ISSN 1344-8803, CSSE-7 January 31, 2000

Generalized State-Dependent Scaling for Local Optimality, Global Inverse Optimality, and Global Robust Stability[1][2]

Hiroshi Ito[†3]

[†]Department of Control Engineering and Science, Kyushu Institute of Technology
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
Phone: (+81)948-29-7717, Fax: (+81)948-29-7709
E-mail: hiroshi@ces.kyutech.ac.jp

Abstract: This paper provides a solution to an inverse optimal robust control problem for uncertain nonlinear systems. A new version of robust backstepping is proposed in which inverse optimality is achieved through the selection of generalized state-dependent scaling factors. Like other robust backstepping methods, this design is always successful for uncertain nonlinear systems in strict-feedback form. The class of cost functionals allowed in the inverse optimal design is such that the uncertainty structure and desired level of global robustness can be prescribed *a priori*. Furthermore, the inverse optimal control law can always be designed such that its linearization is identical to a linear optimal control law for the linearized system with respect to a prescribed quadratic cost functional.

Key Words: Inverse optimal control, Robust backstepping, Global robust stability, State-dependent scaling, Robust control Lyapunov function

[1]Technical Report in Computer Science and Systems Engineering, Log Number CSSE-7, ISSN 1344-8803.

[2]The first version of the paper was completed by November 1, 1998 during a six months visit to Dept. of Electrical and Computer Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3118, USA. This version was completed on March 4, 1999.

[3]Author for correspondence

1 Introduction

Backstepping methods for the design of robustly stabilizing controllers for uncertain nonlinear systems have been evolving over the past several years. Early results in [1, 2, 3] extended the breakthrough [4] in nonlinear adaptive control to a class of strict-feedback systems with nonlinear time-varying uncertainties[5]. More recently, these designs have been refined in the context of an inverse optimal robust control problem in which the resulting controller is optimal with respect to some cost functional belonging to a specified class [6, 7, 8, 9]. A primary benefit of achieving inverse optimality is that, for an appropriately chosen class of cost functionals, it can guarantee additional robustness with respect to certain types of input uncertainties not taken explicitly into account in the robust controller design [10, 7]. The inverse optimal design represents an attractive alternative to the generally infeasible approach of directly computing an optimal controller with respect to a pre-specified cost functional by solving a Hamilton-Jacobi-Bellman (HJB) or Hamilton-Jacobi-Isaacs (HJI) partial differential equation.

A drawback of these inverse optimal backstepping designs is that they do not necessarily provide the same local performance achieved by linear controllers constructed for the linearized system using modern robust control methods such as $\mathcal{H}_\infty$. Ideally one would attempt to design the nonlinear controller not only to achieve global robustness but also to recover the local linear performance. One such design was presented in [8]; here the controllers guarantee local $\mathcal{L}_2$-disturbance attenuation of level γ as well as global inverse optimality and input-to-state stability (ISS). The local property is achieved by forcing the inverse optimal cost functional to agree locally with a pre-specified $\mathcal{H}_\infty$-type cost for the linearized system. However, the parameter γ is not given a *global* interpretation in [8]; the global robustness property is simply stated as input-to-state stability with no reference to γ.

In this paper we consider a global robust stabilization problem in which the parameter γ plays both a local and global role in describing the allowable "size" of the structured uncertainty. In this context global input-to-state stability becomes just a special case of global robustness with respect to a particular class of structured uncertainty. We first generalize the class of cost functionals employed in [8] in the definition of inverse optimality to reflect this new global role of the parameter γ. We then show that for uncertain nonlinear systems in strict feedback form, we can always achieve such extended inverse optimality through the use of generalized state-dependent (SD) scaling. This notion of generalized SD scaling is an extension of the SD scaling developed for robust nonlinear control in [11, 12, 13, 14]. The control system obtained through our generalized SD scaling design is guaranteed to have a global robustness property with respect to structured uncertainty with size described by the parameter γ. Furthermore, the inverse optimal control law can always be designed such that its linearization is identical to a linear optimal control law for the linearized system with respect to a prescribed cost functional (again parameterized by γ). By incorporating the parameter γ into our definition of global robust stability, we achieve a seamless integration of the three desired properties of our nonlinear design: local optimality, global inverse optimality, and global robust stability. Finally, it is worth noting that the generalized SD scaling approach to inverse optimal backstepping is a natural extension of the idea of scaling designs popular in linear robust control theory.

2 Motivation and problem statement

We consider a nonlinear plant with control input u and disturbance w of the form

$$\dot{x} = f(x) + f_1(x)w + f_2(x)u \tag{1}$$

where f, f_1 and f_2 are sufficiently smooth vector fields on $\mathcal{R}^n$ with $f(0)=0$, $f_1(0)=0$ and $f_2(0)=0$. An optimal control problem for (1) is described by a cost functional

$$J_{opt} = \min_u \max_w \int_0^\infty L(x,w,u)dt \tag{2}$$

where L is a smooth function satisfying $L(0,0,0)=0$. The Hamilton-Jacobi(HJ) equation for this optimal control problem is written somewhat informally as

$$\min_u \max_w \left[L(x,w,u) + \dot{V}(x,w,u) \right] = 0, \quad V(0)=0 \tag{3}$$

with a smooth positive semi-definite function $V(x)$. For an appropriate choice of $L(x,w,u)$, the solution V will lead to a control law $u(x)$ which provides optimality, stability, and robustness with respect to the disturbance w.

The disturbance w in (1) usually represents uncertainty arising from various locations in the control system. To specify the location and structure of the uncertainty, the robust control paradigm depicted Fig.1 is popular. The system Δ is an uncertainty and K is a state-feedback controller. The system G not only describes a nominal plant, but also includes information about how the uncertainty affects the nominal plant such as geometrical locations and types of nonlinearities where uncertain parameters are present. This uncertainty structure is described by the input map in G at w and the output map in G at z. The principal problem of robust control for the uncertain system is to construct a controller K achieving global robust stability:

Global robust stability

If the closed-loop system is globally stable in the presence of every Δ belonging to a given family of admissible uncertainties, the system is said to be globally robustly stable.

The structure of the uncertainty defines the set of disturbances w in (1), (2) and (3). It is well known that such a robust control design usually comes at the price of solving a Hamilton-Jacobi partial differential equation (3) with an appropriate function L. Such a task is generally not feasible.

This infeasibility of solving HJ equations motivated the development of the inverse optimality direction in robust nonlinear control [6]. Roughly speaking, an inverse optimal design is an optimal design in which the choice for the function L is left open:

Global inverse optimality

If there exist a function L such that (3) has a positive definite and radially unbounded solution $V(x)$ on $\mathcal{R}^n$, then the control u achieving minimum of (3) is said to be globally inverse optimal.

Inverse optimality exploits the fact that an HJ equation with a fixed function L is only a sufficient condition for achieving robustness. The inverse optimal design seeks a particular solution to an appropriate class of HJ equations which lead to desired robustness of a system by using L as a design parameter. It is important to remember that the inverse approach enables us to design robust controllers without solving HJ equations directly, provided the function L is chosen from an appropriate class.

According to Lyapunov's first method, we can achieve stability robustness in Fig.1 at least locally by using a linearized description of the system. Fortunately, an HJ equation for a linearized system reduces to an algebraic Riccati equation which can be solved easily and efficiently. The optimal controller directly guarantees the stability robustness in Fig.1 for linear G and families of Δ including memoryless and dynamic, structured and unstructured uncertainties. Thus, for achieving robustness at least locally, the following property is desirable:

Local optimality

Suppose that a function L is given *a priori* so that it describes desired stability robustness. If the linearization of a control law u at $x = 0$ is an optimal control associated a value function which solves the linear version of (3) (i.e., Riccati equation), the control law u is said to be locally optimal.

However, a controller that performs well on a linearized model may drastically reduce the stability region of the actual nonlinear system. Local optimality by itself is not sufficient for inherently nonlinear systems.

If maximization with respect to w is dropped in the global inverse optimal design, then global optimality does not necessarily imply global robust stability. For example, consider the following choice for L which is a natural generalization of LQ optimal control:

$$L = q(x) + r(x)u^2, \qquad q(x) \geq 0,\ q(0) = 0, \quad r(x) > 0 \quad \forall x \in \mathcal{R}^n \tag{4}$$

with optimality described by

$$J_{opt} = \min_u \int_0^\infty L(x, w, u)dt \ \text{ subject to } \ w \equiv 0 \tag{5}$$

instead of (2). A global inverse optimal controller with respect to (5) may exhibit robustness properties in the sense of stability margins [10], but it generally only guarantees stability robustness with respect to certain types of input uncertainties. In fact, in general the optimal controller does not achieve robustness for the disturbance w as specified in Fig.1. In contrast, for the same function L in (4), global optimality in terms of (2) implies robust stability of the system in Fig.1 because $J_{opt} < \infty$ implies $x(t) \to 0$ as $t \to \infty$. It is known that robust stabilizability of the system with the family of memoryless uncertainties Δ is equivalent to the existence of a robust control Lyapunov function (rclf). In fact, it has been shown in [6] that every rclf solves an HJ equation associated with the cost functional (2) and the inverse optimal control law is given in terms of the rclf. The inverse optimal control in [6] thus takes robust stability into account. However, the relationship between the global optimality achieved and robust stabilization of the linearized system is not clear. Indeed, in general, the globally inverse optimal control law is not locally optimal.

Another candidate for the function L in (2) for inverse optimal control is as follows:

$$L = q(x) + r(x)u^2 - \gamma^2 w^2, \qquad q(x) \geq 0,\ q(0) = 0, \quad r(x) > 0 \quad \forall x \in \mathcal{R}^n \tag{6}$$

Because this L is a natural generalization of linear $\mathcal{H}^\infty$ optimal control, it has been shown in [8] that the associated globally inverse optimal controllers can be locally $\mathcal{H}^\infty$ optimal for a class of nonlinear systems. However, because of the negative term in L, global inverse optimality does *not* necessarily assure global robust stability in the presence of the disturbance w. In [7], a terminal penalty is introduced which excludes this possibility of destabilizing inverse optimal controllers:

$$J_{opt} = \min_u \max_w \left\{ \lim_{T\to\infty} \left[E(x(T)) + \int_0^T q(x) + r_2(x)u^2 - r_1(x)w^2 dt \right] \right\} \tag{7}$$
$$q(x) \geq 0,\ q(0) = 0, \quad r_1(x) \geq 0, \quad r_2(x) > 0, \quad \forall x \in \mathcal{R}^n$$

Here, E is a positive definite radially unbounded function. It was proved in [7] that for a class of input-to-output stabilizable systems, input-to-output stability can be achieved in the presence of input unmodeled uncertainty. However, this type of global stability is guaranteed for a certain class of unstructured input uncertainties which in general does not conform to the location/structure specified in Fig.1. Furthermore, the inverse optimal control in [7] is not guaranteed to be locally optimal.

The purpose of this paper is to develop a backstepping procedure which meets all the three objectives simultaneously: global inverse optimality, local optimality, and global robust stability. For this purpose, an appropriate choice of L will given by generalized state-dependent scaling.

3 Definitions

Consider the uncertain nonlinear system Σ described by

$$\Sigma : \dot{x} = A(x)x + B(x)w + G(x)u \ . \tag{8}$$

where dimensions of signals are $x(t) \in \mathcal{R}^n$, $u(t) \in \mathcal{R}^1$, $w(t) \in \mathcal{R}^p$. Functions $A(x)$, $B(x)$ and $G(x)$ are assumed to be sufficiently smooth. We make the following structural assumptions on these matrices. First, we assume that A, B and G can be written in the form

$$A(x) = \begin{bmatrix} a_{11}(x) & a_{12}(x) & 0 & \cdots\cdots & 0 \\ a_{21}(x) & a_{22}(x) & a_{23}(x) & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots\ \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n-1,1}(x) & a_{n-1,2}(x) & \cdots & \cdots & a_{n-1,n}(x) \\ a_{n1}(x) & a_{n2}(x) & \cdots & \cdots\cdots & a_{nn}(x) \end{bmatrix} \tag{9}$$

$$B(x) = \begin{bmatrix} B_{11}(x) & 0 & \cdots & 0 \\ B_{21}(x) & B_{22}(x) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ B_{n1}(x) & B_{n2}(x) & \cdots & B_{nn}(x) \end{bmatrix}, \ G(x) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_{n,n+1}(x) \end{bmatrix}. \tag{10}$$

Each scalar-valued function a_{ij} is required to satisfy

$$a_{ij}(x) = a_{ij}(x_1, x_2, \cdots, x_i), \quad 1 \leq i \leq n,\ 1 \leq j \leq i+1 \tag{11}$$
$$a_{i,i+1}(x_1, x_2, \cdots, x_i) \neq 0, \quad 1 \leq i \leq n \tag{12}$$

for all $x \in \mathcal{R}^n$. The dimension for the matrix $B(x)$ is $B_{ij}(x) \in \mathcal{R}^{1\times p_i}$ and $p = \sum_{i=1}^n p_i$, $p_i \geq 0$. Its dependence on x is

$$B_{ij}(x) = B_{ij}(x_1, x_2, \cdots, x_i), \quad 1 \leq i \leq n, \quad 1 \leq j \leq i \ . \tag{13}$$

Definition 1 *Given a positive semi-definite, radially unbounded,* $\mathbf{C}^1$ *function* $V(x)$ *and a positive scalar-valued function* $r(x)$*, the control law*

$$u(x) = -\frac{1}{2} r^{-1} (L_G V)^T \tag{14}$$

is globally inverse optimal with disturbance level γ *when*

(i) *The equilibrium $x = 0$ of Σ is globally asymptotically stabilized by (14) when $w \equiv 0$.*

(ii) *There exist a scalar-valued function $q(x)$ and a $\mathcal{R}^{p\times p}$-valued function $\Theta(x)$ such that*

$$
\begin{aligned}
&L_{Ax}V + \frac{1}{4\gamma^2}L_B V\Theta^{-1}(L_B V)^T - \frac{1}{4}L_G V r^{-1}(L_G V)^T + q = 0 \\
&q(x) \geq 0, \quad r(x) > 0, \quad \Theta(x) = \Theta^T(x) > 0, \quad \forall x \in \mathcal{R}^n \ .
\end{aligned}
\tag{15}
$$

The control law (14) with a solution to (15) minimizes the worst-case value of the cost functional

$$
J(u,w) = \int_0^\infty \left[q(x) + r(x)u^2 - \gamma^2 w^T\Theta(x)w\right] dt \tag{16}
$$

subject to the disturbance w over all stabilizing control laws, i.e., $\min_u \max_{w\in\mathcal{L}_2} J(u,w)$ is achieved. The function $V(x)$ is the optimal value function[15, 16]. The disturbance penalty in the cost functional (16) was refereed to as state-dependent weighting in [7]. It has been proved that if (8) is input-to-state stabilizable, then the inverse optimal problem is solvable with respect to (16)[7]. Note that the function $V(x)$ satisfying the Hamilton-Jacobi-Isaacs(HJI) equation (15) is positive definite if $q(x) \equiv 0$ implies $x \equiv 0$. In the cost functional (16), since Θ can absorb γ, the roll of γ is not clear in Definition 1. The meaning of γ in connection with robustness will be discussed in Section 7. When $p = 0$, the inverse optimal control defined in Definition 1 reduces to the inverse optimality without disturbance [6, 10].

We now consider Jacobian linearization of Σ as follows:

$$
\dot{x} = A_l x + B_l w_l + G_l u_l \tag{17}
$$

$$
A_l = \left.\frac{\partial f}{\partial x}\right|_{x=0} = A(0), \quad B_l = B(0), \quad G_l = G(0) \ . \tag{18}
$$

We assume that $Q_l = Q_l^T > 0$, $r_l > 0$ and $\Theta_l = \Theta_l^T > 0$. Then, since (A_l, G_l) is controllable by our assumptions on $A(x)$ and $G(x)$, there exists a positive number γ_* such that

$$
A_l^T P_l + P_l A_l + P_l\left(\frac{1}{\gamma^2}B_l\Theta^{-1}B_l^T - G_l r_l^{-1}G_l^T\right)P_l + Q_l = 0 \tag{19}
$$

has a unique solution $P_l = P_l^T > 0$ for $\gamma > \gamma_*$ and the control law

$$
u_l = -r_l^{-1}G_l^T P_l x \tag{20}
$$

stabilizes the linear system (17). The control law (20) with the solution $P_l > 0$ to the Riccati equation (19) achieves $\min_{u_l}\max_{w_l\in\mathcal{L}_2} J_l(u_l, w_l)$ for the cost functional

$$
J_l(u_l, w_l) = \int_0^\infty \left[x^T Q_l x + r_l u_l^2 - \gamma^2 w_l^T \Theta_l w_l\right] dt \tag{21}
$$

over all stabilizing control laws u_l. Here, $V_l(x) = x^T P_l x$ is the optimal value function.

Definition 2 *Suppose that $Q_l = Q_l^T > 0$, $r_l > 0$ and $\Theta_l = \Theta_l^T > 0$ are given and γ is chosen as $\gamma > \gamma_*$. The control law (14) is said to be locally optimal and globally inverse optimal with disturbance level γ if*

(i) *the control (14) is globally inverse optimal with disturbance level γ.*

(ii) *the Jacobian linearized control law*

$$u = -\frac{1}{2} r^{-1}(0) G^T(0) \left. \frac{\partial^2 V}{\partial x^2} \right|_{x=0} x \tag{22}$$

stabilizes the linear system (17) and achieves $\min_u \max_{w \in \mathcal{L}_2} J_l(u, w)$.

(iii) $q(x) \geq 0$, $r(x) > 0$ *and* $\Theta(x) = \Theta^T(x) > 0$ *satisfy*

$$2Q_l = \left. \frac{\partial^2 q(x)}{\partial x^2} \right|_{x=0}, \quad r_l = r(0), \quad \Theta_l = \Theta(0) \tag{23}$$

If $\Theta(x)$ is restricted to an identity matrix for all $x \in \mathcal{R}^n$, the above definition reduces to the local optimality and global inverse optimality defined in [8].

Without loss of generality, $\Theta_l = \Theta(0)$ is assumed to be a diagonal matrix throughout this paper. In fact, we can always replace any $\Theta_l > 0$ with an identity matrix as follows. Decompose Θ_l into $\Theta_l = W^T W$ with a lower triangular matrix W by using the Cholesky factorization. By defining $\bar{w}_l = W w_l$, the w_l-term in (21) becomes $\gamma^2 w_l^T \Theta_l w_l = \gamma^2 \bar{w}_l^T \bar{w}_l$. Since W is lower triangular, $B(x)W^{-1}$ is again in the block lower triangular form of (10). Consider the inverse optimal problem of

$$J(u, \bar{w}) = \int_0^\infty \left[q(x) + r(x) u^2 - \gamma^2 \bar{w}^T \bar{\Theta}(x) \bar{w} \right] dt \tag{24}$$

for the system in which $B(x)w$ of Σ is replaced with $B(x)W^{-1}\bar{w}$. This problem is the same as the inverse optimal defined with (16) for the original system Σ with respect to $\Theta(x) = W^T \bar{\Theta}(x) W$, while in the optimal control with respect to (24), we have $\bar{\Theta}(0) = W^{-T}\Theta(0)W^{-1} = I_p$ as desired.

4 Global inverse optimality via generalized SD scaling

We now consider the following weighting functions for the cost J in (16):

$$q(x) = x^T \bar{C}^T(x) \bar{\Lambda}(x) \bar{C}(x) x + x^T \bar{Q}(x) x \tag{25}$$

$$r(x) = \lambda_U(x), \quad \Theta(x) = \Theta^T(x) \tag{26}$$

$$\bar{\Lambda}(x) > 0, \ \bar{Q}(x) = \bar{Q}^T(x) > 0, \ \lambda_U(x) > 0, \ \Theta(x) > 0, \ \forall x \in \mathcal{R}^n$$

The matrix $\bar{C}(x) \in \mathcal{R}^{q \times n}$ is a prescribed function and it is assumed to have the form

$$\bar{C}(x) = \begin{bmatrix} C_{11}(x) & 0 & \cdots & 0 \\ C_{21}(x) & C_{22}(x) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ C_{n1}(x) & C_{n2}(x) & \cdots & C_{nn}(x) \end{bmatrix}, \quad \begin{matrix} C_{ij}(x) = C_{ij}(x_1, x_2, \cdots, x_i) \\ 1 \leq i \leq n, \quad 1 \leq j \leq i \end{matrix} \tag{27}$$

for sufficiently smooth functions $C_{ij}(x) \in \mathcal{R}^{q_i \times 1}$, $q_i \geq 0$ and $q = \sum_{i=1}^n q_i$. The matrices $\bar{\Lambda}(x)$ and $\Theta(x)$ are parameters to be used for achieving the inverse optimality of Definition 1. The matrices are assumed to consist of scalar-valued functions $\lambda_i(x)$ and $\theta_i(x)$, $i = 1, 2, \ldots, n$ with

$$\bar{\Lambda}(x) = \text{block-diag}[\lambda_1(x) I_{q_1}, \lambda_2(x) I_{q_n}, \cdots, \lambda_n(x) I_{q_n}] \tag{28}$$

$$\Theta(x) = \text{block-diag}[\theta_1(x) I_{p_1}, \theta_2(x) I_{q_n}, \cdots, \theta_n(x) I_{p_n}] \ . \tag{29}$$

Here, I_{q_i} denotes a $q_i \times q_i$ identity matrix. Each λ_i or θ_i is a function of x of the form

$$\lambda_i(x) = \lambda_i(x_1, x_2, \cdots, x_i), \;\; \theta_i(x) = \theta_i(x_1, x_2, \cdots, x_i), \;\; 1 \le i \le n$$

The scalar-valued function λ_U is also a parameter to be chosen in the inverse optimal design and its dependence on x is

$$\lambda_U(x) = \lambda_U(x_1, x_2, \cdots, x_n) \;.$$

We assume that

$$\lambda_i(x) > 0, \quad \theta_i(x) > 0, \; \forall i, \quad \lambda_U(x) > 0, \quad \forall x \in \mathcal{R}^n \;.$$

Substituting the above parameters into (16), the cost function $J(u, w)$ becomes

$$J(u, w) = \int_0^\infty \left[z^T \Lambda(x) z + x^T \bar{Q}(x) x - \gamma^2 w^T \Theta(x) w \right] dt \tag{30}$$

with an augmented system

$$\Sigma_a : \begin{cases} \dot{x} = A(x)x + B(x)w + G(x)u \\ z = C(x)x + Hu \end{cases} , \tag{31}$$

where

$$C(x) = \begin{bmatrix} C_{11}(x) & 0 & \cdots & 0 \\ C_{21}(x) & C_{22}(x) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ C_{n1}(x) & C_{n2}(x) & \cdots & C_{nn}(x) \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{32}$$

$$\Lambda(x) = \text{block-diag}[\bar{\Lambda}(x), \lambda_U(x)] \;. \tag{33}$$

The cost functional (30) can be rewritten as

$$J(u, w) = \int_0^\infty \left[\sum_{i=1}^n (\lambda_i(x) z_i^T z_i - \gamma^2 \theta_i(x) w_i^T w_i) + x^T \bar{Q}(x) x + \lambda_U(x) u^2 \right] dt \tag{34}$$

where w and z are partitioned as

$$w = [w_1^T, \cdots, w_n^T]^T, \;\; w_i \in \mathcal{R}^{p_i}, \;\; z = [z_1^T, \cdots, z_n^T, u]^T, \;\; z_i \in \mathcal{R}^{q_i} \;.$$

When we choose $\theta_i = \lambda_i$ for $i = 1, 2, \ldots, n$ in (34), the parameter $\lambda_i(x)$ becomes the state-dependent(SD) scaling [12]. Therefore, in this paper, we call $(\Lambda(x), \Theta(x))$ the generalized state-dependent scaling for inverse optimal control.

Let $x_{[k]}$ denote the states x_1 through x_k.

$$x_{[k]} = [\, x_1 \;\; x_2 \;\; \cdots \;\; x_k \,]^T$$

We consider a diffeomorphism $\chi = S(x)x$ between $x \in \mathcal{R}^n$ and $\chi \in \mathcal{R}^n$. Let $S^{-1}(x)$ denote the inverse map of the diffeomorphism and choose

$$S^{-1}(x) = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ s_1(x) & 1 & 0 & 0 & \cdots & 0 \\ d_{21} & s_2(x) & 1 & 0 & \ddots & 0 \\ d_{31} & d_{32} & s_3(x) & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ d_{n-1,1} & \cdots & \cdots & d_{n-1,n-2} & s_{n-1}(x) & 1 \end{bmatrix} , \tag{35}$$

where the smooth scalar functions $s_1(x_1)$, $s_2(x_{[2]})$, $\cdots$, $s_{n-1}(x_{[n-1]})$ are to be determined in a recursive manner from s_1 through s_{n-1}. Each function s_i depends only on the state components x_1 through x_i. The other scalar constants d_{ij}, $2 \leq i \leq n-1$, $1 \leq j \leq i-1$, are any real numbers. Then, $S(x)$ is

$$S(x) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \triangle_1 & 1 & 0 & \cdots & 0 \\ \triangle_2 & \triangle_2 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \triangle_{n-1} & \cdots & \cdots & \triangle_{n-1} & 1 \end{bmatrix}$$

where $\triangle_j$ denotes any function depending only on s_1 through s_j. This diffeomorphism $S(x)$ is the same as the one in [6] and [14] if we take $d_{ij} = 0$. The time-derivative of χ is

$$\dot{\chi} = \left[\frac{\partial S}{\partial x_1}x, \frac{\partial S}{\partial x_2}x, \cdots, \frac{\partial S}{\partial x_n}x\right]\dot{x} + S(x)\dot{x} = T(x)\dot{x} \ . \tag{36}$$

with a smooth function $T(x)$:

$$T(x) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \star_{1,1} & 1 & 0 & \cdots & 0 \\ \star_{2,2} & \star_{2,2} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \star_{n-1,n-1} & \cdots & \cdots & \star_{n-1,n-1} & 1 \end{bmatrix} .$$

The entries $\star_{i,j}$ depend only on the states x_1 through x_i and the functions s_1 through s_j and their partial derivatives. We now consider a state-feedback law

$$u(x) = s_n(x_{[n]})\chi_n \tag{37}$$

where s_n is another smooth function yet to be determined. Then, the closed-loop system consisting of (31) and the state-feedback law becomes

$$\Sigma_{cl} : \begin{cases} \dot{\chi} = T\left(\hat{A}\hat{S}\chi + Bw\right) \\ z = \hat{C}\hat{S}\chi \end{cases} \tag{38}$$

$$\hat{S} := \left[\begin{array}{c|c} \multicolumn{2}{c}{S^{-1}} \\ \hline 0\cdots 0 & s_n \end{array}\right], \quad \hat{A} := [A \ G], \quad \hat{C} := [C \ H].$$

Theorem 1 *Suppose that $P = \mathrm{diag}[P_1, P_2, \cdots, P_n]$ is a constant diagonal matrix and*

$$M(x) := \begin{bmatrix} \hat{S}^T\hat{A}^T T^T P + PT\hat{A}\hat{S} & \gamma^{-1}PTB & \hat{S}^T\hat{C}^T\Lambda \\ \gamma^{-1}B^T T^T P & -\Theta & 0 \\ \Lambda\hat{C}\hat{S} & 0 & -\Lambda \end{bmatrix} < 0 \tag{39}$$

$$P > 0 \tag{40}$$

are satisfied for all $x \in \mathcal{R}^n$, Then, the nonlinear system Σ is globally asymptotically stabilized by the state-feedback law $u = s_n\chi_n$ and

$$\begin{aligned} J(u, w, T) \ &= \int_0^T \left[z^T\Lambda(x)z + x^T\bar{Q}(x)x - \gamma^2 w^T\Theta(x)w\right] dt \\ &\leq V(x(0)) - V(x(T)) \end{aligned} \tag{41}$$

is achieved for all $w \in \mathcal{L}_2[0,T]$ *and for all* $T \geq 0$ *with*

$$\bar{Q}(x) = -S^T(\hat{S}^T\hat{A}^T T^T P + PT\hat{A}\hat{S} + \frac{1}{\gamma^2}PTB\Theta^{-1}B^T T^T P + \hat{S}^T\hat{C}^T\Lambda\hat{C}\hat{S})S > 0, \quad \forall x \in \mathcal{R}^n \quad (42)$$

$$V(x) = x^T S^T P S x \quad (43)$$

In addition, if $s_n = -P_n a_{n,n+1}/\lambda_U$ *is satisfied for all* $x \in \mathcal{R}^n$*, the state-feedback law*

$$u = -\frac{1}{2}r^{-1}(L_G V)^T = s_n \chi_n \quad (44)$$

is globally inverse optimal with disturbance level γ*. The optimality is achieved with respect to the cost functional (16) for*

$$\begin{aligned} q(x) &= x^T\bar{C}^T(x)\bar{\Lambda}(x)\bar{C}(x)x + x^T\bar{Q}x, \ \forall x \in \mathcal{R}^n & (45)\\ r(x) &= \lambda_U(x) > 0, \quad \forall x \in \mathcal{R}^n & (46)\\ \Theta(x) &= \Theta^T(x) > 0, \quad \forall x \in \mathcal{R}^n \ . & (47) \end{aligned}$$

such that P*,* $\Lambda(x)$ *and* $\Theta(x)$ *satisfy (39).*
Proof : Define a positive definite function $V(x) : \mathcal{R}^n \to [0,\infty)$ by

$$V(x) = \chi^T P \chi \ . \quad (48)$$

which is radially unbounded function of x since S is globally diffeomorphic. The time-derivative of V along the trajectories of the closed-loop system (38) satisfies

$$\frac{d}{dt}V(x) = 2\chi^T PT\left(\hat{A}\hat{S}\chi + Bw\right) \quad (49)$$

From $M(x) < 0$ and $\chi = Sx$ it follows that the system (38) with $w \equiv 0$ satisfies

$$\frac{d}{dt}V(x) = \chi^T(PT\hat{A}\hat{S} + \hat{S}^T\hat{A}^T T^T P)\chi < 0, \quad \forall x \in \mathcal{R}^n \setminus \{0\} \ .$$

Hence, the system (38) is globally asymptotically stable. Now, using the Schur complement formula twice for $M < 0$ in (39), we obtain

$$\hat{S}^T\hat{A}^T T^T P + PT\hat{A}\hat{S} + \frac{1}{\gamma^2}PTB\Theta^{-1}B^T T^T P + \hat{S}^T\hat{C}^T\Lambda\hat{C}\hat{S} < 0 \ .$$

Define $\bar{Q}$ as in (42). Then $\bar{Q}(x) > 0$ holds for all $x \in \mathcal{R}^n$. From (49) we have,

$$\begin{aligned} \frac{d}{dt}V(x) + z^T\Lambda z + x^T\bar{Q}x - \gamma^2 w^T\Theta w \ = \ & \begin{bmatrix} \chi \\ w \end{bmatrix}^T \begin{bmatrix} \hat{S}^T\hat{A}^T T^T P + PT\hat{A}\hat{S} + S^{-T}\bar{Q}S^{-1} & PTB \\ B^T T^T P & 0 \end{bmatrix} \begin{bmatrix} \chi \\ w \end{bmatrix} + \\ & \begin{bmatrix} \chi \\ w \end{bmatrix}^T \begin{bmatrix} 0 & \hat{S}^T\hat{C}^T \\ I_p & 0 \end{bmatrix} \begin{bmatrix} -\gamma^2\Theta & 0 \\ 0 & \Lambda \end{bmatrix} \begin{bmatrix} 0 & I_p \\ \hat{C}\hat{S} & 0 \end{bmatrix} \begin{bmatrix} \chi \\ w \end{bmatrix} \ . \end{aligned} \quad (50)$$

More simply, we write (50) as

$$\frac{d}{dt}V(x) + z^T\Lambda z + x^T\bar{Q}x - \gamma^2 w^T\Theta w = \begin{bmatrix} \chi \\ w \end{bmatrix}^T \hat{M} \begin{bmatrix} \chi \\ w \end{bmatrix} = \begin{bmatrix} \chi \\ w \end{bmatrix}^T \begin{bmatrix} \hat{M}_{11} & \hat{M}_{12} \\ \hat{M}_{12}^T & \hat{M}_{22} \end{bmatrix} \begin{bmatrix} \chi \\ w \end{bmatrix} \ . \quad (51)$$

Note that

$$\hat{M}_{22} = -\gamma^2\Theta < 0, \quad \hat{M}_{11} - \hat{M}_{12}\hat{M}_{22}^{-1}\hat{M}_{12}^T = 0$$

hold for all $x \in \mathcal{R}^n$, where (42) was used. Applying the non-strict inequality version of Schur complement formula to $\hat{M}$, we can prove that $\hat{M} \le 0$ and

$$\frac{d}{dt}V(x) + z^T \Lambda z + x^T \bar{Q} x - \gamma^2 w^T \Theta w \le 0, \quad \forall x \in \mathcal{R}^n \tag{52}$$

Integrating (52) from $t = 0$ to $t = T$, we obtain (41). Furthermore, it is easily seen that

$$\hat{S}^T \hat{A}^T T^T P + PT\hat{A}\hat{S} = S^{-T} A^T T^T P + PTAS^{-1} + \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 2P_n a_{n,n+1} s_n \end{bmatrix}$$

$$\hat{S}^T \hat{C}^T \Lambda \hat{C} \hat{S} = S^{-T} \bar{C}^T \bar{\Lambda} \bar{C} S^{-1} + \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \lambda_U s_n^2 \end{bmatrix}, \quad PTGG^T T^T P = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & P_n^2 a_{n,n+1}^2 \end{bmatrix}.$$

If $s_n = -P_n a_{n,n+1}/\lambda_U$ is satisfied, we have $2P_n a_{n,n+1} s_n + \lambda_U s_n^2 = -P_n^2 a_{n,n+1}^2/\lambda_U$ so that the inequality (42) becomes

$$S^{-T} A^T T^T P + PTAS^{-1} + PT\left(\frac{1}{\gamma^2} B\Theta^{-1} B^T - \frac{1}{\lambda_U} GG^T\right) T^T P + S^{-T} \bar{C}^T \bar{\Lambda} \bar{C} S^{-1} = -S^{-T} \bar{Q} S^{-1} .$$

Using

$$\frac{\partial V}{\partial x} = 2x^T S^T PT, \quad \chi = Sx \tag{53}$$

we arrive at

$$-x^T \bar{Q}(x) x = L_{Ax} V + \frac{1}{4\gamma^2} L_B V \Theta^{-1} (L_B V)^T - \frac{1}{4} L_G V \lambda_U^{-1} (L_G V)^T + x^T \bar{C}^T \bar{\Lambda} \bar{C} x < 0, \quad \forall x \in \mathcal{R}^n \setminus \{0\} .$$

Finally, (53) together with $s_n = -P_n a_{n,n+1}/\lambda_U$ yields

$$-\frac{1}{2} \lambda_U^{-1} (L_G V)^T = -\lambda_U^{-1} G^T T^T P \chi = s_n \chi_n$$

This completes the proof. ∎

Using (42), the function $q(x)$ for the global optimality is alternatively represented by

$$q(x) = -x^T \left[2S^T PTA + S^T PT\left(\frac{1}{\gamma^2} B\Theta^{-1} B^T - \frac{1}{\lambda_U} GG^T\right) T^T PS\right] x > 0, \ \forall x \in \mathcal{R}^n \setminus \{0\} . \tag{54}$$

5 Inverse optimal backstepping

We first investigate the recursive structure of the inequality (39). We define

$$P_{[k]} = \operatorname*{diag}_{i=1}^{k} P_i, \quad 1 \le k \le n$$

We define system matrices for the first k states by

$$\hat{A}_{[k]}(x_{[k]}) = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{k-1,1} & a_{k-1,2} & \cdots & \cdots & a_{k-1,k} & 0 \\ a_{k1} & a_{k2} & \cdots & \cdots & a_{kk} & a_{k,k+1} \end{bmatrix}, \ B_{[k]}(x_{[k]}) = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ B_{21} & B_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ B_{k1} & B_{k2} & \cdots & B_{kk} \end{bmatrix}$$

for $1 \leq k \leq n$. We also define

$$\hat{C}_{[k]}(x_{[k]}) = \begin{bmatrix} C_{11} & 0 & \cdots & 0 \\ C_{21} & C_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ C_{k1} & C_{k2} & \cdots & C_{kk} \end{bmatrix}, \quad \begin{array}{c} 1 \leq k \leq n-1 \\ \\ \hat{C}_{[n]}(x_{[n]}) = \hat{C}(x) \end{array} .$$

In a similar manner, the functions $S_{[k]}$, $S_{[k]}^{-1}$ and $T_{[k]}$ are defined as $k \times k$ upper left parts of S, S^{-1} and T, respectively. Let $\hat{S}_{[k]}$ denote

$$\hat{S}_{[k]}(x_{[k]}) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ s_1 & 1 & 0 & \cdots & 0 \\ d_{21} & s_2 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ d_{k-1,1} & \cdots & d_{k-1,k-2} & s_{k-1} & 1 \\ d_{k,1} & \cdots & d_{k,k-2} & d_{k,k-1} & s_k \end{bmatrix} .$$

Here, we use $d_{n,j} = 0$ for $1 \leq j \leq n-1$ so that $\hat{S}_{[n]}(x_{[n]}) = \hat{S}(x)$. We next define

$$\begin{aligned} \Lambda_{[k]}(x_{[k]}) &= \text{block-diag}[\lambda_1 I_{q_1}, \lambda_2 I_{q_2}, \cdots, \lambda_k I_{q_k}] \\ & \qquad 1 \leq k \leq n-1, \quad \Lambda_{[n]}(x_{[n]}) = \Lambda(x) \\ \Theta_{[k]}(x_{[k]}) &= \text{block-diag}[\theta_1 I_{p_1}, \theta_2 I_{p_2}, \cdots, \theta_k I_{p_k}], \quad 1 \leq k \leq n . \end{aligned}$$

Here, $\Lambda_{[k]}(x_{[k]}) \in \mathcal{R}^{(q_k+l_k)\times(q_k+l_k)}$ and $\Theta_{[k]}(x_{[k]}) \in \mathcal{R}^{p_k \times p_k}$. The integer l_i is defined by $l_n = 1$ and $l_k = 0$ for $1 \leq k \leq n-1$. We also define the matrix $M_{[k]}(x_{[k]})$ by adding subscript $[k]$ to every matrix in (39). Due to the structure of $S(x)$ and Σ_a, we can prove the following.

Theorem 2 *Suppose $2 \leq k \leq n$. $M_{[k]}(x_{[k]}) < 0$ is equivalent to*

$$\begin{bmatrix} M_{[k-1]}(x_{[k-1]}) & \Phi_k(x_{[k]}) \\ \Phi_k^T(x_{[k]}) & \Psi_k(x_{[k]}) \end{bmatrix} < 0 , \tag{55}$$

where Φ_k depends only on λ_k, θ_k, (P_{k-1}, P_k) and $(s_1, \cdots, s_{k-1})$ and their partial derivatives. If $k < n$, the symmetric matrix Ψ_k depends only on λ_k, θ_k, P_k and s_k. Ψ_n depends only on λ_n, θ_n, λ_U, P_n and s_n.

Proof : We first recall that

$$P_{[k]} = \left[\begin{array}{c|c} P_{[k-1]} & 0 \\ \hline 0 & P_k \end{array}\right], \quad T_{[k]}(x_{[k-1]}) = \left[\begin{array}{c|c} T_{[k-1]}(x_{[k-2]}) & 0 \\ \hline \star_{k-1,k-1} & 1 \end{array}\right], \quad \hat{S}_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} \hat{S}_{[k-1]}(x_{[k-1]}) & \begin{array}{c} 0 \\ \vdots \\ 0 \\ 1 \end{array} \\ \hline d_{k,1} \ \cdots \ d_{k,k-1} & s_k \end{array}\right]$$

$$\hat{A}_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} \hat{A}_{[k-1]}(x_{[k-1]}) & 0 \\ \hline \star_{k,0} & a_{k,k+1} \end{array}\right], \quad B_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} B_{[k-1]}(x_{[k-1]}) & 0 \\ \hline \star_{k,0} & B_{kk} \end{array}\right], \ 2 \leq k \leq n$$

$$\hat{C}_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} \hat{C}_{[k-1]}(x_{[k-1]}) & 0 \\ \hline \star_{k,0} & 0 \end{array}\right], \quad \Lambda_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} \Lambda_{[k-1]}(x_{[k-1]}) & 0 \\ \hline 0 & \lambda_k \end{array}\right], 2 \leq k \leq n-1$$

$$\hat{C}_{[n]}(x_{[n]}) = \left[\begin{array}{c|c} \hat{C}_{[n-1]}(x_{[n-1]}) & 0 \\ \hline \star_{n,0} & 0 \\ 0 & 1 \end{array}\right], \quad \Lambda_{[n]}(x_{[n]}) = \left[\begin{array}{c|cc} \Lambda_{[n-1]}(x_{[n-1]}) & 0 & 0 \\ \hline 0 & \lambda_n & 0 \\ 0 & 0 & \lambda_U \end{array}\right]$$

$$\Theta_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} \Theta_{[k-1]}(x_{[k-1]}) & 0 \\ \hline 0 & \theta_k \end{array}\right], 2 \leq k \leq n .$$

Using the following structures

$$P_{[k]}T_{[k]}(x_{[k-1]}) = \left[\begin{array}{c|c} P_{[k-1]}T_{[k-1]}(x_{[k-2]}) & 0 \\ \hline P_k\star_{k-1,k-1} & P_k \end{array}\right], \quad \hat{A}_{[k]}\hat{S}_{[k]} = \left[\begin{array}{c|c} \hat{A}_{[k-1]}(x_{[k-1]})\hat{S}_{[k-1]}(x_{[k-1]}) & \begin{matrix} 0 \\ \vdots \\ 0 \\ a_{k-1,k} \end{matrix} \\ \hline \star_{k,k-1} & a_{kk} + a_{k,k+1}s_k \end{array}\right]$$

we have

$$P_{[k]}T_{[k]}(x_{[k-1]})\hat{A}_{[k]}(x_{[k]})\hat{S}_{[k]}(x_{[k]}) =$$
$$\left[\begin{array}{c|c} P_{[k-1]}T_{[k-1]}(x_{[k-2]})\hat{A}_{[k-1]}(x_{[k-1]})\hat{S}_{[k-1]}(x_{[k-1]}) & \begin{matrix} 0 \\ P_{k-1}a_{k-1,k} \end{matrix} \\ \hline P_k\star_{k,k-1} & P_k(a_{kk} + a_{k,k+1}s_k + \star_{k-1,k-1}) \end{array}\right]$$
$$P_{[k]}T_{[k]}(x_{[k-1]})B_{[k]}(x_{[k]}) =$$
$$\left[\begin{array}{c|c} P_{[k-1]}T_{[k-1]}(x_{[k-2]})B_{[k-1]}(x_{[k-1]}) & 0 \\ \hline \star_{k,k-1} & P_kB_{kk} \end{array}\right] .$$

In the same manner, from

$$\hat{C}_{[k]}(x_{[k]})\hat{S}_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} \hat{C}_{[k-1]}(x_{[k-1]})\hat{S}_{[k-1]}(x_{[k-1]}) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \star_{k,k-1} & C_{kk} \end{array}\right],$$

$$\hat{C}_{[n]}(x_{[n]})\hat{S}_{[n]}(x_{[n]}) = \left[\begin{array}{c|c} \hat{C}_{[n-1]}(x_{[n-1]})\hat{S}_{[n-1]}(x_{[n-1]}) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \star_{n,n-1} & C_{nn} \\ 0 & s_n \end{array}\right],$$

we obtain

$$\Lambda_{[k]}(x_{[k]})\hat{C}_{[k]}(x_{[k]})\hat{S}_{[k]}(x_{[k]}) = \left[\begin{array}{c|c} \Lambda_{[k-1]}(x_{[k-1]})\hat{C}_{[k-1]}(x_{[k-1]})\hat{S}_{[k-1]}(x_{[k-1]}) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \lambda_k\star_{k,k-1} & \lambda_kC_{kk} \end{array}\right]$$

$$\Lambda_{[n]}(x_{[n]})\hat{C}_{[n]}(x_{[n]})\hat{S}_{[n]}(x_{[n]}) = \left[\begin{array}{c|c} \Lambda_{[n-1]}(x_{[n-1]})\hat{C}_{[n-1]}(x_{[n-1]})\hat{S}_{[n-1]}(x_{[n-1]}) & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \lambda_n\star_{n,n-1} & \lambda_nC_{nn} \\ 0 & \lambda_{Un}s_n \end{array}\right] .$$

Now, consider a non-singular matrix:

$$Q_k = \begin{bmatrix} I_{k-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_1 & 0 & 0 \\ 0 & I_{\hat{p}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{p_k} & 0 \\ 0 & 0 & I_{\hat{q}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{q_k+l_k} \end{bmatrix},$$

where I_k denotes a $k \times k$ identity matrix and $\hat{p} := \sum_{i=1}^{k-1} p_i$, $\hat{q} := \sum_{i=1}^{k-1} q_i$. Multiplying $M_{[k]}$ by the nonsingular matrix from both the sides, we obtain

$$Q_k^T M_{[k]}(x_{[k]}) Q_k = \begin{bmatrix} M_{[k-1]}(x_{[k-1]}) & \Phi_k(x_{[k]}) \\ \Phi_k^T(x_{[k]}) & \Psi_k(x_{[k]}) \end{bmatrix} .$$

Here, the matrices Φ_k and Φ_k are

$$\Phi_k(x_{[k]}) = \begin{bmatrix} \star_{k,k-1} & 0 & \star_{k,k-1}\lambda_k \\ \star_{k,k-1} & 0 & 0 \\ 0 & & \\ \vdots & 0 & 0 \\ 0 & & \end{bmatrix}$$

$$\Phi_n(x_{[n]}) = \begin{bmatrix} \star_{n,n-1} & 0 & \star_{n,n-1}\lambda_n & 0 \\ \star_{n,n-1} & 0 & 0 & 0 \\ 0 & & & \\ \vdots & 0 & 0 & 0 \\ 0 & & & \end{bmatrix}$$

$$\Psi_k(x_{[k]}) = \begin{bmatrix} 2P_k(a_{kk} + a_{k,k+1}s_k + \star_{k-1,k-1}) & \gamma^{-1}P_k B_{kk} & C_{kk}^T\lambda_k \\ * & -\theta_k I_{p_k} & 0 \\ * & * & -\lambda_k I_{q_k} \end{bmatrix}$$

$$\Psi_n(x_{[n]}) = \begin{bmatrix} 2P_n(a_{nn} + a_{n,n+1}s_n + \star_{n-1,n-1}) & \gamma^{-1}P_n B_{nn} & C_{nn}^T\lambda_n & s_n\lambda_U \\ * & -\theta_n I_{p_n} & 0 & 0 \\ * & * & -\lambda_n I_{q_n} & 0 \\ * & * & * & -\lambda_U \end{bmatrix} .$$

The claims are obvious from the above matrices. ∎

Next, let $J_k \in \mathcal{R}^{1\times 1}$, $E_k \in \mathcal{R}^{1\times(p_k+q_k+l_k)}$ and $F_k \in \mathcal{R}^{(p_k+q_k+l_k)\times(p_k+q_k+l_k)}$ be defined with

$$\Psi_k(x_{[k]}) - \Phi_k^T(x_{[k]}) M_{[k-1]}^{-1}(x_{[k-1]}) \Phi_k(x_{[k]}) = \begin{bmatrix} J_k & E_k \\ E_k^T & F_k \end{bmatrix} \tag{56}$$

for $2 \le k \le n$. In the $k = 1$ case, $M_{[1]} = \Psi_1(x_1)$ holds with

$$\Psi_1(x_1) = \begin{bmatrix} J_1(x_1) & E_1(x_1) \\ E_1^T(x_1) & F_1(x_1) \end{bmatrix} \tag{57}$$

The matrices J_k, E_k and F_k are obtained as follows:

$$\begin{aligned}
&J_k(x_{[k]}) = 2P_k(a_{kk} + a_{k,k+1}s_k) + \diamondsuit_{k,k-1}, \qquad 2 \le k \le n \\
&E_k = \begin{cases} E_{k1} & \text{for } 1 \le k \le n-1 \\ \begin{bmatrix} E_{k1} & E_{k2} \end{bmatrix} & \text{for } k = n \end{cases} \\
&E_{k1} = \begin{bmatrix} \gamma^{-1}P_k B_{kk} & \diamondsuit_{k,k-1}\lambda_k \end{bmatrix}, 2 \le k \le n, \quad E_{n2} = s_n\lambda_U \\
&F_k = \begin{cases} F_{k1} & \text{for } 1 \le k \le n-1 \\ \begin{bmatrix} F_{k1} & 0 \\ 0 & F_{k2} \end{bmatrix} & \text{for } k = n \end{cases} \\
&F_{k1} = \begin{bmatrix} -\theta_k I_{p_k} & 0 \\ 0 & -\lambda_k I_{q_k} - \lambda_k^2 F_{k3} \end{bmatrix}, 2 \le k \le n, \quad F_{n2} = -\lambda_U, \\
&F_{k3} = \begin{bmatrix} \star_{k,k-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T M_{[k-1]}^{-1} \begin{bmatrix} \star_{k,k-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad 2 \le k \le n \\
&\begin{matrix} J_1(x_1) = 2P_k(a_{11} + a_{1,2}s_1) \\ E_{11} = \begin{bmatrix} \gamma^{-1}P_1 B_{11} & C_{11}^T\lambda_1 \end{bmatrix} \end{matrix}, \; F_{11} = \begin{bmatrix} -\theta_1 I_{p_1} & 0 \\ 0 & -\lambda_1 I_{q_1} \end{bmatrix},
\end{aligned}$$

where $\diamondsuit_{i,j}$ denotes any function depending only on $(x_{[i]}, P_{[i]})$, $\Lambda_{[j]}$ and $(s_1, \cdots, s_j)$ and their partial derivatives. Using the Schur complement of (56) and (57), we have the following.

Corollary 1 *Let $1 \le k \le n$. Assume that $M_{[k-1]}(x_{[k-1]}) < 0$ is satisfied for all $x_{[k-1]} \in \mathcal{R}^{k-1}$ if $k \neq 1$. Then, $M_{[k]}(x_{[k]}) < 0$ holds for all $x_{[k]} \in \mathcal{R}^k$ if and only if*
(i) $1 \le k \le n-1$ *Case:*

$$\begin{aligned} J_k < 0 \quad &\text{for } p_k + q_k = 0 \\ F_{k1} < 0, \quad J_k - E_{k1} F_{k1}^{-1} E_{1k}^T < 0 \quad &\text{for } p_k + q_k \neq 0 \end{aligned}$$

(ii) $k = n$ *Case:*

$$\begin{aligned} F_{n2} < 0, \quad J_n - E_{n2} F_{n2}^{-1} E_{n2}^T < 0 \quad &\text{for } p_n + q_n = 0 \\ F_{n1} < 0, \quad F_{n2} < 0, \quad J_n - E_{n2} F_{n2}^{-1} E_{n2}^T - E_{n1} F_{n1}^{-1} E_{1n}^T < 0 \quad &\text{for } p_n + q_n \neq 0 \end{aligned}$$

are satisfied for all $x_{[k]} \in \mathcal{R}^k$.

Theorem 2 and Corollary 1 tell us that we can decide (λ_k, s_k) and Λ_U recursively from $k = 1$ through $k = n$ in order to try to make $M(x)$ negative.

Theorem 3 *Let $1 \le k \le n$ and $p_k + q_k \neq 0$. Assume that $M_{[k-1]}(x_{[k-1]}) < 0$ holds for all $x_{[k-1]} \in \mathcal{R}^{k-1}$ if $k \neq 1$. The function $\theta_k(x_{[k]})$ is assumed to be an arbitrary function satisfying*

$$\theta_k(x_{[k]}) > 0, \quad \forall x_{[k]} \in \mathcal{R}^k \ . \tag{58}$$

Then, there always exist a scalar-valued smooth function $\lambda_k(x_{[k]})$ such that for any such $\theta_k(x_{[k]})$,

$$\lambda_k(x_{[k]}) > 0, \quad F_{k1}(x_{[k]}) < 0 \tag{59}$$

are satisfied for all $x_{[k]} \in \mathcal{R}^k$.
Proof : If $k = 1$, (59) is satisfied for any functions $\lambda_1(x_1) > 0$ and $\theta_1(x_1) > 0$. We prove the claim for $k \ge 2$. Suppose that $x_{[k]}$ is a vector belonging to $\mathcal{R}^k$. The inequalities (59) at the point $x_{[k]}$ are written as

$$\lambda_k > 0, \quad \lambda_k I_{q_k} + \lambda_k^2 F_{k3} > 0. \tag{60}$$

These two inequalities are independent of θ_k. The second inequality in (60) is implied by

$$-\lambda_{min}(F_{k3})\lambda_k < 1 \ . \tag{61}$$

Here, $\lambda_{min}(\cdot)$ denotes the minimum eigenvalue of a matrix. Since $M_{[k-1]} < 0$ implies $F_{k3} \le 0$, there exists $\lambda_k > 0$ satisfying (60). All entries of F_{k3} are smooth functions defined on $\mathcal{R}^k$ so that $\lambda_{min}(F_{k3})$ is a continuous function of $x_{[k]}$. Hence, there exists a smooth function $\lambda_k(x_{[k]})$ such that inequalities in (60) are satisfied for all $x_{[k]} \in \mathcal{R}^k$ and for any $\theta_k(x_{[k]}) > 0$. ∎

Theorem 4 *Let $1 \le k \le n$. Assume that $M_{[k-1]}(x_{[k-1]}) < 0$ holds for all $x_{[k-1]} \in \mathcal{R}^{k-1}$ if $k \neq 1$.*
(i) $k < n$ *& $p_k + q_k \neq 0$ Case: There always exists a scalar-valued smooth function $s_k(x_{[k]})$ such that*

$$J_k(x_{[k]}) - E_{k1}(x_{[k]}) F_{k1}^{-1}(x_{[k]}) E_{1k}^T(x_{[k]}) < 0 \tag{62}$$

is satisfied for all $x_{[k]} \in \mathcal{R}^k$*.*

(ii) $k < n$ *&* $p_k + q_k = 0$ *Case: There always exists a scalar-valued smooth function* $s_k(x_{[k]})$ *such that*

$$J_k(x_{[k]}) < 0 \tag{63}$$

is satisfied for all $x_{[k]} \in \mathcal{R}^k$*.*

(iii) $k = n$ *&* $p_n + q_n \neq 0$ *Case: There always exist scalar-valued smooth functions* $\lambda_U(x_{[n]})$ *and* $s_n(x_{[n]})$ *such that*

$$\lambda_U(x_{[n]}) > 0, \quad F_{n2}(x_{[n]}) < 0 \tag{64}$$

$$J_n(x_{[n]}) - E_{n2}(x_{[n]})F_{n2}^{-1}(x_{[n]})E_{n2}^T(x_{[n]}) - E_{n1}(x_{[n]})F_{n1}^{-1}(x_{[n]})E_{1n}^T(x_{[n]}) < 0 \tag{65}$$

are satisfied for all $x_{[n]} \in \mathcal{R}^n$*. Furthermore, a solution* s_n *is given by*

$$s_n(x_{[n]}) = \frac{-P_n a_{n,n+1}}{\lambda_U} \tag{66}$$

(iv) $k = n$ *&* $p_n + q_n = 0$ *Case: There always exist scalar-valued smooth functions* $\lambda_U(x_{[n]})$ *and* $s_n(x_{[n]})$ *such that*

$$\lambda_U(x_{[n]}) > 0, \quad F_{n2}(x_{[n]}) < 0 \tag{67}$$

$$J_n(x_{[n]}) - E_{n2}(x_{[n]})F_{n2}^{-1}(x_{[n]})E_{n2}^T(x_{[n]}) < 0 \tag{68}$$

are satisfied for all $x_{[n]} \in \mathcal{R}^n$*. Furthermore, a solution* s_n *is given by (66).*

Proof : **(iii)** The inequality $F_{n2} < 0$ is obvious if $\lambda_U > 0$. The inequality (65) at a point $x_{[n]} \in \mathcal{R}^n$ is

$$2P_n a_{n,n+1} s_n + \diamondsuit_{n,n-1} + \lambda_U s_n^2 - E_{n1}F_{n1}^{-1}E_{1n}^T < 0 \ .$$

We rewrite the above as

$$\lambda_U s_n^2 + 2P_n a_{n,n+1} s_n + P_n^2 \alpha < 0 \ , \tag{69}$$

where a scalar-valued function $\alpha(x_{[n]})$ defined on $\mathcal{R}^n$ is independent of s_n and λ_U. This inequality (69) has a solution $s_n \in \mathcal{R}$ at $x_{[n]}$ if and only if

$$\alpha \lambda_U < a_{n,n+1}^2 \ . \tag{70}$$

Note that $a_{n,n+1}^2(x_{[n]}) \neq 0$ for all $x_{[n]} \in \mathcal{R}^n$ by assumption. Hence, there exists $\lambda_U > 0$ satisfying (70) at $x_{[n]}$. Since $a_{n,n+1}$ and α are smooth functions defined on $\mathcal{R}^n$, there exist smooth functions $\lambda_U(x_{[n]})$ and $s_n(x_{[n]})$ such that (64) and (65) are satisfied for all $x_{[n]} \in \mathcal{R}^n$. Finally, a solution s_n to (69) is (66).

(i) The inequality (62) becomes

$$2P_k a_{k,k+1} s_k + P_k^2 \alpha < 0 \ . \tag{71}$$

This linear inequality always has a solution $s_k \in \mathcal{R}$ for each $x_{[k]}$ in $\mathcal{R}^k$. Smoothness of functions appearing in the inequality guarantees the existence of a smooth function $s_k(x_{[k]})$.

(iv) The inequality (68) is written as

$$2P_n a_{n,n+1} s_n + \diamondsuit_{n,n-1} + \lambda_U s_n^2 < 0 \ .$$

This equality is in the form of (69). Hence, the rest of the proof is the same as the (iii) case.

(ii) The inequality (63) becomes (71) with an appropriate function α. The existence of $s_k(x_{[k]})$ is proved as in the (i) case. ∎

Theorems 3 and 4 have shown that we can always choose appropriate parameters $(\lambda_k, \theta_k, s_k)$ and λ_U recursively from $k = 1$ through $k = n$ such that $M(x)$ is negative. Moreover, the control gain can be selected as $s_n = -P_n a_{n,n+1}/\lambda_U$. By combining these results with Theorem 1, we obtain the following theorem.

Theorem 5 *Consider the nonlinear system Σ. For any $\gamma > 0$, there always exists a control law in the form of (14) with a positive definite solution $V(x)$ to (15) such that the control law is globally inverse optimal with disturbance level γ. Furthermore, in the inverse optimal design, the function $\Theta(x)$ can be any function satisfying (58).*

Although global inverse optimality can be achieved for any $\gamma > 0$, this does not imply that we can make the effect of disturbance w arbitrarily small. Recall the cost function (16) and the concept of inverse optimal control. Since $q(x)$ and $r(x)$ are free, we can reduce the cost $J(u, w)$ by choosing $q(x)$ and $r(x)$ small; moreover, the function $\Theta(x)$ can remove the effect of γ. It is, however, interesting that we can achieve global inverse optimality even if we fix $\Theta(x)$. The freedom of the parameter $\Theta(x)$ in the inverse optimal design will be exploited to achieve robustness of the resulting state-feedback system in Section 7. By using Theorems 3 and 4, we can construct λ_k, θ_k, s_k and λ_U recursively from $k = 1$ through $k = n$. The feedback control law is obtained by (66) and global inverse optimality with disturbance level γ is achieved.

6 Local optimality

Although Theorem 5 guarantees that the backstepping procedure proposed in Section 5 achieves global inverse optimality, it does not assure local optimality. In this section, we show that we can use backstepping to achieve local optimality as well as global inverse optimality. Such a result was first obtained in [8], but here we illustrate how to incorporate the free parameter $\Theta(x)$.

Consider the function

$$x^T \check{C}^T(x) \bar{\Lambda}(x) \check{C}(x) x \tag{72}$$

as a candidate for $q(x)$ in the cost functional (16) for the nonlinear system Σ. Here, we suppose that $\check{C}(x)$ has the triangular structure (27) as $\bar{C}(x)$. The function $\bar{\Lambda}(x)$ is a diagonal matrix as (28) and $\bar{\Lambda}(x) > 0$ for all $x \in \mathcal{R}^n$. Note that any function $x^T Q(x) x$ with $Q(x) = Q^T(x) > 0$ always has the representation (72). In fact, to obtain (72) we can apply the Cholesky factorization to $Q = U^T U$ with a lower triangular matrix U. We now define $Q_l = Q_l^T > 0$ by

$$Q_l = \check{C}^T(0) \bar{\Lambda}(0) \check{C}(0), \quad \bar{\Lambda}_l = \bar{\Lambda}(0) \ .$$

For $r(x)$ and $\Theta(x)$ in (16), we define

$$r_l = r(0), \quad \Theta_l = \Theta(0) = \mathrm{diag}[\theta_{l1}, \theta_{l2}, \cdots, \theta_{ln}] \ .$$

We use the following matrices:

$$\bar{\Lambda}_l = \mathrm{diag}[\lambda_{l1}, \lambda_{l2}, \cdots, \lambda_{ln}], \quad \Lambda_l = \text{block-diag}[\bar{\Lambda}_l, r_l] \ .$$

Let ϵ be any small number satisfying $1 \gg \epsilon > 0$. We choose $\bar{C}(x)$ and $\bar{C}_l$ as

$$\bar{C}(x) = (1-\epsilon)^{1/2} \check{C}(x), \quad \bar{C}_l = \bar{C}(0) \ .$$

The matrix $\hat{C}(x)$ defined in (38) becomes

$$\hat{C}(x) = \begin{bmatrix} \bar{C}(x) & H \\ 0 & \end{bmatrix} .$$

Let $P_l = P_l^T > 0$ be the unique solution to (19). Using the Cholesky factorization, we can decompose P_l as $P_l = S_l^T P S_l$ with

$$S_l = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \odot & 1 & 0 & \ddots & \vdots \\ \odot & \odot & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & 0 \\ \odot & \cdots & \cdots & \odot & 1 \end{bmatrix}, \quad P = \mathrm{diag}[P_1, P_2, \cdots, P_n] > 0 ,$$

where $\odot$ denotes any real constant and P_i, $i = 1, 2, \cdots, n$ are scalar positive numbers. This factorization of P_l is the same as the one used in [8]. Note that the inverse of S_l has the same structure as S_l. We write it as

$$S_l^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ c_1 & 1 & 0 & \ddots & \vdots \\ e_{21} & c_2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & 0 \\ e_{n-1,1} & \cdots & e_{n-1,n-2} & c_{n-1} & 1 \end{bmatrix}, \tag{73}$$

where c_i and e_{ij}, $2 \le i \le n-1$, $1 \le j \le i-1$, are appropriate real numbers. Then, we can prove the following.

Theorem 6 *Suppose that $P_l = P_l^T > 0$ is the solution to (19). Then, the set of parameters*

$$s_i(0) = c_i, \quad 1 \le i \le n-1 \tag{74}$$

$$s_n(0) = -r_l^{-1} P_n a_{n,n+1}(0) \tag{75}$$

$$\lambda_i(0) = \lambda_{li}, \quad 1 \le i \le n \tag{76}$$

$$\lambda_U(0) = r_l \tag{77}$$

$$\theta_i(0) = \theta_{li}, \quad 1 \le i \le n \tag{78}$$

$$d_{ij} = e_{ij}, \quad 2 \le i \le n-1, \quad 1 \le j \le i-1 \tag{79}$$

solves (39) at $x = 0$ and the linear optimal law is represented by $u_l = -r_l^{-1} G^T P_l x = s_n(0) S(0) x$.
Proof : Rearrange (19) as

$$\bar{A}_l^T P_l + P_l \bar{A}_l + \frac{1}{\gamma^2} P_l B_l \Theta_l^{-1} B_l^T P_l + Q_l = 0 \tag{80}$$

$$\bar{A}_l = A_l - \frac{1}{2} G r_l^{-1} G^T P_l .$$

Note here that

$$r_l^{-1} G^T P_l = r_l^{-1} G_l^T S_l^T P S_l = \begin{bmatrix} 0 & \cdots & 0 & r_l^{-1} P_n a_{n,n+1}(0) \end{bmatrix} S_l . \tag{81}$$

Define the following matrices:

$$\hat{S}_l = \left[\begin{array}{c|c} \multicolumn{2}{c}{S_l^{-1}} \\ \hline 0 \cdots 0 & -r_l^{-1} P_n a_{n,n+1}(0) \end{array} \right], \quad \hat{A}_l = [A_l \ G_l] ,$$

Then, (80) is identical with

$$\hat{S}_l^T \hat{A}_l^T S_l^T P + P S_l \hat{A}_l \hat{S}_l + \frac{1}{\gamma^2} P S_l B_l \Theta_l^{-1} B_l^T S_l^T P + \hat{S}_l^T \hat{C}_l^T \Lambda_l \hat{C}_l \hat{S}_l + \epsilon S_l^{-T} Q_l S_l^{-1} = 0 \ .$$

Using the Schur complement formula for the above equation, we arrive at

$$M_l := \begin{bmatrix} \hat{S}_l^T \hat{A}_l^T S_l^T P + P S_l \hat{A}_l \hat{S}_l & \gamma^{-1} P S_l B_l & \hat{S}_l^T \hat{C}_l^T \Lambda_l \\ \gamma^{-1} B_l^T S_l^T P & -\Theta_l & 0 \\ \Lambda_l \hat{C}_l \hat{S}_l & 0 & -\Lambda_l \end{bmatrix} < 0 \ .$$

The choice of parameters (74)-(79) implies

$$S^{-1}(0) = S_l^{-1}, \quad S(0) = S_l, \quad T^{-1}(0) = S_l, \quad \hat{S}(0) = \hat{S}_l \ .$$

Since we also have

$$\hat{A}_l = \hat{A}(0), \quad B_l = B(0), \quad \hat{C}_l = \hat{C}(0) \ ,$$

$M(0) = M_l$ holds. Finally, using (81) the linear optimal control law (20) is obtained as

$$u_l = -r_l^{-1} G_l^T P_l x = \begin{bmatrix} 0 & \cdots & 0 & -r_l^{-1} P_n a_{n,n+1}(0) \end{bmatrix} S_l x \ .$$

∎

This theorem implies that $s_k(0)$, $\lambda_k(0)$, $\theta_k(0)$ and $\lambda_U(0)$ given in (74)-(78) solve the inequalities in Corollary 1 at $x_{[k]} = 0$. Each parameter in (74)-(78) is one of the solutions in Theorems 3 and 4. Note that

$$s_n(0) = -r_l^{-1} P_n a_{n,n+1}(0) = -P_n a_{n,n+1}(0)/\lambda_U \tag{82}$$

does not contradict the solution $s_n(x_{[n]})$ given in Theorem 4. In other words, the requirements (74)-(78) at $x = 0$ can be always met in each step of backstepping in Section 5. Therefore, we arrive at the following statement.

Theorem 7 *Consider the nonlinear system Σ. Suppose that $Q_l = Q_l^T > 0$, $r_l > 0$ and a diagonal matrix $\Theta_l > 0$ are given and that $P_l = P_l^T > 0$ is the solution to (19). Then there always exists a control law in the form of (14) with a positive definite solution $V(x)$ to (15) such that the control law is locally optimal and globally inverse optimal with disturbance level γ. The functions $q(x)$ and $r(x)$ in (15) satisfy*

$$q(x) \ > \ (1-\epsilon) x^T \check{C}^T(x) \bar{\Lambda}(x) \check{C}(x) x, \ \forall x \in \mathcal{R}^n \setminus \{0\} \tag{83}$$

$$r(x) \ = \ \lambda_U(x) \ . \tag{84}$$

Furthermore, in the inverse optimal design, the function $\Theta(x)$ can be any function satisfying (58) and $\Theta(0) = \Theta_l$.

Proof : Since the choice of parameters (74)-(79) is one solution of Theorem 5, the control law (14) with the parameters satisfying (74)-(79) is globally inverse optimal with disturbance attenuation level γ. Since we have

$$-\frac{1}{2} r^{-1}(0) G^T(0) \left. \frac{\partial^2 V}{\partial x^2} \right|_{x=0} x = -r_l^{-1} G_l^T S_l^T P S_l x = -r_l^{-1} G_l^T P x = u_l \ ,$$

the Jacobian linearized control (22) stabilizes the linear system (17) and achieves $\min_{u_l} \max_{w_l \in \mathcal{L}_2} J_l(u_l, w_l)$. Furthermore, according to (54) and (46), the weighting functions of (15) satisfy

$$\begin{aligned}
\left.\frac{\partial^2 q}{\partial x^2}\right|_{x=0} &= -2\left[A_l^T S_l^T P S_l + S_l^T P S_l A_l + S_l^T P S_l \left(\frac{1}{\gamma^2} B_l \Theta_l^{-1} B_l^T - \frac{1}{\lambda_U} G_l G_l^T\right) S_l^T P S_l\right] \\
&= -2\left[A_l^T P_l + P_l A_l + P_l \left(\frac{1}{\gamma^2} B_l \Theta_l^{-1} B_l^T - \frac{1}{\lambda_U} G_l G_l^T\right) P_l\right] \\
r(0) &= \lambda_U(0)
\end{aligned}$$

at $x = 0$. Then, Equation (19) and (77) imply

$$\left.\frac{\partial^2 q}{\partial x^2}\right|_{x=0} = 2Q_l, \quad r(0) = r_l.$$

Finally, from (42) we obtain

$$\begin{aligned}
-\bar{Q} &= A^T T^T P S + S^T P T A + S^T P T \left(\frac{1}{\gamma^2} B \Theta^{-1} B^T - \frac{1}{\lambda_U} G G^T\right) T^T P S + \bar{C}^T \bar{\Lambda} \bar{C} \\
&= A^T T^T P S + S^T P T A + S^T P T \left(\frac{1}{\gamma^2} B \Theta^{-1} B^T - \frac{1}{\lambda_U} G G^T\right) T^T P S + \check{C}^T \bar{\Lambda} \check{C} - \epsilon \check{C}^T \bar{\Lambda} \check{C} \\
&= \hat{Q} - \epsilon \check{C}^T \bar{\Lambda} \check{C} < 0, \ \forall x \in \mathcal{R}^n
\end{aligned}$$

with an appropriate function $\hat{Q}(x) = \hat{Q}^T(x)$. Due to (45), the function $q(x)$ of (15) is given by

$$\begin{aligned}
q(x) &= x^T \bar{C}^T(x) \bar{\Lambda}(x) \bar{C}(x) x + x^T \bar{Q} x \\
&= x^T \check{C}^T(x) \bar{\Lambda}(x) \check{C}(x) x - x^T \hat{Q} x \\
&> (1-\epsilon) x^T \check{C}^T(x) \bar{\Lambda}(x) \check{C}(x) x, \ \forall x \in \mathcal{R}^n \setminus \{0\}
\end{aligned}$$

∎

This theorem includes the result in [8] as a special case with $\Theta(x) = I_p$. Equations (83) and (84) give useful information about the cost functional which is actually minimized by the inverse optimal design.

The backstepping procedure for local optimal and global inverse optimal control is summarized as follows:

Step 1 Let P_1 be any positive number. If $p_1 + q_1 \neq 0$, choose a function $\lambda_1(x_1) > 0$ satisfying (76). Let $\theta_1 > 0$ be any function satisfying (78). Calculate a solution $s_1(x_1)$ to (71) under the constraint (74).

Step k Let P_k be any positive number. If $p_k + q_k \neq 0$, choose a function $\lambda_k(x_{[k]}) > 0$ such that (61) and (76) hold. Let $\theta_k > 0$ be any function satisfying (78). Calculate a solution $s_k(x_{[k]})$ to (71) under the constraint (74).

Step n Let P_n be any positive number. If $p_n + q_n \neq 0$, choose a function $\lambda_n(x_{[n]}) > 0$ such that (61) and (76) hold. Let $\theta_n > 0$ be any function satisfying (78). Find a function $\lambda_U > 0$ satisfying (70) and (77). Choose $s_n(x_{[n]})$ as (66).

By performing the above procedure recursively from $k=1$ through $k=n$, the state-feedback controller $u=s_n(x_{[n]})\chi_n$ is locally optimal and globally inverse optimal with disturbance level γ. If another solution s_n to (65) is chosen instead of the particular solution (66), the control law $u=s_n(x_{[n]})\chi_n$ is neither locally optimal nor globally inverse optimal. However, it still globally asymptotically stabilizes Σ and the nonlinear control law u and the linearized control law u_l achieve

$$\begin{aligned} J(u,w,T) &\leq V(x(0))-V(x(T)), \quad \forall w\in\mathcal{L}_2[0,T] \\ J_l(u_l,w_l,T) &\leq V_l(x(0))-V_l(x(T)), \quad \forall w_l\in\mathcal{L}_2[0,T] \end{aligned}$$

for all $T\geq 0$, respectively.

7 Global robust stability via inverse optimal control

In the previous sections, we considered the cost function $J(u,w)$ in (16) for global inverse optimal control. The motivation of introducing the w-term into the cost functional was to achieve some robustness against disturbances and uncertainty. Indeed, both our design and the one in [8] guarantee local robustness of level γ because the controllers locally solve the disturbance attenuation problem. However, the role γ plays in describing global robust stability has not yet been characterized. Recall that $q(x)$, $r(x)$ and $\Theta(x)$ are free parameters in the inverse optimal design. Since $\Theta(x)$ can absorb γ, it seems that γ does not have any meaning for global robustness. In the class of inverse optimal designs where the function $\Theta(x)$ is restricted to $\Theta(x)=I_p$ as in [8], the level γ might not be meaningful either since small $r(x)$ and $q(x)$ can make $J(u,w)$ small. In this section, we reexamine the particular cost function

$$J(u,w)=\int_0^\infty\left[\sum_{i=1}^{n}(\lambda_i(x)z_i^Tz_i-\gamma^2\theta_i(x)w_i^Tw_i)+x^T\bar{Q}(x)x+\lambda_U(x)u^2\right]dt,\ \bar{Q}(x)>0,\ \forall x\in\mathcal{R}^n \quad (85)$$

with respect to which global inverse optimality is achieved in previous sections. Here, (λ_i,θ_i), $i=1,2,\dots,n$ are called the generalized SD scaling. This section shows that the global inverse optimal control with such scaling indeed achieves global robustness against a class of disturbances w whose sizes are prescribed by γ.

Consider the static uncertain system $\Sigma_\Delta: z\mapsto w$ given by

$$\begin{aligned} w_i &= h_{\Delta_i}(t,z_i), \quad i\in L\subset\{1,2,\dots,n\} \\ 0 &= h_{\Delta_i}(t,0), \quad \forall t\geq 0\ , \end{aligned}$$

where h_{Δ_i} is defined on $\mathcal{R}_+\times\mathcal{R}^{q_i}$. The uncertain system consisting of Σ_a in (31) and the uncertainty Σ_Δ is depicted by Fig.2. Here, z_L and w_L are signals made of all entries z_i and w_i whose indexes i belong to the set L. Signals $z_{\bar{L}}$ and $w_{\bar{L}}$ consist of remaining entries of z and w.

$$\begin{gathered} w_L=\begin{bmatrix} w_i \\ \vdots \end{bmatrix},\ z_L=\begin{bmatrix} z_i \\ \vdots \end{bmatrix},\ w_{\bar{L}}=\begin{bmatrix} w_k \\ \vdots \end{bmatrix},\ z_{\bar{L}}=\begin{bmatrix} z_k \\ \vdots \end{bmatrix} \\ i\in L, \qquad k\in\{1,2,\cdots,n\}\setminus L\ . \end{gathered}$$

If we look at the cost (85) with $\theta_i\leq\lambda_i$, $i=1,2,\dots,n$, then light weighting of z means light weighting of w. The cost J cannot be made small any more by simply letting $q(x)$ be small or letting $\Theta(x)$ be large. Furthermore, the scaling Θ can no longer absorb γ, so γ may retain some meaning in describing the allowable size of the uncertainty. Indeed, the generalized scaling (λ_i,θ_i) in (85) reduces to the SD scaling[12] in the case of $\theta_i=\lambda_i$, and the following can be proved[11, 12]:

Theorem 8 *Suppose that the control law (14) is globally inverse optimal with disturbance level* γ *with respect to the cost (85) with* $\theta_i \leq \lambda_i$ *for all* $i \in L \subset \{1, 2, \ldots, n\}$*. Then, the system* $(\Sigma_a, \Sigma_\Delta)$ *in Fig.2 controlled by the state-feedback law (14) is globally uniformly asymptotically stable for any uncertainty* Σ_Δ *satisfying*

$$\gamma \|w_i\| \leq \|z_i\|, \quad \forall t \in \mathcal{R} \ , \forall i \in L \ . \tag{86}$$

Furthermore, the control law (14) achieves

$$\int_0^T \left[\sum_{i \in \{1,2,\cdots,n\} \setminus L} (\lambda_i(x) z_i^T z_i - \gamma^2 \theta_i(x) w_i^T w_i) + x^T \bar{Q}(x) x + \lambda_U(x) u^2 \right] dt \leq V(x(0)) - V(x(T)), \ \forall w \in \mathcal{L}_2[0, T]$$

for all $T \geq 0$ *and for any uncertainty* Σ_Δ *satisfying (86).*

As shown in Theorems 3 and 4, we can always choose $\theta_i(x) \leq \lambda_i(x)$ for $i = 1, 2, \ldots, n$ in the backstepping design of global inverse optimal controllers. It is important to note that, unlike our notion of global optimality, global robustness is *not* defined here in the inverse sense—the uncertainty structure is prescribed as part of the system through $B(x)$ and $C(x)$ in (31) and is not free to be chosen during the control design. Thus to guarantee such prescribed global robust stability, it is crucial to allow the parameter $\Theta(x)$ in the cost functional (16) to vary with the state x so that we may achieve $\theta_i(x) \leq \lambda_i(x)$.

The type of global robust stability achieved in [8] is input-to-state stability (ISS) with respect to the disturbance w. One can easily verify that if the mapping

$$x \quad \mapsto \quad x^T \, \frac{\bar{C}^T(x) \bar{\Lambda}(x) \bar{C}(x) + \bar{Q}(x)}{\|\Theta(x)\|} \, x \tag{87}$$

is radially unbounded, then the generalized SD scaling design also yields a closed-loop system which is ISS with respect to w. Thus one can view ISS in this context as a consequence of robust stability for a special class of output matrices $\bar{C}$ and scaling matrices $\bar{\Lambda}$ and Θ. Moreover, if we allow the output matrix $\bar{C}$ to be freely chosen as part of the scaling design, then it is always possible to render the mapping (87) radially unbounded.

Theorem 9 *Consider the nonlinear system* Σ *in (8). Suppose that* $Q_l = Q_l^T > 0$*,* $r_l > 0$ *and a diagonal matrix* $\Theta_l > 0$ *are given and that* $P_l = P_l^T > 0$ *is the solution to (19). Then there always exist a control law in the form of (14) with a positive definite solution* $V(x)$ *to (15) and a matrix* $\bar{C}(x) \in \mathcal{R}^{n \times n}$

$$\bar{C}(x) = \begin{bmatrix} C_{11}(x_{[1]}) & 0 & \cdots & 0 \\ C_{21}(x_{[2]}) & C_{22}(x_{[2]}) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ C_{n1}(x_{[n]}) & C_{n2}(x)_{[n]} & \cdots & C_{nn}(x_{[n]}) \end{bmatrix} \tag{88}$$

with smooth functions $C_{ij}(x) \in \mathcal{R}$ *such that the control law is locally optimal and globally inverse optimal with disturbance level* γ *and the resulting closed-loop system is ISS.*

Proof : Let a row vector $\bar{C}_k(x_{[k]}) \in \mathcal{R}^{1 \times k}$ be

$$\bar{C}_k(x_{[k]}) = [C_{k1}(x_{[k]}) \ C_{k2}(x_{[k]}) \ \cdots \ C_{kk}(x_{[k]})]$$

for $1 \leq k \leq n$. By assumption, there exist a positive number ν and a constant matrix $\bar{C}(0)$ in the form of (88) such that

$$M_l := M(0) < -\nu I \tag{89}$$

holds. We first show that $M(x) < -\nu I$ can be achieved for all $x \in \mathcal{R}^n$ with a constant $\nu > 0$ by choosing $\theta_k > 0$, $\lambda_k > 0$ and $\lambda_U > 0$ appropriately under the conditions (74)-(78). Assume that for $k \neq 1$, $M_{[k-1]}(x_{[k-1]}) < -\nu I$ is satisfied for all $x_{[k-1]} \in \mathcal{R}^{k-1}$. According to proofs of Theorem 3 and 4,

$$M_{[k]}(x_{[k]}) < -\nu I, \qquad \forall x_{[k]} \in \mathcal{R}^k$$

is achieved if and only if

$$-\theta_k + \nu < 0 \tag{90}$$

$$\left\{ \begin{array}{cl} -\lambda_k + \nu < 0 & : k = 1 \\ -\lambda_k - \lambda_k^2 F_{k3} + \nu < 0, \quad \lambda_k > 0 & : 2 \leq k \leq n \end{array} \right. \tag{91}$$

$$\left\{ \begin{array}{cl} 2P_k a_{k,k+1} s_k + P_k^2 \alpha < 0 & : 1 \leq k \leq n-1 \\ \lambda_U s_n^2 + 2P_n a_{n,n+1} s_n + P_n^2 \alpha < 0, \quad \lambda_U > 0 & : k = n \end{array} \right. \tag{92}$$

are satisfied for all $x_{[k]} \in \mathcal{R}^k$. In the $k \leq n-1$ case, there always exists a smooth function s_k satisfying (92). As for the $k = n$ case, since a smooth function $\lambda_U > 0$ can be always chosen such that $(\alpha + \nu P_n^{-2})\lambda_U < a_{n,n+1}^2$ is satisfied, (66) is a solution to (92) for such λ_U. Due to (89), these functions can be chosen such that (74)-(78) are satisfied. Obviously, a positive constant $\theta_k(x_{[k]}) = \theta_k(0)$ is a solution to (90). For (91), the existence of a smooth function λ_1 is straightforward. Since (89) is assumed, a solution $\lambda_k(0) > 0$ to (91) at $x_{[k]} = 0$ exists. This implies that $1 + 4F_{k3}(0)^2\nu > 0$ holds. Recall that F_{k3} is represented by

$$F_{k3}(x_{[k]}) = \begin{bmatrix} \hat{S}_{[k-1]}^T \bar{C}_k^T \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T M_{[k-1]}^{-1} \begin{bmatrix} \hat{S}_{[k-1]}^T \bar{C}_k^T \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Thus, a smooth function $\bar{C}(x_{[k]})$ can be chosen such that $F_{k3}(0) = F_{k3}(x_{[k]})$ for all $x_{[k]} \in \mathcal{R}^k$. This implies that the constant $\lambda_k(x_{[k]}) = \lambda_k(0) > 0$ solves (91) for all $x_{[k]} \in \mathcal{R}^k$. Hence, $M(x) < -\nu I$ is proved to be achieved for all $x \in \mathcal{R}^n$ with a constant $\nu > 0$. Now, $M(x) < -\nu I$ is equivalent to

$$\hat{S}^T \hat{A}^T T^T P + PT\hat{A}\hat{S} + \frac{1}{\gamma^2} PTB(\Theta - \nu I)^{-1} B^T T^T P + \hat{S}^T \hat{C}^T \Lambda(\Lambda - \nu I)^{-1} \hat{C}\Lambda\hat{S} + \nu I < 0 \ .$$

Let $\tilde{Q}(x) > 0$ be defined with

$$-S^{-T}\tilde{Q}S^{-1} = \hat{S}^T \hat{A}^T T^T P + PT\hat{A}\hat{S} + \frac{1}{\gamma^2} PTB(\Theta - \nu I)^{-1} B^T T^T P + \hat{S}^T \hat{C}^T \Lambda(\Lambda - \nu I)^{-1} \hat{C}\Lambda\hat{S} + \nu I \ .$$

Using the argument similar to (50)-(52), we obtain

$$\frac{d}{dt}V(x) + z^T \Lambda(\Lambda - \nu I)^{-1}\Lambda z + x^T \tilde{Q} x + \nu \chi^T \chi - \gamma^2 w^T (\Theta - \nu I) w \leq 0, \quad \forall x \in \mathcal{R}^n \ .$$

Since $M(x) < -\nu I$ implies $-\Lambda + \nu I < 0$, we arrive at

$$\frac{d}{dt}V(x) \leq -\nu \chi^T \chi + \gamma^2 w^T (\Theta - \nu I) w, \quad \forall x \in \mathcal{R}^n \ .$$

Since Θ is a constant matrix satisfying $-\Theta + \nu I < 0$, from the global diffeomorphism $S(x)$ it follows that the closed-loop system is ISS. ∎

We therefore recover the ISS robustness result of [8] as a special case of our global robustness result. In fact, the scaling factors $\bar{\Lambda}$ and Θ can be chosen as constants in this case and we achieve global robustness with respect to dynamic as well as memoryless uncertainty[11, 12]. For the more general case in which the output matrix $\bar{C}$ is fixed as part of the uncertainty structure (a case not considered in [8]), the ISS property can be checked by simply examining the mapping (87).

8 Example

We first consider the following nonlinear system from [8]:

$$\Sigma : \begin{cases} \dot{x}_1 = x_1^2 + x_2 + w_p \\ \dot{x}_2 = u \end{cases} \tag{93}$$

with the cost functional

$$J(u, w_p) = \int_0^\infty (q(x) + r(x)u^2 - \gamma^2 w_p^2)dt, \quad \gamma = 5 \ . \tag{94}$$

Let Σ_l denote the Jacobian linearization of Σ. The cost functional for this linear system is chosen as

$$J_l(u_l, w_{pl}) = \int_0^\infty (x_1^2 + x_2^2 + u_l^2 - \gamma^2 w_{pl}^2)dt, \quad \gamma = 5 \ . \tag{95}$$

For the local optimality of the global inverse optimal control, we assume that

$$\left. \frac{1}{2} \frac{\partial^2 q(x)}{\partial x^2} \right|_{x=0} = Q_l = I_2, \quad r(0) = r_l = 1 \ . \tag{96}$$

In this setting, Ezal *et al.* [8] have obtained a control law

$$\begin{aligned} u &= \begin{cases} -1.78(1+\sigma(x))(0.6x_1 + x_2 + x_1^2) & \text{if } \sigma(x) \geq 0 \\ -1.78(0.6x_1 + x_2 + x_1^2) & \text{if } \sigma(x) < 0 \end{cases} \\ \sigma(x) &= 1.8x_1 + 1.05x_1^2 \end{aligned} \tag{97}$$

which is locally optimal and globally inverse optimal. For convenience we will refer to (97) as Ezal's control law.

The generalized SD scaling design developed in this paper allows us to restrict the cost functional (94) for Σ to a more specific one:

$$\begin{aligned} & J(u, w) = \int_0^\infty (\lambda_1(x_1)x_1^2 + \lambda_2(x)x_2^2 + \lambda_U(x)u^2 - \gamma^2 w_p^2 + \hat{q}(x))dt, \quad \gamma = 5 \\ & \lambda_1(x_1) > 0, \quad \lambda_2(x) > 0, \quad \lambda_U(x) > 0, \quad \forall x \in \mathcal{R}^2 \\ & \hat{q}(x) > -\epsilon(\lambda_1(x_1)x_1^2 + \lambda_2(x)x_2^2), \ \forall x \in \mathcal{R}^2 \setminus \{0\} \\ & \hat{q}(0) = 0, \quad \left. \frac{\partial \hat{q}(x)}{\partial x} \right|_{x=0} = 0, \quad \left. \frac{\partial^2 \hat{q}(x)}{\partial x^2} \right|_{x=0} = 0 \end{aligned} \tag{98}$$

for a sufficiently small $1 \gg \epsilon > 0$. By simply forcing $(\lambda_1, \lambda_2, \lambda_U)$ to satisfy

$$\lambda_1(0) = 1, \quad \lambda_2(0) = 1, \quad \lambda_U(0) = 1 \ , \tag{99}$$

the function (98) recovers the cost (95) for the linear system. The inverse optimal cost (98) resembles the linear optimal cost (95) in having a quadratic-like form in x_1 and x_2. We choose $\epsilon = 0.1$. The solution to Riccati equation (19) with $Q_l = I_2$ and $r_l = 1$ is calculated as $P_l = S_l^T P S_l$, where

$$P = \begin{bmatrix} 1.183 & 0 \\ 0 & 1.781 \end{bmatrix}, \quad S_l = \begin{bmatrix} 1 & 0 \\ -0.597 & 1 \end{bmatrix}^{-1} .$$

We now design a global inverse optimal control law. We define $\bar{C}(x) = (0.9)^{1/2} I_2$ and $\Lambda_l = I_3$. To realize the cost (98), we fix Θ as $\Theta(x) = 1$. We carry out the inverse optimal backstepping in Section 6 with the constraints

$$s_1(0) = -0.597,\ \ s_2(0) = -1.781,\ \ \lambda_1(0) = \lambda_2(0) = \lambda_U(0) = 1$$

for local optimality. The feedback law is obtained as

$$u = [-s_2 s_1 \ s_2]x, \quad s_1(x_1) = -x_1 - 0.597, \quad s_2(x) = -11.134((x_1 + 0.1)^2 + 0.15) \tag{100}$$

The generalized scaling parameters are selected as

$$\lambda_1(x_1) = 1, \quad \lambda_2(x) = \frac{0.26}{(x_1 + 0.4)^2 + 0.1}, \quad \lambda_U(x) = \frac{0.16}{(x_1 + 0.1)^2 + 0.15}$$

For brevity, we call (100) the generalized scaling law. The growth order of the generalized scaling law (100) is like x_1^4 which is the same as Ezal's law (97). In Fig.3(a), solid lines represent the state response of the system Σ driven by the generalized scaling law. Dashed lines are the state $x(t)$ for Ezal's law. The response is computed with $w_p \equiv 0$ and $x(0) = [-0.2\ 1]^T$. The performance of the two control laws is almost the same. As shown in Fig.3(b), the control input signals u of both control laws are also almost identical. Dotted lines are for the nonlinear system Σ controlled by the linear optimal law. Thus both the generalized scaling law and Ezal's control law have desirable properties as global inverse optimal controls with linear optimality.

We next consider the same nonlinear system Σ in the presence of an uncertainty:

$$\begin{cases} \dot{x}_1 = x_1^2 + x_2 + w_p \\ \dot{x}_2 = u + \Delta(x)x_1 \end{cases}, \quad \begin{array}{c} |\Delta(x)| < \eta(1-\epsilon)^{1/2} \\ \forall x \in \mathcal{R}^2,\ \eta = 2.5 \end{array} \tag{101}$$

with a sufficiently small $1 \gg \epsilon > 0$. We decompose (101) into a nominal part Σ and an uncertain part Σ_Δ as follows:

$$\Sigma : \begin{cases} \dot{x}_1 = x_1^2 + x_2 + w_p \\ \dot{x}_2 = u + d^{-1/2} w_\delta \\ z_\delta = d^{1/2} x_1 \end{cases}, \quad \Sigma_\Delta \ : \ w_\delta = \Delta(x) z_\delta \tag{102}$$

The positive number d will be chosen later. According to Theorem 8, we can achieve robust stability against the uncertainty $\Delta(x)$ if we modify the cost functional (98) for global inverse optimal control as follows:

$$J(u, w_p, w_\delta) = \int_0^\infty (\lambda_1(x_1)x_1^2 + \lambda_2(x)x_2^2 + \lambda_U(x)u^2 - \gamma^2 w_p^2 + \lambda_2(x) z_\delta^2 - \eta^{-2}\lambda_2(x) w_\delta^2 + \hat{q}(x))dt \tag{103}$$

$$\gamma = 5$$

where

$$\begin{array}{l} \lambda_1(x_1) > 0, \quad \lambda_2(x) > 0, \quad \lambda_U(x) > 0, \quad \forall x \in \mathcal{R}^2 \\ \hat{q}(x) > -\epsilon((1+d)\lambda_1(x_1)x_1^2 + \lambda_2(x)x_2^2),\ \forall x \in \mathcal{R}^2 \setminus \{0\}, \\ \hat{q}(0) = 0, \quad \left.\dfrac{\partial \hat{q}(x)}{\partial x}\right|_{x=0} = 0, \quad \left.\dfrac{\partial^2 \hat{q}(x)}{\partial x^2}\right|_{x=0} = 0\ . \end{array}$$

The cost functional (103) is equivalently represented in the state-dependent scaling form as

$$J(u,w) = \int_0^\infty \left[z^T \Lambda z - w^T \Theta w + \hat{q}(x)\right] dt \tag{104}$$

$$\Lambda = \begin{bmatrix} \bar{\Lambda} & 0 \\ 0 & \lambda_U \end{bmatrix}, \quad \bar{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 I_2 \end{bmatrix}, \quad \Theta = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$\hat{q}(x) > -\epsilon z^T \Lambda z, \quad \forall x \in \mathcal{R}^n \setminus \{0\}, \quad \hat{q}(0) = 0$$

for the augmented system

$$\Sigma_a : \begin{cases} \dot{x}_1 = x_1^2 + x_2 + \gamma^{-1} w_p \\ \dot{x}_2 = u + \eta^{-1} d^{-1/2} w_\delta \\ z_\delta = d^{1/2} x_1 \end{cases}, \quad w = \begin{bmatrix} w_p \\ w_\delta \end{bmatrix}, z = \begin{bmatrix} x_1 \\ x_2 \\ z_\delta \\ u \end{bmatrix}. \tag{105}$$

Let Σ_{al} denote the Jacobian linearization of (105). We now choose the cost functional for Σ_{al} by substituting $\lambda_1(0) = \lambda_2(0) = \lambda_U(0) = 1$ and $\hat{q}(0) = 0$ into (104):

$$J_l(u_l, w_l) = \int_0^\infty \left[z^T \Lambda_l z - w^T \Theta_l w\right] dt, \ \Lambda_l = I_4, \ \Theta_l = I_2 \tag{106}$$

Note that this linear optimal control is a solution to the problem of robustly stabilizing $(\Sigma_l, \Sigma_\Delta)$ and achieving

$$J_l(u_l, w_{pl}) = \int_0^\infty (x_1^2 + x_2^2 + u_l^2 - \gamma^2 w_{pl}^2) dt \leq V(x(0)) \ . \tag{107}$$

with a positive definite function $V(x) = x^T P_l x$ for any disturbance w_p and any admissible uncertainty $\Delta(x)$. Here, Σ_l denotes the Jacobian linearization of the nonlinear nominal system Σ in (102). The constant d is nothing but a scaling factor for robust linear control. Because the linear optimal control problem does not have a solution with $d = 1$, we take $d = 25$. Then, the solution P_l to Riccati equation (19) is

$$P_l = S_l^T P S_l, \quad P = \begin{bmatrix} 10.55 & 0 \\ 0 & 5.076 \end{bmatrix}, \ S_l = \begin{bmatrix} 1 & 0 \\ -1.558 & 1 \end{bmatrix}^{-1} .$$

We now solve the inverse optimal problem for the nonlinear system Σ with respect to (104). We choose $\epsilon = 0.02$. The inverse optimal backstepping is performed under the constraints

$$s_1(0) = -1.558, \ s_2(0) = -5.076, \ \lambda_1(0) = \lambda_2(0) = \lambda_U(0) = 1 \ .$$

The feedback law is found as $u = [-s_2 s_1 \ s_2]x$ with

$$s_1(x_1) = -x_1 - 1.558, \ s_2(x) = -83.35((x_1 + 0.03)^2 + 0.06) \tag{108}$$

The SD scaling parameters are obtained as

$$\lambda_1(x_1) = 1, \quad \lambda_2(x) = \frac{24.25}{(x_1 + 1.5)^2 + 22}, \quad \lambda_U(x) = \frac{0.0609}{(x_1 + 0.03)^2 + 0.06}$$

The scaling control law (108) not only robustly stabilizes (Σ, Σ_Δ), but also it achieves the disturbance attenuation described by

$$J(u,w) = \int_0^\infty ((1 - \epsilon(1+d))(\lambda_1(x_1)x_1^2 + \lambda_2(x)x_2^2) + u^2 - \gamma^2 w_p^2) dt \leq V(x(0)) \tag{109}$$

with $V(x) = x^T S^T P S x$ for any disturbance w_p and any admissible uncertainty $\Delta(x)$. Figure 4(a) depicts state trajectories $x(t)$ of the system Σ driven by the SD scaling control law (108) in the case of $x(0) = [-0.2\ 1]^T$ and $w_p \equiv 0$. Solid lines represent the state response in the presence of the uncertainty $\Delta(x) = 2$. Dashed lines are the state response in the absence of uncertainty. The response for $\Delta(x) = 2$ is almost identical to the uncertainty-free case. The input signals required to stabilize the system against $\Delta(x)$ are shown in Fig.4(b). We see that the SD scaling law (108) uses more control effort than that shown in Fig.3(b). This is to be expected because the control law (108) guarantees the robustness of global asymptotic stability against the uncertainty $\Delta(x)$.

Acknowledgments. The author is grateful to Randy A. Freeman for discussion in this research direction during the author's visit to Northwestern University from May 1998 through November 1998. This present work was motivated by discussions the author and Randy. A. Freeman had with Kenan Ezal, Zigang Pan, and Petar Kokotović concerning their results in [8], and the author is grateful to them for sharing their contributions with us.

References

[1] R. Marino and P. Tomei, "Robust stabilization of feedback linearizable time-varying uncertain nonlinear systems," Automatica, vol. 29, pp. 191–189, 1993.

[2] Z. Qu, "Robust control of nonlinear uncertain systems under generalized matching conditions," Automatica, vol. 29, pp. 985–998, 1993.

[3] R. A. Freeman and P. V. Kokotović, "Backstepping design of robust controllers for a class of nonlinear systems," in *Proceedings of the IFAC Nonlinear Control Systems Design Symposium*, Bordeaux, France, June 1992, pp. 307–312.

[4] I. Kanellakopoulos, P. V. Kokotović, and S. Morse, "Systematic design of adaptive controllers for feedback linearizable systems," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1241–1253, 1991.

[5] M. Krstić, I. Kanellakopoulos and P. V. Kokotović, *Nonlinear and adaptive control design*, John Wiley & Sons, New York, 1995.

[6] R. A. Freeman and P. V. Kokotović, *Robust nonlinear control design: State-space and Lyapunov techniques*, Birkhäuser, Boston, 1996.

[7] M. Krstić and Z. Li, "Inverse optimal design of input-to-state stabilizing nonlinear controllers," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 336–350, 1998.

[8] K. Ezal, Z. Pan and P. V. Kokotović, "Locally optimal backstepping design," in *Proc. 36th IEEE Conf. Decis. Contr.*, Dec. 1997, pp. 1767–1773.

[9] S. Battilotti, "Robust stabilization of nonlinear systems with pointwise norm-bounded uncertainties: a control Lyapunov function approach," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 3–17, 1999.

[10] R. Sepulchre, M. Janković and P. V. Kokotović, *Constructive nonlinear control*, Springer-Verlag, New York, 1997

[11] H. Ito, "Robust control for nonlinear systems with structured $\mathcal{L}_2$-gain bounded uncertainty," *Syst. Contr. Lett.*, vol. 28, pp.167–172, 1996.

[12] H. Ito, "Semi-global robust nonlinear control: state-dependent scaling and computational aspects," in *Proc. Amer. Contr. Conf.*, June 1998, pp. 1663–1668.

[13] H. Ito, "State-dependent scaling for robust nonlinear control: Techniques and effectiveness," in *Proc. 37th IEEE Conf. Decis. Contr.*, Dec. 1998, pp. 4111–4114.

[14] H. Ito and R. A. Freeman, "A new look at robust backstepping through state-dependent scaling design," in *Proc. 1999 Amer. Contr. Conf.*, to appear.

[15] A. van der Schaft, *$\mathcal{L}_2$-gain and passivity techniques in nonlinear control*, Lecture Notes in Control and Information Sciences 218, Springer-Verlag, 1996.

[16] A. Isidori and A. Astolfi, "Disturbance attenuation and $\mathcal{H}^\infty$-control via measurment feedback in nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 1283–1293, 1992.

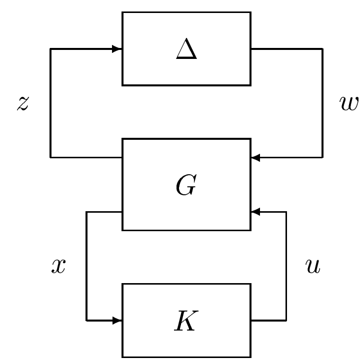


Figure 1: A robust control paradigm.

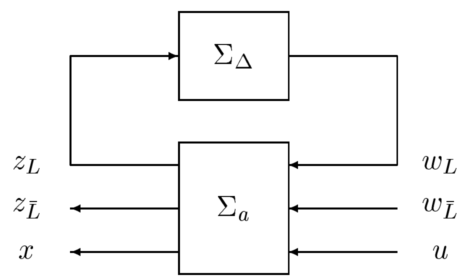
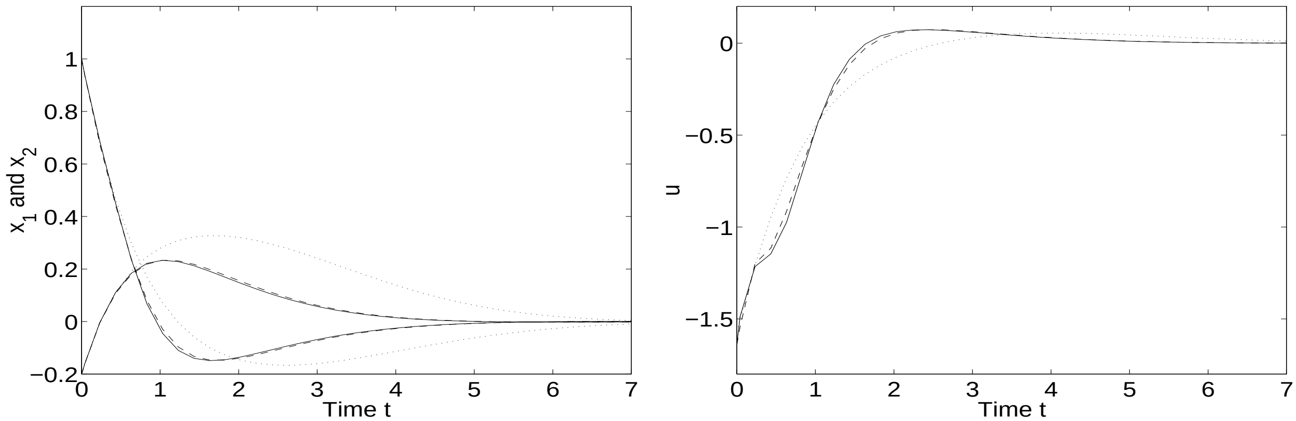
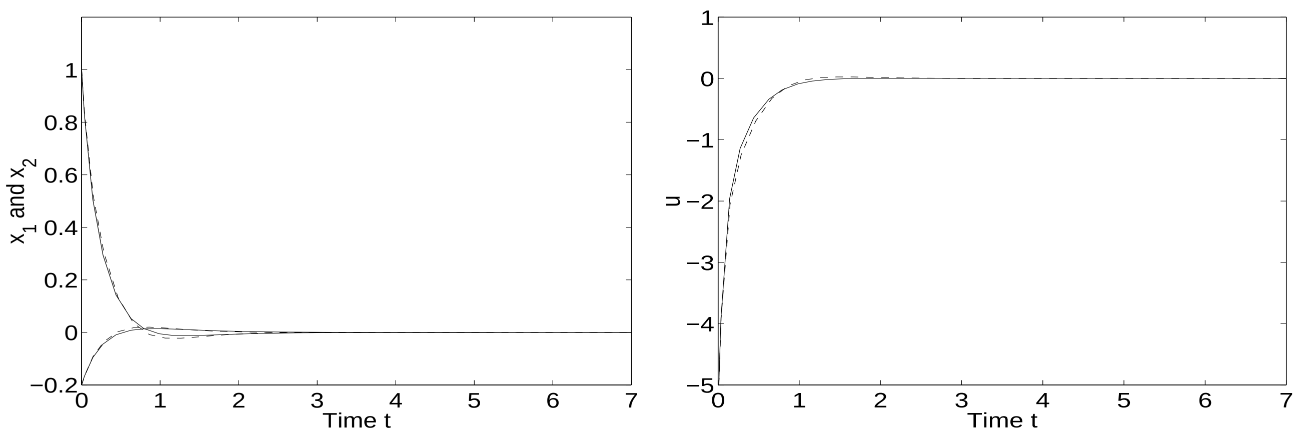


Figure 2: Uncertain nonlinear system



(a) State (b) Input

Figure 3: Response of nominal control system.



(a) State (b) Input

Figure 4: Response with SD scaling control law.