medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

DESY 20-234, HU-EP-20/42

Emergence of universality in the transmission dynamics of COVID-19

Ayan Paul[a,b,†], **Jayanta Kumar Bhattacharjee**[c,†], **Akshay Pal**[c,†], **and Sagar Chakraborty**[d,†]

[a]DESY, Notkestraße 85, D-22607 Hamburg, Germany
[b]Institut für Physik, Humboldt-Universität zu Berlin, D-12489 Berlin, Germany
[c]Indian Institute for Cultivation of Science, Jadavpur, Kolkata 700032, India
[d]Department of Physics, Indian Institute of Technology Kanpur, Uttar Pradesh 208016, India
[†]e-mail: ayan.paul@desy.de, jayanta.bhattacharjee@iacs.ac.in, akshay.pal@iacs.ac.in, sagarc@iitk.ac.in

ABSTRACT

The complexities involved in modeling the transmission dynamics of COVID-19 has been a major roadblock in achieving predictability in the spread and containment of the disease. In addition to understanding the modes of transmission, the effectiveness of the mitigation methods also needs to be built into any effective model for making such predictions. We show that such complexities can be circumvented by appealing to scaling principles which lead to the emergence of universality in the transmission dynamics of the disease. The ensuing data collapse renders the transmission dynamics largely independent of geopolitical variations, the effectiveness of various mitigation strategies, population demographics, etc. We propose a simple two-parameter model—the Blue Sky model—and show that one class of transmission dynamics can be explained by a solution that lives at the edge of a blue sky bifurcation. In addition, the data collapse leads to an enhanced degree of predictability in the disease spread for several geographical scales which can also be realized in a model-independent manner as we show using a deep neural network. The methodology adopted in this work can potentially be applied to the transmission of other infectious diseases and new universality classes may be found. The predictability in transmission dynamics and the simplicity of our methodology can help in building policies for exit strategies and mitigation methods during a pandemic.

Introduction

Despite several warnings of an imminent pandemic, COVID-19 dropped upon the world, quite unexpectedly, out of a blue sky. The spread of SARS-CoV-2 has left significant instabilities in the socioeconomic fabric of the society essentially defenestrating any notion of normality that existed. While the spreading dynamics of the disease is not novel [1], the instabilities it has caused has made different parts of the society, and notably, different governance, respond to containing its spread in very different manners [2–6]. The determination of the optimal strategy, while not dragging what people have known as normality to the ground, has been quite a challenge and highly dependent on the socio-economic condition of the country or region [7–10]. On the other hand, researchers have been struggling to bring some predictability in the spread of the pandemic since even a few weeks of foresight can not only save an economy from being jettisoned but also save a considerable number of lives that needn't be lost. Moreover, given the predator-prey game that is being played out between the human population and the virus, once the game starts showing positive outcomes for the human population, deciding on exit strategies become of prime importance as an immediate resurgence of the pandemic is the least coveted outcome.

Epidemiologists and mathematicians have constructed and resurrected various kinds of models to describe the spread of COVID-19 with varying degree of success. These include several variants of the venerable SIR (Susceptible, Infected , Recovered) model like SEIR [11–14]—E for exposed, SEIRD—D for deceased, and Stochastic SIR [14, 15]. There are more realistic delay differential equation models [16–20], renormalization group equations [21] and the highly computation intensive (and arguably the most detailed) agent-based network models [22–30]. There are even purely data-driven models [31–33]. In all these approaches, sets of parameters have to be tuned precisely to analyse a particular country or region to obtain predictions, since every region seems to have its own special time evolution for disease propagation, making this a Herculean task. Although all models' assumptions cannot strictly describe every detail of any real-world system, it is believed that such systems may possess certain universal behaviours independent of these details [34].

We propose to establish that there is a method in the madness of the spreading of this disease which can be revealed by subjecting the available data of all the countries or regions to a scaling analysis reminiscent of the data collapse in the study of critical phenomena [35, 36] in the sixties and seventies of the previous century. Data collapse leading to universality has been

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

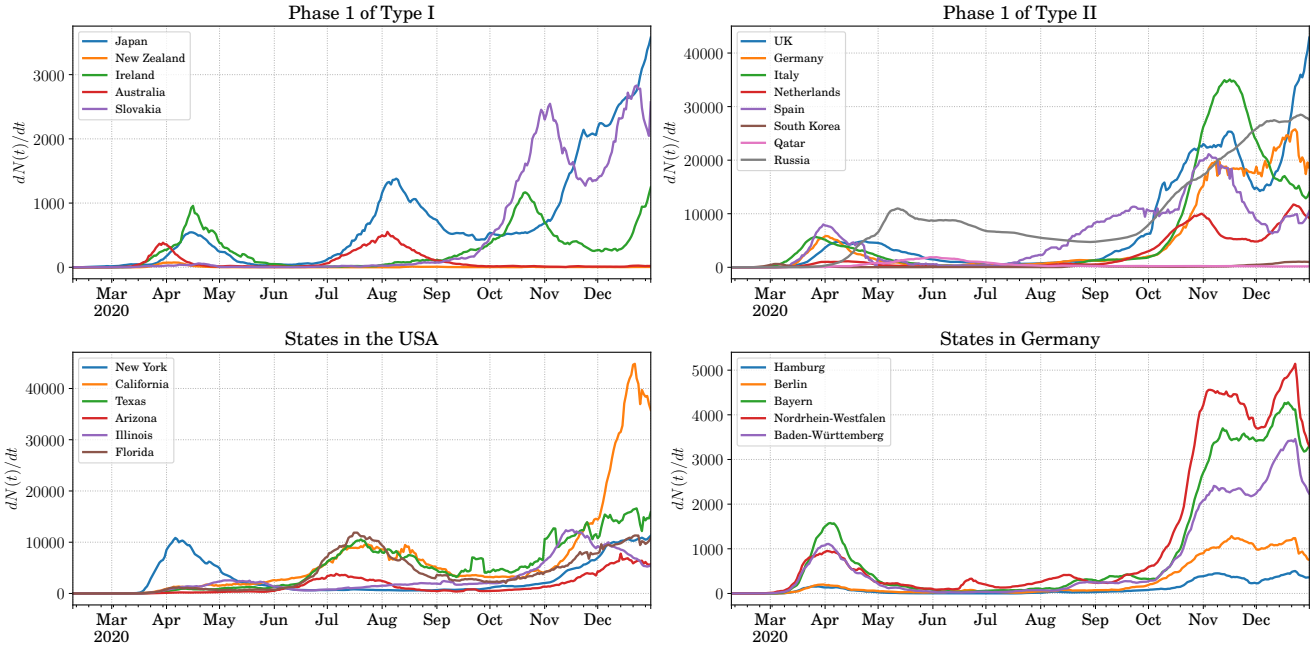

Figure 1. ***The variation in the spreading of COVID-19 in various countries and some states of the USA and Germany.*** *Top left: the countries for which the first phase of spreading falls into the category defined as Type I. Top right: the countries for which the first phase of spreading falls into the category defined as Type II. Bottom left: some states in the USA that show very different spreading patterns for the disease. Bottom right: some states in Germany that show very similar pattern in the spreading of COVID-19.*

observed in a large variety of systems and we provide a short overview of the literature in the Appendix. We emphasize that the procedure we follow is model-agnostic and does not require the assumption of a model for the transmission and mitigation dynamics. In addition, this method lies open to model-based interpretations as we later point out. Almost miraculously, we find a simple yet nontrivial mean-field model where the carrying capacity of infected population depends on the population size. The model—that we call Blue Sky model (BSM)—undergoes a blue sky bifurcation [37] (also called the saddle-node bifurcation or the fold bifurcation) which is a ubiquitous structurally stable bifurcation found in variety of systems, such as power systems [38], viscous profiles [39], neural networks [40], excited molecules [41], plasmas [42], and the Ginzburg–Landau systems [43]. At the edge of the blue sky bifurcation point, appears a universal data collapse. Furthermore, given that the scaling and data collapse have stood well the test of time for a very diverse set of systems and is grounded in well understood physical principles, we are proposing a robust method for making long-term predictions that will help in evaluating disease mitigation efficacies and planning exit strategies for the future. We hope to convince the readers that we have established a method which does not require parameter tuning or model selection to gain a handle over the prediction of disease spread.

The salient features of our work are:

- We show that there exists a universality in the spread of the disease once proper scaling is applied to the data.
- Two universality classes emerge that give us a window into understanding how well a country or region performed in mitigating the disease. These universality classes are independent of geographical scales.
- We propose the BSM—an extension of the logistic growth model—which explains the existence of these two classes, the first, through a solution that lives at the edge of blue sky bifurcation and a second that maps the BSM onto the logistic growth model.
- We show how predictability can be realized from this analysis in a model-agnostic manner relying only on the inference drawn from the first phase of the disease spread. For this we construct a deep neural network (DNN) based predictive model. We also show that an equally good prediction can be made using the BSM. We give a few examples from the ongoing phase of the pandemic in several countries and states.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

DESY 20-234, HU-EP-20/42

Emergence of Universality in Transmission Dynamics

To bring about some semblance of order in what looks like a disparate set of data from several countries, we first focus entirely on phase-1 of the pandemic. We denote the total number of cases (individuals who have tested positive for the COVID-19) at time t as $N(t)$ with t being measured in days. The derivative $dN(t)/dt$ is the number of new cases detected on a given day. Phase-1 refers to the situation, for a given country or region, where the $dN(t)/dt$ increased from zero or a small value, went through a maximum and then decayed to almost zero (Type I transmission) or to a small value (Type II transmission). It stayed near zero or near the small value for a few weeks before starting to rise again at the onset of the next phase. We need to differentiate quantitatively between the phrase "near zero for a few weeks" and "a small value for a few weeks". We do this by inspecting the rise in the total number of cases, $N(t)$, over four weeks past phase-1. If the change is more than 4% of $N(t)$, we mark the transmission as being of Type II, otherwise it is Type I.

Taking a more careful look at Figure 1, we note that the total number of cases and the total duration of phase-1 are very different for each country. This is to be expected as these countries have very different population densities and different patterns of inhomogeneities in the density of population. Moreover, the policies of the government and the acceptance of these by the governed affect the implementation of precautionary measures [8–10, 44]. The policies and exit strategies towards the end of a phase in any country also determines whether the spread of the disease lingers on, albeit at a much reduced rate, or is completely rooted. The latter essentially divides the world into two classes in our formalism. However, looking at bare case numbers, one is bound to conclude that each country or region is a unique case study requiring a detailed modeling of transmission dynamics and incorporating the effects of mitigation and containment measures, a task that seems formidable right from the onset.

The top left and right panels of Figure 1 show some examples of countries that went through a phase-1 transmission of Type I and Type II respectively. The disparity in the spreading patterns is apparent. In the lower panels of Figure 1 we focus on the states of the USA and Germany. In the USA we see that the spreading pattern in the different states was somewhat decorrelated. This is where our formalism is particularly useful. The physical principles that we advocate being independent of geographical scales can be applied to countries, states or even counties and districts given they have seen a significant spread of the disease within the region. On the other hand, the spread of COVID-19 in Germany in the various states has been quite correlated.

Now, let us see how universality can emerge from a scaling of the COVID-19 confirmed case rate data. We introduce a number $N_{\max}$ which is the total number of infected people at the end of phase-1. For Type I transmission the value of $N_{\max}$ is easy to infer. It is the the total number of cases at which the $N(t)$ curve flattens out after the close of the phase. For Type II transmission, where $N(t)$ is never completely flat, we determine the close of phase-1 following the procedure described above. A perfect example is the case of Germany. For the period May 31st, 2020 to July 17th, 2020, the value of $dN(t)/dt$ hovered around 400 which is more than an order of magnitude smaller than the peak value of 6000. In this case $N_{\max}$ corresponds to the number of cases at the end of this period of stagnancy in the growth rate, i.e., the number of cases on July 17th, 2020. Of course, there are some degree of variability associated with the choice of the end-date for a phase. However, an estimation error of a few days does not change the onset of universality. Therefore, we choose the end-date for phase-1 as the first day when the $N(t)$ curve flattens out for Type I transmission and the last day before the onset of the next phase for Type II transmission.

Our conjecture is that universality emerges from the scaling given by

$$\tilde{t} = \frac{t}{t_{1/2}}, \text{ and } \tilde{N}(\tilde{t}) = \frac{N(\tilde{t})}{N_{\max}}. \tag{1}$$

The ratio $\tilde{N}(\tilde{t})$ is relatively insensitive to the testing policies, the inhomogeneity in the population density, the size of the country or region, the details of the mitigation methods advocated, etc. The rescaling of the time by an appropriate time scale for each country or region, irons out the nuances brought about by the government policies like the duration of lockdown, exit strategies, etc., at the lowest order of approximation. Therefore, for each country or region a natural time scale to choose is $t_{1/2}$ which is the time at which $N(t) = N_{\max}/2$.

The choice of the characteristic timescale is subjective. For instance, choosing $t = t_{1/n}$ corresponding to the time when $N(t) = N_{\max}/n$, with n being a positive definite number would not affect the nature of this analysis. However, choosing $n = 2$ allows for an easy interpretation of the characteristic time scale and, as we see later, places the peak of $d\tilde{N}(\tilde{t})/d\tilde{t}$ at $t = t_{1/2}$ for a simple logistic growth. According to our conjecture, $\tilde{N}(\tilde{t})$ plotted against $\tilde{t}$ should show a data collapse opening up avenues for simple predictive modeling and generalization beyond socioeconomic and geopolitical variations.

Data-collapse in transmission dynamics of COVID-19

In the light of the preceding discussion we present the $\tilde{N}(\tilde{t})$ vs. $\tilde{t}$ plots for several countries and states in left-most plots of Figure 2. The data collapse is clearly evident. We see that independent of variabilities in transmission dynamics the universality in the rescaled data dictates that there are two distinct classes. The plots in the middle of Figure 2 show the functional

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

DESY 20-234, HU-EP-20/42

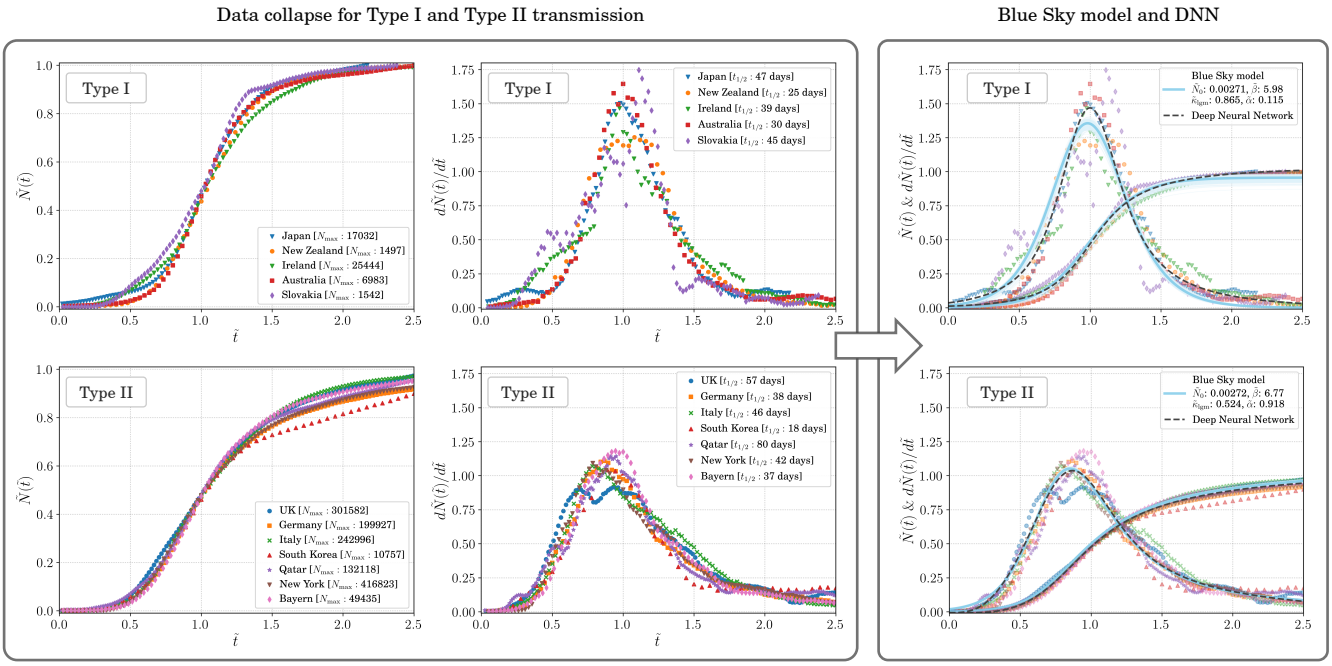

Figure 2. ***The emergence of universality in the transmission dynamics of COVID-19***. *Left panels: The rescaled total number of confirmed cases, $\tilde{N}(\tilde{t}) = N(t)/N_{\max}$, and the its derivative, $d\tilde{N}(\tilde{t})/d\tilde{t}$, with respect to the scaled time $\tilde{t} = t/t_{1/2}$ for Type I transmission where the disease spread was almost completely routed after phase-1 and Type II transmission where the disease spread was not completely routed after phase-1. The data collapse is evident and independent of the country or state being considered showing the emergence of two universality classes in the transmission dynamics of COVID-19 independent of social and geopolitical variations. Right panels: The fit to the rescaled data using the Blue Sky model (solid blue line) and the model agnostic fit (dashed black line) with a deep neural network. Note that $\tilde{\alpha} = 0.918$ in the rightmost bottom panel is practically the value expected at the edge of the blue sky bifurcation at $1/4\tilde{\kappa}^2_{\rm lgm} = 1/(4\times 0.524^2) = 0.910$. The good agreement between the two approached can be clearly seen.*

dependence of $d\tilde{N}(\tilde{t})/d\tilde{t}$ on $\tilde{t}$. For the Type I transmission one clearly sees that tail of the distribution at higher $\tilde{t}$ dips to very low values. This implies that after phase-1 ended, the daily case rate actually went down to a negligible number, which, originally, was our basis for identifying Type I transmission. On the other hand, for the same curve in the lower middle plot of Figure 2, the tail of the distribution at higher $\tilde{t}$ never really goes to as low values as that for Type I transmission. This is because in a transmission of Type II the countries and regions never saw a complete rooting of the spread of the disease. To show the independence of geographical scales of our formalism, we include New York, USA and Bayern, Germany as two states tracking them over phase-1 of the spread of COVID-19. We see that both the states show transmission of Type II.

It is important to point out here that the data collapse that we see effectively filters out the variances in the dynamics of the disease spread and the policies for disease mitigation. It also effectively filters out how quickly the response to the pandemic was orchestrated (a matter of government policy), how drastic these measures were and all other geopolitical demographics that make the disease propagation dynamics in each country or state look very different in Figure 1. The key difference between the two classes reduces to the effectiveness of the countries in fully containing the disease itself, each using their own bouquet of mitigation methods and exit strategies. Indeed, this data collapse removes all necessity of fine-tuning a plethora of parameters to examine each country or state separately and shows that there is a common pattern in the spread of COVID-19.

Interpretation of $t_{1/2}$

The parameter $t_{1/2}$ gives us a distinct insight into how well the spread of the disease was mitigated in a particular phase. A shorter $t_{1/2}$ implies a more effective mitigation strategy. For example, from Figure 2 we see that the $t_{1/2}$ for South Korea is 18 days, which means that the spread of the disease was stopped very rapidly as compared with UK for which $t_{1/2}$ is 57 days. However, despite being effective in slowing down the spread rapidly, South Korea falls into Type II since the spread was not completely stopped but lingered on. On the other hand, Ireland, even being of type I, has a relatively longer $t_{1/2}$ of 39 days showing that while the spread of the disease was almost completely rooted it took a longer period of time to do so hinting as possibilities of drafting better mitigation strategies.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Blue Sky model

Among the four broad classes of models (with possibility of overlap among them) of infectious disease spreading—viz, data-driven phenomenological models [45, 46], compartmental or lumped parameter models [47], agent-based models [48], and stochastic differential equation models [49]—while the former two are mathematically more accessible to policy makers, the latter two are more detailed in their approach. The first two types of models are to the last two types, what thermodynamics is to statistical mechanics. The most simple lumped parameter model is undoubtedly the well-known one dimensional autonomous logistic growth model [50]:

$$\frac{dN(t)}{dt} = \beta N(t)\left(1 - \frac{N(t)}{\kappa}\right), \tag{2}$$

on which we shall focus. Here, $\beta > 0$ is the Malthusian growth rate in the presence of unlimited resource, i.e., $\kappa \to \infty$. Most commonly, the resource is assumed to be unaffected by the population size, $N(t)$, and hence the population grows to settle onto the value $N(\infty) = \kappa$, which is the carrying capacity. We discuss some variants of the the logistic growth model and their fit to the data in the Appendix.

The system of our interest is a population where $N(t)$ represents the number of COVID-19 infected individuals that infect healthy individuals so that the infected population grows at a rate β. All the depletable conditions that are favourable for the infected population to grow is quantified through the positive parameter, κ. One could argue that rather than relying completely on the carrying capacity parameter, one could make this model more elaborate by making it higher dimensional to mathematically incorporate the interaction among infected, susceptible, exposed, asymptomatic, quarantined, recovered, immune, and deceased individuals. However, the fact remains that all such higher dimensional models, despite their share of success, bring in more parameters and more speculations while the only available data is still mostly that of the infected and deceased individuals. Thus, in this paper, we are motivated to take the route of simple two-parameter one dimensional model complimented with the information extracted from the real data.

It is straightforward to realize that the carrying capacity can easily depend on $N(t)$: e.g., government may be practically incapable of quarantining infected individuals as their numbers grow, thereby effectively increasing the carrying capacity because the infected individuals would interact with more and more healthy individuals leading to terms which are higher order in $N(t)$. Therefore, if we were to model the realistic scenarios, we should let κ depend on $N(t)$ in nontrivial ways. In this context, we find it useful to define an order parameter, ρ, as

$$\rho \equiv \frac{1}{\kappa^2}\left(\frac{d\kappa}{dN}\right). \tag{3}$$

Note that lower the value of the order parameter, more is the control on the change of the carrying capacity. The case of the order parameter being zero corresponds to the traditional logistic equation governed system where the carrying capacity is independent of $N(t)$. We would like to mention here that past considerations [16, 51–59] of the logistic growth related compartmental models, as applied to the modeling of COVID-19, do not explicitly incorporate the natural notion of the state-dependent carrying capacity.

In order to account for the aforementioned variation of the carrying capacity with the change in the infected population size, we propose a simple, yet non-trivial, functional form for the carrying capacity:

$$\kappa = \kappa_{\rm bsm} \equiv \frac{\kappa_{\rm lgm}}{1 - \alpha\,\kappa_{\rm lgm} N(t)}, \tag{4}$$

where $\kappa_{\rm lgm}$ is the constant carrying capacity corresponding to the traditional logistic growth model and α is a non-negative parameter. In Equation 2 with $\kappa_{\rm bsm}$ as the carrying capacity, we define the Blue Sky model (BSM) that we use to explain our findings in this paper. In the BSM, we note $\rho = \alpha$ which renders a solid physical meaning to the (otherwise ad hoc) parameter α. Following the transformations, $\tilde{N}(t) = N(t)/N_{\rm max}$ and $\tilde{t} = t/t_{1/2}$, adopted in the previous discourse, we define $\tilde{\beta} \equiv t_{1/2}\beta$, $\tilde{\kappa}_{\rm lgm} \equiv \kappa_{\rm lgm}/N_{\rm max}$, and $\tilde{\alpha} \equiv N_{\rm max}^2\alpha$. Henceforth, we shall consider only the rescaled BSM where these rescaled parameters are used.

In the BSM, the blue sky bifurcation occurs at the nonzero fixed point at $\tilde{\alpha} = \tilde{\alpha}_{\rm edge} \equiv 1/4\tilde{\kappa}_{\rm lgm}^2$—*the edge of blue sky*. Beyond the edge of blue sky, i.e., for $\tilde{\alpha} > 1/4\tilde{\kappa}_{\rm lgm}^2$, the two nonzero fixed points vanish and $\tilde{N}(\tilde{t}) = 0$ is the sole fixed point in the system which is unstable. In other words, the population size of the infected individual grows without any bound, which is a physical impossibility. This sets an important limit that $\tilde{\alpha}$ should respect for the model to be physical: one must impose the condition, $\tilde{\alpha} \in [0, 1/4\tilde{\kappa}_{\rm lgm}^2]$, so that a stable attracting nonzero state of the infected population, viz., $\tilde{N}(\tilde{t}) = (1 - \sqrt{1 - 4\tilde{\alpha}\tilde{\kappa}_{\rm lgm}^2})/(2\tilde{\alpha}\tilde{\kappa}_{\rm lgm})$ exists. The third fixed point, $\tilde{N}(\tilde{t}) = (1 + \sqrt{1 - 4\tilde{\alpha}\tilde{\kappa}_{\rm lgm}^2})/(2\tilde{\alpha}\tilde{\kappa}_{\rm lgm})$, is unstable and not of any practical interest to us.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

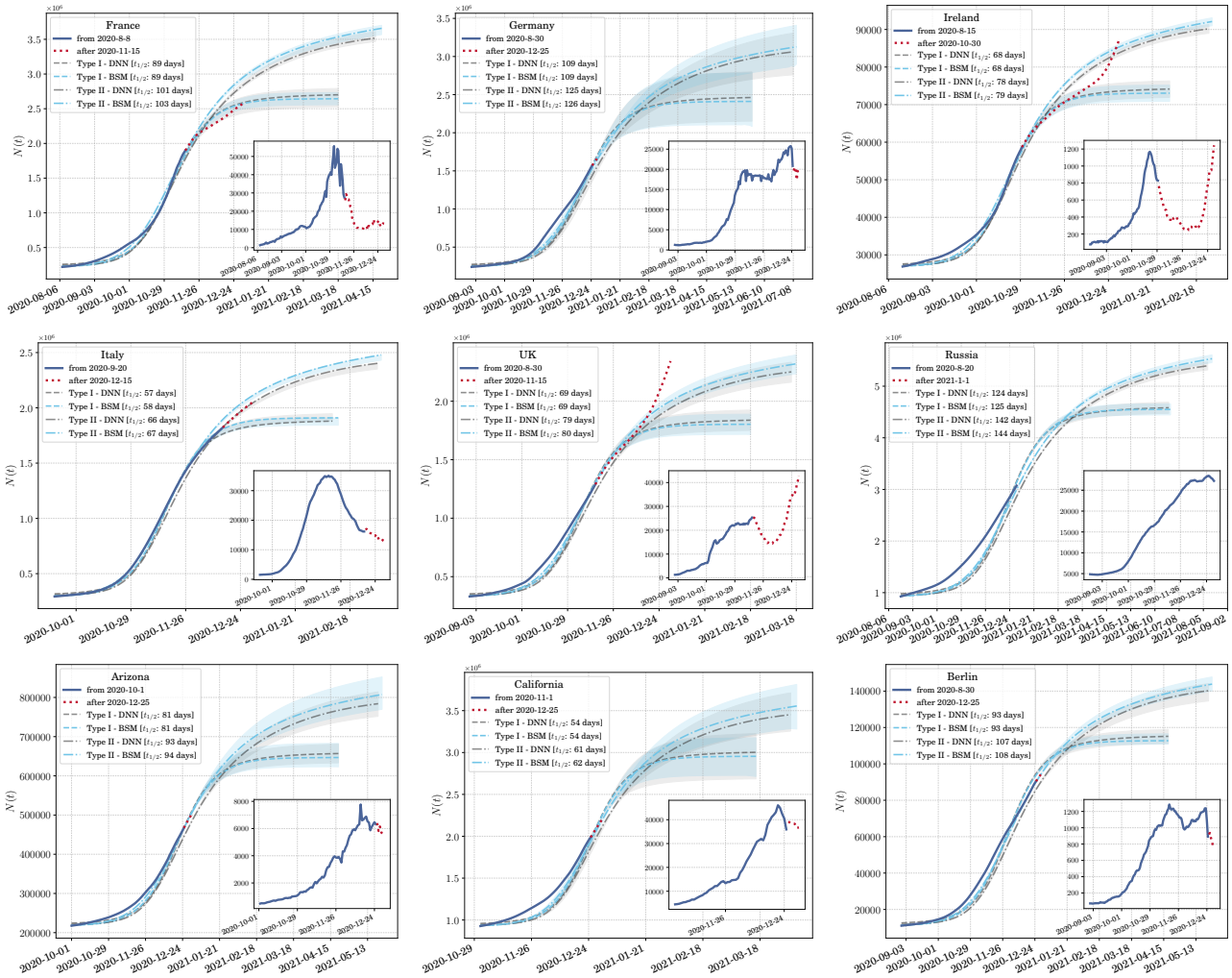

Figure 3. ***A display of predictability from the emergence of universality classes.*** *We show how predictions can be made form scaling data and using knowledge gathered from phase-1 of the pandemic. Once the peak in the daily case rate is reached (or has been assumed to have reached), prediction for several weeks or months can be made. The solid dark blue lines in the insets show the data being used to find t_{peak}, and from it, $t_{1/2}$ and N_{max}. The dotted red lines in the insets are the test data which lie in the future with respect to the data used to determine the peak. The light blue lines and bands are predictions from the model-dependent fit using the Blue Sky model (BSM) to phase-1 data and the gray lines and bands are those from the model-agnostic fit using a deep neural network (DNN). The dashed lines are prediction for transmission of Type I and the dot-dashed lines are for transmission of Type II. All predictions are made till $\tilde{t} = 2.5$ or $t = 2.5t_{1/2}$ starting from the beginning of the ongoing phase of disease spread. The top and middle rows show predictions for various countries while the bottom row shows predictions for two states in the USA and one in Germany.*

As seen in the rightmost column of Figure 2, the BSM remarkably segregates the COVID-19 transmission dynamics in countries and states into two universality classes—Type I and Type II—characterized by two distinct and well separated values of the (rescaled) order parameter, ρ—one close to the logistic growth value ($\tilde{\alpha} \approx 0$) and the other close to the edge of blue sky ($\tilde{\alpha} \approx 1/4\tilde{\kappa}_{\text{lgm}}^2$) respectively. The data collapse onto the respective curves is strongly reminiscent of the similar phenomenon observed in the critical phenomena and their universality across a plethora of physical systems [35, 36]. Thus, we speculate that it may be a potential hint of a hidden renormalization scheme present in the microscopic details of the corresponding agent-based models.

Deep neural network

While the BSM provides a very good and intuitive explanation of the data, we feel that one should not be limited to exploring the universality in the data through specific models only. Hence, we perform a model-agnostic fit using a universal function

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

generator in the form of a deep neural network (DNN). We use a three-layers deep, each layer 16 nodes wide DNN for this purpose. The core of the architechture can be found in the Methods and Materials section and a detailed discussion can be found in the Appendix. The fit to the data using the DNN can be seen in the rightmost panel of Figure 2. The back dashed line corresponds to the DNN fit. We fit to the $\tilde{N}(\tilde{t})$ distribution. We see that the fit of Type II transmission is almost identical to the BSM fit showing that the data-driven and model-dependent methods reach the same conclusion. For transmission of Type I the DNN fit is slightly different from the BSM fit and notably so at the center of the distribution where $d\tilde{N}(\tilde{t})$ is shown to be higher from the DNN fit than from the BSM fit.

Predicting of COVID-19 spread from universality

The data collapse that we show leads us to the ability of making predictions in both a model-dependent and a model-agnostic way. While the BSM shows a clear explanation of the data, the advantage of a data collapse lies in the emergence of universality which allows for model-agnostic predictions from the simple understanding that the transmission dynamics described by the rescaled variables are the same in every phase of the the disease spread. Hence, we would like to emphasize both possibilities although they lead to similar results in predictions.

It should be noted that the only parameter that needs to be estimated for predictions to be made is $t_{1/2}$ for the ongoing phase for a country or region. Once that is determined, the parameters extracted from the fit to the phase-1 data can be used to make predictions for the ongoing phase. For both the approaches we use a common algorithm except for the functional form that we use to fit to the data. For the model-dependent approach we use the BSM and extract its parameters from data, whereas in the model-agnostic approach we use the trained DNN to predict the distribution of $\tilde{N}(\tilde{t})$.

The algorithm used to make the predictions is delineated in Materials and Methods section. To show the extent to which our prediction work, we assume that a country or state is just past the peak of $dN(t)/dt$. The data that we use for determining t_{peak} is represented by the solid dark blue line in the insets of Figure 3. The data past the end of this line, marked with a dotted red line is assumed to be in the future. We use the "past" data to predict for the "future" and then extend our prediction to $\tilde{t} = 2.5$. Since $t_{1/2}$ is unique for each country or state, each approach and each transmission types, the length of time for which predictions are made differ.

Model-dependent prediction

The fitting of the BSM to data from phase-1 paves the way for making model-dependent predictions for future phases of the pandemic. This can be done by identifying $t_{1/2}$ and, from it, N_{max}. This is only possible if, in a current phase of the pandemic, the daily case rate $dN(t)/dt$ has already reached a peak. This peak lies at time $t = \tilde{t}_{\text{peak}} t_{1/2}$ where $\tilde{t}_{\text{peak}}$ can be derived from the fit parameters and is given in Equation 7. By definition, $t = t_{1/2}$ is the time at which the $N = N_{\text{max}}/2$ with $N(t)$ being counted from the beginning of the phase. Once these two parameters have been extracted, one can make two separate predictions for any geographical region: the first assuming a spread of Type I and the second assuming a spread of Type II. These predictions for Type I (dashed blue line) and Type II (dot-dashed blue line) are shown in Figure 3 for several countries and states. The bands in blue represent the error in the prediction propagated from the error in the estimation of $\tilde{t}_{\text{peak}}$.

Model-agnostic prediction

We also show the predictions from the model-agnostic approach using a trained DNN in Figure 3. The logic that we follow to identify the peak of the $dN(t)/dt$ distribution is the same as for the model-dependent predictions. However, $t_{1/2}$ now varies between the two methods since $\tilde{t}_{\text{peak}}$ is different for the model-dependent and the model-agnostic fits. The dashed gray lines represents transmission of Type I and the dot-dashed gray lines represents transmission of Type II. For Type I transmission, the BSM gives predictions similar to the DNN. However, for Type II the BSM mostly predicts somewhat larger number of cases. This is because $\tilde{t}_{\text{peak}} = 0.85$ from the BSM fit of the phase-1 data is a bit lower than that from the DNN fit of $\tilde{t}_{\text{peak}} = 0.87$. The gray bands show the error in prediction propagated from the esrror in estimation of $\tilde{t}_{\text{peak}}$.

In the model-agnostic approach the regression is completely data driven, and hence, does not represent any underlying principle of the disease spread. This can be considered as a drawback of this approach. However, we show that the predictions made through this approach is very similar to the one made with the BSM. In addition the fit to data is also quite robust as can be seen from Figure 3.

Discussions on the predictions

There are several subtleties about the prediction that we should mention here. Firstly, the most important and sensitive part of the prediction is locating the peak of $dN(t)/dt$ during an ongoing phase. For some cases t_{peak} is easy to locate, for instance, for Italy and Ireland in Figure 3. However, for some other countries like Germany, it is somewhat more involved. Secondly, just past the peak it is difficult to gauge whether $N(t)$ will be monotonically decreasing or not. Lastly, in cases like Germany and UK we expect a significant margin of error in determining the peak itself since the $dN(t)/dt$ is not strictly unimodal. This

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

uncertainty in the determination of the peak is what the error bars for the predictions quantify. The error in determining t_{peak} propagates to an error in determining N_{max}. Hence, the error bars are small for Ireland and Italy, while they are quite large for Germany and California—the latter being just past the peak in the dataset we use. The spread of COVID-19 in Russia had just reached the peak while we were doing this analysis and hence, we make pure predictions without any test data.

In all the plots in Figure 3, we see that Type II transmission lasts much longer than Type I transmission. For Ireland we see that it was proceeding along the curve for Type I transmission until the next phase started when it diverged once more. Form this plot we now get an idea of how the pandemic might have been rooted in Ireland had the next wave not started. UK on the other hand seemed to be following the trajectory taken by Type II transmission until the number of new cases per day started rising once more.

Summary

The surfacing of a worldwide phenomenon such as a pandemic is highly disruptive. The social, political, and economic disruption that this pandemic has brought resulted in an essentially chaotic response to its mitigation. Everything from testing strategies to containment procedures followed very different patterns in different nations and variation even at the regional levels were very palpable. This has made the analysis of the data to model the underlying dynamics of disease spread a formidable task that has not yielded to any generalization.

In this paper we show that the emergence of order from within this chaos is made possible by applying well tested procedures drawn from the study of critical phenomenon leading to the formation of universality classes. Such ordering in the data leads to the possibility of characterizing the transmission dynamics into two classes, distinguished by how well the disease has been contained by each country or region. This clearly shows that the variation of the transmission dynamics, and the effect of policies related to disease mitigation, testing guidelines and exit strategies can be scaled away to instigate a data collapse that allows us to bring about remarkable predictability in the spread of the disease.

We reemphasize that without having to go into the rigours of extracting epidemiological parameters and having to formulate specific models for the spread of the disease, the emergence of universality allows for the prediction of future disease spread.

The lack of clear understanding of what model to use and how to extract epidemiological parameters has been a major hurdle to making mathematically sound predictions during the spread of COVID-19. In addition, such methods require that each country and each region be treated separately, quickly leading to an explosion of parameters and systems of equations. While being model-agnostic certainly has several drawbacks, the universality in the spread of the disease that we bring to surface here circumvents many of these problems by exploring the emergence of one single functional form that can be used to make prediction for the spread of the disease once a region has reached the peak of the spread. This clearly allows policy decisions to be made around exit strategies and lets the responsible governance understand how long it might take for a spreading wave of the disease to be contained given the current measures are kept in place.

We would like to point out that the advent of vaccination can be accommodated for in our analysis since it effectively reduces the size of the total susceptible population which, in turn, does not affect the rescaled parameters since they are scaled by N_{max}. Therefore, the validity of our work is not short-lived and have general applicability, regardless of the mitigation methods used.

Lastly, we bring to light the proposition of the BSM which explains the data remarkably well. We wish to highlight that it is an intriguing coincidence that Type II dynamics should live at almost the edge of blue sky bifurcation in the BSM. Looking at the finite nonzero value of the order parameter vis-à-vis the almost vanishing value of order parameter for Type I transmission, one is readily reminded of the phase transition phenomenon in physical systems. It is tempting to interpret the two types of transmission as two phases quantified by an order parameter. Furthermore, it remains to be seen if the newly introduced BSM finds applications in other epidemiological systems.

Acknowledgements

This research was supported in part through the Maxwell computational resources operated at DESY, Hamburg, Germany. We would like to thank Jon Crowcroft, Alejandro de la Puente, Mikhail Prokopenko and Melinda Varga for useful comments.

Appendix: Materials and Methods

All the numerical analysis in this work was performed using Python 3 and several libraries like SciPy, NumPy, Pandas and statsmodels. All the code necessary for reproducing our results, further exploring the data, making predictions and for making the plots are available in a GitHub repository: https://github.com/talismanbrandi/Universality-COVID-19.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Data sources and curation

In this work we used data from publicly available sources. The two primary sources that we used are:

- Data on COVID-19 confirmed case rate obtained from the Johns Hopkins University, Center for Systems Science and Engineering database [60] through their GitHub repository. The time-series for all the countries in the world and for all the counties in the USA are available from here. For the confirmed case rates in each state of the USA we merged the data for the counties in each state.
- For Germany the data is available through the Robert Koch Institute from the RKI COVID19 database.

For all datasets we worked with the seven day averages to remove the fluctuations due to variations in reporting over the period of a week. The last date on the time-series used in this work is 31st of December 2020.

Rescaling of the BSM Differential Equations

The rescaled BSM can be explicitly written as:

$$\frac{d\tilde{N}(\tilde{t})}{d\tilde{t}} = \tilde{\beta}\tilde{N}(\tilde{t})\left(1 - \frac{\tilde{N}(\tilde{t})}{\tilde{\kappa}_{\rm lgm}} + \tilde{\alpha}\tilde{N}(\tilde{t})^2\right). \tag{5}$$

For the purpose of prediction, we need to estimate $\tilde{t} = \tilde{t}_{\rm peak}$ at the peak of the distribution of the first derivative, $d\tilde{N}/d\tilde{t}$. For Type I transmission, $\tilde{t}_{\rm peak} = 1$ to a very good approximation. Numerically, $\tilde{t}_{\rm peak} = 0.98$ from the fit shown in Figure 2 for Type I transmission since the solution lies very close to the pure logistic growth model ($\tilde{\alpha} = 0$). However, for Type II transmission, $\tilde{t}_{\rm peak}$ depends on the fit parameters. Since the transmission happens at the bifurcation point —$\tilde{\alpha} = 1/4\tilde{\kappa}_{\rm lgm}^2$—integration of Equation 5 at the corresponding value of parameters is needed. Integration yields:

$$\tilde{t} = \frac{1}{\tilde{\beta}}\left[\ln\left|\frac{\tilde{N}(\tilde{t})(\tilde{N}_0 - 2\tilde{\kappa}_{\rm lgm})}{\tilde{N}_0(\tilde{N}(\tilde{t}) - 2\tilde{\kappa}_{\rm lgm})}\right| - \left(\frac{2\tilde{\kappa}_{\rm lgm}}{\tilde{N}(\tilde{t}) - 2\tilde{\kappa}_{\rm lgm}} - \frac{2\tilde{\kappa}_{\rm lgm}}{\tilde{N}_0 - 2\tilde{\kappa}_{\rm lgm}}\right)\right]. \tag{6}$$

where $\tilde{N}_0 = \tilde{N}(\tilde{t} = 0)$.

In general, the BSM has three extrema in $d\tilde{N}/d\tilde{t}$ at $\tilde{N}(\tilde{t}) = 0$, $\tilde{N}(\tilde{t}) = 2\tilde{\kappa}_{\rm lgm}/3$, and $\tilde{N}(\tilde{t}) = 2\tilde{\kappa}_{\rm lgm}$. For the solution at the bifurcation point, there are only two extrema which are the first two, and the maximum is then at $\tilde{N}(\tilde{t}) = 2\tilde{\kappa}_{\rm lgm}/3$. Putting all of these together, one finds from Equation 6,

$$\tilde{t}_{\rm peak} = \frac{1}{\tilde{\beta}}\left[\ln\left(\frac{2\tilde{\kappa}_{\rm lgm} - \tilde{N}_0}{2\tilde{N}_0}\right) + \frac{2\tilde{\kappa}_{\rm lgm} - 3\tilde{N}_0}{4\tilde{\kappa}_{\rm lgm} - 2\tilde{N}_0}\right]. \tag{7}$$

On using the fit parameters for transmission of Type II shown in Figure 2, we get $\tilde{t}_{\rm peak} = 0.85$ from Equation 7. Since $\tilde{N}_0 \ll \tilde{\kappa}_{\rm lgm}$, we can simplify the expression for $\tilde{t}_{\rm peak}$ to:

$$\tilde{t}_{\rm peak} = \frac{1}{\tilde{\beta}}\left[\ln\left(\frac{\tilde{\kappa}_{\rm lgm}}{\tilde{N}_0}\right) + \frac{1}{2}\right]. \tag{8}$$

Numerical solution of BSM

The fit to the BSM was performed by numerical integration of the solution to the differential equation given by Equation 5. For this we used the SciPy package [61].

Deep neural network architecture

To construct the DNN used in the model-agnostic fit we used TensorFlow [62]. The following are the details of the DNN architecture:

- depth: 3 layers
- width: 16 nodes in each layer
- optimizer: Adam with a learning rate of 0.01
- loss function: mean square error
- regularization: early stopping
- validation split: 0.2

Since the data for $\tilde{N}(\tilde{t})$ is already scaled between 0 and 1, no scaling of the data was necessary. The DNN took about 100-200 epochs to train with the regression taking only a few seconds to be completed. A more detailed discussion of the architechture can be found in the following sections.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Algorithm for predictions

Once we have the fit to the BSM or have trained the DNN with phase-1 data we can make predictions for the subsequent phases. The procedure we follow for this is given by:

1. Choose a starting date where the cases start to rise (the start-date) from the plot of $dN(t)/dt$ for the second or any subsequent phase.
2. Subtract the number of cases up until the start-date from the total number of cases to get the adjusted number of cases.
3. Locate the time $t = t_{\text{peak}}$ at which the peak of $dN(t)/dt$ occurs with $t = 0$ at the start-date.
4. The time t_{peak} is proportional to $t_{1/2}$ with the constant of proportionality being approximately 1 for Type I transmission and a constant (approximately 0.85) extracted from data for phase 1 of the spread for transmission of Type II.
5. Find the adjusted number of cases on the date corresponding to $t_{1/2}$. Twice that number is the adjusted N_{max}.
6. Use the functions fit from data for phase 1 (the BSM for model-dependent predictions and DNN for model-agnostic predictions) to estimate the future adjusted number of cases.
7. Now add back the number of cases up until the start-date to get the prediction of the disease spread.

Appendix: Universality in physical systems

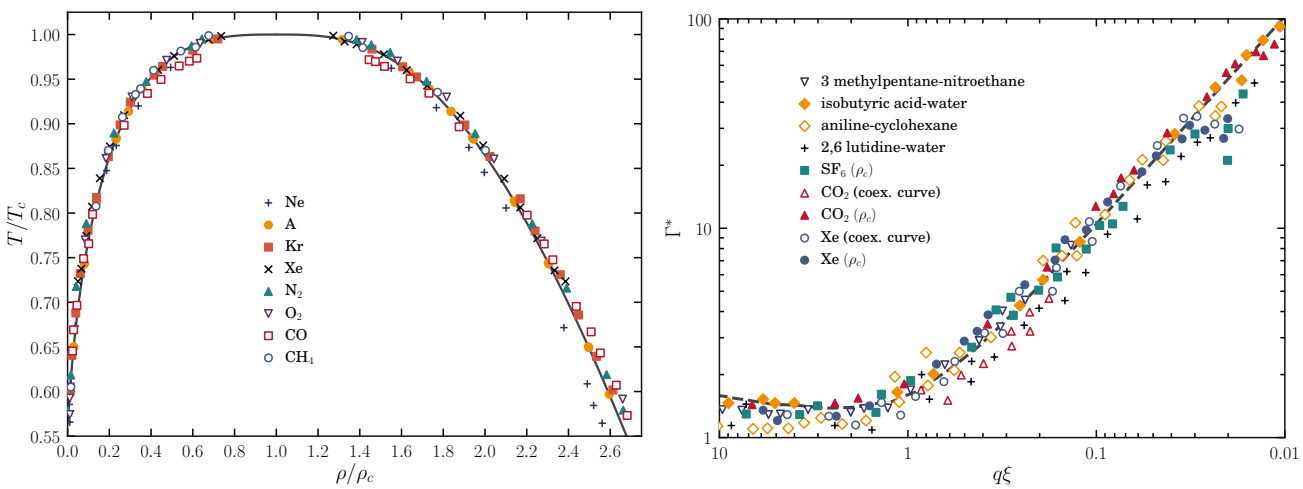

Figure 4. ***The emergence of universality in fluid systems.*** *The left panel is a reproduction of figure 2 of Ref. [63] which shows data collapse in the coexistence of liquid and vapour states for several fluids. The right panel is a reproduction of figure 10 of Ref. [64] and shows dynamical scaling for several fluids.*

In this section we discuss some examples of data collapse seen in physical system. The emergence of universality accompanied by a data collapse is quite common and has been observed for more than half a century. We begin our exposition with data collapse in the study of critical phenomenon [35, 36, 65] in fluids. We intend to keep the discussion only at a representative level to give some introduction to universality and data collapse and not make it an exhaustive survey which is not within the scope of this paper.

It is well known that if a gas is subjected to isothermal compression, then at a certain pressure it undergoes a phase transition to the liquid state while the pressure remains constant. Upon completion of the phase transition, the pressure increases very sharply on further compression while the volume increases too. If the temperature is raised and the process repeated, the same phenomenon occurs except that the parameter space over which the pressure and volume simultaneously rise grows smaller and smaller as the temperature is raised. At a particular temperature, $T = T_c$, this parameter space reduces to a point and at higher temperatures the liquefaction does not occur however high the pressure is. The temperature, T_c, is the critical temperature and in its vicinity the fluid shows a large amount of fluctuation as it lies between a state of reasonable order, the liquid state, and a state of disorder, which is the gaseous state.

The fluctuations in density very close to the critical point actually mean that some short-lived liquid state bubbles are produced in the gas. As the critical point is reached these fluctuations become long lived and infinitely long ranged as one has a liquid-vapour coexistence. The infinite lifetime of fluctuations and the infinite range of correlation mean that this phenomenon

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

cannot be sensitive to which gas is being considered. Whether it be Carbon Dioxide, Oxygen, Nitrogen, Xenon or Sulphur Hexafluoride, they will all behave in the same way. This is known as universality and, hence, means that if we draw the equation of state of any gas near the critical point with pressure scaled by the critical pressure and the volume scaled by the critical volume, then we will get just one curve. The critical volumes, critical temperatures and critical pressures are all different for the different materials but the scaled equation of state is the same as can be seen from Figure 4. If P, V and T are the pressure, volume and temperature of a fluid with a and b being constants specific to the fluid, then

$$\left(p+\frac{a}{V^2}\right)(V-b)=RT \longrightarrow \left(\pi+\frac{3}{\phi^2}\right)(3\phi-1)=8\tau, \tag{9}$$

where $\pi = P/P_c$, $\phi = V/V_c$ and $\tau = T/T_c$ with P_c, V_c and T_c being the pressure volume and temperature at the critical point. The equation of state thus collapses to a single equation independent of the nature of the fluid.

An identical situation holds for the dynamics which probes the lifetime of the fluctuations. The experimental technique for probing fluctuations directly is frequency resolved light scattering. The intensity I as a function of frequency, ω, is what is measured and at a given temperature and has a Lorentzian shape. As the temperature is lowered towards the critical point, both the peak (at $\omega = 0$) and the width at half-maximum increase with decreasing temperature, both headed for infinitely large values at the critical point. Different materials show different Lorentz distributions at different temperatures. If we denote the intensity at a particular frequency and a particular temperature, T, by $I(\omega,\Delta T)$, with $\Delta T = T - T_c$, then for different materials and different sets of ΔT one has different $I(\omega,\Delta T)$ vs ω distributions. However, if one plots $I(\omega,\Delta T)/I(0,\Delta T)$ vs. ω/ω_0 where ω_0 is the characteristic frequency for a given material, then the data for hundreds of individual distributions collapse on a single distribution. This phenomenon is called dynamic scaling [35, 36, 65] and is displayed in Figure 4.

This phenomenon of universal data collapse is ubiquitous: It appears in various forms in network dynamics [66], neuronal network [67], brain [68, 69], supercritical fluids [70], amorphous solids [71], Bose–Einstein condensate [72], jamming transition [73], granular media [74, 75], financial markets [76], metallic liquids [77], etc. Thus, in some sense, it is not surprising that this intriguing phenomenon be present in the transmission of infectious diseases as well.

Appendix: Comparisons between different models

In this section we compare some models that have been used commonly in epidemiology to quantify transmission dynamics of diseases including COVID-19 [53]. All the models that we describe here are variants of the logistic growth model. The logistic growth model (LGM) is defined as:

$$\frac{dN(t)}{dt} = \beta N(t)\left(1-\frac{N(t)}{\kappa}\right). \tag{10}$$

We define $\tilde{t} = t/t_{1/2}$ and $\tilde{N}(t) = N(t)/N_{\max}$, and by setting $\tilde{\beta} = \beta t_{1/2}$ and $\tilde{\kappa} = \kappa/N_{\max}$, we get the rescaled equation:

$$\frac{d\tilde{N}(\tilde{t})}{d\tilde{t}} = \tilde{\beta}\tilde{N}(\tilde{t})\left(1-\frac{\tilde{N}(\tilde{t})}{\tilde{\kappa}}\right). \tag{11}$$

We see that β scales as the time dimension and κ scales as the number of cases. This makes the dimensional quantities, $\tilde{\beta}$ and $\tilde{\kappa}$, independent of any characteristic timescale or size of the system. $\tilde{\kappa} = 1$ in the LGM.

The generalized logistic model (GLM) is given by:

$$\frac{dN(t)}{dt} = \beta N(t)^{\pi}\left(1-\frac{N(t)}{\kappa}\right). \tag{12}$$

Similar rescaling as above leads to the rescaled parameters $\tilde{\beta} = \beta t_{1/2} N_{\max}^{\pi-1}$ and $\tilde{\kappa} = \kappa/N_{\max}$ leading to the rescaled GLM being:

$$\frac{d\tilde{N}(\tilde{t})}{d\tilde{t}} = \tilde{\beta}\tilde{N}(\tilde{t})^{\pi}\left(1-\frac{\tilde{N}(\tilde{t})}{\tilde{\kappa}}\right). \tag{13}$$

The generalized Richards model (GRM) is given by:

$$\frac{dN(t)}{dt} = \beta N(t)^{\pi}\left[1-\left(\frac{N(t)}{\kappa}^{\alpha}\right)\right]. \tag{14}$$

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

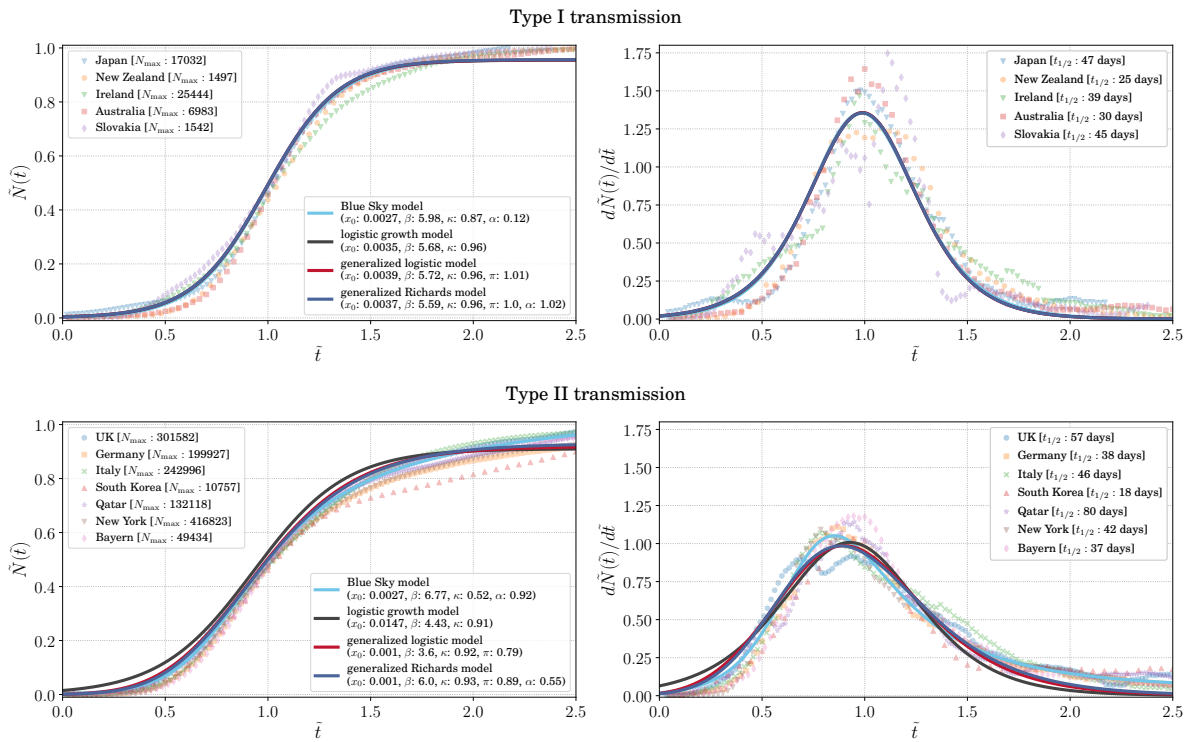

Figure 5. ***A comparison of the variants of the logistic model.*** *The panels on the left show $\tilde{N}(\tilde{t})$ vs. $\tilde{t}$ distribution for Type I and Type II transmission dynamics; and the right panels shows the $d\tilde{N}(\tilde{t})/d\tilde{t}$ vs. $\tilde{t}$ plot for the same. We fit four models to the data – the logistic growth model, the generalized logistic model, the generalized Richards model and the Blue Sky model. While the fit to data for all the models are the same for Type I transmission where all the models go to the limit of being a logistic growth model, for Type II transmission the Blue Sky model can explain the data better than the other models.*

For $\pi = 1$, it reduces to the standard Richards model [78]. We see that the scaling is similar to the GLM with $\tilde{\beta} = \beta t_{1/2} N_{\max}^{\pi-1}$ and $\tilde{\kappa} = \kappa/N_{\max}$; the rescaled GRM is:

$$\frac{d\tilde{N}(\tilde{t})}{d\tilde{t}} = \tilde{\beta}\tilde{N}(\tilde{t})^{\pi}\left[1-\left(\frac{\tilde{N}(\tilde{t})}{\tilde{\kappa}}\right)^{\alpha}\right]. \tag{15}$$

In Figure 5 we show a comparison between all the models we list here and the Blue Sky model (BSM). The two upper panels are for transmission of Type I. The LGM fits this data quite well and all the other models acquire the parameter values that take them to the limit of being an LGM. The two lower panels are for transmission of Type II. Here we clearly see that the BSM fits the data much better that the other models. Especially for the tail of distribution past the peak of the $d\tilde{N}(\tilde{t})/d\tilde{t}$ vs. $\tilde{t}$ plot where $\tilde{t} > \tilde{t}_{\text{peak}}$, the BSM model allows for larger $d\tilde{N}(\tilde{t})/d\tilde{t}$ compared to the other models hence being able to replicate transmission of Type II where the spread of the disease lingers on.

Appendix: The deep neural network architechture

In our model-agnostic approach we used a deep neural network (DNN) [79, 80] to fit the data. A DNN is a universal function generator. The reason we chose a DNN over other machine learning frameworks (i.e. decisions trees, random forests, SVM etc.) is the ease with which it can be use for a regression to noisy data to extract the underlying function. We will describe in brief the choice of the network architechture.

We started with a small DNN with 1 input node for $\tilde{t}$, 2 hidden layers, each with 4 nodes and 1 output layer for $\tilde{N}(\tilde{t})$. All layers were fully connected as shown in Figure 6. The activation function used was the *sigmoid* function. The solution to the LM is a *sigmoid* function and hence the transmission of Type 1 can be fairly explained with a just one sigmoid function that connects the input $\tilde{t}$ to the output $\tilde{N}(\tilde{t})$. This is not possible though for transmission of Type II which shows clear deviation from the LM. However, in the model agnostic approach we did not want to introduce any model biases through assumptions of functional form. Hence we used a larger (and deeper) neural network with 33 parameters to begin with. We scaled up the

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

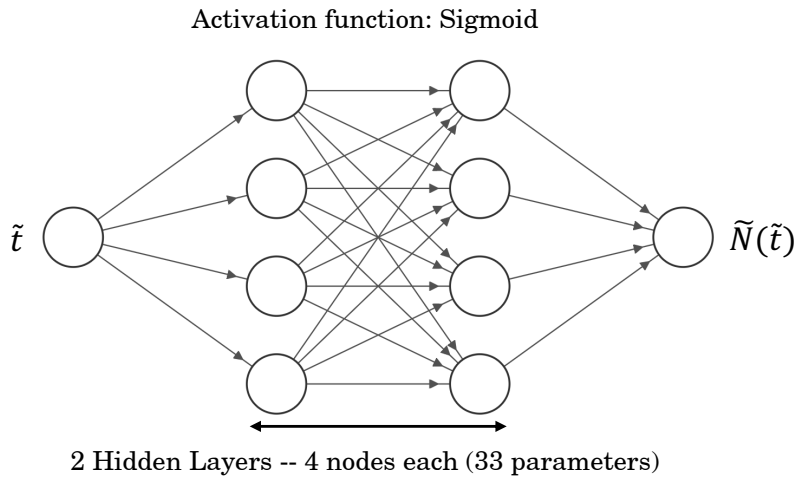

Figure 6. ***A representative architechture of the DNN used for the model-agnostic analysis to fit $\tilde{N}(\tilde{t})$ as a function of $\tilde{t}$.*** *For the final analysis we used a similar architechture with 3 layers each with 16 nodes to keep the analysis as model-agnostic as possible. To regulate the training of the DNN we used an early callback procedure which stops the training when the validation loss stops decreasing thus avoiding training biases.*

network to 3 hidden layers each with 16 nodes to check for the stability of our approach. The regression with the DNN with 2 hidden layers with 4 nodes each gave almost the same fit as the one with 3 hidden layers with 16 nodes each having a total of 593 parameters. To keep our method as general as possible and not introduce model biases from the activation function we decided to use the larger DNN with 3 hidden layers each having 16 nodes.

In addition, we tested the DNN with other activation function like *tanh* and *relu*. For *tanh* a somewhat wider and deeper network was needed to fit to the data. With the *relu* activation the network needed to be very large. Hence, we used the *sigmoid* activation function for the final DNN architecture.

Even though the DNN we used was not so large we used an early callback algorithm to stop the training of the network when the validation loss stopped decreasing to avoid overtraining and chose the epoch where the validation loss was the least. The loss function that we used was a simple mean square error function which measures the L_2 distance of the fit from the data. The training time was a few seconds with the early stopping coming into effect at about 100-150 epochs. The trained DNN was used for making prediction for the ongoing phases

References

1. Li, Q. *et al.* Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *New Engl. J. Medicine* **382**, 1199–1207, DOI: 10.1056/NEJMoa2001316 (2020). PMID: 31995857.
2. Dehning, J. *et al.* Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* **369**, DOI: 10.1126/science.abb9789 (2020).
3. Wilder-Smith, A. & Freedman, D. O. Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. *J. Travel. Medicine* **27**, DOI: 10.1093/jtm/taaa020 (2020).
4. Sohrabi, C. *et al.* World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *Int. J. Surg.* **76**, 71 – 76, DOI: 10.1016/j.ijsu.2020.02.034 (2020).
5. Zhang, S. *et al.* Covid-19 containment: China provides important lessons for global response. *Front. Medicine* **14**, 215–219, DOI: 10.1007/s11684-020-0766-9 (2020).
6. Fisher, D. & Wilder-Smith, A. The global community needs to swiftly ramp up the response to contain COVID-19. *The Lancet* **395**, 1109–1110, DOI: 10.1016/S0140-6736(20)30679-6 (2020).
7. Fraser, C., Riley, S., Anderson, R. M. & Ferguson, N. M. Factors that make an infectious disease outbreak controllable. *Proc. Natl. Acad. Sci.* **101**, 6146–6151, DOI: 10.1073/pnas.0307506101 (2004).
8. Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261, DOI: 10.1038/s41586-020-2405-7 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

9. Perra, N. Non-pharmaceutical interventions during the COVID-19 pandemic: a rapid review. *arXiv e-prints* arXiv:2012.15230 (2020). arXiv:2012.15230.

10. Kühn, M. J. *et al.* Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution. *medRxiv* DOI: 10.1101/2020.12.18.20248509 (2020).

11. Zhou, Y. *et al.* Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. *The Lancet Digit. Heal.* **2**, e417–e424, DOI: 10.1016/S2589-7500(20)30165-5 (2020).

12. Brett, T. S. & Rohani, P. Transmission dynamics reveal the impracticality of COVID-19 herd immunity strategies. *Proc. Natl. Acad. Sci.* **117**, 25897–25903, DOI: 10.1073/pnas.2008087117 (2020).

13. Gatto, M. *et al.* Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc. Natl. Acad. Sci.* **117**, 10484–10491, DOI: 10.1073/pnas.2004978117 (2020).

14. Bertozzi, A. L., Franco, E., Mohler, G., Short, M. B. & Sledge, D. The challenges of modeling and forecasting the spread of COVID-19. *Proc. Natl. Acad. Sci.* **117**, 16732–16738, DOI: 10.1073/pnas.2006520117 (2020).

15. López, L. & Rodó, X. The end of social confinement and COVID-19 re-emergence risk. *Nat. Hum. Behav.* **4**, 746–755, DOI: 10.1038/s41562-020-0908-8 (2020).

16. Shayak, B. & Sharma, M. M. A New Approach to the Dynamic Modeling of an Infectious Disease. *medRxiv* DOI: 10.1101/2020.10.30.20223305 (2020).

17. Dellâ Anna, L. Solvable delay model for epidemic spreading: the case of Covid-19 in Italy. *Sci. Reports* **10**, 15763, DOI: 10.1038/s41598-020-72529-y (2020).

18. Vyasarayani, C. & Chatterjee, A. New approximations, and policy implications, from a delayed dynamic model of a fast pandemic. *Phys. D: Nonlinear Phenom.* **414**, 132701, DOI: 10.1016/j.physd.2020.132701 (2020).

19. Young, L.-S., Ruschel, S., Yanchuk, S. & Pereira, T. Consequences of delays and imperfect implementation of isolation in epidemic control. *Sci. Reports* **9**, 3505, DOI: 10.1038/s41598-019-39714-0 (2019).

20. Kreck, M. & Scholz, E. Studying the course of Covid-19 by a recursive delay approach. *medRxiv* DOI: 10.1101/2021.01.18.21250012 (2021).

21. Cacciapaglia, G., Cot, C. & Sannino, F. Second wave COVID-19 pandemics in Europe: a temporal playbook. *Sci. Reports* **10**, 15514, DOI: 10.1038/s41598-020-72611-5 (2020).

22. Wong, F. & Collins, J. J. Evidence that coronavirus superspreading is fat-tailed. *Proc. Natl. Acad. Sci.* **117**, 29416–29418, DOI: 10.1073/pnas.2018490117 (2020).

23. Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M. & Prokopenko, M. Modelling transmission and control of the COVID-19 pandemic in Australia. *Nat. Commun.* **11**, 5710, DOI: 10.1038/s41467-020-19393-6 (2020).

24. Maziarz, M. & Zach, M. Agent-based modelling for SARS-CoV-2 epidemic prediction and intervention assessment: A methodological appraisal. *J. Eval. Clin. Pract.* **26**, 1352–1360, DOI: 10.1111/jep.13459 (2020).

25. Tatapudi, H., Das, R. & Das, T. K. Impact assessment of full and partial stay-at-home orders, face mask usage, and contact tracing: An agent-based simulation study of COVID-19 for an urban region. *Glob. Epidemiol.* **2**, 100036, DOI: 10.1016/j.gloepi.2020.100036 (2020).

26. Rockett, R. J. *et al.* Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling. *Nat. Medicine* **26**, 1398–1404, DOI: 10.1038/s41591-020-1000-7 (2020).

27. Aleta, A. *et al.* Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat. Hum. Behav.* **4**, 964–971, DOI: 10.1038/s41562-020-0931-9 (2020).

28. Hinch, R. *et al.* OpenABM-Covid19 - an agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. *medRxiv* DOI: 10.1101/2020.09.16.20195925 (2020).

29. Thurner, S., Klimek, P. & Hanel, R. A network-based explanation of why most COVID-19 infection curves are linear. *Proc. Natl. Acad. Sci.* **117**, 22684–22689, DOI: 10.1073/pnas.2010398117 (2020).

30. Thomas, L. J. *et al.* Spatial heterogeneity can lead to substantial local variations in COVID-19 timing and severity. *Proc. Natl. Acad. Sci.* **117**, 24180–24187, DOI: 10.1073/pnas.2011656117 (2020).

31. Merow, C. & Urban, M. C. Seasonality and uncertainty in global COVID-19 growth rates. *Proc. Natl. Acad. Sci.* **117**, 27456–27464, DOI: 10.1073/pnas.2008590117 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

32. Perkins, T. A. *et al.* Estimating unobserved SARS-CoV-2 infections in the United States. *Proc. Natl. Acad. Sci.* **117**, 22597–22602, DOI: 10.1073/pnas.2005476117 (2020).

33. Chatterjee, S. *et al.* Evolution of COVID-19 pandemic: Power-law growth and saturation. *J. Indian Stat. Assoc.* **55**, 1–3 (2020).

34. Siegenfeld, A. F., Taleb, N. N. & Bar-Yam, Y. Opinion: What models can and cannot tell us about COVID-19. *Proc. Natl. Acad. Sci.* **117**, 16092–16095, DOI: 10.1073/pnas.2011542117 (2020).

35. Hohenberg, P. C. & Halperin, B. I. Theory of dynamic critical phenomena. *Rev. Mod. Phys.* **49**, 435–479, DOI: 10.1103/RevModPhys.49.435 (1977).

36. Stanley, H. E. Scaling, universality, and renormalization: Three pillars of modern critical phenomena. *Rev. Mod. Phys.* **71**, S358–S366, DOI: 10.1103/RevModPhys.71.S358 (1999).

37. Abraham, R. & Shaw, C. D. *Dynamics: The Geometry of Behavior. Part 4: Bifurcation Behavior* (Ariel Press, Santa Cruz, CA, 1988).

38. Rao, S. D., Tylavsky, D. J. & Feng, Y. Estimating the saddle-node bifurcation point of static power systems using the holomorphic embedding method. *Int. J. Electr. Power & Energy Syst.* **84**, 1 – 12, DOI: 10.1016/j.ijepes.2016.04.045 (2017).

39. Achleitner, F. & Szmolyan, P. Saddle–node bifurcation of viscous profiles. *Phys. D: Nonlinear Phenom.* **241**, 1703 – 1717, DOI: 10.1016/j.physd.2012.06.008 (2012).

40. Cessac, B. & Samuelides, M. From neuron to neural networks dynamics. *The Eur. Phys. J. Special Top.* **142**, 7–88, DOI: 10.1140/epjst/e2007-00058-2 (2007).

41. Joyeux, M., Farantos, S. C. & Schinke, R. Highly Excited Motion in Molecules: Saddle-Node Bifurcations and Their Fingerprints in Vibrational Spectra. *The J. Phys. Chem. A* **106**, 5407–5421, DOI: 10.1021/jp0131065 (2002).

42. Cassak, P. A., Shay, M. A. & Drake, J. F. A saddle-node bifurcation model of magnetic reconnection onset. *Phys. Plasmas* **17**, 062105, DOI: 10.1063/1.3435269 (2010).

43. Descalzi, O., Argentina, M. & Tirapegui, E. Saddle-node bifurcation: Appearance mechanism of pulses in the subcritical complex Ginzburg-Landau equation. *Phys. Rev. E* **67**, 015601, DOI: 10.1103/PhysRevE.67.015601 (2003).

44. Mitze, T., Kosfeld, R., Rode, J. & Wälde, K. Face masks considerably reduce COVID-19 cases in Germany. *Proc. Natl. Acad. Sci.* **117**, 32293–32301, DOI: 10.1073/pnas.2015954117 (2020).

45. Lega, J. & Brown, H. E. Data-driven outbreak forecasting with a simple nonlinear growth model. *Epidemics* **17**, 19 – 26, DOI: 10.1016/j.epidem.2016.10.002 (2016).

46. Buczak, A. L., Koshute, P. T., Babin, S. M., Feighner, B. H. & Lewis, S. H. A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data. *BMC Med. Informatics Decis. Mak.* **12**, 124, DOI: 10.1186/1472-6947-12-124 (2012).

47. Brauer, F. *Compartmental Models in Epidemiology*, 19–79 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008).

48. Adam, D. Special report: The simulations driving the world's response to COVID-19. *Nature* **580**, 316–318, DOI: 10.1038/d41586-020-01003-6 (2020).

49. Allen, L. J. S. *An Introduction to Stochastic Epidemic Models*, 81–130 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008).

50. Verhulst, P.-F. Notice sur la loi que la population poursuit dans son accroissement. *Corresp. Mathématique et Physique* **10**, 113–121 (1838).

51. Chen, D.-G., Chen, X. & Chen, J. K. Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model. *Glob. Heal. Res. Policy* **5**, 25, DOI: 10.1186/s41256-020-00152-5 (2020).

52. Shen, C. Y. Logistic growth modelling of COVID-19 proliferation in China and its international implications. *Int. J. Infect. Dis.* **96**, 582 – 589, DOI: 10.1016/j.ijid.2020.04.085 (2020).

53. Wu, K., Darcet, D., Wang, Q. & Sornette, D. Generalized logistic growth modeling of the COVID-19 outbreak: comparing the dynamics in the 29 provinces in China and in the rest of the world. *Nonlinear Dyn.* **101**, 1561–1581, DOI: 10.1007/s11071-020-05862-6 (2020).

54. Aviv-Sharon, E. & Aharoni, A. Generalized logistic growth modeling of the COVID-19 pandemic in Asia. *Infect. Dis. Model.* **5**, 502 – 509, DOI: 10.1016/j.idm.2020.07.003 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

55. Jain, M., Bhati, P. K., Kataria, P. & Kumar, R. Modelling Logistic Growth Model for COVID-19 Pandemic in India. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, 784–789, DOI: 10.1109/ICCES48766.2020.9138049 (2020).

56. Malavika, B. *et al.* Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models. *Clin. Epidemiol. Glob. Heal.* **9**, 26 – 33, DOI: 10.1016/j.cegh.2020.06.006 (2021).

57. Zou, Y. *et al.* Outbreak analysis with a logistic growth model shows COVID-19 suppression dynamics in China. *PLoS ONE* **15**, e0235247, DOI: 10.1371/journal.pone.0235247 (2020).

58. Pelinovsky, E., Kurkin, A., Kurkina, O., Kokoulina, M. & Epifanova, A. Logistic equation and COVID-19. *Chaos, Solitons & Fractals* **140**, 110241, DOI: 10.1016/j.chaos.2020.110241 (2020).

59. Wang, P., Zheng, X., Li, J. & Zhu, B. Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. *Chaos, Solitons & Fractals* **139**, 110058, DOI: 10.1016/j.chaos.2020.110058 (2020).

60. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infect. Dis.* **20**, 533–534, DOI: 10.1016/S1473-3099(20)30120-1 (2020).

61. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272, DOI: 10.1038/s41592-019-0686-2 (2020).

62. Abadi, M. *et al.* TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (2015). Software available from tensorflow.org.

63. Guggenheim, E. A. The Principle of Corresponding States. *The J. Chem. Phys.* **13**, 253–261, DOI: 10.1063/1.1724033 (1945).

64. Swinney, H. L. & Henry, D. L. Dynamics of Fluids near the Critical Point: Decay Rate of Order-Parameter Fluctuations. *Phys. Rev. A* **8**, 2586–2617, DOI: 10.1103/PhysRevA.8.2586 (1973).

65. LÜbeck, S. Universal scaling behavior of non-equilibrium phase transitions. *Int. J. Mod. Phys. B* **18**, 3977–4118, DOI: 10.1142/S0217979204027748 (2004).

66. Barzel, B. & Barabási, A.-L. Universality in network dynamics. *Nat. Phys.* **9**, 673–681, DOI: 10.1038/nphys2741 (2013).

67. Friedman, N. *et al.* Universal Critical Dynamics in High Resolution Neuronal Avalanche Data. *Phys. Rev. Lett.* **108**, 208102, DOI: 10.1103/PhysRevLett.108.208102 (2012).

68. Bhattacharya, J. & Petsche, H. Universality in the Brain While Listening to Music. *Proceedings: Biol. Sci.* **268**, 2423–2433 (2001). URL: http://www.jstor.org/stable/3067747.

69. Serafino, M. *et al.* True scale-free networks hidden by finite size effects. *Proc. Natl. Acad. Sci.* **118**, DOI: 10.1073/pnas.2013825118 (2021).

70. Ha, M. Y., Yoon, T. J., Tlusty, T., Jho, Y. & Lee, W. B. Universality, Scaling, and Collapse in Supercritical Fluids. *The J. Phys. Chem. Lett.* **11**, 451–455, DOI: 10.1021/acs.jpclett.9b03360 (2020).

71. Ulrich, S., Upadhyaya, N., Opheusden, B. v. & Vitelli, V. Shear shocks in fragile networks. *Proc. Natl. Acad. Sci.* **110**, 20929–20934, DOI: 10.1073/pnas.1314468110 (2013).

72. Klinder, J., Keßler, H., Wolke, M., Mathey, L. & Hemmerich, A. Dynamical phase transition in the open Dicke model. *Proc. Natl. Acad. Sci.* **112**, 3290–3295, DOI: 10.1073/pnas.1417132112 (2015).

73. Goodrich, C. P., Liu, A. J. & Sethna, J. P. Scaling ansatz for the jamming transition. *Proc. Natl. Acad. Sci.* **113**, 9745–9750, DOI: 10.1073/pnas.1601858113 (2016).

74. Manna, S. S. Two-state model of self-organized criticality. *J. Phys. A: Math. Gen.* **24**, L363–L369, DOI: 10.1088/0305-4470/24/7/009 (1991).

75. Denisov, D. V., Lörincz, K. A., Uhl, J. T., Dahmen, K. A. & Schall, P. Universality of slip avalanches in flowing granular matter. *Nat. Commun.* **7**, 10641, DOI: 10.1038/ncomms10641 (2016).

76. Denys, M., Gubiec, T., Kutner, R., Jagielski, M. & Stanley, H. E. Universality of market superstatistics. *Phys. Rev. E* **94**, 042305, DOI: 10.1103/PhysRevE.94.042305 (2016).

77. Blodgett, M. E., Egami, T., Nussinov, Z. & Kelton, K. F. Proposal for universality in the viscosity of metallic liquids. *Sci. Reports* **5**, 13837, DOI: 10.1038/srep13837 (2015).

78. Richards, F. J. A Flexible Growth Function for Empirical Use. *J. Exp. Bot.* **10**, 290–301, DOI: 10.1093/jxb/10.2.290 (1959).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.29.21250750; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

79. Bengio, Y., Courville, A. & Vincent, P. Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis Mach. Intell.* **35**, 1798–1828, DOI: 10.1109/TPAMI.2013.50 (2013).

80. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Networks* **61**, 85 – 117, DOI: 10.1016/j.neunet.2014.09.003 (2015).