# On Limitations to the Achievable Path Following Performance for Linear Multivariable Plants

Daniel E. Miller, *Senior Member, IEEE*, and Richard H. Middleton, *Fellow, IEEE*

*Abstract*—In this paper, we consider a problem termed "path following". This differs from the common problem of reference tracking, in that here we can adjust the speed at which we traverse the reference trajectory. We are interested in ascertaining the degree to which we can track a given trajectory, and in characterizing the class of paths for which we can generate an appropriate temporal specification so that the path can be tracked arbitrarily well in an $L_2$ sense. We give various bounds on the achievable performance, as well as tight results in special cases. In addition, we give a numerical procedure based on convex optimization for computing the achievable performance. The results demonstrate that there are situations where arbitrarily good $L_2$ performance may be achieved even though the origin is not in the convex hull of the positive limit set of the path to be followed.

*Index Terms*—$L_2$ performance limitations, nonminimum phase systems, path following, tracking.

## I. INTRODUCTION

THE problem of tracking—causing the output of a dynamic system to follow a commanded trajectory—has long been of interest in the control literature. For stably invertible linear systems, a combination of feedback and feedforward actions can be used to ensure that high quality tracking performance is achieved for a broad class of reference signals, e.g. see [22]. This is complemented by the work of [10], where, for a more restrictive class of command signals, tracking can be achieved for a much broader class of systems. Tracking problems have also been considered in a nonlinear systems context, e.g. see [14].

In these tracking problems, it was realized some time ago (see for example [18] for the linear case, or [20] for a nonlinear version of these results) that the lack of a stable inverse (or the existence of unstable zero dynamics) limits the achievable tracking performance. It is therefore of interest to investigate circumstances under which additional information or an alternate problem formulation can be used to enhance the tracking performance.

One means by which tracking performance restrictions might be relaxed is the use of "preview" control, wherein advance knowledge of the trajectory to be tracked permits improved performance. This approach, which can be viewed as a form of

anti-causal (or anticipative) feedforward action, has been investigated in, for example, [5] and [6].

Another means of removing some tracking performance restrictions is to alter the tracking performance specification. In standard trajectory tracking a command signal $r(t)$ is specified at every time instant for the system output $y(t)$. In some situations, such as in the problem of steering an object (such as a ship, a robot, or a cutting tool), the primary objective is to follow a certain path, with the speed at which the path is traversed being of secondary importance. This has lead to an alternate problem formulation called "path following," e.g. see [1], [7], [11], [12].

In this path following formulation one introduces a "path variable" (or "timing signal") $\theta(t)$, which creates an extra degree of freedom, so that the goal is to make $y(t) - r(\theta(t))$ small rather than $y(t) - r(t)$ small. Problems of this form have been considered for some time, with recent work including [1]–[3], [8], [21], and [7], wherein it is shown that such *path following* problems permit significant improvement in the achievable performance. One fundamental objective is to ascertain sufficient conditions on $r$ to guarantee that the path following performance (measured in the 2-norm sense) can be made as small as desired (by a suitable choice of control and path variable or timing signal)—the nonlinear case is considered in [8] and [3] with the linear case considered in [1], [2] and [7]. In particular, in the linear case it is proven in [2] that it is sufficient that $r(t)$ be a finite sum of sinusoids (of non-zero frequency) and it is proven in [7] that (roughly speaking) it is sufficient that the path be repeatable and that the origin lie in the interior of its convex hull; both approaches are constructive, either by fixing a linear feedback structure and adjusting the timing variable in a feedforward way [2], or by regarding the timing variable and the control signal as two separate inputs and recasting the problem into a nonlinear feedback setting [7]. Our approach differs significantly in that we focus purely on analysis rather than synthesis, so we recast the problem as one of optimization, and use real analysis and optimal control theory as tools.

Here we consider the linear case and we are interested in determining the class of trajectories for which we can make the cost as small as desired. In addition, we would like to develop tools to bound and compute the (infimal) optimal cost when it is non-zero. Our focus is analysis, although synthesis is achievable with additional work. We provide conditions, weaker than those previously available in the literature, that ensure that the cost can be made as small as desired, and in a reasonably general situation we show how to compute the optimal cost by solving a finite dimensional convex optimization problem.

Now for a brief outline of the paper. As mentioned above, the main goal is to obtain answers to two key questions, namely (i) For which trajectories can we make the cost as small as desired? and (ii) How do we compute the (infimal) optimal cost when it

is non-zero? In the next section we give a problem statement together with a definition of the cost function to be minimized. In Section III we simplify this cost function in a sequence of three steps, thereby converting it into a classical optimization problem, albeit an infinite dimensional one. In Sections IV and V we focus on answering the first question: in Section IV we use the new problem formulation of Section III to provide upper and lower bounds on the cost function, while in Section V we home in on an important special case (scalar input with zero initial conditions) and provide several necessary and sufficient conditions for which the infimal cost is zero. In Section VI we focus on the second question: under reasonably general conditions we use results from optimal control theory to convert the infinite dimensional optimization problem derived in Section III to a finite-dimensional strictly convex optimization problem, which is readily computable using standard software. In Section VII we present an illustrative example, while in Section VIII we conclude with a summary of the main results.

## A. Preliminary Mathematics

Let $\mathbf{R}$ and $\mathbf{R}^+$ denote the set of real and non-negative real numbers respectively. We use the Hölder 2-norm to measure the size of $x \in \mathbf{R}^n$, which we represent by $\|x\|$; with $A \in \mathbf{R}^{n \times m}$, the corresponding induced norm is denoted by $\|A\|$, and $\sigma_{\min}(A)$ denotes the smallest singular value of $A$. With $A \in \mathbf{R}^{n \times n}$, $sp(A)$ denotes the set of eigenvalues of $A$.

We let $PC(\mathbf{R}^n)$ denote the set of $\mathbf{R}^n$-valued piecewise continuous functions on $\mathbf{R}^+$, let $PC_\infty(\mathbf{R}^n)$ denote the subset of bounded elements, and for $k$ a non-negative integer we let $\mathcal{C}^k(\mathbf{R}^n)$ denote the set of $\mathbf{R}^n$-valued signals on $\mathbf{R}^+$ which are $k$ times continuously differentiable. We let $L_2(\mathbf{R}^n)$ denote the set of $\mathbf{R}^n$-valued, Lebesgue measureable, square integrable functions and use $\|f\|_2$ to denote the norm of $f \in L_2$; we let $L_\infty$ denote the set of $\mathbf{R}^n$-valued, Lebesgue measureable, bounded functions, and use $\|f\|_\infty$ to to denote the norm of $f \in L_\infty$. Henceforth we write $f \in PC$, $f \in PC_\infty$, $f \in \mathcal{C}^k$, $f \in L_2$ and $f \in L_\infty$, as appropriate; the dimension will be clear from the context.

## II. THE SETUP

Here we consider a square plant model of the form

$$\left.\begin{aligned} \dot{x}_p &= A_p x_p + B_p u, \ x_p(0) = x_{p_0} \\ y &= C_p x_p, \end{aligned}\right\} \quad (1)$$

with $x_p(t) \in \mathbf{R}^{n_p}$ and $u(t), y(t) \in \mathbf{R}^m$. We assume that $(A_p, B_p)$ is controllable and $(C_p, A_p)$ is observable. We also impose

*Standing Assumption 1:* The system has no transmission zeros (in the sense of [9]) on the imaginary axis, i.e.,

$$\text{rank} \begin{bmatrix} A_p - j\omega I & B_p \\ C_p & 0 \end{bmatrix} = n_p + m, \ \omega \in \mathbf{R}. \quad (2)$$

## A. Isolating the Zero Dynamics

Following a linear version of [14], under some modest assumptions we can transform the plant to isolate the zero dynamics. For simplicity, we derive this in the case of "uniform vector relative degree"; the more general vector relative degree

case can be carried out with similar analysis to that which follows, but with substantially more complexity.

*Definition 1:* The transfer function $C(sI - A)^{-1}B$ is said to have *uniform vector relative degree* $p$ if

$$CA^j B = 0, \ j = 0, 1, \ldots, p - 2$$

and $CA^{p-1}B$ is nonsingular.

*Standing Assumption 2:* $C_p(sI - A_p)^{-1}B_p$ has uniform vector relative degree $p \in \mathbf{N}$.

With

$$v^T := [y^T \quad (y^{(1)})^T \quad \cdots \quad (y^{(p-1)})^T] \quad (3)$$

it can be shown that we can choose $x = Tx_p$ with $T \in \mathbf{R}^{(n_p - mp) \times n_p}$ so that the plant model can be rewritten in the form

$$\left.\begin{aligned} \begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} &= \begin{bmatrix} A & BC \\ A_x & A_v \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ B_v \end{bmatrix} u \\ y &= [0 \quad C] \begin{bmatrix} x \\ v \end{bmatrix} \end{aligned}\right\} \quad (4)$$

with $B_v$ full column rank. Since this system is obtained by a similarity transformation, it is easy to see that the transmission zeros are identical to those of (1) and are precisely the eigenvalues of the submatrix $A$ in (4). Since we have assumed that there are none on the imaginary axis, i.e., (2) holds, we may as well assume that $T$ has been chosen so that $A$, $B$ and $A_x$ can be partitioned as

$$A = \begin{bmatrix} A^+ & 0 \\ 0 & A^- \end{bmatrix}, B = \begin{bmatrix} B^+ \\ B^- \end{bmatrix}, A_x = [A_1 \quad A_2]$$

with $A^-$ and $-A^+$ Hurwitz. We partition $x$ accordingly as

$$x = \begin{bmatrix} x^+ \\ x^- \end{bmatrix}.$$

The plant model can now be rewritten as

$$\left.\begin{aligned} \dot{x}^+ &= A^+ x^+ + B^+ y, \ x^+(0) = x_0^+ \\ \dot{x}^- &= A^- x^- + B^- y, \ x^-(0) = x_0^- \\ \dot{v} &= A_v v + A_1 x^+ + A_2 x^- + B_v u, \ v(0) = v_0. \\ y &= Cv \end{aligned}\right\} \cdot (5)$$

In this setting the initial condition $x_{p_0}$ on (1) translates to an initial condition of $\begin{bmatrix} x_0^+ \\ x_0^- \\ v_0 \end{bmatrix}$ on (5). The sub-states $x^+$ and $x^-$ are associated with the zero dynamics of the plant; it turns out that $x^+$ plays the critical role here, and we let $n$ denote its dimension.

## B. The Control Objective

In this paper the objective will be to control the system (1), with an initial condition of $x_{p_0}$, in order to follow a pre-specified path $r$, parametrized by the timing signal $\theta$, while requiring that all signals remain bounded; the **path following error** is defined by

$$e(t) := r(\theta(t)) - y(t).$$

The objective is different from the common one of reference tracking: we are allowed to choose $\theta(t)$ to follow the path as

rapidly or as slowly as desired; indeed, in some cases it is permissible to reverse course, although we distinguish between the case where we must always go forward and that for which we need not. Given that the plant is strictly proper with uniform relative degree $p$, it is natural to require that $r \in \mathcal{C}^p$. Furthermore, we will require that the reference signal be bounded, so the class of admissible trajectories is given by

$$\mathcal{R}_p := \{r \in C^p : r^{(i)} \in L_\infty, \ i = 0, 1, \ldots, p\}.$$

*Standing Assumption 3 :* $r \in \mathcal{R}_p$

Now we consider the class of timing signals $\theta$. Since the reference signal will be $r \circ \theta$, we will need $\theta$ to be sufficiently smooth—it must belong to $\mathcal{C}^p$. We will also insist that eventually it traverse the trajectory at the original speed. To this end, we define two different classes of timing signals.

*Definition 1:* $\Theta_p$ is the set of $\theta \in \mathcal{C}^p$ for which $\theta(0) = 0$, $\theta(\cdot) \geq 0$, and for which there exists a $T > 0$ so that

$$\dot{\theta}(t) = 1, \ t \geq T.$$

$\Theta_p^+$ is the subset of those $\theta \in \Theta_p$ for which $\theta(t)$ is monotonically increasing.[1]

*Remark 1:* The constraint on $\theta$ ensures that eventually we will be trying to track the unscaled trajectory. Allowing $\dot{\theta}$ to be negative is akin to allowing one to maneuver a car into a tight parking spot.

*Remark 2:* In practise there are a number of other constraints that may be important for path following problems. These include constraints on the maximum rate of change of $\theta$, actuator constraints, and constraints on the maximum $T$ that can be tolerated. While these features may be important in practice, they add substantial further complications to the analysis presented here. The results in this paper should therefore be interpreted as an analysis of the limits to what can be achieved, in an ideal situation, using the path following paradigm.

*Remark 3:* It is routine to confirm that $\theta \in \Theta_p$ and $r \in \mathcal{R}_p$ implies that $r \circ \theta \in \mathcal{R}_p$.

Given the problem setting, we allow feed-forward control and perfect prior knowledge of $r$. In this context, we will insist that the control signal and state remain bounded, and that the path following error be square integrable.

*Definition 3:* Given $x_{p_0} \in \mathbf{R}^{n_p}$ and $r \in \mathcal{R}_p$, we say that $(u, \theta) \in L_\infty \times \Theta_p$ stabilizes and solves the path following problem for (1) (or (5)) if:

i) $x_p \in L_\infty$, and
ii) $J(r, x_{p_0}, \theta, u) := \|y - r(\theta)\|_2 < \infty$,

in which case we write $(u, \theta) \in S(r, x_{p_0})$; we define $S^+(r, x_{p_0})$ in an analogous way. Here $J(r, x_{p_0}, \theta, u)$ represents the size of the path following error for the given control signal and choice of path variable.

*Remark 4:* In the above definition, condition (i) says that the plant state is well-behaved (bounded), while condition (ii) says that the path following error is square integrable. However, (i)

[1]Here we insist that we have $\theta(0) = 0$, so that we start at the beginning of the path. In the case of $\Theta_p$ this constraint is immaterial, but in the case of $\Theta_p^+$ this constraint is critical.

also implies that $\dot{e} \in L_\infty$, so if we combine this with (ii) and use elementary analysis we conclude that

$$\lim_{t \to \infty} e(t) = 0$$

e.g., see Corollary 2.9 of [17].

Our control problems are two-fold:
i) Given $r \in \mathcal{R}_p$, compute values for, or bounds on, the following optimal costs:

$$J_{opt}(r, x_{p_0}) := \inf_{(u, \theta) \in S(r, x_{p_0})} J(r, x_{p_0}, \theta, u) \qquad (6)$$

and

$$J_{opt}^+(r, x_{p_0}) := \inf_{(u, \theta) \in S^+(r, x_{p_0})} J(r, x_{p_0}, \theta, u). \qquad (7)$$

ii) Characterize the subset of $r \in \mathcal{R}_p$ for which $J_{opt}(r, x_{p_0})$ and $J_{opt}^+(r, x_{p_0})$ are zero, i.e., those trajectories for which we can obtain near optimal path-following.

We will be able to convert the first problem into an unconstrained convex optimization problem, at least to compute $J_{opt}(r, x_{p_0})$; we will also prove that in the important special case in which $r$ is periodic, we have $J_{opt}(r, x_{p_0}) = J_{opt}^+(r, x_{p_0})$. We give various partial solutions to the second problem. More specifically, we give: sufficient conditions under which the infimal performance is zero; sufficient conditions under which the infimal performance can never be zero; a precise characterization in certain special cases of when the infimal cost is zero.

Before proceeding, we briefly examine the system (1) written in zero dynamics form (5). Given $x_{p_0} \in \mathbf{R}^{n_p}$ and $r \in \mathcal{R}_p$, suppose that $(u, \theta) \in S(r, x_{p_0})$. This means, in particular, that $u, y, e, x_p \in L_\infty$. If we solve the differential equation for $x^+$, we obtain

$$\begin{aligned}
x^+(t) &= e^{A^+ t}\left[x_0^+ + \int_0^t e^{-A^+ \tau} B^+ y(\tau) d\tau\right] \\
&= e^{A^+ t}\left[x_0^+ + \int_0^t e^{-A^+ \tau} B^+ r(\theta(\tau)) d\tau \right. \\
&\quad \left. - \int_0^t e^{-A^+ \tau} B^+ e(\tau) d\tau\right].
\end{aligned}$$

Since $r \circ \theta \in L_\infty$, both integrals on the RHS must converge as $t \to \infty$. Since $x^+ \in L_\infty$ it follows from the fact that $-A^+$ is stable that the quantity in square brackets must tend to zero as $t \to \infty$, or equivalently

$$\int_0^\infty e^{-A^+ \tau} B^+ e(\tau) d\tau = x_0^+ + \int_0^\infty e^{-A^+ \tau} B^+ r(\theta(\tau)). \quad (8)$$

This equation, which is associated with the unstable zero dynamics, turns out to be the critical constraint. No such constraint results from examining the stable zero dynamics, since it suffices that $y$ is bounded to ensure that $x^-$ is bounded. Hence, it should not be surprising that in the next section we will prove that the optimal achievable cost depends on $x_{p_0}$ via $x_0^+$—it does not depend on $x_0^-$ nor on $v_0$.

The next step is to convert the nonlinear optimization problems given in (6) and (7) to simpler linear ones. There are two possible approaches: one is to use the "error coordinate form", as in [7], while the other is to use the original representation, which is what we do here.

## III. SIMPLIFYING THE PROBLEM

In this section we use three steps to convert the nonlinear optimization problems given in (6) and (7) to simpler linear ones. In the first step we eliminate $u$ so that we end up with $\theta$ as the only free variable. To proceed, we need some notation, a definition and a preliminary result.

First we define the controllability Grammian corresponding to the unstable zero dynamics sub-system of the plant by

$$W := \int_0^\infty e^{-A^+\tau} B^+ (B^+)^T e^{-(A^+)^T \tau} \, d\tau.$$

Second, motivated by (8):

*Definition 4:* With $T \geq 0$ we define $\rho, \rho_T : L_\infty \times \mathbf{R}^n \to \mathbf{R}^+$ by

$$\rho(f, x_0^+) := \left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+\tau} B^+ f(\tau) \, d\tau \right\|$$

$$\rho_T(f, x_0^+) := \left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^T e^{-A^+\tau} B^+ f(\tau) \, d\tau \right\|.$$

It follows from the triangle inequality and some straight-forward algebra that

$$\rho(f, x_0^+) - \rho(g, x_0^+) \leq \rho(f - g, 0)$$

and a similar property holds for $\rho_T$.

Third, we use the well-known result on optimization.

*Proposition 1:* [4, Appendix C]

Let $w \in \mathbf{R}^n$. Then the equation shown at the bottom of the page holds, and the minimizing $f$ is

$$f(t) = (B^+)^T e^{-A^{+T} t} W^{-1} w, \ t \geq 0.$$

We now turn to the first result, which shows that, in the ideal case considered, we do not have to directly consider the control signal, but instead we can focus solely on the unstable component of the zero dynamics.

*Proposition 2:* Subject to Standing Assumptions 1–3, for every initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ and every reference signal $r \in \mathcal{R}_p$, the infimal costs defined in (6) and (7) satisfy:

$$J_{opt}(r, x_{p_0}) = \inf_{\theta \in \Theta_p} \rho(r \circ \theta, x_0^+),$$

$$J_{opt}^+(r, x_{p_0}) = \inf_{\theta \in \Theta_p^+} \rho(r \circ \theta, x_0^+).$$

*Remark 5:* The equalities in Proposition 2 may look incongruous at first glance due to the $x_{p_0}$ on the LHS and $x_0^+$ on the RHS. However, the initial condition $x_{p_0}$ on (1) induces a natural initial condition on (5), yielding $x_0^+$, $x_0^-$ and $v_0$; only $x_0^+$ enters into the cost.

*Proof of Proposition 2:* To prove this, we will use approximations in $C^p$ to certain discontinuous functions. To this end, we define

$$\delta_\varepsilon(t) = \begin{cases} c_\varepsilon t^{p+1}(t - \varepsilon)^{p+1} & t \in [0, \varepsilon] \\ 0 & \text{otherwise,} \end{cases}$$

with $c_\varepsilon$ chosen so that $\int_0^\varepsilon \delta_\varepsilon(\tau) \, d\tau = 1$. We then define

$$\phi_\varepsilon(t) = \int_0^t \delta_\varepsilon(\tau) \, d\tau, \ t \in \mathbf{R}.$$

Notice that $\delta_\varepsilon, \phi_\varepsilon \in C^p$ with the first $p$ derivatives equal to zero for all $t$ outside $(0, \varepsilon)$. Finally, define

$$W_\varepsilon := \int_0^\infty \phi_\varepsilon(\tau) e^{-A^+\tau} B^+ (B^+)^T e^{-(A^+)^T \tau} \, d\tau > 0;$$

clearly $\lim_{\varepsilon \to 0} \|W - W_\varepsilon\| = 0$.

The equalities to be proved are established by proving non-strict inequalities in both directions.

($\geq$) Suppose that $(u, \theta) \in S(r, x_{p_0})$ or $S^+(r, x_{p_0})$, as appropriate. From the system equation we have that

$$e^{-A^+t} x^+(t) = x_0^+ + \int_0^t e^{-A^+\tau} B^+ y(\tau) \, d\tau. \quad (9)$$

Since $(u, \theta) \in S(r, x_0^+)$, we have that $x^+ \in L_\infty$, so $\|e^{-A^+t} x^+(t)\| \to 0$ as $t \to \infty$, which means that

$$x_0^+ + \int_0^\infty e^{-A^+\tau} B^+ y(\tau) \, d\tau = 0.$$

By Remark 3, $r \circ \theta \in \mathcal{R}_p \subset L_\infty$, which means that

$$w := \int_0^\infty e^{-A^+\tau} B^+ r(\theta(\tau)) d\tau$$

is well-defined. Now $e(t) = r(\theta(t)) - y(t)$, so

$$\int_0^\infty e^{-A^+\tau} B^+ e(\tau) \, d\tau = w + x_0^+.$$

Since $e \in PC$, it follows from Proposition 1 that

$$\|e\|_2 \geq \|W^{-1/2}(w + x_0^+)\|.$$

But $\|e\|_2 = J(r, x_{p_0}, \theta, u)$ and the above inequality holds for every $(u, \theta) \in S(r, x_{p_0})$ (or $S^+(r, x_{p_0})$, as appropriate) so

$$J_{opt}(r, x_{p_0}) \geq \|W^{-1/2}(w + x_0^+)\|$$
$$= \left\| W^{-1/2} x_0^+ + \right.$$

$$\min\{\|f\|_2 : f \in PC \text{ and } \int_0^\infty e^{-A^+\tau} B^+ f(\tau) \, d\tau = w\} = \|W^{-1/2} w\|$$

$$\left\| W^{-1/2} \int_0^\infty e^{-A^+\tau} B^+ r(\theta(\tau)) d\tau \right\|$$
$$= \rho(r \circ \theta, x_0^+),$$

with a similar result holding for $J_{opt}^+(r, x_{p_0})$.

($\leq$) The proof of this part is constructive. Since it is of peripheral interest, it is placed in the Appendix. $\square$

At this point we have eliminated $u$ from the optimization problem. We now show that the optimal cost is unchanged if we restrict our attention to a special class of "almost piecewise constant" $\theta$. To this end, we first define

$$\hat{\Theta} := \{\theta \in PC : \text{ there exists a } T > 0 \text{ so that}$$
$$\theta(t) \text{ is piecewise constant on } [0, T]$$
$$\text{and satisfies } \dot{\theta}(t) = 1 \text{ for } t > T\},$$
$$\hat{\Theta}^+ := \{\theta \in \hat{\Theta} : \theta(t) \text{ is monotonically increasing}\}.$$

Now we can define two new cost functions:

$$\hat{J}_{opt}(r, x_{p_0}) := \inf_{\hat{\theta} \in \hat{\Theta}} \rho(r \circ \hat{\theta}, x_0^+)$$
$$\hat{J}_{opt}^+(r, x_{p_0}) := \inf_{\hat{\theta} \in \hat{\Theta}^+} \rho(r \circ \hat{\theta}, x_0^+).$$

We are now in a position to show that there is no gap between the performance achievable by optimizing over $\theta \in \Theta$ versus optimizing over the "almost piecewise constant" functions in $\hat{\Theta}$.

*Proposition 3:* Subject to Standing Assumptions 1–3, for every initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ and every reference signal $r \in \mathcal{R}_p$:
  i) $\hat{J}_{opt}(r, x_{p_0}) = J_{opt}(r, x_{p_0})$.
  ii) $\hat{J}_{opt}^+(r, x_{p_0}) = J_{opt}^+(r, x_{p_0})$.

*Proof:* We prove (i) by first showing that $\hat{J}_{opt}(r, x_{p_0}) \leq J_{opt}(r, x_{p_0})$ and then showing that the converse also holds. The same style of argument applies to (ii).

In the ($\leq$) direction we start with $\theta \in \Theta_p$ and prove that if we put it through a sampler-and-hold for the first $T$ seconds and let the rest pass through unchanged, then the resulting timing signal yields a level of performance close to that achieved by $\theta$. In the ($\geq$) direction we start with an admissible $\theta$, piecewise constant on the initial interval and a ramp thereafter, and prove that if we convolve it with $\delta_\varepsilon$ with $\varepsilon$ small then the resulting timing signal lies in $\Theta_p$ and yields a level of performance close to that achieved by $\theta$. The details can be found in the Appendix. $\square$

We now demonstrate that in many cases we can simplify the problem even further by eliminating $\theta$ from the optimization problem. To this end, the set of points that $r$ traverses, together with the positive limit set[2] of $r$, play important roles:

$$IM(r) := \{r(t) : t \geq 0\},$$
$$IM^+(r) := \text{ the positive limit set of } r.$$

[2]The positive limit set of $r \in \mathcal{R}_p$ is the set of $\bar{r} \in \mathbf{R}^m$ for which there exists a strictly increasing sequence of times $\{t_i : i \in \mathbf{N}\}$ satisfying $\lim_{i \to \infty} t_i = \infty$ and $\lim_{i \to \infty} r(t_i) = \bar{r}$.

The closure of the convex hulls of these sets also play an important role and we label them $CIM(r)$ and $CIM^+(r)$, respectively. We now define

$$\tilde{J}_{opt}(r, x_{p_0}) := \inf\Big\{\rho(f, x_0^+) : f \in PC_\infty \text{ and}$$
$$f(t) \in CIM(r), t \geq 0\Big\} \quad (10)$$
$$\tilde{J}_{opt}^+(r, x_{p_0}) := \inf\Big\{\rho(f, x_0^+) : f \in PC_\infty \text{ and}$$
$$f(t) \in CIM^+(r), t \geq 0\Big\}. \quad (11)$$

*Remark 6:* In the definition of $\hat{J}_{opt}(r, x_{p_0})$, let $\theta$ be an admissible function over which the optimization is carried out, and define $f(t) := r(\theta(t))$. Then it follows immediately that $f$ is piecewise continuous (indeed, it is piecewise constant for the first interval of time, and continuous thereafter) and takes values in $IM(r) \subset CIM(r)$; furthermore, it is easy to see that $f \in PC_\infty$. Using Proposition 3, this means that

$$\tilde{J}_{opt}(r, x_{p_0}) \leq \hat{J}_{opt}(r, x_{p_0}) = J_{opt}(r, x_{p_0}).$$

However, the same argument does not hold for $\tilde{J}_{opt}^+(r, x_{p_0})$ and $J_{opt}^+(r, x_{p_0})$.

Now we demonstrate that $\tilde{J}_{opt}(r, x_{p_0})$ and $J_{opt}(r, x_{p_0})$ are equal, and that $\tilde{J}_{opt}^+(r, x_{p_0})$ provides an upperbound on $J_{opt}^+(r, x_{p_0})$.

*Proposition 4:* Subject to Standing Assumptions 1–3, for every initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ and every reference signal $r \in \mathcal{R}_p$:
  i) $\tilde{J}_{opt}(r, x_{p_0}) = J_{opt}(r, x_{p_0})$.
  ii) $\tilde{J}_{opt}^+(r, x_{p_0}) \geq J_{opt}^+(r, x_{p_0})$.

*Proof:* The proof uses detailed arguments from real analysis and can be found in the Appendix. $\square$

*Remark 7:* Suppose that $A^+ = \lambda > 0$, $B^+ = 1$, and $CIM(r) = [\underline{b}, \bar{b}]$ with $\underline{b} > 0$. From Proposition 4 (i) and a routine calculation, it follows that

$$J_{opt}(r, x_{p_0}) = \begin{cases} -(\frac{2}{\lambda})^{1/2}[\lambda x_0^+ + \bar{b}] & x_0^+ < \frac{-\bar{b}}{\lambda} \\ 0 & x_0^+ \in \left[\frac{-\bar{b}}{\lambda}, \frac{-\underline{b}}{\lambda}\right] \\ (\frac{2}{\lambda})^{1/2}[\lambda x_0^+ + \underline{b}] & x_0^+ > \frac{-\underline{b}}{\lambda}. \end{cases}$$

We suspect that the inequality in Proposition 4 (ii) cannot be strengthened to equality without additional structure on $r$. One particularly interesting case is that of $r$ being periodic, which is explored in the following result.

*Proposition 5:* Subject to Standing Assumptions 1–3, for every initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ and every periodic reference signal $r \in \mathcal{R}_p$:

$$J_{opt}^+(r, x_{p_0}) = \tilde{J}_{opt}^+(r, x_{p_0}) = \tilde{J}_{opt}(r, x_{p_0}) = J_{opt}(r, x_{p_0}).$$

*Proof:* For every $r \in \mathcal{R}_p$, it is clear from the definition that

$$J_{opt}(r, x_{p_0}) \leq J_{opt}^+(r, x_{p_0}), \quad (12)$$

and from Proposition 4 that

$$J_{opt}^+(r, x_{p_0}) \leq \tilde{J}_{opt}^+(r, x_{p_0}) \tag{13}$$

$$\tilde{J}_{opt}(r, x_{p_0}) = J_{opt}(r, x_{p_0}). \tag{14}$$

If $r \in \mathcal{R}_p$ is periodic, then $CIM^+(r) = CIM(r)$, so

$$\tilde{J}_{opt}(r, x_{p_0}) = \tilde{J}_{opt}^+(r, x_{p_0}). \tag{15}$$

Combining (13), (14) and (15) yields

$$J_{opt}^+(r, x_{p_0}) \leq \tilde{J}_{opt}^+(r, x_{p_0}) = \tilde{J}_{opt}(r, x_{p_0}) = J_{opt}(r, x_{p_0})$$

which, when combined with (12) yields the three equalities. □

In this section, we have therefore shown close links between the original problems of determining $J_{opt}(r, x_{p_0})$ and $J_{opt}^+(r, x_{p_0})$ to those of finding $\tilde{J}_{opt}(r, x_{p_0})$ and $\tilde{J}_{opt}^+(r, x_{p_0})$ as defined in (10) and (11). In particular, if we are not restricted to monotonic $\theta$, or in the case of periodic reference trajectories, Propositions 4 and 5 show that these alternate problems are completely equivalent. We now turn to using these results to give analytic expressions for bounds on the achievable performance.

## IV. BOUNDS ON THE ACHIEVABLE PERFORMANCE

In this section we prove some preliminary results, including some upper and lower bounds on the optimal cost, as well as sufficient conditions for it to be zero. We begin by considering some upper bounds on the achievable performance. To proceed, we choose $c_0 > 0$ and $\lambda_0 < 0$ so that

$$\left\| e^{-A^+ t} \right\| \leq c_0 e^{\lambda_0 t}, \; t \geq 0. \tag{16}$$

*Theorem 1:* Subject to Standing Assumptions 1–3, for every initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ and every reference signal $r \in \mathcal{R}_p$:

i) $J_{opt}(r, x_{p_0}) \leq \|W^{-1/2}\| \inf_{t \geq 0} \|x_0^+ + (A^+)^{-1} B^+ r(t)\|$;

ii) $J_{opt}^+(r, x_{p_0}) \leq \|W^{-1/2}\| \inf_{t \geq 0} \|x_0^+ + (A^+)^{-1} B^+ r(t)\|$;

iii) $J_{opt}(r, x_{p_0}) \leq \|W^{-1/2}\| \times \inf_{w \in CIM(r)} \{\|x_0^+ + (A^+)^{-1} B^+ w\|\}$;

iv) $J_{opt}^+(r, x_{p_0}) \leq \|W^{-1/2}\| \times \inf_{w \in CIM^+(r)} \{\|x_0^+ + (A^+)^{-1} B^+ w\|\}$.

*Proof:* By Proposition 4 we obtain the first equation shown at the bottom of the page, so (i) and (iii) hold.

By Proposition 4, we obtain the second equation shown at the bottom of the page, so (iv) holds. In this case, it could very well be that much of $IM(r)$ does not belong to $CIM^+(r)$, so we cannot immediately obtain (ii) from the above.

To prove (ii), we use Proposition 3. Let $w \in IM(r)$. Then there exist a $\bar{t} \geq 0$ so that $w := r(\bar{t})$; let $T > \bar{t}$. Consider

$$\theta(t) = \begin{cases} \bar{t} & t \in [0, T) \\ t & t \geq T. \end{cases}$$

Since $r \in L_\infty$, there exists a constant $c_1$ so that

$$\|r(t)\| \leq c_1, \; t \geq 0.$$

Then by Proposition 3 it follows that:

$$J_{opt}^+(r, x_{p_0}) \leq \rho(r \circ \theta, x_0^+)$$

$$\leq \left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^T e^{-A^+ \tau} B^+ w \, d\tau \right\|$$

$$+ \left\| W^{-1/2} \int_T^\infty e^{-A^+ \tau} B^+ r(\tau) \, d\tau \right\|$$

$$\leq \|W^{-1/2}\| \times \left\| x_0^+ + (I - e^{-A^+ T})(A^+)^{-1} B^+ w \right\|$$

$$+ \|W^{-1/2}\| \int_T^\infty c_0 e^{\lambda_0 \tau} \|B^+\| c_1 \, d\tau$$

$$\leq \|W^{-1/2}\| \times \left\| x_0^+ + (I - e^{-A^+ T})(A^+)^{-1} B^+ w \right\|$$

---

$$J_{opt}(r, x_{p_0}) = \tilde{J}_{opt}(r, x_{p_0})$$

$$= \inf\{\rho(f, x_0^+) : f \in PC_\infty \text{ and } f(t) \in CIM(r), \; t \geq 0\}$$

$$\leq \inf\left\{ \left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+ \tau} B^+ w \, d\tau \right\| : w \in CIM(r) \right\}$$

$$\leq \inf\left\{ \left\| W^{-1/2} \right\| \times \left\| x_0^+ + (A^+)^{-1} B^+ w \right\| : w \in CIM(r) \right\}$$

$$\leq \|W^{-1/2}\| \inf\left\{ \left\| x_0^+ + (A^+)^{-1} B^+ r(t) \right\| : t \geq 0 \right\}$$

---

$$J_{opt}^+(r, x_{p_0}) \leq \tilde{J}_{opt}^+(r, x_{p_0})$$

$$= \inf\{\rho(f, x_0^+) : f \in PC_\infty \text{ and } f(t) \in CIM^+(r), \; t \geq 0\}$$

$$\leq \inf\left\{ \left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+ \tau} B^+ w \, d\tau \right\| : w \in CIM^+(r) \right\}$$

$$= \|W^{-1/2}\| \inf\left\{ \left\| x_0^+ + (A^+)^{-1} B^+ w \right\| : w \in CIM^+(r) \right\}$$

$$+ \|W^{-1/2}\| \frac{c_0 c_1 e^{\lambda_0 T}}{|\lambda_0|} \|B^+\|;$$

here $\lambda_0$ and $c_0$ are as defined in (16). Since $T > \bar{t}$ is as large as desired, we see that the RHS can be made as close to

$$\|W^{-1/2}\| \times \|x_0^+ + (A^+)^{-1} B^+ w\|$$

as desired; since $w \in IM(r)$ is arbitrary, it follows that (ii) holds. $\quad\square$

The following result is based directly on the results of Theorem 1 and gives several sufficient conditions for the achievable performance to be perfect.

*Corollary 1:* Subject to Standing Assumptions 1–3, for every initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ and every reference signal $r \in \mathcal{R}_p$:
  i) If $\inf_{t \geq 0} \|r(t)\| = 0$ then $J_{opt}^+(r, 0) = J_{opt}(r, 0) = 0$.
  ii) If $\inf_{t \geq 0} \|x_0^+ + (A^+)^{-1} B^+ r(t)\| = 0$ then $J_{opt}^+(r, x_{p_0}) = J_{opt}(r, x_{p_0}) = 0$.
  iii) If $0 \in \{x_0^+ + (A^+)^{-1} B^+ \times CIM(r)\}$ then $J_{opt}(r, x_{p_0}) = 0$.
  iv) If $0 \in \{x_0^+ + (A^+)^{-1} B^+ \times CIM^+(r)\}$ then $J_{opt}^+(r, x_{p_0}) = 0$.

*Proof:* These four results follow directly from the four results listed in Theorem 1. $\quad\square$

In the case of $x_{p_0} = 0$, Corollary 1(iii) says that $J_{opt}(r, 0) = 0$ if $0 \in CIM(r)$; this is consistent with key results in [2] and [7]. It turns out, however, that this condition is not necessary. In the following we provide a weaker sufficient condition for $J_{opt}(r, 0) = 0$; indeed, we prove this in the more general case of $J_{opt}(r, x_{p_0})$. We follow it with a simple example demonstrating the property.

*Theorem 2:* Subject to Standing Assumptions 1–3, for every initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ and every reference signal $r \in \mathcal{R}_p$:
  i) If there exist $w_1, w_2 \in CIM(r)$ and $T \geq 0$ so that

$$x_0^+ + (A^+)^{-1}(I - e^{-2A^+ T})^{-1}(B^+ w_1 + e^{-A^+ T} B^+ w_2) = 0,$$

  then $J_{opt}(r, x_{p_0}) = 0$.
  ii) If there exist $w_1, w_2 \in CIM^+(r)$ and $T \geq 0$ so that

$$x_0^+ + (A^+)^{-1}(I - e^{-2A^+ T})^{-1}(B^+ w_1 + e^{-A^+ T} B^+ w_2) = 0,$$

  then $J_{opt}^+(r, x_{p_0}) = 0$.

*Proof:*
  i) Choose $w_1, w_2 \in CIM(r)$ and $T \geq 0$ to satisfy the stated condition. By Proposition 4

$$J_{opt}(r, x_{p_0}) = \tilde{J}_{opt}(r, x_{p_0})$$
$$:= \inf\{\rho(f, x_0^+) : f \in PC_\infty \text{ and takes values in } CIM(r)\}.$$

Let us construct $f$ to be periodic of period $2T$ and defined by

$$f(t) = \begin{cases} w_1 & t \in [0, T) \\ w_2 & t \in [T, 2T); \end{cases}$$

clearly $f \in PC_\infty$ and takes values in $CIM(r)$ so

$$J_{opt}(r, x_{p_0}) = \tilde{J}_{opt}(r, x_{p_0}) \leq \rho(f, x_0^+)$$
$$= \left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+ \tau} B^+ f(\tau) d\tau \right\|$$

$$= \left\| W^{-1/2} x_0^+ + W^{-1/2} \sum_{k=0}^\infty \int_{2kT}^{2(k+1)T} e^{-A^+ \tau} B^+ f(\tau) d\tau \right\|$$
$$= \left\| W^{-1/2} x_0^+ + W^{-1/2} \left( \sum_{k=0}^\infty e^{-2A^+ kT} \right) \int_0^{2T} e^{-A^+ \theta} B^+ f(\theta) d\theta \right\|$$
$$= \left\| W^{-1/2} [x_0^+ + (A^+)^{-1}(I - e^{-2A^+ T})^{-1}(B^+ w_1 + e^{-A^+ T} B^+ w_2)] \right\|$$
$$= 0,$$

where the last equality follows by the specific construction of $f$. We conclude that $J_{opt}(r, x_{p_0}) = 0$.
  ii) The proof is analogous to that of part (i). $\quad\square$

*Remark 8:* There are cases in which the bound provided by Theorem 1 is weak, but the conditions in Theorem 2 apply. To see this, suppose that

$$x_{p_0} = 0, \quad A^+ = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, B^+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$
$$r(t) = \begin{bmatrix} \frac{1}{6} \\ \frac{5}{18} \end{bmatrix} + \sin(t) \begin{bmatrix} \frac{5}{6} \\ \frac{13}{18} \end{bmatrix},$$

i.e., $r(t)$ oscillates in a straight line between $w_1 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $w_2 := \begin{bmatrix} -2/3 \\ -4/9 \end{bmatrix}$, which means that

$$0 \notin x_0^+ + (A^+)^{-1} B^+ \times CIM(r)$$

so the bound on $J_{opt}(r, 0)$ provided by Theorem 1 (iii) is positive. If we choose $T = \ln 1.5$, it is easy to see that

$$B^+ w_1 + e^{-A^+ T} B^+ w_2 = 0.$$

Hence, Theorem 2 (ii) shows that $J_{opt}(r, 0) = 0$, which means that we do not need $0 \in CIM(r)$ nor $0 \in B^+ \times CIM(r)$ for zero optimal cost. Compare this with the approach of [2] and [7], in which it is shown that if $0 \in CIM(r)$ then $J_{opt}(r, 0) = 0$.

*Remark 9:* Notice that, for a given admissible system, the conditions on $r$ given in Theorem 2 depend solely on the general properties of $r$, namely on $CIM(r)$ or $CIM^+(r)$, as appropriate. In particular, $r$ need not be periodic; the fact that $r$ is periodic in the example above is for simplicity, but this is not essential.

As well as finding upper bounds on the achievable performance and hence giving cases where perfect performance is achievable, we can also find lower bounds on the achievable performance in some cases. This may therefore be used to identify cases where it is assured that perfect performance is not achievable.

*Theorem 3:* Subject to Standing Assumptions 1–3, suppose that the system is such that $A^+$ and $B^+$ are of the form

$$A^+ = \begin{bmatrix} \lambda_1 & 0 \\ 0 & * \end{bmatrix}, B^+ = \begin{bmatrix} B_1^+ \\ * \end{bmatrix},$$

with $\lambda_1 \in \mathbf{R}^+$. With $r \in \mathcal{R}_p$ and the initial condition $x_{p_0} \in \mathbf{R}^{n_p}$ partitioned as

$$x_0^+ = \begin{bmatrix} x_{01}^+ \\ * \end{bmatrix}$$

if $\inf_{t\geq 0}\left|\lambda_1 x_{01}^+ + B_1^+ r(t)\right| > 0$, then

$$J_{opt}^+(r, x_{p_0}) \geq J_{opt}(r, x_{p_0})$$
$$\geq \sigma_{\min}(W^{-1/2})\times$$
$$\frac{1}{\lambda_1}\inf_{t\geq 0}\left|\lambda_1 x_{01}^+ + B_1^+ r(t)\right| > 0.$$

*Proof:* Suppose that $A^+$ and $B^+$ are of the desired form, and let $x_{p_0} \in \mathbf{R}^{n_p}$ and $r \in \mathcal{R}_p$ be such that the stated inequality holds. By Proposition 2 it follows that:

$$J_{opt}(r, x_{p_0}) = \inf_{\theta \in \Theta_p} \rho(r \circ \theta, x_0^+)$$
$$= \inf_{\theta \in \Theta_p} \left\| W^{-1/2}\left[ x_0^+ + \int_0^\infty e^{-A^+\tau} B^+ r(\theta(\tau))d\tau\right]\right\|$$
$$\geq \sigma_{\min}(W^{-1/2})\inf_{\theta \in \Theta_p}\left\|\left[ x_0^+ + \int_0^\infty e^{-A^+\tau} B^+ r(\theta(\tau))d\tau\right]\right\|$$
$$\geq \sigma_{\min}(W^{-1/2})\inf_{\theta \in \Theta_p}\left\|\left[ x_{01}^+ + \int_0^\infty e^{-\lambda_1\tau} B_1^+ r(\theta(\tau))d\tau\right]\right\|$$
$$= \sigma_{\min}(W^{-1/2})\inf_{\theta \in \Theta_p}\left|\int_0^\infty e^{-\lambda_1\tau}[\lambda_1 x_{01}^+ + B_1^+ r(\theta(\tau))]d\tau\right|$$
$$= \sigma_{\min}(W^{-1/2})\inf_{\theta \in \Theta_p}\int_0^\infty e^{-\lambda_1\tau}\left|[\lambda_1 x_{01}^+ + B_1^+ r(\theta(\tau))]\right|d\tau$$

(since $\lambda_1 x_{01}^+ + B_1^+ r(\theta(t))$ never changes sign)

$$\geq \sigma_{\min}(W^{-1/2})\frac{1}{\lambda_1}\inf_{\theta\geq 0}\left|\lambda_1 x_{01}^+ + B_1^+ r(\theta)\right|.$$

$\square$

We now turn from establishing bounds on the achievable performance to special cases where more precise results may be established.

## V. THE SINGLE-INPUT CASE WITH ZERO INITIAL CONDITIONS

Here we examine the special case in which we have a scalar input (which means that we also have a scalar output) and start from zero initial conditions. We are able to obtain easily stated (and checked) necessary and sufficient conditions for the optimal cost to be zero in two cases: when there is at least one real transmission zero in the open RHP, and when there are exactly two transmission zeros in the open RHP—a complex conjugate pair.

### A. At Least One Real Positive Zero

If the system has at least one real positive zero, then the condition of Corollary 1 (i) is tight.

*Theorem 4:* Subject to Standing Assumptions 1–3, if $A^+$ has at least one real eigenvalue then for every reference signal $r \in \mathcal{R}_p$:

$$J_{opt}(r, 0) = 0 \text{ iff } \inf_{t\geq 0}|r(t)| = 0$$

and

$$J_{opt}^+(r, 0) = 0 \text{ iff } \inf_{t\geq 0}|r(t)| = 0.$$

*Proof:* ($\Leftarrow$) This follows immediately from Corollary 1.
($\Rightarrow$) Suppose that $\inf_{t\geq 0}|r(t)| > 0$. Now carry out a similarity transformation, if necessary, so that

$$A^+ := \begin{bmatrix} \lambda & 0 \\ 0 & \star \end{bmatrix}, \ B^+ := \begin{bmatrix} B_1^+ \\ \star \end{bmatrix}$$

with $\lambda$ real and positive; $B_1^+ \neq 0$ since $(A^+, B^+)$ must be controllable. Then $\inf_{t\geq 0}|B_1^+ r(t)| > 0$, so by Theorem 3 we have $J_{opt}^+(r, 0) \geq J_{opt}(r, 0) > 0$. $\square$

*Remark 10:* Remark 8 clearly demonstrates that the result does not generalize to the multi-input case.

*Remark 11:* It would clearly be desireable to be able to obtain an easily stated (and checked) necessary and sufficient conditions for the optimal cost to be zero even when the initial condition is non-zero. Unfortunately, we have not been able to find such a condition in general, but we can provide one in the simplest case, namely that of $n = 1$, i.e., when $A^+ \in \mathbf{R}$. In this situation we can prove that for every reference signal $r \in \mathcal{R}_p$:

$$J_{opt}(r, x_{p_0}) = 0 \text{ iff } x_0^+ \in \frac{B^+}{A^+}[-\sup_{t\geq 0} r(t), -\inf_{t\geq 0} r(t)]$$

and

$$J_{opt}^+(r, x_{p_0}) = 0 \text{ iff } x_0^+ \in \frac{B^+}{A^+}[-\sup_{t\geq 0} r(t), -\inf_{t\geq 0} r(t)].$$

### B. Two Complex Zeros

It turns out that if $A^+$ has exactly two complex eigenvalues then we no longer need $\inf_{t\geq 0}|r(t)| = 0$ to have $J_{opt}(r, 0) = 0$. We begin with an illustrative example.

*Example 1:* Consider the case of

$$A^+ = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}, \ B^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with a reference trajectory periodic of period $T := 2\pi/\beta$ given by

$$r(t) = e^{\alpha t}, \ t \in [0, T).$$

It follows that

$$\rho(r, 0) = \left\| W^{-1/2}\int_0^\infty e^{-A^+\tau} B^+ r(\tau)d\tau\right\|$$
$$= \left\| W^{-1/2}\sum_{k=0}^\infty e^{-\alpha kT}\underbrace{\int_0^T e^{-A^+\tau} B^+ r(\tau)d\tau}_{=0}\right\| = 0,$$

so $J_{opt}(r, 0) = 0$. While $r \notin \mathcal{R}_p$, the trajectory

$$\hat{r}(t) = \frac{e^{\alpha T} + 1}{2} + \frac{e^{\alpha T} - 1}{2}\sin(t)$$

lies in $\mathcal{R}_p$ and $IM(\hat{r}) = IM(r)$ as well, so $J_{opt}(\hat{r}, 0) = 0$.

We now seek, for this situation, necessary and sufficient conditions for $J_{opt}(r,0) = 0$. It is clearly not the same as in the case of having at least one real zero. To proceed, we write

$$[\underline{b}, \overline{b}] := CIM(r) \supset CIM^+(r) =: [\underline{b}^+, \overline{b}^+].$$

Next we need some preliminary technical results. The set which is central to the definition of $\tilde{J}_{opt}(r,0)$ is

$$\mathcal{C}_{[\underline{b}, \overline{b}]} := \left\{ -\int_0^\infty e^{-A^+\tau} B^+ f(\tau)\, d\tau : \right.$$
$$\left. f \in PC \text{ and } f(\tau) \in [\underline{b}, \overline{b}], \tau \geq 0 \right\}, \quad (17)$$

which is simply a controllable set: the closure of the set of all initial conditions $\tilde{x}_0^+$ of

$$\dot{\tilde{x}}^+ = A^+ \tilde{x}^+ + B^+ u, \quad \tilde{x}^+(0) = \tilde{x}_0^+$$

which can be driven to zero in finite time using a control signal $u \in PC$ satisfying $u(\cdot) \in [\underline{b}, \overline{b}]$. This problem was considered in [13], and a description of the surface of $\mathcal{C}_{[\underline{b}, \overline{b}]}$, namely $\partial \mathcal{C}_{[\underline{b}, \overline{b}]}$, is derived. To proceed, define

$$b_{av} := \frac{\overline{b} + \underline{b}}{2}, \quad \Delta := \frac{\overline{b} - \underline{b}}{2}. \quad (18)$$

*Theorem 5:* [13] If $A^+ \in \mathbf{R}^{2\times 2}$ with complex eigenvalues $\alpha \pm j\beta$ and $T := 2\pi/\beta$, then the equation shown at the bottom of the page holds.

We are now in a position to give crisp conditions for when the achievable performance in this case is exactly zero.

*Theorem 6:* Subject to Standing Assumptions 1–3, if $A^+ \in \mathbf{R}^{2\times 2}$ and $sp(A^+) = \alpha \pm j\beta$, then for every $r \in \mathcal{R}_p$:
  i) $J_{opt}(r,0) = 0$ iff $0 \in [\underline{b}, \overline{b}]$ or $e^{\pi\alpha/\beta} \leq \overline{b}/\underline{b}$.
  ii) $J_{opt}^+(r,0) = 0$ if $0 \in [\underline{b}^+, \overline{b}^+]$ or $e^{\pi\alpha/\beta} \leq \overline{b}^+/\underline{b}^+$.

*Proof of Theorem 6:* If $0 \in [\underline{b}, \overline{b}]$ then by Theorem 1 we have $J_{opt}(r,0) = J_{opt}^+(r,0) = 0$, so henceforth we assume that this is not the case. Without loss of generality we assume that $\underline{b} > 0$, for if it is not we simply replace $r$ by $-r$ and use the facts that $J_{opt}(r,0) = J_{opt}(-r,0)$ and $J_{opt}^+(r,0) = J_{opt}^+(-r,0)$. Define $T := 2\pi/\beta$; it is routine to check that

$$e^{A^+T} = e^{\alpha T}I, e^{A^+T/2} = -e^{\alpha T/2}I. \quad (19)$$

$(\Leftarrow)$ We first consider part (i), with part (ii) following similarly. Suppose that $e^{\pi\alpha/\beta} \leq \overline{b}/\underline{b}$. Define $f$ of period $T$ by

$$f(t) = \begin{cases} \underline{b} & t \in \left[0, \frac{T}{2}\right) \\ \underline{b}e^{\alpha T/2} & t \in \left[\frac{T}{2}, T\right). \end{cases}$$

Since $\underline{b}e^{\alpha T/2} = \underline{b}e^{\pi\alpha/\beta}$, $f$ clearly takes values only in $[\underline{b}, \overline{b}]$. It follows from Proposition 4 that:

$$J_{opt}(r,0) \leq \left\| W^{-1/2} \int_0^\infty e^{-A^+\tau} B^+ f(\tau)d\tau \right\|$$
$$= \left\| W^{-1/2} \sum_{k=0}^\infty \int_{kT}^{(k+1)T} e^{-A^+\tau} B^+ f(\tau)d\tau \right\|$$
$$= \| W^{-1/2} \sum_{k=0}^\infty e^{-A^+ kT} \underbrace{\int_0^T e^{-A^+\tau} B^+ f(\tau)d\tau}_{=0 \text{ (using (19) )}} \|$$
$$= 0.$$

The proof of the counterpart in part (ii) follows similar lines.
$(\Rightarrow)$ Suppose that $e^{\pi\alpha/\beta} > \overline{b}/\underline{b}$; we must prove that $J_{opt}(r,0) > 0$. With $\mathcal{C}_{[\underline{b}, \overline{b}]}$ given by (17), it follows from Proposition 4 that

$$J_{opt}(r,0) = \tilde{J}_{opt}(r,0) = 0 \leftrightarrow 0 \in \mathcal{C}_{[\underline{b}, \overline{b}]}.$$

From Theorem 5 the equation shown at the bottom of the page holds, so $0 \in \mathcal{C}_{[\underline{b}, \overline{b}]}$ iff there exists a $\theta \in [0, T/2)$ and $\eta \in [-1, 1]$ so that

$$(A^+)^{-1} B^+ b_{av}$$
$$= \eta \left[ \frac{2}{1 - e^{-\alpha T/2}} e^{-A^+\theta} - I \right] (A^+)^{-1} B^+ \Delta. \quad (20)$$

Since $A^+$ and $e^{-A^+\theta}$ commute, this is equivalent to

$$B^+ b_{av} = \eta \left[ \frac{2}{1 - e^{-\alpha T/2}} e^{-A^+\theta} - I \right] B^+ \Delta$$

so $0 \in \mathcal{C}_{[\underline{b}, \overline{b}]}$ if and only if there exists a $\theta \in [0, T/2)$ and $\eta \in [-1, 1]$ so that

$$[(A^+)^{-1} B^+ \quad B^+] b_{av}$$
$$= \eta \left[ \frac{2}{1 - e^{-\alpha T/2}} e^{-A^+\theta} - I \right] [(A^+)^{-1} B^+ \quad B^+] \Delta.$$

Using the fact that $(A^+, B^+)$ is controllable, this is equivalent to

$$b_{av}I = \eta \left[ \frac{2}{1 - e^{-\alpha T/2}} e^{-A^+\theta} - I \right] \Delta$$
$$\leftrightarrow e^{-A^+\theta} = \frac{(1 - e^{-\alpha T/2})}{2\Delta} \left( \Delta + \frac{1}{\eta} b_{av} \right) I$$
$$\leftrightarrow \theta = 0;$$

the last step follows from the fact that $e^{-A^+\theta}$ is a constant times the identity matrix (for $\theta \in [0, T/2)$) iff $\theta = 0$. If we substitute

---

$$\partial \mathcal{C}_{[\underline{b}, \overline{b}]} = -(A^+)^{-1} B^+ b_{av} \pm \Delta \left\{ \left[ \frac{2}{1 - e^{-\alpha T/2}} e^{-A\theta} - I \right] (A^+)^{-1} B^+ : \theta \in \left[0, \frac{T}{2}\right) \right\}$$

this choice of $\theta$ into the formula (20), we see that $0 \in \mathcal{C}_{[\underline{b}, \overline{b}]}$ iff there exists a $\eta \in [-1, 1]$ so that

$$(A^+)^{-1}B^+b_{av} = \eta \left[ \frac{2}{1 - e^{-\alpha T/2}} I - I \right] (A^+)^{-1}B^+\Delta$$

or equivalently

$$b_{av} = \eta \left[ \frac{2}{1 - e^{-\alpha T/2}} - 1 \right] \Delta$$
$$\leftrightarrow \frac{\overline{b} + \underline{b}}{2} \leq \frac{1 + e^{-\alpha T/2}}{1 - e^{-\alpha T/2}} \times \frac{\overline{b} - \underline{b}}{2}$$
$$\leftrightarrow e^{\alpha T/2} \leq \frac{\overline{b}}{\underline{b}}.$$

Since the last inequality does not hold by hypothesis, it follows that $0 \notin \mathcal{C}_{[\underline{b}, \overline{b}]}$, which means that $J_{opt}(r, 0) > 0$, as desired, which completes the proof of part (i). □

## VI. A Computational Approach

In Section III we simplified the expressions for the optimal costs, in Section IV we provided upper and lower bound on these optimal costs, and in Section V we provided necessary and sufficient conditions for the optimal costs to be zero in several very special cases. If we wish to compute the exact cost in the general case, then we need to solve the optimization problems given in (10) and (11)[3] (also see Propositions 4 and 5), each of which is an infinite dimensional optimization problem, which is difficult to solve directly [16]. In this section, under a mild technical condition we apply results on convex optimization in optimal control to convert each such problem to a much more tractible one, namely a finite dimensional strictly convex optimization problem, which is easy to solve using standard optimization tools.

To this end, consider the problem of computing $J_{opt}(r, x_{p_0})$ and $\tilde{J}^+_{opt}(r, x_{p_0})$ in the general case. In the forth-going, we focus on $J_{opt}(r, x_{p_0})$ (which equals $J^+_{opt}(r, x_{p_0})$ if $r$ is periodic or, more generally, if $CIM(r) = CIM^+(r)$), but computing $\tilde{J}^+_{opt}(r, x_{p_0})$ is similar, with a simple replacement of $CIM(r)$ with $CIM^+(r)$. To proceed, we impose

*Standing Assumption 4:* $CIM(r)$ is strictly convex.

In the general case, we do not have any closed form results. Making use of Proposition 4, we end up with the computational problem of computing

$$J_{opt}(r, x_{p_0})$$
$$= \inf \left\{ \left\| W^{-1/2}x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+\tau}B^+f(\tau)d\tau \right\| : \right.$$
$$\left. f \in PC_\infty \text{ and } f(t) \in CIM(r), t \geq 0 \right\}$$

and its approximation

$$J^T_{opt}(r, x_{p_0}) :=$$
$$\inf \left\{ \left\| W^{-1/2}x_0^+ + W^{-1/2} \int_0^T e^{-A^+\tau}B^+f(\tau)d\tau \right\| : \right.$$
$$\left. f \in PC_\infty \text{ and } f(t) \in CIM(r), t \in [0, T] \right\}.$$

It is clear that $\lim_{T \to \infty} J^T_{opt}(r, x_{p_0}) = J_{opt}(r, x_{p_0})$, indeed the convergence can easily be shown to be exponential as shown in the following result; recall that $c_0$ and $\lambda_0$ satisfy (16).

*Proposition 6:* Subject to Standing Assumptions 1–3, for every $x_0 \in \mathbf{R}^{n_p}$ and $r \in \mathcal{R}_p$:

$$\left| J_{opt}(r, x_{p_0}) - J^T_{opt}(r, x_{p_0}) \right|$$
$$\leq e^{\lambda_0 T} \frac{c_0}{|\lambda_0|} \|W^{-1/2}\| \times \|B^+\| \times \|CIM(r)\|.$$

Hence, it follows that to compute $J_{opt}(r, x_{p_0})$ to within a prescribed bound, it is enough to compute $J^T_{opt}(r, x_{p_0})$. Therefore, we turn our attention to this latter problem, which we can rewrite as a classical control problem.[4] To simplify the notation, in what follows we define

$$U := CIM(r).$$

*Optimization Problem 1 (OPT-1):* Minimize

$$g(x(T)) := \|W^{-1/2}e^{-A^+T}x(T)\|^2$$

subject to

$$\cdot \dot{x} = A^+x + B^+u, x(0) = x_0^+$$
$$\cdot u \in L_2 \text{ satisfies } u(\cdot) \in U.$$

This is simply a classical control problem with an end constraint. To proceed, we define two associated functions:

$$\psi : \mathbf{R}^m \to \mathbf{R}$$
$$\xi \longmapsto \max_{u \in U} \xi^T u$$

and

$$\phi : \mathbf{R}^m \to \mathbf{R}^m$$
$$\xi \longmapsto \arg\max_{u \in U} \xi^T u.$$

The first is well-defined because of the compactness of $U$; indeed, it is convex. The second is well-defined (for $\xi \neq 0$) since $U$ is strictly convex. To illustrate these definitions some examples are in order.

---

[3]Strictly speaking, the optimization problem given in (11) provides an upper bound on $J^+_{opt}(r, x_{p_0})$, though we do obtain equality in some interesting cases, such as when the reference signal is periodic.

[4]Notice that in OPT-1 we are now optimizing over $L_2[0, T]$, namely $L_2$ with its domain restricted to $[0, T]$, rather than the smaller vector space $PC_\infty[0, T]$, namely $PC_\infty$ with its domain restricted to $[0, T]$. However, it will turn out that the optimal $u$ lies in $PC_\infty[0, T]$, as seen in Theorem 7, so this is not an issue.

*Example 2:* In the case of a single input, $U = [\underline{b}, \overline{b}]$; we assume, for simplicity, that $\underline{b} > 0$ (otherwise, replace $U$ by $-U$). It is straight-forward to verify that

$$\psi(\xi) = \max_{u \in U} \xi^T u = \begin{cases} \overline{b}\xi & \xi \geq 0 \\ \underline{b}\xi & \xi < 0 \end{cases}$$

$$\phi(\xi) = \arg\max_{u \in U} \xi^T u = \begin{cases} \overline{b} & \xi > 0 \\ \underline{b} & \xi < 0. \end{cases}$$

*Example 3:* Consider the case of $U$ being an ellipsoid: with $J > 0$ positive definite and symmetric and $u_0 \in \mathbf{R}^m$, consider

$$U = \{u \in \mathbf{R}^m : \|J(u - u_0)\|_2 \leq 1\}.$$

Using a Lagrange Multiplier approach, it is easy to verify that

$$\psi(\xi) = \max_{u \in U} \xi^T u = \xi^T u_0 + \frac{\xi^T J^{-2} \xi}{\|J^{-1}\xi\|},$$

$$\phi(\xi) = \arg\max_{u \in U} \xi^T u = u_0 + \frac{1}{\|J^{-1}\xi\|} J^{-2} \xi.$$

It turns out that OPT-1 is a special case of the classical Meyer Problem. A detailed study of this problem is given in [15], which is based in large part on the results of Rockafellar [19]. In the following we provide a summary of the main steps.

*Proposition 7 (Theorem 1.1 of [15]):* Subject to Standing Assumption 4, every optimal control $u^*$ of OPT-1 is of the form

$$u^*(t) = \phi((B^+)^T \eta(t))$$

subject to

$$\dot{x}^* = A^+ x^* + B^+ u^*, x^*(0) = x_0^+$$
$$\dot{\eta} = -(A^+)^T \eta, \eta(T) = -(\partial_x g)(x^*(T)).$$

So to obtain the optimal cost it is enough to find all triples $(u^*, x^*, \eta)$ which satisfy Proposition 7 and then find the one which is optimal. This constrained optimization problem is difficult so we turn to a dual problem.

*Optimization Problem 2 (OPT-2):* Minimize

$$h(\alpha) := \alpha^T x_0^+ + \frac{1}{4}\alpha^T W \alpha + \int_0^T \psi\left((B^+)^T e^{-(A^+)^T t} \alpha\right) dt$$

subject to $\alpha \in \mathbf{R}^n$.

The first two terms of $h$ form a strictly convex function, while the convexity of $\psi$ ensures that the last term is convex, which means that $h$ is strictly convex. There is a broad range of software tools which can be used to minimize it (assuming that $\psi$ can be easily computed, as it will be in several special cases), yielding the unique minimum; since $\psi(0) = 0$, this minimum is at most zero. It turns out that the solutions of OPT-1 and OPT-2 are related.

*Theorem 7 (Theorem 4.3 of [15]):* Subject to Standing Assumptions 1–4, the optimal solution $\alpha^*$ of OPT-2 yields the optimal solution of OPT-1:
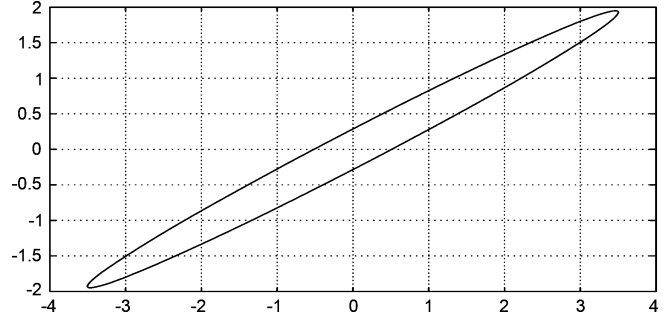
$$J_{opt}^T(r, x_{p_0}) = [-h(\alpha^*)]^{1/2}$$



Fig. 1.   The surface of $CIM(r_0)$.

and the optimal control is

$$u^*(t) = \phi((B^+)^T e^{-(A^+)^T t} \alpha^*), \ t \in [0, T].$$

*Remark 12:* Suppose that $A^+ = \lambda$, $B^+ = 1$, and $CIM(r) = [\underline{b}, \overline{b}]$ with $\underline{b} > 0$. A routine calculation yields

$$h(\alpha) = \alpha x_0^+ + \frac{\alpha^2}{8\lambda} + \begin{cases} \underline{b}(1 - e^{-\lambda T})\frac{\alpha}{\lambda} & \alpha < 0 \\ \overline{b}(1 - e^{-\lambda T})\frac{\alpha}{\lambda} & \alpha \geq 0. \end{cases}$$

Notice that while the last term on the RHS is globally Lipshitz continuous it is not differentiable at zero. After some careful analysis, it can be verified that the optimal cost $J_{opt}^T(r, x_{p_0})$ is given by

$$\frac{\sqrt{2}}{\lambda} \begin{cases} [-\lambda x_0^+ - \overline{b}(1 - e^{-\lambda T})] & x_0^+ < \frac{\overline{b}(e^{-\lambda T} - 1)}{\lambda} \\ 0 & x_0^+ \in \left[\frac{\overline{b}(e^{-\lambda T} - 1)}{\lambda}, \frac{\underline{b}(e^{-\lambda T} - 1)}{\lambda}\right] \\ [\lambda x_0^+ + \underline{b}(1 - e^{-\lambda T})] & x_0^+ > \frac{\underline{b}(e^{-\lambda T} - 1)}{\lambda} \end{cases}$$

which is within $O(e^{-\lambda T})$ of the optimal cost, $J_{opt}(r, x_{p_0})$, which is computed in Remark 6.

We now consider an example from the literature to illustrate several of the results in this paper.

## VII. AN EXAMPLE

Here we consider the example of [2], namely, a vehicle with mass $M$ moving in the plane, on top of which is a mass $m$. This may be modelled by

$$M\ddot{y} = D(\dot{z} - \dot{y}) + u$$
$$m\ddot{z} = D(\dot{y} - \dot{z}) + G(z - y),$$

with $D = \mathrm{diag}(d_1, d_2) > 0$ associated with the viscous friction, $G = \mathrm{diag}(g_1, g_2) > 0$ associated with gravity, $u(t) \in \mathbf{R}^2$ representing the force, while $y(t) \in \mathbf{R}^2$ and $z(t) \in \mathbf{R}^2$ are the positions of the vehicle and mass, respectively. It turns out that the system has two non-minimum phase zeros. For the choice of

$$M = 1, \ m = 0.1, \ d_1 = 1,$$
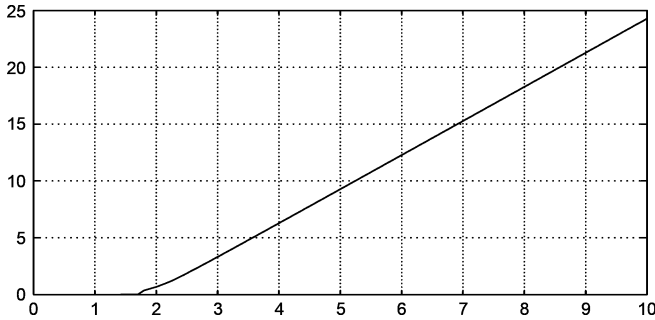$$d_2 = 2, \ g_1 = 1.5, \ g_2 = 1,$$

Fig. 2. A plot of $J_{opt}^T(r_a, 0)$ as a function of $a$.

these zeros are 0.488 and 1.324; the unstable zero dynamics associated with $x^+$ are given by

$$\dot{x}^+ = \underbrace{\begin{bmatrix} 0.488 & 0 \\ 0 & 1.324 \end{bmatrix}}_{=:A^+} x^+ + \underbrace{\begin{bmatrix} 0 & .53 \\ -1.97 & 0 \end{bmatrix}}_{=:B^+} y,$$

$$x^+(0) = x_0^+.$$

With $x_0^+ = 0$, in [2] it is shown that the infimal $L_2$ path following performance can be made as small as desired for all paths arising from a reference trajectory which is a finite sum of sinusoids (with non-zero frequency); this means, in particular, that 0 is contained in $CIM(r)$. As demonstrated in Remark 8, 0 need not lie in $CIM(r)$ for the infimal cost to be zero. To this end, consider the periodic reference signal

$$r_a(t) = \begin{bmatrix} 3.1216 & 1.5884 \\ 1.5884 & 1.2861 \end{bmatrix} \left( \begin{bmatrix} 0 \\ a \end{bmatrix} + \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \right).$$

The surface of $CIM(r_0)$[5] is given in Fig. 1; clearly $CIM(r_a)$ can be obtained by shifting $CIM(r_0)$ by $\begin{bmatrix} 1.5884 \\ 1.2861 \end{bmatrix} a$. From the formula for $r_a$

$$0 \in CIM(r_a) \Leftrightarrow a \in [-1, 1]$$

so it is immediate from Theorem 1 that

$$J_{opt}(r_a, 0) = 0, a \in [-1, 1].$$

Using Theorem 7 we can compute $J_{opt}^T(r_a, 0)$ as a function of $a$—see Fig. 2 (we choose $T = 20$ here). Observe, in particular, that $J_{opt}^T(r_a, 0) = 0$ for $a \in [0, 1.8]$.

## VIII. Conclusion

In this paper we have considered path following performance limitations for linear multivariable non-minimum phase systems. In particular, we have studied problems related to infimal $L_2$ performance. By using the zero dynamics form of a system we are able to derive a number of results, including sufficient conditions for the infimal cost to be zero and other sufficient conditions for it to be strictly greater than zero. In a reasonably general situation, we show that the infimal cost can be computed by solving a finite dimensional convex optimization problem. Our focus has been on analysis of the achievable performance level.

[5]We adopt the natural notation of $r_0(t) = [r_a(t)]|_{a=0}$.

More specifically, we have considered LTI square plants with no transmission zeros on the imaginary axis and with uniform vector relative degree[6], which we label $p$. We require that the path be sufficiently smooth (be continuously differentiable up to the $p$th degree) and be bounded (as well as the derivatives up to the $p$th one). The goal is to to obtain answers to two key questions, namely

i) For which trajectories can we make the cost as small as desired?

ii) How do we compute the (infimal) optimal cost when it is non-zero?

First, we simplify the cost function(s) in a sequence of three steps, thereby converting it into a classical optimization problem, albeit an infinite dimensional one. With respect to the first question: a) we use the new problem formulation to provide upper and lower bounds on the cost function, b) we show that it is not necessary (in the case of zero initial conditions) that zero lie inside the convex hull of the trajectory to have the infimal cost be zero and c) in an important special case (scalar output with zero initial conditions) we provide several necessary and sufficient conditions for which the infimal cost is zero. With respect to the second question: under reasonably general conditions we use results from optimal control theory to show how to carry out computations via a finite-dimensional strictly convex optimization problem, which is readily computable using standard software.

There are a number of feasible research directions. First, it would be desireable to design practical controllers which can provide near optimal performance. Second, it would be interesting to examine the same problem in the context of the $\infty$-norm, which may be more appropriate in certain path following problems.

## Appendix

*Proof of ($\leq$) Part of Proposition 2:* Fix $r \in \mathcal{R}_p$ and $x_{p_0} \in \mathbf{R}^{n_p}$. Let $\theta \in \Theta_p$ and $\varepsilon > 0$. We will construct $u_\varepsilon \in PC$ so that $(u_\varepsilon, \theta) \in S(r, x_{p_0})$ and $\lim_{\varepsilon \to 0} J(r, x_{p_0}, \theta, u_\varepsilon) = \rho(r \circ \theta, x_0^+)$. This will mean that

$$J_{opt}(r, x_{p_0}) \leq \rho(r \circ \theta, x_0^+)$$

since this will be true for every $\theta \in \Theta_p$, the result will follow.

As mentioned above, the proof is constructive. The key is to construct a "near optimal error" $e_\varepsilon$ and then prove that it arises from a $u_\varepsilon \in L_\infty$ satisfying $(u_\varepsilon, \theta) \in S(r, x_{p_0})$. To proceed, observe from (8) that to ensure that we solve the path following problem, we will need

$$\int_0^\infty e^{-A^+\tau} B^+ e_\varepsilon(\tau) d\tau$$

$$= x_0^+ + \underbrace{\int_0^\infty e^{-A^+\tau} B^+ (r \circ \theta)(\tau) d\tau}_{=:w}. \quad (21)$$

From Proposition 1 the choice of $e_\varepsilon$ of smallest norm is exactly

$$(B^+)^T e^{-(A^+)^T t} W^{-1}(x_0^+ + w). \quad (22)$$

Unfortunately, because of a typical mis-match of initial conditions at time zero between this quantity, the plant output, and

[6]This assumption was made for clarity.

$r \circ \theta$, to generate this error signal would, in general, require the use of impulses in the control signal, which is inadmissible. Specifically, since the plant has uniform relative degree $p$, every $u_\varepsilon \in PC$ must yield an error signal $e_\varepsilon$ and output signal $y_\varepsilon$ satisfying the following constraint:

$$
\begin{aligned}
e_\varepsilon^{(i)}(0) &= (r \circ \theta)^{(i)}(0) - y_\varepsilon^{(i)}(0) \\
&= (r \circ \theta)^{(i)}(0) - C_p A_p^i x_{p_0}, \ i = 0, 1, \ldots, p-1. \quad (23)
\end{aligned}
$$

So the idea is to adjust (22) slightly so that this holds. To this end, we use an adjusted version of (22), of the form

$$
\begin{aligned}
e_\varepsilon(t) := &\phi_\varepsilon(t)(B^+)^T e^{-(A^+)^T t} W^{-1}(x_0^+ + w + \Delta_\varepsilon) \\
&+ [1 - \phi_\varepsilon(t)] \times [(r \circ \theta)(t) - C_p e^{A_p t} x_{p_0}] \quad (24)
\end{aligned}
$$

notice that (23) holds but we will need to find $\Delta_\varepsilon$ to ensure that (21) holds. It is easy to check that we should set

$$
\begin{aligned}
\Delta_\varepsilon := &(W - W_\varepsilon) W_\varepsilon^{-1}(x_0^+ + w) \\
&- W W_\varepsilon^{-1} \int_0^\varepsilon [1 - \phi_\varepsilon(\tau)] e^{-A^+ \tau} B^+ \\
&\times [(r \circ \theta)(\tau) - C_p e^{A_p \tau} x_{p_0}] \, d\tau
\end{aligned}
$$

which clearly converges to zero as $\varepsilon \to 0$.

We now define the corresponding output by

$$
\begin{aligned}
y_\varepsilon(t) &= (r \circ \theta)(t) - e_\varepsilon(t) \\
&= (r \circ \theta)(t) - \phi_\varepsilon(t)(B^+)^T e^{-(A^+)^T t} W^{-1} \\
&\quad \times (x_0^+ + w + \Delta_\varepsilon) \\
&\quad - [1 - \phi_\varepsilon(t)] \times [(r \circ \theta)(t) - C_p e^{A_p t} x_{p_0}].
\end{aligned}
$$

It is routine to check that

$$
y_\varepsilon^{(i)} \in L_\infty, i = 0, 1, \ldots, p, \quad (25)
$$

$$
y_\varepsilon^{(i)}(0) = C_p A_p^i x_{p_0}, i = 0, 1, \ldots, p, \quad (26)
$$

$$
\rho(y_\varepsilon, x_0^+) = 0, \quad (27)
$$

$$
y_\varepsilon(t) - r(\theta(t)) \to 0 \text{ as } t \to \infty. \quad (28)
$$

We now define the "near optimal" plant state by

$$
\left.
\begin{aligned}
x_\varepsilon^+(t) &:= e^{A^+ t} x_0^+ + \int_0^t e^{A^+(t-\tau)} B^+ y_\varepsilon(\tau) d\tau, \\
x_\varepsilon^-(t) &:= e^{A^- t} x_0^- + \int_0^t e^{A^-(t-\tau)} B^- y_\varepsilon(\tau) d\tau,
\end{aligned}
\right\} \quad (29)
$$

$$
v_\varepsilon := \begin{bmatrix} y_\varepsilon \\ y_\varepsilon^{(1)} \\ \vdots \\ y_\varepsilon^{(p-1)} \end{bmatrix} \quad (30)
$$

and the corresponding control signal by

$$
u_\varepsilon := (B_v^T B_v)^{-1} B_v^T [\dot{v}_\varepsilon - A_v v_\varepsilon - A_1 x_\varepsilon^+ - A_2 x_\varepsilon^-]. \quad (31)
$$

It is clear by construction that the above state, output and input satisfy the plant equation, and it follows from (29), evaluated at $t = 0$, together with (26), that

$$
\begin{bmatrix} x_\varepsilon^+ \\ x_\varepsilon^- \\ v_\varepsilon \end{bmatrix}(0) = x_{p_0}.
$$

Furthermore, it follows easily from (25), (27), (29), (30), and (31) that

$$
x_\varepsilon^+, x_\varepsilon^-, v_\varepsilon, u_\varepsilon \in L_\infty.
$$

Using the formula for $e_\varepsilon$ given in (24) together with the fact that $\lim_{\varepsilon \to 0} \Delta_\varepsilon = 0$, it follows that the corresponding tracking cost is:

$$
\begin{aligned}
J(r, x_{p_0}, \theta, u_\varepsilon)^2 &= \|e_\varepsilon\|_2^2 \\
&= \int_0^\infty \|\phi_\varepsilon(\tau)(B^+)^T e^{-(A^+)^T \tau} W^{-1}(x_0^+ + w_\varepsilon + \Delta_\varepsilon) \\
&\quad + [1 - \phi_\varepsilon(\tau)] \times [(r \circ \theta)(\tau) - C_p e^{A_p \tau} x_{p_0}]\|^2 \, d\tau \\
&\to \int_0^\infty \|(B^+)^T e^{-(A^+)^T \tau} W^{-1}(x_0^+ + w)\|^2 \, d\tau \ (\text{as } \varepsilon \to 0) \\
&= (x_0 + w)^T W^{-1}(x_0 + w).
\end{aligned}
$$

Combined with (28), it follows that the pair $(u_\varepsilon, \theta)$ stabilizes and solves the path following problem for (1), i.e., $(u_\varepsilon, \theta) \in S(r, x_{p_0})$; furthermore

$$
\begin{aligned}
&(x_0^+ + w)^T W^{-1}(x_0^+ + w) \\
&= \|W^{-1/2} x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+ \tau} B^+ r(\theta(\tau)) d\tau\|^2 \\
&= \rho(r \circ \theta, x_0^+)^2
\end{aligned}
$$

so $J(r, x_{p_0}, \theta, u_\varepsilon) \to \rho(r \circ \theta, x_0^+)$ as $\varepsilon \to 0$, as required.

The proof for $J_{opt}^+(r, x_{p_0})$ is virtually identical, the only change being that now we choose $\theta \in \Theta_p^+$. □

*Proof of Proposition 3:* Let $r \in \mathcal{R}_p$ and $x_{p_0} \in \mathbf{R}^{n_p}$. It follows that:

$$
c_1 := \sup_{\tau \geq 0} \|\dot{r}(\tau)\| < \infty. \quad (32)
$$

We first prove part (i).

($\leq$) Let $\theta \in \Theta_p$ and $\varepsilon > 0$. It follows that there exists a $T > 0$ so that $\dot{\theta}(t) = 1$ for $t > T$. With $\ell \in \mathbf{N}$ define $h := T/\ell$ and the sampled version of $\theta$

$$
\hat{\theta}(t) = \begin{cases} \theta(ih) & t \in [ih, (i+1)h), \ i = 0, 1, \ldots, \ell - 1, \\ \theta(t) & t \geq T = \ell h. \end{cases}
$$

This means that for $t \in [ih, (i+1)h)$ we have $|t - ih| < h$ so

$$
\begin{aligned}
\|r(\theta(t)) - r(\hat{\theta}(t))\| &\leq c_1 |\theta(t) - \hat{\theta}(t)| \leq c_1 \|\dot{\theta}\|_\infty h \\
&= c_1 \|\dot{\theta}\|_\infty \frac{T}{\ell}.
\end{aligned}
$$

Hence, with $c_0$ and $\lambda_0$ chosen so that (16) holds, we have

$$\left\| \int_0^\infty e^{-A^+\tau} B^+ [r(\theta(\tau)) - r(\hat{\theta}(\tau))] d\tau \right\|$$
$$\leq \int_0^T c_0 e^{\lambda_0 \tau} \|B^+\| \|c_1\| \|\dot{\theta}\|_\infty \frac{T}{\ell} d\tau$$
$$\leq \frac{c_0 c_1 T \|B^+\|}{\ell |\lambda_0|} \|\dot{\theta}\|_\infty$$

so for sufficiently large $\ell$ we have

$$\left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+\tau} B^+ r(\hat{\theta}(\tau)) d\tau \right\|$$
$$\leq \left\| W^{-1/2} x_0^+ + W^{-1/2} \int_0^\infty e^{-A^+\tau} B^+ r(\theta(\tau)) d\tau \right\| + \varepsilon.$$

Since the above argument holds for every $\theta \in \Theta_p$ and $\varepsilon > 0$, we see that

$$\hat{J}_{opt}(r, x_{p_0}) \leq J_{opt}(r, x_{p_0}).$$

($\geq$) Now suppose that $\theta(t)$ is piecewise constant, with a finite number of discontinuities on $[0, T]$, and $\dot{\theta}(t) = 1$ for $t \geq T$. So there exist $0 =: t_1 < t_2 < \cdots < t_\ell := T$ and $d_i \in \mathbf{R}$ so that

$$\theta(t) = \begin{cases} d_i & t \in (t_i, t_{i+1}), \ i = 0, \ldots, \ell - 1, \\ d_\ell + t & t > T. \end{cases}$$

Let $\Delta$ denote the smallest time between discontinuities:

$$\Delta := \min_{i \in \{1, 2, \ldots, \ell - 1\}} |t_{i+1} - t_i|.$$

To obtain a smoothed version of $\theta$ we use convolution: with $\varepsilon \in (0, \Delta]$ define

$$\hat{\theta}(t) := (\theta * \delta_\varepsilon)(t) = \int_0^t \theta(t - \tau) \delta_\varepsilon(\tau) d\tau, t \in \mathbf{R}.$$

Clearly $\hat{\theta} \in \mathcal{C}^p$. Indeed, it is routine to confirm that

$$\hat{\theta}(t) = \theta(t), t \in [t_i + \varepsilon, t_{i+1}), i = 1, \ldots, \ell - 1,$$
$$|\hat{\theta}(t)| \leq \underbrace{\max_{\tau \in [0, T+\Delta]} |\theta(\tau)|}_{=:c_2}, t \in [t_i, t_i + \varepsilon], i = 1, \ldots, \ell,$$
$$\hat{\theta}(t) \in [\theta(t) - \varepsilon, \theta(t)], t \geq t_\ell + \varepsilon = T + \varepsilon.$$

It follows from the definition of $c_1$ given in (32) that:

$$\|r(\theta(t)) - r(\hat{\theta}(t))\| \leq c_1 \|\theta(t) - \hat{\theta}(t))\|, t \geq 0,$$
$$\|\theta(t) - \hat{\theta}(t)\| \leq \varepsilon, t \geq T + \Delta.$$

Hence

$$\rho(r \circ \hat{\theta} - r \circ \theta, 0)$$
$$\leq \|W^{-1/2}\| \int_0^{T+\Delta} c_0 e^{\lambda_0 \tau} \|B^+\| c_1 \|\theta(\tau) - \hat{\theta}(\tau)\| d\tau$$
$$+ \|W^{-1/2}\| \int_{T+\Delta}^\infty c_0 e^{\lambda_0 \tau} \|B^+\| \varepsilon c_1 \, d\tau$$
$$\leq \underbrace{c_0 \|W^{-1/2}\| \times \|B^+\| [2c_1 c_2 \ell + \frac{c_1}{|\lambda_0|} e^{\lambda_0 (T+\Delta)}]}_{=:c_3} \varepsilon.$$

Therefore, using a basic property of $\rho$, it follows that:

$$\rho(r \circ \hat{\theta}, x_0^+) \leq \rho(r \circ \theta, x_0^+) + c_3 \varepsilon;$$

since $\hat{\theta} \in \Theta_p$, it follows that:

$$J_{opt}(r, x_{p_0}) \leq \rho(r \circ \theta, x_0^+) + c_3 \varepsilon.$$

Since this is independent of $\theta$ and $\varepsilon > 0$, it follows that:

$$J_{opt}(r, x_{p_0}) \leq \hat{J}_{opt}(r, x_{p_0})$$

as desired.

The proof of (ii) is identical, once it is observed that, in both parts of the proof, if $\theta$ is monotonically increasing then $\hat{\theta}$ is as well. □

*Proof of Proposition 4:* ($\leq$) This proof follows immediately from Remark 6.

($\geq$) Let $x_{p_0} \in \mathbf{R}^{n_p}$ and $r \in \mathcal{R}_p$. Since the proof of i) and ii) are similar, with a difference at one key step, we will do them simultaneously. Suppose that $f \in PC_\infty$ and takes values in $CIM(r)$ or $CIM^+(r)$, as appropriate. Fix $\varepsilon > 0$ and choose $T > 0$ so that

$$\left\| W^{-1/2} \int_T^\infty e^{-A^+\tau} B^+ f(\tau) \, d\tau \right\|$$
$$+ \left\| W^{-1/2} \int_T^\infty e^{-A^+\tau} B^+ r(\tau) \, d\tau \right\| \leq c_0 \|W^{-1/2}\|$$
$$\times \|B^+\| \times \int_T^\infty e^{\lambda_0 \tau} (\|f(\tau)\| + \|r(\tau)\|) \, d\tau \leq \varepsilon. \quad (33)$$

The first step is to sample $f$: with $\ell \in \mathbf{N}$ define

$$\hat{f}(t) := f\left(i\frac{T}{\ell}\right),$$
$$t \in \left[i\frac{T}{\ell}, (i+1)\frac{T}{\ell}\right), \ i = 0, 1, \ldots, \ell - 1.$$

Since $f$ is piecewise continuous on $[0, T]$, it is a basic fact from calculus that we can choose $\ell \in \mathbf{N}$ so that

$$\rho_T(\hat{f} - f, 0) < \varepsilon. \quad (34)$$

Fix $\ell$ with this property and define $h := T/\ell$.

Now $\hat{f}$ takes values in $CIM(r)$ or $CIM^+(r)$, as appropriate. With $\bar{\ell} := (m+1)\ell$, by Carotheodory's Theorem there exists a subset $\{v_j : j = 1, 2, \ldots, \bar{\ell}\}$ of $IM(r)$ or $IM^+(r)$, as appropriate,[7] and constants $c_{ij} \geq 0$ so that

$$\left\| \hat{f}(ih) - \sum_{j=1}^{\bar{\ell}} c_{ij} v_j \right\| \leq \varepsilon, \ \sum_{j=1}^{\bar{\ell}} c_{ij} = 1, i = 0, 1, \ldots, \ell - 1.$$

Hence, define

$$\bar{f}(t) = \sum_{j=1}^{\bar{\ell}} c_{ij} v_j,$$
$$t \in [ih, (i+1)h), \ i = 0, 1, \ldots, \ell - 1$$

[7]If we are dealing with Case (i), then we also choose $\{t_j\}$ so that $v_j = r(t_j)$.

which means that

$$\rho_T(\bar{f} - \hat{f}, 0) < \frac{c_0}{|\lambda_0|} \|W^{-1/2}\| \times \|B^+\|\varepsilon. \qquad (35)$$

The difficulty is that we need a function $\tilde{f}$ which takes values only in $IM(r)$ or $IM^+(r)$, as appropriate. We would like to do so in such a way that $\rho_T(\bar{f} - \tilde{f}, 0)$ is small. The key idea is to split each interval $[ih, (i+1)h)$ up into tiny sub-intervals of length $\bar{h}$, and have $\tilde{f}$ cycle rapidly between elements (or good approximations) of $\{v_j : j = 1, \ldots, \bar{\ell}\}$, spending $c_{ij}\bar{h}$ time units at $v_j$ (or its approximant); in this way, under the integral sign the effect of $\tilde{f}$ over the interval $[ih, (i+1)h)$ should be very close to that of its average, namely $\bar{f}(ih)$. To proceed, we let $q \in \mathbf{N}$ and define $\bar{h} := h/q$. It is convenient at this point to rewrite $\bar{f}$ in terms of $\bar{h}$: for each $i \in \{0, 1, \ldots, q\ell - 1\}$ and $j \in \{1, \ldots, \bar{\ell}\}$ there exists $d_{ij} \in \{c_{i_1 i_2} : i_1 = 0, 1, \ldots, \ell - 1, i_2 = 1, \ldots, \bar{\ell}\}$ so that $\sum_{j=1}^{\bar{\ell}} d_{ij} = 1$ and

$$\bar{f}(t) = \sum_{j=1}^{\bar{\ell}} d_{ij} v_j, t \in [i\bar{h}, (i+1)\bar{h})$$

we simply use the $\{c_{i_1 i_2}\}$'s exactly $q$ times each.

Now we define $\tilde{f}_q$ and $\theta_q$. For each $i \in \{0, 1, \ldots, q\ell - 1\}$, in Case (i) we define

$$(\tilde{f}_q, \theta_q)(i\bar{h} + \tau)$$
$$= \begin{cases} (v_1, t_1) & \tau \in [0, d_{i1}\bar{h}) \\ (v_2, t_2) & \tau \in [d_{i1}\bar{h}, (d_{i1} + d_{i2})\bar{h}) \\ \vdots & \vdots \\ (v_{\bar{\ell}}, t_{\bar{\ell}}) & \tau \in [(d_{i1} + \cdots + d_{i,\bar{\ell}-1})\bar{h}, \bar{h}) \end{cases}$$

with the understanding that some of the intervals may be empty, and define

$$(\tilde{f}_q, \theta_q)(t) = (r(t), t), t \geq T.$$

Hence, $\tilde{f}_q = r \circ \theta_q$, and $\theta_q$ satisfies the constraints imposed on it in the definition of $\hat{J}_{opt}^+(r, x_{p_0})$.

In Case (ii) we first choose the strictly increasing sequence

$$0 < t_{11} < t_{21} < \cdots < t_{\bar{\ell}1} < t_{12} < t_{22}$$
$$< \ldots < t_{\bar{\ell}2} < \cdots < t_{1\bar{\ell}} < t_{2\bar{\ell}} < \cdots < t_{\bar{\ell}\bar{\ell}}$$

satisfying

$$\|v_j - \underbrace{r(t_{ji})}_{=:v_{ji}}\| < \varepsilon, i = 1, 2, \ldots, \ell.$$

For each $i \in \{0, 1, \ldots, q\ell - 1\}$ we then define

$$(\tilde{f}_q, \tilde{\theta}_q)(i\bar{h} + \tau)$$
$$= \begin{cases} (v_{1i}, t_{1i}) & \tau \in [0, d_{i1}\bar{h}) \\ (v_{2i}, t_{2i}) & \tau \in [d_{i1}\bar{h}, (d_{i1} + d_{i2})\bar{h}) \\ \vdots & \vdots \\ (v_{\bar{\ell}i}, t_{\bar{\ell}i}) & \tau \in [(d_{i1} + \cdots + d_{i,\bar{\ell}-1})\bar{h}, \bar{h}) \end{cases}$$

with the understanding that some of the intervals may be empty, and define

$$(\tilde{f}_q, \theta_q)(t) = (r(t + T + t_{\bar{\ell}\bar{\ell}}), t + T + t_{\bar{\ell}\bar{\ell}}), t \geq T.$$

Hence, $\tilde{f}_q = r \circ \theta_q$, and $\theta_q$ satisfies the constraints imposed on it in the definition of $\hat{J}_{opt}^+(r, x_{p_0})$.

We claim that $\rho_T(\bar{f} - \tilde{f}_q, 0)$ becomes small as $q \to \infty$. First observe that there exists a $\bar{c} > 0$ so that

$$\|\underbrace{e^{-A^+\mu} - I}_{=:g(\mu)}\| \leq \bar{c}\mu, \mu \in [0, h]. \qquad (36)$$

It follows that:

$$\rho_T(\bar{f} - \tilde{f}_q, 0)$$
$$= \left\| W^{-1/2} \int_0^T e^{-A^+\tau} B^+ [\bar{f}(\tau) - \tilde{f}_q(\tau)] \, d\tau \right\|$$
$$= \left\| W^{-1/2} \sum_{i=0}^{q\ell-1} \int_{i\bar{h}}^{(i+1)\bar{h}} e^{-A^+\tau} B^+ [\bar{f}(\tau) - \tilde{f}_q(\tau)] \, d\tau \right\|$$
$$\leq \|W^{-1/2}\| \sum_{i=0}^{q\ell-1} \|e^{-A^+ i\bar{h}}\|$$
$$\times \left\| \int_{i\bar{h}}^{(i+1)\bar{h}} e^{-A^+(\tau - i\bar{h})} B^+ [\bar{f}(\tau) - \tilde{f}_q(\tau)] \, d\tau \right\|$$
$$\leq c_0 \|W^{-1/2}\| \sum_{i=0}^{q\ell-1} e^{\lambda_0 i\bar{h}}$$
$$\times \left\| \int_0^{\bar{h}} e^{-A^+\tau} B^+ [\bar{f}(i\bar{h} + \tau) - \tilde{f}_q(i\bar{h} + \tau)] \, d\tau \right\|. \quad (37)$$

With $g$ defined by (36) and using the definition of the $v'_{ij}s$:

$$\left\| \int_0^{\bar{h}} e^{-A^+\tau} B^+ [\bar{f}(i\bar{h} + \tau) - \tilde{f}_q(i\bar{h} + \tau)] \, d\tau \right\|$$
$$\leq \left\| \int_0^{\bar{h}} g(\tau) B^+ [\bar{f}(i\bar{h} + \tau) - \tilde{f}_q(i\bar{h} + \tau)] \, d\tau \right\|$$
$$+ \left\| B^+ \int_0^{\bar{h}} [\bar{f}(i\bar{h} + \tau) - \tilde{f}_q(i\bar{h} + \tau)] \, d\tau \right\|$$
$$\leq \int_0^{\bar{h}} \bar{c}\tau \|B^+\| \left( 2 \max_{i=1,\ldots,\bar{\ell}} \|v_i\| + \varepsilon \right) d\tau$$
$$+ \|B^+\| \times \left\| \int_0^{\bar{h}} [\bar{f}(i\bar{h} + \tau) - \tilde{f}_q(i\bar{h} + \tau)] \, d\tau \right\|$$
$$\leq \bar{c}\frac{\bar{h}^2}{2} \|B^+\| \left( 2 \max_{i=1,\ldots,\bar{\ell}} \|v_i\| + \varepsilon \right) + \bar{h}\varepsilon \|B^+\|.$$

Hence, using this in (37) and recalling that $\bar{h} = h/q$ yields

$$\limsup_{q \to \infty} \rho_T(\bar{f} - \tilde{f}_q, 0) \leq \|W^{-1/2}\| \frac{c_0}{|\lambda_0|} \|B^+\|\varepsilon.$$

Hence, if we combine this with (34) and (35) we have

$$\limsup_{q \to \infty} \rho_T(f - \tilde{f}_q, 0) \leq (1 + \frac{2c_0}{|\lambda_0|} \|W^{-1/2}\| \times \|B^+\|)\varepsilon. \quad (38)$$

From the definition of $\theta_q$ it is clear that $\tilde{f}_q = r \circ \theta_q$ satisfies

$$\left\| W^{-1/2} \int_T^\infty e^{-A^+\tau} B^+ \tilde{f}_q(\tau) \, d\tau \right\|$$
$$\leq c_0 \|W^{-1/2}\| \cdot \|B^+\| \int_T^\infty e^{\lambda_0 \tau} \|r(\tau)\| \, d\tau,$$

so if we combine (33) and (38) we have

$$\limsup_{q \to \infty} \rho(f - \tilde{f}_q, 0) \leq \left[ 2 + \frac{2c_0}{|\lambda_0|} \|W^{-1/2}\| \times \|B^+\| \right] \varepsilon.$$

Since $f \in PC_\infty$ (taking values in $CIM(r)$ or $CIM^+(r)$, as appropriate) and $\varepsilon > 0$ are arbitrary, it follows that:

$$\hat{J}_{opt}(r, x_{p_0}) \leq \tilde{J}_{opt}(r, x_{p_0}),$$
$$\hat{J}_{opt}^+(r, x_{p_0}) \leq \tilde{J}_{opt}^+(r, x_{p_0}).$$

We now apply Proposition 3 to obtain the desired result. $\quad\square$

## REFERENCES

[1] A. P. Aguiar, D. B. Dacic, J. P. Hespanha, and P. Kokotovic, "Path-following or reference tracking?," in *Proc. IAV2004—5th IFAC/EURON Symp. Intell. Autonomous Veh.*, Lisbon, Portugal, Jul. 2004.

[2] A. P. Aguiar, J. P. Hespanha, and P. Kokotovic, "Path-following for non-minimum phase systems removes performance limitations," *IEEE Trans. Automat. Control*, vol. 50, no. 2, pp. 234–239, Feb. 2005.

[3] A. P. Aguiar, J. P. Hespanha, and P. Kokotovic, "Limits of performance in reference-tracking and path-following for nonlinear systems," in *Proc. IFAC World Congress*, Jul. 2005, [CD ROM].

[4] C. T. Chen, *Linear System Theory and Design*. Oxford, U.K.: Oxford Univ. Press, 1984.

[5] D. Chen and B. Paden, "Stable inversion of nonlinear non-minimum phase systems," *Int. J. Control*, vol. 64, no. 1, pp. 81–97, 1996.

[6] J. Chen, Z. Ren, S. Hara, and L. Qiu, "Optimal tracking performance: Preview control and exponential signals," *IEEE Trans. Automat. Control*, vol. 46, no. 10, pp. 1647–1653, Oct. 2001.

[7] D. B. Dacic and P. V. Kokotovic, "Path-following for linear systems with unstable zero dynamics," *Automatica*, vol. 42, pp. 1673–1683, 2006.

[8] D. B. Dacic, M. V. Subbotin, and P. V. Kokotovic, "Path-following for a class of nonlinear systems with unstable zero dynamics," in *Proc. 43rd IEEE Conf. Decision Control*, Atlantis, Bahamas, 2004.

[9] E. J. Davison and S. H. Wang, "Properties and calculation of transmission zeros of linear multivariable systems," *Automatica*, vol. 10, pp. 643–658, 1974.

[10] B. A. Francis and W. M. Wonham, "The internal model principle for linear multivariable regulators," *Appl. Math. Optim.*, vol. 2, pp. 170–194, 1975.

[11] S. Al-Haddabi and N. McClamroch, "Tracking and manuever regulation control for nonlinear nonminimum phase systems: Application to flight control," *IEEE Trans. Control Syst. Techol.*, vol. 10, no. 6, pp. 780–792, Nov. 2002.

[12] J. Hauser and R. Hindman, "Maneuver regulation from trajectory tracking: Feedback linearizable systems," in *Proc. IFAC Symp. Nonlinear Control Syst. Design*, Lake Tahoe, CA, 1995, pp. 595–600.

[13] T. Hu, Z. Lin, and L. Qiu, "An explicit description of null controllable regions of linear systems with saturating actuators," *Syst. Control Lett.*, vol. 47, pp. 65–78, 2002.

[14] A. Isidori, *Nonlinear Control Systems*, 3rd ed. New York: Springer, 1995.

[15] B. Liu, "Methods of Convex Analysis in Optimal Control," M.A.Sc. thesis, Dept. Elect. Comput. Eng., Univ. Waterloo, Waterloo, ON, Canada, 2003.

[16] G. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[17] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[18] L. Qiu and E. J. Davison, "Performance limitations of nonminimum phase systems in the servomechanism problem," *Automatica*, vol. 29, pp. 337–349, 1993.

[19] R. T. Rockafellar, "Conjugate convex functions in optimal control and the calculus of variations," *J. Math. Anal. Appl.*, vol. 32, pp. 174–222, 1970.

[20] M. M. Seron, J. H. Braslavsky, P. V. Kokotovic, and D. Q. Mayne, "Feedback limitations in nonlinear systems: From bode integrals to cheap control," *IEEE Trans. Automat. Control*, vol. 44, no. 4, pp. 829–833, Apr. 1999.

[21] R. Skjetne, T. I. Fossen, and P. V. Kokotovic, "Robust output maneuvering for a class of nonlinear systems," *Automatica*, pp. 373–383, 2004.

[22] W. A. Wolovich, *Linear Multivariable Systems*. New York: Springer-Verlag, 1974.

**Daniel E. Miller** (S'82–M'90–SM'03) received the B.Sc. degree from the University of New Brunswick, Fredericton, NB, Canada, in 1984 and the M.A.Sc. and Ph.D. degrees from the University of Toronto, Toronto, ON, Canada, in 1986 and 1990, respectively, all in electrical engineering.

From 1990 to 1991, he held postdoctoral positions at Cambridge University, Cambridge, U.K., the University of Groningen, and the Massachusetts Institute of Technology (MIT), Cambridge. Since 1991, he has been a faculty member in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, where he currently holds the rank of Professor. His research interests include adaptive control, sampled-data control, time-varying control, and performance limitations.

Dr. Miller was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

**Richard H. Middleton** (F'99) was born in Newcastle, Australia, on December 10, 1961. He received the B.Sc., B.Eng., and Ph.D. degrees from the University of Newcastle, Newcastle, Australia, in 1983, 1984, and 1987, respectively.

He served as an Associate Editor of *Automatica*. He was Head of the Electrical and Computer Engineering Department, University of Newcastle. He is currently a Research Professor at the Hamilton Institute, The National University of Ireland, Maynooth. His research interests include a range of control systems theory and applications.

Dr. Middleton is a Distinguished Lecturer of IEEE Control Systems Society and was a Vice President of the IEEE Control Systems Society (CSS). He served as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY.