Chaos, Percolation and the Coronavirus Spread: a two-step model.

Aldo Bonasera[1,2] & Hua Zheng[3]

[1]Cyclotron Institute, Texas A&M University, College Station, TX 77843-USA.
[2]Laboratori Nazionali del Sud-INFN, v. Santa Sofia 64, 95123 Catania, Italy.
[3] School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710119, China.

ABSTRACT

We discuss a two-step model for the rise and decay of the COVID-19. The first stage is well described by the same equation for turbulent flows and chaotic maps: a small number of infected d_0 grows exponentially to a saturation value d_∞. The typical growth time is given by $\tau = \frac{1}{\lambda}$, where λ is the Lyapunov exponent. After a time t_{crit} determined by social distancing and/or other measures, the spread decreases exponentially as for nuclear decays and non-chaotic maps. A few countries, like China, S. Korea, Italy are in this second stage while other including the USA is near the end of the growth stage. The model predicts 15,000 (±1,500) casualties for the Lombardy region (Italy) at the end of the spreading around May 10,2020. Without the quarantine, the casualties would have been more than 50,000, hundred days after the start of the epidemics. The data from the 50 US states are of very poor quality because of an extremely late response to the epidemics, resulting unfortunately in a large number of casualties, more than 70,000 on May 6,2020. S. Korea, notwithstanding the high population density (511/km^2) and the closeness to China, responded best to the epidemics with 255 deceased as of May 6,2020.

Introduction

Chaotic models have been successfully applied to a large variety of phenomena in physics, economics, medicine and other fields [1-6]. In recent papers [7,8] a model based on turbulent flows and chaotic maps has been applied to the spread of COVID-19 [9]. The model has successfully predicted the rise and saturation of the spreading in terms of probabilities, i.e. the number of infected (or deceased) persons divided by the total number of tests performed. Also a dependence on the number of cases on the population density has been suggested [7] and the different number of fatalities recorded in different countries (or regions of the same country) attributed to hospitals overcrowding [8]. In this paper we would like to extend the model to the second stage, i.e. the decrease of the number of events due to quarantine or other measures. Different fitting parameters of the model are due to the different actions, social behaviors, population densities etc. of each country but there are some features in common and it is opportune to first have a look to some data available beginning of May, 2020.

In the figure 1, we plot the number of positives (top panels) and deceased (bottom panels) as function of time in days from the beginning of the recordings. Some data have been shifted along the abscissa to demonstrate the similar behavior. Different countries are indicated in the figure insets. As we can see all the EU countries display a very similar behavior including the U.K. notwithstanding the Brexit. The USA case has been shifted of 38 days, which is the delay in the response to the epidemics resulting in the large number of fatalities. In contrast, S. Korea reacted promptly and was able to keep the number of positives and more importantly the death

rate down. Among the EU countries, Germany shows the lowest number of deceased cases, which could be due to different ways of counting (for instance performing autopsies to check for the virus like in Italy). In any case, the analysis in ref.[8] shows that different regions of Italy have lower mortality rates (for instance the Veneto region which borders the Lombardy region- the highest hit) compatible to Germany. Thus, similar to [8], we can assume that different overcrowding of health facilities, retirement homes, jails etc. might be the cause for the differences displayed in the figure 1.

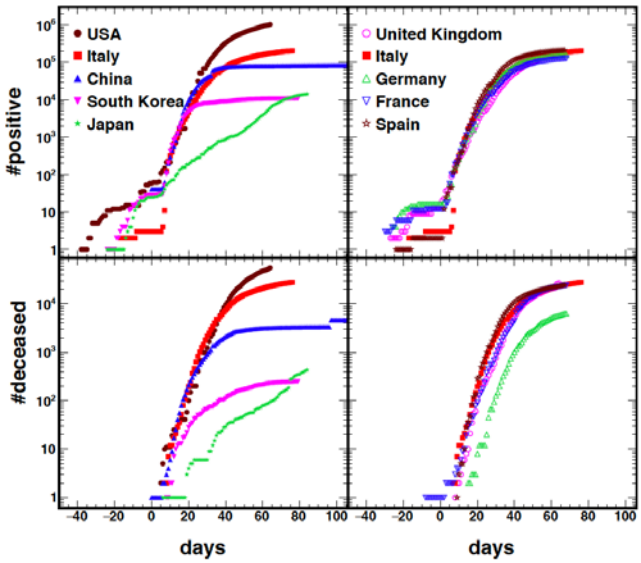

Figure 1. Number of positives (top panels) and deceased (bottom panels) as function of time for different countries indicated in the inset. Time t=0 was suitably chosen to match the exponential growth for the number of positives and it was kept the same for all the other plots, figures 1,3.

To contrast the epidemics, many countries have adopted very strict quarantine measures. Social distancing decreases the probability to remain infected thus we expect that countries with lower population density might have better and faster success. On the other hand, if some country adopts non-effective measures or it is too late in the response, the lower population density might hinder the problem for some time. Thus in order to better stress the efficacy of the quarantine, we have plotted in figure 2 the number of cases DIVIDED by the population density, assuming that it is much easier to perform social distancing if the population density is low. In the figure 2 we see that S. Korea and Japan, even though their densities are rather high, 511/km^2 and 334/km^2 respectively, perform best. We should also consider that S. Korea (or Japan) is '*across the street*' from China, the epicenter of the infection [7,8], while the other countries are located across a continent or an ocean giving further advantages to organize a response which unfortunately turned out to be weak and badly organized. The last points for China reflects an adjustment to the death rate in Wuhan, which probably had similar problems like the Lombardy region in Italy [8]: we will not be surprised to see future corrections.

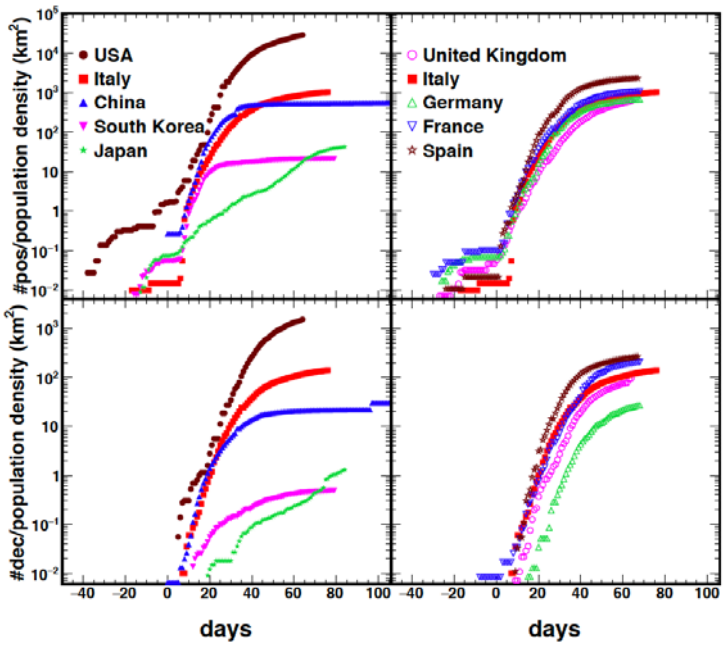

Figure 2. Number of cases divided by the population density of each country vs time, compare to figure 1.

There is a 'hidden' parameter in the figures (1) and (2): the number of tests performed daily. Zero tests, zero cases and no problem but then the hospitals get filled with sick people and we have a pandemic. In order to have realistic information on the time development of the virus, it is better to calculate the total number of cases DIVIDED by the total number of tests, this defines the probability to be infected or the death rate probability due of the virus. We stress that such probability may be biased since often the number of tests is small and administrated to people which are hospitalized or show strong signs of the virus [7,8]. The values we will derive must be considered as upper limits but the time evolution should be realistic.

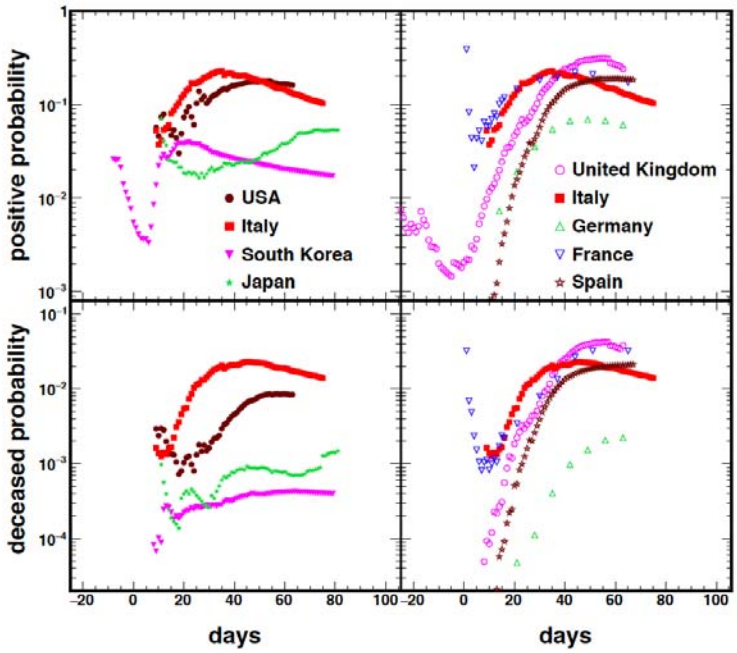

Figure 3. Probabilities vs time for the countries indicated in the inset.

Not all the countries provide the number of tests performed daily (China) or, alternatively, some provide the cumulative number of tests (Spain). In the latter case we have assumed the number of tests performed daily to be constant. In the figure 3, we plot the probabilities vs time for the same countries as in figures (1) and (2). As we can see some cases show a smooth behavior indicating prompt and meaningful data taking. Large fluctuations or missing data are also seen which means that the number of early daily tests was very small. Italy shows a decreasing behavior at long times both for the positive and deceased cases suggesting that the epidemic is getting under control. S. Korea and Japan display a similar behavior but with much lower values. The other countries have not saturated yet or are close to it and the figure 3 (and 2) suggests that the UK, France and Spain will overcome Italy, while Germany performs best among the EU countries analyzed here most importantly regarding the death rate.

The Model

We have discussed and applied the first stage of the model in refs.[7,8]. We briefly recall it and write the number of people (or the probability) positives to the virus (or deceased for the same reason) as:

$$\Pi(d) = \frac{d_0 d_\infty}{d_0 + d_\infty e^{-\lambda d}} \qquad (1).$$

In the equation, d gives the time, in days, from the starting of the epidemic, or the time from the beginning of the tests to isolate the virus. At time d=0, $\Pi(0)=d_0$ which is the very small value (or group of people) from which the infection started. In the opposite limit, $d \to \infty$, $\Pi(\infty) = d_\infty$, the final number of affected people by the virus. Equation (1) has the same form observed in the figures (1) and (2), but in reality it should be applied not the number of positives (or deceased) but to their probabilities, i.e. the number of cases divided by the total number of tests. The main reason for this definition is to avoid the spurious time dependence due to the total number of tests, which varies on a daily basis and very often not in a smooth way [7,8]. In the figure 3 we have plotted the probabilities for different countries since the data is available. It is important to stress that the information on the total number of daily tests is crucial and should be provided also to avoid suspects on data handling. If we treat equation (1) as a probability then we expect to saturate to d_∞ at time t_{crit}. At later times, if social distancing is having an effect, we expect the probability to decrease and eventually tend to zero. In the figure 4, we see exactly such a behavior for the cases of two Italian regions: Lombardy and Sardinia [8], https://github.com/pcm-dpc/COVID-19. For times larger than t_{crit} the decrease is exponential and can be described as for nuclear decays and non-chaotic maps [1,10]:

$$\Pi(d) = d_\infty e^{-\alpha(d - t_{crit})} \qquad (2).$$

α and t_{crit} are fitting parameters. Values for the Lombardy region are α=0.0268(0.025)d^{-1} and t_{crit}=39(42)d for the positives (deceased). We can infer the decay time as $\tau_d=1/\alpha$=37(40)d suggesting that roughly τ_d after the maximum the epidemics should be over, i.e $d_{max} \approx t_{crit}+\tau_d$=76(82) days from February 24,2020.

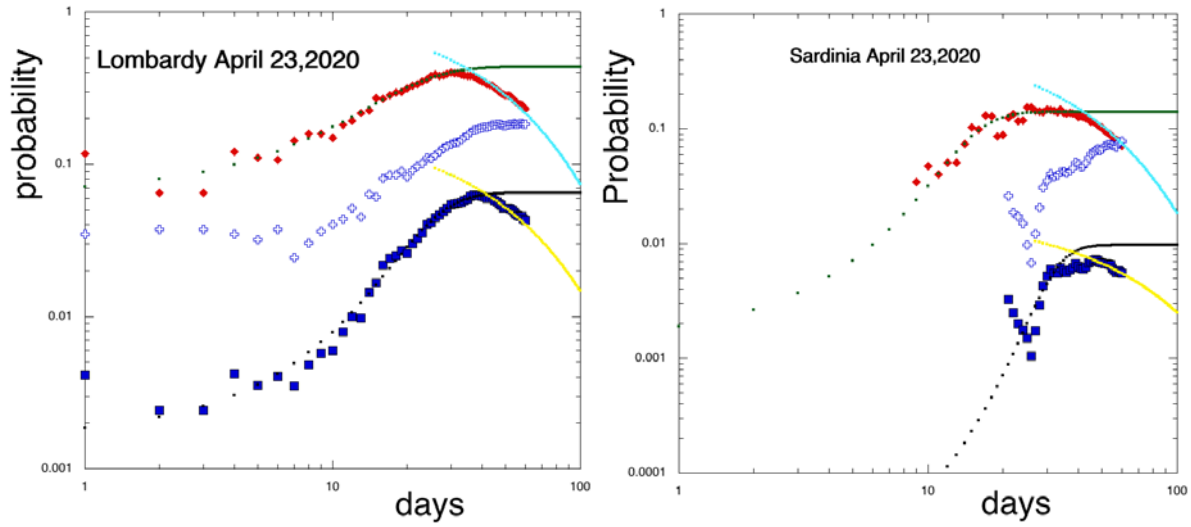

Figure 4. Probability for positives (rhomb symbols) or deceased (square symbols) vs time in days for the Lombardy (left) and Sardinia (right) regions. The open crosses give the ratio positive/deceased and reach almost 20% for Lombardy [8]. The continuous points are obtained from eq.(1) and the exponential decay from eq.(2).

From the figure 4 it is quite easy to derive the value of t_{crit} given by the maximum. This value differs slightly for the positives and the deceased as well as for the different regions. Thus it is important to have enough data to perform best fits using equations (1) and (2). For some of the countries plotted in the figure 3, the first stage only can be reproduced and we can obtain the saturation value d_∞. The value of t_{crit} depends on many factors including the population density, the weather temperature, humidity etc. and especially social distancing or any other measure used to contrast the epidemics. If no measures are adopted (herd immunization or natural selection approach), such as for some countries like Sweden and the UK at first, then we expect the plateaus in figures (3) and (4) to last longer but eventually the process will be described by equations (1) and (2). The herd immunization approach might be reasonable if we do not think we are going to get a vaccine soon. However, in such cases we may also expect to be flooded by positives and deceased persons jeopardizing the health structures and harm the sanitary personnel [8]. A country like Sweden with excellent sanitary structures and low population density (25/km^2) may succeed in this task, but the same attempt in the UK (279/km^2) was a disaster and quickly abandoned as can be seen from the figures (1-3). In particular in figure 3 we see that the UK have the largest probabilities, https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Of course the predictions have validity if the conditions are not changed, for instance relaxing the quarantine too soon. If these conditions are modified then we may have an increase of the cases again and return to the original curve given by equation (1), such a behavior might be noted in figure (3) for Japan. At the same time when the Olympics 2020 were under discussion, Japan interrupted the COVID-19 testing as can be seen from the plateaus in figures (1) and (2) lasting approximately 15 days. Thus it is important to understand when to relax the measures and for this reason we have plotted in figure 4 two cases. Lombardy is the worst case in Italy with more than 14,000 deceased in contrast to Sardinia with about 100 as of May 6,2020. In the figures we can see that the probabilities are much lower for Sardinia, which could be regarded in some sense as the future of what should eventually happen in Lombardy. The

population density of Sardinia is relatively low, 69/km^2, and it is an island away from the mainland. It is in many respects very similar to S. Korea with lower population density. Thus, the measures might be relaxed in Sardinia following the example of S. Korea after careful instructions to the population and random every day testing to search for positives to isolate them. This will provide crucial information on the social behavior and on the virus spread. We will show below that the model predicts a small number of positives and deceased for Lombardy around or after May 10,2020 thus shelter at home might be extended up to that day. It would be important to send some signals to the population of return to normality after months of sheltering by organizing for example sportive events in Sardinia. The Italian national sport, "Serie A", might organize 2-3 games per day in different Sardinian towns, with empty stadiums and broadcasted live. Other limited activities but strongly controlled could be allowed in less affected regions such as Calabria, Abruzzo and other southern Italian regions discussed in ref.[8]. Releasing all measures for the entire country at the same time might be not too wise. Looking at other countries experiences, we would suggest that quarantine should not be released before the probability for positives is less than 4% (the maximum of S. Korea, figure 3). Below such a value, the other countries may follow the S. Korean approach but if they are not organized to do that, reopening too soon may be dangerous.

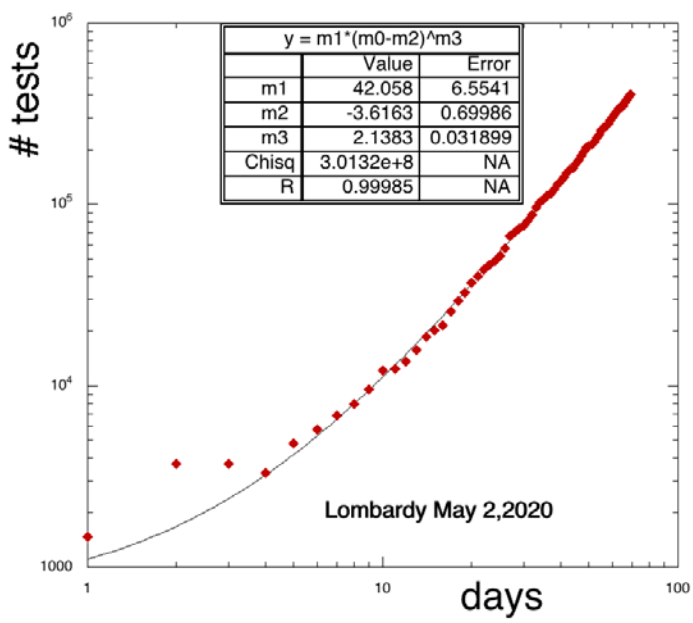

Figure 5. Total number of tests as function of time for Lombardy. The fitting function and its values are displayed in the insets.

The model describes very well the data and might be used for the everyday control on the resurgence of the epidemics. It offers another great advantage: we have described a way to eliminate misleading inputs due to the number of everyday test. We can proceed in the inverse direction in order to predict the total number of deceased and positives cases. The task that we have now is much easier and it is the prediction of the daily tests for each case. As we have seen from figures (1-3), there were some wrong decisions taken by the different countries at the beginning of the epidemics (apart S. Korea and Japan) resulting in a very small number of tests. After 1-2 weeks the number of test per day was increased and eventually become constant. It is this behavior we have to predict in order to extend our model to the total number of cases. In figure (5) we plot the total number of tests vs time in days from the beginning of the recordings

for Lombardy (February 24,2020). We have fitted the data with a power law function as indicated in the figure but any other suitable function f(d) might do as well. As we see from the figure, the Italian data is fitted very well with a small error on the fitting. Fits performed to other countries give a power exponent ranging from 0.73 (S. Korea) to 4.1 (UK). This is also an indication of how well organized the response to the pandemic is. In the ideal case we expect the power to be about 1, the value for the UK suggests some change of strategy (from i.e. herd immunization to quarantine) and because of such high value we are not able to make predictions on the total number of tests say 50 days after May 6,2020.

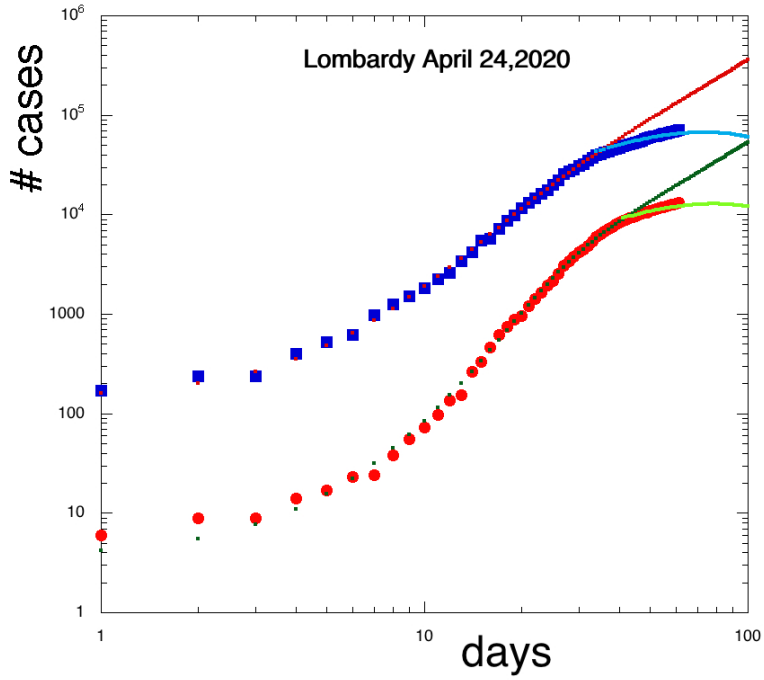

Figure 6. Predicted cases with and without quarantine as function of time, see text. Data for positives and deceased are given by the square and circle symbols respectively.

Multiplying equations (1) or (2) by the predicted number of tests from figure 5, gives the total number of predicted cases and are compared to the data in figure 6. We assume a conservative 10% error in our estimates due to the different fit functions. Without social distancing, using equation (1) gives 360,000 (±36,000) for the positives and 53,000 (±5,300) for the deceased 100 days after the beginning of the epidemics in Lombardy. If the exponential decay given by equation (2) is taken into account (due to the quarantine), the values decrease to 80,000 (±8,000) and 15,000 (±1,500) respectively, thus about 38,000 saved lives in Lombardy alone! There is an important difference between the two stages: if the first stage alone would be at play, the epidemics may continue after the 100 days and eventually slow down at longer times. Recall that the Spanish flu started in 1918 and lasted almost 36 months with an enormous death toll, https://www.washingtonpost.com/graphics/2020/local/retropolis/coronavirus-deadliest-pandemics/.

Because of the second stage, now the predicted values are given by the maxima in figure 6, these occur 76 and 82 days respectively after the start of the epidemics recording, i.e. May 10 and 16,2020 respectively. These values are close to the sum of t_{crit} and τ_d reported above.

If we assume a power law to reproduce the available data for the number of tests, figure 5, then we can write the total number of cases in the second stage as:

$$\#cases = m_1(d - m_2)^{m_3} d_\infty e^{-\alpha(d - t_{crit})} \quad (3).$$

The fitting parameters $m_{1\text{-}3}$ are reported in figure 5 for Lombardy. To find the maximum of equation 3, we simply equate its derivative to zero:

$$d_{max} = \frac{m_3}{\alpha} + m_2 \quad (4).$$

Using the empirical relation above connecting d_{max} and t_{crit} we get:

$$d_{max} \approx \mathrm{t}_{crit} + \frac{1}{\alpha} \rightarrow \mathrm{t}_{crit} = \frac{m_3 - 1}{\alpha} + m_2 \quad (5).$$

This relation is very useful especially when the data does not show the exponential decrease since it reduces the number of free parameters entering equation (2). Similar relations can be derived for different parameterizations for the total number of tests.

The US states

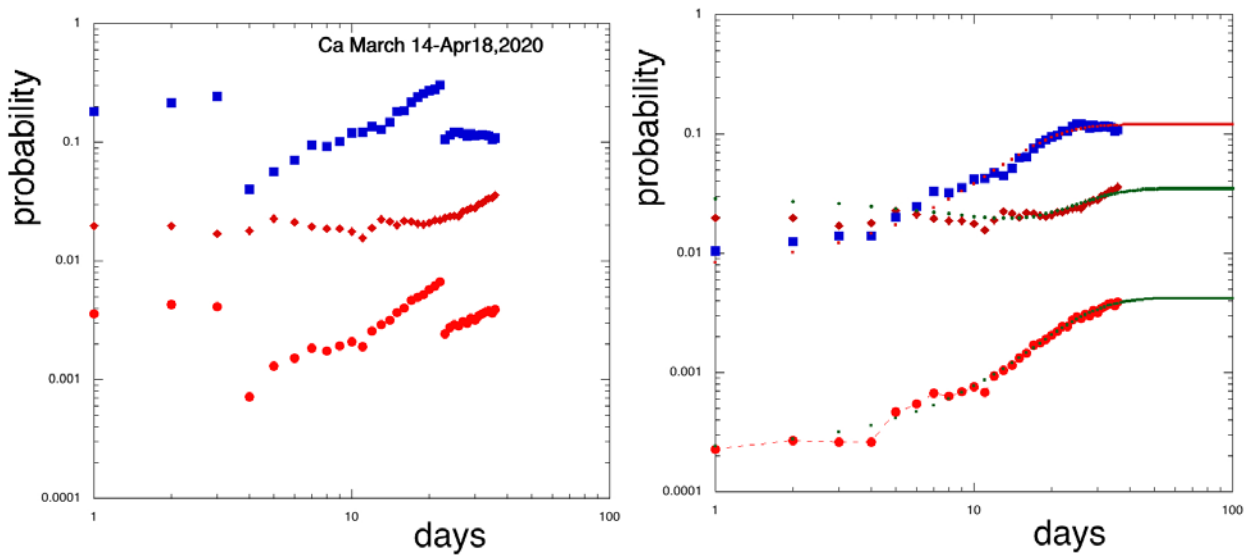

Figure 7. Positive (square symbols) and deceased (circle symbols) probabilities vs time in days. The rhomb symbols represent the ratio deceased/positive independent on the number of tests. The right panel is obtained after renormalization, see text.

Figures 1-3 show that the USA was hit hard by the COVID-19 resulting in different responses from the different states. In this section we will analyze some of these states and more analysis can be found in the supplemental material or available from the authors. In figure 7, we plot the probabilities for the state of California (Ca) for the period indicated in the inset, compare to figure 3 and 4. The discontinuities are due to the change in the number of tests performed daily. Notice that March 14,2020 coincides with the quarantine declaration in Italy, thus it was not a surprise that the virus spread quickly. Fortunately, the San Francisco mayor and the California governor placed strict restriction as early as March 6 without waiting for better testing, https://www.sfdph.org/dph/alerts/files/HealthOfficerLocalEmergencyDeclaration-03062020.pdf. This action saved a large number of lives and kept the ratio deceased/positives very low, compare to figure 4. We can correct in some cases for the low number of tests. Large data taking has a better statistical value thus we can renormalize the data where the jumps occur to the value at later times. In the right panel we display the result of the renormalization together with the fit using equation (1).

The hardest hit state was New York. In the figure 8 we display the probabilities together with the

fits using equations (1) and (2), compare to figures 4,7. The ratio deceased/positives seems smaller than the Lombardy one, however particular attention should be paid to the counting methods and some confusion might arise if the data refer to the state of New York (NY), https://coronavirus.jhu.edu/map.html, or to New York city (NYC), https://covidtracking.com/data/state/new-york#historical, the difference being roughly 5000 deaths since most cases are in NYC. The bending down of the curve is evident and we can make a prediction using equation (2). The resulting fit is displayed in figure 8, it follows well the available points but further confirmation will be given by future data.

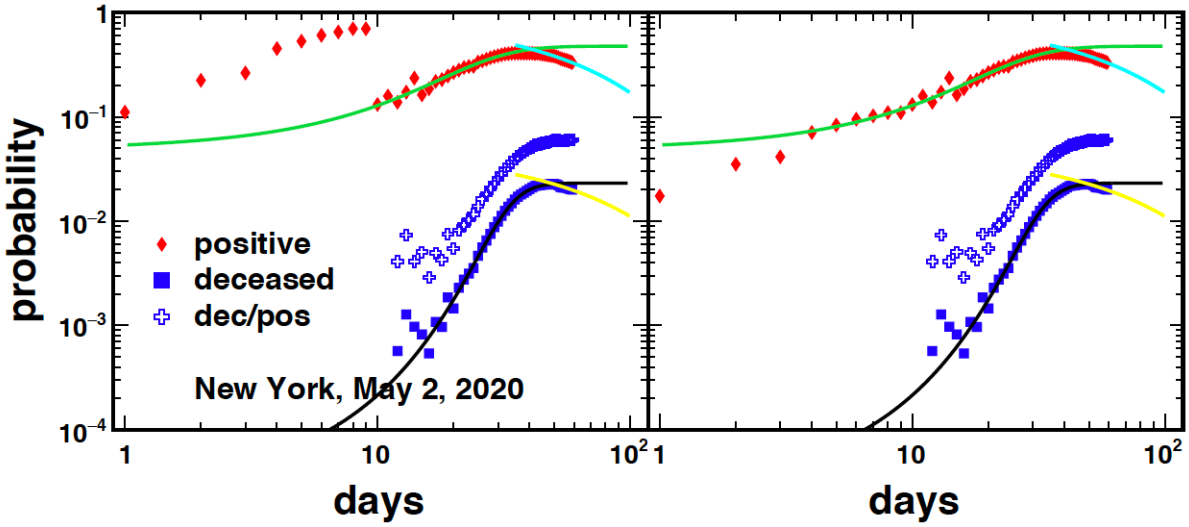

Figure 8. Same as figure 7 for the state of New York.

Using the predicted number of tests for NY given in figure 9, left panel, and the probability fits from equation (1) and (2) displayed in figure 8, we can predict the total number of cases as for Lombardy. The results are plotted in figure 9, right panel, for the first 100 days from the start of the recordings.

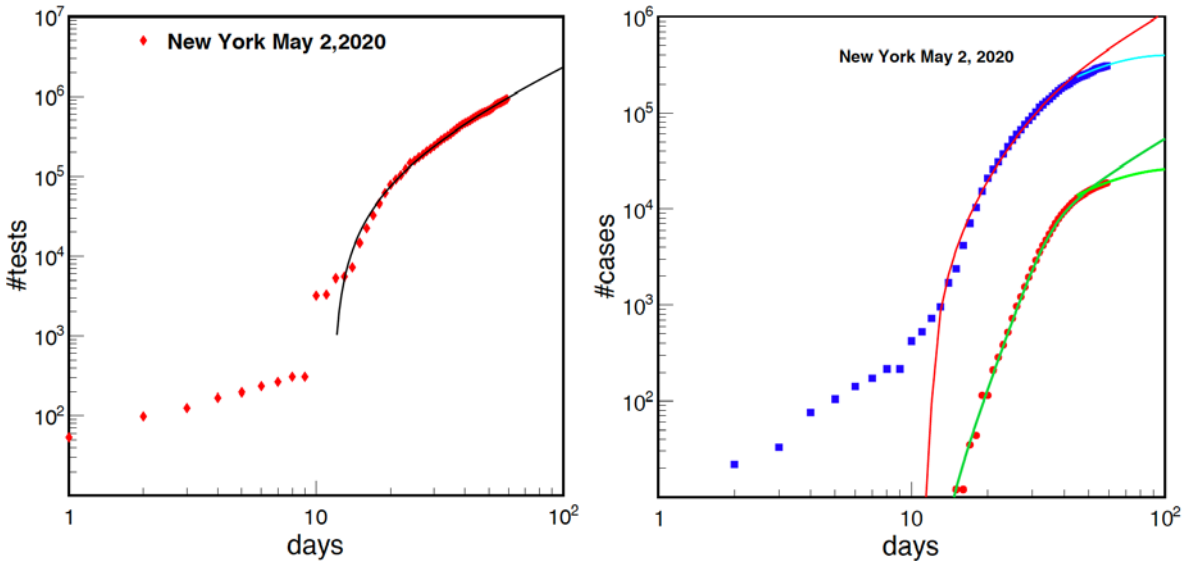

Figure 9. Same as figures 5 and 6 for the state of New York.

Conclusions and outlook.

We have proposed a two-step model to the rise and decay of the epidemics due to the COVID-19. The model needs some input parameters to predict the time evolution up to the saturation of the probability as in equation (1). Once the plateau is reached, given by the d_∞ parameter, the probability remains constant for some time depending on the quarantine measures or other environmental factors. For the Italian case, the first test was published on February 24 and equation (1) was fitted on March 10 before the quarantine was announced, i.e. March 14 [7]. The plateau was reached around March 24 as predicted by the model. These dates suggest that the quarantine was not effective in reducing the maximum probability and the time when this was reached. The quarantine became effective roughly 10 days after saturation. Thus we can estimate that it takes about 3 weeks before the quarantine gives an effect and the probabilities start decreasing, this is the value of t_{crit} entering equation (2). We can suppose that if the quarantine was announced say 14 days earlier, then the exponential decrease, equation (2) would had intercepted the rise, equation (1), earlier resulting in smaller probabilities. This is what happened to S. Korea and Japan, and it would explain the differences among countries: the later and the more feeble the quarantine the higher the probabilities and the longer the time to return to (quasi) normality. From these considerations we can estimate the time it takes for other countries also if the probability decrease is not seen yet. After reaching the top of the probability, see figure 3, it took roughly 10 days for Italy to see the decrease. If the data shows the decreasing part then a fit using equation (2) will be performed, otherwise we use the parameters found for Italy. In figure 10 we plot the predicted total number of positives, in the same figure we plot the same quantities for the deceased. Countries, which did not provide the number of daily tests (Spain, China), were not analyzed including the UK because of the large increase in testing especially at later times in coincidence to their Prime Minister hospitalization.

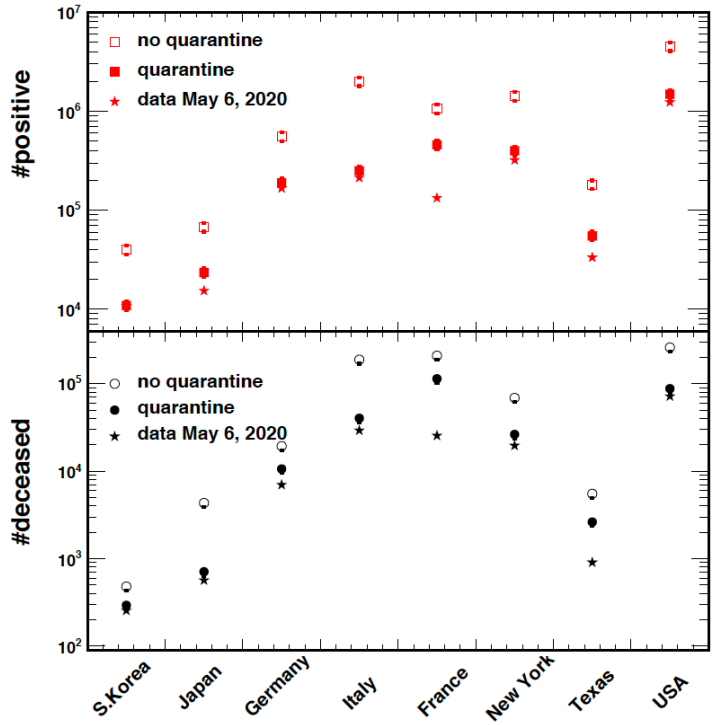

Figure 10. Number of positive (top panel) and deceased (bottom panel) for the countries indicated on the abscissa. Predictions without quarantine measures refer to June 28,2020.

Sweden decided to follow a different path non imposing the quarantine (herd immunization or natural selection), a choice that could be justified under the assumption that the vaccine will not be available soon enough. In figure 11 we plot the probabilities for Sweden, Finland and Norway since they are bordering countries. The probabilities are quite different especially regarding the death rate. We predict for Sweden about 74,750 (±7.5e3), 8,195 (±820) positives and deceased respectively on June 28,2020. Since there is no quarantine we are not able to estimate t_{crit} and the decay rate. On the same day, using equation (1), we predict for the other countries the values 18,271 (±1.8e3) and 1,495 (±149) for Finland, 15,225 (±1.5e3) and 391 (±39) for Norway. Thus we see that herd immunization takes a heavy toll not justified by the larger Sweden population (a factor of 2 respect to the other countries considered) and it will be very difficult to explain this choice to the relatives of the victims and their lawyers. We do not have any explanation for the difference in the number of deceased for Norway and Finland since the number of positives is practically the same. Authorities of those countries should investigate this difference further.

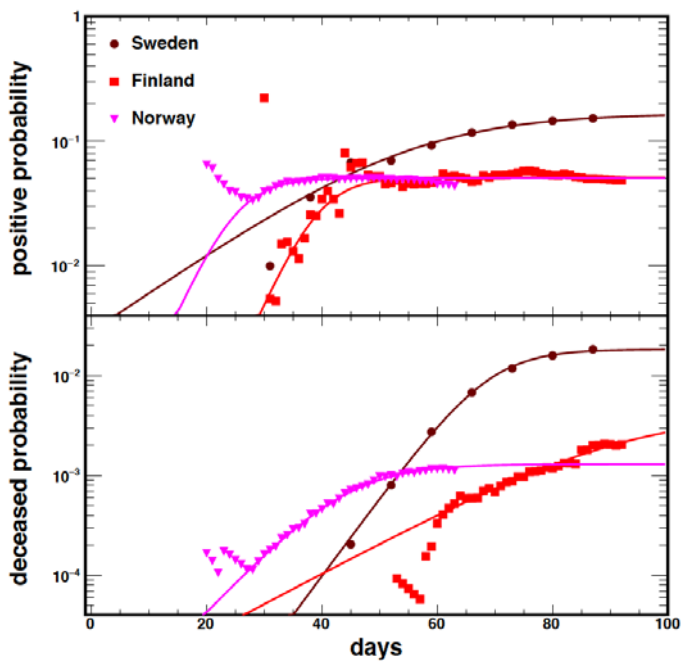

Figure 11. Probabilities as function of time for the countries indicated in the inset. Sweden is adopting the natural selection option resulting in higher probabilities compared to nearby countries. Different starting data depend on which day the complete information needed for the plot was released.

One feature worth noticing from figure 11 is the time delay and the slow spread of the Covid-19, this could be due to the extremely cold weather in the winter and early spring for these countries. There is some hope that the warmer season will help to normalize the situation, as for flu. Other reasons might be put forward, for instance if the virus is somehow adapted to bats, we can naively assume that it will be more deadly for temperatures higher than 10 ^{0}C, since below such value most bats hibernate. Temperature difference might explain the spread delay in countries like France, UK and Germany respect to Italy. Of course, other ingredients must be considered such as people flows from/to infected places, population density etc. No matter what the reasons may be for a temperature dependence of the spread it is clear that some systems perform better if it is not too hot or too cold. We can test these hypotheses using the 50 US states data since they cover a wide range of temperatures in the spring season. In the figure 12, we display the results obtained using the data on May 3,2020. Different states values were averaged if their temperatures differ about 1^{0}C in order to have better statistics. Gaussian fits give 10^0 C at the

maximum and a similar variance, see the inset in the figure 12. The large error bars and the discrepancies respect to the fit at higher temperatures refer to touristic places: Florida, Hawaii and Louisiana, particularly popular during spring break. If we take this result at face value, it predicts about 240 people per million inhabitants to be positive to the virus in the summer with 35 ^{0}C average temperatures. Even if this value seems small, it is a seed d_0 to restart the epidemic. We can already see this in the figure 12 from the large increases over the Gaussian fit corresponding to high population density states and large touristic flows. It might suggest that low temperatures in hospitals may decrease the virus aggressive spreading, keeping in mind that a vaccine is the only definitive solution. Until then we can only aggressively test and isolate positives similarly to the S. Korean approach to the pandemic.

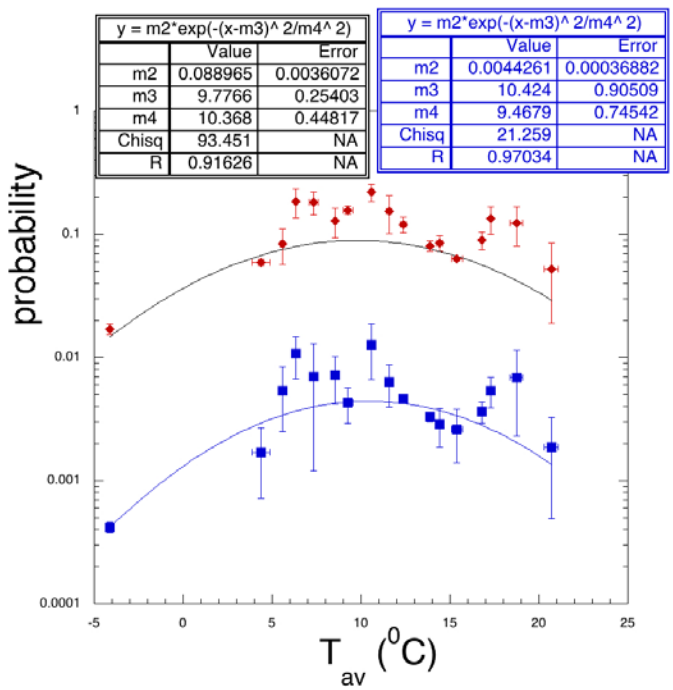

Figure 12. Positive (rhomb symbols) and deceased (square symbols) probabilities vs average temperature of the US states in the spring season. The data to calculate the probabilities were collected on May 3,2020. To get better statistics, averages were performed over states differing about 1 ^{0}C. Gaussian fits are also included and the fit parameters given in the inset.

In conclusion, in this paper we have discussed the predictive power of a two-step model based on chaos theory. A comparison among different countries suggests that it would be safe to release the quarantine when the probability for positive is lower than 4%, the maximum value for S. Korea. This implies that, if the quarantine is dismissed, then the same measures, as for the Koreans, should be followed by the other countries: careful testing, backtracking and isolation of positives. Herd immunization or natural selection is very difficult to justify from the data available so far, especially since we are dealing with thousands of human lives no matter the age or other nonsense.

References

[1] H.G. Schuster, Deterministic Chaos, VCH-1995.
[2] L. D. Landau and E.M. Lifshitz, Fluid Mechanics, Elsevier-2004.
[3] P. Arena, A. Bonasera, C. Brigante, M. Bucolo, F. Di Grazia, D. Lombardo, F. Sapuppo, and M. C. Virzí, in *2008 Mediterr. Conf. Control Autom. - Conf. Proceedings, MED'08* (2008).

[4] V. Baran et al., Rom.Journ.Phys.,60, 1263(2015).
[5] S. Zhang et al., Phys.Rev.C99,044605 (2019).
[6] A. Bonasera et al., Rivista Nuovo Cimento, 23 n.2 (2000)1.
[7] Bonasera A and Zhang S (2020) Chaos, Percolation and theCoronavirus Spread. Front. Phys. 8:171. doi: 10.3389/fphy.2020.00171.
[8] A. Bonasera, G.Bonasera and S.Zhang, 'Chaos, Percolation and the Coronavirus Spread: the Italian case', submitted, https://www.medrxiv.org/content/10.1101/2020.04.10.20060616v1.
[9] M. Perc et al., fphy-08-00127 (2020), https://doi.org/10.3389/fphy.2020.00127 .
[10] Povh, B., Rith, K., Scholz, C., Zetsche, F., Rodejohann, W., Particles and Nuclei, 2006, springer; ISBN 978-3-540-79368-7.