

Iterative Sequential Action Control for Stable, Model-Based Control of Nonlinear Systems

Emmanouil Tzorakoleftherakis, *Student Member, IEEE*, and Todd D. Murphey, *Member, IEEE*

Abstract—This paper presents iterative Sequential Action Control (iSAC), a receding horizon approach for control of nonlinear systems. The iSAC method has a closed-form open-loop solution, which is iteratively updated between time steps by introducing constant control values applied for short duration. Application of a contractive constraint on the cost is shown to lead to closed-loop asymptotic stability under mild assumptions. The effect of asymptotically decaying disturbances on system trajectories is also examined. To demonstrate the applicability of iSAC to a variety of systems and conditions, we employ five different systems, including a 13-dimensional quaternion-based quadrotor. Each system is tested in different scenarios, ranging from feasible and infeasible trajectory tracking, to setpoint stabilization, with or without the presence of external disturbances. Finally, limitations of this work are discussed.

I. INTRODUCTION

Model-predictive control (MPC) has been widely used over the past twenty years for control of linear and nonlinear systems [1], [2]. The rationale behind MPC is the following: at each time step, an L_2 or alternative variation of the cost function is locally optimized over time to obtain the open-loop control as a function of time, only a small portion of which is actually applied to the system. The time horizon is then shifted, and the process is repeated at a later time step based on state feedback. Using this approach, it is clear that optimizing with respect to the entire open-loop control, although often effective, might not be the most efficient way to compute the control since most of the optimizer is typically discarded. In view of this observation, our previous work in [3]–[7] presented Sequential Action Control (SAC), a receding horizon approach for control of nonlinear systems, that exploits elements from hybrid systems theory [8], [9]. In contrast with the aforementioned MPC methods, the open-loop solution in SAC optimizes the needle variation [10] of the cost resulting in a single, constant magnitude action which does not optimize, but rather improves the cost function value across time steps. Since only a single action—obtained in closed-form—is computed at each time step, control calculation is efficient. Our earlier work in [3]–[7] indicates that SAC can drive benchmark and challenging systems—including the cart-pendulum, acrobot and pendubot in [3], hopping and humanoid locomotion in [5] and rotor vehicles with dynamics on Lie groups in [6]—close to a desired equilibrium. Nevertheless, it cannot achieve final stabilization and, as a result, switching to a locally stabilizing controller is often necessary [3], [6].

This paper presents iterative Sequential Action Control (iSAC), an extension of SAC that addresses one fundamental question that was left open in our previous work—how can SAC achieve consistent stabilizing behavior? We provide theoretical and practical results verifying that the iSAC method leads to closed-loop stability, while being efficient enough for application in situations with limited computational resources, e.g., mobile phones [5]. Simulation examples show that iSAC is consistently successful in a variety of control scenarios, ranging from benchmark inversion problems [11] to control of quadrotors performing complex three-dimensional maneuvers like flips. Unlike other system-specific methods (see Section VI), our approach leads to closed-loop stability (when the task is feasible) without modifications across examples; only the dynamic model that is being controlled is changed for each example.

In addition to the aforementioned points, the following novelties of iSAC distinguish this work from alternative MPC methods (see, e.g., [1], [2], [12]–[16] and references therein). In order to solve the open-loop problem, most MPC methods either employ nonlinear programming solvers (see [17] and [18] for a review) or solve a matrix Riccati differential equation as, for example, in [12]–[14]. On the contrary, the solution of the open-loop problem in iSAC has an analytic form which requires only forward simulations of the dynamics and a costate variable, which is computationally efficient and does not depend on nonlinear programming libraries. Moreover, control saturations can be incorporated without additional computational overhead. Finally, as opposed to many MPC alternatives that utilize discrete-time dynamics [1], [2], [14], [15], [19], iSAC uses continuous-time dynamics. As a result, variable-step integration may be used, which, combined with the previous points, significantly speeds-up the solution process.

To achieve closed-loop stability for iSAC, we apply a contractive constraint on the cost [20]–[23]. Contractive constraints have been widely used in the MPC literature to show closed-loop stability as an alternative to methods relying on a terminal (region) constraint [1], [2], [15], [16], [24], [25]. A disadvantage of these methods is that they require the computation of a terminal region, which has to be calculated separately for each system of interest. Many procedures have been proposed on how to find this region; see, e.g., [19], [24], [26]–[28] and references therein. On the contrary, the contractive constraint approach can be applied to a variety of examples without modification and is easily implementable in practice. In our case, a contractive constraint on the cost can be naturally integrated in iSAC, without additional work,

Authors are with the Neuroscience and Robotics Laboratory (N×R) at the Department of Mechanical Engineering, Northwestern University, Evanston, IL. Email: man7therakis@u.northwestern.edu, t-murphey@northwestern.edu.

by exploiting tools from hybrid systems theory, e.g., the mode insertion gradient (see [8], [9]), as will be explained in Section III.

This paper is structured as follows: Section II provides a description of SAC; Section III describes iSAC and the features that differentiate this work from SAC. Sections IV and V analyze local and global stability properties of iSAC respectively. Section VI demonstrates applicability of iSAC to a variety of systems and control scenarios while conclusive remarks are given in Section VII. Finally, all proofs are presented in the Appendix.

II. PRELIMINARIES - SEQUENTIAL ACTION CONTROL

For convenience, we briefly summarize the work presented in [3]–[6]. SAC is a receding horizon method that enables real-time, closed-loop constrained control synthesis for broad and challenging classes of systems, including by following the cyclic diagram in Fig. 1a. We shall consider nonlinear systems with input constraints such that

$$\dot{x} = f(t, x, u) = g(t, x) + h(t, x) u \quad \forall t \quad (1)$$

with $u \in \mathcal{U}$ and

$$\mathcal{U} := \left\{ u \in \mathbb{R}^m : u_{\min} \leq u \leq u_{\max}, u_{\min} \leq 0 \leq u_{\max} \right\},$$

i.e., systems that can be nonlinear with respect to the state vector, $x : \mathbb{R} \rightarrow \mathcal{X}$, but are assumed to be linear with respect to the control vector, $u : \mathbb{R} \rightarrow \mathcal{U}$. The state will sometimes be denoted as $t \mapsto x(t; x(t_i), u(\cdot))$ when we want to make explicit the dependence on the initial state (and time), and corresponding control signal.

The SAC method uses objectives of the form

$$J(x(\cdot)) = \int_{t_i}^{t_i+T} l(t, x(t)) dt + m(t_i + T, x(t_i + T)), \quad (2)$$

with incremental cost $l(t, x(t))$, terminal cost $m(t_i + T, x(t_i + T))$ and time horizon T . Although (2) lacks a norm on control effort, this norm is included in the following steps as shown in (6). The SAC method calculates controls that optimally *improve* (not optimize) the objective (2) even if the objective is non-convex or unbounded. The following definition is necessary before introducing the open-loop problem solved in SAC.

Definition 1: An action A is defined by the triplet consisting of a constant control value, $u_A \in \mathcal{U}$, application duration, $\lambda_A \in \mathbb{R}^+$ and application time, $\tau_A \in \mathbb{R}$, such that $A := \{u_A, \lambda_A, \tau_A\}$.

The open-loop problem \mathcal{P} that is solved in each iteration of the receding horizon strategy can now be defined as follows

$$\mathcal{P}(t_i, x_i) : \quad (3)$$

Find action A such that

$$J(x(t; x_i, u_i^*(\cdot))) < J(x(t; x_i, u_i^{nom}(\cdot))) \quad (4)$$

subject to

$$u_i^*(t) = \begin{cases} u_A & \tau_A \leq t \leq \tau_A + \lambda_A \\ u_i^{nom}(t) & \text{else} \end{cases},$$

and (1) with $t \in [t_i, t_i + T]$ and $x(t_i) = x_i$.

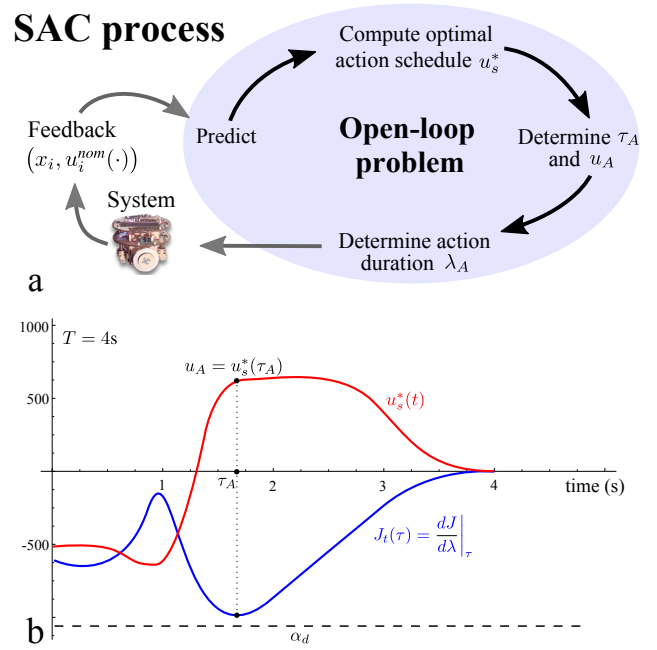


Fig. 1. (a) An overview of the SAC control process. One major advantage of SAC compared to other MPC approaches is that the open-loop problem in SAC—see (3)—can be solved without employing nonlinear programming solvers [17]. In order to solve (3), the SAC method follows four steps as illustrated above. (b) A sample, one-dimensional ($m = 1$) action schedule $u_s^*(t)$ and the corresponding cost (8) used to calculate the application time τ_A for the hypothetical time window $[0, 4]$ s. Every point on $u_s^*(t)$ corresponds to an action of infinitesimal duration, with value and application time as determined by the curve. Because $u_s^*(t)$ is calculated from minimization of (6), all actions in $u_s^*(t)$ would optimally improve (2) relative to only applying nominal control if applied for $\lambda \rightarrow 0^+$. Each point on the mode insertion gradient curve approximates the change in cost (2) achievable by infinitesimal application of the corresponding action in $u_s^*(t)$. By choosing the application time that minimizes this curve, we pick the action that generates the greatest cost reduction in the current time window.

From now on, subscript i will denote the i -th time step, starting from $i = 0$.

Definition 2: Nominal control $u^{nom} : \mathbb{R} \rightarrow \mathcal{U}$ is continuous in t , and is used as a basis when calculating the open-loop solution. It is often $u^{nom}(\cdot) \equiv 0$ so that problem \mathcal{P} outputs the optimal action relative to doing nothing (allowing the system to drift for a horizon into the future). Alternatively, $u^{nom}(\cdot)$ may be an optimized feedforward or state-feedback controller, indicating potential shared control implementations. To evaluate SAC as a standalone controller, the nominal control throughout this paper is considered constant or zero. The system trajectory corresponding to application of nominal control will be denoted as $x(t; x(t_i), u^{nom}(\cdot))$ or $x^{nom}(\cdot)$ for brevity.

The solution $u_i^*(t)$ of problem \mathcal{P} includes a switch to the calculated action A for $t \in [\tau_A, \tau_A + \lambda_A] \subseteq [t_i, t_i + T]$ (Fig. 2a) and thus, it is piecewise continuous in t . When applied to (1), $u_i^*(t)$ generates a (discontinuous) switch of the same duration λ_A in the dynamics from $f(t, x, u_i^{nom}(\cdot))$ to $f(t, x, u_A)$. The condition in (4) highlights a key feature of SAC, e.g., rather than optimize the objective (2), SAC actions *improve* (2) relative to only applying nominal control. As the receding horizon strategy progresses, $\mathcal{P}(t_i, x_i)$ is solved

from the current time t_i using the measured state x_i , and the calculated control $u_i^*(t)$ —corresponding to $x_i^*(t)$ —is applied for t_s seconds (sampling time) to drive the system from x_i to $x(t_i + t_s; t_i, x_i, u_i^*(\cdot))$. The process is then repeated at the next sampling instance, i.e., $t_i \leftarrow t_i + t_s$ and $i \leftarrow i + 1$. The final result of the closed-loop receding horizon strategy is a sequence of actions, forming a piecewise continuous control signal $\bar{u}_{cl}(t)$ with state response $\bar{x}_{cl}(t)$.

Unlike alternative MPC methods [1], [2], [15], [16], [24], [25], the open-loop problem in SAC can be solved in closed form without employing nonlinear programming solvers (see [17] and [18] for a review). In order to solve (3), the SAC method follows four steps as illustrated in Fig. 1a. Beginning with prediction, the steps of the SAC process are described in the subsequent subsections.

Steps for solving the open-loop problem \mathcal{P}

The solution process assumes the following:

Assumption 1: The control objective is to steer the state to the origin. This is not a restrictive assumption as most control scenarios (including trajectory tracking) can satisfy it with a change of coordinates.

Assumption 2: The dynamics f in (1) are continuous in u , piecewise continuous in t and continuously differentiable in x . Also, f is compact, and thus bounded, on compact sets \mathcal{X} and \mathcal{U} . Finally, the system is assumed to be controllable and $f(\cdot, 0, 0) = 0$.

Assumption 3: The terminal cost m is positive definite, continuously differentiable in x and continuous in t . The incremental cost $l(t, x)$ is continuous in t and continuously differentiable in x and $l(\cdot, 0) = 0$. Also, there exists a continuous positive definite, radially unbounded function $M : \mathcal{X} \rightarrow \mathbb{R}^+$ such that $l(t, x) \geq M(x) \forall t$.

Assumption 4: The trajectory $x_i^*(t) \in \mathcal{X}$ corresponding to the solution $u_i^*(t) \in \mathcal{U}$ of $\mathcal{P}(t_i, x_i)$ is absolutely continuous, and thus bounded, in $[t_i, t_i + T]$.

The open-loop problem \mathcal{P} can then be solved by following the four steps presented below.

1) *Predict:* The SAC process begins by predicting the evolution of a system model from current state feedback. In this step, the system (1) is simulated from the current state x_i and time t_i , under $u_i^{nom}(t)$ for $t \in [t_i, t_i + T]$. The sensitivity of (2) to the state variations along the predicted trajectory $x_i^{nom}(\cdot)$ is provided by an adjoint variable, $\rho_i : [t_i, t_i + T] \rightarrow \mathbb{R}^n$, also simulated during the prediction step. The adjoint satisfies

$$\dot{\rho}_i = -D_2 l(t, x_i^{nom})^T - D_2 f(t, x_i^{nom}, u_i^{nom})^T \rho_i \quad (5)$$

$$\text{subject to } \rho_i(t_i + T) = D_2 m(t_i + T, x_i^{nom}(t_i + T))^T.$$

where D_i denotes derivative with respect to i -th argument.

In short, the prediction phase completes upon simulation of the state using (1) and the adjoint system (5) under $u_i^{nom}(\cdot)$ control. The resulting trajectories $x_i^{nom}(\cdot)$, $\rho_i(\cdot)$ will be used in the remaining three steps of the solution process.

2) *Compute optimal action schedule $u_s^*(\cdot)$:* In this step, we compute a schedule $u_s^* : [t_i, t_i + T] \rightarrow \mathbb{R}^m$ which contains candidates for the next action, for all possible times that action may be taken. Specifically, $u_s^*(\cdot)$ contains candidate action

values and their corresponding application times, assuming $\lambda \rightarrow 0^+$ for all (see Fig. 1b for a sample one-dimensional action schedule). The final u_A and τ_A will be selected from these candidates in step three of the solution process such that $u_A = u_s^*(\tau_A)$, while a finite duration λ_A will be selected in step four. The optimal action schedule $u_s^*(\cdot)$ is calculated by minimizing

$$J_{u_s} = \frac{1}{2} \int_{t_i}^{t_i+T} \left[\frac{dJ}{d\lambda}(t) - \alpha_d \right]^2 + \|u_s(t) - u_i^{nom}(t)\|_{\mathbb{R}}^2 dt, \quad (6)$$

$\frac{dJ}{d\lambda}(t) = \rho_i(t)^T [f(t, x_i^{nom}(t), u_s(t)) - f(t, x_i^{nom}(t), u_i^{nom}(t))]$ where the quantity $\frac{dJ}{d\lambda}(\cdot)$, called mode insertion gradient (see [8], [9]), denotes the rate of change of the cost with respect to a switch of infinitesimal duration λ in the dynamics of the system. In this case, $\frac{dJ}{d\lambda}(\cdot)$ shows how the cost will change if we introduce infinitesimal switches from $f(t, x_i^{nom}(t), u_i^{nom}(t))$ to $f(t, x_i^{nom}(t), u_s(t))$ at any time $t \in [t_i, t_i + T]$. Intuitively, minimization of (6) is driving $\frac{dJ}{d\lambda}(\cdot)$ to a negative value $\alpha_d \in \mathbb{R}^-$. As a result, each switch/action value in $u_s^*(t)$ is the single choice that optimally improves (2) (relative to only applying nominal control) if applied for $\lambda \rightarrow 0^+$ at its corresponding application time. The design parameter α_d determines how much the cost is improved by each infinitesimal action in the schedule $u_s^*(t)$.

Based on the simulation of the dynamics (1), and (5) completed in the prediction step (Section II-1), minimization of (6) leads to the following closed-form expression for the optimal action schedule:

$$u_s^*(t) = u_i^{nom}(t) + (\Lambda + R^T)^{-1} h(t, x_i^{nom}(t))^T \rho_i(t) \alpha_d, \quad (7)$$

where $\Lambda \triangleq h(t, x_i^{nom}(t))^T \rho_i(t) \rho_i(t)^T h(t, x_i^{nom}(t))$. The infinitesimal action schedule can then be directly saturated to satisfy any min/max control constraints of the form $u_{min,k} \leq 0 \leq u_{max,k} \forall k \in \{1, \dots, m\}$ such that $u_s^* \in \mathcal{U}$ without additional computational overhead (see [3] for proof).

3) *Determine application time τ_A (and thus u_A value):* The SAC method optimizes a decision variable not normally included in control calculations—the choice of *when* to act. As opposed to always acting at the current time, i.e., $\tau_A = t_i$, the application time of an action is allowed to take values in $\tau_A \in [t_i, t_i + T]$.

Recall that the curve $u_s^*(\cdot)$ provides the values and application times of possible infinitesimal actions that SAC could take at different times to optimally improve (2) from that time. In this step the SAC method chooses one of these actions to apply, i.e., chooses the application time τ_A and thus an action value u_A such that $u_A = u_s^*(\tau_A)$. To do that, $u_s^*(\cdot)$ is searched for a time τ_A that minimizes

$$J_t(\tau) = \frac{dJ}{d\lambda} \Big|_{\tau}, \quad (8)$$

$$\frac{dJ}{d\lambda} \Big|_{\tau} = \rho_i(\tau)^T [f(\tau, x_i^{nom}(\tau), u_s^*(\tau)) - f(\tau, x_i^{nom}(\tau), u_i^{nom}(\tau))] \\ \text{subject to } \tau \in [t_i, t_i + T].$$

Notice that the cost (8) is actually the mode insertion gradient evaluated at the optimal schedule $u_s^*(\cdot)$. Thus, minimization of (8) is equivalent to selecting the infinitesimal action from $u_s^*(\cdot)$ that will generate the greatest cost reduction relative to only applying nominal control. For a sample $J_t(\tau)$ see Fig. 1b. Finally, the selected application time must satisfy the following assumption:

Assumption 5: The selected application time τ_A is such that $h(\tau_A, x(\tau_A))^T \rho(\tau_A) \neq 0 \in \mathbb{R}^{m \times 1}$.

The role of this assumption is explained in the following step.

4) *Determine control duration λ_A :* So far we have selected τ_A and u_A (see Definition 1) from a schedule of possible *infinitesimal* actions, $u_s^*(\cdot)$. The final step in synthesizing a SAC action is to choose how long to act, i.e., a finite control duration λ_A , such that the condition in (4) is satisfied. The following assumption and proposition will facilitate the analysis in the sequel.

Proposition 1: For a choice of $\alpha_d < 0$ in (6), an infinitesimal control action $u_s^*(\tau)$ that is selected according to Assumption 5 will result in $\frac{dJ}{d\lambda}(\cdot) < 0$ and thus, it can reduce (2).

Proof: See Appendix. ■

Proposition 1 proves that if $\alpha_d < 0$, then the infinitesimal $u_s^*(\tau_A)$ will lead to a negative $\frac{dJ}{d\lambda}(\cdot)$. From [9], [29], there is a non-zero neighborhood around $\lambda \rightarrow 0^+$ where the mode insertion gradient models the change in cost in (4) to first order, and thus, a finite duration λ_A exists that satisfies (4). In particular, for *finite* durations λ in this neighborhood we can write

$$\begin{aligned} J(x(t; x_i, u_i^*(\cdot))) - J(x(t; x_i, u_i^{nom}(\cdot))) \\ = \Delta J \approx \left. \frac{dJ}{d\lambda} \right|_{\tau_A} \lambda. \end{aligned} \quad (9)$$

Thus, from Proposition 1 and (9), the condition in (4) is feasible. Then, a finite action duration λ_A can be calculated by employing a *line search* process [29]. Starting with an initial duration λ_0 we simulate the effect of the control action using (1) and (2). If the simulated action satisfies (4), the duration is selected. If this is not the case, the duration is reduced and the process is repeated. By continuity, the final duration will produce a change in cost within tolerance of that predicted from (9).

After computing the duration λ_A , the control action A is fully specified (it has a value, an application time and a duration) and thus the solution $u_i^*(t)$ of problem \mathcal{P} has been determined. By iterating on this process (Section II-1 until Section II-4), SAC synthesizes piecewise continuous, constrained control laws for nonlinear systems. For more information about SAC, the reader is encouraged to consult [3], [5], [6].

III. FROM SAC TO ISAC

In this section we will describe the differences between SAC and iSAC, and the extensions introduced in iSAC that enable us to show closed-loop asymptotic stability. There are two main features introduced in iSAC; the first one is that the open-loop solution $u_i^*(t)$ is iteratively updated as the time horizon progresses. The second is that the cost is decreasing

from time step to time step. These two features are illustrated in Fig. 2. For convenience, the corresponding SAC behavior is also shown in the same figure.

A. Iterative update of $u_i^*(t)$

Figure 2a illustrates how the iSAC method stores actions calculated at previous time steps and modifies $u_i^*(t)$ by introducing a new action at every time step. On the contrary, SAC only keeps a single action in $u_i^*(t)$, regardless of whether that action lies in the current application window $[t_i, t_i + t_s]$ or not. In light of this, the following definition is necessary to distinguish between the nominal control (zero or constant in this paper) from Def. 2, and the default control that iSAC uses to calculate actions.

Definition 3: Default control $u_i^{def} : [t_i, t_i + T] \rightarrow \mathcal{U}$, is piecewise continuous in t and is defined as

$$u_i^{def}(t) = \begin{cases} u_{i-1}^*(t) & t_i \leq t \leq t_i + T - t_s \\ u_i^{nom}(t) & t_i + T - t_s < t \leq t_i + T \end{cases}, \quad (10)$$

with $u_0^{def}(\cdot) \equiv u_0^{nom}(\cdot)$. In expression (10), $u_{i-1}^* : [t_{i-1}, t_{i-1} + T] \rightarrow \mathcal{U}$ is the output of $\mathcal{P}(t_{i-1}, x_{i-1})$ from the previous time step $i-1$, and $t_s = t_i - t_{i-1}$ is the sampling period.

As a result, actions calculated in iSAC improve the cost (2) with respect to applying default control $u_i^{def}(\cdot)$. In fact, as shown in Fig. 2, iSAC actions establish a decreasing behavior¹ for (2) from time step to time step. As will be explained in the following paragraph, this is accomplished by applying a contractive constraint on the cost, while the iterative nature of $u_i^*(t)$ has a major role in the formulation of this constraint.

B. Contractive constraint on cost

In SAC, each action is calculated such that (4) is satisfied. However, in general, this does not ensure that the cost will follow a decreasing trend. The explanation lies in the fact that (4) and (9) only guarantee that the cost will improve relative to a nominal input in the current time step, but not necessarily across time steps. As an example, Fig. 2c shows a general scenario where condition (4) is met. In this case, the system will most likely not meet the desired specifications encoded in the cost, since the latter is not decreasing with time. Thus, even if (4) is satisfied, closed-loop stability cannot be guaranteed.

A solution to this problem is to modify condition (4) such that (2) sufficiently “contracts” between time steps. Specifically, instead of improving the cost relative to only applying nominal control, we could apply

$$J(x_i^*(\cdot)) - J(x_{i-1}^*(\cdot)) \leq - \int_{t_{i-1}}^{t_i} l(t, x_{i-1}^*(t)) dt, \quad (11)$$

as a contractive constraint. Contractive constraints [20]–[23] have been widely used in the MPC literature to show closed-loop stability as an alternative to methods relying on a terminal region constraint [1], [2], [15], [16], [24], [25]. Conditions similar to (11) also appear in terminal region methods, either

¹In Section V we make use of this fact and consider (2) as a candidate Lyapunov function.

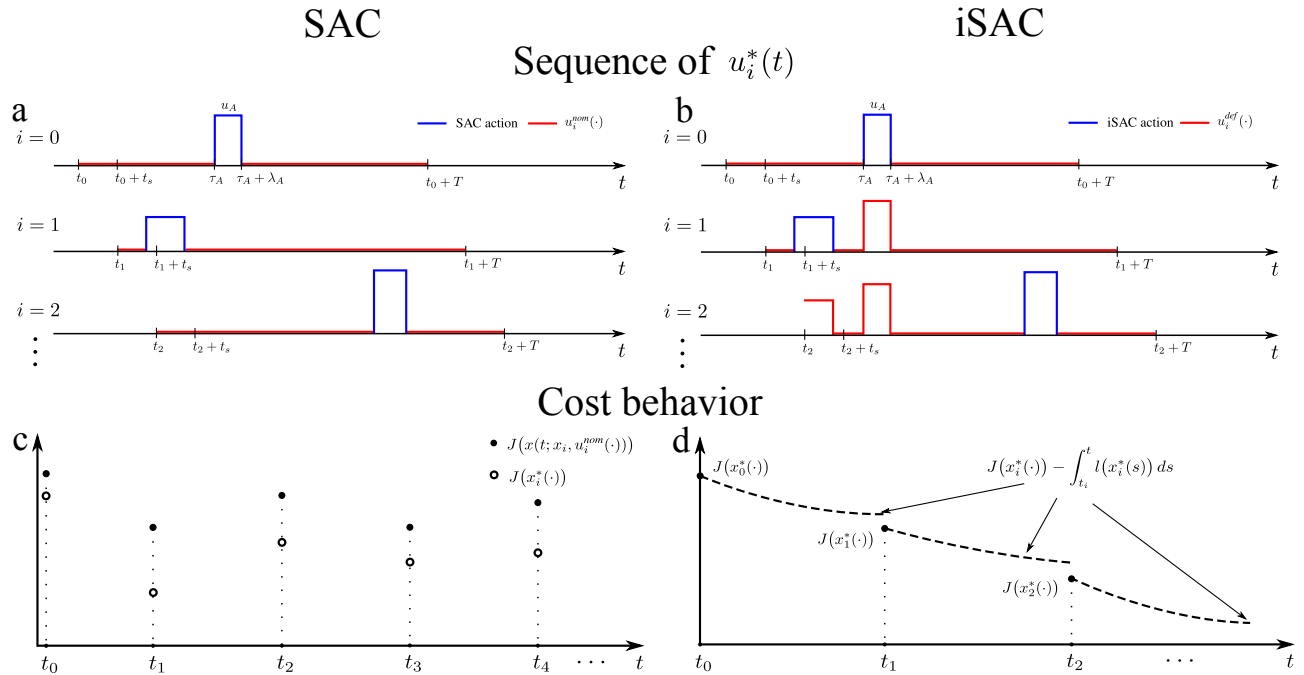


Fig. 2. Summary of differences between SAC and iSAC. Panels a and b underline the fact that iSAC stores previously calculated actions and iteratively improves the open-loop solution $u_i^*(t)$ (assuming zero nominal control). As a result, a contractive constraint can be employed in iSAC to establish a sufficient decrease condition for the cost as shown in panel d. Panel c shows the corresponding graph in SAC as dictated by (4). It can be observed that the behavior of the cost in SAC heavily depends on how $J(x(t); x_i, u_i^{nom}(\cdot))$ evolves in time. Since $J(x(t); x_i, u_i^{nom}(\cdot))$ can take arbitrary values, there is no guarantee that the closed-loop system will be stable, even though condition (4) may be repeatedly satisfied.

in continuous or in discrete time, as an intermediate step used to prove closed-loop stability.

One problem that arises with the new condition (11) in SAC is that the quantities $J(x_i^*(\cdot))$ and $J(x_{i-1}^*(\cdot))$ cannot be related through the mode insertion gradient, unlike, e.g., $J(x_i^*(\cdot))$ and $J(x_i^{nom}(\cdot))$ that appear in the original condition (4) and can be related through (9). Thus, Proposition 1 can no longer be used to ensure that (11) is feasible in SAC. On the contrary, in iSAC, we can use the iterative nature of $u_i^*(t)$ to ensure feasibility of (11) through Proposition 1 for sufficiently small sampling time t_s . In particular, in iSAC we can transform (11) to an equation that is similar to (4) and only includes the terms $J(x_i^*(\cdot))$ and $J(x_i^{def}(\cdot))$, which can be related through the mode insertion gradient $\frac{dJ}{d\lambda}$, as in (9). By definition, (10) leads to $x_i^{def}(t) \equiv x_{i-1}^*(t)$ for $t \in [t_i, t_{i-1} + T]$ (see also Fig. 2b). We can then calculate

$$\begin{aligned}
J(x_{i-1}^*(\cdot)) - J(x_i^{def}(\cdot)) &= \int_{t_{i-1}}^{t_i} l(t, x_{i-1}^*(\cdot)) dt \\
&+ m(t_{i-1} + T, x_{i-1}^*(t_{i-1} + T)) - \int_{t_{i-1} + T}^{t_i + T} l(t, x_i^{def}(t)) dt \\
&- m(t_i + T, x_i^{def}(t_i + T)) = B.
\end{aligned} \tag{12}$$

Combining (12) and (11) and using (10) we get

$$J(x_i^*(\cdot)) - J(x_i^{def}(\cdot)) \leq B - \int_{t_{i-1}}^{t_i} l(t, x_{i-1}^*(t)) dt \tag{13}$$

or equivalently

$$\begin{aligned}
J(x_i^*(\cdot)) - J(x_i^{def}(\cdot)) &\leq m(t_{i-1} + T, x_{i-1}^*(t_{i-1} + T)) \\
&- \int_{t_{i-1} + T}^{t_i + T} l(t, x_i^{def}(t)) dt - m(t_i + T, x_i^{def}(t_i + T)) = C.
\end{aligned} \tag{14}$$

Thus, we were able to transform (11) into a sufficient decrease condition similar to (4), i.e., a condition that involves terms depending on the default and calculated open-loop trajectories. The following Proposition can be used in conjunction with Proposition 1 to ensure that the new condition (14) is (recursively) feasible.

Proposition 2: For $\alpha_d < 0$ there exists a sufficiently small sampling time t_s such that the sufficient cost decrease required by the contractive constraint (14) is attainable.

Proof: See Appendix. ■

The contractive constraint (14) can be applied in the line search process that determines the duration of an action (Section II-4) in lieu of (4). By applying this condition at each time step, we can guarantee that the cost value will decrease across time steps, and using this fact we can prove closed-loop stability (Section V).

Open-loop problem in iSAC

Based on the above, the open-loop problem solved by iSAC at each time step is given by

$$\mathcal{B}(t_i, x_i) : \quad (15)$$

Find action A such that (14) is satisfied subject to

$$u_i^*(t) = \begin{cases} u_A & \tau_A \leq t \leq \tau_A + \lambda_A, \\ u_i^{def}(t) & \text{else} \end{cases},$$

and (1) with $t \in [t_i, t_i + T]$ and $x(t_i) = x_i$.

Similar to the open-loop problem \mathcal{P} in SAC, the solution $u_i^*(t)$ of \mathcal{B} can be obtained in closed form by following the four steps in Fig. 1a and Section II, without relying on nonlinear programming solvers. The only difference is that in the solution process of \mathcal{B} , the superscripts *nom* are replaced by *def*. Our final result in this section utilizes the aforementioned assumptions and propositions to ensure that the open-loop problem \mathcal{B} has a solution:

Proposition 3 (Existence of solution to \mathcal{B}): For sufficiently small sampling time t_s , the solution $u_i^*(t)$ to the open-loop problem $\mathcal{B}(t_i, x_i)$ exists for any x_i, t_i .

Proof: See Appendix. ■

IV. LOCAL STABILITY ANALYSIS

In this section we present local stability properties for iSAC, assuming quadratic cost (2). Note that this section is not a prerequisite for global stability analysis, which is presented in the following section. In other words, the results presented here are independent and not required to show global stability. The purpose of the section is to highlight local properties of iSAC that facilitate analysis of linear or linearized systems with quadratic cost.

Theorem 1: Assume that the cost (2) is quadratic

$$J(x(\cdot)) = \frac{1}{2} \int_{t_i}^{t_i+T} \|x(t)\|_Q^2 dt + \frac{1}{2} \|x(t_i + T)\|_{P_1}^2 \quad (16)$$

with symmetric, positive semidefinite Q and P_1 . Then, there exist neighborhoods around the equilibrium and final time, $\mathcal{N}(0)$ and $\mathcal{N}(t_i + T)$ respectively, where (7) simplifies to a linear feedback regulator

$$u_s^*(t) = u^{def}(t) + R^{-1}h(t, 0)^T P(t)x(t)\alpha_d, \quad (17)$$

where $P(t)$ is a $n \times n$ matrix that can be calculated from a Riccati equation.

Proof: The proof can be found in the Appendix. ■

Similarly, one can also show Theorem 1 yields a feedback expression in error coordinates for trajectory tracking problems. The following Corollary shows that the quantity $P(t)$ is calculated by solving the same Riccati differential equation that runs backward in time as in the LQR solution [30].

Corollary 1: The optimal action schedule (17) is locally equivalent to the LQR solution $u_{LQR}^*(t) = u_d(t) - K(t)x(t)$ for the following problem

$$J(x(\cdot), u(\cdot)) = \frac{1}{2} \int_{t_i}^{t_i+T} \|x(t)\|_Q^2 + \|u(t) - u_d(t)\|_{\bar{R}}^2 dt + \frac{1}{2} \|x(t_i + T)\|_{P_1}^2, \quad (18)$$

with $u_d(t) = u_i^{def}(t)$, $K(t) = \bar{R}^{-1}h(t, 0)^T P(t)$ and $\bar{R} = R \frac{1}{|\alpha_d|}$. Thus, (17) also minimizes (18).

Proof: The proof can be found in the Appendix. ■

Remark 1: Notice, that since the schedule (17) minimizes (18), if (17) is directly applied to the system as a sequence of optimal infinitesimal actions, i.e., set $u_i^*(t) = u_s^*(t)$, the closed loop system will inherit the properties of an LQR response. In this case, one can use LQR theory [30] to prove existence of a Lyapunov function that guarantees local closed-loop stability for iSAC. Additionally, there is no need to search for an application time and duration in this case as (17) models a sequence of optimal infinitesimal actions. Instead, one has to calculate $P(t)$ by solving a Riccati differential equation. Naturally, the contractive constraint (14) does not apply here either since (17) minimizes (18) as opposed to “contracting” it. Switching to this linear, duration-independent equivalent to iSAC near the equilibrium can be useful, e.g., in cases where the duration λ required to satisfy (14) is shorter than the minimum tolerance imposed by numerical software or hardware. Sums-of-Squares (SOS) and the S-procedure [28], [31] can pre-compute regions of attraction for (17) and so can determine when the switch to this duration-independent version of iSAC should occur.

V. GLOBAL STABILITY ANALYSIS

In this section, we provide global stability results for iSAC and discuss how iSAC can be used under a special case of disturbances and to track infeasible trajectories, i.e., trajectories that do not satisfy the dynamics (1).

A. Nominal Case

In this section we are considering nominal stability (not robust stability), i.e., the trajectories of the plant are assumed to coincide with the trajectories predicted by the model. Recall that the iSAC method does not optimize the cost (2) but rather improves it across time steps. It is often the case that minimization of the objective function in the open-loop optimization problem is not required to achieve stability of a model predictive controller. Sometimes, a decrease in the cost at every iteration is sufficient to guarantee stability [19], [24], [32], [33]. In view of these observations, we now present the following Theorem:

Theorem 2 (Asymptotic Stability): For sufficiently small sampling time t_s , the closed-loop system resulting from the receding horizon strategy is asymptotically stable in the sense that $\|\bar{x}_{cl}(t)\| \rightarrow 0$ as $t \rightarrow \infty$.

Proof: See Appendix. ■

Theorem 2 does not imply Lyapunov stability, but rather establishes the usual notion of attractiveness. The importance

of t_s in establishing closed-loop stability is also underlined; if the assumptions of the Theorem hold, by appropriate selection of t_s , we can ensure that the contractive constraint is feasible, and thus, asymptotic stability.

Cost as a Lyapunov function

Now that asymptotic stability has been proved, we will show that, under certain assumptions, the objective (2) is a Lyapunov function for the closed-loop system that decreases at intervals of prediction horizons.

Theorem 3 (Objective Function as a Lyapunov Function): Assume that f is bounded such that $\|f\| \leq \xi\|x\|$ for some finite constant $\xi > 0$, and that the cost (2) is of the form in (16), rewritten as

$$J(t_i, x(\cdot)) = \frac{1}{2} \int_{t_i}^{t_i+T} \|x(t)\|_Q^2 dt + \frac{1}{2} \|x(t_i + T)\|_{P_1}^2. \quad (19)$$

Then, for sufficiently small t_s , (19) is a discrete Lyapunov function for the closed-loop system.

Proof: The proof can be found in the Appendix. ■

Corollary 2: For the assumptions in Theorem 3, if $t_s \rightarrow 0$, then (19) is a continuous Lyapunov function for the closed-loop system, implying asymptotic stability in the classical sense, i.e., Lyapunov stability and attractiveness.

Proof: The proof can be found in the Appendix. ■

B. Asymptotically Decaying Disturbances

This section provides stability results for a special case of disturbances. Here, we assume that the plant is described by the following differential equations (instead of (1))

$$\dot{x}^p = f(t, x^p, u) + \eta(t) = g(t, x^p) + h(t, x^p)u + \eta(t), \quad (20)$$

where $x^p(t)$ represents the trajectory resulting from applying control u and the effect of an unknown additive disturbance $\eta(t)$. The disturbance is bounded and asymptotically decaying; in particular, we assume the following:

Assumption 6: The disturbance $\eta(t)$ is bounded, i.e., $\|\eta_i(t)\| \leq \delta_i < \infty$ for all $t \in [t_i, t_i + T]$ and all $i > 0$. Furthermore, $\eta(t)$ is asymptotically decaying, i.e., $\delta_i \rightarrow 0$ as $i \rightarrow \infty$.

Closed-loop stability under $\eta(t)$ is provided by the following theorem.

Theorem 4 (Asymptotic Stability under Asymptotically Decaying Disturbances): For sufficiently small t_s , the closed-loop system under the disturbance $\eta(t)$ is asymptotically stable in the sense that $\|\bar{x}_{cl}(t)\| \rightarrow 0$ as $t \rightarrow \infty$.

Proof: The proof can be found in the Appendix. ■

As seen in the proof, for as long as the disturbance is acting, the cost is allowed to increase, but the amount of increase is kept sufficiently bounded by iSAC. As time progresses, the disturbance attenuates and the contractive constraint becomes feasible again.

C. Remarks on tracking of non-admissible trajectories

When designing a control task, there are often cases where asymptotic stability cannot be achieved. For example, in trajectory tracking, it is not always easy to identify trajectories that satisfy the dynamics of the system of interest, unless there is special structure one can exploit, e.g., differential flatness [34]. The control goal in this case is to ensure that the deviation of the generated (feasible) open-loop trajectory from the desired one is bounded, i.e., $\|x_i^*(t) - x^d(t)\| \leq \Delta_i$ for $t \in [t_i, t_i + T]$ with $\Delta_i \in [0, \infty)$. Unlike the decaying disturbance case, the upper bound Δ_i does not have any structure that can be exploited. Thus there is no guarantee that $\bar{x}_{cl}(t) \rightarrow x^d(t)$ as $t \rightarrow \infty$.

Similar to the disturbance case, these observations stem from the fact that the feasibility of the contractive constraint, and thus of the open-loop problem, can no longer be ensured when tracking an infeasible trajectory. In these cases, in order to keep the error from the desired behavior bounded, one can set α_d to more negative values such that the control generation is more “aggressive”. By doing so, the selected action leads to smaller cost increase (at a higher control cost) and iSAC generates a feasible response that is close to the desired infeasible objective, as shown in the simulations that follow.

VI. SIMULATION RESULTS

In this section we demonstrate the flexibility and versatility of iSAC by applying the theoretical results presented in the previous sections to a variety of challenging systems. It is important to note that unlike other system-specific methods mentioned in this section, our approach requires no modifications across examples; only the dynamic model that is being controlled and the relevant parameters of iSAC are changed for each example. Also, all the examples that will be presented here run much faster than real time, indicating applicability to hardware implementations. For better visualization of the results, the reader is encouraged to visit <https://vimeo.com/219702474>.

The control objective is encoded using a quadratic cost as in (16) throughout this section (Assumption 3 is thus satisfied). To demonstrate Theorem 1, the control generation in the first example switches to the state feedback version of iSAC in (17) near the equilibrium. The remaining examples use iSAC without switching to state feedback and also include cases of infeasible tracking tasks and disturbance rejection. The sampling time t_s for each example was appropriately selected to achieve closed-loop stability (Theorems 2 and 4) when feasible. Finally, a summary of iSAC-specific parameters for each example is given in Table I, while Fig. 3 shows the resulting trajectories for the examples described below. Our method is shown to be successful in these scenarios, leading to asymptotic stabilization when feasible.

A. Acrobot

The dynamics of the system and the parameters used in the simulations that follow can be found in [35]. The state vector for the system includes the two joint angles and their

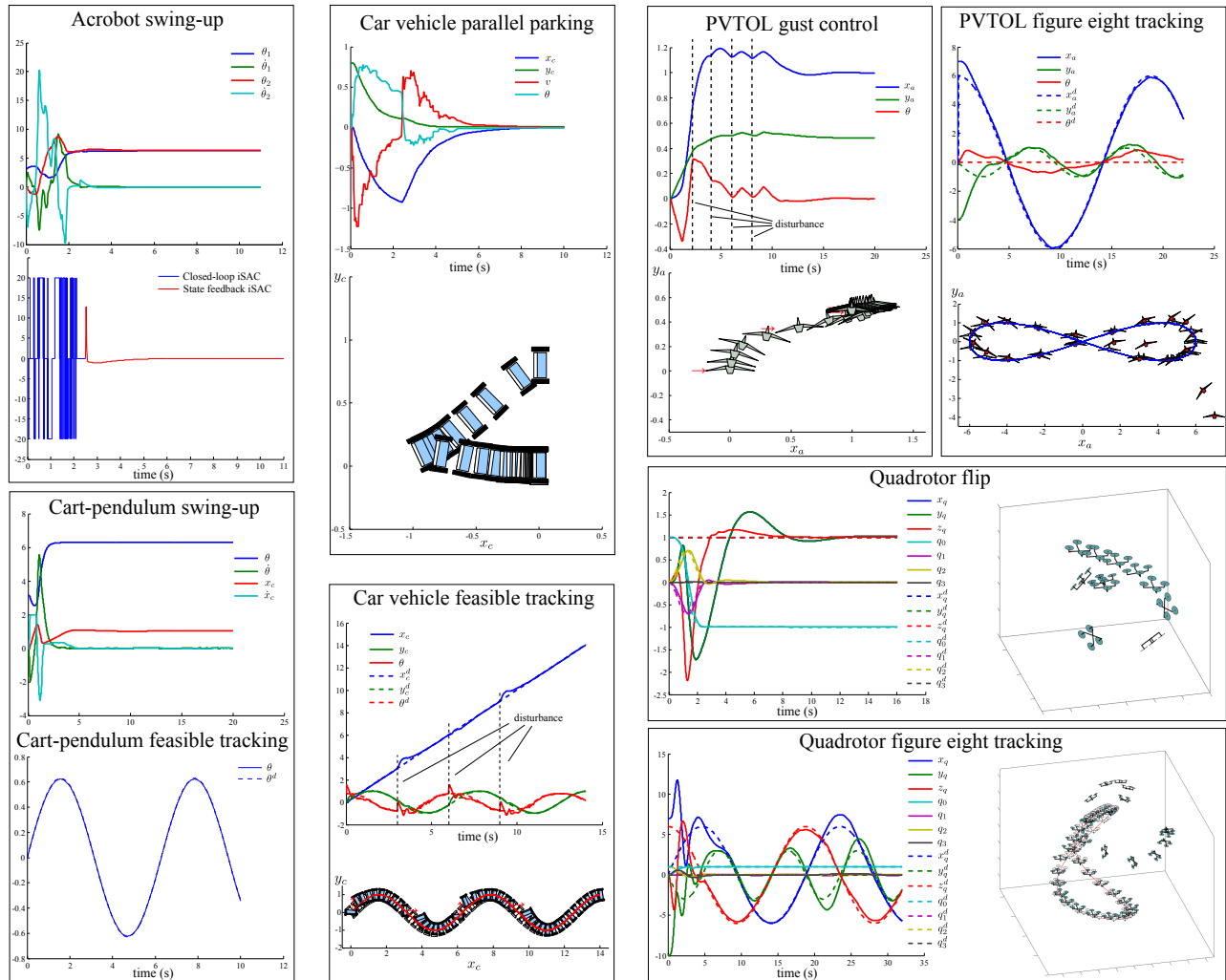


Fig. 3. Summary of results; the iSAC method stabilizes the presented systems successfully. An example of a closed-loop control generated by iSAC is provided in the acrobot panel, that also illustrates the state feedback version of the method. The rest of the examples use only iSAC, without switching to state feedback. Notice that our method is effective even in the presence of decaying disturbances and, also, keeps the deviation from the desired trajectory bounded in infeasible tracking. For better visualization of the results, the reader is encouraged to visit <https://vimeo.com/219702474>.

corresponding velocities, $[\theta_1, \dot{\theta}_1, \theta_2, \dot{\theta}_2]$. The control input is the torque at the distal joint.

Swing-up problem:

The swing-up problem for the acrobot is a benchmark problem in control theory, where the system starts from the downward equilibrium and the control objective is stabilization at the upward equilibrium [36], [37]. The desired setpoint was thus $x^d = [0, 0, 0, 0]$. To illustrate the duration-independent version of iSAC presented in Section IV, control calculation was based on Theorem 1 after system trajectories were within a pre-defined region of the upward equilibrium, found also in [35]. The cost weights were selected as $Q = \text{Diag}[1000, 0, 500, 11]$, and $R = 0.01$.

Figures 3 and 4 show the resulting trajectories, closed-loop control and cost generated by iSAC. In contrast with other methods that switch between separately derived controllers

[38], iSAC pumps energy into the system by appropriate excitation, without explicitly encoding this behavior in the objective. Finally, many methods (see [36], [39]) rely on energy-based objectives to solve the swing-up problem for the acrobot and pendulum systems, primarily because state-tracking objectives result in many local minima that prevent convergence to desirable trajectories. It is noteworthy that, like SAC, iSAC is able to invert the system using a state-tracking objective while bypassing local minima. For a detailed comparison between SAC, sequential quadratic programming (SQP) [40], and iLQG [14] the reader is encouraged to consult [3] and [6].

B. Cart-pendulum

The dynamics of the system and the parameters used in the simulations that follow can be found in [35]. The state vector

for the system includes the angle and angular velocity of the pendulum and the horizontal position and velocity of the cart, $[\theta, \dot{\theta}, x_c, \dot{x}_c]$. The control input is the horizontal acceleration of the cart.

1) *Swing-up problem:*

The cart-pendulum inversion is another benchmark problem in control theory [11], [41]. The desired setpoint in this example was $x^d = [0, 0, 1, 0]$. The cost weights were selected as $Q = \text{Diag}[20, 0, 5, 0]$, $P = \text{Diag}[0.1, 0, 5, 0]$ and $R = 0.3$.

2) *Feasible trajectory tracking:*

The desired trajectory in this example was selected as $\theta^d(t) = \frac{\pi}{5} \sin t$ and the cost weights were $Q = \text{Diag}[5000, 0, 0, 0]$, $P = \text{Diag}[500, 0, 0, 0]$, $R = 0.3$. The desired trajectory was feasible given that no constraints were applied to the cart position and velocity. As the plot in Fig. 3 illustrates, tracking was very accurate, making it hard to distinguish between the desired and actual trajectory followed by the system. The corresponding cost is shown in Fig. 4.

As mentioned in the acrobot example, it is worth noting that iSAC is able to solve the inversion problem by pumping energy into the system without relying on separate controllers. In addition, this was accomplished using the state tracking objective, as opposed to the energy objective commonly employed (see for example [11]), bypassing local minima that cause solutions to converge to undesired trajectories in state tracking costs.

C. Car-like vehicle

The dynamics used in the simulations are

$$\dot{x}_c = v \cos \theta, \dot{y}_c = v \sin \theta, \dot{v} = u_1, \dot{\theta} = u_2, \quad (21)$$

where x_c, y_c denote the position of the car, θ is the angle with respect to the horizontal axis and v is the forward velocity. The iSAC method directly controls the acceleration and the angular velocity of the car. This nonholonomic system violates Brockett's necessary condition for smooth or even continuous stabilization [42], which makes the control design problem challenging. Since iSAC automatically generates a discontinuous control law, it is not subject to this condition.

1) *Parallel parking:*

The desired setpoint for the parallel parking problem was $x^d = [0, 0, 0, 0]$ and the cost weights were selected as $Q = \text{Diag}[1, 15, 0.8, 0.8]$, $P = \text{Diag}[0, 25, 0, 0]$ and $R = \text{Diag}[0.1, 0.1]$. As seen in Fig. 3, iSAC successfully drives the system to the origin. The corresponding cost is in Fig. 4.

2) *Feasible trajectory tracking under disturbance:*

The desired trajectory in this example was chosen as $(x_c^d(t), y_c^d(t), \theta^d(t)) = (t, \sin t, \tan^{-1} \cos t)$ and the cost weights were $Q = \text{Diag}[100, 100, 0, 10]$ and $R = \text{Diag}[0.1, 0.1]$. An external disturbance acted instantaneously on the system on four occasions, i.e., at $t = 0, 3, 6,$ and 9 s, each perturbing v and θ by $+3$ m/s and $+\frac{\pi}{4}$ rad respectively. The effect of the disturbance can be seen in the corresponding plot in Fig. 3; while the car was tracking the desired trajectory, the disturbance pushed the

system away every time it acted on the system. After the last disturbance attenuated the control successfully steered the system back to the desired trajectory. The corresponding cost is in Fig. 4.

A great deal of work on nonholonomically constrained car-like models was completed in the 1990s using nonsmooth or time-varying control laws aiming to overcome stabilizability limitations traditional techniques [43]–[45]. Alternative methods that achieve similar results in the parking problem include fuzzy controllers as in [46] and dynamic feedback linearization [47]. Other approaches that intrinsically lead to piecewise continuous controls like iSAC include MPC as in [20], [48]. The former paper utilizes contractive constraints on the state and, although successful, results in a straight line path between the initial and final configuration. The latter uses the aforementioned terminal region constraints to achieve stabilization, but, unlike iSAC, the method is designed specifically for the car-vehicle system.

D. Planar vertical take-off landing (PVTOL) aircraft

The dynamic model used in this example is given in [49]. The state vector for the system includes the position of the aircraft, the angle with respect to the horizontal axis and the corresponding velocities, i.e., $[x_a, \dot{x}_a, y_a, \dot{y}_a, \theta, \dot{\theta}]$. The control inputs of the system are thrust (directed out the bottom of the aircraft) and the rolling moment. Lastly, the parameter ϵ that appears in the model was chosen as $\epsilon = 0.3$ and $\epsilon = 0.05$ for the two examples that follow respectively. Note that ϵ couples the rolling moment input with the lateral acceleration of the aircraft; a positive value $\epsilon > 0$ means that applying a (positive) moment to roll left produces an acceleration to the right, making this a non-minimum phase system.

1) *Gust control:*

The desired setpoint in this example was $x^d = [1, 0, 0.5, 0, 0, 0]$ and the cost weights were selected as $Q = \text{Diag}[15, 3, 15, 3, 3, 0]$ and $R = \text{Diag}[0.1, 0.1]$. An external disturbance acted instantaneously on the system on five occasions, i.e., at $t = 0, 2, 4, 6,$ and 8 s, each perturbing \dot{x}_a by $+0.1$ m/s. The behavior of the system was similar to the one observed in the car-like vehicle; while the control was steering the system towards the desired setpoint, the disturbance drove it away. After the last disturbance decayed and the setpoint was finally reached.

2) *Infeasible figure eight tracking:*

The desired trajectory for this example was a planar figure eight. To track this trajectory, the aircraft needs make aggressive maneuvers; it is therefore impossible to keep the angle zero and track the desired position at the same time which makes the selected trajectory infeasible. Nevertheless, as seen in Fig. 3, iSAC keeps the deviation from the desired trajectories small. The simulation was based on the following weight values: $Q = \text{Diag}[45, 0, 45, 0, 20, 0]$ and $R = \text{Diag}[0.1, 0.1]$.

Figures 3 and 4 show the resulting trajectories and cost. The control of PVTOL aircrafts has been extensively studied using, e.g., gain-scheduling [50], robust control [51] and input-output linearization [49]. In the latter case, the system is decoupled

into simpler subsystems to facilitate the control process (a similar approach is commonly followed in quadrotor control). It is important to note that iSAC is able to successfully control the system using the full dynamics and without relying on separate controllers, unlike, e.g. in [50]. Finally, other techniques that have been applied on VTOL aircrafts and unmanned aerial vehicles (UAVs) in general include sliding mode control and backstepping as explained in the following example [52].

E. Quadrotor

For this example, a quaternion-based model is used because it leads to polynomial, singularity-free dynamical equations for the quadrotor. The model can be found in [53] and the model parameters we used are the same as in [6]. The 13-dimensional state vector consists of the position and velocity of the center of mass, the angular orientation with respect to the inertial frame expressed in quaternions and the angular velocity expressed in the body frame, i.e., $[x_q, \dot{x}_q, y_q, \dot{y}_q, z_q, \dot{z}_q, q_0, q_1, q_2, q_3, p, q, r]$. The control inputs are the squared angular velocities of the rotors, which are converted to upward thrust for each rotor when multiplied by an appropriately dimensioned constant.

1) Flip:

The desired setpoint for the state vector of the translational dynamics was $x^d = [1, 1, 1, 0, 0, 0]$. To perform a flip, the state vector of the rotational dynamics was set to track a flip trajectory for the first portion of the simulation and then move to x^d . For this example, the cost weights were selected as $Q = \text{Diag}[3, 3, 3, 0, 0, 0, 1500, 1500, 1500, 1500, 0, 0, 0]$ and $R = \text{Diag}[0.1, 0.1, 0.1, 0.1]$. Figures 3 and 4 show the resulting trajectories and cost.

2) Infeasible three dimensional figure eight tracking:

In this example, the quadrotor starts in an upside-down position and is set to track a three-dimensional figure eight while keeping the quaternion states at $(q_0^d(t), q_1^d(t), q_2^d(t), q_3^d(t)) = [1, 0, 0, 0]$. In other words, the quadrotor is requested to track an aggressive 3-dimensional maneuver without changing its angular orientation, which is infeasible. The cost weights were selected as $Q = \text{Diag}[10, 10, 10, 0, 0, 0, 1000, 1000, 1000, 1000, 0, 0, 0]$ and $R = \text{Diag}[0.1, 0.1, 0.1, 0.1]$. As seen in Fig. 3, at the beginning of the simulation the quadrotor reverts its angular orientation and then it starts tracking the desired trajectory keeping the error small. This behavior is automatically generated by our method. The corresponding cost is in Fig. 4.

Many different approaches have been followed for control of quadrotors over the years. Linear control methods like LQ and PID synthesis [54], [55] are popular for their simplicity. They utilize decoupled or simplified dynamics and also exploit differential flatness [34] in order to track feasible trajectories. For example, in [56] a quadrotor flip is implemented using a dedicated attitude controller. In contrast iSAC controls the full nonlinear dynamics of the system, and has good performance even with infeasible trajectories as shown in the simulation examples. No additional steps are required to apply our method to this system, unlike, e.g., feedback linearization [57] or

sliding mode control [52]. The simplicity and versatility of iSAC is one of the main points that this section demonstrates; the same exact approach is being applied to a variety of challenging systems in real time, by replacing only the dynamic model that is being controlled.

F. Discussion and Limitations

In general, iSAC performs well in the variety of scenarios presented. When the task is feasible, trajectories are asymptotically stable and the cost is generally decreasing from time step to time step as a result of applying condition (11). Nevertheless, near the desired equilibrium, the behavior of the cost sometimes deviates from what one would expect, even when the control task is feasible. As an example, in the rightmost panel of Fig. 4 the cost is not continuously decreasing near the equilibrium, although, in both of these examples, the system is effectively stabilized when the unexpected behavior occurs (the cost value is less than 0.01). There are three ways to explain this; first, in order to speed up the solution process, the backtracking process (final step in the solution of the open-loop problem) in all examples presented in this paper was stopped after 10 iterations. While for the majority of cases 10 iterations were sufficient to satisfy the sufficient descent condition (11), on some occasions more iterations were necessary. As a result, a small increase was observed in the cost in these time steps. Another factor to consider is numerical tolerance. For example, from the simplified model (9), depending on the value of the mode insertion gradient, it could be the case that the required λ value that achieves the sufficient decrease descent is smaller than the minimum tolerance of numerical integrators. This issue appears often near the equilibrium where systems are more sensitive to the duration of an action. One way to deal with this issue is to switch to the duration-independent, state feedback version of iSAC near the equilibrium (Theorem 1). Finally, from Proposition 2, it is possible that the selected sampling time t_s is not sufficiently small to achieve (recursive) feasibility of (14). In this case, selecting a smaller t_s would likely solve the problem.

VII. CONCLUSION

In this paper we presented iSAC, a model-predictive approach for control of nonlinear systems. As the time horizon progresses, our method sequences together optimal actions and synthesizes piecewise continuous control laws. Some key characteristics of iSAC include: a) there is no need to rely on nonlinear programming solvers since there is an analytic solution to the open-loop problem, b) it is iterative in the sense that at each time step it updates the open-loop solution of the previous time steps by adding one action and c) it utilizes continuous dynamics while incorporating control constraints without additional overhead. Due to these points, iSAC leads to computationally efficient solutions. To establish closed-loop stability, we applied a contractive constraint on the cost. Compared to methods relying on terminal region constraints, the contractive constraint alleviates the need to calculate a terminal region. We also investigated different control scenarios ranging

TABLE I
ISAC PARAMETER VALUES USED IN SIMULATIONS

Example	VI-A	VI-B1	VI-B2	VI-C1	VI-C2	VI-D1	VI-D2	VI-E1	VI-E2
α_d	$-15 J(x_i^{def}(\cdot))$	$-15 J(x_i^{def}(\cdot))$	-1050	$-100 J(x_i^{def}(\cdot))$	$-10 J(x_i^{def}(\cdot))$	$-10 J(x_i^{def}(\cdot))$			-5000
T	1.1s	1.2s	0.05s	1.2s	0.35s	3s	3s	3s	2s
t_s	0.02s	0.01s	0.005s	0.02s	0.03s	0.02s	0.007s	0.02s	0.05s
$[u_{min}, u_{max}]$	$[-20, 20]$ N m	$[-20, 20]$ m/s ²		$[-10, 10]$ m/s ²		$[0, 5]$ N		$[0, 12]$ rad ² /s ²	
				$[-4, 4]$ rad/s		$[-100, 100]$ N m			for all inputs

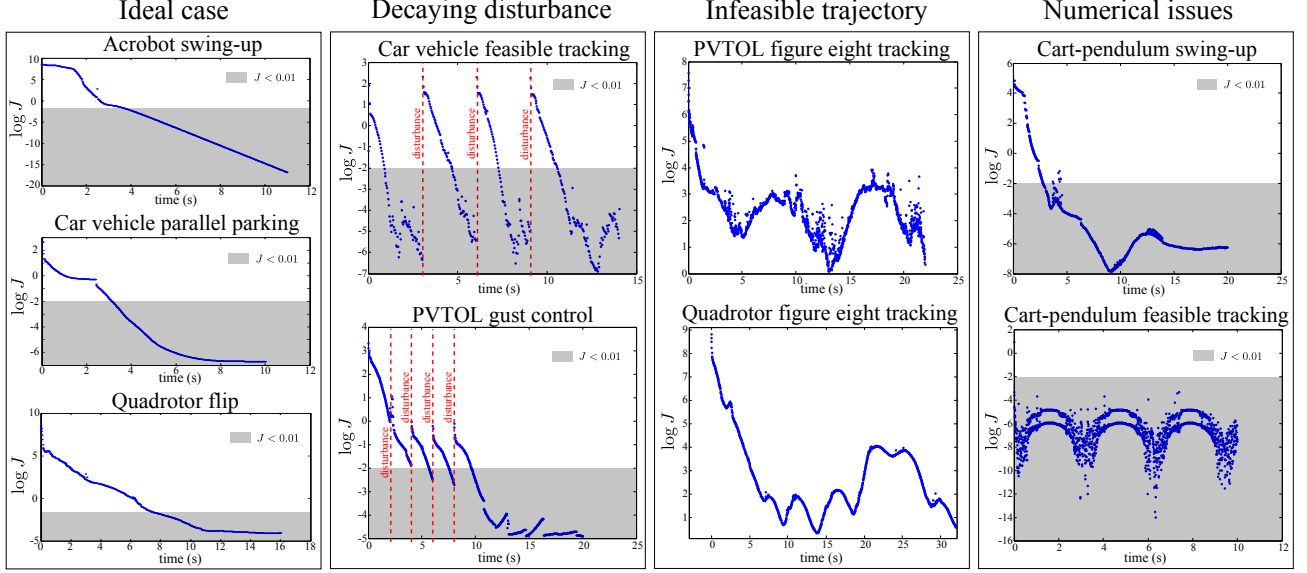


Fig. 4. Cost behavior in the presented examples. As explained in Section VI-F, numerical issues may arise when the system trajectories are near the desired equilibrium (see e.g., gray area in the plots). As a result, the contractive constraint may not be satisfied momentarily leading to a small cost increase. While this may not be a problem since the system is effectively already stabilized at that point (compare corresponding trajectories from Fig. 3), one way to deal with this issue is to switch to the duration-independent, state feedback version of iSAC near the equilibrium (Theorem 1) as demonstrated by the acrobot swing-up panel. Finally, columns three and four show the cost behavior when acting under the presence of an external bounded disturbance and tracking infeasible trajectories respectively. For the former, the unmodeled disturbance leads to a temporary cost increase, but our method responds accordingly as the disturbance attenuates. For the latter, as explained in Section V-C, the contractive constraint is no longer feasible which explains why the cost is not consistently decreasing. Nevertheless, iSAC keeps the deviation from the desired trajectory, e.g., the cost, bounded.

from feasible and infeasible trajectory tracking to set point stabilization with or without external disturbances. Finally, we presented simulation examples using five different systems to demonstrate the applicability and flexibility of our method.

APPENDIX

A. Proof of Proposition 1

Substituting (7) into the mode insertion gradient formula given in (6) we get (we omit superscripts for brevity)

$$\begin{aligned} \frac{dJ}{d\lambda}(t) &= \Gamma(t) (\Gamma(t)^T \Gamma(t) + R^T)^{-1} \Gamma(t)^T \alpha_d \\ &= \alpha_d \|\Gamma(t)\|_{(\Gamma(t)^T \Gamma(t) + R^T)^{-1}}^2 < 0, \end{aligned} \quad (22)$$

where $\Gamma(t)^T = h(t, x(t))^T \rho(t)$. The term $\Gamma(t)^T \Gamma(t)$ produces a positive semi-definite matrix, and adding $R > 0$ yields a positive definite matrix. Because the inverse of a positive definite matrix is positive definite, the quadratic norm $\|\Gamma(t)\|_{(\Gamma(t)^T \Gamma(t) + R^T)^{-1}}^2$ is positive for $\Gamma(t)^T \neq 0$. Therefore, if $\alpha_d < 0$, $\frac{dJ}{d\lambda}(t) < 0$, and thus (7) can reduce (2), e.g., by using (9). ■

B. Proof of Proposition 2

From Proposition 1 and (9), we know that $J(x_i^*(\cdot)) - J(x_i^{def}(\cdot)) = \Delta J < 0$ for each λ in a neighborhood around $\lambda \rightarrow 0^+$ if $\alpha_d < 0$. For the special case of $t_s \rightarrow 0$, it follows directly from Proposition 1 that (14) is always feasible since $C \equiv 0$. For the general case of nonzero sampling time, by continuity assumptions for m and l , C in (14) is continuous with respect to t_s . Thus, it follows that there exists a sufficiently small t_s such that $\Delta J < C$. ■

C. Proof of Proposition 3

The open-loop problem \mathcal{B} is solved by following the same four sequential steps as in the open-loop problem \mathcal{P} in Section II. The only difference is that in the solution process of \mathcal{B} , the superscripts *nom* are replaced by *def*. To show that $u_i^*(t)$ exists, we will show that each of the four solution steps has a solution:

1) The first step in Section II-1 involves calculating the solutions to (1) and (5). Since the default control is in general discontinuous, if solutions are interpreted as sample and hold

(CLSS) solutions (see [58]), existence follows directly from Assumptions 2, 3.

2) In the second step our method in II-2 calculates the optimal action schedule u_s^* by minimizing (6). Because (6) is convex with a continuous first variation from Assumptions 1-3, solutions (7) exist and are unique, which is also both necessary and sufficient for global optimality of (6).

3) In the third step in II-3 the process selects τ_A and u_A by minimizing (8). Because $\tau_A \in [t_i, t_i + T]$ and (8) is in general piecewise continuous (and thus bounded), a solution to this one-dimensional problem exists.

4) Finally, from Proposition 2, the backtracking process in Section II-4 is guaranteed to find a duration that satisfies condition (14) for sufficiently small t_s .

Since all four subproblems have solutions, the open-loop solution $u_i^*(t)$ exists. ■

D. Proof of Theorem 1

At final time, $\rho(t_i + T) = P_1 x(t_i + T)$ from (5) and (16). Due to continuity, the following linear relationship must exist for a nonzero neighborhood of the final time, $\mathcal{N}(t_i + T)$,

$$\rho(t) = P(t)x(t). \quad (23)$$

Ignoring the subscripts and time dependence for brevity, the schedule (7) can be formulated using (23) as

$$u_s^*(t) = u^{def} + (h(t, x)^T P x x^T P^T h(t, x) + R^T)^{-1} h(t, x)^T P x \alpha_d.$$

For $x \in \mathcal{N}(0)$ the quadratic terms in the relationship above go to zero faster than the linear terms, and thus $u_s^*(t)$ converges to the linear feedback regulator

$$u_s^*(t) = u^{def} + R^{-1} h(t, 0)^T P(t)x(t)\alpha_d. \quad \blacksquare$$

E. Proof of Corollary 1

This can be shown by Direct application of the Maximum Principle to (18). ■

F. Proof of Theorem 2

The proof has two parts; Lemma 1 shows that the integral $\int_{t_0}^t M(x(s)) ds$ is bounded for $t \rightarrow \infty$ (see Assumption 3). The latter is used in conjunction with a well-known lemma found, e.g., in [59], [60], to prove asymptotic convergence in the second part of the proof.

First, define

$$V(\alpha, \beta, x(\cdot)) = \int_{\alpha}^{\beta} l(s, x(s)) ds + m(\beta, x(\beta)). \quad (24)$$

Consider the horizon interval $[t_i, t_i + T]$. Let $u_i^*(t)$ be the solution to the open-loop problem $\mathcal{B}(t_i, x_i)$ and $x_i^*(t)$ the corresponding state trajectory. Clearly, $V(t_i, t_i + T, x_i^*(\cdot)) = J(x_i^*(\cdot))$. Then, for $t \in [t_i, t_i + T]$

$$\begin{aligned} V(t, t_i + T, x_i^*(\cdot)) \\ = V(t_i, t_i + T, x_i^*(\cdot)) - \int_{t_i}^t l(s, x_i^*(s)) ds. \end{aligned} \quad (25)$$

Based on the above, we now present the following lemma.

Lemma 1: For small t_s , all $t \in [t_i, t_i + T]$ and all $i \in \mathbb{N}$

$$\begin{aligned} \int_{t_0}^t M(\bar{x}(s)) ds \leq V(t_0, t_0 + T, x_0^*(\cdot)) \\ - V(t, t_i + T, x_i^*(\cdot)), \end{aligned} \quad (26)$$

with

$$\bar{x}(t) = \begin{cases} \bar{x}_{cl}(t) & \text{for } t < t_i \\ x_i^*(t) & \text{else} \end{cases}.$$

Proof: From (11) and Assumption 3 we get that

$$\begin{aligned} V(t_i, t_i + T, x_i^*(\cdot)) - V(t_{i-1}, t_{i-1} + T, x_{i-1}^*(\cdot)) \\ \leq - \int_{t_{i-1}}^{t_i} M(x_{i-1}^*(s)) ds \end{aligned} \quad (27)$$

holds in $[t_i, t_i + T]$ for sufficiently small t_s (Proposition 2). Using the corresponding inequalities from the previous time steps until $[t_0, t_0 + T]$ and the fact that the open-loop solution $u_i^*(\cdot)$ is only applied in $[t_i, t_{i+1}]$ we can write

$$\begin{aligned} \int_{t_0}^{t_i} M(\bar{x}_{cl}(t)) ds \leq V(t_0, t_0 + T, x_0^*(\cdot)) \\ - V(t_i, t_i + T, x_i^*(\cdot)). \end{aligned} \quad (28)$$

Also, from (25) we have

$$\begin{aligned} \int_{t_i}^t M(x_i^*(s)) ds \leq V(t_i, t_i + T, x_i^*(\cdot)) \\ - V(t, t_i + T, x_i^*(\cdot)). \end{aligned} \quad (29)$$

Adding (28) and (29) leads to (26) and Lemma 1 is proved. ■

From Lemma 1, because $V(t, t_i + T, x_i^*(\cdot)) \geq 0$ and M is positive definite, we can deduce that $\int_{t_0}^t M(x(s)) ds$ is bounded for $t \rightarrow \infty$. We also have that $x_i^*(\cdot)$, and thus $\bar{x}_{cl}(\cdot)$, are bounded and from the properties of f , $\dot{x}_i^*(\cdot)$ and $\bar{x}_{cl}(\cdot)$ are bounded as well. These facts combine with the following well-known lemma to prove asymptotic convergence.

Lemma 2: Let $x : \mathbb{R}^+ \rightarrow \mathcal{X}$ be an absolutely continuous function and $M : \mathcal{X} \rightarrow \mathbb{R}^+$ be a continuous, positive definite function ($0 \in \mathcal{X}$). If

$$\|x(\cdot)\|_{L^\infty(\mathbb{R}^+)} < \infty,$$

$$\|\dot{x}(\cdot)\|_{L^\infty(\mathbb{R}^+)} < \infty, \text{ and}$$

$$\lim_{T \rightarrow \infty} \int_0^T M(x(t)) dt < \infty$$

then $x(t) \rightarrow 0$ as $t \rightarrow \infty$.

Proof: The proof can be found, e.g., in [59], [60]. ■

Theorem 2 is proved. ■

G. Proof of Theorem 3

We will show that, for all i , there exist positive constants a, b, c , such that

1. $a\|x_i^*\|^2 \leq J(t_i, x_i^*(\cdot)) \leq b\|x_i^*\|^2$
2. $J(t_i, x_i^*(\cdot)) - J(t_{i-1}, x_{i-1}^*(\cdot)) \leq -c\|x_{i-1}^*\|^2$.

Using Assumption 4, we can find a constant $d > 0$ such that

$$\|x_i^*(t)\| \leq d\|x_i^*\|, \forall t \in [t_i, t_i + T], i \in \mathbb{Z}^+ \quad (30)$$

with $x_i^* = x_i^*(t_i)$. Since u is constrained, this is always true except for systems with finite escape times which are already ruled out from Assumption 4.

• *Upper bound on $J(t_i, x_i^*(\cdot))$* : From (19) and (30) we have

$$\begin{aligned} J(t_i, x_i^*(\cdot)) &\leq \frac{1}{2}T\lambda_{\max}(Q)d^2\|x_i^*\|^2 + \frac{1}{2}\lambda_{\max}(P_1)d^2\|x_i^*\|^2 \\ &= \frac{d^2}{2}(T\lambda_{\max}(Q) + \lambda_{\max}(P_1))\|x_i^*\|^2 = b\|x_i^*\|^2. \end{aligned} \quad (31)$$

• *Lower bound on $J(t_i, x_i^*(\cdot))$* : Using (30), the reverse triangle inequality and $\|f\| \leq \xi\|x\|$ we have

$$\begin{aligned} \|x_i^*(t)\| &\geq \|x_i^*\| - \int_{t_i}^t \|f\| d\tau \\ &\geq \|x_i^*\| - \int_{t_i}^t \xi\|x_i^*(\tau)\| d\tau \\ &\geq [1 - \xi d(t - t_i)]\|x_i^*\|. \end{aligned}$$

It then follows, for example, that

$$\|x_i^*(t)\| \geq \frac{\|x_i^*\|}{2} \text{ for } t \in \left[t_i, t_i + \frac{1}{2\xi d} \right]. \quad (32)$$

Thus, we have two cases to consider:

1) $t_i + T \leq t_i + \frac{1}{2\xi d}$, or $T \leq \frac{1}{2\xi d}$.

In this case,

$$\begin{aligned} J(t_i, x_i^*(\cdot)) &\geq \frac{1}{2} \int_{t_i}^{t_i+T} \lambda_{\min}(Q)\|x_i^*(t)\|^2 dt \\ &\quad + \frac{1}{2}\lambda_{\min}(P_1)\|x_i^*(t_i + T)\|^2 \end{aligned}$$

or

$$J(t_i, x_i^*(\cdot)) \geq \frac{1}{8}T\lambda_{\min}(Q)\|x_i^*\|^2. \quad (33)$$

2) $t_i + T \geq t_i + \frac{1}{2\xi d}$, or $T \geq \frac{1}{2\xi d}$.

In this case

$$\begin{aligned} J(t_i, x_i^*(\cdot)) &\geq \frac{1}{2} \int_{t_i}^{t_i + \frac{1}{2\xi d}} \lambda_{\min}(Q)\|x_i^*(t)\|^2 dt \\ &\quad + \frac{1}{2}\lambda_{\min}(P_1)\|x_i^*(t_i + T)\|^2 \end{aligned}$$

or

$$J(t_i, x_i^*(\cdot)) \geq \frac{1}{16\xi d}\lambda_{\min}(Q)\|x_i^*\|^2. \quad (34)$$

Thus, it follows from (33), (34) that

$$\begin{aligned} J(t_i, x_i^*(\cdot)) &\geq \min \left\{ \frac{1}{8}T\lambda_{\min}(Q), \frac{1}{16\xi d}\lambda_{\min}(Q) \right\} \|x_i^*\|^2 \\ &= a\|x_i^*\|^2. \end{aligned} \quad (35)$$

• *Upper bound on $J(t_i, x_i^*(\cdot)) - J(t_{i-1}, x_{i-1}^*(\cdot))$* : From (11) and (32) we have the following two cases to investigate.

1) $t_i \leq t_{i-1} + \frac{1}{2\xi d}$, or $t_s \leq \frac{1}{2\xi d}$.

In this case,

$$\begin{aligned} &J(t_i, x_i^*(\cdot)) - J(t_{i-1}, x_{i-1}^*(\cdot)) \\ &\leq - \int_{t_{i-1}}^{t_i} \lambda_{\min}(Q)\|x_{i-1}^*(t)\|^2 dt \leq -\frac{1}{4}t_s\lambda_{\min}(Q)\|x_{i-1}^*\|^2. \end{aligned} \quad (36)$$

2) $t_i \geq t_{i-1} + \frac{1}{2\xi d}$, or $t_s \geq \frac{1}{2\xi d}$.

In this case

$$\begin{aligned} &J(t_i, x_i^*(\cdot)) - J(t_{i-1}, x_{i-1}^*(\cdot)) \\ &\leq - \int_{t_{i-1}}^{t_{i-1} + \frac{1}{2\xi d}} \lambda_{\min}(Q)\|x_{i-1}^*(t)\|^2 dt \\ &\leq -\frac{1}{8\xi d}\lambda_{\min}(Q)\|x_{i-1}^*\|^2. \end{aligned} \quad (37)$$

Thus, it follows from (36), (37) that

$$\begin{aligned} &J(t_i, x_i^*(\cdot)) - J(t_{i-1}, x_{i-1}^*(\cdot)) \\ &\leq - \min \left\{ \frac{1}{4}t_s\lambda_{\min}(Q), \frac{1}{8\xi d}\lambda_{\min}(Q) \right\} \|x_{i-1}^*\|^2 = -c\|x_{i-1}^*\|^2. \end{aligned} \quad (38)$$

Theorem 3 is proved. \blacksquare

H. Proof of Corollary 2

We will show that, for all i , there exist positive constant e , such that

$$J(t_i, x_i^*(\cdot)) := \frac{\partial J(t_i, x_i^*(\cdot))}{\partial t_i} \leq -e\|x_i^*\|^2.$$

From (38), we have $t_s \leq \frac{1}{2\xi d}$, and thus

$$J(t_i, x_i^*(\cdot)) - J(t_{i-1}, x_{i-1}^*(\cdot)) \leq -\frac{1}{4}t_s\lambda_{\min}(Q)\|x_{i-1}^*\|^2.$$

Dividing by $t_i - t_{i-1} = t_s$ and taking the limit $t_s \rightarrow 0$ we have

$$\frac{\partial J(t_i, x_i^*(\cdot))}{\partial t_i} \leq -\frac{1}{4}\lambda_{\min}(Q)\|x_{i-1}^*\|^2 = -e\|x_i^*\|^2. \quad (39)$$

Then, combined with (31) and (35), J is a continuous Lyapunov function for the closed-loop system, which implies asymptotic stability in the classical sense, i.e., Lyapunov stability and attractiveness. \blacksquare

I. Proof of Theorem 4

The difference between the dynamics of the plant and the model used by iSAC for the control computation at each time step i and $t \in [t_i, t_i + T]$ is given by

$$\dot{x}_i^p(t) - \dot{x}_i^*(t) = f(t, x^p, u_i^*) - f(t, x_i^*, u_i^*) + \eta_i(t). \quad (40)$$

The states of this model are updated using feedback at every t_i , so $x_i^*(t_i) = x^p(t_i)$ for all i . Then, we can integrate (40) to obtain

$$\begin{aligned} &x_i^p(t) - x_i^*(t) \\ &= \int_{t_i}^t [f(s, x^p(s), u_i^*(s)) - f(s, x_i^*(s), u_i^*(s))] + \eta_i(s) ds. \end{aligned}$$

Therefore, using Assumption 6, we can write

$$\begin{aligned} & \|x_i^p(t) - x_i^*(t)\| \\ & \leq \int_{t_i}^t \|f(s, x^p(s), u_i^*(s)) - f(s, x_i^*(s), u_i^*(s))\| ds \\ & \quad + \delta_i T := \Delta_i(t). \end{aligned} \quad (41)$$

The terminal cost m and the running cost l are Lipschitz continuous from Assumption 3, so we can use this relationship to find the expected cost discrepancy:

$$\begin{aligned} & \|J(x_i^p(t)) - J(x_i^*(t))\| \\ & \leq \int_{t_i}^{t_i+T} \|l(s, x_i^p(s)) - l(s, x_i^*(s))\| ds \\ & \quad + \|m(t_i + T, x_i^p(t_i + T)) - m(t_i + T, x_i^*(t_i + T))\| \\ & \leq \int_{t_i}^{t_i+T} L_1 \Delta_i(t_i + T) ds + L_2 \Delta_i(t_i + T) \\ & \leq \Delta_i(t_i + T)(L_1 T + L_2). \end{aligned} \quad (42)$$

with Lipschitz constants L_1 and L_2 . Using (42) one can show that the upper bound on the contractive constraint (11) will inevitably be higher due to the effect of the disturbance. Specifically, from (11) and (42)

$$\begin{aligned} J(x_i^*(\cdot)) - J(x_{i-1}^*(\cdot)) & \leq - \int_{t_{i-1}}^{t_i} l(t, x_{i-1}^*(t)) dt \\ & \quad + (L_1 T + L_2)[\Delta_i(t_i + T) + \Delta_{i-1}(t_{i-1} + T)]. \end{aligned} \quad (43)$$

This practically means that the feasibility of the contractive constraint cannot be ensured. Thus, the cost is allowed to increase for as long as the disturbance is active and we follow the same procedure as in Section V-C to keep the increase small.

The following Lemma is the last step before proving the theorem.

Lemma 3: As $t \rightarrow \infty$, $\Delta_i(t) \rightarrow 0$ in (41).

Proof: From Assumption 6, $\delta_i \rightarrow 0$ as $i \rightarrow \infty$. Also, using the fact that $x_i^*(t_i) = x^p(t_i)$ and Assumption 2, $x_i^p(t) \rightarrow x_i^*(t)$ as $i \rightarrow \infty$. ■

We can now follow the steps in the proof of Theorem 2 to show asymptotic stability under the disturbance $\eta(t)$. First, we can use Assumption 3, (43) and Lemma 3 to show that $\int_{t_0}^t M(x(s)) ds$ is bounded for $t \rightarrow \infty$ by following the same procedure as in Lemma 1. Asymptotic stability then follows directly from Lemma 2. The process is left as an exercise to the reader.

Theorem 4 is proved. ■

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant CNS 1329891. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Sokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.
- [2] D. Q. Mayne, "Model predictive control: Recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.
- [3] A. R. Ansari and T. D. Murphey, "Sequential action control: Closed-form optimal control for nonlinear and nonsmooth systems," *IEEE Transactions on Robotics*, vol. 32, no. 5, pp. 1196–1214, 2016.
- [4] A. Mavrommati, A. Ansari, and T. Murphey, "Optimal control-on-request: An application in real-time assistive balance control," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 5928–5934.
- [5] E. Tzorakoleftherakis, A. Ansari, A. Wilson, J. Schultz, and T. D. Murphey, "Model-based reactive control for hybrid and high-dimensional robotic systems," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 431–438, 2016.
- [6] T. Fan and T. Murphey, "Online feedback control for input-saturated robotic systems on lie groups," in *Proceedings of Robotics: Science and Systems*, 2016.
- [7] G. Mamakoukas, M. A. MacIver, and T. D. Murphey, "Sequential action control for models of underactuated underwater vehicles in a planar ideal fluid," in *American Control Conference (ACC)*, 2016, 2016, pp. 4500–4506.
- [8] M. Egerstedt, Y. Wardi, and H. Axelsson, "Optimal control of switching times in hybrid systems," in *IEEE International Conference on Methods and Models in Automation and Robotics*, 2003.
- [9] —, "Transition-time optimization for switched-mode dynamical systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 1, pp. 110–115, 2006.
- [10] L. S. Pontryagin, *Mathematical Theory of Optimal Processes*. CRC Press, 1987.
- [11] K. J. Åström and K. Furuta, "Swinging up a pendulum by energy control," *Automatica*, vol. 36, no. 2, pp. 287–295, 2000.
- [12] M. da Silva, Y. Abe, and J. Popović, "Interactive simulation of stylized human locomotion," *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3, p. 82, 2008.
- [13] Y. Tassa, T. Erez, and W. D. Smart, "Receding horizon differential dynamic programming," in *Advances in Neural Information Processing Systems*, 2008, pp. 1465–1472.
- [14] Y. Tassa, T. Erez, and E. Todorov, "Synthesis and stabilization of complex behaviors through online trajectory optimization," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012, pp. 4906–4913.
- [15] H. Chen and F. Allgöwer, "Nonlinear model predictive control schemes with guaranteed stability," in *Nonlinear Model-based Process Control*. Springer, 1998, pp. 465–494.
- [16] L. Grüne and J. Pannek, *Nonlinear Model Predictive Control*. Springer, 2011.
- [17] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2006.
- [18] L. T. Biegler, "A survey on sensitivity-based nonlinear model predictive control," in *Proceedings of the IFAC International Symposium on Dynamics and Control of Process Systems*, 2013, pp. 499–510.
- [19] H. Chen and F. Allgöwer, "A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability," in *European Control Conference (ECC)*, 1997, pp. 1421–1426.
- [20] S. L. de Oliveira Kothare and M. Morari, "Contractive model predictive control for constrained nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1053–1071, 2000.
- [21] E. Camponogara, D. Jia, B. H. Krogh, and S. Talukdar, "Distributed model predictive control," *IEEE Control Systems*, vol. 22, no. 1, pp. 44–52, 2002.
- [22] F. Xie and R. Fierro, "First-state contractive model predictive control of nonholonomic mobile robots," in *American Control Conference (ACC)*, 2008, pp. 3494–3499.
- [23] G. Ferrari-Trecate, L. Galbusera, M. P. E. Marciandi, and R. Scattolini, "Model predictive control schemes for consensus in multi-agent systems with single- and double-integrator dynamics," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2560–2572, 2009.
- [24] F. A. Fontes, "A general framework to design stabilizing nonlinear model predictive controllers," *Systems & Control Letters*, vol. 42, no. 2, pp. 127–143, 2001.

- [25] J. H. Lee, "Model predictive control: review of the three decades of development," *International Journal of Control, Automation and Systems*, vol. 9, no. 3, pp. 415–424, 2011.
- [26] W.-H. Chen, J. O'Reilly, and D. J. Ballance, "On the terminal region of model predictive control for non-linear systems with input/state constraints," *International Journal of Adaptive Control and Signal Processing*, vol. 17, no. 3, pp. 195–207, 2003.
- [27] S. Yu, H. Chen, C. Böhm, and F. Allgöwer, "Enlarging the terminal region of NMPC with parameter-dependent terminal control law," in *Nonlinear Model Predictive Control*. Springer, 2009, pp. 69–78.
- [28] P. A. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," Ph.D. dissertation, California Institute of Technology, 2000.
- [29] T. M. Caldwell and T. D. Murphey, "Projection-based switched system optimization: Absolute continuity of the line search," in *Proceedings of the IEEE Conference on Decision and Control (CDC)*, 2012, pp. 699–706.
- [30] J. P. Hespanha, *Linear systems theory*. Princeton university press, 2009.
- [31] R. Tedrake, I. R. Manchester, M. Tobenkin, and J. W. Roberts, "LQR-trees: Feedback motion planning via sums-of-squares verification," *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 1038–1052, 2010.
- [32] P. O. Scokaert, D. Q. Mayne, and J. B. Rawlings, "Suboptimal model predictive control (feasibility implies stability)," *IEEE Transactions on Automatic Control*, vol. 44, no. 3, pp. 648–654, 1999.
- [33] A. Jadbabaie, J. Yu, and J. Hauser, "Unconstrained receding-horizon control of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 46, no. 5, pp. 776–783, 2001.
- [34] D. Mellinger and V. Kumar, "Minimum snap trajectory generation and control for quadrotors," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 2520–2525.
- [35] E. Tzorakoleftherakis and T. D. Murphey, "Controllers as filters: Noise-driven swing-up control based on maxwell's demon," in *Proceedings of the IEEE Conference on Decision and Control (CDC)*, 2015, pp. 4368–4374.
- [36] X. Xin and T. Yamasaki, "Energy-based swing-up control for a remotely driven acrobot: Theoretical and experimental results," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 4, pp. 1048–1056, 2012.
- [37] M. W. Spong, "The swing up control problem for the acrobot," *IEEE Control Systems*, vol. 15, no. 1, pp. 49–55, 1995.
- [38] T. Albahkali, R. Mukherjee, and T. Das, "Swing-up control of the pendubot: an impulse-momentum approach," *IEEE Transactions on Robotics*, vol. 25, no. 4, pp. 975–982, 2009.
- [39] X. Xin and M. Kaneda, "Analysis of the energy-based swing-up control of the acrobot," *International Journal of Robust and Nonlinear Control*, vol. 17, no. 16, pp. 1503–1524, 2007.
- [40] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*. Academic press, 1981.
- [41] C.-W. Tao, J.-S. Taur, T. W. Hsieh, and C. Tsai, "Design of a fuzzy controller with fuzzy swing-up and parallel distributed pole assignment schemes for an inverted pendulum and cart system," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 6, pp. 1277–1288, 2008.
- [42] R. W. Brockett *et al.*, "Asymptotic stability and feedback stabilization," *Differential Geometric Control Theory*, vol. 27, no. 1, pp. 181–191, 1983.
- [43] C. Samson and K. Ait-Abderrahim, "Feedback control of a nonholonomic wheeled cart in cartesian space," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 1991, pp. 1136–1141.
- [44] C. C. De Wit and O. Sordalen, "Exponential stabilization of mobile robots with nonholonomic constraints," *IEEE Transactions on Automatic Control*, vol. 37, no. 11, pp. 1791–1797, 1992.
- [45] A. Astolfi, "On the stabilization of nonholonomic systems," in *Proceedings of the IEEE Conference on Decision and Control (CDC)*, vol. 4, 1994, pp. 3481–3486.
- [46] C.-S. Chiu, K.-Y. Lian, and P. Liu, "Fuzzy gain scheduling for parallel parking a car-like robot," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 6, pp. 1084–1092, 2005.
- [47] A. De Luca, G. Oriolo, and M. Vendittelli, "Stabilization of the unicycle via dynamic feedback linearization," in *IFAC International Symposium on Robot Control*, 2000, pp. 397–402.
- [48] D. Gu and H. Hu, "A stabilizing receding horizon regulator for nonholonomic mobile robots," *IEEE Transactions on Robotics*, vol. 21, no. 5, pp. 1022–1028, 2005.
- [49] J. Hauser, S. Sastry, and G. Meyer, "Nonlinear control design for slightly non-minimum phase systems: application to v/stol aircraft," *Automatica*, vol. 28, no. 4, pp. 665–679, 1992.
- [50] S.-L. Wu, P.-C. Chen, C.-H. Hsu, and K.-Y. Chang, "Gain-scheduled control of pvtol aircraft dynamics with parameter-dependent disturbance," *Journal of the Franklin Institute*, vol. 345, no. 8, pp. 906–925, 2008.
- [51] F. Lin, W. Zhang, and R. D. Brandt, "Robust hovering control of a pvtol aircraft," *IEEE Transactions on Control Systems Technology*, vol. 7, no. 3, pp. 343–351, 1999.
- [52] S. Bouabdallah and R. Siegwart, "Backstepping and sliding-mode techniques applied to an indoor micro quadrotor," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2005, pp. 2247–2252.
- [53] A. E. C. D. Cunha, "Benchmark: Quadrotor attitude control," in *Proceedings of the International Workshop on Applied Verification for Continuous and Hybrid Systems (ARCH)*, vol. 34, 2015, pp. 57–72.
- [54] P. Pounds, R. Mahony, and P. Corke, "Modelling and control of a large quadrotor robot," *Control Engineering Practice*, vol. 18, no. 7, pp. 691–699, 2010.
- [55] S. Bouabdallah, A. Noth, and R. Siegwart, "PID vs LQ control techniques applied to an indoor micro quadrotor," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, 2004, pp. 2451–2456.
- [56] E. Fresk and G. Nikolakopoulos, "Full quaternion based attitude control for a quadrotor," in *European Control Conference (ECC)*, 2013, pp. 3864–3869.
- [57] D. Lee, H. J. Kim, and S. Sastry, "Feedback linearization vs. adaptive sliding mode control for a quadrotor helicopter," *International Journal of Control, Automation and Systems*, vol. 7, no. 3, pp. 419–428, 2009.
- [58] F. H. Clarke, Y. S. Ledyev, E. D. Sontag, and A. I. Subbotin, "Asymptotic controllability implies feedback stabilization," *IEEE Transactions on Automatic Control*, vol. 42, no. 10, pp. 1394–1407, 1997.
- [59] H. Michalska and R. Vinter, "Nonlinear stabilization using discontinuous moving-horizon control," *IMA Journal of Mathematical Control and Information*, vol. 11, no. 4, pp. 321–340, 1994.
- [60] I. Barbalat, "Systemes d'equations differentielles d'oscillations non lineaires," *Rev. Math. Pures Appl.*, vol. 4, no. 2, pp. 267–270, 1959.