Anomaly-Handling in Lyapunov-Based Economic Model Predictive Control via Empirical Models [⋆]

Helen Durand [*]

[*] *Wayne State University, Detroit, MI 48202 USA (e-mail: helen.durand@wayne.edu).*

Abstract: A question that faces data-driven autonomous systems is verification that they will perform in a safe manner despite changes in the environment on which they act over time or incomplete knowledge of the system model. This work analyzes closed-loop stability of nonlinear systems under Lyapunov-based economic model predictive control (LEMPC) with data-driven models in the case where it is desirable to have the ability to detect when the data-driven model is or becomes insufficiently accurate for maintaining the closed-loop state in an expected region of state-space. Implications of the results for false sensor measurement cyberattacks seeking to impact the fidelity of models derived from model identification are discussed and illustrated through a chemical process example.

Keywords: model predictive control, anomaly response, chemical process control, nonlinear systems, empirical modeling, cybersecurity.

1. INTRODUCTION

Process safety has been an area of significant research attention in recent years (e.g., Ahooyi et al. (2016)). A number of works have looked at how controllers can be enhanced to drive the closed-loop state into safe operating regions Albalawi et al. (2016) or how they can be reconfigured via model updates to handle faults. For example, Armaou and Demetriou (2008) develops a state estimation-based technique for updating a system model as component faults occur that change the process dynamics, and Xue and El-Farra (2018) and Mahmood and Mhaskar (2010) develop techniques for handling actuator faults in Lyapunov-based model predictive control.

In Durand (2020), we explored safety considerations in the sense of anomaly responsiveness for a specific control design known as LEMPC Heidarinejad et al. (2012) (where we define "anomalies" as changes in the dynamic model of the process). LEMPC is an optimization-based controller (a type of economic model predictive controller (EMPC) Ellis et al. (2014); Rawlings et al. (2012) that optimizes a general stage cost) that uses a dynamic process model to aid in selecting optimal control actions. Compared to Alanqar et al. (2017a), where data-driven (empirical) models in EMPC were updated when the error between predictions of the process state and state measurements went above a data-derived threshold, Durand (2020) provided theoretical conditions for closed-loop stability for an EMPC with data-driven models that update over time and suggested a means for triggering model updates based on the control theory. It was suggested in Durand (2020) that the anomaly-handling framework could be a step toward allowing the LEMPC with the data-driven models to be verified to maintain closed-loop stability (safety) over time. In this work, we clarify the method and analyze this consideration in more detail with respect to its benefits and limitations, suggesting potential ramifications for the cyberattack resilience of the strategy.

2. PRELIMINARIES

2.1 Notation

$|\cdot|$ denotes the vector Euclidean norm. $\alpha : [0,a) \to [0,\infty)$ is a class $\mathcal{K}$ function if it is continuous, strictly increasing, and $\alpha(0) = 0$. The notation Ω_ρ defines a level set of a scalar-valued function V (i.e., $\Omega_\rho := \{x \in R^n \ : \ V(x) \leq \rho\}$). The notation $'/'$ signifies set subtraction (i.e., $A/B := \{x \in R^n \ : \ x \in A, \ x \notin B\}$). The transpose of a vector x is represented by x^T. We define a sampling time as $t_k := k\Delta$, $k = 0, 1, \ldots$, where Δ is a sampling period. The "floor" function returns the largest integer less than the argument.

2.2 Class of Systems

This work considers nonlinear systems of the form:

$$\dot{x}_{a,i} = f_i(x_{a,i}(t), u(t), w_i(t)) \tag{1}$$

where the state, input, and disturbance vectors are denoted by $x_{a,i} \in X \subset R^n$, $u \in U \subset R^m$ ($u = [u_1, \ldots, u_m]^T$), and $w_i \in W_i \subset R^z$, respectively, where $W_i := \{w_i \in R^z \ : \ |w_i| \leq \theta_i, \ \theta_i > 0\}$, for $i = 1, 2, \ldots$. When $w_i \equiv 0$, Eq. 1 is referred to as the nominal system. f_i is considered to be a locally Lipschitz function of its arguments with $f_1(0,0,0) = 0$, and $f_i(x_{a,i,s}, u_{i,s}, 0) = 0$ for $i > 1$ (i.e., the steady-state of the nominal i-th model is at $x_{a,i} = x_{a,i,s}$, $u = u_{i,s}$). At a switching time $t_{s,i}$

[⋆] Financial support from the National Science Foundation CBET-1839675, CNS-1932026, Air Force Office of Scientific Research (award number FA9550-19-1-0059), and Wayne State University is gratefully acknowledged.

(not necessarily an integer multiple of t_k) the i-th model begins to be used and $x_{a,i}(t_{s,i+1}) = x_{a,i+1}(t_{s,i+1})$. We assume that state measurements are used by the LEMPC at t_k, but are available for model-building at $\tilde{t}_p = p\tilde{\Delta}$, $p = 1, 2, \ldots$, where $\Delta/\tilde{\Delta}$ is an integer.

We define the deviation variables $\bar{x}_{a,i} = x_{a,i} - x_{a,i,s}$ and $\bar{u}_i = u - u_{i,s}$, where $\bar{f}_i$ is f_i rewritten to have its origin at $\bar{x}_{a,i} = 0$, $\bar{u}_i = 0$. U_i and X_i represent U and X in deviation variable form from $u_{i,s}$ and $x_{a,i,s}$, respectively. We assume that there exist explicit stabilizing (Lyapunov-based) control laws $h_i(\bar{x}_{a,i}) = [h_{i,1}(\bar{x}_{a,i}) \ \ldots \ h_{i,m}(\bar{x}_{a,i})]^T$ that render the origins of the nominal systems of Eq. 1 asymptotically stable such that:

$$\alpha_{1,i}(|\bar{x}_{a,i}|) \le V_i(\bar{x}_{a,i}) \le \alpha_{2,i}(|\bar{x}_{a,i}|) \tag{2}$$

$$\frac{\partial V_i(\bar{x}_{a,i})}{\partial \bar{x}_{a,i}} \bar{f}_i(\bar{x}_{a,i}, h_i(\bar{x}_{a,i}), 0) \le -\alpha_{3,i}(|\bar{x}_{a,i}|) \tag{3}$$

$$\left| \frac{\partial V_i(\bar{x}_{a,i})}{\partial \bar{x}_{a,i}} \right| \le \alpha_{4,i}(|\bar{x}_{a,i}|) \tag{4}$$

$$h_i(\bar{x}_{a,i}) \in U_i \tag{5}$$

for all $\bar{x}_{a,i} \in D_i \subseteq R^n$ and $i = 1, 2, \ldots$, where D_i is an open neighborhood of the origin of $\bar{f}_i$, and V_i is a positive definite, sufficiently smooth Lyapunov function. $\alpha_{1,i}$, $\alpha_{2,i}$, $\alpha_{3,i}$, and $\alpha_{4,i}$ are of class $\mathcal{K}$. The level set $\Omega_{\rho_i} \subset D_i$ (considered to be within X_i) of V_i is referred to as the stability region of the system of Eq. 1 under $h_i(\bar{x}_{a,i})$. The components of $h_i(\cdot)$ are assumed to be Lipschitz continuous. The following hold for $M_i > 0$, for all $x \in \Omega_{\rho_i}$, $u \in U_i$, and $w_i \in W_i$:

$$|\bar{f}_i(x, u, w_i)| \le M_i \tag{6}$$

We consider that the only available model for control design may be an empirical model with the form:

$$\dot{x}_{b,q}(t) = f_{NL,q}(x_{b,q}(t), u(t)) \tag{7}$$

where $f_{NL,q}$ is locally Lipschitz in $x_{b,q} \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ with $f_{NL,1}(0,0) = 0$ and $f_{NL,q}(x_{b,q,s}, u_{q,s}) = 0$ for $q > 1$ (i.e., the steady-state of the updated models is at $x_{b,q} = x_{b,q,s}$, $u = u_{q,s}$). The steady-state of the empirical model in Eq. 7 may not be the same as the steady-state of the process of Eq. 1 which it seeks to approximate. The index q is used instead of i to reflect that the process dynamics may change at times different from the times $t_{s,NL,q}$ when the empirical model is updated to the q-th model ($x_{b,q}(t_{s,NL,q+1}) = x_{b,q+1}(t_{s,NL,q+1})$). We define $\bar{x}_{b,q} = x_{b,q} - x_{b,q,s}$ and $\bar{u}_q = u - u_{q,s}$, with $\bar{f}_{NL,q}$ as $f_{NL,q}$ rewritten to have its origin at $\bar{x}_{b,q} = 0$, $\bar{u}_q = 0$. U_q and X_q represent U and X in deviation variable form from $u_{q,s}$ and $x_{b,q,s}$, respectively. We consider that there exist locally Lipschitz explicit controllers $h_{NL,q}(\bar{x}_{b,q})$ that can render $x_{b,q,s}$, $u_{q,s}$ asymptotically stable in the sense that:

$$\hat{\alpha}_{1,q}(|\bar{x}_{b,q}|) \le \hat{V}_q(\bar{x}_{b,q}) \le \hat{\alpha}_{2,q}(|\bar{x}_{b,q}|) \tag{8a}$$

$$\frac{\partial \hat{V}_q(\bar{x}_{b,q})}{\partial \bar{x}_{b,q}} \bar{f}_{NL,q}(\bar{x}_{b,q}, h_{NL,q}(\bar{x}_{b,q})) \le -\hat{\alpha}_{3,q}(|\bar{x}_{b,q}|) \tag{8b}$$

$$\left| \frac{\partial \hat{V}_q(\bar{x}_{b,q})}{\partial \bar{x}_{b,q}} \right| \le \hat{\alpha}_{4,q}(|\bar{x}_{b,q}|) \tag{8c}$$

$$h_{NL,q}(\bar{x}_{b,q}) \in U_q \tag{8d}$$

for all $\bar{x}_{b,q} \in D_{NL,q}$ ($D_{NL,q}$ is a neighborhood of the origin of $\bar{f}_{NL,q}$ contained in X). $\hat{V}_q : \mathbb{R}^n \to \mathbb{R}_+$ is a sufficiently smooth Lyapunov function. $\hat{\alpha}_{i,q}, i = 1, 2, 3, 4$, are class $\mathcal{K}$ functions. The stability region of Eq. 7 under $h_{NL,q}$ is defined by $\Omega_{\hat{\rho}_q} \subset D_{NL,q}$, for which a superset contained in $D_{NL,q}$ is $\Omega_{\hat{\rho}_{safe,q}}$. The origins of Ω_{ρ_i} and of $\Omega_{\hat{\rho}_{safe,q}}$ may not be the same. There exist $M_{L,q} > 0$ and $L_{L,q} > 0$ $\forall x, x_1, x_2 \in \Omega_{\hat{\rho}_{safe,q}}$, $u \in U_q$, and $q = 1, 2, \ldots$ such that:

$$|\bar{f}_{NL,q}(x, u)| \le M_{L,q} \tag{9a}$$

$$\left| \frac{\partial \hat{V}_q(x_1)}{\partial x} \bar{f}_{NL,q}(x_1, u) - \frac{\partial \hat{V}_q(x_2)}{\partial x} \bar{f}_{NL,q}(x_2, u) \right| \le L_{L,q}|x_1 - x_2| \tag{9b}$$

Furthermore, defining $\bar{x}_{a,i,q} = x_{a,i} - x_{b,q,s}$ and $\bar{f}_{i,q}$ as the right-hand side of Eq. 1 when the model is re-written in terms of $\bar{x}_{a,i,q}$ and $\bar{u}_q$, we consider:

$$\dot{\bar{x}}_{a,i,q} = \bar{f}_{i,q}(\bar{x}_{a,i,q}(t), \bar{u}_q(t), w_i(t)) \tag{10}$$

$x_{b,q,s}$, $u_{q,s}$ is not necessarily a steady-state of this model, but this notation allows both $x_{a,i}$ and $x_{b,q}$ to be recovered by adding $x_{b,q,s}$ to $\bar{x}_{a,i,q}$ and $\bar{x}_{b,q}$. We assume that for all x, x', u', and w such that $x_t + x_{b,q,s} - x_{a,i,s} \in \Omega_{\rho_i}$ ($x_t = x$ or x'), $u' + u_{q,s} \in U$, and $w \in W_i$:

$$|\bar{f}_{i,q}(x, u', w) - \bar{f}_{i,q}(x', u', 0)| \le L_{x,i,q}|x - x'| + L_{w,i,q}|w| \tag{11}$$

$$\left| \frac{\partial \hat{V}_q(x)}{\partial x} \bar{f}_{i,q}(x, u', w) - \frac{\partial \hat{V}_q(x')}{\partial x} \bar{f}_{i,q}(x', u', 0) \right| \le L'_{x,i,q}|x - x'| + L'_{w,i,q}|w| \tag{12}$$

with $L_{x,i,q}$, $L_{w,i,q}$, $L'_{x,i,q}$, and $L'_{w,i,q} > 0$.

Remark 1. We consider that the process dynamics can change over time. Therefore, despite the fact that there is already a model error between $\bar{f}_i$ and $\bar{f}_{NL,q}$ when $w_i \equiv 0$ due to the fact that $\bar{f}_{NL,q}$ is derived from data, we include an additional error component w_i in Eq. 1 that we allow to vary over time to represent potentially time-varying components of the dynamics (e.g., heat exchanger fouling). As the dynamics continue to change over time, they could be represented without using a switched modeling framework, but then the bound on the disturbance may need to be large to reflect that the disturbance continues to grow with respect to a single model over time.

2.3 LEMPC with an Empirical Model

This work considers a control law known as LEMPC Heidarinejad et al. (2012); Alanqar et al. (2015a,b); Giuliani and Durand (2018) defined by:

$$\min_{\bar{u}_q(t) \in S(\Delta)} \int_{t_k}^{t_{k+N}} [L_e(\bar{x}_{b,q}(\tau), \bar{u}_q(\tau))] d\tau \tag{13a}$$

$$\text{s.t.} \quad \dot{\bar{x}}_{b,q} = \bar{f}_{NL,q}(\bar{x}_{b,q}(t), \bar{u}_q(t)) \tag{13b}$$

$$\bar{x}_{b,q}(t_k) = x(t_k) \tag{13c}$$

$$\bar{x}_{b,q}(t) \in X_q, \ \forall\, t \in [t_k, t_{k+N}) \tag{13d}$$

$$\bar{u}_q(t) \in U_q, \ \forall\, t \in [t_k, t_{k+N}) \tag{13e}$$

$$\hat{V}_q(\bar{x}_{b,q}(t)) \le \hat{\rho}_{e,q}, \ \ \forall\, t \in [t_k, t_{k+N}), \ \text{if} \ \ x(t_k) \ \in \ \Omega_{\hat{\rho}_{e,q}} \tag{13f}$$

$$\frac{\partial \hat{V}_q(x(t_k))}{\partial x}(\bar{f}_{NL,q}(x(t_k), \bar{u}_{b,q}(t_k))) \le \frac{\partial \hat{V}_q(x(t_k))}{\partial x}(\bar{f}_{NL,q}(x(t_k), h_{NL,q}(x(t_k)))) \ \text{if} \ \ x(t_k) \notin \Omega_{\hat{\rho}_{e,q}}, \ \text{or} \ \ t_k \ge t' \tag{13g}$$

where $L_e(\cdot,\cdot)$ is the LEMPC stage cost (Eq. 13a), $\bar{u}_q \in S(\Delta)$ signifies that $\bar{u}_q$ is a piecewise-constant input trajectory with period Δ, and the prediction horizon is denoted by N. State predictions from the empirical model (Eq. 13b) are initialized from a state measurement (denoted by $x(t_k)$) from the system of the form of Eq. 1 that describes the process dynamics at t_k. Eqs. 13d and 13e represent state and input constraints, respectively, whereas Eqs. 13f-13g are Lyapunov-based stability constraints.

3. RESPONSE TO ANOMALIES UNDER LEMPC

If the underlying process dynamics change significantly over time, the closed-loop state of the process under control actions computed by the LEMPC of Eq. 13 (without an update to the empirical model) could leave the region where closed-loop stability is guaranteed. Durand (2020) sought to address this by causing an LEMPC to be designed with a sufficiently conservative stability region $\Omega_{\hat{\rho}_q}$ such that there exists a superset $\Omega_{\hat{\rho}_{samp,q}} \subset \Omega_{\hat{\rho}_{safe,q}}$ of the stability region which the closed-loop state will still be within at the first detection that it has left the expected stability region (if the anomaly is not too large). Once $x(t_k) \notin \Omega_{\hat{\rho}_q}$, $h_{NL,q}$ can be used (since the LEMPC of Eq. 13 may then be infeasible) while data is gathered to allow for re-identification of the model before the closed-loop state leaves $\Omega_{\hat{\rho}_{safe,q}}$. This is a control theory-based triggering method for model updates in LEMPC. Under sufficient conditions restated below, the model re-identification can facilitate development of a stabilizing controller that can then drive the closed-loop state toward a neighborhood of the origin of the new empirical model (the model update and controller switch must occur by a sampling time $t_{ID,q}$, which can be no more than $t_{h,q}$ sampling times after $t_{d,q}$).

In this section, we re-examine the stability and feasibility results from Durand (2020) from a run-time safety verification perspective (i.e., to be used by the system in assessing whether it believes it can operate safely in the face of the next potential set of anomalies it may see, and in adjusting its control law to cause it to meet the requirements that would ensure safety as the data-driven model changes). We assume throughout that given the current data-driven model, a decision-making algorithm can suggest a reasonable worst-case set of subsequent models to allow predictions of worst-case behavior (compared to the current condition) if an anomaly occurs to be assessed.

Theoretical results (a proposition and theorem) from Durand (2020) are presented below, where it is assumed that the q-th empirical model was designed using data from the i-th first-principles model, and that the empirical model needs to be updated after i is changed to $i+1$ in Eq. 1, and that both $\Omega_{\hat{\rho}_{safe,q}}$ and $\Omega_{\hat{\rho}_{safe,q+1}}$ are subsets of Ω_{ρ_i} and $\Omega_{\rho_{i+1}}$. The significance of this latter assumption is that it ensures that both before and after the empirical model and therefore its stability region is updated, the closed-loop state is in a region from which the origin of the actual system could be stabilized (Eqs. 2-5).

Proposition 2. Giuliani and Durand (2018) Consider the closed-loop system of Eq. 7 under $h_{NL,q}(\bar{x}_{b,q})$ that satisfies the inequalities of Eq. 8 in sample-and-hold. Let $\Delta > 0$, $\hat{\epsilon}_{W,q} > 0$, and $\hat{\rho}_{safe,q} > \hat{\rho}_q > \hat{\rho}_{e,q} > \hat{\rho}_{\min_q} > \hat{\rho}_{s,q} > 0$ satisfy:

$$-\hat{\alpha}_{3,q}(\hat{\alpha}_{2,q}^{-1}(\hat{\rho}_{s,q})) + L_{L,q}M_{L,q}\Delta \leq -\hat{\epsilon}_{W,q}/\Delta \tag{14}$$

and

$$\hat{\rho}_{\min_q} := \max\{\hat{V}_q(\bar{x}_{b,q}(t{+}\Delta)) \;:\; \hat{V}_q(\bar{x}_{b,q}(t)) \leq \hat{\rho}_{s,q}\}. \tag{15}$$

If $\bar{x}_{b,q}(0) \in \Omega_{\hat{\rho}_{safe,q}}$, then:

$$\hat{V}_q(\bar{x}_{b,q}(t_{k+1})) - \hat{V}_q(\bar{x}_{b,q}(t_k)) \leq -\hat{\epsilon}_{W,q} \tag{16}$$

for $\bar{x}_{b,q}(t_k) \in \Omega_{\hat{\rho}_{safe,q}}/\Omega_{\hat{\rho}_{s,q}}$ and the state trajectory $\bar{x}_{b,q}(t)$ of the closed-loop system is always bounded in $\Omega_{\hat{\rho}_{safe,q}}$ for $t \geq 0$ and is ultimately bounded in $\Omega_{\hat{\rho}_{\min_q}}$.

Theorem 3. Durand (2020) Consider the closed-loop system of Eq. 1 under the LEMPC of Eq. 13 with $h_{NL,q}$ meeting Eq. 8 and Proposition 2. If $x(t_0) \in \Omega_{\hat{\rho}_q}$ and $x_{a,i}(t_{s,i+1}) = x_{a,i+1}(t_{s,i+1}) \in \Omega_{\hat{\rho}_q}$ and after $t_{s,i+1}$, the system of Eq. 1 is controlled by the LEMPC of Eq. 13 until $t_{d,q}$ with $x(t_{d,q}) \in \Omega_{\hat{\rho}_{samp,q}} \subset \Omega_{\hat{\rho}_{safe,q}}$, at which point it is controlled by $h_{NL,q}$ in sample-and-hold until $t_{ID,q} \leq t_{d,q} + t_{h,q}\Delta$, where

$$t_{h,q} = \text{floor}\left(\frac{(\hat{\rho}_{safe,q} - \hat{\rho}_{samp,q})}{\epsilon_{W,i+1,q}}\right) \tag{17}$$

then if the following conditions hold with $\hat{\rho}_{safe,q} > \hat{\rho}_{samp,q} > \hat{\rho}_q > \hat{\rho}_{q,e}$, $\hat{\rho}_{q,e} > \hat{\rho}_{\min,q,i} > \hat{\rho}_{s,q} > 0$, and $\hat{\rho}_{q,e} > \hat{\rho}_{\min,i+1,q} > \hat{\rho}_{s,q} > 0$:

$$\hat{\rho}_{q,e} \leq \hat{\rho}_q - f_{V,q}(f_{W,i,q}(\Delta)) \tag{18}$$

$$\begin{aligned} &-\hat{\alpha}_{3,q}(\hat{\alpha}_{2,q}^{-1}(\hat{\rho}_{s,q})) + \hat{\alpha}_{4,q}(\hat{\alpha}_{1,q}^{-1}(\hat{\rho}_q))M_{err,i,q} \\ &+ L'_{x,i,q}M_i\Delta + L'_{w,i,q}\theta_i \leq -\epsilon_{W,i,q}/\Delta \end{aligned} \tag{19}$$

$$\hat{\rho}_{\min,i,q} := \max\{\hat{V}_q(\bar{x}_{a,i,q}(t+\Delta)) \mid \hat{V}_q(\bar{x}_{a,i,q}(t)) \leq \hat{\rho}_{s,q}\} \tag{20}$$

with $\epsilon_{W,i+1,q}$ and $\hat{\rho}_{samp,q}$ defined to satisfy:

$$\begin{aligned} &-\hat{\alpha}_{3,q}(\hat{\alpha}_{2,q}^{-1}(\hat{\rho}_{s,q})) + \hat{\alpha}_{4,q}(\hat{\alpha}_{1,q}^{-1}(\hat{\rho}_q))M_{err,i+1,q} \\ &+ L'_{x,i+1,q}M_{i+1}\Delta + L'_{w,i+1,q}\theta_{i+1} \leq \epsilon_{W,i+1,q}/\Delta \end{aligned} \tag{21}$$

$$\begin{aligned} &\hat{\rho}_q + f_{V,q}(f_{W,i,q}(\Delta) + (M_{change,i,q})\Delta \\ &+ \frac{L_{w,i,q}\theta_i + M_{err,i,q}}{L_{x,i,q}}(e^{L_{x,i,q}\Delta})) \leq \hat{\rho}_{samp,q} \end{aligned} \tag{22}$$

$$\hat{\rho}_{e,q} + f_{V,q}(f_{W,i+1,q}(\Delta)) \leq \hat{\rho}_{samp,q} \tag{23}$$

$$\hat{\rho}_q + \epsilon_{W,i+1,q} \leq \hat{\rho}_{samp,q} \tag{24}$$

where $f_{W,i,q}$ is defined by

$$f_{W,i,q}(t) := \frac{L_{w,i,q}\theta_i + M_{err,i,q}}{L_{x,i,q}}(e^{L_{x,i,q}t} - 1) \tag{25}$$

with $M_{err,i,q}$ defined by:

$$|\bar{f}_{i,q}(x,u,0) - \bar{f}_{NL,q}(x,u)| \leq M_{err,i,q} \tag{26}$$

for all x contained in $\Omega_{\hat{\rho}_{safe,q}}$ and $\bar{u}_q \in U_q$, $f_{V,q}$ is defined by

$$f_{V,q}(s) := \hat{\alpha}_{4,q}(\hat{\alpha}_{1,q}^{-1}(\hat{\rho}_{safe,q}))s + M_{v,q}s^2 \tag{27}$$

with $M_{v,q}$ is a positive constant, and $M_{change,i,q}$ is defined by

$$\begin{aligned} &|\bar{f}_{i+1,q}(\bar{x}_{a,i+1,q}(s), \bar{u}_q(s), w_{i+1}(s)) \\ &- \bar{f}_{i,q}(\bar{x}_{a,i,q}(s), \bar{u}_q(s), w_i(s))| \leq M_{change,i,q} \end{aligned} \tag{28}$$

for all $\bar{x}_{a,i,q}, \bar{x}_{a,i+1,q} \in \Omega_{\hat{\rho}_{safe,q}}$, $\bar{u}_q \in U_q$, $w_i \in W_i$, and $w_{i+1} \in W_{i+1}$, then $x_{a,i,q}(t) \in \Omega_{\hat{\rho}_{safe,q}}$ for $t \in [t_0, t_{s,i+1}]$ and $x_{a,i+1,q}(t) \in \Omega_{\hat{\rho}_{safe,q}}$ for $t \in [t_{s,i+1}, t_{ID,q}]$.

Theorem 3 above is streamlined in presentation compared to Durand (2020) to better enable discussion of the factors which impact its use for safety verification. $t_{h,q}$ is fixed in Eq. 17 by $\hat{\rho}_{safe,q}$, $\hat{\rho}_{samp,q}$, and $\epsilon_{W,i+1,q}$. Smaller values of $\hat{\rho}_{samp,q}$ and $\epsilon_{W,i+1,q}$, for a fixed value of $\hat{\rho}_{safe,q}$, guarantee a longer time until the closed-loop state leaves $\Omega_{\hat{\rho}_{safe,q}}$ after the change in the process dynamics is detected. Smaller values of $\hat{\rho}_{samp,q}$ are, however, more restrictive for $\hat{\rho}_q$ due to the assumption that $\hat{\rho}_{samp,q} > \hat{\rho}_q$, and may cause Δ to need to be smaller for Eq. 20 to be satisfied with $\hat{\rho}_q > \hat{\rho}_{q,e}$ needing to be larger than both $\hat{\rho}_{\min,q,i}$ and $\hat{\rho}_{\min,i+1,q}$ to meet the assumptions of the theorem. Regardless of how small Δ becomes, however, the plant-model mismatch (which based on the modeling strategy in this paper, comes from both $M_{err,i+1,q}$ and θ_{i+1} in Eq. 21) sets the size of $\epsilon_{W,i+1,q}$, so that if the plant-model mismatch after $t_{s,i+1}$ may be significant, $\hat{\rho}_{samp,q}$ may need to be small to allow there to be a sufficient amount of time after $t_{d,q}$ according to Eq. 17 to obtain data for model re-identification. From the perspective of assuring safety over time, when the model of Eq. 7 changes, parameters such as $\hat{\rho}_q$ need to be selected, and obtaining acceptable values of these parameters may require an on-line adjustment to Δ or to the origin of $\Omega_{\hat{\rho}_{safe,q}}$ and $\Omega_{\hat{\rho}_q}$ to enable handling of future potential anomalies.

A number of parameters and functions in the conditions of Proposition 2 and Theorem 3 rely on characteristics of the underlying process dynamic model of Eq. 1 (e.g., $L_{w,i,q}$, θ_i, $L_{x,i,q}$, $M_{err,i,q}$, and $M_{change,i,q}$), which we assume is not available. These would have to be estimated to utilize the results of Theorem 3 for performing safety verification. Some ideas for attempting to do this would be to make estimates of the potential values of these properties based on the current data-driven model, the potential next set of anomalies from the decision-making algorithm, and safety factors added on to attempt to make those values conservative. In contrast, some of the functions (e.g., $\hat{\alpha}_{j,q}$, $j = 1, 2, 3, 4$) rely only on the knowledge of the data-driven model, and therefore may be estimated from manipulations of or simulations with that model.

Remark 4. Different concepts for handling model changes may vary in ease of verification. For example, consider an alternative method for determining when an empirical model used in EMPC should be updated from Alanqar et al. (2017a), in which the model re-identification is triggered when the error between state predictions from the data-driven model and state measurements exceeds a threshold. One could imagine gaining confidence in safety for this system by simulating various possible models which could give a prediction error within the threshold from many different initial conditions and checking that the closed-loop state does not leave a characterizable region in state-space under any of these conditions for a defined amount of time. However, attempting to devise all tests required to verify that there is no worst-case condition which was not tested has potential to be challenging.

3.1 Handling Anomalies Over Time

The theory developed in Durand (2020) assumes that $\Omega_{\hat{\rho}_{safe,q}} \subset \Omega_{\hat{\rho}_{safe,q+1}}$ when proving that closed-loop stability is maintained after $t_{ID,q}$. It is determined that if Eqs. 18 and 19 are met with q and i replaced by $q+1$ and $i+1$, and $h_{NL,q+1}$ used after $t_{ID,q}$ until the closed-loop state enters $\Omega_{\hat{\rho}_{q+1}}$ (at which point the LEMPC of Eq. 13 with the $q+1$-th data-driven model is used), the closed-loop state is always in $\Omega_{\hat{\rho}_{safe,q+1}}$. Though this may be reasonable for handling a single anomaly, if there are repeated anomalies over time, this condition effectively requires $\Omega_{\hat{\rho}_{safe,q}}$ to grow each time the model is re-identified, which may not be a reasonable assumption based on the requirements of Eq. 8, the assumption that each $\Omega_{\hat{\rho}_{safe,q}}$ is a subset of Ω_{ρ_i} and $\Omega_{\rho_{i+1}}$, and the assumption that the closed-loop state can be maintained within $\Omega_{\hat{\rho}_{safe,q}}$ for a characterizable time period after another anomaly occurs. It may be desirable, for closed-loop stability purposes after $t_{ID,q}$, to instead identify a region $\Omega_{\hat{\rho}_{new,q+1}}$ which is a superset of $\Omega_{\hat{\rho}_{safe,q}}$ and in which Eqs. 18 and 19 are met for the $i+1$ and $q+1$ models (i.e., $h_{NL,q+1}$ can drive the closed-loop state to level sets of $\hat{V}_{q+1}$ with a smaller upper bound within this region), but where $\Omega_{\hat{\rho}_{new,q+1}}$ is not necessarily equal to $\Omega_{\hat{\rho}_{safe,q+1}}$ (selection of $\Omega_{\hat{\rho}_{safe,q}}$ should be made in such a way that $\Omega_{\hat{\rho}_{new,q+1}}$ is expected to exist that contains $\Omega_{\hat{\rho}_{safe,q}}$). To define $\Omega_{\hat{\rho}_{safe,q+1}}$, the conditions of Theorem 3 must be met with respect to an additional anomaly leading to a model update to the $q+2$ and $i+2$ models. To achieve this, it is possible that the origin of $\Omega_{\hat{\rho}_{safe,q+1}}$ may need to be moved compared to the origin of $\Omega_{\hat{\rho}_{new,q+1}}$, or that Δ may need to be adjusted.

The translation of the origin can be performed via a procedure similar to that used in Alanqar et al. (2017b), but here with empirical models. Specifically, a target origin $\bar{x}_{b,o}$, $\bar{u}_{b,o}$ is identified for $\Omega_{\hat{\rho}_{safe,q+1}}$ ($\bar{f}_{NL,q+1}(\bar{x}_{b,o}, \bar{u}_{b,o}) = 0$). If $\Omega_{\hat{\rho}_{q+1}}$ for this $\Omega_{\hat{\rho}_{safe,q+1}}$ includes the neighborhood of the origin $\bar{x}_{b,new}$, $\bar{u}_{b,new}$ of $\Omega_{\hat{\rho}_{new,q+1}}$ that meets Eq. 20 (with i replaced by $i+1$) for f_{NL} rewritten to have its origin at $\bar{x}_{b,new}$, $\bar{u}_{b,new}$, then under the assumption that Eqs. 18-20 hold in $\Omega_{\hat{\rho}_{safe,q+1}}$ (e.g., despite that data may not have been available around this origin when $f_{NL,q+1}$ was originally identified, it remains a sufficiently accurate representation of the process dynamics in $\Omega_{\hat{\rho}_{safe,q+1}}$ and a Lyapunov-based controller meeting Eq. 8 can stabilize $\bar{x}_{b,o}$, $\bar{u}_{b,o}$), the LEMPC of Eq. 13 formulated around $\bar{x}_{b,new}$, $\bar{u}_{b,new}$ can be utilized with the constraint of Eq. 13g always activated until the closed-loop state enters $\Omega_{\hat{\rho}_{q+1}}$; then, the LEMPC of Eq. 13 formulated with respect to $\bar{x}_{b,o}$, $\bar{u}_{b,o}$ can be used to maintain the closed-loop state in $\Omega_{\hat{\rho}_{q+1}}$. If $\Omega_{\hat{\rho}_{q+1}}$ does not contain a neighborhood of $\bar{x}_{b,new}$, $\bar{u}_{b,new}$, then a series of additional origins with stability regions that contain neighborhoods of the origins of the previous stability region, with $\Omega_{\hat{\rho}_{q+1}}$ eventually containing a neighborhood of the origin of one of those regions, can be traversed, if Eqs. 18-20 continue to hold in those regions. With regard to changing Δ to ensure that all conditions in Eq. 3 are met before the next anomaly, this could be done by, at some t_k, changing the LEMPC sampling period (or changing it at $t_{ID,q}$ if it impacts the ability of $h_{NL,q}$ to drive the closed-loop state to lower level sets of $\hat{V}_{q+1}$). This updated implementation strategy again relies on an ability to characterize worst-case conditions, and assumes validity of a data-driven model in state-space regions potentially containing new information about the process dynamics.

3.2 Resilience for LEMPC with Data-Driven Models

An important observation regarding the stability conditions in Theorem 3 is that they depend on $M_{err,i,q}$ and

$M_{err,i+1,q}$. This raises a question with regard to the cyberattack resilience of the anomaly-handling algorithm to false state measurements that could result in inaccurate models being identified for the LEMPC. Furthermore, the model update strategy seems to offer a pathway by which an attacker can, effectively, reprogram an LEMPC through state measurements. In this section, we seek to clarify the concerns, possibilities, and also future work related to this topic through a chemical process example.

The chemical process example consists of a nonisothermal reactor in which an $A \to B$ reaction takes place, and the inputs (reactant inlet concentration C_{A0} and heat rate Q) are adjusted by an LEMPC. The process model is:

$$\dot{C}_A = \frac{F}{V}(C_{A0} - C_A) - k_0 e^{-\frac{E}{R_g T}} C_A^2 \tag{29}$$

$$\dot{T} = \frac{F}{V}(T_0 - T) - \frac{\Delta H k_0}{\rho_L C_p} e^{-\frac{E}{R_g T}} C_A^2 + \frac{Q}{\rho_L C_p V} \tag{30}$$

where the parameters are listed in Alanqar et al. (2015b) and include the reactor volume V, reactant inlet temperature T_0, pre-exponential constant k_0, solution heat capacity C_p and density ρ_L, volumetric flow rate F, gas constant R_g, activation energy E, and heat of reaction ΔH. The states are the reactant concentration C_A and temperature T in the reactor, which can be written in deviation form from the operating steady-state vector $C_{As} = 1.22$ kmol/m^3, $T_s = 438.2$ K, $C_{A0s} = 4$ kmol/m^3, and $Q_s = 0$ kJ/h as $x = [x_1\ x_2]^T = [C_A - C_{As}\ T - T_s]^T$ and $u = [u_1\ u_2]^T = [C_{A0} - C_{A0s}\ Q - Q_s]^T$. To highlight concepts related to resilience without focusing on details of model identification, we consider that the model form in Eqs. 29-30 has been identified without any mismatch and therefore is able to be used for LEMPC design.

The LEMPC is to be designed to maximize the production rate of the desired product with input bounds as follows:

$$L_e = -k_0 e^{-E/(R_g T(\tau))} C_A(\tau)^2 \tag{31}$$

$$0.5 \le C_{A0} \le 7.5 \text{ kmol/m}^3 \tag{32}$$

$$-5 \times 10^5 \le Q \le 5 \times 10^5 \text{ kJ/h} \tag{33}$$

Lyapunov-based stability constraints are also enforced (a constraint of the form of Eq. 13f is enforced at the end of every sampling time if $x(t_k) \in \Omega_{\rho_e}$, and a constraint of the form of Eq. 13g is enforced at t_k when $x(t_k) \in \Omega_{\hat{\rho}}/\Omega_{\hat{\rho}_e}$ with a constraint of the form of Eq. 13f enforced at the end of sampling periods after the first). The Lyapunov function selected was $\hat{V}_q = x^T P x$, with P given as follows:

$$P = \begin{bmatrix} 2000 & -10 \\ -10 & 3 \end{bmatrix} \tag{34}$$

The Lyapunov-based controller $h_{NL,1}(x)$ was designed such that $h_{NL,1,1}(x) = 0$ kmol/m^3 and $h_{NL,1,2}(x)$ is computed via Sontag's formula Lin and Sontag (1991) but saturated at the input bounds of Eq. 33 if they are met. $\hat{\rho}$ and $\hat{\rho}_e$ were taken to be 200 and 150, respectively, and $\hat{\rho}_{safe}$ was set to 400. The process state was initialized at $x_{init} = [0 \text{ kmol/m}^3\ 0 \text{ K}]^T$, with controller parameters $N = 10$ and $\Delta = 0.01$ h. The process model of Eqs. 29-30 was integrated with the Explicit Euler numerical integration method using an integration step size of 10^{-4} h.

To begin to explore resilience of the anomaly-handling method to attacks which modify $M_{err,i,q}$ by changing the data-driven model, we first note that correct state measurements, but a falsified data-driven model (i.e., one that is purposefully identified to be sufficiently inaccurate with respect to the underlying process dynamics) may cause the closed-loop state to leave a bounded region of state-space. For example, if the underlying dynamic model has $k_0 = 8.46 \times 10^6$ m^3/ h kmol, but if the LEMPC is initialized at the steady-state and uses a dynamic model with $k_0 = 10^7$ m^3/ h kmol as a constraint, the closed-loop state exits $\Omega_{\hat{\rho}_q}$ over time. This indicates that the question of whether safety of a process could be compromised if an attacker was able to replace a data-driven model used in an LEMPC with a different one, instead of providing false state measurements to the process or false signals to the actuators, is a valid concern to analyze.

We now conceptually explore how the anomaly-handling algorithm discussed above would work with an attacker seeking to modify the process model to impact safety. In this strategy, if the state measurements to be used for the model identification are the same as those being used by the controller, then there is a barrier to the attacker being able to manipulate the process dynamic model without first manipulating the sensor measurements that here we assume are received by both the controller and the logic system that detects when the model update must be performed. To force a specific model to be identified, the attacker would need to know the model identification algorithm and then to provide false state measurements before the re-identification takes place that would cause this model to be identified. However, the process of doing so may cause the closed-loop state to leave $\Omega_{\hat{\rho}_{safe,q}}$ due to the false state measurements causing inputs to be computed that drive the closed-loop state out of $\Omega_{\hat{\rho}_{safe,q}}$ quickly (i.e., before the time that would have been expected if there were correct measurements and the plant/model mismatch was within the bounds).

As an illustration, we consider the case that it is postulated that k_0 might change over time. If it changes to, for example, $k_0 = 8.48 \times 10^6$ m^3/ h kmol at $t = 0.05$ h and the process is simulated for 0.2 h, initialized at the steady-state, the closed-loop state does not exit $\Omega_{\hat{\rho}_{safe,q}}$ for over two sampling periods after the first sampling time that the closed-loop state has left $\Omega_{\hat{\rho}_q}$ (0.08 h), giving time to detect the model inaccuracy and take action before the closed-loop state leaves $\Omega_{\hat{\rho}_{safe,q}}$. The value of $\hat{V}_q$ does not monotonically increase after $t_{d,q}$ here.

If an attacker would like to cause the model used in the LEMPC to have significant mismatch from the actual model to attempt to cause safety issues, under the strategy developed in this work, the attacker would need to first cause the closed-loop state to leave $\Omega_{\hat{\rho}_q}$ so that a model re-identification is triggered. An attacker could attempt to wait until a re-identification is triggered by an actual change in the process dynamics and then to provide false sensor measurements during the stage in which the gathered data may be considered for model re-identification. Here, we explore a case in which sensor measurements are provided by an attacker in an attempt to indicate that the underlying dynamic model has changed to a new model which the attacker would like to impose in the LEMPC. If there is no change in the underlying process dynamics at $t = 0.05$ h, but after that time false sensor

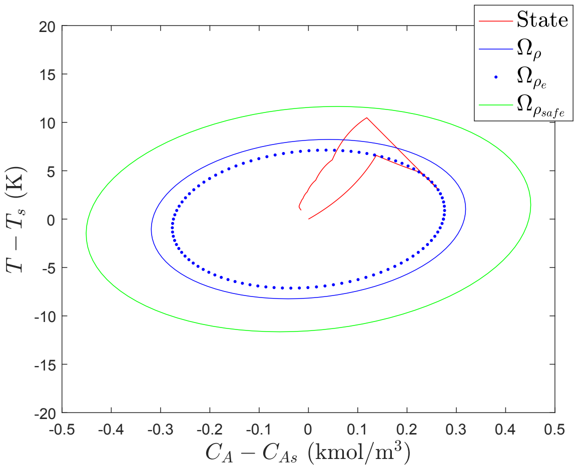

Fig. 1. State-space trajectory when an attack is performed to provide false state measurements that impact the model identification from 0.05 to 0.09 h. The total simulation time is 0.3 h.

readings begin to be presented that follow the trajectory that the closed-loop state would take if the underlying dynamic model had changed at $t = 0.05$ h to one where $k_0 = 8.48 \times 10^6$ m^3/ h kmol, then $t_{d,q}$ would be 0.08 h and over two sampling periods would pass before the closed-loop state leaves $\Omega_{\hat{\rho}_{safe,q}}$. For the purpose of illustration, let us assume that the value of k_0 is identified to be 8.48×10^6 m^3/ h kmol as would be suggested by the (falsified) data between $t_{d,q}$ and $t_{ID,q}$ (which will here be set to 0.10 h to ensure that the closed-loop state in the simulations is still within $\Omega_{\hat{\rho}_{safe,q}}$ at $t_{ID,q}$) for this process. According to the implementation strategy for the anomaly-handling technique, if there is sufficiently small plant-model mismatch after the re-identification, $h_{NL,q+1}$ will be able to drive the closed-loop state to lower level sets of $\hat{V}_q$. Here, however, the model re-identification did not remove plant-model mismatch from the empirical model, but enhanced it. Still, however, in this example, $h_{NL,q+1}$ was able to drive the closed-loop state back into $\Omega_{\hat{\rho}_q}$, as shown in Fig. 1, when correct state measurements began to be again provided to the LEMPC starting at $t_{ID,q}$ (i.e., the attacker was able to remove their presence from the system to potentially aid in avoiding detection). This indicates that the plant-model mismatch induced by the attacker did not cause safety issues for this specific simulation (i.e., not all model updates from false data are problematic for closed-loop stability under the proposed technique).

The above discussion suggests that an attacker with knowledge of how the anomaly-handling strategy works could provide a false sensor measurement at t_k that is outside of $\Omega_{\hat{\rho}_q}$ but still inside $\Omega_{\hat{\rho}_{safe,q}}$ to force the model re-identification to begin (i.e., it is not necessary to provide false state measurements before t_k, reducing the amount of time that the attacker must be engaged). This raises the question of whether attack detection mechanisms could detect targeted attacks intended to impact the controller programming not necessarily through consistent false sensor measurements, but through those at specific times which would be relevant to model identification, and suggests that future research might explore resilient control designs with run-time safety verification properties.

REFERENCES

Ahooyi, T.M., Soroush, M., Arbogast, J.E., Seider, W.D., and Oktem, U.G. (2016). Model-predictive safety system for proactive detection of operation hazards. *AIChE Journal*, 62, 2024–2042.

Alanqar, A., Durand, H., and Christofides, P.D. (2015a). On identification of well-conditioned nonlinear systems: Application to economic model predictive control of nonlinear processes. *AIChE Journal*, 61, 3353–3373.

Alanqar, A., Durand, H., and Christofides, P.D. (2017a). Fault-tolerant economic model predictive control using error-triggered online model identification. *Industrial & Engineering Chemistry Research*, 56, 5652–5667.

Alanqar, A., Durand, H., Albalawi, F., and Christofides, P.D. (2017b). An economic model predictive control approach to integrated production management and process operation. *AIChE Journal*, 63, 1892–1906.

Alanqar, A., Ellis, M., and Christofides, P.D. (2015b). Economic model predictive control of nonlinear process systems using empirical models. *AIChE Journal*, 61, 816–830.

Albalawi, F., Alanqar, A., Durand, H., and Christofides, P.D. (2016). A feedback control framework for safe and economically-optimal operation of nonlinear processes. *AIChE Journal*, 62, 2391–2409.

Armaou, A. and Demetriou, M.A. (2008). Robust detection and accommodation of incipient component and actuator faults in nonlinear distributed processes. *AIChE Journal*, 54, 2651–2662.

Durand, H. (2020). Responsive economic model predictive control for next-generation manufacturing. *Mathematics*, 8, 259.

Durand, H. and Wegener, M. (2020). Mitigating safety concerns and profit/production losses for chemical process control systems under cyberattacks via design/control methods. *Mathematics*, 8(4), 499.

Ellis, M., Durand, H., and Christofides, P.D. (2014). A tutorial review of economic model predictive control methods. *Journal of Process Control*, 24, 1156–1178.

Giuliani, L. and Durand, H. (2018). Data-based nonlinear model identification in economic model predictive control. *Smart and Sustainable Manufacturing Systems*, 2, 61–109.

Heidarinejad, M., Liu, J., and Christofides, P.D. (2012). Economic model predictive control of nonlinear process systems using Lyapunov techniques. *AIChE Journal*, 58, 855–870.

Lin, Y. and Sontag, E.D. (1991). A universal formula for stabilization with bounded controls. *Systems & Control Letters*, 16, 393–397.

Mahmood, M. and Mhaskar, P. (2010). Safe-parking framework for fault-tolerant control of transport-reaction processes. *Industrial & Engineering Chemistry Research*, 49, 4285–4296.

Rawlings, J.B., Angeli, D., and Bates, C.N. (2012). Fundamentals of economic model predictive control. In *Proceedings of the Conference on Decision and Control*, 3851–3861. Maui, Hawaii.

Xue, D. and El-Farra, N. (2018). Forecast-triggered model predictive control of constrained nonlinear processes with control actuator faults. *Mathematics*, 6(6), 104.