2022 American Control Conference (ACC)
Atlanta, USA, June 8-10, 2022

A Learning-Based Model Predictive Control Framework for Real-Time SIR Epidemic Mitigation

Baike She, Shreyas Sundaram, and Philip E. Paré*

***Abstract*— We propose a learning-based model predictive control framework for mitigating the spread of epidemics. We capture the epidemic spreading process using a susceptible-infected-removed (SIR) epidemic model and consider testing for isolation as the control strategy. In the framework, we use a daily testing strategy to remove (isolate) a portion of the infected population. Our goal is to keep the daily infected population below a certain level, while minimizing the total number of tests. Distinct from existing works on leveraging model predictive control in epidemic spreading, we learn the model parameters and compute the feedback control signal simultaneously. We illustrate the results by numerical simulation using COVID-19 data from India.**

I. Introduction

In order to optimally allocate resources for epidemic mitigation while reducing the impact on society, researchers have studied the use of optimal control formulations [1]. However, a gap still exists between the theoretical/numerical results and the implementation of optimal mitigation policies in real-time for mitigation: the open-loop structure of the optimal control framework is not robust to epidemic modeling uncertainties [2]. One way to overcome the challenge is to learn the model and update the optimal control strategy iteratively [3].

Model predictive control (MPC) has demonstrated success in both traditional and modern control systems [4]. One advantage of MPC is the ability to generate policies while considering future performance by solving optimal control problems recursively. Hence, researchers have explored MPC for epidemic control problems [5]–[11]. The authors in [5] leveraged MPC to minimize the cost of social distancing and testing. The authors in [6] considered MPC to minimize the number of fatalities caused by COVID-19, subject to constraints on the economic cost of social distancing, and [7] formulated an optimal on-off (binary) social isolation strategy through the MPC framework to mitigate the COVID-19 contagion in Brazil. The authors in [8] captured logical relations between model variables through temporal logic in constraints for the COVID-19 mitigation framework. In addition, existing literature estimated the model parameters offline through data sets for the MPC implementation [5]–[11]. In real-time epidemic control problems, parameter estimation must be implemented recursively into the MPC framework to learn the model parameters gradually, since the size of the data set can be varying. Therefore, we will adopt a learning-based MPC framework to perform parameter estimation and control design simultaneously.

Our main contribution is to propose a learning-based MPC framework for epidemic mitigation. Specifically, we use a linear regression method to estimate the model parameters. For the MPC framework, we adopt testing for isolation strategies [12], [13]. The testing for isolation strategy aims to remove the infected population from the infected group through uniform random sampling. Similar to the idea of vaccination strategies that remove the susceptible population from the mixed group [14], testing for isolation strategies are another widely adopted method [13]. One disadvantage of vaccination strategies is that the vaccine may be unavailable at the early stages of an outbreak caused by novel viruses like COVID-19. Unlike vaccination, testing for isolation strategies can be implemented much earlier [15].

The paper is organized as follows. In Section II, we introduce the learning-based model predictive control framework that we will leverage for the epidemic mitigation problem. In Section III, we formulate and discuss the linear regression problem for learning the epidemic model parameters for the learning-based MPC framework. Section IV illustrates the learning-based MPC framework through simulations on both simulated data and real COVID-19 data from India. Section V presents the conclusion and future work.

Notation

For a system of three equations captured by A_k, we use $[A]_{S,k}$, $[A]_{I,k}$, $[A]_{R,k}$ to represent the first, second, and third equation in A_k, respectively. We use $E[X]$ to represent the expected value of a random variable X, and $Cov(X,Y)$ to represent the covariance of random variables X and Y. We use $\mathbb{Z}_{>0}$ and $\mathbb{N}$ to denote the set of all positive integers and natural numbers, respectively. For two numbers a and b, we use $a \gg b$ and $a \ll b$ to represent a is much larger and smaller than b, respectively. For any positive integer n, we use $[n]$ to denote the index set $\{1, 2, \ldots, n\}$. We view vectors as column vectors and write $\boldsymbol{x}^\top$ to denote the transpose of a column vector $\boldsymbol{x}$. For any matrix $M \in \mathbb{R}^{n\times n}$, we use $[M]_{i,:}$, $[M]_{:,j}$, $[M]_{ij}$, to denote its ith row, jth column, and ijth entry, respectively. Denote $\mathbf{0}$ as the zero vector with the corresponding dimension in the context.

II. Problem Formulation

In this section, we introduce the control system we will study in this work and formulate the optimization problem. Our goal is to propose a potential way for policy-makers

*Baike She, Shreyas Sundaram, and Philip E. Paré are with the Elmore Family School of Electrical and Computer Engineering at Purdue University. E-mails: {bshe, sundara2, philpare}@purdue.edu. Research supported in part by the C3.ai Digital Transformation Institute sponsored by C3.ai Inc. and the Microsoft Corporation, and in part by the National Science Foundation, grants NSF-CMMI #1635014, NSF-CNS #2028738, and NSF-ECCS #2032258.

978-1-6654-5196-3/$31.00 ©2022 AACC

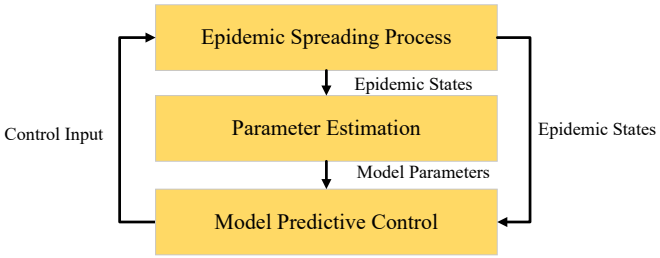

Fig. 1: Learning-Based MPC Framework

to mitigate an epidemic through real-time learning and a feedback testing strategy. Our framework is shown in Fig. 1. In the framework, we obtain epidemic data sets from the real spreading process. We leverage the data sets to estimate the model parameters for the MPC, as illustrated by the arrow from the top block to the middle block in Fig. 1. We also leverage the data sets as the state feedback for the controller design, captured by the right-hand side arrow from the top block to the bottom block in Fig. 1. We implement the control polices generated by the MPC framework to the epidemic spreading process. We perform the process iteratively.

A. Testing for Mitigation

In this subsection, we present the model for the epidemic control problem. We consider the following closed-loop Susceptible-Infected-Recovered/Removed (SIR) model:

$$\dot{S}(t) = -\beta S(t)I(t), \tag{1a}$$
$$\dot{I}(t) = \beta S(t)I(t) - (\gamma + u(t))I(t), \tag{1b}$$
$$\dot{R}(t) = (\gamma + u(t))I(t), \tag{1c}$$

where $S(t) \in [0,1]$, $I(t) \in [0,1]$, $R(t) \in [0,1]$ denote the susceptible, infected, and removed proportion in the population, respectively, at time t, with $S(t) + I(t) + R(t) = 1$, $\forall t \geq 0$. The parameters $\beta > 0$ and $\gamma > 0$ represent the time-invariant transmission rate and removal rate, respectively. In this work, we assume the removal rate captures any processes that separate the infected group from the whole population, which include the recovery process, hospitalization, deaths, etc. We define mitigation as maintaining the *infection level* (the proportion of the infected population $I(t)$) under a certain threshold through control strategies. The control input $u(t)$ captures testing strategies that isolate/remove $u(t) \times 100\%$ of the detected infected population from the infected group. The testing strategy is achieved by uniformly randomly sampling $u(t) \times 100\%$ of population from the mixed susceptible and infected groups. We assume no sampling bias nor testing delay, and that the testing is completely accurate. Through the perfect testing accuracy assumption, we are able to select $u(t) \times 100\%$ of the infected population to be tested positive and thus removed from the infected group, captured by $u(t)I(t)$ in (1). Note that when $u(t) = 0$, the system in (1) becomes the classic SIR model [16].

B. Parameter Learning

In order to formulate the epidemic control problem, we need to learn the model parameters β and γ in (1). Existing literature studying MPC frameworks for epidemic control generates model parameters through numerical optimization methods from data sets. However, none of these works incorporated the parameter learning process with the MPC design simultaneously or considered the external control inputs during the parameter estimation [5]–[11]. In order to implement online parameter estimation with the impact of control inputs, we will formulate the parameter estimation problem for the closed-loop SIR epidemic model as a linear regression problem [17]. We show how to incorporate control inputs into the parameter learning process.

C. Model Predictive Control

In this subsection, we introduce the learning-based MPC framework. Consider the system formulated in (1). The goal for the epidemic mitigation is to optimally allocate the testing resources during the pandemic such that the daily infected population is maintained at/below the desired infection threshold. In this work, we consider mitigating the epidemic by minimizing the total number of tests during the epidemic through the following cost function:

$$J(u(t)) = \int_{t_0}^{t_0+T} u(t)dt, \tag{2}$$

where $[t_0, t_0 + T]$ is the prediction horizon for the MPC framework. In order to obtain the adaptive testing strategy that minimizes the total number of tests needed during each prediction horizon while ensuring that the fraction of infected individuals remains below a desired threshold, we formulate the following optimization problem:

$$\min_{u(t), t_0 \leq t \leq t_0+T} \; J(u(t)) \tag{3a}$$
$$\text{s.t.} \;\; \dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t), u(t)), \tag{3b}$$
$$0 \leq I(t) \leq \bar{I}, \underline{u} \leq u(t) \leq \bar{u}, \forall t \in [t_0, t_0 + T] \tag{3c}$$

where the state constraint $\bar{I}$ describes the *infection threshold* for the fraction of the infected undetected population. The control input constraints $\underline{u}$ and $\bar{u}$ define the upper and lower bounds on the daily testing rates, respectively. The solution for the optimization problem gives the optimal testing strategy $u(t)$ for an epidemic that starts from $\boldsymbol{x}(t_0)$ over the prediction horizon $[t_0, t_0 + T]$. We define $k \in \mathbb{Z}_{>0}$ such that the prediction horizon $[t_0, t_0+T]$ can be partitioned into k equal intervals. In the MPC framework:

1) Compute the optimal solution over the prediction horizon $[t_0, t_0 + T]$ through (3);
2) Implement the testing policy through $[t_0, t_0 + \frac{T}{k}]$ to the closed-loop system in (1) to obtain the final state $\boldsymbol{x}(t_0 + \frac{T}{k})$, and update $t_0 \to t_0 + \frac{T}{k}$;
3) Update the optimization problem in Step 1 by updating the prediction horizon $[t_0, t_0 + T]$ with the updated t_0 from Step 2, and the initial condition $\boldsymbol{x}(t_0) \to \boldsymbol{x}(t_0 + \frac{T}{k})$;
4) Solve the optimization problem and apply the control input to the system by repeating steps 1) - 3).

D. Goals

In this work, we will focus on applying learning-based MPC framework for the epidemic mitigation problem. We

will formulate a parameter learning method for the closed-loop system in (1). We will illustrate the framework through numerical simulations. We aim to show that the proposed learning-based MPC framework provides a potential way for policy-makers to select the optimal resource allocation strategy in real-time epidemic mitigation.

III. Learning-based MPC: Parameter Learning

In this section, we analyze the parameter estimation part proposed in the framework shown in Fig.1. We propose a learning strategy by linear regression to estimate the epidemic parameters recursively.

Note that closed-loop epidemic spreading dynamics in (1) are nonlinear with respect to the dynamic states. However, the equations are linear with respect to the parameters β and γ, which characterize the spreading behavior [18] and are what we want to learn from data. Hence, we discretize the dynamics of the closed-loop SIR model in (1) in order to construct a group of linear predictor functions for the linear regression analysis. Define the sample step size as $h \in \mathbb{R}_{>0}$. Then, $\forall k \in \mathbb{Z}_{>0}$, the discrete dynamics are given by

$$S_k = S_{k-1} - h\beta S_{k-1} I_{k-1}, \tag{4a}$$
$$I_k = I_{k-1} + h\beta S_{k-1} I_{k-1} - h(\gamma + u_{k-1}) I_{k-1}, \tag{4b}$$
$$R_k = R_{k-1} + h(\gamma + u_{k-1}) I_{k-1}. \tag{4c}$$

The matrix form of (4) is given by

$$\underbrace{\begin{bmatrix} S_k \\ I_k \\ R_k \end{bmatrix}}_{Y_k} = Y_{k-1} + h \underbrace{\begin{bmatrix} -S_{k-1}I_{k-1} & 0 \\ S_{k-1}I_{k-1} & -I_{k-1} \\ 0 & I_{k-1} \end{bmatrix}}_{\varphi_{k-1}} \underbrace{\begin{bmatrix} \beta \\ \gamma \end{bmatrix}}_{\theta} + h u_{k-1} \underbrace{\begin{bmatrix} 0 \\ -\hat{I}_{k-1} \\ \hat{I}_{k-1} \end{bmatrix}}_{\tilde{I}_{k-1}}.$$

We define the observed states, $\forall k \in \mathbb{Z}_{>0}$, as $\hat{Y}_k$,

$$\hat{Y}_k = \begin{bmatrix} \hat{S}_k \\ \hat{I}_k \\ \hat{R}_k \end{bmatrix} = \begin{bmatrix} S_k \\ I_k \\ R_k \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon_{S,k} \\ \varepsilon_{I,k} \\ \varepsilon_{R,k} \end{bmatrix}}_{\varepsilon_k}, \tag{5}$$

where $\varepsilon_{j,k}$ are the error terms. We make the following assumptions about the error terms.

Assumption 1. *The error terms $\varepsilon_{S,k}$, $\varepsilon_{I,k}$, $\varepsilon_{R,k}$, $\forall k \in \mathbb{N}$, have the following properties:*

1) $E(\varepsilon_{S,k}) = E(\varepsilon_{I,k}) = E(\varepsilon_{R,k}) = 0$;
2) $Var(\varepsilon_{S,k}) = \sigma_s^2$, $Var(\varepsilon_{I,k}) = \sigma_I^2$, $Var(\varepsilon_{R,k}) = \sigma_R^2$;
3) $\varepsilon_{S,k}$, $\varepsilon_{I,k}$, *and* $\varepsilon_{R,k}$ *are mutually independent;*
4) $\frac{S_k^2}{\sigma_S} \gg 1$, $\frac{I_k^2}{\sigma_I} \gg 1$, $\frac{R_k^2}{\sigma_R} \gg 1$.

To generate the linear relationship between the parameters and the observed states, we provide the following derivation from (4) - (5). First we show that

$$\begin{aligned}\hat{Y}_k &= Y_{k-1} + h\varphi_{k-1}\theta + hu_{k-1}\tilde{I}_{k-1} \\ &\quad + \varepsilon_k + \varepsilon_{k-1} - \varepsilon_{k-1} + hu_{k-1}\Upsilon_{I,k-1} - hu_{k-1}\Upsilon_{I,k-1}, \\ &= \hat{Y}_{k-1} + h\varphi_{k-1}\theta + hu_{k-1}\hat{\tilde{I}}_{k-1} + \varepsilon_k - \varepsilon_{k-1} - hu_{k-1}\Upsilon_{I,k-1},\end{aligned}$$

where $\Upsilon_{I,k-1} = [0 \ -\varepsilon_{I,k-1} \ \varepsilon_{I,k-1}]^\top$, $\hat{\tilde{I}}_{k-1} = [0 \ -\hat{I}_{k-1} \ \hat{I}_{k-1}]^\top$, and the fact $\hat{\tilde{I}}_{k-1} = \tilde{I}_{k-1} + hu_{k-1}\Upsilon_{I,k-1}$ is used. To replace φ_{k-1} with $\hat{\varphi}_{k-1}$, being composed of $\hat{S}_{k-1}$ and $\hat{I}_{k-1}$,

$$\begin{aligned}&\hat{\varphi}_{k-1} \\ &= \begin{bmatrix} -(S_{k-1}+\varepsilon_{S,k-1})(I_{k-1}+\varepsilon_{I,k-1}) & 0 \\ (S_{k-1}+\varepsilon_{S,k-1})(I_{k-1}+\varepsilon_{I,k-1}) & -(I_{k-1}+\varepsilon_{I,k-1}) \\ 0 & (I_{k-1}+\varepsilon_{I,k-1}) \end{bmatrix} \\ &= \varphi_{k-1} + \underbrace{\begin{bmatrix} -(S_{k-1}\varepsilon_{I,k-1} + I_{k-1}\varepsilon_{S,k-1} + \varepsilon_{S,k-1}\varepsilon_{I,k-1}) & 0 \\ S_{k-1}\varepsilon_{I,k-1} + I_{k-1}\varepsilon_{S,k-1} + \varepsilon_{S,k-1}\varepsilon_{I,k-1} & -\varepsilon_{I,k-1} \\ 0 & \varepsilon_{I,k-1} \end{bmatrix}}_{\Delta_{k-1}}.\end{aligned}$$

Then, we have

$$\begin{aligned}\hat{Y}_k &= \hat{Y}_{k-1} + h\varphi_{k-1}\theta + h\Delta_{k-1}\theta - h\Delta_{k-1}\theta \\ &\quad + hu_{k-1}\hat{\tilde{I}}_{k-1} + \varepsilon_k - \varepsilon_{k-1} - hu_{k-1}\Upsilon_{I,k-1} \\ &= \hat{Y}_{k-1} + h\hat{\varphi}_{k-1}\theta + hu_{k-1}\hat{\tilde{I}}_{k-1} + \Sigma_{k-1},\end{aligned}$$

where the error vector term Σ_{k-1} is defined as

$$\Sigma_{k-1} = \varepsilon_k - \varepsilon_{k-1} - hu_{k-1}\Upsilon_{I,k-1} - h\Delta_{k-1}\theta.$$

Therefore, we can construct a group of the following linear equations at time step k, $\forall k \in \mathbb{Z}_{>0}$, as:

$$\underbrace{\hat{Y}_k - \hat{Y}_{k-1} - hu_{k-1}\hat{\tilde{I}}_{k-1}}_{\hat{\Omega}_k} = h\hat{\varphi}_{k-1}\theta + \Sigma_{k-1}. \tag{6}$$

Note that the observed states and control inputs in $\hat{\Omega}_k$ capture the changes for S, I, and R, from time step $k-1$ to k, $\forall k \in \mathbb{Z}_{>0}$. The observed states from (5) can be formulated as a group of linear equations with error terms by

$$\hat{\Omega} = h\hat{\varphi}\theta + \Sigma, \tag{7}$$

where $\hat{\Omega} = [\hat{\Omega}_1^\top \ \hat{\Omega}_2^\top \ \dots \ \hat{\Omega}_n^\top]^\top$, $\hat{\varphi} = [\hat{\varphi}_0^\top \ \hat{\varphi}_1^\top \ \dots \ \hat{\varphi}_{n-1}^\top]^\top$, and $\Sigma = [\Sigma_0^\top \ \Sigma_1^\top \ \dots \ \Sigma_{n-1}^\top]^\top$.

For simplicity, we analyze the n equations from (7) that contain only β to estimate $\hat{\beta}$ while using the n equations that only include $\hat{\gamma}$ to estimate γ. Thus, we have

$$\hat{\Omega}_S = h\hat{\varphi}_S\beta + \Sigma_S, \tag{8}$$
$$\hat{\Omega}_R = h\hat{\varphi}_R\gamma + \Sigma_R, \tag{9}$$

where $\hat{\Omega}_S = [\hat{\Omega}_{S,1}^\top \ \dots \ \hat{\Omega}_{S,n}^\top]^\top$, $\hat{\Omega}_R = [\hat{\Omega}_{R,1}^\top \ \dots \ \hat{\Omega}_{R,n}^\top]^\top$, $\hat{\varphi}_S = [\hat{\varphi}_{S,0}^\top \ \dots \ \hat{\varphi}_{S,n-1}^\top]^\top$, $\hat{\varphi}_R = [\hat{\varphi}_{R,0}^\top \ \dots \ \hat{\varphi}_{R,n-1}^\top]^\top$, $\Sigma_S = [\Sigma_{S,0}^\top \ \dots \ \Sigma_{S,n-1}^\top]^\top$, and $\Sigma_R = [\Sigma_{R,0}^\top \ \dots \Sigma_{R,n-1}^\top]^\top$. We formulate the two groups of linear equations to estimate $\hat{\beta}$ and $\hat{\gamma}$ separately. Recall from Section II that the goal of the linear regression is to find $\hat{\beta}$ and $\hat{\gamma}$ which are the solutions for the following optimization problems, respectively:

$$\hat{\beta} = \underset{\hat{\beta}>0}{\arg\min}\, J_\beta \quad \text{and} \quad \hat{\gamma} = \underset{\hat{\gamma}>0}{\arg\min}\, J_\gamma,$$

where

$$J_\beta = \sum_{k=1}^{n} (\hat{\Omega}_{S,k} + h\beta(\hat{S}_{k-1}\hat{I}_{k-1}))^2 = ||\hat{\Omega}_S + h\beta\hat{Z}||^2,$$

$$J_\gamma = \sum_{k=1}^{n} (\hat{\Omega}_{R,k} - h\gamma\hat{I}_{k-1})^2 = ||\hat{\Omega}_R - h\gamma\hat{I}||^2,$$

where $\hat{Z} = [\hat{S}_0\hat{I}_0 \;\; \hat{S}_1\hat{I}_1 \cdots \hat{S}_{n-1}\hat{I}_{n-1}]^T$.

Note that (8) and (9) are errors-in-variables models, where the input states are corrupted with noise. In addition, the observation noise Σ_S and Σ_R are dependent on the noisy input states $\hat{\varphi}_S$ and $\hat{\varphi}_R$, respectively. Hence, to simplify the regression analysis, we have the following assumption.

Assumption 2. *The estimated states* $\hat{I}_{k-1} = I_{k-1}$ $\forall k \in \mathbb{Z}_{>0}$, *i.e.,* $\varepsilon_{I,k} = 0$, $\forall k \in \mathbb{Z}_{>0}$.

Recall from Section II we adopt uniform random sampling to estimate the daily infected population, assume there is no testing delay, and assume the testing results are accurate. Therefore, through the testing rate $u(k-1)$ $\forall k \in \mathbb{Z}_{>0}$, we assume we can obtain accurate daily infected population measurements with $\varepsilon_{I,k} = 0$, $\forall k \in \mathbb{Z}_{>0}$, including both symptomatic and asymptomatic infections. Thus, we have Assumption 2. Note that, compared to estimating the infected population, estimating the removed and the susceptible populations is more challenging under the settings in our framework; thus we keep the corresponding errors for both groups in our estimation framework nonzero.

In order to select the appropriate least squares estimator to solve the linear regression problem, we now explore the properties of the error terms Σ_S and Σ_R from (8) and (9).

Lemma 1. *Under Assumption 1 and 2,* $E(\Sigma_S) = \mathbf{0}$, $E(\Sigma_R) = \mathbf{0}$, *and* $E(\Sigma_R|\hat{\varphi}_R) = \mathbf{0}$.

Proof. We check the kth error terms of $E(\Sigma_{S,k-1})$ and $E(\Sigma_{R,k-1})$ by examining $E(\Sigma_{k-1})$ $\forall k \in \mathbb{Z}_{>0}$:

$$E(\Sigma_{k-1}) = h\begin{bmatrix} -\beta I_{k-1}E(\varepsilon_{S,k-1}) \\ \hline \beta I_{k-1}E(\varepsilon_{S,k-1}) \\ \hline 0 \end{bmatrix} = \mathbf{0}.$$

In the above derivation, we use Assumption 1 that the errors terms ε_S, ε_R are mutually independent and zero mean, and Assumption 2 that $\varepsilon_{I,k} = 0$ $\forall k \in \mathbb{Z}_{>0}$. Since $E(\Sigma_{k-1}) = \mathbf{0}$, $\forall k \in \mathbb{Z}_{>0}$, we have $E(\Sigma_S) = \mathbf{0}$, and $E(\Sigma_R) = \mathbf{0}$. Now we check the conditional mean $E(\Sigma_R|\hat{\varphi}_R)$. Based on Assumption 2, the regression model in (9) has no errors in infection states and thus the observation error is independent from the infection states. Therefore, $\hat{\varphi}_I = \varphi_I$ is deterministic, and $\Sigma_{R,k-1}$ is independent from $\varphi_{R,k-1}$ $\forall k \in \mathbb{Z}_{>0}$. Hence, $E(\Sigma_{R,k-1}|\hat{\varphi}_{R,k-1}) = E(\Sigma_{R,k-1}) = \mathbf{0}$, $\forall k \in \mathbb{Z}_{>0}$. Further, $E(\Sigma_R|\hat{\varphi}_R) = \mathbf{0}$. □

After discussing the (conditional) expectations of the error terms Σ_S and Σ_R, we explore the variances of the error terms at time step k to generate the covariance matrices of Σ_S and Σ_R. Define $Var(\Sigma_k) = [Var(\Sigma_{S,k})\;\; Var(\Sigma_{I,k})\;\; Var(\Sigma_{R,k})]^\top$, and the covariance matrices of Σ_S and Σ_R as $C(\Sigma_S)$ and $C(\Sigma_R)$, respectively.

Lemma 2. *Under Assumptions 1 and 2, the covariance matrices of* Σ_S *and* Σ_R *are given by*

$$C(\Sigma_S)_{ij} = \begin{cases} \sigma_S^2 + (h\beta I_{i-1} - 1)^2\sigma_S^2, & i = j, \\ (h\beta I_i - 1)\sigma_S^2, & i - j = 1, \\ (h\beta I_j - 1)\sigma_S^2, & i - j = -1, \\ 0, & \textit{otherwise,} \end{cases}$$

$$C(\Sigma_R)_{ij} = \begin{cases} 2\sigma_R^2, & i = j, \\ -\sigma_R^2, & i - j = 1, \\ -\sigma_R^2, & i - j = -1, \\ 0, & \textit{otherwise,} \end{cases}$$

respectively, and $i \in \mathbb{N}$, $j \in \mathbb{N}$. *The conditional covariance matrix* $C(\Sigma_R|\hat{\varphi}_R) = C(\Sigma_R)$.

Proof. The variance of the error terms at time step $k-1$ is given by

$$\begin{aligned} Var(\Sigma_{k-1}) &= \begin{bmatrix} Var(\varepsilon_{S,k} + (h\beta I_{k-1} - 1)\varepsilon_{S,k-1}) \\ \hline Var(\varepsilon_{I,k} - h\beta I_{k-1}\varepsilon_{S,k-1}) \\ \hline Var(\varepsilon_{R,k} - \varepsilon_{R,k-1}) \end{bmatrix} \\ &= \begin{bmatrix} Var(\varepsilon_{S,k}) + (h\beta I_{k-1} - 1)^2 Var(\varepsilon_{S,k-1}) \\ \hline Var(\varepsilon_{I,k}) - (h\beta I_{k-1})^2 Var(\varepsilon_{S,k-1}) \\ \hline Var(\varepsilon_{R,k}) + Var(\varepsilon_{R,k-1}) \end{bmatrix}. \end{aligned}$$

Now we have that the diagonal entries of the covariance matrices $C(\Sigma_S)$ and $C(\Sigma_R)$ are given by $C(\Sigma_S)_{ii} = Var(\Sigma_{S,i-1})$, $C(\Sigma_R)_{ii} = Var(\Sigma_{R,i-1})$, $\forall i \in \mathbb{Z}_{>0}$. It can be observed that the error terms $\Sigma_{S,i}$ and $\Sigma_{S,j}$ (same for $\Sigma_{R,i}$ and $\Sigma_{R,j}$) are correlated if and only if $|i-j| \leq 1$, $\forall i, j \in \mathbb{Z}_{>0}$. Hence, for the off-diagonal entries of the covariance matrices $C(\Sigma_S)$ and $C(\Sigma_R)$, we have $C(\Sigma_S)_{ij} = 0$, $C(\Sigma_R)_{ij} = 0$, if $|i-j| > 1$,$\forall i, j \in \mathbb{Z}_{>0}$. In addition, if $i - j = 1$, $\forall i, j \in \mathbb{Z}_{>0}$,

$$\begin{aligned} C(\Sigma_S)_{k+1,k} &= Cov(\Sigma_{S,k}, \Sigma_{S,k-1}) = (h\beta I_k - 1)Var(\varepsilon_{S,k}), \\ C(\Sigma_R)_{k+1,k} &= Cov(\Sigma_{R,k}, \Sigma_{R,k-1}) = -Var(\varepsilon_{R,k}). \end{aligned}$$

Now we have showed the construction of $C(\Sigma_S)$ and $C(\Sigma_R)$ to prove the lemma. Recall from the proof of Lemma 1 that there is no correlation between $\hat{\varphi}_{R,k-1}$ and $\Sigma_{R,k-1}$ $\forall k \in \mathbb{Z}_{>0}$, and thus $C(\Sigma_R|\hat{\varphi}_R) = C(\Sigma_R)$. □

Lemma 2 establishes the (conditional) covariance matrix of the error Σ_R. Note that although all the error terms are independent from Assumption 1, the error terms in Σ_S are correlated, and the same for Σ_R. Hence, the ordinary least square (OLS) estimator for (8) or (9) is not suitable due to the requirement for the uncorrelated error terms [19]. Hence, we leverage the generalized least square (GLS) method, where the correlation of the error terms are allowed [17][pg. 222].

Recall that $\hat{Z} = [\hat{S}_0\hat{I}_0 \;\; \hat{S}_1\hat{I}_1 \;\cdots\; \hat{S}_{n-1}\hat{I}_{n-1}]^T$.

Lemma 3. *Under Assumptions 1 and 2, the solutions for the GLS problem of* (8) *and* (9) *with the objective functions given by* J_β *and* J_γ *are given by*

$$\hat{\beta} = -\frac{1}{h}(\hat{Z}^\top[C(\Sigma_S)]^{-1}\hat{Z})^{-1}\hat{Z}^\top[C(\Sigma_S)]^{-1}\hat{\Omega}_S, \tag{10}$$

$$\hat{\gamma} = \frac{1}{h}(\hat{I}^\top[C(\Sigma_R)]^{-1}\hat{I})^{-1}\hat{I}^\top[C(\Sigma_R)]^{-1}\hat{\Omega}_R, \tag{11}$$

respectively. Further, $E(\hat{\gamma}|\hat{\varphi}_R) = \gamma$.

Proof. Lemma 3 can be obtained by following the result from GLS [17][Eq. 9.2.10] under the conditions from Lemma 1 and Lemma 2. In particular, since $E(\Sigma_R|\hat{\varphi}_R) = \mathbf{0}$ and $C(\Sigma_R|\hat{\varphi}_R) = C(\Sigma_R)$, we have $E(\hat{\gamma}|\hat{\varphi}_R) = \gamma$, and $Var(\hat{\gamma}|\hat{\varphi}_R) = \frac{1}{h}(\hat{I}^\top[C(\Sigma_R)]^{-1}\hat{I})^{-1}$ [17][Eq. 9.2.13]. □

Lemma 3 indicates that the estimated parameter $\hat{\gamma}$ obtained from the GLS is unbiased. However, we cannot guarantee whether the estimated parameter $\hat{\beta}$ is biased or not. Therefore, we explore the performance of the parameter learning process in the following section through simulations.

IV. Simulations

In this section, we illustrate the proposed learning-based MPC framework through simulations. First, we employ the learning-based MPC framework on a simulated SIR system with additive Gaussian measurement noise. Then we implement the framework using real COVID-19 data from India to reconstruct the epidemic spreading process and thus to demonstrate the potential application of the learning-based MPC framework for real-time epidemic mitigation.

A. Learning-based MPC: Simulation

Consider an epidemic spreading process in (1) with $\beta = 0.035$ and $\gamma = 0.010$. The goal of the framework is to minimize the total number of tests used during the epidemic given by (3) recursively. We maintain each testing policy for a week ($k = 7$), under the condition that the daily upper and lower bounds on the testing rate are $\frac{30\%}{7}$ and $\frac{5\%}{7}$. We set the infection threshold on the daily infection level as $\bar{I} = 4\%$. For the MPC framework, we set the prediction horizon as four weeks ($T = 28$) when computing the optimal testing policy. After obtaining the optimal control inputs by minimizing the total number of tests during the first four-week prediction horizon, we implement the first-week testing policy, and continue the process recursively. We implement the MPC framework through CasADi tools [20].

We assume the observed data is corrupted with additive Gaussian noise, and the signal-noise ratio is $60dB$. From Fig. 1, the observed data will impact both the recursive parameter learning and the computation of the control input. For the epidemic spreading process, the population size is $N = 10000$ and the initial conditions are $I(0) = 10/N$, $R(0) = 10/N$, and $S(0) = 1 - I(0) - R(0)$.

Fig. 2 shows the learning process of the transmission and healing rates. It takes around 20 weeks (20 steps) of data for $\frac{|\hat{\beta}-\beta|}{\beta} \times 100\% < 1.14\%$. It takes around 40 weeks (40 steps) for $\frac{|\hat{\gamma}-\gamma|}{\gamma} \times 100\% < 3\%$. In order to match the one week policy updating period, the estimated parameters $\hat{\beta}$ and $\hat{\gamma}$ are updated weekly through the learning process for the learning-based MPC framework.

Fig. 3 shows the results for the epidemic mitigation problem by leveraging the learning-based MPC framework, where $\hat{I}(t)$ and $I(t)$ represent the observed and true infection states, respectively. The first row of Fig. 3 illustrates the infection level is controlled around the 4% infection threshold throughout the outbreak. Note that $\hat{I}(t) > 0.04$ for several time steps t, which is caused by noisy feedback data and estimated parameters for the MPC framework. The oscillation in $I(t)$ is generated by the spikes from the control inputs, since the parameter learning with the noisy data will have an impact on the system and the initial conditions when solving the optimization problem of the MPC framework given in (3).

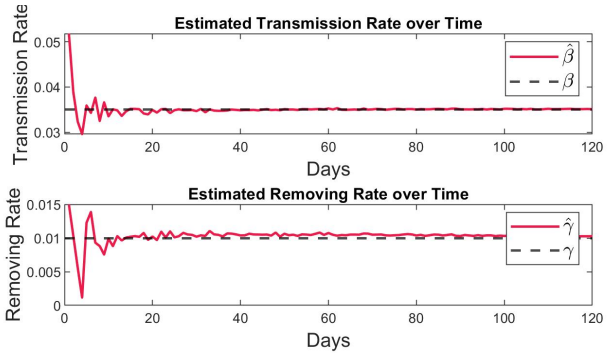

Fig. 2: Parameter Learning Process

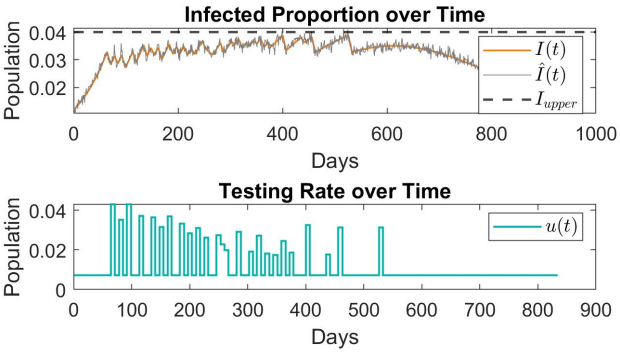

Fig. 3: Epidemic Mitigation with Learning-Based MPC

B. Learning-based MPC: Case Study

We illustrate the proposed learning-based MPC framework using COVID-19 data from India [21]. We leverage 420 days of daily recorded infected population, and the total daily recorded recovered and deceased population collected from March 20, 2020. We assume that the total population of India is $N = 1.38 \times 10^9$. Unlike the simulated SIR model in the previous section, the true epidemic spreading process was time-varying, since the transmission and healing rates were affected by different interventions implemented over different stages of the pandemic. Hence, in order to implement the learning-based MPC framework, we assume that the epidemic spreading process and the estimation error over a relatively short period were time-invariant [22].

First, we leverage the parameter estimation method proposed in Lemma 3 to estimate the time-varying transmission rate $\hat{\beta}(t)$ and removal rate $\hat{\gamma}(t)$. We assume that the epidemic spreading process did not change drastically during any four-week window, and we estimate each pair of time-invariant parameters over each four-week window. In addition, starting from day 1 (i.e., March 20, 2020), we move the estimation window by sliding it for 7 days, iteratively. The parameters of the epidemic spreading process estimated from the data are given in Fig. 4. Further, we leverage the estimated parameters to reconstruct the time-varying epidemic spreading process through a time-varying SIR model, shown in Fig. 5. In Fig. 5, the solid lines $I(t)$ and $R(t)$ represent the recorded infected and removed proportions, respectively. The dashed lines $\hat{I}(t)$ and $\hat{R}(t)$, represent the infected and recovered proportions reconstructed through-time varying transmission rate $\hat{\beta}(k)$ and removal rate $\hat{\gamma}(k)$ in Fig. 4.

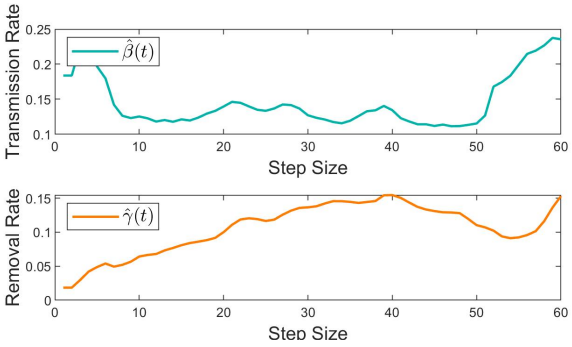

Fig. 4: Estimated Time-Varying Parameters

After discussing the parameter learning process, now we consider implementing the learning-based MPC framework. We assume that the bounds on the testing rate satisfy $\bar{u} = \frac{30\%}{7}$, $\underline{u} = \frac{5\%}{7}$, and the infection threshold $\bar{I} = 5 \times 10^{-6}$. For any prediction horizon $[t_0, t_0 + 4 \times T]$, from (3), the MPC framework will minimize the total number of tests used during the four-week prediction horizon. We set the length of the prediction horizon $T = 28$ days (i.e., four weeks).

We illustrate the performance of the learning-based MPC framework, when the parameter estimation and control design are performed simultaneously and iteratively. The simulation results in Fig. 6 show that the framework achieves the goal of maintaining the infection level under $\bar{I}$. Hence, we have illustrated the proposed learning-based MPC is a potential method for real-time epidemic mitigation.

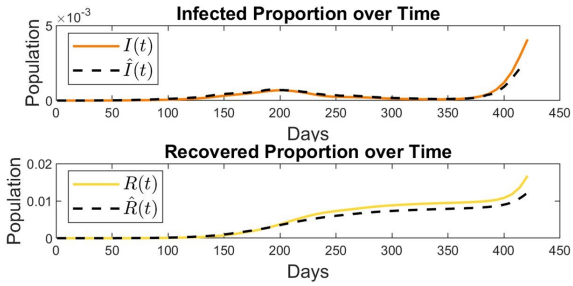

Fig. 5: Fitting SIR Model to Data

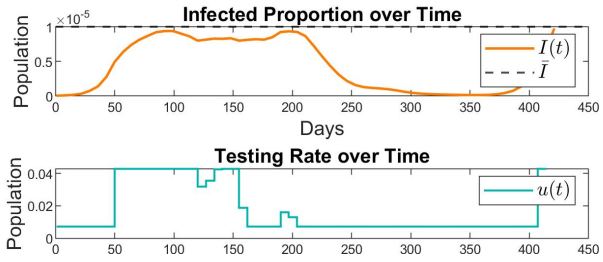

Fig. 6: Learning-Based MPC: Case Study

V. Conclusion

We proposed a learning-based model predictive control framework for epidemic mitigation problems. We formulated the parameter learning process as a linear regression problem and illustrated the efficiency of the framework through both simulated and real epidemic data. Future works include exploring model learning processes to generate more accurate models, and developing theoretical foundations for learning-based MPC for real-time epidemic mitigation problems.

References

[1] O. Sharomi and T. Malik, "Optimal control in epidemiology," *Ann. of Opera. Research*, vol. 251, no. 1-2, pp. 55–71, 2017.

[2] D. H. Morris, F. W. Rossine, J. B. Plotkin, and S. A. Levin, "Optimal, near-optimal, and robust epidemic control," *Comm. Phys.*, vol. 4, no. 1, pp. 1–8, 2021.

[3] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, "Learning-based model predictive control: Toward safe learning in control," *Annu. Rev. of Control, Robot., and Auto. Syst.*, vol. 3, pp. 269–296, 2020.

[4] F. Allgöwer and A. Zheng, *Nonlinear Model Predictive Control*. Birkhäuser, 2012, vol. 26.

[5] C. Tsay, F. Lejarza, M. A. Stadtherr, and M. Baldea, "Modeling, state estimation, and optimal control for the US COVID-19 outbreak," *Sci. Rep.*, vol. 10, no. 1, pp. 1–12, 2020.

[6] J. Köhler, L. Schwenkel, A. Koch, J. Berberich, P. Pauli, and F. Allgöwer, "Robust and optimal predictive control of the COVID-19 outbreak," *Annu. Rev. in Control*, 2020.

[7] M. M. Morato, S. B. Bastos, D. O. Cajueiro, and J. E. Normey-Rico, "An optimal predictive control strategy for COVID-19 (SARS-CoV-2) social distancing policies in Brazil," *Annu. Rev. in Control*, vol. 50, pp. 417–431, 2020.

[8] T. Péni, B. Csutak, G. Szederkényi, and G. Röst, "Nonlinear model predictive control with logic constraints for COVID-19 management," *Nonlinear Dynam.*, vol. 102, no. 4, pp. 1965–1986, 2020.

[9] R. Carli, G. Cavone, N. Epicoco, P. Scarabaggio, and M. Dotoli, "Model predictive control to mitigate the COVID-19 outbreak in a multi-region scenario," *Annu. Rev. in Control*, vol. 50, pp. 373–393, 2020.

[10] S. Grundel, S. Heyder, T. Hotz, T. K. Ritschel, P. Sauerteig, and K. Worthmann, "How much testing and social distancing is required to control COVID-19? Some insight based on an age-differentiated compartmental model," *arXiv preprint arXiv:2011.01282*, 2020.

[11] J. Sereno, A. Jorge, A. Ferramosca, E. Hernandez-Vargas, and A. Gonzalez, "Model predictive control for optimal social distancing in a type SIR-switched model," *arXiv preprint arXiv:2106.02093*, 2021.

[12] D. Acemoglu, A. Fallah, A. Giometto, D. Huttenlocher, A. Ozdaglar, F. Parise, and S. Pattathil, "Optimal adaptive testing for epidemic control: Combining molecular and serology tests," *arXiv preprint arXiv:2101.00773*, 2021.

[13] F. Casella, "Can the COVID-19 epidemic be controlled on the basis of daily test reports?" *IEEE Control Syst. Letters*, vol. 5, no. 3, pp. 1079–1084, 2020.

[14] S. M. Grundel, S. Heyder, T. Hotz, T. K. Ritschel, P. Sauerteig, and K. Worthmann, "How to coordinate vaccination and social distancing to mitigate SARS-CoV-2 outbreaks," *SIAM J. on Appl. Dynam. Syst.*, vol. 20, no. 2, pp. 1135–1157, 2021.

[15] IDA+A at Purdue University, "Ongoing COVID-19 results for testing administered by the Protect Purdue Health Center," https://protect.purdue.edu/dashboard/.

[16] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proc. of the Royal Soc. A*, vol. 115, no. 772, pp. 700–721, 1927.

[17] N. R. Draper and H. Smith, *Applied Regression Analysis*. John Wiley & Sons, 1998, vol. 326.

[18] W. Mei, S. Mohagheghi, S. Zampieri, and F. Bullo, "On the dynamics of deterministic epidemic propagation over networks," *Annu. Rev. in Control*, vol. 44, pp. 116–128, 2017.

[19] R. Davidson, J. G. MacKinnon *et al.*, *Econometric Theory and Methods*. Oxford University Press New York, 2004, vol. 5.

[20] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi – A software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.

[21] "India COVID-19 data," 2021, https://api.rootnet.in/covid19-in/stats/history.

[22] H. Zhao, N. N. Merchant, A. McNulty, T. A. Radcliff, M. J. Cote, R. S. Fischer, H. Sang, and M. G. Ory, "COVID-19: Short term prediction model using daily incidence data," *PloS One*, vol. 16, no. 4, p. e0250110, 2021.