medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 .

Noname manuscript No.
(will be inserted by the editor)

Evolution of COVID-19 pandemic: Power-law growth and saturation

Soumyadeep Chatterjee* · **B. Shayak** · **Ali Asad** · **Shashwat Bhattacharya** · **Shadab Alam** · **Mahendra K. Verma**

Received: date / Accepted: date

Abstract In this paper, we analyze the real-time infection data of COVID-19 epidemic for 21 nations up to May 18, 2020. For most of these nations, the total number of infected individuals exhibits a succession of exponential growth and power-law growth before the flattening of the curve. In particular, we find a universal $\sqrt{t}$ growth before they reach saturation. India, Singapore, and Sri Lanka have reached up to linear growth ($I(t) \sim t$), and they are yet to flatten their curves. Russia and Brazil are still in the power-law (t^2) growth regime. Thus, the polynomials of the $I(t)$ curves provide valuable information on the stage of the epidemic evolution. Besides these detailed analyses, we compare the predictions of an extended SEIR model and a delay differential equation-based model with the reported infection data and observed good agreement among them, including the $\sqrt{t}$ behaviour.

Keywords Covid-19 · Epidemic · SEIR model · Power law

Soumyadeep Chatterjee*
Department of Physics, Indian Institute of Technology Kanpur, Kanpur 208016, India
E-mail: soumyade@iitk.ac.in

B. Shayak
Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, USA
E-mail: shayak.2015@iitkalumni.org

Ali Asad
Department of Physics, Indian Institute of Technology Kanpur, Kanpur 208016, India
E-mail: aliasad@iitk.ac.in

Shashwat Bhattcharya
Department of Mechanical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India
E-mail: shabhatt@iitk.ac.in

Shadab Alam
Department of Mechanical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India
E-mail: shadab@iitk.ac.in

Mahendra K. Verma
Department of Physics, Indian Institute of Technology Kanpur, Kanpur 208016, India
Tel.: +512-2597396
E-mail: mkv@iitk.ac.in

1 Introduction

As of May 23, 2020, COVID-19 pandemic has infected more than 5.3 million of the human population and caused 0.34 million deaths. The world economy is in tatters. Therefore, understanding the progression of the pandemic is extremely crucial. In the present paper, we analyze the publicly-available national COVID-19 infection data [37] up to May 18, 2020. We observe that the COVID-19 infection curves for many nations exhibit power-law growth after exponential growth. We compare the reported data with model predictions and observe a good agreement among them.

To understand and forecast epidemics, epidemiologists have made many models [15, 4, 10]. One of the first models is called the SIR model, where the variables S and I describe the numbers of susceptible and infected individuals, respectively. The third variable R represents the removed individuals who have either recovered or died. An advanced model, called SEIR model, includes *exposed* individuals, E, who are infected but not yet infectious [4, 10].

SARS-CoV-2 is one of the seven human coronaviruses which have been identified so far. It is the most dangerous among all of these because of its highly infectious nature and its lethality. *Asymptomatic carriers*, individuals who do not exhibit any symptoms, have carried the virus to far off places where it has spread

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

rapidly [19]. Even symptomatic patients manifest symptoms two to three days after turning transmissible. To stop the spread of the deadly virus, various nations have employed lockdowns, mandatory social distancing, quarantines for the affected, etc.

Despite the difficulties stated above, many models are able to describe the COVID-19 pandemic data quite well. Peng et al. [27] constructed a generalized SEIR model with seven-variables (including quarantined and death variables) for the epidemic spread in China. Their predictions are in good agreement with the present data. López and Rodo [20] formulated an extended version of this model to analyse the spread of the pandemic in Spain and Italy. Earlier, Cheynet [7,8] had developed a code to simulate this model. Hellewell et al. [13] studied the effects of isolation on controlling the COVID-19 epidemic. Chinazzi et al. [9] analyzed the effects of travel restrictions on the spread of COVID-19 in China and in the world using the global metapopulation disease transmission model. Mandal et al. [23] constructed model for devising intervention strategies in India. Shayak et al. [34] have constructed delay differential equation (DDE) model for the spread of COVID-19; this model takes into account the pre-symptomatic period and predicts a path to the end of the epidemic.

Due to the above complex issues in the epidemic models of COVID-19, many researchers have consciously focussed on the data and attempted to extract useful information from them. It has been observed that the analysis of the pandemic provides important clues that may be useful for its forecast. In particular, Ziff and Ziff [39], Komarova and Wodarz [16], Manchein et al. [22], Blasius [5], Marsland and Mehta [25], Li et al. [18], Singer [35], Beare and Toda [2], and Cherednik and Hill [6] analyzed the reported count of total infections ($I(t)$) in various nations and observed power-law growth after the exponential regime. Verma et al. [36] analyzed the data of 9 nations up to April 7, 2020 and showed that the $I(t)$ goes through power laws, t^3, t^2, t and $\sqrt{t}$, in temporal sequence before flattening out. By April 7, 2020, China and South Korea had flattened the $I(t)$ curves, but other nations were either in the exponential regime or in the power-law regime. However, by May 18, many nations are close to the flattening of their epidemic curves. Besides the above results, Prakash et al. [31] reported a linear growth of $I(t)$ after early exponential growth.

Schüttler et al. [33] analyzed the daily death counts for various nations and observed that their probability distributions appear to follow a Gaussian profile. Marsland and Mehta [25] observed that the error function provides best fit to the total count, $I(t)$; this observation follows from Schüttler et al. [33]'s analysis.

There are epidemic growth models based on population growth [10,38]. COVID-19 spread via asymptomatic carriers leads to a network formation. Hence, network-based epidemic growth models may be useful for modelling COVID-19 pandemic. Marathe and Vullikanti [24] review computational epidemiology with a focus on epidemic spread over a network.

In this paper, we analyzed the COVID-19 infection data up to May 18, 2020 for 21 nations and observed that all the nations are following transition from exponential to power-law growth in infection counts. Many of the 21 nations are close to flattening their curves with several exceptions (for example, Russia). We also showed that three epidemics—Ebola, COVID-19, MERS—have similar evolution: exponential growth, power-law growth, and then flattening of the curve. In addition, we compared the predictions of an extended SEIR model [20] and a delay-differential equation model [34] with the real-time data and observed good agreement among them.

The structure of the paper is as follows: in Sec. 2 we analyse the COVID-19 data for 21 leading nations and observe power law growth for them after the exponential growth. The evolutions of Ebola, MERS, and COVID-19 are compared in Sec. 3. The predictions of two models of COVID-19 pandemic are compared with the observed data in Sections 4 and 5. We conclude in Sec. 6.

2 Data analysis of COVID-19 epidemic

In this section, we present our results based on a comprehensive data analysis of COVID-19 cases for 21 countries (see Table 1) up to May 18, 2020. The majority of the countries in our analysis include those with a large number of COVID-19 cases, including USA, Italy, Germany, China, and India. For a complete study, we also include countries with a relatively smaller number of cases such as Sri Lanka and Hong Kong. We used the real-time data available at *worldOmeter* [37] and chose the starting date (see Table 1) as the one from which the number of infected cases increased rapidly. Corona Resource Center [14] too is an important repository for COVID-19 data.

We analyze the evolution of cumulative number of infected cases, which is denoted by $I(t)$, with time in days, denoted by t. For all $I(t)$ curves, we compute the derivatives $\dot{I}(t)$ using Python's *gradient* function. These derivatives indicate the daily count of the infected cases. Note that $\dot{I}(t)$ exhibit lower fluctuations than the measured daily counts due to smoothing. In Fig. 1, we exhibit the plots of $I(t)$ (red curves) and $\dot{I}(t)$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

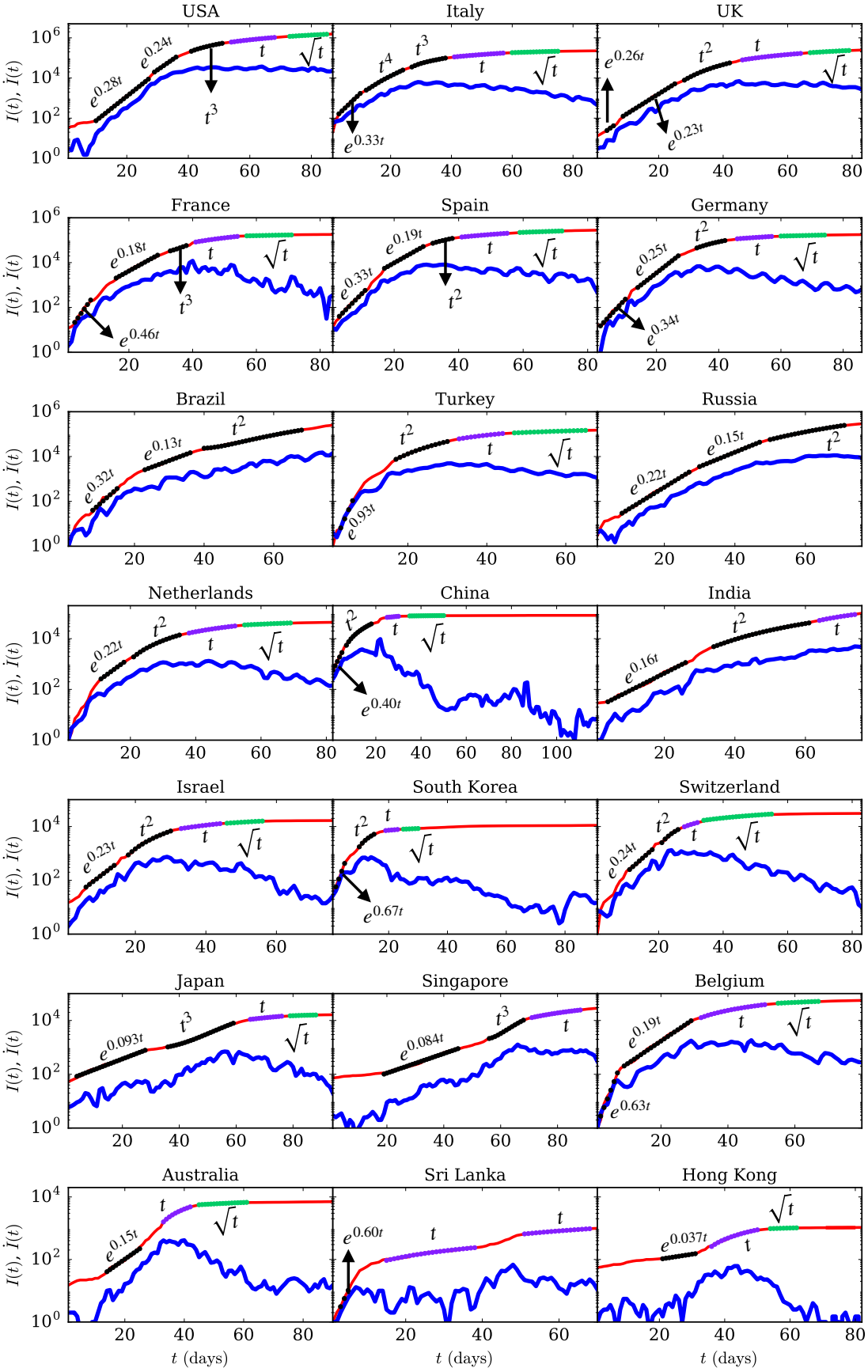

Fig. 1 (color online) For the COVID-19 epidemic, the *semi-logy* plots of total infected individuals ($I(t)$) vs. time (t) (red curves) for 21 countries. We also plot $\dot{I}(t)$ vs. t (blue curves). The dotted curves represent the best-fit curves, which are exponential or polynomial functions.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

Table 1 The best-fit curves along with their relative errors for the COVID-19 data for 21 countries.. The best-fit curves are shown in Figs. 1. The start date for the plots of Fig. 1 are listed below the names of the countries.

Country (Start Date)	Best-fit functions and errors	Country (Start Date)	Best-fit functions and errors
USA (February 21)	1) $4.6e^{0.28t}$ ($\pm4.1\%$) 2) $17e^{0.24t}$ ($\pm3.5\%$) 3) $8.5t^3 - 970t^2 + 67000t - 15 \times 10^5$ ($\pm0.46\%$) 4) $30000t - 1 \times 10^6$ ($\pm0.44\%$) 5) $48 \times 10^4\sqrt{t} - 29 \times 10^5$ ($\pm0.2\%$)	Italy (February 21)	1) $60e^{0.33t}$ ($\pm3.2\%$) 2) $0.51t^4 - 30t^3 + 770t^2 - 8500t + 35000$ ($\pm1.5\%$) 3) $7.8t^3 - 760t^2 + 30000t - 38 \times 10^4$ ($\pm0.61\%$) 4) $3800t - 45000$ ($\pm0.61\%$) 5) $35000\sqrt{t} - 87000$ ($\pm0.46\%$)
UK (February 26)	1) $8.5e^{0.26t}$ ($\pm8.1\%$) 2) $16e^{0.23t}$ ($\pm5.1\%$) 3) $130t^2 - 5200t + 56000$ ($\pm1.7\%$) 4) $4900t - 14 \times 10^4$ ($\pm0.56\%$) 5) $74000\sqrt{t} - 43 \times 10^4$ ($\pm0.51\%$)	France (February 24)	1) $5.4e^{0.46t}$ ($\pm11\%$) 2) $110e^{0.18t}$ ($\pm6.5\%$) 3) $6.9t^3 - 360t^2 + 4200t + 40000$ ($\pm1.6\%$) 4) $4900t - 12 \times 10^4$ ($\pm1.3\%$) 5) $17000\sqrt{t} + 24000$ ($\pm0.44\%$)
Spain (February 26)	1) $15e^{0.33t}$ ($\pm8.1\%$) 2) $230e^{0.19t}$ ($\pm3.6\%$) 3) $27t^2 + 5900t - 14 \times 10^4$ ($\pm0.34\%$) 4) $4600t - 50000$ ($\pm0.64\%$) 5) $45000\sqrt{t} - 12 \times 10^4$ ($\pm0.39\%$)	Germany (February 23)	1) $7.6e^{0.34t}$ ($\pm11\%$) 2) $24e^{0.25t}$ ($\pm2.9\%$) 3) $22t^2 + 4100t - 11 \times 10^4$ ($\pm1.2\%$) 4) $2800t - 14000$ ($\pm0.58\%$) 5) $20000\sqrt{t} + 1100$ ($\pm0.38\%$)
Brazil (March 03)	1) $3.2e^{0.32t}$ ($\pm11\%$) 2) $120e^{0.14t}$ ($\pm3.6\%$) 3) $150t^2 - 12000t + 25 \times 10^4$ ($\pm3.2\%$)	Turkey (March 12)	1) $0.40e^{0.93t}$ ($\pm9.1\%$) 2) $110t^2 - 2300t + 14000$ ($\pm0.85\%$) 3) $4100t - 73000$ ($\pm0.75\%$) 4) $28000\sqrt{t} - 75000$ ($\pm0.47\%$)
Russia (March 04)	1) $5.1e^{0.22t}$ ($\pm5.8\%$) 2) $43e^{0.15t}$ ($\pm1.3\%$) 3) $190t^2 - 14000t + 29 \times 10^4$ ($\pm1.0\%$)	Netherlands (February 27)	1) $24e^{0.22t}$ ($\pm2.3\%$) 2) $27t^2 - 640t + 3600$ ($\pm2.6\%$) 3) $1100t - 24000$ ($\pm0.77\%$) 4) $7600\sqrt{t} - 21000$ ($\pm0.73\%$)
China (January 22)	1) $380e^{0.40t}$ ($\pm2.2\%$) 2) $89t^2 + 750t - 4100$ ($\pm2.8\%$) 3) $1400t - 34000$ ($\pm0.64\%$) 4) $2200\sqrt{t} + 65000$ ($\pm0.36\%$)	India (March 04)	1) $17e^{0.16t}$ ($\pm7.1\%$) 2) $29t^2 - 1400t + 19000$ ($\pm1.1\%$) 3) $3700t - 19 \times 10^4$ ($\pm0.54\%$)
Israel (March 04)	1) $14e^{0.23t}$ ($\pm6.0\%$) 2) $19t^2 - 440t - 2500$ ($\pm2.5\%$) 3) $400t - 4800$ ($\pm0.6\%$) 4) $3800\sqrt{t} - 12000$ ($\pm0.79\%$)	South Korea (February 18)	1) $15e^{0.67t}$ ($\pm2.2\%$) 2) $9.7t^2 + 430t - 3500$ ($\pm1.5\%$) 3) $160t + 4000$ ($\pm0.67\%$) 4) $880\sqrt{t} + 3600$ ($\pm0.085\%$)
Switzerland (February 26)	1) $17e^{0.24t}$ ($\pm4.2\%$) 2) $31t^2 - 440t - 2200$ ($\pm5.6\%$) 3) $1000t - 19000$ ($\pm0.47\%$) 4) $7500\sqrt{t} - 27000$ ($\pm2.4\%$)	Japan (February 15)	1) $58e^{0.093t}$ ($\pm4.6\%$) 2) $0.38t^3 - 38t^2 + 1400t - 16000$ ($\pm1.9\%$) 3) $310t - 9200$ ($\pm1.4\%$) 4) $2100\sqrt{t} - 4100$ ($\pm0.23\%$)
Singapore (February 15)	1) $21e^{0.08t}$($\pm5.6\%$) 2) $0.4t^3 - 32t^2 - 440t + 50000$ ($\pm1.4\%$) 3) $680t - 36000$ ($\pm0.71\%$)	Belgium (February 29)	1) $0.63e^{0.74t}$ ($\pm18\%$) 2) $36e^{0.19t}$ ($\pm7.6\%$) 3) $1400t - 31000$ ($\pm0.8\%$) 4) $9500\sqrt{t} - 26000$ ($\pm0.74\%$)
Australia (February 19)	1) $4.7e^{0.15t}$ ($\pm6.1\%$) 2) $360t - 10000$ ($\pm1.3\%$) 3) $980\sqrt{t} - 960$ ($\pm1.2\%$)	Sri Lanka (March 10)	1) $0.51e^{0.60t}$ ($\pm4.0\%$) 2) $6.3t - 1.0$ ($\pm3.1\%$) 3) $18t - 260$ ($\pm1.1\%$)
Hong Kong (February 15)	1) $49e^{0.037t}$ ($\pm1.5\%$) 2) $45t - 10000$ ($\pm3.0\%$) 3) $100\sqrt{t} + 200$ ($\pm0.55\%$)		

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

Table 2 The best-fit curves along with their relative errors for the COVID-19 data for the world. The best-fit curves are shown in Fig. 2. Part 1 of the curve (Fig. 2(a)) is from January 22 to March 01, while Part 2 (Fig. 2(b)) is from March 02 to May 18. For part 2, $\bar{t} = t - 40$ and $\bar{I}(\bar{t}) = I(t) - I(t_0) = I(t) - 88000$, that is, time is measured from March 02.

Part (Start Date)	Best-fit functions and errors
Part 1 (January 22)	1) $390e^{0.40t}$ ($\pm 2.3\%$) 2) $90t^2 + 750t - 4100$ ($\pm 2.7\%$) 3) $1800t + 25000$ ($\pm 0.89\%$) 4) $13000\sqrt{t} + 6500$ ($\pm 0.29\%$)
Part 2 (March 02)	1) $4000e^{0.23\bar{t}}$ ($\pm 6.9\%$) 2) $1700\bar{t}^2 - 29000\bar{t} + 95000$ ($\pm 0.83\%$) 3) $83000\bar{t} - 18 \times 10^5$ ($\pm 0.56\%$)

(blue curves) in *semi-logy* format for all the 21 countries.

We find that a single function does not describe the $I(t)$ curves; hence, we compute best-fit curves for different parts of $I(t)$ by employing exponential and polynomial functions. We used Python's *polyfit* function to compute the best-fit curves. These curves are listed in Table 1 along with the relative errors between the original data and the fitted data. However, we exhibit only the leading power laws of the polynomials in the plots of the figures.

Initially, all the countries exhibit exponential growth ($I(t) = A\exp(\beta t)$), which is expected. It is worth mentioning that the $I(t)$ plots for USA, UK, France, Spain, Germany, Russia, Belgium, and Brazil have two exponential functions for the fit. For example, the $I(t)$ curve of UK is described by two exponential functions, $\sim \exp(0.26t)$ and $\sim \exp(0.23t)$. The quantity β is proportional to the growth rate. The value of β varies for different countries as it depends on factors such as population density, immunity level of the population, climate, local policy decisions (social distancing, lockdowns, testing capacity), etc.

In the exponential regime, the daily infection count is directly proportional to the cumulative count of cases, that is, $\dot{I} \approx \beta I$. The cumulative case count doubles in time $T = (\log 2)/\beta$ in this regime, For Italy, $\beta = 0.33$, resulting in $T \approx 2$ days, which means Italy's $I(t)$ doubled every two days in the early phase (February 22 to March 01).

Next, the curves transition to the regimes that are best described by polynomials and can be approximated as power laws. In Fig. 1, we report the leading terms of the best-fit polynomials as power laws (also see Table 1). The $I(t)$ curves for South Korea, China, Spain, Germany, Israel, Netherlands, and Switzerland exhibit three power law regimes—t^2, t, and $\sqrt{t}$—before flattening. Similarly, $I(t)$ for Australia, Belgium and Hong Kong saturate after t and $\sqrt{t}$ regimes. As predicted in our earlier work [36], countries such as USA, France, Italy, Spain, and Germany transition to a linear and then to $\sqrt{t}$ regime after going through regimes of t^4, or t^3, or t^2. These nations are close to flattening their $I(t)$ curves. UK, Turkey, Israel and Netherlands exhibit similar transitions. However, $I(t)$ for countries like Hong Kong, Sri Lanka, Australia, and Belgium directly transitioned to the linear regime from an exponential phase. We make a cautionary remark that the coefficients of the polynomials depend quite critically on the choice of endpoints of the fit. Our observations of power-law growth are consistent with earlier results [39, 16, 22, 5, 25, 18, 35, 2, 36].

In Fig. 1, in the exponential regime, the $\dot{I}(t)$ curves (daily counts) are nearly parallel to the $I(t)$ curves. It means that $\dot{I}$ increases exponentially in the beginning, similar to $I(t)$. Subsequently, the curves transition to power-law regimes. As discussed by Verma et al. [36], the power law can be approximated as $I(t) \sim At^n$ and $\dot{I}(t) \sim I^{1-1/n}$, which is slower than $\dot{I}$ for the exponential regime. We also remark that for large n, $\dot{I}(t) \sim I$, similar to exponential function.

The linear growth regime has an interesting property. In this regime, $\dot{I} \approx$ const., that is, constant daily infection count. The daily infection count starts to decrease after the linear regime; hence linear regime is the transition point.

We also analyze the data of cumulative infected individuals in the entire world. In Fig. 2, we plot $I(t)$ and $\dot{I}(t)$ versus time in *semi-logy* format. Note that the initial epicenter of the COVID-19 outbreak was in China, and then it shifted to Europe and then to USA. Therefore, we divide the plot in two parts. In the first part [Fig. 2(a)], we illustrate cases that belong mostly to China. After approximately thirty days of outbreak (around February 20), $I(t)$ for China starts to saturate. In Fig. 2(b), we exhibit the $\bar{I}(\bar{t})$ curve after $t = 41$ (March 02) when China had achieved flattening of the curve. In Fig. 2(b), $\bar{t} = t - 40$ and $\bar{I}(\bar{t}) = I(t) - I(t_0) = I(t) - 88000$ due to the coordinate shifts. Both the plots exhibit exponential and power law regimes, but $\bar{I}(\bar{t})$ is yet to flatten (see Table 2). We hope that there is no third part to this curve, which is possible if the unaffected countries remain so.

The transition from the exponential to power-law behaviour is expected from the nature of the $I(t)$ curve. The $I(t)$ curve is convex during the exponential growth phase, that is, its center of curvature is upward. However, the curve must turn concave for it to flatten. This transformation occurs via a sequence of growth phases:

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

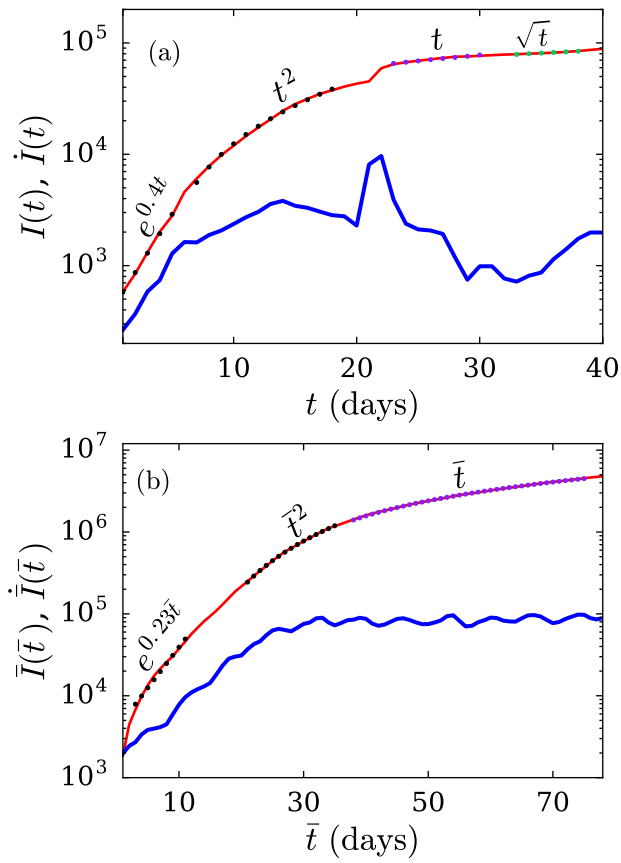

Fig. 2 (color online) For the COVID-19 epidemic, the *semi-logy* plots of total infected individuals ($I(t)$) vs. time (t) (red curves) for (a) Part 1: from January 22 to March 01 (comprising mostly of China), (b) Part 2: from March 02 to May 18 (world other than China). In Part 2, $\bar{t} = t - 40$ and $\bar{I}(\bar{t}) = I(t) - I(t_0) = I(t) - 88000$. The blue curves in the plot describe the derivatives of $I(t)$. The dotted curves represent the best-fit curves.

power-law, linear, square-root, and then flat. The curve transitions from convex to concave in the linear regime for which the radius of curvature is infinite. In Sec. 6 we argue that the power-law behaviour is possibly due to lockdown and social distancing.

We remark that the death count due to COVID-19 also exhibits similar behaviour as the infection count $I(t)$. It is expected because a fraction of infected individuals, unfortunately, die. However, we expect a small time delay between the death time series and the infection time series. Some researchers have attempted to fit the $I(t)$ and death counts with error functions [25,33].

The above analysis shows that we can track the development of the epidemic locally in time. The best-fit curves in particular segments provide the status of the epidemic. For example, if we have reached the linear regime, then we are not far from flattening the curve. Similarly, a $\sqrt{t}$ regime indicates that the flattening of the curve has begun. Thus, simple data analytics described above has significant predictive power.

In the next section we compare the functional behaviour of COVID-19's $I(t)$ curve with those of other major epidemics.

Table 3 For the cumulative infected cases data for MERS, Ebola, and COVID-19, the best-fit functions and the respective relative errors for various stages of evolution shown in Fig. 3.

Epidemic	Best fit functions and errors
MERS $t_{max} = 720$	1) $60e^{0.005t}$ ($\pm 3.2\%$) 2) $0.03t^2 - 24t + 5100$ ($\pm 3.4\%$) 3) $1.5t + 38$ ($\pm 1.2\%$)
Ebola $t_{max} = 360$	1) $14e^{0.03t}$ ($\pm 7.5\%$) 2) $0.7t^3 - 570t^2 + 14\times 10^4 t - 12\times 10^6$ ($\pm 1.4\%$) 3) $135t - 24000$ ($\pm 0.71\%$) 4) $2500\sqrt{t} - 24000$ ($\pm 0.48\%$)
Covid-19 $t_{max} = 118$	See Table 2

3 Comparisonof COVID-19 with other epidemics

A natural question is whether the epidemic evolution of COVID-19 differs from the spread of Ebola and MERS (Middle Eastern Respiratory Syndrome). In this section we perform a comparative study of Ebola, MERS, and Covid-19 epidemics. We digitized data for these epidemics for their respective time periods: MERS [29] from May 01, 2013 to April 30, 2015; Ebola [21] from May 01, 2014 to April 30, 2015; and Covid-19 [37] from January 22, 2020 to May 18, 2020. In Fig. 3, we plot $I(t)$ vs. normalized time, t/t_{max}, in a *semi-logy* format for all three epidemics. Here, t_{max} (see Table 3) is the time span of the epidemic, except for COVID-19 for which t_{max} is taken up to May 18.

In Fig. 3 we present the best-fit curves as dotted lines. Clearly, the three curves look similar, with regimes exhibiting exponential, power-law, and linear growth before flattening. A major difference is that COVID-19 has two subparts, which is essentially due to the spread of COVID-19 by asymptomatic carriers. As shown in Table 2 and Fig. 2, the epidemic first spread in China and then in rest of the world. In contrast, the other two epidemics, Ebola and MERS, were somewhat confined.

In the next section, we present a model for COVID-19 whose predictions match several countries.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

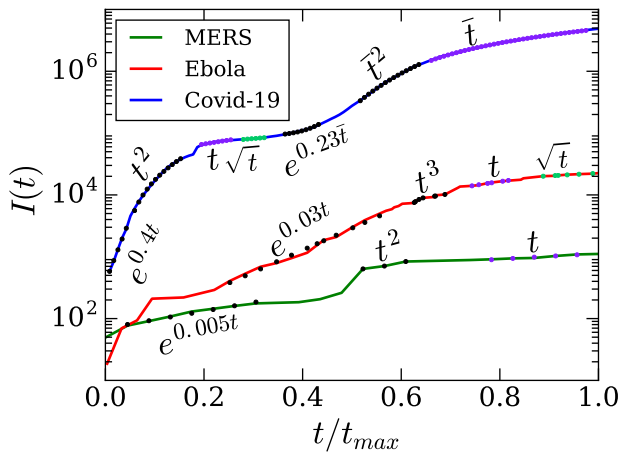

Fig. 3 (color online) For the COVID-19, Ebola and, MERS epidemics, the *semi-logy* plots of total infected individuals ($I(t)$) vs. normalized time (t/t_{max}). The dotted curves represent the best-fit curves using the exponential and polynomial functions. For COVID-19 curve, refer to Fig. 2.

4 SEIR model for COVID-19 epidemic

In Sec. 2, we analyzed the COVID-19 infection data and observed a power-law growth (followed by a linear regime near saturation) after an exponential growth. In this section, we attempt to get some insights about this transition using SEIR model [4,10,20,17,28]. Note the other important epidemic models are regression models [30,12], ARIMA forcasting model [1,3,11], SIR model [15,32], etc. All these models have been frequently and successfuly used to analyse the transmission dynamics of COVID-19. For example, Labadin and Hong [17] used this model to predict the second confirmed case in Malaysia.

Recently, Peng et al. [27] constructed a generalised SEIR model for the spread of SARS-Cov-2 virus in China. López and Rodo [20] modified Peng et al. [27]'s model to analyze the data of Spain and Italy up to the end of March. In this section, we will discuss a simplified version of López and Rodo [20]'s SEIR model and fit it with the real-time data of USA, Italy, Spain and Japan till 18th May 2020.

In the model, we assume the disease transmission to take place only among humans. Further, the natural birth and death rates are assumed to be negligible. We divide the total population (N) at a certain place at time t into seven categories: Susceptible ($S(t)$), Exposed ($E(t)$), Infected ($I(t)$), Recovered ($R(t)$), Insusceptible ($P(t)$), Quarantined ($Q(t)$) and Dead ($D(t)$). Here, $Q(t)$ is the number of confirmed infected cases at time t. The evolution equations of the seven categories are:

$$\frac{dS(t)}{dt} = -\beta\frac{S(t)I(t)}{N} - \alpha S(t), \quad (1)$$

$$\frac{dP(t)}{dt} = \alpha S(t), \quad (2)$$

$$\frac{dE(t)}{dt} = \beta\frac{S(t)I(t)}{N} - \gamma E(t), \quad (3)$$

$$\frac{dI(t)}{dt} = \gamma E(t) - \delta I(t), \quad (4)$$

$$\frac{dQ(t)}{dt} = \delta I(t) - \lambda(t)Q(t) - \kappa(t)Q(t), \quad (5)$$

$$\frac{dR(t)}{dt} = \lambda(t)Q(t), \quad (6)$$

$$\frac{dD(t)}{dt} = \kappa(t)Q(t), \quad (7)$$

where $\beta, \alpha, \delta, \lambda(t)$, and $\kappa(t)$ are the infection, protection, average quarantine, recovery and mortality rates respectively; and γ^{-1} is the average latency period for COVID-19. The protection rate α is governed by the intensity of contact tracing, lockdown policies, and improvement of health facilities. The time-dependent parameters $\lambda(t)$ and $\kappa(t)$ are modeled as follows [20]:

$$\lambda(t) = \lambda_0(1 - exp(-\lambda_1 t)), \quad (8)$$

$$\kappa(t) = \kappa_0 exp(-\kappa_1 t), \quad (9)$$

where λ_0, κ_0, λ_1, and κ_1 are constants. The functional forms in Eqs. (8-9) are chosen in such a way that the recovery rate saturates and the death rate vanishes with time. Note that the cumulative number of reported infected cases (denoted by $I(t)$ in Sec. 2) is the sum of $Q(t)$, $R(t)$, and $D(t)$. For further details, refer to Peng et al. [27], and Lopez and Rodo [20].

We compare the model predictions [Eqs. (1-9)] with the available data [37] for USA, Italy, Spain and Japan. For Spain and Japan, $t = 0$ is taken to be the starting date shown in Table 1. For USA and Italy, $t = 0$ corresponds to 29th and 24th February respectively. The end date for all the four countries is 18th May. The initial values of Q, R, and D are taken to be the total active cases, recovered cases and deaths respectively at $t = 0$ for each country. The number of initial insusceptible cases ($P(t = 0)$) is assumed to be zero. We adjust the parameters $\{\alpha, \beta, \gamma, \delta, \{\lambda_0, \lambda_1\}, \{\kappa_0, \kappa_1\}\}$, $E(t = 0)$ and $I(t = 0)$ such that the relative error between the model and actual data is minimized. Note that the initial condition satisfies the relation

$$S(t = 0) = N - E(t = 0) - I(t = 0) - Q(t = 0) - R(t = 0) - D(t = 0) - P(t = 0), \quad (10)$$

where N is the total population of the country.

In Fig. 4, we present the best-fit curves from the SEIR model along with the actual real-time data. This model fits well with the data for Italy and Spain. In

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

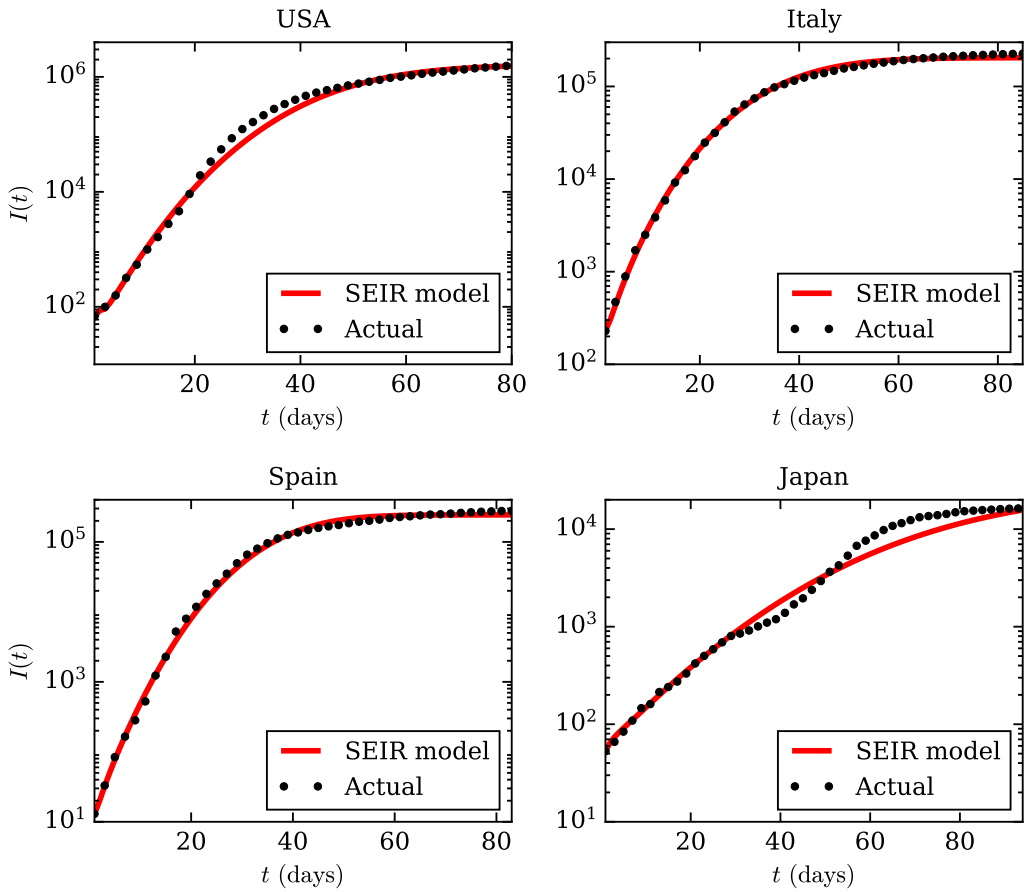

Fig. 4 (color online) For the COVID-19 pandemic, *semi-logy* plots of total infected individuals ($I(t)$) vs. time (t) (black dotted curves) for USA, Italy, Spain and Japan from the available data at *worldOmeter* [37] till May 18. The red curves are the best-fit curves using the SEIR model described in Sec. 4.

Table 4 Best-fit numerical values of $\{\alpha, \beta, \gamma, \delta, \{\lambda_0, \lambda_1\}, \{\kappa_0, \kappa_1\}\}$ of SEIR model for four countries with relative errors between the best-fit curves and actual data [37]. The best-fit curves are shown in Fig. 4.

	Best-fit parameters of $SEIR$ model								
Country	α	β	γ	δ	λ_0	λ_1	κ_0	κ_1	Errors
USA	0.027	8.0	0.12	1.6	0.14	0.099	1.9	2.0	15.7%
Italy	0.037	1.3	0.24	0.41	0.85	0.054	0.31	0.045	5.0%
Spain	0.017	1.8	1.1	0.99	1.3	0.059	0.38	0.16	10.3%
Japan	0.02	1.6	0.05	0.28	0.07	0.3	0.34	0.089	17.9%

Table 4 we list the numerical values of the best-fit parameters and the relative errors between the predictions and data. Note that for Spain and Italy, López and Rodo [20] considered natural birth and death rates in their model and obtained fits for Q, R and D seperately. In contrast, we stick to the fundamental assumption regarding natural birth and death rates of the basic SEIR model [17] and obtain the fits for $Q(t)+R(t)+D(t)$ till May 18.

Our best-fit values of parameters for Spain and Italy are nearly consistent with those of López and Rodo [20]. The model shows that high infection rates (β) and small average latency periods γ^{-1} try to push the cumulative number of infected cases (reported) to a large

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

saturation value via an exponential growth. On the other hand, high protection and quarantine rates, α and δ, slow down the growth and minimize the saturation level of the cumulative infected (reported) cases. Thus, the values of the control parameter set $\{\beta, \gamma, \alpha, \delta\}$ in Table 4 determine the nature of the power-law after the exponential growth. On the other hand, the linear regime (for Italy, Spain) near the saturation is determined well by the removal rate set $\{\lambda_0, \lambda_1, \kappa_0, \kappa_1\}$. Thus, the present model is consistent with the results presented in Section. 2.

In the next section, we will present another model which is based on delayed differential equations.

5 Model based on delayed-differential equations

In this section we consider a class of models based on delay differential equations (DDE), which are different from SEIR model. Here, the equations often look simpler than their SEIR counterparts since delay can be used to account for multiple features without increasing the number of variables. The flip side, however, is that delays can be analytically intractable.

In this section, we focus on one particular delayed model [34], which uses delays to account for the pre-symptomatic period and the infection period. This model has been used to track the evolution of the epidemic, especially in the post-linear regime. It describes a potential new route to the end of the pandemic through a combination of social distancing, sanitization, contact tracing and preventive testing. In the controlled endgame phase of the epidemic, which we call self-burnout, we have a slightly different equation. In this phase, there is extensive enforcement of separation minima (a term we prefer to social distancing as it does not carry connotations of emotional isolation) so the rate of new cases does not depend on the number of healthy and susceptible people at large (i.e. not in quarantine). Rather, we assume that each sick person spreads the disease at a constant rate m_0. Under these conditions, the dynamic model for the spread of cases (y, which is same as $I(t)$) is

$$\frac{\mathrm{d}y}{\mathrm{d}t} = m_0[y(t) - (1-\mu_3)\, y(t-\tau_2/2) - (1-\mu_1)\,\mu_3 y(t-\tau_2) - \mu_1\mu_3 y(t-\tau_1)], \qquad (11)$$

where $\mu_1, \mu_2, \mu_3, \tau_1, \tau_2$ are parameter. In our model, the contact tracing manages to capture a fraction $1-\mu_3$ of all the sick patients and places them into quarantine.

A solution to the above equation is $y = const$. It has been shown in [34] that this solution is stable if and only if

$$m_0\left(\frac{1+\mu_3-2\mu_1\mu_3}{2}\tau_2 + \mu_1\mu_3\tau_1\right) < 1 \qquad (12)$$

This identifies a maximum value of m_0 for which the solution is stable, i.e., the epidemic gets over in time. For the plausible parameter values $\tau_1 = 7$, $\tau_2 = 3$, $\mu_1 = 1/5$ and $\mu_3 = 1/2$, the critical value of m_0 turns out to be 20/53. Here we assume that the test results are instantaneous due to high testing capacity present in the region.

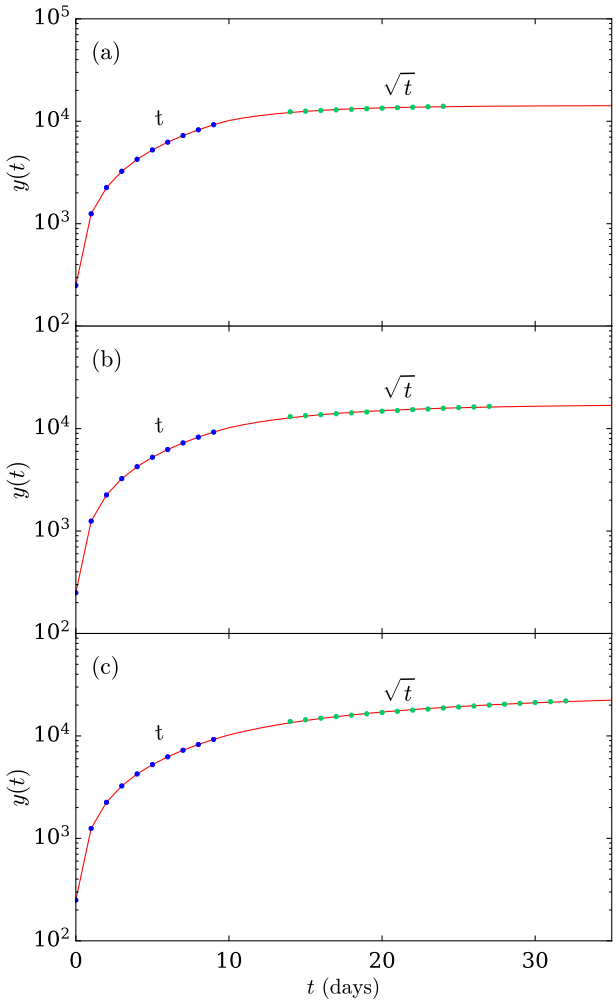

Fig. 5 (color online) Simulation of the DDE-based model (Eq. (11)) with m_0 being (a) 70%, (b) 80%, (c) 90% of the critical value, as explained in the text.

We perform simulation runs of Eq. (11) with the above parameter values and m_0 having the values 70, 80 and 90 percent of the critical limit. We seed the equation with the linear function $y = 1000t$ for the first ten days. We find a considerable region thereafter where the case histories show a $\sqrt{t}$ profile before saturating.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

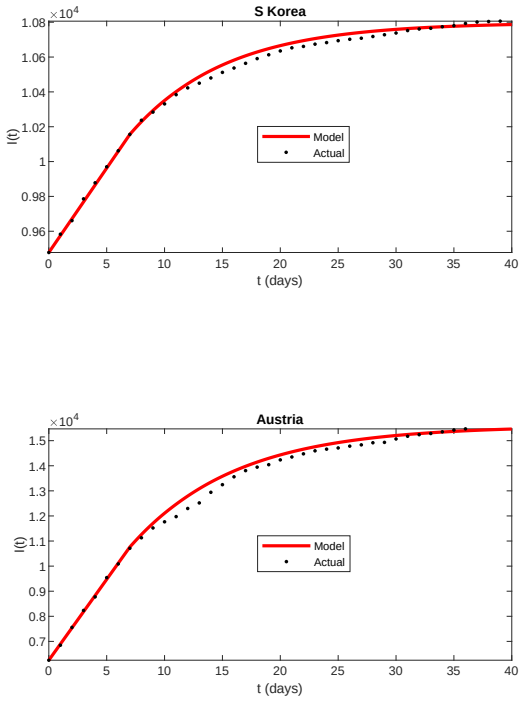

Fig. 6 (color online) For the COVID-19 pandemic, *semi-logy* plots of total infected individuals ($I(t)$) vs. time (t) (black dotted curves) for South Korea and Austria from the available data at *worldOmeter* [37]. The red curves are the best-fit curves obtained using the DDE-based model including self-burnout (see Sec. 5).

The errors between the observed data and the best-fit curves are less than 1 percent in each case. This explains why the countries which are achieving saturation are showing a pronounced $\sqrt{t}$ phase after the linear phase.

To further bolster the validity of our model, we consider two countries (South Korea and Austria) which have shown a very good linear region followed by saturation. South Korea showed linear regime from March 28 to April 04 (with 9478 and 10156 cases respectively), after which it enters the burnout phase. Using this as the seeding data and taking the parameter values mentioned above, we find the best fit for the next 20 days for m_0 equal to 77 percent of the critical. The error between the best-curve and the data is 0.34 percent. Note that Shayak and Rand [34] has found an m_0 of 75 and not 77 percent of critical, because the fit was upto a smaller duration.

Austria showed linear regime from March 28 to April 01 (with 6250 and 10711 cases respectively) before entering self-burnout phase. We find the best fit for m_0 to be 79 percent of the critical. The error is 2.6 percent. However, the actual data for the 8th to the 15th day appears to be too low—the curve has a convex profile which is probably unrealistic. If we consider the error from the 16th to the 30th day then we find a value of 1.1 percent only. We present these best-fit curves in Fig. 6. Other regions which are in the self-burnout phase are Vietnam, Australia, New Zealand, and Goa, Kerala and Odisha in India. We have chosen South Korea and Austria since their data shows the smoothest profile on account of high testing capacity.

Both SEIR and the DDE models describe the evolution of COVID-19 epidemic quite well for many countries. A detailed comparison between the two models will be performed in future. Also, we plan to employ the two models to understand the epidemic evolution for many nations.

6 Discussions and Conclusions

COVID-19 pandemic involves many factors, for example, asymptomatic carriers, lockdown, social distancing, quarantine, etc. Considering these complex issues, we focus on data analysis. In particular, we analyze the real-time infection data of COVID-19 epidemic for 21 nations up to May 18, 2020. Our analysis shows that many nations are close to flattening the epidemic curve.

A key feature of our analysis is the emergence of power-law behavior after an exponential growth, which has also been observed by other researchers [39, 16, 22, 5, 25, 18, 35, 2]. The exponential growth is easily explained using $\dot{I} \propto \beta I$ relation, which arises due to the spread by contact. For power-law growth, $I(t) \sim t^n$, the above relation is modified to $\dot{I} \sim I^{1-1/n}$. The suppression of $I^{-1/n}$ in $\dot{I}$ could be attributed to lockdowns and social distancing etc. A careful analysis of the epidemic models should yield this feature. Interestingly, Ebola and MERS also exhibit similar behavior. This generic feature is very useful for the forecast of the epidemic evolution.

Note that the $I(t)$ curve needs to turn from convex (during the exponential growth) to concave for flattening. Hence, a transition from exponential growth to a power-law growth is expected. The lockdowns and social distancing are likely to make the transition earlier, thus suppressing the exponential growth to some degree. Earlier, Verma et al. [36] had conjectured that the power-law growth might occur due to asymptomatic carriers and/or community spread. This conjecture needs a closer examination.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

In this paper, we only studied the infection counts. However, it is evident that during the growth phase, the active cases and death counts, would follow similar pattern as $I(t)$. The total death count too flattens along with the infection count, but the active cases decreases with time during the saturation.

Prakash et al. [31] studied the phase space portraits, that is, $\dot{I}$ vs. I plots. They observed the phase-space curves to be linear. This is natural for the exponential growth ($\dot{I} \propto \beta I$), as well as for the power-law growth with large exponent n because $\dot{I} \sim I^{1-1/n}$. In another interesting analysis of COVID-19 epidemic, Schüttler et al. [33] and Marsland and Mehta [25] argued that $I(t)$ or total death count could be modelled using error function. Using this result, we may be able to predict the asymptotic behavior of $I(t)$ that may yield valuable clues regarding the extent and duration of the epidemic.

Epidemic spread has similarities with rumor spread and the growth of a network [10, 26]. A comparison of the power-law growth in these systems will yield fruitful results for the epidemic forecast.

In summary, COVID-19 epidemic data reveal interesting properties that can be used for its forecast. The $\sqrt{t}$ dependence of total infections after the linear regime is a striking feature that might have connections with other aspects of physics and mathematics. We leave these considerations for future studies.

Acknowledgements The authors thank Santosh Ansumali for useful discussions. We also thank the *worldOmeter* for the data that made this work possible. Ali Asad is supported by Indo-French (CEFIPRA) project 6104-1, and Soumyadeep Chatterjee is supported by INSPIRE fellowship (IF180094) of Department of Science & Technology, India.

7 Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Anne, R.: Arima modelling of predicting covid-19 infections. medrxiv.org (doi:10.1101/2020.04.18.20070631) (2020)
2. Beare, B.K., Toda, A.A.: On the Emergence of a Power Law in the Distribution of COVID-19 Cases. arXiv p. arXiv:2004.12772 (2020)
3. Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., Ciccozzi, M.: Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in Brief **29**, 105340 (2020)
4. Bjørnstad, O.N.: Epidemics: Models and Data using R. Springer (2018)
5. Blasius, B.: Power-law distribution in the number of confirmed COVID-19 cases. arXiv p. arXiv:2004.00940 (2020)
6. Cherednik, I., Hill, U.C.: Momentum managing epidemic spread and Bessel functions. arXiv.org p. arXiv:2004.06021 (2020)
7. Cheynet, E.: ECheynet/SEIR: Generalized SEIR Epidemic Model (fitting and computation) (2020). DOI 10.5281/zenodo.3819519. URL `https://doi.org/10.5281/zenodo.3819519`
8. Cheynet, E: (2020). URL `https://github.com/ECheynet/SEIR`
9. Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore y Piontti, A., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M.E., Longini Jr., I.M., Vespignani, A.: The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science **(preprint)** (2020)
10. Daley, D.J., Gani, J.: Epidemic Modelling: An Introduction. Cambridge University Press (2001)
11. Deb, S., Majumdar, M.: A time series method to analyze incidence pattern and estimate reproduction number of COVID-19. arXiv.org p. arXiv:2003.10655 (2020)
12. Gupta, R., Pandey, G., Chaudhary, P., Pal, S.K.: Seir and regression model based covid-19 outbreak predictions in india. medrxiv.org (doi:10.1101/2020.04.01.20049825) (2020)
13. Hellewell, J., Abbott, S., Gimma, A., Bosse, N.I., Jarvis, C.I., Russell, T.W., Munday, J.D., Kucharski, A.J., Edmunds, J.: Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. The Lancet Global Health **8**(4), DOI: 10.1016/S2214–109X(20)30074–7 (2020)
14. Johns Hopkins University, Corona Resource Center: (2020). URL `https://coronavirus.jhu.edu/map.html`
15. Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society A **115**(772), 700–721 (1927)
16. Komarova, N.L., Wodarz, D.: Patterns of the COVID19 epidemic spread around the world: exponential vs power laws. medrxiv.org (DOI: 10.1101/2020.03.30.20047274) (2020)
17. Labadin, J., Hong, B.H.: Transmission Dynamics of 2019-nCoV in Malaysia. medrxiv.org (doi: 0.1101/2020.02.07.20021188) (2020)
18. Li, M., Chen, J., Deng, Y.: Scaling features in the spreading of COVID-19. arXiv p. arXiv:2002.09199 (2020)
19. Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., Shaman, J.: Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science **6**, DOI: 10.1126/science.abb3221 (2020)
20. López, L.R., Rodo, X.: A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: simulating control scenarios and multi-scale epidemics. medarxiv.org (doi: 10.1101/2020.03.27.20045005) (2020)
21. Ma, J.: Estimating epidemic exponential growth rate and basic reproduction number. Infectious Disease Modelling **5**, 129–141 (2020)
22. Manchein, C., Brugnago, E.L., da Silva, R.M., Mendes, C.F.O., Beims, M.W.: Strong correlations between power-law growth of COVID-19 in four continents and the inefficiency of soft quarantine strategies. Chaos: An Interdisciplinary Journal of Nonlinear Science **30**(4), 041102 (2020)
23. Mandal, S., Bhatnagar, T., Arinaminpathy, N., Agarwal, A., Chowdhury, A., Murhekar, M., Gangakhedkar, R.,

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091389; this version posted June 1, 2020. 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 .

Sarkar, S.: Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach. Indian Journal of Medical Research **(preprint)** (2020)
24. Marathe, M., Vullikanti, A.K.S.: Computational epidemiology. Commun. ACM **56**(7), 88–96 (2013)
25. Marsland, R., Mehta, P.: Data-driven modeling reveals a universal dynamic underlying the COVID-19 pandemic under social distancing. arXiv p. arXiv:2004.10666 (2020)
26. Newman, M.: The physics of networks. Phys. Today **61**(11), 33–38 (2008)
27. Peng, L., Yang, W., Zhang, D., Zhuge, C., Hong, L.: Epidemic analysis of COVID-19 in China by dynamical modeling. medarxiv.org (doi: 10.1101/2020.03.14.20036202) (2020)
28. Peng, L., Yang, W., Zhang, D., Zhuge, C., Hong, L.: Epidemic analysis of COVID-19 in China by dynamical modeling. arXiv.org (2020)
29. Phoenix7777 on Wikipedia: URL `https://upload.wikimedia.org/wikipedia/commons/7/7d/World_MERS_outbreak.svg`
30. Pirouz, B., Shaffiee Haghshenas, S., Shaffiee Haghshenas, S., Piro, P.: Investigating a Serious Challenge in the Sustainable Development Process: Analysis of Confirmed cases of COVID-19 (New Type of Coronavirus) Through a Binary Classification Using Artificial Intelligence and Regression Analysis. Sustainability **12**, 2427 (2020)
31. Prakash, M.K., Kaushal, S., Bhattacharya, S., Chandran, A., Kumar, A., Ansumali, S.: A minimal and adaptive prediction strategy for critical resource planning in a pandemic. medrxiv.org (DOI:10.1101/2020.04.08.20057414) (2020)
32. Ranjan, R.: Estimating the final epidemic size for covid-19 outbreak using improved epidemiological models. medrxiv.org (doi:10.1101/2020.04.12.20061002) (2020)
33. Schüttler, J., Schlickeiser, R., Schlickeiser, F., Kröger, M.: Covid-19 Predictions Using a Gauss Model, Based on Data from April 2. preprints.org (2020)
34. Shayak, B., Rand, R.H.: Self-burnout - a new path to the end of covid-19. medrxiv.org (doi:10.1101/2020.04.17.20069443) (2020)
35. Singer, H.M.: The COVID-19 pandemic: growth patterns, power law scaling, and saturation. arXiv p. arXiv:2004.03859 (2020)
36. Verma, M.K., Asad, A., Chatterjee, S.: COVID-19 epidemic: Power law spread and flattening of the curve. Transactions of the Indian National Academy of Engineering **Manuscript accepted for publication.**
37. WorldOMeter: URL `https://www.worldometers.info/coronavirus/`
38. Wu, K., Darcet, D., Wang, Q., Sornette, D.: Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. arXiv.org (2020)
39. Ziff, A.L., Ziff, R.M.: Fractal kinetics of COVID-19 pandemic. medrxiv.org (DOI: 10.1101/2020.02.16.20023820) (2020)