Test-Based Stabilizing Control of COVID-19 Transmission

Nader Sadegh[1] and Payman Sadegh[2]

[1] Georgia Institute of Technology, Atlanta, GA 30332, USA,
sadegh@gatech.edu,
https://www.me.gatech.edu/faculty/sadegh
[2] ChainSquared, Alpharetta, GA 30009,
payman@chainsquared.com
https://www.linkedin.com/in/payman-sadegh/

Abstract. The COVID-19 pandemic has confronted the world with a health and an economic crisis not seen since 1918. Appearing first in Wuhan, China, in late 2019, the virus has now spread to over 150 countries. At the time of this writing, hundreds of thousands have already been infected and tens of thousands have died as a result. Given the virus' high degree of contagion, a question that occupies everyone's mind is whether a strategy could be conceived to suppress or at least substantially slow the spread of the virus until a vaccine becomes widely available. Using an extended SEIR model with COVID-19's available epidemiological parameters, this paper studies a feedback control solution to stabilize the virus' transmission dynamic based on a test-and-isolate strategy. The derived theoretical results combined with simulation analyses are used to answer key questions around testing strategies that will be required for suppression of the virus.

1 Introduction

COVID-19 has posed an unparalleled challenge to nations around the globe. Appearing first in Wuhan, China, in late 2019, the virus has now spread to over 150 countries. The disease prompted China to enforce extreme measures to stop the virus. It has ravaged Italy's healthcare system and caused thousands of deaths. Millions of people around the world are either living in self-quarantine or government-imposed lock-downs. If the virus is allowed to run its course, healthcare systems will be overloaded, millions will die, and economies will be devastated. Social distancing measures, such as school closings and other bans on gatherings beyond a certain size, have been widely adopted including in the United States. Although social distancing may be both proper and necessary as part of a broader strategy, when taken alone it can only mitigate the spread, lasting as long as the related measures are in place. Additionally, a singular focus on social distancing and expecting the related measures to bear the brunt of the effort will come at an enormous economic cost, as already evident from various economic indicators. Even if the growth of the disease is temporarily halted, there is the risk of additional waves of the pandemic as was the case with the Spanish Flu of 1918.

Our goal in this paper is to investigate the viability of a test-and-isolate strategy for suppression of COVID-19 spread. The strategy works as follows. A set of criteria are developed according to which members of the population qualify for testing. These criteria could include not only the disease symptoms such as coughs or fever but also criteria such as social interaction, geography or age. Once the results are available, all positively tested individuals are subsequently isolated to prevent further transmission of the virus. Of interest in this paper is establishing a testing strategy, including a minimum number of tests that need to be administered within a given period of time (daily) such that the test-and-isolate approach would lead to suppression of the virus.

2 Problem Formulation

In the analysis that follows, we utilize the widely-known SEIR disease transmission model (SEIR stands for Susceptible-Exposed-Infectious-Recovered). SEIR extends the basic SIR model [1] by adding an *exposed* state to account for the incubation period of a disease, during which the individual does not transmit the disease. In the case of COVID-19, the transmission within the incubation period can in fact take place, although it is thought to be infrequent [2]. If needed, the assumption of no transmission from the exposed state can be relaxed in a variety of ways by either removing the exposed state altogether (i.e. using a basic SIR model as has been commonly used for COVID-19 modeling), assigning a small transmission rate from the exposed state, or simply using coefficients that assume a shorter incubation period for the disease than generally reported. None of these relaxations will be material to the methodology presented here as our primary goal is to demonstrate how the model may be utilized for development of effective test-and-isolate strategies. Similar to the standard model, we assume once individuals recover from the infection, they will become immune to the disease.

The following equations describe the dynamic of the SEIR model:

$$\begin{aligned} \dot{S} &= \epsilon N - \frac{\beta I S}{N} - \epsilon S \\ \dot{E} &= \frac{\beta I S}{N} - (\alpha + \epsilon) E \\ \dot{I} &= \alpha E - (\mu + \gamma) I \\ \dot{R} &= \gamma I - \epsilon R \end{aligned} \tag{1}$$

where

- S, E, I, R represent population sizes that are in the susceptible, exposed, infected, and recovered states, $N = S + E + I + R$.
- α is the fraction of the exposed population that becomes infected (i.e. can start transmitting the disease) over the unit time.
- β is the percentage of the susceptible population S that transitions to exposed E over the unit time for every one percentage of infection among the population N. β can also be interpreted as the number of people exposed to the virus by one individual in I per unit time at the beginning of the outbreak when the entire population is susceptible to the disease, i.e $N \approx S$.
- γ is the fraction of the infected population who recover over the unit time.
- ϵ is the general population birth and death rate, which are assumed to be equal during the study period.
- μ is the death rate among the infected population ($\mu \geq \epsilon$).

The model coefficients α-γ, ϵ, and μ are not necessarily constant. They may in fact vary with time during the course of study, especially as various social distancing measures are introduced or lifted. The only difference between the system of equations in (1) and the standard SEIR model is that the death-rate from the disease is larger than the natural population death rather, a relaxation introduced given the short time scales of COVID-19 propagation versus generally longer time frames of the standard SEIR model.

Once an individual enters the infected state, the average time of exit from the state through either recovery or death is equal to $(\mu+\gamma)^{-1}$. Recall when $S \approx N$, β is equal to the average number of individuals exposed to the virus by an individual in I *per unit time*. But ϵE of the exposed individuals die per unit time and eventually only $\alpha(\epsilon + \alpha)^{-1}$ fraction of them enter the infected

state. Therefore, the average number of patients that eventually become infected by an individual in I, aka Basic Reproduction Number (BRN), is given by

$$\rho_0 = \frac{\alpha}{\epsilon + \alpha} \frac{\beta}{\gamma + \mu} \tag{2}$$

The BRN is believed to be in the 2-3 range for COVID-19 [3]. This will be utilized later when parameters are set for control design and simulations.

Although through measures such as social distancing it may be possible to affect the disease transmission coefficient such as β, the system of equations in (1) otherwise describes the growth of the virus without any intervention. The test-and-isolate strategy will introduce a controlled input U and a quarantine (isolate) state Q as follows.

$$\begin{aligned}
\dot{S} &= \epsilon N - \frac{\beta I S}{N} - \epsilon S \\
\dot{E} &= \frac{\beta I S}{N} - (\alpha + \epsilon) E \\
\dot{I} &= \alpha E - (\mu + \gamma) I - U \\
\dot{Q} &= U - (\mu + \gamma) Q \\
\dot{R} &= \gamma I + \gamma Q - \epsilon R \\
N &= S + E + I + Q + R
\end{aligned} \tag{3}$$

The introduced variable and state are related to the control strategy in the following way. At any given instance t, an intervention is made by identifying and selecting $U(t)$ individuals for removal from I and moving them into isolation Q. This implies that they will no longer be able to transmit the disease. The exit from the quarantine takes place after recovery or death at the same rate as recovery and death from the infected state takes place.

3 Control Design

In this section we are concerned with the design of an output feedback control strategy. The control law we consider has the form

$$U = KI \tag{4}$$

as the infected individuals I are the ones that can be at least partially identified by testing.

In an unreal universe where every infected person could be identified with ease and moved to quarantine, we would obviously be able to suppress the virus in no time. In the real world, however, we only have partial knowledge about who may be infected and therefore are bounded by the level of intervention that could be made. To reflect that constraint, we seek the smallest gain K that stabilizes (3), hence flattening E, I, and Q or suppressing them toward zero at a rate depending on how aggressive the control law is enforced. An estimate of such a K can be found by linearizing (3) at $t = 0$ about $(N_0, 0, 0, 0, 0)$ where N_0 is the initial population. Letting $x = [S - N_0 \ E \ I \ Q \ R]^T$, the linearized system may be expressed by $\dot{x} = \mathbf{A}x$ where

$$\mathbf{A} = \begin{bmatrix}
0 & \epsilon & \epsilon - \beta & \epsilon & \epsilon \\
0 & -\alpha - \epsilon & \beta & 0 & 0 \\
0 & \alpha & -(\gamma + \mu + K) & 0 & 0 \\
0 & 0 & K & -(\gamma + \mu) & 0 \\
0 & 0 & \gamma & \gamma & -\epsilon
\end{bmatrix} \tag{5}$$

The eigenvalues of $\mathbf{A}$ are the roots of its characteristic equation, $|\lambda\mathbf{I} - \mathbf{A}|$ given by $|\lambda\mathbf{I} - \mathbf{A}| = \lambda(\lambda+\epsilon)(\lambda+\gamma+\mu)C(\lambda)$ where

$$C(\lambda) = \lambda^2 + (\alpha+\epsilon+\gamma+\mu+K)\lambda + (\alpha+\epsilon)(\gamma+\mu+K) - \alpha\beta \tag{6}$$

The eigenvalues of $\mathbf{A}$ are guaranteed to be in the left half of the complex plane (i.e., stable) if and only if $(\alpha+\epsilon)(\gamma+\mu+K) \geq \alpha\beta$ or equivalently $K \geq (\rho_0 - 1)(\gamma+\mu)$. The following Theorem proves that if K is chosen slightly larger than this minimum K, both E and I of the full nonlinear SEIR controlled system converge to zero exponentially starting from arbitrary positive initial conditions.

Theorem 1. *Consider the controlled SEIR system* (3)*subject to control law* $U(t) = K(t)I(t)$*. The state variables* $E(t)$*,* $I(t)$*, and* $Q(t)$ *converge to zero exponentially starting form any* $S(0)$*,* $E(0)$*,* $I(0)$*,* $Q(0)$*,* $R(0) > 0$ *provided the time-varying gain* $K(t) \geq K_{\min}(t) + \delta$ *for some fixed* $\delta > 0$ *where*

$$K_{\min} := (\rho - 1)(\gamma + \mu) \;\; \textit{and} \;\; \rho = \frac{(\alpha + \beta S N^{-1})^2}{4(\alpha+\epsilon)(\mu+\gamma)}$$

In particular, there exists a constant $\sigma > 0$ *such that*

$$E^2(t) + I^2(t) \leq e^{-\sigma t}(E^2(0) + I^2(0)),\ \forall t \geq 0$$

Proof. See Appendix A.

The initial value of $K_{\min}$ assuming $S(0)/N(0) = 1$ based on the assumed parameter values for COVID-19 is 0.0898, which is fairly close to 0.0879 based on the linearized system. Setting $K > K_{\min}$ guarantees exponential convergence for the actual nonlinear SEIR system. It is also worth mentioning that $K_{\min}(t)$ typically decreases with t as $r(t) = S(t)/N(t)$ has its highest value initially. This means that about 9% of the infected population needs to be tested and isolated daily in order for the infected population to decrease. In the next section, we explore the number of tests that need to be performed in order to reach a stabilizing threshold.

4 Computing the Size of the Testing Population

In the previous section, we computed a control law for stabilization of the disease transmission dynamic, expressed as the percentage of the infected population that ought to be tested on a given cadence (e.g. daily). The problem is that while we have indication as to the probability of infection based on symptoms or other factors, the infected population is not directly observed. The question we address in this section is how many people need to be sampled for testing such that a sufficient number of infected individuals are identified for the test-and-isolate strategy.

Let the sets $\mathit{\Pi}$ and Z denote the infected group and the group of individuals that meet the testing criteria, respectively, as shown in Figure 1. We also refer to Z as the *sampling universe*. Each sample from Z returns an infected individual, i.e. a member of $\mathit{\Pi}$, with the probability π where

$$\pi = \frac{|\mathit{\Pi} \cap Z|}{|Z|} = \Pr(\mathit{\Pi}|Z)$$

and $|.|$ denotes set cardinality. The higher the *specificity* of the test selection criteria, the larger π will be.

Similarly, let $\zeta = \Pr(Z|\mathit{\Pi})$, i.e. the probability that the test selection criteria are manifest in the infected individuals. The higher the *sensitivity* of the selection criteria, the larger ζ will be. If the testing criteria are too restrictive, for example allowing tests for severe symptoms only, they

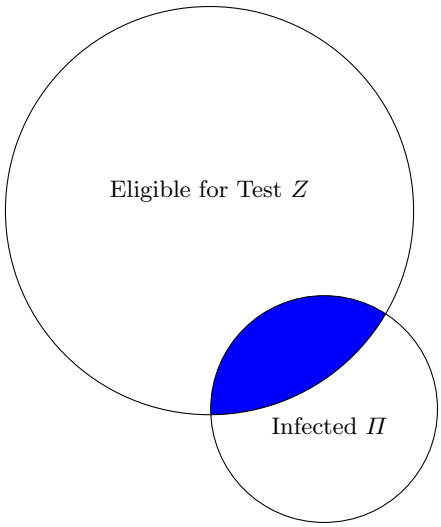

Fig. 1. The sets of Z versus Π

will have a high specificity value π but at the expense of lower sensitivity ζ since they would likely leave out many infected individuals who demonstrate only mild symptoms. The maximum number of individuals that can be sampled from Π via sampling from Z is the entire shaded area in Figure 1 calculated as:

$$|\Pi \cap Z| = \frac{|\Pi \cap Z|}{|\Pi|}|\Pi| = \zeta|\Pi|$$

Now suppose M samples are selected from Z. While M can be controlled subject to resource constraints, the number of infected individuals the tests return, X, is a random variable, which follows a Hypergeometric distribution [4] with total size equal to the size of the sampling universe $|Z|$,

$$|Z| = \frac{|\Pi \cap Z)|}{\pi} = \frac{\zeta}{\pi}|\Pi|,$$

and the size of the sub-population with desired characteristic (infected individuals) equal to $|\Pi \cap Z|$, i.e.

$$\begin{aligned} X &\sim \text{Hyp}(M, |\Pi \cap Z|, |Z|) \\ X &\sim \text{Hyp}(M, \zeta|\Pi|, \pi^{-1}\zeta|\Pi|) \end{aligned} \tag{7}$$

The number of the infected population that can be identified via sampling from the above distribution is bounded by $|\Pi \cap Z| = \zeta|\Pi|$. Recalling the notation used for minimum stabilizing gain, $K_{\min}$, then a necessary condition to stabilize the system will be:

$$K_{\min}|\Pi| < \zeta|\Pi| \iff K_{\min} < \zeta \tag{8}$$

Lastly, when the sizes of the sampling universe Z and the intersection $\Pi \cap Z$ are large, the Hypergeometrical distribution simplifies to

$$\text{Binomial}(M, \frac{\zeta|\Pi|}{\zeta|\Pi|\pi^{-1}}) = \text{Binomial}(M, \pi)$$

which in turn may be approximated by a Normal distribution with mean $M\pi$ and variance $M\pi(1-\pi)$.

Consider a confidence level $1-\phi$, where $\phi \geq 0$ is a small number (e.g. 0.05). The distribution described in (7), returns a number of samples from Π that exceeds $K_{\min}\Pi$, and hence stabilizing, with probability $\geq 1-\phi$, only if $M \geq M_{\min}(\phi, K_{\min})$ where:

$$M_{\min}(\phi, K\text{min}) = \max\{K|\Pr(X \geq K_{\min}|\Pi|) \geq 1-\phi\}$$

In closing this section, it is important to note that we have not explicitly included the testing error in our calculations, although for a more complete analysis, they may be rolled into the overall quantities for test specificity and sensitivity.

5 Numerical and Simulation Analysis

In order to compute the required number of tests, we need values for the parameters involved in the derived distributions and verify the condition stated in inequality (8). Unfortunately, these parameters are not all directly observable. One such parameter is ζ. Although, ζ disappears from the distribution after simplification to Binomial, information about its value is necessary to assess among others whether a minimum number of infected cases can ever be drawn for the implementation of the control strategy. The same is true for the number of infected population. There are also additional constraints that further complicate the computations. For example, in an ideal world, the percentage of the infected population that is tested would only be a function of qualifying symptoms or indicators. In the real world, however, that number could be severely limited by availability, access, or logistical and economical considerations as was the case with the testing situation at the early stages of the outbreak in U.S. Since we do not have access to detailed testing data, the numerical results presented below are based purely *on assumptions* and should be taken and interpreted as such. It may be possible to develop a methodology that uses the system dynamics along with observations to estimate the unobserved states (e.g. the size of the infected population) and other parameters involved. Such analysis is not discussed in this paper.

Recall from the previous section that the sampling universe is represented by Z. The sampling universe has two parts. One that includes individuals from the infected population $\Pi \cap Z$ and the other that does not $Z_0 = Z - (\Pi \cap Z)$. In the numerical analysis below, we set $|Z_0| = 300$K. This is the size of the uninfected population that the testing criteria (symptoms, social contact, etc.) select as eligible for testing on any given day. This figure is perhaps a conservative estimate, since it amounts to 109 million people over one year (one-third of the U.S. population), or slightly over the size of the U.S. population that develops flu or allergies in a given year.

The subset out of Z that includes the infected individuals is given by $|\Pi \cap Z| = \zeta|\Pi|$. As discussed earlier ζ is a property of the test selection criteria: the higher the sensitivity of the selection, the higher its value. The values of Z and ζ are not independent. If the universe is selected too small, then its intersection with Π and therefore the value of ζ will be too small as well. On the other hand, if the test selection criteria are smartly designed, the size of Z may be reduced with little effect on ζ. Knowing the relationship between Z and ζ for various selection criteria will allow design of testing strategies that achieve the desired outcome without exorbitant testing.

If the daily tests were performed at their minimum stabilizing levels, the number of true infections times the minimum stabilizing gain, $K_{\min} = 0.09$, would amount to the observed daily confirmed cases (DCC) we receive from websites such as COVID-19 Tracking, or conversely the total number of infections would amount to $K_{\min}^{-1}$DCC. Had we known the value of ζ and the total size of the sampling universe, we would have been able to determine if the tests are run at their minimum

stabilizing level. The plot in Figure 2 provides an easy to use tool to find the minimum required number of tests for any given DCC under the assumption that $|Z - (\Pi \cap Z)| = 300\text{K}$ (there are 300K individuals on any given day that qualify for test for any reason despite not being infected).

As of March 26, 2020, U.S. registered an additional 17K confirmed cases of COVID-19 infections while running at 101K tests. The good news is that for any ζ value larger than 0.6, this number of tests would be able to stabilize the system. Given the assumption on $|Z_0|$, we may derive an approximation for ζ. Assuming that π may be estimated by the ratio of positively tested to total tested individuals, and using the numbers reported on the COVID-19 Tracking Project for March 26, we arrive at an estimate of π equal to 0.16. On the other hand, 17K DCC would only be stabilizing if the true number of infections were less than or equal to $0.09^{-1} \times 17\text{K} = 189\text{K}$. Using

$$\pi\,(300\text{K} + |\Pi \cap Z|) = |\Pi \cap Z|$$

allows us to estimate $|\Pi \cap Z| \approx 57\text{K}$. Dividing this value by 189K provides the estimate $\zeta \approx 0.3$. But with that estimate for ζ, we should be running at minimum level above 150K tests a day to stabilize the system. Compare this to data from March 24, 2020 when DCC and the number of tests stood at 9806 and 65K, respectively. The plot informs that even in the most optimistic scenario ($\zeta = 0.8$), those number of tests were not adequate to suppress the spread of the virus. A confidence level of 99% ($\phi = 0.01$) has been used to generate the plot.

It may seem counter-intuitive at first that at the very low end of the range, the minimum number of stabilizing tests shoots up to include the entire sampling universe of eligible individuals ($\geq 300\text{K}$). Remember, however, that as the number of infections dwindle, the only way that the very few remaining infected individuals may be identified will be by sampling the entire universe. This is akin to trying to "find the needle in the haystack" as the expression goes. In the early stages of the outbreak when for example the virus is limited to certain geographic areas, the universe may be selected much smaller in size without much impact on the value of ζ. When the virus is broadly spread, on the other hand, the testing selection criteria may have to be broadened else ζ will be too small. In reality, it may not be possible to test vast numbers, say a third of the U.S. population, on a sustained long term basis, say over a year or more. Therefore it is realistic to assume that the level of testing will have to be capped. This implies that even after the numbers of the infections begin to go down, the required number of stabilizing tests will at some point cross the cap. When that happens, we anticipate the total number of infections to stabilize around an equilibrium dictated by the cap. An indicator of having reached such state is that the new confirmed cases (as reported by various websites) will begin to grow only *linearly* rather than *exponentially*. Reducing the growth to linear rather than complete eradication will likely be the best we could accomplish until a vaccine becomes widely available, in light of practical, economical and resource constraints.

The discussions above reveal multiple opportunities for future investigation. One is finding smart ways of reducing the size of the testing universe without impacting the ζ, or the question of optimizing the *testing-mix*, which will be critical, and an area where machine learning and artificial intelligence (AI) technologies would be instrumental. Another important direction is development of estimation strategies using the SEIR dynamic model where the *measurement equation* used for state and parameter estimation will be derived using the analysis presented in the previous section.

Lastly, let us analyze, via simulation studies, the expected outcomes under four scenarios: do nothing, social distancing alone, test-and-isolate strategy without social distancing, and test-and-isolate with social distancing.

The first two scenarios, i.e. do-nothing and introduction of social distancing measures alone, are tantamount to open-loop simulations with higher and lower ρ_0 parameters. The plots in Fig. 3 depict the expected outcome along with assumptions behind them. On the top left, Figure 3.a

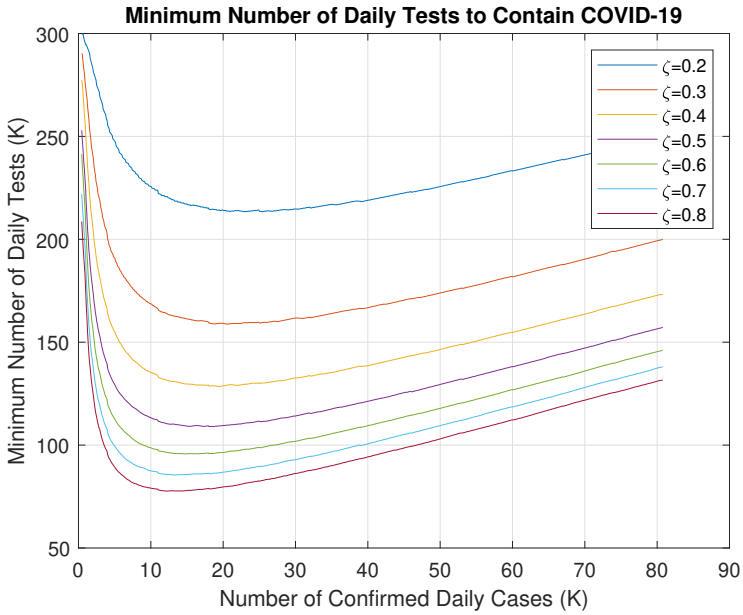

Fig. 2. Stabilizing Tests versus Number Confirmed Daily Infections

depicts the spread of the disease in the United States without any significant restrictions. The results shown in Fig. 3.b, on the other hand, imposes restrictions such as adherence to personal hygiene and social distancing guidelines effectively lowering ρ_0 from 2.2 to 1.28. When left unrestricted, the disease completes its course in about 50 weeks after infecting 2/3 of the country, resulting in millions of fatalities considering a death rate even as low as 1%. Restrictions such as social distancing could mitigate the situation somewhat by stretching the duration of the disease and reducing the peak number of infections by half. However, it does not contain the spread of the disease. The parameters used in the simulations are as follows: $\alpha^{-1} = 5$ days, $\epsilon = 2.5\%$/year, $\mu = 1.25\%$/week, and $\beta = \rho_0(\alpha+\epsilon)(\mu+\gamma)/\alpha = 0.1245$. The open-loop simulation results presented here are consistent with the publicly accessible COVID-19 simulators (e.g., [5])

To avoid such catastrophic outcomes other measures such as test-and-isolate will be unavoidable. Moreover, the economic cost of extreme social distancing for over a year will be incalculable. A closed-loop strategy such as test-and-isolate will allow the country to return to normalcy by placing the restrictions primarily on the infected individuals, whose numbers will decline over time. The closed-loop simulation results in Fig. 3.b and 3.c which use a Hypergeometric random generator to isolate the infected individuals based on caps on daily testings of 80K and 110K, respectively. The daily testing of 80K is not enough to contain the virus and will result in a similar number of fatalities and total infection to that of the open-loop case. Daily testing of 110K contains the spread of COVID-19 resulting is a steady-state infection of roughly 90K, which is fairly close to the confirmed U.S. cases as of March 27, 2020. These results confirm that by testing and isolation according to the limits established in Fig. 2, the spread of COVID-19 can be completely contained.

6 Discussion and Concluding Remarks

In this paper, we derived an output feedback control law for stabilization of COVID-19 transmission dynamic based on a slightly modified SEIR model, which accounts for a quarantine or isolation

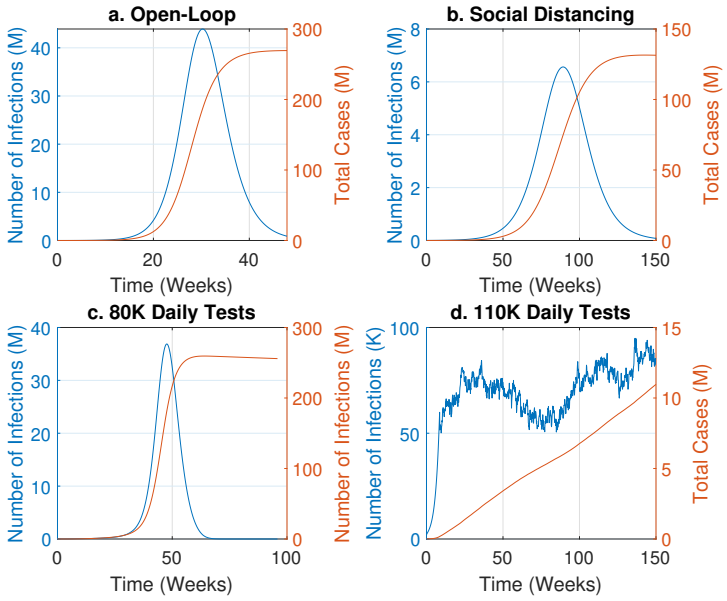

Fig. 3. Open and Closed-Loop Simulation Results

state. The feedback control operates on the basis of a test-and-isolate strategy: individuals are selected based on certain criteria and isolated if tested positive. Using COVID-19 epidemiological parameters, we demonstrated that daily identification and isolation of approximately 10% of the infected population will theoretically suppress the growth of the disease and ultimately stabilize it. Identifying the 10% of the population through testing, requires sampling at much greater numbers. In the early stages of outbreak, the challenge will be in the identification of the required number of infected out of a potentially very large population, or finding the needle in the haystack as the expression goes. As the disease expands, the challenge will be of a different kind as finding the infected individuals will become easier but managing them in and out of quarantine will be an enormous undertaking. If the disease expands beyond a threshold, the sampling universe will be so vast that we will be forced to either let the disease run its course or introduce draconian nation-wide measure such as complete lock-downs to bring it under control. Which way we go will depend on how thoughtfully and smartly we design the testing strategy. Should we be able to bring the disease under control, it will be very unlikely to completely eradicate it without a vaccine. In the most likely scenario, the testing has to continue.

From an applied research point of view, there are many questions that still need to be answered. One key application of the model will involve design and development of smart testing methodologies, both in terms of size and mix, down to the state and even county and city level. The model we presented was intentionally kept simple and high level to allow for a core understanding of the presented concepts. The model among others assumes that the population is overall closed. This may be a reasonable approximation for the current conditions at the national level in U.S. as travel in and out of the country has slowed, if not completely ceased. However, as we go more granular and as the traveling returns, not only the number of compartments in the model should be expanded to account for the higher granularity, but the movement between the compartments of the SEIR model must also be accounted for. Expanding the model along the geographical dimension and development

of testing strategies at the local level using the expanded model will present an important next step. The model can also be used to help the health systems at the local level predict and plan.

References

[1] W. O. Kermack and A. G. McKendrick, "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*, vol. 115, no. 772, pp. 700–721, 1927.

[2] M. Cascella, M. Rajnik, A. Cuomo, S. C. Dulebohn, and R. Di Napoli, "Features, evaluation and treatment coronavirus (covid-19)," in *StatPearls [Internet]*. StatPearls Publishing, 2020.

[3] S. Zhang, M. Diao, W. Yu, L. Pei, Z. Lin, and D. Chen, "Estimation of the reproductive number of novel coronavirus (covid-19) and the probable outbreak size on the diamond princess cruise ship: A data-driven analysis," *International Journal of Infectious Diseases*, 2020.

[4] "Hypergeometric distribution," https://mathworld.wolfram.com/HypergeometricDistribution.html, accessed: 2020-03-24.

[5] "Covidsim: A covid-19 simulator," http://covidsim.eu/, accessed: 2020-03-27.

A Proof of Theorem 1

Proof. The positive initial conditions guarantees that the ratio $r = S/N$ is continuously differentiable and lies between 0 and 1. To prove the exponential convergence of E and I consider the subsystem

$$\begin{bmatrix} \dot{E} \\ \dot{I} \end{bmatrix} = \begin{bmatrix} -(\alpha+\epsilon) & \beta r \\ \alpha & -\gamma-\mu-K \end{bmatrix} \begin{bmatrix} E \\ I \end{bmatrix} \tag{9}$$

and define $V(t) = E^2 + I^2$ as a Lyapunov function candidate for (9). The derivative of V along (9) is seen to be $dV/dt = 2E\dot{E} + 2I\dot{I} = -z^T\mathbf{P}z$ where $z = [E\ I]^T$ and

$$\mathbf{P} = \begin{bmatrix} 2(\alpha+\epsilon) & -(\alpha+\beta r) \\ -(\alpha+\beta r) & 2(\gamma+\mu+K) \end{bmatrix}$$

Matrix $\mathbf{P}$ is positive definite if and only if its determinant, i.e.,

$$4(\alpha+\epsilon)(K+\gamma+\mu) - (\alpha+\beta r)^2 > 0$$

Now choosing $K(t) \geq K_{\min}(t) + \delta$ as specified implies that $\mathbf{P} \geq \sigma I$ for some $\sigma \geq 2\min(\delta, \alpha+\epsilon) > 0$. Thus $\dot{V} \leq -\sigma V \Rightarrow \dot{V} + \sigma V \leq 0$. Multiplying both sides of the preceding inequality by $e^{\sigma t} > 0$, we get $\frac{d}{dt}(V(t)e^{\sigma t}) \leq 0$, which upon integration from 0 to a t implies

$$\int_0^t \frac{d}{d\tau}(V(\tau)e^{\sigma\tau})d\tau = V(t)e^{\sigma t} - V(0) \leq 0 \Rightarrow V(t) \leq e^{-\sigma t}V(0)$$

proving that $V(t)$ and consequently E and I converge to zero exponentially. The exponential convergence of Q follows from $\dot{Q} = KI - (\gamma+\mu)Q$ and exponential convergence of I.