ISSN 1344-8803, CSSE-3 June 1, 1999

State-Dependent Scaling Design for Robust Backstepping via Output Feedback[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

Extended Abstract: This paper considers global robust stabilization of a class of nonlinear systems via output feedback. A new approach to output-feedback backstepping is proposed. The approach provides us with a systematic design procedure which can handle output-feedback stabilization problems of strict-feedback nonlinear systems in a unified way. More importantly, the approach by itself has a mechanism of achieving robust stabilization against a general class of structured uncertainties in the procedure. Compared with the state-feedback global stabilization, the the class of uncertainties which has been treated by the literature of global robust stabilization problems via output feedback is quite restricted in spite of the practical importance of considering various locations and structure of uncertainties. The approach presented in this paper can be considered as an successful extension of the author's state-dependent design for state-feedback backstepping to the output feedback case. Thereby, this paper shows the power of the general concept of state-dependent scaling design for nonlinear systems control by looking at output-feedback stabilization problems, especially in a backstepping manner. The scaling approach allows us to treats both static and dynamic uncertainty in an unified way and , in addition, be able to clarify the difference between their

[1]Technical Report in Computer Science and Systems Engineering, Log Number CSSE-3, ISSN 1344-8803.

[2]The current version of the paper was completed by January 20, 1999 when the author visited to Dept. of Appl. Mech. & Eng. Sci., University of California at San Diego, La Jolla, CA 92093-0411, USA.

[3]Author for correspondence

consequences of stabilization in a simple way. The output feedback design proposed also inherits advantages of SD scaling design such as automatic computation of backstepping based on optimization. Controllers in this paper are dynamic feedback which consists of observer and feedback gain(or controller). The essential difference between nominal stabilization and robust stabilization is described. It is shown that observer design cannot be separated globally from controller design. The observer should be designed strong enough to compensate "nonlinear size" of the uncertainty on the entire state-space. The coupling is natural and inevitable in robust stabilization as it is for linear systems. In addition, for nonlinear systems, nonlinearity of the coupling is crucial for global stabilization which cannot be compensated globally by either feedback-gain or observer-gain independently. This fact contrasts with nominal stabilization in which it is possible to stabilize the whole system globally by designing controller strong enough whenever the observer dynamics by itself design to be only stable(or, vice versa). Strong observers required for robust stabilization may not exist unless the output have the full information of the state. If the nonlinear size of uncertainties are small enough, the global robust stabilization can be certainly achieved. This paper shows the condition of allowable size and nonlinearity of uncertainties for which robust stabilization can be done via backstepping. The condition is considered as the index γ which describes the largest allowable size of uncertainty in robust stabilization via linear $\mathcal{H}^\infty$ control. Indeed, for linear systems, the condition of γ has coupling between feedback gain and observer design(or Riccati inequalities). In addition to the coupling, the condition of the uncertainty size in this paper exhibits a recursive form because of backstepping. Another feature of the output backstepping procedures in this paper is that it does not require Young's inequality. Instead, the paper uses the Schur complements formula which gives a necessary and sufficient condition for negativity of a quadratic form. This paper also proposes a novel recursive procedure of robust observer design, which resembles backstepping or forwarding for controller design.

Key Words: robust backstepping; state-dependent scaling; global robust stability; output feedback; observer design; input-to-state stability; matrix inequality; convex optimization.

1 Introduction

For global stabilization of uncertain nonlinear systems in the so-called strict-feedback form, backstepping requires domination of uncertain nonlinearities at each step of its recursive procedure(Krstić *et al.*, 1995; Freeman and Kokotović, 1996). Such domination is achieved through the choice of appropriate dominating functions which satisfy certain inequalities in the Lyapunov derivative corresponding to the locations and characteristics of uncertain components in the system. Ito and Freeman (1998*a*) has shown that state-dependent scaling provides us with a systematic and unified method for constructing suitable dominating functions in robust backstepping design for state-feedback.

The idea of state-dependent(SD) scaling design was proposed in Ito (1998*a*), which was motivated by the fact that scaling factors of small-gain conditions are allowed to be functions of state variables conditions(Ito, 1996). The drawback of nonlinear $\mathcal{H}_\infty$ control as a nonlinear design tool can be overcome in the sense that SD scaling has the ability to enlarge stability regions in small-gain type robust stabilization(Ito, 1998*b*). The methodology of SD scaling design is applicable not only to strict feedback systems, but also to other general classes of nonlinear systems(Ito, 1998*a*; Ito, 1998*b*). The concept of SD scaling design is general enough to formulate a wide variety of robust nonlinear control problems and it is amenable to computational optimization techniques. When it comes to uncertain systems in the strict-feedback form, the SD scaling design is considered as a unified robust backstepping procedure which encompasses a large class of uncertain nonlinear systems with structured, memoryless and dynamic uncertainties. It has been shown that robust backstepping can be described as recursive selection of appropriate scaling factors(Ito and Freeman, 1998*a*; Ito and Freeman, 1998*b*). The backstepping can be performed by computational optimization as well.

If the state variables are not available for feedback, one may simply give up seeking global stabilization and settle for semi-global stabilization. On this standpoint, there are a lot of paper dealing with semi-global stabilization by output feedback. The idea of input saturation and high-gain observer has been successful for such semi-global stabilization (Esfandiari and Khalil, 1992; Khalil and Esfandiari, 1993; Lin and Saberi, 1995). Teel and Praly (1995) and Teel and Praly (1994) proposed a useful semi-global backstepping lemma and high-gain observers with saturating control for dynamic output feedback. By using these semi-global techniques, a robust stabilization problem was also considered intensively for a certain type and location of unstructured uncertainty, namely, robustness against unknown stable zero dynamics. It is possible to deal with unknown parameters in such a semi-global stabilization as well, e.g.(Lin and Qian, 1998). However, from anther view point, given an uncertain system, semi-global stabilization using high-gain and saturation may be meaningful only if the system cannot be globally stabilized.

There are also global results for output-feedback stabilization of nonlinear systems in a form of strict-feedback or chain of integrators. However, the typical results, e.g.(Krstić *et al.*, 1995) are applicable only to nonlinear systems whose nonlinearities in the system equation do not depend on the states that are not measured. It is, however, not clear what is the essential ingredient of this assumption, apart from their technique of constructing observers and controllers. Aside from inverse optimality, discussion about robust global stabilization via this type of output feedback is absent in spite of their practical importance.

The first objective of this paper is to propose a unified procedure to achieve robust and global stability via output feedback for the class of uncertainty which is as large as the uncertainty tackled in the state-feedback control literature, namely, uncertain systems in the strict-feedback form with nonlinearly bounded uncertainties. In other words, this objective it to enlarging the class of nominal models and especially the structure of uncertainty, that can be globally stabilized under output feedback, In order to accomplish this first objective, this paper successfully extends the author's state-dependent design for state-feedback backstepping to the output feedback case. By doing that, the power of the general concept of state-dependent scaling design for nonlinear systems control is shown as well. The author's position is seeking global stabilization in stead of settling for semi-global stabilization from the beginning. Thereby, the paper clarifies essential points required to make global stabilization robust as desired. Thus, the second objective is to characterize the essential difference between nominal global stabilization and robust global stabilization in output feedback control. The robustness in this paper is more desirable in that the size and location of uncertainty is prescribed *a priori*, which is completely different from the inverse optimal type of robustness. The backstepping is developed without the assumption that requires the nominal system to have nonlinearities depending only on measured states, i.e., $\phi(y)$ where y is the output. This paper clarifies that the feedback-gain part of output feedback design by itself does not need to exclude nonlinearities such as $\phi(y)x$, where x is unnecessarily measured. This paper describes what kind of task is essentially required for observer design in such a case. The paper also shows a condition on which robust stabilization can be achieved globally for a prescribed class of uncertainties. It will be shown that, exclusively for nonlinear systems, "nonlinear size" of uncertainties, appeared as coupling, is crucial for global robust stabilization, which cannot compensated globally by either feedback-gain or observer-gain independently.

The idea of SD scaling approach to backstepping in this paper is as follows:

- characterize robustness analysis by SD scaling
- introduce coordinate transformation to the entire closed-loop system in order to create a freedom in choosing Lyapunov functions
- use the Schur complements formula to extract a recursive structure of design
- solve the design problem by selecting SD scaling and coordinate transformation recursively.

Furthermore, it is completed by

- show that the design problem is recursively linear in the parameters of SD scaling and coordinate transformation
- show the existence of solutions
- provide computational formulas and analytical solutions

One feature of the backstepping proposed in this paper is that the procedures are amenable to automated *numerical* computation based on convex optimization. Since the backstepping is performed by domination, it is unnecessary to use precise parameters of systems in the control law, which prevents the controller from having long and complicated terms. Another important feature of this paper is that the output backstepping is shown to be feasible without using Young's inequality. Instead, the paper uses the Schur complements formula which gives a necessary and sufficient condition for negativity of a quadratic form.

The author needs to explain the standpoint of this paper since it is quite different from those of nonlinear adaptive control and many of backstepping papers. The author's point of view is similar to that of linear robust stabilization via $\mathcal{H}^\infty$ control. The roles of $\mathcal{H}^\infty$ types robust control are

- provide a method of solving(more precisely, trying to solve) the problem
- characterize a condition under which the robust stabilization is solvable
- provide information about how large size of uncertainty is allowable.

The latter two roles are necessary since the problem by itself does not always have the solution. The reason why this situation occurs is that we specify the nominal system and structure and size of uncertainties *a priori*. A robust stabilization problem is solvable obviously if the uncertainty is sufficiently small. This type of robust control is attractive in the sense that it provides us with a way to obtain a control law even if it is not as good as we originally desired. It theoretically persuades us to give up seeking unreasonably large uncertainty, in particular, since the constantly scaled $\mathcal{H}^\infty$ control is necessary and sufficient for achieving robust stabilization against time-varying uncertainty. This paper looks at global robust stabilization of nonlinear systems from the same point of view. In addition, this paper demonstrates a class of uncertain systems which is always robustly stabilizable for arbitrarily large size and arbitrarily fast growth order uncertainties. The latter standpoint is more common in the nonlinear literature. In this way, this paper takes practically good positions from both the sides.

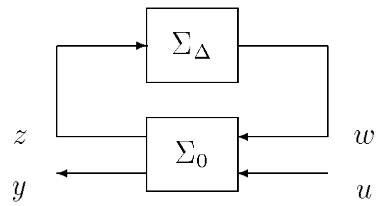

Figure 1: Uncertain nonlinear plant Σ_P

2 Uncertain nonlinear systems

Consider the uncertain nonlinear system Σ_P shown in Fig.1. Here, Σ_0 denotes a nominal plant and Σ_Δ represents the uncertainty and modeling error of the plant. We assume that the nominal part Σ_0 is described by

$$\Sigma_0 : \begin{cases} \dot{x} = A(y)x + B(y)w + G(y)u & x(t) \in \mathcal{R}^n, u(t) \in \mathcal{R}^1 \\ z = C(y)x & , w(t) \in \mathcal{R}^p, z(t) \in \mathcal{R}^p \\ y = C_y(y)x & y(t) \in \mathcal{R}^r \end{cases} \tag{1}$$

The matrix-valued functions A, B, C, G and C_y are assumed to be $\mathbf{C}^0$ functions. The vectors w and z are defined as

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}, \quad \begin{array}{c} w_i(t) \in \mathcal{R}^{p_i} \\ z_i(t) \in \mathcal{R}^{p_i} \\ p_i \geq 0, \quad p = \sum_{i=1}^{n} p_i \\ i = 1, 2, \cdots, n \end{array} \tag{2}$$

Suppose that the uncertain system Σ_Δ has the following structure of nonlinear mappings $\Delta : z \mapsto w$.

$$\Delta = \text{block-diag}[\Delta_1, \Delta_2, \cdots, \Delta_n], \tag{3}$$

where some of the mappings $\Delta_i : z_i \mapsto w_i$, $i = 1, 2, \ldots, n$ can be zero in vector size. Each uncertainty Δ_i is allowed to have three types of components:

$$\Delta_i : z_i = \begin{bmatrix} z_{id} \\ z_{is} \\ z_{ir} \end{bmatrix} \mapsto w_i = \begin{bmatrix} w_{id} \\ w_{is} \\ w_{ir} \end{bmatrix}, \quad w_i = \begin{bmatrix} \Delta_{id} & 0 & 0 \\ 0 & \Delta_{is} & 0 \\ 0 & 0 & \Delta_{ir} \end{bmatrix} z_i \ . \tag{4}$$

Here, Δ_{id} represents a dynamic system. Δ_{is} and Δ_{ir} denote full static and repeated static scalar systems, respectively. It is unnecessary for Δ_i to have all types of uncertainty. The dynamic uncertainty Δ_{id} is defined by

$$\Delta_{id} : \begin{cases} \dot{x}_{\Delta_i} = f_{\Delta_{id}}(x_{\Delta_i}, z_{id}, t) \\ w_{id} = h_{\Delta_{id}}(x_{\Delta_i}, z_{id}, t) \end{cases}, \tag{5}$$

where $f_{\Delta_{id}}(0,0,t) = 0$ and $h_{\Delta_{id}}(0,0,t) = 0$ are satisfied for all $t \geq 0$ and $f_{\Delta_{id}}$ and $h_{\Delta_{id}}$ are vector-valued $\mathbf{C}^0$ functions. The full static part Δ_{is} is described by

$$\Delta_{is} : w_{is} = h_{\Delta_{is}}(z_{is}, t), \tag{6}$$

where $h_{\Delta_{is}}$ is a vector-valued $\mathbf{C}^0$ function and $h_{\Delta_{is}}(0,t) = 0$ for all $t \geq 0$. The repeated static part Δ_{ir} is defined with $r_i > 1$ copies of a static scalar system δ_{ir} :

$$\Delta_{ir} = \operatorname*{diag}_{j=1}^{r_i} \delta_{ir} = \delta_{ir} I_{r_i}, \qquad \delta_{ir} : \ w_{\delta_{ir}} = h_{\delta_{ir}}(t) z_{\delta_{ir}} \tag{7}$$

where $h_{\delta_{ir}}$ is a scalar-valued $\mathbf{C}^0$ function. We consider the following class of uncertainty Σ_Δ.

Definition 1 *The uncertainty Σ_Δ is said to be admissible if* **(i)**-**(iii)** *are satisfied for $i = 1, 2, \ldots, n$:* **(i)** *Δ_{id} has $\mathcal{L}_2 -$ gain less than or equal to 1 with a radially unbounded $\mathbf{C}^1$ storage function $V_{\Delta i}(x_{\Delta_i})$ satisfying $V_{\Delta i}(0) = 0$.* **(ii)** *Δ_{is} satisfies $\|z_{\Delta_{is}}\|^2 \geq \|w_{\Delta_{is}}\|^2$ for all $t \in [0, \infty)$.* **(iii)** *δ_{ir} satisfies $\|z_{\delta_{ir}}\|^2 \geq \|w_{\delta_{ir}}\|^2$ for all $t \in [0, \infty)$.*

The uncertain system Σ_P has an equilibrium point at $x_{cl} = 0$ when $u \equiv 0$. Roughly speaking, the gain of each uncertainty is assumed to be less than or equal to unity. Uncertainty having super-linear growth (and thus unbounded gain) can still be included by a judicious choice of the nonlinear weights $B(x)$ and $C(x)$. Indeed, Σ_0 not only describes a nominal plant, but also can include information about input-output nonlinearities of uncertainty. The manipulation to choose an appropriate pair of $(\Sigma_0, \Sigma_\Delta)$ taking nonlinearity into account is essentially similar to the idea of introducing functions of signal norms (Sontag and Wang, 1996; Sontag, 1998; Mareels and Hill, 1992). Note that Σ_0 also describes how the uncertainty affects the nominal plant such as geometrical locations, structures of uncertainties where uncertain parameters are present. Remember that $B(x)$ and $C(x)$ specify the "nonlinear size"(including size, nonlinearity, location and structure) of uncertainties.

3 SD scaling analysis for observer-feedback control

With definitions of the uncertainty Σ_Δ in mind, several sets of real-valued scaling matrices will be defined. For notational simplicity, we assume that Δ_{id} and Δ_{is} are square in size of input and output vectors for all $i = 1, 2, \ldots, n$. For the dynamic uncertainty Δ_{id}, we define

$$\mathbf{L}_{id} := \{L_{id} = \lambda_{id} I_{id} : \lambda_{id} > 0\}. \tag{8}$$

Here, I_{id} denotes an identity matrix which is compatible in size with the vector z_{id}. For the full static uncertainty Δ_{is}, a set of scaling is defined by

$$\mathbf{L}_{is} := \{L_{is} = \lambda_{is}(y, \hat{x}) I_{is} : \lambda_{is}(y, \hat{x}) > 0 \ \forall (y, \hat{x}) \in \mathcal{R}^r \times \mathcal{R}^n\}. \tag{9}$$

In the case of the repeated static uncertainty Δ_{ir}, we define two sets of scaling matrices.

$$\mathbf{L}_{ir} := \{L_{ir} : L_{ir}^T(y, \hat{x}) = L_{ir}(y, \hat{x}), \ L_{ir}(y, \hat{x}) > 0 \ \forall (y, \hat{x}) \in \mathcal{R}^r \times \mathcal{R}^n\}. \tag{10}$$

$$\mathbf{R}_{ir} := \{R_{ir} : R_{ir}^T(y, \hat{x}) = -R_{ir}(y, \hat{x}) \ \forall (y, \hat{x}) \in \mathcal{R}^r \times \mathcal{R}^n\}. \tag{11}$$

Here, both L_{ir} and R_{ir} are square matrices whose size is the same as the dimension of z_{is}. These scaling matrices are used to estimate the worst case value of the time-derivative of Lyapunov functions(Ito, 1998*b*; Ito and Freeman, 1998*a*). Let $L_i(y, \hat{x})$ and $R_i(y, \hat{x})$ be defined by

$$\mathbf{L}_i := \left\{ L_i(y, \hat{x}) = \begin{bmatrix} L_{id} & 0 & 0 \\ 0 & L_{is}(y, \hat{x}) & 0 \\ 0 & 0 & L_{ir}(y, \hat{x}) \end{bmatrix} : \begin{matrix} L_{id} \in \mathbf{L}_{id} \\ L_{is} \in \mathbf{L}_{is} \\ L_{ir} \in \mathbf{L}_{ir} \end{matrix} \right\} \tag{12}$$

$$\mathbf{R}_i := \left\{ R_i(y, \hat{x}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & R_{ir}(y, \hat{x}) \end{bmatrix} : R_{ir} \in \mathbf{R}_{ir} \right\}, \tag{13}$$

for $i = 1, 2, \ldots, n$. Note that a constant $\lambda > 0$ satisfies $\lambda I \in \mathbf{L}_i$ and $0 \in \mathbf{R}_i$. Now define two sets of scaling matrices for the whole Σ_Δ as follows:

$$\mathbf{L} := \left\{ L = \operatorname*{block-diag}_{i=1}^{n} L_i(y, \hat{x}), \quad L_i \in \mathbf{L}_i \right\} \tag{14}$$

$$\mathbf{R} := \left\{ R = \operatorname*{block-diag}_{i=1}^{n} R_i(y, \hat{x}), \quad R_i \in \mathbf{R}_i \right\}. \tag{15}$$

These scaling matrices are functions of y and $\hat{x}$. The situation contrasts sharply with the linear systems case where constant scalings are used for time-varying uncertainty. Scaling matrices for static uncertainty are chosen as functions of output and state estimate, while static uncertainty arising in a linear system is usually not distinguished from dynamic uncertainty(Ito, 1996).

We next consider robust stabilization of the uncertain nonlinear system Σ_P by dynamic output feedback. We employ the full order observer

$$\left\{ \begin{array}{l} \dot{\hat{x}} = A(y)\hat{x} + Y(y,\hat{x})(y-\hat{y}) + G(y)u \\ \hat{y} = C_y(y)\hat{x} \end{array} \right. \tag{16}$$

By using the signal $\hat{x}$ estimated by the observer, we choose a dynamic output feedback law as

$$u = K(y,\hat{x})\hat{x} \ . \tag{17}$$

Then, the closed-loop system is written as

$$\frac{d}{dt}\begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} A & GK \\ YC_y & A - YC_y + GK \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} w \tag{18}$$

We now characterize robust stabilization of Σ_P using SD scaling, quadratic Lyapunov functions and a diffeomorphic coordinate change. Consider the diffeomorphism between $\hat{x} \in \mathcal{R}^n$ and $\hat{\chi} \in \mathcal{R}^n$ as follows:

$$\hat{\chi} = S(y,\hat{x})\hat{x} \ . \tag{19}$$

The time-derivative of $\hat{\chi}$ is obtained as

$$\dot{\hat{\chi}} = \left[\frac{\partial S}{\partial y_1}\hat{x}, \frac{\partial S}{\partial y_2}\hat{x}, \cdots, \frac{\partial S}{\partial y_n}\hat{x} \right] C_y \dot{x} + \left[\frac{\partial S}{\partial \hat{x}_1}\hat{x}, \frac{\partial S}{\partial \hat{x}_2}\hat{x}, \cdots, \frac{\partial S}{\partial \hat{x}_n}\hat{x} \right] \dot{\hat{x}} + S(y,\hat{x})\dot{\hat{x}} = V(y,\hat{x})\dot{x} + T(y,\hat{x})\dot{\hat{x}} \ . \tag{20}$$

Define $\tilde{x} = \hat{x} - x$. Then,

$$\frac{d}{dt}\begin{bmatrix} \hat{\chi} \\ \tilde{x} \end{bmatrix} = \begin{bmatrix} V(y,\hat{x}) & T(y,\hat{x}) \\ -I & I \end{bmatrix} \frac{d}{dt}\begin{bmatrix} x \\ \hat{x} \end{bmatrix}, \quad \begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} S(y,\hat{x})^{-1} & -I \\ S(y,\hat{x})^{-1} & 0 \end{bmatrix} \begin{bmatrix} \hat{\chi} \\ \tilde{x} \end{bmatrix} \tag{21}$$

The closed-loop system becomes

$$\frac{d}{dt}\begin{bmatrix} \hat{\chi} \\ \tilde{x} \end{bmatrix} = \begin{bmatrix} (V+T)(A+GK)S^{-1} & -(VA+TYC_y) \\ 0 & A - YC_y \end{bmatrix} \begin{bmatrix} \hat{\chi} \\ \tilde{x} \end{bmatrix} + \begin{bmatrix} VB \\ -B \end{bmatrix} w \tag{22}$$

We also introduce another coordinate transformation to $\tilde{x}$:

$$\eta = W\tilde{x} \tag{23}$$

where W is a constant non-singular matrix. The closed-loop system on the coordinate $(\hat{\chi}, \eta)$ is

$$\frac{d}{dt}\begin{bmatrix} \hat{\chi} \\ \eta \end{bmatrix} = \begin{bmatrix} (V+T)(A+GK)S^{-1} & -(VA+TYC_y)W^{-1} \\ 0 & W(A - YC_y)W^{-1} \end{bmatrix} \begin{bmatrix} \hat{\chi} \\ \eta \end{bmatrix} + \begin{bmatrix} VB \\ -WB \end{bmatrix} w \tag{24}$$

$$z = C\begin{bmatrix} S^{-1} & -W^{-1} \end{bmatrix} \begin{bmatrix} \hat{\chi} \\ \eta \end{bmatrix} \tag{25}$$

The following describes the main idea of the SD scaling approach to the output feedback problem.

Theorem 1 *(i) Suppose that there exist constant symmetric matrices P and $\tilde{P}$ such that*

$$N(y,\hat{x}) = \begin{bmatrix} S^{-T}(A+GK)^T(V+T)^T P + P(V+T)(A+GK)S^{-1} & -P(VA+TYC_y)W^{-1} \\ -W^{-T}(VA+TYC_y)^T P & W^{-T}(A-YC_y)^T W^T \tilde{P} + \tilde{P}W(A-YC_y)W^{-1} \end{bmatrix} < 0 \quad (26)$$

$$P > 0, \quad \tilde{P} > 0 \quad (27)$$

are satisfied for all $(y,\hat{x})$ in $\mathcal{R}^r \times \mathcal{R}^n$, then the nominal nonlinear system Σ_0 is globally uniformly asymptotically stabilized by the dynamic output feedback (16-17). Furthermore, a Lyapunov function is given by $V(x,\hat{x}) = \hat{\chi}^T P \hat{\chi} + \eta^T \tilde{P} \eta$.
(ii) Suppose that there exist constant symmetric matrices P, $\tilde{P}$ and scaling functions $L \in \mathbf{L}$ and $R \in \mathbf{R}$ such that

$$M(y,\hat{x}) = \begin{bmatrix} \left\{ \begin{matrix} S^{-T}(A+GK)^T(V+T)^T P + \\ P(V+T)(A+GK)S^{-1} \end{matrix} \right\} & PVB + S^{-T}C^T R^T & S^{-T}C^T L & -P(VA+TYC_y)W^{-1} \\ B^T V^T P + RCS^{-1} & -L & 0 & -B^T W^T \tilde{P} - RCW^{-1} \\ LCS^{-1} & 0 & -L & -LCW^{-1} \\ -W^{-T}(VA+TYC_y)^T P & -\tilde{P}WB - W^{-T}C^T R^T & -W^{-T}C^T L & \left\{ \begin{matrix} W^{-T}(A-YC_y)^T W^T \tilde{P} + \\ \tilde{P}W(A-YC_y)W^{-1} \end{matrix} \right\} \end{bmatrix} < 0 \quad (28)$$

$$P > 0, \quad \tilde{P} > 0 \quad (29)$$

are satisfied for all $(y,\hat{x})$ in $\mathcal{R}^r \times \mathcal{R}^n$, then the uncertain nonlinear system Σ_P is globally uniformly asymptotically stabilized by the dynamic output feedback (16-17) for any admissible uncertainty Σ_Δ. Furthermore, a Lyapunov function is given by $V(x,\hat{x}) = \hat{\chi}^T P \hat{\chi} + \eta^T \tilde{P} \eta + \sum_{i=1}^n \lambda_{id} V_{\Delta i}(x_{\Delta_i})$.

The analysis problem of robust stability is reduced into the existence of scaling matrices which make M negative. This is actually considered as the definition of the state-dependent scaling approach to output feedback control with full-order observers.

Although the representation (22) may seem to allow us to use a sort of separation between state-feedback stabilization and observer design somehow at a glance, it is certainly not true for nonlinear systems stabilization. To explain this point, we need the following lemma.

Lemma 1 *Consider a symmetric matrix*

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix} \quad (30)$$

(i) *Schur complements formula : $M < 0$ is equivalent to*

$$M_{22} < 0, \quad M_{11} - M_{12} M_{22}^{-1} M_{12}^T < 0 \quad (31)$$

(ii) *Young's inequality : $M < 0$ is satisfied if*

$$M_{22} + \Gamma^{-1} < 0, \quad M_{11} + M_{12} \Gamma M_{12}^T < 0 \quad (32)$$

$$\Gamma = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \gamma_n \end{bmatrix} > 0 \quad (33)$$

Proof : (ii) The inequalities can be derived easily by using elementary linear algebra as follows:

$$0 > M_{11} + M_{12}\Gamma M_{12}^T > M_{11} + M_{12}\Gamma M_{12}^T - M_{12}(\Gamma + M_{22}^{-1})M_{12}^T = M_{11} - M_{12}M_{22}^{-1}M_{12}^T \tag{34}$$

Instead, the purpose is to show that the pair of inequalities in (32) is an alternative expression of Young's inequality:

$$2y^Tz \leq y^T\Gamma y + z^T\Gamma^{-1}z \tag{35}$$

where y and z are vectors. It is easily verified that

$$\begin{bmatrix} x \\ z \end{bmatrix}^T \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = x^TM_{11}x + 2x^TM_{12}z + z^TM_{22}z \tag{36}$$

$$\leq x^TM_{11}x + x^TM_{12}\Gamma M_{12}^Tx + z^T\Gamma^{-1}z + z^TM_{22}z \tag{37}$$

$$= \begin{bmatrix} x \\ z \end{bmatrix}^T \begin{bmatrix} M_{11} + M_{12}\Gamma M_{12}^T & 0 \\ 0 & M_{22} + \Gamma^{-1} \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} \tag{38}$$

∎

The inequalities (32) of matrices are not the formula which is usually called Young's inequality. However, this paper refers to that as Young's inequality in order to distinguish that from the Schur complements formula. It is also true that the inequalities (32) has appeared as an ordinary Young's inequality of vectors or scalars in nonlinear systems control. A common role of Young's inequality is to get rid of products of two vectors in the Lyapunov derivative and to get a decoupled quadratic expression. This is explained in the proof of the above theorem. The Schur complements formula looks at the negativity in terms of matrices in stead of the scalar value of quadratic forms. The Schur complements formula gives a necessary and sufficient condition while Young's inequality is only sufficient. The idea of Young's inequality is to replace the full information of the matrix $-M_{22}^{-1}$ with simple scalar parameters γ_i at a price of loosing necessity. Actually, an alternative statement of the Schur complements formula is as follows: $M < 0$ holds if and only if there exists a diagonal matrix $\Gamma > 0$ such that

$$M_{22} + \Gamma^{-1} < 0, \quad M_{11} - M_{12}M_{22}^{-1}M_{12}^T < 0 \tag{39}$$

are satisfied. Compared with the Schur complements, Young's inequality is conservative. The Schur complements is superior to Young's inequality in this sense although Young's inequality is a common tool in nonlinear systems design(Krstić *et al.*, 1995; Freeman and Kokotović, 1996; Sepulchre *et al.*, 1997). From this standpoint, this paper replaces the task of Young's inequality with the Schur complements. The output feedback design will be shown to have recursive structures for backstepping even if Young's inequality is not used. In other words, this paper proposes backstepping procedures without introducing any conservatism *in solving problems recursively* except that Theorem 1 is a sufficient condition (note that a recursive structure of solution by itself may have unnecessary conservatism). This may not only allows the design to tolerate large size of uncertainties, but also prevent controllers from having unnecessary high gain and harmfully first or slow growth order.

Corollary 1 *Assume that there exists a constant matrix $\tilde{P} > 0$ such that*

$$H(y, \hat{x}) := W^{-T}(A - YC_y)^TW^T\tilde{P} + \tilde{P}W(A - YC_y)W^{-1} < 0 \tag{40}$$

holds for all $(y, \hat{x}) \in \mathcal{R}^r \times \mathcal{R}^n$
(i) Suppose that there exists a constant matrix $P > 0$ such that the inequality

$$\bar{N}(y, \hat{x}) := N_{11}(y, \hat{x}) - N_{12}(y, \hat{x})H^{-1}(y, \hat{x})N_{12}^T(y, \hat{x}) < 0 \tag{41}$$

$$N_{11}(y, \hat{x}) := S^{-T}(A + GK)^T(V + T)^TP + P(T + V)(A + GK)S^{-1}, \ N_{12}(y, \hat{x}) := P(VA + TYC_y)W^{-1} \tag{42}$$

is satisfied for all $(y,\hat{x})$ in $\mathcal{R}^r\times\mathcal{R}^n$, then the nominal nonlinear system Σ_0 is globally uniformly asymptotically stabilized by the dynamic output feedback (16-17). Moreover, if Σ_0 is a linear system and if S is constant, the set of conditions (41) and (40) is equivalent to the existence of $P>0$ and $\tilde{P}>0$ satisfying

$$N_{11} := S^{-T}(A+GK)^T(V+T)^TP+P(V+T)(A+GK)S^{-1}<0 \tag{43}$$
$$H := W^{-T}(A-YC_y)^TW^T\tilde{P}+\tilde{P}W(A-YC_y)W^{-1}<0 \tag{44}$$

(ii) Suppose that there exist a constant matrix $P>0$ and scaling functions $L\in\mathbf{L}$ and $R\in\mathbf{R}$ such that the inequality

$$\bar{M}(y,\hat{x}) := M_{11}(y,\hat{x})-M_{12}(y,\hat{x})H^{-1}(y,\hat{x})M_{12}^T(y,\hat{x})<0 \tag{45}$$
$$M_{11}(y,\hat{x}) := \begin{bmatrix} \left\{\begin{matrix} S^{-T}(A+GK)^T(V+T)^TP+ \\ P(T+V)(A+GK)S^{-1}\end{matrix}\right\} & PVB+S^{-T}C^TR^T & S^{-T}C^TL \\ B^TV^TP+RCS^{-1} & -L & 0 \\ LCS^{-1} & 0 & -L \end{bmatrix} \tag{46}$$
$$M_{12}(y,\hat{x}) := \begin{bmatrix} P(VA+TYC_y)W^{-1} \\ B^TW^T\tilde{P}+RCW^{-1} \\ LCW^{-1}\end{bmatrix} \tag{47}$$

is satisfied for all $(y,\hat{x})$ in $\mathcal{R}^r\times\mathcal{R}^n$, then the uncertain nonlinear system Σ_P is globally uniformly asymptotically stabilized by the dynamic output feedback (16-17) for any admissible uncertainty Σ_Δ.
Proof : (i) The conditions (41) and (40) are straightforward from (26) by using the Schur complements formula. Obviously, (41) and (40) imply (43) and (44). Now, suppose that $P>0$ is a solution to (43) with a constant S for a linear system Σ_0. Let $\tilde{P}$ be a solution to (44). If $\tilde{P}$ in (41) is replaced by $\beta\tilde{P}$, the inequality (41) is satisfied for a sufficient large constant $\beta>0$.
(ii) It is straightforward from the Schur complements formula. ∎

The two inequalities (43) and (44) in this corollary merely represent the separation principle for linear systems. The conditions (43) and (44), however, do not guarantee global stability for nonlinear Σ_0. If Σ_0 is nonlinear, β in the above proof may be required to be unbounded as y or $\hat{x}$ goes to $\pm\infty$. If β is a function of $(y,\hat{x})$, there is no guarantee that there exists a Lyapunov function V which is consistent with

$$\frac{\partial V}{\partial[\hat{\chi}^T,\eta^T]^T} = 2\left[\,\hat{\chi}^TP,\ \eta^T\beta(y,\hat{x})\tilde{P}\,\right]$$

for (43) and (44). It is, however, true that (41) can be satisfied semi-globally by a sufficient large constant β. We may achieve semi-global stabilization by using the separation (43-44) and taking into account the level set of the Lyapunov function $V(x,\hat{x})=\chi^TP\chi+\eta^T\beta\tilde{P}\eta$ deformed by β. This paper does not pursue this obvious direction of semi-global stabilization since it does not capture essential points required for global and nonlinear stabilization. This paper, instead, is focused on global stabilization and characterizes requirements for global stabilization. As for robust stabilization, we cannot separate observer design completely from robust stabilization in a global sense. The separation argument in (i) of Corollary 1 is not applicable to (ii) either even for linear Σ_0 because of the coupling term M_{12}(especially the term $B^TW^T\tilde{P}$) between feedback and observer in $M<0$. In fact, linear robust control theory tells us that observer design must be coupled with robustification of stabilization against uncertainties. In other words, the observer should be designed strong enough by taking into account the effect of uncertainty and robustness objectives.

4 A robust strict-feedback form and observers

This section defines the class of uncertain nonlinear systems to which output backstepping design via SD scaling will apply. The output equation of the system is supposed to given by

$$y = x_1 \tag{48}$$

or equivalently

$$C_y = [\,1\ \ 0\ \cdots\ 0\,] \tag{49}$$

This case is sometimes called output feedback in the nonlinear control literature(Krstić *et al.*, 1995). This paper deals with the uncertain nonlinear system Σ_P under the following structural assumptions. First, we assume that A and G can be written in the form

$$A(x_1)=\begin{bmatrix} a_{11} & a_{12} & 0 & \cdots\cdots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots\ \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n-1,1} & a_{n-1,2} & \cdots\cdots & & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots\cdots & & a_{n,n} \end{bmatrix}, G(x_1)=\begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_{n,n+1} \end{bmatrix}. \tag{50}$$

with $\mathbf{C}^0$ scalar functions a_{ij} of the measured state x_1. The function $a_{ij}(x_1)$ is required to satisfy

$$a_{i,i+1}(x_1) \neq 0, \quad 1 \leq i \leq n, \quad \forall x_1 \in \mathcal{R} \tag{51}$$

As for functions B and C, we assume

$$B(x_1) = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ B_{21} & B_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ B_{n,1} & \cdots & B_{n,n-1} & B_{n,n} \end{bmatrix}, \qquad C(x_1) = \begin{bmatrix} C_{11} & 0 & \cdots & 0 \\ C_{21} & C_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ C_{n,1} & \cdots & C_{n,n-1} & C_{n,n} \end{bmatrix} \tag{52}$$

where $B_{ij}(x_1) \in \mathcal{R}^{1\times p_i}$ and $C_{ij}(x_1) \in \mathcal{R}^{p_i\times 1}$. Then, the uncertainty affects the system as

$$B(x)w = \begin{bmatrix} B_{11}\Delta_1 C_{11} & 0 & 0 & \cdots \\ B_{21}\Delta_1 C_{11} + B_{22}\Delta_2 C_{21} & B_{22}\Delta_2 C_{22} & 0 & \ddots \\ \vdots & \vdots & \ddots & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{bmatrix} \tag{53}$$

This expression is only an aid for illustrating the structure of the uncertainty and is mathematically ambiguous. The operator Δ_i in the above equation does not represent matrix multiplication but nonlinear mappings which can have dynamics with initial conditions. For simplicity, this paper assumes that the system does not have any uncertainties in the virtual control coefficients which appear in the backstepping procedure. It is certainly possible to extend the idea of SD scaling easily to the uncertain system which has Δ blocks in a more general manner as in Ito and Freeman (1998*a*). Because each entry $B_{ii}\Delta_j C_{ji}$ above is scalar, a repeated static uncertainty can always be represented by a scalar full static uncertainty. However, we include the repeated representation here because it allows more degrees of freedom in the scaling design and it also prepares the way for a multivariable version of our results.

Two types of properties of observers will be used in this paper.

Ordinary observer : *The observer-gain $Y(x_1)$ is chosen as a $\mathbf{C}^0$ function matrix such that there exist a constant matrix $\tilde{P}$ and a $\mathbf{C}^0$ function matrix $Q_y(x_1)$ satisfying*

$$H(x_1) := (A - YC_y)^T\tilde{P} + \tilde{P}(A - YC_y) < -Q_y \tag{54}$$

$$\tilde{P} > 0, \quad Q_y > 0 \tag{55}$$

hold for all $x_1 \in \mathcal{R}$.

Robust observer : *Given a matrix-valued function* $\Gamma(x_1) > 0$. *The* $\mathbf{C}^0$ *observer-gain function* $Y(x_1)$ *and the constant matrix* W *are chosen such that there exists a constant diagonal matrix* $\tilde{P}$ *satisfying*

$$H(x_1) := W^{-T}(A - YC_y)^T W^T \tilde{P} + \tilde{P}W(A - YC_y)W^{-1} < -\Gamma^{-1} \tag{56}$$

$$\tilde{P} > 0 \tag{57}$$

hold for all $x_1 \in \mathcal{R}$. Note that $H < -\Gamma^{-1} < 0$ is equivalent to $0 < -H^{-1} < \Gamma$.

The requirement of robust observer is stronger than that of ordinary observer. A robust observer is an ordinary observer. The converse is not true. Suppose that $\tilde{P} > 0$ is a solution to (54). We can decompose the matrix into $\tilde{P} = W^T \Lambda W$ with a lower triangular W and a diagonal matrix Λ. This means that (56) is satisfied by replacing Γ^{-1} with $W^{-T}Q_y W^{-1}$. However, $\Gamma^{-1} \leq W^{-T}Q_y W^{-1}$ is not guaranteed at all. The first two terms on the left hand side of (56) correspond to the Lyapunov derivative of the observer error system. The robust observer requires that the observer error system is stable to a degree prescribed by Γ. That is why the robust observer can be used for making a control system robust against uncertainties. The function Γ is considered as an index of robustness. The smaller $\Gamma > 0$ is, the more robust the resulting observer is. This will be explained later on. Note that for a certain class of systems, it is always possible to construct observer gains required for ordinary observers and robust observers. The observer design will be explained in Section 8.

5 Backstepping design for output feedback

We now direct our attention to a diffeomorphism $S(y, \hat{x})$ in a special form. The diffeomorphism will lead us to a recursive structure with which output feedback backstepping is proposed. We thereby extend the robust backstepping procedure presented in Ito and Freeman (1998*a*) for output feedback design. The backstepping is carried out successfully by selecting SD scaling matrices recursively.

Let $\hat{x}_{[k]}$ denote the state of the observer $\hat{x}_1$ through $\hat{x}_k$:

$$\hat{x}_{[k]} = [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_k]^T . \tag{58}$$

Consider smooth scalar-valued functions $s_1(x_1)$, $s_2(x_1, \hat{x}_1)$, $\cdots$, $s_{n-1}(x_1, \hat{x}_{[n-2]})$ which are to be determined in a recursive manner from s_1 through s_{n-1}. We define a diffeomorphism $S(x_1, \hat{x})$ as follows:

$$S^{-1}(x_1, \hat{x}_{[n-2]}) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ s_1 & 1 & 0 & \cdots & 0 \\ 0 & s_2 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & s_{n-1} & 1 \end{bmatrix} \tag{59}$$

$$S(x_1, \hat{x}_{[n-2]}) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -s_1 & 1 & 0 & \cdots & 0 \\ s_1 s_2 & -s_2 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ (-1)^{n-1} s_1 \cdots s_{n-1} & \cdots & s_{n-2}s_{n-1} & -s_{n-1} & 1 \end{bmatrix} . \tag{60}$$

The smooth function $V(x_1, \hat{x}_{[n-1]})$ and $T(x_1, \hat{x}_{[n-1]})$ in (20) are obtained as

$$V(x_1, \hat{x}_{[n-1]}) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \star_{1,1,1} & 0 & 0 & \cdots & 0 \\ \star_{1,2,2} & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \star_{1,n-1,n-1} & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{61}$$

$$T(x_1,\hat{x}_{[n-1]}) = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \star_{1,0,1} & 1 & 0 & 0 & \cdots & 0 \\ \star_{1,2,2} & \star_{1,1,2} & 1 & 0 & \ddots & 0 \\ \star_{1,3,3} & \star_{1,3,3} & \star_{1,2,3} & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \star_{1,n-1,n-1} & \cdots & \cdots & \star_{1,n-1,n-1} & \star_{1,n-2,n-1} & 1 \end{bmatrix}, \tag{62}$$

where $\star_{1,i,j}$ denotes any function depending only on $(x_1, \hat{x}_{[i]})$ and the functions s_1 through s_j and their partial derivatives. We choose a feedback gain (17) in the form of

$$K = \begin{bmatrix} (-1)^{n-1} s_1 \cdots s_n & \cdots & -s_{n-1}s_n & s_n \end{bmatrix} \tag{63}$$

where $s_n(x_1, \hat{x}_{[n-1]})$ is another smooth function yet to be determined. Then, the matrices in (26) and (28) for the closed-loop system become

$$N := \begin{bmatrix} \hat{S}^T \hat{A}^T (V+T)^T P + P(V+T)\hat{A}\hat{S} & -P(VA+TYC_y)W^{-1} \\ -W^{-T}(VA+TYC_y)^T P & W^{-T}(A-YC_y)^T W^T \tilde{P} + \tilde{P}W(A-YC_y)W^{-1} \end{bmatrix} \tag{64}$$

$$M := \begin{bmatrix} \hat{S}^T \hat{A}^T (V+T)^T P + P(V+T)\hat{A}\hat{S} & PVB + S^{-T}C^T R^T & S^{-T}C^T L & -P(VA+TYC_y)W^{-1} \\ B^T V^T P + RCS^{-1} & -L & 0 & -B^T W^T \tilde{P} - RCW^{-1} \\ LCS^{-1} & 0 & -L & -LCW^{-1} \\ -W^{-T}(VA+TYC_y)^T P & -\tilde{P}WB - W^{-T}C^T R^T & -W^{-T}C^T L & \left\{ \begin{matrix} W^{-T}(A-YC_y)^T W^T \tilde{P} + \\ \tilde{P}W(A-YC_y)W^{-1} \end{matrix} \right\} \end{bmatrix} \tag{65}$$

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

We now restrict P to diagonal:

$$P = \operatorname*{diag}_{i=1}^{n} P_i, \qquad P_i > 0, \quad P_{[k]} = \operatorname*{diag}_{i=1}^{k} P_i \tag{66}$$

We also consider the coordinate transformation W of $\tilde{x}$ in a lower triangular form:

$$W = \begin{bmatrix} W_{11} & 0 & 0 & \cdots & 0 \\ W_{21} & W_{22} & 0 & \cdots & 0 \\ W_{31} & W_{32} & W_{33} & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ W_{n1} & \cdots & W_{n,n-2} & W_{n,n-1} & W_{n,n} \end{bmatrix} \tag{67}$$

Define system matrices for the first k state and input components by

$$\hat{A}_{[k]}(x_1) = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{k-1,1} & a_{k-1,2} & \cdots & \cdots & a_{k-1,k} & 0 \\ a_{k1} & a_{k2} & \cdots & \cdots & a_{kk} & a_{k,k+1} \end{bmatrix} \tag{68}$$

$$B_{[k]}(x_1) = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ B_{21} & B_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ B_{k1} & \cdots & B_{k,k-1} & B_{kk} \end{bmatrix}, \quad C_{[k]}(x_1) = \begin{bmatrix} C_{11} & 0 & \cdots & 0 \\ C_{21} & C_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ C_{k1} & \cdots & C_{k,k-1} & C_{kk} \end{bmatrix} \tag{69}$$

In a similar manner, $S_{[k]}(x_1,\hat{x}_{[k-2]})$, $S_{[k]}^{-1}(x_1,\hat{x}_{[k-2]})$, $V_{[k]}(x_1,\hat{x}_{[k-1]})$, $T_{[k]}(x_1,\hat{x}_{[k-1]})$, $W_{[k]}$ and $W_{[k]}^{-1}$ are defined as $k\times k$ upper left parts of S, S^{-1}, V, T, W and W^{-1}, respectively. Let

$$\hat{S}_{[k]}(x_1,\hat{x}_{[k-1]}) = \left[\begin{array}{c|c} \multicolumn{2}{c}{S_{[k]}^{-1}} \\ \hline 0\cdots 0 & s_k \end{array}\right], \quad \tilde{P}_{[k]} = \left[\begin{array}{c|c} \tilde{P}_{[k]} & \star_{0,0,0} \\ \hline \star_{0,0,0} & \tilde{P}_{kk} \end{array}\right] \tag{70}$$

$$A_{[k]} = \left[\begin{array}{c|c} & 0 \\ A_{[k-1]} & \vdots \\ & 0 \\ & a_{k-1,k} \\ \hline a_{k,*} & a_{kk} \end{array}\right], \quad Y_{[k]} = \left[\begin{array}{c} Y_{[k-1]} \\ \hline Y_k \end{array}\right] \tag{71}$$

$$C_{y[k]} = [\,1\;\,0\;\cdots\;0\,|\,0\,] = [\,C_{y[k-1]}\,|\,0\,], \quad C_{y[1]} = 1 \tag{72}$$

Scaling matrices are also defined recursively as

$$\mathbf{L}_{[k]} := \left\{ L_{[k]} = \operatorname*{block\text{-}diag}_{i=1}^{k} L_i(x_1,\hat{x}_{[i-2]}), \quad L_i \in \mathbf{L}_i \right\} \tag{73}$$

$$\mathbf{R}_{[k]} := \left\{ R_{[k]} = \operatorname*{block\text{-}diag}_{i=1}^{k} R_i(x_1,\hat{x}_{[i-2]}), \quad R_i \in \mathbf{R}_i \right\}\ . \tag{74}$$

Now, we define $N_{[k]}(x_1,\hat{x}_{[k-1]})$ by adding subscript $[k]$ to every matrix in the right hand side of (64). By using

$$[H]_{[k]}(x_1) = \left[\begin{array}{c|c} [H]_{[k-1]} & \star_{1,0,0} \\ \hline \star_{1,0,0} & [H]_{kk} \end{array}\right], \quad [H]_{[n]} = H \tag{75}$$

the matrix $N_{[k]}$ can be represented by

$$N_{[k]}(x_1,\hat{x}_{[k-1]}) = \begin{bmatrix} N_{[k]11}(x_1,\hat{x}_{[k-1]}) & N_{[k]12}(x_1,\hat{x}_{[k-1]}) \\ N_{[k]12}^T(x_1,\hat{x}_{[k-1]}) & H_{[k]}(x_1) \end{bmatrix}, \quad N_{[n]} = N \tag{76}$$

$$N_{[k]11} := \hat{S}_{[k]}^T\hat{A}_{[k]}^T(V_{[k]}+T_{[k]})^T P_{[k]} + P_{[k]}(V_{[k]}+T_{[k]})\hat{A}_{[k]}\hat{S}_{[k]},\ N_{[k]12} := P_{[k]}(V_{[k]}A_{[k]}+T_{[k]}Y_{[k]}C_{y[k]})W_{[k]}^{-1} \tag{77}$$

We also define $\tilde{M}_{[k]}$ as

$$\tilde{M}_{[k]}(x_1,\hat{x}_{[k-1]}) = \begin{bmatrix} M_{[k]11}(x_1,\hat{x}_{[k-1]}) & Q_k^T M_{12}(x_1,\hat{x}_{[n-1]}) \\ M_{12}^T(x_1,\hat{x}_{[n-1]})Q_k & H(x_1) \end{bmatrix}, \quad \tilde{M}_{[n]} = M \tag{78}$$

$$M_{[k]11} := \begin{bmatrix} \hat{S}_{[k]}^T\hat{A}_{[k]}^T(V_{[k]}+T_{[k]})^T P_{[k]} + P_{[k]}(V_{[k]}+T_{[k]})\hat{A}_{[k]}\hat{S}_{[k]} & P_{[k]}V_{[k]}B_{[k]} + S_{[k]}^{-T}C_{[k]}^T R_{[k]}^T & S_{[k]}^{-T}C_{[k]}^T L_{[k]} \\ B_{[k]}^T V_{[k]}^T P_{[k]} + R_{[k]}C_{[k]}S_{[k]}^{-1} & -L_{[k]} & 0 \\ L_{[k]}C_{[k]}S_{[k]}^{-1} & 0 & -L_{[k]} \end{bmatrix} \tag{79}$$

$$M_{12} := \begin{bmatrix} P(VA+TYC_y)W^{-1} \\ B^T W^T \tilde{P} + RCW^{-1} \\ LCW^{-1} \end{bmatrix}, \quad \bar{Q}_k = \begin{bmatrix} I_k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & I_{\bar{q}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{\bar{q}} \end{bmatrix}, \quad \bar{Q}_n = I_{n+2P} \tag{80}$$

where I_k denotes a $k\times k$ identity matrix and $\bar{q} := \sum_{i=1}^k p_i$. Note that $M_{[k]11} = \bar{Q}_k^T M_{11}\bar{Q}_k$ holds. We can verify the following.

Theorem 2 *Suppose* $1\le k\le n$.

(i-a) $N_{[k]}$ *does not include* $\{s_{k+1}, s_{k+2},\cdots,s_n\}$.

(i-b) *Every entry of* $N_{[k]}$ *is affine in* s_k.

(i-c) *Every entry of* $N_{[k]}$ *is simultaneously affine in all the entries of* $P_{[k]}$.

(i-d) $N_{[k]} < 0$ *implies* $N_{[k-1]} < 0$ *unless* $k = 1$.

(ii-a) $\tilde{M}_{[k]}$ *does not include either* $\{s_{k+1}, s_{k+2}, \cdots, s_n\}$, $\{L_{k+1}, L_{k+2}, \cdots, L_n\}$ *or* $\{R_{k+1}, R_{k+2}, \cdots, R_n\}$.

(ii-b) *Every entry of* $\tilde{M}_{[k]}$ *is simultaneously affine in* L_k, R_k *and* s_k.

(ii-c) *Every entry of* $\tilde{M}_{[k]}$ *is simultaneously affine in all the entries of* $L_{[k]}$, $R_{[k]}$ *and* $P_{[k]}$.

(ii-d) $\tilde{M}_{[k]} < 0$ *implies* $\tilde{M}_{[k-1]} < 0$ *unless* $k = 1$.

Although the system is nonlinear in state variables, the above theorem shows that the problem of SD scaling is recursively linear in decision variables(or design parameters). This suggests that the nonlinearity of the system essentially do not make the problem seriously difficult. The character of the problem still remains the same as that of robust linear design in this sense.

On the basis of Theorem 2, this paper proposes the following procedures of backstepping for feedback gain design.

Nominal backstepping : Solve

$$N_{[k]}(x_1, \hat{x}_{[k-1]}) < 0, \quad \forall (x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1} \tag{81}$$

for s_k from $k = 1$ through $k = n$.

Robust backstepping : Solve

$$\tilde{M}_{[k]}(x_1, \hat{x}_{[k-1]}) < 0, \quad \forall (x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1} \tag{82}$$

for $\{s_k, L_k, R_k\}$ from $k = 1$ through $k = n$.

Both the procedures suppose that P, $\tilde{P}$ and Y are given. The above procedures can be carried out recursively since the process of finding decision parameters at Step k does not require any decision parameters to be found at Step $k+1, k+2, \ldots, n$. The recursive procedures can be also justified in that Step k is a necessary step for accomplishing Step $k+1, k+2, \ldots, n$.

Theorem 3 *(i) If the whole procedure of nominal backstepping is completed from* $k = 1$ *through* $k = n$ *properly, the parameters* $\{s_1, s_2, \ldots, s_n\}$ *solve*

$$N(x_1, \hat{x}_{[n-1]}) < 0, \quad \forall (x_1, \hat{x}_{[n-1]}) \in \mathcal{R} \times \mathcal{R}^{n-1} \tag{83}$$

(i) If the whole procedure of robust backstepping is completed from $k = 1$ *through* $k = n$ *properly, the parameters* $\{s_1, s_2, \ldots, s_n\}$, $\{L_1, L_2, \ldots, L_n\}$ *and* $\{R_1, R_2, \ldots, R_n\}$ *solve*

$$M(x_1, \hat{x}_{[n-1]}) < 0, \quad \forall (x_1, \hat{x}_{[n-1]}) \in \mathcal{R} \times \mathcal{R}^{n-1} \tag{84}$$

For instance, the problem of finding $\{L_k, R_k, s_k\}$ satisfying $\tilde{M}_{[k]} < 0$ is a convex optimization problem. The nominal backstepping and the robust backstepping for output feedback design via SD scaling are amenable to computation based on optimization as it has been shown for state-feedback design in Ito and Freeman (1998*a*). It is ready for automated *numerical* calculation by computer. The recursive design proposed in this section does not require precise knowledge of each system parameter since the design is based on domination instead of cancelation. An exactly canceling formula is considered as one special solution to the domination. Moreover, the domination approach can be exploited to get rid of the propagation of complicated and long terms in the control law K.

The subsequent sections investigate whether the solutions exist or not in the recursive procedures. In other words, a condition of the allowable size and nonlinearity of uncertainty will be derived. Existence conditions and analytical solutions will be developed. Furthermore, a class of systems which can be always robustly stabilizable against arbitrarily large uncertainties by output feedback via the robust backstepping will be shown.

6 Equivalent recursive procedures

This section transforms the nominal backstepping and robust backstepping into problems which are suitable for finding analytical solutions. No conservatism will be introduced in this section. Solving a transformed problem is equivalent to performing the backstepping in Section 5.

Define the following two functions.

$$\bar{N}_{[k]}(x_1, \hat{x}_{[k-1]}) := N_{[k]11}(x_1, \hat{x}_{[k-1]}) - N_{[k]12}(x_1, \hat{x}_{[k-1]})H_{[k]}^{-1}(x_1)N_{[k]12}^T(x_1, \hat{x}_{[k-1]}) \tag{85}$$

$$\bar{M}_{[k]}(x_1, \hat{x}_{[k-1]}) := M_{[k]11}(x_1, \hat{x}_{[k-1]}) - \bar{Q}_k^T M_{12}(x_1, \hat{x}_{[n-1]})H^{-1}M_{12}^T(x_1, \hat{x}_{[n-1]})\bar{Q}_k \tag{86}$$

From the Schur complements, the equivalence

$$\bar{N}_{[k]} < 0 \Leftrightarrow N_{[k]} < 0 \tag{87}$$

$$\bar{M}_{[k]} < 0 \Leftrightarrow \tilde{M}_{[k]} < 0 \tag{88}$$

are obviously true on the assumption that $H < 0$ holds. Due to structures of S, W and the strict-feedback form of Σ_0 , we can prove the following for $\bar{N}_{[k]}$ and $\bar{M}_{[k]}$.

Theorem 4 *Suppose* $2 \le k \le n$.
(i) The symmetric matrix

$$\bar{N}_{[1]}(x_1) = \tilde{\Psi}_1(x_1) \tag{89}$$

depends only on s_1. $\bar{N}_{[k]}(x_1, \hat{x}_{[k-1]}) < 0$ *is equivalent to*

$$\begin{bmatrix} \bar{N}_{[k-1]}(x_1, \hat{x}_{[k-2]}) & \tilde{\Phi}_k(x_1, \hat{x}_{[k-1]}) \\ \tilde{\Phi}_k^T(x_1, \hat{x}_{[k-1]}) & \tilde{\Psi}_k(x_1, \hat{x}_{[k-1]}) \end{bmatrix} < 0 \ , \tag{90}$$

where $\tilde{\Phi}_k$ *depends only on* $(s_1, \cdots, s_{k-1})$ *and their partial derivatives. The symmetric matrix* $\tilde{\Psi}_k$ *depends on* s_k.
(ii) Assume that $\tilde{P}$ *is diagonal:*

$$\tilde{P} = \operatorname*{diag}_{i=1}^{n} \tilde{P}_i, \qquad \tilde{P}_i > 0, \quad \tilde{P}_{[k]} = \operatorname*{diag}_{i=1}^{k} \tilde{P}_i \tag{91}$$

Then, the symmetric matrix

$$\bar{M}_{[1]}(x_1) = \Psi_1(x_1) \tag{92}$$

depends only on (L_1, R_1) *and* s_1. $\bar{M}_{[k]}(x_1, \hat{x}_{[k-1]}) < 0$ *is equivalent to*

$$\begin{bmatrix} \bar{M}_{[k-1]}(x_1, \hat{x}_{[k-2]}) & \Phi_k(x_1, \hat{x}_{[k-1]}) \\ \Phi_k^T(x_1, \hat{x}_{[k-1]}) & \Psi_k(x_1, \hat{x}_{[k-1]}) \end{bmatrix} < 0 \ , \tag{93}$$

where Φ_k *depends only on* $(L_{[k]}, R_{[k]})$ *and* $(s_1, \cdots, s_{k-1})$ *and their partial derivatives. The symmetric matrix* Ψ_k *depends on* (L_k, R_k) *and* s_k.
Proof : Recall that

$$\hat{S}_{[k]}(x_1, \hat{x}_{[k-1]}) = \left[\begin{array}{c|c} \hat{S}_{[k-1]}(x_1, \hat{x}_{[k-2]}) & \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \end{matrix} \\ \hline 0 & s_k \end{array}\right], \quad S_{[k]}^{-1}(x_1, \hat{x}_{[k-2]}) = \left[\begin{array}{c|c} S_{[k-1]}^{-1}(x_1, \hat{x}_{[k-3]}) & 0 \\ \hline \star_{1,k-2,k-1} & 1 \end{array}\right]$$

$$\hat{A}_{[k]}(x_1) = \left[\begin{array}{c|c} \hat{A}_{[k-1]}(x_1) & 0 \\ \hline \star_{1,0,0} & a_{k,k+1} \end{array}\right], \quad B_{[k]}(x_1) = \left[\begin{array}{c|c} B_{[k-1]}(x_1) & 0 \\ \hline \star_{1,0,0} & B_{kk} \end{array}\right]$$

$$C_{[k]}(x_1) = \left[\begin{array}{c|c} C_{[k-1]}(x_1) & 0 \\ \hline \star_{1,0,0} & C_{kk} \end{array}\right] = \left[\begin{array}{c|c} C_{[k-1]}(x_1) & 0 \\ \hline C_{k,*} & C_{kk} \end{array}\right] = \left[\begin{array}{c|c} C_{[k-1]}(x_1) & 0 \\ \hline \multicolumn{2}{c}{C_{k,-}} \end{array}\right]$$

$$V_{[k]}(x_1, \hat{x}_{[k-1]}) = \left[\begin{array}{c|c} V_{[k-1]}(x_1, \hat{x}_{[k-2]}) & 0 \\ \hline \star_{1,k-1,k-1} & 0 \end{array}\right], \quad T_{[k]}(x_1, \hat{x}_{[k-1]}) = \left[\begin{array}{c|c} T_{[k-1]}(x_1, \hat{x}_{[k-2]}) & 0 \\ \hline \star_{1,k-1,k-1} & 1 \end{array}\right]$$

$$L_{[k]}(x_1, \hat{x}_{[k-2]}) = \left[\begin{array}{c|c} L_{[k-1]}(x_1, \hat{x}_{[k-3]}) & 0 \\ \hline 0 & L_k \end{array}\right], \quad R_{[k]}(x_1, \hat{x}_{[k-2]}) = \left[\begin{array}{c|c} R_{[k-1]}(x_1, \hat{x}_{[k-3]}) & 0 \\ \hline 0 & R_k \end{array}\right]$$

Then, we have the following.

$$P_{[k]}(V_{[k]} + T_{[k]})\hat{A}_{[k]}\hat{S}_{[k]} =$$
$$\left[\begin{array}{c|c} \left\{\begin{array}{r} P_{[k-1]}(V_{[k-1]}(x_1, \hat{x}_{[k-2]}) + T_{[k-1]}(x_1, \hat{x}_{[k-2]})) \\ \hat{A}_{[k-1]}(x_1)\hat{S}_{[k-1]}(x_1, \hat{x}_{[k-2]}) \end{array}\right\} & \begin{array}{c} 0 \\ P_{k-1}a_{k-1,k} \end{array} \\ \hline P_k \star_{1,k-1,k-1} & P_k(a_{kk} + a_{k,k+1}s_k + \star_{1,k-1,k-1}) \end{array}\right]$$

$$P_{[k]}V_{[k]}B_{[k]} = \left[\begin{array}{c|c} P_{[k-1]}V_{[k-1]}(x_1, \hat{x}_{[k-2]})B_{[k-1]}(x_1) & 0 \\ \hline \star_{1,k-1,k-1} & 0 \end{array}\right]$$

$$P_{[k]}V_{[k]}A_{[k]}W_{[k]}^{-1} = \left[\begin{array}{c|c} P_{[k-1]}V_{[k-1]}(x_1, \hat{x}_{[k-2]})A_{[k-1]}(x_1) & 0 \\ \hline P_k \star_{1,k-1,k-1} & P_k \star_{1,k-1,k-1} \end{array}\right]$$

$$P_{[k]}T_{[k]}Y_{[k]}C_{y,[k]}W_{[k]}^{-1} = \left[\begin{array}{c|c} P_{[k-1]}T_{[k-1]}(x_1, \hat{x}_{[k-2]})Y_{[k-1]}(x_1)W_{kk}^{-1} \;\; 0 \cdots 0 & 0 \\ \hline P_k \star_{1,k-1,k-1} \qquad\qquad 0 \cdots 0 & 0 \end{array}\right] .$$

Thus, we obtain

$$N_{[k]12}H_{[k]}^{-1}N_{[k]12}^T = \begin{bmatrix} N_{[k-1]12}H_{[k-1]}^{-1}N_{[k-1]12}^T & \star_{1,k-1,k-1}P_k \\ P_k \star_{1,k-1,k-1} & P_k^2 \star_{1,k-1,k-1} \end{bmatrix} \tag{94}$$

$$\bar{N}_{[k]} = \begin{bmatrix} \bar{N}_{[k-1]} & \star_{1,k-1,k-1} \\ \star_{1,k-1,k-1} & 2P_k(a_{kk} + a_{k,k+1}s_k + \star_{1,k-1,k-1}) \end{bmatrix} = \begin{bmatrix} \bar{N}_{[k-1]} & \tilde{\Phi}_k \\ \tilde{\Phi}_k^T & \tilde{\Psi}_k \end{bmatrix} \tag{95}$$

To prove the claim for $\bar{M}_{[k]}$, we need

$$\tilde{P}_{[k]}W_{[k]}B_{[k]} = \left[\begin{array}{c|c} \tilde{P}_{[k-1]}W_{[k-1]}B_{[k-1]}(x_1) & 0 \\ \hline \star_{1,0,0} & \tilde{P}_{kk}W_{kk}B_{kk} \end{array}\right] \tag{96}$$

$$L_{[k]}C_{[k]}S_{[k]}^{-1} = \left[\begin{array}{c|c} L_{[k-1]}(x_1, \hat{x}_{[k-3]})C_{[k-1]}(x_1)S_{[k-1]}^{-1}(x_1, \hat{x}_{[k-3]}) & 0 \\ \hline L_k(C_{k,*}\star_{1,k-3,k-2} + C_{k,k}\star_{1,k-2,k-1}) & L_kC_{kk} \end{array}\right] \tag{97}$$

$$R_{[k]}C_{[k]}S_{[k]}^{-1} = \left[\begin{array}{c|c} R_{[k-1]}(x_1, \hat{x}_{[k-3]})C_{[k-1]}(x_1)S_{[k-1]}^{-1}(x_1, \hat{x}_{[k-3]}) & 0 \\ \hline R_k(C_{k,*}\star_{1,k-3,k-2} + C_{k,k}\star_{1,k-2,k-1}) & R_kC_{kk} \end{array}\right] \tag{98}$$

$$L_{[k]}C_{[k]}W_{[k]}^{-1} = \left[\begin{array}{c|c} L_{[k-1]}(x_1, \hat{x}_{[k-3]})C_{[k-1]}(x_1)W_{[k-1]}^{-1} & 0 \\ \hline L_kC_{k,-}\star_{0,0,0} & L_kC_{kk}W_{kk}^{-1} \end{array}\right] \tag{99}$$

$$R_{[k]}C_{[k]}W_{[k]}^{-1} = \left[\begin{array}{c|c} R_{[k-1]}(x_1, \hat{x}_{[k-3]})C_{[k-1]}(x_1)W_{[k-1]}^{-1} & 0 \\ \hline R_kC_{k,-}\star_{0,0,0} & R_kC_{kk}W_{kk}^{-1} \end{array}\right] \tag{100}$$

Now, let U denote

$$U(x_1, \hat{x}_{[n-2]}) = -(B^TW^T\tilde{P} + RCW^{-1})H^{-1}(\tilde{P}WB + W^{-T}C^TR^T) \tag{101}$$

The following recursive notation is used.

$$U_{[k]}(x_1, \hat{x}_{[k-2]}) = \begin{bmatrix} U_{[k-1]}(x_1, \hat{x}_{[k-3]}) & U_{*,k}(x_1, \hat{x}_{[k-2]}) \\ U_{*,k}^T(x_1, \hat{x}_{[k-2]}) & U_{kk}(x_1, \hat{x}_{[k-2]}) \end{bmatrix}, \quad U_{kk}(x_1, \hat{x}_{[k-2]}) \in \mathcal{R}^{p_k \times p_k} \tag{102}$$

The definition of these matrices are given by

$$B^TW^T\tilde{P}+RCW^{-1}=\begin{bmatrix}[B^TW^T\tilde{P}+RCW^{-1}]_1\\ [B^TW^T\tilde{P}+RCW^{-1}]_2\\ \cdots\\ [B^TW^T\tilde{P}+RCW^{-1}]_n\end{bmatrix},[B^TW^T\tilde{P}+RCW^{-1}]_{[k]}=\begin{bmatrix}[B^TW^T\tilde{P}+RCW^{-1}]_1\\ [B^TW^T\tilde{P}+RCW^{-1}]_2\\ \cdots\\ [B^TW^T\tilde{P}+RCW^{-1}]_k\end{bmatrix}$$

$$U_{kk}=-[B^TW^T\tilde{P}+RCW^{-1}]_kH^{-1}[B^TW^T\tilde{P}+RCW^{-1}]_k^T$$
$$U_{*,k}=-[B^TW^T\tilde{P}+RCW^{-1}]_{[k-1]}H^{-1}[B^TW^T\tilde{P}+RCW^{-1}]_k^T$$

Let $[H^{-1}]_{[k]}$ be defined as

$$[H^{-1}]_{[k]}(x_1)=\left[\begin{array}{c|c}[H^{-1}]_{[k-1]} & \star_{1,0,0}\\ \hline \star_{1,0,0} & [H^{-1}]_{kk}\end{array}\right],\quad [H^{-1}]_{[n]}=H^{-1} \tag{103}$$

Consider a non-singular matrix.

$$Q_k=\begin{bmatrix}I_{k-1}&0&0&0&0&0\\0&0&0&I_1&0&0\\0&I_q&0&0&0&0\\0&0&0&0&I_{p_k}&0\\0&0&I_q&0&0&0\\0&0&0&0&0&I_{p_k}\end{bmatrix},$$

where $q:=\sum_{i=1}^{k-1}p_i$. Then, we have

$$Z_{[k]}(x_1,\hat{x}_{[n-1]})=-Q_k^T\bar{Q}_k^TM_{12}(x_1,\hat{x}_{[n-1]})H^{-1}(x_1)M_{12}^T(x_1,\hat{x}_{[n-1]})\bar{Q}_kQ_k \tag{104}$$

$$=\left[\begin{array}{ccc|ccc} & Z_{[k-1]} & & \star_{1,k-1,k-1} & \star_{1,k-2,k-2}R_k^T & \star_{1,k-2,k-2}C_{k,-}^TL_k\\ & & & \Diamond_{1,k-1,k-1} & U_{*,k} & \Diamond_{1,k-3,k-1}C_{k,-}^TL_k\\ & & & \Diamond_{1,k-1,k-1} & \Diamond_{1,k-3,k-1}R_k^T & \Diamond_{1,k-3,k-1}C_{k,-}^TL_k\\ \hline * & * & * & P_k^2\star_{1,k-1,k-1} & \star_{1,k-1,k-1}+\star_{1,k-1,k-1}R_k^T & \star_{1,k-1,k-1}C_{k,-}^TL_k\\ * & * & * & * & U_{kk} & (R_k\star_{1,0,0}+\star_{1,0,0})C_{k,-}^TL_k\\ * & * & * & * & * & -L_kC_{k,-}W_{[k]}^{-1}[H^{-1}]_{[k]}W_{[k]}^{-T}C_{k,-}^TL_k\end{array}\right] \tag{105}$$

$$Q_k^TM_{[k]11}Q_k=\left[\begin{array}{ccc|ccc} & M_{[k-1]11} & & \star_{1,k-1,k-1} & \star_{1,k-2,k-1}R_k^T & \star_{1,k-2,k-1}C_{k,-}^TL_k\\ & & & \star_{1,k-1,k-1} & 0 & 0\\ & & & 0 & 0 & 0\\ \hline * & * & * & 2P_k(a_{kk}+a_{k,k+1}s_k+\star_{1,k-1,k-1}) & C_{kk}^TR_k^T & C_{kk}^TL_k\\ * & * & * & * & -L_k & 0\\ * & * & * & * & * & -L_k\end{array}\right] \tag{106}$$

where $\Diamond_{1,i,j}$ denotes any function depending only on $(x_1,\hat{x}_{[i]})$, $(L_{[j]},R_{[j]})$ and $(s_1,\cdots,s_j)$ and their partial derivatives. By using $Q_k^T\bar{M}_{[k]}Q_k=Q_k^TM_{[k]11}Q_k+Z_{[k]}$, we arrive at

$$Q_k^T\bar{M}_{[k]}(x_1,\hat{x}_{[k-1]})Q_k=\begin{bmatrix}\bar{M}_{[k-1]}(x_1,\hat{x}_{[k-2]}) & \Phi_k(x_1,\hat{x}_{[k-1]})\\ \Phi_k^T(x_1,\hat{x}_{[k-1]}) & \Psi_k(x_1,\hat{x}_{[k-1]})\end{bmatrix}, \tag{107}$$

where the functions Φ_k and Φ_k are obtained as

$$\Phi_k(x_1,\hat{x}_{[k-1]})=\begin{bmatrix}\star_{1,k-1,k-1} & \star_{1,k-2,k-1}R_k^T & \star_{1,k-2,k-1}C_{k,-}^TL_k\\ \Diamond_{1,k-1,k-1} & U_{*,k} & \Diamond_{1,k-3,k-1}C_{k,-}^TL_k\\ \Diamond_{1,k-1,k-1} & \Diamond_{1,k-3,k-1}R_k^T & \Diamond_{1,k-3,k-1}C_{k,-}^TL_k\end{bmatrix} \tag{108}$$

$$\Psi_k(x_1,\hat{x}_{[k-1]})=$$
$$\begin{bmatrix}2P_k(a_{kk}+a_{k,k+1}s_k+\star_{1,k-1,k-1}) & \star_{1,k-1,k-1}+(\star_{1,k-1,k-1}+C_{kk}^T)R_k^T & (\star_{1,k-1,k-1}+C_{kk}^T)L_k\\ * & -L_k+U_{kk} & (R_k\star_{1,0,0}+\star_{1,0,0})C_{k,-}^TL_k\\ * & * & -L_k-L_kC_{k,-}W_{[k]}^{-1}[H^{-1}]_{[k]}W_{[k]}^{-T}C_{k,-}^TL_k\end{bmatrix} \tag{109}$$

Hence, the claim is established. ∎

The task of the SD scaling backstepping is to show how to achieve $\bar{M}_{[k-1]} < 0$ ($\bar{N}_{[k-1]} < 0$) by choosing s_k and scaling functions L_k and R_k if $\bar{M}_{[k-1]} < 0$ ($\bar{N}_{[k-1]} < 0$, respectively) is guaranteed at the previous step. The converse directions follows from the above theorem.

Now, the negativity problems in Theorem 4 are equivalently transformed into problems in smaller size of matrices again. Let $\tilde{J}_k(x_1, \hat{x}_{[k-1]}) \in \mathcal{R}^{1\times 1}$ be defined with

$$\begin{array}{ll} \tilde{\Psi}_k - \tilde{\Phi}_k^T \bar{N}_{[k-1]}^{-1} \tilde{\Phi}_k = \tilde{J}_k & \text{for } k \geq 2 \\ \tilde{\Psi}_1 = \tilde{J}_1 & \text{for } k = 1 \end{array} \tag{110}$$

We also define $J_k(x_1, \hat{x}_{[k-1]}) \in \mathcal{R}^{1\times 1}$, $E_k(x_1, \hat{x}_{[k-1]}) \in \mathcal{R}^{1\times 2p_k}$ and $F_k(x_1, \hat{x}_{[k-1]}) \in \mathcal{R}^{2p_k\times 2p_k}$ as

$$\begin{array}{ll} \Psi_k - \Phi_k^T \bar{M}_{[k-1]}^{-1} \Phi_k = \begin{bmatrix} J_k & E_k \\ E_k^T & F_k \end{bmatrix} & \text{for } k \geq 2 \\ \Psi_1 = \begin{bmatrix} J_1 & E_1 \\ E_1^T & F_1 \end{bmatrix} & \text{for } k = 1 \end{array} \tag{111}$$

Using the Schur complements of (110) and (111) , we have the following.

Corollary 2 *Let k is any integer belonging to $[1, n]$.*
(i) Assume that $\bar{N}_{[k-1]}(x_1, \hat{x}_{[k-2]}) < 0$ is satisfied for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$ unless $k = 1$. Then, $\bar{N}_{[k]}(x_1, \hat{x}_{[k-1]}) < 0$ holds for all $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1}$ if and only if

$$\tilde{J}_k < 0 \tag{112}$$

is satisfied for all $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1}$.
(ii) Assume that $\bar{M}_{[k-1]}(x_1, \hat{x}_{[k-2]}) < 0$ is satisfied for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$ unless $k = 1$. Then, $\bar{M}_{[k]}(x_1, \hat{x}_{[k-1]}) < 0$ holds for all $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1}$ if and only if

$$F_k < 0, \;\; J_k - E_k F_k^{-1} E_k^T < 0, \quad \textit{when } p_k \neq 0 \tag{113}$$

$$J_k < 0, \qquad \textit{when } p_k = 0 \tag{114}$$

are satisfied for all $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1}$.

7 Existence and analytical solution

This section derives the condition of allowable uncertainty under which the robust stabilization problem can be always solved by robust backstepping proposed in Section 5. As for nominal stabilization, It will be shown that the stabilization problem is always solvable on the assumption that nominal observers exist. The section also provides us with analytical versions of nominal and robust backstepping and their analytical solutions. The backsteping procedure based on materials in this section can be done by numerical calculation again. In the robust stabilization case, the problem is no longer affine in decision parameters. Although it may be difficult to solve it as an optimization problem, the backstepping can be easily performed only by curve fitting of a real-valued function whose range is specified by intervals. Analytical solutions of such functions are also available.

First, nominal stabilization is briefly explained. From (95), the function $\tilde{J}_k$ is given by

$$\tilde{J}_k = 2P_k(a_{kk} + a_{k,k+1}s_k + \star_{1,k-1,k-1}) \tag{115}$$

for $k = 1, 2, \ldots, n$. We can prove the following.

Theorem 5 *Let k is any integer belonging to $[1,n]$. Let $\gamma_k(x_1)$ be any $\mathbf{C}^0$ function. There always exist a scalar-valued smooth function $s_k(x_1,\hat{x}_{[k-1]})$ such that $\tilde{J}_k(x_1,\hat{x}_{[k-1]})<0$ is satisfied for all $(x_1,\hat{x}_{[k-1]})\in\mathcal{R}\times\mathcal{R}^{k-1}$.*

Proof : Remember that $a_{k,k+1}(x_1)$ is non-zero for all $x_1\in\mathcal{R}$. Since $a_{k,k+1}$ and other functions in (115) are $\mathbf{C}^0$ functions defined on $\mathcal{R}\times\mathcal{R}^{k-1}$, there exist a smooth function $s_k(x_1,\hat{x}_{[k-1]})$ such that

$$2P_k(a_{kk}+a_{k,k+1}s_k+\star_{1,k-1,k-1})<0 \tag{116}$$

is satisfied for all $(x_1,\hat{x}_{[k-1]})\in\mathcal{R}\times\mathcal{R}^{k-1}$. ∎

This theorem leads us to the following statement for nominal stabilization.

Theorem 6 *Given an ordinary observer, the nominal nonlinear system Σ_0 can be always globally uniformly asymptotically stabilized by the dynamic output-feedback law (16-17) with a smooth function K.*

Proof : Let $W=I$. Theorem 5 guarantees that $\tilde{J}_1<0$ can be achieved, which turns out to be $\bar{N}_{[1]}(x_1)<0$ for all $x_1\in\mathcal{R}$ by Corollary 2. Suppose that $\bar{N}_{[k-1]}(x_1,\hat{x}_{[k-2]})<0$ is satisfied for all $(x_1,\hat{x}_{[k-2]})\in\mathcal{R}\times\mathcal{R}^{k-2}$. Then, we can obtain $\tilde{J}_k<0$ again by using Theorem 5 and Corollary 2 implies $\bar{N}_{[k]}(x_1,\hat{x}_{[k-1]})<0$ for all $(x_1,\hat{x}_{[k-1]})\in\mathcal{R}\times\mathcal{R}^{k-1}$. Finally, Corollary 1 proves the claim. ∎

We now move onto the robust stabilization problem. From (108) and (109), the matrices J_k is given as follows:

$$J_1(x_1)=2P_1(a_{11}+a_{12}s_1+\star_{1,0,0})\ \text{ for } k=1 \tag{117}$$

$$J_k(x_1,\hat{x}_{[k-1]})=2P_k(a_{kk}+a_{k,k+1}s_k+\star_{1,k-1,k-1})+\diamondsuit_{1,k-1,k-1}\ \text{ for } k\geq 2 \tag{118}$$

The following can be proved.

Theorem 7 *Let k is any integer belonging to $[1,n]$. Suppose that $\mathbf{C}^0$ function matrices $L_k(x_1,\hat{x}_{[k-2]})$ and $R_k(x_1,\hat{x}_{[k-2]})$ belong to $\mathbf{L}_k$ and $\mathbf{R}_k$, respectively. Then, there always exist a scalar-valued smooth function $s_k(x_1,\hat{x}_{[k-1]})$ such that $J_k-E_kF_k^{-1}E_k^T<0$ is satisfied for all $(x_1,\hat{x}_{[k-1]})\in\mathcal{R}\times\mathcal{R}^{k-1}$.*

Next, the existence of L_k and R_k is investigated. Due to (108) and (109), the matrices E_k and F_k defined with $R_k=0$ are

$$E_1(x_1)=\begin{bmatrix}\star_{1,0,0} & (C_{11}^T+\star_{1,0,0})L_1\end{bmatrix},\ F_1(x_1)=\begin{bmatrix}-L_1+U_{11} & \star_{1,0,0}C_{11}^TL_1\\ L_1C_{11}\star_{1,0,0} & -L_1-L_1C_{11}W_{11}^{-1}[H^{-1}]_{11}W_{11}^{-1}C_{11}^TL_1\end{bmatrix} \tag{119}$$

$$E_k(x_1,\hat{x}_{[k-1]})=\begin{bmatrix}\diamondsuit_{1,k-1,k-1} & (C_{kk}^T+\diamondsuit_{1,k-1,k-1})L_k\end{bmatrix}\ \text{ for } k\geq 2 \tag{120}$$

$$F_k(x_1,\hat{x}_{[k-2]})=\qquad\text{for } k\geq 2$$

$$\begin{bmatrix}-L_k-B_{-,k}^TW_{\langle k\rangle}^T\tilde{P}_{\langle k\rangle}([H^{-1}]_{\langle k\rangle}+\bar{F}_{k11})\tilde{P}_{\langle k\rangle}W_{\langle k\rangle}B_{-,k} & (\star_{1,0,0}-B_k^TW^T\tilde{P}\bar{F}_{k12})C_{k,-}^TL_k\\ L_kC_{k,-}(\star_{1,0,0}-\bar{F}_{k12}^T\tilde{P}WB_k) & -L_k-L_kC_{k,-}(W_{[k]}^{-1}[H^{-1}]_{[k]}W_{[k]}^{-T}+\bar{F}_{k22})C_{k,-}^TL_k\end{bmatrix} \tag{121}$$

$$\bar{F}_k(x_1,\hat{x}_{[k-2]})=\begin{bmatrix}0 & \star_{1,k-2,k-1}\\ \bar{U}_k & \diamondsuit_{1,k-3,k-1}\\ 0 & \diamondsuit_{1,k-3,k-1}\end{bmatrix}^T\bar{M}_{[k-1]}^{-1}\begin{bmatrix}0 & \star_{1,k-2,k-1}\\ \bar{U}_k & \diamondsuit_{1,k-3,k-1}\\ 0 & \diamondsuit_{1,k-3,k-1}\end{bmatrix} \tag{122}$$

$$\bar{F}_k(x_1,\hat{x}_{[k-2]})=\begin{bmatrix}\bar{F}_{k11}(x_1,\hat{x}_{[k-2]}) & \bar{F}_{k12}(x_1,\hat{x}_{[k-2]})\\ \bar{F}_{k12}^T(x_1,\hat{x}_{[k-2]}) & \bar{F}_{k22}(x_1,\hat{x}_{[k-2]})\end{bmatrix},\quad\text{for } k\geq 2 \tag{123}$$

$$\bar{U}_k(x_1,\hat{x}_{[k-3]})=-[B^TW^T\tilde{P}+RCW^{-1}]_{[k-1]}H^{-1}\begin{bmatrix}0\\ I_{n-k+1}\end{bmatrix} \tag{124}$$

Here, the following expressions are used.

$$B = \left[\begin{array}{c|c|c|c} & 0 & \cdots & 0 \\ \cline{2-2} B_{-,1} & & \ddots & \vdots \\ & B_{-,2} & & 0 \\ \cline{3-4} & & \cdots & B_{-,n} \end{array}\right], \quad B_{-,n} = B_{nn} \tag{125}$$

$$[H^{-1}]_{\langle k\rangle} = \left[\begin{array}{c|c} [H^{-1}]_{kk} & \star_{1,0,0} \\ \hline \star_{1,0,0} & [H^{-1}]_{\langle k-1\rangle} \end{array}\right], \quad [H^{-1}]_{\langle 1\rangle} = H^{-1}, \quad [H^{-1}]_{\langle n\rangle} = [H^{-1}]_{nn} \tag{126}$$

$$\tilde{P}_{\langle k\rangle} = \left[\begin{array}{c|c} \tilde{P}_k & 0 \\ \hline 0 & \tilde{P}_{\langle k+1\rangle} \end{array}\right], \quad \tilde{P}_{\langle 1\rangle} = \tilde{P}, \quad \tilde{P}_{\langle n\rangle} = \tilde{P}_n \tag{127}$$

$$W_{\langle k\rangle} = \left[\begin{array}{c|c} W_{kk} & 0 \\ \hline \star_{0,0,0} & W_{\langle k+1\rangle} \end{array}\right], \quad W_{\langle 1\rangle} = W, \quad W_{\langle n\rangle} = W_{nn} \tag{128}$$

Note that $\bar{F}_k \leq 0$ holds if $\bar{M}_{[k-1]} < 0$ is satisfied.

Theorem 8 *Let $1 \leq k \leq n$. Suppose that $R_k = 0$ and $p_k \neq 0$. Assume that $H(x_1) < 0$ and $\bar{M}_{[k-1]}(x_1, \hat{x}_{[k-2]}) < 0$ hold for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$ unless $k = 1$. There exists a scalar-valued $\mathbf{C}^0$ function $\lambda_k(x_1, \hat{x}_{[k-2]})$ such that*

$$\lambda_k(x_1, \hat{x}_{[k-2]}) > 0, \quad F_k(x_1, \hat{x}_{[k-2]}) < 0 \tag{129}$$

are satisfied for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$ with $L_k(x_1, \hat{x}_{[k-2]}) = \lambda_k(x_1, \hat{x}_{[k-2]}) I_{p_k}$ if

$$\begin{aligned} \lambda_{max}\left(-B_{-,k}^T W_{\langle k\rangle}^T \tilde{P}_{\langle k\rangle}([H^{-1}]_{\langle k\rangle} + \bar{F}_{k11})\tilde{P}_{\langle k\rangle} W_{\langle k\rangle} B_{-,k}\right) \times & \\ \lambda_{max}\left(-C_{k,-}(W_{[k]}^{-1}[H^{-1}]_{[k]} W_{[k]}^{-T} + \bar{F}_{k22})C_{k,-}^T\right) &\leq \frac{1}{4} \end{aligned} \tag{130}$$

holds for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$.
Proof : Define the following matrices

$$Z_a = -B_{-,k}^T W_{\langle k\rangle}^T \tilde{P}_{\langle k\rangle}([H^{-1}]_{\langle k\rangle} + \bar{F}_{k11})\tilde{P}_{\langle k\rangle} W_{\langle k\rangle} B_{-,k} \tag{131}$$

$$Z_b = (\star_{1,0,0} - \bar{F}_{k12}^T \tilde{P} W B_{-,k})^T C_{k,-}^T \tag{132}$$

$$Z_c = -C_{k,-}(W_{[k]}^{-1}[H^{-1}]_{[k]} W_{[k]}^{-T} + \bar{F}_{k22})C_{k,-}^T \tag{133}$$

with which F_k is represented as

$$F_k = \begin{bmatrix} -\lambda_k I + Z_a & \lambda_k Z_b \\ \lambda_k Z_b^T & -\lambda_k I + \lambda_k^2 Z_c \end{bmatrix} \tag{134}$$

Looking at the definition of F_k in terms of Ψ_k and Φ_k carefully, $\bar{M}_{[k-1]} < 0$ and $H < 0$ imply that

$$\begin{bmatrix} Z_a & Z_b \\ Z_b^T & Z_c \end{bmatrix} \geq 0 \tag{135}$$

holds for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$. Now, define

$$\bar{a} = \lambda_{max}(Z_a) \tag{136}$$

$$\bar{b} = \lambda_{max}(Z_b^T Z_b) \tag{137}$$

$$\bar{c} = \lambda_{max}(Z_c) \tag{138}$$

Here, $\lambda_{max}(\cdot)$ denotes the maximum eigenvalue of a matrix. The inequality (135) directly proves that these non-negative numbers $\bar{a}$, $\bar{b}$ and $\bar{c}$ satisfy $\bar{b} \leq \bar{a}\bar{c}$ for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$. Now, we consider F_k at an arbitrary point $(x_1, \hat{x}_{[k-2]})$ in $\mathcal{R} \times \mathcal{R}^{k-2}$. From Young's inequality, the inequality $F_k < 0$ for the matrix (134) is implied by the existence of $q > 0$ satisfying

$$-\lambda_k I + Z_a + q^{-1} I < 0, \quad \lambda_k^2 Z_c - \lambda_k I + q\lambda_k^2 Z_b^T Z_b < 0 \tag{139}$$

Obviously, (139) is met if the inequalities

$$q^{-1} < \lambda_k - \bar{a}, \quad \bar{b} < q^{-1}(\lambda_k^{-1} - \bar{c}) \tag{140}$$

are satisfied. The existence of $q > 0$ in (140) is guaranteed by

$$\lambda_k - \bar{a} > 0, \quad \bar{b} < (\lambda_k^{-1} - \bar{c})(\lambda_k - \bar{a}) \tag{141}$$

Thus, it has been shown that $F_k < 0$ holds if

$$\lambda_k > \bar{a} \tag{142}$$

$$\lambda_k^2 \bar{c} + (\bar{b} - \bar{a}\bar{c} - 1)\lambda_k + \bar{a} < 0 \tag{143}$$

are satisfied. By manipulating the determinant of (143) together with the condition (142), it is verified that there exists a real number λ_k such that (142) and (143) are satisfied if and only if

$$\bar{b} \leq \bar{a}\bar{c} + 1 - 2\sqrt{\bar{a}\bar{c}}, \quad \bar{a}\bar{c} < 1 \tag{144}$$

hold. Moreover, the solution $\lambda_k \in \mathcal{R}$ to (143) automatically satisfies (142). Recall that the triplet $(\bar{a}, \bar{b}, \bar{c})$ satisfies $\bar{b} \leq \bar{a}\bar{c}$. It is obvious that the two condition in (144) are met if

$$1 - 4\bar{a}\bar{c} > 0 \tag{145}$$

is satisfied. Hence, if (130) holds, then, any real number belonging to

$$\left(\frac{(1 + \bar{a}\bar{c} - \bar{b}) - \sqrt{(1 + \bar{a}\bar{c} - \bar{b})^2 - 4\bar{a}\bar{c}}}{2\bar{c}}, \frac{(1 + \bar{a}\bar{c} - \bar{b}) + \sqrt{(1 + \bar{a}\bar{c} - \bar{b})^2 - 4\bar{a}\bar{c}}}{2\bar{c}} \right) \tag{146}$$

achieves $F_1 < 0$. Note that (146) becomes $(\bar{a}, +\infty)$ as $\bar{c}$ goes to 0. Since all functions $\bar{a}$, $\bar{b}$ and $\bar{c}$ are $\mathbf{C}^0$ functions defined on $\mathcal{R}^k \times \mathcal{R}^{k-2}$, there exits $\mathbf{C}^0$ function $\lambda_k(x_1, \hat{x}_{[k-2]})$ such that the two inequalities in (129) hold for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R}^k \times \mathcal{R}^{k-2}$ under the assumption (130). ∎

The condition (130) is only sufficient for existence of L_k in the backstepping procedure. For instance, Young's inequality is sufficient when $p_k > 1$. The sufficiency is only for the purpose of obtaining a simple and explanatory condition. It should be emphasized that the backstepping procedure by itself does not need any conservatism introduced in the proof of the above theorem. To check whether the backstepping selects $\{L_k, R_k, s_k\}$ properly or not, one only has to compute $\tilde{M}_{[k]} < 0$ or $\bar{M}_{[k]} < 0$. Since the entries of B and C matrices represent the nonlinear bounds of uncertainties, the condition (130) is considered as the upper bound and nonlinearity of tolerable uncertainties. In order to make this point clear, we temporarily suppose that C and B are block diagonal matrices and $R = 0$, $W = I$. In this simple case, uncertainties appear in Σ_P as $B_{ii}\Delta_i C_{ii}$ and the condition (130) becomes

$$\lambda_{max}(B_{kk}^T B_{kk})\lambda_{max}(C_{kk}C_{kk}^T) \leq \frac{1}{4\gamma_k(\gamma_k - \bar{f}_k)\tilde{P}_k^2} \tag{147}$$

$$\bar{f}_1 = 0, \quad \bar{f}_k(x_1, \hat{x}_{[k-2]}) = \begin{bmatrix} \star_{1,k-2,k-1} \\ \diamondsuit_{1,k-3,k-1} \\ \diamondsuit_{1,k-3,k-1} \end{bmatrix}^T \bar{M}_{[k-1]}^{-1} \begin{bmatrix} \star_{1,k-2,k-1} \\ \diamondsuit_{1,k-3,k-1} \\ \diamondsuit_{1,k-3,k-1} \end{bmatrix} \leq 0 \ \text{ for } k \geq 2$$

where $\gamma(x_1)$ is any function satisfying

$$\begin{bmatrix} \gamma_1 & 0 & \cdots \\ 0 & \ddots & \ddots \\ \vdots & \ddots & \gamma_n \end{bmatrix} > -H^{-1} > 0 \tag{148}$$

Although Theorem 8 is derived with $R_k = 0$ for simplicity, the parameters R_k are not useless. One advantage of the SD scaling approach in Section 5 is that we can introduce the additional parameters R_k into each step of backstepping to improve controller performance other than mere robust stabilization (a controller obtained without R_k may not be the best one).

Theorems 7 and 8 show that $\bar{M}_{[k]} < 0$ can be achieved globally for $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1}$ by using scaling matrices in the form of $\lambda_k(x_1, \hat{x}_{[k-2]})I_{p_k} > 0$. If $\lambda_k(x_1, \hat{x}_{[k-2]}) > 0$ is replaced with a positive constant number λ_k, the inequalities (129) might not be met globally. The size of region where these inequalities are satisfied depends on parameters of $(\hat{A}_{[k]}, B_{[k]}, C_{[k]})$ and $(s_1, \cdots, s_{k-1}, L_{[k-1]}, R_{[k-1]})$. In order to achieve the robust stabilization of Σ_P with output feedback via SD scaling, according to Corollary 1, $\lambda_k(x_1, \hat{x}_{[k-2]})I$ should be in the scaling set $\mathbf{L}_k$. Although the SD scaling $\lambda_k(x_1, \hat{x}_{[k-2]})I$ belongs to the scaling sets for both the full and repeated static uncertainties, it does not belong to the scaling set for dynamic uncertainties. Note that 0 is a member of the scaling sets $\mathbf{R}_k$. Thus, we obtain the following.

Theorem 9 *Suppose that a robust observer is chosen such that*

$$\lambda_{max}\left(-B_{-,k}^T W_{\langle k\rangle}^T \tilde{P}_{\langle k\rangle}([H^{-1}]_{\langle k\rangle} + \bar{F}_{k11})\tilde{P}_{\langle k\rangle} W_{\langle k\rangle} B_{-,k}\right) \times$$
$$\lambda_{max}\left(-C_{k,-}(W_{[k]}^{-1}[H^{-1}]_{[k]}W_{[k]}^{-T} + \bar{F}_{k22}\right) C_{k,-}^T) \le \frac{1}{4}, \quad \forall (x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2} \tag{149}$$

is satisfied for all $k = 1, 2, \ldots, n$.

(i) *Assume that the uncertainty Σ_Δ only has static uncertain components Δ_{is} and Δ_{ir}. The system Σ_P can be globally uniformly asymptotically stabilized for any admissible uncertainty by the dynamic output-feedback law (16-17) with a smooth function K.*

(ii) *Assume that the uncertainty Σ_Δ has dynamic uncertain components Δ_{id}. If there exists a constant λ_k belonging to (146) for each $k = 1, 2, \ldots, n$. then, the system Σ_P can be globally uniformly asymptotically stabilized for any admissible uncertainty by the dynamic output-feedback law (16-17) with a smooth function K.*

Proof : Omitted. ∎

The condition (149) may be satisfied for any $\mathbf{C}^0$ functions B and C by taking sufficiently small $\mathbf{C}^0$ functions $\gamma_i(x_1) > 0$, $i = 1, 2, \ldots, n$ since $-H^{-1} < \Gamma$ holds for a robust observer. This is always the case when $n \le 3$(The $n = 3$ case is proved by making s_1 and Y depend on $\hat{x}_1$ as well as x_1.). However, this argument is valid only if an observer is constructed for such a large Γ^{-1}. The smaller γ_k puts a heavier burden on the observer. The required *strong* observers may not always exist unless the full information of the state x is available for feedback. Even if an observer exists, the magnitude of Y and W may become very large when γ_i is too small. Recall that $\bar{F}_k$ depends on Y and W. In this way, the condition (149) shows that there is a strong coupling between observer-gain design and the feedback-gain design. The coupling is due to the existence of the term $\tilde{P}WB$ as expected in Section 3. This fact reveals that the output-feedback robust stabilization problem is not always solvable globally in a backstepping manner for arbitrarily large uncertainties. The condition (149) actually describes the tolerable uncertainties in a recursive manner. This constants sharply with state-feedback control by which global stabilization can be always achieved for arbitrarily large uncertainties(Ito and Freeman, 1998a).

8 Recursive global observer design for output feedback

As it has been shown in previous sections, to design an output-feedback control law via backstepping, one is required to find an observer whose error dynamics is stable to a certain degree. This section shows how to construct such observers in the output feedback case $y = x_1$.

The ordinary observer defined in Section 4 can be constructed whenever the $\mathbf{C}^0$ function $A(x_1)x$ satisfies

$$A(x_1)x = A_1 x + \psi(x_1)$$

with a constant matrix A_1 (Krstić *et al.*, 1995). To explain this fact, we rewrite $A(x_1)x$ as

$$A(y)x = A_1 x + A_2(x_1)y = A_1 x + A_2(x_1)C_y x$$

The observer gain $Y = -Y_1 + A_2(y)$ with a constant vector Y_1 yields

$$A - YC_y = A_1 - Y_1 C_y$$

The right hand side of the above equation is obviously a constant matrix which can be always made stable by choosing Y_1. This implies that (54) is satisfied with a constant matrix $Q_y > 0$. However, this observer-gain is not enough for creating a robust observer. We need to develop a method of constructing robust observer gain.

Again, the matrices A, Y and Γ are supposed to be $\mathbf{C}^0$ functions of y, namely, they are represented by $A(x_1)$, $Y(x_1)$ and $\Gamma(x_1)$, respectively. The matrix W is constant and non-singular. Given $\Gamma(x_1)$, it is required to find the coordinate transformation W and the observer gain $Y(x_1)$ such that (56) is satisfied for all $x_1 \in \mathcal{R}$ with a diagonal Lyapunov matrix $\tilde{P} > 0$.

First, we choose W as

$$W = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ w_2 & 1 & 0 & \cdots & 0 \\ 0 & w_3 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & w_n & 1 \end{bmatrix} \tag{150}$$

$$W^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -w_2 & 1 & 0 & \cdots & 0 \\ w_2 w_3 & -w_3 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ (-1)^{n-1} w_2 \cdots w_n & \cdots & w_{n-1}w_n & -w_n & 1 \end{bmatrix} \tag{151}$$

These entries w_i for $2 \le i \le n$ are constant numbers. Now define

$$\hat{W} = \left[\begin{array}{c} w_1(x_1)\ 0 \cdots 0 \\ \hline W^T \end{array} \right] \tag{152}$$

where $w_1(x_1)$ is a $\mathbf{C}^0$ function defined on $x_1 \in \mathcal{R}$ yet to be determined. Let the observer gain be

$$Y(x_1) = -W^{-1} \begin{bmatrix} w_1(x_1) \\ 0 \end{bmatrix} = - \begin{bmatrix} w_1 \\ -w_1 w_2 \\ \vdots \\ (-1)^{n-1} w_1 w_2 \cdots w_n \end{bmatrix} \tag{153}$$

Then, we obtain

$$\begin{bmatrix} C_y^T & A^T \end{bmatrix} \hat{W} = -C_y^T Y^T W^T + A^T W^T = (A^T - C_y^T Y^T) W^T \tag{154}$$

The inequality (56) is equivalent to

$$\bar{H}(x_1) := \hat{W}^T \bar{A}^T W^{-1} \tilde{P}^{-1} + \tilde{P}^{-1} W^{-T} \bar{A} \hat{W} + \tilde{P}^{-1} \Gamma^{-1} \tilde{P}^{-1} < 0 \tag{155}$$

$$\bar{A} := \begin{bmatrix} C_y^T & A^T \end{bmatrix} \tag{156}$$

Now, let $\tilde{P}$ be any diagonal matrix with positive entries. We now pay attention to the structure of (155) which is actually the same as that of feedback gain design via backstepping except that the lower triangular structure is replaced by the upper triangular one. That is why the parameters of W can be determined recursively from w_n to w_1. To explain this, the following notation is needed.

$$\bar{H}_{\langle k \rangle}(x_1) := \hat{W}_{\langle k \rangle}^T \bar{A}_{\langle k \rangle}^T W_{\langle k \rangle}^{-1} \tilde{P}_{\langle k \rangle}^{-1} + \tilde{P}_{\langle k \rangle}^{-1} W_{\langle k \rangle}^{-T} \bar{A}_{\langle k \rangle} \hat{W}_{\langle k \rangle} + \tilde{P}_{\langle k \rangle}^{-1} \Gamma_{\langle k \rangle}^{-1} \tilde{P}_{\langle k \rangle}^{-1} \tag{157}$$

for $k = 1, 2, \ldots, n$, where

$$\bar{A}_{\langle k \rangle}(x_1) = \left[\begin{array}{c|c} a_{k-1,k} & \star \\ \hline 0 & \bar{A}_{\langle k+1 \rangle} \end{array}\right], \quad \bar{A}_{\langle 1 \rangle} = \bar{A}, \quad \bar{A}_{\langle n \rangle} = [\, a_{n-1,n} \ a_{n,n} \,] \tag{158}$$

$$\Gamma_{\langle k \rangle}(x_1) = \left[\begin{array}{c|c} \gamma_k & 0 \\ \hline 0 & \Gamma_{\langle k+1 \rangle} \end{array}\right], \quad \Gamma_{\langle 1 \rangle} = \Gamma, \quad \Gamma_{\langle n \rangle} = \gamma_n \tag{159}$$

$$\tilde{P}_{\langle k \rangle} = \left[\begin{array}{c|c} \tilde{P}_k & 0 \\ \hline 0 & \tilde{P}_{\langle k+1 \rangle} \end{array}\right], \quad \tilde{P}_{\langle 1 \rangle} = \tilde{P}, \quad \tilde{P}_{\langle n \rangle} = \tilde{P}_n \tag{160}$$

$$W_{\langle k \rangle}^{-1} = \left[\begin{array}{c|c} 1 & 0 \\ \hline \star & W_{\langle k+1 \rangle}^{-1} \end{array}\right], \quad W_{\langle 1 \rangle}^{-1} = W^{-1}, \quad W_{\langle n \rangle}^{-1} = 1 \tag{161}$$

$$\hat{W}_{\langle k \rangle} = \left[\begin{array}{c} w_k \ 0 \ \cdots \ 0 \\ \hline \hat{W}_{\langle k \rangle}^T \end{array}\right], \quad \hat{W}_{\langle 1 \rangle} = \hat{W}, \quad \hat{W}_{\langle n \rangle} = \begin{bmatrix} w_1 \\ 1 \end{bmatrix} \tag{162}$$

Obviously, $\bar{H} = \bar{H}_{\langle 1 \rangle}$. Recall that $a_{k-1,k} \neq 0$ holds for all $x_1 \in \mathcal{R}$ for every $k = 2, 3, \ldots, n$ by assumption. The following can be proved.

Theorem 10 *Suppose that $\bar{A}_{\langle 3 \rangle}$ and $\Gamma_{\langle 3 \rangle}$ are constant matrices. Given an integer $k \in [1, n]$, assume that $\bar{H}_{\langle k+1 \rangle}(x_1) < 0$ holds for all $x_1 \in \mathcal{R}$ unless $k = n$.*
(i) For $k = n, n-1, \ldots 3$
There always exists a constant w_k such that $\bar{H}_{\langle k \rangle} < 0$ is satisfied.
(ii) For $k = 2$
There always exists a constant w_2 such that $\bar{H}_{\langle 2 \rangle}(x_1) < 0$ is satisfied for all $x_1 \in \mathcal{R}$ if there exists positive constants c_i such that

$$\left| \frac{a_{i2}(x_1)}{a_{12}(x_1)} \right| \leq c_i, \quad i = 2, 3, \ldots, n, \quad \left| \frac{1}{\gamma_2(x_1) a_{12}(x_1)} \right| \leq c_0 \tag{163}$$

hold for all $x_1 \in \mathcal{R}$.
(iii) For $k = 1$
There always exists a smooth function $w_1(x_1)$ such that $\bar{H}_{\langle 1 \rangle}(x_1) < 0$ is satisfied for all $x_1 \in \mathcal{R}$.

Because of space limitation, the proof is omitted. Explicit and simple formulas of constructing w_k are available. The constant and growth requirement of A and Γ guarantees w_k to be constant for $2 \leq k \leq n$. The conditions in (163) are met automatically if $\bar{A}_{\langle 2 \rangle}$ is constant. The matrix $\bar{A}_{\langle 3 \rangle}$ is constant if and only if the $\mathbf{C}^0$ function $A(x_1)x$ satisfies

$$A(x_1)x = A_0 x + A_1(x_1) x_1 + A_1(x_1) x_2 \tag{164}$$

where A_0 is a constant matrix. The condition (163) is introduced to guarantee that $\bar{H} < 0$ can be *globally* achieved. It is important that $\bar{H} < 0$ is always semi-globally achievable without the condition (163).

This recursive design of observers resembles backstepping very much. The design starts with a parameter away from observer-gain and back to the actual observer-gain. The recursive structure is in an upper triangular form which is similar to forwarding(or a dual procedure of backstepping). This type of design procedure for observer is a unique feature of this paper. It is ready for automated *numerical* calculation by computer as well. Another feature of the recursive design of observers in this section is that precise knowledge of the system equation is not required for calculating w_i since the design is based on domination instead of cancelation. The approach is amenable to robustification in that the entries of A are allowed to be uncertain. Since observers can be always designed strong enough to an arbitrary degree for linear systems, we can prove the following.

Theorem 11 *Consider the uncertain system Σ_P in the strict-feedback form defined as in Section 4. Suppose that the uncertain system Σ_P is linear. Then, Σ_P is always robustly stabilizable for arbitrarily large static uncertainties by the dynamic output-feedback law (16-17) with constant K and Y.*

Proof : Due to the block lower triangular structure of B and C, the closed-loop system Σ_P with memoryless uncertainties can be described as

$$\frac{d}{dt}\begin{bmatrix}\hat{\chi}\\ \eta\end{bmatrix} = \begin{bmatrix} S(A_\delta + GK)S^{-1} & -SYC_yW^{-1} \\ 0 & W(A_\delta - YC_y)W^{-1}\end{bmatrix}\begin{bmatrix}\hat{\chi}\\ \eta\end{bmatrix} \tag{165}$$

with the uncertain matrix:

$$A_\delta = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots\cdots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots\ \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n-1,1} & a_{n-1,2} & \cdots\cdots & & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots\cdots & & a_{n,n}\end{bmatrix} + \begin{bmatrix} \delta_{11} & 0 & 0 & \cdots\cdots & 0 \\ \delta_{21} & \delta_{22} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots\ \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \delta_{n-1,1} & \delta_{n-1,2} & \cdots\cdots & & 0 \\ \delta_{n,1} & \delta_{n,2} & \cdots\cdots & & \delta_{n,n}\end{bmatrix} \tag{166}$$

where each δ_{ij} is a uniformly bounded function of t. By using the observer design in this section,

$$W^{-T}(A_\delta - YC_y)^TW^T\tilde{P} + \tilde{P}W(A_\delta - YC_y)W^{-1} < 0 \tag{167}$$

can be achieved uniformly in t for all admissible uncertainties δ_{ij}. It is also true that

$$S^{-T}(A_\delta + GK)^TS^TP + PS(A_\delta + GK)S^{-1} < 0 \tag{168}$$

can be satisfied by constant S and K (Freeman and Kokotović, 1996; Ito and Freeman, 1998*a*). According to the argument in (i) of Corollary 1, $N < 0$ is satisfied for A_δ. Finally, Theorem 1 completes the proof. ∎

This theorem reveals that for strict-feedback linear systems, independent robustification of observer design and feedback-gain design can result in robust stabilization of the whole system. This separation of robust observer design from robust feedback-gain design in robustly achieving $N < 0$ *semi-globally* is valid even for a nonlinear system Σ_P. However, for *global* stabilization of nonlinear systems, independent robust observer design is not enough. The robust observer design should be coupled with the robust feedback-gain design to compensate the nonlinear size of uncertainties together (see Section 7, 9 and 10).

Finally, let the author mention some remarks about output-feedback nonlinear design using this recursive observer design and the recursive feedback-gain design proposed in Section 5. Let $N_{\langle k\rangle}$ and $M_{\langle k\rangle}$ be defined by adding subscript $\langle k\rangle$ to every matrix in the right hand side of (64) and (65), respectively. The observer design has a property which is similar to the feedback gain design.

Theorem 12 *Suppose* $1 \leq k \leq n$.

(i-a) $N_{\langle k\rangle}$ *does not include* $\{w_1, w_2, \cdots, w_{k-1}\}$.

(i-b) *Every entry of* $N_{\langle k\rangle}$ *is affine in* w_k.

(i-c) $N_{\langle k\rangle} < 0$ *implies* $N_{\langle k+1\rangle} < 0$ *unless* $k = n$.

(ii-a) $M_{\langle k\rangle}$ *does not include either* $\{w_1, w_2, \cdots, w_{k-1}\}$, $\{L_1, L_2, \cdots, L_{k-1}\}$ *or* $\{R_1, R_2, \cdots, R_{k-1}\}$.

(ii-b) *Every entry of* $M_{\langle k\rangle}$ *is simultaneously affine in* L_k, R_k *and* w_k.

(ii-c) $M_{\langle k\rangle} < 0$ *implies* $M_{\langle k+1\rangle} < 0$ *unless* $k = n$.

In general, the two recursive designs cannot interlace with each other. For example, the observer design must be completed before performing feedback design. There are three reasons. Design parameters of two recursive designs are coupled in M (or N). The feedback-gain design augments the system from top to bottom, and the observer-gain design does from bottom to top. For instance, w_1 is not available when one want to design s_1. In order to carry out the observer design to achieve $M_{\langle k\rangle} < 0$ (or $N_{\langle k\rangle} < 0$), the parameter matrix W should be a function instead of a constant. Further research is needed in this direction.

9 Robust stabilization problems with guaranteed solutions

This section focuses on a special class of uncertain systems considered in previous sections. It will be shown that the class of systems is always robustly stabilized by output feedback for arbitrarily large uncertainties.

The class of systems is described in the following assumptions.

Assumption 1 *The state* x_1 *is available for feedback control. The system matrices* A, B *and* C *depend only on* x_1. *The* B *and* C *matrices satisfy*

$$B(x_1) = \begin{bmatrix} B_{11}(x_1) \\ B_{21}(x_1) \\ \vdots \\ B_{n1}(x_1) \end{bmatrix}, \quad C(x_1) = [\, C_{11}(x_1) \ 0 \ \cdots \ 0 \,] \tag{169}$$

where $B_{i1}(x_1) \in \mathcal{R}^{1 \times p_1}$, $C_{11}(x_1) \in \mathcal{R}^{p_1 \times 1}$ *and* $p_1 = p$.

Assumption 2 *The function* $A(x_1)x$ *satisfies*

$$A(x_1)x = A_0 x + \psi(x_1) + \phi(x_1)x_2 \tag{170}$$

with a constant matrix A_0 *and* $\mathbf{C}^0$ *functions* ψ *and* ϕ. *There exists positive constants* c_i *such that*

$$\left| \frac{a_{i2}^2(x_1)}{a_{12}(x_1)} \right| \leq c_i, \quad i = 2, 3, \ldots, n \tag{171}$$

hold for all $x_1 \in \mathcal{R}$.

Since $A(x_1)x$ is zero at $x = 0$, $\psi(x_1)$ satisfies $\psi(0) = 0$. This implies that (170) is equivalent to

$$A(x_1)x = A_0 x + A_1(x_1)x_1 + A_2(x_1)x_2 \tag{172}$$

with $\mathbf{C}^0$ functions $A_1(x_1)$ and $A_2(x_1)$. This assumption is weaker than a common assumption

$$A(x_1)x = A_0x + \psi(x_1), \quad \psi(0) = 0 \tag{173}$$

in which the nonlinearities are allowed to depend only on the measured state. Note that if $\phi_1(x_1)$(the first row of the vector ϕ) is a constant, we do not need the constraint (171) since such a system can be transformed to a system with $\phi = 0$ by using coordinate transformation.

The coordinate transformation W of $\tilde{x}$ is a constant matrix represented as

$$W = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ w_2 & 1 & 0 & \cdots & 0 \\ 0 & w_3 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & w_n & 1 \end{bmatrix} = \left[\begin{array}{c} 1\ 0\ \cdots\ 0 \\ \hline W_l \end{array}\right] \tag{174}$$

$$W^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -w_2 & 1 & 0 & \cdots & 0 \\ w_2w_3 & -w_3 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ (-1)^{n-1}w_2 \cdots w_n & \cdots & w_{n-1}w_n & -w_n & 1 \end{bmatrix} \tag{175}$$

We also use the following recursive representation.

$$W_{[k]}^{-1} = \left[\begin{array}{c|c} W_{[k-1]}^{-1} & 0 \\ \hline \star_{0,0,0} & 1 \end{array}\right], \quad B_{[k]1} = \begin{bmatrix} B_{11} \\ B_{21} \\ \vdots \\ B_{k1} \end{bmatrix} \tag{176}$$

Let $\tilde{P}$ be a diagonal matrix. Then, we have

$$[P_{[k]}(V_{[k]} + T_{[k]})\hat{A}_{[k]}\hat{S}_{[k]}](x_1, \hat{x}_{[k-1]}) =$$
$$\left[\begin{array}{c|c} \left\{ \begin{array}{r} P_{[k-1]}(V_{[k-1]}(x_1, \hat{x}_{[k-2]}) + T_{[k-1]}(x_1, \hat{x}_{[k-2]})) \\ \hat{A}_{[k-1]}(x_1)\hat{S}_{[k-1]}(x_1, \hat{x}_{[k-2]}) \end{array} \right\} & \begin{array}{c} 0 \\ P_{k-1}a_{k-1,k} \end{array} \\ \hline P_k\star_{1,k-1,k-1} & P_k(a_{kk} + a_{k,k+1}s_k + \star_{1,k-1,k-1}) \end{array}\right]$$
$$[P_{[k]}V_{[k]}B_{[k]1}](x_1, \hat{x}_{[k-1]}) = \left[\begin{array}{c} P_{[k-1]}V_{[k-1]}(x_1, \hat{x}_{[k-2]})B_{[k-1]1}(x_1) \\ \hline \star_{1,k-1,k-1} \end{array}\right]$$
$$[LCS^{-1}](x_1) = [\, L_1(x_1)C_{11}(x_1)\ 0\ \cdots\ 0\,], \quad [RCS^{-1}](x_1) = [\, R_1(x_1)C_{11}\ 0\ \cdots\ 0\,]$$
$$[LCW^{-1}](x_1) = [\, L_1(x_1)C_{11}(x_1)\ 0\ \cdots\ 0\,], \quad [RCW^{-1}](x_1) = [\, R_1(x_1)C_{11}(x_1)\ 0\ \cdots\ 0\,]$$
$$[P_{[k]}V_{[k]}A_{[k]}W_{[k]}^{-1}](x_1, \hat{x}_{[k-1]}) = \left[\begin{array}{c|c} P_{[k-1]}V_{[k-1]}(x_1, \hat{x}_{[k-2]})A_{[k-1]}(x_1)W_{[k-1]}^{-1} & 0 \\ \hline P_k\star_{1,k-1,k-1} & P_k\star_{1,k-1,k-1} \end{array}\right]$$
$$[P_{[k]}T_{[k]}Y_{[k]}C_{y,[k]}W_{[k]}^{-1}](x_1, \hat{x}_{[k-1]}) = \left[\begin{array}{cc|c} P_{[k-1]}T_{[k-1]}(x_1, \hat{x}_{[k-2]})Y_{[k-1]} & 0\ \cdots\ 0 & 0 \\ \hline P_k\star_{1,k-1,k-1} & 0\ \cdots\ 0 & 0 \end{array}\right]$$
$$[B^TW^T\tilde{P}](x_1) = \left[\, B_{11}^T(x_1)\tilde{P}_1\ \ B(x_1)^TW_l^T\tilde{P}_{\langle 2\rangle}\,\right]$$

By using these structures, we obtain the following.

Theorem 13 *(i) For $k = 1$: The symmetric matrix*

$$\bar{M}_{[1]}(x_1) = \Psi_1(x_1) \tag{177}$$

depends only on (L_1, R_1) and s_1.
(ii) For $2 \le k \le n$: $\bar{M}_{[k]}(x_1, \hat{x}_{[k-1]}) < 0$ is equivalent to

$$\begin{bmatrix} \bar{M}_{[k-1]}(x_1, \hat{x}_{[k-2]}) & \Phi_k(x_1, \hat{x}_{[k-1]}) \\ \Phi_k^T(x_1, \hat{x}_{[k-1]}) & \Psi_k(x_1, \hat{x}_{[k-1]}) \end{bmatrix} < 0\ , \tag{178}$$

where Φ_k depends only on (L_1, R_1) and $(s_1, \cdots, s_{k-1})$ and their partial derivatives. The symmetric matrix Ψ_k depends on s_k.

The functions Φ_k and Ψ_k are obtained as follows:

$$\Psi_1(x_1) = \tag{179}$$

$$\begin{bmatrix} 2P_1(a_{11} + a_{12}s_1 + \star_{1,0,0}) & \star_{1,0,0} + (\star_{1,0,0} + C_{11}^T)R_1^T & (\star_{1,0,0} + C_{11}^T)L_1 \\ * & -L_1 - (\tilde{P}WB + W^{-1}C^T R_1^T)^T H^{-1}(\tilde{P}WB + W^{-1}C^T R_1^T) & (\star_{1,0,0} + R_1\star_{1,0,0})C_{11}^T L_1 \\ * & * & -L_1 - L_1 C_{11}[H^{-1}]_{11}C_{11}^T L_1 \end{bmatrix}$$

$$\Phi_k(x_1, \hat{x}_{[k-1]}) = \begin{bmatrix} \star_{1,k-1,k-1} \\ \diamondsuit_{1,k-1,k-1} \\ \diamondsuit_{1,k-1,k-1} \end{bmatrix}, \quad \text{for } 2 \le k \le n \tag{180}$$

$$\Psi_k(x_1, \hat{x}_{[k-1]}) = 2P_k(a_{kk} + a_{k,k+1}s_k + \star_{1,k-1,k-1}), \quad \text{for } 2 \le k \le n \tag{181}$$

where $\star_{1,i,j}$ denotes any function depending only on $(x_1, \hat{x}_{[i]})$ and $(s_1, \cdots, s_j)$ and their partial derivatives. If a function $\star_{1,i,j}$ also depends on (L_1, R_1), it is denoted by $\diamondsuit_{1,i,j}$.

Next, we define $J_k(x_1, \hat{x}_{[k-1]}) \in \mathcal{R}^{1\times 1}$, $E_1(x_1) \in \mathcal{R}^{1\times 2p_1}$ and $F_1(x_1) \in \mathcal{R}^{2p_1\times 2p_1}$ as

$$\Psi_1 = \begin{bmatrix} J_1 & E_1 \\ E_1^T & F_1 \end{bmatrix}, \quad \text{for } k = 1 \tag{182}$$

$$\Psi_k - \Phi_k^T \bar{M}_{[k-1]}^{-1}\Phi_k = J_k, \quad \text{for } k \ge 2 \tag{183}$$

Using the Schur complements formula, we have the following.

Corollary 3 *Let k is any integer belonging to $[1, n]$. Assume that $\bar{M}_{[k-1]}(x_1, \hat{x}_{[k-2]}) < 0$ is satisfied for all $(x_1, \hat{x}_{[k-2]}) \in \mathcal{R} \times \mathcal{R}^{k-2}$ unless $k = 1$. Then, $\bar{M}_{[k]}(x_1, \hat{x}_{[k-1]}) < 0$ holds for all $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1}$ if and only if*

$$F_1 < 0, \ J_1 - E_1 F_1^{-1} E_1^T < 0, \quad \textit{when } k = 1 \tag{184}$$

$$J_k < 0, \quad \textit{when } k \ge 2 \tag{185}$$

are satisfied for all $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R}^k \times \mathcal{R}^{k-1}$.

Theorem 14 *Let k is any integer belonging to $[1, n]$. Given arbitrary $L_1 \in \mathbf{L}_1$ and $R_1 \in \mathbf{R}_1$, there always exist a scalar-valued smooth function $s_k(x_1, \hat{x}_{[k-1]})$ such that*

$$J_1 - E_1 F_1^{-1} E_1^T < 0, \quad \textit{when } k = 1 \tag{186}$$

$$J_k < 0, \quad \textit{when } k \ge 2 \tag{187}$$

are satisfied for all $(x_1, \hat{x}_{[k-1]}) \in \mathcal{R}^k \times \mathcal{R}^{k-1}$.

Theorem 15 *Let $R_1 = 0$. Assume that $H(x_1) < 0$ holds for all $x_1 \in \mathcal{R}$. There exists a scalar-valued* $\mathbf{C}^0$ *function $\lambda_1(x_1)$ such that*

$$\lambda_1(x_1) > 0, \quad F_1(x_1) < 0 \tag{188}$$

are satisfied for all $x_1 \in \mathcal{R}$ with $L_1(x_1) = \lambda_1(x_1)I_{p_1}$ if

$$-[H^{-1}]_{11}\lambda_{max}\left(-B^T W^T \tilde{P} H^{-1} \tilde{P} W B\right)\lambda_{max}\left(C_{11}C_{11}^T\right) \le \frac{1}{4} \tag{189}$$

hold for all $x_1 \in \mathcal{R}$.

I should be noted that the simple form $L_1 = \lambda_1 I$ is used in the above theorem only for the purpose of deriving a simple condition like (189). Remember that the actual backstepping does not need to use the above theorem. Although $L_1 = \lambda_1 I$ is enough to show the existence, one had better exploit the freedom allowed in $\mathbf{L}_i$ to avoid unnecessary high-gain and growth order of control laws. Indeed, this is an advantage of the scaling approach. For example, if uncertain parameters appear in the system equation as scalar-valued functions, the scaling L_1 can be a diagonal matrix with independent entries:

$$L_1 = \begin{bmatrix} \lambda_{11} & 0 & \ddots & \vdots \\ 0 & \lambda_{12} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \cdots & \cdots & 0 & \lambda_{1,p_1} \end{bmatrix} \tag{190}$$

The condition (189) can be always satisfied for any B and C by a sufficiently small and positive function $-[H^{-1}]_{11}(x_1)$. If $H < -\Gamma^{-1}$ is satisfied by a robust observer, $[-H^{-1}]_{11} < \gamma_1$ holds. The smaller γ_1 puts a heavier burden on the observer. However, in Section 8, it has been shown that such a strong observer can be always constructed for arbitrary function $\gamma_1(x_1) > 0$ and arbitrary constants γ_k, $k = 3, 4, \ldots, n$ under Assumption 2. Recall that W is independent of γ_1 in the observer design. In this way, Assumption 1 allows us to have the following.

Theorem 16 *Consider the uncertain system Σ_P under Assumption 1 and Assumption 2.*

(i) *Assume that the uncertainty Σ_Δ only has static uncertain components Δ_{is} and Δ_{ir}. The system Σ_P can be globally uniformly asymptotically stabilized for any admissible uncertainty by the dynamic output-feedback law (16-17) with a smooth function K.*

(ii) *Assume that the uncertainty Σ_Δ has dynamic uncertain components Δ_{id}. If there exists a constant λ_1 belonging to (146) for $k = 1$. then, the system Σ_P can be globally uniformly asymptotically stabilized for any admissible uncertainty by the dynamic output-feedback law (16-17) with a smooth function K.*

Proof : Since $a_{12}(x_1) \neq 0$ for all $x_1 \in \mathcal{R}$ by assumption, we can define a $\mathbf{C}^0$ function as follows:

$$\gamma_2(x_1) = \frac{1}{\sqrt{a_{12}^2(x_1)}}, \quad \left| \frac{1}{\gamma_2(x_1) a_{12}(x_1)} \right| = 1, \quad \gamma_2(x_1) > 0, \quad \forall x_1 \in \mathcal{R} \tag{191}$$

Let $\gamma_1(x_1)$ be defined such that

$$\gamma_1 \lambda_{max}\left(B^T W^T \tilde{P} \Gamma \tilde{P} W B\right) \lambda_{max}\left(C_{11} C_{11}^T\right) \leq \frac{1}{4} \tag{192}$$

is satisfied for all $x_1 \in \mathcal{R}$. Choose γ_i, $i = 3, 4, \ldots, n$ as any positive numbers. Due to Theorem 10 and Assumption 2, there always exists $Y(x_1)$ such that $\bar{H}(x_1) < 0$ holds for all $x_1 \in \mathcal{R}$. Theorem 14 and 15 also guarantees $\bar{M}(x_1, \hat{x}_{[k-1]}) < 0$ for all $(x_1, x_{[k-1]}) \in \mathcal{R} \times \mathcal{R}^{k-1}$ on the Assumption 1, Thus, Corollary 1 proves the claim. ∎

Remember that Assumption 2 only comes from observer design to ensure that the obsever is *globally* strong enough.

As seen in the above theorem, global robust stabilizability against dynamic uncertainties via output feedback is not always achievable if the uncertainty structure and size of uncertainty is prescribed *a priori*. However, if we relax the robustness requirement, a stability robustness in terms of Input-to-State Stability(ISS) can be obtained.

Theorem 17 *Assume that the system Σ_0 satisfies Assumption 1 and Assumption 2. The system Σ_0 can be ISS stabilized by the dynamic output-feedback law (16-17) with a smooth function K.*
Proof : First, set $R_1 = 0$. Let F_1 be represented as

$$F_1 = \begin{bmatrix} -\lambda_1 I + Z_a & \lambda_1 Z_b \\ \lambda_1 Z_b^T & -\lambda_1 I + \lambda_1^2 Z_c \end{bmatrix} \geq 0 \tag{193}$$

$$Z_a = -B^T W^T \tilde{P} H^{-1} \tilde{P} W B, \quad , \quad Z_b = \star_{1,0,0} C_{11}^T, \quad Z_c = -C_{11}[H^{-1}]_{11} C_{11}^T \tag{194}$$

Maximum eigenvalues of the above matrices are denoted by

$$\bar{a} = \lambda_{max}(Z_a), \quad \bar{b} = \lambda_{max}(Z_b^T Z_b), \quad \bar{c} = \lambda_{max}(Z_c) \tag{195}$$

Choose constants

$$\alpha_1 > 0, \quad \nu > 0, \quad \epsilon > 0 \tag{196}$$

arbitrarily. Let $\gamma_1(x_1)$ be a $\mathbf{C}^0$ function such that

$$\gamma_1(x_1)\lambda_{max}(B_{11}^T(x_1)\tilde{P}_1^2 B_{11}(x_1)) + \lambda_{max}(B^T(x_1)W_l^T \tilde{P}_{\langle 2\rangle}\Gamma_{\langle 2\rangle}\tilde{P}_{\langle 2\rangle}W_l B(x_1)) = \alpha_1 \tag{197}$$

holds for all $x_1 \in \mathcal{R}$, where $\Gamma_{\langle 2\rangle} > 0$ is any constant matrix. The existence of a robust observer with respect to these γ_1 and $\Gamma_{\langle 2\rangle}$ is guaranteed by Theorem 10. From $0 < -H^{-1} < \Gamma$ it follows that

$$\alpha_1 \geq \lambda_{max}(B^T W^T \tilde{P} \Gamma \tilde{P} W B) \geq \lambda_{max}(-B^T W^T \tilde{P} H^{-1} \tilde{P} W B) = \bar{a} \tag{198}$$

Let C_{11} be

$$C_{11}(x_1) = \beta_1(x_1)\tilde{C}_{11} \tag{199}$$

where $\tilde{C}_{11}$ is a constant matrix satisfying $\lambda_{max}(\tilde{C}_{11}\tilde{C}_{11}^T) = 1$. Choose a $\mathbf{C}^0$ function $\beta_1(x_1)$ such that

$$(-\epsilon[H^{-1}]_{11}\beta_1^2(x_1) + \bar{b}(x_1)) < \frac{\epsilon(\alpha_1 + \epsilon)}{(\alpha_1 + \nu + \epsilon)^2} \tag{200}$$

holds for all $x_1 \in \mathcal{R}$. There exists such a function β_1 since $\bar{b}(x_1) = \lambda_{max}(C_{11} \star_{1,0,0}^2 C_{11}^T) \geq 0$ and $\epsilon[H^{-1}]_{11} < 0$. Now, choose λ_1 as a positive constant defined by

$$\lambda_1 = \alpha_1 + \nu + \epsilon \tag{201}$$

From $\bar{c} = -[H^{-1}]_{11}\beta_1^2$, the inequality (200) is rewritten as

$$\bar{c} < (\lambda_1^{-1} - \nu\lambda_1^{-2}) - \bar{b}(\lambda_1 - \alpha_1 - \nu)^{-1}, \quad \lambda_1 - \alpha_1 - \nu > 0 \tag{202}$$

The condition (202) is equivalent to

$$\lambda_1 - \alpha_1 - \nu > 0, \quad \bar{b} < (\lambda_1^{-1} - \bar{c} - \lambda_1^{-2}\nu)(\lambda_1 - \alpha_1 - \nu) \tag{203}$$

This implies that there exists a function $q(x_1) > 0$ such that

$$q^{-1} < \lambda_1 - \alpha_1 - \nu, \quad \bar{b} < q^{-1}(\lambda_1^{-1} - \bar{c} - \lambda_1^{-2}\nu) \tag{204}$$

Since $\alpha_1 \geq \bar{a} \geq 0$, it is also true that

$$q^{-1} < \lambda_1 - \bar{a} - \nu, \quad \bar{b} < q^{-1}(\lambda_1^{-1} - \bar{c} - \lambda_1^{-2}\nu) \tag{205}$$

By using Young's inequality, we have

$$\begin{bmatrix} -\lambda_1 I + Z_a & \lambda_1 Z_b \\ \lambda_1 Z_b^T & -\lambda_1 I + \lambda_1^2 Z_c \end{bmatrix} + \nu I < 0 \tag{206}$$

It has been verified that the positive constant $\lambda_1 = \alpha_1 + \nu + \epsilon$ solves $F_1(x_1) + \nu I < 0$ for all $x_1 \in \mathcal{R}$. Since J_k, $k = 1, 2, \ldots, n$ is affine in s_k, it is always possible to achieve

$$M(x_1, \hat{x}_{[n-1]}) + \nu I < 0 \tag{207}$$

for all $(x_1, \hat{x}_{[n-1]}) \in \mathcal{R} \times \mathcal{R}^{n-1}$ by selecting s_k. Using the Schur complements formula of M, we arrive at

$$\frac{d}{dt} V(x, \hat{x}) \leq -\nu \begin{bmatrix} \hat{\chi} \\ \eta \end{bmatrix}^T \begin{bmatrix} \hat{\chi} \\ \eta \end{bmatrix} + w_1^T (L_1 - \nu I) w_1 \tag{208}$$

for all $(x, \hat{x}) \in \mathcal{R}^n \times \mathcal{R}^n$. Here, $L_1 - \nu I = \lambda_1 I - \nu I$ is a positive definite constant matrix. Since S and W are global diffeomorphism it follows that the closed-system is ISS. ∎

Note that the signal z is not required for ISS since ISS is a property defined only with state and disturbance input signals. That is why the matrix C is considered as a free parameter to prove the above theorem.

The class of systems Σ_0 which satisfy Assumption 1 and Assumption 2 includes

$$\begin{aligned} \dot{x}_1 &= x_2 + \psi_1(x_1) + \phi_1(x_1) x_2 + b_1(x_1) w && (209) \\ \dot{x}_2 &= x_3 + \psi_2(x_1) + \phi_2(x_1) x_2 + b_2(x_1) w && (210) \\ \vdots \ & \ \ \vdots && (211) \\ \dot{x}_n &= g(x_1) u + \psi_n(x_1) + \phi_n(x_1) x_2 + b_n(x_1) w && (212) \\ y &= x_1 && (213) \end{aligned}$$

where ϕ has positive constants c_i, $i = 2, 3, \ldots, n$ such that

$$\left| \frac{\phi_i(x_1)}{1 + \phi_1(x_1)} \right| \leq c_i \tag{214}$$

holds for all $x_1 \in \mathcal{R}$. As for uncertain systems, the class of uncertain systems Σ_P which satisfy Assumption 1 and Assumption 2 includes

$$\begin{aligned} \dot{x}_1 &= x_2 + \psi_1(x_1) + \phi_1(x_1) x_2 + \delta_1(x_1, t) && (215) \\ \dot{x}_2 &= x_3 + \psi_2(x_1) + \phi_2(x_1) x_2 + \delta_2(x_1, t) && (216) \\ \vdots \ & \ \ \vdots && (217) \\ \dot{x}_n &= g(x_1) u + \psi_n(x_1) + \phi_n(x_1) x_2 + \delta_n(x_1, t) && (218) \\ y &= x_1 && (219) \end{aligned}$$

where ϕ_i satisfies (214). For each $i \in [1, n]$, the uncertain part $\delta_i(x_1, t)$ is any function whose absolute value is bounded above by a $\mathbf{C}^0$ function, namely, there exists a $\mathbf{C}^0$ function f_i such that

$$|\delta_i(x_1, t)| \leq |f_i(x_1)|, \quad f_i(0) = 0 \tag{220}$$

holds for all $x_1 \in \mathcal{R}$ and $t \geq 0$. In fact, since the function $f_i(x_1)$ can be decomposed as $f_i(x_1) = m_i(x_1)x_1$, the uncertain system can be represented by Σ_0 with

$$C_{11} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad B = \begin{bmatrix} m_1 & 0 & \cdots & 0 \\ 0 & m_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & m_n \end{bmatrix}, \quad p_1 = n \tag{221}$$

The uncertain block consists of $w_i = h_{\Delta_i}(z_i, t)$, $i = 1, 2, \ldots, n$ which have instantaneous gain less than equal to one.

10 Enlarging the class of nonlinearities

This section shows how to remove the constraint (171). To overcome the limitation, we need to modify the coordinate change for the recursive observer design. Consider the diffeomorphism between $\tilde{x}$ and η as a function of $\hat{x}_1$ and $\tilde{x}_1$.

$$\eta = W(\hat{x}_1, \tilde{x}_1)\tilde{x} \tag{222}$$

Since the matrix W depends on $\hat{x}_1$ as well as $\tilde{x}_1$, we have

$$\dot{\eta} = \tilde{W}(\hat{x}_1, \tilde{x})\dot{\hat{x}} + \bar{W}(\hat{x}_1, \tilde{x})\dot{\tilde{x}} \tag{223}$$

with appropriate matrices $\tilde{W}$ and $\bar{W}$. Now we choose $\bar{W}$ as

$$\bar{W}(\hat{x}_1, \tilde{x}_1) = \bar{W}(x_1) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \bar{w}_2(x_1) & 1 & 0 & \cdots & 0 \\ 0 & \bar{w}_3 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \bar{w}_n & 1 \end{bmatrix} \tag{224}$$

Here, the parameter $\bar{w}_2$ is a $\mathbf{C}^0$ function of x_1 instead of a constant. Let W have the following structure:

$$W(x_1, \hat{x}_1) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ w_2(\hat{x}_1, \tilde{x}_1) & 1 & 0 & \cdots & 0 \\ 0 & w_3 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & w_n & 1 \end{bmatrix} \tag{225}$$

Then, we have

$$w_2(\hat{x}_1, \tilde{x}_1)\tilde{x}_1 = \int_0^{\tilde{x}_1} \bar{w}_2(\hat{x}_1 - \xi)d\xi, \quad w_i = \bar{w}_i, \; i = 3, 4, \ldots, n \tag{226}$$

$$\tilde{W}(\hat{x}_1, \tilde{x}_1) = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \tilde{w}_2(\hat{x}_1, \tilde{x}_1) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \tilde{w}_2(\hat{x}_1, \tilde{x}_1) = \frac{dw_2(\hat{x}_1, \tilde{x}_1)}{d\hat{x}_1}\tilde{x}_1 \tag{227}$$

From $\dot{\hat{\chi}} = (V + T)\dot{\hat{x}}$ and $\dot{\hat{\chi}}_1 = \dot{\hat{x}}_1$ it follows that

$$\dot{\eta} = \tilde{W}(\hat{x}_1, \tilde{x}_1)\dot{\hat{\chi}} + \bar{W}(\hat{x}_1, \tilde{x}_1)\dot{\tilde{x}} \tag{228}$$

The closed-loop system becomes

$$\frac{d}{dt}\begin{bmatrix}\hat{\chi}\\ \eta\end{bmatrix} = \begin{bmatrix}(V+T)(A+GK)S^{-1} & -(VA+TYC_y)W^{-1}\\ \tilde{W}T(A+GK)S^{-1} & \bar{W}(A-YC_y)W^{-1}-\tilde{W}TYC_y\end{bmatrix}\begin{bmatrix}\hat{\chi}\\ \eta\end{bmatrix} + \begin{bmatrix}VB\\ -B\end{bmatrix} w \tag{229}$$

Here, $\tilde{W}V = 0$ is used. The observer inequality becomes

$$H := (W^{-T}(A-YC_y)^T\bar{W}^T - C_y^T Y^T T^T \tilde{W}^T)\tilde{P} + \tilde{P}(\bar{W}(A-YC_y)W^{-1} - \tilde{W}TYC_y) < -\Gamma^{-1} \tag{230}$$

The matrix $\bar{M}$ is obtained as

$$\bar{M}(x_1,\hat{x}) := M_{11}(x_1,\hat{x}) - M_{12}(x_1,\hat{x})H^{-1}(x_1,\hat{x}_1)(x_1,\hat{x})M_{12}^T(x_1,\hat{x}) \tag{231}$$

$$M_{11}(x_1,\hat{x}) := \begin{bmatrix}\left\{\begin{matrix} S^{-T}(A+GK)^T(V+T)^TP+\\ P(T+V)(A+GK)S^{-1}\end{matrix}\right\} & PVB+S^{-T}C^TR^T & S^{-T}C^TL\\ B^TV^TP+RCS^{-1} & -L & 0\\ LCS^{-1} & 0 & -L\end{bmatrix} \tag{232}$$

$$M_{12}(x_1,\hat{x}) := \begin{bmatrix}P(VA+TYC_y)W^{-1} - S^{-T}(A+GK)^T T^T \tilde{W}^T\tilde{P}\\ B^T\tilde{W}^T\tilde{P}+RCW^{-1}\\ LCW^{-1}\end{bmatrix} \tag{233}$$

We next show that it is always possible to construct an observer-gain such that (230) holds. By using

$$\bar{W}^{-1}\tilde{W}T = \bar{W}^{-1}\tilde{W}, \quad C_yW = C_y \tag{234}$$

we have

$$\bar{W}(A-YC_y)W^{-1} - \tilde{W}TYC_y = \bar{W}(A-YC_y-\bar{W}^{-1}\tilde{W}YC_y)W^{-1} \tag{235}$$

Let the modified observer-gain $\bar{Y}$ be defined by

$$\bar{Y} = (I+\bar{Q})Y, \quad \bar{Q} = \bar{W}^{-1}\tilde{W} = \begin{bmatrix}0 & 0 & \cdots & 0\\ \tilde{w}_2 & 0 & \cdots & 0\\ \star_{0,0,0} & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ \star_{0,0,0} & 0 & \cdots & 0\end{bmatrix} \tag{236}$$

where $(I+\bar{Q})$ is invertible. The actual observer-gain can be recovered by $Y = (I+\bar{Q})^{-1}\bar{Y}$. Then, we obtain

$$\bar{W}(A-YC_y)W^{-1} - \tilde{W}TYC_y = \bar{W}(A-\bar{Y}C_y)W^{-1} \tag{237}$$

Choose the modified observer gain as

$$\bar{Y}(x_1,\hat{x}_1) = -\bar{W}^{-1}(x_1)\begin{bmatrix}\bar{w}_1(x_1,\hat{x}_1)\\ 0\end{bmatrix} = -\begin{bmatrix}\bar{w}_1\\ -\bar{w}_1\bar{w}_2\\ \vdots\\ (-1)^{n-1}\bar{w}_1\bar{w}_2\cdots\bar{w}_n\end{bmatrix} \tag{238}$$

Then, the inequality (230) is equivalent to

$$\bar{H}(x_1,\hat{x}_1) := \hat{W}^T\bar{A}^TW^{-1}\tilde{P}^{-1} + \tilde{P}^{-1}W^{-T}\bar{A}\hat{W} + \tilde{P}^{-1}\Gamma^{-1}\tilde{P}^{-1} < 0 \tag{239}$$

$$\bar{A} := \begin{bmatrix} C_y^T & A^T\end{bmatrix}, \quad \hat{W}(x_1,\hat{x}_1) = \left[\frac{\bar{w}_1(x_1,\hat{x}_1)|0}{W^T}\right] \tag{240}$$

where $\bar{w}_1(x_1,\hat{x}_1)$ is a $\mathbf{C}^0$ function defined on $(x_1,\hat{x}_1)\in\mathcal{R}\times\mathcal{R}$ yet to be determined. Now, let $\tilde{P}$ be a diagonal matrix with positive entries. Then, the structure of (239) is the same as that of (155) so that the parameters of W can be always determined recursively from $\bar{w}_n$ to $\bar{w}_1$.

Theorem 18 *Suppose that $\bar{A}_{\langle 3\rangle}$ and $\Gamma_{\langle 3\rangle}$ are constant matrices. The parameter γ_2 is allowed to depend only on x_1, and γ_1 is allowed to depend only on $(x_1, \hat{x}_1)$. Given an integer $k \in [1, n]$, assume that $\bar{H}_{\langle k+1\rangle}(x_1) < 0$ holds for all $x_1 \in \mathcal{R}$ unless $k = n$.*
(i) For $k = n, n-1, \ldots 3$
There always exists a constant $\bar{w}_k$ such that $\bar{H}_{\langle k\rangle} < 0$ is satisfied.
(ii) For $k = 2$
There always exists a smooth function $\bar{w}_2(x_1)$ such that $\bar{H}_{\langle 2\rangle}(x_1) < 0$ is satisfied for all $x_1 \in \mathcal{R}$.
(iii) For $k = 1$
There always exists a smooth function $\bar{w}_1(x_1, \hat{x}_1)$ such that $\bar{H}_{\langle 1\rangle}(x_1, \hat{x}_1) < 0$ is satisfied for all $(x_1, \hat{x}_1) \in \mathcal{R} \times \mathcal{R}$.

Next, we consider the feedback-gain design on Assumption 1. The procedure of backstepping can be carried out in the same way as Section 5 or 9. The rest of section investigates the existence of solutions $\{s_1, s_2, \ldots, s_n, L_1, R_1\}$ in the recursive design. According to (231), the term which is structurely different from the constant W case appears only in M_{12} as

$$-S^{-T}(A+GK)^T T^T \tilde{W}^T \tilde{P} = \begin{bmatrix} 0 & \tilde{w}_2\tilde{P}_2(a_{11}+a_{12}s_1) & 0 \cdots 0 \\ 0 & \tilde{w}_2\tilde{P}_2 a_{12} & 0 \cdots 0 \\ 0 & 0 & 0 \cdots 0 \\ \vdots & \vdots & \vdots \ \ddots \ \vdots \\ 0 & 0 & 0 \cdots 0 \end{bmatrix} \tag{241}$$

Thus, Φ_k $(2 \leq k \leq n)$ and Ψ_k $(1 \leq k \leq n)$ are the same as (180) and (181). This implies that the recursive design yields a solution $(s_1, s_2, \ldots, s_n)$ to $\bar{M} < 0$ whenever there exists s_1 solving $\Psi_1 < 0$. We shall calculate an analytical expression of the existence condition for s_1. Assume that a diagonal matrix Γ satisfies $-H^{-1} < \Gamma$. Only for the purpose of showing an explanatory analytical solution, $-H^{-1}$ is temporarily replaced with Γ in the definition of $\bar{M}_{[1]}$, i.e., Ψ_1. This replacement is valid since

$$M_{[1]11} - \bar{Q}_1^T M_{12} H^{-1} M_{12}^T \bar{Q}_1 \leq M_{[1]11} + \bar{Q}_1^T M_{12} \Gamma M_{12}^T \bar{Q}_1 \tag{242}$$

We now define Ψ_1 as the right hand side of the above equation instead of the left hand side. Then, the matrix Ψ_1 is

$$\Psi_1(x_1, \hat{x}_1) =$$
$$\begin{bmatrix} 2P_1(a_{11}+a_{12}s_1+\rho_1) & C_{11}^T R_1^T + \rho_2 & (C_{11}^T + \gamma_1\star_{1,0,0})L_1 \\ * & \left\{ \begin{array}{l} -L_1 + \gamma_1(B_{11}^T\tilde{P}_1 + R_1C_{11})(\tilde{P}_1B_{11} + C_{11}^T R_1^T) \\ +B^T W_l^T \tilde{P}_{\langle 2\rangle}\Gamma_{\langle 2\rangle}\tilde{P}_{\langle 2\rangle}W_l B \end{array} \right\} & \gamma_1(B_{11}^T\tilde{P}_1 + R_1C_{11})C_{11}^T L_1 \\ * & * & -L_1 + \gamma_1 L_1 C_{11} C_{11}^T L_1 \end{bmatrix} \tag{243}$$
$$\rho_1 = \gamma_1 P_1^2 \bar{w}_1^2 + \gamma_2 \tilde{P}_2^2 \tilde{w}^2 (a_{11}+a_{12}s_1)^2 \tag{244}$$
$$\rho_2 = -\gamma_1 P_1 \bar{w}_1(\tilde{P}B_{11}^T + R_1C_{11}) + \gamma_2\tilde{P}_2^2\tilde{w}_2(a_{11}+a_{12}s_1)(\bar{w}_2B_{11} + B_{21}) \tag{245}$$

The function $F_1 \in \mathcal{R}^{2p_i \times 2p_i}$ is identical with that in Section 9(replacing $-H^{-1}$ with Γ). This implies that Theorem 15 is valid for Ψ_1 in (243). As for Theorem 14, the inequality

$$J_1 - E_1 F_1^{-1} E_1^T < 0 \tag{246}$$

becomes

$$\gamma_2\tilde{w}_2^2\tilde{P}_2^2(1+\gamma_2\star_{1,0,0}^2)(a_{11}+a_{12}s_1)^2 + (2P_1 + \gamma_1\gamma_2\bar{w}_1\tilde{w}_2\star_{1,0,0})(a_{11}+a_{12}s_1) + \gamma_1\bar{w}_1^2P_1^2(1+\gamma_1\star_{1,0,0}^2) < 0 \tag{247}$$

Since this inequality is quadratic in s_1, the existence of solution s_1 is not guaranteed. Note that $\tilde{w}_2$ and $\bar{w}_1$ depend on $\hat{x}_1$. If solutions s_1 exist, they are functions of $\hat{x}_1$ as well as x_1. By making γ_1 small, the inequality (247) seems to have a solution $s_1(x_1, \hat{x}_1)$. However, this is not always the case since the magnitude of $\bar{w}_1$ may become large when γ_1 is too small. Let D_s denote

$$D_s = (2P_1 + \gamma_1\gamma_2\bar{w}_1\tilde{w}_2\star_{1,0,0})^2 - 4\gamma_1\gamma_2\bar{w}_1^2\tilde{w}_2^2P_1^2\tilde{P}_2^2(1+\gamma_1\star_{1,0,0}^2)(1+\gamma_2\star_{1,0,0}^2) \tag{248}$$

Then, we have the following theorem.

Theorem 19 *Consider the uncertain system Σ_P satisfies Assumption 1 and (170). If there exists a* $\mathbf{C}^0$ *function $\gamma_1(x_1, \hat{x}_1) > 0$ such that $D_s(x_1, \hat{x}_1) > 0$ and*

$$\gamma_1\lambda_{max}\left(-B^TW^T\tilde{P}\Gamma^{-1}\tilde{P}WB\right)\lambda_{max}\left(C_{11}C_{11}^T\right) \leq \frac{1}{4} \tag{249}$$

are satisfied for all $(x_1, \hat{x}_1) \in \mathcal{R}\times\mathcal{R}$, then, the system Σ_P can be globally uniformly asymptotically stabilized for any admissible static uncertainty by the dynamic output-feedback law (16-17) with a smooth function K.

Note that W in (189) is independent of γ_1. Recall that γ_1 is an index of how much the observer is made robust against uncertainty. Thus, the observer-gain parameter $\bar{w}_1$ depends on γ_1. There is inevitable interplay between observer design for choosing $\bar{w}_1$ and feedback-gain design for s_1. The existence of γ_1 in Theorem 19 for the class of systems has not been proved at this time. This sort of coupling exists even in stabilization of a nominal system. If the system does not have any uncertainties, the design inequality of s_1 is

$$J_1 = \gamma_2\tilde{w}_2^2\tilde{P}_2^2(a_{11}+a_{12}s_1)^2 + 2P_1(a_{11}+a_{12}s_1) + \gamma_1\bar{w}_1^2P_1^2 < 0\ . \tag{250}$$

Although (249) disappears in nominal stabilization, the solvability of (250) is apparently coupled with observer design.

Acknowledgments

The author is grateful to Mirsolav Krstić for discussion and constructive comments on the direction of this research during the author's visit to University of California, San Diego from November 1998 through March 1999.

References

Esfandiari, F. and H.K. Khalil (1992). Output feedback stabilization of fully linearizable systems. *Int. J. Contr* **56**, 1007–1037.

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

Ito, H. (1996). Robust control for nonlinear systems with structured $\mathcal{L}_2$-gain bounded uncertainty. *Syst. Contr. Lett.* **28**, 167–172.

Ito, H. (1998*a*). Semi-global robust nonlinear control: State-dependent scaling and computational aspects. In: *Proc. Amer. Contr. Conf.*. pp. 1663–1668.

Ito, H. (1998*b*). State-dependent scaling for robust nonlinear control: Techniques and effectiveness. In: *Proc. 37th IEEE Conf. Decis. Contr.*. pp. 4111–4114.

Ito, H. and R. Freeman (1998*a*). A new look at robust backstepping through state-dependent scaling design. submitted to 1999 Amer. Contr. Conf.

Ito, H. and R.A. Freeman (1998*b*). Generalized state-dependent scaling: Backstepping for local optimality, global inverse optimality, and global robust stability. submitted to 1999 European Contr. Conf.

Khalil, H.K. and F. Esfandiari (1993). Semi-global stabilization of a class of nonlinear systems using output feedback. *IEEE Trans. Autom. Contr.* **38**, 1412–1415.

Krstić, M., I. Kanellakopoulos and P.V. Kokotović (1995). *Nonlinear and adaptive control design*. John Wiley & Sons. Now York.

Lin, W. and C Qian (1998). Semi-global robust stabilization of nonlinear systems by partial state and output feedback. pp. 3105–3110.

Lin, Z. and A. Saberi (1995). Robust semi-global stabilization of minimum-phase input-output linearizable systems via partial state and output feedback. *IEEE Trans. Autom. Contr.* **40**, 1029–1041.

Mareels, I.M. and D.J. Hill (1992). Monotone stability of nonlinear feedback systems. *J. Math. Syst. Estimation and Contr.* **2**, 275–291.

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

Sontag, E. (1998). Notions of integral input-to-state stability. In: *Proc. Amer. Contr. Conf.*. pp. 3210–3214.

Sontag, E. and Y. Wang (1996). New characterizations of input to state stability. *IEEE Trans. Automat. Control* **41**, 1283–1294.

Teel, A.R. and L. Praly (1994). Global stabilizability and observability imply semi-global stabilizability by output feedback. *Syst. Contr. Lett.* **22**, 313–325.

Teel, A.R. and L. Praly (1995). Tools for semi-global stabilization by partial state feedback and output feedback. *SIAM J. Contr. and Optim.* **33**, 1443–1488.