Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 56-2 (2023) 3978–3985

Indirect Model Predictive Control with Sparse Nonlinear Regression on Erdös-Rényi-generated Bernoulli-SIR Network Models

Jonas Hjulstad * **Morten Hovd** **

* *Norwegian University of Science and Technology, (e-mail: jonashj@ntnu.no).*
** *Norwegian University of Science and Technology (e-mail: morten.hovd@ntnu.no)*

Abstract: Epidemiological modeling is important in order to be able to predict and mitigate the consequences of epidemics. Disease transmission network models can be used to model epidemics on a detailed level, which in turn can yield better predictions, at the cost of being more difficult to analyze and control. This paper demonstrates methods that enable simple network models to be controlled with conventional indirect optimal control methods, through model simplification via Monte-Carlo simulations and sparse nonlinear regression.

Copyright © 2023 The Authors. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: identification for control, optimization and control of large-scale network systems, Monte Carlo methods

1. INTRODUCTION

1.1 Background

The present world situation is heavily impacted by the COVID-19 pandemic. With over 613 millon stated cases and 6.52 million confirmed deaths worldwide(Ritchie et al. (2020)), COVID-19 is ranks as one of the deadliest epidemics in history. In order to forecast and prevent viral and bacterial spread in epidemics, it is important to develop methods that are able to accurately predict infections occurring in a future time horizon.

Epidemics on large populations are frequently modeled by dividing the population into compartments. One of the most well-known compartmental models is the Kermack-Mckendrick (SIR) model (Kermack and McKendrick (1927)), which captures the dynamics of infections and recovery from diseases. Under the assumption of homogenous dynamics, these models can yield accurate predictions for large populations.

One challenge with the compartmental models is the ability to validate its parameters using other sources of data than population counts. Many real-life social networks exhibit small-world properties(Watts and Strogatz (1998)), indicating that the true dynamics of disease transmission depend on heterogenous relations in a population. For Agent-Based Models(ABMs, Bissett et al. (2021)), disease transmission networks are able to relate infection probabilities to individuals locations and activities, which could reduce the uncertainty of an epidemics outcome.

Unfortunately, the increased complexity of network models complicates the search for epidemic control policies like social distancing, vaccination and quarantine. For compartmental deterministic models like the SIR-model, control strategies are solvable numerically using Model Predictive Control (MPC) strategies(Sereno et al. (2021)). Despite the long history of epidemiological Models and control theory, there is a limited amount of work related to the unification of the two(Nowzari et al. (2016), Bussell et al. (2019)). Applying MPC directly to network models is computationally expensive. While the local dynamics between nodes in a network are solvable, global graph optimization is in general NP-hard.

One potential solution to avoid the infeasibility of graph optimization is to treat network models as black-box models, which only considers input signals and output measurements of the network. Black-box models lack interpretability and transparency, which are important properties for control strategies that potentially have fatal consequences. A remedy for the interpretability in black-box models is sparse nonlinear regression methods, which have gained traction over the recent years(Brunton et al. (2016)). Sparse nonlinear regression enable the identification of models based on input signals and output measurements, such that the resulting models consists of as few parameters as possible.

The stochastic nature of epidemics is frequently modeled with chain-binomial differential equations(MAIA (1952)), whose posterior distribution is computationally intractable. Instead, Monte Carlo (MC) methods are used with Bayesian inference in order to approximate the posterior. Importance sampling(Kloek and van Dijk (1978)) is frequently used to reduce the dimensionality of the sample space. Other approaches utilize Sequential Monte Carlo(Liu and Chen (1998)) and Approximate Bayesian Computation(Sunnåker et al. (2013)) for likelihood-free problems. These methods are frequently used for parame-

2405-8963 Copyright © 2023 The Authors. This is an open access article under the CC BY-NC-ND license.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2023.10.1361

ter estimation used in prediction and long-term forecasting, and enable simpler epidemiological models to be inferred from ABMs and transmission networks. The methods used in this paper also address parameter inference on transmission networks, but with the goal of identifying simplified, controllable systems.

1.2 Overview

The theory and methods used are separated into four sections, where each section aims to provide brief but detailed descriptions of its relevance and implementation. First, the epidemiological models used in the paper are presented, followed by the methods used to sample data from the models. The next section describes a method for transforming the generated data to models with sparse nonlinear regression. The last theoretical section presents an optimal control problem, and describes how the simplified models can be used to approximate solutions for it. Parameters are then assigned for numerical simulations, which results in the simulation of four different uncontrolled and controlled epidemic scenarios which are presented in the results. Lastly, the performance and limitations of the simulations, regression method and MPC are discussed.

2. EPIDEMIOLOGICAL MODELS

2.1 SIR Model

SIR-dynamics on a completely homogenous population N_{pop} can be expressed with three discrete, deterministic differential equations.

$$N_{S,t+1} = N_{S,t} - \beta \frac{N_{S,t} N_{I,t}}{N_{pop}} \tag{1}$$

$$N_{I,t+1} = N_{I,t} + \beta \frac{N_{S,t} N_{I,t}}{N_{pop}} - \alpha N_{I,t} \tag{2}$$

$$N_{R,t+1} = N_{R,t} + \alpha N_{I,t} \tag{3}$$

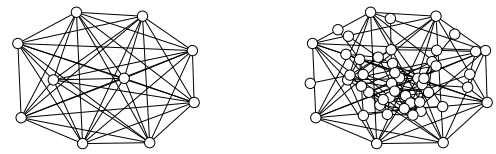
S, I, R denotes the whether an individual is susceptible, infectious or recovered, and $N_{(.)}$ denotes the compartmental population sizes. β is proportional to the number of contacts for each person between each time instance t, and α denotes the average recovery rate. This paper aims to identify parameters similar to α, β by using sparse nonlinear regression on network models.

2.2 Erdös-Rényi Graph Model

Erdös-Rényi models $G_{ER}(N_{pop}, p_{ER})$ (Erdös and Rényi (1959)) are used to generate graph structures for the epidemiological network model. Population count N_{pop} gives the number of vertices N_V, while p_{ER} is used as connection probability to generate N_E edges from all possible node connections in the graph. The degree k for the vertices in the graphs follow binomial distributions.

$$p_k = \binom{N_{pop} - 1}{k} p_{ER}^k (1 - p_{ER})^{N_{pop}-1-k} \tag{4}$$

The variance of the binomial distribution indicates that the graph structure will be heterogenous (high degree variance) when connection probability p_{ER} is close to 0.5, and homogenous (low degree variance) when p_{ER} is close to 0 or 1.



(a) Complete Graph $N_{pop} = 10$, $p_{ER} = 1.0$

(b) Sparsely connected ER-model $N_{pop} = 40$, $p_{ER} = 0.1$

2.3 Bernoulli-SIR Network Model

Each vertex $v_i \in \{0, \dots N_V - 1\}$ is assigned a state $V_i(t) \in \{S = 0, I = 1, R = 2\}$. Heterogeneity in the disease transmission network is fully determined by p_{ER}, while disease transmission parameters are homogenously assigned. Infection probabilities between individuals in the network are uniformly Bernoulli-distributed with $p_{I,t} \in [p_{I,min}, p_{I,max}]$, while recovery probability p_R is fixed. p_I is an idealized parameter which in practice could be related to targeted location and activity control strategies in an ABM, like school/workplace control policies and quarantine recommentdations for subpopulations. These mitigation strategies are likely to have a discrete relation to p_I, but for relatability to (1-3) it is given a continous range. Initial infections are assigned to vertices with fixed probability p_{I_0}, while remaining vertices are assigned susceptible. Under this configuration, a fully connected graph will converge towards (1-3) in its large graph limit.

The temporal dynamics of each individual in the network can be formulated under Markov assumptions.

$$V_{t+1}^{(i)} = \begin{cases} \mathbb{1}\Big(\sum_{j \in \mathcal{N}_{I,t}^{(i)}} U_{I,t}^{(ij)}(\omega) \leq p_{I,t}\Big), & V_t^{(i)} = S \quad (5) \\ 1 + \mathbb{1}(U_{R,t}^{(i)}(\omega) \leq p_R), & V_t^{(i)} = I \quad (6) \end{cases}$$

$\mathcal{N}_I^{(i)}(t)$ denotes the directly connected neighboring vertices v_j of v_i infected at time $t \in [0, N_t]$. $U_{I,t}^{(ij)}, U_{R,t}^{(i)} \sim \text{Uniform}(0, 1)$ denotes the realizations of uniform random samples used to determine infections from v_j to v_i, and infection recoveries. Equations (5) and (6) are spatial dynamics which resembles the infection and recovery dynamics of the original, deterministic SIR-model (Kermack et al. (1927)). The remaining transition probabilities are trivial. $y_t = [N_{S,t}, N_{I,t}, N_{R,t}]^T$ denotes the total count of individuals in each state at time t.

3. MONTE-CARLO IMPORTANCE SAMPLING

Monte-Carlo (MC) importance sampling will be used to approximate the marginal likelihood for the state trajectories. Let $p_{I,t} \sim U(p_{I,min}, p_{I,max})$, and let ξ denote the vector of all random variables $\boldsymbol{U}_{I,0:N_t}, \boldsymbol{U}_{R,0:N_t}, p_{I,0:N_t-1}$ in the network models trajectory. Let $\theta_G = \{G_{ER}, p_{I,0}, p_R\}$ denote all fixed parameters required for initialization of the network. A marginal distribution for the state trajectory can be found by performing an integration over all independent random variables ξ.

$$p(y_{1:N_t}|\theta_G) = \int_{\xi} p(y_{1:N_t}, \xi|\theta_G)p(\xi)d\xi \tag{7}$$

An importance distribution $\pi(.)$ is introduced, which advances (5) and (6) in order to generate trajectories with reduced variance.

$$\mu = \int_{y_{1:N_t}} p(y_{1:N_t}|\theta_G)\pi(y_{1:N_t}|\theta_G)dy_{1:N_t} \tag{8}$$

Statistical moments of interest μ can be approximated by drawing samples of $\{y^{(k)}_{1:N_t}\}^{N_{MC}}_{k=1}$ from $\pi(.)$, where N_{MC} denotes the total number of MC-simulations. The regression algorithm in this paper will approximate the expectation values for the number of infected individuals in these network models.

$$\mathbb{E}_N(N_{I,1:N_t}|\theta_G) = \sum_{k=0}^{N_{MC}-1} N^{(k)}_{I,1:N_t} p(N^{(k)}_{I,1:N_t}|\theta_G) \tag{9}$$

Algorithm 1 Bernoulli-Network Monte-Carlo Sampling

Require: $G_{ER}, p_{I_0}, \boldsymbol{p}_I, p_R,$
  for $v_i \in G_{ER}$ **do**
    $V^{(i)}_0 \leftarrow I$ with probability p_{I_0}
  end for
  for $k \leftarrow 0$ to $N_{MC} - 1$ **do**
    $y^{(k)}_0 \leftarrow [N_{S,0}, N_{I,0}, N_{R,0}]$
    for $t \leftarrow 0$ to $N_t - 1$ **do**
      $p_{I,t} \leftarrow U(p_{I,min}, p_{I,max})$
      for $\{\forall v_i \in G_{ER}|X^{(i)}_t \neq R\}$ **do**
        Assign $V^{(i)}_{t+1}$ with (5), (6)
        Increment total count $N_{S,t+1}, N_{I,t+1}$ or $N_{R,t+1}$
      end for
      $y^{(k)}_{t+1} \leftarrow [N_{S,t+1}, N_{I,t+1}, N_{R,t+1}]$
      if $N_{I,t+1} < N_{I,min}$ **then**
        break
      end if
    end for
  end for
  return $\boldsymbol{y}_{MC} = \begin{bmatrix} \boldsymbol{y}^{(0)} \\ \vdots \\ \boldsymbol{y}^{(N_{MC}-1)} \end{bmatrix}$

4. REGRESSION

4.1 Feature Selection

It is desirable to identify simple, sparse models in order to keep the following optimal control problem interpretable and computationally tractable. Model features are selected from a library of monomials $\mathcal{D}$ with increasing degree order.

$$\boldsymbol{X}^{(k)} = [N^{(k)}_{S,0:N_t-1}, N^{(k)}_{I,0:N_t-1}, p^{(k)}_{I,0:N_t-1}] \tag{10}$$

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}^{(0)} \\ \vdots \\ \boldsymbol{X}^{(N_{MC}-1)} \end{bmatrix}, \quad \boldsymbol{y}_r = \begin{bmatrix} \Delta N^{(0)}_{I,0:N_t-1} \\ \vdots \\ \Delta N^{(N_{MC}-1)}_{I,0:N_t-1} \end{bmatrix} \tag{11}$$

$$P_{\boldsymbol{d}}(\boldsymbol{X}) = \boldsymbol{X}[:,0]^{d_0} \odot \boldsymbol{X}[:,1]^{d_1} \odot \boldsymbol{X}[:,2]^{d_2} \tag{12}$$

$$\mathcal{D}_{\boldsymbol{d}_{max}}(\boldsymbol{X}) = \{P_{[\mathbf{0,0,1}]}(\boldsymbol{X}), P_{[\mathbf{0,1,0}]}(\boldsymbol{X}), \ldots, P_{\boldsymbol{d}_{max}}(\boldsymbol{X})\} \tag{13}$$

$\odot$ denotes the Hadamard product, and element-wise operations are performed on the columns of $\boldsymbol{X}$ for powers $\boldsymbol{d}$ in equation 12. $[:,n]$ denotes the nth column in $\boldsymbol{X}$.

$\boldsymbol{X} \in \mathbb{R}^{(N_t-1)N_{MC}}$ and $\boldsymbol{y}_r \in \mathbb{R}^{(N_t-1)N_{MC}}$ denotes the explanatory and response variable data which will be used to fit regression models for each MC-simulation.

4.2 Polynomial NARX-Model

The proposed nonlinear system identification method aims to fit a Nonlinear AutoRegressive eXogenous model (NARX) which accounts for dynamics caused by control inputs and the previous timesteps state. Let $\mathcal{S} \subset \mathcal{D}_{\boldsymbol{d}_{max}}$ denote a subset selection with a maximum of N_s monomial features. The structure of a polynomial NARX-model $F_{\mathcal{S}} : (\mathbb{R}^3, \mathbb{R}) \rightarrow \mathbb{R}$ can be composed from the monomial subset of features.

$$F_{\mathcal{S}}(x_t, u_t|\boldsymbol{\theta}) = \theta_0\mathcal{S}\{0\}(x_t, u_t) + \theta_1\mathcal{S}\{1\}(x_t, u_t) + \cdots + \theta_{N_s-1}\mathcal{S}\{N_s - 1\}(x_t, u_t) \tag{14}$$

The rate of change in the number of infected ($\boldsymbol{y}_r$) is considered for regression in this paper, which is denoted $F_{\mathcal{S},I}$.

4.3 Sparse Nonlinear Regression

The Forward Regression Orthogonal Least Squares algorithm (FROLS) is used draw and evaluate models from $\mathcal{D}$ in order to obtain $\mathcal{S}$ for (14). FROLS is a greedy algorithm that iteratively orthogonalize and fit coefficients for nonlinear features, resulting in sparse nonlinear models with interpretable features. This paper presents an applied version of FROLS with similar notation to the more generic algorithm presented in Billings (2013).

Let $\boldsymbol{p}_m$ denote the m'th feature drawn from $\mathcal{D}$, and let $\boldsymbol{q}_{m,s}$ denote the m'th orthogonalized feature in feature-selection iteration s for MC-simulation k. $g_{m,s}$ is the covariance-to-variance ratio between the response variable $\boldsymbol{y}_r$ and orthogonalized feature $\boldsymbol{q}_{m,s}$.

$$g_{m,s} = \frac{(\boldsymbol{y}_r)^T\boldsymbol{q}_{m,s}}{(\boldsymbol{q}_{m,s})^T\boldsymbol{q}_{m,s}} \tag{15}$$

Fitness of each feature is determined by the Error Reduction Ratio (ERR), which quantifies how the variance of each feature contributes to the explanation of $\boldsymbol{y}_r$'s total variance.

$$ERR_{m,s} = (g_{m,s})^2 \frac{(\boldsymbol{q}_{m,s})^T(\boldsymbol{q}_{m,s})}{(\boldsymbol{y}_r)^T(\boldsymbol{y}_r)} \tag{16}$$

Algorithm 2 Monomial-Feature FROLS

Require: $\boldsymbol{y}_r, \boldsymbol{X}, \mathcal{D}_{\boldsymbol{d_{max}}}$
  Zero-initialize $\boldsymbol{Q}, \boldsymbol{Q}_s \in \mathbb{R}^{N_t-1\times N_s}$,
  $\boldsymbol{ERR} \in \mathbb{R}^{N_s\times N_m-1}, \boldsymbol{\ell} \in \mathbb{Z}^{+,N_s}, \boldsymbol{g}_s \in \mathbb{R}^{N_m}$
  $\boldsymbol{g} \in \mathbb{R}^{N_s}$
  for $s \leftarrow 0$ to $N_s - 1$ **do**
    for $m \leftarrow 0$ to N_m **do**
      $p_m \leftarrow \mathcal{D}_{\boldsymbol{d_{max}}}(\boldsymbol{X})\{m\}$
      $\boldsymbol{Q}_s \leftarrow$ repeated orthogonalization of p_m on $\boldsymbol{Q}[:,:s]$
      Compute $\boldsymbol{g}_s[m]$ and $\boldsymbol{ERR}[m,s]$
    end for
    $\boldsymbol{\ell}[s] \leftarrow$ index of $\max(\boldsymbol{ERR}[s,:])$
    $\boldsymbol{g}[s] \leftarrow \boldsymbol{g}_s[\boldsymbol{\ell}[s]]$
    $\boldsymbol{Q}[:,s] \leftarrow \boldsymbol{Q}_s[:,\boldsymbol{\ell}[s]]$
    $\boldsymbol{A}[m,s] \leftarrow \frac{\boldsymbol{Q}[:,s]^T\boldsymbol{X}[\ell[s]]}{\boldsymbol{X}[\ell[s]]^T\boldsymbol{X}[\ell[s]]}$
    if $\text{sum}(\boldsymbol{ERR}[\boldsymbol{\ell}]) >$ ERR_threshold **then**
      break
    end if
  end for
  $\boldsymbol{\theta} \leftarrow A[:s,:s]^{-1}\boldsymbol{g}[:s]$
  return $\boldsymbol{\theta}$

Since total population N_{pop} is conserved, remaining NARX-models for the number of susceptibles ($F_{\mathcal{S},S}$) and recovered ($F_{\mathcal{S},R}$) can be retrieved from $F_{\mathcal{S},I}$, yielding full system dynamics $\boldsymbol{F}_{\mathcal{S}}$.

$$\boldsymbol{F}_{\mathcal{S}} = \begin{bmatrix} F_{\mathcal{S},S} \\ F_{\mathcal{S},I} \\ F_{\mathcal{S},R} \end{bmatrix} = \begin{bmatrix} -(\text{positive terms of } F_{\mathcal{S},I}) \\ F_{\mathcal{S},I} \\ -(\text{negative terms of } F_{\mathcal{S},I}) \end{bmatrix} \tag{17}$$

5. OPTIMAL CONTROL PROBLEM

5.1 Robust Model Predictive Control

$$\arg\min_{\boldsymbol{x},\boldsymbol{u}} \quad \Phi(\boldsymbol{x},\boldsymbol{u}) \tag{18}$$
$$s.t \quad x_0 = \bar{x}_0 \tag{19}$$
$$x_{t+1} = \boldsymbol{F}(x_k, u_k|\boldsymbol{\theta}) \quad \forall t\backslash N_t \tag{20}$$
$$\tau \leq P(g(x_t) \leq 0) \quad \forall t\backslash 0 \tag{21}$$
$$u_{\min} \leq u_t \leq u_{\max} \quad \forall t\backslash N_t \tag{22}$$

Equation 21 are *single chance constraints*(Shapiro et al. (2009)) which relate probability distributions to Optimal Control Problems (OCPs) as inequality constraints. These are the constraints targeted for approximation by MC-simulations and regression. Equation 20 ensures that the input variables of objective function $\Phi(.)$ are constrained to follow the true, stochastic system dynamics $\boldsymbol{F}(.)$ at every timestep of the control horizon. Trajectories generated by the MPC are denoted with $\boldsymbol{x}$, which in this application is the predicted population counts $(N_{S,t}, N_{I,t}, N_{R,t})$ over the control horizon. Infection probabilities $\boldsymbol{p}_I$ acts as control inputs over the control horizon, denoted with $\boldsymbol{u}$.

5.2 Regression Model Predictive Control

The control strategy proposed in this paper is kept simple in accordance with the model itself.

$$\arg\min_{\boldsymbol{x},\boldsymbol{u}} \quad \sum_{t=1}^{N_t}\{\frac{N_{I,t}}{N_{pop}} - \frac{W_u}{N_{pop}}(u_t - u_{\max})\} \tag{23}$$
$$s.t \quad x_0 = \bar{x}_0 \tag{24}$$
$$x_{t+1} = F_{\mathcal{S}}(x_k, u_k|\boldsymbol{\theta}) \quad \forall t\backslash N_t \tag{25}$$
$$x_t[1] \leq I_{max} \quad \forall t\backslash 0 \tag{26}$$
$$u_{\min} \leq u_t \leq u_{\max} \quad \forall t\backslash N_t \tag{27}$$

The motivation for the proposed optimal control objective is to "flatten the curve" within the given control horizon N_t while the impact of social distancing measures are minimized. W_u is used to quantify the cost of social distancing restrictions. As a replacement for the chance constraint in equation 21, the regression MPC is instead constrained to follow an identified NARX-model $F_{\mathcal{S}}(.)$ at every timestep. In order to utilize FROLS-regression the chance constraint is replaced with (9).

$$\mathbb{E}_N(N_{I,1:N_t}|\theta_G) \approx \begin{bmatrix} x_0 + \sum_{t=0}^{0} \boldsymbol{F}_{\mathcal{S},I}(x_t, u_t) \\ \vdots \\ x_0 + \sum_{t=0}^{N_t-1} \boldsymbol{F}_{\mathcal{S},I}(x_t, u_t)] \end{bmatrix} \tag{28}$$

6. SIMULATIONS

6.1 Monte-Carlo Sampling

Models are generated using different population sizes N_{pop} and connection probabilities p_{ER}, with the intention of demonstrating system identification and predictive control for both heterogenous and homogenous structures. Population sizes are kept small due to the computational constraints caused by the number of edges in the networks. The MC-sampling is implemented in C++ with CPU-parallel execution on single networks.

Table 1. Graph structure parameter sets for ER-models

N_{pop}	p_{ER}	$p_{I,max}$
50	0.1	0.05
50	1.0	0.005
100	0.1	0.001
100	1.0	0.3
200	0.1	0.2
200	1.0	0.1

Other parameters related to MC-simulations, regression and optimal control remain fixed in order to generate comparable results. $p_{I,t}$ is for the MC-simulations constrained to change at the same time instances as the control input u_t, which in this situation changes every 7 timesteps. Simulations are terminated early when $N_I < N_{I,min}$.

Table 2. Fixed parameter set for the simulation scenarios

p_{I0}	0.1	p_{R0}	0.05
u_{min}	10^{-3}	u_{max}	0.3
p_R	0.1	$p_{I,min}$	0
N_{MC}	500	N_t	70
$N_{I,min}$	2	W_u	10^3

6.2 Regression

FROLS algorithm is implemented in C++ with CPU-parallel evaluation of features. Maximum polynomial order d_{max} is fixed to 2, and the maximum number of selected features N_s is constrained to 2.

6.3 Optimal Control Problem

OCPs are constrained with the regression NARX-models from result table 3 as $F_{\mathcal{S}}$, and constructed symbolically using CasADI(Andersson et al. (2019)) in Python. The NARX-model constraints cause nonlinearities in OCPs constraints, which requires a nonlinear numerical solver in order to find local optimal solutions. The dimensionality of the problem makes the Interior-Point OPTimizer (IPOPT)(Wächter and Biegler (2004)) a good solver to use. For numerical stability, the integration of $\boldsymbol{F}_{\mathcal{S}}$ in constraints (25) are enforced with a direct multiple-shooting integration scheme (Stoer and Bulirsch (2002)). Control inputs u_t are limited to change every 7 timesteps over a control horizon of 56 timesteps.

7. RESULTS

Running N_{MC} simulations over N_t timesteps on the graphs generated by table 1 resulted in the trajectories shown in figure 2. Monomial-Feature FROLS on the MC-data resulted in NARX-models shown in Table 3.

Table 3. Identified NARX parameters

N_{pop}	p_{ER}	NARX-Model
50	0.1	$\Delta S_{t+1} = -0.024 S_t I_t p_I$ $\Delta I_{t+1} = -0.092 I_t + 0.024 S_t I_t p_I$ $\Delta R_{t+1} = 0.092 I_t$
50	1.0	$\Delta S_{t+1} = -0.301 S_t I_t p_I$ $\Delta I_{t+1} = -0.093 I_t + 0.301 S_t I_t p_I$ $\Delta R_{t+1} = 0.093 I_t$
100	0.1	$\Delta S_{t+1} = -0.028 S_t I_t p_I$ $\Delta I_{t+1} = -0.096 I_t + 0.028 S_t I_t p_I$ $\Delta R_{t+1} = 0.096 I_t$
100	1.0	$\Delta S_{t+1} = -0.004 S_t^2 I_t p_I$ $\Delta I_{t+1} = -0.096 I_t + 0.004 S_t^2 I_t p_I$ $\Delta R_{t+1} = 0.096 I_t$
200	0.1	$\Delta S_{t+1} = -0.042 S_t I_t p_I$ $\Delta I_{t+1} = -0.098 I_t + 0.042 S_t I_t p_I$ $\Delta R_{t+1} = 0.098 I_t$
200	1.0	$\Delta S_{t+1} = -0.002 S_t^2 I_t p_I$ $\Delta I_{t+1} = -0.096 I_t + 0.002 S_t^2 I_t p_I$ $\Delta R_{t+1} = 0.096 I_t$

Solving OCP (18)-(22) with the NARX models in table 3 resulted in trajectories shown in figures 3-8.

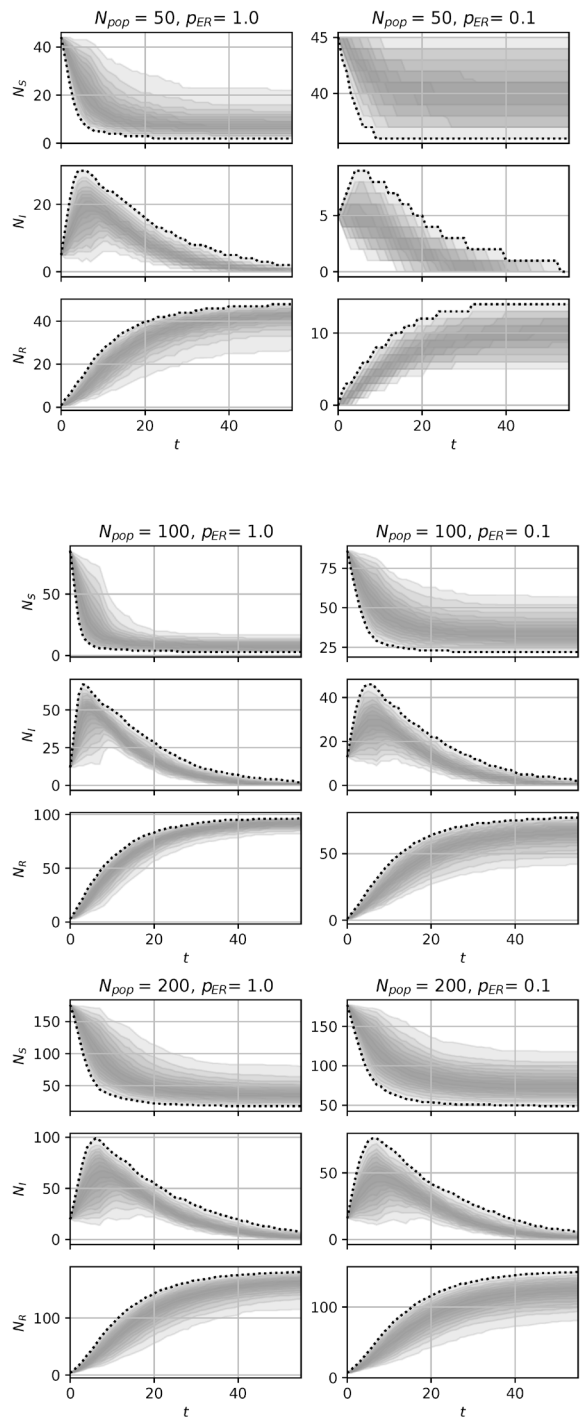


Fig. 2. Trajectories for uncontrolled MC simulations. (95-percentile shown with dotted lines)

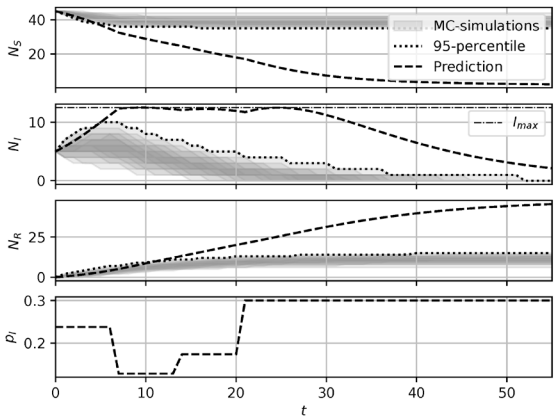


Fig. 3. Optimal control solution for $N_{pop} = 50$, $p_{ER} = 0.1$.

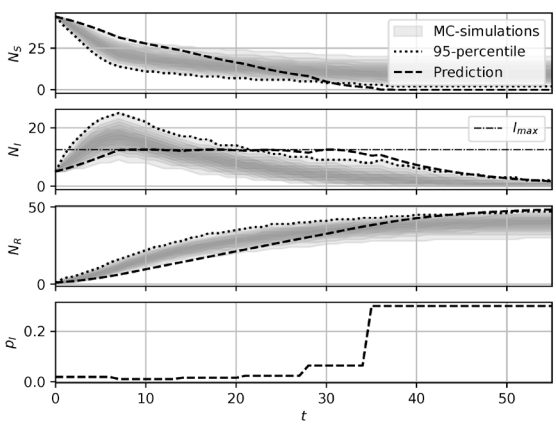


Fig. 4. Optimal control solution for $N_{pop} = 50$, $p_{ER} = 1.0$.

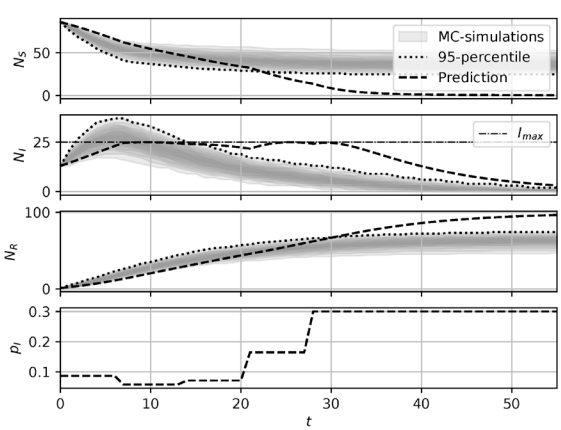


Fig. 5. Optimal control solution for $N_{pop} = 100$, $p_{ER} = 0.1$.

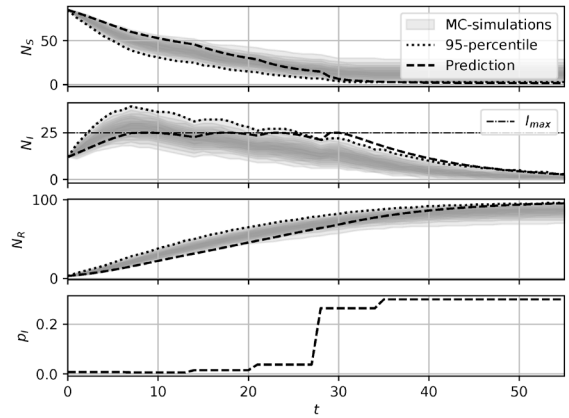


Fig. 6. Optimal control solution for $N_{pop} = 100$, $p_{ER} = 1.0$.

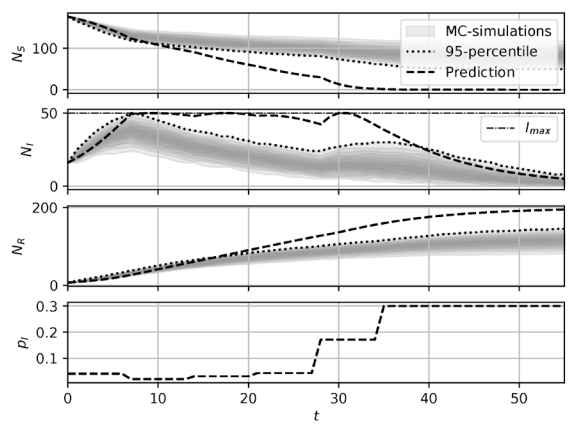


Fig. 7. Optimal control solution for $N_{pop} = 200$, $p_{ER} = 0.1$.

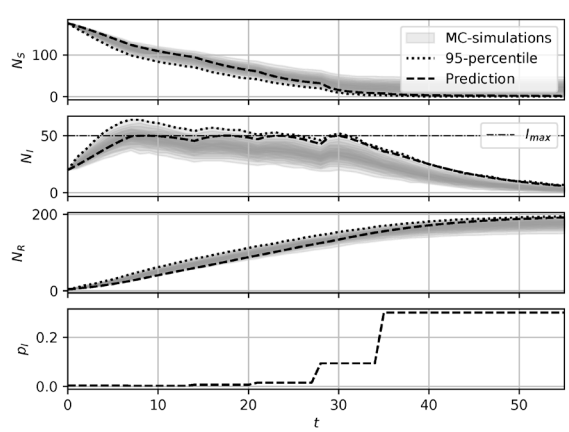


Fig. 8. Optimal control solution for $N_{pop} = 200$, $p_{ER} = 1.0$.

8. DISCUSSION

The solutions robustness with respect to (21) have many dependencies. The model simplification requires MC-simulations that provide a sufficient amount of excitation of the networks with respect to the control input p_I, whose input space will be harder to determine beforehand for heteregenous networks. Given a sufficiently excited network with constrained timeseries, the FROLS regression algorithm is able to identify the same features used by the structureless SIR-model.

Performing model predictive control on these models requires the features identified by FROLS to be feasible with respect to the optimization objective. False correlations identified by FROLS may result in gradient search directions in the optimization problem that is wrong with respect to the underlying network dynamics. This, combined with computational constraints was found to be very restrictive for the size of monomial dictionary $\mathcal{S}$, and the maximum number of terms N_s that could be used.

The accuracy of the regression models increases with the population size of the network and the number of edges connecting its individuals. This indicates that homogenous network clusters is well approximated by compartmental models, while smaller clusters with uneven distributions of edges strongly depends on the network structure. Due to the difference in accuracy, optimal control on large, fully connected networks enforce the infected capacity constraint (26) better than the smaller networks with fewer edges.

The proposed control strategy and outcome for networks with $p_{ER} = 0.1$ are similar for all population sizes. These strategies underestimate impact of the early infection spread, resulting in infection capacity violations during $t \in [3, 10]$ for the majority of the MC-simulations. In contrast, the fully connected larger populations upholds the capacity constraint for the majority of its MC-simulations.

The replacement of (21) with (25) and (26) only approximate the first order momentum of the distribution, which can become very inaccurate for distributions with high variance. In the case of infection capacity constraints, introducing a lower I_{max} could bring the worst-case MC-scenarios down to the desired capacity.

The long-term validity of the identified sparse nonlinear models enable the MPC to evaluate long-term consequences for the epidemics. Feedback control with offline identification of models can be used to improve MPC performance in cases where this does not hold.

While the practical applications for a network with these small population sizes is limited, a metapopulation model with community-level nodes(Bailey (1986)) could be used in practice. Approximating such a network is likely to require larger MC-simulation datasets to capture its dynamics, which in turn requires efficient algorithms for MC-simulations and regression. While this papers implementation is C++-based and targeted towards CPUs, a SYCL-based (Maria Rovatsou and Keryell (2022)) implementation is planned for acceleration in future projects.

REFERENCES

Andersson, J.A.E., Gillis, J., Horn, G., Rawlings, J.B., and Diehl, M. (2019). CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1), 1–36. doi: 10.1007/s12532-018-0139-4.

Bailey, N.T. (1986). Macro-modelling and prediction of epidemic spread at community level. *Mathematical Modelling*, 7(5), 689–717. doi:https://doi.org/10.1016/0270-0255(86)90128-4.

Billings, S. (2013). Nonlinear system identification: Narmax methods in the time, frequency, and spatio-temporal domains. *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. doi:10.1002/9781118535561.

Bissett, K.R., Cadena, J., Khan, M., and Kuhlman, C.J. (2021). Agent-based computational epidemiological modeling. *Journal of the Indian Institute of Science*, 101(3), 303–327. doi:10.1007/s41745-021-00260-2.

Brunton, S.L., Proctor, J.L., and Kutz, J.N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937. doi:10.1073/pnas.1517384113.

Bussell, E.H., Dangerfield, C.E., Gilligan, C.A., and Cunniffe, N.J. (2019). Applying optimal control theory to complex epidemiological models to inform real-world disease management. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 374(1776), 20180284.

Erdös, P. and Rényi, A. (1959). On random graphs i. *Publicationes Mathematicae Debrecen*, 6, 290.

Kermack, W.O., McKendrick, A.G., and Walker, G.T. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772), 700–721. doi: 10.1098/rspa.1927.0118.

Kermack, W.O. and McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772), 700–721.

Kloek, T. and van Dijk, H.K. (1978). Bayesian estimates of equation system parameters: An application of integration by monte carlo. *Econometrica*, 46(1), 1–19.

Liu, J.S. and Chen, R. (1998). Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93(443), 1032–1044. doi: 10.1080/01621459.1998.10473765.

MAIA, J.D.O.C. (1952). Some mathematical developments on the epidemic theory formulated by reed and frost. *Human Biology*, 24(3), 167–200.

Maria Rovatsou, L.H. and Keryell, R. (2022). Sycl specification. *Khronos Registry*.

Nowzari, C., Preciado, V.M., and Pappas, G.J. (2016). Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems Magazine*, 36(1), 26–46. doi: 10.1109/MCS.2015.2495000.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., and Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*.

Https://ourworldindata.org/coronavirus.

Sereno, J., D'Jorge, A., Ferramosca, A., Hernandez-Vargas, E., and González, A. (2021). Model predictive control for optimal social distancing in a type sir-switched model. *IFAC-PapersOnLine*, 54(15), 251–256. doi:https://doi.org/10.1016/j.ifacol.2021.10.264. 11th IFAC Symposium on Biological and Medical Systems BMS 2021.

Shapiro, A., Dentcheva, D., and Ruszczyński, A. (2009). *Lectures on stochastic programming. Modeling and theory.* Society for Industrial and Applied Mathematics. doi:10.1137/1.9780898718751.

Stoer, J. and Bulirsch, R. (2002). *Ordinary Differential Equations*, 465–618. Springer New York, New York, NY. doi:10.1007/978-0-387-21738-3_7.

Sunnåker, M., Busetto, A.G., Numminen, E., Corander, J., Foll, M., and Dessimoz, C. (2013). Approximate bayesian computation. *PLoS Comput Biol*, 9(1), e1002803.

Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442. doi:10.1038/30918.

Wächter, A. and Biegler, L.T. (2004). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming.