$, $\varepsilon_1 = 0.1$, $\varepsilon_{11} = 0.29$, $\kappa_1 = 0.05$, $\beta_1 = 40$; the stochastic disturbance $\frac{\mathrm{d}w}{\mathrm{d}t}$ is chosen to be Gaussian white noise with power 1; and the initial conditions are simply set to $x_1(0) = 0.8$, $x_2(0) = 0$, $\widehat{x}_1(0) = 0$, $\widehat{x}_2(0) = 0$.

The simulation results are shown in Figures 1–4 given below. In particular, Figure 1 is about x_1 (solid line) and its estimation $\widehat{x}_1$ (dashdotted line); Figure 2 is about x_2 (solid line) and its estimation $\widehat{x}_2$ (dashdotted line); Figure 3 is about desired output y_d (solid line), system output y (dashdotted line), and tracking error $y - y_d$ (dashed line); Figure 4 is about control input u; Figure 5 gives a diagram of $\frac{1}{t}\int_0^t (y(s) - y_d(s))^2 \mathrm{d}s$, used to demonstrate the validity of the design. From Figure 3 and Figure 5 we can see that the system output tracks the desired output ideally.

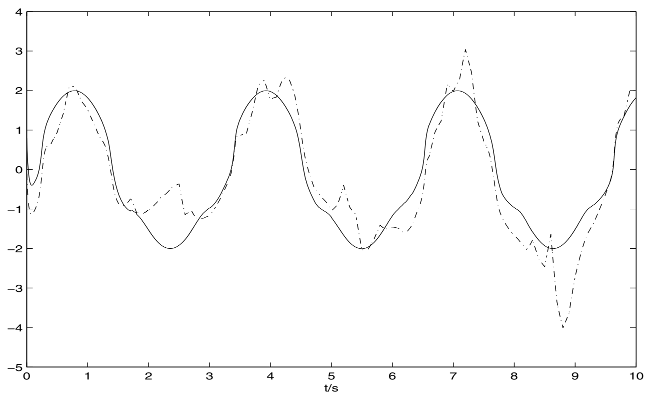

FIG. 1. *System state* x_1 *(solid) and observer state* $\hat{x}_1$ *(dashdotted).*

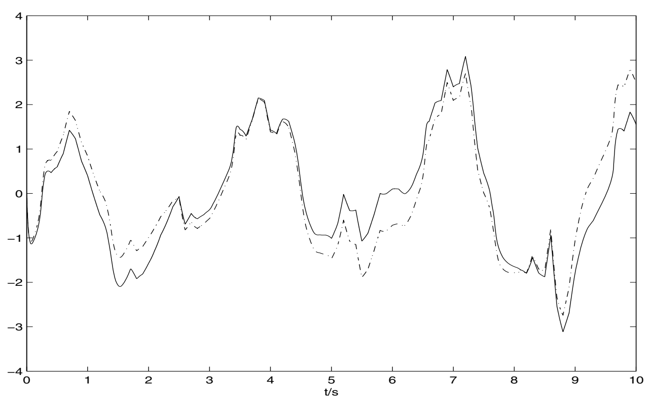

FIG. 2. *System state* x_2 *(solid) and observer state* $\hat{x}_2$ *(dashdotted).*

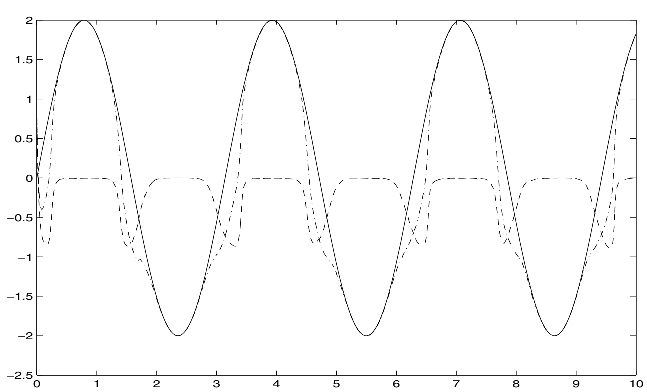

FIG. 3. *Desired output* y_d *(solid), system output* y *(dashdotted), and tracking error* $y - y_d$ *(dashed).*

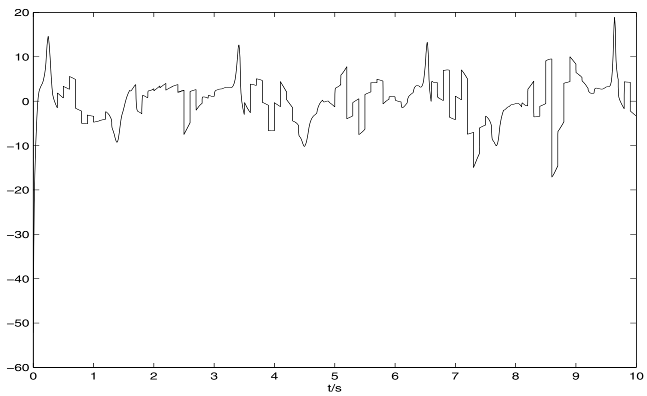

FIG. 4. *Control input* u.

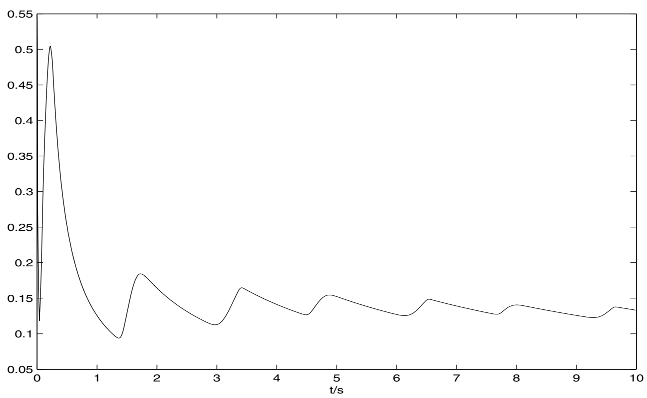

FIG. 5. $\frac{1}{t}\int_0^t (y(s) - y_d(s))^2 \mathrm{d}s$.

7. Concluding remarks. In this paper, the practical output-feedback control design problem of stochastic nonlinear strict-feedback systems in observer canonical form with stable zero-dynamics under a long-term tracking risk-sensitive cost criterion is investigated. A state observer is designed to guarantee an exponentially convergent state estimate when there is no disturbance. By introducing a state-transformation, we transform the system with the state observer in the loop into a lower triangular structure. And then, for any given risk-sensitivity parameter and desired cost value, by using an integrator backstepping method, we present constructively the output-feedback control design algorithm. The cost function adopted here is of quadratic form usually encountered in practice, rather than the quartic one used to avoid difficulty on controller design and performance analysis of the closed-loop systems. It is shown that under our control design (a) the closed-loop system is bounded in probability, and (b) the long-term average risk-sensitive cost of the closed-loop systems is upper bounded by the desired value. Besides, the value range of the characteristic parameters of the value function is investigated. As a special case when system vector nonlinearity and stochastic disturbance vector field vanish at the desired output y_d, it can be expected that there exists a control such that the closed-loop system is asymptotically stable in the large and admits a zero risk-sensitive cost. This question is now under study.

Appendix A. Preliminary results. In this appendix, we give the definitions of *bounded in probability* and *asymptotically stable in the large*, as well as a key theorem to present the sufficient conditions for these two stability notions.

For a general control-free stochastic nonlinear system,

$$\mathrm{d}x = f(t,x)\,\mathrm{d}t + h(t,x)\mathrm{d}w, \quad x(t_0) = x_0, \tag{A.1}$$

where x is an n-dimensional state vector, $n \in \mathbb{N}$; $f : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ and $h : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times s}$, $s \in \mathbb{N}$, are assumed to be continuous in t and locally Lipschitz in x; w is an s-dimensional vector-valued Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$; $t_0 \geq 0$ and $x_0 \in \mathbb{R}^n$. Denote the solution to (A.1) by $x_{t_0,x_0}(t)$.

DEFINITION A.1. *The solution process* $\{x_{t_0,x_0}(t),\ t \geq t_0\}$ *is said to be* bounded in probability *if*

$$\lim_{\varepsilon\to\infty} \sup_{t\in[t_0,\,\infty)} \mathcal{P}\left\{\|x_{t_0,x_0}(t)\| > \varepsilon\right\} = 0.$$

DEFINITION A.2. *Consider the system* (A.1), *with* $f(t, \mathbf{0}_{n\times1}) = \mathbf{0}_{n\times1}$ *and* $h(t, \mathbf{0}_{n\times1}) = \mathbf{0}_{n\times s}$ $\forall t \geq 0$. *The identically zero solution process is said to be* asymptotically stable in the large *if* $\forall\varepsilon > 0$, $t_0 \in [0, \infty)$,

$$\lim_{\|x_0\|\to0^+} \mathcal{P}\left\{\sup_{t\geq0} \|x_{t_0,x_0}(t)\| \geq \varepsilon\right\} = 0$$

and $\forall x_0 \in \mathbb{R}^n$, $\forall t_0 \in [0, \infty)$,

$$\mathcal{P}\left\{\lim_{t\to\infty} x_{t_0,x_0}(t) = \mathbf{0}_{n\times1}\right\} = 1.$$

The following theorem gives the sufficient conditions for the above two stochastic stability concepts.

THEOREM A.1. *Consider stochastic nonlinear system* (A.1) *and the following risk-sensitive cost criterion:*

$$J_\theta = \limsup_{T\to\infty} \frac{1}{T}\,\frac{2}{\theta}\ln\left(E\left(\exp\left(\frac{\theta}{2}\int_0^T q(t, x_{0,x_0}(t))\,\mathrm{d}t\right)\right)\right), \tag{A.2}$$

where $\theta > 0$ *is the risk-sensitive parameter and* $q : [0, \infty)\times\mathbb{R}^n \to \mathbb{R}$ *is a nonnegative continuous function. For any* $\theta > 0$ *and any desired cost value* $R_l > 0$, *if there exists a nonnegative value function* $V : [0, \infty)\times\mathbb{R}^n \to \mathbb{R}$, *which is* $\mathcal{C}^1$ *in the first argument and* $\mathcal{C}^2$ *in the second argument; a continuous function* $\sigma : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^{1\times s}$; *a nonnegative continuous function* $l : [0, \infty)\times\mathbb{R}^n \to \mathbb{R}$; *and a nonnegative, continuous, and radially unbounded function* $W_1 : \mathbb{R}^n \to \mathbb{R}$ *such that*

$$W_1(x) \leq V(t, x) \quad \forall(t, x) \in [0, \infty) \times \mathbb{R}^n, \tag{A.3}$$

$$\begin{aligned} \mathrm{d}V(t, x) = {}& \sigma(t, x)\,\mathrm{d}w - \frac{\theta}{4}\sigma(t, x)(\sigma(t, x))^\top\,\mathrm{d}t - l(t, x)\,\mathrm{d}t \\ & - q(t, x)\,\mathrm{d}t + R_l\,\mathrm{d}t \qquad \forall(t, x) \in [0, \infty) \times \mathbb{R}^n, \end{aligned} \tag{A.4}$$

then the following statements hold:

1. *The system* (A.1) *has a unique solution on* $[t_0, \infty)$ *almost surely* $\forall t_0 \geq 0$.

2. $J_\theta(x_0) \le R_l \ \forall x_0 \in \mathbb{R}^n$.
3. *If, in addition, there are constants* $c_{l1} \in (0, \infty)$ *and* $c_{l2} \in (0, \infty)$ *such that*

$$\mathcal{L}V(t, x) \le -c_{l1}V(t, x) + c_{l2} \quad \forall (t, x) \in [0, \infty) \times \mathbb{R}^n, \tag{A.5}$$

then the solution of system (A.1) *is bounded in probability.*

4. *If, in addition,* $f(t, \mathbf{0}_{n\times 1}) = \mathbf{0}_{n\times 1}$, $h(t, \mathbf{0}_{n\times 1}) = \mathbf{0}_{n\times s}$ $\forall t \ge 0$, W_1 *is positive definite,* $R_l = 0$, *and there exist a continuous and positive definite function* $W_2 :\ \mathbb{R}^n \to \mathbb{R}$ *and a positive definite function* $W_3 :\ \mathbb{R}^n \to \mathbb{R}$ *such that*

$$V(t, x) \le W_2(x) \quad \forall (t, x) \in [0, \infty) \times \mathbb{R}^n, \tag{A.6}$$

$$\frac{\theta}{4}\sigma(t, x)(\sigma(t, x))^\top + l(t, x) + q(t, x) \ge W_3(x) \quad \forall (t, x) \in [0, \infty) \times \mathbb{R}^n, \tag{A.7}$$

then the zero solution of the system (A.1) *is asymptotically stable in the large.*

Proof. Define $\widehat{V} \stackrel{\triangle}{=} V + R_l$. Clearly, $\widehat{V}$ is nonnegative and satisfies

$$\mathcal{L}\widehat{V} = -\frac{\theta}{4}\sigma(t, x)(\sigma(t, x))^\top - l(t, x) - q(t, x) + R_l \le \widehat{V}$$

and

$$\lim_{r\to\infty} \inf_{\|x\|>r} \widehat{V}(t, x) \ge \lim_{r\to\infty} \inf_{\|x\|>r} W_1(x) = \infty \quad \forall t \in [0, \infty).$$

Then, by Theorem 4.1 of Chapter III of [21], statement 1 follows.

For statement 2, fix $t_0 = 0$ and $x_0 \in \mathbb{R}^n$. By (A.4), we have

$$\begin{aligned}
&V(T, x_{0,x_0}(T)) + \int_0^T (q(t, x_{0,x_0}(t)) + l(t, x_{0,x_0}(t)))\,\mathrm{d}t \\
&\le V(0, x_0) + \int_0^T \sigma(t, x_{0,x_0}(t))\,\mathrm{d}w \\
&\quad - \frac{\theta}{4}\int_0^T \sigma(t, x_{0,x_0}(t))(\sigma(t, x_{0,x_0}(t)))^\top\,\mathrm{d}t + R_l T \quad \forall T \ge 0.
\end{aligned}$$

This implies that

$$\begin{aligned}
&\frac{1}{T}\frac{2}{\theta}\ln\bigg(E\bigg(\exp\bigg(\frac{\theta}{2}\int_0^T q(t, x_{0,x_0}(t))\,\mathrm{d}t\bigg)\bigg)\bigg) \\
&\le \frac{1}{T}\frac{2}{\theta}\ln\bigg(E\bigg(\exp\bigg(\frac{\theta}{2}\bigg(V(T, x_{0,x_0}(T)) + \int_0^T \big(q(t, x_{0,x_0}(t)) + l(t, x_{0,x_0}(t))\big)\,\mathrm{d}t\bigg)\bigg)\bigg)\bigg) \\
&\le \frac{V(0, x_0)}{T} + \frac{1}{T}\frac{2}{\theta}\ln\bigg(E\bigg(\exp\bigg(\frac{\theta}{2}\bigg(\int_0^T \sigma(t, x_{0,x_0}(t))\,\mathrm{d}w \\
&\qquad\qquad - \frac{\theta}{4}\int_0^T \sigma(t, x_{0,x_0}(t))(\sigma(t, x_{0,x_0}(t)))^\top\,\mathrm{d}t\bigg)\bigg)\bigg)\bigg) + R_l.
\end{aligned}$$

Let, $\forall T \ge 0$,

$$\zeta(T) \stackrel{\triangle}{=} \exp\left(\int_0^T \frac{\theta}{2}\sigma(t, x_{0,x_0}(t))\,\mathrm{d}w - \frac{\theta^2}{8}\int_0^T \sigma(t, x_{0,x_0}(t))(\sigma(t, x_{0,x_0}(t)))^\top\,\mathrm{d}t\right).$$

Then $\zeta(T)$ is a supermartingale (see [33]), and $E(\zeta(T)) \le E(\zeta(0)) = 1\ \forall T \ge 0$. Thus, we have

$$J_\theta(x_0) \le \limsup_{T\to\infty}\left(\frac{V(0,\,x_0)}{T} + \frac{1}{T}\,\frac{2}{\theta}\ln\left(E\left(\zeta(T)\right)\right) + R_l\right) \le R_l.$$

This establishes statement 2.

For statement 3, fix $t_0 \in [0,\,\infty)$ and $x_0 \in \mathbb{R}^n$. Let

$$\alpha(c) = \inf_{\|x\|>c,\,x\in\mathbb{R}^n} W_1(x) \quad \forall c \in [0,\,\infty).$$

Then, by (A.3) and (A.5), we have, for sufficiently large $c \in [0,\,\infty)$ and any $t \ge t_0$,

$$\begin{aligned}
\mathcal{P}(\|x_{t_0,x_0}(t)\| > c) &= \int_\Omega I_{\{\|x_{t_0,x_0}(t)\|>c\}}(\omega)\mathcal{P}(d\omega)\\
&= \int_\Omega \frac{I_{\{\|x_{t_0,x_0}(t)\|>c\}}W_1(x_{t_0,x_0}(t))}{W_1(x_{t_0,x_0}(t))}\mathcal{P}(d\omega)\\
&\le \frac{E(W_1(x_{t_0,x_0}(t)))}{\alpha(c)} \le \frac{E(V(t,\,x_{t_0,x_0}(t)))}{\alpha(c)}\\
&\le \frac{V(t_0,\,x_0) + c_{l2}/c_{l1}}{\alpha(c)}.
\end{aligned}$$

Since the fact that W_1 is radially unbounded implies that $\alpha(c) \to \infty$ as $c \to \infty$, then statement 3 follows.

For statement 4, we note that V is clearly positive definite,

$$0 \le \lim_{x\to\mathbf{0}_{n\times 1}} \sup_{t\ge 0} V(t,\,x) \le \lim_{x\to\mathbf{0}_{n\times 1}} W_2(x) = 0,$$

which implies that V has infinitesimal upper limit, $\mathcal{L}V(t,\,x) = -\frac{\theta}{4}\sigma(t,\,x)(\sigma(t,\,x))^\top - l(t,\,x) - q(t,\,x) \le -W_3(x)\ \forall (t,\,x) \in [0,\,\infty)\times\mathbb{R}^n$, which is negative definite, and

$$\lim_{R\to\infty}\ \inf_{\|x\|>R,\,x\in\mathbb{R}^n}\ \inf_{t>0} V(t,\,x) \ge \lim_{R\to\infty}\ \inf_{\|x\|>R,\,x\in\mathbb{R}^n} W_1(x) = \infty.$$

By Theorem 4.4 in Chapter V of [21], the zero solution of system (A.1) is asymptotically stable in the large. □

Appendix B. Technical lemmas.

LEMMA B.1. *Let $n \in \mathbb{N}$, P be an $n\times n$-dimensional symmetric positive definite matrix, $\gamma \in (\frac{1}{2},\,1)$,*

$$\Pi_\gamma(x,c) = \|x\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + x^\top Px)^{\gamma-1}\|x\|^2 \qquad \forall x \in \mathbb{R}^n,\ \forall c \in (0,\,\infty),$$

and

$$\mathcal{M}_\gamma(c) = \sup_{x\in\mathbb{R}^n} \Pi_\gamma(x,c) \quad \forall c \in (0,\,\infty).$$

Then $\Pi_\gamma(x,c) \ge 0\ \forall x \in \mathbb{R}^n$, $\forall c \in (0,\,\infty)$; and $\mathcal{M}_\gamma$ is strictly increasing on $(0,\,\infty)$; and

$$\lim_{c\to 0^+} \mathcal{M}_\gamma(c) = 0, \qquad \lim_{c\to+\infty} \mathcal{M}_\gamma(c) = +\infty.$$

Proof. Clearly, for any nonzero vector $x \in \mathbb{R}^n$, we have

$$\Pi(x,\, c) > \Pi(x,\, 0) = 0 \qquad \forall c \in (0,\, \infty).$$

From this together with $\Pi(\mathbf{0}_{n\times 1},\, c) = 0$ $\forall c \in (0,\, \infty)$, it follows that $\Pi_\gamma(x, c) \geq 0$ $\forall x \in \mathbb{R}^n$, $\forall c \in (0,\, \infty)$.

Let us next prove the properties of $\mathcal{M}_\gamma$. From the definitions of $\lambda_{\max}(P)$ and $\lambda_{\min}(P)$, it follows that for any nonzero vector $x \in \mathbb{R}^n$,

$$\lambda_{\min}(P) \;\leq\; \frac{x^\top P x}{\|x\|^2} \;\leq\; \lambda_{\max}(P).$$

Further, by $\gamma < 1$ we have for any nonzero vector $x \in \mathbb{R}^n$,

$$\begin{aligned}
\Pi_\gamma(x, c) \;&\leq\; \|x\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + \lambda_{\max}(P)\|x\|^2)^{\gamma-1}\|x\|^2 \\
&=\; \lambda_{\max}^{\gamma-1}(P)\|x\|^{2\gamma}\left(1 - \left(1 + \frac{c}{\lambda_{\max}(P)\|x\|^2}\right)^{\gamma-1}\right) \\
&=\; \frac{\lambda_{\max}^{\gamma-1}(P)\|x\|^{2\gamma}}{\left(1 + \frac{c}{\lambda_{\max}(P)\|x\|^2}\right)^{\gamma-1}}\left(\left(1 + \frac{c}{\lambda_{\max}(P)\|x\|^2}\right)^{1-\gamma} - 1\right) \\
&\leq\; \lambda_{\max}^{\gamma-1}(P)\|x\|^{2\gamma}\left(1 + \frac{c(1-\gamma)}{\lambda_{\max}(P)\|x\|^2} - 1\right) \\
&=\; c(1-\gamma)\lambda_{\max}^{\gamma-2}(P)\|x\|^{2(\gamma-1)} \\
&\longrightarrow 0 \qquad \text{as} \quad \|x\| \to \infty,
\end{aligned}$$

where we have used the following inequality: $(1+a)^r \leq 1 + ar$ $\forall a \in [0,\, \infty)$, $\forall r \in (0,\, 1)$.

Let

$$\mathcal{X}_0 = \{x \in \mathbb{R}^n : Px = \lambda_{\max}(P)x\}\,.$$

Then it can be shown that for any constant $c > 0$,

$$\mathcal{M}_\gamma(c) \;\geq \sup_{x \in \mathcal{X}_0, \|x\|=1} \Pi_\gamma(x, c) = \lambda_{\max}^{\gamma-1}(P) - (c + \lambda_{\max}(P))^{\gamma-1} \;>\; 0.$$

Therefore, there is a nonzero $x_1 \in \mathbb{R}^n$ at which $\Pi_\gamma(x, c)$ reaches its maximum. Furthermore, we can show that $x_1 \in \mathcal{X}_0$, since otherwise there would be $x_1^\top P x_1 < \lambda_{\max}(P)\|x_1\|^2$. Take $x_0 \in \mathcal{X}_0$ such that $\|x_0\| = \|x_1\|$. Then,

$$\begin{aligned}
\Pi_\gamma(x_0, c) &= \|x_0\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + x_0^\top P x_0)^{\gamma-1}\|x_0\|^2 \\
&= \|x_0\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + \lambda_{\max}(P)\|x_0\|^2)^{\gamma-1}\|x_0\|^2 \\
&= \|x_1\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + \lambda_{\max}(P)\|x_1\|^2)^{\gamma-1}\|x_1\|^2 \\
&> \|x_1\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + x_1^\top P x_1)^{\gamma-1}\|x_1\|^2.
\end{aligned}$$

This contradicts the fact that x_1 is the maximum point of $\Pi_\gamma(x, c)$.

Thus, there must be

$$\begin{aligned}
\mathcal{M}_\gamma(c) &= \sup_{x \in \mathbb{R}^n} \Pi_\gamma(x, c) = \sup_{x \in \mathcal{X}_0} \Pi_\gamma(x, c) \\
&= \sup_{x \in \mathcal{X}_0} \left[\|x\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + \lambda_{\max}(P)\|x\|^2)^{\gamma-1}\|x\|^2\right].
\end{aligned}$$

In other words, the vector maximization problem has been transformed into a scalar one in α of the following two-variable function $f(\alpha, c)$:

$$f(\alpha, c) = \alpha^{\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + \lambda_{\max}(P)\alpha)^{\gamma-1}\alpha.$$

That is, for any given $c \geq 0$, we have

$$\mathcal{M}_{\gamma}(c) = \sup_{\alpha \geq 0} f(\alpha, c).$$

Noticing that

$$\frac{\partial f(\alpha, c)}{\partial c} = (1-\gamma)(c + \lambda_{\max}(P)\alpha)^{\gamma-2}\alpha \, > \, 0 \quad \forall \alpha > 0, \ \forall c > 0,$$

we know that $f(\alpha, c)$ is a strictly increasing function of c for fixed $\alpha \in (0, \infty)$. It can also be shown that $\mathcal{M}_{\gamma}(c)$ is strictly increasing; i.e., for any $0 < c_1 < c_2$ there is always $0 < \mathcal{M}_{\gamma}(c_1) < \mathcal{M}_{\gamma}(c_2)$ due to the following argument. Let α_{c_1} maximize $f(\alpha, c_1)$, or $\mathcal{M}_{\gamma}(\alpha_{c_1}, c_1) = f(\alpha_{c_1}, c_1) = \sup_{\alpha \geq 0} f(\alpha, c_1)$. Then by the monotonicity of $f(\alpha, \cdot)$, we have

$$\mathcal{M}_{\gamma}(c_1) = f(\alpha_{c_1}, c_1) \, < \, f(\alpha_{c_1}, c_2) \, \leq \, \sup_{\alpha \geq 0} f(\alpha, c_2) = \mathcal{M}_{\gamma}(c_2).$$

Note that, $\forall c \in (0, \infty)$ and $\forall x \in \mathbb{R}^n$ with $x \neq \mathbf{0}_{n\times 1}$,

$$\begin{aligned}
\Pi_{\gamma}(x, c) &\leq \|x\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P) - (c + \lambda_{\max}(P)\|x\|^2)^{\gamma-1}\|x\|^2 \\
&= \frac{\|x\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P)}{(c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma}}\left((c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma} - \lambda_{\max}^{1-\gamma}(P)\|x\|^{2(1-\gamma)}\right) \\
&= \frac{\|x\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P)}{(c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma}}\left(\left(1 + \frac{c}{\lambda_{\max}(P)\|x\|^2}\right)^{1-\gamma}\lambda_{\max}^{1-\gamma}(P)\|x\|^{2(1-\gamma)}\right. \\
&\qquad \left. - \lambda_{\max}^{1-\gamma}(P)\|x\|^{2(1-\gamma)}\right) \\
&= \frac{\|x\|^2}{(c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma}}\left(\left(1 + \frac{c}{\lambda_{\max}(P)\|x\|^2}\right)^{1-\gamma} - 1\right) \\
&< \frac{\|x\|^2}{(c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma}}\,\frac{c(1-\gamma)}{\lambda_{\max}(P)\|x\|^2} \\
&= \frac{c(1-\gamma)(\lambda_{\max}(P))^{-1}}{(c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma}}.
\end{aligned}$$

Clearly, we have $\Pi_{\gamma}(\mathbf{0}_{n\times 1}, c) = 0 < \frac{c(1-\gamma)(\lambda_{\max}(P))^{-1}}{(c)^{1-\gamma}}$ $\forall c \in (0, \infty)$. Then,

$$\Pi_{\gamma}(x, c) < \frac{c(1-\gamma)(\lambda_{\max}(P))^{-1}}{(c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma}} \quad \forall c \in (0, \infty), \ \forall x \in \mathbb{R}^n.$$

This implies that, $\forall c \in (0, \infty)$,

$$\mathcal{M}_{\gamma}(c) \; = \; \sup_{x \in \mathbb{R}^n} \Pi_g(x, c) \leq \sup_{x \in \mathbb{R}^n} \frac{c(1-\gamma)(\lambda_{\max}(P))^{-1}}{(c + \lambda_{\max}(P)\|x\|^2)^{1-\gamma}} \; = \; c^{\gamma}(1-\gamma)(\lambda_{\max}(P))^{-1}.$$

Clearly, $\mathcal{M}_{\gamma}(c) > 0$ $\forall c \in (0, \infty)$. Moreover, we have

$$0 \leq \lim_{c \to 0^+} \mathcal{M}_g(c) \leq \lim_{c \to 0^+} c^{\gamma}(1-\gamma)(\lambda_{\max}(P))^{-1} = 0.$$

We now show

$$\lim_{c\to+\infty}\mathcal{M}_\gamma(c)=+\infty, \tag{B.1}$$

since otherwise, by the fact that $\mathcal{M}_\gamma(c)$ is strictly increasing in $[0,\,\infty)$, $\lim_{c\to+\infty}\mathcal{M}_\gamma(c)$ would be existent and finite. Let $\Upsilon=\lim_{c\to+\infty}\mathcal{M}_\gamma(c)$. Then for $c=\left(1-\frac{1}{2^{1-\gamma}}\right)^{-\frac{1}{\gamma}}\cdot(2\lambda_{\max}(P)\Upsilon)^{\frac{1}{\gamma}}$, we would have $\mathcal{M}_\gamma(c)\geq f(c\lambda_{\max}^{-1}(P),c)=2\Upsilon>\Upsilon$. This contradicts $\Upsilon=\lim_{c\to+\infty}\mathcal{M}_\gamma(c)$ and the fact that $\mathcal{M}_\gamma(c)$ is strictly increasing in $[0,\infty)$. Thus, (B.1) is true. □

Remark B.1. Lemma B.1 means that the difference between term $\|x\|^{2\gamma}\lambda_{\max}^{\gamma-1}(P)$ and term $(c+x^\top Px)^{\gamma-1}\,\|x\|^2$ is less than or equal to $\mathcal{M}_\gamma(c)$, which is arbitrarily close to zero as constant c is.

LEMMA B.2. *Let $n\in\mathbb{N}$, $P\in\mathbb{R}^{n\times n}$ be symmetric and positive definite, $c\in[0,\,\infty)$, and $\gamma\in(\frac{1}{2},\,1)$. Define the function*

$$\Delta_\gamma(x,c)=(c+x^\top Px)^\gamma-c^\gamma-\lambda_{\max}^\gamma(P)\|x\|^{2\gamma},\quad \forall c\geq 0,\ \forall x\in\mathbb{R}^n.$$

Then we have

$$\Delta_\gamma(x,c)\leq 0\quad \forall x\in\mathbb{R}^n,\ \forall c\geq 0. \tag{B.2}$$

Proof. Let $z=\|x\|^2$ and $\overline{\Delta}_\gamma(z,c)=(c+\lambda_{\max}(P)z)^\gamma-c^\gamma-\lambda_{\max}^\gamma(P)z^\gamma$. Then, $\overline{\Delta}_\gamma(z,c)\geq\Delta_\gamma(x,c)\,\forall\,c\geq 0$, and for any $z>0$ and $c\geq 0$,

$$\begin{aligned}\frac{\partial\overline{\Delta}_\gamma(z,\,c)}{\partial z}&=\frac{\gamma\lambda_{\max}(P)}{(c+\lambda_{\max}(P)z)^{1-\gamma}}-\frac{\gamma\lambda_{\max}^\gamma(P)}{z^{1-\gamma}}\\&\leq\frac{\gamma\lambda_{\max}^\gamma(P)}{(c\lambda_{\max}^{-1}(P)+z)^{1-\gamma}}-\frac{\gamma\lambda_{\max}^\gamma(P)}{z^{1-\gamma}}\leq 0.\end{aligned}$$

This together with $\overline{\Delta}_\gamma(0,c)=\Delta_\gamma(0,c)=0$ gives (B.2). □

Remark B.2. From Lemmas B.1 and B.2, we know that for any $\chi\in\mathbb{R}^{n+m}$ there exist the following inequalities:

$$\begin{aligned}\|\chi\|^{2\gamma}&\leq\lambda_{\max}^{1-\gamma}(P)\left(\mathcal{M}_\gamma(c)+\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\right),\\(c+\chi^\top P\chi)^\gamma-c^\gamma&\leq\lambda_{\max}(P)\left(\mathcal{M}_\gamma(c)+\frac{\|\chi\|^2}{(c+\chi^\top P\chi)^{1-\gamma}}\right).\end{aligned} \tag{B.3}$$

LEMMA B.3. *For any given constants a_1 and a_2 satisfying $1\leq a_1\leq a_2$, set $f_{a_1,\,a_2}(x)=x^{a_1}-x^{\alpha_2}\ \forall x\geq 0$, and*

$$\begin{cases}\mathcal{K}(a_1,a_2)=\frac{a_2-a_1}{a_2}\left(\frac{a_1}{a_2}\right)^{\frac{a_1}{a_2-a_1}} & \text{if } a_1<a_2,\\ \mathcal{K}(a_1,a_2)=0 & \text{if } a_1=a_2.\end{cases}$$

Then

$$\sup_{x\geq 0}f_{a_1,\,a_2}(x)=\mathcal{K}(a_1,a_2). \tag{B.4}$$

Proof. When $1\leq a_1<a_2$, by $\frac{\mathrm{d}}{\mathrm{d}x}f_{a_1,\,a_2}(x)=a_1x^{a_1-1}-a_2x^{\alpha_2-1}=0$, we see that $f_{a_1,\,a_2}(x)$ achieves its maximum at $x_m=(\frac{a_1}{a_2})^{\frac{1}{a_2-a_1}}$. Substituting x_m into $f_{a_1,\,a_2}(x)$ leads to (B.4).

When $a_2 = a_1 \geq 1$, (B.4) is obvious since, in this case, $f_{a_1, a_2}(x) \equiv 0$ and $\mathcal{K}(a_1, a_2) = 0$. ∎

LEMMA B.4. *For positive definite and radially unbounded functions $V_1 : \mathbb{R}^m \to \mathbb{R}$ and $V_2 : \mathbb{R} \to \mathbb{R}$, and a positive continuous function $\Xi : \mathbb{R}^m \to \mathbb{R}^+$, $m \in \mathbb{N}$, define $V(X, x) = V_1(X) + \Xi(X)V_2(x)$. Then, for any $[X^\top, x]^\top \neq \mathbf{0}_{(m+1)\times 1}$, $V(X, x) > 0$, and in addition, if $\sqrt{\|X\|^2 + x^2} \to \infty$, $V(X, x) \to \infty$. That is, $V(X, x)$ is positive definite and radially unbounded.*

Proof. We first show the positive definiteness of $V(X, x)$. From $[X^\top, x]^\top \neq 0_{(m+1)\times 1}$ we have either $X \neq 0_{m\times 1}$ or $X = 0_{m\times 1}$ and $x \neq 0$. If $X \neq 0_{m\times 1}$, then we have $V(X, x) \geq V_1(X) > 0$. If $X = 0_{m\times 1}$ and $x \neq 0$, then by the positiveness of Ξ we have $V(0_{m\times 1}, x) \geq \Xi(0_{m\times 1})V_2(x) > 0$. Clearly, $V(\mathbf{0}_{m\times 1}, 0) = 0$. Thus, $V(X, x)$ is positive definite.

Let us next show the radial unboundedness of $V(X, x)$ by contradiction. Suppose there were a sequence of $\{X_k, x_k, k \in \mathbb{N}\}$ satisfying $\lim_{k\to\infty}(\|X_k\| + \|x_k\|) = \infty$ and a constant $C > 0$ such that $V(X_k, x_k) \leq C < \infty$ $\forall k \in \mathbb{N}$. Then, there would be $V_1(X_k) \leq C$ $\forall k \in \mathbb{N}$ and $\Xi(X_k)V_2(x_k) \leq C$ $\forall k \in \mathbb{N}$. Noticing the positive definiteness and radial unboundedness of V_1, one can show that there is a constant $\delta_1(C) > 0$ such that

$$\|X_k\| \leq \delta_1(C) \quad \forall k \in \mathbb{N}. \tag{B.5}$$

Let $M = \min_{\|X\|\leq\delta_1(C)} \Xi(X)$. Then, by the positiveness and continuity of Ξ we have $M > 0$. Thus, $V_2(x_k) \leq \frac{C}{M}$ $\forall k \in \mathbb{N}$. This together with the positive definiteness and radial unboundedness of V_2 in turn implies that there exists a constant $\delta_2(C/M) > 0$ such that

$$\|x_k\| \leq \delta_2(C/M) \quad \forall k \in \mathbb{N}. \tag{B.6}$$

From this and (B.5) we have $\|X_k\| + \|x_k\| \leq \delta_1(C) + \delta_2(C/M) < \infty$ $\forall k \in \mathbb{N}$, which contradicts $\lim_{k\to\infty}(\|X_k\| + \|x_k\|) = \infty$. Thus, $V(X, x)$ is radially unbounded. ∎

Acknowledgments. The authors wish to thank Professor Zigang Pan for helpful discussions on this topic. The authors are also indebted to the anonymous referees for their valuable comments.

REFERENCES

[1] G. Arslan and T. Başar, *Risk-sensitive adaptive trackers for strict-feedback systems with output measurements*, IEEE Trans. Automat. Control, 47 (2002), pp. 1754–1758.
[2] A. Bensoussan, *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge, UK, 1992.
[3] A. Bensoussan, *General finite-dimensional risk-sensitive problems and small noise limits*, IEEE Trans. Automat. Control, 41 (1996), pp. 210–215.
[4] A. Bensoussan and J. H. van Schuppen, *Optimal control of partially observable stochastic systems with an exponential-of-integral performance index*, SIAM J. Control Optim., 23 (1985), pp. 599–613.
[5] H. Deng and M. Krstić, *Output-feedback stochastic nonlinear stabilization*, IEEE Trans. Automat. Control, 44 (1999), pp. 328–333.
[6] H. Deng and M. Krstić, *Output-feedback stabilization of stochastic nonlinear systems driven by noise of unknown covariance*, Systems Control Lett., 39 (2000), pp. 173–182.
[7] H. Deng, M. Krstić, and R. Williams, *Stabilization of stochastic nonlinear systems driven by noise of unknown covariance*, IEEE Trans. Automat. Control, 46 (2001), pp. 1237–1253.
[8] K. Ezal, Z. Pan, and P. V. Kokotović, *Locally optimal backstepping design*, IEEE Trans. Automat. Control, 45 (2000), pp. 260–271.

[9] W. H. Fleming and W. M. McEneaney, *Risk-sensitive control and differential games*, in Stochastic Theory and Adaptive Control, Lecture Notes in Control and Inform. Sci. 184, Springer-Verlag, New York, 1992, pp. 185–197.

[10] W. H. Fleming and W. M. McEneaney, *Risk-sensitive control on an infinite time horizon*, SIAM J. Control Optim., 33 (1995), pp. 1881–1915.

[11] P. Florchinger, *Lyapunov-like techniques for stochastic stability*, SIAM J. Control Optim., 33 (1995), pp. 1151–1169.

[12] J. W. Helton and M. R. James, *Extending H^{∞} Control to Nonlinear Systems: Control of Nonlinear Systems to Achieve Performance Objectives*, Adv. Des. Control 1, SIAM, Philadelphia, 1999.

[13] A. Isidori, *Nonlinear Control Systems*, 3rd ed., Springer-Verlag, London, 1995.

[14] D. H. Jacobson, *Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games*, IEEE Trans. Automat. Control, 18 (1973), pp. 167–172.

[15] S. Jain and F. Khorrami, *Application of a decentralized adaptive output feedback based on backstepping to power systems*, in Proceedings of the 34th IEEE Conference on Decision and Control, New Orleans, LA, 1995, IEEE Press, Piscataway, NJ, pp. 1585–1590.

[16] M. R. James and J. S. Baras, *Partially observed differential games, infinite-dimensional Hamilton–Jacobi–Isaacs equations, and nonlinear H_{∞} control*, SIAM J. Control Optim., 34 (1996), pp. 1342–1364.

[17] M. R. James, J. Baras, and R. J. Elliott, *Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems*, IEEE Trans. Automat. Control, 39 (1994), pp. 780–792.

[18] Z. P. Jiang and H. Nijmeijer, *A recursive technique for tracking control of nonholonomic systems in chained form*, IEEE Trans. Automat. Control, 44 (1999), pp. 265–279.

[19] Z. P. Jiang and J. B. Pomet, *Backstepping-based adaptive controllers for uncertain nonholonomic systems*, in Proceedings of the 34th IEEE Conference on Decision and Control, New Orleans, LA, 1995, IEEE Press, Piscataway, NJ, pp. 1573–1578.

[20] I. Kanellakopoulos, P. V. Kokotović, and A. S. Morse, *Systematic design of adaptive controllers for feedback linearizable systems*, IEEE Trans. Automat. Control, 36 (1991), pp. 1241–1253.

[21] R. Z. Kha'minslii, *Stochastic Stability of Differential Equations*, S & N International Publishers, Rockville, MD, 1980.

[22] P. Kokotović and M. Arcak, *Constructive nonlinear control: A historical perspective*, Automatica, 37 (2001), pp. 637–662.

[23] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović, *Nonlinear design of adaptive controllers for linear systems*, IEEE Trans. Automat. Control, 39 (1994), pp. 738–752.

[24] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović, *Nonlinear and Adaptive Control Design*, John Wiley & Sons, New York, 1995.

[25] W. Lin and R. Pongvuthithum, *Adaptive output tracking of inherently nonlinear systems with nonlinear parameterizations*, IEEE Trans. Automat. Control, 48 (2003), pp. 1737–1749.

[26] Y. Liu, Z. Pan, and S. Shi, *Output feedback control design for strict-feedback stochastic nonlinear systems under a risk-sensitive cost criterion*, IEEE Trans. Automat. Control, 48 (2003), pp. 509–513.

[27] Y. Liu and J. Zhang, *Reduced-order observer-based control design for stochastic nonlinear systems*, Systems Control Lett., 52 (2004), pp. 123–135.

[28] Y. Liu and J. Zhang, *Minimal-order observer and output-feedback stabilization control design of stochastic nonlinear systems*, Science in China (Series F), 47 (2004), pp. 527–544.

[29] Y. Liu, J. Zhang, and Z. Pan, *Design of satisfaction output feedback controls for stochastic nonlinear systems under quadratic tracking risk-sensitive index*, Sciences in China (Series F), 46 (2003), pp. 126–145.

[30] R. Marino and P. Tomei, *Global adaptive output-feedback control of nonlinear systems, Part* I: *Linear parameterization*, IEEE Trans. Automat. Control, 38 (1993), pp. 17–32.

[31] H. Nagai, *Bellman equations of risk-sensitive control*, SIAM J. Control Optim., 34 (1996), pp. 74–101.

[32] Z. Pan and T. Başar, *Adaptive controller design for tracking and disturbance attenuation in parametric-feedback nonlinear systems*, IEEE Trans. Automat. Control, 43 (1998), pp. 1066–1083.

[33] Z. Pan and T. Başar, *Backstepping controller design for nonlinear stochastic systems under a risk-sensitive cost criterion*, SIAM J. Control Optim., 37 (1999), pp. 957–995.

[34] Z. Pan, K. Ezal, A. Krener, and P. V. Kokotović, *Backstepping design with local optimality matching*, IEEE Trans. Automat. Control, 46 (2001), pp. 1014–1027.

[35] Z. Pan, Y. Liu, and S. Shi, *Output feedback stabilization for stochastic nonlinear systems in observer canonical form with stable zero-dynamics*, Science in China (Series F), 44 (2001), pp. 292–308.

[36] Z. Pan, Y. Liu, and S. Shi, *Output feedback stabilization for stochastic nonlinear systems in observer canonical form with stable zero-dynamics*, in Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, NV, 2002, IEEE Press, Piscataway, NJ, pp. 1392–1397.

[37] T. Runolfsson, *The equivalence between infinite horizon control of stochastic systems with exponential-of-integral performance index and stochastic differential games*, IEEE Trans. Automat. Control, 39 (1994), pp. 1551–1563.

[38] D. Seto, A. M. Annaswamy, and J. Baillieul, *Adaptive control of nonlinear systems with triangular structure*, IEEE Trans. Automat. Control, 39 (1994), pp. 1411–1428.

[39] P. Whittle, *Risk-sensitive linear-quadratic-Gaussian control*, Adv. Appl. Probab., 13 (1981), pp. 764–777.

[40] P. Whittle, *Risk-Sensitive Optimal Control*, John Wiley & Sons, New York, 1990.