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

Modeling the role of clusters and diffusion in the evolution of COVID-19 infections during lock-down

Wouter J. T. Bos · Jean-Pierre Bertoglio · Louis Gostiaux

February 19, 2021

Abstract The dynamics of the spread of epidemics, such as the recent outbreak of the SARS-CoV-2 virus, is highly nonlinear and therefore difficult to predict. As time evolves in the present pandemic, it appears more and more clearly that a clustered dynamics is a key element of the description. This means that the disease rapidly evolves within spatially localized networks, that diffuse and eventually create new clusters. We improve upon the simplest possible compartmental model, the SIR model, by adding an additional compartment associated with the clustered individuals. This sophistication is compatible with more advanced compartmental models and allows, at the lowest level of complexity, to leverage the well-mixedness assumption. The so-obtained SBIR model takes into account the effect of inhomogeneity on epidemic spreading, and compares satisfactorily with results on the pandemic propagation in a number of European countries, during and immediately after lock-down. Especially, the decay exponent of the number of new cases after the first peak of the epidemic is captured without the need to vary the coefficients of the model with time. We show that this decay exponent is directly determined by the diffusion of the ensemble of clustered individuals and can be related to a global reproduction number, that overrides the classical, local reproduction number.

1 Lock-down and modeling of the COVID-19 epidemic

Since the beginning of the COVID-19 epidemic in December 2019, many countries have chosen severe restrictions on mobility and transport. It has become a controversial question in public debate whether these decisions were necessary and inevitable, and whether the lock-down was maintained for an unnecessary long period or, on the contrary, was relaxed too soon in an unsafe way. In order to evaluate retrospectively such decisions, and to anticipate similar measures in the future, an in-depth understanding of epidemic propagation is needed. In the present investigation we focus on the seemingly universal way in which the SARS-CoV-2 disease is controlled by lock-down.

Indeed, in a number of countries in which lock-down was applied, the number of new infections decays exponentially after the peak of the epidemic with a very similar timescale (see Fig. 4). In the present investigation we shed light on the role of the finite size of spatially localized clusters on this phenomenon, by improving upon a simple paradigm model, the classical SIR model, and by assessing the resulting SBIR model against realistic data.

The SIR model in its simplest form contains three compartments, corresponding to susceptible, infected and recovered or removed individuals, and the dynamics are governed by two model parameters γ (recovery rate) and β (infection rate). A wide variety of SIR modifications have seen the light since its original formulation [13]. Many of them propose additional compartments (see for instance Massonis, Banga, and Villaverde [17] for an analysis of a great number of compartmental models), nonlinear incidence rates [7], and non-constant model coefficients [3, 8]. An overview of these models is beyond the scope of the present work and we refer to the review articles on the subject [12, 6].

In a large number of these epidemiologic studies using compartmental models, the dynamics are extremely sensitive on the temporal variations of the model parameters, and saturation of the epidemic at a realistic level is only obtained if the main parameters, such as the infection rate (or reproduction number $\mathscr{R}_0 = \beta/\gamma$) are changed during the analysis. Sophisticated approaches have been developed to esti-

W. J. T. Bos, J.-P. Bertoglio and L. Gostiaux
Univ Lyon, École Centrale de Lyon, INSA Lyon, Université Claude Bernard Lyon 1, CNRS, Laboratoire de Mécanique des Fluides et d'Acoustique, UMR 5509, 36 Avenue Guy de Collongue, F-69134, ECULLY, France

mate the temporal evolution of β by assuming for instance a prescribed shape of the time-evolution of the reproduction number [15] or by adjusting the value of β instantaneaously to the observations [8]. We show that even using constant coefficients, saturation of the epidemic is obtained when the local nature of interactions within clusters is taken into account.

Indeed, perhaps the largest known flaw of the SIR model is the assumption of well mixedness. By assimilating all susceptible individuals into one single compartment, it is tacitly assumed that all individuals are in contact with all others, and the spatial structure of the epidemic is thereby ignored. An approach to overcome this weakness is to use either networks, self-propelled interacting agents, nested or coupled local SIR-type models [18] or diffusion equations [19, 14, 16]. Such modifications change the level of complexity of the model considerably requiring either the solution of a two-dimensional reaction-diffusion problem, tracking a large number of propelled particles, or the solution of integro-differential equations. We refer to Rahmandad and Sterman [21] for a discussion of the differences between agent-based and compartmental models, and to Zohdi [26] for an agent-based model of epidemic spreading applied to a planetary domain.

The model we introduce in the present work is much simpler and uses a two-scale approach: either individuals are far enough away from the individuals not to be in contact, or they pertain to a cluster of individuals in contact with the infected. This approach allows to take into account spatial structuring at the lowest level of complexity, i.e., by adding one single compartment. We sub-divide the group of susceptibles into B, those in a so-called "blob", who are or have been in contact with the infected, and S, those who are (spatially) distant from the infected. We will show, by comparison with data for a variety of European countries, that this model reproduces the exponential decay of the number of newly infected after its peak-value, with an exponent determined by the spatial diffusion coefficient of the blob.

An important simplification of reality which we will assume throughout this investigation is the constancy of the model-coefficients. All countries introduced different measures which change the effective transmission rate by physical effects such as quarantine-measures [5], or by changing the awareness of the individuals [25]. The present investigation aims primarily at the assessment of a novel approach to take into account inhomogeneity. The lock-down period for a number of similar occidental countries seems the correct starting point, since despite differences in detailed measures, the parameters governing the dynamics of the epidemic are expected to become constant during lock-down. The universal features captured by the model can therefore be disentangled from other effects, and the results can be compared to the data within the limits of the quality of the measurements.

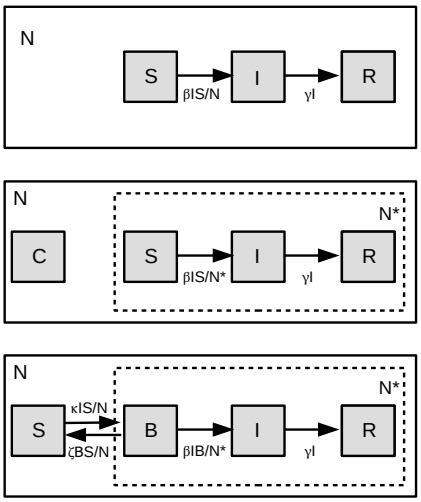


Fig. 1 Diagrammatic representation of three compartmental models, SIR, CSIR and SBIR. In SIR and CSIR three compartments interact, with the only difference in CSIR that those three compartments do not contain the entire population, since part of the population is supposed to be protected by cross-immunity. In the SBIR model, the susceptible part of the population is subdivided into individuals B inside the blob, and S, outside the blob, and the four compartments interact.

In the following we will first present the SBIR model and its difference with the classical SIR approach (section 2). We will analytically determine the exponential decay time-scale for the SIR model with and without cross-immunity, as well as for the SBIR model. We show, by comparison with data from the Johns Hopkins University data-base [10], that the SBIR model allows to simultaneously fit the data in the increasing and decreasing phase of the epidemic without the need of adjusting the model-parameters during the epidemic (section 3). We further estimate the reproduction number and show that it reaches a stable minimal value during lock-down. The appendices contain the mathematical procedure to analytically estimate the decay exponents of the different models and the initial and final reproduction numbers.

2 A cluster-based compartmental model

Spatially resolved epidemiological models are elegant and physically better justified than the SIR model but have the disadvantage that they are harder to analyze analytically, which makes insights on the effect of the different parameters more complicated. Indeed the complexity increases in general the descriptive capacities of a model, but decreases the insights which can be obtained.

We propose here a new model, that does only introduce one new compartment, representing the individuals who are or have been spatially close to the infected people. Thereby our model remains based on global parameters and does

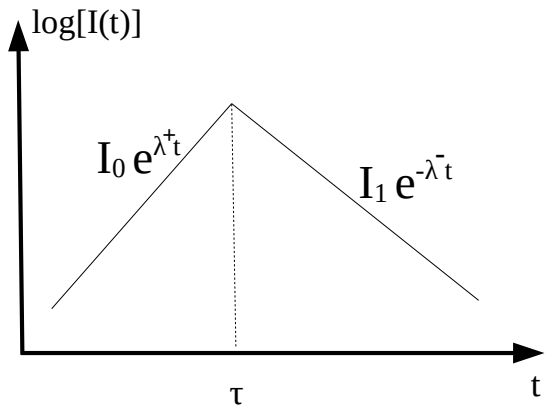


Fig. 2 Schematic representation of the evolution of the number of infected individuals in SIR and SBIR under the assumption of constant model-parameters. Time τ represents the peak-value of the number of infected.

not introduce the need to solve a two-dimensional diffusion problem. This simplicity allows the analytical derivation of the long-time behavior and thereby elucidates the role of local cluster-saturation on global decay-rates.

We will in this section first recall the SIR model, give an analytical estimate of the decay-exponent in the final phase of the epidemic, and show the influence of cross-immunity in a simple CSIR model. After that we introduce the SBIR model and discuss its physical and some of its mathematical properties.

2.1 The decay exponent in SIR and cross-immunity

The simplest model for macroscopic modeling of an epidemy is the so-called SIR model [13, 9]. The model describes a population divided in three categories, Susceptible S, Infected I and Removed or Recovered R (Fig. 1, top). The total population $N = S+I+R$ is supposed constant on the timescale of the epidemic. The evolution of, and the interaction between the three compartments is given by a set of three coupled ordinary differential equations,

$$\begin{aligned}\frac{dS}{dt} &= -\beta\frac{IS}{N}\\ \frac{dI}{dt} &= \beta\frac{IS}{N} - \gamma I\\ \frac{dR}{dt} &= \gamma I.\end{aligned} \tag{1}$$

In the following it is convenient to normalize these equations by N, so that R,S,I all take values between one and zero and $S+R+I = N = 1$. Initially $R(0) = 0$, $I(0) = I_0 \ll 1$ and $S(0) = 1-I(0) \approx 1$. The dynamics of the SIR model are characterized by an exponential growth with exponent $\lambda^+ = \beta - \gamma$, upto a time τ where saturation of the infected is obtained, by collective, or herd immunity, obtained when $S < \gamma/\beta$. Note that in the present epidemic of SARS-CoV-2, we remain very far below such collective immunity at the time of the first peak. Long enough after this $t = \tau$, exponential decay is observed until all infected have disappeared. The behavior is sketched in Fig. 2. We show in appendix A a simple and straightforward way to determine the approximate exponential growth and decay rate of the epidemic. This analysis, assuming continuity of two exponentials (as in Fig. 2), yields the decay-exponent

$$\lambda^- = \frac{\gamma\lambda^+}{\beta} \equiv \frac{\gamma\lambda^+}{\lambda^+ + \gamma}, \tag{2}$$

which gives a direct relation between the exponential during the initial, growing phase of the epidemic, and the decaying phase. We note that an exact analytical solution of the system exists [11], but the extension of the method to more complicated systems might not be straightforward. The present approach yields less accurate but readily interpretable results, allowing direct insights in the effect of the parameters on the evolution of the epidemic.

Furthermore, the peak-value of the number of infections is given by

$$I(\tau) = \left(\frac{\lambda^+}{\beta}\right)^2, \tag{3}$$

and at this time τ, the rate of infection $P = \beta SI$ is close to its maximum and is given by (see appendix A)

$$P(\tau) = \gamma\left(\frac{\lambda^+}{\beta}\right)^2. \tag{4}$$

An important question is whether cross-immunity could explain the saturation of an epidemic such as the ongoing one at the observed levels. We consider thereto a population where a part of the individuals has attained cross-immunity - the precise source of it, past contact with similar viruses or any other hidden mechanism being beyond the scope of this study. This part of the individuals is called C (Fig. 1, center). To obtain the CSIR equations, the SIR-equations are now only changed on the level of the global community size, where we replace

$$N = C+S+I+R = 1 \rightarrow N^* = S+I+R = 1-C. \tag{5}$$

The growth-exponent λ^+ is not changed by this modification for given γ,β, but the peak-values of I and P are multiplied by a factor $(1-C)$ as explained in appendix B. For instance

$$P_{\mathrm{CSIR}}(\tau) = \gamma(1-C)\left(\frac{\lambda^+}{\beta}\right)^2 = (1-C)P_{\mathrm{SIR}}(\tau). \tag{6}$$

In the long-time however, the decay-exponent λ^- is the same for the SIR and the CSIR model.

2.2 SBIR: SIR in a blob

We consider a cluster-based compartmental model, where only a sub-ensemble of individuals $N^* < N$ will interact with infected individuals. Via social interaction, a set of clusters diffusing around infected individuals will emerge, this diffusion process being distinct from virus transmission, and having its own dynamics. We define the number of persons belonging to one of these diffusing clusters as $N^* = B+I+R \equiv 1-S$, where I and R have the same interpretation as before, i.e., infected and recovered/removed, respectively, and B is referred to as the blob, namely the people who joined a diffusing cluster through social interaction and who are not yet infected or recovered. These individuals can be seen as local susceptibles, and in the following the B, I and R compartments are considered globally, independently of the number of clusters they correspond to. The persons in S are susceptible individuals, that need first to join the blob before being possibly infected. The 4 compartments, as illustrated in Fig. 1, are now related by the following system of ordinary differential equations,

$$\frac{dS}{dt} = -\kappa\frac{SI}{N} + \zeta\frac{SB}{N} \tag{7}$$

$$\frac{dB}{dt} = \kappa\frac{SI}{N} - \zeta\frac{SB}{N} - \beta\frac{BI}{N^*} \tag{8}$$

$$\frac{dI}{dt} = \beta\frac{BI}{N^*} - \gamma I \tag{9}$$

$$\frac{dR}{dt} = \gamma I. \tag{10}$$

If the growth exponent $\lambda^+ = \beta - \gamma$ has the same expression as in the SIR model, the final decay exponent, if $\beta > \kappa$, is given by (see appendix C),

$$\lambda^- = \gamma - \kappa \tag{11}$$

which is different from the expressions for SIR and CSIR, and now depends on the spatial diffusion-coefficient of the blob related to κ.

In the limit of very fast diffusion, B will approach S, and the SIR-dynamics is observed subsequently. Furthermore, mathematical robustness of the formulation is ensured, since none of the quantities S, B, I, R can become negative.

2.3 The physics behind SBIR

Before integrating the SBIR model, it is insightful to show what the physical picture behind the model is. This will help to interpret the results and to guide the choice of the parameters. The parameters of the model are γ, β, κ and ζ. Whereas γ is an intrinsic parameter characterizing the disease, β, κ and ζ depend on physical, biological and psychological influences. The β parameter, representing the infection rate is already present in the classical SIR model. It will be obtained here from an estimation of λ^+ in the early stage of the epidemics, which is known to be barely affected by the spatial structure of the population [24].

In order to understand the dynamics of the blob, let us consider the coupling between the S and B compartments, and first neglect the term proportional to ζ. We have then

$$\frac{dN^*}{dt} = \kappa\frac{SI}{N}, \tag{12}$$

If we assume $S/N \approx 1$ (as expected in early stages) and that I varies slowly, we have

$$N^* \approx \kappa I t. \tag{13}$$

When a constant population density σ_P is introduced, the surface of the blob N^*/σ_P corresponds to the increase of a surface explored by a random walk. Indeed, the separation of a random walker from its initial position is

$$r(t) \equiv \sqrt{\langle(\boldsymbol{x}(t) - \boldsymbol{x}(0))^2\rangle} \sim \sqrt{Dt}, \tag{14}$$

where D is the diffusion coefficient, and the surface explored is proportional to the square of this length. The interaction between the two compartments is therefore governed by a Brownian-like diffusion process, with a diffusion coefficient $D = \kappa I/\sigma_P$. In most epidemics, I rapidly grows so that the time-varying effective diffusion coefficient increases as well, and super-diffusive growth of the blob is observed.

We can ask how this diffusion is related to the movement of individuals. In the kinetic theory of gases, the macroscopic diffusion coefficient is related to microscopic motion by

$$D_m \sim ul, \tag{15}$$

where in this expression u is the velocity magnitude of the gas-molecules and l the mean-free-path. In our description, where the blob diffuses with a typical timescale κ^{-1}, the effective mean free path of individuals in contact with infected is proportional to the number of infected I, assuming that their velocity, which is related to the displacement of individuals, is constant (and reduced) during lock-down. Even though this is obviously a very rough approximation, we could use this idea to estimate the change of diffusion after lock-down using smartphone mobility data (as for instance [1, 2]). In the present investigation we will use constant model-coefficients and show that this approximation, during confinement, gives good agreement with data. The influence of non-constant coefficients is left for future investigation.

We still have the parameter ζ to discuss. The evolution equation of S can be written as

$$\frac{dS}{dt} = -\kappa\frac{SI}{N}\left(1 - \Gamma^{-1}\frac{B}{I}\right), \tag{16}$$

where

$$\Gamma = \kappa/\zeta. \tag{17}$$

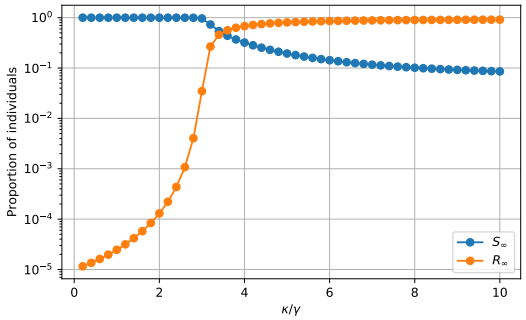


Fig. 3 Influence of the diffusion rate κ on the final number of susceptible individuals $S(\infty) = 1 - R(\infty)$ and recovered or removed individuals $R(\infty)$ for given constant coefficients.

The ratio B/I indicates how many individuals in the cluster are typically interacting with one infected. Indeed, if we set ζ to zero, there is no bound on the blob-size, and a blob of one-million individuals in contact with a single infected case would not make sense physically. This is precisely the flaw in the SIR model (the well-mixedness hypothesis) that we want to improve upon in the SBIR model. The combination of the κ and ζ term will allow the system to relax to an equilibrium value Γ of the typical infection density per cluster. In the present investigation we will use Γ not only to set the ratio of κ and ζ, but also to set the initial condition for the ratio B/I.

2.4 Dependence of epidemic spreading on the diffusion rate

We have thus given physical interpretations of the model parameters κ and ζ. Before showing how the SBIR model compares to realistic data for constant coefficients, we will determine the long-time evolution of the system for well defined initial conditions. The values we used are $\beta = 3\gamma$, $\zeta = \kappa/10$, $\gamma = 0.1$, $B(0) = R(0) = 0$, $I(0) = 1 \cdot 10^{-5}$. The temporal evolution of the different quantities is documented in Appendix D. We have observed that at very long times, the only stable solution of the system is $I = 0, B = 0, S = 1 - R$. The final state of the population is shown in Fig. 3, where the values of $S(\infty)$ and $R(\infty)$ are plotted as a function of κ/γ. Values of $R(\infty)$ can range from $R(\infty) = I(0)$ for $\kappa = 0$ to $R(\infty) = \mathcal{O}(1)$ for $\kappa \gg 1$, where the SIR limit is attained. It is clear from this plot that the final state of the SBIR-population is very importantly determined by the diffusion rate. This allows to model the evolution of the COVID-19 epidemic during its first months, without the need to change the model coefficients during a simulation.

3 Comparison with data on confirmed infections

In this section we compare the results of our model to recent data on the epidemic. The number of daily confirmed new infected cases is obtained from the Johns Hopkins University data-base [10]. This quantity is related to the source term in equation (9),

$$P_{\mathrm{SBIR}} \Delta t = \beta \frac{BI}{N^*} \Delta t. \tag{18}$$

With $N^* = B + I + R$ and $\Delta t = 1$ day. We have chosen to compare to several European countries where the lock-down is applied and Sweden, a country who has only applied "mild" restrictions on its population.

Our procedure is the following. We use for the parameter $\gamma = 0.1$ and determine in the data the time $t = 0$ where the prevalence of daily confirmed new infected cases exceeds $P(0) = 1 \cdot 10^{-6}$. The value γ^{-1} determines the typical time-interval during which an individual is infectious and can thereby transmit the disease to other individuals. The exact length of this time interval can vary between 3 and 18 days [15], and our choice of 10 days is in agreement with recent investigations [4, 8, 23]. The value $\gamma = 0.1$ corresponds to the case where from a group of isolated infected individuals, 50% will recover after approximately 7 days. Choosing a different value for this parameter would not dramatically change the results, but would directly change the choice of β used in the remainder. This latter parameter is determined as follows. During the early growth phase of the epidemic, the data is fitted by an exponential, thereby determining $\lambda^+ = \beta - \gamma$, and thus also β. We use $I(0) = P(0)/\beta$ and the value of $B(0)$ is adjusted to match the height of the peak $P(\tau)$. The value of the cluster diffusion κ is varied to get a best fit on the decay phase. We note that in the phase of the epidemic we focus on, the influence of the parameter ζ on the results is weak compared to the influence of the choice of κ. It is chosen so that the initial ratio $B(0)/I(0)$ is equal to κ/ζ. Finally, we impose $R(0) = 0$.

The results of the integration of the SBIR model are compared in Fig. 4 to the data, and the parameters are indicated on each plot. The general agreement is rather encouraging. It is observed that for all countries the value of β is of order 0.3. The value of the diffusivity varies within the subset of considered countries in the range $\kappa \in [4.5 \cdot 10^{-2}, 7.5 \cdot 10^{-2}]$ and the ζ parameter is typically 2 to 3 orders of magnitude smaller. For all countries a revival of the epidemic starts some time after the lock-down period, which is not reproduced by the model in its present form, and the reason for this is discussed in the conclusion. The lock-down period is indicated in Fig. 4 by a shaded area. We have used an objective definition for the dates of this lock-down period using mobility data from Google [1], given in appendix E.

Also shown in the figure for comparison are the CSIR results. We have again fitted the initial growth exponent,

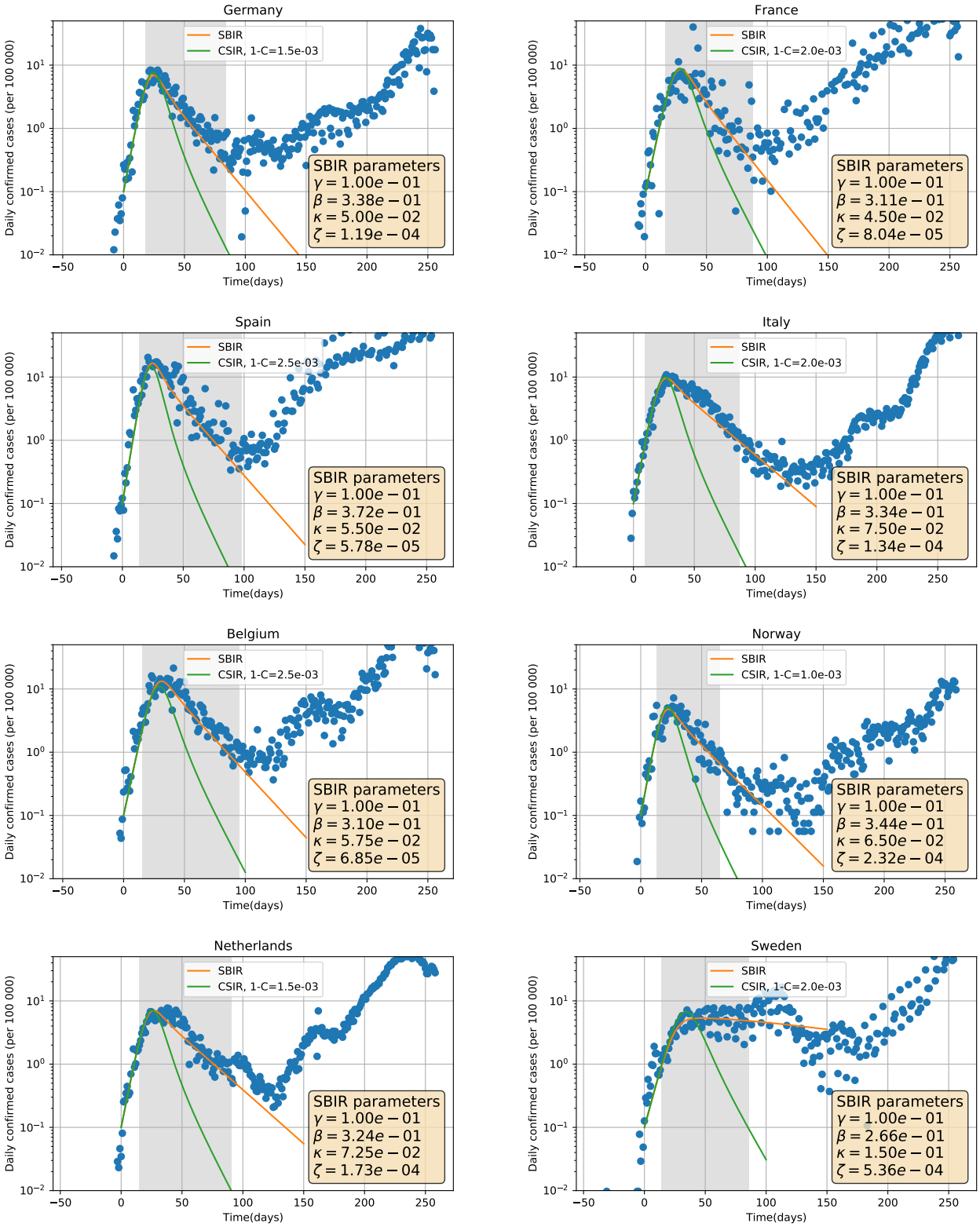


Fig. 4 Comparison of SBIR and CSIR results, with data from JHU [10] for 8 European countries. The data represent the reported daily new infections. The shaded region corresponds to the lock-down period in the different countries.

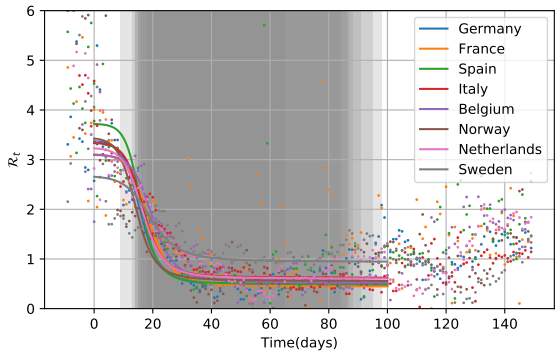


Fig. 5 Comparison of the SBIR results for $\mathscr{R}_t$ for 8 European countries. The shaded region is the superposition in transparency of the lock-down periods in the considered countries.

and we have adjusted the value of the cross-immunity C to reproduce the peak of infection. To obtain a correct fit we need to use values for the part C of individuals having cross-immunity in the interval $C \in [0.997, 0.999]$. The cross-immunity needed to attain a saturation of the number of newly infected is therefore unrealistic. Furthermore, the decay exponent λ^- in the CSIR model $\lambda^-_{\mathrm{SIR}} \approx 0.07$ is too large, which largely underestimates the duration of the epidemic. The SBIR model, with the introduction of the diffusivity parameter κ, gives fairly good agreement in the decay regime with $\lambda^-_{\mathrm{SBIR}} \approx 0.04$. This ability to correctly represent and explain the fast increase and the slow decrease of the epidemic is the major advantage of the SBIR model.

We also compare fhe SBIR model to a country where lock-down restrictions were very light, Sweden. It is observed that the model also describes the behaviour of the data in this country, but with a value of κ more than twice larger than in countries with severe lock-down restrictions. The spatial diffusion of the blob in which the ensemble of cluster evolves is in Sweden thus, according to our model, twice as large as in the other considered countries.

4 Estimating the reproduction number

The evolution of the number of new infected individuals can also be quantified by the reproduction number $\mathscr{R}_t$. We recall that the initial value of the reproduction number within the SIR or SBIR models is defined by $\mathscr{R}_0 = \beta/\gamma$.

We can rewrite the I-equation of the SBIR model [Eq. (9)] as

$$\frac{1}{I}\frac{dI}{dt} = \gamma(\mathscr{R}_t - 1) \tag{19}$$

where the instantaneous reproduction number is given by the expression

$$\mathscr{R}_t = \frac{\beta}{\gamma}\frac{B(t)}{N^*(t)}. \tag{20}$$

Since both B and N^* evolve in time, this quantity will also vary. Initially, since $R, I \ll B$, we have $\mathscr{R}_t \approx \mathscr{R}_0$.

In Fig. 5, we show the evolution of $\mathscr{R}_t$ obtained from the SBIR-integration. To compare with data we use the local logarithmic derivative of the newly infected cases P to estimate $\mathscr{R}_t$

$$\mathscr{R}_t = 1 + \frac{1}{\gamma P}\frac{dP}{dt}. \tag{21}$$

The use of P instead of I in Eq. (19) is justified as long as B/N^* is a slowly varying quantity. In order to reduce the noise in the value of the derivative, data is preprocessed using a two-week moving average.

We see that the values of $\mathscr{R}_t$ from SBIR do agree relatively well with the data. Around $t = 0$ the scatter is very large, which is in part caused by the quality of the data which depends among others on the number of tests. During lockdown, the value of $\mathscr{R}_t$ drops from its initial value which varies for the considered countries around $3-4$, to a value around 0.5. The rapid decay of $\mathscr{R}_t$ is present in both the data and the model estimates. The shape of the temporal evolution is close to a hyperbolic tan shape, as used in a recent modeling investigation of the COVID-19 epidemic [15]. However, there seems to be a time-delay between the two values. This might stem from the modeling approach, as no delay is introduced here between exposition and infection. The SBIR model, which is a direct improvement of the SIR model could possibly reproduce this when an E-compartment is added as in classical SEIR models.

Finally, the relation between the value of the decay exponent λ^- observed during lockdown in Fig. 4 and $\mathscr{R}$ can be derived when a pure exponential is assumed,

$$\lambda^- = \gamma(1 - \mathscr{R}_t). \tag{22}$$

Therefore, using Eq. (11), we have for SBIR when the blob-size is determined by the diffusion of B,

$$\mathscr{R}_t = \frac{\kappa}{\gamma} \equiv \mathscr{R}_\kappa. \tag{23}$$

We could call $\mathscr{R}_\kappa$ the macroscopic or global reproduction number, whereas the classical quantity $\mathscr{R}_0 = \beta/\gamma$ is the microscopic or local reproduction number. When $\mathscr{R}_0 > \mathscr{R}_\kappa$, the number of infected individuals in the blob will increase faster than new susceptibles can enter the blob, so that diffusion will sooner or later become the limiting factor in the spreading of the epidemic. During lock-down it is this factor which will determine the decay of the number of new infections.

5 Discussion

We have shown that the introduction of an additional compartment in the SIR model allows to take into account the presence of clusters in the evolution of an epidemic. We have

shown that the SBIR model reproduces the decay-exponent of new infections during lockdown, without the need to vary in time the model parameters over a duration exceeding 100 days.

The results in the present paper have to be understood as an illustration of the potential of the approach. A more comprehensive study should take into account the increase in the number of tests and its consequences on the number of new infected cases [22]. Also, we have deliberately chosen in the present work to consider the simplest model, SIR, and to add the effect of clustering upon this model. This allows to assess most easily the refinement. However, the nature of the present subdivision of compartments is so simple that most existing compartmental models (such as the SEIR model), can be modified in a similar manner. Effects such as vaccination, a finite incubation period, structuring of the compartments by age, and incomplete immunity could be taken into account in this manner.

The SBIR model shows that the close-to-universal decay exponent of the number of newly infected during lockdown, as observed in the data, can be linked to the cluster-diffusion. We also show that, during lockdown, a highly infectuous disease can be characterized by a global reproduction number $\mathscr{R}_\kappa = \kappa/\gamma$, which overrides the classical local reproduction number $\mathscr{R}_0 = \beta/\gamma$.

We also show that the SBIR model in its present form reproduces observations in countries during lock-down, but no longer reproduces observations long time after restrictions are weakened. We trace back the origin of this different behavior to two reasons. The main reason is that we have assumed a constant diffusion coefficient κ during lock-down and we have not increased this coefficient at the end of the lock-down period. Using realistic data on mobility from mobile network providers (as used here to determine the period of lockdown) to modulate κ, would allow the blob to increase more rapidly after lockdown, and to allow the epidemic to rapidly generate new clusters. Another reason is the fact that in the SBIR model, as formulated here, all the recovered individuals R are counted in the blob-population N^*. This is plausible in countries during lock-down, where clusters are fixed in space, hence will saturate and suffocate when R increases. In countries where the clusters not only expand, but also relocate, the epicenter of the epidemic will move spatially and leave behind regions where a large part of the population has been touched by the epidemic, but where the disease is less active. The way in which we can model these two distinct phenomena will be discussed and assessed in a future investigation.

Data Availability Statements

The data that support the findings of this study are openly available from the Johns Hopkins University data-base [10]. Mobility data used in appendix E are openly available from [1].

Acknowledgments

The authors thank Samuel Alizon (MIGEVEC - CNRS) and Loïc Méès (LMFA - CNRS) for stimulating discussions, as well as Benjamin Pillot (UMR Espace-Dev 228) for developing the Pycovid-19 package [20] used to access data of the JHU-database and for useful advices. We also thank Pascal Roy (LBBE UMR 5558) and Vitaly Volpert (ICJ UMR 5208) for their helpful advices.

References

1. (2020) COVID19 mobility reports Google. https://www.google.com/covid19/mobility, accessed: 2020-07-03
2. (2020) COVID19 mobility trends Apple Maps. https://www.apple.com/covid19/mobility, accessed: 2020-07-03
3. Bacaër N (2020) Un modèle mathématique des débuts de l'épidémie de coronavirus en france. Mathematical Modelling of Natural Phenomena 15:29
4. Bar-On YM, Sender R, Flamholz AI, Phillips R, Milo R (2020) A quantitative compendium of covid-19 epidemiology. arXiv preprint arXiv:200601283
5. Berger DW, Herkenhoff KF, Mongey S (2020) An seir infectious disease model with testing and conditional quarantine. Tech. rep., National Bureau of Economic Research
6. Brauer F, Castillo-Chavez C, Castillo-Chavez C (2012) Mathematical models in population biology and epidemiology, vol 2. Springer
7. Capasso V, Serio G (1978) A generalization of the kermack-mckendrick deterministic epidemic model. Mathematical Biosciences 42(1-2):43–61
8. Derrode S, Gauchon R, Ponthus N, Rigotti C, Pothier C, Volpert V, Loisel S, Bertoglio JP, Roy P (2020) Piecewise estimation of R0 by a simple SEIR model. Application to COVID-19 in French regions and departments until June 30, 2020, URL https://hal.archives-ouvertes.fr/hal-02910202
9. Diekmann O, Heesterbeek JAP, Metz JA (1995) The legacy of kermack and mckendrick. Publications of the Newton Institute 5:95–115
10. Dong E, Du H, Gardner L (2020) An interactive web-based dashboard to track covid-19 in real time. The Lancet infectious diseases 20(5):533–534
11. Harko T, Lobo FS, Mak M (2014) Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death

and birth rates. Applied Mathematics and Computation 236:184–194
12. Hethcote HW (2000) The mathematics of infectious diseases. SIAM review 42(4):599–653
13. Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proceedings of the royal society of london Series A, Containing papers of a mathematical and physical character 115(772):700–721
14. Kuperman M, Abramson G (2001) Small world effect in an epidemiological model. Physical Review Letters 86(13):2909
15. Linka K, Peirlinck M, Kuhl E (2020) The reproduction number of covid-19 and its correlation with public health interventions. Comput Mech 66:1035
16. Martcheva M (2015) An introduction to mathematical epidemiology, vol 61. Springer
17. Massonis G, Banga JR, Villaverde AF (2020) Structural identifiability and observability of compartmental models of the covid-19 pandemic. arXiv preprint arXiv:200614295
18. Mideo N, Alizon S, Day T (2008) Linking within-and between-host dynamics in the evolutionary epidemiology of infectious diseases. Trends in ecology & evolution 23(9):511–517
19. Murray JD, Stanley EA, Brown DL (1986) On the spatial spread of rabies among foxes. Proceedings of the Royal society of London Series B Biological sciences 229(1255):111–150
20. Pillot B, Gostiaux L (2020) Pycovid-19 python library. https://framagit.org/benjaminpillot/covid-19
21. Rahmandad H, Sterman J (2008) Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. Management Science 54(5):998–1014
22. Sauter T, Pires Pacheco M (2020) Testing informed sir based epidemiological model for covid-19 in luxembourg, DOI 10.1101/2020.07.21.20159046
23. Tian H, Liu Y, Li Y, Wu CH, Chen B, Kraemer MU, Li B, Cai J, Xu B, Yang Q, et al. (2020) An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china. Science 368(6491):638–642
24. Trapman P, Ball F, Dhersin JS, Tran VC, Wallinga J, Britton T (2016) Inferring R_0 in emerging epidemics—the effect of common population structure is small. Journal of The Royal Society Interface 13(121):20160288, DOI 10.1098/rsif.2016.0288
25. Wu Q, Fu X, Small M, Xu XJ (2012) The impact of awareness on epidemic spreading in networks. Chaos: an interdisciplinary journal of nonlinear science 22(1):013101
26. Zohdi T (2020) An agent-based computational framework for simulation of global pandemic and social response on planet x. Computational Mechanics 66(5):1195–1209

A Analytical determination of the decay exponent in SIR

The decay exponent in the SIR model can be approximated analytically. We define $t = \tau$ the time where the number of infected I peaks ($dI/dt = 0$). We observe exponential increase and decay before and after this time, respectively (see Fig. 2),

$$I(t \ll \tau) = I_0 \exp(\lambda^+ t), \quad I(t \gg \tau) = I_1 \exp(-\lambda^- t). \tag{24}$$

For short times $S \approx 1$ and for very long times, $I \ll R, S$. Therefore

$$\lambda^+ = \beta - \gamma, \tag{25}$$
$$-\lambda^- = \beta S_\infty - \gamma \tag{26}$$
$$= \lambda^+ - \beta R_\infty. \tag{27}$$

In order to determine λ^-, we need therefore the total number of recovered, R_∞, given by

$$R_\infty = \gamma \int_0^\infty I(t) dt. \tag{28}$$

Since we assume exponential growth for short time, and exponential decay for long times, respectively, continuity at $t = \tau$ allows to write,

$$R_\infty = \gamma \int_0^\tau I(t) dt + \gamma \int_\tau^\infty I(t) dt \tag{29}$$
$$= \gamma \left(\frac{1}{\lambda^+} + \frac{1}{\lambda^-} \right) I(\tau), \tag{30}$$

which expresses R_∞ as a function of a new unknown, $I(\tau)$, the peak-value of I. Considering the I-equation,

$$d_t I = (S\beta - \gamma) I, \tag{31}$$

the term in brackets vanishes for $S = \gamma/\beta$. We know that $S = 1 - I - R$, we also have that in the initial exponential phase

$$R = \frac{\gamma}{\lambda^+} I \tag{32}$$

and therefore

$$\gamma/\beta = 1 - (1 + \gamma/\lambda^+) I(\tau) \tag{33}$$
$$= 1 - (\beta/\lambda^+) I(\tau). \tag{34}$$

so that

$$I(\tau) = \left(\frac{\lambda^+}{\beta} \right)^2 \tag{35}$$

Which yields an expression for $I(\tau)$. We have the relations (27), (30), (34) for the unknown quantities, I_τ, R_∞ and λ^-, and solving this system yields

$$\lambda^- = \frac{\gamma \lambda^+}{\beta} = \frac{\gamma(\beta - \gamma)}{\beta}. \tag{36}$$

Note that for typical values, for instance $\gamma = 0.2$; $\beta = 0.4$, this yields $\lambda^- = 0.1$. Numerical integration yields $\lambda^- = 0.119$. This error of 20% is due to the assumptions of the two exponentials near the peak, where the rounding of the integrals significantly contributes to the error.

The quantity that we will consider in the comparison with data is the daily confirmed new infected $P\Delta t$, which is associated with the source term in the I-equation,

$$P = \beta S I, \tag{37}$$

where in our system, where the time-unity is days, $\Delta t = 1$.

Interesting relations, when substituting this value for λ^- in the equations are

$$I(\tau) = \left(\frac{\lambda^+}{\beta}\right)^2, \quad P(\tau) = \gamma\left(\frac{\lambda^+}{\beta}\right)^2 \tag{38}$$

$$S(\tau) = \frac{\gamma}{\beta}, \quad R(\tau) = \frac{\lambda^+}{\beta}\left(1-\frac{\lambda^+}{\beta}\right), \tag{39}$$

$$R_\infty = 1-\frac{\gamma^2}{\beta^2}, \quad S_\infty = \frac{\gamma^2}{\beta^2}. \tag{40}$$

B Exponent in SIR with Cross immunity (CSIR)

Let us now consider the case of a partially immune population,

$$1 = C+S+I+R, \tag{41}$$

where $I(0)/S(0)$ and $R(0)/S(0)$ are small (as in the previous case), so that $S(0) \approx 1-C$. The I-equation reads

$$\frac{dI}{dt} = \beta\frac{IS}{1-C} - \gamma I. \tag{42}$$

For short times, we have thus, since $S \approx 1-C$,

$$\lambda^+ = \beta - \gamma, \tag{43}$$

as for SIR. The growth of I halts at time τ when $dI/dt = 0$, so that we find for $S(\tau)$

$$S(\tau) = (1-C)\frac{\gamma}{\beta}. \tag{44}$$

We find then for the different quantities, following the same reasoning as in the previous section,

$$I(\tau) = (1-C)\left(\frac{\lambda^+}{\beta}\right)^2, \quad P(\tau) = \gamma(1-C)\left(\frac{\lambda^+}{\beta}\right)^2 \tag{45}$$

and for the exponential decay-rate,

$$\lambda^- = \frac{\gamma\lambda^+}{\beta}, \tag{46}$$

as for *SIR*.

What we learn from this is that, cross-immunity or not, if we measure λ^+, initially, λ^- is not influenced for a given γ. Only the height of the peak of the number of infected I, and of the number of daily confirmed cases P will be multiplied by $1-C$.

C Exponent in SBIR

Let us now consider the SBIR model. The I equation of SBIR reads,

$$\frac{dI}{dt} = \left(\beta\frac{B}{B+I+R} - \gamma\right)I \tag{47}$$

Let us consider the situation where the number of infected individuals grows faster that the number of individuals in the blob. The growth of the number of infections is then bounded by the source-term in the blob-evolution equation, and $\beta IB/N* \approx \kappa IS/N$. In this saturated phase, the I-equation becomes then,

$$\frac{dI}{dt} = \kappa\frac{IS}{N} - \gamma I. \tag{48}$$

If, as during the first wave of the COVID-19 epidemic, the global infection rate is still low and $S/N \approx 1$, it is immediately found that

$$\frac{dI}{dt} \approx (\kappa-\gamma)I, \tag{49}$$

so that exponential decay is observed with a decay exponent $\lambda^- = \gamma - \kappa$ for $\kappa < \gamma$, or a slow increase if $\beta < \kappa < \gamma$. The reproduction number is in this phase therefore determined by the spatial spreading, so that the value of $\mathscr{R}_t$ evolves from its local, microscopic value $\mathscr{R}_0 = \beta/\gamma$ to its diffusion determined value $\mathscr{R}_\kappa = \kappa/\gamma$,

$$\frac{dI}{dt} \approx \gamma(\mathscr{R}_\kappa - 1)I. \tag{50}$$

Let us evaluate when this is the case. Let us rewrite the I-equation as

$$\frac{dI}{dt} = \left(\beta\frac{1}{1+\frac{I+R}{B}} - \gamma\right)I. \tag{51}$$

We will again assume exponentials as in Fig. 2. At short times, since $B \gg I+R$, we have the same increasing exponential as in SIR and CSIR, with $\lambda^+ = \beta - \gamma$. This increase will continue as long as $I+R \ll B$. Since R is in the exponential phase proportional to I (as for SIR), we have

$$\frac{I+R}{B} \approx \left(1+\frac{\gamma}{\lambda^+}\right)\frac{I}{B}. \tag{52}$$

Let us therefore focus on the quantity I/B to see whether this quantity will be negligible for long times,

$$\frac{d(B/I)}{dt} = \frac{1}{I}\frac{dB}{dt} - \frac{B}{I^2}\frac{dI}{dt} \tag{53}$$

$$= \frac{B}{I}\left[\kappa\frac{SI}{BN} - \zeta\frac{S}{N} - \beta\frac{I}{N^*} - \beta\frac{B}{N^*} + \gamma\right]. \tag{54}$$

Considering the initial phase, where $I/B \ll 1$, and estimating the order of magnitude of the different terms, it is observed that the leading order contributions of this equation are the last two terms in brackets, yielding,

$$\frac{B}{I} \sim \exp\left[-\lambda^+ t\right]. \tag{55}$$

Within a finite time, the blob will thus contain a considerable number of infected and recovered individuals, so that there is no more room for additional infected.

D Numerical integration of the model

We present here an analysis of the behavior of the model by numerical integration of the ODEs (7)-(10) constituting the SBIR system. Since the principal difference of SBIR with respect to SIR is the diffusion from the S to the B compartment, we will assess the sensitivity of the model on a change of the diffusion coefficient. For this, we fix the parameters as follows $\beta = 3\gamma$, $\zeta = \kappa/10$, $\gamma = 0.1$, $B(0) = R(0) = 0$, $I(0) = 1\cdot 10^{-5}$ and $S(0) = 1-I(0) \approx 1$. Using these parameters we vary the value of the diffusion rate in the interval $\kappa \in [0, 10\gamma]$.

In Fig. 6 (a) we show the susceptible individuals as a function of time for different values of κ/γ. As expected, for small values of κ, S remains close to unity, whereas for large values the S-compartment is emptied almost completely. For long times the S value is slowly increasing, an effect associated with the flux of individuals from the B to the S compartment when the number of infected becomes very small. We have observed that this latter effect can be very slow for certain sets of parameters.

Fig. 6 (b) shows the temporal behavior of the individuals contained in the blob. For small values of the diffusion, B reaches rapidly a local maximum, and the number of individuals in the blob does not exceed the value 10^{-5}. The epidemic is therefore rapidly limited in size. The value of the local maximum increases when κ increases and the blob will contain almost all susceptible individuals after around 100 days for the largest values of the diffusivity. For these values, the SIR behavior is recovered.

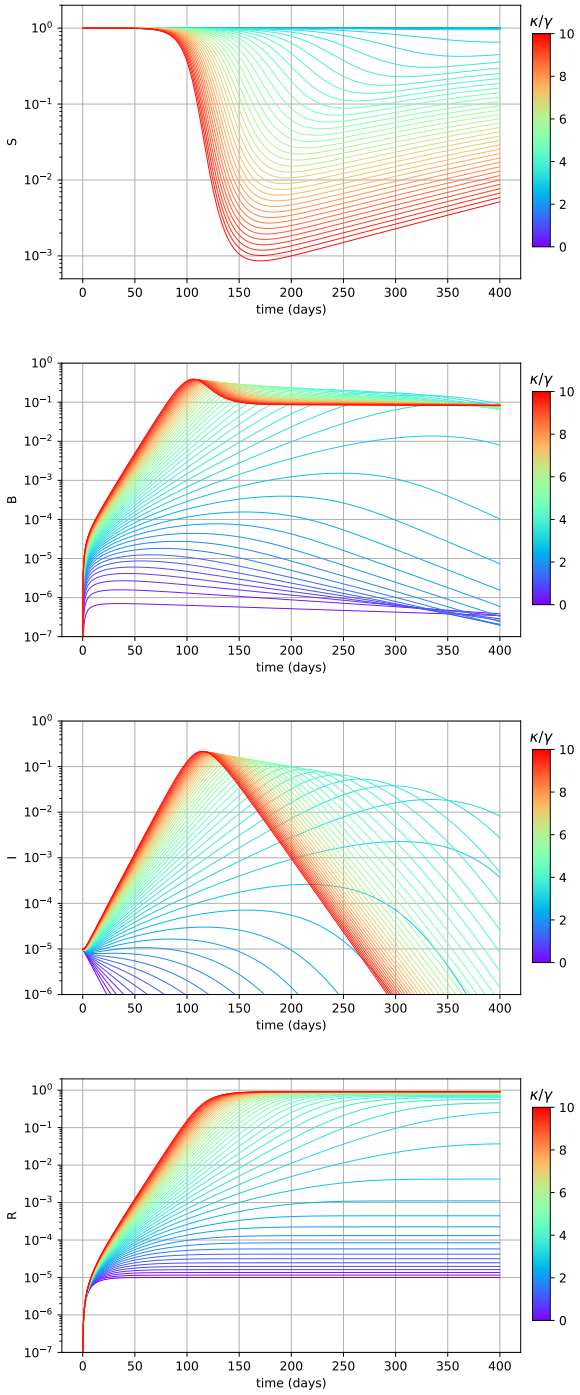


Fig. 6 Temporal evolution of (a) susceptible individuals S, (b) individuals in the blob B, (c) the number of infected individuals I, (d) recovered/removed individuals R. The different colors indicate different values of the diffusion rate in the range $\kappa \in [0, 10\gamma]$.

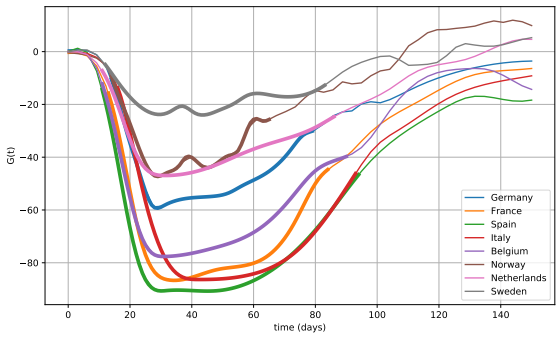


Fig. 7 Estimation of the lock-down dates from Google mobility reports. Thick lines represent the lock-down period determined for each country.

In Fig. 6 (c) we display the temporal behavior of the number of infected individuals for different values of κ. It is observed that for $\kappa < \gamma$, the number of infected decreases almost immediately: the epidemic has no time to spread, since the infected individuals remain confined locally around the source of the epidemic outbreak. On the contrary, when diffusion is strong, the epidemic will spread and rapidly all susceptible individuals will be part of the blob. The system behaves in this limit of strong diffusion as a well-mixed system, and the classic SIR-behavior is observed for the highest values of κ/γ considered here.

Finally, Fig. 6 (d) shows all recovered and removed individuals as a function of time. At the end of the evaluated time-interval the value of R is a continuously increasing function, ranging from $R \approx I(0)$ for $\kappa \downarrow 0$ to a value of order (but somewhat below) unity, corresponding to the asymptotic SIR value for $\kappa \to \infty$.

E Use of Google mobility-data to determine lock-down dates

We use the mobility data provided by Google [1] to determine the dates of the lock-down in the different countries. Fig. 7 shows the variations of mobility compared to a reference-value measured mid-january for the six countries considered in this work. The data used here corresponds to the retail and recreation data. Several other quantities are reported by Google, but most reproduce the same trend. Since we only use this data to illustrate the duration of the lockdown and do not use this to change the model-coefficients, the precise choice does not seem very important.

We have smoothened the data to get rid of high-frequency variations. We call the smoothened variable $\mathbf{G}(t)$. The general trend shows that during the epidemic the mobility drops to a value depending on the restrictions in the different countries. An approximate plateau value is observed. We take for each country the minimum value of this plateau as the reference mobility-level during lock-down, indicated by $\mathbf{G}_0$. The beginning of lock-down is then determined for the time when for the first time $\mathbf{G}(t) < 1.15\,\mathbf{G}_0$ and the end, the first later date where $\mathbf{G}(t) > 1.5\,\mathbf{G}_0$, as can be seen on Fig. 7. These begin and end dates are used to indicate the lock-down periods as a shaded region in Fig. 4.

We note that this approach could allow to modulate the model coefficients. Since this will hinder the analysis of the present approach and since the assessment of the SBIR model is the main goal of the present work, this modulation is not implemented here. However, if a second wave of infections is to be reproduced, changing the model coefficients, or the dynamics of the R-individuals will become com-

pulsory. Realistic mobility data can allow to get a handel on the way in which to modify the coefficients.