Covid-19: an analysis of a modified SEIR model and a comparison of different intervention strategies.

Arghya Das, Abhishek Dhar, Srashti Goyal, Anupam Kundu
International Center for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore-560089, India

We present a careful analysis of an extended SEIR model, that takes into account the presence of aymptomatic and presymtomatic populations, and explore the effects of different intervention strategies such as (a) social distancing (SD) and (b) testing-quarantining (TQ). These two strategies try to reduce the disease reproductive number ($R_0 > 1$) to a target value [$R_0(\text{target}) < 1$], but in distinct ways, which we implement in the SEIR model equations. We find that for the same target reproductive number $R_0(\text{target})$, testing-quarantining appears to be more efficient in controlling the pandemic than lockdowns (which only implement SD). However, for TQ to be effective, the number of tests/per day in a given region (say, a city) has to be scaled with the number of new cases detected and has to be based on contact tracing of the new detected cases. A combination of SD and TQ might be the most effective and practical strategy. As has been pointed out in other studies, we too find that removal of interventions before the disease has been eliminated would result in a second wave of infections. Weak extended intervention strategies (that reduce R_0 but not to a value < 1) can reduce the peak values of the infected number (related to required hospital beds) and the asymptotic affected population and we provide simple expressions for these in terms of the disease parameters. Looking at real data we find that for many countries, several broad qualitative features are captured surprisingly well by the model. Finally, we propose an accurate way of specifying initial conditions for the numerics (from insufficient data) using the fact that the early time exponential growth is well-described by linearized equations and the dynamics quickly moves along the eigenvector corresponding to the largest eigenvalue. It is then sufficient to have knowledge of any one single dynamical variable describing the motion at some early time. This also suggests a possibile way to collapse data for different countries and we test this.

I. INTRODUCTION

There are two main class of models that have analyzed mathematical models in an attempt to understand and sometimes make predictions about the growth and spread of Covid-19: (i) compartmentalized models and (ii) agent-based models. The former divides the entire population into blocks and then considers the dynamics of how the populations of these blocks evolve, either stochastically or (when numbers are large) through deterministic ordinary differential equations for the mean populations in each block. On the other hand, agent based models follow the evolution of individuals in a population. The so-called SEIR model is an example of the compartmentalized models with four classes of susceptible (S), exposed (E), Infected (I) and Recovered (R) individuals with $S+E+I+R=N$ being the total population. The SEIR model is parameterized by the rates β (infectivity), σ, specifying $E \to I$ transitions and γ specifying $I \to R$ transitions. In terms of the data that is typcially measured and reported, R corresponds to the total number of cases till the present date, while γI would be the number of new cases/ per day. The numner of deaths would be some fraction ($\approx 5\%$) of R while the number of hospital beds reqired at any time would be new cases $\times$ typical days to recovery.

The two main intervention schemes for controlling the pandemic are social distancing (SD) and testing-quarantining (TQ). Lockdowns (LD) impose social distancing and effectively reduce contacts between the susceptible and infected populations, while testing-quarantining means that there is an extra channel to remove people from the infectious population. These two intervention schemes can be implemented in the model, respectively by either introducing a time-dependence for the infectivity parameter β or by introducing a time-dependence for the recovery parameter γ. An important parameter characterizing the disease growth is the reproductive number R_0 — when this has a value > 1, the disease grows exponentially. Typical values reported in the literature for Covid-19 are in the range $R_0 = 2-7$ [1]. Intervention schemes attempt to reduce this to a value < 1. For the basic SEIR model [see App. (B)] one has $R_0 = \beta/\gamma$ and it is clear that we can reduce R_0 by either decreasing β or by increasing γ. The introduction of two time-dependent parameters has been discussed in earlier papers, e.g [2].

In the present work, we analyze intervention strategies in a modified version of the SEIR model which incorporates the fact that asymptomatic or mildly symptomatic individuals [3] are believed to play a significant role in the transmission of Covid-19. In the modified SEIR model, the infected individuals are divided into two classes of asymptomatics (I_a) and presymptomatics (I_p). This or similar type of comparmentalization has been discussed in several recent work [4]. Since we do not attempt to make direct comparisons with data, we do not include separate compartments for the number of hospitalized and dead. For this modified SEIR model we discuss the performance of two different intervention strategies (namely SD and TQ) in the disease dynamics and control.

Our main results can be summarized as follows:

- Two important parameters characterizing the disease are the reproductive number, R_0, and the exponential growth rate μ [related to the "doubling time" given by $\ln(2)/\mu$]. The rate μ is obtained as the largest eigenvalue of the linearized dynamics. We point out some useful general results on the peak value of infections (which is proportional to the number of hospitalizations required) and the number of days to reach this peak value. For both the basic and modified SEIR models we find that these are given by simple general relations:

$$I^{(m)} \approx \frac{\sigma}{\gamma+\sigma}\left(1-\frac{1+\ln R_0}{R_0}\right)N. \tag{1}$$

$$t^{(m)} \approx \frac{\ln N}{\mu}. \tag{2}$$

  The fraction of population, $\bar{x} = R(t\to\infty)/N$, that is eventually affected is given by the solution of the equation

$$1-\bar{x}-e^{-R_0\bar{x}} = 0, \tag{3}$$

  this result being valid for both the simple SEIR and the modified version. These relations are useful — for example they give good estimates for the required affected population before herd immunity is attained and the typical numbers for peak infections and when they happen (see below). We also provide relations for estimating the number of asymptomatic infected and recovered individuals.

- Interventions either through social distancing (SD) or testing-quarantining (TQ) effectively reduce the reproductive number (in different ways) and one can talk of a time-dependent reproductive number $R^{\rm eff}(t)$ with a targeted long time value $R^{\rm eff}(t=\infty) = R_0(\text{target})$. The exponential growth will stop around the time $t^{(\rm int)}$ when $R^{\rm eff}(t)$ crosses the value 1. After this time, the infection numbers would start decaying exponentially. For the case $R_0(\text{target}) \lesssim 1$ we expect a very slow decay.

- We classify intervention strategies by the targeted $R_0(\text{target})$ value. A strong intervention is one where $R_0(\text{target}) < 1$ while a weak intervention is one with $R_0(\text{target}) \gtrsim 1$.

- *Weak interventions*: In this case the results in Eqs. (1-3) can be used by replacing R_0 by $R_0(\text{target})$ in these equations, and one can then get good estimates for the peak infection numbers and total asymptotic infected population. For typical parameters, say $\beta = 0.35, \sigma = 0.333, \gamma = 0.1$, one finds $R_0 = \beta/\gamma = 3.5$, $\mu = 0.144$ which gives a peak infection of 27% of the population after around 110 days (for $N = 10^7$) while eventually about 96% of the population is affected. On the other hand, consider interventions which reduce R_0 by a factor of 3 to a value which is still > 1. If we do this through social distancing (reduce β by a factor of 1/3), then we get $\bar{x} = 27\%, I^{(m)} = 0.8\%, t^{(m)} \approx 3.5$ years, while a testing-quaranting protocol (increase γ by a factor of 3) gives the same $\bar{x} = 27\%$ but $I^{(m)} = 0.5\%$ and $t^{(m)} \approx 1.7$ years.

  We note that while weak interventions can slow down and reduce the impact of the pandemic, they do not lead to development of herd immunity of the population. Herd immunity in the above example would require that 96% of the population be affected.

- For the strong intervention case our main conclusions from numerical studies with various intervention protocols of the modified SEIR model are:

  1. We show that for TQ to be effective, *the number of tests/per day in a given region (say a city) has to be scaled with the number of new cases detected/per day and has to be based on contact tracing of the new detected cases.* It is necessary to increase testing numbers at an early stage when the number of new cases is still small and one has not hit community spreading. For random testing to be effective the required number of tests per day would be impractical.
  2. If the interventions are completely removed before the infected population drops to zero, it is clear that the disease will continue to the same peak and saturation values as one would have got in the absence of any interventions. We only delay the process.
  3. Comparing different intervention strategies that aim to completely end the pandemic, we find that with the same target $R_0 < 1$, *TQ* is most effective while a combination of SD and TQ is more effective in controlling the disease than just SD. Mathematically we understand this by again looking at the maximal eigenvalue μ (now negative and quantifying the rate at which the infection numbers decay) which is larger in magnitude for the case with TQ than with SD, for the same R_0. The expected time for the pandemic to die would be roughly given by

$$t_{\text{end}} \sim \frac{\ln(\text{Peak infection number})}{|\mu(\text{post} - \text{intervention})|}, \tag{4}$$

     and so it is important that intervention schemes are implemented early and as strongly as possible.

  Given the huge economic and social costs of implemeting SD and the difficulties (perhaps technical and economic) in implementing TQ, we believe that a combination of the two would be the most efficient way of controlling the pandemic. A sustained and targeted testing and quarantining strategy (assuming community spreading is still limited), combined with some level of social-distancing has to b implemented to the fullest extent. Indeed it seems that countries which have been most successful at control have used this strategy.

- We do not attempt a detailed comparison of the model predictions with real data since there are too many poorly known parameters and possibly quite inaccurate knowledge of the initial conditions of the variables themselves. We make some overall qualitative observations relating real data to the predictions from SEIR-type models and find that in many cases, several broad qualitative features are remarkably well captured by the model.

- An observation that we make is that, independent of initial conditions, the vector describing all the system variables will quickly point along the direction of the eigenvector corresponding to the largest eigenvalue. Hence (at such longish times) if we know one variable (or a linear combination), then the full vector is completely specified. This implies that different initial conditions (such as different seed infections) will only cause a temporal shift of the observed evolution. This means that if we plot data for different countries, starting from the same initial value of say the confirmed number of cases (normalized by the population), we should see a collapse of the data. We test this idea and find that indeed an approximate collapse of data is obtained for a number of countries.

The paper is structured as follows — in Sec. (II) we present the analysis of the modified SEIR model with and without interventions. We present some useful analytic results and expressions and numerical results on different

intervention protocols. In Sec. (III) we make qualitative comparisons of the predictions of the SEIR model with real data on confirmed number of cases. We summarize our results in Sec. (IV). Technical details and various analytical results are presented in two appendices (Sec. (A) and Sec. (B)).

II. A MODIFIED SEIR MODEL WITH LOCKDOWN AND TESTING INTERVENTIONS

Definition of the modified SEIR model: We consider a population of size N that is divided into eight compartments of

1. S = Susceptible individuals.
2. E = Exposed but not yet contagious individuals.
3. I_a = Asymptomatic, either develop no symptoms or mild symptoms.
4. I_p = Presymptomatic, those who would eventually develop strong symptoms.
5. U_a = Undetected asymptomatic individuals who have recovered.
6. D_a = Asymptomatic individulas who are detected because of directed testing-quaranting, may have mild symptoms, and have been placed under home isolation (few in India).
7. U_p = Presymptomatic individuals who are detected at a late stage after they develop serious symptoms and report to hospotals.
8. D_p = Presymptomatic individuals who are detected because of directed testing-quaranting.

We have the constraint that $N = S + E + I_a + I_p + U_a + D_a + U_p + D_p$. A standard dynamics for the population classes is given by the following set of equations:

$$\frac{dS}{dt} = -\frac{u(\beta_a I_a + \beta_p I_p)}{N} S \tag{5}$$

$$\frac{dE}{dt} = \frac{u(\beta_a I_a + \beta_p I_p)}{N} S - \sigma E \tag{6}$$

$$\frac{dI_a}{dt} = \alpha\sigma E - \gamma_a I_a - r\nu_a I_a \tag{7}$$

$$\frac{dI_p}{dt} = (1-\alpha)\sigma E - \gamma_p I_p - r\nu_p I_p \tag{8}$$

$$\frac{dU_a}{dt} = \gamma_a I_a \tag{9}$$

$$\frac{dD_a}{dt} = r\nu_a I_a \tag{10}$$

$$\frac{dU_p}{dt} = \gamma_p I_p \tag{11}$$

$$\frac{dD_p}{dt} = r\nu_p I_p. \tag{12}$$

The parameters in the above equation correspond to

- α: fraction of asymptomatic carriers.
- β_a: infectivity of asymptomatic carriers.
- β_p: infectivity of presymptomatic carriers.
- σ: transition rate from exposed to infectious.
- γ_a: recovery rate of asymptomatic carriers.
- γ_p: recovery rate of presymptomatics.
- ν_a, ν_p: detection probabilities of asymptomatic carriers and symptomatic carriers. Here we choose $\nu_a = 1/3, \nu_p = 1/2$,

- u: intervention factor due to social distancing (time dependence specified below).
- r: intervention factor due to testing/quarantining (time dependence specified below). This depends on testing-quarantining rates.

With our definitions, the total number of confirmed cases, C, and the number of daily recorded new cases D would be

$$C = D_a + D_p + U_p, \quad D = r\nu_a I_a + (\gamma_a + r\nu_p) I_p \ . \tag{13}$$

Note that we include U_p because these are people who are not detected through directed tests but eventually get detected (after $1/\gamma_p$ days) when they get very sick and go to hospitals. On the other hand the class D_p gets detected earlier and infect less people.

A. Linear analysis of the dynamical equations

Since at early times $S \approx N$ and all the other populations $E, I_a, I_p, D_a, D_p, U_a, U_p << N$, one can perform a linearlization of the above equations and this tells us about the early times growth of the pandemic, in particular the exponential growth rate. For the present let us ignore the time dependence of the SD factor u and the TQ factor r. As shown in App. (A 2), the system has three non-zero eigenvalues given by the roots of the cubic equation:

$$\lambda^3 + (\tilde{\gamma}_a + \tilde{\gamma}_p + \sigma)\lambda^2 + [\tilde{\gamma}_a\tilde{\gamma}_p + \sigma(\tilde{\gamma}_a + \tilde{\gamma}_p)(1 - \tilde{R}_0)]\lambda + \sigma\tilde{\gamma}_a\tilde{\gamma}_p(1 - R_0) = 0, \tag{14}$$

where $\tilde{\beta}_a = u\beta_a, \tilde{\beta}_p = u\beta_p, \tilde{\gamma}_a = \gamma_a + r\nu_a, \tilde{\gamma}_p = \gamma_p + r\nu_p$, $\tilde{R}_0 = \alpha\tilde{\beta}_a/(\tilde{\gamma}_a + \tilde{\gamma}_p) + (1 - \alpha)\tilde{\beta}_p/(\tilde{\gamma}_a + \tilde{\gamma}_p)$ and

$$R_0 = \alpha\frac{\tilde{\beta}_a}{\tilde{\gamma}_a} + (1 - \alpha)\frac{\tilde{\beta}_p}{\tilde{\gamma}_p} = \alpha\frac{u\beta_a}{\gamma_a + r\nu_a} + (1 - \alpha)\frac{u\beta_p}{\gamma_p + r\nu_p} \tag{15}$$

is the expected form for the reproductive number for the disease. Noting the fact that $\tilde{R}_0 < R_0$, it follows that the condition for at least one positive eigenvalue is

$$R_0 > 1. \tag{16}$$

We denote the largest eigenvalue by μ. At early times the number of cases detected would grow as $\sim e^{\mu t}$.

In our numerical study we choose the parameter set $\alpha = 0.67$ while the rates $\beta_a = 0.333, \beta_p = 0.5, \sigma = 1/3, \gamma_a = 1/8, \gamma_p = 1/12$ all in units of day^{-1}. For the specified choice of parameter values (free case with $u = 1.0, r = 0.0$) we get $\mu = 0.158$ which is close to the initial Indian observed value. The corresponding free value of R_0 is 3.7665. Note that μ is not uniquely fixed by R_0 and different choices of parameters can give the same observed μ but different values of R_0.

Role of initial conditions: In Appendix. (A 2)) we explain the fact that all initial conditions (which satisfy the condition $S(0) \approx N$) will quickly move along the direction of the largest eigenvector and so all the trajectories for different initial conditions *are identical upto a time translation.* We illustrate this idea in Fig. (1) where we show plots of $I(t) = I_a(t) + I_p(t)$ and $C(t) = D_a(t) + U_p(t) + D_p(t)$ for 5 different initial conditions. The right panel shows a collapse of all the trajectories by an appropriate time translation of the different trajectories. In this sense the system is not sensitive to initial conditions.

Interventions: The idea of interventions is clear from Eq. (15) — we want to decrease R_0 to a new target value $R_0(\text{target}) < 1$ by either decreasing u (through SD) or increasing r (through TQ). Once we achieve this, how fast the disease dies depends on the magnitude of μ (now negative) and as we demonstrate numerically in Sec. (II D), TQ could be more effective than SD.

B. Final infected population and peak infections

Let us define the asymptotic populations (i.e the populations at very long times) in the different compartments as $\bar{U}_a, \bar{D}_a, \bar{U}_p, \bar{U}_p$, and let $\bar{R}_a = \bar{U}_a + \bar{D}_a$, $\bar{R}_p = \bar{U}_p + \bar{D}_p$, $\bar{R} = \bar{R}_a + \bar{R}_p$. The total population that would eventually be affected by the disease (and either recovered or died) is given by $\bar{R}$ and would have developed immunity. A fraction $\bar{U}_a$ (see below) would be undetected and uncounted.

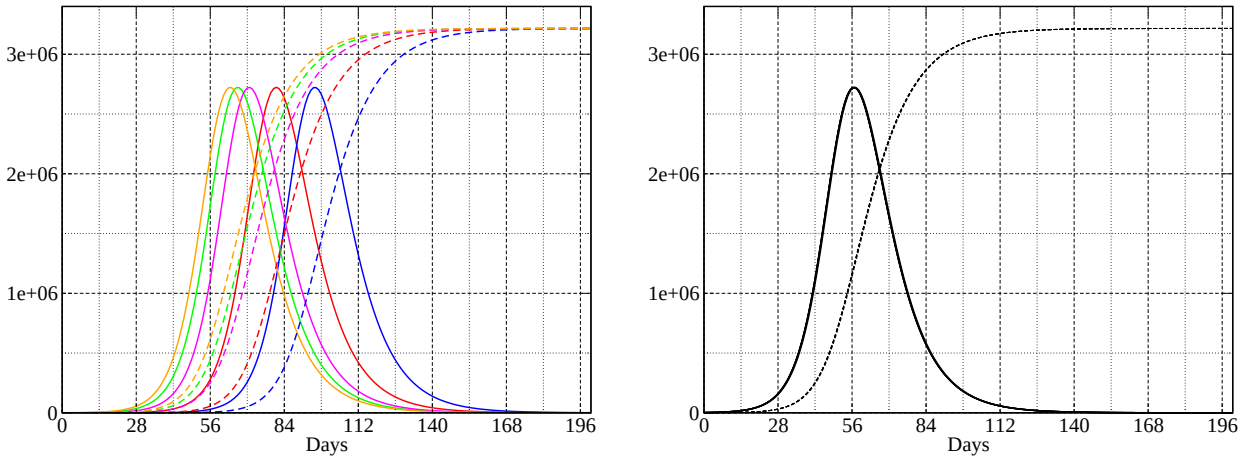

FIG. 1. **Role of initial conditions**: (left) Plots showing $I(t)$ and $C(t)$ for 5 very different initial conditions: (1)$E(0) = 100, I_a(0) = 0, I_p(0) = 0$, (2)$E(0) = 10, I_a(0) = 0, I_p(0) = 0$, (3)$E(0) = 1000, I_a(0) = 0, I_p(0) = 0$, (4) $E(0) = 233, I_a(0) = 100, I_p(0) = 75$, (5) $E(0) = 233, I_a(0) = 1000, I_p(0) = 75$. (right) A collapse of all the curves obtained by translating all the trajectories so that they start with the same value of I.

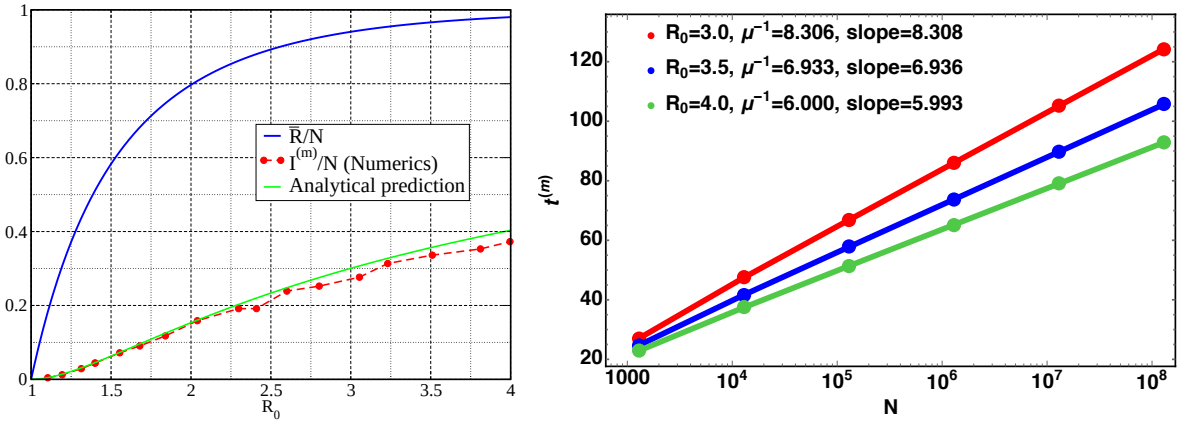

FIG. 2. Plot of the asymptotic total affected population fraction, $\bar{R}/N$, as a function of the reproductive number R_0. We also plot the quantity $(I^{(m)}/N)(\sigma+\gamma)/\sigma$, obtained numerically from many different parameter sets, and compare it with the theoretical predicted curve $1-(1+\ln R_0)/R_0$ (black line). (right) Verification of the $\ln(N)$ dependence of $t_{\max}$ in Eq. (20) for different choices of R_0. The slopes of the straight lines compares well with μ^{-1} as stated in Eq. (20).

Here for the moment let us assume that u and r do not have any time dependence. As shown in App. (A) the asymptotic fraction $\bar{x} = \bar{R}/N$ is simply given by the solution of the equation

$$1 - \bar{x} = e^{-R_0 \bar{x}}, \tag{17}$$

with R_0 being the reproductive number given by Eq. (15) and Eq. (17) has a non-zero solution only when $R_0 > 1$.

The asymptotic population of the individual populations are then given by

$$\begin{aligned} R_a &= \alpha R,\ R_p = (1-\alpha)R \\ U_a &= \frac{\gamma_a}{\gamma_a + r\nu_a} R_a, D_a = \frac{r\nu_a}{\gamma_a + r\nu_a} R_a, \\ U_a &= \frac{\gamma_a}{\gamma_p + r\nu_p} R_p, D_p = \frac{r\nu_p}{\gamma_a + r\nu_p} R_p. \end{aligned} \tag{18}$$

As shown in App. (B 1) for the SEIR model, the peak value of the infection number ($I = I_a + I_p$) can be found

from a heuristic argument and is very accurately given by the formula

$$I^{(m)} = \frac{\sigma}{\gamma+\sigma}\left(1 - \frac{1+\ln R_0}{R_0}\right) N. \tag{19}$$

We find that this also describes accurately the peak value for the modified SEIR dynamics with γ now interpreted as $[\alpha\gamma_a^{-1} + (1-\alpha)\gamma_p^{-1}]^{-1}$. In Fig. (2) we show the dependence of $\bar{x}$ on R_0 (as obtained from a numerical solution of Eq.(17) and provide a numerical verification of the result in Eq. 19. The peak values of the asymptomatic and presymptomatic populations are given by $I_a^{(m)} = \alpha\gamma/\gamma_a$ and $I_p^{(m)} = (1-\alpha)\gamma/\gamma_p$ respectively.

An estimate of the time to reach this peak value can be obtained by noting that we can use the linearized dynamics (see previous section) till the time $I(t)$ reaches its peak $I_{\max}$. Hence we write $I^{(m)} = I(t^{(m)}) = I(0)\ e^{\mu t^{(m)}}$ which provides $t^{(m)} = \frac{\ln[I^{(m)}/I(0)]}{\mu}$. We naturally expect that $I^{(m)}$ is of the order $O(N)$ which implies

$$t^{(m)} \sim \frac{\ln[I^{(m)}/I(0)]}{\mu} \sim \frac{\ln N}{\mu}. \tag{20}$$

A verification of this result, obtained by solving the basic SEIR equations numerically, is provided in Fig. (2).

C. Interventions: Social distancing and Testing-Quarantining

We discuss here the choices of the intervention functions u and r introduced in the dynamical equations in Sec. (II). Note that u is a dimensionless number quantifying the level of social contacts, while r is a rate which, as we will see, is closely related to the testing rate.

Social distancing (SD): We multiply the constant factors $\beta_{a,p}$ by the time dependent function, $u(t)$, the "lock-down" function that incorporates the effect of a social distancing, i.e reducing contacts between people. A reasonable form is one where $u(t)$ has the constant value $(=1)$ before the beginning of any interventions, and then from time t_{on} it changes to a value $0 < u_l < 1$, over a characteristic time scale $\sim t_w$. Thus we take a form

$$\begin{aligned} u(t) &= 1 \quad t < t_{\rm on}, \\ &= u_l + (1-u_l)e^{-(t-t_{\rm on})/t_w}, \quad t > t_{\rm on}. \end{aligned} \tag{21}$$

The number u_l indicates the lowering of social contacts.

Testing-quarantining (TQ): We expect that testing and quarantining will take out individuals from the infectious population and so this is captured by the terms $r\nu_a I_a$ and $r\nu_p I_p$ in the dynamical equations. A reasonable choice for the TQ function is perhaps to take

$$\begin{aligned} r(t) &= 0 \quad t < t'_{\rm on}, \\ &= r_l - r_l e^{-(t-t'_{\rm on})/t_w}, \quad t > t'_{\rm on}. \end{aligned} \tag{22}$$

where we one needs a final rate $r_l > 0$. In general the time at which the TQ begins to be implemented $t'_{\rm on}$ and the time required for it to be effective t'_w could be different from those used for SD.

A useful quantity to characterize the system with interventions is the time-dependent effective reproductive number given by

$$R_0^{\rm eff}(t) = \alpha\frac{u(t)\beta_a}{\gamma_a + r(t)\mu_a} + (1-\alpha)\frac{u(t)\beta_p}{\gamma_p + r(t)\nu_p}. \tag{23}$$

At long times this goes to the targeted reproduction number

$$R_0(\text{target}) = R_0^{\rm eff}(t\to\infty) = \alpha\frac{u_l\beta_a}{\gamma_a + r_l(t)\mu_a} + (1-\alpha)\frac{u_l\beta_p}{\gamma_p + r_l\nu_p}. \tag{24}$$

The time scale for the intervention target to be achieved is given by t_w and t'_w.

Relation of the TQ function $r(t)$ to the number of tests done per day:

Let us suppose that the number of tests per person per day is given by T_r. We show in Fig. (3) the data for the number of tests per 1000 people per day across a set of countries and see that this is around 0.05 for India which means that $T_r = 0.00005$. If tests are done completely randomly, then the number of infected people (assuming that

*

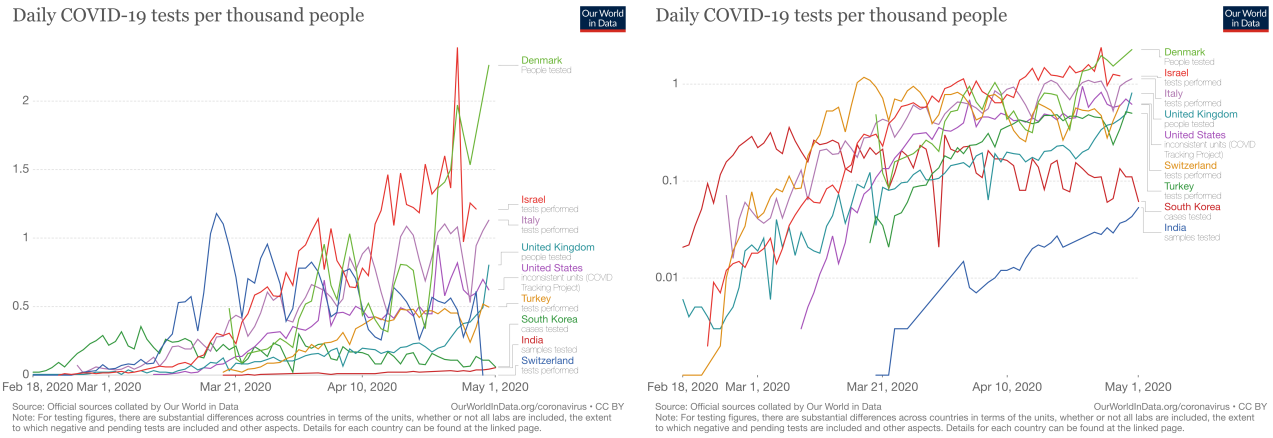

FIG. 3. Data of number of tests per day per thousand in several countries on a (left) linear scale and (right) on a log-scale. Notice in particular the large testing numbers in the early period in countries such as Korea and Switzerland (also Germany not shown in figure) which have been more successful at controlling the disease. In comparison the numbers for India appear to be highly inadequate.

the tests are perfect) would be $T_r \times I$ and so it is clear that we can identify $r(t) = T_r(t)$. It is then clear that this would have no effect on the pandemic control. To have any effect we would need $r \gtrsim \gamma_p \approx 0.1$ which means around 200 tests per 1000 people per day which is clearly not practical.

However a better strategy is to do focused tests on the contacts of all those who have been detected on a given day. In our modified SEIR model the number of detected cases per day is given by $D(t) = r\nu_a I_a + (\gamma_p + r\nu_p)I_p$. Then number of contacts of these individuals would be $AD(t)$ where A is the number of contacts a given infected person made. A good assumption is to say that the infected people are from this pool. Hence, if we make a total of $T = NT_r$ tests per day *on only this set of people*, then the number of detected cases (through contact tracing) would be

$$D_{a,p} = \frac{NT_r}{AD(t)}\nu_{a,p}I_{a,p} = \frac{T}{AD(t)}\nu_{a,p}I_{a,p}. \tag{25}$$

We can then make the identification that a good control strategy is to set

$$r(t) = \frac{T}{AD(t)}, \tag{26}$$

which means that we need $T(t) \sim r_l AD(t)$, with $r_l \approx \gamma_p$, that is *the number of tests/per day has to be proportional to number of new detections/per day*. This is clearly achievable especially when we note that enforcing social distancing in parallel would have reduced the value of A. A noteworthy case that we see in Fig. (3) is the plot for South Korea where we see the large testing rate at early days of the pandemic. Perhaps this explains the quick control of the pandemic in that country.

D. Numerical results: a comparison of different intervention protocols

We work with a population $N = 10^7$ and initial conditions $E(0) = 100$, $I_a(0) = I_p(0) = U_a(0) = D_a(0) = U_p(0) = D_p(0) = 0$ and $S(0) = N - E - I_a - I_p - U_a - D_a - U_p - D_p$. In all cases, we will assume that intervention strategies are switched on when the confirmed number of cases reaches 50 and after that interventions are attained over a time scale $t_w = t'_w = 5$ days.

Parameter set I [$R_0(\text{target}) = 0.667$]: We choose three social distancing and testing-quarantining strengthss as (i) SD: $u_l = 0.177, r_l = 0$, (ii) TQ: $u_l = 1, r_l = 1.2$ and (iii) SD-TQ: $u_l = 0.461, r_l = 0.4$. This choice correspond to changing the free value of $R_0 = 3.766$ to a target value $R_0(\text{target}) = 0.667$, for all the three different strategies. The largest eigenvalue μ changes from the free value $\mu = 0.158$ to the values (i) $\mu = -0.027$, (ii) $\mu = -0.077$ (iii) $\mu = -0.0546$ respectively.

Parameter set II [$R_0(\text{target}) = 0.904$]: We choose three SD and TQ amplitudes as (i) SD: $u_l = 0.24, r_l = 0$, (ii) TQ: $u_l = 1, r_l = 0.809$ and (iii) SD-TQ: $u_l = 0.6, r_l = 0.373$. This choice correspond to changing the free value of

*

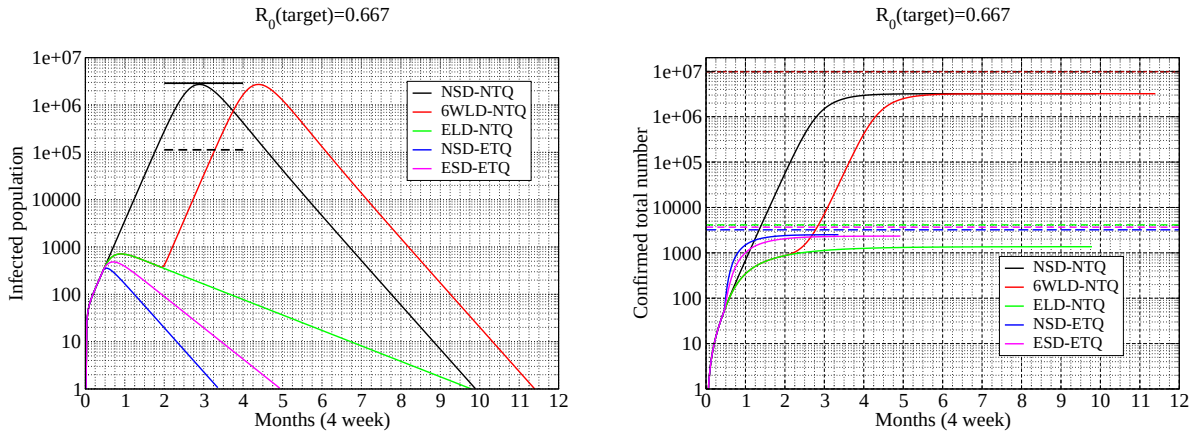

FIG. 4. **Parameter set I**: (left) Total number of infected cases $I = I_a + I_p$ for different intervention strategies. The solid and dashed black lines indicate the peak infected cases ($I^{(m)}$) as given by Eq. (1) and the corresponding value of $I_p^{(m)}$. (right) Total number of confirmed cases $C = U_p + D_a + D_p$. The dashed lines indicate the total affected population $R = C + U_a$ at the end of one year, for the different strategies. In the absence of interventions this is close to 96% and is given by Eq. (3). The total population was taken as $N = 10^7$.

*

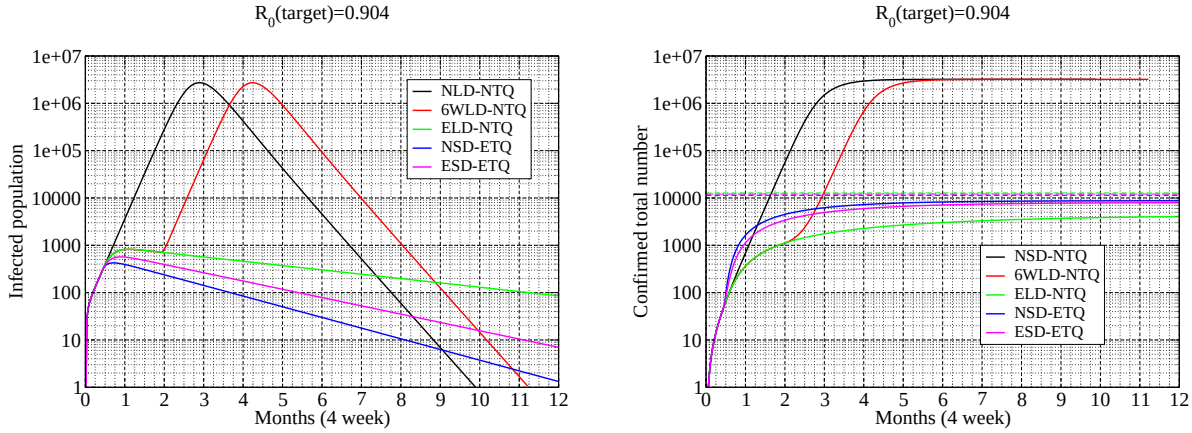

FIG. 5. **Parameter set II**: (left) Total number of infected cases $I = I_a + I_p$ for different intervention strategies. (right) Total number of confirmed cases $C = U_p + D_a + D_p$. The dashed lines indicate the total affected population $R = C + U_a$ at the end of one year, for the different strategies. Total population was taken as $N = 10^7$.

$R_0 = 3.766$ to a fixed target value $R_0(\text{target}) = 0.904$ for all the three different strategies. The largest eigenvalue μ changes from the free value $\mu = 0.158$ to the values (i) $\mu = -0.00744$, (ii) $\mu = -0.0184$ (iii) $\mu = -0.0142$ respectively.

Parameter set III [$R_0(\text{target}) = 1.205$]: We choose three SD and TQ amplitudes as (i) SD: $u_l = 0.32, r_l = 0$, (ii) TQ: $u_l = 1, r_l = 0.536$ and (iii) SD-TQ: $u_l = 0.634, r_l = 0.24$. This choice corresponds to changing the free value of $R_0 = 3.766$ to a fixed target value $R_0(\text{target}) = 1.205$ for all the three different strategies. The largest eigenvalue μ remains positive and changes from the free value $\mu = 0.158$ to the values (i) $\mu = 0.0152$, (ii) $\mu = 0.032$ (iii) $\mu = 0.0248$ respectively.

In Figs. (4,5,6) we show the number of infected and confirmed cases for the above parameter sets and five different intervention schemes:

(1) NSD-NTQ: No social distancing and no testing-quarantining.

(2) 6WLD-NTQ: Six weeks lockdown (strong value of SD parameter) and no testing-quarantining.

(3) ELD-NTQ: Extended lockdown and no testing-quarantining.

*

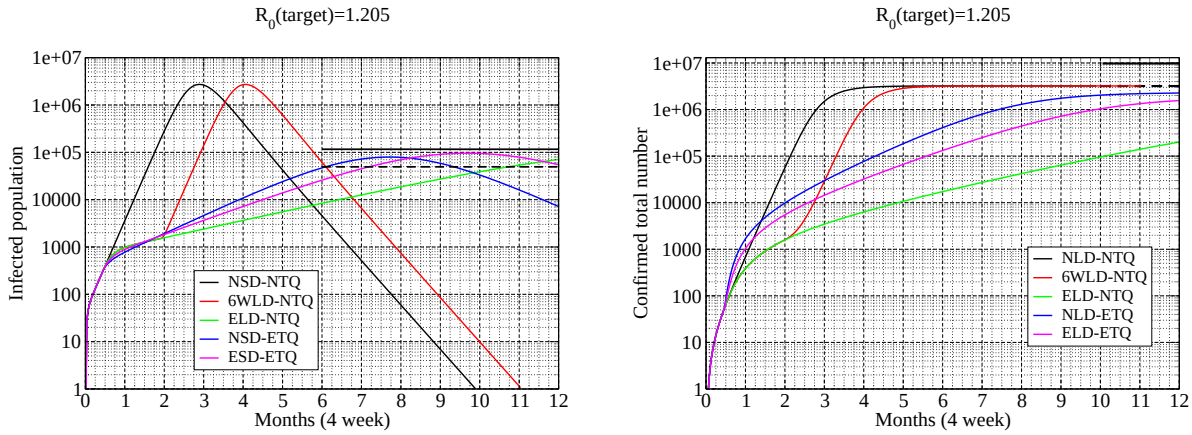

FIG. 6. **Parameter set III**: (left) Total number of infected cases $I = I_a + I_p$ for different intervention strategies. (right) Total number of confirmed cases $C = U_p + D_a + D_p$. The dashed lines indicate the total affected population $R = C + U_a$ at the end of one year, for the different strategies. Total population was taken as $N = 10^7$.

(4) NSD-ETQ: No social distancing and extended testing-quarantining.
(5) ESD-ETQ: Extended social distancing and extended testing-quarantining.

Main observations: A six weeks (or eight week) lockdown is insufficient to end the pandemic and will lead to a second wave. If the interventions are carried on indefinitely, we find that even though the SD and TQ strategies have the same R_0 values, the TQ strategy is more effective in controlling the pandemic.

- **Strong interventions** - parameter sets **(I),(II)**:

  1. In these cases, the pandemic is controlled and only affects a very small fraction of the population (less than 1%). We can understand all features of the dynamics from the linear theory. In Fig. (4) intervention is switched on after ≈ two weeks and the peak in infections shows roughly after a period of $t_w = 5$ days. Thereafter however, the decay in the number of infections occurs slowly, depending on how successful one has been in reducing R_0. In the next section we show data of several countries where the slow decay is clearly onserved.

  2. For the case of the strongest intervention **(I)**, we see that ELD-NTQ ends the pandemic in about 10 months while NSD-ETQ would take around 3.5 months. This can be understood from the fact that the corresponding μ values (post-intervention) are given by $\mu = -0.077$ and $\mu = -0.027$ respectively, i.e, they differ by a factor of about 3. With a mixed strategy where one allows almost three times more social contacts than for LD case and that requires 3 times less testing, we see that the disease is controlled in about 5 months.

  3. For the case of the less stronger intervention **(II)**, we see that NSD-ETQ ends the pandemic in about a years time while the other strategies would take much longer. The post-intervention μ values indicate that ELD-NTQ would take about 28 months while the ESD-ETQ combination would take about 15 months.

- **Weak intervention** - parameter set **(III)**:

  1. In this case, a finite fraction of the population is affected. The intervention succeeds in reducing the fraction that is eventually affected and the peak number of infections and in delaying the date at which the peak occurs. These modified values can be obtained from the expressions in Eqs. (1,2,3), using the post-intervention values of R_0 and μ.

  2. We find that the peak infection numbers are smallest for the case with ELD-NTQ and occur at a later stage. Again these results can be understood mathematically from the expressions in Eq. (1) and Eq. (2) using the post-intervention values of γ and μ.

*

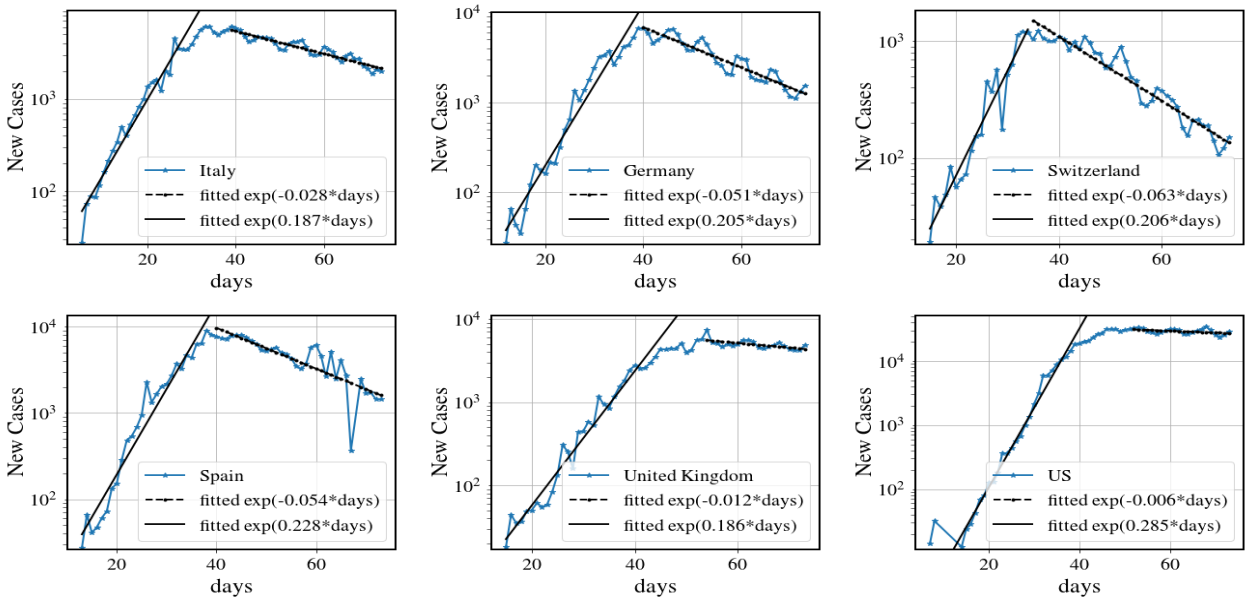

FIG. 7. Number of new cases per day for six different countries. We note that the data exhibits the same broad features that we see for the model predictions in Fig. (4,5). In particular we see the fast exponential growth and slow exponential decrease in new cases (following interventions). The two countries UK and US show a very slow decay rate, indicating that either SD or TQ measures have not been sufficiently implemented.

III. SOME COMMENTS ON COMPARISON WITH REAL DATA

We do not attempt to make a detailed comparison of the real data of the Covid-19 pandemic with the numerical solution of the SEIR equations since there are too many uncertain parameters and it is not clear what this data fitting exercise would teach us. However we try to see if some overall general features can be seen.

In Figs. (7) we give some examples of data for number of new cases for six countries where we see that some of the qualitative features of Figs. (4,5). In particular we see the fast exponential growth phase and then a much slower decay phase. There are several countries which do not show this clear correspondence with the SEIR model predictions (see Fig. (8) and other examples in [6]).

One issue is that different countries start with different initial conditions (for example the seed exposed population could be very different between countries). As discussed in Sec. (A 2), as long as the number of confirmed cases is much smaller than the population size, a description in terms of the linearized dynamics is accurate. This would predict an initial exponential growth and then as intervention schemes begin to operate, the reproductive number and the corresponding growth exponent would decrease till eventually one is able to achieve $R_0 < 1$ and correspondingly $\mu < 0$. In Fig. (8) we show data for the reported number of new cases in 12 different countries and approximately see these features. Most countries show a slow decay phase. A few Asian countries (India, Pakistan, Indonesia) have not yet entered the decaying phase.

The linearized SEIR dynamics also predicts that (see Sec. (A 2)), if one uses similar parameters and intervention parameters, then all countries should follow the same trajectory provided they start with the same value for the normalized fraction of confirmed new cases (D_0/N). In the right panel of Fig. (8) we plot the data with this normalization and initial condition and see a rough collapse for several countries. We notice in particular that three of the Asian countries (India, Pakistan, Indonesia) follow a distinctly different trajectory and similarly for Korea — this could indicate either that the disease parameters are different or that the intervention strategies have been different, or the reporting of cases is inaccurate. In Fig. (9) we see a similar approximate collapse of the data for the total number of confirmed cases.

*

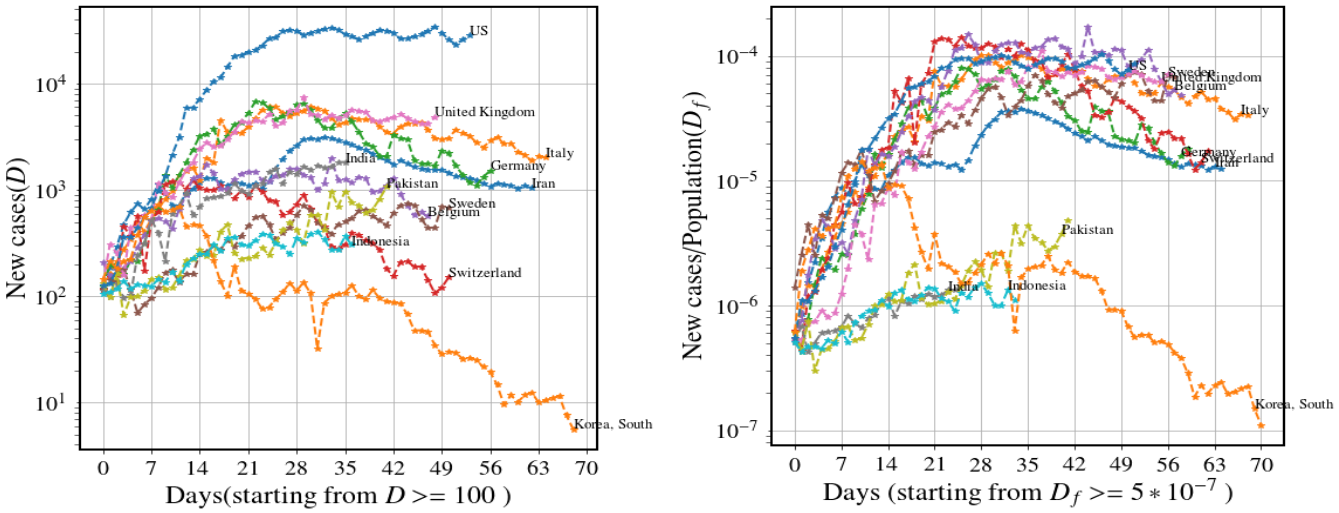

FIG. 8. (left) Number of new cases per day for different countries. (right) Number of new cases normalized by the total population, with the time axis shifted so that every country starts with the same normalized value.

*

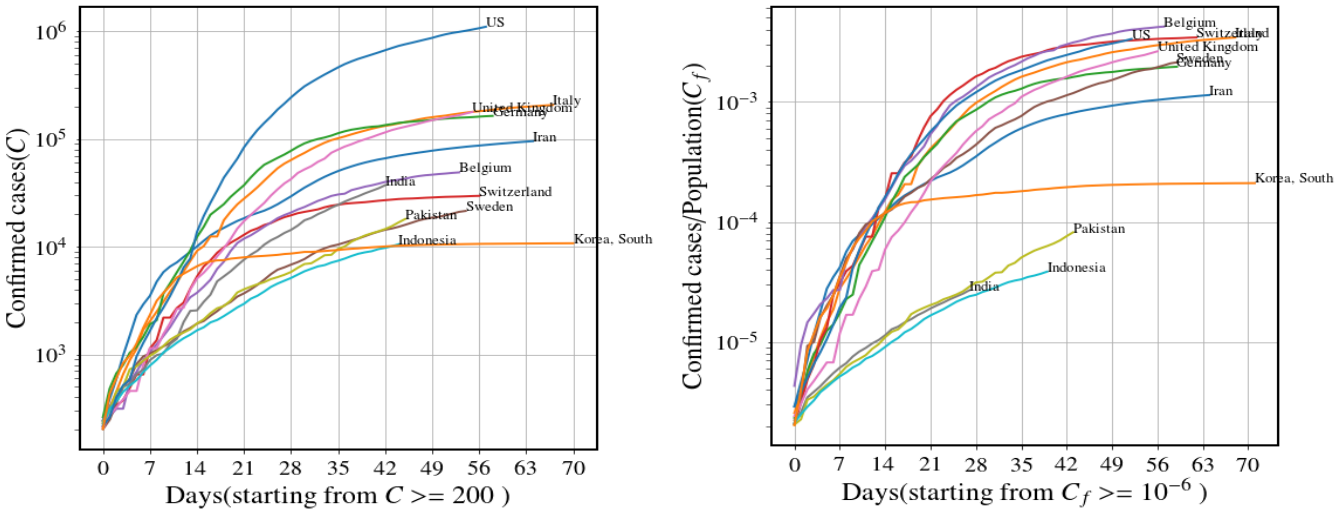

FIG. 9. (left) Total number of confirmed cases for different countries. (right) Total number of cases normalized by the total population, with the time axis shifted so that every country starts with the same normalized value.

IV. CONCLUSIONS

A modified version of the SEIR model, incorporating asymptomatic individuals, was used for analyzing the effectiveness of different intervention protocols in controlling the growth of the Covid-19 pandemic. Non-clinical interventions can be either through social distancing or through testing-quarantining. Our results indicate that a combination of both, implemented over an extended period may be the most effective and practical strategy. We point out that short-term lock-downs cannot stop a recurrence of the pandemic if interventions are completely relaxed and developing herd immunity is not a practical solution either since this would affect a very large fraction of the population.

We have provided numerical examples to illustrate the basic ideas and in addition have stated a number of analytical results which can be useful in making empirical estimates of various important quantities that provide information on the disease progression. Looking at real data for new Covid-19 cases in several countries, we find that the SEIR model captures some important qualitative features and hence could provide some guidance for policy-makers.

ACKNOWLEDGMENTS

We thank Jitendra Kethepalli and Kanaya Malakar for very helpful discussions and Ranjini Bandyopadhyay for a careful reading of the draft.

[1] Ying Liu, Albert A Gayle, Annelies Wilder-Smith, Joacim Rocklöv, *The reproductive number of COVID-19 is higher compared to SARS coronavirus*, Journal of Travel Medicine, Volume 27, Issue 2, March 2020, `https://doi.org/10.1093/jtm/taaa021`.

[2] B. Tang, N. Luigi Bragazzi, Qian Li, S. Tang, Y. Xiao, Jianhong Wu, *An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov)*, Infectious Disease Modelling **5**, 248 (2020).

[3] K. Y. Leung, P. Trapman, T. Britton, *Who is the infector? Epidemic models with symptomatic and asymptomatic cases*, Mathematical Biosciences, **301**, 190 (2018).

[4] M. Gatto, E. Bertuzzo, L. Mari, S. Miccoli, L. Carraro, R. Casagrandi, A. Rinaldo, *Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures*, Proceedings of the National Academy of Sciences, 202004978. DOI:10.1073/pnas.2004978117.

[5] F. Ball and D. Clancy, *The final outcome of an epidemic model with several different types of infective in a large population*, J. Appl. Prob. **32**, 579 (1995).

[6] Covid-19 data analysis: `https://www.icts.res.in/COVID-data-commentary`

Appendix A: Modified SEIR model

1. Asymptotic total affected population

Let us consider a more general form of the SEIR equations with n compartments for the infectious population with $I_1, I_2, \ldots, I_n$, n compartments for the recovered population with $R_1, R_2, \ldots, R_n$ and the other 2 compartments of S, E, R with the following dynamics

$$\frac{dS}{dt} = -\sum_{i=1}^{n} \beta_i \frac{I_i S}{N}, \tag{A1}$$

$$\frac{dE}{dt} = \sum_{i=1}^{n} \beta_i \frac{I_i S}{N} - \sigma E, \tag{A2}$$

$$\frac{dI_i}{dt} = \sigma_i E - \gamma_i I_i, \quad i = 1, 2, \ldots, n, \tag{A3}$$

$$\frac{dR_i}{dt} = \gamma_i I_i, \quad i = 1, 2, \ldots, n, \tag{A4}$$

where $\sigma = \sum_{i=1}^{n} \sigma_i$.

Let us assume that $R_i(0) = 0$ for all i, and $S(0) \approx N$. Then solving Eq. (A1)), we get

$$\bar{S} = N e^{-\sum_{i=1}^{n} \beta_i \int_0^\infty dt I_i(t)/N}. \tag{A5}$$

Multiplying Eqs. (A4) by β_i/γ_i, summing over i and integating time from 0 to ∞, we get $\sum_{i=1}^{n} \beta_i \int_0^\infty dt I_i(t) = (\beta_i/\gamma_i)\bar{R}_i$. Plugging this into the previous equation then gives

$$\bar{S} = N e^{-\sum_{i=1}^{n} \frac{\beta_i}{\gamma_i} \bar{R}_i/N}. \tag{A6}$$

Next we note that $(d/dt)(I_i + R_i) = \sigma_i E$. Hence for the initial condition $I_i = R_i = 0$ we find that the ratio $[I_i(t) + R_i(t)]/[I_j(t) + R_j(t)] = \sigma_i/\sigma_j$ at all times. Since at large times $I_i \to 0$, this means that the asymptic values of R_is are given by

$$\bar{R}_i = \frac{\sigma_i}{\sigma} \bar{R}. \tag{A7}$$

Using this in Eq. (A6), noting that $\bar{S}+\bar{R}=N$ and defining $x=\bar{R}/N$, we then get the following simple equation that determines the asymptotic total affected population:

$$1-x=e^{-R_0 x}, \tag{A8}$$

$$\text{where}\quad R_0=\sum_{i=1}^{n}\frac{\beta_i\sigma_i}{\gamma_i\sigma} \tag{A9}$$

is the reproductive number.

Note that replacing S by $e^{-\sum_{i=1}^{n}\frac{\beta_i}{\gamma_i}\bar{R}_i/N}$ and E by $(N-S-\sum_i I_i-\sum_i R_i)$, we get the following equations for $x_i=R_i/N$ and $v_i=\gamma_i I_i/N$:

$$\frac{dx_i}{dt}=v_i \tag{A10}$$

$$\frac{dv_i}{dt}=-(\gamma_i+\sigma_i)v_i+\sigma_i\gamma_i(1-x_i-e^{\sum_i R_0^{(i)}x_i}), \tag{A11}$$

with $R_0^{(i)}=\beta_i/\gamma_i$. We see that in this case the non-dissipative force is non-gradient and cannot be expressed in terms of a potential.

2. Linear analysis of modified SEIR model

We now again focus on the special case with the $n=8$ variable dynamics described by Eqs. (5-12). Let us denote the variables by $x_1=S-N, x_2=E, x_3=I_a, x_4=I_p, x_5=U_a, x_6=D_a, x_7=U_p, x_8=D_p$. At early times when $x_i << N$, the dynamics is captured by linear equations

$$\frac{dx}{dt}=Mx, \tag{A12}$$

$$\text{with}\quad M=\begin{pmatrix} 0 & 0 & -\tilde{\beta}_a & -\tilde{\beta}_p & 0 & 0 & 0 & 0\\ 0 & -\sigma & \tilde{\beta}_a & \tilde{\beta}_p & 0 & 0 & 0 & 0\\ 0 & \alpha\sigma & -\tilde{\gamma}_a & 0 & 0 & 0 & 0 & 0\\ 0 & (1-\alpha)\sigma & 0 & -\tilde{\gamma}_p & 0 & 0 & 0 & 0\\ 0 & 0 & \gamma_a & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & r\nu_a & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & \gamma_p & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & r\nu_p & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{A13}$$

where $\tilde{\beta}_a=u\beta_a,\ \tilde{\beta}_p=u\beta_p, \tilde{\gamma}_a=\gamma_a+r\nu_a, \tilde{\gamma}_p=\gamma_p+r\nu_p$. This has 5 zero eigenvalues while the remaining 3 ones are given by the roots of the cubic equation for λ:

$$\lambda^3+(\tilde{\gamma}_a+\tilde{\gamma}_p+\sigma)\lambda^2+[\tilde{\gamma}_a\tilde{\gamma}_p+\tilde{\gamma}_a\sigma+\tilde{\gamma}_p\sigma-\alpha\tilde{\beta}_a\sigma-(1-\alpha)\tilde{\beta}_p\sigma]\lambda+\sigma\left[\tilde{\gamma}_a\tilde{\gamma}_p-(1-\alpha)\tilde{\beta}_p\tilde{\gamma}_a-\alpha\tilde{\beta}_a\tilde{\gamma}_p\right]=0. \tag{A14}$$

This can be written in the form

$$\lambda^3+(\tilde{\gamma}_a+\tilde{\gamma}_p+\sigma)\lambda^2+[\tilde{\gamma}_a\tilde{\gamma}_p+\sigma(\tilde{\gamma}_a+\tilde{\gamma}_p)(1-\tilde{R}_0)]\lambda+\sigma\tilde{\gamma}_a\tilde{\gamma}_p(1-R_0)=0, \tag{A15}$$

$$\text{where}\quad \tilde{R}_0=\alpha\frac{\tilde{\beta}_a}{\tilde{\gamma}_a+\tilde{\gamma}_p}+(1-\alpha)\frac{\tilde{\beta}_p}{\tilde{\gamma}_a+\tilde{\gamma}_p} \tag{A16}$$

$$\text{and}\quad R_0=\alpha\frac{u\beta_a}{\tilde{\gamma}_a}+(1-\alpha)\frac{u\beta_p}{\tilde{\gamma}_p}. \tag{A17}$$

We identify R_0 with the reproductive number of the disease. Noting the fact that $\tilde{R}_0<R_0$, it is easy to prove that the necessary condition for at least one positive eigenvalue is

$$R_0>1. \tag{A18}$$

Let us denote the largest eigenvalue by μ. For $R_0\approx 1$, we expect that the largest eigenvalue is close to zero and from Eq. (A17) we can read off the value as

$$\mu\approx\frac{\sigma(R_0-1)}{1+\sigma(\tilde{\gamma}_a^{-1}+\tilde{\gamma}_p^{-1})(1-\tilde{R}_0)}. \tag{A19}$$

We denote the right and left eigenvectors corresponding to the eigenvalue μ by $\phi_m(i)$ and $\chi_m(i)$ respectively. The time evolution of the vector $X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$ is given by

$$\begin{aligned} x_i(t) &= \sum_j \sum_q \phi_q(i)\chi_q(j)e^{\lambda_q t}x_j(0) \\ &\approx \sum_j \phi_m(i)\chi_m(j)e^{\mu t}x_j(0), \end{aligned} \tag{A20}$$

where the last line is true at sufficiently large times when only one eigenvalue λ dominates. Let us consider the initial condition $X = (-\epsilon, 0, 0, \epsilon, 0, 0, 0, 0)$ so that (noting that $\chi_m(1) = 0$)

$$x_i(t) \approx \epsilon\phi_m(i)\chi_m(4)e^{\mu t} = a_i\epsilon e^{\mu t}, \tag{A21}$$

where $a_i = \phi_m(i)\chi_m(4)$. At a sufficiently large time t_l we equate the observed confirmed number C_0 on some day to $x_6(t_l) + x_7(t_l) + x_8(t_l)$ which therefore gives us the relation

$$\epsilon e^{\mu t_l} = \frac{C_0}{a_6 + a_7 + a_8}. \tag{A22}$$

This then tells us that we should start with the following initial conditions, counting now time from $t = 0$:

$$x_i(0) = \frac{\phi_m(i)}{\phi_m(6) + \phi_m(7) + \phi_m(8)}C_0. \tag{A23}$$

The crucial point is that the leading eigenvector fixes the direction of the growth and then knowledge of linear combination fixes all the other coordinates.

Appendix B: Analysis of the basic SEIR model

In the standard SEIR model one divides a population of size N into four compartments of

1. S = Number of Susceptible individuals.
2. E = Number of Exposed but not yet contagious individuals.
3. I = Number of Infected contagious individuals
4. R = Number of Recovered, hospitalized or dead individuals.

The dynamics of this model can be described as follows:

- The infected individuals, I, come in contact with the susceptible population , S, and cause transitions $S \to I$.
- People who are Exposed carry the virus, do not yet show symptons and cannot infect others.
- After a latency period T_L the Exposed people become Infected and can now infect others, so $E \to I$ happens at a rate $\sigma = 1/T_L$. These people could either be symptomatic or asymptomatic and their diseases are yet un-detected.
- We assume that infected people typically either recover or get detected after T_R days, so $I \to R$ happens at a rate $\sigma = 1/T_R$.

We then have the following equations for the dynamics for the system

$$\frac{dS}{dt} = -\frac{\beta I}{N}S, \tag{B1}$$

$$\frac{dE}{dt} = \frac{\beta I}{N}S - \sigma E, \tag{B2}$$

$$\frac{dI}{dt} = \sigma E - \gamma I, \tag{B3}$$

$$\frac{dR}{dt} = \gamma I. \tag{B4}$$

In this case the reproductive number is simply given by $R_0 = \beta/\gamma$.

1. Final fraction of infected population in the absence of intervention

We ask as to what fraction of the population would be affected finally if there was no intervention. To answer this question, we first note from Eqs. (B1,B4) that

$$\frac{dS}{dR} = -\frac{\beta}{N\gamma}S = -\frac{R_0}{N}S, \tag{B5}$$

hence

$$S(t) = S(0)e^{-R_0 R(t)/N}, \tag{B6}$$

where we have assumed $R(0) = 0$. In the steady state we should have $E = I = 0$ while $\bar{S} = S(\infty), \bar{R} = R(\infty)$ are determined from the condition $N = S + E + I + R$, which gives

$$N = \bar{S} + \bar{R} = S(0)e^{-R_0\bar{R}/N} + \bar{R}. \tag{B7}$$

Denoting the fraction of initially infected by $\epsilon = I(0)/N$ and the fraction of total eventually affected fraction by $\bar{x} = \bar{R}/N$, we see that x can be determined from solution the following equation

$$1 = (1-\epsilon)e^{-R_0\bar{x}} + \bar{x}. \tag{B8}$$

Typically $\epsilon << 1$ and so see that the final fraction of affected population is given by the solution of the equation

$$1 - \bar{x} - e^{-R_0\bar{x}} = 0. \tag{B9}$$

2. Size of infection population peak $I^{(m)}$ and the number of days to reach the peak

We now evaluate the peak value $I_{\max}$ of the infected population in the course of the outbreak. We first note that the equation (B6) allows us to express the susceptible population at any time t as a function of $R(t)$. In fact, one can express all the other populations in terms of $R(t)$ or its time derivatives, such as

$$\begin{aligned} I &= \frac{1}{\gamma}\frac{dR}{dt} \\ E &= \frac{1}{\sigma}\frac{dI}{dt} + \frac{\gamma}{\sigma}I = \frac{1}{\gamma\sigma}\frac{d^2R}{dt^2} + \frac{1}{\sigma}\frac{dR}{dt}. \end{aligned} \tag{B10}$$

Hence, after expressing all variables in terms of R and inserting them in the constraint equation $S + E + I + R = N$ we get

$$\begin{aligned} & Ne^{-R_0\frac{R}{N}} + \frac{1}{\gamma\sigma}\frac{d^2R}{dt^2} + \frac{1}{\sigma}\frac{dR}{dt} + \frac{1}{\gamma}\frac{dR}{dt} + R = N \\ \Rightarrow \quad & \frac{d^2x}{dt^2} + (\gamma+\sigma)\frac{dx}{dt} + \gamma\sigma\left(x + e^{-R_0 x} - 1\right) = 0, \end{aligned} \tag{B11}$$

where, $x = R/N$. Defining $v = \frac{dx}{dt} = \gamma I/N$, we see that the four dimensional SEIR-dynamics is equivalent to a two-dimensional dynamical system specified by the equations

$$\frac{dx}{dt} = v \tag{B12}$$

$$\frac{dv}{dt} = -(\gamma+\sigma)v - \gamma\sigma\left(x + e^{-R_0 x} - 1\right). \tag{B13}$$

The above equation resembles a damped oscillator constrained to move in the positive half and in a potential $U(x) = \gamma\sigma(x^2/2 - x) + (\gamma^2\sigma/\beta)e^{-R_0 x}$ so that $F(x) = -U'(x) = -\gamma\sigma(x + e^{-R_0 x} - 1)$. The nontrivial fixed point, which is the steady state, is given by the zero of $F(x)$, as already obtained in earlier section.

On the other hand, the peak of the infected population is given by setting $dI/dt = 0$ or $dv/dt = 0$, which implies $v^{(m)} = -\frac{\gamma\sigma}{\gamma+\sigma}\left(x_m + e^{-R_0 x_m} - 1\right)$, where $x^{(m)}$, $v^{(m)}$ denote the values of x, v at the time when I peaks. To determine $(x^{(m)}, v^{(m)}$ we need another equation which could be obtained for example from a solution of the equation for dv/dx.

This is difficult to obtain exactly. However we can obtain a second equation if we make the reasonable assumption that I and E peak at around the same time, hence we get a second equation which simply gives $S^{(m)}/N = 1/R_0$. On the other hand From Eq. (B6) we get $R^{(m)} = -R_0^{-1} \ln(S/N) = R_0^{-1} \ln R_0$. Then using the overall constraint $N = S + E + I + R$ we finally obtain

$$I^{(m)} = \frac{\sigma}{\gamma + \sigma} \left(1 - \frac{1 + \ln R_0}{R_0} \right) N. \tag{B14}$$

An estimate of the time to reach this peak value can be obtained by noting that we can use the linearized dynamics (see previous section) till the time $I(t)$ reaches its peak $I^{(m)}$. Hence we write $I^{(m)} = I(t_{\max}) = I(0)\ e^{\mu t_{\max}}$ which provides $t_{\max} = \frac{\ln[I_{\max}/I(0)]}{\mu}$. We naturally expect that $I_{\max}$ is of the order $O(N)$ which implies

$$t_m \sim \frac{\ln[I^{(m)}/I(0)]}{\mu} \sim \frac{\ln N}{\mu}. \tag{B15}$$

3. Linear analysis of the SEIR model and fixing of initial conditions

To get the growth at early time regime let us define the variables $S = N + s, E = e, I = i, R = r$. Inserting these in Eqs. (B1, B2, B3) and (B4), and then expanding the right hand sides of each equations to linear order, we get

$$\frac{dX}{dt} = M\ X, \quad \text{where,} \quad X = \begin{pmatrix} s \\ e \\ i \\ r \end{pmatrix}, \quad \text{and} \quad M = \begin{pmatrix} 0 & 0 & -\beta & 0 \\ 0 & -\sigma & \beta & 0 \\ 0 & \sigma & -\gamma & 0 \\ 0 & 0 & \gamma & 0 \end{pmatrix}. \tag{B16}$$

This set of linear equations can be solved by diagonalising the matrix M. It has eigenvalues

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= 0, \\ \lambda_3 &= \frac{-(\sigma+\gamma) - \sqrt{(\sigma-\gamma)^2 + 4\beta\sigma}}{2} \\ \lambda_4 &= \frac{-(\sigma+\gamma) + \sqrt{(\sigma-\gamma)^2 + 4\beta\sigma}}{2}. \end{aligned} \tag{B17}$$

Let us denote the right and left eigenvectors corresponding to the eigenvalue λ_q by $\phi_q(i)$ and $\chi_q(i)$ respectively. The right eigenvectors are given by

$$\begin{aligned} \phi_1 &= (0, 0, 0, 1), \\ \phi_2 &= (1, 0, 0, 0), \\ \phi_3 &= (-\beta/\gamma, \lambda_3(\lambda_3 + \gamma)/(\sigma\gamma), \lambda_3/\gamma, 1), \\ \phi_4 &= (-\beta/\gamma, \lambda_4(\lambda_4 + \gamma)/(\sigma\gamma), \lambda_4/\gamma, 1). \end{aligned} \tag{B18}$$

We denote the largest eigenvalue $\lambda_4 \equiv \mu = [-(\sigma+\gamma) + \sqrt{(\sigma-\gamma)^2 + 4\beta\sigma}]/2$ and it is easy to see that this is positive for $\beta/\gamma = R_0 > 1$.

It is instructive to examine the structure of μ near $R_0 = 1$. For this we rewrite this in the form

$$\mu = \frac{-(\sigma+\gamma) + \sqrt{(\sigma+\gamma)^2 + 4\gamma\sigma(R_0 - 1)}}{2} \approx \frac{\gamma\sigma}{\gamma+\sigma}(R_0 - 1). \tag{B19}$$

One qualitative aspect that this equation tells us is the following. Suppose we start with free parameters β, γ such that $R_0 = 1.8$ and want to change (through interventions) the reproductive number to a target value $R_0(\text{target}) = R_0/2 = 0.9$. We can do this either (a) by decreasing β to $\beta' = \beta/2$ or (b) by increasing γ to a value $\gamma' = 2\gamma$. It is clear from the above expression that (b) would lead to a negative eigenvalue of larger magnitude and so a faster decay of the disease.