A simple SIR model with a large set of asymptomatic infectives

Giuseppe Gaeta
Dipartimento di Matematica,
Università degli Studi di Milano,
via Saldini 50, I-20133 Milano (Italy)
and
SMRI, 00058 Santa Marinella (Italy)
(Dated: 24/05/2020– revised & augmented version)

There is increasing evidence that one of the most difficult problems in trying to control the ongoing COVID-19 epidemic is the presence of a large cohort of asymptomatic infectives. We develop a SIR-type model taking into account the presence of asymptomatic, or however undetected, infective, and the substantially long time these spend being infective and not isolated. We discuss how a SIR-based prediction of the epidemic course based on early data but not taking into account the presence of a large set of asymptomatic infectives would give wrong estimate of very relevant quantities such as the need of hospital beds, the time to the epidemic peak, and the number of people which are left untouched by the first wave and thus in danger in case of a second epidemic wave. In the second part of the note, we apply our model to the COVID-19 epidemics in Northern Italy. We obtain a good agreement with epidemiological data; according to the best fit of epidemiological data in terms of this model, only 10% of infectives in Italy is symptomatic.

I. INTRODUCTION

There is increasing evidence that one of the main difficulties in trying to control the ongoing COVID-19 epidemic is the presence of a large cohort of *asymptomatic infectives* [1–16]. This feature was first noted in the analysis of passengers of evacuation flights from Wuhan, see e.g. [17], and on the Crown Princess cruise ship [18], but in these circumstances the asymptomatic cases could easily be thought to be just cases which had not yet developed symptoms.

So the first confirmation of this characteristic of COVID was obtained in the large scale study of one of the first infection foci in Italy, that of Vò Euganeo near Padua; here the whole population (about 3,000 people) of the village were tested twice – at one week distance – for the virus, and a significant number of asymptomatic infectives was detected [19].

We stress that it is correct to speak of asymptomatic *infectives* and not just *carriers*. In fact, it is by now thought that they are as infective as symptomatic ones in terms of viral charge [1, 13–15] – and of course potentially even more dangerous as the absence of symptoms leads to a low level of precautions.

One of the consequences of this fact is that the *registered infectives*, those known to the national health systems and thus isolated and monitored, are only a part of the total pool of infectives.

The initial estimates were that registered infectives would be between 1/3 and 1/4 of the actual infectives [20]; there have been early claims by the British Government scientific advisers [21] that this ratio could be as little as 1/10. In a recent contribution [22] Li *et al.* estimate that only about 1/7 of the infections are detected and can thus be isolated. More recently, other studies have suggested that the fractions of undetected infections could be even higher [23, 24].

This is obviously a very relevant matter, both to understand the epidemic dynamic and to design concrete actions to counter the epidemic spreading.

The goal of this note is two-fold: on the one hand we want to develop a simple general model to take into account the presence of asymptomatic infectives; on the other hand, we want to apply this to the ongoing COVID-19 epidemic, in particular considering the situation in Italy, i.e. in the first European country to be heavily struck by COVID.

One of the problems faced in this situation is that at the start of the epidemic due to a *new* pathogen, one has to estimate the infectivity of this, and more generally the parameters in any model used to describe the epidemic. This will have to be considered in some detail.

The estimate of the fraction of asymptomatic infectives is one of the problems encountered in this sense. In the case of COVID in Italy, our estimate – based on a fit of the epidemiological data by the model, see Sections VII and VIII below – is that only 10% of infectives are symptomatic; this also means that in the early phase of the epidemic, in particular before the relevance of asymptomatic transmission was fully understood and asymptomatic infectives were searched for, only 10% of the infectives (at best, i.e. assuming that symptoms were properly related to COVID) were registered by the health system.

The *plan of the paper* is as follows. We start by recalling some basic facts about the well known SIR model (Section II), and discuss in more detail how this can be fitted against the data available in the first phase of an epidemic (Section II E).

Beside all the obvious limitations of SIR-type models, the standard SIR – as recalled above – does not take into account the special features introduced by the presence of a large set of asymptomatic infectives, i.e. the aspect we want to focus on in this contribution. We will therefore

develop a SIR-type model taking into account the presence of asymptomatic infectives, and the substantially long time these spend being infective and not isolated; this is called A-SIR, the A standing indeed for asymptomatic (Section III). We also repeat in this context the discussion on how the model parameters can be estimated on the basis of the early stages of the epidemics in that context, see Sections III C and III D (we will find that parameters present in both the two models are fitted in the same way from available data).

We will then come, in Section IV, to one of the main points of our study, i.e. comparing the different predictions of the standard SIR and of the new A-SIR model for a given set of initial-stage epidemiological data (which are coded by the coefficients of a function fitting them). This comparison is made by means of numerical simulations for realistic values of the parameters, but with no reference yet to any concrete case.

One of our main interests is in understanding how relevant it can be to uncover asymptomatic infectives and promptly isolate them; we then study (numerically) how the dynamics is affected by a reduction of the removal time for this class (Section V), as could be obtained by a contact tracing and testing campaign [19].

The discussion so far considered a *generic* infection (providing long-term immunity to recovered patients) with a large set of asymptomatic infectives. In the second part of the paper, we apply our model to the COVID-19 epidemics in Italy (Section VI). In order to do this we first of all have to estimate the model parameters from data in the initial phase; to this aim we use the data from the first half of March and determine our best fit of the epidemiological data through a two-step procedure. This first determines relations between the parameters, reducing these to expressions in terms of the the removal times for symptomatic and asymptomatic patients, i.e. on parameters β and η (see below), and this is obtained based on data for the first week; then data for the second week are used to fit these latter parameters. We do of course also compare these fits with the data in the initial phase, see Section VII. The A-SIR model outperforms the standard SIR model in this respect.

We will naturally also consider the subsequent development of the COVID-19 epidemics, but in order to do this we will have to consider the different sets of measures taken on March 8 and March 23 by the Italian Government and based on *social distancing*; these will be taken into account by means of a *reduction of the contact rate* parameter in the model. The amount r of reduction will be a fitted parameter (Section VIII).

With a suitable choice of r, we get a rather good agreement between the dynamic of our model and epidemiological data up to mid-May (i.e. the time of writing of this paper), see Figures 9 and 10.

We will end the paper by a discussion of some relevant general points and by conclusions (Section IX); we anticipate here the main ones:

(i) there is a marked difference between the standard SIR dynamics and the dynamics of the A-SIR model, i.e. the one taking into account the presence of a large class of asymptomatic infectives;

(ii) in the case of the COVID-19 epidemic in Italy, assuming a ratio of symptomatic to total infections of $\xi = 1/10$ yields a good agreement between the model and epidemiological data.

We stress that our estimate of the ratio ξ is – as far as we know – the first one given on the basis of a theoretical model and not just of statistics. On the other hand, it agrees with the current estimates given by health agencies and physicians.

It is maybe useful to also anticipate what are our conclusions about the consequences of the "marked difference" mentioned in item (i) above. We find – in Section IV – that a standard SIR analysis in a situation where the A-SIR model applies, would make three substantial errors: (a) The number of (symptomatic) infectives needing Hospital care would be over-estimated; (b) The time before the epidemic peak so the time available to prepare the health system to face it would be over-estimated; (c) The number of people not touched by the epidemic wave, so still in danger if a second wave arises, would be over-estimated. It is rather clear that each of these errors would have substantial practical consequences.

In an Appendix, we will also discuss how the presence of asymptomatic infectives affects our estimate of the *basic reproduction number* (usually denoted as R_0); this might explain why many national health agencies were caught short by the rapid rise in COVID-19 cases.

Our discussion will include a number of small deviations from the central development; these will be given in the form of Remarks. The symbol $\odot$ will signal the end of a Remark.

II. THE SIR MODEL

The classical SIR model for the dynamics of an infective epidemic providing permanent immunity to those who have already been infected and recovered [25–31] describes a homogeneous and isolated population of N individuals by partitioning them into three classes: each individual can be either susceptible (S), infected and infective (I), or removed (R) from the epidemic dynamics (that is, either recovered, dead, or isolated). We denote by $S(t)$, $I(t)$ and $R(t)$ the populations of these classes at time t; by assumption, $S(t) + I(t) + R(t) = N$ for all t.

The model is described by the equations

$$\begin{aligned} dS/dt &= -\alpha S I \\ dI/dt &= \alpha S I - \beta I \\ dR/dt &= \beta I \, . \end{aligned} \tag{1}$$

In the following, the parameter

$$\gamma = \beta/\alpha \tag{2}$$

will have a special relevance.

This model is well known, but we recall here some of its features both for the sake of completeness and with the purpose of comparing these with those for the new model to be introduced below. Further detail on the SIR model can be found in textbooks [27–31].

We would like to stress a relevant general point. As mentioned above, the SIR model (and more generally SIR type models) assume that we have an isolated population, and that individuals react to the pathogen and interact socially in a homogeneous way. In physicists' language, this is a *mean field* theory, i.e. in common language all individuals characteristics are erased, and real individuals are replaced by an "ideal" type corresponding to the average over the population. It goes without saying that these assumptions are not only radical, but also non realistic: individuals in any population differ for age, health state, and contact network. On the other hand, when – as in a starting epidemic which is overwhelming the sanitary system – we have scarce data (and available data are not organized in the way they should be in a laboratory experiment) it has the advantage of providing a qualitative description, whose quantitative predictions can be compared with experiment also if fed with the scarce and rough available data.

In other words, we agree that e.g. an epidemic model on a network [32–34] would be more realistic, but on the one hand we doubt that the available data allow to identify the existing network at this stage[53], and on the other hand we would not like to have predictions which depend on the (unknown) features of this network. We will thus be satisfied with working with SIR-type models.

Remark 1. Note that according to eqs.(1), an infected individual is immediately infective. For most infections this is not realistic, of course, but if the delay is substantially smaller that the characteristic removal time β^{-1}, we get a good approximation and still keep to a very simple model with all its advantages for qualitative discussion. The same remark will also apply to the A-SIR model, to be introduced and discussed in Section III below.

Note also that the SIR equations (1) stipulate that we have a constant population; in practical terms, this means we are considering an epidemic developing over a short enough timespan, i.e. such that one can disregard deaths and new births. In particular the latter provide new fuel to the susceptible class and thus if the epidemic goes on for a long time related terms should be included in the model; this may led to the presence of an *endemic state* [27]. ⊙

A. Epidemic dynamics, epidemic peak and total number of infections

It is immediately apparent that in the SIR model the number of infected will grow as long as

$$S \ > \ \gamma \ ; \tag{3}$$

thus γ is also known as the *epidemic threshold*. The epidemic can develop only if the population is above the epidemic threshold.

The ratio

$$R_0 \ := \ \frac{S_0}{\gamma} \tag{4}$$

is known as the *basic reproduction number* (BRN), or also *basic reproduction rate*, and is an estimate of how many new infections are originated from a single infective in the initial phase of the epidemic [35–37]. (In fact, the first equation in (1) says that there are $\alpha SI(\delta t)$ new infections in a time interval δt, and each infective is such for an average time β^{-1}.)

The parameters α and β describe the contact rate and the removal rate; they depend both on the characteristics of the pathogen and on social behavior. For example, a prompt isolation of infected individuals is reflected in raising β, a reduction of social contacts is reflected in lowering α, and both these actions raise the epidemic threshold γ. If this is raised above the level of the total population N, the epidemic stops (which means the number of infected individuals starts to decrease, albeit new individuals will still be infected). The same effect can be obtained by reducing the population N (keeping α and β constant), i.e. by partitioning it into non-communicating compartments, each of them with a population below the epidemic threshold.

Remark 2. Albeit strictly speaking these predictions only hold within the SIR model, and surely the exact value of the threshold refers to this model only, the mechanism at play is rather general, and similar behaviors are indeed predicted by all kind of epidemic models. ⊙

One can easily obtain the relation between I and S by considering the equations governing their evolution in (1) and eliminating dt; this provides

$$dI/dS \ = \ -1 \ + \ \gamma/S \ . \tag{5}$$

Upon elementary integration this yields (note we always write "log" for the natural logarithm)

$$I \ = \ I_0 \ + \ (S_0 \ - \ S) \ + \ \gamma \log(S/S_0) \ ; \tag{6}$$

with I_0, S_0 the initial data for $I(t)$ and $S(t)$; unless there are naturally immune individuals (which is not the case for new infections), $S_0 = N - I_0 \simeq N$.

As we know (see above) that the maximum I_* of I will be reached when $S = \gamma$, we obtain from (6) an estimate

of the level of this maximum; i.e. writing $\gamma = \sigma N$ (with $\sigma < 1$) this reads

$$I_* \;=\; [1 \;-\; \sigma \;-\; \sigma \log(1/\sigma)] \;\; N \; . \qquad (7)$$

Note that we do *not* have an analytical estimate of the time needed to reach this maximum; see below.

It follows from (7) that increasing γ, even if we do not manage to take it above the population N, leads to a reduction of the epidemic peak; if we are sufficiently near to the epidemic threshold, this reduction can be rather relevant also for a relatively moderate reduction of α and thus increase of γ.

The formula (6) also allows to obtain an estimate for another parameter describing the severity of the epidemics, i.e. the total number of individuals R_∞ which are infected over the whole span of the epidemics. In fact, the epidemic is extinct (at an unknown time $t = T_0$) when $I = 0$; the number of susceptibles S_∞ at this stage is provided there by the (lower) root of the equation

$$I_0 \;+\; (S_0 \;-\; S) \;+\; \gamma \log(S/S_0) \;=\; 0 \; ;$$

as noted above $I_0 \simeq 0$, $S_0 \simeq N$, and we can simply look at

$$(N \;-\; S_\infty) \;+\; \gamma \log(S_\infty/N) \;=\; 0 \; . \qquad (8)$$

This is a transcendental equation, but it is easily solved numerically if γ is known. The sought for number of overall infected individuals R_∞ is of course provided by

$$R_\infty \;=\; N \;-\; S_\infty \; . \qquad (9)$$

Remark 3. In the case of a small epidemic, we have $S_\infty/N \simeq 1$, and we can expand the logarithm in a Taylor series; in this way we get $S_\infty = (3 - 2N/\gamma)N$. See also Section II C below. ⊙

Another key quantity is the *speed* at which the epidemic dynamics develops, and in particular the time t_* at which I reaches its maximum value I_*, and the time t_∞ needed for I to get to zero (and $S(t_\infty) = S_\infty$, of course). In this case one can not get an analytical estimate, but it is possible to describe how this depends on the values of α and β for a given population level N and initial conditions $\{S_0, I_0, R_0\}$. In fact, the equations (1) are invariant under the scaling

$$\alpha \;\to\; \lambda\,\alpha \; , \quad \beta \;\to\; \lambda\,\beta \; , \quad t \;\to\; \lambda^{-1}\,t \; . \qquad (10)$$

(Note that the inverse scaling of β and t is enforced by the very physical meaning of β, which is the inverse of the characteristic time for the removal of infectives.)

The meaning of (10) is that if we manage to reduce α by a factor λ, even in the case β is also reduced and thus γ remains unchanged, the speed of the epidemic dynamics is also reduced by a factor λ. On the other hand, it is clear from the equation for dS/dt in (1) that reducing α reduces the speed at which new infective appear; if the removal rate β is unchanged, this will make that I grows slower and reaches a lower level. See Figure 1 in this regard.

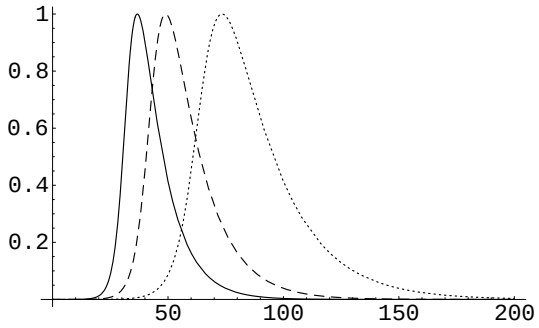


FIG. 1: Effect of changing the parameters α and β with constant γ on the SIR dynamics. Numerical solutions to the SIR equations with given initial conditions are computed for different choices of α and β with constant ratio $\gamma = \beta/\alpha$. In particular we have considered $N = S_0 = 5{*}10^7$, $I_0 = 10^2$, and $\alpha = \alpha_0 = 10^{-8}$, $\beta = \beta_0 = 10^{-1}$ (solid curve); $\alpha_1 = (3/4)\alpha_0$, $\beta_1 = (3/4)\beta_0$ (dashed curve); $\alpha_2 = (1/2)\alpha_0$, $\beta_2 = (1/2)\beta_0$ (dotted curve). The vertical scale is in terms of the maximum $I_* \approx 2.39 * 10^7$ attained by $I(t)$. This is attained, in the three runs, at times $t_0 \simeq 36.70$, $t_1 = (4/3)t_0 \approx 48.94$ and $t_2 = 2t_0 \approx 73.41$ respectively.

B. Early dynamics

The SIR equations are nonlinear, and an analytical solution of them turns out to be impossible; they can of course be numerically integrated with any desired precision *if* the initial conditions *and* the value of the parameters are known.

In the case of well known infective agents (e.g. for the influenza virus) the parameters are known with good precision, and indeed Health Agencies are able to forecast the development of seasonal epidemics with good precision. Unfortunately this is not the case when we face a new virus, as for COVID-19.

Moreover, when we first face a new virus we only know, by definition, the early phase of the dynamics, so parameters should be extracted from such data.[54] We will thus concentrate on this initial phase, with $I(t)$ and $R(t)$ rather small, and try to obtain approximate analytical expressions for the dynamics; the purpose will be to estimate the parameters α and β – and thus also the epidemic threshold γ – in this case.

Remark 4. Albeit we do *not* expect, for various reasons, the SIR model to provide a good description of the dynamics when the infection produces a large number of asymptomatic carriers, having an estimate of these parameters will be needed to compare the predictions which one would extract from the standard SIR model in such circumstances with those obtained by the modified model we will consider later on, see Section III. As for the reasons to expect the SIR model to perform poorly in the presence of a large set of asymptomatic carriers, these have to do both with the practical implementation and with intrinsic limitations of the model. As for the first type, in practice we estimate the model's parameters

by epidemiological data based only on *registered* (thus mostly symptomatic) infectives; if these are only a small part of the infectives, the estimates of the parameters will be grossly different from the true ones. As for the intrinsic limitations of standard SIR in this setting, we will come back to this point at the end of Section III A, in Remark 12. ⊙

C. KMK approximate equations and their exact solution

In the case of "small epidemics" there is a way to obtain an analytical expression for the solutions to the SIR equations or more precisely to the approximate equations valid in the limit of small R/γ; this is associated to the names of Kermack and McKendrick [25], and we will therefore refer to it as the KMK method. (This is very classical, and is discussed here for the sake of completeness.)

What matters more here, the expression obtained in this way is also an analytical expression holding *in the initial phase* of any epidemics, small or large, i.e. – as we will discuss in a moment – until $R(t) \ll \gamma$.

Thus such an analytical expression can be compared to *early* epidemiological data and used to estimate the unknown parameters α and β, and hence the fundamental parameter γ. Once this is done, the model can be studied numerically (or, if we are – as has to be hoped – in the favorable situation where $N \simeq \gamma$, one can set predictions on the basis of the "small SIR epidemic" model) – recalling of course that the SIR model itself is far too simple to be reliable in a situation where the actions undertaken have heavy consequences on public health – in order to have some kind of estimate of the length of the epidemics and of other relevant outcomes, such as the numbers I_* and J_∞ considered above.

It should be noted that we do not have full knowledge about the number of infective people at each time; the best we can have is the number of people who are hospitalized or however registered by the health system. Assuming that infective people are immediately isolated, this provides an estimate (actually from below) of $R(t)$. Thus we should be able to compare the predictions for the removed class with epidemiological data, and in order to do this we should focus on $R(t)$. We stress that his problem was already clear to Kermack and McKendrick [25], see e.g. the discussion in Murray [27], and that we will basically follow their idea albeit with a relevant difference, which will allow for a simpler fit of the data.

Putting together the equations for S and for R in (1), we have $dS/dR = -S/\gamma$, which of course provides

$$S(R) \;=\; S_0\, e^{-(R-R_0)/\gamma}\;. \tag{11}$$

We can proceed similarly with the equations for I and for R, getting $dI/dR = -1 + S/\gamma$, where now S should be thought of as a function of R through (11). Solving this equation we get

$$I(R) \;=\; I_0 \,+\, S_0\,\left(1 \,-\, \exp[-(R-R_0)/\gamma]\right) \,-\, (R-R_0)\;. \tag{12}$$

We are however interested in the temporal dynamics of the model. In order to do this, we can substitute for $I = I(R)$ using (12) in the equation for dR/dt in (1); moreover we will look at the variable

$$P(t) \;:=\; R(t) \;-\; R_0\;, \tag{13}$$

which of course satisfies $P(0) = 0$ and $dP/dt = dR/dt$. In this way we have

$$I(P) \;=\; I_0 \;+\; S_0\,\left(1 \,-\, e^{-P/\gamma}\right) \;-\; P\;. \tag{14}$$

Plugging now this into the third equation of (1), we finally get

$$\frac{dP}{dt} \;=\; \beta\,\left[I_0 \;+\; S_0\,\left(1 \,-\, e^{-P/\gamma}\right) \;-\; P\right]\;. \tag{15}$$

This is a transcendental equation and can *not* be solved exactly. However, as long as $P/\gamma \ll 1$, i.e. as long as $R(t)$ is well below the epidemic threshold, we can replace the exponential by (a suitable truncation of) its Taylor series expansion.

Remark 5. In textbook discussions, it is usually required to consider a *second order* Taylor expansion; this guarantees that counter-terms preventing the exponential explosion of $R(t)$ (and thus the violation of the condition $R(t) \ll \gamma$) are present, and allows to obtain an analytical expression for $R(t)$ – solution of the approximate SIR equations for $R/\gamma \ll 1$ – valid at *all times*. This is, more precisely, in the form

$$R(t) \;=\; \frac{\alpha^2}{S_0}\,\left[\phi \;+\; k_1\,\tanh\left[\frac{k_1\beta}{2}\,t \;-\; k_2\right]\right]\;, \tag{16}$$

where we have written $\phi := (S_0/\gamma - 1)$ and k_1 and k_2 are explicitly given by

$$k_1 = \sqrt{\phi^2 + 2(S_0/\gamma^2)(N - S_0)};\; k_2 = k_1^{-1}\,\mathrm{arctanh}(\phi)\;. \tag{17}$$

As we assume there is no natural immunity, we can take $S_0 \approx N$, obtaining $k_1 \approx \phi$ and hence slightly simpler complete expressions. ⊙

Remark 6. In particular, in this case the maximum of $R'(t)$ – and hence of $I(t)$, see (1) – is obtained at time

$$t_* \;=\; \frac{2\;\mathrm{arctanh}(\phi)}{\beta\;\phi^2}\;;$$

as our result holds for the "small epidemics", ϕ is small and we can write

$$\widetilde{t_*} \;\simeq\; \frac{2}{\beta\,\phi} \;+\; \frac{2}{3}\frac{\phi}{\beta}\;.$$

Note that t_* is therefore rapidly decreasing with ϕ (for small ϕ). On the other hand, looking back at (7), and

noticing that in terms of the notation used there $\sigma = 1/(1+\phi)$, we obtain immediately that I_* grows with ϕ. ⊙

Remark 7. This also means that if one would be able to tune the parameters α and β (and hence ϕ) there would be a contrast between trying to have a low I_* and hence a small ϕ, and trying not to have the epidemic running for too long – which can be devastating on social and economic grounds. If, on the other way, the priority from the temporal point of view is on slowing down the epidemic, e.g. to have the time to prepare the health system facing the peak, a small ϕ should be pursued. ⊙

D. Small time solution of the KMK equations

We have seen in Section II C that a full solution of the KMK model is equivalent to a full solution of eq. (15), and if we have a small epidemic then the solution is well approximated – *at all times* – by (16).

Here we are less ambitious: when the epidemic is just starting, we can in any case only fit the initial phase of the epidemic, which shows an exponential increase of $R(t)$, and correspondingly we can expand the exponential in (15) at *first* order in P/γ. This yields the equation

$$\frac{dP}{dt} \;=\; \beta\,\left[I_0 \;+\; \left(\frac{S_0}{\gamma} \;-\; 1\right)\;P\right]\;, \tag{18}$$

with initial condition $P(0) = 0$. This is immediately solved to give

$$P(t) \;=\; I_0\;\frac{\exp[\beta\,(S_0/\gamma - 1)\,t] \;-\; 1}{(S_0/\gamma \;-\; 1)}\;. \tag{19}$$

Introducing the parameter, which we stress is now *not* assumed to be small,

$$\phi \;:=\; \frac{S_0}{\gamma} \;-\; 1\;, \tag{20}$$

the above is more simply written as

$$P(t) \;=\; \frac{I_0}{\phi}\;\left[e^{\beta\,\phi\,t} \;-\; 1\right]\;, \tag{21}$$

and finally we get

$$R(t) \;=\; R_0 \;+\; \frac{I_0}{\phi}\;\left[e^{\beta\,\phi\,t} \;-\; 1\right]\;. \tag{22}$$

As expected this – at difference with (16) – is not saturating but just expanding exponentially, and thus cannot be valid for all times, but only for t sufficiently small (even for small ϕ).

The expression (22) can then be expanded in series to give the small t expression of the solution, which can be fitted against experimental data thus determining (some of) the parameters, see Section II E below.

Remark 8. It is relevant for the following of our discussion to note that the solution (22) can be obtained also in a different way, i.e. noticing that in the initial phase of the epidemic the number of susceptibles vary very little and can thus be considered as constant, $S(t) \simeq S_0$. Within his approximation, and writing again $\phi = (S_0/\gamma - 1)$, the SIR equations reduce to

$$\begin{cases} dI/dt \;=\; \beta\,\phi\,I \\ dR/dt \;=\; \beta\,I \end{cases}\;; \tag{23}$$

this is a linear system, and it is promptly solved to yield indeed (22).

Note this approach is actually simpler than the one followed above (so the reader may wonder why we have not taken this immediately), but on the one hand the procedure given above is along the lines of the tradition of SIR analysis, and on the other hand having seen that derivation gives us more confidence that a rough approach as this one provides the same results as a more refined one; this will be of use dealing with more complex models, where the Kermack-McKendrick approach can not be followed, see Section III C below. ⊙

E. Fitting the SIR parameters

Note that the solution (22) depends on three parameters, i.e. β, I_0 and ϕ, which in turn depends on the known number $S_0 \simeq N$ and on the unknown parameter γ. None of the parameters $\{\beta, \phi, I_0\}$ is known, but β can be directly estimated on the basis of medical data as it corresponds to the inverse of the typical removal time (for trivial infections, this corresponds to the time of healing; in the case of COVID it is the time from infection to isolation), and similarly once we fix a time $t = 0$ the number I_0 can be estimated *a posteriori* looking at epidemiological data for the next few days and depending on our estimate of β.

In order to estimate the parameters on the basis of the measurements of R, we can work either on R itself, or on its logarithm. That is, we have two alternative – but obviously equivalent – ways to proceed.

(a) *Working on the time series for $R(t)$.*
We fit the time series around t_0 by

$$R(t) \;=\; r_0 \;+\; r_1\,t \;+\; \frac{1}{2}\,r_2 t^2\;; \tag{24}$$

Having these coefficients r_k, we can compare with the series expansion for $R(t)$ given by (22), which is just

$$R(t) \;=\; R_0 \;+\; (\beta\,I_0)\;t \;+\; \frac{1}{2}\;\left(\beta^2\,I_0\,\phi\right)\;t^2\;. \tag{25}$$

We obtain easily that – using also the definition of ϕ (20) – our parameters and the associated parameter γ

are given by

$$R_0 \;=\; r_0 \;,\quad I_0 \;=\; \frac{r_1}{\beta} \;,\quad \phi \;=\; \frac{r_2}{r_1\,\beta} \;;$$
$$\gamma \;=\; \frac{\beta S_0 r_1}{\beta r_1 + r_2} \;. \tag{26}$$

(b) *Working on the time series for* $\log[R(t)]$.
As $R(t)$ grows – in the early phase – in a substantially (but not exactly, see above) exponential way, one usually deals with data in logarithmic form; that is one has a fit for $\log[R(t)]$, say of the form

$$\log[R(t)] \;=\; A \;+\; B\,t \;+\; \frac{1}{2}\,C\;t^2 \;. \tag{27}$$

Comparing with the series expansion of $\log[R(t)]$ for R as in (22), i.e.

$$\log[R(t)] \;=\; \log(R_0) \;+\; \beta\;\frac{I_0}{R_0}\;t \\ -\;\frac{1}{2}\;\left(\frac{\beta\,I_0}{R_0}\right)^2\left(1-\phi\frac{R_0}{I_0}\right)\;t^2 \;, \tag{28}$$

we obtain that the I_0, ϕ and γ parameters can be estimated as

$$R_0 \;=\; e^A \;,\quad I_0 \;=\; \frac{B}{\beta}\;e^A \;,\quad \phi \;=\; \frac{B^2 \;+\; 2\,C}{\beta\;B} \;;$$
$$\gamma \;=\; \frac{\beta S_0 B}{\beta B + B^2 + 2C} \;. \tag{29}$$

We stress again that our estimates for the parameters depend on our estimate for the leading parameter β.

This concludes our discussion of the standard SIR model.

III. A MODEL WITH ASYMPTOMATIC INFECTIVES

It may happen to have an epidemic such that a rather large fraction of infected people are actually asymptomatic, but still fully infective; this appears to be the case for COVID-19.[55]

A little reflection shows that the presence of a large population of asymptomatic infectives, or however of infectives which show only very mild symptoms, easily thought not to be related with the concerned infective agent, changes the dynamics in two – contrasting – ways:

- On the one hand, asymptomatic infectives are a formidable vehicle of contagion, as they have no reason to take special precautions, and get in contact with a number of people which themselves do not take the due precautions (which would be taken in the case of an individual with evident symptoms);
- On the other hand, assuming once the infection is ceased the infected (including of course asymptomatic ones) have acquired permanent immunity, they contribute to group immunity reached once the population of susceptibles falls below the epidemic threshold.

We are thus going to study how the SIR dynamics is altered by the presence of a large class of asymptomatic infectives.

Remark 9. An obvious but important Remark is in order here. If we find out that known infectives are only a fraction $\xi < 1$ of the total infectives, this means that on the one hand the mortality rate (number of deceased over number of infected) is actually smaller by the same factor. On the other hand, the total number of infected persons is increased by a factor ξ^{-1}, so that it looks more difficult to stop the spreading of the epidemics, and the final number of infected will be quite large.

In this respect, one should however remember that the total number of casualties does not depend only on the total number of individuals with symptoms but also on the number of patients needing Intensive Care (for the COVID epidemic in China this was estimated at 20 % of hospitalized patients [38, 39]; in Italy this fraction reached 10% at peak time, albeit in the most affected regions the number was higher [40–42]) and on the availability of such care; from this point of view, slowing down the pace of the epidemics can substantially lower the death toll even if the total number of affected individuals remains the same. ⊙

A. The A-SIR model

We will formulate a very simplified model of SIR type, where infective people are either symptomatic or asymptomatic. A more refined subdivision of their state would be more realistic, but the discussion of this simple case will suffice to show how to proceed in a more general setting.[56]

In our model we still assume permanent immunity of individuals who have been infected and recovered, and constant population. We also assume that – as in the classical SIR – infected individuals are immediately infective (see Remark 1 in this regard). We will have susceptibles $S(t)$ in a unique class, but two classes of infected and infective people: symptomatic $I(t)$ and asymptomatic $J(t)$; and similarly two classes of removed people: formerly symptomatic removed $R(t)$ and formerly asymptomatic removed (mostly passing unnoticed through the infection) $U(t)$. Symptomatic infectives are removed by the epidemic dynamics through isolation (in hospital or at home) at a removal rate β (thus with typical delay β^{-1}, while asymptomatic people are removed from the epidemic dynamics mostly through spontaneous recovery, at a recovery rate $\eta \ll \beta$, thus after a typical time $\eta^{-1} \gg \beta^{-1}$; detecting asymptomatic infectives leads to

their isolation before healing, and thus to a reduction in η^{-1}.

We assume that both classes of infected people are infective in the same way, and that an individual who gets infected belongs with probability ξ to the class I and with probability $(1-\xi)$ to the class J.

Our model, which we will call A-SIR (Asymptomatic-SIR) will then be

$$\begin{aligned} dS/dt &= -\alpha\, S\,(I+J) \\ dI/dt &= \alpha\,\xi\, S\,(I+J) \;-\; \beta\, I \\ dJ/dt &= \alpha\,(1-\xi)\, S\,(I+J) \;-\; \eta\, J \\ dR/dt &= \beta\, I \\ dU/dt &= \eta\, J\ . \end{aligned} \tag{30}$$

Note that the last two equations amount to an integral, i.e. are solved by

$$\begin{aligned} R(t) &= R_0 \;+\; \beta \int_0^t I(\tau)\; d\tau\ , \\ U(t) &= U_0 \;+\; \eta \int_0^t J(\tau)\; d\tau\ ; \end{aligned} \tag{31}$$

thus they can be considered as "reconstruction equations", and we will focus on the first three equations in (30). Moreover, the total population $N = S + I + J + R + U$ is constant.

Remark 10. The fact that asymptomatic infectives are as infective as symptomatic ones is not at all obvious. In the case of COVID-19 it appears we are exactly in this situation [1, 13–15]; in any case this will be our working hypothesis. ⊙

Remark 11. An alternative writing for the equations (30) is also possible, and in some case convenient. We denote by

$$K(t) \;=\; I(t) \;+\; J(t) \tag{32}$$

the total number of infectives, and by

$$x(t) \;=\; I(t)/K(t) \tag{33}$$

the fraction of symptomatic infectives; obviously the fraction of asymptomatic ones will be $y(t) = J(t)/K(t) = 1 - x(t)$.

By standard algebra the (first three) equations (30) for the A-SIR dynamics are then rewritten in terms of these variables as

$$\begin{aligned} dS/dt &= -\alpha\, S\, K \\ dK/dt &= \alpha\, S\, K \;-\; [\beta\, x \;+\; \eta\,(1-x)]\; K \\ dx/dt &= \alpha\, S\,(\xi - x) \;-\; (\beta - \eta)\,(1-x)\ . \end{aligned} \tag{34}$$

This writing also stresses that albeit ξ is a constant, the fraction x of symptomatic infectives is a *dynamical variable*. ⊙

Remark 12. We can now better discuss, in the light of the previous Remark 11, what are the intrinsic reasons which make that we expect the standard SIR model to perform poorly in describing an epidemic with a large number of asymptomatic infectives (see Remark 4 above). As already recalled, one would expect β to correspond to the inverse of the average removal time for registered infectives; however, once the presence of asymptomatic – and in particular undetected – infectives is ascertained, this could be corrected by accepting as β a weighted average B of the removal times for symptomatic and asymptomatic (that is, of β and η in the A-SIR model). But as we have just seen in Remark 11, the fraction x of symptomatic infectives is a dynamical variable, hence the parameter B should vary with time depending on the internal dynamics of the system (which is described by A-SIR, not by SIR), and hence we would be outside the proper SIR framework. ⊙

B. A-SIR dynamics

Some general considerations on eqs.(30) can be made immediately. First of all, we note that $I(t)$ will increase as far as the condition $\alpha\xi S(I+J) > \beta I$ is satisfied; that is, as far as

$$S \;>\; \gamma_1 \;:=\; \frac{1}{\xi}\,\frac{\beta}{\alpha}\,\frac{I}{I+J} \;=\; \frac{x}{\xi}\,\frac{\beta}{\alpha}\ . \tag{35}$$

Thus the epidemic threshold (for symptomatic patients) γ_1 depends both on the fixed parameters ξ, α, β *and* on the variable ratio $x(t)$ of symptomatic infectives over total infectives, see (33).

Similarly, the number of asymptomatic infectives $J(t)$ will grow as far as $\alpha(1-\xi)S(I+J) > \eta J$ is satisfied, i.e. as far as

$$S \;>\; \gamma_2 \;:=\; \frac{1}{1-\xi}\,\frac{\eta}{\alpha}\,\frac{J}{I+J} \;=\; \frac{1-x}{1-\xi}\,\frac{\eta}{\alpha}\ . \tag{36}$$

Again the epidemic threshold (for asymptomatic patients) γ_2 depends both on the fixed parameters ξ, α, η *and* on the variable ratio $y(t) = 1 - x(t)$ of asymptomatic – and thus "hidden" – infectives over total infectives.

Note that

$$\frac{\gamma_1}{\gamma_2} \;=\; \left(\frac{1-\xi}{\xi}\right)\left(\frac{\beta}{\eta}\right)\left(\frac{I}{J}\right) \;=\; \left(\frac{1-\xi}{\xi}\right)\left(\frac{\beta}{\eta}\right)\left(\frac{x}{1-x}\right)\ .$$

As we expect on the one hand to have $\xi < 1/2$ and $\beta > \eta$, but on the other hand $I < J$ and hence $x < 1-x$, we cannot claim there is a definite ordering between γ_1 and γ_2; this means that we will have situations where I declines and J is still growing, but the opposite is also possible.

We expect that in the very first phase –when the different removal times have not yet shown their effects – we have $J \simeq [(1-\xi)/\xi]I$ and hence $(1-x) \simeq [(1-\xi)/\xi]x$;

under this condition, we get

$$\frac{\gamma_1}{\gamma_2} \simeq \frac{\beta}{\eta} > 1 .$$

Remark 13. In the case of COVID-19, it is known that the incubation time is about 5.1 days; assuming that symptomatic infection is promptly recognized and swiftly treated, epidemiological and clinical data suggest the approximate values (note that asymptomatic removal time η^{-1} includes both the incubation time and the healing time) $\beta^{-1} \simeq 5-7$, $\eta^{-1} \simeq 14-21$ for the removal and recovery rates; the value of ξ is more controversial, as mentioned in the Introduction, but it is presently believed that $\xi \simeq 1/10$. We will come back to these matters when dealing specifically with COVID, but wanted to give immediately an idea of what "realistic" values can be for the parameters appearing the A-SIR model. ⊙

C. Early dynamics

It is quite clear that we can not go through the Kermack-McKendrick procedure to obtain approximate equations valid in the case of "small epidemics", not even through the simplified (first rather than second order) procedure we have used above. We can however go through the even simpler approach mentioned in Remark 8 (and which we have seen there produces the same results as the KMK procedure).

With $S(t) \simeq S_0$, the equations (30) reduce to a *linear* system of four equations with constant coefficients, or more precisely to a "master" system of two equations

$$\frac{dI}{dt} = (\alpha\,\xi\,S_0 - \beta)\,I + (\alpha\,\xi\,S_0)\,J \quad (37)$$

$$\frac{dJ}{dt} = [\alpha\,(1-\xi)\,S_0]\,I + [\alpha\,(1-\xi)\,S_0 - \eta]J \quad (38)$$

plus the two direct integrations (31).[57]

As for the two equations, (37) and (38), we can get their solution in explicit form by means of some standard algebra; they are slightly involved when written in fully explicit form, and we do not report them here.

With these, we can compute $R(t)$ and $U(t)$; their explicit expressions are also quite involved, and again we do not report them here.

Remark 14. As already remarked – and as well known from textbooks, see e.g. Murray [27] – we can only access $R(t)$ from epidemiological data.In our case, this is the number of symptomatic patients which have been registered and thus isolated, i.e. in particular removed from the epidemic dynamic. As for $U(t)$, this time series is basically unattainable: some of the asymptomatic infectives will be discovered and registered, but many – and probably most – of them will go unnoticed. Thus we should base our considerations *only* on $R(t)$. Unfortunately, the data published by health agencies and by WHO do not distinguish between symptomatic and asymptomatic patients. ⊙

D. Fitting the parameters

We can now proceed as in Section II E, i.e. series expand $R(t)$ in order to fit the parameters. From the explicit expression of $R(t)$ we get

$$R(t) \simeq R_0 + \beta\,I_0\,t + \frac{1}{2}\,\beta\,[\alpha(I_0+J_0)S_0\xi - \beta I_0]\;t^2 ; \quad (39)$$

$$\log[R(t)] \simeq \log(R_0) + \beta\,\frac{I_0}{R_0}\,t + \frac{1}{2}\,\frac{\beta[\alpha(I_0+J_0)R_0S_0\xi - \beta I_0(I_0+R_0)]}{R_0^2}\,t^2 \quad (40)$$

Comparing these with the generic form[58] of the fits (24) and (27), we can express the parameters I_0 and $\gamma = \beta/\alpha$. Note that we can not express both γ and J_0 with the same fitting, as both of them only appear in the coefficient of the quadratic term. Note also that in this context γ is *not* any more the epidemic threshold, as discussed in Section III A; the time-varying epidemic threshold $\gamma_1(t) = [x(t)/\xi]\gamma$ is however expressed in terms of γ, so that it still makes sense to fit it.

Actually, since new infected are with probability ξ in the class I and with probability $(1-\xi)$ in the class J, it is natural to set as initial condition

$$J_0 = \left(\frac{1-\xi}{\xi}\right)\,I_0 ; \quad (41)$$

with this assumption, we have

$$\gamma_1(t_0) = \gamma . \quad (42)$$

Remark 15. It should be noted that actually if we want to fit γ we need to have some estimate on J_0 (while I_0 can be fitted from first order coefficient in the series for $R(t)$ or $\log[R(t)]$); to this aim we will use consistently (41). ⊙

In particular, using the fit of R we get (through this assumption)

$$R_0 = r_0 , \quad I_0 = \frac{r_1}{\beta} , \quad \gamma = \frac{\beta\,r_1\,S_0}{(\beta\,r_1 + r_2)} . \quad (43)$$

Using instead the fit of $\log[R(t)]$, and again the assumption (41), we get

$$R_0 = e^A , \quad I_0 = \frac{B\,e^A}{\beta} ;$$
$$\gamma = \frac{\beta\,S_0\,B}{\beta\,B + B^2 + 2\,C} . \quad (44)$$

Once the parameters are estimated, the nonlinear equations (30) can be solved numerically.

It is immediate to check that the expressions (43) and (44) – which we recall were obtained under the assumption (41) for J_0 – are *exactly* the same as for the SIR model; see (26) and (29).

IV. COMPARING SIR AND A-SIR DYNAMICS

Summarizing our discussion so far, we have considered both the standard SIR model and a variant of it, the A-SIR model, taking into account the presence of large set of asymptomatic infectives. We have discussed in particular how the parameters for the models can be estimated on the basis of the time series for $R(t)$ in the early stage of the epidemic.

We want now to discuss how the prediction extracted from a given time series for $R(t)$ in the beginning of an epidemic differ if these are analyzed using the SIR or the A-SIR models.

As discussed above, we are able to extract only limited analytical information from the nonlinear SIR equations (1), and no relevant analytical predictions at all from the nonlinear A-SIR equations (30). Thus the only way to compare the predictions of the latter model with those of a standard SIR model (or variations on it, such as the SEIR model [27]), and hence see how the presence of a large class of asymptomatic infectives affects the dynamics within a SIR-type framework, is at present by running numerical simulations, i.e. numerically integrate the SIR and the A-SIR equations for *coherent* sets of parameters.

Here by "coherent" we will mean "extracted from the same time series for $R(t)$ in the early phase of the epidemics".

In practice, in view of the discussion above and depending on our choice to use the (43) or the (44) fit, this means "with given $\{r_0, r_1, r_2; \beta\}$ or given $\{A, B, C; \beta\}$ coefficients".

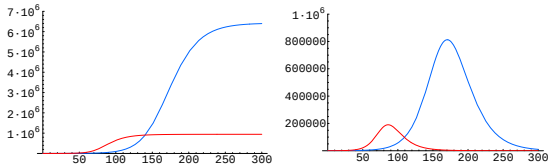
In the following, we will analyze the concrete epidemiological data for the COVID-19 epidemics in Italy, but here we would like to compare the SIR and the A-SIR predictions in a less concrete case, so that one can focus on general features rather than having a concrete case (with all the intricacies of real cases, see below) in mind.

We will thus choose an arbitrary (but realistic) set of parameters, which should be thought of as being extracted from the time series for the early phase of the epidemic. We consider a total population $N = 10^7$, $S_0 = N$, and parameters

$$A = 5\,, \quad B = 0.1\,, \quad C = -0.002\,; \\ \beta = \frac{1}{10}\,, \quad \eta = \frac{1}{20}\,; \quad \xi = \frac{1}{10}\,. \tag{45}$$

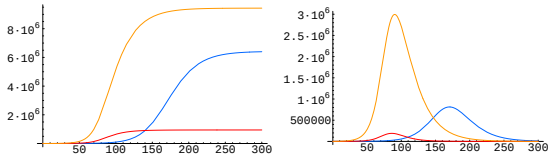
Moreover, we will assume that asymptomatic infectives (in the A-SIR modeling) are totally undetected; that is, we compare predictions which would be made by a mathematician using the SIR model and totally unaware of the existence of asymptomatic infectives with those made by a mathematician using the A-SIR model and aware of the relevance of asymptomatic infectives.

In Fig.2 we compare the predictions issued by the SIR and the A-SIR models for the cumulative number of (detected) removed, i.e. for $R(t)$, and for the number of (detected, hence symptomatic) infectives $I(t)$.



$R(t)$ $\qquad$ $I(t)$

FIG. 2: The predictions for the cumulative number of removed $R(t)$ and for the (detected) infectives $I(t)$ provided by the SIR model (blue curves) and by the A-SIR model (red curves) for a total population $N = 10^7$ and the parameters given in eq.(45).



$R(t)$ and $[R(t) + U(t)]$ $\qquad$ $I(t)$ and $[I(t) + J(t)]$

FIG. 3: The predictions for the cumulative number of removed $R(t)$ and for the (detected) infectives $I(t)$ provided by the SIR model (blue curves) and those by the A-SIR model (red curves) together with the predictions for $R(t)+U(t)$ and for $I(t) + J(t)$ provided by the A-SIR model (orange curves) for a total population $N = 10^7$ and the parameters given in eq.(45).

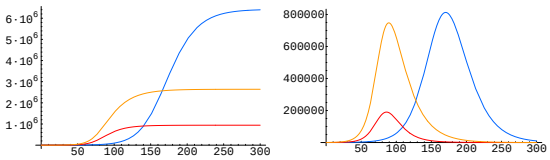
Note that not only the SIR predicts a higher infective peak, but it also expects it to occur at a later time; moreover for the initial phase, in particular after the A-SIR peak, the prediction of the number of infectives issued by the SIR model are *lower* than those issued by the A-SIR model. Note also that the SIR model expects a much greater part of the population to go through symptomatic infection, as seen from the $R(t)$ plot.

These discrepancies are reduced, and somehow reversed, if in the A-SIR modeling we consider *both* symptomatic and asymptomatic infected. These graphs are plotted in Fig.3.

In a realistic situation one should expect that at least apart of the asymptomatic infectives are anyway detected (e.g. being tested due to contacts with known infectives). We have thus plotted the equivalent of Fig.3 in Fig.4, supposing that a fraction (1/5) of asymptomatic infectives is detected. Note that in this case the height of the epidemic peak is about the same as for the SIR model, but it occurs at an earlier time, while the number of detected removed when the epidemic is over is still substantially lower (about half) of the one predicted by the SIR model.

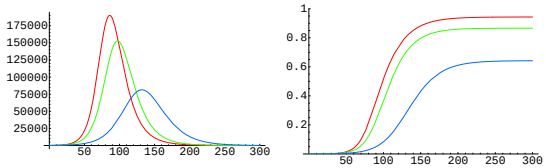
In other words, a modeler issuing his/her advice based on a SIR analysis in a situation where the A-SIR model applies, would make three substantial errors:

- The number of (symptomatic) infectives needing Hospital care would be over-estimated;
- The time before the epidemic peak – so the time



$R(t)$ and $[R(t)+\sigma U(t)]$ $I(t)$ and $[I(t)+\sigma J(t)]$

FIG. 4: The predictions for the cumulative number of removed $R(t)$ and for the (detected) infectives $I(t)$ provided by the SIR model (blue curves) and those by the A-SIR model (red curves) together with the predictions for detected removed $R(t)+\sigma U(t)$ and detected infectives $I(t)+\sigma J(t)$ provided by the A-SIR model (orange curves) if a fraction $\sigma = 1/5$ of asymptomatic infectives is detected, for a total population $N = 10^7$ and the parameters given in eq.(45).



$I(t)$ $R(t)+U(t)$

FIG. 5: The predictions for the number of symptomatic infectives $I(t)$ and for the total number of removed (and thus in the end immune) patients $R(t)+U(t)$ provided by the A-SIR model for a total population $N = 10^7$ and the parameters given in eq.(45), with $\eta = 1/20$ (red curves), $\eta = 1/15$ (green curves) and $\eta = 1/10 = \beta$ (blue curves).

available to prepare the health system to face it – would be over-estimated;

- The number of people not touched by the epidemic wave, so still in danger if a second wave arises, would be over-estimated.

It is rather clear that all of these errors would have substantial consequences. We stress that one could naively think that the presence of asymptomatic infections is essentially of help, in that a large part of the population getting in contact with the virus will not need any medical help. This is true, but on the other hand we have seen that it will also make that the curve raises more sharply than it would be expected on the basis of the SIR model (see also the discussion in Appendix A), and that in all the first phase of the epidemic – until the true epidemic peak and also for a period after this – the number of symptomatic infectives is higher than what would be foreseen in the basis of the SIR model.

We conclude that it is absolutely essential to take into account the presence of asymptomatic infectives.

V. TRACING, TESTING, TREATING

In the previous Section IV we have seen how the presence of asymptomatic infectives changes the predictions which would be done on the basis of the standard SIR model. In the present Section we want to consider a problem which naturally arises when consider an infection with a large number of asymptomatic infectives; that is, how the epidemiological dynamics is changed by an extensive campaign of *testing* (maybe *tracing* contacts of known infectives), followed of curse by *treating* – which in the case of asymptomatic means essentially just isolating the infectives so they they do not spread the infection.[59]

To this aim, we will consider again – obviously just in the context of the A-SIR model – the situation seen in the previous Section, but will take into account the effects of such a "triple T" action by a reduction of the average removal time $\tau = \eta^{-1}$ for asymptomatic infectives. (A more detailed study is given in our related paper [44].)

We will thus run numerical simulations with the same parameters (except for η), total population and initial conditions as in Section IV above, and choose three different values for $\eta \leq \beta$, i.e. $\beta = 1/20, 1/15, 1/10$. The result of these simulations is displayed in Fig.5.

We see from these that a "Tracing, Testing, Treating" campaign can reduce substantially the epidemic peak[60]; on the other hand this leads to a slowing down of the epidemics, and to a smaller number of total infected individuals, i.e. to a smaller immune population at the end of the epidemic wave – and hence smaller resistance in the case of a second wave.

This concludes our general discussion. In the following, we will apply our formalism to the analysis of the COVID-19 epidemics in Italy.

VI. THE COVID-19 EPIDEMICS IN ITALY

In the previous Sections, we have conducted our analysis in general terms. This referred to a generic infection with a large number of asymptomatic infectives, but of course the motivation for it was provided by the ongoing COVID-19 pandemic.

We will study in particular the situation in Italy; as well known, this was the first European country heavily struck by COVID, and data for it are widely available. Moreover, when COVID landed in Europe there was already an alert, so that transient phenomena due to late recognition of the problem – which were unavoidable in Eastern Asia – were of a much smaller size. Last but not least, in this case we are confident to have all the relevant information.

A. Epidemiological data

The data for the cumulative number of registered infected communicated by the Italian Health System [41] – and also available through the WHO reports [40] – are given in Table I. There we give data for February 21 to

May 15, albeit we will for the moment only use those for the first half of March; later data will be of use in the following. (We prefer not to use the data for the first few days, as too many spurious effects an affect these.)

One should note, in this respect, that the first cases in Italy (apart from sporadic and promptly isolated cases of tourists) were discovered on February 21. The public awareness campaign started immediately – and actually the public was already alert, e.g. it was impossible to find face masks since some weeks – and the first local mild restrictive measures were taken a few days later (February 24), and more restrictive measures involving the most affected areas[61] were taken on March 1. A more stringent set of measures went into effect for the whole nation on March 8, and still more strict measures on March 22.

Thus the epidemics developed with varying parameters. Moreover, as the incubation time for COVID ranges from 2 to 10 days, with a mean time of 5.1 days [20], there is a notable delay in the effect of any measure. Note that most countries have a quarantine length of 14 days; thus we expect that any measure will show their effect with a delay of one to two weeks.

B. Fit of the data

For our fitting in the early phase, we will consider the data of the period March 1 through March 10, denoted in the following as t_i and t_f respectively; this leaves us some later days to compare the functions obtained through the fit with subsequent evolution, before the effect of the first set of measure can show up (March 15).

The best direct fit of $R(t)$ through a quadratic function, see (24), is obtained with the constants

$$r_0 = 3862.32\,, \quad r_1 = 966.54\,, \quad r_2 = 80.35\,. \tag{46}$$

The fit is reasonably good in the considered time interval (t_i, t_f), but fails completely for $t < t_i$ (see Figure 6, where we use data from February 24 on) and is rather poor for $t > t_f$. This is not surprising, as we know that $R(t)$ is, in this early phase, growing through a slightly corrected exponential law, see (22).[62]

Let us then look at the fit of $\log[R(t)]$ as in (27). In this case the best fit is obtained with the constants

$$A = 8.26648\,, \quad B = 0.221083\,, \quad C = -0.00430354\,. \tag{47}$$

In this case the fit is very good not only within (t_i, t_f) but also outside it, at least for the period until March 15. We will thus work only with this (exponential) fit.

We will consider these values for the coefficients $\{r_0, r_1, r_2\}$ or for the coefficients $\{A, B, C\}$ as experimental measurements.

We can now use the formulas obtained before, both for the SIR and the A-SIR model, to estimate the parameters of these models in terms of these fits following the discussion in Sections II E and III D.

day	Feb 21	Feb 22	Feb 23	Feb 24	Feb 25
R	20	77	146	229	322
day	Feb 26	Feb 27	Feb 28	Feb 29	Mar 1
R	400	650	888	1128	1694
day	Mar 2	Mar 3	Mar 4	Mar 5	Mar 6
R	1835	2502	3089	3858	4636
day	Mar 7	Mar 8	Mar 9	Mar 10	Mar 11
R	5883	7375	9172	10149	12462
day	Mar 12	Mar 13	Mar 14	Mar 15	Mar 16
R	15113	17660	21157	24747	27980
day	Mar 17	Mar 18	Mar 19	Mar 20	Mar 21
R	31506	35713	41035	47021	53578
day	Mar 22	Mar 23	Mar 24	Mar 25	Mar 26
R	59138	63927	69176	74386	80539
day	Mar 27	Mar 28	Mar 29	Mar 30	Mar 31
R	86498	92472	97689	101739	105792
day	Apr 1	Apr 2	Apr 3	Apr 4	Apr 5
R	110574	115242	119827	124632	128948
day	Apr 6	Apr 7	Apr 8	Apr 9	Apr 10
R	132547	135586	139422	143626	147577
day	Apr 11	Apr 12	Apr 13	Apr 14	Apr 15
R	152271	156363	159516	162488	165155
day	Apr 16	Apr 17	Apr 18	Apr 19	Apr 20
R	168941	172434	175925	178972	181228
day	Apr 21	Apr 22	Apr 23	Apr 24	Apr 25
R	183957	187327	189973	192994	195351
day	Apr 26	Apr 27	Apr 28	Apr 29	Apr 30
R	197675	199414	201505	203591	205463
day	May 1	May 2	May 3	May 4	May 5
R	207428	209328	210717	211938	213013
day	May 6	May 7	May 8	May 9	May 10
R	214457	215858	217185	218268	219070
day	May 11	May 12	May 13	May 14	May 15
R	219814	221216	222104	223096	223885

TABLE I: Cumulative number $R(t)$ of COVID-19 registered infect in Italy [41, 42]. In our fits for the early phase of the epidemic, $t = t_0$ corresponds to March 5 and (t_i, t_f) to the period March 1 through March 10.

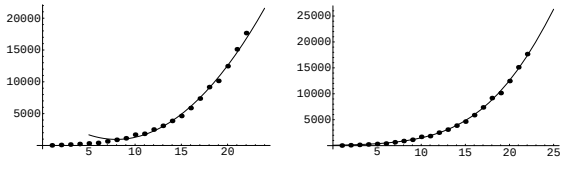


FIG. 6: Data for $R(t)$ in the COVID epidemics in Northern Italy from February 24 to March 13, with fits obtained using data for March 1 through March 10. Time is measured in days, with day zero being February 20. Left: polynomial (quadratic) fit (24); Right: corrected exponential fit (27).

QF	β	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10
	I_0	2900	3866	4833	5799	6766	7732	8699	9665
	ϕ	0.25	0.33	0.42	0.50	0.58	0.67	0.75	0.83
	S_0/γ	1.25	1.33	1.42	1.50	1.58	1.67	1.75	1.83
EF	β	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10
	I_0	2581	3441	4301	5162	6022	6882	7743	8602
	ϕ	0.55	0.73	0.91	1.09	1.28	1.46	1.64	1.82
	S_0/γ	1.55	1.73	1.91	2.09	2.28	2.46	2.64	2.82

TABLE II: Upper part (quadratic fit – QF): Parameters for the SIR and A-SIR models obtained through the SIR quadratic local fit of $R(t)$; see (26), (43). Lower part (exponential fit – EF): Parameters for the SIR and A-SIR models obtained through the SIR modified exponential local fit of $R(t)$; see (29), (44).

C. Estimate of SIR and A-SIR parameters

We have remarked in Section III D that the SIR and A-SIR models (in the latter case, under the assumption (41) for J_0) yield exactly the same values for the I_0 and γ parameters. Now we want to estimate these values for the data given in Section VI B; this amounts to a direct application of formulas (43) and (44) – or equivalently (26) and (29), as already remarked.

As already remarked, these formulas depend on the estimated value of the parameter β. The values obtained using the direct fit of R are tabulated for different values of β in the upper part of Table II.

We can also proceed by using the fit of $\log[R(t)]$; the values obtained in this way are tabulated for different values of β in the lower part of Table II.

We remind that the delay time δ from infection to arise of symptoms is estimated to be around $\delta \simeq 5.2$ [20]; thus albeit we have tabulated several options for β, values greater than 1/5 are not realistic, at least in the absence of an aggressive contact tracing and tracking campaign.

On the other hand, albeit we expect β^{-1} to be around one week for *symptomatic* patients, it should be recalled that in the standard SIR framework this parameter refers to the average removal time for *all* patients, symptomatic and asymptomatic. Thus the presence of asymptomatic ones could make β^{-1} to be substantially larger. We will leave this remark aside for the time being; see the discussion in the Appendix, in particular Remark A1.

In all cases, S_0/γ is quite far from one and ϕ from zero, so one can not rely on the "small epidemic" formulas [27] discussed in Section II C. We will thus resort to numerical integration, see next Section.

This concludes our estimation of the SIR or A-SIR parameters from the epidemiological data referring to the early phase of the epidemic.

VII. NUMERICAL SIMULATIONS. TIMESCALE OF THE EPIDEMIC

We have now determined the parameters for both the SIR and the A-SIR model in the very early phase of the epidemic; this determination depends actually on the value of β, so from now on we will consider β as the only parameter to be fitted.

We can now run numerical simulations with the SIR and the A-SIR equations; we can vary β (and thus implicitly also the other parameters, maintaining the relation between these and β) and fit β according to the agreement between the outcome of the numerical simulations and the epidemiological data.

In all of our simulations, day one is February 21, so the fitting period (t_i, t_f) corresponding to the first decade of March used to determine the relation between β and the other parameters is centered around day 14; but this does not determine β itself. We will now examine the fit of data in the first period outside this window in order to determine β.

A. General study

Note that while the discussion of the previous Section VI is complete for what concerns the SIR parameters, in the A-SIR equation we also need to introduce the removal rate for asymptomatic individuals, i.e. η; this is related to the time length $\delta = \eta^{-1}$ of their infective period, which is equal to the incubation time plus the spontaneous healing time[63]. While the former is around $\delta \simeq 5$, the latter is generally considered to be around 14 days, albeit we know that for hospitalized patients (which however are by definition symptomatic) this may be longer. Thus, on these medical grounds, we expect η^{-1} to be between 14 and 21 days.

1. *Determining β*

We ran a number of simulations, both for the SIR and the A-SIR dynamics, with varying β and – for the A-SIR model – with varying η; the parameter β was varied in the range $(1/10 - 1/5)$, while η in the range $(1/25, 1/10)$.

The best results in terms of agreement with epidemiological data for the period March 1 through March 15 were obtained for

$$\beta \;=\; \frac{1}{7}\;, \quad \eta \;=\; \frac{1}{21}\;.$$

The plot for these values of the parameters are displayed in Figs.7 and 8; note that for the SIR case the fit is rather poor.

In fact, it should be stressed that the situation is quite different in the cases of SIR and of A-SIR dynamics. In particular, the SIR dynamics does not fit the data for the

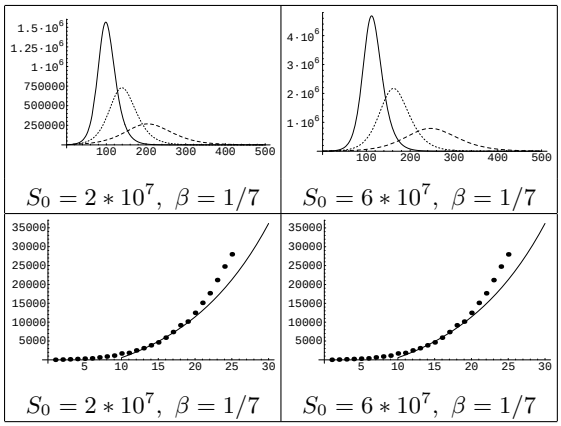


FIG. 7: Upper row: numerical solution of the SIR equations for $\beta = 1/7$ and total population $S_0 = 2 * 10^7$ (left) and $S_0 = 6 * 10^7$ (right), corresponding to the overall population of the three Northern Italy regions most affected by COVID and to the population of all of Italy respectively, using for the parameters I_0 and γ the fit of eqs. (43), (44) on the basis of the data for March 1 through March 10, see Table I. The plots of $I(t)$ – where t is measured in days – are shown for: $r = 1$ (solid curve), $r = 0.85$ (dotted curve) and $r = 0.75$ (dashed curve). Lower row: plot of the data for the COVID epidemics in Italy for March 1 through March 17 (hence outside the fitting region) versus the numerical integration of the SIR model (with $r = 1$). The SIR model is not properly describing the dynamics.

week after the fitting window, while the A-SIR dynamics fits these quite well.

We stress that the fit $\beta \simeq 1/7$ is coherent with medical data: in fact, as already recalled, the average delay from infection to the first symptoms is about 5 days, and there should be some further delay for recognizing these symptoms – which in the first phase can be rather trivial, such as cold, cough, or light fever – as due to COVID and leading to hospitalization or anyway to isolation.[64]

Remark 16. A relevant point should be made here. In the first phase of the epidemic, and in particular in the period covered by Figs.7 and 8, the COVID diffusion was essentially limited to Northern Italy, and in particular to the three regions of Lombardia, Veneto and Emilia-Romagna; the total population of Italy is about 60 millions, while that of the three mentioned regions totals about 20 millions. This is why simulations with both $S_0 = 2 * 10^7$ and $S_0 = 6 * 10^7$ are shown in these Figures. Actually, the comparison with the epidemiological data *in the early stage* should be better done with the simulation with $S_0 = 2 * 10^7$ (we are giving the other one for the SIR model only to show that even considering the overall population we do not get a reasonable fit). ⊙

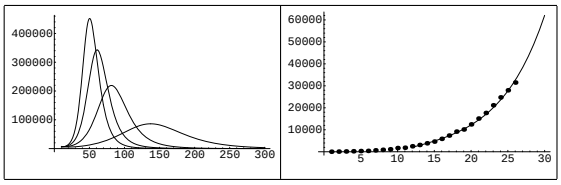


FIG. 8: Numerical solution of the A-SIR equations for $S_0 = 2 * 10^7$, $\beta = 1/7$, $\eta = 1/21$, $\xi = 1/10$, using for the parameters I_0 and γ the fit of eqs. (43), (44) on the basis of the data of Table I for the period 1-10 March. Left: plots of $I(t)$ – where t is measured in days – are shown for: $r = 1$, $r = 0.8$, $r = 0.6$ and $r = 0.4$ (the curves for higher r are those with higher peak). Right: Plot of the data for the COVID epidemics in Italy versus the numerical integration of the A-SIR model (with $r = 1$); plotted data go until March 17.

2. *Timescale of the epidemic*

In Section IV above we have discussed the different timescale of an epidemic with given initial time series predicted by the SIR and by the A-SIR models. It is natural to wonder how long the COVID epidemic will last according to these.

This is not a well posed question, because there are restrictive measures being taken (or relaxed at a later stage, when the situation improves) which reduce the contact rate and thus the spread of the epidemic.

So, these simulations can at their best show what would be the behavior (of the system described by the SIR or A-SIR equations, which do not necessarily describe correctly the COVID epidemics) *with constant coefficients*. On the other hand, they can give an idea of what should be expected in case of no action.

More generally, and in line with our general discussion of Sections IV and V, they show how the dynamics predicted by the SIR and the A-SIR models with coherent sets of parameters differ from each other in general (albeit using the concrete COVID parameters to fix ideas).

It should be stressed in this context that the containment measures based on *social distancing* do not act on β, but on α; albeit in general $\alpha = \beta/\gamma$, in studying the effect of restrictive measures it is more convenient to consider the reduction factor r. That is, if the fit of the initial phase of the epidemic yields $\alpha_0 = \beta_0/\gamma_0$ (where γ_0 is determined through the formulas of Sections II E and III D), we consider in later phases a contact rate

$$\alpha \ = \ r\ \alpha_0\ , \quad 0 < r < 1\ . \tag{48}$$

We will see that after the first set of measures in Italy we got $r \simeq 0.5$ (compared with the initial period of the epidemic [45]), and after the second set of measures this went to $r \simeq 0.2$, and to $r \simeq 0.08$ at a still later stage.

Remark 17. It should be noted however that the studies suggesting the presence of a *very* large fraction of undetected infections [23] are also implicitly suggesting that the reduction in the epidemic growth could be due not

only – or not so much – to the containment measures, but rather to the intrinsic dynamics of the system, and on the rapid depletion of the susceptibles reservoir. This point should be carefully considered. We find that in the case of COVID-19 in Italy our model suggests $\xi = 0.1$, and discuss this point in view of that estimate. In order to fully explain the decrease of the contact rate by this mechanism – thus essentially to the often mentioned "herd immunity" – however, one would need $\xi \simeq 0.01$, which appears to be non realistic (on the other hand, the suggestion $\xi = 0.02$ [24] is still very low but cannot be discarded *apriori*) . ⊙

Remark 18. The effect of different types of measures, let us say those acting on social distance and hence on α on the one hand, and those acting on early isolation and hence on β (and on η for the A-SIR model) on the other hand, on the epidemic timescale is specially relevant in view of the economic and social costs of a generalized lockdown as the one adopted in many European countries. This is discussed in detail in related publications [46–48], both for the SIR and the A-SIR models. ⊙

However, the real concrete interest of this study is in a different point. That is, there is considerable debate on the most appropriate way to use laboratory exams, and in particular if there should be a generalized COVID testing, at least of those having had contacts with known infects, or if only clinically suspect cases should be tested. We are of course aware that the real obstacle to a generalized testing – which should however be repeated over and over to be sure the individual has not been infected since the last test – is of practical nature, as testing a population of several tens of million people (not to say about China or India) is unfeasible, so that this alternative is a concrete one only in small communities; these could be isolated areas or also e.g. the community of people working in a Hospital.[65]

On the other hand, a strategy aiming at generalized testing of those who have been in contact with known infectives is more feasible; actually this strategy was conducted with remarkable success in at least one Italian region, i.e. Veneto.

In any case, we want to study what the impact of reducing $\tau = \eta^{-1}$, thus raising η, would be on the development of the A-SIR dynamics. This is illustrated in Tables III and IV below.

B. More detailed study with selected parameters

As mentioned above, our numerical simulations suggest that the epidemic in Northern Italy is – or more precisely, was before the restrictive measures of March 8 went into operation – better described, in terms of our model, by the situation with

$$\beta = 1/7 \ , \quad \eta = 1/21 \ ; \quad \xi = 1/10 \ . \tag{49}$$

We will thus devote further analysis to this setting.[66]

When discussing if and how we can change these parameters, it is essential to state clearly what the two classes I and J (and hence also R and U) represent when we act on the system.

What we mean here is that in the "natural" situations J represents the class of asymptomatic *and hence* undetected infectives[67]; on the other hand, when we start chasing for asymptomatic infectives these two characteristics are *not* equivalent.

Here we will understand that J represents *asymptomatic* infectives, detected or undetected as they are.

It should be stressed that – at least in this framework, see below – these parameters cannot be altered: indeed, ξ depends on the interaction of the virus with human bodies and is thus fixed by Nature, while the removal time $\beta^{-1} \simeq 7$ can hardly be compressed considering that typically the first symptoms arise after 5 days, but these are usually weak and thus receive attention (especially in a difficult situation like the present one) only after some time.

Remark 19. We said above that the parameters β and ξ can not be altered "in this framework"; this requires a brief explanation. As for ξ this statement is rather clear[68], and is connected to considering J as the class of asymptomatic infectives, while if they were meant to be the class of registered infectives, ξ could be altered by a large scale campaign of tests on asymptomatic population. On the other hand, for what concerns β the impossibility of compressing the time from infection to isolation is "in this framework" in the sense that in it isolation depends on the display of symptoms (plus some unavoidable reaction delay). The time $\delta = \beta^{-1}$ could instead be compressed by a campaign of *tracing contacts* of known (symptomatic or asymptomatic) infectives, as done in the field at Vò Euganeo [19]. These points will not be discussed here, but will be considered in a related paper [44]. ⊙

On the other hand, it is conceivable that $\tau = \eta^{-1}$ could somehow be compressed if a general screening was conducted. At the same time, the contact rate α can be reduced by a more or less rigorous lockdown; in our discussion, this reduction is encoded in the reduction parameter r, which yields the ratio of the achieved contact rate over the "natural" one – i.e. the one measured at the beginning of the epidemic.

We have thus ran several numerical simulations for these values of β and ξ, both for a total population of $N = 2 * 10^7$ (the total population of the initially most affected regions) and for $N = 6 * 10^7$ (the total population of Italy).

These give of course very similar results – if referred to the total population – as our estimates for the parameters, and in particular for the one leading the dynamics, i.e. γ, depend themselves on S_0.

Moreover, the questions discussed in this subsection do not concern the early phase of the epidemics (which was

	r	I_*	t_*	R_∞/S_0	U_∞/S_0	S_∞/S_0
$\xi = 1/10$	1.0	$1.3 * 10^6$	57	0.10	0.89	0.01
$\eta = 1/21$	0.8	$1.0 * 10^6$	70	0.10	0.87	0.03
	0.6	$6.5 * 10^5$	95	0.09	0.82	0.09
	0.4	$2.5 * 10^5$	167	0.07	0.65	0.28

TABLE III: Simulations for the A-SIR model on a population of $S_0 = 6 * 10^7$, with $\beta = 1/7$ and for the fitted initial conditions discussed in Section VI B, for $\xi = 1/10$ and $\eta = 1/21$, for various values of the reduction factor r. We report the maximum of the (registered) infectives I_*, the time t_* at which this maximum is reached, and the fraction of the initial population which passed through the infection being registered (R_∞/S_0) or unknowingly (U_∞/S_0); the remaining fraction of population S_∞/S_0 remains not covered by immunity.

then limited to Northern Italy), but its future development. We will thus present the results directly for the case $N = S_0 = 6 * 10^7$.

We have investigated two questions:

(A) How a reduction in the removal time for asymptomatic infectives, i.e. in $\tau = \eta^{-1}$, would affect – according to the A-SIR model – the dynamics and the basic epidemiological outcomes of it in the regime where the epidemic is taking place;

(B) In the case r is small enough to make the population below the epidemic threshold, what are the basic epidemiological outcomes predicted by the model, again depending on various parameters including η.

The results of these numerical investigations are summarized in Table III and Table IV respectively. We have also studied, for comparison, question (B) in the framework of the standard SIR model (question (A) can not be set in this framework). The outcomes of this study are summarized in Table V.

These Tables show that a reduction of $\tau = \eta^{-1}$ can have a significant impact – more or less relevant depending on the r parameter – in the main epidemiological parameters, such as the infection peak, the epidemic time-span, and the fraction of the population which goes through infection with or without symptoms. (These results should be seen as preliminary, see the companion paper [44] for a more detailed study.)

The point we want to stress here is that even in this concrete case *the predictions of the A-SIR model*, taking into account the presence of a large set of asymptomatic infectives, *differ from those of the standard SIR model*; this difference is in some cases quite significant.

Remark 20. As mentioned above, these results are *not* a forecast of the development of the COVID epidemics, as they are obtained under the hypothesis of constant parameters, while political action will drive modification of the effective parameters in one way or the other. ⊙

	r	t_e	R_∞/S_0	U_∞/S_0
$\xi = 1/10$	0.2	539	$1.02 * 10^{-4}$	$9.17 * 10^{-3}$
$\eta = 1/21$	0.1	107	$2.67 * 10^{-4}$	$2.40 * 10^{-3}$
	0.05	66	$2.09 * 10^{-4}$	$1.88 * 10^{-3}$
	0.02	50	$1.88 * 10^{-4}$	$1.69 * 10^{-3}$
	0.01	46	$1.82 * 10^{-4}$	$1.64 * 10^{-3}$

TABLE IV: Simulations for the A-SIR model on a population of $S_0 = 6 * 10^7$, with $\beta = 1/7$ and for the fitted initial conditions discussed in Section VI B, for $\xi = 1/10$ and $\eta = 1/21$, for various values of the reduction factor r such that the population is below the epidemic threshold. We report the time t_e at which there are less than 100 known infectives, and the fraction of the initial population which passed through the infection being registered (R_∞/S_0) or unknowingly (U_∞/S_0).

r	t_e	R_∞/S_0	r	t_e	R_∞/S_0
0.20	57	$2.29 * 10^{-4}$	0.20	41	$1.77 * 10^{-4}$
0.10	49	$1.98 * 10^{-4}$	0.10	37	$1.58 * 10^{-4}$
0.05	46	$1.87 * 10^{-4}$	0.05	35	$1.51 * 10^{-4}$
0.02	44	$1.81 * 10^{-4}$	0.02	34	$1.47 * 10^{-4}$
0.01	44	$1.79 * 10^{-4}$	0.01	34	$1.46 * 10^{-4}$

TABLE V: Simulations for the standard SIR model on a population of $S_0 = 6 * 10^7$, with $\beta = 1/7$ (left hand side) and – for comparison – also for $\beta = 1/5$ (right hand side), and for the fitted initial conditions discussed in Section VI B, for various values of the reduction factor r such that the population is below the epidemic threshold. We report the time t_e at which there are less than 100 known infectives, and the fraction of the initial population which passed through the infection (R_∞/S_0).

VIII. COVID-19 IN ITALY AND MITIGATION MEASURES

Our study in the previous Sections went until March 17 (see Figs.7 and 8); this is the time interval in which it makes sense to consider *constant* parameters, as the first set of governmental measures did not show its effect yet, and we got a good fit of epidemiological data by our A-SIR model.

A. Mitigation measures and reduction factors

After (or shortly before) this date, the first set of restrictive measures, taken on March 8 and gradually enforced, should have shown their effect; thus – in terms of the model – the parameter $\alpha = \alpha_0$ changed to a different value $\alpha = \alpha_1 = r_1\alpha_0$. Similarly, more restrictive measures took effect on March 23 (this time in a sharper way), and this should have shown their effect about one week later, thus changing again the parameter α, say taking it to be $\alpha = \alpha_2 = r_2\alpha_0$. It should be mentioned that albeit no official change took place, it was generally remarked that a different attitude in enforcing the mea-

sures appeared after Easter (April 11); at about the same time individual protective device became easily available. These changes are somehow reflected in our fitting, see below, by the introduction of a different reduction parameter r_3 after April 25.

Thus in considering if our model can describe the actual development of the epidemic in Italy, one should take into account these changes of parameters, i.e. fit also the constants r_1, r_2 and r_3 introduced above. It should be stressed that we do not have an analytical formula involving some parameters which can then be fitted against experimental data; we are instead studying numerical solution of the A-SIR dynamical system (30) and checking how this fits the data. On the other hand, we are assuming that nothing changes for the β, η and ξ parameters.

Data until May 15 are given in Table I. We found a good agreement keeping the value $\xi = 1/10$, and setting

$$r_1 \;=\; 0.50\;, \quad r_2 \;=\; 0.20\;, \quad r_3 \;=\; 0.08\;. \tag{50}$$

In other words, we have set (recall time is measured in days, day 1 being February 21)

$$\alpha \;=\; \alpha(t) \;=\; \begin{cases} 1.00 * \alpha_0 & t \;\leq\; 25 \;\;, \\ 0.50 * \alpha_0 & 25 < t \leq 35 \;\;, \\ 0.20 * \alpha_0 & 35 < t \leq 63 \;\;, \\ 0.08 * \alpha_0 & 63 < t \;\;. \end{cases} \tag{51}$$

Remark 21. Not that here we are considering the full Italian population, i.e. $S_0 = 6 * 10^7$; this introduces a factor 3 compared with the setting used in Sections VI and VII to fit α_0 on the basis of the early data which concerned essentially three regions with a total population of about $2 * 10^7$; we have correspondingly to divide the α_0 determined in there by a factor three, thus getting the value

$$\alpha_0 \;=\; 3.77 * 10^{-9}\;; \tag{52}$$

this is to be used in (51) above. ⊙

Proceeding in this way, we obtain the curve plotted in Fig.9 against the epidemiological data; this curve shows a good agreement with the data. See also Fig.10 for a different representation in terms of the infectives $I(t)$ (in one of the panels we have adopted a smoothing procedure consisting in mobile average over five days).

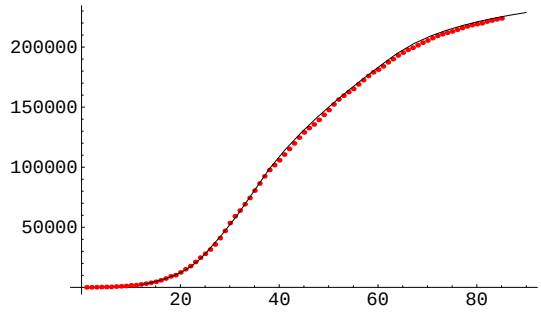


FIG. 9: Solution of the A-SIR equation, taking into account the changes in the contact rate α determined by governmental measures, against epidemiological data for Italy. See text.

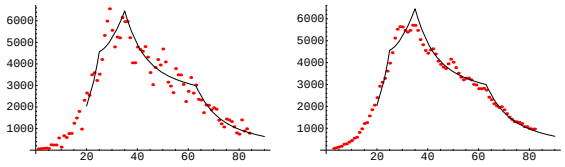


FIG. 10: Data for $I(t)$ as obtained from the model equation, i.e. as $\beta I(t) = R'(t)$, against data for the $R(t)$ increment over one day. Left: raw data; Right: data smoothed by averaging over five days, centered at plotted day. We are actually plotting daily increments of R, i.e. the A-SIR estimate for $\beta I(t)$.

B. COVID-19, Italy, and mitigation measures: a brief discussion

We give here a brief discussion of several aspects of the application of our theory to the specific situation of the COVID-19 epidemics in Italy, and of the results obtained above. This also involve non-mathematical matters. Moreover, the points touched upon here are not applying to the general theory, so we have preferred to keep this discussion separate from the short general one given in the next Section IX.

First of all, we note that the model allows to give an estimate also for the lockdown duration (assuming no further measures or relaxation of measures is adopted in the meanwhile, which is of course improbable in practical terms). E.g., if the lockdown should go on until the level of registered infectives present when the first measures were adopted on day 17, i.e. $I(t) \simeq 7,000$, then according to the model and disregarding the third reduction, it should have gone on until day 129, i.e. the end of June. Taking into account also the third reduction in α, it should have gone until day 79, so May 19, in quite good agreement with the actual choice of the Italian Government to reduce limitations starting on May 18. We stress that according to our model and more generally to SIR-type models, a strategy based also on *tracing contacts* and *early detection* is much more effective (also in stopping the epidemic, as shown in the field by the Padua

team [19], besides in keeping restrictions within an affordable time), as discussed in a companion paper [44]; see also [46–48].

Estimating the number of asymptomatic infectives is of course relevant in choosing measures to counteract the COVID-19 epidemics; as far as we know this is the first estimate based on a theoretical model and not just on statistical considerations. It follows from our model that the fraction x of symptomatic infectives – and the fraction $y = 1 - x$ of asymptomatic ones – are *dynamical* variables, and change with time depending on other features of the system, in particular its total population. Thus a purely statistical evaluation appears to be *necessarily misleading*. See also Fig.11 in this regard.

The previous observation is specially significant when we look at the tail of the epidemic. In fact, while in this limit the ratio of removed infectives $U(t)/R(t)$ goes to the natural limit $(1-\xi)/\xi$, the ratio of *active* infectives $J(t)/I(t)$ goes to a much larger number, see Fig.11. This means that when restrictions are removed, it is absolutely essential to be able to track asymptomatic infectives, to avoid these can spark new fires of infection.

It should also be mentioned that our analysis suggests that the restrictive measures adopted in Italy were quite successful in reducing the contact rate α. This is in a way surprising, given their mild nature if compared to those adopted in China. A couple of facts should be recalled in this respect: (a) on the one hand a large part (possibly larger than in other European countries with a similar demographic profile) of infections and casualties is related to hospitals or senior citizens residences, so that once these were put under control the transmission rate was substantially lowered; the latter are much more common in the Northern regions; (b) a reduction on α has an impact on the (only) quadratic term in the SIR or A-SIR equations; thus a reduction in mobility by a factor μ could result in a reduction by a factor μ^2 for the contact rate α.[69] $\odot$

Finally, we note that while the first two steps in the α reduction are temporally correlated to the restrictive measures adopted by the Government and are thus thought to be due to these, the third step cannot be explained in such a way; several tentative explanations can be put forward (but with available data not scientifically tested) for this. These include relaxation of "stay at home" campaign (home is second only to senior citizens residence for frequency of contagion), availability and thus generalized use of individual protection device, increased solar UV radiation, and reduction of ambient viral load [50]. No final word on this matter can of course be said in terms of mathematical models alone.

IX. CONCLUSIONS

Motivated by the peculiar features of the COVID epidemics, we have considered a SIR-type model, called A-SIR model, taking into account the presence of a large set of *asymptomatic infectives*.

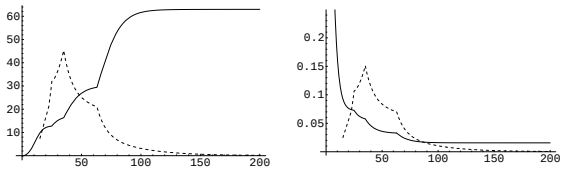


FIG. 11: The balance between symptomatic and asymptomatic infectives. Left: the ratio $J(t)/I(t)$ (solid) and $I(t)$ (dashed, different scale) on the numerical solution to A-SIR equations. Right: the ratio $x(t) = I(t)/[I(t) + J(t)]$ (solid) and $I(t)$ (dashed, different scale) on the numerical solution to A-SIR equations.

We have shown that the dynamics of the SIR and the A-SIR models, for parameters fitted from the *same set of data* available in the early phase of an epidemic, differ significantly; this is not surprising, but is in itself a significant conclusion when we have to deal with a concrete epidemics with this characteristic. This is our first result.

We have analyzed the available data for the COVID-19 epidemics in Northern Italy in terms of the SIR and of the A-SIR models; in particular we have fitted the model parameters based on the period 1-10 March, and considered how these models with such parameters are performing in predicting the evolution for the subsequent week, 11-17 March. As shown by Figure 7 on the one hand, and by Figure 8 on the other hand, it appears that the A-SIR model is much better in predicting such (admittedly short time) evolution. In particular, this is the case with the estimate $\xi = 1/10$ for the ratio of clearly symptomatic versus total infections (this is slightly smaller that the Li *et al.* [22] estimate $\xi = 1/7$); and with the reasonable estimates $\beta^{-1} = 7$ days for the time from infection to isolation for symptomatic infectives, and $\eta^{-1} = 21$ for the time from infection to healing of asymptomatic infectives.

We have studied in more detail the case which best fits the epidemiological data outside the period used to fix the model parameters. This study was conducted in two directions.

(A) On the one hand, we have considered what would be the effect of a reduction of the time spent by asymptomatic being infective and non-isolated. In this framework, two cases are possible: either the restrictive measures are only mitigating the epidemic, or they are capable of stopping it by raising the epidemic threshold above the population level. In the first case, a reduction of η^{-1} from 21 to 14 days produce a substantial lowering of the epidemic peak and also substantially postpones its occurrence; in the second case, the effect of such a reduction may be quite relevant if the population remains just under the threshold (see the case with $r = 0.2$ in Tables IV), or not so relevant if the reduction of the contact rate is taking the epidemic threshold well above the population level (see the cases with lower r in Table IV).

In all cases, *there is a marked difference with the behav-*

ior of a standard SIR model with equivalent parameters.

(B) On the other hand, we have then considered how our model can describe the COVID-19 epidemic dynamic in Italy outside the time window used to fit the parameters. In this context, restrictive measures adopted in different stages have reduced the epidemic development, hence also altered the parameters – in particular, as the measures were essentially based on *social distancing*, the contact rate α. We have seen that assuming the measures showed their effect after one week (in line with our estimate $\beta^{-1} \simeq 7$) there is an estimate of their effect on the contact rate α which produces *a good agreement between the model and the data*.

This agreement depends on the chosen value for ξ, i.e. the ratio between symptomatic infections and total infections, and is quite good for $\xi \simeq 1/10$ (confirming the early estimate by Vallance and Whitty [21]). It was stressed that this estimate of ξ is not of statistical nature, but follows instead from a theoretical model; this is specially relevant in that the same model shows that albeit the probability of symptomatic infection ξ is a constant, the different times for which symptomatic and asymptomatic infectives do take part in the epidemic dynamic makes that the fraction x of symptomatic over total infectives is a dynamical variable, and changes – even substantially – with time.[70]

We trust that our work shows convincingly the need to take into account the presence of asymptomatic infectives – and the longer time they spend before going out of the infective dynamics – when they are a substantial number; see also the Appendix and [51].

The ongoing COVID-19 epidemic taught us that there can be relevant epidemics with a large number of asymptomatic infectives; more detailed models can surely be cast, but the very simple model presented here is already sufficient to show that the standard SIR model is not a good guide in this case, as it leads to overestimate certain very relevant parameters and underestimate others.

A more detailed study of how different strategies to mitigate the epidemic may affect the A-SIR dynamics is presented in a related paper [44]; see also [47, 48].

Acknowledgements

I thank L. Peliti (SMRI), M. Cadoni (Cagliari) and E. Franco (Roma) for useful discussions. The paper was prepared over a (locked-in) stay at SMRI. The author is also a member of GNFM-INdAM.

APPENDIX A: A-SIR VERSUS SIR DYNAMICS, AND R_0 ESTIMATES

The *basic reproduction number* (BRN) is a popular indicator for the speed of diffusion of an epidemic in its first stages [35–37]. This is usually denoted as R_0, but which here will be denoted as ρ_0 (to avoid any confusion with the value of $R(t)$ at $t = 0$).

The BRN represents the average number of infections caused by a single infective; we want to discuss how the estimate of this can be flawed if one works with the SIR formalism in a situation where a large number of asymptomatic infectives is present, and hence better described by the A-SIR model; our discussion reproduces that in [51], and is given here for the sake of completeness.

1. The reproduction number and the SIR framework

The extraction of the BRN from epidemiological data can be tricky, but its evaluation in terms of a model is much simpler. In fact, in this case $S(t)$ can be considered, for small t, as constant, i.e. $S(t) \approx S_0$. The SIR equations reduce to linear ones, and in the very beginning (before any infective can be removed) the new infectives grow as $dI/dt = \alpha S_0 I$. Thus any infective produces $\alpha S_0(\delta t)$ new infectives in a short time length δt. As each infective is active, on the average, for a time β^{-1}, this yields a first rough estimate

$$\rho_0 \;\simeq\; (\alpha/\beta)\, S_0 \;=\; S_0/\gamma \;. \tag{A1}$$

Obviously, as β^{-1} is not so small (of the order of days), we should look more carefully at the equation for $I(t)$, which can be solved and yields $I(t) = \exp\left[\alpha S_0 t\right] I(0)$. As each infective is active for an average time $\delta = \beta^{-1}$, this just reads

$$I(\delta) \;=\; \exp\left[S_0/ga\right]\; I(0) \;. \tag{A2}$$

The situation is slightly different if we take into account also the removal of infectives. With the same approximation $S \approx S_0$ but with the full (1) equations, we get $dI/dt = \alpha(S_0 - \gamma)I$ and hence

$$I(t) \;\approx\; \exp[\alpha(S_0 - \gamma)\, t]\; I_0 \;:=\; H(t)\, I_0 \;.$$

Considering again that an infective remains active for a time $\delta = \beta^{-1}$ on average, the previous equation means that (on average) there will be $\rho_0 = H(\delta)$ new infectives originating from a single one. From the previous expression for $H(t)$ we easily get $\rho_0 = \exp[(S_0/\gamma) - 1]$.

If we are facing a pathogen for which there is no natural immunity in the population (e.g. a new virus), $S_0 = N$ (the total population) and we get

$$\rho_0 \;=\; \exp\left[\frac{N}{\gamma} \;-\; 1\right] \;. \tag{A3}$$

Thus, evaluating R_0 is in this case immediate if we know γ; and conversely if we are able to evaluate ρ_0 by epidemiological data, then it is immediate to evaluate γ as

$$\gamma \;=\; \frac{N}{1 + \log(\rho_0)} \tag{A4}$$

(or a simpler formula if we adopt the earlier and simpler definitions for ρ_0, e.g. (A1) above).

The parameter ρ_0, and more generally $\rho(t)$, met a great favor in the general press, as it gives the impression to describe a complex phenomenon by a single number. As we mentioned above, evaluating $\rho(t)$ from epidemiological data can be a tricky matter [35, 36], in particular if this is attempted without resorting to a specific model.

In this case too, using the SIR model without taking into account the presence of a large set of asymptomatic infectives leads to a wrong estimation of ρ, as we will see in a moment.

2. The reproduction number and the A-SIR framework

With the A-SIR equations (30), the number of new infected per unit of time is $\alpha S(I+J)$, and in the early phase of the epidemic we can assume $S \approx S_0$, and infectives will grow as

$$K(t+\delta t) \;=\; K(t)\;\exp[\alpha\; S_0\;\delta t]\;.$$

Thus each infective will give origin in the time span δt to $\exp[\alpha S_0 \delta t]$ new infectives. In order to know the BRN, we we should look at the *average infective time* τ in the early phase of the epidemic, and choose $\delta t = \tau$. At the beginning of the epidemic the ratio between registered and total infectives is simply

$$x_0 \;:=\; \frac{I_0}{I_0+J_0} \;\simeq\; \xi\;, \tag{A5}$$

while – as discussed above – in later stages the proportion between I and J changes, as individuals stay longer in the J class than in the I class.

The average removal rate for $t \simeq 0$ is thus

$$B_0 \;=\; \xi\,\beta \;+\; (1-\xi)\,\eta\;. \tag{A6}$$

This means that each (symptomatic or asymptomatic) infective individual will give direct origin, across its infective and non-isolation period, not to $\rho_0 = \exp[\alpha S_0/\beta] = \exp[S_0/\gamma]$ new infectives, but instead to

$$\widehat{\rho}_0 \;=\; \exp[(\alpha/B)S_0] \;=\; (\rho_0)^{\beta/B} \tag{A7}$$

new infectives. As $\beta > B$, this means that the actual basic reproduction number $\widehat{\rho}_0$ is larger – and possibly substantially larger – than the value ρ_0 which would be estimated solely on the basis of registered infections.

A trivial computation on the basis of the values given above – i.e. $\beta^{-1} \simeq 7$, $\eta^{-1} \simeq 21$, $\xi \simeq 1/10$ – provides

$$\widehat{\rho}_0 \;=\; (\rho_0)^{5/2}\;. \tag{A8}$$

Recalling that the estimates of the COVID ρ_0 on the basis of registered infectives are in the range $\rho_0 \in (2.5-3.0)$, this means that the actual BRN turns out to be instead in the range

$$\widehat{\rho}_0 \;\in\; (10-15)\;. \tag{A9}$$

This could explain why all Health Systems were surprised by the rapid growth of the number of COVID-19 infections; in fact, the presence of a large set of asymptomatic infectives was not realized when the epidemic attacked the first countries, and is becoming clearly established only now.

Remark A.1. In terms of the standard SIR model, our discussion shows that one should consider B_0 – see eq.(A6) – rather than β in computing the BRN. This also explains the poor performance of the SIR model in fitting Italian epidemiological data with "medically reasonable" estimates for β; in fact, considering the B_0 value instead of β produces a good enough fit [48]. See also Remark 12 in this context. ⊙

Remark A.2. On the other hand, it should be noted that according to Britton, Ball and Trapman [52] the herd immunity level for COVID-19 is substantially lower than the classical herd immunity level; this yields a nicer perspective for the future. ⊙

Remark A.3. Note that eq.(A8) is obtained considering the A-SIR equivalent of eq.(A4). Had we stayed with the rough estimate (A1) (as in [51]), its A-SIR equivalent would have given $\widehat{\rho}_0 = (5/2)\rho_0$, which means that the actual BRN is in the range $\widehat{\rho}_0 \in (6.25-7.5)$, i.e. still substantially higher that the SIR estimation of ρ_0. ⊙

[1] M. Gandhi, D.S. Yokoe and D.V. Havlir, "Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19" (Editorial), *N. Engl. J. Med.* 2020, doi: 10.1056/NEJMe2009758

[2] M.M. Arons *et al.*, "Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility", *N. Engl. J. Med.* 2020, doi: 10.1056/NEJMoa2008457

[3] Y. Bai *et al.*, "Presumed Asymptomatic Carrier Transmission of COVID-19", *JAMA* **323** (2020), 1406-1407

[4] H. Mason, "Pre- and asymptomatic individuals contribute up to 80% of COVID-19 transmission", *Univadis Medical News* 07/05/2020

[5] N.W. Furukawa, J.T. Brooks and J. Sobel, "Evidence supporting transmission of severe acute respiratory syndrome coronavirus 2 while presymptomatic or asymptomatic", *Emerg. Infect. Dis.* 2020, https://doi.org/10.3201/eid2607.201595

[6] Q. Li *et al.*, "Early transmission dynamics in Wuhan,

China, of novel coronavirus-infected pneumonia", *N. Engl. J. Med.* 2020; doi:10.1056/NEJMoa2001316

[7] X. Yu and R. Yang, "COVID-19 transmission through asymptomatic carriers is a challenge to containment", *Influenza Other Respir Viruses* 2020, doi: 10.1111/irv.12743

[8] S. W. Park, D.M Cornforth, J. Dushoff and J.S. Weitz, "The time scale of asymptomatic transmission affects estimates of epidemic potential in the COVID-19 outbreak", medRxiv 2020.03.09.20033514; doi: https://doi.org/10.1101/2020.03.09.20033514

[9] Y. Yang Yu *et al.*, "COVID-19 Asymptomatic Infection Estimation", doi: https://doi.org/10.1101/2020.04.19.20068072

[10] M. Day, "Covid-19: four fifths of cases are asymptomatic, China figures indicate", *BMJ* **369** (2020), m1375; doi: 10.1136/bmj.m1375

[11] T.A. Treibel *et al.*, "COVID-19: PCR screening of asymptomatic health-care workers at London hospital", *Lancet* May 08, 2020, doi: https://doi.org/10.1016/S0140-6736(20)31100-4

[12] J. Zhang, S. Tian, J. Lou and Y. Chen, "Familial cluster of COVID-19 infection from an asymptomatic", *Critical Care* **24** (2020), 119; https://doi.org/10.1186/s13054-020-2817-7

[13] G. Gronvall, "Asymptomatic spread makes COVID-19 tough to contain", (John Hopkins University COVID resource); https://hub.jhu.edu/2020/05/12/gigi-gronvall-asymptomatic-spread-covid-19-immunity-passports/

[14] R. Zhou *et al.*, "Viral dynamics in asymptomatic patients with COVID-19", *Int. J. Inf. Dis.* (2020), at press; doi: https://doi.org/10.1016/j.ijid.2020.05.030

[15] X. He, "Temporal dynamics in viral shedding and transmissibility of COVID-19", *Nature Medecine (Brief Communications)* **26** (2020), 672-675; https://doi.org/10.1038/s41591-020-0869-5

[16] https://timesofindia.indiatimes.com/india/how-asymptomatic-cases-fuelled-spread-of-coronavirus/articleshow/75292506.cms

[17] H. Nishiura *et al.*, "The rate of underascertainment of novel coronavirus (2019-nCoV) infection: estimation using Japanese passengers data on evacuation flights", *J. Clin. Med.* **9** (2020), 419; "Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19)", at press in *Int. J. Infect. Dis.* (2020)

[18] K. Mizumoto, K. Kagaya, A. Zarebski and G. Chowell, "Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020", *Euro Surveill.* **25** (2020), 2000180

[19] E. Lavezzo *et al.*, "Suppression of COVID-19 outbreak in the municipality of Vò, Italy", `medRxiv 2020.04.17.20053157`

[20] http://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/

[21] See e.g.: https://www.theguardian.com/world/2020/mar/12/uk-governments-coronavirus-advice-and-why-it-gave-it ; https://www.ft.com/content/38a81588-6508-11ea-b3f3-fe4680ea68b5 ; https://www.theguardian.com/politics/live/2020/mar/05/chief-medical-officer-chris-whitty-questioned-by-mps-live-news

[22] R. Li et al., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2), *Science* (2020) 10.1126/science.abb3221

[23] J. Lourenco *et al.*, "Fundamental principles of epidemic spread highlight the immediate need for large-scale serological survey to assess the stage of the SARS-CoV-2 epidemic", https://www.medrxiv.org/content/10.1101/2020.03.24.20042291v1

[24] S. Flaxmann *et al.*, "Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries", preprint http://www.imperial.ac.uk/ mrc-global-infectious-disease-analysis/covid-19/report-13-europe-npi-impact/

[25] W.O. Kermack and A.G. McKendrick, "Contributions to the Mathematical Theory of Epidemics", *Proc. R. Soc. Lond. A* **138** (1932), 55-83; *Proc. R. Soc. Lond. A* **141** (1933), 94-122

[26] H.W. Hethcote, "The Mathematics of Infectious Diseases", *SIAM Review* **42** (2000), 599-653

[27] J.D. Murray, *Mathematical Biology. I: An Introduction*, Springer (Berlin) 2002

[28] L. Edelstein-Keshet, *Mathematical Models in Biology*, SIAM 2005

[29] N.F. Britton, *Essential Mathematical Biology*, Springer 2003

[30] O. Diekmann and H. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases:Model Building, Analysis and Interpretation*, Wiley 2000

[31] R.M. Anderson and R.M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford UP 1992

[32] M.E.J. Newman, "Spread of epidemic disease on networks", *Phys. Rev. E* **66** (2002), 016128

[33] A. Majumdar and A.Mehta, "Heterogeneous contact networks in COVID-19 spreading: the role of social deprivation", `arXiv:2005.00491`

[34] M. Gatto *et al.*, "Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures", *PNAS* **117** (2020), 10484-10491; https://doi.org/10.1073/pnas.2004978117

[35] O. Diekmann, J.A.P. Heesterbeek and J.A.J. Metz, "On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations", *J. Math. Biol.* **28** (1990), 356-382

[36] K. Dietz, "The estimation of the basic reproduction number for infectious diseases", *Stat. Meths. Med. Res.* **2** (1993), 23-41

[37] J.A.P. Heesterbeek, "A brief history of R_0 and a recipe for its calculation", *Acta Biotheoretica* **50** (2002), 189-204

[38] The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team: "Vital Surveillances: The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) – China, 2020", *China CDC Weekly 2020, 2(8): 113-122.*

[39] https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on- covid-19-final-report.pdf

[40] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

[41] http://www.protezionecivile.gov.it/ media-comunicazione/comunicati-stampa

[42] http://www.salute.gov.it/nuovocoronavirus

[43] G. Giordano *et al.*, "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy", Nature Medicine Letters, https://doi.org/10.1038/s41591-020-0883-7 (see also

arXiv:2003.09861)
[44] G. Gaeta, "Social distancing versus early detection and contacts tracing in epidemic management", arXiv:2003.14102
[45] G. Gaeta, "Data analysis for the COVID-19 early dynamics in Northern Italy", arXiv:2003.02062; "Data Analysis for the COVID-19 early dynamics in Northern Italy. The effect of first restrictive measures", arXiv:2003.03775
[46] M. Cadoni, "How to reduce epidemic peaks keeping under control the time-span of the epidemic", arXiv:2004.02189
[47] M. Cadoni and G. Gaeta, "How long does a lockdown need to be?", arXiv:2004.11633
[48] M. Cadoni and G. Gaeta, "Size and timescale of epidemics in the SIR framework", arXiv:2004.13024
[49] https://www.google.com/covid19/mobility/
[50] M. Banerjee, A. Tokarev and V. Volpert, "Immuno-epidemiological model of two-stage epidemic growth", *Math. Model. Nat. Phenom.* **15** (2020), 27; https://doi.org/10.1051/mmnp/2020012 ; see also arXiv:2003.14152
[51] G. Gaeta, "Asymptomatic infectives and R_0 for COVID", arXiv:2003.14098
[52] T. Britton, F. Ball and P. Trapman. "The disease-induced herd immunity level for Covid-19 is substantially lower than the classical herd immunity level", arXiv:2005.03085
[53] The health agencies have studied the diffusion of influenza over many years, and this presumably allows to reconstruct some features of the diffusion network (e.g. at department level); but surely the extraordinary alarm caused by COVID emergence has modified all kinds of interpersonal relations and all usual travel patterns among countries or among regions in the same country.
[54] We stress that the parameters depend not only on the pathogen agent, but also on the social structure and organization of the country, e.g. on its population density, restrictive measures, and sanitary system; thus they are different in each country, and data for different countries can not be pooled together to have a wider statistics.
[55] Albeit in the first stage it was believed that asymptomatic were not fully infectives, now they are thought to be so [1–16].
[56] A notable example of these more detailed models for COVID is provided by [43]. In our opinion keeping to a simple setting will help focusing on the essential mechanism at hand, albeit when facing a concrete epidemic the health authorities would probably prefer to follow a less generic model – if available, and if the available data are sufficient to reliably fit all the many parameters appearing in it.
[57] Note that if we use the alternative formulation (34), setting $S \simeq S_0$ still leaves us with nonlinear equations.
[58] Which we repeat here for convenience of the reader: $R(t) \simeq r_0 + r_1 t + (1/2) r_2 t^2$, $\log[R(t)] \simeq A + Bt + (1/2)Ct^2$.
[59] The title of the Section refers to the fortunate slogan proposed in the COVID-19 context. Much before this slogan was proposed, this approached was followed (in a low-tech mode) by the Padua group to tackle the COVID-19 epidemics in the field in Vò Euganeo and then in Veneto, with remarkable success [19].
[60] More precisely, Fig.5 displays a reduction of the *symptomatic* epidemic peak – which is the relevant one in terms of charge on the health system – but the plot for the total infected would be analogous, except of course for a different scale.
[61] Due to a leak of information, a number of people fled from the most affected area before the prohibition to do so went into effect; this has most probably pushed the spreading of the infection in different regions, but luckily the Southern regions, which were expected to be heavily affected by this, in the end were only marginally touched.
[62] Note however that here the fit has not the goal to provide an analytical description of $R(t)$ for a larger interval of time, but only to estimate some parameters for the nonlinear SIR or A-SIR equations.
[63] More precisely, this would be the time until the viral load becomes too low to infect other individuals. From this point of view, different attitudes exist in different countries, also concerning isolation requirements. E.g., in Italy patients who tested positive but with little or light symptoms and are thus isolated at their home, are required to remain isolated until testing negative in *two* swab tests, while in Germany and other countries, they are automatically authorized to leave home after 14 days.
[64] This even disregarding the fact that at the peak of the epidemic the medical system was completely overwhelmed in several Departments – especially in Lombardia – and action was delayed.
[65] We recall anyway that when this was conducted in the small community of Vo' Euganeo, the study led to a change in our basic understanding of the COVID spreading [19].
[66] We have also conducted a detailed study of the case $\xi = 1/7$, as suggested by Li *et al.* [22]; but the agreement with epidemiological data is definitely less good than for $\xi = 1/10$ and hence these simulations will not be shown.
[67] This was in particular true at the beginning of the COVID epidemic in Italy, when the amplitude of the asymptomatic infections was not realized; in fact,the WHO report on China was not alerting about the relevance of the phenomenon, and this was realized only after the study of the Padua group [19]. See also [1, 4, 8–10, 13]
[68] It could depend on genetics, and thus differ in different continents, but for a given country it cannot be altered, as far as we know at the moment. Needless to say, if a medical treatment preventing the insurgence of symptom would become available, it would alter ξ but at the same time it would to a great extent solve the COVID problem.
[69] In this respect, it should be mentioned that Google has provided data analyzing the reduction of mobility in different contexts for different countries [49]. These are difficult to compare across countries, as the restrictions have not been the same in different countries; there is however a homogeneous sector, as in all countries the access to food and medicine shops remained unrestricted. In these sector, according to the Google data for March, Italy is – among the large European countries severely affected by COVID – the one where the reduction in this sector was more marked. We have indeed -85% for Italy, -76% for Spain, -62% for France, -51% for Germany, - 46% for Great Britain. Thus the reduction of α could be not only, or not so much, due to the governmental measures, but to the more prudent attitude of citizens.
[70] In this respect, it should be mentioned that recently several Groups have suggested the fraction of detected in-

fectives could be, for COVID, even smaller [23, 24]; this would by all means make even greater the differences between the dynamic predictions by a standard SIR model or by a model, like the simple A-SIR model we propose here, taking into account the peculiar feature of the presence of a large class of asymptomatic infectives.