2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Asymptotic Optimal Tracking Control for an Uncertain Nonlinear Euler-Lagrange System: A RISE-based Closed-Loop Stackelberg Game Approach

T. Hiramatsu, M. Johnson, N. G. Fitz-Coy, and W. E. Dixon

Abstract— **Differential game theory is used to develop controllers for an uncertain nonlinear Euler-Lagrange system. A closed-loop Stackelberg strategy based on hierarchical characteristics of the system is employed. A Robust Integral Sign of the Error (RISE) controller is used to partially cancel uncertain nonlinearities in the system first, and the residual system is modeled as an infinite-horizon two-person Stackelberg differential game. Although the game is linear-quadratic (LQ) not all the nonlinearities are lost since the residual system is linear in errors but not in the original states. To alleviate time inconsistency a closed-loop strategy is sought such that the controller assumes the potential perturbation to the system and computes its strategy accordingly. An analytical solution is presented to allow of a real-time controller implementation. A Lyapunov analysis is provided to examine the stability of the developed controller.**

I. Introduction

Noncooperative differential game theory has been applied to a variety of control problems [1]–[14]. While zero-sum differential games have been heavily exploited in nonlinear H_∞ control theory, nonzero-sum differential games have had limited application in feedback control. In particular, Stackelberg differential games, which is based on a hierarchical relationship between the players, have been utilized in a decentralized control system [5], hierarchical control problems [3], [4], [12], and nonclassical control problems [6]. Differential games, as well as optimal control, are difficult tools to apply because of the challenges associated with determining analytical solutions for real-time implementation, with a few exceptions such as the linear quadratic structure. One way to incorporate optimal control and differential game structures is to formulate a system composed of control terms to feedback linearize and additional control terms to optimize the residual system. For example, optimal controller are developed with feedback linearization with exact model knowledge assumption [15] and via neural networks [16]–[18]. In [19] an open-loop Stackelberg game-based controller is developed based on the Robust Integral of the Sign of the Error (RISE) [20]–[22] technique.

The solution to a differential game consists of the optimal strategy (i.e., control input) of each player, the state trajectory propagated based on the players' strategies, and the corresponding costs. Ideally, the solution is good for the entire horizon. However, any changes in the game, such as changes in the objective of each player and different behavior of the system due to uncertainty and disturbances, compromise the optimality of the game strategies. This phenomenon is known as "time inconsistency" or "subgame-imperfection" in game theory [23]. There are strong and weak time inconsistencies, where the distinction is based on the initial conditions. The type of inconsistency addressed in this paper is associated with strong time inconsistency.

To account for time inconsistency, this paper extends our previous open-loop work in [19] to design a game-theoretic controller using a closed-loop Stackelberg strategy. In the differential game-theoretic sense, "open-loop" refers to a decision making of each player based on the initial condition, and "closed-loop" refers to the ability of the players to change their decisions based on current information. In Stackelberg games there is also a distinction between "closed-loop" and "feedback" strategies, where the former corresponds to the ability of the follower to change its strategy based on current information, and the latter corresponds to the ability of the leader to further change its strategy in reaction to the follower's closed-loop strategy.

This paper is aimed at investigating the development of the RISE controller in conjunction with a differential game-based controller with a closed-loop Stackelberg strategy, for uncertain Euler-Lagrange systems with additive disturbances. While the RISE controller partially feedback linearizes the Euler-Lagrange system, a closed-loop Stackelberg strategy is employed to minimize cost functionals associated with the residual dynamics. A Lyapunov analysis is used to prove semi-global asymptotic tracking. The result is described as an asymptotic optimal tracking result because the RISE controller asymptotically compensates for uncertainties and disturbances, eventually yielding a residual system for which the Stackelberg-derived component of the controller is used to minimize the corresponding cost functionals.

II. Dynamic Model and Properties

The class of nonlinear dynamic systems considered in this paper is assumed to be modeled by the following Euler-Lagrange formulation:

$$M(q)\ddot{q} + V_m(q,\dot{q})\dot{q} + G(q) + F(\dot{q}) + \tau_d(t) = \tau_L + \tau_F, \quad (1)$$

where $M(q) \in \mathbb{R}^{n\times n}$ denotes the generalized inertia matrix, $V_m(q,\dot{q}) \in \mathbb{R}^{n\times n}$ denotes the generalized gravity vector, $F(\dot{q}) \in \mathbb{R}^n$ denotes the generalized friction vector, $\tau_d \in \mathbb{R}^n$

This research is supported in part by the NSF award number 0901491.

T. Hiramatsu and M. Johnson are graduate students, and N. G. Fitz-Coy and W. E. Dixon are faculty members in the Dept. of Mechanical and Aerospace Engineering, University of Florida, Gainesville, Florida 32611, USA {takashi,marc1518,nfc,wdixon}@ufl.edu.

978-1-61284-799-3/11/$26.00 ©2011 IEEE

denotes a general uncertain disturbance, τ_L, $\tau_F \in \mathbb{R}^n$ denote the control input vectors, and $q(t)$, $\dot{q}(t)$, $\ddot{q}(t) \in \mathbb{R}^n$ denote the generalized position, velocity, and acceleration vectors, respectively. It is assumed that $q(t)$ and $\dot{q}(t)$ are measurable, and that $M(q)$, $V_m(q,\dot{q})$, $G(q)$, $F(\dot{q})$, and τ_d are unknown. In addition, the following assumptions are exploited in the subsequent development.

Assumption 1: The inertia matrix $M(q)$ is symmetric, positive-definite, and satisfies the following inequality:

$$m_1 \left\|\xi\right\|^2 \leq \xi^T M(q)\xi \leq \bar{m}(q) \left\|\xi\right\|^2, \quad \forall \xi(t) \in \mathbb{R}^n \tag{2}$$

where $m_1 \in \mathbb{R}$, $\bar{m}(q) \in \mathbb{R}$, and $\|\cdot\|$, respectively, denote a known positive constant, a known positive functions, and the standard Euclidean norm.

Assumption 2: The following skew-symmetric relationships are satisfied for $\forall \xi \in \mathbb{R}^n$:

$$\begin{aligned} \xi^T \left(\dot{M}(q) - 2V_m(q,\dot{q})\right)\xi = 0, \\ \xi^T \left(\dot{M}(q) - \left(V_m(q,\dot{q}) + V_m^T(q,\dot{q})\right)\right)\xi = 0. \end{aligned} \tag{3}$$

Note that in general $V_m \neq V_m^T$.

Assumption 3: If $q(t)$, $\dot{q}(t) \in \mathcal{L}_\infty$, then $V_m(q,\dot{q})$, $F(\dot{q})$, and $G(q)$ are bounded. Moreover, if $q(t)$, $\dot{q}(t) \in \mathcal{L}_\infty$, then the first and the second partial derivatives of the elements of $M(q)$, $V_m(q,\dot{q})$, and $G(q)$ with respect to $q(t)$ exist and are bounded, and the first and second partial derivatives of the elements of $V_m(q,\dot{q})$ and $F(\dot{q})$ with respect to $\dot{q}(t)$ exist and are bounded.

Assumption 4: The desired trajectory is assumed to be designed such that $q_d(t)$, $\dot{q}_d(t)$, $\ddot{q}_d(t)$, $\dddot{q}_d(t)$, and $\ddddot{q}_d(t) \in \mathbb{R}^n$ exist and are bounded.

Assumption 5: The disturbance term and its first two time derivatives (i.e., $\tau_d(t)$, $\dot{\tau}_d(t)$, $\ddot{\tau}_d(t)$) are bounded by known constants.

III. Error System Development

The control objective is to ensure that the system tracks a desired time-varying trajectory, denoted by $q_d(t) \in \mathbb{R}^n$, despite uncertainties in the dynamic model, while minimizing a given performance index. To quantify the tracking objective, a position tracking error, denoted by $e_1(t) \in \mathbb{R}^n$, is defined as

$$e_1 \triangleq q_d - q. \tag{4}$$

To facilitate the subsequent analysis, filtered tracking errors, denoted by $e_2(t)$, $r(t) \in \mathbb{R}^n$, are also defined as

$$e_2 \triangleq \dot{e}_1 + \alpha_1 e_1, \tag{5}$$

$$r \triangleq \dot{e}_2 + \alpha_2 e_2, \tag{6}$$

where $\alpha_1, \alpha_2 \in \mathbb{R}^{n\times n}$, are positive definite constant gain matrices. The filtered tracking error $r(t)$ is not measurable since the expressions in (6) depend on $\ddot{q}(t)$. The error systems are based on the assumptions that the generalized coordinates of the Euler-Lagrange dynamics allow additive errors instead of multiplicative errors (e.g., error quaternions).

A state-space model can be developed based on the tracking errors in (4) and (5). For this model, a controller is developed that minimizes a quadratic performance index under the temporary assumption that the dynamics in (1) are known. The feedback controller of interest is the solution to a two-person nonzero-sum differential game using a closed-loop Stackelberg strategy. The subsequent analysis then uses a robust controller to identify the unknown dynamics and additive disturbance, thereby relaxing the temporary assumption that these dynamics are known.

To develop a state-space model for the tracking errors in (4) and (5), the inertia matrix is premultiplied to the time derivative of (6), and substitutions are made from (1) and (4) to obtain

$$M\dot{e}_2 = -V_m e_2 - (\tau_L + \tau_F) + h + \tau_d, \tag{7}$$

where the nonlinear function $h(q,\dot{q},t) \in \mathbb{R}^n$ is defined as

$$h \triangleq M(\ddot{q}_d + \alpha_1 \dot{e}_1) + V_m(\dot{q}_d + \alpha_1 e_1) + G + F. \tag{8}$$

Under the temporary assumption that the dynamics in (1) are known, the control input is designed as

$$\tau_L + \tau_F \triangleq h + \tau_d - (u_L + u_F), \tag{9}$$

where u_L denotes the leader's input and u_F denotes the follower's input that will be respectively designed to minimize their performance indices. By substituting (9) into (7) the closed-loop error system for $e_2(t)$ can be obtained as

$$M\dot{e}_2 = -V_m e_2 + u_L + u_F. \tag{10}$$

From (5) and (10), a state-space model for e_1 and e_2 are developed as

$$\dot{z} = Az + Bu_F + Bu_L, \tag{11}$$

where

$$\begin{aligned} A(q,\dot{q}) &\triangleq \begin{bmatrix} -\alpha_1 & I_{n\times n} \\ 0_{n\times n} & -M^{-1}V_m \end{bmatrix}, \\ B(q) &\triangleq \begin{bmatrix} 0_{n\times n} & M^{-1} \end{bmatrix}^T, \\ z(t) &\triangleq \begin{bmatrix} e_1^T & e_2^T \end{bmatrix}^T, \end{aligned} \tag{12}$$

where $I_{n\times n}$ and $0_{n\times n}$ denote the $n \times n$ identity matrix and the matrix of zeros, respectively.

IV. Closed-Loop Stackelberg Game Control

Stackelberg games provide a framework for systems that operate on different levels with a prescribed hierarchy of decisions. For a two-person Stackelberg game where the system is affected by two decision makers, the problem is cast in two solution spaces: the leader and the follower, where each player tries to minimize their respective cost functionals. The leader is a decision maker that can enforce its strategy to minimize its objective metric over the follower. For example, when two inputs affect the behavior of a system, the one with more rapid dynamics can be considered the leader in Stackelberg structure; since the system responds more rapidly to the leader's control input, it is reasonable to put more weight in optimizing the leader's control strategy while making the follower compromise.

A Stackelberg differential game problem, with u_F as the follower and u_L as the leader, is formulated by a differential constraint and the cost functionals $J_1(z, u_F, u_L)$, $J_2(z, u_F, u_L) \in \mathbb{R}$ as

$$\begin{aligned} \dot{z} &= Az + Bu_F + Bu_L, \\ J_F &= \frac{1}{2}\int_{t_0}^{\infty} \left(z^T Q z + u_F^T R_{11} u_F + u_L^T R_{12} u_L\right) dt, \\ J_L &= \frac{1}{2}\int_{t_0}^{\infty} \left(z^T N z + u_F^T R_{21} u_F + u_L^T R_{22} u_L\right) dt, \end{aligned} \tag{13}$$

where Q, $N \in \mathbb{R}^{2n\times 2n}$ are symmetric constant matrices defined as

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & N_{12} \\ N_{12}^T & N_{22} \end{bmatrix}, \tag{14}$$

and Q_{ij}, $N_{ij} \in \mathbb{R}^{n\times n}$ are positive definite and symmetric constant matrices $\forall i,\ j = 1,\ 2$. A closed-loop solution is sought by extending [19]. Unlike the open-loop case, the follower assumes that the leader's strategy explicitly affects the system. With the game being of linear-quadratic structure, the following assumption is made.

Assumption 6: In computing its strategy, the follower assumes that the leader's strategy is linear in the states such that

$$u_L = F_2 z,$$

where $F_2(t) \in \mathbb{R}^{2n\times 2n}$ such that the follower's problem is written as

$$\begin{aligned} \dot{z} &= (A + BF_2)z + Bu_F, \\ J_F &= \frac{1}{2}\int_{t_0}^{\infty} \left(z^T \left(Q + F_2^T R_{12} F_2\right) z + u_F^T R_{11} u_F\right) dt. \end{aligned}$$

The Hamiltonian of the follower is

$$\begin{aligned} H_F = &\frac{1}{2}\left(z^T (Q + F_2^T R_{12} F_2) z + u_F^T R_{11} u_F\right) \\ &+ \lambda_1^T \left((A + BF_2)x + Bu_F\right), \end{aligned}$$

where the optimal control strategy and the costate equation of the follower are obtained as

$$u_F = -R_{11}^{-1} B^T \lambda_1, \tag{15}$$

$$\dot{\lambda}_1 = -(Q + F_2^T R_{12} F_2)^T z - (A + BF_2)^T \lambda_1. \tag{16}$$

Substituting (15) and (16) into the dynamics and $J_2(z, u_F, u_L, t)$ yields an optimal control problem of the leader

$$\begin{aligned} \dot{z} &= Az - BR_{11}^{-1} B^T \lambda_1 + Bu_L, \\ J_L &= \frac{1}{2}\int_{t_0}^{\infty} \left(z^T N z + u_L^T R_{22} u_L \right. \\ &\quad \left. + \lambda_1^T B R_{11}^{-1} R_{21} R_{11}^{-1} B^T \lambda_1\right) dt, \end{aligned}$$

where the Hamiltonian of the leader is constructed as

$$\begin{aligned} H_L = &\frac{1}{2}\left(z^T N z + \lambda_1^T B R_{11}^{-1} R_{21} R_{11}^{-1} B^T \lambda_1 + u_L^T R_{22} u_L\right) \\ &+ \lambda_2^T \left(Az - BR_{11}^{-1} B^T \lambda_1 + Bu_L\right) \\ &+ \psi^T \left(-(Q + F_2^T R_{12} F_2)^T z - (A + BF_2)^T \lambda_1\right), \end{aligned}$$

where

$$u_L = -R_{22}^{-1} B^T \lambda_2, \tag{17}$$

$$\dot{\lambda}_2 = -N^T x - A^T \lambda_2 + (Q + F_2^T R_{12} F_2)\psi, \tag{18}$$

$$\begin{aligned} \dot{\psi} = &-BR_{11}^{-1} R_{21} R_{11}^{-1} B^T \lambda_1 + BR_{11}^{-1} B^T \lambda_2 \\ &+ (A + BF_2)\psi. \end{aligned} \tag{19}$$

The expressions derived in (15)-(19) define the solution to the differential game. The subsequent analysis aims at developing an expression for the costate variables ($\lambda_1(t)$, $\lambda_2(t)$, $\psi(t)$) which can be implemented by the controllers $u_F(t)$ and $u_L(t)$. Suppose that the costates are linear in the state:

$$\lambda_1 = Kz, \tag{20}$$

$$\lambda_2 = Pz, \tag{21}$$

$$\psi = Sz, \tag{22}$$

where $K(t)$, $P(t)$, $S(t) \in \mathbb{R}^{2n\times 2n}$ are time-varying positive definite diagonal matrices. Given these assumed solutions, conditions and constraints are developed to ensure (20)-(22) satisfy (16), (18), and (19). Differentiating (20)-(22) and substituting the dynamic constraint in (13) along with (15)-(19) yields three differential Riccati equations

$$\begin{aligned} 0 = &\dot{K} + KA - KBR_{11}^{-1} B^T K - KBR_{22}^{-1} B^T P + Q \\ &+ PBR_{22}^{-1} R_{12} R_{22}^{-1} B^T P + A^T K - PBR_{22}^{-1} B^T K, \end{aligned} \tag{23}$$

$$\begin{aligned} 0 = &\dot{P} + PA - PBR_{11}^{-1} B^T K - PBR_{22}^{-1} B^T P + N \\ &+ A^T P - QS - PBR_{22}^{-1} R_{12} R_{22}^{-1} B^T PS, \end{aligned} \tag{24}$$

$$\begin{aligned} 0 = &\dot{S} + SA - SBR_{11}^{-1} B^T K - SBR_{22}^{-1} B^T P \\ &+ BR_{11}^{-1} R_{21} R_{11}^{-1} B^T K - BR_{11}^{-1} B^T P \\ &- AS + BR_{22}^{-1} B^T PS. \end{aligned} \tag{25}$$

Equations (23)-(25) can be expressed as open-loop Riccati equations plus additional terms. From [19] the open-loop Riccati equations are

$$\begin{aligned} 0 = &\dot{K} + KA + A^T K - KBR_{11}^{-1} B^T K \\ &- KBR_{22}^{-1} B^T P + Q^T, \end{aligned} \tag{26}$$

$$\begin{aligned} 0 = &\dot{P} + PA + A^T P - PBR_{11}^{-1} B^T K \\ &- PBR_{22}^{-1} B^T P + N^T - QS, \end{aligned} \tag{27}$$

$$\begin{aligned} 0 = &\dot{S} + SA - AS - SBR_{11}^{-1} B^T K - SBR_{22}^{-1} B^T P \\ &+ BR_{11}^{-1} R_{21} R_{11}^{-1} B^T K - BR_{11}^{-1} B^T P. \end{aligned} \tag{28}$$

Let the subscripts CRE and ORE denote the closed-loop (23)-(25) and open-loop (26)-(28) Ricatti equations, respectively. Then the closed-loop Riccati equations can be written as

$$K_{CRE} = K_{ORE} + PBR_{22}^{-1}R_{12}R_{22}^{-1}B^TP - PBR_{22}^{-1}B^TK = 0, \tag{29}$$

$$P_{CRE} = P_{ORE} - PBR_{22}^{-1}R_{12}R_{22}^{-1}B^TPS = 0, \tag{30}$$

$$S_{CRE} = S_{ORE} + BR_{22}^{-1}B^TPS = 0. \tag{31}$$

In (29)-(31), $K(t)$ and $P(t)$ correspond to $u_F(t)$ and $u_L(t)$ respectively, while $S(t)$ places constraints between the $K(t)$ and $P(t)$. Equations (29)-(31) must be solved simultaneously to yield Stackelberg control strategies for the leader and the follower. If $P(t)$, $K(t)$, and $S(t)$ are selected as

$$P = \begin{bmatrix} P_{11} & 0_{n\times n} \\ 0_{n\times n} & M \end{bmatrix}, \tag{32}$$

$$K = \begin{bmatrix} K_{11} & 0_{n\times n} \\ 0_{n\times n} & M \end{bmatrix}, \tag{33}$$

$$S = \begin{bmatrix} S_{11} & 0_{n\times n} \\ 0_{n\times n} & -2I_{n\times n} \end{bmatrix}, \tag{34}$$

where K_{11} and P_{11} satisfy

$$K_{11} = -\frac{1}{2}\left(Q_{12} + Q_{12}^T\right), \quad P_{11} = -\frac{1}{2}\left(N_{12} + N_{12}^T\right) + 2K_{11}, \tag{35}$$

then (23)-(25) are solved with the following constraints on the cost functionals:

$$\begin{aligned} \frac{1}{2}\left[\left(Q_{12} + Q_{12}^T\right)\alpha_1 + \alpha_1^T\left(Q_{12} + Q_{12}^T\right)\right] + Q_{11} = 0, \\ \frac{1}{2}\left[\left(N_{12} + N_{12}^T\right)\alpha_1 + \alpha_1^T\left(N_{12} + N_{12}^T\right)\right] + N_{11} = 0, \\ Q_{22} + R_{22}^{-1} + R_{11}^{-1}R_{21}R_{11}^{-1} = 0, \\ -R_{11}^{-1} - 2R_{22}^{-1} + Q_{22} + R_{22}^{-1}R_{12}R_{22}^{-1} = 0, \\ Q_{22} + N_{22} = 0. \end{aligned} \tag{36}$$

From (15), (17), (20), (21), (32), and (33), the closed-loop Stackelberg game-based controllers are obtained as

$$u_F = -R_{11}^{-1}e_2, \tag{37}$$

$$u_L = -R_{22}^{-1}e_2. \tag{38}$$

Note that the solution has the same form as the open-loop problem except that more conservative constraints are placed on the relationship among the gain matrices. In particular, constraints in (36) include R_{12}, which affects the decision of $u_F(t)$ due to the decision of $u_L(t)$ for the closed-loop case. Therefore, a closed-loop strategy for the follower is better than an open-loop strategy in addressing the time inconsistency.

V. RISE FEEDBACK CONTROL DEVELOPMENT

In general, the bounded disturbance $\tau_d(t)$ and the nonlinear dynamics given in (8) are unknown, so the controller given in (9) cannot be implemented. However, if the control input can identify and cancel these effects, then $z(t)$ will converge to the state space model in (11) such that that u_F and u_L minimize the respective performance index J_F and J_L. In this section, a control input is developed that exploits RISE feedback to identify the nonlinear effects and bounded disturbances thus enabling $z(t)$ to asymptotically converge to the state space model in (11).

To develop the control input, the error system in (6) is premultiplied by $M(q)$ and the expressions in (1), (4), and (5) are utilized to obtain

$$Mr = -V_me_2 + h + \tau_d + \alpha_2Me_2 - (\tau_F + \tau_L). \tag{39}$$

Based on the open-loop error system in (39), the control input is composed of the game theoretic controllers developed in (37) and (38), plus a subsequently designed auxiliary control term $\mu(t) \in \mathbb{R}^n$ as

$$(\tau_F + \tau_L) \triangleq \mu - (u_F + u_L). \tag{40}$$

The closed-loop tracking error system can be developed by substituting (40) into (39) as

$$Mr = -V_me_2 + h + \tau_d + \alpha_2Me_2 + (u_F + u_L) - \mu. \tag{41}$$

To facilitate the subsequent stability analysis the auxiliary function $f_d(t) \in \mathbb{R}^n$, which is defined as

$$f_d \triangleq M(q_d)\ddot{q}_d + V_m(q_d, \dot{q}_d)\dot{q}_d + G(q_d) + F(\dot{q}_d), \tag{42}$$

is added and subtracted to (41) to yield

$$Mr = -V_me_2 + \bar{h} + f_d + \tau_d + (u_F + u_L) - \mu + \alpha_2Me_2, \tag{43}$$

where $\bar{h} \in \mathbb{R}^n$ is defined as

$$\bar{h} \triangleq h - f_d.$$

Substituting (38) into (43), taking the time derivative, and manipulating with (6) yields

$$M\dot{r} = -\frac{1}{2}\dot{M}r + \tilde{N} + N_D - e_2 - \left(R_{11}^{-1} + R_{22}^{-1}\right)r - \dot{\mu}, \tag{44}$$

after strategically grouping specific terms. In (44), the unmeasurable auxiliary terms $\tilde{N}(e_1, e_2, r, t)$, $N_D(t) \in \mathbb{R}^n$ are defined as

$$\begin{aligned} \tilde{N} &\triangleq -\dot{V}_me_2 - V_m\dot{e}_2 - \frac{1}{2}\dot{M}r + \dot{\bar{h}} + \alpha_2\dot{M}e_2 \\ &\quad + \alpha_2M\dot{e}_2 + e_2 + \left(R_{11}^{-1} + R_{22}^{-1}\right)\alpha_2e_2, \\ N_D &\triangleq \dot{f}_d + \dot{\tau}_d. \end{aligned}$$

The Mean Value Theorem and Assumptions 3, 4, and 5 can be used to upper bound the auxiliary terms as

$$\left\|\tilde{N}(t)\right\| \leq \rho\left(\|y\|\right)\|y\|, \; \|N_D\| \leq \zeta_1, \; \left\|\dot{N}_D\right\| \leq \zeta_2, \tag{45}$$

where $y(t) \in \mathbb{R}^{3n}$ is defined as

$$y(t) \triangleq \begin{bmatrix} e_1^T & e_2^T & r^T \end{bmatrix}^T,$$

the bounding function $\rho\left(\|y\|\right) \in \mathbb{R}$ is a positive globally invertible nondecreasing function, and $\zeta_i \in \mathbb{R}$, $i = 1, 2$, denote known positive constants. Based on (44), the control term $\mu(t)$ is designed as the generalized solution to

$$\dot{\mu}(t) \triangleq k_sr(t) + \beta_1\mathrm{sgn}(e_2), \tag{46}$$

where k_s, $\beta_1 \in \mathbb{R}$ are positive constant control gains. The closed-loop error systems for $r(t)$ can now be obtained by substituting (46) into (44) as

$$M\dot{r} = -\frac{1}{2}\dot{M}r + \tilde{N} + N_D - e_2 - \left(R_{11}^{-1} + R_{22}^{-1}\right) r - k_s r - \beta_1 \mathrm{sgn}(e_2). \tag{47}$$

VI. Stability Analysis

It can be shown that the controller given by (37), (38), (40), and (46) ensures that all system signals are bounded under closed-loop operation, and the tracking error are regulated in the sense that (see [19] for similar details)

$$\|e_1(t)\|, \|e_2(t)\|, \|r(t)\| \to 0 \quad \text{as} \quad t \to \infty. \tag{48}$$

The boundedness of the closed-loop signals and the result in (48) can be obtained provided the control gain k_s introduced in (46) is selected sufficiently large (see the subsequent stability analysis), and α_1, α_2 are selected according to the sufficient conditions

$$\lambda_{\min}(\alpha_1) > \frac{1}{2}, \quad \lambda_{\min}(\alpha_2) > 1, \tag{49}$$

where $\lambda_{\min}(\alpha_1)$ and $\lambda_{\min}(\alpha_2)$ are the minimum eigenvalues of α_1 and α_2, respectively. The gain β_1 is selected according to the following sufficient condition:

$$\beta_1 > \zeta_1 + \frac{\zeta_2}{\lambda_{\min}(\alpha_2)}. \tag{50}$$

Let a Lyapunov function $V_L(\Phi, t) : \mathcal{D} \times [0, \infty) \to \mathbb{R}$ be a continuously differentiable positive definite function defined in [19] as

$$V_L(\Phi, t) \triangleq \|e_1\|^2 + \frac{1}{2}\|e_2\|^2 + \frac{1}{2} r^T M r + O. \tag{51}$$

where the auxiliary function $O(t) \in \mathbb{R}$ is the solution to (see [19] for further details)

$$\begin{aligned} \dot{O} &\triangleq -r^T \left(N_D - \beta_1 \mathrm{sgn}(e_2)\right), \\ O(0) &= \beta_1 \sum_{i=1}^{n} |e_{2i}(0)| - e_2(0)^T N_D(0). \end{aligned} \tag{52}$$

Taking the time derivative of (51) yields

$$\dot{V}_L = 2e_1^T \dot{e}_1 + e_2^T \dot{e}_2 + \frac{1}{2} r^T \dot{M} r + r^T M \dot{r} + \dot{O}.$$

Utilizing (5), (6), (47), and (52), the Lyapunov derivative is rewritten as

$$\begin{aligned} \dot{V}_L(\Phi, t) \leq& -2e_1^T \alpha_1 e_1 + 2e_2^T e_1 + r^T \tilde{N} \\ &- \left\{k_s + \lambda_{\min}\left(R_{11}^{-1} + R_{22}^{-1}\right)\right\} \|r\|^2 \\ &- \lambda_{\min}(\alpha_2) \|e_2\|^2. \end{aligned} \tag{53}$$

Utilizing (45), (53) can be further simplified as

$$\dot{V}_L \leq -\lambda_3 \|y\|^2 - \left[k_s \|r\|^2 - \rho(\|y\|) \|r\| \|y\|\right], \tag{54}$$

where

$$\lambda_3 \triangleq \min \left\{ \begin{array}{c} 2\lambda_{\min}(\alpha_1) - 1 \\ \lambda_{\min}(\alpha_2) - 1 \\ \lambda_{\min}\left(R_{11}^{-1} + R_{22}^{-2}\right) \end{array} \right\}. \tag{55}$$

Completing the squares for the terms inside the brackets in (54) yields

$$\dot{V}_L \leq -\lambda_3 \|y\|^2 + \frac{\rho^2(\|y\|) \|y\|^2}{4k_s} \leq -U(\Phi), \tag{56}$$

where $U(\Phi) = c\|y\|^2$ for some positive constant c. The function $U(\Phi)$ is a continuous, positive semi-definite and defined within the closed set:

$$\mathcal{D} \triangleq \left\{ \Phi \in \mathbb{R}^{3n+1} \middle| \ \|\Phi\| \leq \rho^{-1}\left(2\sqrt{\lambda_3 k_s}\right) \right\}.$$

The inequality in (56) can be used to show that $V_L(\Phi, t) \in \mathcal{L}_\infty$ in $\mathcal{D}$; hence, $e_1(t)$, $e_2(t)$, and $r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Then standard linear analysis methods can be used to prove that $\dot{e}_1(t)$, $\dot{e}_2 \in \mathcal{L}_\infty$ in $\mathcal{D}$ from (5)-(6). Since $e_1(t)$, $e_2(t)$, $r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, Assumption 4 is used along with (5)-(6) to conclude that $q(t)$, $\dot{q}(t)$, $\ddot{q}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, which is then combined with Assumption 3 to conclude that $M(q)$, $V_m(q, \dot{q})$, $G(q)$, and $F(\dot{q}) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Thus, from (1) and Assumption 4, it can be shown that $\tau_L(t)$, $\tau_F(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. With $r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, it can be shown that $\dot{\mu}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$; hence, (47) can be used to show that $\dot{r}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. From $\dot{e}_1(t)$, $\dot{e}_2(t)$, $\dot{r} \in \mathcal{L}_\infty$ in $\mathcal{D}$, the definitions for $U(y)$ and $z(t)$ can be used to prove that $U(y)$ is uniformly continuous in $\mathcal{D}$.

Using similar arguments as given in [19] it can be shown that

$$c\|y(t)\|^2 \to 0 \quad \text{as} \quad t \to \infty \quad \forall y(0) \in \mathcal{S}. \tag{57}$$

Based on the definition of $y(t)$, (57) can be used to conclude that Theorem 1 holds for all $y(0) \in \mathcal{S}$.

Since $u_F(t)$, $u_L(t) \to 0$ as $e_2(t) \to 0$ from (37) and (38), then (43) can be used to conclude that

$$\mu \to \bar{h} + f_d + \tau_d \quad \text{as} \quad r(t),\ e_2(t) \to 0. \tag{58}$$

Equation (58) indicates that the dynamics in (1) converge to the state-space model in (11). Hence, $u_F(t)$ and $u_L(t)$ converges to an optimal controller to solve the game defined in (13), provided the gain constraints in (36) are satisfied.

VII. Conclusion

A closed-loop Stackelberg-based feedback controller for an Euler-Lagrange system subject to state dependent and bounded disturbances is developed to alleviate limitations due to time inconsistency. Asymptotic optimality of the proposed controller is achieved through a two-level architecture: the RISE controller yields a residual dynamical model by compensating for nonlinear uncertainties, and then the closed-loop Stackelberg-based controller minimizes cost functionals for the residual hierarchical system. Using a Lyapunov stability analysis and a Stackelberg game development, sufficient gain conditions were derived to ensure asymptotic tracking while minimizing the cost functionals. The developed controller provides a more robust solution than our previous open-loop method. The advantages of the closed-loop method will be examined through experimental results in future development. Furthermore, future efforts will also focus on the development of feedback solutions which will enable the leader to respond to changes in the follower's policy.

References

[1] T. Basar and G. Olsder, *Dynamic Noncooperative Game Theory*. SIAM, PA, 1999.

[2] M. Bloem, T. Alpcan, and T. Başar, "A Stackelberg game for power control and channel allocation in cognitive radio networks," in *Proc. Workshop on Game Theory in Communication Networks (GameComm)*, Nantes, France, October 2007.

[3] T. Basar and H. Selbuz, "Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems," *IEEE Trans. Autom. Control*, vol. 24 no. 2, pp. 166–179, 1979.

[4] J. Medanic, "Closed-loop Stackelberg strategies in linear-quadratic problems," *IEEE Trans. Autom. Control*, vol. 23 no. 4, pp. 632–637, 1978.

[5] M. Simaan and J. Cruz, J., "A Stackelberg solution for games with many players," *IEEE Trans. Autom. Control*, vol. 18, no. 3, pp. 322 – 324, 1973.

[6] G. Papavassilopoulos and J. Cruz, "Nonclassical control problems and Stackelberg games," *IEEE Trans. Autom. Control*, vol. 24 no. 2, pp. 155–166, 1979.

[7] A. Gambier, A. Wellenreuther, and E. Badreddin, "A new approach to design multi-loop control systems with multiple controllers," in *Proc. IEEE Conf. Decis. Control*, 13-15 2006, pp. 1828 –1833.

[8] J. Hongbin and C. Y. Huang, "Non-cooperative uplink power control in cellular radio systems," *Wireless Networks*, vol. 4 no. 3, pp. 233–240, 1998.

[9] T. Basar and P. Bernhard, *H-infinity Optimal Control and Related Minimax Design Problems*. Boston: Birkhuser, 2008.

[10] A. Isidori and A. Astolfi, "Disturbance attenuation and H-infinity-control via measurement feedback in nonlinear systems," *IEEE Trans. Autom. Control*, vol. 37, no. 9, pp. 1283–1293, Sept. 1992.

[11] L. Pavel, "A noncooperative game approach to OSNR optimization in optical networks," *IEEE Trans. Autom. Control*, vol. 51 no. 5, pp. 848–852, 2006.

[12] C. Tomlin, J. Lygeros, and S. Sastry, *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 1998, ch. 27, pp. 360–373.

[13] T. Basar and G. Olsder, "Team-optimal closed loop Stackelberg strategies in hierarchical control problems," *Automatica*, vol. 16 no. 4, pp. 409–414, 1980.

[14] M. Jungers, E. Trelat, and H. Abou-Kandil, "Min-max and min-min Stackelberg strategy with closed-loop information," *HAL Hyper Articles en Ligne*, vol. 3, 2010.

[15] R. Johansson, "Quadratic optimization of motion coordination and control," *Automatic Control, IEEE Transactions on*, vol. 35, no. 11, pp. 1197–1208, 1990.

[16] Y. Kim and F. Lewis, "Optimal design of CMAC neural-network controller for robot manipulators," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 30, no. 1, pp. 22 –31, feb 2000.

[17] Y. Kim, F. Lewis, and D. Dawson, "Intelligent optimal control of robotic manipulators using neural networks," *AUTOMATICA-OXFORD-*, vol. 36, pp. 1355–1364, 2000.

[18] M. Abu-Khalaf and F. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779–791, 2005.

[19] M. Johnson, T. Hiramatsu, N. Fitz-Coy, and W. E. Dixon, "Asymptotic stackelberg optimal control design for an uncertain euler-lagrange system," in *Proc. IEEE Conf. Decis. Control*, Atlanta, GA, 2010, pp. 6686–6691.

[20] Z. Cai, M. S. de Queiroz, and D. M. Dawson, "Robust adaptive asymptotic tracking of nonlinear systems with additive disturbance," *IEEE Trans. Autom. Control*, vol. 51, pp. 524–529, 2006.

[21] C. Makkar, G. Hu, W. G. Sawyer, and W. E. Dixon, "Lyapunov-based tracking control in the presence of uncertain nonlinear parameterizable friction," *IEEE Trans. Autom. Control*, vol. 52, no. 10, pp. 1988–1994, 2007.

[22] P. M. Patre, W. MacKunis, K. Kaiser, and W. E. Dixon, "Asymptotic tracking for uncertain dynamic systems via a multilayer neural network feedforward and RISE feedback control structure," *IEEE Trans. Autom. Control*, vol. 53, no. 9, pp. 2180–2185, 2008.

[23] J. Engwerda, *LQ dynamic optimization and differential games*. Wiley, 2005.