BACK TO THE ROOTS: A DISCRETE KERMACK-MCKENDRICK MODEL ADAPTED TO COVID-19

MATTHIAS KRECK[†], ERHARD SCHOLZ [‡]

ABSTRACT. A widely used tool for analysing the Covid-19 pandemic is the standard SIR model. It seems often to be used as a black box, not taking into account that this model was derived as a special case of the seminal Kermack-McKendrick theory from 1927. This is our starting point. We explain the setup of the Kermack-McKendrick theory (passing to a discrete approach) and use medical information for specializing to a model which we call *adapted K-McK-model*. This includes effects of vaccination, mass testing and mutants. We demonstrate the use of the model by applying it to the development in Germany. As a striking application we demonstrate that a comparatively mild intervention reducing the time until quarantine by one day leads to a drastic improvement. A similar effect can be obtained by certain mass testings as we will demonstrate. We discuss possibilities to apply the model both for predictions and as an analysis tool. We compare the adapted K-McK-model to the standard SIR model and observe condiderable differences if the contact rates are not constant. Finally we compare the model reproduction rate with the empirical reproduction rate determined by the Robert Koch-Institut.

1. INTRODUCTION

> *As early as 1927, Kermack and McKendrick published a paper "A contribution to the mathematical theory of epidemics" in the Proceedings of the Royal Society London Ser. A. The paper became a classic in infectious disease epidemiology and has been cited innumerable times. But how often is it actually read? Judging from an incessant misconception of its contents, one is inclined to conclude: hardly ever!* [2, p. 105]

The authors of the cited paper continue with the observation that even experienced experts often believe that the paper is just about the SIR or SEIR equations. Taking this criticism seriously, the present paper proposes a Kermack-McKendrick type model adapted to Covid-19, which differs from the widely used S(E)IR models and which we believe has a higher plausibility and applicability.

Date: April 26, 2021.

[†]Mathematisches Institut der Universität Bonn and Mathematisches Institut der Universität Frankfurt, Germany, kreck@math.uni-bonn.de

[‡]University of Wuppertal, Faculty of Math./Natural Sciences, and Interdisciplinary Centre for History and Philosophy of Science, scholz@math.uni-wuppertal.de

We do this with a slight modification of the original idea. We work with a discrete time parameter. After all, people are not participating in the spread of the disease during all parts of a day, e.g.. when they are asleep. And also the data are communicated on a daily basis. In addition discrete models are easy to program. So we adapt the Kermack-McKendrick approach and replace functions depending on a continuous time t by a functions depending on integers k.

We will describe the model in the next section and then explain the role of its parameters (sections 1 and 2). Then we describe how one applies the model starting from the empirical data and demonstrate the result for Germany (sections 3 and 4). We give a striking application showing what happens if people are sent earlier to quarantine or hospital (section 5), before we discuss how vaccination, the rise of new mutants and mass testing can be implemented in the model by a refinement and extension of the recursion presented in the first section (section 6). Finally we add some remarks on the relationship between the reproduction numbers as they appear in our model to the ones given by the Robert-Koch-Institut for Germany (section 9). At the end we discuss some aspects of our model and formulate a few clear messages.

2. The model

We recall the often neglected fundamental idea fundamental idea of Kermack and McKendrick following [8, 1, 2, 4]. As mentioned before we use the discrete setting of the Kermack-McKendrick approach, which the original paper started with before the continuous version was introduced as a limiting case. A central input of a Kermack/McKendrick model is a function $\gamma(j)$ which measures the medical strength of infectivity of a person at day j counted from the moment on of infectivity. On this basis the Kermack-McKendrick approach allows to derive equations which express how many people are susceptible to the virus if one knows the mean contact rate of the interaction of the population. To emphasize that γ is the central input we call such a model a *discrete γ-K-McK model*.

If you want to construct a γ-K-McK-model the first thing to do is to look for virological studies which determine $\gamma(j)$ as closely as possible. This will depend on the type of an epidemic. The closer the input function $\gamma(j)$ is to the medical facts, the more reliable is the model. As mentioned above, this function is all one needs to model the number of susceptibles. But one is interested in other data like for instance how many people are newly infected at day k. This is not determined by γ alone. For this we look what happens when a person is infected in an epidemic. After infection there are a few days where people are exposed, meaning that they are already infected but not yet infectious. Then the period begins where a person is infectious (we call them propagators) and after that in general may be considered as immune. The period of both lengths is not sharp, varies from person to person

a bit. Our assumption in this paper is that this variation is small, meaning that the functions which describe these lengths are nearly a jump function. With other words we assume that after infection there are e days, where people are exposed, meaning that they are already infected but not yet infectious. After that there are p_d days where a person is infectious and after that in general may be considered as immune. There are indications that this is the case for Covid-19. If there is an epidemic where this is not the case and one knows the functions which describe the rate of the lengths for being exposed or propagating one can easily adapt our model, but we don't do this in this paper. We will discuss a special case of such a function for the propagating people when we compare our model to the standard SIR model. Amongst those who are infectious there are those who are positively tested after $p_c < p_d$ days on average, and then are confined to home isolation or hospital. (We will explain the notation p_d and p_c later, for an example see fig. 1.) Functions γ with only a finite number of days where they are non-zero are called *functions with finite support*.

The γ-K-McK-models are flexible enough to allow for adding important additional features of an epidemic like, e.g., unreported cases, vaccination, effects of testing and of the rise of new mutants etc. Here we take up the first point only and discuss the other points later.

The assumption of a finite support of γ fits with the following picture: At each day the population can, in the model, be divided into disjoint subsets, called compartments, as follows:

- The *compartment S of susceptible people* to the disease with cardinality $S(k)$ at the day k. Here susceptible means not immunised, i.e. either recovered or vaccinated.
- The *compartment E of exposed people* whose cardinality at the day k we denote by $E(k)$. By this we mean people who are infected but are not yet infectious. The function γ determines the number e of days people stay in compartment E as

$$e := \min\{j \,|\gamma(j) \neq 0\} - 1 \; .$$

- The *compartment P of propagators*. These are people who are infectious (they propagate the virus) but not yet in quarantine or isolation. We will divide this into the following subcompartments:
  The *compartement P_c of "counted" propagators* with cardinality $P_c(k)$, these are the infectious who will later be diagnosed, recorded and counted in the statistics. As mentioned above the number of days people stay in this compartment is denoted by p_c.
  And the *compartement P_d of "dark sector" propagators* who are never reported as infected, *the dark sector* with cardinality $P_d(k)$. The number of

days people stay in this compartment is again derived from γ as the cardinality of its support:

$$p_d := \sharp\{j \,|\gamma(j) \neq 0\}$$

- The *compartment* Q of persons who are diagnosed and set under *quarantine* or isolated until recovery or death. The cardinality of this compartment is $Q(k)$ and the mean residence time is denoted by q. The members of Q are essentially *no longer infecting others*, although they are still infectious in the medical sense. The number $Q_{\text{new}}(k)$ of daily new registered persons in Q is a central datum for the epidemic.
- The *compartment* R *of removed* from the epidemic (recovered or dead) with cardinality $R(k)$.

For modelling an epidemic realistically in the γ-K-McK framework which also includes the dark sector and accounts for changing contact rates one needs other input parameters which have to be estimated from empirical data:

- We assume that for the infected persons leaving compartment E at any day k there is a a certain fraction $\alpha(k)$ of people from E who move to compartment P_c at the day k, whereas the fraction $1-\alpha(k)$ of people moves to the compartment P_d of people in the dark sector.
- It could be that people in the dark sector are less infectious by a factor $\xi \leq 1$, so this has to be taken into consideration. A lot of them have no symptoms which might indicate that they have lower viral load.
- Besides the medical function $\gamma(j)$ for a certain epidemic one needs to know a daily proportionality factor g called for simplicity the *contact rate* $\kappa(k)$.

Now we explain the *dynamics of the model* which demonstrates the central role of the function γ and the daily contact rate $\kappa(k)$. The idea is that all people in compartments P_c and P_d infect people from compartment S proportionally to the strength of the infection given by γ and the contact rate $\kappa(k)$. Here we note that both functions γ and κ are dependent on the choice of a unit, but if we replace γ by $c\,\gamma$ we have to compensate this by replacing κ by $\frac{\kappa}{c}$. Another convention influencing κ is to introduce $s(k) := \frac{S(k)}{N}$, where N denotes the size of the population, which in this paper we consider as constant.

To explain the dynamics in detail we denote the number of people who newly entered compartment E at day k by $E_{\text{new}}(k)$. In the situation when no vaccination is available (we shall discuss the variants with vaccination in Section 7) it can be expressed in terms of the function S as:

$$E_{\text{new}}(k) = S(k-1) - S(k)\,. \tag{1}$$

Similarly we denote the number of people who newly entered compartment P_c at day k by $P_{c,new}(k)$ and similarly those who newly entered compartment P_d at day

k by $P_{d,new}(k)$. We first explain the dynamics of those who are in compartment P_c at the day $k-1$. They are the disjoint union of people who newly entered this compartment during the p_c days before. So for each $1 \leq j \leq p_c$ there are $P_{c,new}(k-j)$ people in P_c, who infect people who are in S at the day $k-1$ with strength $\gamma(j)$ and proportional to the contact rate $\kappa(k-1)$.

From κ and γ one can read off the probability that a person who entered compartment P_c at day $k-j$ infects a susceptible at day $k-1$, namely as $\frac{\kappa(k-1)\gamma(j)}{N}$. This implies that the overall probability involving all people who entered compartment P_c at day $k-j$ to infect a susceptible at day $k-1$ is $1-(1-\frac{\kappa(k-1)\gamma(j)}{N})^{P_{c,new}(k-j)}$ [4, p. 4]. Often the linear approximation is used instead, which is $\frac{\kappa(k-1)\gamma(j)}{N} P_{c,new}(k-j)$. In our situation the difference is negligible and so we use the linearisation instead. So the contribution to the newly infected at day $k-1$ by those people who entered compartment P_c at day $k-j$ is (for $1 \leq j \leq p_c$).

$$s(k-1)\kappa(k-1)\gamma(j)P_{c,new}(k-j).$$

Similarly for the compartment P_d (for $1 \leq j \leq p_d$) the contribution is

$$\xi s(k-1)\kappa(k-1)\gamma(j)P_{d,new}(k-j).$$

Summing up we obtain the following equation expressing the dynamics of the epidemic:

$$E_{\text{new}}(k) = s(k-1)\,\kappa(k-1)\left[\sum_{j=1}^{p_c}\gamma(j)P_{c,new}(k-j) + \xi\sum_{j=1}^{p_d}\gamma(j)P_{d,new}(k-j)\right] \tag{2}$$

If we were working in a continuous model, the sum would be an integral from 0 to p_c.

Our aim is to derive from these inputs a single recursion equation for the number of susceptible people at day k. For this we also express $P_{c,new}$ and $P_{d,new}$ in terms of S. Namely after e days the newly exposed people at day k move to compartments P_c and P_d with ratio α respectively $1-\alpha$ leading to the equations

$$P_{c,new}(k+e) = \alpha(k+e)E_{\text{new}}(k) = \alpha(k+e)\big(S(k-1)-S(k)\big)\,, \tag{3}$$

$$P_{d,new}(k+e) = (1-\alpha(k+e))E_{\text{new}}(k) = \big(1-\alpha(k+e)\big)\big(S(k-1)-S(k)\big)\,. \tag{4}$$

To simplify the presentation we introduce two functions, which summarise the two expressions on the right hand side of equation (2.

$$\mathbb{P}_c(k-1) := \sum_{j=1}^{p_c}\gamma(j)\,\alpha(k-j)\,\big[S(k-j-e-1)-S(k-j-e)\big]\,, \tag{5}$$

$$\mathbb{P}_d(k-1) := \sum_{j=1}^{p_d}\gamma(j)\,\big(1-\alpha(k-j)\big)\,\big[S(k-j-e-1)-S(k-j-e)\big]\,. \tag{6}$$

Using this and equations (1) - (6) we have finished the derivation of our model equation.

Proposition. *Given the functions γ, κ, α and the integer p_c. Then if the dynamics of an epidemic is given by equation (2), the single recursive equation for the number of susceptible people is*

$$S(k-1) - S(k) = s(k-1)\,\kappa(k-1)\left[\mathbb{P}_c(k-1) + \xi\,\mathbb{P}_d(k-1)\right]. \tag{7}$$

The functions E, P_c, P_d, and Q are expressed in terms of S as follows:

$$E(k) = S(k-e) - S(k) \tag{8}$$

$$P_c(k) = \sum_{j=0}^{p_c-1} \alpha(k-j)\left(S(k-e-j-1) - S(k-e-j)\right) \tag{9}$$

Similarly

$$P_d(k) = \sum_{j=0}^{p_d-1} \big(1-\alpha(k-j)\big)\left(S(k-e-j-1) - S(k-e-j)\right) \tag{10}$$

and

$$Q(k) = \sum_{j=0}^{q-1} \alpha(k-p_c-j)\left(S(k-e-p_c-j-1) - S(k-e-p_c-j)\right). \tag{11}$$

The number of reported as newly infected people, denoted by $Q_{new}(k)$ is

$$Q_{new}(k) = \alpha(k-p_c)\left(S(k-e-p_c-1) - S(k-e-p_c)\right). \tag{12}$$

Proof. We have given the proof of the recursion equation 7 before the proposition.

Since people stay in compartment E for e days before they move on to the compartment P we obtain:

$$E(k) = E_{\text{new}}(k) + E_{\text{new}}(k-1) + \cdots + E_{\text{new}}(k-e+1)$$
$$= -\big(S(k) - S(k-1)\big) + \big(S(k-1) - S(k-2)\big) + \cdots + \big(S(k-e+1) - S(k-e)\big)$$
$$= S(k-e) - S(k)$$

We apply the same summation to the recognized propagating people,

$$P_c(k) = P_{c,new}(k) + P_{c,new}(k-1) + \cdots + P_{c,new}(k-p_c+1),$$

then one obtains (9), and the proof of (10) is the same. The proof of (11) is also the same, once we have formula (12). For this we note that people newly entering

compartment Q at day k are the new people entering the compartment P_c at day $k - p_c$ the compartment P_c:

$$Q_{\text{new}}(k) = P_{c,new}(k - p_c)$$

and so by formula (3) we have

$$Q_{\text{new}}(k) = \alpha(k - p_c)\big(S(k - e - p_c - 1) - S(k - e - p_c)\big)$$

□

We will give a more general version of our model involving the effects of vaccination, testing a certain percentage of the population each day and of arising of new mutants in section 7 (formula 23).

An important key figure of an epidemic is the daily reproduction number $\rho(k)$. This is defined as the average number of secondary infected people by one typical primary infected (averaged over all infectious at the day k). Since there are no data available from which one can directly read off the reproduction number the best thing one can do is to derive it from a reliable model.

The reproduction numbers in the adapted K-McK-model are:

$$\begin{aligned}(13)\quad \rho(k) \;=\; & \sum_{j=0}^{p_c-1} s(k-1+j)\kappa(k-1+j)\alpha(k-1+j)\,\gamma(j+1) \\ & + \sum_{j=0}^{p_d-1} s(k-1+j)\kappa(k-1+j)(1-\alpha(k-1+j))\,\gamma(j+1)\end{aligned}$$

For approximately constant values of $\alpha(k), s(k), \kappa(k)$ in the interval $[k, k + p_d)$ this boils down to a weighted average of the number of secondary infected by a person entering P_c at day k and the analogous number for those entering P_d at the same day.

Remark: *The model is characterized by a delay structure which appears typically in Covid-19. For example a change in the contact rates κ at a day k will be observable in the numbers of daily new recorded infected Q_{new} only $e + p_c$ later (a sharp eye will notice the time delay in fig. 3).*

Similar models without dark sector and with a simple box-like function γ were constructed and used in [11, 5, 15, 9].

3. The role of the parameters

If one wants to model an epidemic one has to find the appropriate parameters. If for the sake of argument we abstract from the dark sector for a moment, there are two types of parameters in the model: parameters that are of a medical nature like the duration of the exposed people period e and the infectivity function $\gamma(k)$,

and those that can be influenced by the politicians, like the duration p_c between the onset of infectivity and the beginning of quarantine or isolation, or the daily contact rate $\kappa(k)$. Coming back to the dark sector, one has to know in addition the ratio α between the later recorded people and the rest, and the duration p_d of the infectious phase of the dark people. Amongst these parameters we assume that e and the strength function γ are essentially unchanged during the course of the epidemic, as well as p_c, p_d and even α if we abstract from changes in the ratio due to rising unspecific testing. The parameter which definitely changes over time is $\kappa(k)$.
Observation: *If we start with data for $S(k)$ and keep all parameters fixed except $\kappa(k)$, then $\kappa(k)$ is determined by equation (7) as the quotient of the left hand side by the factor of $\kappa(k)$ on the right side. With these $\kappa(k)$ the data are identically reproduced by the model. In turn if we start with arbitrary values of $\kappa(k)$ this gives us values $S(k)$ if we assume constant values for α. So if we allow to read $\kappa(k)$ off from equation (7) we obtain a bijection between values $S(k)$ and values $\kappa(k)$. In other words, if one allows daily changing values of $\kappa(k)$ derived from the empirical data the model is a tautology.*

This holds for any input function γ whatsoever, also for continuous time t, in particular for

$$\gamma_{SIR}(t) = \lambda e^{-\mu t} \tag{14}$$

which is the input for the SIR model [2]. This changes if one assumes constant contact rates, or contact rates which are constant for some period and then are changed by a certain factor to a new constancy level. Then the models cease to be tautological. In the section 6 we will demonstrate this difference by comparing our model with the SIR model.

Thus the choice of $\gamma(k)$ is essential for obtaining a realistic model for Covid-19. In principle it is given by medical data, more precisely by studies in which virologists have measured the viral load of infected in the course of time, e.g. [16]. Simplifying a bit, this leads to the following picture. The length of e is approximately 2 days, so $\gamma(j) = 0$ for $1 \leq j \leq 2$. Within another 2 days it reaches a maximum close to the moment where symptoms show up, before it starts to decrease, slowly at the beginning and then faster until, after about further 8 days, it reaches a value where people are practically no longer infectious. This indicates that p_d is 10. Curves which show this increase and decrease for different infected persons are shown in [16]; they allow to characterize a typical course of the strength of infectivity $\gamma(k)$ for Covid-19. A rough picture of the curve is given in [16] and from this we derive the following table:

k	0	1	2	3	4	5	6	7	8	9	10	11
$\gamma(e+k)$	0	0,5	0,9	0,9	0,85	0,8	0,7	0,6	0,45	0,15	0,05	0

This function is shown in fig. 1, left, and can be compared with the input function for the standard SIR model on the right hand side.

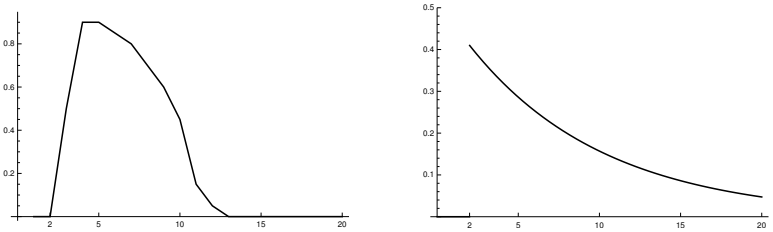

FIGURE 1. Left: $\gamma(j)$ of the adapted model, here linearly interpolated. Right: Graph of γ_{SIR} with values for $\lambda = 0.41, \mu = 0.12$ close to beginning of Covid-19 in Germany according to [3, fig. 1]. For both cases cut-off after 20 days.

We remark that this is the point where one can see why in [2] the authors speak about an"incessant misconception" if people apply the SIR-model indiscriminately (and the situation is not so different for the SEIR model). The function $\gamma(t) = \lambda e^{-\mu t}$ – at least for Covid 19 – has nothing to do with the medical data observed by the virologists. This may have considerable consequences for conditional predictions, see section 8.

A model using the biological input function γ above will be called a K-McK-model adapted to Covid-19, or for short an *adapted K-McK-model*. To specify such a model one has to estimate three further parameters p_c, $\alpha(k)$ and ξ.

The next parameter one has to chose is p_c. It is difficult to estimate and depends on the country and the action mode of the health institutions. For Germany it is estimated between 5 and 8; where for the year 2020 officials of the health system estimate it in the upper range. So we work with $p_c = 7$. There are data that indicate that the ratio of recorded infected α was about 0.25 for Germany at the beginning of the epidemic, but changed to approximately 0.5 during the course of the year, probably due to the increasing test rates [7]. Since the influence of the dark sector was negligible in the first months we haven chosen $\alpha = 0.5$ for the whole period. Finally, it is difficult to estimate the relative strength of infectivity ξ of the unrecorded infected from empirical data, so we take $\xi = 1$, i.e., we assume that people in the dark sector (including the asymptomatic ones) are basically as infectious as those in P_c.

4. From data to model curves

Now we have determined all parameters except the contact rate κ. In this paper we are mainly interested in reproducing the curves, so we start from the data given by the *Humdata* project of the Johns Hopkins University for 200 world countries and

regions.[1] More precisely, we estimate the values for $S(k)$ from the listed confirmed cases and the ratio α. In this step the estimated time translation $e + p_c$ between the times of infection and recording has to be taken into account. Moreover, there are weekly fluctuations of the values of newly recorded cases (the first differences of the data lists of confirmed cases) due to weekends where data taking and transmission is usually slowed down considerably. Since we are interested in the central development of the epidemic we replace the daily values of the newly recorded by 7-days averages and also the values of the susceptible people $S_7(k)$.

In the next step we compute the daily contact rates $\kappa(k)$ according to the observation above. This allows one to give the tautological model picture of the values of $S_7(k)$ according to the observation above; but of course this is not what we want. We therefore investigate the values of $\kappa(k)$ and look for periods where they are more or less stationary and can be approximated by a constant. Such time intervals may be interpreted as periods where the behaviour of the population is roughly the same and no crucial change of the virus occurs. The transition times between such stationary phases seem to characterize periods in which the contact behaviour of the population changes. This is often the result of political measures (non-pharmaceutical interventions) reducing the contacts or allowing more contacts, which in many cases may lead to new levels where the contact rates can again be approximated by a constant. Besides this, eyeballing will indicate further levels of near-constancy that then need to be justified. For example, this explanation might be that the changed contact rate is the result of a climate change from summer to winter where people meet more in closed rooms, or the other way round. It could also be that the infectivity of the virus changes (e.g. by the rise of a more aggressive mutant). In our view the most important criterion for the applicability of a Kermack/McKendrick type model to a real life epidemic boils down to the question whether or not replacing the contact rates during periods where they stay nearly constant by their averages, or a value close to it inside the respective σ-interval, leads to a good approximation of the data curves.

In the case of Germany one observes seven rather obvious periods between March 2020 and mid January 2021 in which the κ can be approximated by a constant, ignoring fluctuations (see fig. 2). Some of these constancy intervals can be interpreted as results of interventions, others not: There was a series of three interventions in March 2020 resulting in the first constancy interval from March 24, 2020, ($= t_1$) to April 26 ($= t_2$), a rather short interval. Around April 26 the interventions were reduced and the result is a long period until July 3 ($= t_3$). Already in this phase people were presumably getting more careless, the vacation period started and the contacts increased, which lead to the third period, lasting until September 27 (t_4).

[1] `https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases`

In this phase the mean reproduction rate rose noticeably above 1. With the beginning of the cooler temperature, life moving more to closed rooms, the reproduction number rose to about 1.5 in the fourth interval until October 31 (t_5). This provoked new containment measures, at first at Oct. 16 and Nov. 2 which are reflected by the fourth interval lasting until Nov. 26, in which the daily new infections went down. They started to rise again, probably because of early Christmas shopping at Nov. 26 (t_6 =). A partial lockdown (schools, restaurants, cultural activities) at December 16 (our t_7) brought the reproduction rate below 1 and resulted in, on average, falling numbers of new infections. This period lasts until the end of the data considered here in mid January (Jan. 15).

The model values of the contact rates κ_j in the intervals $[t_j,\, t_{j+1} - \Delta_{j+1}]$ ($1 \le j \le 7$), with varying durations Δ_j of the transition periods, are chosen close to the mean values, with small deviations inside the σ-interval allowed if this improves the fit considerably. The model κ_j and the (effective) model reproduction numbers ρ_j at the left edge of the intervals are given in the following table:

κ_0 and model κ_j, ρ_j, t_j for intervals J_j for Germany							
	J_1	J_2	J_3	J_4	J_5	J_6	J_7
κ_j	0.131	0.162	208	0.271	0.180	0.207	0.164
ρ_j	0.73	0.90	1.16	1.50	0.99	1.12	0.88
t_j (M/D/2020)	03/24	04/26	07/03	09/27	10/31	11/26	12/16

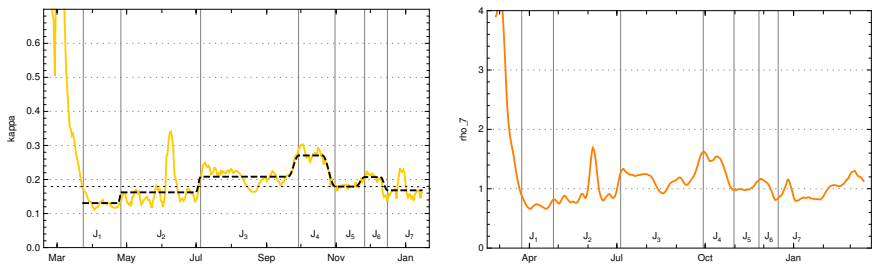

FIGURE 2. Left: Relative contact rates for Germany $\kappa(k)$ until mid January 2021 (yellow) and model values in constancy intervals (black dashed), critical value corresponding to reproduction rate = 1 (dotted straight line). Right: Corresponding reproduction rates.

Using this input our recursion formula determines the model values for Q_{new} as shown in figure 3.

5. A striking application

So far we have estimated all parameters from empirical information available to us except for the contact rate $\kappa(k)$, which we derive from the JHU data. Most of the

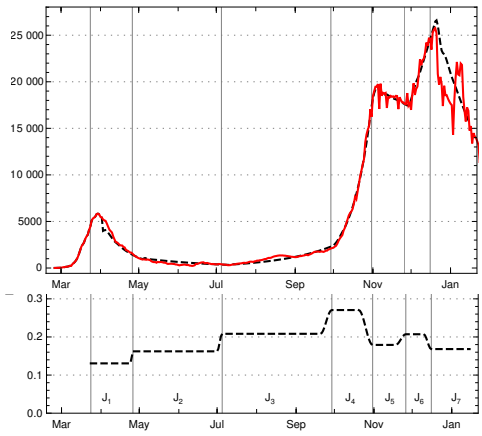

FIGURE 3. Top: JHU data (solid red) and model values (black dashed) for the number of daily new recorded cases Q_{new} (7-day averages, red) in Germany. Bottom: κ used in model, like in fig. 2 left.

parameters cannot be influenced by politicians, but the contact rates can. Moreover there is another one that is highly dependent on social rules, namely the time p_c between the beginning of infectivity and quarantine. One could make attempts to reduce this and check what happens if *all other parameters are unchanged, including the contact rate*. If we assume that the contact rate is unchanged and we change the value of p_c this will change the model curve. This is clear without any model. But it is one of the strengths of our model that the input parameters have a direct relation to data; in particular p_c shows up as a central parameter. For this reason a change of p_c results in a *quantitative effect* for the model curves. This effect is drastic. The following graphic shows what, according to our model, would have happened in Germany if from June 2020 on one would have reduced the time until quarantine by even one day, from 7 to 6.

Whereas the actual number of new infections reported per day (7-day average) rose to more than 25000 in December, the model predicts that the maximum would have been a bit less than 5000 if the time until quarantine had been reduced in this way.

6. Comparison with the standard SIR model: An example

As memtioned in the beginning of section 2 we want to return to the paper of Kermack and McKendrick and discuss a case, where there is not a sharp border between being infectious (propagating) and recovery. This situation can be modelled by a function $b(j)$, the *ratio of infectious persons* at day j in relation to the first day.

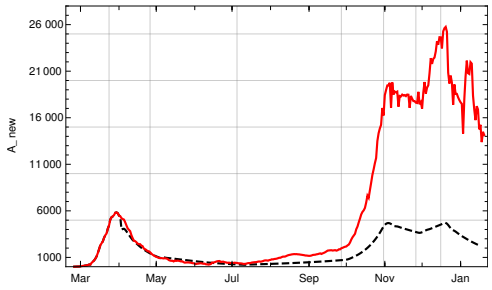

FIGURE 4. Comparison of 7-day averages of recorded daily newly sent to quarantine Q_{new} in Germany (JHU data, solid red) with model values (black dashed) assuming a reduction of p_c from 7 to 6; dark sector with $\alpha = 0.5$.

In addition one has the function $\varphi(j)$ which measures the mean *medical strength of infectivity* of a person still infectious at day j counted from the moment of infectivity on ("rate of infectivity" in the terminology of Kermack-McKendrick). The product of these functions $\gamma(j) = \varphi(j)b(j)$ is in this context the function which measures the strength of the infection. There is a similar version for continuous time t (cf. [1, eq. (1)]). Kermack-McKendrick called $\psi(t) = -b'(t)/b(t)$ the "rate of recovery".

In the widely used SIR model both, the rate of infectivity and the rate of recovery, are assumed to be constant, $\varphi(t) = \alpha$, and $\psi(t) = \beta$. This results in a pure exponential decay for the function $\gamma(t) = \alpha e^{-\beta t}$. For Kermack-McKendrick this was a very special case introduced close to the end of the paper only [8, p. 713]. It leads to the well known standard SIR differential equations. For Covid 19, however, none of the two the assumptions for the standard SIR is satisfied.

Alone this should make people hesitate to apply the standard SIR model to Covid-19. But there are cases where a "bad" model is useful, perhaps even more useful than a "better" model. We discuss now why we doubt this by comparing the model curves using the standard SIR model with those of our model for Covid-19. To do this we chose model parameters for both models so that at the beginning the two models are approximatively equal. The standard SIR model is based on $\gamma_2(t) = \lambda e^{-\mu t}$ (resulting in the well known differential equations for SIR). We choose the start conditions of the our γ-K-McK-model for Covid-19 such that they agree (approximately) with the SIR-model during the first days (exact only at the origin). By construction of the SIR-model the function $I(t)$ of actually infected at time t is nearly an exponential function as long as the function S is nearly constant and the contact rates are unchanged.

But in reality jumps of the contact rates take place, for example as the result of interventions. Models which are close to reality should behave equally if the contact rate (which for the SIR-model corresponds to λ, and for the adapted model to κ) is

changing by a certain percentage. Figure 5 shows that there is a drastic difference between the functions I and Q which describe the number of infected respectively counted as infected in the two models.

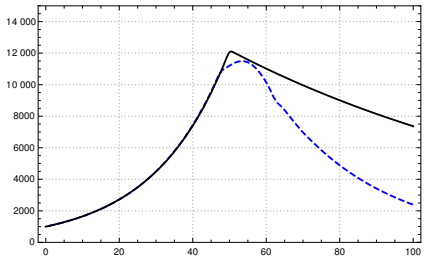

FIGURE 5. Comparison of $I(t)$ for SIR (black) and Q(k) of the present γ-model (dashed blue) with vanishing dark sector, $\alpha = 1$, with identical exponential increase in the initial upswing. Parameter values: $N = 8 \cdot 10^6$, $I(0) = 1000$, $\lambda_1 = 0.15$, $\mu_1 = 0.1$, $\kappa_1 = 0.256$. Reduction of reproduction rate by 40 % in both cases between days 49 and 51. Constant coefficients for $t \leq 29$ and $t \geq 31$, beyond $t = 31$ change to $\lambda_2 = 0.6 \cdot \lambda_1 = 0.09, \mu_1 = \mu_2 = 0.1$; $\kappa_2 = 0.6 \cdot \kappa_1 = 0.15361$.

These considerations show that the choice of the model may result in important differences for modelling the development of the epidemic. This is already indicated by the simpler γ-K-McK model using box function considered in [5, figs. 5, 6] and [9, fig. 3]. Needless to add that we consider our γ-K-McK model more realistic for Covid-19 than either the standard SIR approach or the box-shaped γ-K-McK model.

In adddition the figure shows clearly that the delay structure of the γ-K-McK model leads to a smooth transition between the phases before and after the change point, while for the SIR model the change is nearly a cusp.

7. Vaccination, new mutants, and mass testing

In this section we extend our model step by step to include the effects of vaccination, in addition new mutants and in further addition mass testing. Please note that all our functions $S(k)$, $E(k)$ etc. have a new definition in this section and should not be mixed up with the previous formulas.

Once the past and the expected future rates of daily vaccination are known or estimated, it is simple to extend the model to take vaccination into account. If $V_{\text{new}}(k)$ is the number of vaccinated persons on day k and $V(k)$ the total number of immunized by vaccination on day k,

$$V_{\text{new}}(k) = V(k) - V(k-1)\,,$$

equ. (1) turns into

$$E_{\text{new}}(k) = S(k-1) - S(k) - V_{\text{new}}(k)\,. \tag{15}$$

The recursion equation (7) has to be amended accordingly by the additional summand $V_{\text{new}}(k)$ on the right hand side (see the Theorem below in the special case $\zeta = 1$ and $\beta = 0$) and the terms $\mathbb{P}_c$ and $\mathbb{P}_d$ have to be adapted:

$$\mathbb{P}_c(k-1) := \sum_{j=1}^{p_c} \gamma(j)\alpha(k-j)\big[S(k-j-e-1) - S(k-j-e) - V_{\text{new}}(k-j-e)\big] \tag{16}$$

$$\mathbb{P}_d(k-1) := \sum_{j=1}^{d} \gamma(j)\big(1-\alpha(k-j)\big)\big[S(k-j-e-1) - S(k-j-e) - V_{\text{new}}(k-j-e)\big] \tag{17}$$

Making provision for the appearance of a new mutant of the virus with stronger or weaker infectivity is slightly more involved. If we assume that the function of infectivity changes only up to a time dependent scalar factor $\zeta(k)$, with k the epidemic time scale, $\gamma(j)$ (as above j denotes here the day after the onset of infectivity) has to be replaced by

$$\zeta(k)\,\gamma(j)\,.$$

In this way the coefficient of the right hand side of eq. 7 is enriched by the factor $\zeta(k)$ (see eq. 23).

Now we consider (in addition to vaccination and new mutants) reliable unspecific daily mass testing of a specified subset M_{t} (with cardinality N_{t}) of the total population M (with cardinality N). Such testing may be realistic if based on *new generation sequencing* proposed in [10, 14]. We denote the complement of M_{t} by $M_{\text{n-t}}$ and assume that the decomposition $M = M_{\text{t}} \cup M_{\text{n-t}}$ is kept fix over the whole period of testing. To simplify the considerations we assume that testing takes place all the times every day and is error free (sensitivity 100% and specificity 100%). If in applications the testing starts later, or takes place on selected days a week only, one has to modify the algorithm appropriately. We call $\beta := N_{\text{t}}/N$ the *test ratio*. Assuming that the tests are reliable implies that there is no dark sector within M_{t}. Moreover we assume that the test is sensitive from day l of the period of infection on. A positively tested person is assumed to be sent to quarantine on the *following* day.

We further assume perfect contact mixing between M_{t} and $M_{\text{n-t}}$. If we denote the number of newly infected at day k amongst people in the tested group M_{t} by $E_{\text{t,new}}(k)$ and in the non-tested group by $E_{\text{n-t,new}}(k)$, we have:

$$E_{\text{t,new}}(k) = \beta\, E_{\text{new}}(k)\,, \qquad E_{\text{n-t,new}}(k) = (1-\beta)\, E_{\text{new}}(k)\,, \tag{18}$$

where $E_{\text{new}}(k)$ is given by formula (15).

Now we count the number $P_{\text{t}}(k)$ of persons from $M_{\text{t}}(k)$ who are effectively propagating the virus under the test regime on day k. Since it takes at most l days until

a newly infectious person is either positively or negatively tested we obtain:

$$P_{\mathrm{t}}(k) = \sum_{j=1}^{l} E_{\mathrm{t,new}}(k-e-j+1) \tag{19}$$

This formula has to be amended for weekly tests cycles in which tests are not take daily but only on specified week days.

We define $Q_{\mathrm{t}}(k)$ as the number of persons from $M_{\mathrm{t}}(k)$ who are in quarantine or isolation on day k. Observing that the quarantine starts the day after positive testing we find for the number of persons newly sent to quarantine on day k:

$$Q_{\mathrm{t,new}}(k) \;=\; E_{\mathrm{t,new}}(k-e-l) \tag{20}$$

The corresponding number for the people who are not tested is

$$Q_{\text{n-t,new}}(k) = \alpha(k-p_c)(1-\beta)\, E_{\mathrm{new}}(k-e-pc)\,, \tag{21}$$

where again $E_{\mathrm{new}}(k)$ is given by formula (15).

Next we define

$$\mathbb{P}_{\mathrm{t}}(k-1) \;=\; \beta\Big(\sum_{j=1}^{l} \gamma(j)\, E_{\mathrm{new}}(k-e-j)\Big) \tag{22}$$

,

where $E_{\mathrm{new}}(k)$ is given by formula (15).

This number counts only the propagating people who are tested (those in M_{t}); the corresponding number for those who are are not tested (those from $M_{\text{n-t}}$) has to be added. It is essentially given by 16 and 17, but the expressions have to be multiplied by the factor $(1-\beta)$, i.e. the percentage of non-tested.

This allows to formulate our final result including the effects of vaccination, new mutants and mass testing as described above

Theorem. *If we take vaccination, the rise of new mutants and unspecific mass testing like above into account, the central recursion of our model becomes*

$$\begin{aligned} S(k-1) - S(k) \;&=\; V_{new}(k-1) + \\ &s(k-1)\,\kappa(k-1)\,\zeta(k-1)\,\Big(\mathbb{P}_t(k-1) + (1-\beta)[\mathbb{P}_c(k-1) + \mathbb{P}_d(k-1)]\Big)\,, \end{aligned} \tag{23}$$

where $\mathbb{P}_t$ is given by formula (22), $\mathbb{P}_c$ by (16), and $\mathbb{P}_d$ by (17).

The number of daily new detected infected by testing is given by equation (20), the ones in the non-tested ensemble by equation (21). So the total number of new detected infected is

$$Q_{new}(k) = \beta\, E_{new}(k-e-l) + \alpha(k-p_c)(1-\beta)\, E_{new}(k-e-pc)\,, \tag{24}$$

where $E_{new}(k)$ is given by formula (15).

Equation (7) is, of course, the special case of equation 23 with $\zeta(k) = 1$, $\beta = 0$, *and* $V_{new}(k) = 0$ *(for all* k*).*

In this case the method for determining the empirical values for the $\kappa(k)$ described in the *Observation* above finds the values of the product $\kappa(k)\zeta(k)$ rather than $\kappa(k)$ itself. The latter can be isolated by correcting for the factor $\zeta(k)$, which is to be determined from the data.

In the following we apply this theorem and study the effect of mutants, of vaccination and a scenario of mass testing for Germany.

In the German case, we know that the share of infections carrying the mutant B.1.1.7, determined by genetic sequencing, rose approximately in a linear progression from roughly 6% in calendar week (CW) 4, the last week of January 2021, to approximately 70% in CW 10 and 88% in CW 12 [13]. As we understand, the sequencing data evaluate the probes of infections which occurred about a week earlier; we therefore time translate the sequencing data 7 days backward (fig 6, blue dots). We may safely assume that about 2 or 3 weeks after CW 10, i.e. in early April, B.1.1.7 will be dominating the infections observed in Germany.

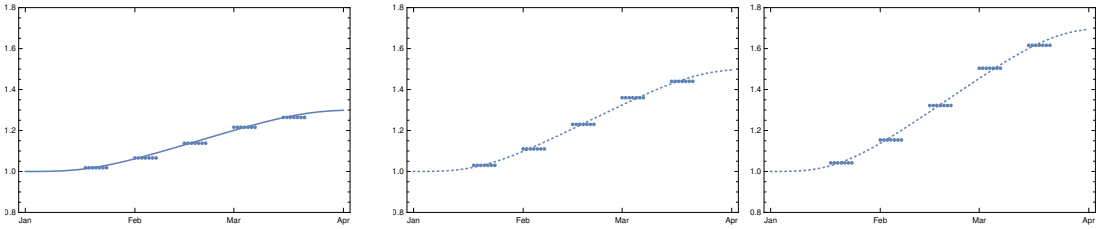

FIGURE 6. Left: Scaling function $\zeta(k)$ (blue line) with $c = 1.3$ (left), $c = 1.5$ (middle) and $c = 1.7$ (right)) for the infectivity of the mixture of "old" virus and new mutants, mainly B.1.1.7 in Germany. The ratio of B.1.1.7 as given in [13], backward time translated by 7 days and weighted by c, is shown by series of blue dots.

We look for a datum where the transition begins, a duration of the transition and an upscaling factor c which describes the new value of γ after the transition and a "natural" function which describes the transition in such a way that, after a time where it is nearly constant and equal to 1, it changes quickly into a nearly linear function and then becomes nearly constant equal to c. There is a well known C^∞ function, the so called transition function,

$$g(x) := \frac{f(x)}{f(x) + f(1-x)},$$

where

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{for} \quad x > 0 \\ 0 & \text{for} \quad x \leq 0\,. \end{cases}$$

The characteristic of this function is that it is 0 for $x \leq 0$ and 1 for $x \geq 1$. It stays close to 0 for a while for x near 0 and close to 1 for a while for x near 1 and in between it changes quickly into an almost linear function. So it fits well with the observation of the progression of the mutant. This function describes a transition which during a time interval of length 1 raises the level from 0 to 1. Adding 1 and adjusting the time duration from 1 to d and the up-scaling factor from 1 to c leads to the function

$$\zeta_{c,d}(x) = 1 + (c-1)g\big(\frac{x}{d}\big)\,. \tag{25}$$

with the desired properties of rescaling γ by the factor c within d days. The date of the beginning of the transition and the duration d as well as the up-scaling factor c are roughly known from sequencing data. B.1.1.7 was first identified in Germany in late December 2020 and the sequencing data of [13] show that the transition ends in early April 2021. According to the virological estimates the final scaling factor is constrained by $1.3 \leq c \leq 1.7$. For constraining it more narrowly we investigate the effects of the choice of c on the rise of $\kappa(k)$ in early March (see fig. 7). For $c = 1$ or $c = 1.3$ the resulting rise of $\kappa(k)$ (fig. 7, top) is too high to be plausible, as no relaxing was announced by the German authorities and the majority of the population accepted the restriction measures issued in December 2020 quite patiently. In early March a partial opening of shops was permitted and a slightly disorganized discourse started on possible relaxation of restrictions, which might be acceptable if combined with an increase of unspecific testing. This is sufficient for explaining a moderate rise of κ at the beginning of March by 5 - 10 %, as we find it for $c \approx 1.5$ (fig. 7, bottom left). The hypothesis $c \approx 1.7$ appears again implausible as it would indicate that the early March discourse was without any effect on the contact rate (fig. 7, bottom right).

Summing up, we obtain a consistent and plausible choice for the beginning of the transition at December 28, 2020, a duration $d \approx 100$ and an up-scaling factor $c = 1.5$.

If we now factor in the available data on vaccinations in Germany,[2] we can analyse the impact of vaccination on the course of the epidemic in Germany. Figure 8 shows the numbers of daily new infections (7-day averages); red the JHU data, black dashed the model reconstruction taking the vaccinations into account, pink dotted the model values under the counterfactual assumption that no vaccination had taken place in early 2021.

In fig. 9 one can inspect the change of the effective reproduction numbers for Germany in early 2021 (data available until end of March) and an outlook on the future months April and May under the assumption of no essential changes in the contact rates (orange – empirical values, black dashed and dotted model with vaccination,

[2] `https://impfdashboard.de/static/data/germany_vaccinations_timeseries_v2.tsv`

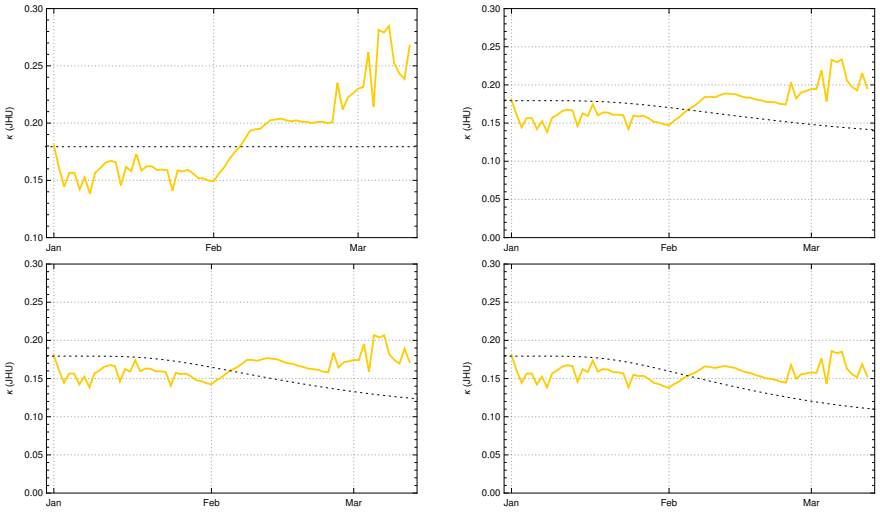

FIGURE 7. $\kappa(k)\zeta(k)$ for Germany (yellow) in the first three months of 2021. Top left $c = 1$, right $c = 1.3$. Bottom: left $c = 1.5$, $c = 1.7$. The dotted downward swung line (top left dotted straight line) shows the critical value for $\kappa(k)$ (corresponding to the basis reproduction rate = 1).

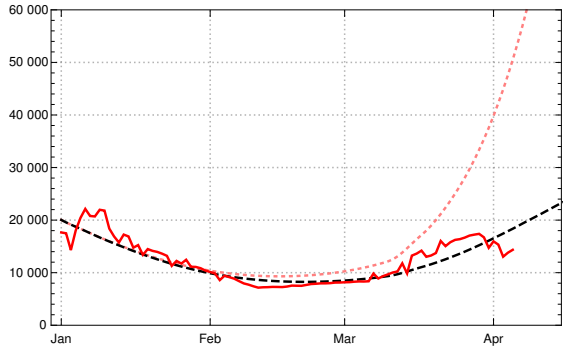

FIGURE 8. Daily new infections for Germany based on data available on April 09, 2021) Red JHU data, black dashed model, pink dotted model with the same contact rates $\kappa(k)$ and new mutant B 1.1.7 under the (counterfactual) assumption that no vaccinations had taken place.

pink dotted model without vaccination). Under this assumption the vaccinations will push the effective reproduction rate below the critical value 1 at the turn from April to May 2021. If no vaccination would take place this would happen only in late May – after a catastrophic increase of daily new infected by herd immunization induced by the infections.

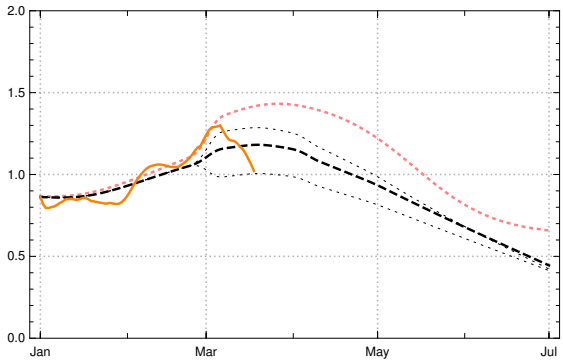

FIGURE 9. Reproduction number $\rho(k)$ for the model scenario above in Germany. Orange: empirical values (using JHU data available until April 09, 2021). Black: model $\rho(k)$; dashed basic scenario, dotted the boundaries of 1 sigma domain of empirical $\kappa(k)$ in the last constancy interval. Pink dotted: model $\rho(k)$ if no vaccination had happened.

We would like to demonstrate the strong effect of the vaccination by showing the catastrophic rise of daily new infections without vaccination, assuming no essential change of contact behaviour of the German population (in comparison with March 2021). During April the number of daily new infections would rise above 100.000, see fig. 10, dotted pink curve. But because of the ongoing vaccinations our model lets us expect the local maximum of the third wave for daily new infections at the turn from April to May 2021 with values (7-day averages) about 25 000 (black dashed; dotted the boundaries of the 1 sigma interval for the empirical $\kappa(k)$ in the last constancy interval). Note that the data available on April 09, 2021, are subject to an artificial drop of $Q_{\mathrm{new}}(k)$ induced by the Easter days (April 2–5); this leads to additional uncertainties in the data evaluation and conditional predictions. In any case the figures 9 and 10 demonstrate the important role of the vaccinations for turning the tide of the epidemic in Germany during the months April and May 2021.

Finally we apply our theorem for studying the effectiveness of mass testing techniques, e.g., by *new generation sequencing* proposed in [10, 14]. The question is how fast testing on a daily basis of a certain subset M^{t} can be expected to suppress an epidemic of the Covid-19 type. We model here an ideal mass test (sensitivity and specificity both 100%) of a fixed subset M^{t} containing 60% of the total population on a daily basis, starting on October 15, 2020. Moreover we even assume (counterfactually) that the *contact regime* of early October is *not being changed* by further restrictive measures later in the year, while the effects of vaccinations and the rise of the new mutant B.1.1.7 since late December 2020 have been included.

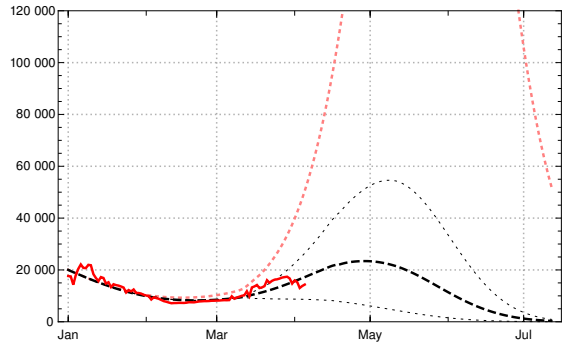

FIGURE 10. Future scenario of daily new infections (7-day averages) for Germany, based on data available on April 09, 2021, with parameters described in the main text. Red: JHU data. Black dashed: model scenario based on last available for empirical contact coefficients $\kappa(k)$; black dotted: model scenario boundaries of 1 sigma domain of empirical $\kappa(k)$. Pink dotted: model scenario under assumption of no vaccination.

In other words, we use the right hand side (rhs) of eq. 7 for $t <$ Oct 15 with the contact coefficients of the German model and the rhs of eq. (23) with constant contact coefficient κ_4 (see sec. 4). Using the characteristic function $f_{\text{test}}(t) = \chi[t, J]$ of the interval $J = [t_a, \infty)$ the recursion is then

$$
\begin{aligned}
(26) \qquad S(k-1) - S(k) \;&=\; V_{\text{new}}(k-1) \;+ \\
& s(k-1)\,\kappa(k-1)\,\zeta(k-1)\,\Big[(1 - f_{\text{test}}(k-1))\big(\mathbb{P}_c(k-1) + \mathbb{P}_d(k-1)\big) \\
& + f_{\text{test}}(k-1)\Big(\mathbb{P}_{\text{t}}(k-1) + (1-\beta)[\mathbb{P}_c(k-1) + \mathbb{P}_d(k-1)]\Big)\Big]
\end{aligned}
$$

Figure 11 shows the suppression of $Q_{\text{new}}(k)$ which are to be expected in terms of our model calculation.

It demonstrates a high in-principle effectiveness of such an approach, although in an idealised scenario. Under a test regime similar to the one above the rise of the new mutant would only lead to a slight increase of registered new infections starting in early March 2021. But it would be held below roughly 1000 by the increase of vaccinations in April and May (with actual data on vaccination for Germany). We have checked that for a moderately reduced reliability of the test with a ratio of 0.9 recognized infectious persons (sensitivity 90%) and daily testing the picture would not change significantly (the number of false positive tested for genetic screening tests is expected to be extremely low; specificity 1 with extremely high precision).

We would like to remark that the effect of mass testing has a similar effect as the shortening of the time between the offset of symptoms and sending people to

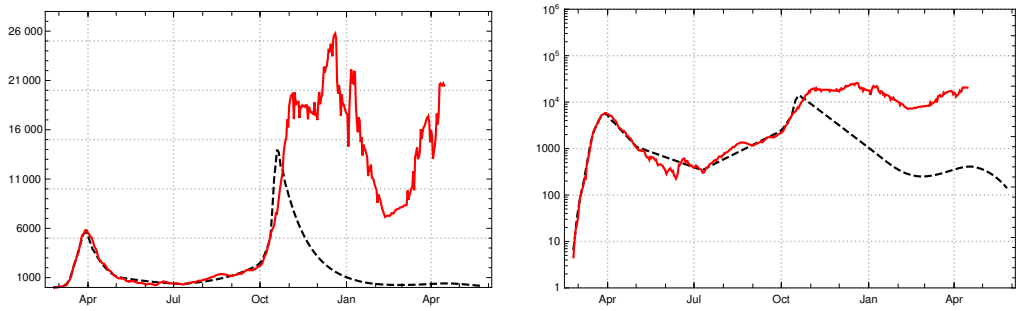

FIGURE 11. Comparison of development of registered daily new infections $Q_{\text{new}}(k)$ (7-day averages) of JHU data for Germany (red) with expectation under a test regime with daily testing of 60% of the population and *no* further *contact restrictions* than those already operative in late September 2020 (black dashed). Left: linear scale. Right: logarithmic scale.

quarantine or isolation as discussed in our striking application. Mass testing is a particular effective way to shorten this period for a certain part of the population.

8. Comparison of reproduction numbers

The Robert Koch-Institut (RKI), Berlin, publishes the daily data for an epidemic in Germany and determines the daily reproduction numbers from these data. The RKI calculation uses a method described in [6] for a stochastic estimation of the numbers of newly infected, called $E(t)$, from the raw data of newly reported cases,. The calculation of the reproduction numbers works with these $E(t)$ and assumes constant generation time and serial intervals of equal lengths $\tau_g = \tau_s = 4$ [12]. Two versions of reproduction numbers are being used, a day-sharp and therefore "sensitive" one $\rho_{rki,1}(t) = \frac{E(t)}{E(t-4)}$, and a weekly averaged one,

$$\rho_{rki,7}(t) = \frac{\sum_{j=0}^{6} E(t-j)}{\sum_{j=0}^{6} E(t-4-j)}\,,$$

which we refer to in the following simply as $\rho_{\text{RKI}}(t)$.

The paper remarks that the RKI reproduction numbers ("R-values") $\rho_{\text{RKI}}(u)$ indexed by the date u of calculation refer to a period of infection which, after taking the incubation period ι between 4 and 6 days into account, lies between $u-16,\ldots,u-8$ (with central day $u-12$ in the interval). The reproduction numbers calculated in the adapted K-McK approach are very close to the RKI reproduction numbers. The main differences lie in the usage of different raw data bases (RKI versus JHU) and the adjustment of the raw data (stochastic redistribution $E(t)$ versus sliding 7-day

averages $\hat{Q}_{new,7}$). After a reasonable time shift the agreement between the reproduction numbers of the adapted γ-K-McK approach and the $\rho_{\mathrm{RKI}}(k)$ are close. The stochastic smoothing of the RKI data seem to lead to smaller amplitudes of the fluctuations which otherwise are in striking coincidence (fig. 12).

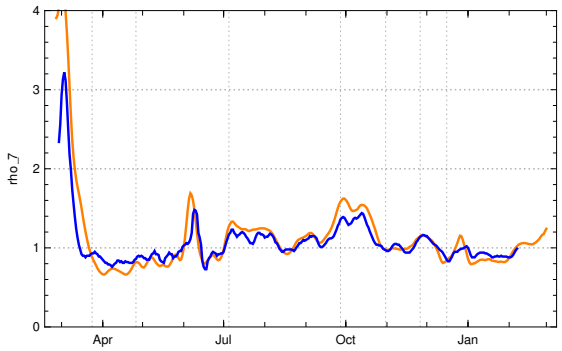

FIGURE 12. Empirical reproduction numbers $\rho(k)$ of the adapted γ-K-McK model for Germany (orange) and reproduction numbers $\rho_{\mathrm{RKI}}(k-9)$ of the RKI (blue); both of them 7-day averages.

9. Discussion

We formulate four points with regard to the adapted K-McK-model, which at present seem the most important ones to us.

(1) A model has to be useful for analysing an epidemic on the basis of the available data. The most important data for this purpose are the daily newly infected; but one is interested in other numbers also, like the number of actually counted as infected or those in hospital, or how many people are actually infectious or how many people are already immune (keyword: herd immunity). In particular one is interested in the daily reproduction number (see formula (13). All these numbers are given by our model. Knowing these numbers is an important information for the politicians as well as for citizens. For politicians such numbers are the basis to think about regulations and for citizens this information is a helpful guide for how to behave.

> *The present model is a useful tool for determining key figures of an epidemic like the reproduction number and the number of people in quarantine or hospital.*

(2) A reliable model can do more. The optimum would be that it allows to predict the future. But in the strict sense this is impossible, since important parameters like the contact rate changes depending on the behaviour of the population. If one knows the development of the epidemic in the past (in terms of a model) and assumes that the contact rate is not changing drastically, it is possible to extrapolate

the development for some time. We indicate how this may be done by looking at the situation in Germany at the beginning of April 2020. In a series of preliminary steps a lockdown was imposed on March 22. Only a few days later it pushed the reproduction number below 1, and about a week later down to about $\rho = 0.7$ (see fig. 2). It fluctuated about that number for roughly a month. This was observable in the data about a fortnight later, so that around April 10 the number was stable already for about a week. If at that moment one would assume that the behaviour of the population is not changing very much, one could use the model for predicting the future during the next weeks quite successfully until the contact rate was changing substantially. Figure 3 shows how well this worked until the end of April when the restrictions were partially relaxed and the reproduction number rose. In other words:

> *The adapted K-McK-model is a reliable tool for extrapolating the development in times where the reproduction number is stable and there are reasons to assume that this stays so for a while.*

At the beginning of the epidemic there was no experience what the effect of interventions would be. But over the time there was a chance of using the model for observing the effects and, if possible, to draw realistic conclusions. A quantitative estimate of the effectiveness of non-pharmaceutical interventions in terms of the change of contact rates (or reproduction numbers) determined in the model framework is a delicate point. But keeping this in mind, a realistic model like the adapted K-McK-model may be used heuristically for deriving useful information about how certain interventions will influence the development by comparison with other ones in the past which took place under similar conditions. Like any realistic model the present one may be used as a tool for learning how certain interventions under specified conditions (like climate) influence the development. This was also tried in [3] for the three interventions in Germany in March 2020, claiming a direct relation between each of these and a new level of the reproduction number. In the light of the data evaluation represented in figure (2) we are not able to confirm this result. In the decline of the contact rates, respectively reproduction numbers, in Germany during March 2020 no clear intermediate steps can be discerned which could be read as the signature of the first two interventions. To assume the existence of constancy intervals between two non-pharmaceutical interventions per se seems doubtful to us. In this phase we only see a cumulative effect of all three interventions.

On the other hand, if the data evaluation shows a relative stability of the contact rate $\kappa(k)$, a conditional prediction for the coming development is possible, sometimes even for several weeks. "Conditional" means here under the assumption of no essential change of the behaviour of the people, relative stability of climatic conditions and no mutation of the virus. In section 4 time intervals in which such conditional predictions could be made have been discussed. For example one could

recognize a stable trend of increasing numbers of daily new infected for Germany already in late July and early August 2020 and could have foreseen the approach of the second wave long before autumn (see fig. 3, bottom, and table in Section 4).

(3) As conditional predictions with the present model turn out to be convincing and reliable we also consider future scenarios for the overall development of the epidemic for up to a few months as informative. For an example scenario which can be checked against the real course of the epidemic see end of section 7. Such scenarios may be helpful for exploring alternatives of actions (e.g. non-pharmaceutical interventions) and for identifying necessary steps to avoid the breakdown of the health system.

> *The study of future scenarios with the γ-K-McK model may be useful for persons or institutions looking for orientation with regard to alternatives for actions and/or external changes of condition (rise of new mutants, climate change).*

(4) As we have seen there is a parameter which can, in principle, be influenced by socio-political decisions comparatively easily, the time between infection and the day people go to quarantine. We therefore emphasize again:

> *The model shows that reducing the time until entering quarantine by one day leads to a drastic improvement.*

Whether one can reduce it depends on various factors, in particular the infrastructure and effectiveness of the health system. Discussions with experts about the German health system have convinced us that there are good chances for lowering the time until quarantine by at least one day. Of course this needs a great effort: One has to enable the health system to carry it out and one has to convince the population to follow the corresponding rules. But the latter should have good chances since this is a restriction which hits only a small number of the population: those who show first symptoms or are identified as Covid-positive in an unspecified test.

(5) A comparison between different models is difficult. If models are based on clear principles the first thing one could do is to compare the principles and discuss their strength and weakness. If one wants to compare the standard SIR model with our adapted model this is simplified by the fact that they are both based on the same principle, the Kermack-McKendrick idea of a central input parameter, a function γ which measures the medical strength of the infection. So this comparison boils down to comparing the two input functions. And there it is obvious that both input functions differ very much (see figure 1), and since the γ for the adapted model is based on virological studies it has to be considered more realistic. But it could be that the models are so robust that this difference doesn't play a big role. This is actually the case as long as one keeps the contact rate as well as the number of susceptible people approximately constant. Then for both models the number of people counted as infected are approximately exponential functions and so after

adjusting the parameters the models are essentially equivalent. But this changes if the contact rate jumps in a short time to a new constancy level. Then both models differ drastically as we demonstrate in figure 5. And then the model with the more realistic input function γ is clearly the first choice. The same happens if one looks at the long term behaviour where the number of susceptible people changes considerably.

Saying this, one should not forget that no model is a mirror of reality. It can even happen that a model based on less realistic principles may work better, but so far we don't see a sign of this in the comparison of the adapted model with the standard SIR model. So we conclude:

> *Our comparison of the adapted model with the standard SIR model indicates that the former one is based on more realistic assumptions. Since the differences in the long run or when the contact rates jump are large, the adapted model should be the first choice.*

Acknowledgments

We thank Odo Diekmann for discussing our thoughts. His help was invaluable to us, non-experts in the field, for coming to a proper understanding of compartment models. Moreover, we thank Martin Bootsma, Robert Feßler, Jan Mohring, Robert Schaback and Hans Lehrach for hints and discussions, and in particular Don Zagier for his intensive support in the final preparation of the manuscript. Calculations and graphics were made with Mathematica 12.

References

[1] Brauer, Fred. 2017. "Mathematical epidemiology: Past, present, and future." *Infectious Disease Modelling* 2(2):113–127.

[2] Breda, Dimitri, Odo Diekmann and W.F. de Graaf. 2012. "On the formulation of epidemic models (an appraisal of Kermack and McKendrick)." *Journal of Biological Dynamics* 6, Suppl 2:103–117.

[3] Dehning, Jonas, Johannes Zierenberg, Paul Spitzner, Joao Pinhero Neto, Michael Wilczek and Viola Priesemann et al. 2020. "Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions." *Science* 369:eabb9789.

[4] Diekmann, Odo, Hans Ohmer, Robert Planqué and Martin Bootsma. 2020. "On discrete time epidemic models." In preparation.

[5] Feßler, Robert. 2020. "A general integral equation model for epidemics." Kaiserslautern: Fraunhofer Institut für Techno- und Wirtschaftsmathematik ITWM `https://www.itwm.fraunhofer.de/content/dam/itwm/de/documents/`

anwendungsfelder/20201201_fessler_IntegralEquationModelOfEpidemics.pdf.

[6] an der Heiden, Matthias and Osamah Hamouda. 2020. "Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting 23." *RKI Bulletin* pp. 10–16.

[7] HelmholtzZentrum, München. 2020. "Die Dynamik der COVID-19 Pandemie im Blick. Zwischenergebnisse zur zweiten Runde der KoCo19-Antikörperstudie. Pressemitteilung (23. 12. 2020): Prospektive COVID-19 Kohorte München (KoCo19)." www.klinikum.uni-muenchen.de/Abteilung-fuer-Infektions-und-Tropenmedizin/download/de/KoCo19/2020_12_23-PM_KoCo19_Runde2_Zwischenergebnisse.pdf.

[8] Kermack, William O. and Anderson G. McKendrick. 1927. "A contribution to the mathematical theory of epidemics." *Proceedings of the Royal Society of London* 115(772):700–721.

[9] Kreck, Matthias and Erhard Scholz. 2020. "Studying the course of Covid-19 by a recursive delay approach." arXiv:2101.08660, https://www.medrxiv.org/content/10.1101/2021.01.18.21250012v2.

[10] Lehrach, Hans; Church, George. 2021. "Der Ausweg aus der Pandemie." White paper http://www.alacris.de/wp-content/uploads/2020/05/Der-Ausweg-aus-der-Pandemie.pdf retrieved 04/15/2021.

[11] Mohring, Jan, Raimund Wegener, Simone Gramsch and Anita Schöbel. 2020. "Prognosemodelle für die Corona-Pandemie." Kaiserslautern: Fraunhofer Institut für Techno- und Wirtschaftsmathematik ITWM https://www.itwm.fraunhofer.de/de/suche.html?_charset_=UTF-8&numberResults=10&page=1&scope=ITWM&lang=de&queryString=Mohring+prognosemodelle.

[12] Robert Koch-Institut, RKI. 2020. "Erläuterung der Schätzung der zeitlich variierenden Reproduktionszahl R.". https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/Nowcasting.html.

[13] Robert Koch-Institut, RKI. 2021. "Bericht zu Virusvarianten von SARS-CoV-2 in Deutschland, insbesondere zur *Variant of Concern (VOC)* B.1.1.7." https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/DESH/Bericht_VOC_2021-03-31.pdf?__blob=publicationFile.

[14] Schmidt-Burgk, Jonathan, Ricarda M. Schmithausen and David Li et al. 2020. "LAMP-Seq: Population-scale Covid-19 diagnostics using combinatorial barconding." https://www.biorxiv.org/content/10.1101/2020.04.06.025635v2.

[15] Shayak, B. and Mohit M. Sharma. 2020. "A new approach to the dynamic modeling of an infectious disease." medRxiv https://www.medrxiv.org/content/10.1101/2020.10.30.20223305v1.

[16] Wölfel, R., V.M. Corman and W. Guggemos et al. 2020. "Virological assessment of hospitalized patients with COVID-2019." *Nature* (581):465–469.