Backstepping Design with Local Optimality Matching

Zigang Pan, *Member, IEEE*, Kenan Ezal, *Member, IEEE*, Arthur J. Krener, *Fellow, IEEE*, and Petar V. Kokotović, *Fellow, IEEE*

Abstract— **In this study of the nonlinear H^∞-optimal control design for strict-feedback nonlinear systems our objective is to construct globally stabilizing control laws to match the optimal control law up to any desired order, and to be inverse optimal with respect to some computable cost functional. Our recursive construction of a cost functional and the corresponding solution to the Hamilton-Jacobi-Isaacs equation employs a new concept of nonlinear Cholesky factorization. When the value function for the system has a nonlinear Cholesky factorization, we show that the backstepping design procedure can be tuned to yield the optimal control law.**

Keywords— **Nonlinear Cholesky factorization, backstepping, local optimality, inverse optimality, disturbance attenuation.**

I. Introduction

After a successful solution of the linear H^∞-optimal control problem, recent research attention has been focused on the robust control design of nonlinear systems [2], [3], [4], [5], [6]. The dynamic game approach [7] provides a natural setting for worst-case designs and requires the solution of a Hamilton-Jacobi-Isaacs (HJI) equation.

Although a general solution method is not available for HJI equations, a local solution exists when the nonlinear system has a controllable Jacobi linearization and the cost functional has a Taylor series expansion with quadratic leading terms. Similar to the solution of the Hamilton-Jacobi-Bellman equation [8], [9], the solution to the HJI equation can be found by computing the coefficients of its Taylor series expansion [10], [11], and [12]. This Taylor series solution may provide an adequate approximation to the optimal control law in a small neighborhood of the origin but it may be unsatisfactory, or even unstable, in a larger region.

Among the recent advances in nonlinear feedback design surveyed in [13], a systematic procedure, known as integrator backstepping [14], is applicable to nonlinear systems in strict-feedback form. This procedure has been perfected in a recent book [15]. The backstepping design offers a lot of flexibility at each step of the control design, but it does not clarify what choice will lead to a better overall design.

In this paper, we approach the H^∞-optimal control problem for strict-feedback nonlinear systems by combining the Taylor series approach with the backstepping design. Our objective is to construct globally stabilizing control laws which match the optimal control law up to any desired order and are inverse optimal with respect to some computable cost functional. Our recursive procedure employs a new concept of nonlinear Cholesky factorization, such that the given positive definite function is equal to the sum of squares of the state variables when expressed in appropriate coordinates. We derive conditions under which a given positive definite nonlinear function has a nonlinear Cholesky factorization. When the optimal value function has a nonlinear Cholesky factorization, we show that the backstepping procedure can be tuned to yield the optimal control design. We develop a recursive computation scheme for an inverse optimal controller that matches the optimal solution up to any desired order of Taylor series expansion. Using this approximating function and its nonlinear Cholesky factorization, we recursively construct the desired matching controller by another backstepping procedure. A simulation example illustrates the theoretical findings. A detailed discussion of the first order matching design can be found in [16].

Inverse optimal design for nonlinear systems has been investigated in earlier studies [17], [18]. In [17], robust control problems for systems subject to bounded disturbances are studied, for which robust inverse optimal controllers that guarantee input-to-state stability of the closed-loop systems are obtained. In [18], it is shown that input-to-state stability is both necessary and sufficient for inverse optimality of nonlinear systems. The objective of this paper, on the other hand, is to design inverse optimal controllers which achieve a desired level of $\mathcal{L}_2$ disturbance attenuation with any desired order of local optimality.

We formulate the problem in Section 2. In Section 3, we present nonlinear Cholesky factorization, which is utilized for the backstepping design in Section 4. A numerical example in Section 5 illustrates the theory. We close with concluding remarks in Section 6.

Manuscript received November 2, 1999;revised October 30, 2000. Recommended by Associate Editor P. Tomei. This work was supported in part by the National Science Foundation under grant ECS-9812346 and the Air Force Office of Scientific Research under grant F49620-00-1-0358. This research first appeared as a technical report [1].

Z. Pan is with Department of Electrical and Computer Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221, USA. E-mail: zpan@ececs.uc.edu.

K. Ezal is with Toyon Research Corporation, 75 Aero Camino, Suite A, Goleta, CA 93117, USA.

A. J. Krener is with the Department of Mathematics, University of California, Davis, CA 95616, USA.

P. V. Kokotović is with the Center for Control Engineering and Computation, Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106, USA.

II. Problem Formulation

We consider the following strict-feedback nonlinear system with disturbance inputs:

$$\begin{aligned}\dot{x}_1 &= x_2 + f_1(x_1) + h_1'(x_1)w && \text{(1a)}\\ \dot{x}_2 &= x_3 + f_2(x_1,x_2) + h_2'(x_1,x_2)w && \text{(1b)}\\ &\vdots \qquad \vdots \\ \dot{x}_n &= u + f_n(x_1,\cdots,x_n) + h_n'(x_1,\ldots,x_n)w && \text{(1c)}\end{aligned}$$

where $x = (x_1, \ldots, x_n)'$ is the state variable with $x(0) = 0$; u is the scalar control input; and w is the q-dimensional disturbance input. We will denote the set of locally Lipschitz feedback control laws $u = \mu(x)$ by $\mathcal{M}_u$. We consider that the disturbance w is generated by some adversary player according to $w(t) = \nu(t,x)$, where $\nu : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^p$ is piecewise continuous in t and locally Lipschitz in x. We will denote the set of of disturbance strategies $\nu(t,x)$ by $\mathcal{M}_w$.

This nonlinear system can be compactly written as $\dot{x} = f(x) + Bu + H(x)w$, where $B = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}'$. Associated with it, we introduce a cost functional $J = \int_0^\infty l(x,u)\,dt$, where $l(x,u) = q(x) + r(x)u^2$.[1] The design objective is to minimize the worst case disturbance attenuation level $\inf_{\mu\in\mathcal{M}_u}\sup_{\nu\in\mathcal{M}_w} J^{1/2}/\|w\| =: \gamma^*$, where $\|.\|$ denotes the $\mathcal{L}_2$ norm of a signal and γ^* is the optimal performance. The nonlinear H^∞ optimal control problem then consists of the evaluation of γ^* and finding a controller $\mu \in \mathcal{M}_u$ that guarantees any desired disturbance attenuation level $\gamma > \gamma^*$.

This nonlinear H^∞ problem has been shown [7] to be closely related to a class of zero-sum differential games with the cost functional indexed by the desired attenuation level γ,

$$J_\gamma = J - \gamma^2\|w\|^2 = \int_0^\infty (l(x,u) - \gamma^2 w'w)\,dt, \tag{2}$$

where the control is the minimizer and the disturbance is the maximizer. We are particularly interested in the upper value of the game: $\inf_{\mu\in\mathcal{M}_u}\sup_{\nu\in\mathcal{M}_w} J_\gamma$. For any $\gamma > \gamma^*$ the upper value of this zero-sum game is *zero* when $x(0) = 0$ [2]. On the other hand, for any $\gamma < \gamma^*$ the upper value is strictly positive. Because of the above stated equivalence, we will focus on the zero-sum differential game problem (2) for a given value of γ.

The following two assumptions characterize the functions in this study.

Assumption A1: The nonlinear functions f_i, h_i, $i = 1,\ldots,n$, q and r are $\mathcal{C}^\infty$ (or simply smooth) in all of their arguments. The open loop system of (1) has the origin as an equilibrium, $f_i(x_1,\ldots,x_i)|_{x=0} = 0$, $i = 1,\ldots,n$. ◇

[1] We assume that the weighting function is quadratic in u. For the more general case $l(x,u) = q(x) + p(x)'u + r(x)u^2$, we can easily remove the cross term by a simple redefinition of the control variable $\tilde{u} = u + r^{-1}p(x)$.

[2] In the case when $x(0)$ is not zero but fixed, we have the following inequality if $\gamma > \gamma^*$, $\int_0^T l(x,u)\,dt \le \gamma^2\int_0^T w'w\,dt + C$, for some constant $C > 0$, which depends only on the initial condition, and for any terminal time $T > 0$.

Assumption A2: The weighting function $l(x,u)$ is positive definite and radially unbounded in x and u, and $l(0,0) = 0$. Furthermore, the Hessian of $q(x)$ evaluated at 0 is a positive definite matrix Q_l. The function $r(x)$ is positive for any x. In particular $r(0) = R_l > 0$. ◇

The solution to the nonlinear zero-sum differential game problem satisfies the HJI equation:

$$V_x(x)f(x) + \frac{1}{4\gamma^2}V_x(x)H(x)H'(x)V_x'(x) - \frac{1}{4}V_x(x)Br^{-1}(x)B'V_x' + q(x) = 0 \tag{3}$$

and, when the solution is available, the optimal control law is given by

$$u = \mu^*(x) = -\frac{1}{2}r^{-1}(x)B'V_x'. \tag{4}$$

While it is very difficult to obtain $V(x)$ in an explicit form, it is relatively simple to compute the Taylor series expansions of $V(x)$ and $\mu^*(x)$ around $x = 0$, see, for example, [8], [9], [10], [2], [11], [12]. Truncated Taylor series expansions yield near-optimal controllers when x is close to the origin, but, when $|x|$ is large, they may not guarantee stability for the nonlinear system.

Our design objective is not only to match the optimal control design up to any desired order, but also to guarantee global stability and inverse robust optimality of the closed-loop system. This objective is made precise with the following two definitions.

Definition 1: A smooth control law $\mu(x)$ is *locally optimal matching to the mth order*, if the truncation of the Taylor series expansion of $\mu(x)$ up to the mth order is equal to that of the worst-case optimal control law μ^*. ◇

Definition 2: A smooth control law $\mu(x)$ is *inversely robust optimal*, if there exists a nonnegative function $\check{l}(x,u) = \check{q}(x) + \check{r}(x)u^2$ such that the zero-sum game with cost functional

$$J_{\gamma i} = \int_0^\infty (\check{l}(x,u) - \gamma^2|w|^2)\,dt \tag{5}$$

admits $\mu(x)$ as its minimax control law. Furthermore, the value function $V_{\gamma i}$ associated with the cost functional $J_{\gamma i}$ and $\check{q}$ are radially unbounded, and the function $\check{r}$ is positive for any $x \in \mathbb{R}^n$. ◇

Throughout the paper, any function with an "over bar" will denote a function defined in terms of the transformed state variables, such as $\bar{a}$ denoting $a(x)$ expressed in terms of a new state variable z. Given any polynomial function V, we will denote its mth order homogeneous terms by $V_{[m]}$, and its homogeneous terms up to mth order by $V^{[m]}$, so that $V^{[m]} = \sum_{i=0}^m V_{[i]}$. For a smooth function V, $V_{[m]}$ denotes the mth order homogeneous terms in the Taylor series expansion of V around the origin, and $V^{[m]}$ denotes the homogeneous terms up to mth order in the Taylor series expansion. The variable $z = (z_1,\ldots,z_n)$ denotes the transformed coordinates for the nonlinear system.

Using this notation, we make the following basic assumption.

Assumption A3: The Taylor series expansion of the optimal value function is given up to $(m+1)$st order $V^{[m+1]}(x) = V_{[2]}(x) + V_{[3]}(x) + \cdots + V_{[m+1]}(x)$, and the Taylor series expansion of the minimax control law $\mu^*(x)$ is given up to mth order $\mu^{*[m]}(x) = \mu^*_{[1]}(x) + \cdots + \mu^*_{[m]}(x) = -\frac{1}{2}(r^{-1}(x)B'V_x')^{[m]}$. ◇

III. A Nonlinear Cholesky Factorization

In the earlier linear-quadratic matching backstepping design [16], the key step that prescribes all the backstepping coefficients is the Cholesky factorization of the quadratic value function. In this section, we study the nonlinear version of the Cholesky factorization.

For any given positive definite and radially unbounded nonlinear function $V(x)$, we are interested in the necessary and sufficient conditions under which there exists global upper triangular diffeomorphism $z = \Phi(x)$ such that $V(x) = \Phi'(x)\Phi(x)$, where Φ is called upper triangular if it can be written as

$$z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} \phi_1(x_1, \ldots, x_n) \\ \vdots \\ \phi_n(x_n) \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{6}$$

Without loss of generality, we will limit our attention to zero preserving transformations, that is, $0 = \Phi(0)$. While a lower triangular transformation is important for the backstepping control design, we pursue a upper triangular transformation just to parallel the well known Cholesky factorization for positive definite matrices. The result obtained below has a direct counterpart for lower triangular transformations if we reverse the ordering of the state variables $(x_1, \ldots, x_n)$.

The Jacobian of the mapping Φ is

$$\frac{\partial \Phi}{\partial x} = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & * & \cdots & * \\ 0 & \frac{\partial \phi_2}{\partial x_2} & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial \phi_n}{\partial x_n} \end{bmatrix}$$

where $*$ denotes terms of no particular interest. Since the mapping Φ is a diffeomorphism, the Jacobian is always nonsingular and the diagonal partial derivatives are nonzero, $\partial\phi_i/\partial x_i \neq 0$, $\forall x \in \mathbb{R}^n$. Therefore, these quantities are sign definite. Without loss of generality, we consider only Φ where the diagonal terms of $\dfrac{\partial \Phi}{\partial x}$ are positive.

Definition 3: An upper triangular diffeomorphism $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *positive* if the diagonal partial derivatives are positive, $\dfrac{\partial \phi_i}{\partial x_i} > 0$, $i = 1, \ldots, n$ for any $x \in \mathbb{R}^n$. ◇

The Cholesky factorization of a nonlinear radially unbounded function $V(x)$, where $x \in \mathbb{R}^n$, is defined as follows.

Definition 4: A radially unbounded $\mathcal{C}^\infty$ function $V(x)$ is said to admit a nonlinear Cholesky factorization if $V(x) = \Phi'(x)\Phi(x)$ for some global positive upper triangular diffeomorphism with $\Phi(0) = 0$. ◇

It will be shown later that the nonlinear Cholesky factorization is unique within the class of positive diffeomorphisms.

Given a positive upper triangular state transformation $z = \Phi(x)$, it is still quite difficult to obtain the inverse transformation of Φ. To simplify this, we introduced the notion of the simple upper triangular diffeomorphism.

Definition 5: An upper triangular diffeomorphism $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ is said to be *simple* if $\dfrac{\partial \phi_i}{\partial x_i} \equiv 1$, $i = 1, \ldots, n$. Then, the individual coordinate map is given by $\phi_i(x_i, \ldots, x_n) = x_i - \alpha_i(x_{i+1}, \ldots, x_n)$, $i = 1, \ldots, n$, where α_i, $i = 1, \ldots, n$ (we have set $\alpha_n = 0$) are arbitrary smooth functions such that $\alpha_i(0) = 0$. ◇

There exists an equivalence relationship between the positive upper triangular diffeomorphisms and simple ones in terms of yet another factorization.

Proposition 1: An upper triangular diffeomorphism Φ is positive if and only if it can be uniquely factored as $\Phi = \Delta\Phi_s$, where Φ_s is a simple upper triangular diffeomorphism; and Δ is a diagonal matrix

$$\Delta = \text{diag}\,\{\delta_1(x_1, \ldots, x_n), \delta_2(x_2, \ldots, x_n), \ldots, \delta_n(x_n)\} \tag{7}$$

such that, for each $i = 1, \ldots, n$,

$$\delta_i(x_i, \ldots, x_n) > 0 \tag{8}$$

$$\pi_{\delta i}(x_i, \ldots, x_n) := \delta_i(x_i, \ldots, x_n) + (x_i - \alpha_i(x_{i+1}, \ldots, x_n))\frac{\partial \delta_i}{\partial x_i}(x_i, \ldots, x_n) > 0 \tag{9}$$

$$\lim_{|x_i| \to \infty} |(x_i - \alpha_i(x_{i+1}, \ldots, x_n))\delta_i(x_i, \ldots, x_n)| = \infty \tag{10}$$

for all $x \in \mathbb{R}^n$.

Proof: The sufficiency part is straightforward by observing the diagonal of the Jacobian of $\Delta\Phi_s$ is exactly diagonal$\{\pi_{\delta 1}, \ldots, \pi_{\delta n}\}$, which is nonsingular. Furthermore, the unboundedness condition (10) implies that the diffeomorphism Φ is global.

For the necessity part, we observe that $\phi_i(x_i, \ldots, x_n) = 0$, for any $i = 1, \ldots, n$, has a unique root $x_i = \alpha_i(x_{i+1}, \ldots, x_n)$ when $i = 1, \ldots, n-1$, or $x_n = 0 =: \alpha_n$. The functions $\alpha_1, \ldots, \alpha_n$, are smooth by the positiveness of $\dfrac{\partial \phi_i}{\partial x_i}$. This uniquely defines a simple upper triangular diffeomorphism Φ_s:

$$\eta = \Phi_s(x) = \begin{bmatrix} x_1 - \alpha_1(x_2, \ldots, x_n) \\ \vdots \\ x_n \end{bmatrix}$$

Then, the positive and smooth diagonal matrix Δ is defined by

$$\delta_i(x_i,\dots,x_n) := \frac{\phi_i(x_i,\dots,x_n)}{x_i-\alpha_i(x_{i+1},\dots,x_n)},\ i=1,\dots,n.$$

Evaluating $\delta_i + \eta_i \frac{\partial \delta_i}{\partial x_i} = \frac{\partial \phi_i}{\partial x_i} > 0$, $i=1,\dots,n$, where $\eta_i = x_i - \alpha_1(x_{i+1},\dots,x_n)$, proves that Δ satisfies (9). The property (10) is satisfied since $\lim_{|x_i|\to\infty} |\phi_i| = \infty$. This completes the proof.

To obtain the nonlinear Cholesky factorization of $V(x)$ we will go through n steps of a recursive construction. As shown in the following lemma, the objective of each step is to find one coordinate transformation.

Lemma 1: Consider a nonlinear function $V(\boldsymbol{a},\boldsymbol{b})$, where $\boldsymbol{a}$ is a scalar and $\boldsymbol{b}$ is a vector. Assume the following conditions hold:

C1*: V is $\mathcal{C}^\infty$ in all of its arguments.*

C2*: There is a constant $C \geq 0$ such that, for each fixed $\boldsymbol{b}$, $V(\boldsymbol{a},\boldsymbol{b}) \geq C$, $\forall \boldsymbol{a} \in \mathbb{R}$.*

C3*: For each fixed $\boldsymbol{b}$, V is radially unbounded in $\boldsymbol{a}$, i. e., $\lim_{|\boldsymbol{a}|\to\infty} V(\boldsymbol{a},\boldsymbol{b}) = +\infty$.*

Then, there exists a unique mapping $\phi(\boldsymbol{a},\boldsymbol{b})$ such that

1. ϕ is $\mathcal{C}^\infty$ in all of its the arguments.

2. For each fixed $\boldsymbol{b}$, $\frac{\partial \phi}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}) > 0$, $\forall \boldsymbol{a}$.

3. For each fixed $\boldsymbol{b}$, ϕ maps $\mathbb{R}$ onto $\mathbb{R}$, i. e., $\lim_{|\boldsymbol{a}|\to\infty} |\phi(\boldsymbol{a},\boldsymbol{b})| = \infty$.

4. The function V can be decomposed as

$$V(\boldsymbol{a},\boldsymbol{b}) = \hat{V}(\boldsymbol{b}) + \phi^2(\boldsymbol{a},\boldsymbol{b}) \tag{11}$$

for some $\mathcal{C}^\infty$ function $\hat{V}$, such that $\hat{V}(\boldsymbol{b}) \geq C$ for any $\boldsymbol{b}$.
if and only if the following condition holds:

E1*: For each fixed $\boldsymbol{b}$, $\frac{\partial V}{\partial \boldsymbol{a}} = 0$ has a unique root $\boldsymbol{a} = \alpha(\boldsymbol{b})$, which is simple, i. e.,*

$$\frac{\partial V}{\partial \boldsymbol{a}}(\alpha(\boldsymbol{b}),\boldsymbol{b}) \equiv 0; \qquad \frac{\partial^2 V}{\partial \boldsymbol{a}^2}(\alpha(\boldsymbol{b}),\boldsymbol{b}) \neq 0 \qquad \forall \boldsymbol{b}$$

Proof: To prove sufficiency, we let conditions **E1**, **C1**–**C3** hold. Conditions **C2** and **C3** imply that, for each fixed $\boldsymbol{b}$, the function V has at least a local minimum in $\boldsymbol{a}$. This, coupled with the smoothness assumption **C1** and uniqueness assumption **E1**, implies that the the minimum is achieved at $\boldsymbol{a} = \alpha(\boldsymbol{b})$, so that $\frac{\partial^2 V}{\partial \boldsymbol{a}^2}(\alpha(\boldsymbol{b}),\boldsymbol{b}) > 0$, $\forall \boldsymbol{b}$. The nonsingularity of $\frac{\partial^2 V}{\partial \boldsymbol{a}^2}(\alpha(\boldsymbol{b}),\boldsymbol{b})$ guarantees that the root function $\alpha(\boldsymbol{b})$ is $\mathcal{C}^\infty$ by the implicit function theorem.

Let $\hat{V}$ and π be defined as

$$\hat{V}(\boldsymbol{b}) := V(\alpha(\boldsymbol{b}),\boldsymbol{b}) \tag{12}$$

$$(\boldsymbol{a}-\alpha(\boldsymbol{b}))\pi(\boldsymbol{a},\boldsymbol{b}) := \frac{\partial V}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}). \tag{13}$$

Then it follows from **C2** that $\hat{V}(\boldsymbol{b}) \geq C$ for any $\boldsymbol{b}$. The function π is well defined because $\boldsymbol{a} = \alpha(\boldsymbol{b})$ is the root of the $\mathcal{C}^\infty$ function $\frac{\partial V}{\partial \boldsymbol{a}}$. By **E1** and the fact that $\alpha(\boldsymbol{b})$ achieves the minimum, we have that $\partial V/\partial \boldsymbol{a} > 0$, $\forall \boldsymbol{a} > \alpha(\boldsymbol{b})$; $\partial V/\partial \boldsymbol{a} < 0$, $\forall \boldsymbol{a} < \alpha(\boldsymbol{b})$. Therefore, $\pi(\boldsymbol{a},\boldsymbol{b}) > 0$, $\forall (\boldsymbol{a},\boldsymbol{b})$. This allows the following definition of a $\mathcal{C}^\infty$ positive function $m(\boldsymbol{a},\boldsymbol{b})$:

$$V(\boldsymbol{a},\boldsymbol{b}) - V(\alpha(\boldsymbol{b}),\boldsymbol{b}) =: m(\boldsymbol{a},\boldsymbol{b})(\boldsymbol{a}-\alpha(\boldsymbol{b}))^2 \tag{14}$$

since

$$V(\boldsymbol{a},\boldsymbol{b}) - V(\alpha(\boldsymbol{b}),\boldsymbol{b}) = \int_{\alpha(\boldsymbol{b})}^{\boldsymbol{a}} (s-\alpha(\boldsymbol{b}))\pi(s,\boldsymbol{b})\,ds. \tag{15}$$

The mapping ϕ is then defined by

$$\phi(\boldsymbol{a},\boldsymbol{b}) := (\boldsymbol{a}-\alpha(\boldsymbol{b}))\delta(\boldsymbol{a},\boldsymbol{b});\ \ \delta(\boldsymbol{a},\boldsymbol{b}) := \sqrt{m(\boldsymbol{a},\boldsymbol{b})} \tag{16}$$

Statement 1 follows from the definition of ϕ and δ. Statement 3 holds by the assumption **C3** and the definition of ϕ, while statement 4 holds by the definition of $\hat{V}$ and ϕ.

For statement 2, we evaluate the partial derivatives:

$$\begin{aligned}
\frac{\partial \phi}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}) &= \delta(\boldsymbol{a},\boldsymbol{b}) + (\boldsymbol{a}-\alpha(\boldsymbol{b}))\frac{\partial \delta}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}) =: \pi_\delta(\boldsymbol{a},\boldsymbol{b})\\
\frac{\partial V}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}) &= \frac{\partial}{\partial \boldsymbol{a}}(\delta^2(\boldsymbol{a},\boldsymbol{b})(\boldsymbol{a}-\alpha(\boldsymbol{b}))^2)\\
&= 2(\boldsymbol{a}-\alpha(\boldsymbol{b}))\delta(\boldsymbol{a},\boldsymbol{b})(\delta(\boldsymbol{a},\boldsymbol{b})\\
&\qquad +(\boldsymbol{a}-\alpha(\boldsymbol{b}))\frac{\partial \delta}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}))\\
&= (\boldsymbol{a}-\alpha(\boldsymbol{b}))\pi(\boldsymbol{a},\boldsymbol{b})
\end{aligned}$$

Hence, the identity: $\frac{\partial \phi}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}) = \frac{\pi(\boldsymbol{a},\boldsymbol{b})}{2\delta(\boldsymbol{a},\boldsymbol{b})} > 0$, $\forall (\boldsymbol{a},\boldsymbol{b})$, which proves statement 2.

Next, we show that the mapping $\phi(\boldsymbol{a},\boldsymbol{b})$ is unique, which further implies the function $\hat{V}$ is unique. We will prove this by contradiction. Suppose there exists another mapping ϕ_*, which corresponds to a function $\hat{V}_*$, that also satisfies the statements 1–4, and $\phi(\boldsymbol{a}_0,\boldsymbol{b}_0) \neq \phi_*(\boldsymbol{a}_0,\boldsymbol{b}_0)$, for some $(\boldsymbol{a}_0,\boldsymbol{b}_0)$. Because $V = \hat{V} + \phi^2 = \hat{V}_* + \phi_*^2$, we have *either* (a) $\hat{V}(\boldsymbol{b}_0) \neq \hat{V}_*(\boldsymbol{b}_0)$ *or* (b) $\phi(\boldsymbol{a}_0,\boldsymbol{b}_0) = -\phi_*(\boldsymbol{a}_0,\boldsymbol{b}_0) \neq 0$.

In case of (b), we observe the following equality:

$$\begin{aligned}
\frac{\partial V}{\partial \boldsymbol{a}}(\boldsymbol{a}_0,\boldsymbol{b}_0) &= 2\phi(\boldsymbol{a}_0,\boldsymbol{b}_0)\frac{\partial \phi}{\partial \boldsymbol{a}}(\boldsymbol{a}_0,\boldsymbol{b}_0)\\
&= 2\phi_*(\boldsymbol{a}_0,\boldsymbol{b}_0)\frac{\partial \phi_*}{\partial \boldsymbol{a}}(\boldsymbol{a}_0,\boldsymbol{b}_0)
\end{aligned}$$

This is a contradiction because both $\frac{\partial \phi}{\partial \boldsymbol{a}}(\boldsymbol{a}_0,\boldsymbol{b}_0)$ and $\frac{\partial \phi_*}{\partial \boldsymbol{a}}(\boldsymbol{a}_0,\boldsymbol{b}_0)$ are positive by statement 2.

In case of (a), we conclude that $\phi(\boldsymbol{a},\boldsymbol{b}_0) \neq \phi_*(\boldsymbol{a},\boldsymbol{b}_0)$, $\forall \boldsymbol{a}$. In particular, $0 = \phi(\alpha(\boldsymbol{b}_0),\boldsymbol{b}_0) \neq \phi_*(\alpha(\boldsymbol{b}_0),\boldsymbol{b}_0)$. This leads to the contradiction:

$$\begin{aligned}
0 &= 2\phi(\alpha(\boldsymbol{b}_0),\boldsymbol{b}_0)\frac{\partial \phi}{\partial \boldsymbol{a}}(\alpha(\boldsymbol{b}_0),\boldsymbol{b}_0) = \frac{\partial V}{\partial \boldsymbol{a}}(\alpha(\boldsymbol{b}_0),\boldsymbol{b}_0)\\
&= 2\phi_*(\alpha(\boldsymbol{b}_0),\boldsymbol{b}_0)\frac{\partial \phi_*}{\partial \boldsymbol{a}}(\alpha(\boldsymbol{b}_0),\boldsymbol{b}_0) \neq 0
\end{aligned}$$

where the last inequality follows from statement 2. Consequently, the hypothesis is not valid, and the mapping ϕ is uniquely defined. This completes the sufficiency proof.

To prove necessity, we assume that there exists a mapping ϕ satisfying the statements 1–4. Then, $\frac{\partial V}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b}) = 2\phi(\boldsymbol{a},\boldsymbol{b})\frac{\partial \phi}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b})$. It is concluded that $\frac{\partial V}{\partial \boldsymbol{a}}(\boldsymbol{a},\boldsymbol{b})$ vanishes if and only if $\phi(\boldsymbol{a},\boldsymbol{b}) = 0$, because of statement 2. Again by statement 2, ϕ is strictly increasing in $\boldsymbol{a}$. Coupled with radially unboundedness statement 3, we conclude that there is a unique root $\boldsymbol{a} = \alpha(\boldsymbol{b})$ for the equation $\phi(\boldsymbol{a},\boldsymbol{b}) = 0$. Along with conditions **C2** and **C3**, this implies that, for each fixed $\boldsymbol{b}$, the function V has a unique minimum in $\boldsymbol{a}$. The result **E1** follows from $\frac{\partial^2 V}{\partial \boldsymbol{a}^2}(\alpha(\boldsymbol{b}),\boldsymbol{b}) = 2\left(\frac{\partial \phi}{\partial \boldsymbol{a}}(\alpha(\boldsymbol{b}),\boldsymbol{b})\right)^2 > 0$, and consequently, the necessity part of the lemma.

We observe that the construction of the mapping ϕ in the above lemma involves the following steps. First, we solve for the unique smooth root function $\alpha(\boldsymbol{b})$, which is guaranteed to exist by the condition **E1**. Next, we can define the function $\hat{V}$ by (12). Then, mapping ϕ is given by (16).

Now, we are in the position to present the n step recursive construction for the nonlinear Cholesky factorization of a given radially unbounded and smooth function V.

Step 1: For notational consistency we let $V(x) =: V_1(x)$. Since $V(x)$ is $\mathcal{C}^\infty$ and radially unbounded, the conditions **C1**–**C3** of Lemma 1 are satisfied for V, with $\boldsymbol{a} = x_1$ and $\boldsymbol{b} = (x_2, \dots, x_n)'$.

As delineated in Lemma 1, in order for a smooth transformation $\phi_1(x)$ to exist we make the following assumption:
B1*:* For any fixed $n-1$ dimensional vector $(x_2, \dots, x_n)'$, the algebraic equation $\frac{\partial V_1}{\partial x_1} = 0$ has a unique root $x_1 = \alpha_1(x_2, \dots, x_n)$, which is simple,

$$\frac{\partial^2 V_1}{\partial x_1^2}(\alpha_1(x_2, \dots, x_n), x_2, \dots, x_n) \neq 0.$$

Under this assumption, there exists a smooth mapping $z_1 = \phi_1(x_1, \dots, x_n)$ and a smooth nonlinear function $V_2(x_2, \dots, x_n)$ such that

R1.1*:* For any fixed $(x_2, \dots, x_n)'$, $\frac{\partial \phi_1}{\partial x_1}(x_1, \dots, x_n) > 0$, $\forall x_1 \in \mathbb{R}$.
R1.2*:* For each fixed $(x_2, \dots, x_n)'$, ϕ_1 maps $\mathbb{R}$ onto $\mathbb{R}$, i. e., $\lim_{|x_1|\to\infty} |\phi_1(x_1, \dots, x_n)| = \infty$.
R1.3*:* The function V_1 can be decomposed as

$$V_1(x_1, \dots, x_n) = V_2(x_2, \dots, x_n) + \phi_1^2(x_1, \dots, x_n).$$

We note here that the assumption **B1** is also necessary for the existence of the transformation ϕ_1 as proved in Lemma 1. Because the zero vector is the global minimum of the function V, we have:
R1.4*:* The mapping ϕ_1 is zero preserving, $\phi_1(0) = 0$.

Because V_1 is radially unbounded and smooth, we can conclude that the nonlinear function V_2 is also radially unbounded (when restricted to the $n-1$ dimensional subspace $(x_2, \dots, x_n)'$) and smooth, based on the equality: $V_1(\alpha_1(x_2, \dots, x_n), x_2, \dots, x_n) \equiv V_2(x_2, \dots, x_n)$. This nonlinear function V_2 is uniquely defined by Lemma 1 and, hence, the construction in the subsequent steps is *independent* of the construction at this step.

Step i, $2 \le i \le n$: We assume inductively that the mappings $\phi_j(x_j, \dots, x_n)$ are constructed from the previous steps $1, \dots, i-1$, and we are given a radially unbounded and smooth nonlinear function $V_i(x_i, \dots, x_n)$, that satisfies $V(x) = \phi_1^2 + \dots + \phi_{i-1}^2 + V_i(x_i, \dots, x_n)$.

The radially unboundedness and smoothness of V_i implies conditions **C1**–**C3** of Lemma 1. We make the following assumption, in order for a smooth mapping $\phi_i(x)$ to exist,
Bi*:* For any fixed $n-i$ dimensional vector $(x_{i+1}, \dots, x_n)'$, the algebraic equation $\frac{\partial V_i}{\partial x_i} = 0$ has a unique root $x_i = \alpha_i(x_{i+1}, \dots, x_n)$, which is simple,

$$\frac{\partial^2 V_i}{\partial x_i^2}(\alpha_i(x_{i+1}, \dots, x_n), x_{i+1}, \dots, x_n) \neq 0.$$

Under this assumption, there exists a smooth transformation $z_i = \phi_1(x_i, \dots, x_n)$ and a smooth nonlinear function $V_{i+1}(x_{i+1}, \dots, x_n)$ such that

Ri**.1***:* For any fixed $(x_{i+1}, \dots, x_n)'$, $\frac{\partial \phi_i}{\partial x_i}(x_i, \dots, x_n) > 0$, $\forall x_i \in \mathbb{R}$.
Ri**.2***:* For each fixed $(x_i, \dots, x_n)'$, ϕ_i maps $\mathbb{R}$ onto $\mathbb{R}$, i. e., $\lim_{|x_i|\to\infty} |\phi_i(x_i, \dots, x_n)| = \infty$.
Ri**.3***:* The function V_i can be decomposed as $V_i(x_i, \dots, x_n) = V_{i+1}(x_{i+1}, \dots, x_n) + \phi_i^2(x_i, \dots, x_n)$.
Again, the assumption **B**i is necessary for the existence of the transformation ϕ_i as shown in Lemma 1. The mapping ϕ_i is also zero preserving, since V_i admits a global minimum at $(x_i, \dots, x_n) = (0, \dots, 0)$.
Ri**.4***:* The mapping ϕ_i is zero preserving, i. e., $\phi_i(0) = 0$.

The following equality holds, as prescribed by the construction in Lemma 1:

$$V_i(\alpha_i(x_{i+1}, \dots, x_n), x_{i+1}, \dots, x_n) \equiv V_{i+1}(x_{i+1}, \dots, x_n)$$

Therefore, from the fact that V_i is radially unbounded and smooth, we can conclude that the nonlinear function V_{i+1} is also radially unbounded and smooth. Furthermore, this nonlinear function V_{i+1} is uniquely defined by Lemma 1 and, hence, the construction of the subsequent steps is *independent* of the construction at this step.

Furthermore, at the nth step, we have $V_{n+1} \equiv 0$, which is equivalent to

$$V(x) \equiv \phi_1^2 + \dots + \phi_n^2 \tag{17}$$

This completes the n-step recursive construction of the nonlinear Cholesky factorization, which is summarized as follows:

Theorem 1: A positive definite, radially unbounded and smooth function $V(x)$ *with* $x \in \mathbb{R}^n$ *has a unique nonlinear*

Cholesky factorization $z = \Phi(x)$ if and only if the assumptions **B1**–**B***n are satisfied. Furthermore, under the same assumptions, $V(x)$ can be factored as*

$$V(x) = \Phi_s'(x)\Delta(x)\Delta(x)\Phi_s(x) \tag{18}$$

where $\Phi_s(x)$ a simple upper triangular diffeomorphism and Δ is a diagonal matrix that satisfies (7)–(10).

Proof: To prove sufficiency let assumptions **B1**–**B***n* hold. By the recursive construction in Lemma 1 there exists triangular mappings $\Phi = (\phi_1, \ldots, \phi_n)'$ such that the conditions **R1.1**–**R1.4**, ..., **R***n***.1**–**R***n***.4** are satisfied and (17) holds. These conditions imply that Φ is a positive upper triangular diffeomorphism.

To prove the necessity part of the theorem, we assume that V admits a nonlinear Cholesky factorization. We first consider the mapping ϕ_1 and define $V_2(x_2, \ldots, x_n) = z_2^2 + \ldots + z_n^2$. Since V is radially unbounded and smooth, it satisfies conditions **C1**–**C3** of Lemma 1. The definition of the nonlinear Cholesky factorization implies that ϕ_1 and V_2 satisfy the statements 1–4 of Lemma 1 so that **E1** holds, which is equivalent to **B1**. Therefore, **B1** is necessary for the existence of a nonlinear Cholesky factorization. When **B1** is satisfied, ϕ_1 is uniquely defined, and so is V_2.

Arguing inductively for $i = 2, \ldots, n$, we determine for each i, that the function V_i is independent of the construction of the previous mapping. The function V_i is radially unbounded and smooth. Then, by Lemma 1, the unique mapping ϕ_i exists only if assumption **B***i* holds. Therefore, assumptions **B1**–**B***n* are necessary for the existence of a nonlinear Cholesky factorization.

By the uniqueness of the mappings ϕ_i, at each step $i = 1, \ldots, n$ the diffeomorphism Φ is uniquely defined whenever it exists.

The second part of the theorem follows directly from the Proposition 1.

The motivation for the introduction of the simple diffeomorphism (18) is that the computation of the inverse mapping Φ_s^{-1} involves only substitutions.

Next, we consider two illustrative examples. The first example is for quadratic functions. The second example involves a 4th order polynomial and shows that the ordering of the coordinate variables $(x_1, \ldots, x_n)$ in Definition 4 is critical for the existence of the factorization.

Example 1. For $V = x'Px$, with positive definite P, the nonlinear Cholesky factorization mapping Φ is $z = Ux$, where the upper triangular matrix U is the Cholesky factor of the matrix P, $P = U'U$.

Example 2. Smooth and radially unbounded function $V(x_1, x_2) = (x_1 + x_2^2)^2 + x_2^2$ admits a nonlinear Cholesky factorization in the ordering (x_1, x_2) with Φ given by $\begin{bmatrix} z_1 & z_2 \end{bmatrix}' := \begin{bmatrix} x_1 + x_2^2 & x_2 \end{bmatrix}'$. If we switch the ordering, we obtain the function $W(x_1, x_2) := V(x_2, x_1) = (x_2 + x_1^2)^2 + x_1^2$. Applying our result, we observe that the equation $\dfrac{\partial W}{\partial x_1} = 4x_1^3 + 4x_1x_2 + 2x_1 = 0$ has three distinct roots when $x_2 = -1$. By Theorem 4 the function W does not admit a nonlinear Cholesky factorization.

A question of practical interest is whether assumptions **B1**–**B***n* can be checked without actually carrying out the recursive construction. It turns out that this is possible when V is strictly convex. To prove this proposition, we need the following corollary to Lemma 1.

Corollary 1: For a nonlinear function $V(\boldsymbol{a}, \boldsymbol{b})$ that satisfies **C1**–**C3** *assume that*

C4*: V is strictly convex so that $\dfrac{\partial^2 V}{\partial(\boldsymbol{a}, \boldsymbol{b})^2}(\boldsymbol{a}, \boldsymbol{b}) > 0, \forall(\boldsymbol{a}, \boldsymbol{b})$.*

Then, **E1** *is satisfied and statements 1–4 of the Lemma 1 hold. Furthermore, the nonlinear function $\hat{V}(\boldsymbol{b})$ of the decomposition (11) is strictly convex.*

Proof: To show that **E1** holds under the convexity assumption, we note that **C4** implies that V is strictly convex in $\boldsymbol{a}$, so that there is at most one root for the equation $\dfrac{\partial V}{\partial \boldsymbol{a}} = 0$. By **C2** and **C3**, there exists at least one root because V admits a global minimum for any fixed $\boldsymbol{b}$. Therefore, this root is unique and simple because $\dfrac{\partial^2 V}{\partial \boldsymbol{a}^2} > 0$. Hence, **E1** holds.

To prove that $\dfrac{\partial^2 \hat{V}}{\partial \boldsymbol{b}^2}(\boldsymbol{b}) > 0 \; \forall \boldsymbol{b}$, we show that $p' \frac{\partial^2 \hat{V}}{\partial \boldsymbol{b}^2}(\boldsymbol{b}) p > 0$, for any nonzero vector p and any fixed vector $\boldsymbol{b}$. For each $\boldsymbol{b}$ and nonzero p, there exists a scalar δ such that $\bar{p} := \begin{bmatrix} \delta & p' \end{bmatrix}'$ satisfies $\dfrac{\partial \phi}{\partial(\boldsymbol{a}, \boldsymbol{b})}(\alpha(\boldsymbol{b}), \boldsymbol{b})\bar{p} = 0$, where $\boldsymbol{a} = \alpha(\boldsymbol{b})$ is the unique root of $\phi(\boldsymbol{a}, \boldsymbol{b}) = 0$. Thus, we have

$$\begin{aligned} 0 &< \bar{p}' \frac{\partial^2 V}{\partial(\boldsymbol{a}, \boldsymbol{b})^2}(\alpha(\boldsymbol{b}), \boldsymbol{b})\bar{p} \\ &= p' \frac{\partial^2 \hat{V}}{\partial \boldsymbol{b}^2}(\boldsymbol{b}) p \\ &\quad + 2\bar{p}' \left(\frac{\partial \phi}{\partial(\boldsymbol{a}, \boldsymbol{b})}(\alpha(\boldsymbol{b}), \boldsymbol{b}) \right)' \left(\frac{\partial \phi}{\partial(\boldsymbol{a}, \boldsymbol{b})}(\alpha(\boldsymbol{b}), \boldsymbol{b}) \right) \bar{p} \\ &= p' \frac{\partial^2 \hat{V}}{\partial \boldsymbol{b}^2}(\boldsymbol{b}) p. \end{aligned}$$

This proves the strict convexity of $\hat{V}(\boldsymbol{b})$.

For a strictly convex function V, a nonlinear Cholesky factorization exists for any ordering of the coordinate variables $(x_1, \ldots, x_n)$. This is an consequence of the following corollary to Theorem 1, obtained by a recursive application of Corollary 1.

Corollary 2: Consider a positive definite, radially unbounded and smooth nonlinear function $V(x)$, where $x \in \mathbb{R}^n$. The function $V(x)$ admits a nonlinear Cholesky factorization $z = \Phi(x)$, as defined in Definition 4, if it is strictly convex, $\partial^2 V/\partial x^2 > 0$, $\forall x \in \mathbb{R}^n$.

So far, we have presented a procedure to obtain a nonlinear Cholesky factorization for a given function. For the nonlinear robust control problem formulated in Section 2, the value function V is known only approximately via its Taylor series expansion. In the remainder of this section the objective is to derive a factorizable function W that matches the function V up to any given order of accuracy. We assume that the quadratic approximation $V^{[2]}(x) = x'Px$ of V is positive definite. Our goal is to find, for any positive integer m, a radially unbounded and smooth function $W(x)$ which has a nonlinear Cholesky factorization and such that $W^{[m]}(x) = V^{[m]}(x)$. An approximate nonlinear Cholesky factorization can be easily shown to exist by adding higher order terms V_Δ to make $W = V^{[m]} + V_\Delta$ strictly convex. However, this simplistic convexification approach would cause serious computational difficulties. Instead, we present an explicit construction for the approximate function W and the factor $\Phi(x)$ with milder nonlinear growth than the factor obtained via convexification.

Our recursive construction consists of n steps. The following illustrates a typical step. Consider a polynomial function $V(\boldsymbol{a}, \boldsymbol{b})$ of order m, where $\boldsymbol{a}$ is a scalar and $\boldsymbol{b}$ is a vector. Assume that $V^{[2]}(\boldsymbol{a}, \boldsymbol{b}) = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b}' \end{bmatrix} P \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b}' \end{bmatrix}'$ and $P > 0$. We partition the matrix P conformal with $(\boldsymbol{a}, \boldsymbol{b})$,

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}' & P_{22} \end{bmatrix}$$

where P_{11} is a positive scalar. Our goal is to find $\phi(\boldsymbol{a}, \boldsymbol{b})$ which satisfies the statements 1–4 of Lemma 1, and $\hat{V}(\boldsymbol{b})$ in (11) is such that $\hat{V}^{[2]}(\boldsymbol{b}) = \boldsymbol{b}'\hat{P}\boldsymbol{b}$ with $\hat{P} > 0$.

First, we evaluate the root of $\dfrac{\partial V}{\partial \boldsymbol{a}} = 0$ approximately. Let $V_p(\boldsymbol{a}, \boldsymbol{b}) := \dfrac{\partial V^{[m]}}{\partial \boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{b})$. Then, V_p is an $(m-1)$th order polynomial and $V_p^{[1]} = 2 \begin{bmatrix} \boldsymbol{a} & \boldsymbol{b}' \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \end{bmatrix}'$. This implies that the equation $V_p^{[1]}(\boldsymbol{a}, \boldsymbol{b}) = 0$ admits a unique root $\boldsymbol{a} = -\dfrac{1}{P_{11}} P_{12}\boldsymbol{b}$.

Although, there may be several roots to the equation

$$\frac{\partial V^{[m]}}{\partial \boldsymbol{a}} = 0 \tag{19}$$

we are only interested in the root around $\boldsymbol{a} = -\dfrac{1}{P_{11}} P_{12}\boldsymbol{b}$. The Taylor series expansion of the root function $\alpha(\boldsymbol{b})$ up to $(m-1)$st order can be computed by equating homogeneous terms of the same order. Starting with $\alpha^{[1]}(\boldsymbol{b}) = -\dfrac{1}{P_{11}} P_{12}\boldsymbol{b}$, we obtain higher-order terms

$$\alpha_{[i]}(\boldsymbol{b}) = -\frac{1}{2P_{11}} \left(\sum_{j=2}^{i} V_{p[j]}(\alpha^{[i-1]}(\boldsymbol{b}), \boldsymbol{b}) \right)_{[i]}, \tag{20}$$

for $i = 2, \ldots, m-1$. Therefore, we have

$$\left(V_p(\alpha^{[m-1]}(\boldsymbol{b}), \boldsymbol{b}) \right)^{[m-1]} = 0 \tag{21}$$

Next, by adding higher order correction terms we enforce $\alpha^{[m-1]}(\boldsymbol{b})$ to be the unique root of (19), and to satisfy **C1**–**C3** and **E1**. Toward this end, we let $\eta := \boldsymbol{a} - \alpha^{[m-1]}(\boldsymbol{b})$ and define a function $\bar{\Pi}(\eta, \boldsymbol{b})$ through

$$\eta\bar{\Pi}(\eta, \boldsymbol{b}) := V_p(\eta + \alpha^{[m-1]}(\boldsymbol{b}), \boldsymbol{b}) - V_p(\alpha^{[m-1]}(\boldsymbol{b}), \boldsymbol{b}) \tag{22}$$

From (21), we get

$$\left((\boldsymbol{a} - \alpha^{[m-1]}(\boldsymbol{b}))\bar{\Pi}(\boldsymbol{a} - \alpha^{[m-1]}(\boldsymbol{b}), \boldsymbol{b}) \right)^{[m-1]} = \frac{\partial V^{[m]}}{\partial \boldsymbol{a}}(\boldsymbol{a}, \boldsymbol{b})$$

and

$$\left(\int_0^\eta (s\bar{\Pi}(s, \boldsymbol{b}))^{[m-1]} \, ds \bigg|_{\eta = \boldsymbol{a} - \alpha^{[m-1]}(\boldsymbol{b})} \right)^{[m]} = V^{[m]}(\boldsymbol{a}, \boldsymbol{b}) - (V(\alpha^{[m-1]}(\boldsymbol{b}), \boldsymbol{b}))^{[m]} \tag{23}$$

Next, we proceed to modify the high order polynomial terms of the function $\bar{\Pi}$ to guarantee that it is larger than a positive constant for all $(\boldsymbol{a}, \boldsymbol{b})$ so that **C2**, **C3**, and **E1** are satisfied.

To modify the high-order terms of $\bar{\Pi}$ in the $(\eta, \boldsymbol{b})$ coordinates we need a polynomial $\bar{\Pi}_s(\eta, \boldsymbol{b})$ of order $(m-2)$ such that

$$\left(\bar{\Pi}_s^2(\eta, \boldsymbol{b}) \right)^{[m-2]} + C = \bar{\Pi}^{[m-2]}(\eta, \boldsymbol{b}) \tag{24}$$

for some constant $C > 0$. From $\bar{\Pi}(0,0) = 2P_{11}$, for the root function $\bar{\Pi}_s$ to exist around the origin, we must choose C to be a fraction of P_{11}, say $\dfrac{1}{m}P_{11}$. The substitution of $\bar{\Pi}_s$ into (24), where $\bar{\Pi}_s$ is expanded as

$$\bar{\Pi}_s(\eta, \boldsymbol{b}) = \sum_{i=0}^{m-2} \bar{\Pi}_{s[i]}(\eta, \boldsymbol{b}), \tag{25}$$

and equating polynomials of the same order, results in

$$\begin{aligned} \bar{\Pi}_{s[0]} &= \sqrt{2P_{11} - C} \\ \bar{\Pi}_{s[i]} &= \frac{1}{2\sqrt{2P_{11} - C}} (\bar{\Pi}_{[i]} - \sum_{j=1}^{i-1} \bar{\Pi}_{s[j]}\bar{\Pi}_{s[i-j]}) \end{aligned}$$

for $i = 1, \ldots, m-2$.

Therefore, by (23) and (24), we have

$$\begin{aligned} V^{[m]}(\boldsymbol{a}, \boldsymbol{b}) = & \left(\int_0^\eta s(\bar{\Pi}_s^2(s, \boldsymbol{b}) + C) \, ds \bigg|_{\eta = \boldsymbol{a} - \alpha^{[m-1]}(\boldsymbol{b})} \right)^{[m]} \\ & + \left(V(\alpha^{[m-1]}(\boldsymbol{b}), \boldsymbol{b}) \right)^{[m]} \end{aligned} \tag{26}$$

For large values of m, $\bar{\Pi}_s^2(\boldsymbol{a}, \boldsymbol{b})$ and $\alpha^{[m-1]}(\boldsymbol{b})$ involve high-order polynomials that grow rapidly in magnitude as x deviates from the origin. We alleviate this difficulty with two smooth scaling functions $S(\boldsymbol{b})$ and $T(\boldsymbol{a}, \boldsymbol{b})$, which satisfy

$$S^{[m-1]}(\boldsymbol{b}) \equiv 1 \tag{27}$$

$$T^{[m-2]}(\boldsymbol{a}, \boldsymbol{b}) \equiv 1, \qquad T(\boldsymbol{a}, \boldsymbol{b}) > 0 \quad \forall (\boldsymbol{a}, \boldsymbol{b}). \tag{28}$$

Using these scaling functions in (26), we have the following identity

$$\begin{aligned} V^{[m]}(\boldsymbol{a}, \boldsymbol{b}) = & \left(V(\alpha^{[m-1]}(\boldsymbol{b}), \boldsymbol{b}) \right)^{[m]} \\ & + \left(\int_0^\eta s(\bar{\Pi}_s^2(s, \boldsymbol{b}) + C)\bar{T}^2(s, \boldsymbol{b}) \, ds \bigg| \right)^{[m]} \end{aligned} \tag{29}$$

where $\eta = \boldsymbol{a} - \alpha^{[m-1]}(\boldsymbol{b})S(\boldsymbol{b})$. We note that multiplier S has been removed in the term $\left(V(\alpha^{[m-1]}(\boldsymbol{b}),\boldsymbol{b})\right)^{[m]}$ because $\left(V(\alpha^{[m-1]}(\boldsymbol{b}),\boldsymbol{b})\right)^{[m]} = \left(V(\alpha^{[m-1]}(\boldsymbol{b})S(\boldsymbol{b}),\boldsymbol{b})\right)^{[m]}$.

Define $\bar{\delta}(\eta,\boldsymbol{b}) = \sqrt{\frac{1}{\eta^2}\int_0^\eta s(\bar{\Pi}_s^2(s,\boldsymbol{b}) + C)\bar{T}^2(s,\boldsymbol{b})\,ds}$ and $\hat{V}(\boldsymbol{b}) = (V(\alpha^{[m-1]}(\boldsymbol{b}),\boldsymbol{b}))^{[m]}$. We note that $\bar{\delta}$ is a smooth positive function of $(\boldsymbol{a},\boldsymbol{b})$, and $\hat{V}$ is an mth order polynomial of $\boldsymbol{b}$. Then, $V^{[m]}(\boldsymbol{a},\boldsymbol{b}) = W^{[m]}(\boldsymbol{a},\boldsymbol{b})$, where

$$W(\boldsymbol{a},\boldsymbol{b}) := \left(\eta^2\bar{\delta}^2(\eta,\boldsymbol{b})\big|_{\eta=\boldsymbol{a}-\alpha^{[m-1]}(\boldsymbol{b})S(\boldsymbol{b})}\right)^{[m]} + \hat{V}(\boldsymbol{b}) \tag{30}$$

Since $\eta = 0$ is the unique root of $\eta(\bar{\Pi}_s^2(\eta,\boldsymbol{b}) + C)\bar{T}^2(\eta,\boldsymbol{b}) = 0$ and $\frac{\partial}{\partial\eta}(\eta(\bar{\Pi}_s^2(\eta,\boldsymbol{b}) + C)\bar{T}^2(\eta,\boldsymbol{b}))\Big|_{\eta=0} = (\bar{\Pi}_s^2(0,\boldsymbol{b}) + C)\bar{T}^2(0,\boldsymbol{b}) > 0$, condition **E1** is satisfied for the function W. By inspection, condition **C2** is satisfied. For condition **C3** to hold, it is sufficient that $\bar{T}^2 \geq \hat{C}\frac{1}{(1+\eta^2)(C + \bar{\Pi}_s^2(\eta,\boldsymbol{b}))}$, $\forall(\eta,\boldsymbol{b})$, for some positive constant $\hat{C}$.

With the scaling function S, we can shape the growth rate of the function $\hat{\alpha}(\boldsymbol{b}) := \alpha^{[m-1]}(\boldsymbol{b})S(\boldsymbol{b})$, which corresponds to the virtual control law in the backstepping design. With the scaling function $\bar{T}$, we can indirectly shape the growth rate of the function $\delta(\boldsymbol{a},\boldsymbol{b})$, which further determines the growth rate of the approximate value function and the virtual control law in the backstepping design. For induction purpose, we verify that the quadratic part of the function $\hat{V}$ is $\hat{V}^{[2]}(\boldsymbol{b}) = \boldsymbol{b}'\hat{P}\boldsymbol{b}$, where $\hat{P} = P_{22} - P_{12}'P_{11}^{-1}P_{12}$ is positive definite due to the positive definiteness of P.

By repeated application of the above process, we can prove the following theorem.

Theorem 2: Assume that a smooth function $V(x)$ with $x \in \mathbb{R}^n$ is locally quadratic, $V^{[2]}(x) = x'Px$, and P is positive definite. Then, for any desired matching order $m \geq 2$, there exists a radially unbounded and smooth function $W(x)$ which admits a nonlinear Cholesky factorization such that $W^{[m]}(x) = V^{[m]}(x)$ and $W = \Phi'\Phi = \Phi_s'\Delta\Delta\Phi_s$, where Φ_s, a simple upper triangular diffeomorphism, and Δ, a diagonal matrix, are recursively constructed by the above procedure.

When $V(x)$ is an analytic function and admits a Cholesky factorization $V = \Phi_{Vs}'\Delta_V\Delta_V\Phi_{Vs}$, then there exists an $\epsilon_0 > 0$ such that, $\forall|x| \leq \epsilon_0$, $\lim_{m\to\infty} W(x) = V(x)$, where the function W is constructed as above under the additional assumption that as $m \to \infty$ the scaling functions S and T, introduced at each step of the construction, converge to 1 on the set $\{x:\ |x| \leq \epsilon_0\}$.

Proof: The first part follows from the construction preceding the theorem. When V is positive definite and analytic, the functions Δ_V and Φ_{Vs} are analytic because the elements of Φ_{Vs} are roots of n analytic equations. By the analytic version of the implicit function theorem, they are also analytic under the necessary and sufficient conditions of Theorem 1.

For the second part of the theorem, we first consider the case when the scaling functions S and T are set to 1 at each step of recursive construction. Then, the simple diffeomorphism Φ_s is the truncated Taylor series approximation to Φ_{Vs} up to the order of $m-1$. By the analyticity of Φ_{Vs}, it converges to Φ_s as $m \to \infty$ within a radius of convergence. The nonlinear function $\bar{\Pi}_s$ in (25) is also the truncated Taylor series approximation for the function V. It converges, again due to the analyticity of Δ_V, as $m \to \infty$ within a radius of convergence. This further implies that Δ_{Vs} converges to Δ_V as $m \to \infty$ within the same radius of convergence. Hence, the function W converges to V as $m \to \infty$ within the common radius of convergence.

For the general case, when the scaling factors S and T converges to 1 as $m \to \infty$, we can conclude the same convergence result because of the fact that a product converges when all its factors converge.

In general, the function generated by the recursive procedure is not convex. Therefore the approximate function W includes lower order polynomial terms as compared with the approximation resulting from the simplistic convexification approach.

Equipped with this high-order matching result we proceed to present our design procedure for an inverse optimal control law that matches the optimal design up to any given order.

IV. Higher Order Matching Control Design

Applying the nonlinear Cholesky factorization result to the robust optimal control problem for system (1) and cost functional (2), we can prove that a backstepping procedure can produce the optimal solution as long as the value function solution is factorizable in the lower triangular fashion.

Theorem 3: Consider the nonlinear system (1) and cost functional (2) under Assumptions **A1** *and* **A2**. *Assume that there exists a smooth value function $V(x)$ satisfying the HJI equation (3), which has a nonlinear Cholesky factorization $V = \Phi_s'\Delta\Delta\Phi_s$ in the reverse order of state variables, $(x_n, \ldots, x_1)$. Then, the value function, minimax control law (4) and the corresponding worst-case disturbance can be obtained by a backstepping procedure.*

The proof is straightforward and can be found in [1].

Theorem 3 assumes the knowledge of the value function $V(x)$. Instead, we pursue the inverse optimal design with locally optimal matching as prescribed in Definitions 1 and 2.

Let $m \geq 1$ be the desired matching order. By a result of [8], we obtain the Taylor series expansion of V up to $(m+1)$st order. Using Theorem 2, we can find a function $W_l(x)$ that matches V up to $(m+1)$st order and can be factored as $W_l(x) = \Phi_s(x)'\Delta(x)\Delta(x)\Phi_s(x)$. Clearly, W_l satisfies the HJI equation (3) up to $(m+1)$th order. Introducing $z_l = \Phi_s(x)$, we further assume that the Φ_s and Δ are given by

$$\Phi_s(x) = \begin{bmatrix} x_1 \\ x_2 - \alpha_{l1}(x_1) \\ \vdots \\ x_n - \alpha_{l\,n-1}(x_1,\ldots,x_{n-1}) \end{bmatrix} \quad (31a)$$

$$\Delta(x) = \mathrm{diag}\{\delta_{l1}(x_1),\ldots,\delta_{ln}(x)\} \quad (31b)$$

where the subscript l indicates that this nonlinear function is to be matched in the backstepping design. System (1) in the z_l coordinate is

$$\begin{aligned} \dot z_{l1} &= z_{l2} + \bar{\hat f}_{l1}(z_{l1}) + \bar{\hat h}'_{l1}(z_{l1})w \\ \vdots &= \vdots \\ \dot z_{ln} &= u + \bar{\hat f}_{ln}(z_{l1},\ldots,z_{ln}) + \bar{\hat h}'_{ln}(z_{l1},\ldots,z_{ln})w \\ J_\gamma &= \int_0^\infty (\bar q_l(z_l) + \bar r_l(z_l)u^2 - \gamma^2 w'w)\,dt \end{aligned}$$

or, in compact form,

$$\dot z_l = \bar{\hat f}_l(z_l) + Bu + \bar{\hat H}_l(z_l)w.$$

Then, the function W_l can be expressed as $\bar W_l(z_l) = \sum_{i=1}^n \bar\delta_{li}^2 z_{li}^2$ and satisfies the HJI equation (3) up to $(m+1)$th order:

$$\left(\frac{\partial \bar W_l}{\partial z_l}\bar{\hat f}_l - \frac14 \frac{\partial \bar W_l}{\partial z_l}(B\bar r_l^{-1}B' - \frac{1}{\gamma^2}\bar{\hat H}_l\bar{\hat H}_l')\left(\frac{\partial \bar W_l}{\partial z_l}\right)' + \bar q_l\right)^{[m+1]} = 0 \quad (32)$$

As a consequence of this construction, we have $\bar\pi_i(z_{l\{1\}}^{\{i\}}) := \bar\delta_{li}^2 + z_{li}\bar\delta_{li}\frac{\partial \bar\delta_{li}}{\partial z_{li}} > 0$, for any z_l and all $i = 1,\ldots,n$. Setting $z_{l\,i+1} = \cdots = z_{ln} = 0$ yields the following approximate HJI equation for $\bar W_{li}$, for any $i = 1,\ldots,n-1$,

$$\left(\frac{\partial \bar W_{li}}{\partial z_{l\{1\}}^{\{i\}}}\bar{\hat f}_{l\{1\}}^{\{i\}} + \frac{1}{4\gamma^2}\frac{\partial \bar W_{li}}{\partial z_{l\{1\}}^{\{i\}}}\bar{\hat H}_{l\{1\}}^{\{i\}}\,\bar{\hat H}_{l\{1\}}^{\{i\}\,\prime}\,\frac{\partial \bar W_{li}}{\partial z_{l\{1\}}^{\{i\}}}' + \bar q_{l\{1\}}^{\{i\}}\right)^{[m+1]} = 0 \quad (33)$$

where $z_{l\{1\}}^{\{i\}} = \begin{bmatrix} z_{l1} & \ldots & z_{li}\end{bmatrix}'$ and

$$\begin{aligned} \bar W_{li} &:= \sum_{j=1}^i \bar\delta_{lj}^2 z_{lj}^2 \\ \bar q_{l\{1\}}^{\{i\}} &:= \bar q_l(z_{l1},\ldots,z_{li},0,\ldots,0) \\ \bar{\hat f}_{l\{1\}}^{\{i\}} &:= \begin{bmatrix} z_{l2} + \bar{\hat f}_{l1}(z_{l1}) \\ \vdots \\ z_{l\,i} + \bar{\hat f}_{li-1}(z_{l\{1\}}^{\{i-1\}}) \\ \bar{\hat f}_{li}(z_{l\{1\}}^{\{i\}}) \end{bmatrix} \\ \bar{\hat H}_{l\{1\}}^{\{i\}} &:= \begin{bmatrix} \bar{\hat h}_{l1}(z_{l1}) & \ldots & \bar{\hat h}_{li}(z_{l\{1\}}^{\{i\}}) \end{bmatrix}'. \end{aligned}$$

We are now ready to present the backstepping procedure for high-order optimal matching design with global inverse optimality.

Step 1: Define $z_1 := z_{l1} = x_1$ and $\bar W_1(z_{l1}) := \bar\delta_{l1}^2(z_1)z_1^2$. The virtual control law prescribed by the nonlinear Cholesky factorization of W_l is $x_2 = \bar\alpha_{l1}(z_1)$. Under this control law, the value function $\bar W_1$ satisfies the HJI equation (33), with $i=1$. Therefore, the derivative of $\bar W_1$ is given by

$$\begin{aligned} \dot{\bar W}_1 &= -\bar q_{l\{1\}}^{\{1\}} + \gamma^2 w'w - \gamma^2\left|w - \frac{1}{2\gamma^2}\,\bar{\hat H}_{l\{1\}}^{\{1\}\,\prime}\left(\frac{\partial \bar W_1}{\partial z_1}\right)'\right|^2 \\ &\quad + \bar\xi_1(z_1) + 2z_1\bar\pi_1(z_1)z_{l2} \end{aligned}$$

where $\bar\xi_1$ is the remainder of the approximate HJI equality (33),

$$\begin{aligned} \bar\xi_1(z_1) &= \frac{\partial \bar W_1}{\partial z_1}\bar{\hat f}_{l\{1\}}^{\{1\}} + \frac{1}{4\gamma^2}\frac{\partial \bar W_1}{\partial z_1}\bar{\hat H}_{l\{1\}}^{\{1\}}\,\bar{\hat H}_{l\{1\}}^{\{1\}\,\prime}\left(\frac{\partial \bar W_1}{\partial z_1}\right)' \\ &\quad + \bar q_{l\{1\}}^{\{1\}}. \end{aligned}$$

This suggests the following smooth virtual control law $z_{l2} = -\frac{1}{2\bar\pi_1(z_1)z_1}\bar\xi_1(z_1)$, which is equivalent to

$$\begin{aligned} x_2 &= \bar\alpha_{l1}(z_1) - \frac{1}{2\bar\pi_1(z_1)z_1}\bar\xi_1(z_1) \\ &=: \bar\alpha_{l1}(z_1) + \tilde{\bar\alpha}_1(z_1) := \bar\alpha_1(z_1). \end{aligned} \quad (34)$$

Because $\bar\xi_1^{[m+1]} = 0$ and $\bar\pi_1 > 0$, we conclude that $\tilde{\bar\alpha}_1^{[m]} = 0$, and hence is higher-order than the desired matching order. Under this control law, we have

$$\begin{aligned} \dot{\bar W}_1 &= -\bar q_{l\{1\}}^{\{i\}} + \gamma^2 w'w - \gamma^2\left|w - \frac{1}{2\gamma^2}\,\bar{\hat H}_{l\{1\}}^{\{i\}\,\prime}\left(\frac{\partial \bar W_1}{\partial z_1}\right)'\right|^2 \\ &\quad + 2z_1\bar\pi_1(z_1)(x_2 - \bar\alpha_1(z_1)). \end{aligned}$$

Step i, $1 < i < n$: From the preceding step we have a value function $\bar W_{i-1} = \sum_{j=1}^{i-1}\bar\delta_{lj}^2 z_j^2$ and a virtual control law $\bar\alpha_{i-1}$ for x_i, which can be decomposed into a matching part and a high-order part as follows: $\bar\alpha_{i-1}(z_{\{1\}}^{\{i-1\}}) = \bar\alpha_{l\,i-1}(z_{\{1\}}^{\{i-1\}}) + \tilde{\bar\alpha}_{i-1}(z_{\{1\}}^{\{i-1\}})$. The derivative of $\bar W_{i-1}$ is

$$\begin{aligned} \dot{\bar W}_{i-1} &= -\bar q_{l\{1\}}^{\{i-1\}}(z_{\{1\}}^{\{i-1\}}) + \gamma^2 w'w - \gamma^2\left|w - \bar\nu_{i-1}(z_{\{1\}}^{\{i\}})\right|^2 \\ &\quad + 2z_{i-1}\bar\pi_{i-1}(x_i - \bar\alpha_{i-1}(z_{\{1\}}^{\{i-1\}})) \end{aligned} \quad (35)$$

where $\bar\nu_{i-1}$ is the corresponding worst-case disturbance,

$$\bar\nu_{i-1} = \frac{1}{2\gamma^2}\left(\bar{\hat H}_{l\{1\}}^{\{i-1\}} + \bar{\hat N}_{\{1\}}^{\{i-1\}}\right)'\left(\frac{\partial \bar W_{i-1}}{\partial z_{\{1\}}^{\{i-1\}}}\right)' \quad (36)$$

$$\bar{\hat N}_{\{1\}}^{\{i-1\}} = \begin{bmatrix} \bar{\hat N}_1(z_1) \\ \vdots \\ \bar{\hat N}_{i-1}(z_{\{1\}}^{\{i-1\}}) \end{bmatrix}.$$

The dynamics of the state variable $z_{\{1\}}^{\{i-1\}}$ are

$$\begin{aligned}\dot{z}_1 &= z_2 + \bar{\hat{f}}_{l1}(z_1) + \bar{\hat{h}}'_{l1}(z_1)w + \bar{\hat{M}}_1(z_1) \\ &\quad + \bar{\hat{N}}_1(z_1)w \qquad (37a)\\ \vdots &= \quad \vdots \\ \dot{z}_{i-1} &= x_i - \bar{\alpha}_{i-1}(z_{\{1\}}^{\{i-1\}}) \\ &\quad + \bar{\hat{f}}_{l\,i-1}(z_{\{1\}}^{\{i-1\}}) + \bar{\hat{h}}'_{l\,i-1}(z_{\{1\}}^{\{i-1\}})w \\ &\quad + \bar{\hat{M}}_{i-1}(z_{\{1\}}^{\{i-1\}}) + \bar{\hat{N}}_{i-1}(z_{\{1\}}^{\{i-1\}})w \qquad (37b)\end{aligned}$$

where $\bar{\hat{M}}_j(z_{\{1\}}^{\{j\}})$ and $\bar{\hat{N}}_j(z_{\{1\}}^{\{j\}})$, $j = 1, \ldots, i-1$ are high-order nonlinear (possibly vector-valued) functions.

By hypothesis, the transformation z_j matches z_{lj}, $j = 1, \ldots, i-1$, up to mth order. At the ith step, we define the new coordinate $z_i := x_i - \bar{\alpha}_{i-1}(z_{\{1\}}^{\{i-1\}})$. Because of the matching between $\bar{\alpha}_{i-1}$ and $\bar{\alpha}_{l\,i-1}$, and the matching between $z_{\{1\}}^{\{i-1\}}$ and $z_{l\{1\}}^{\{i-1\}}$, the dynamics of z_i is given by

$$\begin{aligned}\dot{z}_i &= \bar{\hat{f}}_{li}(z_{\{1\}}^{\{i\}}) + z_{l\,i+1} + \bar{\hat{h}}'_{li}(z_{\{1\}}^{\{i\}})w \\ &\quad + \bar{\hat{M}}_{bi}(z_{\{1\}}^{\{i\}}) + \bar{\hat{N}}_i(z_{\{1\}}^{\{i\}})w\end{aligned}$$

where $\bar{\hat{M}}_{bi}$ and $\bar{\hat{N}}_i$ are high-order nonlinear functions.

Again, based on the approximate factorization, we recursively define the value function for this step as: $\bar{W}_i(z_{\{1\}}^{\{i\}}) := \bar{W}_{i-1}(z_{\{1\}}^{\{i-1\}}) + \bar{\delta}_{li}^2(z_{\{1\}}^{\{i\}})z_i^2$. Using the relationship (33), the derivative of $\bar{W}_i$ can be expressed as follows:

$$\begin{aligned}\dot{\bar{W}}_i &= -\bar{q}_{l\{1\}}^{\{i\}}(z_{\{1\}}^{\{i\}}) + \gamma^2 w'w - \gamma^2 \left| w - \bar{\nu}_i(z_{\{1\}}^{\{i\}})\right|^2 \\ &\quad + 2z_i\bar{\pi}_i z_{l\,i+1} + \bar{\xi}_i(z_{\{1\}}^{\{i\}})\end{aligned}$$

where

$$\begin{aligned}\bar{\xi}_i(z_{\{1\}}^{\{i\}}) &= \bar{q}_{l\{1\}}^{\{i\}} - \bar{q}_{l\{1\}}^{\{i-1\}} + \gamma^2\bar{\nu}_i'\bar{\nu}_i - \gamma^2\bar{\nu}_{i-1}'\bar{\nu}_{i-1} \\ &\quad + \sum_{j=1}^{i-1} 2z_i^2\bar{\delta}_{li}\frac{\partial\bar{\delta}_{li}}{\partial z_j}(z_{j+1} + \bar{\hat{f}}_{lj} + \bar{\hat{M}}_j) \\ &\quad + 2z_i\bar{\pi}_i(\bar{\hat{f}}_{li} + \bar{\hat{M}}_{bi}) + 2z_{i-1}\bar{\pi}_{i-1}z_i \qquad (38)\\ \bar{\nu}_i(z_{\{1\}}^{\{i\}}) &= \frac{1}{2\gamma^2}\left(\bar{\hat{H}}_{l\{1\}}^{\{i\}} + \bar{\hat{N}}_{\{1\}}^{\{i\}}\right)'\left(\frac{\partial\bar{W}_i}{\partial z_{\{1\}}^{\{i\}}}\right)' \qquad (39)\\ \bar{\hat{N}}_{\{1\}}^{\{i\}}(z_{\{1\}}^{\{i\}}) &= \left[\begin{array}{ccc}\bar{\hat{N}}_1(z_1)' & \ldots & \bar{\hat{N}}_i(z_{\{1\}}^{\{i\}})'\end{array}\right]'.\end{aligned}$$

By the approximate HJI equality (33), we conclude that $\left(\bar{\xi}_i(z_{\{1\}}^{\{i\}})\right)^{[m+1]} = 0$. From the expressions (38), (39) and (36), we also conclude that $\bar{\xi}_i(z_{\{1\}}^{\{i\}})$ contains z_i as a factor. Hence, the smooth virtual control law for $z_{l\,i+1}$ is $z_{l\,i+1} = -\frac{1}{2z_i\bar{\pi}_i}\bar{\xi}_i(z_{\{1\}}^{\{i\}})$.

This leads to the satisfaction of a dissipation inequality for the $x_{\{1\}}^{\{i\}}$ dynamics with supply rate, $\bar{q}_{l\{1\}}^{\{i\}}(z_{\{1\}}^{\{i\}}) - \gamma^2 w'w$.

The equivalent virtual control law for x_{i+1} is

$$\begin{aligned}x_{i+1} &= \bar{\alpha}_i(z_{\{1\}}^{\{i\}}) := \bar{\alpha}_{li}(z_{\{1\}}^{\{i\}}) - \frac{1}{2z_i\bar{\pi}_i}\bar{\xi}_i(z_{\{1\}}^{\{i\}}) \\ &=: \bar{\alpha}_{li}(z_{\{1\}}^{\{i\}}) + \tilde{\bar{\alpha}}_i(z_{\{1\}}^{\{i\}}).\end{aligned}$$

Therefore, the matching property of the virtual control law is verified. Under this control law, the derivative of $\bar{W}_i$ is

$$\begin{aligned}\dot{\bar{W}}_i &= -\bar{q}_{l\{1\}}^{\{i\}}(z_{\{1\}}^{\{i\}}) + \gamma^2 w'w - \gamma^2\left|w - \bar{\nu}_i(z_{\{1\}}^{\{i\}})\right|^2 \\ &\quad + 2z_i\bar{\pi}_i(x_{i+1} - \bar{\alpha}_i(z_{\{1\}}^{\{i\}})).\end{aligned}$$

This completes the ith design step, up to $i = n-1$.

Step n: For the final nth step, we define $z_n := x_n - \bar{\alpha}_{n-1}(z_{\{1\}}^{\{n-1\}})$. Now the actual control variable u appears in the derivative of z_n:

$$\dot{z}_n = \bar{\hat{f}}_{ln}(z) + u + \bar{\hat{h}}'_{ln}(z)w + \bar{\hat{M}}_n(z) + \bar{\hat{N}}_n(z)w$$

where $\bar{\hat{M}}_n$ and $\bar{\hat{N}}_n$ are present due to the higher-order mismatches between z and z_l.

The value function for this final step, which becomes the value function for the complete system, is $\bar{W}_n(z) := \bar{W}_{n-1}(z_{\{1\}}^{\{n-1\}}) + \bar{\delta}_{ln}^2(z)z_n^2 =: \bar{W}(z)$. Because of the approximate HJI equality (32), the derivative of $\bar{W}$ is equal to the following expression, after two "completion of squares" arguments:

$$\begin{aligned}\dot{\bar{W}} &= -\bar{q}_l(z) - \bar{\check{r}}u^2 + \gamma^2 w'w + \bar{\check{r}}\left|u + \bar{\check{r}}^{-1}\bar{\pi}_n(z)z_n\right|^2 \\ &\quad - \gamma^2\left|w - \bar{\nu}_n(z)\right|^2 - (\bar{\check{r}}^{-1}(z) - \bar{r}_l^{-1}(z))\bar{\pi}_n^2(z)z_n^2 \\ &\quad + \bar{\xi}_n(z)\end{aligned}$$

where

$$\begin{aligned}\bar{\xi}_n(z) &= \bar{q}_l - \bar{q}_{l\{1\}}^{\{n-1\}} + \gamma^2\bar{\nu}_n'\bar{\nu}_n - \gamma^2\bar{\nu}_{n-1}'\bar{\nu}_{n-1} \\ &\quad + \sum_{j=1}^{n-1} 2z_n^2\bar{\delta}_{ln}\frac{\partial\bar{\delta}_{ln}}{\partial z_j}(z_{j+1} + \bar{\hat{f}}_{lj} + \bar{\hat{M}}_j) \\ &\quad + 2z_n\bar{\pi}_n(\bar{\hat{f}}_{ln} + \bar{\hat{M}}_n) + 2z_{n-1}\bar{\pi}_{n-1}z_n \\ &\quad - \bar{r}_l^{-1}(z)\bar{\pi}_n^2(z)z_n^2 \\ \bar{\nu}_n(z) &= \frac{1}{2\gamma^2}\left(\bar{\hat{H}}_l' + \bar{\hat{N}}\right)'\left(\frac{\partial\bar{W}}{\partial z}\right)' \\ \bar{\hat{N}}(z) &= \left[\begin{array}{ccc}\bar{\hat{N}}_1(z_1)' & \ldots & \bar{\hat{N}}_n(z)'\end{array}\right]'.\end{aligned}$$

Because the approximating value function W satisfies (32), the function $\bar{\xi}_n$ is higher-order such that $\bar{\xi}_n^{[m+1]} = 0$. Furthermore, from the definition of $\bar{\xi}_n$, we conclude that $\bar{\xi}_n$ contains z_n as a factor, and $\bar{\xi}_f(z) := \bar{\xi}_n(z)/z_n$ is a smooth function such that $\bar{\xi}_f^{[m]} = 0$.

As $\check{r}$ we choose the locally Lipschitz continuous function

$$\bar{\check{r}}(z) = \begin{cases} \bar{\check{r}}_1(z) & \text{if } \bar{q}_l(z)/4 < \bar{\xi}_f(z)z_n \\ \bar{r}_l(z) & \text{if } \bar{q}_l(z)/2 - \bar{r}_l^{-1}(z)\bar{\pi}_n^2(z)z_n^2 - \epsilon \\ & \quad \le \bar{\xi}_f(z)z_n \le \bar{q}_l(z)/4 \\ \bar{\check{r}}_2(z) & \text{if } \bar{q}_l(z)/2 - \bar{r}_l^{-1}(z)\bar{\pi}_n^2(z)z_n^2 - \epsilon \\ & \quad > \bar{\xi}_f(z)z_n \end{cases} \qquad (40)$$

where

$$\bar{\bar{r}}_1(z) = \frac{2\bar{q}_l(z)\bar{\pi}_n^2(z)z_n^2\bar{r}_l(z)}{(\bar{q}_l(z)/4 - \bar{\xi}_f(z)z_n)^2\bar{r}_l(z) + 2\bar{q}_l(z)\bar{\pi}_n^2(z)z_n^2}$$
$$\bar{\bar{r}}_2(z) = \bar{r}_l(z)(1 + (\bar{q}_l/2 - \bar{r}_l^{-1}\bar{\pi}_n^2 z_n^2 - \epsilon - \bar{\xi}_f z_n)^2\bar{\pi}_n^2 z_n^2),$$

and ϵ is any positive constant.

Therefore, the control law,

$$\bar{\mu}(z) = -\bar{\bar{r}}^{-1}\bar{\pi}_n(z)z_n \tag{41}$$

is the desired inverse optimal matching control design, and it is inverse robust optimal with respect to the cost function $\check{q} + \check{r}u^2$, where

$$\bar{\bar{q}} := \bar{q}_l(z) + (\bar{\bar{r}}^{-1}(z) - \bar{r}_l^{-1}(z))\bar{\pi}_n^2(z)z_n^2 - \bar{\xi}_n(z). \tag{42}$$

With this choice of $\check{r}$, we note that, in a neighborhood of the origin, $\bar{q}_l(z)/2 \geq c_1|z|^2 \geq c_2|z|^3 \geq \bar{\xi}_f(z)z_n$ for some positive constants c_1 and c_2, and $\epsilon \geq \bar{q}_l(z)/2 - \check{r}_l^{-1}(z)\bar{\pi}_n^2(z)z_n^2 - \bar{\xi}_f(z)z_n$. Therefore, in a neighborhood of the origin, we have $\bar{\bar{r}}(z) = \bar{r}_l(z)$. This implies that the control law (41) satisfies the matching requirement. In the case when $\bar{q}_l(z)/2 - \bar{r}_l^{-1}(z)\bar{\pi}_n^2(z)z_n^2 - \epsilon > \bar{\xi}_f(z)z_n$, we have $\bar{\bar{q}}(z) > \bar{q}_l(z)/2 + \epsilon$. In the case when $\bar{q}_l(z)/2 < \bar{\xi}_f(z)z_n$, we have $|z| > 0$, which implies that $\bar{q}_l(z)/2 > 0$, and therefore $z_n \neq 0$. Then, $\bar{\bar{r}}(z) > 0$ and

$$\bar{\bar{r}}^{-1}(z) - \bar{r}_l^{-1}(z))\bar{\pi}_n^2(z)z_n^2 - \bar{\xi}_f(z)z_n + \frac{3}{4}\bar{q}_l(z)$$
$$= \frac{1}{2}\bar{q}_l(z)\left|1 - \frac{z_n\bar{\xi}_f - \bar{q}_l(z)/4}{\bar{q}_l(z)}\right|^2 \geq 0$$

This completes the control design. This result is summarized in the following theorem.

Theorem 4: Consider the nonlinear system (1) and cost functional (2) with Assumptions **A1**, **A2**, *and* **A3**. *Let the function W_l be any radially unbounded and smooth function that has the same Taylor series expansion as the value function V up to $(m+1)$th order and admits a nonlinear Cholesky factorization in the reverse order of state variables $(x_n, \ldots, x_1)$. Then, the control law (41) is locally optimal matching up to mth order and is inverse robust optimal with respect to the cost functional (5), where $\check{q}$ is a positive definite and radially unbounded function defined by (42), and $\check{r}$ is a positive function defined by (40).*

Proof: We note that the constructed value function $\bar{W}$ is positive definite and radially unbounded because W_l is assumed to be positive definite and radially unbounded, and $\bar{W}(z) = \bar{W}_l(z)$. The function $\check{q}$ is positive definite and radially unbounded because it is larger than or equal to $q(x)/4$. Furthermore, the function $\bar{W}$ satisfies the HJI equation and the matching requirements by the recursive construction.

In the above recursive construction, we can start from any given factorizable function W_l that approximates the value function V. Clearly, for different W_l's, the construction leads to, in general, different controller designs, that all satisfy the inverse robust optimality requirement and the local matching property. A question that remains to be answered is whether there is any criterion to distinguish which W_l is more appropriate for a particular problem.

V. Example

We consider a second order nonlinear system:

$$\dot{x}_1 = x_1^2 + x_2 + w$$
$$\dot{x}_2 = u$$

with cost functional

$$J_\gamma = \int_0^\infty (x_1^2 + x_2^2 + u^2 - \gamma^2 w^2)\, dt$$

where we fix the desired disturbance attenuation level to be $\gamma = \sqrt{2}$.

The HJI equation associated with this zero-sum differential game is

$$V_{x_1}(x_1^2 + x_2) + \frac{1}{4\gamma^2}V_{x_1}^2 - \frac{1}{4}V_{x_2}^2 + x_1^2 + x_2^2 = 0 \tag{43}$$

where $V(x_1, x_2)$ is the value function for the game, if it exists. The Taylor series expansion of the function $V(x_1, x_2)$ around the origin up to 3rd order is given by [2]

$$V^{[3]}(x_1, x_2) = 6.680x_1^2 + 9.657x_1x_2 + 4.724x_2^2 + 89.92x_1^3$$
$$+189.4x_1^2x_2 + 134.7x_1x_2^2 + 32.50x_2^3.$$

1st order matching

For $m = 1$ we are only concerned with the 2nd order Taylor series expansion of the value function

$$V^{[2]}(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}' \begin{bmatrix} 6.680 & 4.828 \\ 4.828 & 4.724 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

The controller design starts by first obtaining a factorizable function W that matches the value function V up to 2nd order. It turns out that the function W is exactly $V^{[2]}$. The nonlinear Cholesky factorization is equivalent to the Cholesky factorization for positive definite matrices:

$$V^{[2]}(x_1, x_2) = 1.323^2x_1^2 + 2.174^2(x_2 + 1.022x_1)^2.$$

Hence, the simple diffeomorphism is given by $\begin{bmatrix} z_{l1} & z_{l2} \end{bmatrix}' = \begin{bmatrix} x_1 & x_2 + 1.022x_1 \end{bmatrix}'$. In the $z_l = (z_{l1}, z_{l2})'$ coordinates, the original system is given by

$$\dot{z}_{l1} = -1.022z_{l1} + z_{l1}^2 + z_{l2} + w$$
$$\dot{z}_{l2} = -1.045z_{l1} + 1.022z_{l1}^2 + 1.022z_{l2} + u + 1.022w,$$

and the cost functional is

$$J_\gamma = \int_0^\infty (2.045z_{l1}^2 - 2.044z_{l1}z_{l2} + z_{l2}^2 + u^2 - 2w^2)\, dt.$$

Next, we proceed through the two steps of backstepping. In the first step, we set $z_1 = z_{l1}$, and choose a value function $V_1(z_1) = 1.323^2z_1^2 = 1.751z_1^2$, so that

$$\dot{V}_1 = 3.502z_1(z_{l2} - 1.022z_1 + z_1^2 + w)$$
$$= -2.045z_1^2 + 3.502z_1(z_{l2} + z_1^2) + 2w^2$$
$$-2(w - 0.8755z_1)^2.$$

Hence, the desired virtual control law for z_{l2} is $-z_1^2$. The worst case disturbance for this step is $\bar{\nu}_1^* = 0.8755z_1$.

In the second step, which is also the last step, we define the state variable $z_2 = z_{l2} + z_1^2$ and obtain

$$\begin{aligned}\dot{z}_2 &= u - 0.1502z_1 + 1.022z_2 + 2z_1z_2 - 0.2940z_1^2 \\ &\quad +(1.022 + 2z_1)(w - 0.8755z_1).\end{aligned}$$

The value function for this step is $V_2(z_1, z_2) = V_1 + 2.174^2 z_2^2$, and, after some algebraic manipulations, we get

$$\begin{aligned}\dot{V}_2 &= -2.045z_1^2 + 2.044z_1z_2 - z_2^2 + 2w^2 \\ &\quad -2(w - 0.8755z_1 - 2.415z_2 - 4.728z_1z_2)^2 \\ &\quad -\bar{\bar{r}}u^2 + \bar{\bar{r}}(u + 4.726\bar{\bar{r}}^{-1}z_2)^2 - 22.34z_2^2(\bar{\bar{r}}^{-1} - 1) \\ &\quad +64.58z_1z_2^2 + 44.71z_1^4 - 2.779z_1^2z_2.\end{aligned}$$

Therefore, we have $\bar{\xi}_2 = 64.58z_1z_2^2 + 44.71z_1^4 - 2.779z_1^2z_2 =: \bar{\xi}_f z_2$, and define $\bar{\bar{r}}(z_1, z_2)$ as in (40) with $\bar{q}_l(z) = 2.045z_1^2 - 2.044z_1z_2 + z_2^2$, $\bar{r}_l(z) = 1$, $\bar{\pi}_2^2 = 2.174^4$, and $\epsilon = 0.1$. Then, the inverse optimal control law with optimality matching up to 1st order is

$$u^* = -4.726\check{r}^{-1}(x)(x_2 + 1.022x_1 + x_1^2).$$

The worst case disturbance is $\bar{\nu}_2^* = 0.8755z_1 + 2.415z_2 + 4.728z_1z_2$.

2nd order matching

For $m = 2$, we are concerned with the 3rd order Taylor series expansion of the value function V. Again, we start by obtaining a factorizable function W that matches the value function V up to 3rd order. We follow the recursive procedure of Section 3 to construct the function W that is factorizable in the order (x_2, x_1). Let

$$\begin{aligned}V_p(x_2, x_1) = \frac{\partial V^{[3]}}{\partial x_2} &= 9.657x_1 + 9.448x_2 + 189.4x_1^2 \\ &\quad +269.4x_1x_2 + 97.50x_2^2.\end{aligned}$$

Using (20), the root function for x_2 up to second order is $\alpha(x_1) = -1.022x_1 - 1.684x_1^2$. In terms of $\eta := x_2 + 1.022x_1 + 1.684x_1^2$ and x_1, the function $\bar{\Pi}$ is obtained as $\bar{\Pi}(\eta, x_1) = 9.448 + 70.11x_1 + 97.50\eta - 328.4x_1^2$. Choosing $C = 0.448$, we obtain $\bar{\Pi}_s(\eta, x_1) = 3 + 11.69x_1 + 16.25\eta$. Then, we have the following identity

$$\begin{aligned}V^{[3]}(x_1, x_2) &= 1.745x_1^2 + 2.351x_1^3 \\ &\quad + \left(\int_0^\eta s((3 + 11.69x_1 + 16.25s)^2 + 0.448)\bar{T}_1^2\, ds\right)^{[3]}\end{aligned}$$

where $\eta = x_2 + 1.022x_1 + 1.684x_1^2$, $\bar{T}_1^2(s, x_1) := \frac{1}{1+p_1s^2}$ and $p_1 \geq 0$ is a design parameter.

Hence, the function to be matched at the second step of the nonlinear Cholesky factorization is $\hat{V}(x_1) = 1.745x_1^2 + 2.351x_1^3$.

Let $V_p(x_1) = \frac{\partial \hat{V}}{\partial x_1} = 3.490x_1 + 7.053x_1^2$. For this last step, the root function is fixed to be $x_1 = 0$. Then, we have $\bar{\Pi}(x_1) = 3.490 + 7.053x_1$. With $C = 0.250$, $\bar{\Pi}_s$ is given by $\bar{\Pi}_s(x_1) = 1.8 + 1.959x_1$. Thus, $\hat{V}$ matches the following function up to 3rd order.

$$\begin{aligned}\hat{V}(x_1) &= \left(\int_0^\eta s((1.8 + 1.959s)^2 + 0.250)\bar{T}_2^2(s)\, ds\right)^{[3]} \\ &= 5p_2^{-1}\left((0.349 - 0.384p_2^{-1})\ln(1 + p_2x_1^2)\right. \\ &\quad \left.-1.41\tan^{-1}(p_2^{0.5}x_1)p_2^{-0.5} + 1.41x_1 + 0.384x_1^2\right)^{[3]}\end{aligned}$$

where $\eta = x_1$, $\bar{T}_2^2 := \frac{1}{1+p_2s^2}$ and $p_2 \geq 0$ is a design parameter.

The scaling functions $\bar{T}_1$ and $\bar{T}_2$ are chosen to reduce the control effort for the inverse optimal matching design. Compared with a design using $\bar{T}_1 \equiv 1$ and $\bar{T}_2 \equiv 1$, the control effort is dramatically reduced. A choice for $\bar{T}_1^2$ which could further reduce the control effort, but which leads to a more complicated control law, is $\frac{1}{1+p_{11}x_1^2+p_{12}x_1s+p_{13}s^2}$.

This completes the construction of the matching function

$$W(x_1, x_2) = \hat{W}_1(\eta)\Big|_{\eta=x_1} + \hat{W}_2(x_1, \eta)\Big|_{\eta=x_2+1.022x_1+1.684x_1^2}$$

where

$$\begin{aligned}\hat{W}_1 &:= \int_0^\eta s((1.8 + 1.959s)^2 + 0.250)T_2^2\, ds \\ \hat{W}_2 &:= \int_0^\eta s((3 + 11.69x_1 + 16.25s)^2 + 0.448)T_1^2\, ds\end{aligned}$$

with the properties $\frac{\partial \hat{W}_1}{\partial x_1}(x_1) = 2x_1\bar{\pi}_1(x_1)$, $\frac{\partial \hat{W}_2}{\partial \eta}(x_1, \eta) = 2\eta\bar{\pi}_2(x_1, \eta)$, $\frac{\partial \hat{W}_2}{\partial x_1}(x_1, \eta) = I(x_1, \eta)$, and where

$$\begin{aligned}\bar{\pi}_1(x_1) &:= 0.5\left((1.8 + 1.959x_1)^2 + 0.25\right)/(1 + p_2x_1^2) > 0 \\ \bar{\pi}_2(x_1, \eta) &:= 0.5\left((3 + 11.69x_1 + 16.25\eta)^2 + 0.448\right)/(1 + p_1\eta^2) > 0 \\ I(x_1, \eta) &:= -5.0\left[-75.985\eta + 75.985\tan^{-1}(p_1^{0.5}\eta)p_1^{-0.5}\right. \\ &\quad \left.-7.014\ln(1 + p_1\eta^2) - 27.33\ln(1 + p_1\eta^2)z_1\right]p_1^{-1}.\end{aligned}$$

This function W matches the value function V up to the 3rd order.

The simple diffeomorphism Φ_s in the nonlinear Cholesky factorization for the function W is

$$\begin{bmatrix} z_{l1} \\ z_{l2} \end{bmatrix} = \Phi_s(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 + 1.022x_1 + 1.684x_1^2 \end{bmatrix}$$

and, in the z_l coordinates, the original system becomes

$$\begin{aligned}\dot{z}_{l1} &= -1.022z_{l1} - 0.684z_{l1}^2 + z_{l2} + w \\ \dot{z}_{l2} &= -1.044z_{l1} - 4.141z_{l1}^2 - 2.304z_{l1}^3 + 3.368z_{l1}z_{l2} \\ &\quad +1.022z_{l2} + u + (1.022 + 3.368z_{l1})w.\end{aligned}$$

The approximate value function W can be expressed as

$$\bar{W} = \hat{W}_1(z_{l1}) + \hat{W}_2(z_{l1}, z_{l2})$$

and the cost functional is

$$\begin{aligned}J_\gamma &= \int_0^\infty (2.044z_{l1}^2 + 3.442z_{l1}^3 + 2.836z_{l1}^4 + \\ &\quad (-2.044z_{l1} - 3.368z_{l1}^2)z_{l2} + z_{l2}^2 + u^2 - 2w^2)\, dt.\end{aligned}$$

With the function W available, we now proceed to the two steps of backstepping for the construction of an inversely robust optimal controller that matches the optimal control law up to 2nd order.

For the first step, we define $z_1 = z_{l1}$, $V_1(z_1) = \hat{W}_1(z_1)$, and get

$$\begin{aligned}\dot{V}_1 &= 2w^2 - 2(w - 0.5\bar{\pi}_1 z_1)^2 - 2.044z_1^2 - 3.442z_1^3 \\ &\quad -2.836z_1^4 + 2z_1\bar{\pi}_1(z_{l2} + \bar{\alpha}(z_1)),\end{aligned}$$

where $(\bar{\alpha}(z_1))^{[2]} = 0$. Hence, the desired virtual control is $z_{l2} = -\bar{\alpha}(z_1)$, and the worst case disturbance for this step is $\bar{\nu}_1^* = 0.5\bar{\pi}_1(z_1)z_1$.

For the second step, we define a new state variable $z_2 = z_{l2} + \bar{\alpha}(z_1)$ which satisfies

$$\begin{aligned}\dot{z}_2 &= -1.044z_1 - 4.141z_1^2 + 3.368z_1z_2 + 1.022z_2 + u \\ &\quad +H.O.T^{[3]} + (1.022 + 3.368z_1 + H.O.T^{[2]})w\end{aligned}$$

where $H.O.T^{[i]}$ denotes the terms of ith order and above.

The value function for this step is

$$V_2(z_1, z_2) = V_1(z_1) + \hat{W}_2(z_1, z_2),$$

and its derivative is given by

$$\begin{aligned}\dot{V}_2 &= 2.044z_1^2 - 3.442z_1^3 - z_2^2 - 2.836z_1^4 \\ &\quad +2w^2 - 2\Bigg(w - 0.5\bar{\pi}_1 z_1 - \frac{1}{4}I(z_1, z_2) \\ &\quad -\frac{1}{2}(1.022 + 3.368z_1 + H.O.T^{[2]})\bar{\pi}_2 z_2\Bigg)^2 \\ &\quad +(2.051z_1 + 3.368z_1^2)z_2 + z_2\xi_f(z_1, z_2) \\ &\quad -\bar{\check{r}}u^2 + \bar{\check{r}}(u + \bar{\check{r}}^{-1}\bar{\pi}_2 z_2)^2 - (\bar{\check{r}}^{-1} - 1)\bar{\pi}_2^2 z_2^2\end{aligned}$$

where $\xi_f^{[2]}(z_1, z_2) = 0$. Now $\check{r}$ can be chosen according to (40). The inverse robust optimal control law is given by

$$u^* = -\bar{\check{r}}^{-1}, \bar{\pi}_2(z_1, z_2)z_2$$

and the corresponding worst-case disturbance is

$$\begin{aligned}\bar{\nu}_2^* &= 0.5\bar{\pi}_1 z_1 + \frac{1}{4}I(z_1, z_2) \\ &\quad +\frac{1}{2}(1.022 + 3.368z_1 + H.O.T^{[2]})\bar{\pi}_2(z_1, z_2)z_2.\end{aligned}$$

Simulations are used to illustrate the theoretical findings. The design parameters $p_1 = 2000$ and $p_2 = 4$ are chosen to reduce the control magnitude. For comparison, the 2nd-order local controller has also been simulated with the control law $u_{local} = -V_p/2$. The phase-portraits for the disturbance-free closed-loop systems under the two controllers are shown in Figure 1. While the manifold M_{s2} is the stability boundary for the local controller, the matching controller guarantees global asymptotic stability. Starting at point A_2, we have depicted the state trajectories in Figure 2 and the control trajectories in Figure 3. We observe that initially the control magnitude for the matching controller is somewhat larger than that for the local controller. This is the price to be paid for global robustness with respect to the disturbance input.

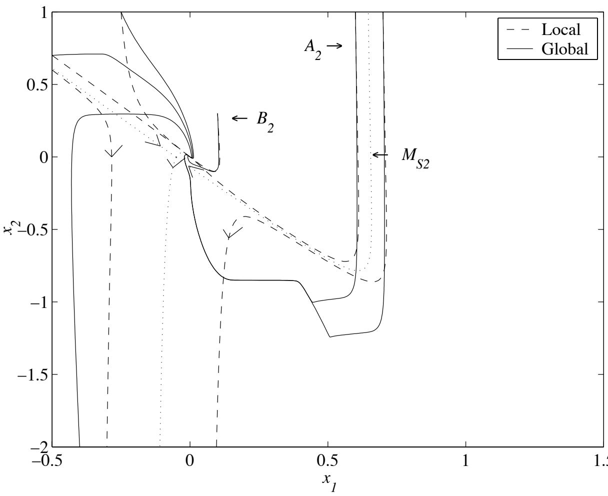

Fig. 1. Phase-portraits.

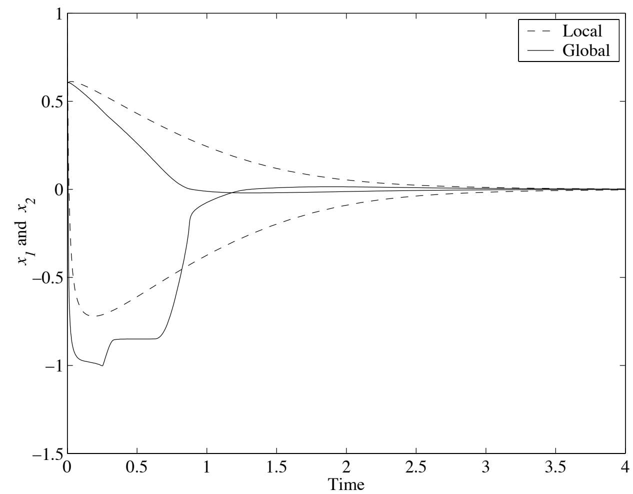

Fig. 2. State trajectories.

VI. Conclusion

We have developed a design procedure which combines the Taylor series expansion and integrator backstepping to solve the problem of nonlinear H^∞ optimal control for strict-feedback nonlinear systems. The procedure recursively constructs a cost functional and the corresponding solution to the HJI equation for a strict-feedback nonlinear system such that the optimal performance is matched up to *any* desired order of the Taylor series expansion. Moreover, this procedure is also applicable to the nonlinear regulator problem. What lies at the heart of the recursive construction is the new concept of nonlinear Cholesky factorization. The nonlinear Cholesky factorization for a given positive definite function is defined as an upper triangular coordinate transformation such that in the new coordinates the given function is equal to the sum of squares. We have obtained precise conditions under which a given nonlinear function has a nonlinear Cholesky factorization.

When the value function for the optimal control problem has a nonlinear Cholesky factorization, we have shown that

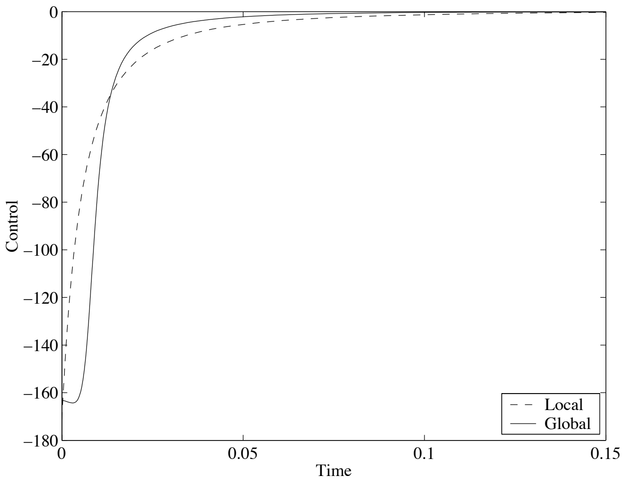

Fig. 3. Control efforts.

the backstepping procedure can be tuned to result in the optimal control design, thus justifying backstepping as an optimal design method. Making use of the Taylor series expansion for the value function, we have developed explicit recursive computation schemes for a globally stable and inversely optimal controller that matches the optimal solution up to any desired order. A simulation example has been included to illustrate the theoretical findings.

References

[1] Z. Pan, K. Ezal, A. J. Krener, and P. V. Kokotović, "Backstepping design with local optimality matching," Technical Report CCEC 96-0810, University of California at Santa Barbara, Santa Barbara, CA, August 1996.

[2] A. J. van der Schaft, "L_2-gain analysis of nonlinear systems and nonlinear H_∞ control," *IEEE Transactions on Automatic Control*, vol. 37, pp. 770–784, 1992.

[3] A. Isidori and W. Kang, "H_∞ control via measurement feedback for general nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 40, pp. 466–472, March 1995.

[4] J. A. Ball, J. W. Helton, and M. Walker, "H^∞ control for nonlinear systems with output feedback," *IEEE Transactions on Automatic Control*, vol. 38, pp. 546–559, April 1993.

[5] R. Marino, W. Respondek, A. J. van der Schaft, and P. Tomei, "Nonlinear H_∞ almost disturbance decoupling," *Systems and Control Letters*, vol. 23, pp. 159–168, 1994.

[6] Z. Pan and T. Başar, "Robustness of minimax controllers to nonlinear perturbations," *Journal of Optimization Theory and Applications*, vol. 87, pp. 631–678, December 1995.

[7] T. Başar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Boston, MA: Birkhäuser, 2nd ed., 1995.

[8] D. L. Lukes, "Optimal regulation of nonlinear dynamical systems," *SIAM Journal on Control*, vol. 7, pp. 75–100, Febuary 1969.

[9] E. G. Al'brecht, "On the optimal stabilization of nonlinear systems," *PMM – J. Appl. Math. Mech.*, vol. 25, pp. 1254–1266, 1961.

[10] W. Kang, P. K. De, and A. Isidori, "Flight control in a windshear via nonlinear H_∞ method," in *Proceedings of the 31st IEEE Conference on Decision and Control*, (Tucson, AZ), pp. 1135–1142, 1992.

[11] A. J. Krener, "The construction of optimal linear and nonlinear regulators," in *Systems, Models and Feedback, Theory and Applications* (A. Isidori and T. J. Tarn, eds.), Boston: Birkhäuser, 1992.

[12] A. J. Krener, "Optimal model matching controllers for linear and nonlinear systems," in *Proceedings of the 2nd IFAC Symposium* (M. Fliess, ed.), (Bordeaux, France), pp. 209–214, Pergamon Press, 1992.

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

[14] I. Kanellakopoulos, P. V. Kokotović, and A. S. Morse, "Systematic design of adaptive controllers for feedback linearizable systems," *IEEE Transactions on Automatic Control*, vol. 36, pp. 1241–1253, 1991.

[15] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović, *Nonlinear and Adaptive Control Design*. New York, NY: Wiley, 1995.

[16] K. Ezal, Z. Pan, and P. V. Kokotović, "Locally optimal and robust backstepping design," *IEEE Transactions on Automatic Control*, vol. 45, pp. 260–271, February 2000.

[17] R. A. Freeman and P. V. Kokotović, *Robust Nonlinear Control Design State-Space and Lyapunov Techniques*. Boston, MA: Birkhäuser, 1996.

[18] M. Krstić and Z. Li, "Inverse optimal design of input-to-state stabilizing nonlinear controllers," *IEEE Transactions on Automatic Control*, vol. 43, pp. 336–350, March 1998.

Zigang Pan (S'92–M'96) received his B.S. degree in Automatic Control from Shanghai Jiao Tong University in 1990, and the M.S. and Ph. D. degrees in Electrical Engineering from University of Illinois at Urbana-Champaign in 1992 and 1996, respectively. In 1996, he was a Research Engineer at the Center for Control Engineering and Computation at University of California, Santa Barbara. In the same year, he joined the Polytechnic University, Brooklyn, NY, as an Assistant Professor in the Department of Electrical Engineering for two years. Afterwards, he joined the faculty of Shanghai Jiao Tong University, Shanghai, P. R. China, in the Department of Automation. In December 2000, he joined the Department of Electrical and Computer Engineering and Computer Science of the University of Cincinnati, where he is currently an Assistant Professor. His current research interests include perturbation theory for robust control systems, optimality-guided robust parameter identification, optimality-guided robust adaptive control, control design for stochastic nonlinear systems, and applications of adaptive control systems. He was a coauthor of a paper (with T. Başar) that received the 1995 George Axelby Best Paper Award. He is a member of IEEE and a member of the Phi Kappa Phi Honor Society.

Kenan Ezal (S'89–M'99) received the B.S. degree in electrical engineering from the University of California at Los Angeles in 1989 and the M.S.E. degree from the University of Michigan at Ann Arbor in 1990. In 1991, he joined Delco Systems Operations in Santa Barbara, California, where he was involved in the development of guidance and control algorithms for military aircraft, as well as the control and calibration of resonating gyroscopes. He obtained the M.S. and Ph.D. degrees in electrical engineering from the University of California at Santa Barbara in 1996 and 1998, respectively. Dr. Ezal was the recipient of the Best Student Paper Award at the 1997 Conference on Decision and Control. Currently, he is a Senior Analyst at Toyon Research Corporation in Santa Barbara, where he is researching automatic target recognition algorithms, foliage penetrating radar, and bearings-only multi-target multi-sensor tracking algorithms. His research interests also include robust and adaptive control of nonlinear systems, antilock braking, and adaptive linearization of nonlinear amplifiers utilized in communications.

Arthur J. Krener (M'77–SM'82–F'90) was born in Brooklyn, NY on October 8, 1942. He received the BS degree from Holy Cross College in 1964 and the MA and PhD degrees from the University of California, Berkeley in 1967 and 1971, all in mathematics. Since 1971 he has been at the University of California, Davis where he is currently Professor of Mathematics. He has held visiting positions at Harvard University, the University of Rome, Imperial College of Science and Technology, NASA Ames Research Center, the University of California, Berkeley, the University of Paris IX, the University of Maryland, University of Newcastle and the University of Padua. His research interests are in developing methods for the control and estimation of nonlinear dynamical systems and stochastic processes. Professor Krener is a Fellow of the IEEE. His 1981 TAC paper with Isidori, Gori-Giorgi and Monaco won a Best Paper Award. His 1977 TAC paper with Robert Hermann was recently chosen as one of 25 Seminal Papers in Control.

Petar V. Kokotović (SM'74–F'80) has been active for more than thirty years as a control engineer, researcher and educator, first in his native Yugoslavia and then, from 1966 through 1990, at the University of Illinois, where he held the endowed Grainger Chair. In 1991 he joined the University of California, Santa Barbara, where he directs the Center for Control Engineering and Computation. He has co-authored eight books and numerous articles contributing to sensitivity analysis, singular perturbation methods, and robust adaptive and nonlinear control. Professor Kokotović is also active in industrial applications of control theory. As a consultant to Ford he was involved in the development of the first series of automotive computer controls and at General Electric he participated in large scale systems studies. Professor Kokotović is a Fellow of IEEE, and a member of the National Academy of Engineering. He received the 1983 and 1993 Outstanding IEEE Transactions Paper Awards and presented the 1991 Bode Prize Lecture. He is the recipient of the 1990 IFAC Quazza Medal and the 1995 IEEE Control Systems Award.