Model-based Prediction and Optimal Control of Pandemics by Nonpharmaceutical Interventions

Reza Sameni[*]

Alphanumerics Lab, Department of Biomedical Informatics, Emory University School of Medicine

Abstract—The XPRIZE Pandemic Response Challenge addressed the problem of predicting future trends of the pandemic in different regions and countries under non-pharmaceutical interventions (NPI), and prescribing Pareto efficient NPI that compromise between the number of new cases and the cost of intervention. Herein, we present the models and algorithms developed by the Alphanumerics Team during the Pandemic Response Challenge. The algorithms are based on an *optimal state estimation* and *finite horizon optimal control* formulation of the problem. Our major contributions were to develop a predictor and optimal prescriptor over an *ad hoc* compartmental model for tracking the variations of the new cases with an *extended Kalman filter/smoother*. The algorithm provides both prediction and prescription, provides performance bounds and requires extremely low computational resources. The solution is theoretically proved to be an optimal balance between the NPI, within the scope of the presented compartmental model. The NPI found by this algorithm are on the *Pareto front* of the bi-objective optimization space and enable strategists to select the desired balance between the human factor and NPI cost by modifying a single parameter. The parameters of the model, which correspond to the human interaction model and recovery rates, can be trained by using machine learning techniques. The algorithm is applied on historic global data from the Oxford COVID-19 Government Response Tracker project and compared versus random and fixed interventions. MATLAB implementations of the developed algorithms are available online.

I. INTRODUCTION

The mathematical modeling of pandemic waves has significant importance for governments and decision-makers. Non-pharmaceutical interventions (NPIs) refer to actions and policies adopted by individuals, authorities or governments that help slowing down the spread of epidemic diseases. NPIs are among the best ways of controlling pandemic diseases when vaccines or medications are not yet available[1]. During the COVID-19 pandemic, several attempts have been made to categorize and quantify the various NPIs of different regions and nations. The quantification of the NPI are very helpful for comparing the effectiveness of regional policies in containing the pandemic spread. By using machine learning techniques, the quantified NPI can be used to forecast the future trends of the pandemic and to simulate *"what if scenarios"* for the better management of human and medical resources, and to eventually prescribe appropriate NPI for controlling the

pandemic [1]. The Oxford COVID-19 Government Response Tracker (OxCGRT) is one of the NPI tracking projects, which were launched and regularly updated during the COVID-19 pandemic [2]. Most recently, this project has been used in the machine learning community to launch data challenges for NPI-based prediction and prescription plans. Specifically the XPRIZE Pandemic Response Challenge addressed the problem of predicting future trends of the pandemic in different regions and countries under non-pharmaceutical interventions (NPI), and prescribing Pareto efficient NPI that compromise between the number of new cases and the weighted-cost of intervention [3]. The challenge was motivated by the fact that due to the variations in infrastructures, available resources and priorities, policymakers worldwide may tend to give different weights to each NPI and are interested to know the impacts and consequences of each policy in advance.

During this challenge, the Alphanumerics Team from the Department of Biomedical Informatics at Emory University, adopted a model-based *estimation theoretical* and *finite horizon optimal control* strategy to address the problem of weighted NPI prescription. The idea of bi-objective finite horizon pandemic control has also been considered by other researchers in simulated scenarios, including [4] and the SEEPIA project [5].

Considering that the only available data for the challenge were the total confirmed cases, the total confirmed deaths, and the daily NPIs, we have adopted an extension of a generic susceptible-infected (SI) compartmental model from our previous work [6], as the base model for all regions/countries. During Phase I, the proposed model parameters were trained on historic data and used to predict future trends from input NPI by using an extended Kalman filter. In Phase II, the model was integrated with a finite horizon optimal controller to find the optimal daily NPIs with arbitrary NPI cost weight vectors.

II. HIGHLIGHTS OF THE DEVELOPED ALGORITHM

The proposed method is summarized in Algorithm 1. The major elements of this algorithm are explained in the sequel. Before going into the details of the model and algorithm, we highlight some of its important features, which are justified in the sequel and our online source-codes[2]:

- The proposed method is based on theoretical derivations and within the scope of the proposed compartmental model accuracy (which is asymptotically accurate for region/country-

---

[*]R. Sameni is an Associate Professor of Biomedical Engineering and the director of the Alphanumerics Lab at the Department of Biomedical Informatics, Emory University School of Medicine, 101 Woodruff Circle, Atlanta, GA 30322, US (e-mail: rsameni@dbmi.emory.edu).

[1]See Centers for Disease Control and Prevention guidelines on NPIs: https://www.cdc.gov/nonpharmaceutical-interventions/.

[2]Cf. https://github.com/alphanumericslab/EpidemicModeling

Algorithm 1 Summary of the proposed algorithm

Input: Historic case reports and NPI files (or an arbitrary scenario file from the *standard predictor model*)
Input: The NPI weights $\mathbf{w}(t)$, per region/country
Input: The Pareto front tuning parameter $\epsilon \in [0, 1]$.

1: **for all** Regions **do**
2:     Train the compartmental model parameters over historic NPI and case reports (or the standard predictor scenario file).
3:     Use EKF and EKS for prediction and prescription of finite horizon optimal control inputs $\mathbf{u}^*(t)$.
4: **end for**

---

level population sizes), it gives accurate *Pareto efficient* solutions.

- The desired point on the Pareto front can be selected by a single parameter $\epsilon \in [0, 1]$, where the corner case $\epsilon = 0$ neglects the NPI cost and $\epsilon = 1$ neglects the human factor.
- The prediction and prescription problems are integrated in a unified framework. Nevertheless, the method is applicable to both real-world data and the standard challenge predictor.
- The method can be used for *targeted pandemic control*, where strategists can target specific infection bounds that match the medical resources of a country/region, over a fixed or maximally bounded period of time. Therefore, within the proposed framework, apart from the optimal NPI and fatality rate objectives, scenarios such as this are also possible: *"how to bring the pandemic reproduction rate below 0.8, or the new cases below 100 per day in less than two months?"*
- Both the model parameters and NPI cost weights can be updated over time. Specifically, we do not assume the NPI cost weights to be constant over time. Therefore, unprecedented events such as vaccination or virus mutation effects can be integrated in the model without additional training. Following the so-called *principle of optimality*, any portion of an optimal control trajectory is optimal [7, Sec 6.4]. Therefore, the prescribed optimal control strategy may be adopted at any point, regardless of the past actions of a region/nation.
- Since the Pareto front solutions are found by mathematical derivation (rather than trial and error or cumbersome searches), the run-time computational load is minimal.
- We have developed a novel technique for solving the numerical problem of finite horizon optimal control, using an extended Kalman filter (EKF) and a finite-interval extended Kalman smoother (EKS).
- The proposed framework is extendable to pharmaceutical intervention plans and vaccinations, whenever available.
- The predictor part of the model gives confidence intervals during both the prediction and prescription steps of the algorithm. Therefore, the performance and well-function of the algorithm can be continuously monitored and adapted.

III. THE DATA MODEL

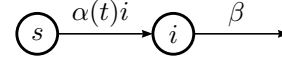The two major classes of methods for epidemic disease spread modeling are:

Fig. 1. The base susceptible-infected compartmental model with NPI-controlled infection rate

1) *Compartmental models*, which split the total population of a region into various compartments (groups) such as susceptibles, exposed, infected, recovered, vaccinated, diseased, etc. These compartments are used to form differential/difference equations, which are fit onto real data and are analytically/numerically solved to predict future trends of the disease spread.
2) *Agent-based models*, which model the behaviors of individuals and their interactions at a simplified level of abstraction. Using these models, large groups of agents are generated in stochastic simulated environments as they randomly move, interact and probabilistically pass the infection to one another, recover, pass away, etc. The population-level properties are calculated by ensemble averaging over the entire population.

Each approach has its advantages and limitations. For large population sizes at regional or national levels— which was the scope of the Pandemic Response Challenge— the first approach is asymptotically accurate and is more advantageous as it can be analytically studied in a rigorous mathematical framework and combined with state estimation techniques for forecasting [6], and optimal control theories for NPI prescription. Based on this assumption, we adopted the first approach, using a contact-controlled time-variant version of the so-called *susceptible-infected (SI)* compartmental model shown in Fig. 1, which is a simplified variant of the general multi-compartment models studied in out previous work [6]. Apparently, more accurate models can be used for the regions that further data such as the number of recovered, hospitalized, vaccinated, or the age pyramid of the population are available. However, for the current study, since the only global data provided in the Oxford dataset were the number of daily confirmed cases, total death cases, and the regional NPIs, the same SI model is used for all regions and countries.

With this background, the nonlinear dynamic equations corresponding to the proposed model are:

$$\begin{aligned}
\dot{s}(t) &= -\alpha(t)s(t)i(t) \\
\dot{i}(t) &= \alpha(t)s(t)i(t) - \beta i(t) \\
\dot{\alpha}(t) &= -\gamma\alpha(t) + \gamma h[\mathbf{u}(t)]
\end{aligned} \tag{1}$$

where

- $s(t)$ is the fraction of population in a region/country that is susceptible at time t (i.e., the susceptible population divided by the population size N);
- $i(t)$ is the fraction of population that is infected and contagious at t (i.e., the infected contagious population divided by the regional population size N);
- $\mathbf{u}(t) \in \mathbb{R}^L$ is the NPI vector considered as an exogenous control input ($L = 12$ during Phase II of the XPRIZE challenge). The exact definition of the Oxford data NPI set, and the subset used during the XPRIZE Challenge are detailed in [2] and [3];

- $\alpha(t)$ is the time-variant contagion rate with inverse time unit;
- $h[\mathbf{u}(t)]$ is a causal monotonic function of the NPI, which maps the NPI to the contagion rate;
- β is the rate of elimination from the contagious group (through quarantine, recovery, or death), assumed to be constant in the simplified case;
- γ is the *action to effect rate* (or the inverse of the NPI to individual contact rate lag), which accounts for the delay between adopting an NPI policy and the onset of its piratical effectiveness. The third equation in (1) is equivalent to $\alpha(t) = \gamma \exp[-\gamma(t - t_0)] * h[\mathbf{u}(t)]$ (for $t \geq t_0$), which is a smoothed version of $h[\mathbf{u}(t)]$. As a corner case, $\gamma \to \infty$ represents zero latency between action and effect, resulting in $\alpha(t) = h[\mathbf{u}(t)]$.

The parameters β, γ and the function $h[\mathbf{u}(t)]$ require learning using the observed variables, as detailed in Section VI. Furthermore, in [6] we showed how the infection *reproduction rate* $\mathcal{R}_t$ can be calculated from $\alpha(t)$ and β. Specifically, using the eigenanalysis-based definition of the reproduction rate proposed in [6], during the pandemic outbreak (when only several percents of the population are infected and *herd immunity* has not been reached), we have:

$$\mathcal{R}_t = \exp[\Delta(\alpha(t) - \beta)] \qquad (2)$$

where Δ is the reproduction rate generation time unit.

It is later shown that the model parameters and variables in (1) can be estimated from (noisy) observations of the *new cases* fraction:

$$n(t) = \alpha(t)s(t)i(t) + v(t), \qquad (3)$$

or through the total *confirmed cases* fraction:

$$c(t) = s(t_0) - s(t) + v(t), \qquad (4)$$

where $v(t)$ is measurement noise due to case report errors, and $s(t_0)$ is the initial susceptible population fraction at the beginning of the pandemic.

IV. FINITE HORIZON OPTIMAL NPI CONTROL

A. Cost function and problem statement

From (1), the total number of new infections over an arbitrary time window $[t_0, t_1]$ is:

$$J_0(\mathbf{u}) = \int_{t=t_0}^{t_1} \alpha(t)s(t)i(t)dt, \qquad (5)$$

the total weighted-cost of NPIs over the same time period is

$$J_1(\mathbf{u}) = \int_{t=t_0}^{t_1} \mathbf{w}(t)^T \mathbf{u}(t)dt \qquad (6)$$

where $\mathbf{w}(t)$ is the NPI weight vector given as input (cf. [3] for the definition of the NPI weights), and the set of admissible inputs for the system is

$$\Gamma = \{\mathbf{u} | \mathbf{u}^{\min} \leq \mathbf{u}(t) \leq \mathbf{u}^{\max}, \forall t \in [t_0, t_1]\} \qquad (7)$$

where $\mathbf{u}^{\min}$ and $\mathbf{u}^{\max}$ are (element-wise) the minimum and maximum ranges of each NPI from the Oxford dataset [3].

With the above definitions, the optimization problem can be formulated as a bi-objective optimization problem, with total cost:

$$J(\mathbf{u}) = (1 - \epsilon)J_0(\mathbf{u}) + \epsilon J_1(\mathbf{u}) \quad \text{s.t. } \mathbf{u} \in \Gamma \qquad (8)$$

where $\epsilon \in [0, 1]$ is a free parameter that compromises between the human factor (J_0) and the NPI cost (J_1). For a given weight pair $\{\epsilon, \mathbf{w}(t)\}$, the objective is to find $\mathbf{u}^*(t)$ for all $t \in [t_0, t_1]$, such that:

$$J(\mathbf{u}^*) = \min_{\Gamma}(J(\mathbf{u})) \qquad (9)$$

B. The Pareto optimal solution

The inputs which satisfy (9) are *Pareto optimal (efficient)*. The problem (9) can be solved by *finite horizon optimization* [7]. In fact, for an arbitrary weight vector $\mathbf{w}(t)$, by sweeping ϵ over $[0, 1]$, the *Pareto-optimal front* of the optimization problem is found, from which strategists can select the desired free parameter ϵ and its corresponding optimal input $\mathbf{u}^*(t)$.

In order to solve (9), first the corresponding *Hamiltonian* function is formed [7, Ch. 2]:

$$\begin{aligned} \mathcal{H} = \quad & (1 - \epsilon)\alpha(t)s(t)i(t) + \epsilon\mathbf{w}(t)^T\mathbf{u}(t) \\ & -\lambda_1(t)\alpha(t)s(t)i(t) \\ & +\lambda_2(t)[\alpha(t)s(t)i(t) - \beta i(t)] \\ & -\gamma\lambda_3(t)\{\alpha(t) - h[\mathbf{u}(t)]\} \end{aligned} \qquad (10)$$

where $\lambda_1(t)$, $\lambda_2(t)$ and $\lambda_2(t)$ are known as *co-states*. According to *Pontryagin's minimum principle*, the co-states and the optimal solution $\mathbf{u}^*$ satisfy [7, Ch. 6]:

$$\begin{aligned} \dot{\lambda}_1(t) &= -\frac{\partial \mathcal{H}}{\partial s} = [\lambda_1(t) - \lambda_2(t) - (1 - \epsilon)]\alpha(t)i(t) \\ \dot{\lambda}_2(t) &= -\frac{\partial \mathcal{H}}{\partial i} = [\lambda_1(t) - \lambda_2(t) - (1 - \epsilon)]\alpha(t)s(t) + \beta\lambda_2(t) \\ \dot{\lambda}_3(t) &= -\frac{\partial \mathcal{H}}{\partial \alpha} = [\lambda_1(t) - \lambda_2(t) - (1 - \epsilon)]s(t)i(t) + \gamma\lambda_3(t) \\ \mathcal{H}(\mathbf{u}^*) &\leq \mathcal{H}(\mathbf{u}), \quad \forall \mathbf{u} \in \Gamma \end{aligned}$$
$$(11)$$

When the inputs are unconstrained, the Hamiltonian minimizer input $\mathbf{u}^*$, in the last condition of (11) can be found by solving

$$\nabla_{\mathbf{u}}\mathcal{H}(\mathbf{u}^*) = \mathbf{0}, \qquad (12)$$

where $\nabla_{\mathbf{u}}\mathcal{H}$ denotes the Hamiltonian gradient with respect to the input vector $\mathbf{u}$, and the condition should hold element-wise. In this case, a sufficient condition for the existence of a solution is to have $\nabla_{\mathbf{u}}^2\mathcal{H}(\mathbf{u}^*) \succ 0$ (where $\nabla_{\mathbf{u}}^2$ denotes the Hessian operator with respect to the input vector $\mathbf{u}$, and $\succ 0$ denotes positive-definiteness). In the constrained input case—as in our problem— while the solution of (12) might not exist or belong to the admissible set (7), a Hamiltonian minimizer optimal input still exists. In either case, the optimal input is found as a parametric function of the costate $\lambda_3(t)$ and the other model parameters.

The resulting parametric optimal input is next combined with (11) and (1) to calculate the states, using the initial conditions and appropriate boundary conditions (also known as the *transversality conditions*) on the co-states and the

Hamiltonian. The desired boundary conditions, which satisfy the Challenge scenario are:

$$\lambda_1(t_1) = 0, \quad \lambda_2(t_1) = 0, \quad \lambda_3(t_1) = 0. \quad (13)$$

The conditions in (13) are the general free end-point conditions of finite horizon optimization problems, which match the Challenge scenario during the prescription phase. Alternative transversality conditions that can be studied within the proposed framework are [7, Section 2.7]:

1) When the end-time t_1 is not fixed, but we require that $i(t_1)$ reaches below $i_{\max}$ by the end of the control period (*infinite horizon* scenario). This requires the additional condition: $\mathcal{H}(t_1) = 0$.
2) Assuming that the objective of any NPI policy over a reasonable time period $[t_0, t_1]$ (long enough to make the NPIs effective) is to bring the number of active cases down to $i(t_1) \leq i_{\max}$, where $i_{\max}$ is some target fraction of active cases (ideally zero), the second condition in (13) can be replaced with: $\lambda_2(t_1)[i(t_1) - i_{\max}] = 0$.
3) We require that $i(t_1)$ drops below $i_{\max}$ any time before a maximum end time t_f, which requires $(t_1 - t_f)\mathcal{H}(t_1) = 0$.

C. The NPI to inter-human contact map

The solution of the NPI optimization problem depends on the choice of $h[\mathbf{u}(t)]$, i.e. the NPI to inter-human contact mapping model. We study the following two cases, which lead to closed form solutions for the optimal input as functions of the model co-states:

1) *Linear regression model*: $h[\mathbf{u}(t)] = b + \mathbf{a}^T[\mathbf{u}^{\max} - \mathbf{u}(t)]$, where $\mathbf{a}$ is a vector of *input influence weights* and b is a constant bias (intercept value). The LASSO model falls into this category. Adding the constraint $\mathbf{a} \geq \mathbf{0}$ guarantees a monotonic relationship between the NPI and α. The interpretation of these assumptions is that more stringent NPI policies have a non-increasing effect on the human interactions α (i.e. the NPI do not have any counter-impacts on the contact rates). In this case:

$$\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{H}(\mathbf{u}) = \epsilon \mathbf{w}(t) - \gamma\lambda_3(t)\mathbf{a} \quad (14)$$

Since $\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{H}(\mathbf{u})$ is independent of $\mathbf{u}$, the Hamiltonian minimum occurs at one of the extreme ends, depending on the sign of $\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{H}(\mathbf{u})$. This gives

$$u_k^*(t) = \begin{cases} u_k^{\min}: & \epsilon w_k(t) > \gamma\lambda_3(t)a_k \\ u_k^{\max}: & \epsilon w_k(t) < \gamma\lambda_3(t)a_k \end{cases} \quad (15)$$

for $k = 1, \ldots, L$.

2) *Quadratic regression*: $h[\mathbf{u}(t)] = b + \mathbf{a}^T[\mathbf{u}^{\max} - \mathbf{u}(t)] + \frac{1}{2}[\mathbf{u}^{\max} - \mathbf{u}(t)]^T\mathbf{S}[\mathbf{u}^{\max} - \mathbf{u}(t)]$, where $\mathbf{a} \geq \mathbf{0}$ and $\mathbf{S} \in \mathbb{R}^L$ is a positive definite matrix (to guarantee the monotonic relationship between the NPI and $\alpha(t)$). Multivariate polynomial fitting can be used to find $\mathbf{S}$, $\mathbf{a}$ and b from historic NPI and case-report data. Note that despite the quadratic form of $h[\mathbf{u}(t)]$, the model is linear in terms of the unknown parameters b, $\mathbf{a}$ and $\mathbf{S}$. Therefore, the required constraint polynomial fitting is straightforward to

implement by conventional least squares solvers. In this case:

$$\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{H}(\mathbf{u}) = \epsilon \mathbf{w}(t) - \gamma\lambda_3(t)\{\mathbf{a} + \mathbf{S}[\mathbf{u}^{\max} - \mathbf{u}(t)]\} \quad (16)$$

and setting $\boldsymbol{\nabla}_{\mathbf{u}}\mathcal{H}(\tilde{\mathbf{u}}) = \mathbf{0}$ gives

$$\tilde{\mathbf{u}} = \mathbf{u}^{\max} - \mathbf{S}^{-1}[\frac{\epsilon \mathbf{w}(t)}{\gamma\lambda_3(t)} - \mathbf{a}] \quad (17)$$

From the *second partial derivative test*, since $\boldsymbol{\nabla}_{\mathbf{u}}^2\mathcal{H} = \gamma\lambda_3(t)\mathbf{S}$, and the fact that $\mathbf{S}$ is assumed to be positive definite: 1) if $\lambda_3(t) > 0$, $\tilde{\mathbf{u}}$ is a local minimum, 2) if $\lambda_3(t) < 0$, it is a local maximum, and 3) $\lambda_3(t) = 0$ results in a saddle point. Therefore, by applying Pontryagin's minimum principle and considering the admissible input range $[\mathbf{u}^{\min}, \mathbf{u}^{\max}]$, after some basic algebra we find

$$u_k^*(t) = \begin{cases} u_k^{\min}: & \epsilon w_k(t) > \gamma\lambda_3(t)(a_k + s_k) \\ \tilde{u}_k: & \gamma\lambda_3(t)a_k < \epsilon w_k(t) < \gamma\lambda_3(t)(a_k + s_k) \\ u_k^{\max}: & \epsilon w_k(t) < \gamma\lambda_3(t)a_k \end{cases}$$
$$(18)$$

where $\tilde{u}_k$ and s_k are the kth entries of the vectors $\tilde{\mathbf{u}}$ and $\mathbf{S}(\mathbf{u}^{\max} - \mathbf{u}^{\min})$, respectively. Comparing (18) and (15), it is clear how the quadratic case simplifies to the linear case when $\mathbf{S} \to 0$. It is also seen that in the linear case, the "optimal NPI" is always one of the extreme cases $u_k^{\min}$ (no action) or $u_k^{\max}$ (maximum stringency). But when the NPI to contact rate map $h(\cdot)$ is nonlinear, intermediate interventions may also be in the optimal NPI set.

V. A UNIFIED PREDICTOR AND PRESCRIPTOR

For an ideal model, the state and co-state dynamic equations detailed in Section IV can be solved with numerical toolboxes for finite horizon control (cf. [8] for a MATLAB-based solution). However, in practice, there are some major issues, which limit the numerical performance, including: 1) model inaccuracies, 2) noisy observations (inaccurate case reports), 3) missing reports (due to holidays), 4) unknown or variable parameters, 5) the difficulty of incorporating the start- and end-point boundary conditions from (13).

Due to these issues, we have developed a novel technique, based on optimal state estimation. Accordingly, we have integrated the finite horizon NPI optimizer and the new-case predictor in a classical EKF and EKS scheme [9]. Using (1), (3) and (11), the unified dynamic equations for the EKF are as follows:

$$\begin{aligned}
\dot{s}(t) &= -\alpha(t)s(t)i(t) + w_s(t) \\
\dot{i}(t) &= \alpha(t)s(t)i(t) - \beta i(t) + w_i(t) \\
\dot{\alpha}(t) &= -\gamma\alpha(t) + \gamma h[\mathbf{u}^*(t)] + w_\alpha(t) \\
\dot{\lambda}_1(t) &= [\lambda_1(t) - \lambda_2(t) - (1-\epsilon)]\alpha(t)i(t) + \eta_1(t) \\
\dot{\lambda}_2(t) &= [\lambda_1(t) - \lambda_2(t) - (1-\epsilon)]\alpha(t)s(t) + \beta\lambda_2(t) + \eta_2(t) \\
\dot{\lambda}_3(t) &= [\lambda_1(t) - \lambda_2(t) - (1-\epsilon)]s(t)i(t) + \gamma\lambda_3(t) + \eta_3(t) \\
n(t) &= \alpha(t)s(t)i(t) + v(t)
\end{aligned}$$
$$(19)$$

where the first six equations are the state and co-state dynamics, the last equation is the observation equation and $h[\mathbf{u}^*(t)]$ is the impact of the optimal control calculated from (15) or (18). The terms $w_s(t)$, $w_i(t)$, $w_\alpha(t)$, $\eta_1(t)$, $\eta_2(t)$ and $\eta_3(t)$ in (19) represent process noises, and $v(t)$ is observation noise.

For an estimation based on the total number of confirmed cases, the last observation equation in (19) can be replaced with (4). The discretized versions of (19), which are required for the discrete-dime implementation of the EKF and EKS are detailed in the Appendix. The MATLAB implementation of the developed algorithm is available in our online repository.

VI. MODEL TRAINING

The model parameters $h[\mathbf{u}(t)]$, β, γ and the EKF/EKS parameters (initial/final states and covariance matrices) require region-wise training or fact-based selection. For this study, we used classical techniques for *Kalman filter engineering* [10, Ch. 8], based on monitoring the properties of the *innovations process* of the Kalman filter and adapting them automatically over time.

The mapping $h[\mathbf{u}(t)]$ was trained over the Oxford dataset historic cases and NPIs starting from January 1, 2020 to Feb 7, 2021 [2]. For this, the developed EKS was first applied to the historic data by assuming $h[\mathbf{u}(t)] = 0$. Referring to (19), this assumption is practically equivalent with considering the input-driven fluctuations of $\alpha(t)$ inside the process noise $w_\alpha(t)$. This gives us a primary estimate of $\alpha(t)$ over the training period, which is given to a constrained LASSO or quadratic polynomial fitter (for the linear and quadratic forms presumed in Section IV, respectively), to estimate $h[\mathbf{u}(t)]$ using the historic NPI data. The trained model $h[\mathbf{u}(t)]$, together with the historic data is used in a second round of EKS; this time by using the historic NPI and apparently a smaller *a priori* assumption for the variance of $w_\alpha(t)$. After the secondary EKS, the new estimates of $\alpha(t)$ are once more used to refine the model parameters of $h[\mathbf{u}(t)]$. The refined parameters are stored per country/region for utilization during the prescription phase (over real or synthetic scenarios).

The *action to effect rate* parameter γ was selected intuitively. From different social experiences, it is reasonable to expect a smooth transition in $\alpha(t)$ due to any change in the NPI. This is based on the social experience that imposing any policy on a complex social system is rarely abrupt. Although the transition is region and NPI dependent, in order to reduce the model complexity, we have fixed $\gamma=1/(7 \text{ days})=0.1429 \text{ days}^{-1}$, for all regions/countries.

The recovery parameter β was selected by educated guesses from the CDC reports regarding recovery and contagion periods (https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html). Accordingly, multiple scientific studies worldwide have reported that an exposed subject is no longer *infectious* after three to four weeks. This is evidently a stochastic range. To clarify, with an exponential model such as the SI mode, in absence of new infected cases ($\alpha = 0$), we find $l \triangleq i(t_0 + T)/i(t_0) = \exp[-\beta T]$, which can be interpreted as an exponential law for the probability of infectiousness after T time units (days). Combining the model with the CDC reports, we derive the following rule for setting β:

$$\beta = \frac{-\log(\text{probability of contagion after } T \text{ time units})}{T} \tag{20}$$

For the later presented results, we have set the probability of contagion to 0.01 and T=21 days, resulting in β=0.2193 days^{-1}.

Following recent studies [11], the reproduction rate of the pandemic during outbreak was taken to be $\mathcal{R}_0$=2.5, which following (2) together with β were used to initialize $\alpha(t_0)$, which is the contact rate during outbreak.

Note that one of the advantages of the EKF/EKS framework is that the model parameters can also be considered as state variables and be *state augmented* with the other equations to be estimated (or updated over time). This approach can be used for both γ and β to refine the initial educated guesses.

Finally, the regional/national population sizes, as required for normalizing the total and new contaminated cases to the normalized variables of the SI model were obtained from public global population datasets and assumed to remain fix over the study (i.e., immigration, inter-border travels, natural birth/deaths have been neglected in the current model).

VII. RESULTS

We trained and applied the proposed finite horizon optimal controller to all regions/countries with arbitrary NPI cost weights. As proof of concept, examples of the bi-objective optimization space are shown in Fig. 2 for several countries, and compared with 1) minimal stringency $\mathbf{u}(t) = \mathbf{u}^{\min}$, 2) maximal stringency $\mathbf{u}(t) = \mathbf{u}^{\max}$, 3) random constant stringency $\mathbf{u}(t) = \boldsymbol{\kappa}$, $\boldsymbol{\kappa} \in [\mathbf{u}^{\min}, \mathbf{u}^{\max}]$, and 4) random variable stringency $\mathbf{u}(t) \in [\mathbf{u}^{\min}, \mathbf{u}^{\max}]$. Each point in Fig. 2 corresponds to one NPI scenario for the designated country starting from Feb 7, 2021 and the next 120 days. Accordingly, the Pareto efficient front comprises of the NPI points which either have a smaller value of J_0 or J_1, while the non-efficient solutions are the ones for which there exists at least a point that gives a smaller cost of both J_0 and J_1. Therefore the Pareto efficient NPI policies should be close to the origin or along one of the left or bottom axes. It is seen that despite the differences between the different countries, the Pareto optimal points all belong to the proposed algorithm and the optimal point is clearly a function of the bi-objective parameter ϵ.

A. Processing load

The MATLAB version of the codes applied to all regions and countries (235 in total), take less than 30 s to train over the historic cases and NPI on a MacBook Pro laptop with 2.3 GHz Quad-Core Intel Core i7 and 32GB of memory, without notable optimizations. The run-time on the test scenarios takes about 15 s in total (since it contains only one EKS stage during the test period). Therefore, the proposed framework is extremely efficient, which permits the combination of the proposed method with other machine learning methods to reduce the search space and to improve the accuracy on other datasets and under more complicated models such as the Long short-term memory (LSTM).

VIII. CONCLUSION AND FUTURE WORK

In this study, a model-based approach was used for the prediction and prescription of NPI that best balance between
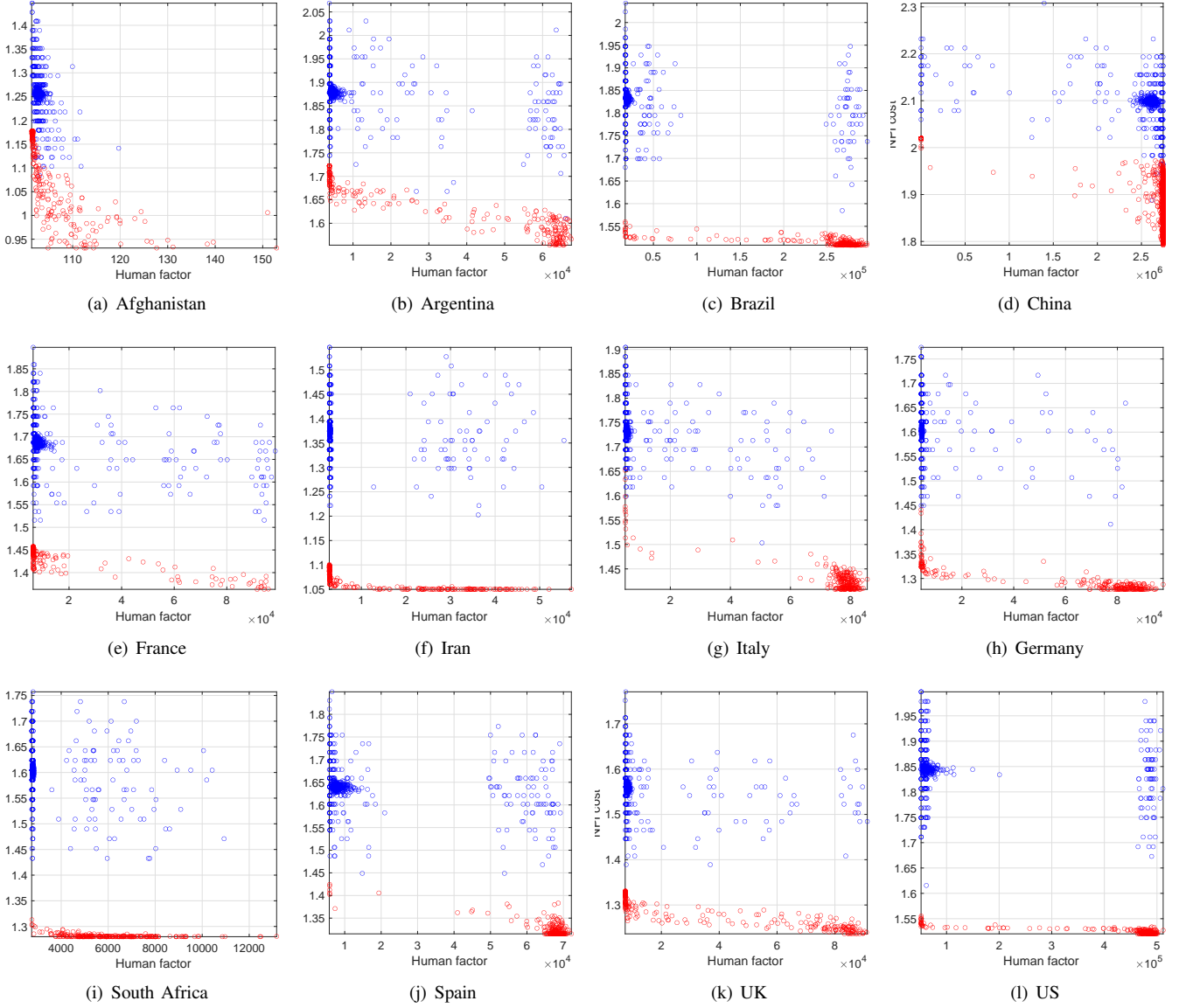
Fig. 2. Biobjective optimization space for sample countries. Blue: random inputs; Red: optimal input for 2000 $\epsilon \in [0, 1]$; Green: Full stringency; Black: No stringency. $h[\mathbf{u}(t)]$ was found by linear regression over historic NPI from Jan 1, 2020 to Feb 7, 2021, using a LASSO with positive coefficients constraint.

an arbitrary weighted-cost of interventions and the number of new cases. The proposed algorithm and the prescribed NPI are proved to be Pareto optimal, to the extent of the accuracy of the utilized compartmental model. Software implementations of the proposed algorithms are online available at: https://github.com/alphanumericslab/EpidemicModeling.

In future studies, different aspects of the method can be extended and improved, including:

- Using advanced machine learning algorithms for learning the NPI to contact rate function $\mathbf{h}(\cdot)$. The LSTM is specifically a promising approach.
- For regions which have access to additional data (e.g., the number of hospitalized, number of vaccinated, fatality rate of the virus, the population age pyramid, etc.), more accurate models such as the fatal susceptible-exposed-infected-recovered (SEIR) can be used [6]. Other indexes such as the daily death reports can be augmented as additional

observation equations in the dynamic model (19), and will help to increase the EKF/EKS accuracies.

- Theoretical aspects of the proposed EKF/EKS frameworks, including stability conditions, parameter identifiability and robustness to parameter and modeling errors require further studies.

APPENDIX
THE DISCRETE-TIME MODEL

For a discrete-time implementation of the EKF/EKS, the discrete form of the dynamic system (19) is required. Accordingly, we define

$$\boldsymbol{s}_k = [s(k\Delta), i(k\Delta), \alpha(k\Delta)]^T$$
$$\boldsymbol{w}_k = [w_s(k\Delta), w_i(k\Delta), w_\alpha(k\Delta), \eta_1(k\Delta), \eta_2(k\Delta), \eta_3(k\Delta)]^T$$
$$n_k = n(k\Delta), \quad c_k = c(k\Delta), \quad v_k = v(k\Delta)$$

where Δ is the discretization time unit. Assuming that Δ is small as compared with the variations of the system, a first order discrete approximation of (19) is found as follows:

$$s_{k+1} = s_k - \Delta\alpha_k s_k i_k + \Delta w_{sk}$$
$$i_{k+1} = i_k + \Delta\alpha_k s_k i_k - \Delta\beta i_k + \Delta w_{ik}$$
$$\alpha_{k+1} = \alpha_k - \Delta\gamma\alpha_k + \Delta\gamma h[\mathbf{u}_k^*] + \Delta w_{\alpha k}$$
$$\lambda_{1,k+1} = \lambda_{1k} + \Delta[\lambda_{1k} - \lambda_{2k} - (1-\epsilon)]\alpha_k i_k + \Delta\eta_{1k}$$
$$\lambda_{2,k+1} = \lambda_{2k} + \Delta[\lambda_{1k} - \lambda_{2k} - (1-\epsilon)]\alpha_k s_k + \Delta\beta\lambda_{2k} + \Delta\eta_{2k}$$
$$\lambda_{3,k+1} = \lambda_{3k} + \Delta[\lambda_{1k} - \lambda_{2k} - (1-\epsilon)]s_k i_k + \Delta\gamma\lambda_{3k} + \Delta\eta_{3k}$$
$$n_k = \alpha_k s_k i_k + v_k$$

$$\tag{21}$$

which can be formulated in a compact form:

$$\boldsymbol{s}_{k+1} = \mathbf{f}(\boldsymbol{s}_k, \boldsymbol{w}_k; h(\mathbf{u}_k))$$
$$n_k = \mathbf{g}(\boldsymbol{s}_k) + v_k \tag{22}$$

where $\mathbf{f}(\cdot)$ and $\mathbf{g}(\cdot)$ represent the nonlinear equations in (21). Note that following (4), if the number of confirmed cases is used as the observation, the second equation in (22) is replaced with

$$c_k = s_0 - s_k + v_k \tag{23}$$

which is a linear function of the state vector.

It is straightforward to linearize the discrete-time dynamic model (21), as required for the implementation of the EKF/EKS. An alternative approach is to use a *continuous-dynamics discrete-observations* approach, which is a well-known method in optimal state estimation. Accordingly, for implementing the EKF/EKS, the state equations can be updated by using the continuous version of the dynamic model (19), while the observations are only updated on discrete-time intervals (e.g., on a daily basis).

REFERENCES

[1] R. Miikkulainen, O. Francon, E. Meyerson, X. Qiu, E. Canzani, and B. Hodjat, "From Prediction to Prescription: Evolutionary Optimization of Non-Pharmaceutical Interventions in the COVID-19 Pandemic," 2020. [Online]. Available: https://arxiv.org/abs/2005.13766

[2] Thomas Hale and Sam Webster and Anna Petherick and Toby Phillips and Beatriz Kira, *Oxford COVID-19 Government Response Tracker*, 2020, Blavatnik School of Government. [Online]. Available: https://github.com/OxCGRT/covid-policy-tracker

[3] XPRIZE, *The XPRIZE Pandemic Response Challenge*, Oct 2020 – Feb 2021. [Online]. Available: https://xprize.org/pandemicresponse

[4] A. Mallela, "Optimal control applied to a SEIR model of 2019-nCoV with social distancing," *medRxiv*, apr 2020.

[5] L. Guan, C. Prieur, L. Zhang, C. Prieur, D. Georges, and P. Bellemain, "Transport effect of covid-19 pandemic in france," *Annual Reviews in Control*, vol. 50, pp. 394–408, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578820300663

[6] R. Sameni, "Mathematical Modeling of Epidemic Diseases; A Case Study of the COVID-19 Coronavirus," 2020. [Online]. Available: https://arxiv.org/abs/2003.11371

[7] D. S. Naidu, *Optimal control systems*. CRC press, 2003.

[8] X. Wang, "Solving optimal control problems with MATLAB: Indirect methods," *ISE Dept., NCSU, Raleigh, NC*, vol. 27695, 2009. [Online]. Available: http://solmaz.eng.uci.edu/Teaching/MAE274/SolvingOptContProb_MATLAB.pdf

[9] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Dover Publications, Inc., 1979.

[10] M. S. Grewal, L. R. Weill, and A. P. Andrews, *Global positioning systems, inertial navigation, and integration*. John Wiley & Sons, 2007.

[11] E. Petersen, M. Koopmans, U. Go, D. H. Hamer, N. Petrosillo, F. Castelli, M. Storgaard, S. A. Khalili, and L. Simonsen, "Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics," *The Lancet Infectious Diseases*, vol. 20, no. 9, pp. e238–e244, sep 2020.