TWCCC ⋆ Texas – Wisconsin – California Control Consortium

Technical report number 2024-04

Offset-free model predictive control: stability under plant-model mismatch*

Steven J. Kuntz and James B. Rawlings
Department of Chemical Engineering
University of California, Santa Barbara

December 10, 2024

Abstract

We present the first general stability results for nonlinear offset-free model predictive control (MPC). Despite over twenty years of active research, the offset-free MPC literature has not shaken the assumption of closed-loop stability for establishing offset-free performance. In this paper, we present a nonlinear offset-free MPC design that is robustly stable with respect to the tracking errors, and thus achieves offset-free performance, despite plant-model mismatch and persistent disturbances. Key features and assumptions of this design include quadratic costs, differentiability of the plant and model functions, constraint backoffs at steady state, and a robustly stable state and disturbance estimator. We first establish nominal stability and offset-free performance. Then, robustness to state and disturbance estimate errors and setpoint and disturbance changes is demonstrated. Finally, the results are extended to sufficiently small plant-model mismatch. The results are illustrated by numerical examples.

1 Introduction

Offset-free model predictive control (MPC) is a popular advanced control method for offset-free tracking of setpoints despite plant-model mismatch and persistent disturbances. This is accomplished by combining regulation, estimation, and steady-state target problems,

*This report is an extended version of a submitted paper. This work was supported by the National Science Foundation (NSF) under Grant 2138985. (e-mail: skuntz@ucsb.edu; jbraw@ucsb.edu)

each designed with a state-space model that is augmented with uncontrollable integrating modes, called *integrating disturbances*, that provide integral action through the estimator. Despite over twenty years of applied use and active research, there are no results on the stability of nonlinear offset-free MPC.

Sufficient conditions for which linear offset-free MPC stability implies offset-free performance were first established by Muske and Badgwell (2002); Pannocchia and Rawlings (2003). While Muske and Badgwell (2002); Pannocchia and Rawlings (2003) do not explicitly mention control of nonlinear plants, the results are widely applicable to both linear and nonlinear plants with asymptotically constant disturbances, as controller stability is assumed rather than explicitly demonstrated. In fact, Pannocchia and Rawlings (2003) demonstrate offset-free control on a highly nonlinear, non-isothermal reactor model.

Offset-free MPC designs with *nonlinear models* and tracking costs were first considered by Morari and Maeder (2012). For the special case of state feedback, Pannocchia et al. (2015) give a disturbance model and estimator design for which the offset-free MPC is provably asymptotically stable and offset-free. In Pannocchia et al. (2015), the state-feedback observer design is generalized to economic cost functions, and convergence to the optimal steady state is demonstrated. A general, output-feedback offset-free *economic* MPC was first proposed by Vaccari and Pannocchia (2017), who use gradient correction strategies to ensure the economic MPC, if it converges, achieves the optimal steady-state performance. For further developments of offset-free economic MPC, we refer the reader to Pannocchia (2018); Faulwasser and Pannocchia (2019); Vaccari et al. (2021).

To the best of our knowledge, there are no stability results for offset-free MPC in the intended setting: persistent disturbances and plant-model mismatch. The results discussed thus far have assumed closed-loop stability rather than proven it. Some authors have proposed provably stable nonlinear MPC designs for output tracking (Falugi, 2015; Limon et al., 2018; Köhler et al., 2020; Berberich et al., 2022; Galuppini et al., 2023; Soloperto et al., 2023), but access to the plant dynamic equations is assumed and process and measurement disturbances are not considered.

In this paper, we propose a nonlinear offset-free MPC design that has offset-free performance and asymptotic stability subject to plant-model mismatch, persistent disturbances, and changing references. Based on the results in Kuntz and Rawlings (2024), we use positive definite quadratic costs and assume differentiability of the plant and model equations to ensure the plant-model mismatch does not prevent stability with respect to the steady-state targets. To ensure the controller is robustly feasible, we soften any output constraints in the regulator using an exact penalty method, and to guarantee nominal regulator stability, we apply constraint backoffs to the steady-state target problem. Lipschitz continuity of the steady-state target problem solutions is required to guarantee robustness to estimate errors and setpoint and disturbance changes.

The remainder of this section outlines the paper and establishes notation, definitions, and basic facts used throughout. In Section 2, the offset-free MPC design is presented. In Section 4, we establish asymptotic stability of the nominal system. In Section 5, we establish robust performance with respect to estimate errors, setpoint changes, and disturbance changes. In Section 6, we extend these results to the mismatched system using the approach in Kuntz and Rawlings (2024). Finally, in Section 8, we conclude the paper with a discussion of limitations and future work.

Notation Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{>0}$ denote the real, nonnegative real, and positive real numbers, respectively. Let $\mathbb{I}$, $\mathbb{I}_{\geq 0}$, $\mathbb{I}_{>0}$, and $\mathbb{I}_{m:n}$ denote the integers, nonnegative integers, positive integers, and integers from m to n (inclusive), respectively. Let $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote real n-vectors and $n \times m$ matrices, respectively. For any matrix $A \in \mathbb{R}^{n\times n}$, we denote by $\lambda(A)$ the set of eigenvalues of A, we call $\rho(A) := \max_{\lambda\in\lambda(A)} |\lambda|$ the spectral radius of A. We say A is Schur stable if $\rho(A) < 1$. We denote by $\underline{\sigma}(A)$ and $\overline{\sigma}(A)$ the smallest and largest singular values of $A \in \mathbb{R}^{n\times n}$. We say a symmetric matrix $P = P^\top \in \mathbb{R}^{n\times n}$ is positive definite (semidefinite) if $x^\top Px > 0$ ($x^\top Px \geq 0$) for all nonzero $x \in \mathbb{R}^n$. We define the Euclidean and Q-weighted norms by $|x| := \sqrt{x^\top x}$ and $|x|_Q := \sqrt{x^\top Qx}$, for each $x \in \mathbb{R}^n$, where Q is positive definite. For any positive definite $Q \in \mathbb{R}^{n\times n}$, we have $\underline{\sigma}(Q)|x|^2 \leq |x|_Q^2 \leq \overline{\sigma}(Q)|x|^2$ for all $x \in \mathbb{R}^n$. For any function $V : X \to \mathbb{R}$ and $\rho > 0$, we define $\text{lev}_\rho V := \{\, x \in X \mid V(x) \leq \rho \,\}$. For any signal $a(k)$, we denote both infinite and finite sequences in bold font as $\mathbf{a} := (a(0), \ldots, a(k))$ or $\mathbf{a} := (a(0), a(1), \ldots)$. A subsequence of $\mathbf{a}$ is denoted $\mathbf{a}_{i:j} = (a(i), \ldots, a(j))$ where $i \leq j$. We define the infinite and length-k signal norm as $\|\mathbf{a}\| := \sup_{k\geq 0} |a(k)|$ and $\|\mathbf{a}\|_{0:k} := \max_{0\leq i\leq k} |a(i)|$. Let $\mathcal{K}$ be the class of functions $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ that are strictly increasing and $\alpha(0) = 0$. Let $\mathcal{K}_\infty$ be the class of unbounded class-$\mathcal{K}$ functions. Let $\mathcal{KL}$ be the class of functions $\beta : \mathbb{R}_{\geq 0} \times \mathbb{I}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\beta(\cdot, k) \in \mathcal{K}$, $\beta(r, \cdot)$ is nonincreasing, and $\lim_{i\to\infty} \beta(r, i) = 0$, for all $(r, k) \in \mathbb{R}_{\geq 0} \times \mathbb{I}_{\geq 0}$. Let $\text{ID}(\cdot) := (\cdot) \in \mathcal{K}_\infty$ denote the identity map.

2 Problem statement

2.1 System of interest

Consider the following discrete-time plant:

$$x_\text{P}^+ = f_\text{P}(x_\text{P}, u, w_\text{P}) \tag{1a}$$
$$y = h_\text{P}(x_\text{P}, u, w_\text{P}) \tag{1b}$$

where $x_\text{P} \in \mathbb{X} \subseteq \mathbb{R}^n$ is the *plant* state, $u \in \mathbb{U} \subseteq \mathbb{R}^{n_u}$ is the input, $y \in \mathbb{Y} \subseteq \mathbb{R}^{n_y}$ is the output, and $w_\text{P} \in \mathbb{W} \subseteq \mathbb{R}^{n_w}$ is the *plant* disturbance. The functions f_P and h_P are not known. Instead, we assume access to a model of the plant,

$$x^+ = f(x, u, d) \tag{2a}$$
$$y = h(x, u, d) \tag{2b}$$

where $x \in \mathbb{X} \subseteq \mathbb{R}^n$ is the *model* state and $d \in \mathbb{D} \subseteq \mathbb{R}^{n_d}$ is the *model* disturbance. Without loss of generality, we assume the nominal plant and model functions are consistent, i.e.,

$$f(x, u, 0) = f_\text{P}(x, u, 0), \qquad h(x, u, 0) = h_\text{P}(x, u, 0) \tag{3}$$

for all $(x, u) \in \mathbb{X} \times \mathbb{U}$. The plant disturbance w_P may include process and measurement noise, exogenous disturbances, parameter errors, discretization errors, and even unmodeled dynamics. The purpose of the model disturbance d is to correct for steady-state output errors introduced by the plant disturbance w_P. The model disturbance d may include any of the plant disturbances and/or fictitious signals accounting for the effect of the plant disturbances on the steady-state output.

Example 1. Consider a single-state linear plant with parameter errors,

$$\begin{aligned} f_\mathrm{P}(x_\mathrm{P}, u, w_\mathrm{P}) &= (\hat{a} + (w_\mathrm{P})_1)x_\mathrm{P} + (\hat{b} + (w_\mathrm{P})_2)u \\ h_\mathrm{P}(x_\mathrm{P}, u, w_\mathrm{P}) &= x_\mathrm{P} + (w_\mathrm{P})_3 \end{aligned}$$

and a single-state linear model with an input disturbance:

$$f(x, u, d) = \hat{a}x + \hat{b}(u + d), \qquad h(x, u, d) = x.$$

For this example, the plant disturbance w_P includes both parameter errors and measurement noise, whereas the model disturbance only provides the means to shift the model steady states in response to plant disturbances.

The control objective is to drive the reference signal,

$$r = g(u, y) \tag{4}$$

to the setpoint r_sp using only knowledge of the model (2), past (u, y) data, and auxiliary setpoints $(u_\mathrm{sp}, y_\mathrm{sp})$ (to be defined). The setpoints $s_\mathrm{sp} := (r_\mathrm{sp}, u_\mathrm{sp}, y_\mathrm{sp})$ are possibly time-varying, but only the current value is available at a given time. The controller should be *offset-free* when the setpoint and plant disturbances are asymptotically constant, i.e.,

$$(\Delta s_\mathrm{sp}(k), \Delta w_\mathrm{P}(k)) \to 0 \qquad \Rightarrow \qquad r(k) - r_\mathrm{sp}(k) \to 0$$

where $\Delta s_\mathrm{sp}(k) := s_\mathrm{sp}(k) - s_\mathrm{sp}(k-1)$ and $\Delta w_\mathrm{P}(k) := w_\mathrm{P}(k) - w_\mathrm{P}(k-1)$. Otherwise, the amount of offset should be robust to setpoint and disturbance *increments* $(\Delta s_\mathrm{sp}, \Delta w_\mathrm{P})$.

Remark 1. To achieve the nominal consistency assumption (3) and track the reference (4), we typically need the dimensional constraints $n_y \leq n_d$ and $n_r \leq n_u$, respectively. Otherwise their are insufficient degrees of freedom to manipulate the output and reference at steady state with the disturbance and input, respectively.

Remark 2. We do not strictly require an asymptotically constant disturbance. For example, if $r_\mathrm{sp}(k) = \sin(1/k)$ and $w_\mathrm{P} \equiv 0$, then the setpoint increments go to zero $\Delta r_\mathrm{sp}(k) = \sin(1/k) - \sin(1/(k-1)) = O(1/k^2)$. But the setpoint signal becomes approximately constant as $k \to \infty$, so we should expect the offset-free MPC to be approximately offset-free.

Throughout, we make the following assumptions on plant, model, and reference functions.

Assumption 1 (Continuity). The functions $g : \mathbb{U} \times \mathbb{Y} \to \mathbb{R}^{n_r}$, $(f_\mathrm{P}, h_\mathrm{P}) : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \to \mathbb{X} \times \mathbb{Y}$, and $(f, h) : \mathbb{X} \times \mathbb{U} \times \mathbb{D} \to \mathbb{X} \times \mathbb{Y}$ are continuous, and $f(0,0,0) = 0$, $h(0,0,0) = 0$, $g(0,0) = 0$, and (3) holds for all $(x, u) \in \mathbb{X} \times \mathbb{U}$.

2.2 Constraints

The sets $(\mathbb{X}, \mathbb{Y}, \mathbb{D}, \mathbb{W})$ are physical constraints (e.g., nonnegativity of chemical concentrations, temperatures, pressures, etc.) that the systems (1)–(3) automatically satisfy. These are hard constraints enforced only during state estimation. On the other hand, we enforce the hard constraint $u \in \mathbb{U}$ during both regulation and target selection. Additionally, we enforce soft joint input-output constraints of the form

$$\mathbb{Z}_y := \{ (u, y) \in \mathbb{U} \times \mathbb{Y} \mid c_i(u, y) \leq 0 \; \forall \, i \in \mathbb{I}_{1:n_c} \}$$

where $c : \mathbb{U} \times \mathbb{Y} \to \mathbb{R}^{n_c}$ is the soft constraint function. Having active constraints at steady state is problematic, so the constraints are sometimes tightened as follows:

$$\overline{\mathbb{Z}}_y := \{ (u, y) \mid c_i(u, y) + b_i \leq 0 \; \forall \, i \in \mathbb{I}_{1:n_c} \}$$

where $b \in \mathbb{R}^{n_c}_{>0}$ is the vector of back-off constants. No such constraint tightening is required for the input constraints. We assume the constraints and the back-off constant satisfy the following properties throughout.

Assumption 2 (Constraints). The sets $(\mathbb{X}, \mathbb{Y})$ are closed, $(\mathbb{U}, \mathbb{W}, \mathbb{D})$ are compact, and all contain the origin. The soft constraint function $c : \mathbb{U} \times \mathbb{Y} \to \mathbb{R}^{n_c}$ is continuous and

$$0 < b_i < -c_i(0, 0), \qquad \forall \, i \in \mathbb{I}_{1:n_c}.$$

2.3 Offset-free model predictive control

Offset-free MPC consists of three parts or subroutines: target selection, regulation, and state estimation.

2.3.1 Steady-state target problem

Given a *model* disturbance $d \in \mathbb{D}$ and setpoint $r_{\text{sp}} \in \mathbb{R}^{n_r}$, we define the set of offset-free steady-state pairs by

$$\mathcal{Z}_O(r_{\text{sp}}, d) := \{ (x, u) \in \mathbb{X} \times \mathbb{U} \mid x = f(x, u, d), \; y = h(x, u, d), \; (u, y) \in \overline{\mathbb{Z}}_y, \; r_{\text{sp}} = g(u, y) \} . \tag{5}$$

To pick the best steady-state pair among members of $\mathcal{Z}_O(r_{\text{sp}}, d)$, it is customary to optimize the steady state with respect to some auxiliary setpoint pair $z_{\text{sp}} := (u_{\text{sp}}, y_{\text{sp}}) \in \overline{\mathbb{Z}}_y$ (typically chosen such that $r_{\text{sp}} = g(u_{\text{sp}}, y_{\text{sp}})$). For each $(r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, d) \in \mathbb{R}^{n_r} \times \overline{\mathbb{Z}}_y \times \mathbb{D}$, we define the steady-state target problem (SSTP) by

$$V_s^0(\beta) := \min_{(x,u) \in \mathcal{Z}_O(r_{\text{sp}}, d)} \ell_s(u - u_{\text{sp}}, h(x, u, d) - y_{\text{sp}}) \tag{6}$$

where $\beta := (r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, d)$ are the *SSTP parameters* and $\ell_s : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}_{\geq 0}$ is a steady-state cost function, typically a positive definite quadratic. We define the set of feasible SSTP parameters as

$$\mathcal{B} := \{ (r_{\text{sp}}, z_{\text{sp}}, d) \in \mathbb{R}^{n_r} \times \overline{\mathbb{Z}}_y \times \mathbb{D} \mid \mathcal{Z}_O(r_{\text{sp}}, d) \neq \varnothing \} . \tag{7}$$

To guarantee the existence of solutions to the SSTP (6), the following assumption is required.

Assumption 3. The function $\ell_s : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}_{\geq 0}$ is continuous and, for each $\beta = (r_{\rm sp}, u_{\rm sp}, y_{\rm sp}, d) \in \mathcal{B}$, at least one of the following properties holds:

(i) $\mathcal{Z}_O(r_{\rm sp}, d)$ is compact;

(ii) with $V_s(x, u, \beta) := \ell_s(u - u_{\rm sp}, h(x, u, d) - y_{\rm sp})$, the function $V_s(\cdot, \beta)$ is coercive in $\mathcal{Z}_O(r_{\rm sp}, d)$, i.e., for any sequence $\mathbf{z} \in (\mathcal{Z}_O(r_{\rm sp}, d))^\infty$ such that $|z(k)| \to \infty$, we have $V_s(z(k), \beta) \to \infty$.

Under Assumptions 1 to 3, $\mathcal{B}$ is nonempty and the SSTP (6) has solutions for all $\beta \in \mathcal{B}$. The solution to (6) may not be unique. Throughout, we assume some selection rule has been applied and denote the functions returning solutions to (6) by $z_s(\cdot) := (x_s(\cdot), u_s(\cdot)) : \mathcal{B} \to \mathbb{X} \times \mathbb{U}$.

2.3.2 Regulator

Given the SSTP parameters $\beta \in \mathcal{B}$, the regulator is defined as a finite horizon optimal control problem (FHOCP) with the steady-state targets $(x_s(\beta), u_s(\beta))$. We consider a FHOCP with a horizon length $N \in \mathbb{I}_{>0}$, stage cost $\ell : \mathbb{X} \times \mathbb{U} \times \mathcal{B} \to \mathbb{R}_{\geq 0}$, terminal cost $V_f : \mathbb{X} \times \mathcal{B} \to \mathbb{R}_{\geq 0}$, and terminal constraint $\mathbb{X}_f(\beta) \subseteq \mathbb{X}$ (to be defined). For each $\beta = (r_{\rm sp}, u_{\rm sp}, y_{\rm sp}, d) \in \mathcal{B}$, we define the terminal constraint (8), feasible initial state and input sequence pairs (9), feasible input sequences at $x \in \mathbb{X}$ (10), feasible initial states (11), and feasible state-parameter pairs (12) by the sets

$$\mathbb{X}_f(\beta) := \operatorname{lev}_{c_f} V_f(\cdot, \beta) \tag{8}$$

$$\mathcal{Z}_N(\beta) := \{\, (x, \mathbf{u}) \in \mathbb{X} \times \mathbb{U}^N \mid \phi(N; x, \mathbf{u}, d) \in \mathbb{X}_f(\beta) \,\} \tag{9}$$

$$\mathcal{U}_N(x, \beta) := \{\, \mathbf{u} \in \mathbb{U}^N \mid (x, \mathbf{u}) \in \mathcal{Z}_N(\beta) \,\} \tag{10}$$

$$\mathcal{X}_N(\beta) := \{\, x \in \mathbb{X} \mid \mathcal{U}_N(x, \beta) \neq \varnothing \,\} \tag{11}$$

$$\mathcal{S}_N := \{\, (x, \beta) \in \mathbb{X} \times \mathcal{B} \mid \mathcal{U}_N(x, \beta) \neq \varnothing \,\} \tag{12}$$

where $c_f > 0$ and $\phi(k; x, \mathbf{u}, d)$ denotes the solution to (2a) at time k given an initial state x, constant disturbance d, and sufficiently long input sequence $\mathbf{u}$. For each $(x, \mathbf{u}, \beta) \in \mathbb{X} \times \mathbb{U}^N \times \mathcal{B}$, we define the FHOCP objective by

$$V_N(x, \mathbf{u}, \beta) := V_f(\phi(N; x, \mathbf{u}, d), \beta) + \sum_{k=0}^{N-1} \ell(\phi(k; x, \mathbf{u}, d), u(k), \beta). \tag{13}$$

For each $(x, \beta) \in \mathcal{S}_N$, we define the FHOCP by

$$V_N^0(x, \beta) := \min_{\mathbf{u} \in \mathcal{U}_N(x,\beta)} V_N(x, \mathbf{u}, \beta). \tag{14}$$

Using the convention of Rockafellar and Wets (1998) for infeasible problems, we take $V_N^0(x, \beta) := \infty$ for all $(x, \beta) \notin \mathcal{S}_N$.

To guarantee closed-loop stability and robustness, we consider the following assumptions.

Assumption 4 (Terminal control law). There exists a function $\kappa_f : \mathbb{X} \times \mathcal{B} \to \mathbb{U}$ such that

$$V_f(f(x, \kappa_f(x, \beta), d), \beta) - V_f(x, \beta) \le -\ell(x, \kappa_f(x, \beta), \beta)$$

for all $x \in \mathbb{X}_f(\beta)$ and $\beta := (r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, d) \in \mathcal{B}$.

Assumption 5 (Quadratic costs). The stage and terminal costs take the form

$$\begin{aligned}\ell(x, u, \beta) &= |x - x_s(\beta)|_Q^2 + |u - u_s(\beta)|_R^2 + \sum_{i=1}^{n_c} w_i \max\{0, c_i(u, h(x, u, d))\} \\ V_f(x, \beta) &= |x - x_s(\beta)|_{P_f(\beta)}^2\end{aligned}$$

for each $(x, u) \in \mathbb{X} \times \mathbb{U}$ and $\beta := (r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, d) \in \mathcal{B}$, where Q, R, and $P_f(\beta)$ are positive definite matrices for each $\beta \in \mathcal{B}$, the function P_f is continuous, and $w_i > 0$ for each $i \in \mathbb{I}_{1:n_c}$.

Remark 3. With $\beta = (s_{\text{sp}}, d) \in \mathcal{B}$, Assumption 4 and the terminal set definition (8) imply $V_f(f(x, \kappa_f(x, \beta), d), \beta) \le V_f(x, \beta) \le c_f$ for all $x \in \mathbb{X}_f(\beta)$ and therefore $\mathbb{X}_f(\beta)$ is positive invariant for $x^+ = f(x, \kappa_f(x, \beta), d)$.

Assumptions 1 to 3 and 5 guarantee the existence of solutions to (14) for all $(x, \beta) \in \mathcal{S}_N$ (Rawlings et al., 2020, Prop. 2.4). We denote any such solution by $\mathbf{u}^0(x, \beta) = (u^0(0; x, \beta), \dots, u^0(N-1; x, \beta))$, and define the corresponding optimal state $x^0(k; x, \beta) := \phi(k; x, \mathbf{u}^0(x, \beta), d)$ and optimal state sequence $\mathbf{x}^0(x, \beta) := (x^0(0; x, \beta), \dots, x^0(N; x, \beta))$. We define the FHOCP control law by $\kappa_N(x, \beta) := u^0(0; x, \beta)$.

Remark 4. Given Assumptions 1 to 3 and 5, it may be impossible to satisfy Assumption 4 without constraint back-offs, i.e., $b = 0$. This is because the terminal cost difference $V_f(f(x, \kappa_f(x, \beta), d)) - V_f(x)$ is, at best, negative definite with quadratic scaling (regardless of the target value), whereas the stage cost $\ell(x, \kappa_f(x, \beta), \beta)$ has quadratic scaling when the soft constraint is satisfied but linear scaling when the soft constraint is violated. Thus, if the constraints are active at the targets, the stage cost will always exceed the decrease in terminal cost if the state violates the constraints and is sufficiently small.

Example 2. Consider the scalar linear system $x^+ = x + u + d$, $y = x$, and $r = y$ with stage costs of the form Assumption 5 and the soft constraint function $c(u, y) = y - 1$. Let $b = 0$ and $\beta = (1, 0, 1, 0)$. Clearly the target is reachable, and we can take the SSTP (6) solution $(x_s(\beta), u_s(\beta)) = (1, 0)$. Then we have stage costs of the form $\ell(x, u, \beta) = q(x-1)^2 + ru^2 + w \max\{0, x-1\}$ and $V_f(x, \beta) = p_f x^2$, where $q, r, w, p_f > 0$. Assumption 4 is not satisfied if there exists $x \in \mathbb{R}$ such that

$$\mathcal{F}(x, u) := p_f(x + u - 1)^2 - p_f(x-1)^2 + q(x-1)^2 + ru^2 + w \max\{0, x-1\} > 0$$

for all $u \in \mathbb{R}$. Completing the squares gives

$$\begin{aligned}\mathcal{F}(x, u) &= (\tilde{a}u + \tilde{b}(x-1))^2 + \tilde{c}(x-1)^2 + w \max\{0, x-1\} \\ &\ge \tilde{c}(x-1)^2 + w \max\{0, x-1\}\end{aligned}$$

for all $x \in \mathbb{R}$ and $u \in \mathbb{R}$, where $\tilde{a} := \sqrt{r + p_f}$, $\tilde{b} := \frac{p_f}{2\tilde{a}}$, and $\tilde{c} := q - \tilde{b}^2$. Ideally, we would have chosen (q, r, p_f) so that $\tilde{c} < 0$. But this means we can still take $0 < x - 1 < \sqrt{\frac{w}{\tilde{c}}}$ to give

$$\mathcal{F}(x, u) \geq \tilde{c}(x-1)^2 + w(x-1) > 0$$

for all $u \in \mathbb{R}$, no matter the chosen $w > 0$.

On the other hand, let $b = 1$ and $\beta = (0, 0, 0, 0)$. Again, the target is reachable and we can take the SSTP solution $(x_s(0), u_s(0)) = (0, 0)$. Notice that for both problems the backed-off constraint $c(u, y) + b$ is active at the solution. This time, however, we have

$$\begin{aligned}\mathcal{F}(x, u) &:= p_f(x+u)^2 - p_f x^2 + qx^2 + ru^2 + w \max\{0, x-1\} \\ &= (\tilde{a}u + \tilde{b}x)^2 + \tilde{c}x^2 + w \max\{0, x-1\}\end{aligned}$$

and with $\kappa_f(x, 0) := -\frac{\tilde{b}}{\tilde{a}}x$, we have

$$\mathcal{F}(x, \kappa_f(x, 0)) = \tilde{c}x^2 + w \max\{0, x-1\}$$

for all $x \in \mathbb{R}$. Let $c_f = p_f$ and suppose $\tilde{c} < 0$. Then, for each $x \in \mathbb{X}_f(0)$, we have $|x| \leq 1$ and therefore

$$\mathcal{F}(x, \kappa_f(x, 0)) = \tilde{c}x^2 \leq 0.$$

2.3.3 State estimation

In practice, the SSTP and FHOCP are implemented with state and disturbance estimates rather than the true values. To this end, we consider any estimator that estimates both plant and disturbance states.

Definition 1. A *joint state and disturbance estimator* is a sequence of functions $\Phi_k : \mathbb{X} \times \mathbb{D} \times \mathbb{U}^k \times \mathbb{Y}^k \to \mathbb{X} \times \mathbb{D}$ defined for each $k \in \mathbb{I}_{\geq 0}$. For each $k \in \mathbb{I}_{\geq 0}$, we define the *state and disturbance estimates* by

$$(\hat{x}(k), \hat{d}(k)) := \Phi_k(\overline{x}, \overline{d}, \mathbf{u}_{0:k-1}, \mathbf{y}_{0:k-1}) \tag{15}$$

where $(\overline{x}, \overline{d}) \in \mathbb{X} \times \mathbb{D}$ is the initial guess at time $k = 0$, $\mathbf{u} \in \mathbb{U}^\infty$ is the input data, and $\mathbf{y} \in \mathbb{Y}^\infty$ is the output data.

Remark 5. Since the regulator requires a state estimate to compute, and the input directly affects the output, the current state and disturbance estimates $(\hat{x}(k), \hat{d}(k))$ must be functions of past data, not including the current measurement $y(k)$. Therefore, at time $k = 0$, there is no data available to update the prior guess, and most estimator designs will take Φ_0 as the identity map, i.e.,

$$(\hat{x}(0), \hat{d}(0)) := \Phi_0(\overline{x}, \overline{d}) = (\overline{x}, \overline{d}).$$

However, we can also consider models without direct feedthrough effects (i.e., $y = h(x, d)$) in which case Definition 1 can be modified so the estimator functions also take $y(k)$ as an argument.

The estimator (15) is designed according to the model (2) and thus has no knowledge of the plant state x_{P} or plant disturbance w_{P}. To analyze its performance and state the assumptions needed to establish offset-free performance, we consider the following noise model:

$$x^+ = f(x, u, d) + w \tag{16a}$$
$$d^+ = d + w_d \tag{16b}$$
$$y = h(x, u, d) + v \tag{16c}$$

where $\tilde{w} := (w, w_d, v) \in \tilde{\mathbb{W}}(x, u, d) \subseteq \mathbb{R}^{n_{\tilde{w}}}$ are the process, disturbance, and measurement noises, $n_{\tilde{w}} := n + n_d + n_y$, and

$$\tilde{\mathbb{W}}(x, u, d) := \{\, (w, w_d, v) \mid (x^+, d^+, y) \in \mathbb{X} \times \mathbb{D} \times \mathbb{Y}, (16) \,\}$$

is a constraint set that ensures all quantities remain physical. We define the set of feasible trajectories by

$$\tilde{\mathbb{Z}}_e := \{\, (\mathbf{x}, \mathbf{u}, \mathbf{d}, \mathbf{y}, \tilde{\mathbf{w}}) \in \mathbb{X}^\infty \times \mathbb{U}^\infty \times \mathbb{D}^\infty \times \mathbb{Y}^\infty \times (\mathbb{R}^{n_{\tilde{w}}})^\infty \mid (16) \text{ and } \tilde{w} = (w, w_d, v) \in \tilde{\mathbb{W}}(x, u, d) \,\}.$$

Finally, denoting the state, disturbance, and errors by

$$e_x(k) := x(k) - \hat{x}(k), \qquad e_d(k) := d(k) - \hat{d}(k), \tag{17a}$$
$$e(k) := \begin{bmatrix} e_x(k) \\ e_d(k) \end{bmatrix}, \qquad \overline{e} := \begin{bmatrix} x(0) - \overline{x} \\ d(0) - \overline{d} \end{bmatrix}, \tag{17b}$$

we define robust stability of the estimator (15) as follows.

Definition 2. The estimator (15) is *robustly globally exponentially stable* (RGES) for the system (16) if there exist constants $c_{e,1}, c_{e,2} > 0$ and $\lambda_e \in (0, 1)$ such that

$$|e(k)| \leq c_{e,1} \lambda_e^k |\overline{e}| + c_{e,2} \sum_{j=1}^{k} \lambda_e^{j-1} |\tilde{w}(k - j)|$$

for each $k \in \mathbb{I}_{\geq 0}$, prior guess $(\overline{x}, \overline{d}) \in \mathbb{X} \times \mathbb{D}$, and trajectories $(\mathbf{x}, \mathbf{u}, \mathbf{d}, \mathbf{y}, \tilde{\mathbf{w}}) \in \tilde{\mathbb{Z}}_e$, given definitions (15) and (17).

For the case with plant-model mismatch, the estimator (15) is not only assumed to be RGES for the system (16), but is also assumed to admit a robust global Lyapunov function.

Assumption 6. The initial estimator Φ_0 is the identity map. There exists a function $V_e : \mathbb{X} \times \mathbb{D} \times \mathbb{X} \times \mathbb{D} \to \mathbb{R}_{\geq 0}$ and constants $c_1, c_2, c_3, c_4, \delta_w > 0$ such that

$$c_1 |e(k)|^2 \leq V_e(k) \leq c_2 |e(k)|^2 \tag{18a}$$
$$V_e(k + 1) \leq V_e(k) - c_3 |e(k)|^2 + c_4 |\tilde{w}(k)|^2 \tag{18b}$$

for all $(\overline{x}, \overline{d}) \in \mathbb{X} \times \mathbb{D}$, $(\mathbf{x}, \mathbf{u}, \mathbf{d}, \mathbf{y}, \tilde{\mathbf{w}}) \in \tilde{\mathbb{Z}}_e$, and $k \in \mathbb{I}_{\geq 0}$, where (15), (17), and $V_e(k) := V_e(x(k), d(k), \hat{x}(k), \hat{d}(k))$.

The following theorem establishes that Assumption 6 implies RGES of the estimator (15) for the system (16) (see Appendix A.1 for proof).

Theorem 1. *Suppose the estimator* (15) *for the system* (16) *satisfies Assumption* 6*. Then the estimator is RGES under Definition* 2*.*

Remark 6. In Assumption 6, we assume Φ_0 is the identity map, and therefore $e(0) = \overline{e}$. However, as mentioned in Remark 5, if we consider models without direct input-output effects (i.e., $y = \hat{h}(x, d)$), then the estimator functions Φ_k may become a function of the current output $y(k)$ and it is no longer reasonable to assume Φ_0 is the identity map. Then $e(0) \neq \overline{e}$ in general. However, we can modify Definition 1 to include robustness to the current noise $\tilde{n}(k)$, and we can modify Assumption 6 to include a linear bound of the form $|e(0)| \leq \overline{a}_1|\overline{e}| + \overline{a}_2|\tilde{w}(0)|$, for some $\overline{a}_1, \overline{a}_2 > 0$, to again imply RGES of the estimator.

While Assumption 6 is satisfied for stable full-order observers of (16),[1] we know of no nonlinear results that guarantee a Lyapunov function characterization of stability (i.e., Assumption 6) for the full information estimation (FIE) or moving horizon estimation (MHE) algorithms. FIE and MHE were shown to be RGES for exponentially detectable and stabilizable systems by Allan and Rawlings (2021), but they use a Q-function to demonstrate stability. To the best of our knowledge, the closest construction is the N-step Lyapunov function of Schiller et al. (2023). If we treat the disturbance as a parameter, rather than an uncontrollable integrator, there are FIE and MHE algorithms for combined state and parameter estimation that could also be used to estimate the states and disturbances (Muntwiler et al., 2023; Schiller and Müller, 2023).[2]

3 Robust stability for tracking and estimation

In this section, we consider stabilization of the system,[3]

$$\xi^+ = F(\xi, u, \omega), \qquad \omega \in \Omega(\xi, u). \tag{19}$$

The system (19) represents the evolution of an *extended* plant state $\xi \in \Xi \subseteq \mathbb{R}^{n_\xi}$ subject to the input $u \in \mathbb{U}$ and *extended* disturbance $\omega \in \Omega(\xi, u) \subseteq \mathbb{R}^{n_\omega}$ (to be defined). Greek letters are used for the extended state and disturbance (ξ, ω) to avoid confusion with the states and disturbances of (1), (2), and (16). Throughout, we assume Ξ is closed and $0 \in \Omega(\xi, u)$ and $F(\xi, u, \omega) \in \Xi$ for all $(\xi, u) \in \Xi \times \mathbb{U}$ and $\omega \in \Omega(\xi, u)$.

---

[1] A full-order state observer of (16) is a dynamical system, evolving in the same state space as (16), stabilized with respect to x by output feedback.

[2] The estimation algorithms of Muntwiler et al. (2023) produce RGES state estimates, but it is not shown the parameter estimates are RGES. The estimation algorithm of Schiller and Müller (2023) produces RGES state and parameter estimates, but only under a persistence of excitation condition.

[3] To ensure unphysical states are not produced by additive disturbances, we let the disturbance set be a function of the state and input. However, we can convert (19) to a standard form by taking $\xi^+ = \tilde{F}(\xi, u, \omega)$, $\omega \in \Omega$ where $\tilde{F}(\xi, u, \omega) = F(\xi, \mathrm{proj}_{\Omega(\xi,u)}(\omega))$, $\Omega := \bigcup_{(\xi,u)\in\Xi\times\mathbb{U}} \Omega(\xi, u)$, and $\mathrm{proj}_{\Omega(\xi,u)}(\omega) = \mathrm{argmin}_{\omega'\in\Omega(\xi,u)} |\omega - \omega'|$.

3.1 Robust stability with respect to two outputs

We first consider stabilization of (19) under state feedback,

$$\xi^+ = F_c(\xi, \omega), \qquad\qquad \omega \in \Omega_c(\xi) \tag{20}$$

where $\kappa : \Xi \to \mathbb{U}$ is the control law, $F_c(\xi, \omega) := F(\xi, \kappa(\xi), \omega)$, and $\Omega_c(\xi) := \Omega(\xi, \kappa(\xi))$. We define robust positive invariance for the system (20) as follows.

Definition 3 (Robust positive invariance). A closed set $X \subseteq \Xi$ is *robustly positive invariant* (RPI) for the system (20) if $\xi \in X$ and $\omega \in \Omega_c(\xi)$ imply $F_c(\xi, \omega) \in X$.

Robust target- and setpoint-tracking stability are defined under the umbrella of input-to-state stability (ISS) with respect to two measurement functions (Tran et al., 2015). We slightly modify their definition by considering measurement functions of (ξ, ω) (rather than just ξ) and structuring the measurement functions as norms of the outputs $\zeta_1 \in \mathbb{R}^{n_{\zeta_1}}$ and $\zeta_2 \in \mathbb{R}^{n_{\zeta_2}}$, where

$$\zeta_1 = G_1(\xi, \omega), \qquad\qquad \zeta_2 = G_2(\xi, \omega). \tag{21}$$

The definition of Tran et al. (2015) can be reconstructed by taking G_1 and G_2 as scalar-valued, positive semidefinite functions of ξ.

Definition 4 (Robust stability w.r.t. two outputs). We say the system (20) (with outputs (21)) is *robustly asymptotically stable* (RAS) (*on a RPI set* $X \subseteq \Xi$) *with respect to* (ζ_1, ζ_2) if there exist $\beta_\zeta \in \mathcal{KL}$ and $\gamma_\zeta \in \mathcal{K}$ such that

$$|\zeta_1(k)| \leq \beta_\zeta(|\zeta_2(0)|, k) + \gamma_\zeta(\|\boldsymbol{\omega}\|_{0:k}) \tag{22}$$

for each $k \in \mathbb{I}_{\geq 0}$ and trajectories $(\boldsymbol{\xi}, \boldsymbol{\omega}, \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2)$ satisfying (20), (21), and $\xi(0) \in X$. We say (20) is *robustly exponentially stable* (RES) w.r.t. (ζ_1, ζ_2) if it is RAS w.r.t. (ζ_1, ζ_2) with $\beta_\zeta(s, k) := c_\zeta \lambda_\zeta^k s$ for some $c_\zeta > 0$ and $\lambda_\zeta \in (0, 1)$.

For the nominal case (i.e., $\Omega(\xi, u) \equiv \{ 0 \}$), we drop the word *robust* from Definitions 3 and 4 and simply write *positive invariant*, *asymptotically stable* (AS), and *exponentially stable* (ES). Moreover, if (20) is RAS (RES) w.r.t. (ζ, ζ), where $\zeta = G(\xi, \omega)$, we simply say it is RAS (RES) w.r.t. ζ.

In Sections 4 and 5, we demonstrate nominal stability and robustness to estimate error, noise, and SSTP parameter changes. The following cases of the system (19), control law $u = \kappa(\xi)$, and outputs (21) are considered.

1. *Nominal stability*: Let $\xi := x$, $u = \kappa(\xi) := \kappa_N(x, \beta)$, $\omega := 0$, $\zeta_1 := g(u, h(x, u, d)) - r_{\mathrm{sp}}$, and $\zeta_2 := x - x_s(\beta)$. Then, for each *fixed* $\beta = (r_{\mathrm{sp}}, u_{\mathrm{sp}}, y_{\mathrm{sp}}, d) \in \mathcal{B}$, the closed-loop system has dynamics (20) and outputs (21) with

$$\begin{aligned} F(\xi, \omega) &:= f(x, \kappa_N(x, \beta), \beta) \\ G_1(\xi) &:= g(x, h(x, \kappa_N(x, \beta), d)) - r_{\mathrm{sp}} \\ G_2(\xi) &:= x - x_s(\beta) \end{aligned}$$

for each $\xi \in \mathcal{X}_N^\rho := \text{lev}_\rho V_N^0$ and $\omega = 0$. AS (ES) w.r.t. ζ_2 corresponds to (exponential) target-tracking stability, and AS (ES) w.r.t. (ζ_1, ζ_2) corresponds to (exponential) setpoint-tracking stability.

2. *Robust stability (w.r.t. estimate error, noise, SSTP parameter changes)*: Let $\xi := (\hat{x}, \hat{\beta})$, $\kappa(\xi) := \kappa_N(\xi)$, $\omega := (e, e^+, \Delta s_{\text{sp}}, \tilde{w})$, $\zeta_1 := r - r_{\text{sp}}$, $\zeta_2 := \hat{x} - x_s(\hat{\beta})$, where $r := g(u, h(\hat{x} + e_x, u, \hat{d} + e_d) + v)$ and $\hat{\beta} := (s_{\text{sp}}, \hat{d})$. Then the closed-loop system has dynamics (20) and outputs (21) with

$$F(\xi, \omega) := \begin{bmatrix} f(\hat{x} + e_x, \kappa_N(\hat{x}, \hat{\beta}), \hat{d} + e_d) + w - e_x^+ \\ s_{\text{sp}} + \Delta s_{\text{sp}} \\ \hat{d} + e_d + w_d - e_d^+ \end{bmatrix}$$
$$G_1(\xi) := g(x, h(\hat{x} + e_x, \kappa_N(\hat{x}, \hat{\beta}), \hat{d} + e_d) + v) - r_{\text{sp}},$$
$$G_2(\xi) := \hat{x} - x_s(\hat{\beta})$$

for each $\xi = (\hat{x}, \hat{\beta})$ in a to-be-defined RPI set $\hat{\mathcal{S}}_N^\rho$ and $\omega \in \Omega_c(\xi)$ (to be defined). RAS (RES) of (20) w.r.t. ζ_2 alone corresponds to robust (exponential) target-tracking stability, and RAS (RES) w.r.t. (ζ_1, ζ_2) corresponds to robust (exponential) setpoint-tracking stability.

Remark 7. If (20) is RAS on $X \subseteq \Xi$ w.r.t. (ζ_1, ζ_2), then $\omega(k) \to 0$ implies $\zeta_1(k) \to 0$ so long as $\xi(0) \in X$.

Remark 8. Definition 4 generalizes many ISS and input-to-output stability (IOS) definitions originally posed for continuous-time systems by Sontag and Wang (1995, 1999, 2000). However, only Definition 4 is suitable for analyzing both target- and setpoint-tracking performance of the offset-free MPC. ISS is not appropriate as the SSTP parameters β are often part of the extended state ξ. IOS and robust output stability allow the tracking performance to degrade with the magnitude of the SSTP parameters. While state-independent IOS (SIIOS) coincides with the special case of $\zeta = G_1(\xi) \equiv G_2(\xi)$ (e.g., for target-tracking), we find the setpoint-tracking error is more tightly bounded by the initial target-tracking error.

Next, we define an (exponential) ISS Lyapunov function with respect to the noise-free outputs

$$\zeta_1 = G_1(\xi), \qquad \zeta_2 = G_2(\xi) \tag{23}$$

and show its existence implies RAS (RES) of (20) with respect to (ζ_1, ζ_2) (see Appendix A.2 for proof).

Definition 5. Consider the system (20) with outputs (23). We call $V : \Xi \to \mathbb{R}_{\geq 0}$ an *ISS Lyapunov function* (*on a RPI set* $X \subseteq \Xi$) *with respect to* (ζ_1, ζ_2) if there exist $\alpha_i \in \mathcal{K}_\infty, i \in \mathbb{I}_{1:3}$ and $\sigma \in \mathcal{K}$ such that, for each $\xi \in X$ and $\omega \in \Omega_c(\xi)$,

$$\alpha_1(|G_1(\xi)|) \leq V(\xi) \leq \alpha_2(|G_2(\xi)|) \tag{24a}$$
$$V(F_c(\xi, \omega)) \leq V(\xi) - \alpha_3(V(\xi)) + \sigma(|\omega|). \tag{24b}$$

We say V is an *exponential ISS Lyapunov function* with respect to (ζ_1, ζ_2) if it is an ISS Lyapunov function with respect to (ζ_1, ζ_2) with $\alpha_i(\cdot) = a_i(\cdot)^b$ for some $a_i, b > 0, i \in \mathbb{I}_{1:3}$.

Theorem 2. *If the system* (20) *with outputs* (23) *admits an (exponential) ISS Lyapunov function* $V : \Xi \to \mathbb{R}_{\geq 0}$ *on an RPI set* $X \subseteq \Xi$ *with respect to* (ζ_1, ζ_2)*, then it is RAS (RES) on* X *with respect to* (ζ_1, ζ_2)*.*

Similarly to Definitions 3 and 4, we call V a *Lyapunov function* or *exponential Lyapunov function* w.r.t. (ζ_1, ζ_2) if it satisfies Definition 5 in the nominal case (i.e., $\Omega(\xi, u) \equiv \{ 0 \}$). Moreover, we note that the proof of Theorem 2 trivially extends to the nominal case by setting $\omega = 0$ throughout.

Remark 9. If $\zeta = G_1(\xi) \equiv G_2(\xi)$, then it suffices to replace (24b) with $V(F_c(\xi, \omega)) \leq V(\xi) - \tilde{\alpha}_3(|G_1(\xi)|) + \sigma(|\omega|)$ to establish ISS with respect to ζ, where $\tilde{\alpha}_3 \in \mathcal{K}_\infty$. Then (24b) holds with $\alpha_3 := \tilde{\alpha}_3 \circ \alpha_2^{-1}$.

3.2 Combined controller-estimator robust stability

In applications without plant-model mismatch, it suffices to consider RES of each of the controller and estimator subsystems to establish RES of the combined system. This is because the controller and estimator error systems are connected *sequentially*, with the target- and setpoint-tracking errors having no influence on the estimation errors. However, as we show in Section 6, plant-model mismatch makes this a *feedback interconnection*, with the tracking errors influencing the state estimate errors and vice versa. Therefore it is necessary to analyze stability of the combined system.

We define the *extended* sensor output $\upsilon \in \Upsilon \subseteq \mathbb{R}^{n_\upsilon}$ by

$$\upsilon = H(\xi, u, \omega). \tag{25}$$

Assume Υ is closed and $H(\xi, u, \omega) \in \Upsilon$ for all $(\xi, u) \in \Xi \times \mathbb{U}$ and $\omega \in \Omega(\xi, u)$. We consider the extended state estimator

$$\hat{\xi}(k) := \Phi_k^\xi(\overline{\xi}, \mathbf{u}_{0:k-1}, \boldsymbol{\upsilon}_{0:k-1}) \tag{26}$$

where $\overline{\xi} \in \hat{\Xi} \subseteq \mathbb{R}^{n_{\hat{\xi}}}$ is the prior guess and $\Phi_k^\xi : \hat{\Xi} \times \mathbb{U}^k \times \Upsilon^k \to \hat{\Xi}, k \in \mathbb{I}_{\geq 0}$. The set $\hat{\Xi}$ is closed but is not necessarily the same, let alone of the same dimension, as Ξ. We consider stabilization via state estimate feedback,

$$u = \hat{\kappa}(\hat{\xi}) \tag{27}$$

where $\hat{\kappa} : \hat{\Xi} \to \mathbb{U}$. Finally, we define a RPI set as follows.

Definition 6. A closed set $S \subseteq \Xi \times \hat{\Xi}$ is RPI for the system (19) and (25)–(27) if $(\xi(k), \hat{\xi}(k)) \in \mathcal{S}$ for all $k \in \mathbb{I}_{\geq 0}$ and $(\boldsymbol{\xi}, \mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\upsilon})$ satisfying (19), (25)–(27), and $(\xi(0), \overline{\xi}) \in \mathcal{S}$.

With plant-model mismatch, the *extended* plant and model states to evolve on different spaces. Thus, we define the estimator error $\varepsilon \in \mathbb{R}^{n_{\hat{\xi}}}$ as the deviation of the estimate $\hat{\xi}$ from an arbitrary function $G_\varepsilon : \Xi \to \hat{\Xi}$ of the state ξ,

$$\varepsilon(k) = G_\varepsilon(\xi(k)) - \hat{\xi}(k), \qquad \overline{\varepsilon} := G_\varepsilon(\xi(0)) - \overline{\xi}. \tag{28}$$

Finally, we define robust stability with respect to the outputs

$$\zeta_1 = G_1(\xi, \hat{\xi}, u, \omega), \qquad \zeta_2 = G_2(\xi, \hat{\xi}, u, \omega) \tag{29}$$

similarly to Definition 4.

Definition 7. The system (19) and (25)–(27) (with outputs (29)) is RAS in a RPI set $\mathcal{S} \subseteq \mathbb{X} \times \hat{\mathbb{X}}$ with respect to (ζ_1, ζ_2) if there exist functions $\beta_\zeta, \gamma_\zeta \in \mathcal{KL}$ such that

$$|(\zeta_1(k), \varepsilon(k))| \leq \beta_\zeta(|(\zeta_2(0), \overline{\varepsilon})|, k) + \sum_{i=0}^{k} \gamma_\zeta(|\omega(k-i)|, i) \tag{30}$$

for all $k \in \mathbb{I}_{\geq 0}$ and all trajectories $(\boldsymbol{\xi}, \mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\upsilon}, \boldsymbol{\varepsilon}, \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2)$ satisfying (19), (25)–(29), and $(\xi(0), \overline{\xi}) \in \mathcal{S}$. We say (19) and (25)–(27) is RES w.r.t. (ζ_1, ζ_2) if it is RAS w.r.t. (ζ_1, ζ_2) with $\beta_\zeta(s, k) := c_\zeta \lambda_\zeta^k s$ and $\gamma_\zeta(s, k) := \lambda_\zeta^k \sigma_\zeta(s)$ for some $c_\zeta > 0$, $\lambda_\zeta \in (0, 1)$, and $\sigma_\zeta \in \mathcal{K}$.

As in Section 3.1, we say (19) and (25)–(27) is RAS (RES) w.r.t. $\zeta = G(\xi, \omega)$ if it is RAS (RES) w.r.t. (ζ, ζ).

In Section 6, we establish robustness of offset-free MPC with plant-model mismatch in terms of Definition 7, using the following definition of the system (19) and (25)–(27), estimate errors (28), and outputs (29):

3. *With mismatch*: Let $\xi := (x_{\rm P}, \alpha)$, $\hat{\xi} := (\hat{x}, \hat{\beta})$, $u := \kappa_N(\hat{\xi})$, $\omega := (\Delta s_{\rm sp}, \Delta w_{\rm P})$, $\upsilon := (y, \Delta s_{\rm sp})$, $\varepsilon := (x_{\rm P} + \Delta x_s(\alpha), s_{\rm sp}, d_s(\alpha)) - \hat{\xi}$, $\zeta_1 := r - r_{\rm sp}$, $\zeta_2 := \hat{x} - x_s(\hat{\beta})$, where $r := g(u, h_{\rm P}(x, u, w_{\rm P}))$, $\alpha := (s_{\rm sp}, w_{\rm P})$, $\hat{\beta} := (s_{\rm sp}, \hat{d})$, and $(\Delta x_s(\alpha), d_s(\alpha))$ are to be defined. Then the closed-loop system has dynamics (19) and (25)–(27), errors (28), and outputs (29) with

$$F(\xi, u, \omega) := \begin{bmatrix} f_{\rm P}(x_{\rm P}, u, w_{\rm P}) \\ s_{\rm sp} + \Delta s_{\rm sp} \\ w_{\rm P} + \Delta w_{\rm P} \end{bmatrix}, \qquad H(\xi, u, \omega) := \begin{bmatrix} h_{\rm P}(\xi, u, w_{\rm P}) \\ \Delta s_{\rm sp} \end{bmatrix},$$

$$\Phi_k^\xi(\overline{\xi}, \mathbf{u}_{0:k-1}, \boldsymbol{\upsilon}_{0:k-1}) := (\hat{x}(k), s_{\rm sp}(k), \hat{d}(k)), \qquad G_\varepsilon(\xi) := \begin{bmatrix} x_{\rm P} + \Delta x_s(\alpha) \\ d_s(\alpha) \end{bmatrix}$$

$$G_1(\xi, u, \omega) := g(u, h_{\rm P}(x_{\rm P}, u, w_{\rm P})) - r_{\rm sp}, \qquad G_2(\hat{\xi}) := \hat{x} - x_s(\hat{\beta})$$

for each $(\xi, \hat{\xi}) = (x, \beta, \hat{x}, \hat{\beta})$ in a to-be-defined RPI set $\mathcal{S}_N^{\rho,\tau}$ and $\omega \in \Omega_c(\xi)$ (to be defined), where $(\hat{x}(k), \hat{d}(k)) := \Phi_k(\overline{x}, \overline{d}, \mathbf{u}_{0:k-1}, \mathbf{y}_{0:k-1})$ as in Definition 1.

As in Section 3.1, RAS (RES) w.r.t. ζ_2 corresponds to robust (exponential) target-tracking stability, and RAS (ES) w.r.t. (ζ_1, ζ_2) corresponds to robust (exponential) setpoint-tracking stability.

Remark 10. If (19) and (25)–(27) is RAS on a RPI set $\mathcal{S} \subseteq \Xi \times \hat{\Xi}$ w.r.t. (ζ_1, ζ_2), then $\omega(k) \to 0$ implies $(\zeta_1(k), \varepsilon(k)) \to 0$ so long as $(\xi(0), \overline{\xi}) \in \mathcal{S}$ (cf. (Allan and Rawlings, 2021, Prop. 3.11)).

To analyze stability of the system (19) and (25)–(27), we use the following theorem (see Appendix A.3 for proof).

Theorem 3. *Consider the system* (19) *and* (25)–(27) *with errors* (28) *and output* $\zeta = G(\hat{\xi})$*. Suppose* Φ_0^ξ *is the identity map and there exist constants* $a_i, b_i > 0, i \in \mathbb{I}_{1:4}$*, a RPI set* $\mathcal{S} \subseteq \mathbb{X} \times \hat{\mathbb{X}}$*, and functions* $V : \hat{\Xi} \to \mathbb{R}_{\geq 0}$*,* $V_\varepsilon : \Xi \times \hat{\Xi} \to \mathbb{R}_{\geq 0}$*, and* $\sigma, \sigma_\varepsilon \in \mathcal{K}$ *such that* $\frac{a_4 c_4}{a_3 c_1} < 1$*,* $\frac{a_4 c_4}{a_3 c_3} < \frac{c_1}{c_1 + c_2}$*, and, for all trajectories* $(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}, \mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\upsilon}, \boldsymbol{\varepsilon}, \boldsymbol{\zeta})$ *satisfying* (19) *and* (25)–(28)*,* $\zeta = G(\hat{\xi})$*, and* $(\xi(0), \overline{\xi}) \in \mathcal{S}$*, we also satisfy*

$$a_1|\zeta|^2 \leq V(\hat{\xi}) \leq a_2|\zeta|^2 \tag{31a}$$

$$V(\hat{\xi}^+) \leq V(\hat{\xi}) - a_3|\zeta|^2 + a_4|(\varepsilon, \varepsilon^+)|^2 + \sigma(|\omega|) \tag{31b}$$

$$c_1|\varepsilon|^2 \leq V_\varepsilon(\xi, \hat{\xi}) \leq c_2|\varepsilon|^2 \tag{31c}$$

$$V_\varepsilon(\xi^+, \hat{\xi}^+) \leq V_\varepsilon(\xi, \hat{\xi}) - c_3|\varepsilon|^2 + c_4|\zeta|^2 + \sigma_\varepsilon(|\omega|). \tag{31d}$$

Then the system (19) *and* (25)–(27) *is RES in* $\mathcal{S}$ *w.r.t.* ζ*.*

4 Nominal offset-free performance

In this section, we consider the application of offset-free MPC to the model (2) in the *nominal* case (i.e., without estimate errors or setpoint and disturbance changes). Consider the following *modeled* closed-loop system:

$$x^+ = f_c(x, \beta) := f(x, \kappa_N(x, \beta), d) \tag{32a}$$

$$y = h_c(x, \beta) := h(x, \kappa_N(x, \beta), d) \tag{32b}$$

$$r = g_c(x, \beta) := g(\kappa_N(x, \beta), h_c(x, \beta)) \tag{32c}$$

where $(x, \beta) := (x, r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, d) \in \mathcal{S}_N$. For each $\rho > 0$ and $\beta \in \mathcal{B}$, we define the candidate domain of stability

$$\mathcal{X}_N^\rho(\beta) := \text{lev}_\rho V_N^0(\cdot, \beta). \tag{33}$$

In the following theorem, we establish *nominal* stability and offset-free performance of the modeled closed-loop system (32), under Assumptions 1 to 5 and with constant, known setpoints $s_{\text{sp}} = (r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}})$ and disturbance d.

Theorem 4. *Suppose Assumptions* 1 *to* 5 *hold. Let* $\rho > 0$*.*

(a) For each compact $\mathcal{B}_c \subseteq \mathcal{B}$*, there exist constants* $a_1, a_2, a_3 > 0$ *such that*

$$a_1|x - x_s(\beta)|^2 \leq V_N^0(x, \beta) \leq a_2|x - x_s(\beta)|^2 \tag{34a}$$

$$V_N^0(f_c(x, \beta), \beta) \leq V_N^0(x, \beta) - a_3|x - x_s(\beta)|^2 \tag{34b}$$

for all $x \in \mathcal{X}_N^\rho(\beta)$ *and* $\beta \in \mathcal{B}_c$*.*

(b) For each $\beta \in \mathcal{B}$, the system (32a) *is ES on $\mathcal{X}_N^\rho(\beta)$ with respect to the target-tracking error $\delta x := x - x_s(\beta)$.*

(c) For each $\beta = (r_{\rm sp}, u_{\rm sp}, y_{\rm sp}, d) \in \mathcal{B}$, the system (32a) *is AS on $\mathcal{X}_N^\rho(\beta)$ with respect to $(\delta r, \delta x)$, where $\delta r := g_c(x, \beta) - r_{\rm sp}$ is the setpoint-tracking error.*

(d) If g and h are Lipschitz continuous on bounded sets, then part (c) can be upgraded to ES.

We include a proof of Theorem 4 in Appendix B.1. Two details of the proof are required for the subsequent results. First, from (Rawlings et al., 2020, Prop. 2.4), we have

$$V_N(f_c(x,\beta), \tilde{\mathbf{u}}(x,\beta), \beta) \le V_N^0(x,\beta) - \ell(x, \kappa_N(x,\beta), \beta) \tag{35}$$

for all $(x, \beta) \in \mathcal{S}_N$, where

$$\tilde{\mathbf{u}}(x,\beta) := (u^0(1; x, \beta), \dots, u^0(N-1; x, \beta), \kappa_f(x^0(N; x, \beta), \beta)) \tag{36}$$

is a suboptimal sequence for $x^+ := f_c(x, \beta)$. Second, for each $(x, \beta) \in \mathcal{S}_N$, the suboptimal sequence $\tilde{\mathbf{u}}(x,\beta)$ steers the system from $f_c(x,\beta)$ to the terminal constraint $\mathbb{X}_f(\beta)$ in $N-1$ moves and keeps it there (by Assumption 4). Therefore $\tilde{\mathbf{u}}(x,\beta) \in \mathcal{U}_N(f_c(x,\beta), \beta)$ and $f_c(x,\beta) \in \mathcal{X}_N(\beta)$.

Remark 11. Theorem 4(a) provides Lyapunov bounds that are *uniform* in the SSTP parameters β on compact subsets $\mathcal{B}_c \subseteq \mathcal{B}$. This implies a guaranteed decay rate $\lambda \in (0, 1)$ for the deviation of the state from its target $x - x_s(\beta)$, although this guaranteed rate may become arbitrarily close to 1 as we expand the size of the compact set $\mathcal{B}_c$.

5 Offset-free performance without mismatch

In this section, we show offset-free MPC (without plant-model mismatch) is robust to estimate errors and setpoint and disturbance changes. We assume the actual plant evolves according to the noisy model equations (16). We assume the setpoints evolve according to

$$s_{\rm sp}^+ = s_{\rm sp} + \Delta s_{\rm sp} \tag{37}$$

where $s_{\rm sp} := (r_{\rm sp}, u_{\rm sp}, y_{\rm sp})$ and $\Delta s_{\rm sp} := (\Delta r_{\rm sp}, \Delta u_{\rm sp}, \Delta y_{\rm sp})$. At each time, we define $\beta := (s_{\rm sp}, d)$ and $\Delta\beta := (\Delta s_{\rm sp}, w_d)$ and sometimes write $\beta^+ = \beta + \Delta\beta$. Taking the approach of (Rawlings et al., 2020, Sec. 4.6), the estimate error system evolves as

$$\hat{x}^+ = f(\hat{x} + e_x, u, \hat{d} + e_d) + w - e_x^+ \tag{38a}$$

$$\hat{d}^+ = \hat{d} + e_d + w_d - e_d^+ \tag{38b}$$

$$y = h(\hat{x} + e_x, u, \hat{d} + e_d) + v. \tag{38c}$$

We lump the perturbation terms from (37) and (38) into a single disturbance variable, defined as $\tilde{d} := (e, e^+, \Delta s_{\rm sp}, \tilde{w})$. To ensure the noise does not result in unphysical states, disturbances, or measurements, we define the set of admissible perturbations as

$$\tilde{\mathbb{D}}(\hat{x}, u, \hat{d}) := \{\ \tilde{d} = (e_x, e_d, e_x^+, e_d^+, \Delta s_{\rm sp}, \tilde{w}) \mid (38), \\ (\hat{x}^+, \hat{d}^+) \in \mathbb{X} \times \mathbb{D}, \tilde{w} \in \tilde{\mathbb{W}}(\hat{x} + e_x, u, \hat{d} + e_d)\ \}$$

for each $(\hat{x}, u, \hat{d}) \in \mathbb{X} \times \mathbb{U} \times \mathbb{D}$. The closed-loop estimate error system, defined by (6), (14), (15), (37), and (38), evolves as

$$\hat{x}^+ = \hat{f}_c(\hat{x}, \hat{\beta}, \tilde{d}) := f(\hat{x} + e_x, \kappa_N(\hat{x}, \hat{\beta}), \hat{d} + e_d) + w - e_x^+ \tag{39a}$$

$$\hat{\beta}^+ = \hat{f}_{\beta,c}(\hat{\beta}, \tilde{d}) := \begin{bmatrix} s_{\rm sp} + \Delta s_{\rm sp} \\ \hat{d} + e_d + w_d - e_d^+ \end{bmatrix} \tag{39b}$$

$$y = \hat{h}_c(\hat{x}, \hat{\beta}, \tilde{d}) := h(\hat{x} + e_x, \kappa_N(\hat{x}, \hat{\beta}), \hat{d} + e_d) + v$$

$$r = \hat{g}_c(\hat{x}, \hat{\beta}, \tilde{d}) := g(\kappa_N(\hat{x}, \hat{\beta}), h_c(\hat{x}, \hat{\beta}, \tilde{d}))$$

where $\hat{\beta} := (s_{\rm sp}, \hat{d})$.

5.1 Steady-state target problem assumptions

Even with bounds on the estimate errors and setpoint and disturbance changes, there are no guarantees the SSTP (6) is feasible at all times. Moreover, there is no guarantee the SSTP solutions themselves are robust to disturbance estimate errors. To guarantee robust feasibility of the SSTP (6) and robustness of the targets themselves, we make the following assumption.

Assumption 7. There exists a compact set $\mathcal{B}_c \subseteq \mathcal{B}$ and constant $\delta_0 > 0$ such that

(i) $\hat{\mathcal{B}}_c := \{\ (s, \hat{d}) \mid (s, d) \in \mathcal{B}_c,\ |e_d| \leq \delta_0,\ \hat{d} := d - e_d \in \mathbb{D}\ \} \subseteq \mathcal{B}$; and

(ii) z_s is continuous on $\hat{\mathcal{B}}_c$.

Assumption 7(i) guarantees robust feasibility of the SSTP so long as $\boldsymbol{\beta} \in \mathcal{B}_c^\infty$ and $\|\mathbf{e}_d\| \leq \delta_0$. Whenever Assumption 7(i) is satisfied, it is convenient to define

$$\tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) := \{\ \tilde{d} \in \tilde{\mathbb{D}}(\hat{x}, \kappa_N(\hat{x}, \hat{\beta}), \hat{\beta}) \mid \hat{f}_{\beta,c}(\hat{\beta}, \tilde{d}) \in \hat{\mathcal{B}}_c\ \}$$

for each $(\hat{x}, \hat{\beta}) \in \mathcal{S}_N$. As long as the disturbance always lies in $\tilde{\mathbb{D}}(\hat{x}, \hat{\beta})$, the SSTP is feasible at all times.

In the following lemma, we show Assumption 7 holds for some $\mathcal{B}_c = \delta\mathbb{B}^{n_\beta}$ when a rank condition is satisfied by the system linearized at the origin (see Appendix C for proof).

Lemma 1. *Suppose Assumptions* 1 *and* 2 *hold, the sets* $\mathbb{X}, \mathbb{U}, \mathbb{D}$ *contain neighborhoods of the origin, the functions* f, g, h, ℓ_s *are twice continuously differentiable,* $\ell_s(0,0) = 0$, $\partial_{(u,y)}\ell_s(0,0) = 0$, $\partial^2_{(u,y)}\ell_s(0,0)$ *is positive definite,*

$$M_1 := \begin{bmatrix} A - I & B \\ H_y C & H_y D + H_u \end{bmatrix} \tag{40a}$$

is full row rank, and (A, C) *is detectable, where*

$$A := \partial_x f(0,0,0), \qquad B := \partial_u f(0,0,0), \tag{40b}$$
$$C := \partial_x h(0,0,0), \qquad D := \partial_u h(0,0,0), \tag{40c}$$
$$H_y := \partial_y g(0,0), \qquad H_u := \partial_u g(0,0). \tag{40d}$$

Then there exists a compact set $\mathcal{B}_c \subseteq \mathcal{B}$ *and a function* $z_s : \mathcal{B} \to \mathbb{X} \times \mathbb{U}$ *satisfying all parts of Assumption* 7*. Moreover,* $z_s(\beta)$ *uniquely solves* (6) *for all* $\beta \in \hat{\mathcal{B}}_c$*.*

5.2 Robust stability

In Proposition 1, we establish recursive feasibility of the FHOCP given feasibility of the SSTP at each time for sufficiently small $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta})$. For brevity, we defer the proof to Appendix B.2. However, we sketch the proof as follows. First, we show the suboptimal input sequence $\tilde{\mathbf{u}}(x, \hat{\beta})$ is recursively feasible. Second, we establish a cost decrease of the form

$$V_N(\hat{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{\beta}^+) \leq V_N^0(\hat{x}, \hat{\beta}) - a_3|\delta\hat{x}|^2 + \sigma_r(|\tilde{d}|) \tag{41}$$

where $a_3 > 0$, $\sigma_r \in \mathcal{K}_\infty$, and $\delta\hat{x} := \hat{x} - x_s(\hat{\beta})$ is the target-tracking error. Third, we use this cost decrease to show the FHOCP is recursively feasible.

Proposition 1. *Suppose Assumptions* 1 *to* 5 *and* 7 *hold and let* $\rho > 0$*. There exists* $\sigma_r \in \mathcal{K}_\infty$ *and* $a_3, \delta > 0$ *such that*

(a) $\tilde{\mathbf{u}}(\hat{x}, \hat{\beta}) \in \mathcal{U}_N(\hat{x}^+, \hat{\beta}^+)$*,*

(b) (41) *holds, and*

(c) $\hat{x}^+ \in \mathcal{X}_N^\rho(\hat{\beta}^+)$*,*

for all $\hat{\beta} \in \hat{\mathcal{B}}_c$*,* $\hat{x} \in \mathcal{X}_N^\rho(\hat{\beta})$ *and* $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$*, where* $(\hat{x}^+, \hat{\beta}^+)$ *are defined as in* (39)*.*

Finally, we present the main result of this section.

Theorem 5. *Suppose Assumptions* 1 *to* 5 *and* 7 *hold and let* $\rho > 0$*. There exists* $\delta > 0$ *such that*

(a) the following set is RPI for the closed-loop system (39) *with disturbance* $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$*:*

$$\hat{\mathcal{S}}_N^\rho := \{\, (\hat{x}, \hat{\beta}) \in \mathcal{S}_N \mid \hat{x} \in \mathcal{X}_N^\rho(\hat{\beta}), \hat{\beta} \in \hat{\mathcal{B}}_c \,\}; \tag{42}$$

(b) there exist $a_i > 0, i \in \mathbb{I}_{1:3}$ *and* $\sigma_r \in \mathcal{K}_\infty$ *such that*

$$a_1|\delta\hat{x}|^2 \leq V_N^0(\hat{x}, \hat{\beta}) \leq a_2|\delta\hat{x}|^2 \tag{43a}$$
$$V_N^0(\hat{x}^+, \hat{\beta}^+) \leq V_N^0(\hat{x}, \hat{\beta}) - a_3|\delta\hat{x}|^2 + \sigma_r(|\tilde{d}|) \tag{43b}$$

for all $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N^\rho$ *and* $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$*, given* (39) *and* $\delta\hat{x} := \hat{x} - x_s(\hat{\beta})$*;*

(c) *the closed-loop system* (39) *with disturbance* $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$ *is RES on* $\hat{\mathcal{S}}_N^\rho$ *with respect to the target-tracking error* $\delta\hat{x} := \hat{x} - x_s(\hat{\beta})$*;*

(d) *the closed-loop system* (39) *with disturbance* $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$ *is RAS on* $\hat{\mathcal{S}}_N^\rho$ *with respect to* $(\delta r, \delta\hat{x})$*, where* $\delta r := \hat{g}_c(\hat{x}, \hat{\beta}, \tilde{d}) - r_{\rm sp}$ *is the setpoint-tracking error and* $\hat{\beta} = (r_{\rm sp}, u_{\rm sp}, y_{\rm sp}, \hat{d})$*; and*

(e) *if* g *and* h *are Lipschitz continuous on bounded sets, then part (d) can be upgraded to RES.*

To prove Theorem 5(d,e), we require the following proposition (see Appendix B.3 for proof).

Proposition 2. *Let Assumptions* 1 *to* 5 *hold,* $\rho, \delta > 0$*, and* $\mathcal{B}_c \subseteq \mathcal{B}$ *be compact. There exist* $\sigma_r, \sigma_g \in \mathcal{K}_\infty$ *such that*

$$|g_c(\hat{x}, \hat{\beta}) - r_{\rm sp}| \leq \sigma_r(|\hat{x} - x_s(\hat{\beta})|) \tag{44a}$$

$$|\hat{g}_c(\hat{x}, \hat{\beta}, \tilde{d}) - r_{\rm sp}| \leq |g_c(\hat{x}, \hat{\beta}) - r_{\rm sp}| + \sigma_g(|\tilde{d}|) \tag{44b}$$

for all $\hat{x} \in \mathcal{X}_N^\rho(\beta)$*,* $\hat{\beta} = (r_{\rm sp}, z_{\rm sp}, d) \in \mathcal{B}_c$*, and* $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$*. If* g *and* h *are Lipschitz on bounded sets, then we can take* $\sigma_r(\cdot) := c_r(\cdot)$ *and* $\sigma_g(\cdot) := c_g(\cdot)$ *for some* $c_r, c_g > 0$*.*

Proof of Theorem 5. (a)—If $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N$ and $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta})$, then $\hat{\beta}^+ := \hat{f}_{\beta,c}(\hat{\beta}, \tilde{d}) \in \hat{\mathcal{B}}_c$ by construction of $\tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta})$, and by Proposition 1(c), there exists $\delta > 0$ such that $\hat{x}^+ := \hat{f}_c(\hat{x}, \hat{\beta}, \tilde{d}) \in \mathcal{X}_N^\rho(\hat{\beta})$ so long as $|\tilde{d}| \leq \delta$.

(b)—Theorem 4 gives (43a), and Proposition 1(a,b) and the principle of optimality give (43b).

(c)—This follows from part (b) due to Theorem 2.

(d)—Let $(\hat{\mathbf{x}}, \hat{\boldsymbol{\beta}}, \tilde{\mathbf{d}}, \mathbf{r})$ satisfy (39), $(\hat{x}(0), \hat{\beta}(0)) \in \hat{\mathcal{S}}_N^\rho$, $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$, and $r = \hat{g}_c(\hat{x}, \hat{\beta}, \tilde{d})$. Define $\delta r := r - r_{\rm sp}$ and $\delta\hat{r} = g_c(\hat{x}, \hat{\beta}) - r_{\rm sp}$ where $\hat{\beta} = (r_{\rm sp}, z_{\rm sp}, \hat{d})$. Then

$$\alpha_1(|\delta\hat{r}|) := a_1[\sigma_r^{-1}(|\delta\hat{r}|)]^2 \leq a_1|\delta\hat{x}|^2 \leq V_N^0(\hat{x}, \hat{\beta})$$

by Proposition 2 and part (b). Moreover, V_N^0 is an ISS Lyapunov function on $\hat{\mathcal{S}}_N^\rho$ with respect to $(\delta\hat{r}, \delta\hat{x})$, and RAS on $\hat{\mathcal{S}}_N^\rho$ with respect to $(\delta\hat{r}, \delta\hat{x})$ follows by Theorem 2. Then RAS w.r.t. $(\delta\hat{r}, \delta\hat{x})$ and (Rawlings and Ji, 2012, Eq. (1)) gives

$$\begin{aligned}
|\delta r(k)| &\leq \sigma_r(|\delta\hat{r}(k)|) + \sigma_g(|\tilde{d}(k)|) \\
&\leq \sigma_r(c\lambda^k|\delta\hat{x}(0)| + \gamma(\|\tilde{\mathbf{d}}\|_{0:k-1})) + \sigma_g(|\tilde{d}(k)|) \\
&\leq \sigma_r(2c\lambda^k|\delta\hat{x}(0)|) + \sigma_r(2\gamma(\|\tilde{\mathbf{d}}\|_{0:k-1})) + \sigma_g(|\tilde{d}(k)|) \\
&\leq \sigma_r(2c\lambda^k|\delta\hat{x}(0)|) + (\sigma_r \circ 2\gamma + \sigma_g)(\|\tilde{\mathbf{d}}\|_{0:k}) \\
&=: \beta_r(|\delta\hat{x}(0)|, k) + \gamma_r(\|\tilde{\mathbf{d}}\|_{0:k})
\end{aligned} \tag{45}$$

for all $k \in \mathbb{I}_{\geq 0}$ and some $c > 0$, $\lambda \in (0, 1)$, and $\gamma \in \mathcal{K}$.

(e)—If g and h are Lipschitz continuous on bounded sets, then by Proposition 2, we can repeat part (d) with $\sigma_r(\cdot) := c_r(\cdot)$ and some $c_r > 0$. □

6 Offset-free MPC under mismatch

In this section, we show offset-free MPC, *despite (sufficiently small) plant-model mismatch*, is robust to setpoint and disturbance changes. We consider the plant (1), setpoint dynamics (37), and plant disturbance dynamics

$$w_{\mathrm{P}}^{+} = w_{\mathrm{P}} + \Delta w_{\mathrm{P}}. \tag{46}$$

With $\alpha := (s_{\mathrm{sp}}, w_{\mathrm{P}})$ and $\Delta\alpha := (\Delta s_{\mathrm{sp}}, \Delta w_{\mathrm{P}})$, we have the relationship $\alpha^{+} = \alpha + \Delta\alpha$. The SSTP and regulator are designed with the model (2), and the estimator is designed with the noisy model (16).

6.1 Target selection under mismatch

With plant-model mismatch, the connection between the steady-state targets and plant steady states becomes more complicated. To guarantee there is a plant steady state providing offset-free performance and that we can align the plant and model steady states using the disturbance estimate, we make the following assumptions about the SSTP.

Assumption 8. There exist compact sets $\mathcal{A}_c \subseteq \mathbb{R}^{n_r} \times \overline{\mathbb{Z}}_y \times \mathbb{W}$ and $\mathcal{B}_c \subseteq \mathcal{B}$ containing the origin, continuous functions $(x_{\mathrm{P,s}}, d_s) : \mathcal{A}_c \to \mathbb{X} \times \mathbb{D}$, and a constant $\delta_0 > 0$ for which

(a) $\hat{\mathcal{B}}_c$ (as defined in Assumption 7) is contained in $\mathcal{B}$;

(b) z_s is Lipschitz continuous on $\hat{\mathcal{B}}_c$;

(c) for each $\alpha = (s_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c$, the pair $(x_{\mathrm{P,s}}, d_s) = (x_{\mathrm{P,s}}(\alpha), d_s(\alpha))$ is the unique solution to

$$x_{\mathrm{P,s}} = f_{\mathrm{P}}(x_{\mathrm{P,s}}, u_s(s_{\mathrm{sp}}, d_s), w_{\mathrm{P}}) \tag{47a}$$
$$y_s(s_{\mathrm{sp}}, d_s) := h_{\mathrm{P}}(x_{\mathrm{P,s}}, u_s(s_{\mathrm{sp}}, d_s), w_{\mathrm{P}}) \tag{47b}$$

where $y_s(s_{\mathrm{sp}}, d_s) := h(x_s(s_{\mathrm{sp}}, d_s), u_s(s_{\mathrm{sp}}, d_s), d_s)$;

(d) $(s_{\mathrm{sp}}, d_s(s_{\mathrm{sp}}, w_{\mathrm{P}})) \in \mathcal{B}_c$ for all $(s_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c$; and

(e) $(s_{\mathrm{sp}}, 0) \in \mathcal{A}_c$ for all $(s_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c$.

For each $\alpha = (s_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c$, Assumption 8 guarantees there is a unique *model* disturbance $d_s(\alpha)$ to estimate and the SSTP (6) is robustly feasible at $\beta = (s_{\mathrm{sp}}, d_s(\alpha))$. Of course, the system cannot be stabilized for unbounded plant-model mismatch. Given Assumption 8, we define

$$\mathcal{A}_c(\delta_w) := \{\, (s_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c \mid |w_{\mathrm{P}}| \leq \delta_w \,\}$$
$$\mathbb{A}_c(\alpha, \delta_w) := \{\, \Delta\alpha \in \mathbb{R}^{n_\alpha} \mid \alpha + \Delta\alpha \in \mathcal{A}_c(\delta_w) \,\}.$$

Then $\mathcal{A}_c(\delta_w)$ is RPI for the system $\alpha^{+} = \alpha + \Delta\alpha, \Delta\alpha \in \mathbb{A}(\alpha, \delta_w)$, and if $\|\mathbf{e}_d\| \leq \delta_0$, then $\hat{\beta} = (s_{\mathrm{sp}}, d_s(\alpha) - e_d) \in \hat{\mathcal{B}}_c$ and the SSTP is feasible at all times.

Assumption 8 can be verified through a linearization analysis that is similar to the standard linear offset-free conditions (Muske and Badgwell, 2002; Pannocchia and Rawlings, 2003) (see Appendix C for proof).

Lemma 2. *Suppose the conditions of Lemma* 1 *hold,* $f_{\mathrm{P}}, h_{\mathrm{P}}$ *are twice continuously differentiable, and*

$$M_2 := \begin{bmatrix} A - I & B_d \\ C & C_d \end{bmatrix} \tag{48}$$

is invertible, given the definitions (40)*,* $B_d := \partial_d f(0,0,0)$*, and* $C_d := \partial_d h(0,0,0)$*. Then there exist compact sets* $\mathcal{A}_c \subseteq \mathbb{R}^{n_r} \times \overline{\mathbb{Z}}_y \times \mathbb{W}$ *and* $\mathcal{B}_c \subseteq \mathcal{B}$ *containing neighborhoods of the origin and functions* $z_s : \mathcal{B} \to \mathbb{X} \times \mathbb{U}$ *and* $(x_{\mathrm{P,s}}, d_s) : \mathcal{A}_c \to \mathbb{X} \times \mathbb{D}$ *satisfying all parts of Assumption* 8*. Moreover,* $z_s(\beta)$ *and* $(x_{\mathrm{P,s}}(\alpha), d_s(\alpha))$ *are the unique solutions to* (6) *and* (47) *for all* $\alpha = (s_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c$ *and* $\beta := (s_{\mathrm{sp}}, d_s(\alpha))$*.*

6.2 State estimation and regulation under mismatch

Given Assumption 8, we can define a "true" model state as $x := x_{\mathrm{P}} - \Delta x_s(\alpha)$ where $\Delta x_s := x_{\mathrm{P,s}}(\alpha) - x_s(s_{\mathrm{sp}}, d_s(\alpha))$ and $\alpha = (s_{\mathrm{sp}}, w_{\mathrm{P}})$. Then the plant (1) can be rewritten in terms of the model state x as

$$x^+ = f_{\mathrm{P}}(x + \Delta x_s(\alpha), u, w_{\mathrm{P}}) - \Delta x_s(\alpha^+) \tag{49a}$$

$$y = h_{\mathrm{P}}(x + \Delta x_s(\alpha), u, w_{\mathrm{P}}). \tag{49b}$$

Alternatively, the plant can be written as (16), where

$$w := f_{\mathrm{P}}(x + \Delta x_s(\alpha), u, w_{\mathrm{P}}) - f(x, u, d_s(\alpha)) - \Delta x_s(\alpha^+) \tag{50a}$$

$$w_d := d_s(\alpha^+) - d_s(\alpha) \tag{50b}$$

$$v := h_{\mathrm{P}}(x + \Delta x_s(\alpha), u, w_{\mathrm{P}}) - h(x, u, d_s(\alpha)). \tag{50c}$$

Clearly $\tilde{w} := (w, w_d, v) \in \mathbb{W}(x, u, d)$ by construction. Under Assumption 6, the state and disturbance estimator (15) is RGES for the constructed model state x and noise vector $\tilde{w}$.

The noise vector $\tilde{w}$ is still a function of the model state x, input u, and steady-state parameters α. Therefore, we bound it by more manageable variables, i.e., the tracking error $z - z_s(\beta)$, estimate errors e, plant disturbance w_{P}, and changes to the plant steady-state parameters $\Delta\alpha$. To this end, the following differentiability assumption is required.

Assumption 9. The derivatives $\partial_{(x,u)} f_{\mathrm{P}}$ and $\partial_{(x,u)} h_{\mathrm{P}}$ exist and are continuous on $\mathbb{X} \times \mathbb{U} \times \mathbb{W}$. The functions f, h and g are continuously differentiable on $\mathbb{X} \times \mathbb{U} \times \mathbb{D}$ and $\mathbb{U} \times \mathbb{Y}$.

Remark 12. Assumption 9 implies f, h are Lipschitz continuous on bounded sets.

Consider the closed-loop system

$$x^+ = f_{\mathrm{P}}(x + \Delta x_s(\alpha), \kappa_N(\hat{x}, \hat{\beta}), w_{\mathrm{P}}) - \Delta x_s(\alpha^+) \tag{51a}$$

$$\alpha^+ = \alpha + \Delta\alpha \tag{51b}$$

$$y = h_{\mathrm{P}}(x + \Delta x_s(\alpha), \kappa_N(\hat{x}, \hat{\beta}), w_{\mathrm{P}}). \tag{51c}$$

In the following propositions, we establish cost decreases for estimator and regulator Lyapunov functions for (51) (see Appendices B.4 and B.5 for proofs).

Proposition 3. *Suppose Assumptions* 1 *to* 6*,* 8 *and* 9 *hold. Let* $\rho > 0$*. There exist* $\hat{c}_3, \delta_w > 0$ *and* $\hat{\sigma}_w, \hat{\sigma}_\alpha \in \mathcal{K}_\infty$ *such that*

$$V_e^+ \leq V_e - \hat{c}_3|e|^2 + \hat{\sigma}_w(|w_{\rm P}|)|\delta\hat{x}|^2 + \hat{\sigma}_\alpha(|\Delta\alpha|) \tag{52}$$

so long as $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N^\rho$*,* $x \in \mathbb{X}$*,* $\alpha = (s_{\rm sp}, w_{\rm P}) \in \mathcal{A}_c(\delta_w)$*,* $\Delta\alpha = (\Delta s_{\rm sp}, \Delta w_{\rm P}) \in \mathbb{A}_c(\alpha, \delta_w)$*, and* $|e|, |e^+| \leq \delta_0$*, where* $V_e(k) := V_e(x(k), d_s(\alpha(k)), \hat{x}(k), \hat{d}(k))$*,* (17)*,* (50)*, and* (51)*.*

Proposition 4. *Let Assumptions* 1 *to* 5*,* 8 *and* 9 *hold and* $\rho > 0$*. There exist* $\tilde{a}_3, \tilde{a}_4, \delta, \delta_w > 0$ *and* $\tilde{\sigma}_\alpha \in \mathcal{K}_\infty$ *such that*

$$V_N^0(\hat{x}^+, \hat{\beta}^+) \leq V_N^0(\hat{x}, \hat{\beta}) - \tilde{a}_3|\delta\hat{x}|^2 + \tilde{a}_4|(e, e^+)|^2 + \sigma_\alpha(|\Delta\alpha|) \tag{53}$$

so long as $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N^\rho$*,* $x \in \mathbb{X}$*,* $\alpha = (s_{\rm sp}, w_{\rm P}) \in \mathcal{A}_c(\delta_w)$*,* $\Delta\alpha = (\Delta s_{\rm sp}, \Delta w_{\rm P}) \in \mathbb{A}_c(\alpha, \delta_w)$*, and* $|\tilde{d}| \leq \delta$*, where* $\tilde{d} := (e, e^+, \Delta s_{\rm sp}, \tilde{w})$*,* (17)*,* (50)*, and* (51)*.*

6.3 Main result

Finally, we state the main result of this section.

Theorem 6. *Suppose Assumptions* 1 *to* 6*,* 8 *and* 9 *hold and let* $\rho > 0$*. There exists* $\tau, \delta_w, \delta_\alpha > 0$ *such that, with*

$$\mathcal{S}_N^{\rho,\tau} := \{\, (x, \alpha, \hat{x}, \hat{\beta}) \in \mathbb{X} \times \mathcal{A}_c \times \hat{\mathcal{S}}_N^\rho \mid V_e(x, d_s(\alpha), \hat{x}, \hat{d}) \leq \tau,\ \alpha = (s_{\rm sp}, w_{\rm P}),\ \hat{\beta} = (s_{\rm sp}, \hat{d}) \,\}$$

the following statements hold:

- *(a)* $\mathcal{S}_N^{\rho,\tau}$ *is RPI for the closed-loop system* (15) *and* (51) *with the disturbance* $\Delta\alpha \in \mathbb{A}_c(\alpha, \delta_w) \cap \delta_\alpha\mathbb{B}^{n_\alpha}$*;*
- *(b) the closed-loop system* (15) *and* (51) *with the disturbance* $\Delta\alpha \in \mathbb{A}_c(\alpha, \delta_w) \cap \delta_\alpha\mathbb{B}^{n_\alpha}$ *is RES on* $\mathcal{S}_N^{\rho,\tau}$ *with respect to the target-tracking error* $\delta\hat{x} := \hat{x} - x_s(\hat{\beta})$*; and*
- *(c) the closed-loop system* (15) *and* (51) *with the disturbance* $\Delta\alpha \in \mathbb{A}_c(\alpha, \delta_w) \cap \delta_\alpha\mathbb{B}^{n_\alpha}$ *is RES on* $\mathcal{S}_N^{\rho,\tau}$ *with respect to* $(\delta r, \delta\hat{x})$*, where* $\delta r := r - r_{\rm sp}$ *is the setpoint-tracking error,* $\alpha = (r_{\rm sp}, u_{\rm sp}, y_{\rm sp}, w_{\rm P})$*,* $r = g(\kappa_N(\hat{x}, \hat{\beta}), y)$*, and* (51c)*.*

Proof. (a)—We already have that $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N^\rho$ and $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_d}$ implies $(\hat{x}^+, \hat{\beta}^+) \in \hat{\mathcal{S}}_N^\rho$ for some $\delta > 0$. To keep the trajectory of $(x, \alpha, \hat{x}, \hat{\beta})$ in $\mathcal{S}_N^{\rho,\tau}$ at all times, it suffices to show there exist $\tau, \delta_w, \delta_\alpha > 0$ such that $\alpha \in \mathcal{A}_c(\delta_w)$, $\Delta\alpha \in \mathbb{A}_c(\alpha, \delta_w) \cap \delta_\alpha\mathbb{B}^{n_\alpha}$, and $V_e := V_e(x, d_s(\alpha), \hat{x}, \hat{d}) \leq \tau$ implies $V_e^+ := V_e(x^+, \hat{x}^+) \leq \tau$ and $|(e, e^+, w)| \leq \delta$.

By Propositions 3 and 10 (in Appendix B.5), there exist constants $\hat{c}_3, \tilde{c}_e, \delta_w > 0$ and functions $\hat{\sigma}_w, \hat{\sigma}_\alpha, \tilde{\sigma}_w, \tilde{\sigma}_\alpha \in \mathcal{K}_\infty$ satisfying (52) and

$$|\tilde{d}|^2 \leq \tilde{c}_e|(e, e^+)|^2 + \tilde{\sigma}_w(|w_{\rm P}|)|\delta\hat{x}|^2 + \tilde{\sigma}_\alpha(|\Delta\alpha|) \tag{54}$$

so long as $\alpha = (s_{\rm sp}, w_{\rm P}) \in \mathcal{A}_c(\delta_w)$ and $\Delta\alpha \in \mathbb{A}_c(\alpha, \delta_w)$.

Assume, without loss of generality, that $\delta_w < (\frac{4c_2\tilde{c}_3}{a_1c_1\hat{c}_3}\hat{\sigma}_w + \tilde{\sigma}_w)^{-1}(\frac{a_1\delta^2}{\rho})$, which implies $\frac{2c_2\hat{\sigma}_w(\delta_w)\rho}{a_1\hat{c}_3} < \left(\delta^2 - \frac{\tilde{\sigma}_w(\delta_w)\rho}{a_1}\right)\frac{c_1}{2\tilde{c}_e}$ and $\frac{\tilde{\sigma}_w(\delta_w)\rho}{a_1} < \delta^2$. Then we can take

$$\tau \in \left(\frac{2c_2\hat{\sigma}_w(\delta_w)\rho}{a_1\hat{c}_3}, \left(\delta^2 - \frac{\tilde{\sigma}_w(\delta_w)\rho}{a_1}\right)\frac{c_1}{2\tilde{c}_e}\right)$$

which implies $\frac{\tau\hat{c}_3}{2c_2} > \frac{\hat{\sigma}_w(\delta_w)\rho}{a_1}$ and $\delta^2 > \frac{2\tilde{c}_e\tau}{c_1} + \frac{\tilde{\sigma}_w(\delta_w)\rho}{a_1}$.

From (52), we have

$$V_e^+ \leq \begin{cases} \frac{\tau}{2} + \frac{\hat{\sigma}_w(\delta_w)\rho}{a_1} + \hat{\sigma}_\alpha(|\Delta\alpha|), & V_e \leq \frac{\tau}{2} \\ \tau - \frac{\tau\hat{c}_3}{2c_2} + \frac{\hat{\sigma}_w(\delta_w)\rho}{a_1} + \hat{\sigma}_\alpha(|\Delta\alpha|), & \frac{\tau}{2} < V_e \leq \tau. \end{cases}$$

But $\hat{c}_3 \leq c_2$ (otherwise we could show $V_e < 0$ with $w_{\rm P} = 0$, $\Delta\alpha = 0$, and $e \neq 0$) so $\frac{\tau}{2} \geq \frac{\tau\hat{c}_3}{2c_2} > \frac{\hat{\sigma}_w(\delta_w)\rho}{a_1}$ and we have $V_e^+ \leq \tau$ so long as $|\Delta\alpha| \leq \delta_{\alpha,1} := \hat{\sigma}_\alpha^{-1}(\frac{\tau\hat{c}_3}{2c_2} - \frac{\hat{\sigma}_w(\delta_w)\rho}{a_1})$, which is positive by construction. Moreover, $V_e, V_e^+ \leq \tau$ implies $|(e,e^+)|^2 = |e|^2 + |e^+|^2 \leq \frac{2\tau}{c_1}$ and by (54),

$$\begin{aligned} |\tilde{d}|^2 &\leq \tilde{c}_e|(e,e^+)|^2 + \tilde{\sigma}_w(|w_{\rm P}|)|\hat{x} - x_s(\hat{\beta})|^2 + \tilde{\sigma}_\alpha(|\Delta\alpha|) \\ &\leq \frac{2\tilde{c}_e\tau}{c_1} + \tilde{\sigma}_w(\delta_w)\rho^2 + \tilde{\sigma}_\alpha(\delta_\alpha) \\ &\leq \delta^2 \end{aligned}$$

so long as $|\Delta\alpha| \leq \delta_{\alpha,2} := \tilde{\sigma}_\alpha^{-1}(\delta^2 - \frac{2\tilde{c}_e\tau}{c_1} - \frac{\tilde{\sigma}_w(\delta_w)\rho}{a_1})$, which exists and is positive by construction. Finally, we can take $\delta_\alpha := \min\{\delta_{\alpha,1}, \delta_{\alpha,2}\}$ to achieve $(x,\alpha,\hat{x},\hat{\beta}) \in \mathcal{S}_N^{\rho,\tau}$ at all times.

(b)—From part (a), we already have $\tau, \delta_w, \delta_\alpha > 0$ such that $\mathcal{S}_N^{\rho,\tau}$ is RPI. By Assumption 6 and Theorem 5 we have (18a) and (43a) at all times for some $a_1, a_2, c_1, c_2 > 0$. By Propositions 3 and 4, there exist $\hat{c}_3, \tilde{a}_3, \tilde{a}_4 > 0$ and $\hat{\sigma}_w, \hat{\sigma}_\alpha, \sigma_\alpha \in \mathcal{K}_\infty$ such that (52) and (53) at all times. Assume, without loss of generality, that $\delta_w < \hat{\sigma}_w^{-1}(\min\{\frac{c_1\tilde{a}_3}{\tilde{a}_4}, \frac{a_3\hat{c}_3}{a_4}\frac{c_1}{c_1+c_2}\})$. By Theorem 3, the system is RES on $\mathcal{S}_N^{\rho,\tau}$ w.r.t. $\delta\hat{x}$.

(c)—By Proposition 2, there exist $c_r, c_g > 0$ such that $|\delta r| \leq c_r|\delta\hat{x}| + c_g|\tilde{d}|$ where $\tilde{d} := (e, e^+, \Delta s_{\rm sp}, \tilde{w})$. Combining this inequality with (18a), (52), and (54) gives

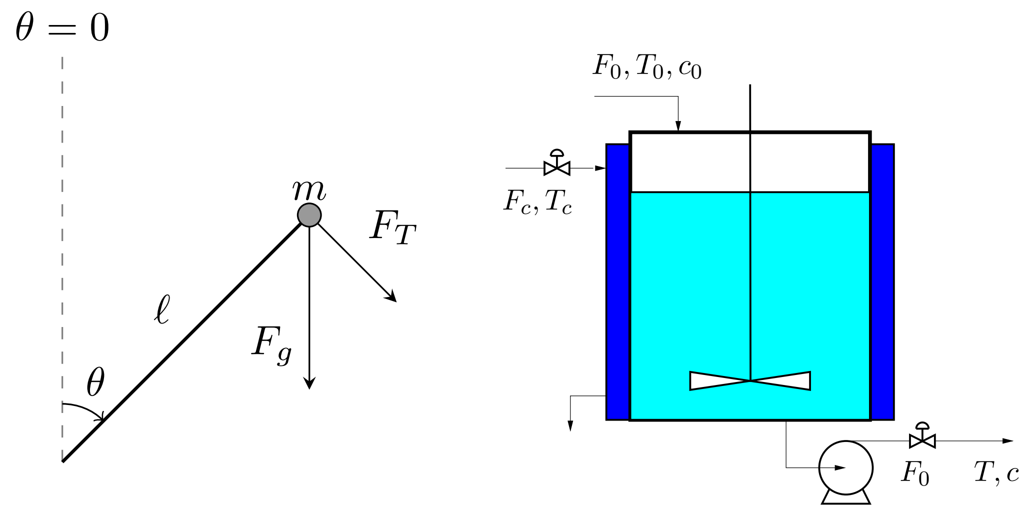
$$|\delta r| \leq c_{r,x}|\delta\hat{x}| + c_{r,e}|e| + \tilde{\gamma}_r(|\Delta\alpha|)$$

where $c_{r,x} := c_r + c_g(\sqrt{\tilde{\sigma}_\alpha(\delta_w)} + \sqrt{\tilde{c}_e\hat{\sigma}_\alpha(\delta_w)})$, $c_{r,e} := c_g\sqrt{\tilde{c}_e}(1 + \sqrt{c_2 - \hat{c}_3})$, and $\tilde{\gamma}_r := c_g(\sqrt{\tilde{\sigma}_\alpha} + \sqrt{\tilde{c}_e\hat{\sigma}_\alpha})$. Then

$$|(\delta r, e)| \leq \tilde{c}_r|(\delta\hat{x}, e)| + \tilde{\gamma}_r(|\Delta\alpha|)$$

where $\tilde{c}_r := c_{r,x} + c_{r,e} + 1$. Finally, RES w.r.t. $\delta\hat{x}$ gives

$$|(\delta\hat{x}(k), e(k))| \leq \tilde{c}\lambda^k|(\delta\hat{x}(0), \overline{e})| + \sum_{j=0}^{k} \lambda^j\tilde{\gamma}(|\Delta\alpha(k-j)|)$$



(a) Simple pendulum (b) Continuous stirred-tank reactor

Figure 1: Example systems.

for some $\tilde{c} > 0$, $\lambda \in (0,1)$, and $\tilde{\gamma} \in \mathcal{K}$, and therefore

$$\begin{aligned}|(\delta r(k), e(k))| &\leq \tilde{c}_r |(\delta\hat{x}(k), e(k))| + \tilde{\gamma}_r(|\Delta\alpha(k)|) \\ &\leq c\lambda^k |(\delta\hat{x}(0), \overline{e})| + \sum_{j=0}^{k} \lambda^j \gamma(|\Delta\alpha(k-j)|)\end{aligned}$$

where $c := \tilde{c}_r \tilde{c} > 0$ and $\gamma := \tilde{c}_r \tilde{\gamma} + \tilde{\gamma}_r \in \mathcal{K}_\infty$. □

7 Examples

In this section, we illustrate the main results using the example systems depicted in Figure 1. We compare two MPCs in our experiments.

Offset-free MPC The offset-free MPC (OFMPC) uses (6) and (14) and the following MHE problem:

$$\min_{(\mathbf{x},\mathbf{d}) \in \mathbb{X}^{T_k+1} \times \mathbb{D}^{T_k+1}} V_T^{\text{MHE}}(\mathbf{x}, \mathbf{d}, \mathbf{u}, \mathbf{y}) \tag{55}$$

where $T_k := \min\{k, T\}$, $T \in \mathbb{I}_{>0}$, and

$$\begin{aligned}V_T^{\text{MHE}}(\mathbf{x}, \mathbf{d}, \mathbf{u}, \mathbf{y}) := \sum_{j=0}^{T_k-1} &|x_{j+1} - f(x_j, u(j), d_j)|^2_{Q_w^{-1}} \\ &+ |d_{j+1} - d_j|^2_{Q_d^{-1}} + |y(j) - h(x_j, u(j), d_j)|^2_{R_v^{-1}}.\end{aligned} \tag{56}$$

For simplicity, a prior term is not used. Let $\hat{x}(j;\mathbf{u},\mathbf{y})$ and $\hat{d}(j;\mathbf{u},\mathbf{y})$ denote solutions to the above problem, and define the estimates by

$$\hat{x}(k) := \hat{x}(k;\mathbf{u}_{k-T_k:k-1},\mathbf{y}_{k-T_k:k-1}), \qquad \hat{d}(k) := \hat{d}(k;\mathbf{u}_{k-T_k:k-1},\mathbf{y}_{k-T_k:k-1}).$$

Tracking MPC The nominal tracking MPC (TMPC) uses (6) and (14) and the following MHE problem:

$$\min_{\mathbf{x}\in\mathbb{X}^{T_k+1}} V_T^{\mathrm{MHE}}(\mathbf{x},0,\mathbf{u}_{k-T_k:k-1},\mathbf{y}_{k-T_k:k-1}) \tag{57}$$

With solutions denoted by $\hat{x}(j;\mathbf{u},\mathbf{y})$, we define the estimates by estimates by

$$\hat{x}(k) := \hat{x}(k;\mathbf{u}_{k-T_k:k-1},\mathbf{y}_{k-T_k:k-1}), \qquad \hat{d}(k) := 0.$$

We also construct, in the proof of the following lemma, terminal ingredients satisfying Assumption 4.

Lemma 3. *Suppose Assumptions* 1 *to* 3 *and* 7 *hold with* $\mathcal{B}=\hat{\mathcal{B}}_c$ *and* $n_c=0$*, let* $Q\in\mathbb{R}^{n\times n}$ *and* $R\in\mathbb{R}^{n_u\times n_u}$ *be positive definite, and* $\partial^2_{(x,u)}f_i$*,* $i\in\mathbb{I}_{1:n}$ *exist and are bounded on* $\mathbb{X}\times\mathbb{U}\times\mathbb{D}$*. For each* $\beta=(r_{\mathrm{sp}},u_{\mathrm{sp}},y_{\mathrm{sp}},d)\in\mathcal{B}$*, let*

$$A(\beta) := \partial_x f(z_s(\beta),d), \qquad B(\beta) := \partial_u f(z_s(\beta),d).$$

If $(A(\beta),B(\beta))$ *is stabilizable for each* $\beta\in\mathcal{B}$*, then there exist functions* $\kappa_f:\mathbb{X}\times\mathcal{B}$ *and* $P_f:\mathcal{B}\to\mathbb{R}^{n\times n}$*, and a constant* $c_f>0$ *satisfying Assumptions* 4 *and* 5*.*

Proof. Throughout this proof, we let $\beta=(r_{\mathrm{sp}},u_{\mathrm{sp}},y_{\mathrm{sp}},\beta)\in\mathcal{B}$. Since $(A(\beta),B(\beta))$ is stabilizable, there exists a positive definite $P(\beta)$ that uniquely solves the following discrete algebraic Riccati equation,

$$P = A^\top PA + Q - A^\top PB(B^\top PB+R)^{-1}B^\top PA$$

where dependence on β has been suppressed for brevity. The solution P is continuous at each (A,B,Q,R) such that (A,B) is stabilizable and (Q,R) are positive definite (Sun, 1998).[4] Moreover, since f is twice differentiable and (x_s,u_s) are continuous on $\mathcal{B}$, then (A,B) must be continuous on $\mathcal{B}$. Therefore P is continuous on $\mathcal{B}$ and Assumption 5 holds for $P_f(\beta) := 2P(\beta)$.

Next, with $K := PB(B^\top PB+R)^{-1}$, $A_K := A-BK$, and $Q_K := Q+K^\top RK$, we have $A_K^\top P_f A_K - P_f = -2Q_K$, where dependence on β has been suppressed for brevity. Then

$$V_f(\overline{x}^+,\beta) - V_f(x,\beta) \le -2|\delta x|^2_{Q_K(\beta)} \tag{58}$$

where $\overline{x}^+ := A_K(\beta)\delta x + x_s(\beta)$ and $\delta x := x - x_s(\beta)$. Since the second derivatives $\partial^2_{(x,u)}f_i$, $i\in\mathbb{I}_{1:n}$ are bounded, there exists $\overline{c}>0$ (independent of β) such that $|x^+ - \overline{x}^+| \le \overline{c}|\delta x|^2$

[4]In fact, Sun (1998) only needed $(A,Q^{1/2})$ detectable to derive perturbation bounds. However, Assumption 5 guarantees positive definiteness of Q, so we get this automatically.

where $x^+ := f(x, \kappa_f(x,\beta), d)$ and $\kappa_f(x,\beta) := -K(\beta)\delta x + u_s(\beta)$.[5] Therefore, with $a(\beta) := 2\overline{c}\overline{\sigma}([A_K(\beta)]^\top P_f(\beta))$ and $b(\beta) := \overline{c}^2\overline{\sigma}(P_f(\beta))$, we have

$$|V_f(x^+,\beta) - V_f(\overline{x}^+,\beta)| \leq a(\beta)|\delta x|^3 + b(\beta)|\delta x|^4 \tag{59}$$

and combining (58) with (59), we have

$$\begin{aligned} &V_f(x^+,\beta) - V_f(x,\beta) + \ell(x, \kappa_f(x,\beta), \beta) \\ &\quad \leq -|\delta x|^2_{Q_K(\beta)} + V_f(x^+,\beta) - V_f(\overline{x}^+,\beta) \\ &\quad \leq -[c(\beta) - b(\beta)|\delta x| - a(\beta)|\delta x|^2]|\delta x|^2 \end{aligned} \tag{60}$$

where $c(\beta) := \underline{\sigma}(Q_K(\beta))$. The polynomial $p_\beta(s) = c(\beta) - b(\beta)s - a(\beta)s^2$ has roots at $s_\pm(\beta) = \frac{-b(\beta)\pm\sqrt{[b(\beta)]^2+4a(\beta)c(\beta)}}{2a(\beta)}$ and is positive in between. Moreover, $s_\pm$ are continuous over $\mathcal{B}$ because (a,b,c) are as well, and $s_\pm(\beta)$ are positive and negative, respectively. Define $c_f := \min_{\beta\in\mathcal{B}} \underline{\sigma}(P_f(\beta))[s_+(\beta)]^2$ which exists and is positive due to continuity and positivity of x_+ and $\underline{\sigma}(P_f(\cdot))$ and compactness of $\mathcal{B}$. Finally, we have that $V_f(x,\beta) \leq c_f$ implies $\underline{\sigma}(P_f(\beta))|\delta x|^2 \leq V_f(x,\beta) \leq c_f$ and therefore $|\delta x| \leq \sqrt{\frac{c_f}{\underline{\sigma}(P_f(\beta))}} \leq s_+(\beta)$ and (60) implies Assumption 4 with $P_f(\beta)$ and $c_f > 0$ as constructed. □

7.1 Simple pendulum

Consider the following nondimensionalized pendulum system (Figure 1a):

$$\dot{x} = F_{\mathrm{P}}(x,u,w_{\mathrm{P}}) := \begin{bmatrix} x_2 \\ \sin x_1 - (w_{\mathrm{P}})_1^2 x_2 + (\hat{k} + (w_{\mathrm{P}})_2)u + (w_{\mathrm{P}})_3 \end{bmatrix} \tag{61a}$$

$$y = h_{\mathrm{P}}(x,u,w_{\mathrm{P}}) := x_1 + (w_{\mathrm{P}})_4 \tag{61b}$$

$$r = g(u,y) := y \tag{61c}$$

where $(x_1,x_2) \in \mathbb{X} := \mathbb{R}^2$ are the angle and angular velocity, $u \in \mathbb{U} := [-1,1]$ is the (dimensionless) motor voltage, $\hat{k} = 5$ rad/s^2 is the estimated motor gain, $(w_{\mathrm{P}})_1$ is an air resistance factor, $(w_{\mathrm{P}})_2$ is the error in the motor gain estimate, $(w_{\mathrm{P}})_3$ is an externally applied torque, and $(w_{\mathrm{P}})_4$ is the measurement noise. Let $\psi(t;x,u,w_{\mathrm{P}})$ denote the solution to (61) at time t given $x(0) = x$, $u(t) = u$, and $w_{\mathrm{P}}(t) = w_{\mathrm{P}}$. We model the discretization of (61) by

$$x^+ = f_{\mathrm{P}}(x,u,w_{\mathrm{P}}) := x + \Delta F_{\mathrm{P}}(x,u,w_{\mathrm{P}}) + (w_{\mathrm{P}})_5 r_d(x,u,w_{\mathrm{P}}) \tag{62a}$$

where $(w_{\mathrm{P}})_5$ scales the discretization error, r_d is a residual function given by

$$r_d(x,u,w_{\mathrm{P}}) := \int_0^\Delta [F_{\mathrm{P}}(x(t),u,w_{\mathrm{P}}) - F_{\mathrm{P}}(x,u,w_{\mathrm{P}})]dt \tag{62b}$$

[5]This follows by applying Taylor's theorem to $e(x,\beta) := x^+ - \overline{x}^+$ at $(x_s(\beta), d)$ and noting the intercept and first derivative (in x) is zero.

and $x(t) = \psi(t; x, u, w_{\mathrm{P}})$. Assuming a zero-order hold on the input u and disturbance w_{P}, the system (61) is discretized (exactly) as (62) with $(w_{\mathrm{P}})_5 \equiv 1$. We model the system with $w_{\mathrm{P}} = w(d) := (0, 0, d, 0, 0)$, i.e.,

$$x^+ = f(x, u, d) := f_{\mathrm{P}}(x, u, w(d)) = x + \Delta \begin{bmatrix} x_2 \\ \sin x_1 + \hat{k}u + d \end{bmatrix} \tag{63a}$$

$$y = h(x, u, d) := h_{\mathrm{P}}(x, u, w(d)) = x_1 \tag{63b}$$

and therefore we do not need access to the residual function r_d to design the offset-free MPC.

For the following simulations, assume $w_{\mathrm{P}} \in \mathbb{W} := [-3, 3]^3 \times [-0.05, 0.05] \times \{0, 1\}$, and let the sample time be $\Delta = 0.1$ s. Regardless of objective ℓ_s, the SSTP (6) is uniquely solved by

$$x_s(\beta) := \begin{bmatrix} r_{\mathrm{sp}} \\ 0 \end{bmatrix}, \qquad u_s(\beta) := -\frac{1}{\hat{k}}(\sin r_{\mathrm{sp}} + d)$$

for each $\beta = (r_{\mathrm{sp}}, u_{\mathrm{sp}}, y_{\mathrm{sp}}, d) \in \mathcal{B}_c$, where

$$\mathcal{B}_c := \{ (r, u, y, d) \in \mathbb{R}^4 \mid |\sin r + d|, |\sin y + d| \leq \hat{k}, |u| \leq 1 \}$$

and $\delta_0 > 0$. Likewise, the solution to (47) is

$$x_{\mathrm{P,s}}(\alpha) := \begin{bmatrix} r_{\mathrm{sp}} \\ 0 \end{bmatrix}, \qquad d_s(\alpha) := \frac{\hat{k}(w_{\mathrm{P}})_3 - (w_{\mathrm{P}})_2 \sin r_{\mathrm{sp}}}{\hat{k} + (w_{\mathrm{P}})_2} \tag{64}$$

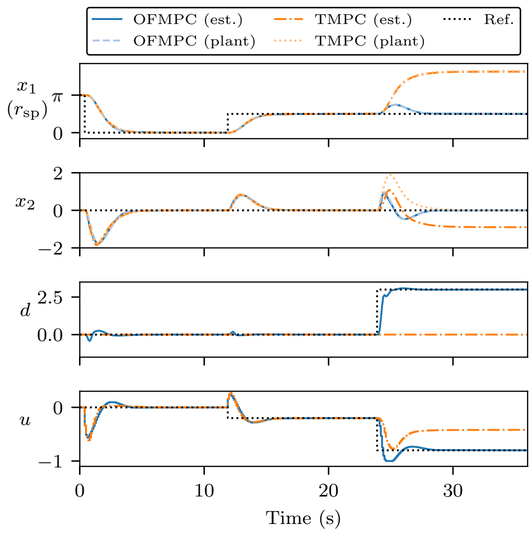
for each $\alpha = (r_{\mathrm{sp}}, u_{\mathrm{sp}}, y_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c$, where

$$\mathcal{A}_c := \{ (r, u, y, w) \in \mathbb{R}^3 \times \mathbb{W} \mid |\sin r + (w_{\mathrm{P}})_3|, |\sin y + (w_{\mathrm{P}})_3| \leq \hat{k} + (w_{\mathrm{P}})_2, |u| \leq 1 \}.$$

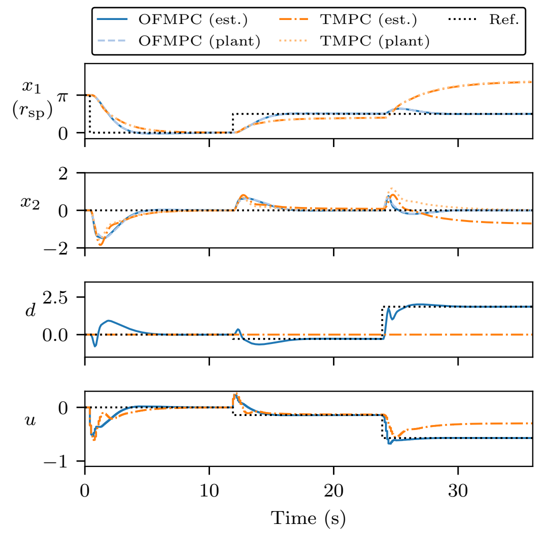
Notice that $\mathcal{A}_c$ and $\mathcal{B}_c$ are compact and satisfy Assumption 8. We define a regulator with $N := 20$, $\mathbb{U} := [-1, 1]$, $\ell_s(u, y) = |u|^2 + |y|^2$, $\ell(x, u, \Delta u, \beta) := |x - x_s(\beta)|^2 + 10^{-2}(u - u_s(\beta))^2 + 10^2(\Delta u)^2$,[6] $V_f(x, \beta) := |x - x_s(\beta)|^2_{P_f(\beta)}$, and $\mathbb{X}_f := \mathrm{lev}_{c_f} V_f$, where $P_f(\beta)$ and $c_f \approx 0.4364$ are chosen according to the proof of Lemma 3 to satisfy Assumptions 4 and 5. Assumption 2 is satisfied trivially and Assumptions 1, 8 and 9 are satisfied since smoothness of F implies that ψ, r, and f are smooth (Hale, 1980, Thm. 3.3). Finally, we use MHE designs (55) and (57) for the offset-free MPC and tracking MPC, respectively, where $T = 5$, $Q_w := \begin{bmatrix} 10^{-3} & \\ & 10^{-6} \end{bmatrix}$, and $Q_d := R_v := 1$. While the estimators defined by (55) and (57) should be RGES (Allan and Rawlings, 2021), it is not known if they satisfy Assumption 6. If Assumption 6 is satisfied, then Theorem 6 gives robust stability with respect to the tracking errors.

We present the results of numerical experiments in Figure 2. To ensure numerical accuracy, the plant (61) is simulated by four 4th-order Runga-Kutta steps per sample

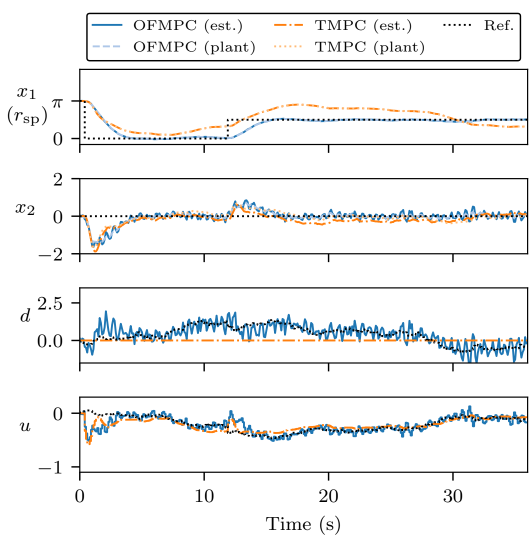
[6] The $\Delta u(k) := u(k) - u(k-1)$ penalty is a standard generalization used by practitioners to "smooth" the closed-loop response in a tuneable fashion.



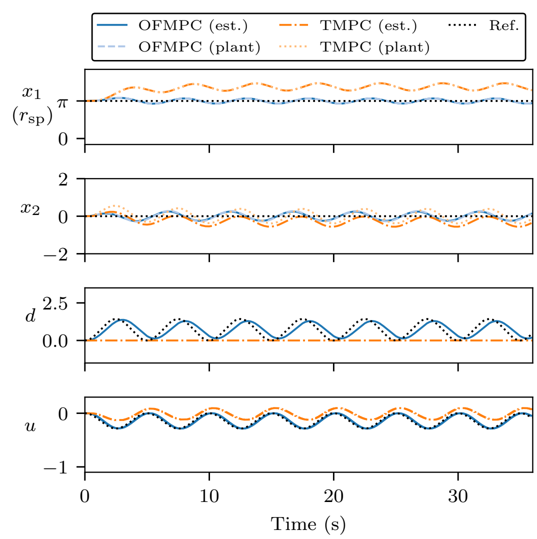
(a) No mismatch: $(w_P)_1 \equiv 0$ and $(w_P)_2 \equiv 0$.



(b) Mismatch: $(w_P)_1 \equiv 1$ and $(w_P)_2 \equiv 2$.



(c) Noise and mismatch: $(w_P)_3^+ = (w_P)_3 + (\Delta w_P)_3$, $(\Delta w_P)_3 \sim \mathrm{N}(0, 10^{-2})$, and $(w_P)_4 \sim \mathrm{N}(0, 10^{-4})$.



(d) Oscillating disturbance and mismatch: $(w_P)_3(k) = 1 - \cos(\frac{2\pi k}{50})$ and $r_{sp}(k) \equiv \pi$.

Figure 2: Simulated closed-loop trajectories for the offset-free MPC and tracking MPC of (61). Solid blue and dot-dashed orange lines represent the closed-loop estimates and inputs $(\hat{x}, \hat{d}, u)$ for the offset-free MPC and tracking MPC simulations, respectively. Dashed blue and dotted orange lines represent the closed-loop plant states x_P for the offset-free MPC and tracking MPC simulations, respectively. Dotted black lines represent the intended steady-state targets and disturbance values $(x_{P,s}, d_s, u_s)$ found by solving (6) and (47). We set $(w_P)_1 \equiv 1$, $(w_P)_2 \equiv 2$, $(w_P)_3(k) = 3H(k-240)$, $(w_P)_4 \equiv 1$, $(w_P)_5 \equiv 0$, and $r_{sp}(k) = \pi H(5-k) + \frac{\pi}{2} H(k-120)$, unless otherwise specified.

time. Unless otherwise specified, we consider, in each simulation, unmodeled air resistance $(w_\mathrm{P})_1 \equiv 1$, motor gain error $(w_\mathrm{P})_2 \equiv 2$, an exogenous torque $(w_\mathrm{P})_3(k) = 3H(k-240)$, the discretization parameter $(w_\mathrm{P})_4 \equiv 1$, no measurement noise $(w_\mathrm{P})_5 \equiv 0$, and a reference signal $r_\mathrm{sp}(k) = \pi H(5-k) + \frac{\pi}{2}H(k-120)$, where H denotes the unit step function. The setpoint brings the pendulum from the resting state $x_1 = \pi$, to the upright position $x_1 = 0$, to the half-way position $x_1 = \frac{\pi}{2}$.

In the first experiment, we consider the case without plant-model mismatch, i.e., $(w_\mathrm{P})_1 \equiv 0$ and $(w_\mathrm{P})_2 \equiv 0$ (Figure 2a). Both offset-free and tracking MPC remove offset after the setpoint changes. However, only offset-free MPC removes offset after the disturbance is injected. Without a disturbance model, the tracking MPC cannot produce useful steady-state targets, and the pendulum drifts far from the setpoint. Moreover, the tracking MPC produces pathological state estimates, with nonzero velocity at steady state.

The second experiment considers plant-model mismatch $(w_\mathrm{P})_1 \equiv 1$ and $(w_\mathrm{P})_2 \equiv 2$ (Figure 2b). As in the first experiment, both the tracking MPC and offset-free MPC bring the pendulum to the upright position $x_1 = 0$, without offset. However, only the offset-free MPC brings the pendulum to the half-way position $x_1 = \frac{\pi}{2}$. The tracking MPC, not accounting for motor gain errors, provides an insufficient force and does not remove offset. Note the intended disturbance estimate $d_s = \frac{13}{7}$ is a smaller value that the actual injected disturbance $(w_\mathrm{P})_3 = 3$, as underestimation of the motor gain necessitates a smaller disturbance value to be corrected. Again, the tracking MPC produces pathological state estimates.

The third experiment follows the second, except the exogenous torque is an integrating disturbance $(w_\mathrm{P})_3^+ = (w_\mathrm{P})_3 + (\Delta w_\mathrm{P})_3$ where $(w_\mathrm{P})_3 \sim \mathrm{N}(0, 10^{-2})$, and we have measurement noise $(w_\mathrm{P})_5 \sim \mathrm{N}(0, 10^{-4})$ (Figure 2c). In this experiment, we see the remarkable ability of offset-free MPC to track a reference subject to random disturbances. While the tracking MPC is robust to the disturbance $(w_\mathrm{P})_3$, it is not robust to the disturbance changes $(\Delta w_\mathrm{P})_3$ and wanders far from the setpoint as a result. On the other hand, offset-free MPC is robust to both and exhibits practically offset-free performance. We remark that, while the example is mechanical in nature, we are illustrating a behavior that is often desired in chemical process control, where process specifications must be met despite constantly, but slowly varying upstream conditions.

In the fourth and final experiment, the pendulum maintains the resting position $r_\mathrm{sp} = \pi$ subject to an oscillating torque $(w_\mathrm{P})_3(k) = 1 - \cos(\frac{2\pi k}{50})$ (Figure 2d). Tracking MPC wanders away from the setpoint, whereas offset-free MPC oscillates around it with small amplitude. We note the disturbance estimate $\hat{d}$ does not ever "catch" the intended value d_s as the disturbance model has no ability to match its *velocity* or *acceleration*. More general integrator schemes (e.g., double or triple integrators) could provide more dynamic tracking performance at the cost of a higher disturbance dimension (c.f., Maeder and Morari (2010) or (Zagrobelny, 2014, Ch. 5)).

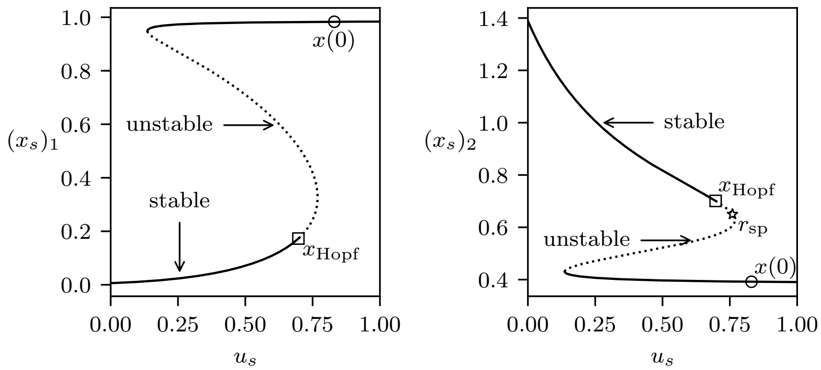


Figure 3: Nominal steady states for the CSTR (65).

7.2 Continuous stirred-tank reactor

Consider the following nonisothermal continuous stirred-tank reactor (CSTR) (Hicks and Ray, 1971; Kameswaran and Biegler, 2006) (Figure 1b):

$$\dot{x} = F_{\mathrm{P}}(x, u, w_{\mathrm{P}})$$
$$:= \begin{bmatrix} \theta^{-1}(1 + (w_{\mathrm{P}})_1 - x_1) - ke^{(w_{\mathrm{P}})_2 - M/x_2} x_1 \\ \theta^{-1}(x_f - x_2) + ke^{(w_{\mathrm{P}})_2 - M/x_2} x_1 - \gamma u (x_2 - x_c - (w_{\mathrm{P}})_3) \end{bmatrix} \tag{65a}$$
$$y = h_{\mathrm{P}}(x, u, w_{\mathrm{P}}) := x_2 + (w_{\mathrm{P}})_4 \tag{65b}$$
$$r = g(u, y) := y \tag{65c}$$

where $(x_1, x_2) \in \mathbb{X} := \mathbb{R}^2_{\geq 0}$ are the dimensionless concentration and temperature, $u \in \mathbb{U} := [0, 2]$ is the dimensionless coolant flowrate, $\theta = 20$ s is the residence time, $k = 300$ s^{-1} is the rate coefficient, $M = 5$ is the dimensionless activation energy, $x_f = 0.3947$ and $x_c = 0.3816$ are dimensionless feed and coolant temperatures, $\gamma = 0.117$ s^{-1} is the heat transfer coefficient, $(w_{\mathrm{P}})_1$ is a kinetic modeling error, $(w_{\mathrm{P}})_2$ is a change to the coolant temperature, and $(w_{\mathrm{P}})_4$ is the measurement noise. Again, we discretize the system (65) via the equations (62), where the continuous system is recovered with $(w_{\mathrm{P}})_5 = 1$ and zero-order holds on u and w_{P}. The system is modeled with $w_{\mathrm{P}} = w(d) := (0, d, 0, 0, 0)$, i.e.,

$$x^+ = f(x, u, d) := x + \Delta \begin{bmatrix} \theta^{-1}(1 - x_1) - ke^{-M/x_2} x_1 \\ \theta^{-1}(x_f - x_2) + ke^{-M/x_2} x_1 - \gamma u (x_2 - x_c - d) \end{bmatrix} \tag{66a}$$
$$y = h(x, u, d) := x_2. \tag{66b}$$

The goal in the following simulations is to control the CSTR (65) from a nominal steady state $(x(0), u(-1)) \approx (0.9831, 0.3918, 0.8305)$ to a temperature setpoint $r_{\mathrm{sp}} \in [0.6, 0.7]$. In this range the nominal steady states are unstable, with a nearby Hopf bifurcation at $(x_{\mathrm{Hopf}}, u_{\mathrm{Hopf}}) \approx (0.1728, 0.7009, 0.6973)$. We plot the nominal steady states (i.e., $w_{\mathrm{P}} = 0$) along with the initial steady state $x(0)$ and the Hopf bifurcation x_{Hopf} in Figure 3.

For the following simulations, assume disturbance set is $w_{\mathrm{P}} \in \mathbb{W} := [-0.05, 0.05]^4 \times \{0, 1\}$, and let the sample time be $\Delta = 1$ s. Regardless of objective ℓ_s, the SSTP (6) is uniquely solved by

$$x_s(\beta) := \begin{bmatrix} \frac{1}{1+\theta k e^{-M/r}} \\ r_{\mathrm{sp}} \end{bmatrix}, \qquad u_s(\beta) := \frac{x_f - r + 1 - (x_s(\beta))_1}{\theta\gamma(r - x_c - d)}$$

for each $\beta = (r_{\mathrm{sp}}, u_{\mathrm{sp}}, y_{\mathrm{sp}}, d) \in \mathcal{B}_c$, where

$$\mathcal{B}_c := [0.6, 0.7] \times \mathbb{U} \times [0.6, 0.7] \times [-0.1, 0.1]$$

and $\delta_0 > 0$. Likewise, the solution to (47) is

$$x_{\mathrm{P,s}}(\alpha) := \begin{bmatrix} r_{\mathrm{sp}} \\ 0 \end{bmatrix}, \qquad d_s(\alpha) \tag{67}$$

for each $\alpha = (r_{\mathrm{sp}}, u_{\mathrm{sp}}, y_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c$, where

$$\mathcal{A}_c := [0.6, 0.7] \times \mathbb{U} \times [0.6, 0.7] \times \mathbb{W}.$$

It is straightforward to verify $\mathcal{A}_c$ and $\mathcal{B}_c$ are compact and satisfy Assumption 8.

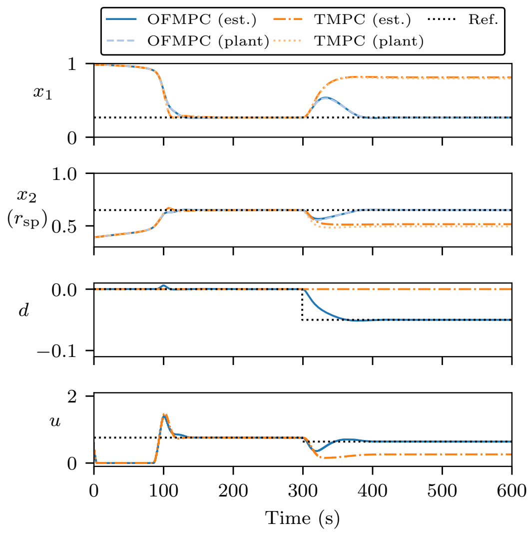
We define a regulator with $N := 150$, $\ell(x, u, \Delta u, \beta) := |x - x_s(\beta)|_Q^2 + 10^{-3}(u - u_s(\beta))^2 + (\Delta u)^2$, $Q = \begin{bmatrix} 10^{-3} & \\ & 1 \end{bmatrix}$, $V_f(x, \beta) := |x - x_s(\beta)|^2_{P_f(\beta)}$, and $\mathbb{X}_f := \mathrm{lev}_{c_f} V_f$, where $P_f(\beta)$ and $c_f \approx 6.7031 \times 10^{-16}$ are chosen according to the proof of Lemma 3 to satisfy Assumption 4.[7] Finally, we use MHE designs (55) and (57) for the offset-free MPC and tracking MPC, respectively, where $T := N$, $Q_w := 10^{-4} I$, $Q_d := 10^{-2}$, and $R_v := 1$. As in the simple pendulum example, if Assumption 6 is satisfied, then Theorem 6 implies the offset-free MPC can robustly track setpoints despite plant-model mismatch.

The results of the CSTR experiments are presented in Figure 4. Throughout these experiments, the plant (65) is simulated by ten 4th-order Runga-Kutta steps per sample time. Unless otherwise specified, each simulation is carried out with error in the feed concentration $(w_{\mathrm{P}})_1 \equiv -0.05$, error in the activation energy $(w_{\mathrm{P}})_2 \equiv -0.05$, a step in the coolant temperature $(w_{\mathrm{P}})_3(k) = -0.05 H(k - 300)$, no measurement noise $(w_{\mathrm{P}})_4 \equiv 0$, the discretization parameter $(w_{\mathrm{P}})_5 \equiv 1$, and a constant reference signal $r_{\mathrm{sp}} \equiv 0.65$.

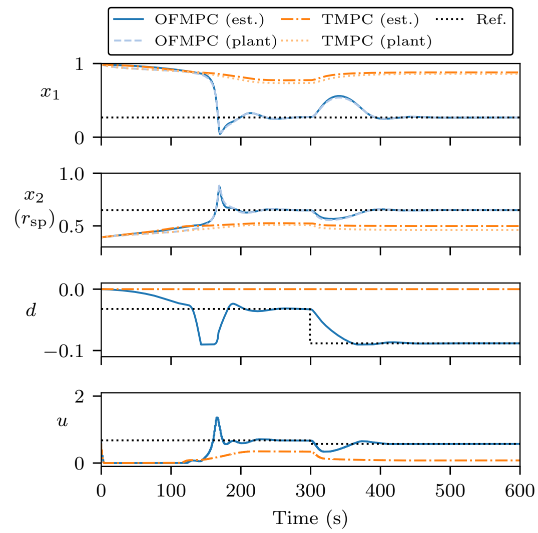
In the first experiment, we consider the case without plant-model mismatch, i.e., $(w_{\mathrm{P}})_1 \equiv 0$ and $(w_{\mathrm{P}})_2 \equiv 0$ (Figure 4a). As in the pendulum experiment, both offset-free and tracking MPC remove offset after the setpoint changes, but only offset-free MPC removes offset after the disturbance is injected. We also note that, after the disturbance is injected, the tracking MPC state estimates are slightly different than the plant states.

We consider plant-model mismatch $(w_{\mathrm{P}})_1 \equiv -0.05$ and $(w_{\mathrm{P}})_2 \equiv -0.05$ in the second experiment (Figure 2b). The offset-free MPC is able to track the reference and reject the disturbance despite mismatch, this time at the cost of a significant temperature spike

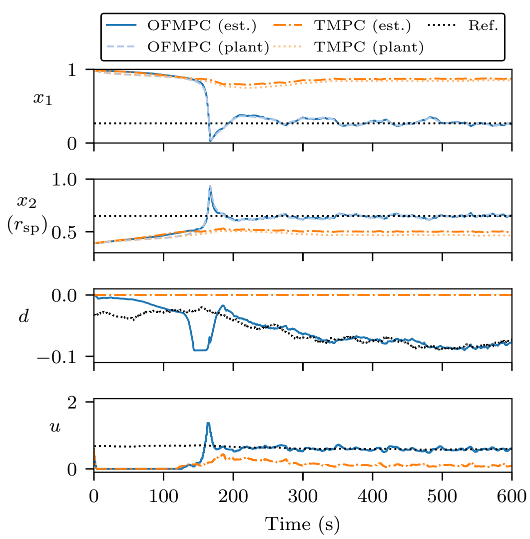
[7]While c_f was chosen near machine precision, the CSTR tends to evolve to the nearest stable steady state, and the horizon is chosen long enough to easily achieve this steady state to a high degree of precision. Thus, the system remains robust despite the tight terminal constraint.



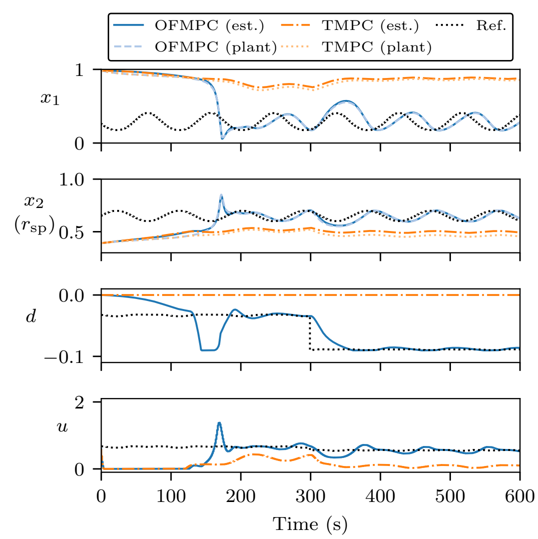
(a) No mismatch: $(w_P)_1 \equiv 0$ and $(w_P)_2 \equiv 0$.



(b) Mismatch: $(w_P)_1 \equiv -0.05$ and $(w_P)_2 \equiv -0.05$.



(c) Noise and mismatch: $(w_P)_3^+ = (w_P)_3 + (\Delta w_P)_3$, $(\Delta w_P)_3 \sim \mathrm{N}(0, 10^{-6})$, and $(w_P)_4 \sim \mathrm{N}(0, 10^{-4})$.



(d) Oscillating setpoint: $r_{sp}(k) = 0.05 \sin(\frac{2\pi k}{90}) + 0.65$.

Figure 4: Simulated closed-loop trajectories for the offset-free MPC and tracking MPC of the CSTR (65). Solid blue and dot-dashed orange lines represent the closed-loop estimates and inputs $(\hat{x}, \hat{d}, u)$ for the offset-free MPC and tracking MPC simulations, respectively. Dashed blue and dotted orange lines represent the closed-loop plant states x_P for the offset-free MPC and tracking MPC simulations, respectively. Dotted black lines represent the intended steady-state targets and disturbance values $(x_{P,s}, d_s, u_s)$ found by solving (6) and (47). We set $(w_P)_1 \equiv -0.05$, $(w_P)_2 \equiv -0.05$, $(w_P)_3(k) = -0.05H(k-300)$, $(w_P)_4 \equiv 0$, $(w_P)_5 \equiv 1$, and $r_{sp} \equiv 0.65$ unless otherwise specified.

around $k = 170$. On the other hand, the tracking MPC fails to bring the temperature above $x_2 = 0.5$, far from the setpoint $r_{\text{sp}} = 0.65$.

In the third experiment, the coolant temperature is an integrating disturbance $(w_{\text{P}})_3^+ = (w_{\text{P}})_3 + (\Delta w_{\text{P}})_3$, $(\Delta w_{\text{P}})_3 \sim \text{N}(0, 10^{-6})$, and we have measurement noise $(w_{\text{P}})_4 \sim \text{N}(0, 10^{-4})$ (Figure 4c). As in the corresponding pendulum experiment, offset-free MPC tracks the reference despite the randomly drifting disturbance. Here we are illustrating a behavior that is often desired in chemical process control, where process specifications must be met despite constantly, but slowly varying upstream conditions. We remark that, while the pendulum example is mechanical in nature, it illustrated the same property. The tracking MPC, on the other hand, still cannot handle the plant-model mismatch and fails to bring the temperature up to the setpoint.

In the fourth and final experiment, the setpoint follows an oscillating pattern $r_{\text{sp}}(k) = 0.05\sin(\frac{2\pi k}{90}) + 0.65$. Tracking MPC again fails bring the temperature up to the setpoint. Offset-free MPC closely follows the setpoint, substantially deviating from it only at the start-up phase and when the coolant temperature disturbance is injected. Again, we note that a precise tracking of this disturbance and reference signal could be accomplished by more general integrator schemes. (c.f., Maeder and Morari (2010) or (Zagrobelny, 2014, Sec. 5.3, 5.4)).

8 Conclusions

In this paper, we presented a nonlinear offset-free MPC design that is robustly stable with respect to setpoint- and target-tracking errors, despite persistent disturbances and plant-model mismatch. Our results are significantly stronger than the standard offset-free sufficient conditions that can be found in the literature. Notably, we do not assume stability of the closed-loop system to guarantee offset-free performance. The results are illustrated in numerical experiments.

These results form a foundation on which offset-free performance guarantees can be established on a wider class of MPC designs and applications. The results without mismatch (Theorem 5) should also extend to the control of plants with parameter drifts. A few limitations of this work, notably the requirement of a Lyapunov function for the estimator (Assumption 6), and the necessity of quadratic costs (Assumption 5), are also possible areas of future research.

A Proofs of robust estimation and tracking stability

A.1 Proof of Theorem 1

Proof of Theorem 1. First, note that $c_3 \leq c_2$, as otherwise, this would imply $V_e(k+1) \leq 0$ whenever $\tilde{w}(k) = 0$. We combine the upper bound (18a) and bound on the difference (18b) to give

$$V_e(k+1) \leq \lambda V_e(k) + c_4|\tilde{w}(k)|^2$$

where $\lambda := 1 - \frac{c_3}{c_2} \in (0,1)$. Recursively applying the above inequality gives

$$\begin{aligned} V_e(k) &\leq \lambda^k V_e(0) + \sum_{j=1}^{k} c_4 \lambda^{j-1} |\tilde{w}(k-j)|^2 \\ &\leq c_2 \lambda^{k+1} |\overline{e}|^2 + \sum_{j=1}^{k} c_4 \lambda^{j-1} |\tilde{w}(k-j)|^2 \end{aligned}$$

noting that $e(0) = \overline{e}$ because Φ_0 is the identity map. Finally,

$$|e(k)| \leq \sqrt{\frac{V_e(k)}{c_1}} \leq c_{e,1} \lambda_e^k |\overline{e}| + c_{e,2} \sum_{j=1}^{k+1} \lambda_e^{j-1} |\tilde{w}(k-j)|$$

where $c_{e,1} := \sqrt{\frac{c_2}{c_1}}$, $c_{e,2} := \sqrt{\frac{c_4}{c_1}}$, and $\lambda_e := \sqrt{\lambda}$. □

A.2 Proof of Theorem 2

Proof of Theorem 2. Suppose $X \subseteq \Xi$ is RPI for (20). Let the functions $V : \Xi \to \mathbb{R}_{\geq 0}$ and $\alpha_i, \sigma \in \mathcal{K}_\infty, i \in \mathbb{I}_{1:3}$ satisfy (24) for all $\xi \in X$ and $\omega \in \Omega_c(\xi)$. Let $(\boldsymbol{\xi}, \boldsymbol{\omega}, \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2)$ satisfy (20) and $\xi(0) \in X$.

Asymptotic case. The proof of this part follows similarly to (Jiang and Wang, 2001, Lem. 3.5) and (Tran et al., 2015, Thm. 1). We start by noting (24b) can be rewritten

$$V(F_c(\xi,\omega)) \leq (\textsc{id} - \alpha_4)(V(\xi)) + \sigma(|\omega|) \tag{68}$$

where $\alpha_4 := \alpha_3 \circ \alpha_2^{-1} \in \mathcal{K}_\infty$. Without loss of generality, we can assume $\textsc{id} - \alpha_4 \in \mathcal{K}$ (Jiang and Wang, 2001, Lem. B.1). Let $\rho \in \mathcal{K}_\infty$ such that $\textsc{id} - \rho \in \mathcal{K}_\infty$.

Let $b := \alpha_4^{-1}(\rho^{-1}(\sigma(\|\boldsymbol{\omega}\|)))$ and $D := \{\, \xi \in \Xi \mid V(\xi) \leq b \,\}$. The following intermediate result is required.

Lemma 4. *If there exists $k_0 \in \mathbb{I}_{\geq 0}$ such that $\xi(k_0) \in D$, then $\xi(k) \in D$ for all $k \geq k_0$.*

Proof. Suppose $k \geq k_0$ and $\xi(k) \in D$. Then $V(\xi(k)) \leq b$ and by (68),

$$\begin{aligned} V(\xi(k+1)) &\leq (\textsc{id} - \alpha_4)(V(\xi(k))) + \sigma(\|\boldsymbol{\omega}\|) \\ &\leq (\textsc{id} - \alpha_4)(b) + \sigma(\|\boldsymbol{\omega}\|) \\ &= \underbrace{-(\textsc{id} - \rho)(\alpha_4(b))}_{\leq 0} + b \underbrace{-\rho(\alpha_4(b)) + \sigma(\|\boldsymbol{\omega}\|)}_{=0} \leq b. \end{aligned}$$

The result follows by induction. □

Next, let $j_0 := \min\{\, k \in \mathbb{I}_{\geq 0} \mid \xi(k) \in D \,\}$. The above lemma gives $V(\xi(k)) \leq \gamma(\|\boldsymbol{\omega}\|)$ for all $k \geq j_0$, where $\gamma := \alpha_4^{-1} \circ \rho^{-1} \circ \sigma$. On the other hand, if $k < j_0$, then we have $\rho(\alpha_4(V(\xi(k)))) > \sigma(\|\boldsymbol{\omega}\|)$ and therefore

$$\begin{aligned} V(\xi(k+1)) - V(\xi(k)) &\leq -\alpha_4(V(\xi(k))) + \sigma(\|\boldsymbol{\omega}\|) \\ &= -\alpha_4(V(\xi(k))) + \rho(\alpha_4(V(\xi(k)))) - \rho(\alpha_4(V(\xi(k)))) + \sigma(\|\boldsymbol{\omega}\|) \\ &\leq -\alpha_4(V(\xi(k))) + \rho(\alpha_4(V(\xi(k)))). \end{aligned}$$

By (Jiang and Wang, 2002, Lem. 4.3), there exists $\beta \in \mathcal{KL}$ such that

$$\alpha_1(|\zeta_1(k)|) \leq V(\xi(k)) \leq \beta(V(\xi(0)), k) \leq \beta(\alpha_2(|\zeta_2(0)|), k).$$

Combining the above inequalities gives

$$|\zeta_1(k)| \leq \max\{\beta_\zeta(|\zeta_2(0)|, k),\ \gamma_\zeta(\|\boldsymbol{\omega}\|)\} \leq \beta_\zeta(|\zeta_2(0)|, k) + \gamma_\zeta(\|\boldsymbol{\omega}\|)$$

where $\beta_\zeta(s, k) := \alpha_1^{-1}(\beta(\alpha_2(s), k))$ and $\gamma_\zeta := \alpha_1^{-1} \circ \gamma$. Finally, causality lets us drop future terms of ω from the signal norm in the above inequality and simply write (22).

Exponential case. Suppose, additionally, that $\alpha_i(\cdot) := a_i(\cdot)^b, i \in \mathbb{I}_{1:3}$. Without loss of generality, we can assume $\lambda := 1 - a_3 \in (0, 1)$. Recursively applying (24b) gives

$$\begin{aligned} V(\xi(k)) &\leq \lambda^k V(\xi(0)) + \sum_{i=1}^{k} \lambda^{i-1} \sigma(|\omega(k-i)|) \\ &\leq \lambda^k a_2 |\zeta_2(0)|^b + \frac{\sigma(\|\boldsymbol{\omega}\|_{0:k-1})}{1-\lambda}. \end{aligned}$$

Applying (24a), we have

$$|\zeta_1(k)| \leq \left(\frac{a_2}{a_1} \lambda^k |\zeta_2(0)|^b + \frac{\sigma(\|\boldsymbol{\omega}\|_{0:k-1})}{a_1(1-\lambda)} \right)^{1/b}.$$

If $b \geq 1$, the triangle inequality gives

$$|\zeta_1(k)| \leq c_\zeta \lambda_\zeta^k |\zeta_2(0)| + \gamma_\zeta(\|\boldsymbol{\omega}\|_{0:k-1}) \tag{69}$$

with $c_\zeta := \left(\frac{a_2}{a_1}\right)^{1/b}$, $\lambda_\zeta := \lambda^{1/b}$, and $\gamma_\zeta(\cdot) := \left(\frac{\sigma(\cdot)}{a_1(1-\lambda)}\right)^{1/b}$. Otherwise, if $b < 1$, then convexity gives (69) with $c_\zeta := \frac{1}{2}\left(\frac{2a_2}{a_1}\right)^{1/b}$, $\lambda_\zeta := \lambda^{1/b}$, and $\gamma_\zeta(\cdot) := \frac{1}{2}\left(\frac{2\sigma(\cdot)}{a_1(1-\lambda)}\right)^{1/b}$. □

A.3 Proof of Theorem 3

Proof of Theorem 3. Throughout, we fix $k \in \mathbb{I}_{\geq 0}$ and drop dependence on k when it is understood from context. Let the trajectories $(\boldsymbol{\xi}, \hat{\boldsymbol{\xi}}, \mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\upsilon}, \boldsymbol{\varepsilon}, \boldsymbol{\zeta})$ satisfy (19) and (25)–(28), $\zeta = G(\hat{\xi})$, and $(\xi(0), \overline{\xi}) \in \mathcal{S}$, where $\mathcal{S}$ is RPI. Suppose Φ_0^ξ is the identity map. Let $a_i, b_i > 0, i \in \mathbb{I}_{1:4}$, $V : \hat{\Xi} \to \mathbb{R}_{\geq 0}$, $V_\varepsilon : \Xi \times \hat{\Xi} \to \mathbb{R}_{\geq 0}$, and $\sigma, \sigma_\varepsilon \in \mathcal{K}$ satisfy $\frac{a_4 c_4}{a_3 c_1} < 1$, $\frac{a_4 c_4}{a_3 c_3} < \frac{c_1}{c_1 + c_2}$, and (31).

Joint Lyapunov function Combining the fact $|(\varepsilon,\varepsilon^+)|^2 = |\varepsilon|^2 + |\varepsilon^+|^2$ with the inequalities (31), we have

$$\begin{aligned} V(\hat{\xi}^+) - V(\hat{\xi}) &\overset{(31b)}{\leq} -a_3|\zeta|^2 + a_4|\varepsilon|^2 + a_4|\varepsilon^+|^2 + \sigma(|\omega|) \\ &\overset{(31c)}{\leq} -a_3|\zeta|^2 + a_4|\varepsilon|^2 + \frac{a_4}{c_1} V_\varepsilon(\xi^+, \hat{\xi}^+) + \sigma(|\omega|) \\ &\overset{(31d)}{\leq} -\tilde{a}_3|\zeta|^2 + a_4\left(1 - \frac{c_3}{c_1}\right)|\varepsilon|^2 + \frac{a_4}{c_1} V_\varepsilon(\xi, \hat{\xi}) + \tilde{\sigma}(|\omega|) \\ &\overset{(31c)}{\leq} -\tilde{a}_3|\zeta|^2 + \tilde{a}_4|\varepsilon|^2 + \tilde{\sigma}(|\omega|) \end{aligned}$$

where $\tilde{a}_3 := a_3 - \frac{a_4 c_4}{c_1}$, $\tilde{a}_4 := a_4\left(1 + \frac{c_2 - c_3}{c_1}\right)$, and $\tilde{\sigma} := \frac{a_4}{c_1}\sigma_\varepsilon + \sigma \in \mathcal{K}$. Note that $\tilde{a}_3 = a_3\left(1 - \frac{a_4 c_4}{a_3 c_1}\right) > 0$ by assumption, and $\tilde{a}_4 > 0$ since $c_2 > c_3$.

Let $W(\xi, \hat{\xi}) := V(\hat{\xi}) + qV_\varepsilon(\xi, \hat{\xi})$ where $q > 0$. With $b_1 := \min\{a_1, qc_1\}$, we have the lower bound,

$$b_1|(\zeta,\varepsilon)|^2 = b_1|\zeta|^2 + b_1|\varepsilon|^2 \leq a_1|\zeta|^2 + qc_1|\varepsilon|^2 \leq V(\hat{\xi}) + qV_\varepsilon(\xi,\hat{\xi}) =: W(\xi,\hat{\xi}). \tag{70}$$

With $b_2 := \max\{a_2, qc_2\}$, we have the upper bound

$$W(\xi,\hat{\xi}) := V(\hat{\xi}) + qV_\varepsilon(\xi,\hat{\xi}) \leq a_2|\zeta|^2 + qc_2|\varepsilon|^2 \leq b_2|\zeta|^2 + b_2|\varepsilon|^2 = b_2|(\zeta,\varepsilon)|^2. \tag{71}$$

For the cost decrease, we first note that $\frac{a_4 c_4}{a_3 c_3} < \frac{c_1}{c_1 + c_2}$ implies

$$\tilde{a}_4 c_4 = a_4\left(\frac{c_1 + c_2}{c_1} - \frac{c_3}{c_1}\right)c_4 < a_4\left(\frac{a_3 c_3}{a_4 c_4} - \frac{c_3}{c_1}\right)c_4 = a_3 c_3 - \frac{a_4 c_3 c_4}{c_1} = \tilde{a}_3 c_3$$

and therefore $\frac{\tilde{a}_4}{c_3} < \frac{\tilde{a}_3}{c_4}$. With $q \in \left(\frac{\tilde{a}_4}{c_3}, \frac{\tilde{a}_3}{c_4}\right)$, we have

$$W(\xi^+,\hat{\xi}^+) \leq V(\hat{\xi}^+) + qV_\varepsilon(\xi^+,\hat{\xi}^+) \leq W(\xi,\hat{\xi}) - b_3|(\zeta,\varepsilon)|^2 + \sigma_W(|\omega|) \tag{72}$$

where $b_3 := \min\{\tilde{a}_3 - qc_4, qc_3 - \tilde{a}_4\} > 0$ and $\sigma_W := \tilde{\sigma} + q\sigma_\varepsilon \in \mathcal{K}$ by construction.

Robust exponential stability Substituting the upper bound (71) into the cost decrease (72) gives

$$W(\xi^+,\hat{\xi}^+) \leq \lambda W(\xi,\hat{\xi}) - b_3|(\zeta,\varepsilon)|^2 + \sigma_W(|\omega|) \tag{73}$$

where $\lambda := 1 - \frac{b_3}{b_2}$ and we can assume $\lambda \in (0,1)$ since

$$b_2 \geq qc_2 > qc_3 > qc_3 - \tilde{a}_4 \geq b_3.$$

Recursively applying (73) gives

$$\begin{aligned} W(\xi(k),\hat{\xi}(k)) &\leq \lambda^k W(\xi(0),\hat{\xi}(0)) + \sum_{i=1}^{k} \lambda^{i-1}\sigma(|\omega(k-i)|) \\ &\leq b_2\lambda^k|(\zeta(0),\varepsilon(0))|^2 + \sum_{i=1}^{k} \lambda^{i-1}\sigma(|\omega(k-i)|) \end{aligned}$$

where the second inequality follows from (71). Finally, by (70) and the triangle inequality, we have

$$|(\zeta(k), e(k))| \leq c_\zeta \lambda_\zeta^k |(\zeta(0), \varepsilon(0))| + \sum_{i=1}^{k} \gamma_\zeta(|\omega(k-i)|, i)$$

where $c_\zeta := \sqrt{\frac{b_2}{b_1}}$, $\lambda_\zeta := \sqrt{\lambda}$, and $\gamma_\zeta(s,k) := \lambda_\zeta^{k-1}\sqrt{\frac{\sigma(s)}{b_1}}$. □

B Proofs of offset-free MPC stability

B.1 Proof of Theorem 4

We begin by proving Theorem 4(a,b).

Proof of Theorem 4(a,b). (a)—Suppose $x \in \mathcal{X}_N^\rho(\beta)$ and $\beta \in \mathcal{B}_c$. From the main text, $\tilde{\mathbf{u}}(x,\beta)$ is feasible, so

$$V_N^0(f_c(x,\beta),\beta) \leq V_N(f_c(x,\beta), \tilde{\mathbf{u}}(x,\beta), \beta)$$

and, applying the inequality (35), we have

$$V_N^0(f_c(x,\beta),\beta) \leq V_N^0(x,\beta) - \ell(x, \kappa_N(x,\beta), \beta).$$

But

$$\underline{\sigma}(Q)|x - x_s(\beta)|^2 \leq \ell(x, \kappa_N(x,\beta), \beta) \leq V_N^0(x,\beta)$$

so the lower bound (34a) and the cost decrease (34b) both hold with $a_1 = a_3 = \underline{\sigma}(Q)$.

To establish the upper bound of (34a), we first note that since $P_f(\cdot)$ is continuous and positive definite, and $\mathcal{B}_c$ is compact, the maximum $\gamma := \max_{\beta\in\mathcal{B}_c} \overline{\sigma}(P_f(\beta)) > 0$ exists. Then $|x - x_s(\beta)| \leq \varepsilon := \sqrt{\frac{c_f}{\gamma}}$ implies

$$V_f(x,\beta) \leq \overline{\sigma}(P_f(\beta))|x - x_s(\beta)|^2 \leq \gamma|x - x_s(\beta)|^2 \leq c_f$$

and therefore $x \in \mathbb{X}_f(\beta)$. By monotonicity of the value function (Rawlings et al., 2020, Prop. 2.18) we have $V_N^0(x,\beta) \leq V_f(x,\beta)$ whenever $x \in \mathbb{X}_f(\beta)$, and therefore

$$V_N^0(x,\beta) \leq V_f(x,\beta) \leq \gamma|x - x_s(\beta)|^2$$

whenever $|x - x_s(\beta)| \leq \varepsilon$. On the other hand, if $|x - x_s(\beta)| > \varepsilon$, then

$$V_N^0(x,\beta) \leq \rho \leq \frac{\rho}{\varepsilon^2}|x - x_s(\beta)|^2.$$

Finally, we have the upper bound (34a) with $a_2 := \max\{\gamma, \frac{\rho}{\varepsilon^2}\}$.

(b)—Let $\beta \in \mathcal{B}$. We already have that $V_N^0(\cdot,\beta)$ is a Lyapunov function (for the system (32), on $\mathcal{X}_N^\rho(\beta)$) with respect to $x - x_s(\beta)$, and $f_c(x,\beta) \in \mathcal{X}_N(\beta)$ for all $x \in \mathcal{X}_N^\rho(\beta)$ by recursive feasibility. We can choose any compact set $\mathcal{B}_c \subseteq \mathcal{B}$ containing β to achieve the descent property (34b). Then, for each $x \in \mathcal{X}_N^\rho(\beta)$, we have

$$V_N^0(f_c(x,\beta),\beta) \leq V_N^0(x,\beta) - a_1|x - x_s(\beta)|^2 \leq \rho$$

and therefore $f_c(x,\beta) \in \mathcal{X}_N^\rho(\beta)$. In other words, $\mathcal{X}_N^\rho(\beta)$ is positive invariant for the system (32a). Finally, ES in $\mathcal{X}_N^\rho(\beta)$ w.r.t. $x - x_s(\beta)$ follows from Theorem 2. □

To prove Theorem 4(c,d), we need a few preliminary results.

Proposition 5 ((Allan et al., 2017, Prop. 20)). *Let $C \subseteq D \subseteq \mathbb{R}^m$, with C compact, D closed, and $V : D \to \mathbb{R}^p$ continuous. Then there exists $\alpha \in \mathcal{K}_\infty$ such that $|V(x) - V(y)| \leq \alpha(|x - y|)$ for all $x \in C$ and $y \in D$.*

Proposition 6. *Suppose Assumptions 1 to 5 hold. Let $\rho > 0$ and $\mathcal{B}_c \subseteq \mathcal{B}$ be compact. There exist $c_x, c_u > 0$ such that*

$$|x^0(j; x, \beta) - x_s(\beta)| \leq c_x |x - x_s(\beta)| \tag{74a}$$

$$|u^0(k; x, \beta) - u_s(\beta)| \leq c_u |x - x_s(\beta)| \tag{74b}$$

for each $x \in \mathcal{X}_N^\rho(\beta)$, $\beta \in \mathcal{B}_c$, $j \in \mathbb{I}_{1:N}$, and $k \in \mathbb{I}_{1:N-1}$.

Proof. Throughout, we fix $x \in \mathcal{X}_N^\rho(\beta)$ and $\beta \in \mathcal{B}_c$. Unless otherwise specified, the constructed constants and functions are independent of (x, β). By Theorem 4(a), there exists $a_2 > 0$ satisfying the upper bound (43a). Since P_f is continuous and positive definite and $\mathcal{B}_c$ is compact, the minimum $\gamma := \min_{\beta \in \mathcal{B}_c} \underline{\sigma}(P_f(\beta))$ exists and is positive. Moreover, since Q, R are positive definite, we have $\underline{\sigma}(Q), \underline{\sigma}(R) > 0$. For each $k \in \mathbb{I}_{0:N-1}$,

$$\begin{aligned}
\underline{\sigma}(Q)|x^0(k; x, \beta) - x_s(\beta)|^2 &\leq |x^0(k; x, \beta) - x_s(\beta)|_Q^2 \\
&\leq V_N^0(x, \beta) \leq a_2 |x - x_s(\beta)|^2 \\
\gamma |x^0(N; x, \beta) - x_s(\beta)|^2 &\leq |x^0(N; x, \beta) - x_s(\beta)|_{P_f(\beta)}^2 \\
&\leq V_N^0(x, \beta) \leq a_2 |x - x_s(\beta)|^2 \\
\underline{\sigma}(R)|u^0(k; x, \beta) - u_s(\beta)|^2 &\leq |u^0(k; x, \beta) - u_s(\beta)|_R^2 \\
&\leq V_N^0(x, \beta) \leq a_2 |x - x_s(\beta)|^2.
\end{aligned}$$

Thus, (74) holds for all $j \in \mathbb{I}_{1:N}$ and $k \in \mathbb{I}_{1:N-1}$ with $c_x := \max\{\sqrt{\frac{a_2}{\underline{\sigma}(Q)}}, \sqrt{\frac{a_2}{\gamma}}\}$ and $c_u := \sqrt{\frac{a_2}{\underline{\sigma}(R)}}$. □

Proposition 7. *Suppose Assumptions 1 to 5 hold. Let $\rho > 0$, $\mathcal{B}_c \subseteq \mathcal{B}$ be compact. There exists $\sigma_r \in \mathcal{K}_\infty$ such that*

$$|g_c(x, \beta) - r_{\text{sp}}| \leq \sigma_r(|x - x_s(\beta)|) \tag{75}$$

for each $x \in \mathcal{X}_N^\rho(\beta)$ and $\beta = (r_{\text{sp}}, z_{\text{sp}}, d) \in \mathcal{B}_c$. Moreover, if g and h are Lipschitz continuous on bounded sets, then (75) holds on the same sets with $\sigma_r(\cdot) := c_r(\cdot)$ and some $c_r > 0$.

Proof. By Proposition 5, there exists $\tilde{\sigma}_r \in \mathcal{K}_\infty$ such that

$$|g(u, h(x, u, d)) - g(\tilde{u}, h(\tilde{x}, \tilde{u}, \tilde{d}))| \leq \tilde{\sigma}_r(|(x - \tilde{x}, u - \tilde{u}, \beta - \tilde{\beta})|)$$

for all $x, \tilde{x} \in \mathcal{X}_N^\rho$, $u, \tilde{u} \in \mathbb{U}$, and $\beta = (r, z, d), \tilde{\beta} = (\tilde{r}, \tilde{z}, \tilde{d}) \in \mathcal{B}_c$. Fix $x \in \mathcal{X}_N^\rho(\beta)$ and $\beta \in \mathcal{B}_c$. The following constructions are independent of (x, β) unless otherwise specified. By Proposition 6, there exists $c_u > 0$ such that

$$|\kappa_N(x, \beta) - u_s(\beta)| \leq c_u |x - x_s(\beta)|$$

Combining these inequalities gives

$$\begin{aligned} |g_c(x, \beta) - r_{\mathrm{sp}}| &\leq \tilde{\sigma}_r(|(x - x_s(\beta), \kappa_N(x, \beta) - u_s(\beta))|) \\ &\leq \tilde{\sigma}_r((1 + c_u)|x - x_s(\beta)|) \\ &\leq \sigma_r(|x - x_s(\beta)|) \end{aligned}$$

where $\sigma_r(\cdot) := \tilde{\sigma}_r((1 + c_u)(\cdot)) \in \mathcal{K}_\infty$. If we also have that g and h are Lipschitz on bounded sets, then we can take $\sigma_r(\cdot) := c_r(\cdot)$ and $c_r := L_r(1 + c_u) > 0$, where $L_r > 0$ is the Lipschitz constant for $g(u, h(x, u, d))$ over $\mathcal{X}_N^\rho \times \mathbb{U} \times \mathcal{B}_c$. □

Proof of Theorem 4(c,d). Fix $x \in \mathcal{X}_N^\rho(\beta)$ and $\beta \in \mathcal{B}$. Let $\mathcal{B}_c \subseteq \mathcal{B}$ be compact, containing β. Define $\delta r := g_c(x, \beta) - r_{\mathrm{sp}}$ and $\delta x := x - x_s(\beta)$.

(c)—By Proposition 7, there exists $\sigma_r \in \mathcal{K}_\infty$ such that (75) holds. Then

$$\alpha_1(|\delta r|) := a_1[\sigma_r^{-1}(|\delta r|)]^2 \leq a_1|\delta x|^2 \leq V_N^0(x, \beta)$$

so $V_N^0(\cdot, \beta)$ is a Lyapunov function on $\mathcal{X}_N^\rho(\beta)$ w.r.t. $(\delta r, \delta x)$, and AS on $\mathcal{X}_N^\rho(\beta)$ w.r.t. $(\delta r, \delta x)$ follows by Theorem 2.

(d)—If g and h are Lipschitz continuous on bounded sets, then by Proposition 7, we can repeat part (c) with $\alpha_1(\cdot) := a_1 c_r^{-2}(\cdot)^2$ and some $c_r > 0$. Then $V_N^0(\cdot, \beta)$ is an exponential Lyapunov function on $\mathcal{X}_N^\rho(\beta)$ w.r.t. $(\delta r, \delta x)$, and ES on $\mathcal{X}_N^\rho(\beta)$ w.r.t. $(\delta r, \delta x)$ follows by Theorem 2. □

B.2 Proof of Proposition 1

To establish Proposition 1, we require the following result.

Proposition 8. *Suppose Assumptions* 1 *to* 5 *and* 7 *hold and let* $\rho > 0$. *The set*

$$\hat{\mathcal{X}}_N^\rho := \bigcup_{\hat{\beta} \in \hat{\mathcal{B}}_c} \mathcal{X}_N^\rho(\hat{\beta})$$

is compact, where $\hat{\mathcal{B}}_c$ *is defined as in Assumption* 7*(i).*

Proof. Consider the lifted set

$$\mathcal{F} := \{ (\hat{x}, \mathbf{u}, \hat{\beta}) \in \mathbb{X} \times \mathbb{U}^N \times \hat{\mathcal{B}}_c \mid V_f(\phi(N; \hat{x}, \mathbf{u}, \hat{\beta})) \leq c_f,\ V_N(\hat{x}, \mathbf{u}, \hat{\beta}) \leq \rho \}.$$

Notice $\hat{\mathcal{X}}_N^\rho$ is equivalent to the projection of $\mathcal{F}$ onto the first coordinate, i.e., $\hat{\mathcal{X}}_N^\rho = P(\mathcal{F})$ where $P(\hat{x}, \mathbf{u}, \hat{\beta}) = \hat{x}$. Since P is continuous, the image $\hat{\mathcal{X}}_N^\rho = P(\mathcal{F})$ is compact whenever $\mathcal{F}$ is compact. Thus, it suffices to show $\mathcal{F}$ is compact.

The set $\mathcal{F}$ is closed because $(\mathbb{X}, \mathbb{U}, \hat{\mathcal{B}}_c)$ are closed, and continuity of (f, x_s, u_s, ℓ, V_f) implies continuity of $V_f(\phi(N; \cdot, \cdot, \cdot))$ and $V_N(\cdot, \cdot, \cdot)$. Next, we show $\mathcal{F}$ is bounded. Since x_s is continuous and $\hat{\mathcal{B}}_c$ is compact, the maximum $\rho_s := \max_{\hat{\beta} \in \hat{\mathcal{B}}_c} |x_s(\hat{\beta})|$ exists and is finite. For each $(\hat{x}, \mathbf{u}, \hat{\beta}) \in \mathcal{F}$, we have $V_N^0(\hat{x}, \hat{\beta}) \leq V_N(\hat{x}, \mathbf{u}, \hat{\beta}) \leq \rho$ by construction. But $V_N^0(\hat{x}, \hat{\beta}) \geq \underline{\sigma}(Q)|\hat{x} - x_s(\hat{\beta})|^2$, so this implies $|\hat{x} - x_s(\hat{\beta})| \leq \sqrt{\frac{\rho}{\underline{\sigma}(Q)}}$ and therefore $|\hat{x}| \leq \sqrt{\frac{\rho}{\underline{\sigma}(Q)}} + \rho_s$. But $\mathbf{u}$ and $\hat{\beta}$ always lie in compact sets, so $\mathcal{F}$ must be bounded. □

Proof of Proposition 1. Let $\hat{\beta} \in \hat{\mathcal{B}}_c$, $\hat{x} \in \mathcal{X}_N^\rho(\hat{\beta})$, and $|\tilde{d}| \leq \delta_0$ such that $\hat{\beta}^+ := \hat{f}_{\beta,c}(\hat{\beta}, \tilde{d}) \in \hat{\mathcal{B}}_c$. For brevity, let

$$\overline{x}^+ := f_c(\hat{x}, \hat{\beta}), \qquad \overline{x}^+(N) := \phi(N; \overline{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{d}), \qquad \overline{x}(N) := x^0(N; \hat{x}, \hat{\beta}),$$
$$\hat{x}^+ := \hat{f}_c(\hat{x}, \hat{\beta}, \tilde{d}), \qquad \hat{x}^+(N) := \phi(N; \hat{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{d}^+).$$

Recall $\tilde{d} := (e, e^+, \Delta\beta, w, v)$, $e := (e_x, e_d)$, $e^+ := (e_x^+, e_d^+)$, and $\Delta\beta := (\Delta s_{\text{sp}}, w_d)$.

From Proposition 8, the set $\hat{\mathcal{X}}_N^\rho$ is compact. Since the functions (f, x_s, u_s, P_f) are continuous, so are (V_f, V_N). By Proposition 5, there exist $\sigma_f, \sigma_{V_f}, \sigma_{V_N} \in \mathcal{K}_\infty$ such that

$$|f(x_1, u_1, \hat{d}_1) - f(x_2, u_2, \hat{d}_2)| \leq \sigma_f(|(x_1 - x_2, u_1 - u_2, \hat{d}_1 - \hat{d}_2)|) \tag{76}$$

$$|V_f(\phi(N; x_1, \mathbf{u}_1, \hat{d}_1), \hat{\beta}_1) - V_f(\phi(N; x_2, \mathbf{u}_2, \hat{d}_2), \hat{\beta}_2)| \leq \sigma_{V_f}(|(x_1 - x_2, \mathbf{u}_1 - \mathbf{u}_2, \hat{\beta}_1 - \hat{\beta}_2)|) \tag{77}$$

$$|V_N(x_1, \mathbf{u}_1, \hat{\beta}_1) - V_N(x_2, \mathbf{u}_2, \hat{\beta}_2)| \leq \sigma_{V_N}(|(x_1 - x_2, \mathbf{u}_1 - \mathbf{u}_2, \hat{\beta}_1 - \hat{\beta}_2)|) \tag{78}$$

for all $x_1 \in \mathbb{X}$, $x_2 \in \hat{\mathcal{X}}_N^\rho$, $u_1, u_2 \in \mathbb{U}$, $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^N$, and $\hat{\beta}_1 = (s_1, \hat{d}_1), \hat{\beta}_2 = (s_2, \hat{d}_2) \in \hat{\mathcal{B}}_c$.

Substituting $x_1 = \hat{x} + e_x$, $x_2 = \hat{x}$, $u_1 = u_2 = \kappa_N(\hat{x}, \hat{\beta})$, $\hat{d}_1 = \hat{d} + e_d$, and $\hat{d}_2 = \hat{d}$ into (76), we have $|\hat{x}^+ - \overline{x}^+| \leq \sigma_f(|e|) + |w| + |e_x^+|$. But $|\hat{\beta}^+ - \hat{\beta}| \leq |\Delta\beta| + |e_d| + |e_d^+|$, so

$$|(\hat{x}^+ - \overline{x}^+, \hat{\beta}^+ - \hat{\beta})| \leq \sigma_f(\tilde{d}) + 5|\tilde{d}|. \tag{79}$$

Substituting $x_1 = \hat{x}^+$, $x_2 = \hat{f}_c(\hat{x}, \hat{\beta})$, $\mathbf{u}_1 = \mathbf{u}_2 = \tilde{\mathbf{u}}(\hat{x}, \hat{\beta})$, $\hat{\beta}_1 = \hat{\beta}^+$, and $\hat{\beta}_2 = \hat{\beta}$ into (77) and (78) gives

$$\begin{aligned} |V_f(\hat{x}^+(N), \hat{\beta}^+) - V_f(\overline{x}^+(N), \hat{\beta})| &\leq \sigma_{V_f}(|(\hat{x}^+ - \overline{x}^+, \hat{\beta}^+ - \hat{\beta})|) \\ &\leq \tilde{\sigma}_{V_f}(|\tilde{d}|) \end{aligned} \tag{80}$$

$$\begin{aligned} |V_N(\hat{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{\beta}^+) - V_N(\overline{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{\beta})| &\leq \sigma_{V_N}(|(\hat{x}^+ - \overline{x}^+, \hat{\beta}^+ - \hat{\beta})|) \\ &\leq \sigma_r(|\tilde{d}|) \end{aligned} \tag{81}$$

where $\tilde{\sigma}_{V_f}(\cdot) := \sigma_{V_f}(\sigma_f(\cdot) + 5(\cdot)) \in \mathcal{K}_\infty$, $\sigma_r(\cdot) := \sigma_{V_N}(\sigma_f(\cdot) + 5(\cdot)) \in \mathcal{K}_\infty$, and the second and fourth inequalities follow from (79).

Part (a). By definition (9) and (10), we have $\tilde{\mathbf{u}}(\hat{x}, \hat{\beta}) \in \mathcal{U}_N(\hat{x}^+, \hat{\beta}^+)$ if and only if $V_f(\hat{x}^+(N), \hat{\beta}^+) \leq c_f$. Thus, it suffices to construct $\delta_1 > 0$ (independently of $\hat{\beta}$ and $\tilde{d}$)

for which $\hat{x} \in \mathcal{X}_N(\hat{\beta})$ implies $V_f(\hat{x}^+(N), \hat{\beta}^+) \leq c_f$. Since $\hat{x} \in \mathcal{X}_N(\hat{\beta})$, we already have $V_f(\overline{x}(N), \hat{\beta}) \leq c_f$, and by Assumptions 4 and 5,

$$\begin{aligned} V_f(\overline{x}^+(N), \hat{\beta}) &\leq V_f(\overline{x}(N), \hat{\beta}) - \ell(\overline{x}(N), \kappa_f(\overline{x}(N), \hat{\beta}), \hat{\beta}) \\ &\leq V_f(\overline{x}(N), \hat{\beta}) - \underline{\sigma}(Q)|\overline{x}(N) - x_s(\hat{\beta})|^2. \end{aligned}$$

Since $\hat{\mathcal{B}}_c$ is compact and $\overline{\sigma}, P_f$ are continuous functions, the maximum

$$a_{f,2} := \max_{\hat{\beta} \in \hat{\mathcal{B}}_c} \overline{\sigma}(P_f(\hat{\beta}))$$

exists and is finite, so

$$\frac{c_f}{2} \leq V_f(\overline{x}(N), \hat{\beta}) \leq a_{f,2}|\overline{x}(N) - x_s(\hat{\beta})|^2.$$

Then $|\overline{x}(N) - x_s(\hat{\beta})| \geq \sqrt{\frac{c_f}{2a_{f,2}}}$ and

$$V_f(\overline{x}^+(N), \hat{\beta}) \leq c_f - \frac{c_f \underline{\sigma}(Q)}{2a_{f,2}}. \tag{82}$$

On the other hand, if $V_f(\overline{x}(N), \hat{\beta}) \leq \frac{c_f}{2}$, then we have

$$V_f(\overline{x}^+(N), \hat{\beta}) \leq \frac{c_f}{2}. \tag{83}$$

Finally, combining (80), (82), and (83), we have

$$V_f(\hat{x}^+(N), \hat{\beta}^+) \leq c_f - \gamma_f + \tilde{\sigma}_{V_f}(|\tilde{d}|)$$

where $\gamma_f := \min\{ \frac{c_f}{2}, \frac{c_f \underline{\sigma}(Q)}{2a_{f,2}} \}$ was defined independently of $(\hat{\beta}, \tilde{d})$. Finally, taking $\delta_1 := \min\{ \delta_0, \tilde{\sigma}_{V_f}^{-1}(\gamma_f) \}$, we have $V_f(\hat{x}^+(N), \hat{\beta}^+) \leq c_f$ and $\tilde{\mathbf{u}}(\hat{x}, \hat{\beta}) \in \mathcal{U}_N(\hat{x}^+, \hat{\beta}^+)$.

Part (b). By (35), we have

$$\begin{aligned} V_N(\overline{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{\beta}) &\leq V_N^0(\hat{x}, \hat{\beta}) - \ell(\hat{x}, \kappa_N(\hat{x}, \hat{\beta}), \hat{\beta}) \\ &\leq V_N^0(\hat{x}, \hat{\beta}) - \underline{\sigma}(Q)|\overline{x}(N) - x_s(\hat{\beta})|^2. \end{aligned} \tag{84}$$

Combining (81) and (84) gives (41) with $a_3 := \underline{\sigma}(Q)$, which is positive since Q is positive definite.

Part (c). The proof of this part follows similarly that of part (a). Since $\hat{x} \in \mathcal{X}_N^\rho(\hat{\beta})$, we have $V_N^0(\hat{x}, \hat{\beta}) \leq \rho$. If $V_N^0(\hat{x}, \hat{\beta}) \geq \frac{\rho}{2}$, then, by Theorem 4(a), we have

$$\frac{\rho}{2} \leq V_N^0(\hat{x}, \hat{\beta}) \leq a_2|\hat{x} - x_s(\hat{\beta})|^2$$

for some $a_2 > 0$. Therefore $|\hat{x} - x_s(\hat{\beta})| \leq \sqrt{\frac{\rho}{2a_2}}$ and

$$V_N(\overline{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{\beta}) \leq \rho - \frac{\rho \underline{\sigma}(Q)}{2a_2}. \tag{85}$$

On the other hand, if $V_N^0(\hat{x},\hat{\beta}) \leq \frac{\rho}{2}$, then

$$V_N(\overline{x}^+, \tilde{\mathbf{u}}(\hat{x},\hat{\beta}),\hat{\beta}) \leq \frac{\rho}{2}. \tag{86}$$

Combining (41), (85), and (86) gives

$$V_N(\hat{x}^+, \tilde{\mathbf{u}}(\hat{x},\hat{\beta}),\hat{\beta}) \leq \rho - \gamma + \tilde{\sigma}_{V_N}(|\tilde{d}|)$$

where $\gamma := \min\{ \frac{\rho}{2}, \frac{\rho\underline{\sigma}(Q)}{2a_2} \}$. But $\tilde{\mathbf{u}}(\hat{x},\hat{\beta})$ is feasible by part (a), so by optimality, we have

$$V_N^0(\hat{x}^+,\hat{\beta}^+) \leq V_N(\hat{x}^+, \tilde{\mathbf{u}}(\hat{x},\hat{\beta}),\hat{\beta}) \leq \rho - \gamma + \tilde{\sigma}_{V_N}(|\tilde{d}|).$$

Thus, as long as $|\tilde{d}| \leq \delta := \min\{ \delta_1, \tilde{\sigma}_{V_N}^{-1}(\gamma) \}$, we have $V_N^0(\hat{x}^+,\hat{\beta}^+) \leq \rho$ and $\hat{x}^+ \in \mathcal{X}_N^\rho(\hat{\beta}^+)$.□

B.3 Proof of Proposition 2

Proof of Proposition 2. Proposition 7 gives (44a). By Proposition 5, there exists $\sigma_g \in \mathcal{K}_\infty$ such that

$$|g(u_1, h(x_1,u_1,d_1) + v_1) - g(u_2, h(x_2,u_2,d_2) + v_2)| \\ \leq \sigma_g(|(x_1 - x_2, u_1 - u_2, d_1 - d_2, v_1 - v_2)|) \tag{87}$$

for all $x_1, x_2 \in \mathcal{X}_N^\rho(\beta)$, $u_1, u_2 \in \mathbb{U}$, $d_1, d_2 \in \mathbb{D}_c$, and $v_1 \in \mathbb{V}_c(x_1,u_1,d_1)$, and $v_2 \in \mathbb{V}_c(x_2,u_2,d_2)$, where

$$\mathbb{D}_c := \{ d \in \mathbb{D} \mid (s_{\text{sp}}, d) \in \mathcal{B}_c \}$$
$$\mathbb{V}_c(x,u,d) := \{ v \in \delta\mathbb{B}^{n_y} \mid h(x,u,d) + v \in \mathbb{Y} \}$$

Fix $\hat{x} \in \mathcal{X}_N^\rho(\hat{\beta})$, $\hat{\beta} = (s_{\text{sp}}, \hat{d}) \in \mathcal{B}_c$, and $\tilde{d} = (e, e^+, \Delta s_{\text{sp}}, \tilde{w}) \in \tilde{\mathbb{D}}_c(\hat{x},\hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$, where $e = (e_x, e_d)$ and $\tilde{w} = (w, w_d, v)$. Substituting $x_1 = \hat{x} + e_x$, $x_2 = \hat{x}$, $u_1 = u_2 = \kappa_N(\hat{x},\hat{\beta})$, $d_1 = \hat{d} + e_d$, $d_2 = \hat{d}$, $v_1 = v$, and $v_2 = 0$ into (87) gives, independently of $(\hat{x},\hat{\beta},\tilde{d})$,

$$|\hat{g}_c(\hat{x},\hat{\beta},\tilde{d}) - g_c(\hat{x},\hat{\beta})| \leq \sigma_g(|(e_x, e_d, v)|) \leq \sigma_g(|\tilde{d}|).$$

Then (44b) follows by the triangle inequality. Finally, if g and h are Lipschitz continuous on bounded sets, we can take $\sigma_g(\cdot) := c_g(\cdot)$ where $c_g > 0$ is the Lipschitz constant for $g(u, h(x,u,d) + v)$. □

B.4 Proof of Proposition 3

To prove Proposition 3, we derive a bound on $|\tilde{w}|$.

Proposition 9. *Suppose Assumptions* 1 *to* 3*,* 8 *and* 9 *hold. For any compact* $X \subseteq \mathbb{X}$ *and* $\mathcal{A}_c \subseteq \mathbb{R}^{n_r} \times \overline{\mathbb{Z}}_y \times \mathbb{W}$ *such that* $(s_{\text{sp}}, w_{\text{P}}) \in \mathcal{A}_c$ *implies* $(s_{\text{sp}}, 0) \in \mathcal{A}_c$*, there exist functions* $\sigma_w, \sigma_\alpha \in \mathcal{K}_\infty$ *for which*

$$|\tilde{w}| \leq \sigma_w(|w_{\text{P}}|)|z - z_s(\beta)| + \sigma_\alpha(|\Delta\alpha|) \tag{88}$$

for all $z \in X \times \mathbb{U}$ *and* $\alpha = (s_{\text{sp}}, w_{\text{P}}), \alpha^+ \in \mathcal{A}_c$*, where* $\beta := (s_{\text{sp}}, d_s(\alpha))$*,* $\tilde{w} := (w, w_d, v)$*,* $\Delta\alpha := \alpha^+ - \alpha$*, and* (50)*.*

Proof. For ease of notation, we let $z = (x, u) \in X \times \mathbb{U}$, $\alpha = (s_{\text{sp}}, w_{\text{P}}) \in \mathcal{A}_c$, $\beta := (s_{\text{sp}}, d_s(\alpha))$, $\tilde{w} := (w, w_d, v)$, and

$$\Delta\tilde{w}(x, u, \alpha) := \begin{bmatrix} f_{\text{P}}(x + \Delta x_s(\alpha), u, w_{\text{P}}) - f(x, u, \hat{d}_s(\alpha)) - \Delta x_s(\alpha) \\ h_{\text{P}}(x + \Delta x_s(\alpha), u, w_{\text{P}}) - h(x, u, \hat{d}_s(\alpha)) \end{bmatrix}$$

throughout. We also note that, by definition of the SSTP (6) and the nominal model assumption (3), we have

$$\Delta\tilde{w}(z_s(\beta), \alpha) = 0, \qquad \partial_z \Delta\tilde{w}(z, s_{\text{sp}}, 0) = 0. \tag{89}$$

Assume all functions continuously differentiable on $\mathbb{X} \times \mathbb{U}$ have been extended continuously differentiable functions on all of $\mathbb{R}^{n+n_u}$ using appropriately defined partitions of unity (cf. (Lee, 2012, Lem. 2.26)).

Let Z_c denote the convex hull of $X \times \mathbb{U}$. For each $i \in \mathbb{I}_{1:n+n_y}$, $\partial_z \Delta\tilde{w}_i$ is continuous, and by Proposition 5, there exists $\sigma_i \in \mathcal{K}_\infty$ such that

$$|\partial_z \Delta\tilde{w}_i(z_1, \alpha_1) - \partial_z \Delta\tilde{w}_i(z_2, \alpha_2)| \le \sigma_i(|(z_1 - z_2, \alpha_1 - \alpha_2)|)$$

for all $z_1, z_2 \in Z_c$ and $\alpha_1, \alpha_2 \in \mathcal{A}_c$. Substituting $z_1 = z_2 = z$, $\alpha_1 = \alpha$, and $\alpha_2 = (s_{\text{sp}}, 0)$ into the above inequality, we have

$$|\partial_z \Delta\tilde{w}_i(z, \alpha)| = |\partial_z \Delta\tilde{w}_i(z, \alpha) - \partial_z \Delta\tilde{w}(z, s_{\text{sp}}, 0)| \le \sigma_i(|w_{\text{P}}|) \tag{90}$$

where the equality follows by (89). By Taylor's theorem (Apostol, 1974, Thm. 12.14), for each $i \in \mathbb{I}_{1:n+n_y}$, there exist $z_i(z, \alpha) \in Z_c$ and $t_i(z, \alpha) \in (0, 1)$ such that

$$\Delta\tilde{w}_i(z, \alpha) = \partial_z \Delta\tilde{w}_i(\tilde{z}_i(z, \alpha), \alpha)(z - z_s(\beta)) \tag{91}$$

where $\tilde{z}_i(z, \alpha) := t_i(z, \alpha) z_s(\beta) + (1 - t_i(z, \alpha)) z_i(z, \alpha) \in Z_c$ by convexity of Z_c, and the zero-order term drops by (89). Combining (90) and (91),

$$|\Delta\tilde{w}(z, \alpha)| \le \sum_{i=1}^{n+n_y} |\Delta\tilde{w}_i(z, \alpha)| \le \sum_{i=1}^{n+n_y} \sigma_i(|w_{\text{P}}|)|z - z_s(\beta)| = \sigma_w(|w_{\text{P}}|)|z - z_s(\beta)| \tag{92}$$

where $\sigma_w := \sum_{i=1}^{n+n_y} \sigma_i$. By Proposition 5, since $x_{\text{P},s}, x_s, d_s$ are continuous, there exist $\sigma_x, \sigma_d \in \mathcal{K}_\infty$ such that

$$|\Delta x_s(\alpha_1) - \Delta x_s(\alpha_2)| \le \sigma_x(|\alpha_1 - \alpha_2|) \tag{93a}$$

$$|d_s(\alpha_1) - d_s(\alpha_2)| \le \sigma_d(|\alpha_1 - \alpha_2|) \tag{93b}$$

for all $\alpha_1, \alpha_2 \in \mathcal{A}_c$. Finally, using (92) and (93) with $\alpha_1 = \alpha$ and $\alpha_2 = \alpha^+$ gives

$$\begin{aligned} |\tilde{w}| &\le |\Delta\tilde{w}(z, \alpha)| + |\Delta x_s(\alpha^+) - \Delta x_s(\alpha)| + |d_s(\alpha^+) - d_s(\alpha)| \\ &\le \sigma_w(|w_{\text{P}}|)|z - z_s(\beta)| + \sigma_\alpha(|\Delta\alpha|) \end{aligned}$$

with $\sigma_\alpha := \sigma_x + \sigma_d \in \mathcal{K}_\infty$. □

Proof of Proposition 3. With $\delta_w \in (0, \sigma_w^{-1}(\sqrt{\frac{c_3}{4c_4 L_s^2}}))$, we can combine (18b), (74b), and (88) (from Assumption 6 and Propositions 6 and 9, respectively) and the identity $(a+b)^2 \leq 2a^2 + 2b^2$ to give

$$\begin{aligned}
|\tilde{w}|^2 &\leq [\sigma_w(|w_{\mathrm{P}}|)|z - z_s(\beta)| + \sigma_\alpha(|\Delta\alpha|)]^2 \\
&\leq 2[\sigma_w(|w_{\mathrm{P}}|)]^2|z - z_s(\beta)|^2 + 2[\sigma_\alpha(|\Delta\alpha|)]^2 \\
&\leq 2[\sigma_w(|w_{\mathrm{P}}|)]^2[(1+c_u)|\hat{x} - x_s(\hat{\beta})| + L_s|e|]^2 + 2[\sigma_\alpha(|\Delta\alpha|)]^2 \\
&\leq 4[\sigma_w(|w_{\mathrm{P}}|)]^2(1+c_u)^2|\hat{x} - x_s(\hat{\beta})|^2 + 4[\sigma_w(|w_{\mathrm{P}}|)]^2 L_s^2|e|^2 + 2[\sigma_\alpha(|\Delta\alpha|)]^2
\end{aligned}$$

and therefore (52), where $\hat{c}_3 := c_3 - 4c_4[\sigma_w(\delta_w)]^2 L_s^2 > 0$, $\hat{\sigma}_w(\cdot) := 4c_4[\sigma_w(\cdot)]^2(1+c_u)^2$, $\hat{\sigma}_\alpha(\cdot) := 2c_4[\sigma_\alpha(\cdot)]^2$, and $L_s > 0$ is the Lipschitz constant for z_s. □

B.5 Proof of Proposition 4

To establish Proposition 4, we require two preliminary results.

Proposition 10. *Suppose Assumptions* 1 *to* 5, 8 *and* 9 *hold. Let* $\rho, \delta_w > 0$. *There exist* $\tilde{c}_e > 0$ *and* $\tilde{\sigma}_w, \tilde{\sigma}_\alpha \in \mathcal{K}_\infty$ *such that*

$$|\tilde{d}|^2 \leq \tilde{c}_e|(e, e^+)|^2 + \tilde{\sigma}_w(|w_{\mathrm{P}}|)|\hat{x} - x_s(\hat{\beta})|^2 + \tilde{\sigma}_\alpha(|\Delta\alpha|) \tag{94}$$

so long as $\alpha = (s_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathcal{A}_c(\delta_w)$, $\Delta\alpha = (\Delta s_{\mathrm{sp}}, \Delta w_{\mathrm{P}}) \in \mathbb{A}_c(\alpha, \delta_w)$, $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N^\rho$, $\tilde{d} = (e, e^+, \Delta s_{\mathrm{sp}}, \tilde{w}) \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta_0 \mathbb{B}^{n_{\tilde{d}}}$, *and* $\hat{\beta} = (s_{\mathrm{sp}}, \hat{d})$, *given* (17) *and* (50).

Proof. From Propositions 6 and 9 and (Rawlings and Ji, 2012, Eq. (1)),

$$\begin{aligned}
|\tilde{w}|^2 &\leq [\sigma_w(|w_{\mathrm{P}}|)|z - z_s(\beta)| + \sigma_\alpha(|\Delta\alpha|)]^2 \\
&\leq [\sigma_w(|w_{\mathrm{P}}|)|z - z_s(\hat{\beta})| + L_s\sigma_w(|w_{\mathrm{P}}|)|e| + \sigma_\alpha(|\Delta\alpha|)]^2 \\
&\leq [\sigma_w(|w_{\mathrm{P}}|)|x - x_s(\hat{\beta})| + \sigma_w(|w_{\mathrm{P}}|)|u - u_s(\hat{\beta})| + L_s\sigma_w(|w_{\mathrm{P}}|)|e| + \sigma_\alpha(|\Delta\alpha|)]^2 \\
&\leq [(1+c_u)\sigma_w(|w_{\mathrm{P}}|)|\hat{x} - x_s(\hat{\beta})| + (L_s+1)\sigma_w(|w_{\mathrm{P}}|)|e| + \sigma_\alpha(|\Delta\alpha|)]^2 \\
&\leq 9(1+c_u)^2[\sigma_w(|w_{\mathrm{P}}|)]^2|\hat{x} - x_s(\beta)|^2 + 9(L_s+1)^2[\sigma_w(|w_{\mathrm{P}}|)]^2|e|^2 + 9[\sigma_\alpha(|\Delta\alpha|)]^2
\end{aligned}$$

where $L_s > 0$ is the Lipschitz constant for z_s and $c_u > 0$ and $\sigma_w, \sigma_\alpha \in \mathcal{K}_\infty$ satisfy (74b) and (88). Therefore

$$\begin{aligned}
|\tilde{d}|^2 &= |(e, e^+)|^2 + |\Delta s_{\mathrm{sp}}|^2 + |\tilde{w}|^2 \\
&\leq 9(1+c_u)^2(\sigma_w(|w_{\mathrm{P}}|))^2|\hat{x} - x_s(\beta)|^2 \\
&\qquad + (1 + 9(L_s+1)^2(\sigma_w(\delta_w))^2)|(e, e^+)|^2 + |\Delta\alpha|^2 + 9\sigma_\alpha(|\Delta\alpha|))^2
\end{aligned}$$

so (94) holds with $\tilde{c}_e := 1 + 9(L_s+1)^2[\sigma_w(\delta_w)]^2 > 0$ and $\tilde{\sigma}_w := 9(1+c_u)^2\sigma_w^2, \sigma_\alpha := \mathrm{ID}^2 + 9\sigma_\alpha \in \mathcal{K}_\infty$. □

Proposition 11. *Suppose Assumptions 1 to 5, 8 and 9 hold and let $\rho > 0$. There exist $a_{V_N,1} \in (0, \underline{\sigma}(Q))$ and $a_{V_N,2}, \delta > 0$ and $\sigma_{V_N} \in \mathcal{K}_\infty$ such that*

$$|V_N(\hat{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{\beta}^+) - V_N(\overline{x}^+, \tilde{\mathbf{u}}(\hat{x}, \hat{\beta}), \hat{\beta})| \leq a_{V_N,1}|\hat{x} - x_s(\hat{\beta})|^2 + a_{V_N,2}|\tilde{d}|^2 \tag{95}$$

for all $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N^\rho$ and $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$, where $\overline{x}^+ := f_c(\hat{x}, \hat{\beta})$, $\hat{x}^+ := \hat{f}_c(\hat{x}, \hat{\beta}, \tilde{d})$, and $\hat{\beta}^+ := \hat{f}_{\beta,c}(\hat{\beta}, \tilde{d})$.

Proof. By continuity of P_f, there exists $\sigma_{P_f} \in \mathcal{K}_\infty$ such that

$$\|P_f(\beta_1) - P_f(\beta_2)\| \leq \sigma_{P_f}(|\beta_1 - \beta_2|) \tag{96}$$

for all $\beta_1, \beta_2 \in \hat{\mathcal{B}}_c$. Moreover, since $\hat{\mathcal{B}}_c$ is compact and $P_f(\cdot)$ is continuous and positive definite, the maximum $\gamma := \max_{\hat{\beta}\in\hat{\mathcal{B}}_c} \overline{\sigma}(P_f(\hat{\beta}))$ exists and is finite and the minimum $\gamma_0 := \max_{\hat{\beta}\in\hat{\mathcal{B}}_c} \underline{\sigma}(P_f(\hat{\beta}))$ exists and is positive. Let $L_s > 0$ denote the Lipschitz constant for z_s on $\hat{\mathcal{B}}_c$. Throughout, we let $(\hat{x}, \hat{\beta}) \in \hat{\mathcal{S}}_N^\rho$, $\delta\hat{x} := \hat{x} - x_s(\hat{\beta})$, $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta})$, $\overline{x}^+ := f_c(\hat{x}, \hat{\beta})$, $\hat{x}^+ := \hat{f}_c(\hat{x}, \hat{\beta}, \tilde{d})$, $\hat{\beta}^+ := \hat{f}_{\beta,c}(\hat{\beta}, \tilde{d})$, $\tilde{\mathbf{u}} := \tilde{\mathbf{u}}(\hat{x}, \hat{\beta})$, $\overline{x}^+(k) := \phi(k; \overline{x}^+, \tilde{\mathbf{u}}, \hat{\beta})$, and $\hat{x}^+(k) := \phi(k; \hat{x}^+, \tilde{\mathbf{u}}, \hat{\beta}^+)$.

By Assumption 9, we have

$$|\overline{x}^+ - \hat{x}^+| \leq L_f|e| + |w| + |e_x^+| \leq L_f'|\tilde{d}| \tag{97}$$

where $L_f > 0$ is the Lipschitz constant for f and $L_f' := L_f + 2$, by Assumption 8(b), we have

$$|z_s(\hat{\beta}^+) - z_s(\hat{\beta})| \leq L_s|\hat{\beta}^+ - \hat{\beta}| \leq L_s(|\Delta\beta| + |e_d| + |e_d^+|) \leq 3L_s|\tilde{d}| \tag{98}$$

and by Proposition 6, we have $c_x, c_u > 0$ such that

$$|\overline{x}^+(j) - x_s(\hat{\beta})| \leq c_x|\delta\hat{x}| \tag{99}$$

$$|\tilde{u}(k) - u_s(\hat{\beta})| \leq c_u|\delta\hat{x}| \tag{100}$$

for each $j \in \mathbb{I}_{0:N-1}$ and $k \in \mathbb{I}_{0:N-2}$.

By Assumptions 4 and 5, we have

$$\begin{aligned}
\gamma_0|\overline{x}^+(N) - x_s(\hat{\beta})|^2 &\leq V_f(\overline{x}^+(N-1), \hat{\beta}) \\
&\leq V_f(\overline{x}^+(N-1), \hat{\beta}) - \underline{\sigma}(Q)|\overline{x}^+(N-1) - x_s(\hat{\beta})|^2 \\
&\leq [\gamma - \underline{\sigma}(Q)]|\overline{x}^+(N-1) - x_s(\hat{\beta})|^2 \\
&\overset{(99)}{\leq} [\gamma - \underline{\sigma}(Q)]c_x^2|\delta\hat{x}|^2.
\end{aligned}$$

Therefore

$$|\overline{x}^+(N) - x_s(\hat{\beta})| \leq \gamma_f c_x|\delta\hat{x}| \tag{101a}$$

where $\gamma_f := \sqrt{\frac{\gamma-\underline{\sigma}(Q)}{\gamma_0}}$. Similarly, using the fact that $V_f(\overline{x}^+(N),\hat{\beta}) \geq 0$, we have

$$\begin{aligned}\underline{\sigma}(R)|\tilde{u}(N-1)-u_s(\hat{\beta})|^2 &\leq V_f(\overline{x}^+(N-1),\hat{\beta}) - \underline{\sigma}(Q)|\overline{x}^+(N-1)-x_s(\hat{\beta})|^2 \\ &\leq (\gamma-\underline{\sigma}(Q))|\overline{x}^+(N-1)-x_s(\hat{\beta})|^2 \\ &\overset{(99)}{\leq} (\gamma-\underline{\sigma}(Q))c_x^2|\delta\hat{x}|^2\end{aligned}$$

and therefore

$$|\tilde{u}(N-1)-u_s(\hat{\beta})| \leq c_{u,f}|\delta\hat{x}| \tag{101b}$$

with $c_{u,f} := c_x\sqrt{\frac{\gamma-\underline{\sigma}(Q)}{\underline{\sigma}(R)}}$.

Due to continuous differentiability of f, we have

$$\begin{aligned}|\hat{x}^+(k)-\overline{x}^+(k)| &= |f(\hat{x}^+(k-1),\tilde{u}(k),\hat{d}^+) - f(\overline{x}^+(k-1),\tilde{u}(k),\hat{d})| \\ &\leq L_f|\hat{x}^+(k-1)-\overline{x}^+(k-1)| + L_f|\hat{d}^+-\hat{d}|\end{aligned}$$

where $L_f > 0$ is the Lipschitz constant for f. Applying this inequality recursively, for all $k \in \mathbb{I}_{0:N}$, we have

$$|\hat{x}^+(k)-\overline{x}^+(k)| \leq L_f^k|\hat{x}^+-\overline{x}^+| + L_f(k)|\hat{d}^+-\hat{d}| \leq L_f'(k)|\tilde{d}| \tag{102}$$

where $L_f(k) := \sum_{i=1}^k L_f^i$ and $L_f'(k) := L_f^k L_f' + 3L_f(k)$, and we have used (97) and the fact that $|\hat{d}^+-\hat{d}| \leq |w_d| + |e_d| + |e_d^+| \leq 3|\tilde{d}|$. Moreover,

$$|\hat{x}^+(k)-x_s(\hat{\beta})| \overset{(99),(102)}{\leq} c_x|\delta\hat{x}| + L_f'(k)|\tilde{d}| \tag{103}$$

and

$$|\hat{x}^+(N)-x_s(\hat{\beta})| \overset{(101),(102)}{\leq} c_x\gamma_f|\delta\hat{x}| + L_f'(N)|\tilde{d}|. \tag{104}$$

Using the inequalities, $||\xi|_{M_1}^2 - |\xi|_{M_2}^2| \leq \|M_1-M_2\||\xi|^2$, (96), and $|\hat{\beta}^+-\hat{\beta}| \leq |\Delta\beta| + |e_d| + |e_d^+| \leq 3|\tilde{d}|$, we have

$$V_f(\hat{x}^+(N),\hat{\beta}^+) \leq |\hat{x}^+(N)-x_s(\hat{\beta}^+)|_{P_f(\hat{\beta})}^2 + \sigma_{P_f}(3|\tilde{d}|)|\hat{x}^+(N)-x_s(\hat{\beta}^+)|^2.$$

Using the identity $|\xi_1+\xi_2|^2 \leq 2|\xi_1|^2 + 2|\xi_2|^2$, we have

$$V_f(\hat{x}^+(N),\hat{\beta}^+) \leq |\hat{x}^+(N)-x_s(\hat{\beta}^+)|_{P_f(\hat{\beta})}^2 + \sigma_{P_f,x}(|\tilde{d}|)|\delta\hat{x}|^2 + \sigma_{P_f,d}(|\tilde{d}|)|\tilde{d}|^2. \tag{105}$$

where $\sigma_{P_f,x}(\cdot) := 2c_x^2\gamma_f^2\sigma_{P_f}(3(\cdot))$ and $\sigma_{P_f,d}(\cdot) := 2(L_f'(N))^2\sigma_{P_f}(3(\cdot))$.

For the remainder of this part, we let $\lambda > 0$ (to be defined) and use the identity $2ab \leq \lambda a^2 + \lambda^{-1}b^2$. Expanding quadratics and using the identities (98)–(100), we have

$$\begin{aligned}&||\hat{x}^+(N)-x_s(\hat{\beta}^+)|_{P_f(\hat{\beta})}^2 - |\hat{x}^+(N)-x_s(\hat{\beta})|_{P_f(\hat{\beta})}^2| \\ &\quad\leq 6\gamma L_s|\hat{x}^+(N)-x_s(\hat{\beta})||\tilde{d}| + 9\gamma L_s^2|\tilde{d}|^2 \\ &\quad\leq 6\gamma L_s c_x\gamma_f|\delta\hat{x}||\tilde{d}| + (6\gamma L_s L_f'(N) + 9\gamma L_s^2)|\tilde{d}|^2 \\ &\quad\leq 3\gamma\lambda L_s c_x\gamma_f|\delta\hat{x}|^2 + (6\gamma L_s L_f'(N) + 9\gamma L_s^2 + 3\lambda^{-1}\gamma L_s c_x\gamma_f)|\tilde{d}|^2 \\ &\quad\leq \lambda\hat{L}_1(N)|\delta\hat{x}|^2 + \hat{L}_2(N,\lambda)|\tilde{d}|^2\end{aligned} \tag{106}$$

where $\hat{L}_1(N) := 3\gamma L_s c_x \gamma_f$ and $\hat{L}_2(N,\lambda) := 6\gamma L_s L_f'(N) + 9\gamma L_s^2 + 3\lambda^{-1}\gamma L_s c_x \gamma_f$. Similarly, for each $k \in \mathbb{I}_{0:N-1}$,

$$\begin{aligned}
&||\hat{x}^+(k) - x_s(\hat{\beta}^+)|_Q^2 - |\hat{x}^+(k) - x_s(\hat{\beta})|_Q^2| \\
&\qquad \leq 6\underline{\sigma}(Q)L_s|\hat{x}^+(k) - x_s(\hat{\beta})||\tilde{d}| + 9\underline{\sigma}(Q)L_s^2|\tilde{d}|^2 \\
&\qquad \leq 6\underline{\sigma}(Q)L_s c_x|\delta\hat{x}||\tilde{d}| + (6\underline{\sigma}(Q)L_s L_f'(k) + 9\underline{\sigma}(Q)L_s^2)|\tilde{d}|^2 \\
&\qquad \leq 3\lambda\underline{\sigma}(Q)L_s c_x|\delta\hat{x}|^2 + (6\underline{\sigma}(Q)L_s L_f'(k) + 9\underline{\sigma}(Q)L_s^2 + 3\lambda^{-1}\gamma L_s c_x)|\tilde{d}|^2 \\
&\qquad \leq \lambda\hat{L}_1(k)|\delta\hat{x}|^2 + \hat{L}_2(k,\lambda)|\tilde{d}|^2
\end{aligned} \tag{107}$$

where $\hat{L}_1(k) := 3\underline{\sigma}(Q)L_s c_x$ and $\hat{L}_2(k,\lambda) := 6\underline{\sigma}(Q)L_s L_f'(k) + 9\underline{\sigma}(Q)L_s^2 + 3\lambda^{-1}\gamma L_s c_x$, and

$$\begin{aligned}
&||\tilde{u}(k) - u_s(\hat{\beta}^+)|_R^2 - |\tilde{u}(k) - u_s(\hat{\beta})|_R^2| \\
&\qquad \leq 6\underline{\sigma}(R)L_s|\tilde{u}(k) - u_s(\hat{\beta})||\tilde{d}| + 9\underline{\sigma}(R)L_s^2|\tilde{d}|^2 \\
&\qquad \leq 6\underline{\sigma}(R)L_s c_u(k)|\delta\hat{x}||\tilde{d}| + 9\underline{\sigma}(R)L_s^2|\tilde{d}|^2 \\
&\qquad \leq 3\lambda\underline{\sigma}(R)L_s c_u(k)|\delta\hat{x}|^2 + (9\underline{\sigma}(R)L_s^2 + 3\lambda^{-1}\underline{\sigma}(R)L_s c_u(k))|\tilde{d}|^2 \\
&\qquad \leq \lambda\tilde{L}_1(k)|\delta\hat{x}|^2 + \tilde{L}_2(k,\lambda)|\tilde{d}|^2
\end{aligned} \tag{108}$$

where $\tilde{L}_1(k) := 3\underline{\sigma}(R)L_s c_u(k)$, $\tilde{L}_2(k,\lambda) := 9\underline{\sigma}(R)L_s^2 + 3\lambda^{-1}\underline{\sigma}(R)L_s c_u(k)$, and $c_u(k) = c_u$ if $k \in \mathbb{I}_{0:N-2}$ and $c_u(N-1) = c_{u,f}$.

For the uniform $\hat{\beta}$ bound, we have

$$\begin{aligned}
&|V_N(\hat{x}^+, \tilde{\mathbf{u}}, \hat{\beta}) - V_N(\overline{x}^+, \tilde{\mathbf{u}}, \hat{\beta})| \\
&\qquad \leq \sum_{k=0}^{N-1} 2\overline{\sigma}(Q)|\hat{x}^+(k) - \overline{x}^+(k)||\overline{x}^+(k) - x_s(\hat{\beta})| + \overline{\sigma}(Q)|\hat{x}^+(k) - \overline{x}^+(k)|^2 \\
&\qquad\quad + 2\gamma|\hat{x}^+(N) - \overline{x}^+(N)||\overline{x}^+(N) - x_s(\hat{\beta})| + \gamma|\hat{x}^+(N) - \overline{x}^+(N)|^2 \\
&\qquad \leq \sum_{k=0}^{N-1} 2\overline{\sigma}(Q)c_x L_f'(k)|\delta\hat{x}||\tilde{d}| + \overline{\sigma}(Q)(L_f'(k))^2|\tilde{d}|^2 \\
&\qquad\quad + 2\gamma c_x\gamma_f L_f'(N)|\delta\hat{x}||\tilde{d}| + \gamma(L_f'(N))^2|\tilde{d}|^2 \\
&\qquad \leq \sum_{k=0}^{N-1} \lambda\overline{\sigma}(Q)c_x L_f'(k)|\delta\hat{x}|^2 + (\overline{\sigma}(Q)(L_f'(k))^2 + \lambda^{-1}\overline{\sigma}(Q)c_x L_f'(k))|\tilde{d}|^2 \\
&\qquad\quad + \lambda\gamma c_x\gamma_f L_f'(N)|\delta\hat{x}|^2 + (\gamma(L_f'(N))^2 + \lambda^{-1}\gamma c_x\gamma_f L_f'(N))|\tilde{d}|^2 \\
&\qquad \leq \sum_{k=0}^{N-1} \lambda L_1(k)|\delta\hat{x}|^2 + L_2(k,\lambda)|\tilde{d}|^2 + \lambda L_1(N)|\delta\hat{x}|^2 + L_2(N,\lambda)|\tilde{d}|^2
\end{aligned}$$

where $L_1(k) := \overline{\sigma}(Q)c_x L_f'(k)$ and $L_2(k,\lambda) := \overline{\sigma}(Q)(L_f'(k))^2 + \lambda^{-1}\overline{\sigma}(Q)c_x L_f'(k)$ for each $k \in \mathbb{I}_{0:N-1}$, and $L_1(N) := \gamma c_x\gamma_f L_f'(N)$ and $L_2(N,\lambda) := \gamma(L_f'(N))^2 + \lambda^{-1}\gamma c_x\gamma_f L_f'(N)$.

Finally, we compile the above results,

$$\begin{aligned}&\left||\hat{x}^+(N) - x_s(\hat{\beta}^+)|^2_{P_f(\hat{\beta}^+)} - |\overline{x}^+(N) - x_s(\hat{\beta})|^2_{P_f(\hat{\beta})}\right| \\ &\overset{(105)}{\leq} \left||\hat{x}^+(N) - x_s(\hat{\beta}^+)|^2_{P_f(\hat{\beta})} - |\overline{x}^+(N) - x_s(\hat{\beta})|^2_{P_f(\hat{\beta})}\right| + \sigma_{P_f,x}(|\tilde{d}|)|\delta\hat{x}|^2 + \sigma_{P_f,d}(|\tilde{d}|)|\tilde{d}|^2 \\ &\overset{(106)}{\leq} \left||\hat{x}^+(N) - x_s(\hat{\beta}^+)|^2_{P_f(\hat{\beta})} - |\overline{x}^+(N) - x_s(\hat{\beta})|^2_{P_f(\hat{\beta})}\right| \\ &\qquad + (\sigma_{P_f,x}(|\tilde{d}|) + \lambda\hat{L}_1(N))|\delta\hat{x}|^2 + (\sigma_{P_f,d}(|\tilde{d}|) + \hat{L}_2(N,\lambda))|\tilde{d}|^2 \end{aligned} \tag{109}$$

and therefore

$$\begin{aligned}&|V_N(\hat{x}^+, \tilde{\mathbf{u}}, \hat{\beta}^+) - V_N(\hat{x}^+, \tilde{\mathbf{u}}, \hat{\beta})| \\ &\overset{(107)\text{–}(109)}{\leq} \sum_{k=0}^{N-1} \lambda(\hat{L}_1(k) + \tilde{L}_1(k))|\delta\hat{x}|^2 + (\hat{L}_2(k,\lambda) + \tilde{L}_2(k,\lambda))|\tilde{d}|^2 \\ &\qquad + (\sigma_{P_f,x}(|\tilde{d}|) + \lambda\hat{L}_1(N))|\delta\hat{x}|^2 + (\sigma_{P_f,d}(|\tilde{d}|) + \hat{L}_2(N,\lambda))|\tilde{d}|^2\end{aligned}$$

Finally (95) holds so long as $|\tilde{d}| \leq \delta$, with

$$\begin{aligned}a_{V_N,1} &:= \sigma_{P_f,x}(\delta) + \lambda\left(L_1(N) + \hat{L}_1(N) + \sum_{k=0}^{N-1} \overline{L}_1(k)\right) \\ a_{V_N,2} &:= \sigma_{P_f,d}(\delta) + L_2(N,\lambda) + \hat{L}_2(N,\lambda) + \sum_{k=0}^{N-1} \overline{L}_2(k,\lambda)\end{aligned}$$

where $\overline{L}_1(k) := L_1(k) + \hat{L}_1(k) + \tilde{L}_1(k)$ and $\overline{L}_2(k,\lambda) := L_2(k,\lambda) + \hat{L}_2(k,\lambda) + \tilde{L}_2(k,\lambda)$. To ensure $a_{V_N,1} < \underline{\sigma}(Q)$, we can simply choose $\lambda \in \left(0, \frac{\underline{\sigma}(Q) - \sigma_{P_f,x}(\delta)}{L_1(N) + \hat{L}_1(N) + \sum_{k=0}^{N-1} \overline{L}_1(k)}\right)$ and $\delta \in (0, \sigma^{-1}_{P_f,x}(\underline{\sigma}(Q)))$. □

Proof of Proposition 4. For convenience, we define $\tilde{\mathbf{u}} := \tilde{\mathbf{u}}(\hat{x}, \hat{\beta})$. From Propositions 10 and 11, we have $a_{V_N,1} \in (0, \underline{\sigma}(Q))$, $a_{V_N,2}, \tilde{c}_e, \delta, \delta_w > 0$, and $\tilde{\sigma}_w, \tilde{\sigma}_\alpha \in \mathcal{K}_\infty$ such that

$$|V_N(\hat{x}^+, \tilde{\mathbf{u}}, \hat{\beta}^+) - V_N(\overline{x}^+, \tilde{\mathbf{u}}, \hat{\beta})| \leq (a_{V_N,1} + \tilde{\sigma}_w(|w_{\mathrm{P}}|))|\delta\hat{x}|^2 + a_{V_N,2}c_e|(e, e^+)|^2 + \tilde{\sigma}_\alpha(|\Delta\alpha|)$$

so long as $\tilde{d} \in \tilde{\mathbb{D}}_c(\hat{x}, \hat{\beta}) \cap \delta\mathbb{B}^{n_{\tilde{d}}}$, $\alpha \in \mathcal{A}_c(\delta_w)$, and $\Delta\alpha \in \mathbb{A}_c(\alpha, \delta_w)$. Without loss of generality, assume $\delta_w < \tilde{\sigma}_w^{-1}(\underline{\sigma}(Q) - a_{V_N,1})$. By Proposition 1, we can choose $\delta > 0$ such that $\tilde{\mathbf{u}} \in \mathcal{U}_N(\hat{x}^+, \hat{\beta}^+)$, so

$$\begin{aligned}V_N^0(\hat{x}^+, \hat{\beta}^+) &\leq V_N(\hat{x}^+, \tilde{\mathbf{u}}, \hat{\beta}^+) \\ &\leq V_N(\overline{x}^+, \tilde{\mathbf{u}}, \hat{\beta}) + (a_{V_N,1} + \tilde{\sigma}_w(\delta_w))|\delta\hat{x}|^2 + a_{V_N,2}c_e|(e, e^+)|^2 + \tilde{\sigma}_\alpha(|\Delta\alpha|) \\ &\leq V_N^0(\hat{x}, \hat{\beta}) - (\underline{\sigma}(Q) - a_{V_N,1} - \tilde{\sigma}_w(\delta_w))|\delta\hat{x}|^2 + a_{V_N,2}c_e|(e, e^+)|^2 + \tilde{\sigma}_\alpha(|\Delta\alpha|).\end{aligned}$$

where the first inequality follows by optimality and the third inequality follows by (35). Thus, (53) holds with $\tilde{a}_3 := \underline{\sigma}(Q) - a_{V_N,1} - \tilde{\sigma}_w(\delta_w) > 0$ and $\tilde{a}_4 := a_{V_N,2}c_e > 0$. □

C Establishing steady-state target problem assumptions

C.1 Proof of Lemma 1

Proof of Lemma 1. First, note that M_1 full row rank implies $n_r \leq n_u$. Consider the function

$$\mathbf{f}_1(z_s, \beta) := \begin{bmatrix} f(x_s, u_s, d) - x_s \\ g(u, h(x_s, u_s, d)) - r_{\text{sp}} \end{bmatrix}$$

and define the objective and Lagrangian

$$\begin{aligned} \phi(z_s, \beta) &:= \ell_s(u_s - u_{\text{sp}}, y_s(z_s, \beta) - y_{\text{sp}}) \\ \mathcal{L}(z_s, \beta, \lambda) &:= \phi(z_s, \beta) + \lambda^\top \mathbf{f}_1(z_s, \beta) \end{aligned}$$

where $z_s := (x_s, u_s)$, $y_s(z_s, \beta) := h(x_s, u_s, d)$, and $\beta := (r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, d)$. The first-order derivatives of the Lagrangian are

$$\begin{aligned} \partial_{z_s} \mathcal{L}(z_s, \beta, \lambda) &= \partial_{z_s} \phi(z_s, \beta) + [\partial_{z_s} \mathbf{f}_1(z_s, \beta)]^\top \lambda \\ \partial_\lambda \mathcal{L}(z_s, \beta, \lambda) &= \mathbf{f}_1(z_s, \beta). \end{aligned}$$

The goal of the proof is to use the implicit function theorem on $\partial_{(z_s,\lambda)} \mathcal{L}(z_s, \beta, \lambda)$ to establish Lipschitz continuity of the SSTP solution map $z_s(\cdot)$. We already have $\partial_{(z_s,\lambda)} \mathcal{L}(0,0,0) = 0$ by assumption. Next, we aim to show $\partial_{(z_s,\lambda)} \mathcal{L}(z_s, \beta, \lambda) = 0$ is a necessary and sufficient condition for solving (6).

First, we have the partial derivatives $\partial_{z_s} \mathbf{f}_1(0,0) = M_1$, which is full row rank by assumption. By continuity of $\partial_{z_s} \mathbf{f}_1$, there exist constants $\varepsilon_1, \delta_1 > 0$ such that $\partial_{z_s} \mathbf{f}_1(z_s, \beta)$ is full row rank for all $|z_s| \leq \varepsilon_1$ and $|\beta| \leq \delta_1$. Then, so long as (z_s, β) are kept sufficiently small, the linear independence constraint qualification holds, and $\partial_{(z_s,\lambda)} \mathcal{L}(z_s, \beta, \lambda) = 0$ is a necessary condition for solving (6).

Consider the following second-order derivatives:

$$\begin{aligned} \partial_{z_s}^2 \mathcal{L}(0,0,0) &= M_3^\top \partial_{(u,y)}^2 \ell_s(0,0) M_3 \\ \partial_{z_s} \partial_\lambda \mathcal{L}(0,0,0) &= \partial_{z_s} \mathbf{f}_1(0,0) = M_1 \\ \partial_\lambda^2 \mathcal{L}(0,0,0) &= 0 \end{aligned}$$

where $M_3 := \left[\begin{smallmatrix} 0 & I \\ C & D \end{smallmatrix}\right]$.[8] We have $\partial_{(z_s,\lambda)} \mathcal{L}(z_s, \beta, \lambda) = 0$ is a sufficient condition for solving (6) if

$$d^\top \partial_{z_s}^2 \mathcal{L}(z_s, \beta, \lambda) d > 0$$

for all $d \in \mathcal{N}(\partial_{z_s} \mathbf{f}_1(z_s, \beta)) \setminus \{ 0 \}$. We require the following intermediate result.

Lemma 5. *For each $A = A^\top \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times n}$, we have $x^\top A x > 0$ for all $x \in \mathcal{N}(B) \setminus \{ 0 \}$ if and only if $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right]$ is full column rank.*

[8] The second-order derivatives of $y_s(z_s, \beta)$ and $\mathbf{f}_1(z_s, \beta)$ vanish since $\partial_{(u,y)} \ell_s(0,0)$ and $y_s(0,0) = 0$ (by assumption) and we have set $\lambda = 0$.

Proof. First, note that $\mathcal{N}(A + B^\top B) = \mathcal{N}(\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right])$, so $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right]$ is full column rank is equivalent to $A + B^\top B$ being positive definite.

($\Rightarrow$) Suppose $x^\top A x > 0$ for all $x \in \mathcal{N}(B) \setminus \{0\}$. Then $x^\top(A + B^\top B)x \geq x^\top A x > 0$ for all $x \in \mathcal{N}(B) \setminus \{0\}$ and $x^\top(A + B^\top B)x \geq x^\top B^\top B x > 0$ for all $x \notin \mathcal{N}(B)$, so $A + B^\top B$ is positive definite.

($\Leftarrow$) Suppose $A + B^\top B$ is positive definite. Then $x^\top A x = x^\top(A + B^\top B)x > 0$ for all $x \in \mathcal{N}(B) \setminus \{0\}$. ☐

Thus, it suffices to show

$$\begin{bmatrix} \partial^2_{z_s} \mathcal{L}(z_s, \beta, \lambda) \\ \partial_{z_s} \mathbf{f}_1(z_s, \beta) \end{bmatrix} \tag{110}$$

is full column rank. Since $\partial^2_{(u,y)} \ell_s(0,0)$ is positive definite, $\mathcal{N}(\partial^2_{z_s} \mathcal{L}(0,0,0)) = \mathcal{N}(M_3)$. Then with $M_4 := \left[\begin{smallmatrix} \partial^2_{z_s} \mathcal{L}(0,0,0) \\ \partial_{z_s} \mathbf{f}_1(0,0) \end{smallmatrix}\right]$ we have

$$\begin{aligned} \mathcal{N}(M_4) &= \mathcal{N}(\partial^2_{z_s} \mathcal{L}(0,0,0)) \cap \mathcal{N}(\partial_{z_s} \mathbf{f}_1(0,0)) \\ &= \mathcal{N}(M_3) \cap \mathcal{N}(M_1) \\ &= \mathcal{N}\left(\begin{bmatrix} M_3 \\ M_1 \end{bmatrix}\right) = \{0\} \end{aligned}$$

where the last equality follows from the fact that

$$\begin{bmatrix} M_3 \\ M_1 \end{bmatrix} = \left[\begin{smallmatrix} 0 & I \\ C & D \\ A-I & B \\ H_y C & H_u + H_y D \end{smallmatrix}\right]$$

is full column rank, as it is the row permutation of a block triangular matrix with full column rank diagonal blocks I and $\left[\begin{smallmatrix} A-I \\ C \end{smallmatrix}\right]$.[9] Therefore M_4 is full column rank, and because (110) is continuous, there exist $\varepsilon_2, \delta_2, \gamma_2 > 0$ for which (110) is full column rank for all $|z_s| \leq \varepsilon_2$, $|\beta| \leq \delta_2$, and $|\lambda| \leq \gamma_2$. Therefore, so long as (z_s, β, λ) are kept sufficiently small, $\partial_{(z_s,\lambda)} \mathcal{L}(z_s, \beta, \lambda) = 0$ is in fact a necessary and sufficient condition for solving (6).

Now we are able to solve (6). We have the derivatives

$$\partial^2_{(z_s,\lambda)} \mathcal{L}(0,0,0) = \begin{bmatrix} M_3^\top \partial^2_{(u,y)} \ell_s(0,0) M_3 & M_1^\top \\ M_1 & 0 \end{bmatrix}.$$

According to (Magnus and Neudecker, 2019, Thm. 3.21), we have the nullspace relationship

$$\mathcal{N}(\partial^2_{(z_s,\lambda)} \mathcal{L}(0,0,0)) = \mathcal{N}\left(\begin{bmatrix} V_0 & \\ & W_0 \end{bmatrix}\right) \tag{111}$$

where

$$\begin{aligned} V_0 &:= M_3^\top \partial^2_{(u,y)} \ell_s(0,0) M_3 + M_1^\top M_1 = \begin{bmatrix} M_3 \\ M_1 \end{bmatrix}^\top \begin{bmatrix} \partial^2_{(u,y)} \ell_s(0,0) & \\ & I \end{bmatrix} \begin{bmatrix} M_3 \\ M_1 \end{bmatrix} \\ W_0 &:= M_1 V_0^+ M_1^\top. \end{aligned}$$

[9] Full column rank of $\left[\begin{smallmatrix} A-\lambda I \\ C \end{smallmatrix}\right]$ for all $|\lambda| \geq 1$ follows from detectability of (A, C).

Recall $\begin{bmatrix} M_3 \\ M_1 \end{bmatrix}$ is full column rank and $\partial^2_{(u,y)}\ell_s(0,0)$ is invertible, so V_0 is invertible. Likewise, M_1 full row rank and V_0 invertible implies that W_0 is invertible. Finally, $\begin{bmatrix} V_0 & \\ & W_0 \end{bmatrix}$ is invertible, and by (111), $\partial^2_{(z_s,\lambda)}\mathcal{L}(0,0,0)$ is invertible. By the implicit function theorem (Rudin, 1976, Thm. 9.24) there exist $\delta_3 > 0$ and continuously differentiable functions $\mathbf{g}_1 : \mathbb{R}^{n_\beta} \to \mathbb{R}^{n+n_u}$ and $\mathbf{g}_\lambda : \mathbb{R}^{n_\beta} \to \mathbb{R}^{n+n_r}$ such that $\mathbf{g}_1(0) = 0$, $\mathbf{g}_\lambda(0) = 0$, and $\partial_{(\alpha,\lambda)}\mathcal{L}(\mathbf{g}_1(\beta), \beta, \mathbf{g}_\lambda(\beta)) = 0$ for all $|\beta| \leq \delta_3$.

For convenience, we define the functions

$$\begin{aligned} \mathbf{g}_1(\beta) &=: (x_s(\beta), u_s(\beta)) \\ \tilde{c}(\beta) &:= \max_{1\leq i\leq n_c} c_i(u_s(\beta), h(x_s(\beta), u_s(\beta), d)) + b_i \end{aligned}$$

for each $\beta = (r_{\text{sp}}, z_{\text{sp}}, d) \in \mathcal{B}$, which are continuous because $\mathbf{g}_1$, h, and c are continuous. Moreover, $\mathbb{X}, \mathbb{U}$ contain neighborhoods of the origin and $\tilde{c}(0) < 0$ by assumption, so there exists $\delta_3 > 0$ for which $z_s(\beta) \in \mathbb{X} \times \mathbb{U}$ and $\tilde{c}(\beta) \leq 0$ for all $|\beta| \leq \delta_3$. Let $\delta < \delta_4 := \min\{\delta_1, \delta_2, \delta_3\}$, $\delta_0 := \delta_4 - \delta$, $\mathcal{B}_c := \delta\mathbb{B}^{n_\beta}$, and $\overline{\mathcal{B}}_c := \delta_4\mathbb{B}^{n_\beta}$. Defining $\hat{\mathcal{B}}_c$ as in Assumption 7(i), we have $|\hat{\beta}| \leq |\beta| + |e_d| \leq \delta + \delta_0 = \delta_4$ for each $\hat{\beta} = (s_{\text{sp}}, \hat{d}) \in \hat{\mathcal{B}}_c$, and therefore $\mathcal{B}_c \subseteq \hat{\mathcal{B}}_c \subseteq \overline{\mathcal{B}}_c \subseteq \mathcal{B}$. Moreover, $(x_s(\hat{\beta}), u_s(\hat{\beta})) \in \mathcal{Z}_O(r_{\text{sp}}, \hat{d})$ and $(x_s(\hat{\beta}), u_s(\hat{\beta}))$ uniquely solve (6) and are continuously differentiable for each $\hat{\beta} = (s_{\text{sp}}, \hat{d}) \in \hat{\mathcal{B}}_c$. Finally, Assumption 7 is satisfied by z_s, $\mathcal{B}_c \subseteq \mathcal{B}$, and $\delta_0 > 0$. □

C.2 Proof of Lemma 2

Proof of Lemma 2. Recall from the proof of Lemma 1 that M_1 full row rank implies $n_r \leq n_u$. Moreover, M_2 invertible implies $n_d = n_y$. Consider the functions

$$\begin{aligned} \mathbf{f}_2(z_s, \mathbf{x}_s, \alpha) &:= \begin{bmatrix} f_{\text{P}}(x_{\text{P,s}}, u_s, w_{\text{P}}) - x_{\text{P,s}} \\ h_{\text{P}}(x_{\text{P,s}}, u_s, w_{\text{P}}) - h(x_s, u_s, d_s) \end{bmatrix} \\ \mathbf{f}(z_s, \mathbf{x}_s, \alpha) &:= \begin{bmatrix} \mathbf{f}_1(z_s, \beta) \\ \mathbf{f}_2(z_s, \mathbf{x}_s, \alpha) \end{bmatrix} \end{aligned}$$

where $z_s := (x_s, u_s)$, $\mathbf{x}_s := (x_{\text{P,s}}, d_s)$, $\alpha := (r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, w_{\text{P}})$, $\beta := (r_{\text{sp}}, u_{\text{sp}}, y_{\text{sp}}, d_s)$, and $\mathbf{f}_1$ is defined in the proof of Lemma 1. Defining ϕ and $\mathcal{L}$ as in the proof of Lemma 1, we seek to use the implicit function theorem on

$$\begin{aligned} \mathbf{h}(z_s, \mathbf{x}_s, \lambda, \alpha) &:= \begin{bmatrix} \partial_{(z_s,\lambda)}\mathcal{L}(z_s, \mathbf{x}_s, \beta, \lambda) \\ \mathbf{f}_2(z_s, \mathbf{x}_s, \beta) \end{bmatrix} \\ &= \begin{bmatrix} \partial_{z_s}\phi(z_s, \mathbf{x}_s, \beta) + [\partial_{z_s}\mathbf{f}_1(z_s, \mathbf{x}_s, \beta)]^\top \lambda \\ \mathbf{f}(\alpha, \beta) \end{bmatrix} \end{aligned}$$

which is the combination of the stationary point condition for the Lagrangian of (6) with the steady-state disturbance problem (47). We already have $\mathbf{h}(0,0,0,0) = 0$ by assumption. From the proof of Lemma 1, there exists $\delta_1 > 0$ such that, for all $|(\mathbf{x}_s, \alpha)| \leq \delta_1$, $\partial_{(z_s,\lambda)}\mathcal{L}(z_s, \mathbf{x}_s, \alpha, \lambda) = 0$ is a necessary and sufficient condition for solving (6). Thus, if we keep $|(\mathbf{x}_s, \alpha)| \leq \delta_1$ sufficiently small, then $\mathbf{h}(z_s, \mathbf{x}_s, \lambda, \alpha) = 0$ is necessarily and sufficient for simultaneously solving (6) and (47).

Defining the invertible matrices

$$T_1 := \begin{bmatrix} I_n & 0 & 0 & 0 \\ 0 & I_{n_r} & 0 & 0 \\ I_n & 0 & -I_n & 0 \\ 0 & 0 & 0 & I_{n_y} \end{bmatrix}, \qquad T_2 := \begin{bmatrix} I_n & 0 & 0 & 0 \\ 0 & I_{n_u} & 0 & 0 \\ I_n & 0 & -I_n & 0 \\ 0 & 0 & 0 & I_{n_d} \end{bmatrix},$$

We have

$$T_1 \partial_{(z_s, \mathbf{x}_s)} \mathbf{f}(0,0,0) T_2 = \begin{bmatrix} M_1 & * \\ 0 & M_2 \end{bmatrix}.$$

We can write the derivatives

$$\partial_{(z_s, \mathbf{x}_s, \lambda)} \mathbf{h}(0,0,0,0) = \begin{bmatrix} M_3^\top \partial^2_{(u,y)} \ell_s(0,0) M_5 & M_1^\top \\ \partial_{(z_s, \mathbf{x}_s)} \mathbf{f}(0,0,0) & 0 \end{bmatrix}$$

where M_3 is defined as in the proof of Lemma 1, and $M_5 := \left[\begin{smallmatrix} 0 & I & 0 & 0 \\ C & D & 0 & C_d \end{smallmatrix}\right]$. Note that $M_5 T_2 = M_5$ and $M_5 = \begin{bmatrix} M_3 & * \end{bmatrix}$. Define the invertible matrices

$$T_3 := \begin{bmatrix} I_{n+n_u} & \\ & T_1 \end{bmatrix}, \qquad T_4 := \begin{bmatrix} T_2 & \\ & I_{n+n_d} \end{bmatrix}, \qquad P := \begin{bmatrix} I_{n+n_u} & 0 & 0 \\ 0 & 0 & I_{n+n_d} \\ 0 & I_{n+n_r} & 0 \end{bmatrix}.$$

Then we can write

$$T_3 \partial_{(z_s, \mathbf{x}_s, \lambda)} \mathbf{h}(0,0,0,0) T_4 P = \begin{bmatrix} M_3^\top \partial^2_{(u,y)} \ell_s(0,0) M_3 & M_1^\top & * \\ M_1 & 0 & * \\ 0 & 0 & M_2 \end{bmatrix}. \tag{112}$$

But M_2 is invertible by assumption, and $\left[\begin{smallmatrix} M_3^\top \partial^2_{(u,y)} \ell_s(0,0) M_3 & M_1^\top \\ M_1 & 0 \end{smallmatrix}\right]$ was shown to be invertible in the proof of Lemma 1, so $\partial_{(z_s, \mathbf{x}_s, \lambda)} \mathbf{h}(0,0,0,0)$ must be invertible. By the implicit function theorem (Rudin, 1976, Thm. 9.24) there exist $\delta_2 > 0$ and continuously differentiable functions $\mathbf{g} : \mathbb{R}^{n_\alpha} \to \mathbb{R}^{2n+n_u+n_d}$ and $\mathbf{g}_\lambda : \mathbb{R}^{n_\alpha} \to \mathbb{R}^{2n+n_r+n_y}$ (where $\mathcal{A} := \mathbb{R}^{n_r} \times \overline{\mathbb{Z}}_y \times \mathbb{W}$) such that $\mathbf{g}(0) = 0$, $\mathbf{g}_\lambda(0) = 0$, and $\partial_{(z_s, \mathbf{x}_s, \lambda)} \mathcal{L}(\mathbf{g}(\alpha), \alpha, \mathbf{g}_\lambda(\alpha)) = 0$ for all $|\alpha| \leq \delta_2$.

As in the proof of Lemma 1, we define the functions

$$\begin{aligned} \mathbf{g}(\alpha) &=: (x_s(\alpha), u_s(\alpha), x_{\mathrm{P,s}}(\alpha), d_s(\alpha)) \\ \tilde{c}(\alpha) &:= \max_{1 \leq i \leq n_c} c_i(u_s(\alpha), h_{\mathrm{P}}(x_{\mathrm{P,s}}(\alpha), u_s(\alpha), w_{\mathrm{P}})) + b_i \end{aligned}$$

for each $\alpha = (r_{\mathrm{sp}}, z_{\mathrm{sp}}, w_{\mathrm{P}}) \in \mathbb{R}^{n_\alpha}$, which are continuous because $\mathbf{g}$, h_{P}, and c are continuous. From Lemma 1, we already have a set $\mathcal{B}_c \subseteq \mathcal{B}$ containing a neighborhood of the origin and continuously differentiable functions (with a slight abuse of notation) $(x_s, u_s) : \mathcal{B} \to \mathbb{X} \times \mathbb{U}$ that uniquely solve (6) (and satisfies Assumption 7). Since $\mathbb{X}, \mathbb{U}, \mathbb{D}, \mathcal{B}_c$ contain neighborhoods of the origin, there must exist $\delta_3 > 0$ such that $\mathbf{g}(\alpha) \in \mathbb{X} \times \mathbb{U} \times \mathbb{X} \times \mathbb{D}$, $\beta = (r_{\mathrm{sp}}, z_{\mathrm{sp}}, d_s(\alpha)) \in \mathcal{B}_c$, $|(\mathbf{x}_s(\alpha), \alpha)| \leq \delta_2$, and $\tilde{c}(\beta) \leq 0$ for all $|\alpha| \leq \delta_3$. Therefore $(x_s(\alpha), u_s(\alpha))$ are also the unique solutions to (6) with $\beta = (r_{\mathrm{sp}}, z_{\mathrm{sp}}, d_s(\alpha))$, i.e., $(x_s(\alpha), u_s(\alpha)) = (x_s(\beta), u_s(\beta))$, and all parts of Assumption 8 are satisfied with $(x_s, u_s) : \mathcal{B} \to \mathbb{X} \times \mathbb{U}$, $(x_{\mathrm{P,s}}, d_s) : \mathcal{A}_c \to \mathbb{X} \times \mathbb{D}$, $\mathcal{A}_c := \delta \mathbb{B}^{n_\alpha}$, and $\delta := \min \{ \delta_1, \delta_2, \delta_3 \} > 0$. □

References

K. R. Muske and T. A. Badgwell, "Disturbance modeling for offset-free linear model predictive control," *J. Proc. Cont.*, vol. 12, no. 5, pp. 617–632, 2002.

G. Pannocchia and J. B. Rawlings, "Disturbance models for offset-free MPC control," *AIChE J.*, vol. 49, no. 2, pp. 426–437, 2003.

M. Morari and U. Maeder, "Nonlinear offset-free model predictive control," *Automatica*, vol. 48, no. 9, pp. 2059–2067, 2012.

G. Pannocchia, M. Gabiccini, and A. Artoni, "Offset-free MPC explained: novelties, subtleties, and applications," *IFAC–P. Online*, vol. 48, no. 23, pp. 342–351, 2015.

M. Vaccari and G. Pannocchia, "A Modifier-Adaptation Strategy Towards Offset-Free Economic MPC," *Processes*, vol. 5, no. 1, p. 2, 2017.

G. Pannocchia, "An economic MPC formulation with offset-free asymptotic performance," *IFAC–P. Online*, vol. 51, no. 18, pp. 393–398, 2018.

T. Faulwasser and G. Pannocchia, "Toward a Unifying Framework Blending Real-Time Optimization and Economic Model Predictive Control," *Ind. Eng. Chem. Res.*, vol. 58, no. 30, pp. 13 583–13 598, 2019.

M. Vaccari, D. Bonvin, F. Pelagagge, and G. Pannocchia, "Offset-Free Economic MPC Based on Modifier Adaptation: Investigation of Several Gradient-Estimation Techniques," *Processes*, vol. 9, no. 5, p. 901, 2021.

P. Falugi, "Model predictive control for tracking randomly varying references," *Int. J. Control*, vol. 88, no. 4, pp. 745–753, 2015.

D. Limon, A. Ferramosca, I. Alvarado, and T. Alamo, "Nonlinear MPC for Tracking Piece-Wise Constant Reference Signals," *IEEE Trans. Auto. Cont.*, vol. 63, no. 11, pp. 3735–3750, 2018.

J. Köhler, M. A. Müller, and F. Allgöwer, "A nonlinear tracking model predictive control scheme for dynamic target signals," *Automatica*, vol. 118, p. 109030, 2020.

J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Linear Tracking MPC for Nonlinear Systems—Part I: The Model-Based Case," *IEEE Trans. Auto. Cont.*, vol. 67, no. 9, pp. 4390–4405, 2022.

G. Galuppini, L. Magni, and A. Ferramosca, "Nonlinear MPC for Tracking Piecewise-Constant Reference Signals: the Positive Semidefinite Stage Cost Case," *IFAC–P. Online*, vol. 56, no. 1, pp. 210–215, 2023.

R. Soloperto, J. Köhler, and F. Allgöwer, "A Nonlinear MPC Scheme for Output Tracking Without Terminal Ingredients," *IEEE Trans. Auto. Cont.*, vol. 68, no. 4, pp. 2368–2375, 2023.

S. J. Kuntz and J. B. Rawlings, “Beyond inherent robustness: asymptotic stability of MPC despite plant-model mismatch,” 2024, arXiv:2411.15452 [eecs, math]. [Online]. Available: https://arxiv.org/abs/2411.15452

R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis.* Springer-Verlag, 1998.

J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, *Model Predictive Control: Theory, Design, and Computation*, 2nd ed. Santa Barbara, CA: Nob Hill Publishing, 2020, 770 pages, ISBN 978-0-9759377-5-4.

D. A. Allan and J. B. Rawlings, “Robust stability of full information estimation,” *SIAM J. Cont. Opt.*, vol. 59, no. 5, pp. 3472–3497, 2021.

J. D. Schiller, S. Muntwiler, J. Köhler, M. N. Zeilinger, and M. A. Müller, “A Lyapunov Function for Robust Stability of Moving Horizon Estimation,” *IEEE Trans. Auto. Cont.*, vol. 68, no. 12, pp. 7466–7481, 2023.

S. Muntwiler, J. Köhler, and M. N. Zeilinger, “MHE under parametric uncertainty – Robust state estimation without informative data,” 2023, arXiv:2312.14049 [cs, eess]. [Online]. Available: http://arxiv.org/abs/2312.14049

J. D. Schiller and M. A. Müller, “A moving horizon state and parameter estimation scheme with guaranteed robust convergence,” *IFAC–P. Online*, vol. 56, no. 2, pp. 6759–6764, 2023.

D. N. Tran, C. M. Kellett, and P. M. Dower, “Input-to-state stability with respect to two measurement functions: Discrete-time systems,” in *2015 54th IEEE Conference on Decision and Control (CDC)*, 2015, pp. 1817–1822.

E. D. Sontag and Y. Wang, “On the characterization of the input to state stability property,” *Sys. Cont. Let.*, vol. 24, pp. 351–359, 1995.

——, “Notions of input to output stability,” *Sys. Cont. Let.*, vol. 38, no. 4-5, pp. 235–248, 1999.

E. Sontag and Y. Wang, “Lyapunov Characterizations of Input to Output Stability,” *SIAM J. Cont. Opt.*, vol. 39, no. 1, pp. 226–249, 2000.

J. B. Rawlings and L. Ji, “Optimization-based state estimation: Current status and some new results,” *J. Proc. Cont.*, vol. 22, pp. 1439–1444, 2012.

J.-G. Sun, “Perturbation Theory for Algebraic Riccati Equations,” *SIAM J. Matrix Anal. and Appl.*, vol. 19, no. 1, pp. 39–65, 1998.

J. Hale, *Ordinary Differential Equations*, 2nd ed. Robert E. Krieger Publishing Company, 1980.

U. Maeder and M. Morari, “Offset-free reference tracking with model predictive control,” *Automatica*, vol. 46, no. 9, pp. 1469–1476, 2010.

M. A. Zagrobelny, "MPC performance monitoring and disturbance model identification," Ph.D. dissertation, University of Wisconsin–Madison, December 2014. [Online]. Available: https://engineering.ucsb.edu/~jbraw/jbrweb-archives/theses/zagrobelny.pdf

G. A. Hicks and W. H. Ray, "Approximation methods for optimal control synthesis," *Can. J. Chem. Eng.*, vol. 49, pp. 522–528, Aug 1971.

S. Kameswaran and L. T. Biegler, "Simultaneous dynamic optimization strategies: Recent advances and challenges," *Comput. Chem. Eng.*, vol. 30, pp. 1560–1575, Sep 2006.

Z.-P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, pp. 857–869, 2001.

——, "A converse Lyapunov theorem for discrete-time systems with disturbances," *Sys. Cont. Let.*, vol. 45, pp. 49–58, 2002.

D. A. Allan, C. N. Bates, M. J. Risbeck, and J. B. Rawlings, "On the inherent robustness of optimal and suboptimal nonlinear MPC," *Sys. Cont. Let.*, vol. 106, pp. 68 – 78, 2017.

J. M. Lee, *Introduction to Smooth Manifolds*, ser. Graduate Texts in Mathematics. New York, NY: Springer, 2012, vol. 218.

T. M. Apostol, *Mathematical analysis.* Addison-Wesley, 1974.

J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 3rd ed. Hoboken, NJ: John Wiley & Sons Ltd., 2019.

W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.