# Social contacts and the spread of infectious diseases

Giacomo Dimarco, Benoît Perthame, Giuseppe Toscani, Mattia Zanella

## HAL Id: hal-02934939
## https://hal.sorbonne-universite.fr/hal-02934939v1

Submitted on 9 Sep 2020

# Social contacts and the spread of infectious diseases

G. Dimarco[*1], B. Perthame[†2], G. Toscani[‡3], and M. Zanella[§3]

[1]Mathematics and Computer Science Department,
University of Ferrara, Italy.
[2]Sorbonne Université, CNRS, Université de Paris, Inria,
Laboratoire Jacques-Louis Lions, 75005 Paris, France.
[3] Mathematics Department, University of Pavia, Italy.

## Abstract

Motivated by the COVID-19 pandemic, we introduce a mathematical description of the impact of sociality in the spread of infectious diseases by integrating an epidemiological dynamic with a kinetic modeling of population-based contacts. The kinetic description leads to study the evolution over time of Boltzmann type equations describing the number densities of social contacts of susceptible, infected and recovered individuals, whose proportions are driven by a classical compartmental model in epidemiology. Explicit calculations show that the spread of the disease is closely related to the mean number of contacts, thus justifying the lockdown strategies assumed by governments to prevent them. Furthermore, the kinetic model allows to clarify how a selective control can be assumed to achieve a minimal lockdown strategy by only reducing individuals undergoing a very large number of daily contacts. This, in turns, could permit to maintain at best the economic activities which would seriously suffer from a total closure policy. Numerical simulations confirm the ability of the model to describe different phenomena characteristic of the rapid spread of an epidemic. A last part is dedicated to fit numerical solutions of the proposed model with experimental data coming from different European countries.

# Contents

[*]giacomo.dimarco@unife.it
[†]benoit.perthame@ljll.math.upmc.fr
[‡]giuseppe.toscani@unipv.it
[§]mattia.zanella@unipv.it

# 1 Introduction

The recent spreading of the COVID-19 epidemic led in many countries to heavy lockdown measures assumed by the governments with the aim to control and limit its effects on population. It is clear that these severe social distancing measures have the main consequence to produce dramatic damages in the real economy and society, even if their precise quantification is an extremely difficult problem which also requires the knowledge/estimation of a large number of variables and relationships among them [5]. In this direction, a recent study of the economic consequences of this new epidemic and the related lockdown measures on western economies has been done by resorting on a coupling of a kinetic model for the wealth with a SIR model for the infectious disease [15]. This model, while relying on few evident universal features, can be easily implemented to furnish answers on possible scenarios as consequence of epidemic and permitted to show that indeed lockdown measures may trigger an economic crisis leading to a drastic reduction of the middle class.

A very interesting aspect of mathematical modeling of infectious diseases is that, unlike other classical phenomenological approaches, it allows a direct comparison with real data. This consequently permits evaluation of control and prevention strategies by comparing their cost with effectiveness and support to public health decisions [18,44]. Most of the models present in literature make assumptions on transmission parameters [6,13] which are considered the only responsible of the spread of the infection. However, special attention was recently paid by the scientific community to the role and the estimate of the distribution of contacts between individuals as also a relevant cause of the potential pathogen transmission (cf. [3,16,20,38] and the references therein).

In this direction, the results reported in [3] can be of great help when designing partial lockdown strategies. In fact, and optimal limitation of the pathogen transmission of the epidemic, which limitates the number of daily contacts among people taking into account contact number heterogeneity, permits probably to reduce the severe impact on economy and society that equal distributed closure policies generate. In particular, the detailed analysis performed in [3] put into evidence that the number of social contacts in the population is in general well-fitted by a Gamma distribution, even if this distribution is not uniform with respect to age, sex and wealth. Gamma distributions belong to the wide class of generalized Gamma distributions [35,46], which have been recently connected to the statistical study of social phenomena [33,43], and fruitfully described as steady states of kinetic models in [14,47], aiming to describe the formation of these profiles in consequence of repeated elementary interactions.

Starting from the above consideration and inspired by the recent development of kinetic models for describing human behaviors [14,47], in this paper we develop a mathematical framework to connect the distribution of social contacts with the spreading of a disease in a multi agent system. This result is obtained by integrating an epidemiological dynamics given by a classical compartmental model [6,13,30] with an approach based on kinetic equations determining the formation of social contacts. In this paper, from the epidemiological point of view we concentrate on the classical so-called SIR dynamic. However, we stress that the ideas here described are clearly not reduced to this model which represent an example, instead the methodology here discussed can be easily extended to incorporate more complex epidemiological dynamics, like for instance the classical endemic models discussed in [6,13,30]. Resorting to SIR model means for instance that we abstract away from disease related mortality. Clearly, this in some situation can be a strong hypothesis, as it excludes demographic interaction. However, it is currently believed that the mortality rate of COVID-19 is likely low enough to justify this assumption. Moreover, we are particularly interested in how the epidemic spreads in a country and not on studying the consequences of this diffusion.

Thus, in this context, mortality has a low impact on the dynamics of the number of infected people and can be disregarded in first approximation. Other aspects, which certainly have a stronger impact on how a virus spreads, are related to the presence of asymptomatic individuals [25] as well as to a time delay between contacts and outbreak of the disease [10] which helps in the diffusion of the illness. These possible modeling improvements are the subject of future investigations and they will not be treated in this work. Besides this simplifying assumptions, we will show that the basic features considered and detailed in the rest of the article are sufficient in many cases to construct a model which well fits with the experimental data.

An easy way to understand epidemiology models is that, given a population composed of agents, they prescribe movements of individuals between different states based on some matching functions or laws of motion. According to the classical SIR models [30], agents in the system are split into three classes, the susceptible, who can contract the disease, the infected, who have already contracted the disease and can transmit it, and the recovered who are healed, immune or isolated. Inspired by the model considered in [14] for describing a social attitude and making use of the SIR dynamics, we present here a model composed by a system of three kinetic equations, each one describing the time evolution of the distribution of the number of contacts for the subpopulation belonging to a given epidemiological class. These three equations are further coupled by taking into account the movements of agents from one class to the other as a consequence of the infection. The interactions which describe the social contacts of individuals are based on few rules and can be easily adapted to take into account the behavior of agents in the different classes in presence of the infectious disease. Our joint framework is consequently based on two models which can be considered classical in the respective fields. From the epidemic side, the SIR and its related models are widely used in most applications (cf. [30] and the references therein). From the side of multi-agent systems in statistical mechanics, the linear kinetic model introduced in [14, 47] has been shown to be flexible and able to describe, with opportune modifications, different problems in which human behavior plays an essential role, like the formation of social contacts.

The study of the effects of the intensity of social contacts in the epidemic diffusion, among other consequences, allows to obtain in a rigorous way a variety of non-linear incidence rates of the infectious disease as for instance recently considered in [34]. It is also interesting to remark that the presence of non-linearity in the incidence rate function, and in particular, the concavity condition with respect to the number of infected has been considered in [34] as a consequence of *psychological* effects. Namely, the authors observed that in the presence of a very large number of infected, the probability for an infective to transmit the virus further may decrease because the population *tend to naturally reduce the number of contacts.* This fact will be embedded in the kinetic model producing a change in the mean value of the number of daily contacts. The importance of reducing at best the social contacts to countering the advance of a pandemic is a well-known and studied phenomenon [18]. While in normal life activity, it is commonly assumed that a large part of agents behaves in a similar way, in presence of an extraordinary situation like the one due to a pandemic, it is highly reasonable to conjecture that the social behavior of individuals is strictly affected by their personal feeling in terms of safeness. In this work, we focus on the assumption that it is the degree of diffusion of the disease that changes people's behavior in terms of social contacts, in view both of the personal perception and/or of the external government intervention. More generally, the model can be clearly extended to consider more realistic dependencies between an epidemic disease and the social contacts of individuals. However, this does not change the essential conclusions of our analysis, namely that there is a close interplay between the spread of the disease and the distribution of contacts, that the kinetic description is able to quantify. The encouraging results described in the rest of the article, finally suggest that a similar analysis can be carried out, at the price of an increasing difficulty in computations, in more complex epidemiological models like the SIDARTHE model [26, 27], recently used to simulate the COVID-19 epidemic in Italy to validate and improve the eventual partial lockdown strategies of the government and to suggest future measures.

The rest of the paper is organized as follows. Section 2 introduces the system of three SIR-type kinetic equations combining the dynamics of social contacts with the spread of infectious

disease in a system of interacting individuals. Then, in Section 3 we show that in a suitable asymptotic procedure the solution to the kinetic system tends towards the solution of a system of three SIR-type Fokker-Planck type equations with local equilibria of Gamma-type [3]. Once the system of Fokker-Planck type equations has been derived, in Section 4 we close the SIR-type system of kinetic equations around the Gamma-type equilibria to obtain a SIR model in which the presence and consequently the evolution of the social contacts leads to a non-linear incidence rate of the infectious disease satisfying the compatibility conditions introduced in [34]. Last, in Section 5, we investigate at a numerical level the relationships between the solutions of the kinetic system of Boltzmann type, its Fokker-Planck asymptotics and the SIR model. These simulations confirm the ability of the model to describe different phenomena characteristic of the trend of social contacts in situations compromised by the rapid spread of an epidemic, and the consequences of various lockdown action in its evolution. A last part is sacred to a fitting of the model with the experimental data by first estimating the parameters of the epidemic through data and successively by using these parameters in the kinetic model.

## 2 The joint action of social contacts and epidemic

Our first goal is to build a kinetic system which suitably describes the spreading of an infectious disease under a dependence of the contagiousness parameters on the individual number of social contacts of the agents. As in classical SIR models [30], the entire population is divided into three classes: susceptible, infected and recovered individuals.

Aiming to understand social contacts effects on the dynamics, we will not consider in the sequel the role of heterogeneity in the disease parameters (such as the personal susceptibility to a given disease), which could be derived from the classical SIR model, suitably adjusted to account for new information [12, 39, 49]. Consequently, agents in the system are considered indistinguishable [41]. This means that the state of a person in each class at any instant of time $t \geq 0$ is completely characterized by the number of contacts $x \geq 0$, measured in some unit.

While $x$ is a natural positive number at the individual level, without loss of generality we will consider $x$ in the rest of the paper to be a nonnegative real number, $x \in \mathbb{R}_+$, at the population level. We denote then by $f_S(x,t)$, $f_I(x,t)$ and $f_R(x,t)$, the distributions at time $t > 0$ of the number of social contacts of the population of susceptible, infected and recovered individuals, respectively. The distribution of contacts of the whole population is then recovered as the sum of the three distributions

$$f(x,t) = f_S(x,t) + f_I(x,t) + f_R(x,t).$$

As outlined in the Introduction, we do not consider for simplicity of presentation disease related mortality as well as the presence of asymptomatic individuals. Therefore, we can fix the total distribution of social contacts to be a probability density for all times $t \geq 0$

$$\int_{\mathbb{R}_+} f(x,t)\,dx = 1.$$

As a consequence, the quantities

$$J(t) = \int_{\mathbb{R}^+} f_J(x,t)\,dx, \quad J \in \{S, I, R\} \tag{1}$$

denote the fractions of susceptible, infected and recovered respectively. For a given constant $\alpha > 0$, we also denote with $m^\alpha(t)$ the moment of order $\alpha$ of the distribution of the number of contacts

$$m^\alpha(t) = \int_{\mathbb{R}^+} x^\alpha f(x,t)\,dx,$$

and the moments of order $\alpha$ for the distributions of the number of contacts for each class

$$m_J^\alpha(t) = \int_{\mathbb{R}^+} x^\alpha f_J(x,t)\,dx, \quad J \in \{S, I, R\}.$$

4

Unambiguously, we will indicate the mean values, corresponding to $\alpha = 1$, by $m(t)$ and, respectively, $m_J(t)$, $J \in \{S, I, R\}$.

In what follows, we assume that the evolution of the densities is built according to the classical SIR model [30], and that the various classes in the model could act differently in the social process constituting the contact dynamics. The kinetic model then follows combining the epidemic process with the contact dynamics, as modeled in [14, 47]. This gives the system

$$\begin{aligned} \frac{\partial f_S(x,t)}{\partial t} &= -K(f_S, f_I)(x,t) + \frac{1}{\tau} Q_S(f_S)(x,t) \\ \frac{\partial f_I(x,t)}{\partial t} &= K(f_S, f_I)(x,t) - \gamma f_I(x,t) + \frac{1}{\tau} Q_I(f_I)(x,t) \\ \frac{\partial f_R(x,t)}{\partial t} &= \gamma f_I(x,t) + \frac{1}{\tau} Q_R(f_R)(x,t) \end{aligned} \quad (2)$$

where $\gamma$ is the constant recovery rate while the transmission of the infection is governed by the function $K(f_S, f_I)$, the local incidence rate, expressed by

$$K(f_S, f_I)(x,t) = f_S(x,t) \int_{\mathbb{R}^+} \kappa(x,y) f_I(y,t) \, dy. \quad (3)$$

In (3) the contact function $\kappa(x,y)$ is a nonnegative function growing with respect to the number of contacts $y$ of the population of infected, such that $\kappa(x,0) = 0$. Then, the spreading of the epidemic depends heavily on the function $\kappa(\cdot, \cdot)$ used to quantify the rate of possible contagion in terms of the number of social contacts between the classes of susceptible and infected. The evolution of the mass fractions $J(t)$, $J \in \{S, I, R\}$ then obeys to the classical SIR model by choosing $\kappa(x,y) \equiv \beta > 0$ [30], thus considering the spreading of the disease independent of the intensity of social contacts. A leading example is obtained by choosing a rank 1 matrix

$$\kappa(x,y) = \beta \, x^\alpha y^\alpha,$$

where $\alpha, \beta$ are positive constants, that is by taking the incidence rate directly proportional to the product of the number of contacts of susceptible and infected people. When $\alpha = 1$, for example, the incidence rate takes the simpler form

$$K(f_S, f_I)(x,t) = \beta \, x f_S(x,t) m_I(t). \quad (4)$$

In equations (2) the operators $Q_J$, $J \in \{S, I, R\}$ characterize the thermalization of the distribution of social contacts in the various classes in terms of repeated interactions [14, 47]. The $Q_J$, $J \in \{S, I, R\}$ are integral operators that modify the distribution of contacts $f_J(x,t)$, $J \in \{S, I, R\}$. Their action on observable quantities is given by

$$\int_{\mathbb{R}_+} \varphi(x) \, Q_J(f_J)(x,t) \, dx = \left\langle \int_{\mathbb{R}_+} B(x) \big( \varphi(x_J^*) - \varphi(x) \big) f_J(x,t) \, dx \right\rangle.$$

where $B(x)$ measures the interaction frequency, $\langle \cdot \rangle$ denotes mathematical expectation with respect to a random quantity, and $x_J^*$ denote the post-interaction value of the number $x$ of social contacts of the $J$-th population. Last, the constant $\tau$ in front of the interaction integral measures the time scale of the thermalization of the distribution of social contacts.

**Remark 1.** *The relaxation constant $\tau$ plays an important role in the time evolution of the model (2), since it relates the time scale of the epidemic evolution with that of the statistical formation of social contacts. Small values of the constant $\tau$ will correspond to a fast adaptation of people to a steady situation. Hence, since it is reasonable to assume that adaptation of people is faster that the dynamics of epidemic, in what follows we will assume $\tau \ll 1$.*

The process of formation of the social contacts distribution is obtained by taking into account the typical aspects of human being, in particular the search, in absence of epidemics, of opportunities for socialization. In addition to that, social contacts are due to the common use of public

transportations to reach schools, offices and, in general, places of work and to basic needs of inter-actions due to work duties. As shown in [3], this leads individuals to stabilize on a characteristic number of daily contacts depending on the social habits of the different countries, (represented by a suitable value $\bar{x}_M$ of the variable $x$, which can be considered as the mean number of contacts relative to the population under investigation).

Then, following a consolidated path in the kinetic theory of social phenomena, one aims in obtaining the characterization of the distribution of social contacts in a multi-agent system (the macroscopic behavior), starting from some universal assumption about the personal behavior of agents (the microscopic behavior). Indeed, as in many other human activities, the daily amount of social contacts is the result of a repeated upgrading based on well-established rules. To this extent, it is enough to recall that the daily life of each person is based on a certain number of activities, and each activity carries a certain number of contacts. Moreover, for any given activity, the number of social contacts varies in consequence of the personal choice. The typical example is the use or not of public transportations to reach the place of work or the social attitudes which scales with the age. Clearly, independently of the personal choices or needs, the number of daily social contacts contains a certain amount of randomness, that has to be taken into account. Also, while it is very easy to reach a high number of social contacts attending crowded places for need or will, going below a given threshold is very difficult, since various contacts are forced by working or school activities, among others causes. This asymmetry between growth and decrease, as exhaustively discussed in [14, 28, 29], can be suitably modeled by resorting to the value function description of the elementary variation of the $x$ variable.

**Remark 2.** *It is important to outline that, in the presence of an epidemic, the characteristic mean number of daily contacts $\bar{x}_M$ reasonably changes in time, even in absence of an external lockdown intervention, in reason of the perception of danger linked to social contacts that people manifest. Consequently, even if not always explicitly indicated, we will assume $\bar{x}_M = \bar{x}_M(t)$. Also, this time dependent value can be different depending on the class to which agents belong.*

A further non secondary aspect of the formation of the number of social contacts is that their frequency is not uniform with respect to the values of $x$. Indeed, it is reasonable to assume that the frequency of interactions is inversely proportional to the number of contacts $x$. This relationship takes into account that it is highly probable to have at least some contacts, and the rare situation in which one reaches a very high values of contacts $x$.

The choice of a variable interaction frequency has been fruitfully applied in a different context in a different situation [23]. This was done to better describe the evolution in time of the wealth distribution in a western society. There, the frequency of the economic transactions has been proportionally related to the values of the wealth involved, to take into account the low interest of trading agents in transactions with small values of the traded wealth. As discussed in [23], the introduction of a variable kernel into the kinetic equation does not modify the shape of the equilibrium density, but it allows a better physical description of the phenomenon under study, including an exponential rate of relaxation to equilibrium for the underlying Fokker-Planck type equation.

Following [14, 28, 29, 47], we will now illustrate the mathematical formulation of the previously discussed behavior. In full generality, we assume for the moment that individuals in different classes can have a different mean number of contacts. The microscopic updates of social contacts of individuals in the class $J$, $J \in \{S, I, R\}$ will be taken of the form

$$x_J^* = x - \Phi^\epsilon(x/\bar{x}_J)x + \eta_\epsilon x. \tag{5}$$

In a single update (interaction), the number $x$ of contacts can be modified for two reasons, expressed by two terms, both proportional to the value $x$. In the first one, the coefficient $\Phi^\epsilon(\cdot)$, which can assume both positive and negative values, characterizes the typical and predictable variation of the social contacts of agents, namely the personal social behavior of agents. The second term takes into account a certain amount of unpredictability in the process. The usual choice is to assume that the random variable $\eta$ is of zero mean and bounded variance, expressed by $\langle \eta \rangle = 0$, $\langle \eta^2 \rangle = \lambda$, with

$\lambda > 0$. Small random variations of the interaction (5) will be expressed simply by multiplying $\eta$ by a small positive constant $\sqrt{\epsilon}$, with $\epsilon \ll 1$, which produces the new (small) variance $\epsilon\lambda$. Notice that this formulation makes an essential use of the choice $x \in \mathbb{R}_+$. This scaled variable is the variable $\eta_\epsilon$ appearing in (5).

The function $\Phi^\epsilon$ plays the role of the *value function* in the prospect theory of Kahneman and Twersky [31, 32], and contains the mathematical details of the expected human behavior in the phenomenon under consideration. In particular, the hypothesis on which it is built is that it is normally easier to increase the value of $x$ than to decrease it, in relationship with the mean value $\bar{x}_J$, $J \in \{S, I, R\}$. In terms of the variable $s = x/\bar{x}_J$ we consider as in [14] the class of value functions given by

$$\Phi_\delta^\epsilon(s) = \mu \frac{e^{\epsilon(s^\delta-1)/\delta} - 1}{e^{\epsilon(s^\delta-1)/\delta} + 1}, \quad s \geq 0, \tag{6}$$

where $0 < \delta \leq 1$ and $0 < \mu < 1$ are suitable constants characterizing the intensity of the individual behavior, while the constant $\epsilon > 0$ is related to the intensity of the interaction. Hence, the choice $\epsilon \ll 1$ corresponds to small variations of the mean difference $\langle x_J^* - x \rangle$ and making it common to both effects, randomness and adaptation, permits to equilibriate their effects as can be seen in Section 3. In (7), the value $\mu$ denotes the maximal amount of variation of $x$ that agents will be able to obtain in a single interaction. Note indeed that the value function $\Phi_\delta^\epsilon(s)$ is such that

$$-\mu \leq \Phi_\delta^\epsilon(s) \leq \mu.$$

Clearly, the choice $\mu < 1$ implies that, in absence of randomness, the value of $x_J^*$ remains positive if $x$ is positive. As proven in [14], the value function (6) satisfies

$$-\Phi_\delta^\epsilon(1 - s) > \Phi_\delta^\epsilon(1 + s),$$

and

$$\frac{d}{ds}\Phi_\delta^\epsilon(1 + s) < \frac{d}{ds}\Phi_\delta^\epsilon(1 - s).$$

These properties are in agreement with the expected behavior of agents, since deviations from the reference point ($s = 1$ in our case), are bigger below it than above. Letting $\delta \to 0$ in (6) allows to recover the value function

$$\Phi_0^\epsilon(s) = \mu \frac{s^\epsilon - 1}{s^\epsilon + 1}, \quad s \geq 0. \tag{7}$$

introduced in [28, 29] to describe phenomena characterized by the lognormal distribution [2, 36].

Once the *elementary interaction* (5) is given, for any choice of the value function, the study of the time-evolution of the distribution of the number $x$ of social contacts follows by resorting to kinetic collision-like models [8, 41], that quantify at any given time the variation of the density of the contact variable $x$ in terms of the interaction operators. For a given density $f_J(x, t)$, $J \in \{S, I, R\}$, the action on the density of the interaction operators $Q_J(f)(x, t)$ in equations (2) is fruitfully written in weak form. The weak form corresponds to say that for all smooth functions $\varphi(x)$ (the observable quantities) the action of the operator $Q_J$ on $\varphi$ is given by

$$\int_{\mathbb{R}_+} \varphi(x) \, Q_J(f)(x, t) \, dx = \Big\langle \int_{\mathbb{R}_+} B(x)\big(\varphi(x_J^*) - \varphi(x)\big) f_J(x, t) \, dx \Big\rangle. \tag{8}$$

Here expectation $\langle \cdot \rangle$ takes into account the presence of the random parameter $\eta_\epsilon$ in the microscopic interaction (5). The function $B(x)$ measures the interaction frequency. The right-hand side of equation (8) quantifies the variation in density, at a given time $t > 0$, of individuals in the class $J \in \{S, I, R\}$ that modify their value from $x$ to $x_J^*$ (loss term with negative sign) and agents that change their value from $x_J^*$ to $x$ (gain term with positive sign). In [14], the interaction kernel $B(x)$ has been assumed constant. This simplifying hypothesis, which is not always well justified from a modeling point of view, is the common assumption in the Boltzmann-type description of socio-economic phenomena [22, 41]. In particular, the role of a non constant collision kernel $B(x)$ has been analyzed in its critical aspects in [23].

Following the approach in [23, 47], we express the mathematical form of the kernel $B(x)$ by assuming that the frequency of changes which leads to increase the number $x$ of social contacts is inversely proportional to $x$. Hence, we consider collision kernels in the form

$$B(x) = ax^{-b},$$

for some constants $a > 0$ and $b > 0$. This kernel assigns a low probability to interactions in which individuals have already a large number of contacts and assigns a high probability to interactions in which the value of the variable $x$ is small.

The constants $a$ and $b$ can be suitably chosen by resorting to the following argument [47]. For small values of the $x$ variable, the rate of variation of the value function (6) is given by

$$\frac{d}{dx}\Phi_\delta^\epsilon\left(\frac{x}{\bar{x}_J}\right) \approx \mu^2\epsilon\,\bar{x}_J^{-\delta}\,x^{\delta-1}. \tag{9}$$

Hence, for small values of $x$, the mean individual rate predicted by the value function is proportional to $x^{\delta-1}$. Then, the choice $b = \delta$ would correspond to a collective rate of variation of the system independent of the parameter $\delta$ which instead characterizes the individual rate of variation of the value function.

A second important fact is that the individual rate of variation (9) depends linearly on the positive constant $\epsilon$, and it is such that the intensity of the variation decreases as $\epsilon$ decreases. Then, the choice

$$a = \frac{1}{\epsilon},$$

is such that the collective rate of variation remains bounded even in presence of very small values of the constant $\epsilon$.

With these assumptions, the weak form of the interaction integrals in (8), is given by

$$\int_{\mathbb{R}_+} \varphi(x)\,Q_J(f)(x,t)\,dx = \frac{1}{\epsilon}\left\langle \int_{\mathbb{R}_+} x^{-\delta}\big(\varphi(x_J^*) - \varphi(x)\big)f(x,t)\,dx\right\rangle. \tag{10}$$

Note that, in consequence of the choice made on the interaction kernel $B(\cdot)$, the evolution of the density $f(x,t)$ is tuned by the parameter $\epsilon$, which characterizes both the intensity of interactions and the interaction frequency. This choice implies

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Phi_\delta^\epsilon\left(\frac{x}{\bar{x}_J}\right) = \frac{\mu}{2\delta}\left[\left(\frac{x}{\bar{x}_J}\right)^\delta - 1\right], \quad \lim_{\epsilon\to 0}\frac{1}{\epsilon}\langle\eta_\epsilon^2\rangle = \lambda. \tag{11}$$

Consequently, the actions of both the value function and the random part of the elementary interaction in (5) survive in the limit $\epsilon \to 0$.

**Remark 3.** *The approach just described can be easily adapted to other compartmental models in epidemiology like SEIR, MSEIR [6, 13, 30] and/or SIDHARTE [26, 27]. For all these cited models, the fundamental aspects of the interaction between social contacts and the spread of the infectious disease we expect not to change in a substantial way.*

## 3  Asymptotic scaling and steady states

The limit procedure induced by (11) corresponds to the situation in which elementary interactions (5) which produce an extremely small modification of the number of social contacts (quasi-invariant interactions) are prevalent in the dynamics. At the same time, the frequency of these interactions is suitably increased to still permitting to see their effect.

In kinetic theory, this is a well-known procedure with the name of *grazing limit*. We point the interested reader to [11, 21, 41] for further details. Expanding the difference $\varphi(x_J^*) - \varphi(x)$ in Taylor series, and then passing to the limit into the right-hand side of (10), one obtains, for $J \in \{S, I, R\}$

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left\langle \int_{\mathbb{R}_+} x^{-\delta}\big(\varphi(x_J^*) - \varphi(x)\big)f_J(x,t)\,dx\right\rangle = \int_{\mathbb{R}_+} Q_J^\delta(f_J)(x,t)\,\varphi(x)\,dx =$$

$$\int_{\mathbb{R}_+} \left\{ -\varphi'(x) \frac{\mu \, x^{1-\delta}}{2\delta} \left[ \left( \frac{x}{\bar{x}_J} \right)^\delta - 1 \right] + \frac{\lambda}{2} \varphi''(x) \, x^{2-\delta} \right\} f_J(x,t) \, dx$$

If we impose at $x = 0$ the boundary conditions

$$\frac{\partial}{\partial x} (x^{2-\delta} f_J(x,t)) \Big|_{x=0} = 0, \qquad f_J(0,t) = 0, \quad J \in \{S, I, R\}, \tag{12}$$

and a suitable rapid decay of the densities $f_J(x,t)$ at $x = +\infty$ [22], the limit operators $Q_J^\delta$, $J \in \{S, I, R\}$, coincide with the Fokker–Planck type operators

$$Q_J^\delta(f_J)(x,t) = \frac{\mu}{2\delta} \frac{\partial}{\partial x} \left\{ x^{1-\delta} \left[ \left( \frac{x}{\bar{x}_J} \right)^\delta - 1 \right] f_J(x,t) \right\} + \frac{\lambda}{2} \frac{\partial^2}{\partial x^2} (x^{2-\delta} f_J(x,t)), \tag{13}$$

characterized by a variable diffusion coefficient and a time-dependent drift term. Indeed, in view of Remark 2 we have $\bar{x}_J = \bar{x}_J(t)$. Thus, system (2), in the limit $\epsilon \to 0$ corresponds to the following Fokker-Planck system in strong form

$$\begin{aligned}
\frac{\partial f_S(x,t)}{\partial t} &= -K(f_S, f_I)(x,t) + \frac{1}{\tau} Q_S^\delta(f_S)(x,t), \\
\frac{\partial f_I(x,t)}{\partial t} &= K(f_S, f_I)(x,t) - \gamma f_I(x,t) + \frac{1}{\tau} Q_I^\delta(f_I)(x,t), \\
\frac{\partial f_R(x,t)}{\partial t} &= \gamma f_I(x,t) + \frac{1}{\tau} Q_R^\delta(f_R)(x,t).
\end{aligned} \tag{14}$$

The Fokker-Planck system (14) is complemented with the boundary conditions (12) at $x = 0$.

Clearly, the steady distributions satisfy the ordinary differential equations corresponding to equations (14) with time derivatives set equal to zero. It is interesting to remark that the explicit form of the equilibria is easily obtained when both $\beta$ and $\gamma$ are set equal to zero, and the contact function $\kappa(x,y) = 0$, so that the incidence rate $K(f_S, f_I)$ vanishes. In this simple case, by assuming that the mean values $\bar{x}_J$, $J \in \{S, I, R\}$ are constant, and by setting $\nu = \mu/\lambda$, the equilibria are given by the functions

$$f_J^\infty(x) = f_J^\infty(\bar{x}_J) \bar{x}_J^{2-\delta} x^{\nu/\delta + \delta - 2} \exp \left\{ -\frac{\nu}{\delta^2} \left( \left( \frac{x}{\bar{x}_J} \right)^\delta - 1 \right) \right\}, \tag{15}$$

$J \in \{S, I, R\}$. By fixing the mass of the steady state (15) equal to one, in agreement with [3], the consequent probability densities are generalized Gamma $f_\infty(x; \theta, \chi, \delta)$ defined by [35, 46]

$$f_\infty(x; \theta, \chi, \delta) = \frac{\delta}{\theta^\chi} \frac{1}{\Gamma(\chi/\delta)} x^{\chi-1} \exp \left\{ -(x/\theta)^\delta \right\}. \tag{16}$$

characterized in terms of the shape $\chi > 0$, the scale parameter $\theta > 0$, and the exponent $\delta > 0$ that in the present situation are given by

$$\chi = \frac{\nu}{\delta} + \delta - 1, \quad \theta = \bar{x}_J \left( \frac{\delta^2}{\nu} \right)^{1/\delta}. \tag{17}$$

It has to be remarked that the shape $\chi$ is positive, only if the constant $\nu = \mu/\lambda$ satisfies the bound

$$\nu > \delta(1 - \delta). \tag{18}$$

Note that condition (18) holds, independently of $\delta$, when $\mu \geq 4\lambda$, namely when the variance of the random variation in (5) is small with respect to the maximal variation of the value function. Note moreover that for all values $\delta > 0$ the moments are expressed in terms of the parameters denoting respectively the mean $\bar{x}_J$, $J \in \{S, I, R\}$, the variance $\lambda$ of the random effects and the values $\delta$ and $\mu$ characterizing the value function $\phi_\delta^\epsilon$ defined in (6).

The standard Gamma and Weibull distributions are included in (15), and are obtained by choosing $\delta = 1$ and, respectively $\delta = \chi$. In both cases, the shape $\chi = \nu$, and no conditions are required for its positivity.

# 4 A SIR-type evolution of spreading

The Fokker-Planck system (14) contains all the information on the spreading of the epidemic in terms of the distribution of social contacts. Indeed, the knowledge of the densities $f_J(x,t)$, $J \in \{S, I, R\}$, allows to evaluate by integrations all moments of interest. However, in reason of the presence of the incidence rate $K(f_S, f_I)$, as given by (3), the time evolution of the moments of the distribution functions is not explicitly computable, since the evolution of a moment of a certain order depends on the knowledge of higher order moments, thus producing a hierarchy of equations, like in classical kinetic theory of rarefied gases [8]. We will come back to this question later on.

Before discussing the closure, we establish some choices in the model in order to obtain results which are in agreement with the ones reported in [3]. We fix the value $\delta = 1$, and the incidence rate as in (4). In particular, the choice $\delta = 1$ gives, for $J \in \{S, I, R\}$

$$Q_J^1(f)(x) = \frac{\mu}{2} \frac{\partial}{\partial x} \left[ \left( \frac{x}{\bar{x}_J} - 1 \right) f(x) \right] + \frac{\lambda}{2} \frac{\partial^2}{\partial x^2}(xf(x)).$$

In this case (17) imply $\chi = \nu$ and $\theta = \bar{x}_J/\nu$, and the steady states of unit mass, for $J \in \{S, I, R\}$, are the Gamma densities

$$f_J^\infty(x; \theta, \nu) = \left( \frac{\nu}{\bar{x}_J} \right)^\nu \frac{1}{\Gamma(\nu)} x^{\nu-1} \exp \left\{ -\frac{\nu}{\bar{x}_J} x \right\}. \tag{19}$$

With this particular choice, the mean value and the variance of the densities (19), $J \in \{S, I, R\}$, are given by

$$\int_{\mathbb{R}^+} x f_J^\infty(x; \theta, \nu) \, dx = \bar{x}_J; \qquad \int_{\mathbb{R}^+} x^2 f_J^\infty(x; \theta, \nu) \, dx = \frac{\nu+1}{\nu} \bar{x}_J^2.$$

We now go back to the problem of modeling the time evolution of the densities $f_J(x,t)$. We assume that due to the presence of the epidemic, the population tends to reduce the typical average number of contacts which exhibits in standard situations. This can happen due to two main reasons: on voluntary basis for preventing contagion or by authorities decision through lockdown measures. This average reduction effect can be introduced by assuming that the mean number of daily contacts $\bar{x}_J(t)$, $J \in \{S, I\}$ depends on the proportion of infected

$$\bar{x}_J(t) = x_J H(I(t)), \tag{20}$$

where the function $H(r)$ is decreasing with respect to $r$, $0 \leq r \leq 1$, starting from $H(0) = 1$. In particular, in our model, we do not consider a time-dependent mean number of social contacts for the class of recovered. This is due to the fact that the behavior of epidemic does not depend on it. For $J \in \{S, I, R\}$, and $\bar{x}_J(t)$ as in (20), we now define

$$\bar{Q}_J(f_J(x,t)) = \frac{\mu}{2} \frac{\partial}{\partial x} \left[ \left( \frac{x}{\bar{x}_J(t)} - 1 \right) f(x) \right] + \frac{\lambda}{2} \frac{\partial^2}{\partial x^2}(xf(x))$$

With these notations, system (14) with $\delta = 1$ reads

$$\begin{aligned}
\frac{\partial f_S(x,t)}{\partial t} &= -\beta x \, f_S(x,t) m_I(t) + \frac{1}{\tau} \bar{Q}_S(f_S)(x,t), \\
\frac{\partial f_I(x,t)}{\partial t} &= \beta x \, f_S(x,t) m_I(t) - \gamma f_I(x,t) + \frac{1}{\tau} \bar{Q}_I(f_I)(x,t), \\
\frac{\partial f_R(x,t)}{\partial t} &= \gamma f_I(x,t) + \frac{1}{\tau} \bar{Q}_R(f_R)(x,t).
\end{aligned} \tag{21}$$

Integrating both sides of equations in (21) with respect to $x$, and recalling that the Fokker-Planck type operators are mass-preserving, we obtain the system for the evolution of the proportions $J(t)$

defined in (1), $J \in \{S, I, R\}$

$$\frac{\partial S(t)}{\partial t} = -\beta\, m_S(t) m_I(t),$$

$$\frac{\partial I(t)}{\partial t} = \beta\, m_S(t) m_I(t) - \gamma I(t), \tag{22}$$

$$\frac{\partial R(t)}{\partial t} = \gamma I(t).$$

As anticipated, unlike the classical SIR model, system (22) is not closed, since the evolution of the mass fractions $J(t)$, $J \in \{S, I, R\}$ depends on the mean values $m_J(t)$.

Similarly to the derivation of macroscopic equations in kinetic theory, the closure of system (22) can be obtained by resorting, at least formally, on a limit procedure. In fact, as outlined in the Introduction, the typical time scale involved in the social contact dynamic is $\tau \ll 1$ which identifies a faster adaptation of individuals to social contacts with respect to the evolution time of the epidemic disease. The choice of the value $\tau \ll 1$ pushes the distribution function $f_S(x, t)$ towards the Gamma equilibrium density with a mass fraction $S(t)$ and momentum

$$m_S(t) = \bar{x}_S S(t) H(I(t)), \tag{23}$$

as it can be easily verified from the differential expression of the interaction operator $\bar{Q}_S$.

Indeed, if $\tau \ll 1$ is sufficiently small, one can easily argue from the exponential convergence of the solution of the following expression

$$\frac{\partial f_S(x, t)}{\partial t} = \frac{1}{\tau} \bar{Q}_S(f_S)(x, t) \tag{24}$$

towards the equilibrium $f_S^\infty(x; \theta, \nu)$ (see [48] for details), that the solution $f_S(x, t)$ remains sufficiently close to the Gamma density with mass $S(t)$ and momentum given by (23) for all times. This equilibrium distribution $f_S^\infty(x; \theta, \nu)$ can be plugged into the first equation of (21) and then one can integrate it with respect to the variable $x$. This procedure gives formally the first equation of (22). Analogous analysis and formal limit procedure can be done with the second equation in system (21), which leads to relaxation towards a Gamma density with mass fraction $I(t)$ and momentum given by

$$m_I(t) = \bar{x}_I I(t) H(I(t)). \tag{25}$$

Consequently, we obtain the closure

$$\frac{\partial S(t)}{\partial t} = -\bar{\beta}\, S(t)\, I(t)\, H^2(I(t)),$$

$$\frac{\partial I(t)}{\partial t} = \bar{\beta}\, S(t)\, I(t)\, H^2(I(t)) - \gamma I(t), \tag{26}$$

$$\frac{\partial R(t)}{\partial t} = \gamma I(t).$$

In system (26) we defined $\bar{\beta} = \beta \bar{x}_I \bar{x}_S$. This identifies the classical transmission parameter of the SIR model, where however now the difference is that this quantities are not postulated but instead derived starting from microscopic considerations. In the following, we refer to this model as the social contact-SIR model (S-SIR).

**Remark 4.** *The outlined closure strategy can be obtained also by resorting to the splitting method, a very popular numerical approach for the Boltzmann equation [24, 40]. If at each time step $(t, t + \Delta t)$ we consider sequentially the population-based interaction and relaxation operators in the first equation in (14), during this short time interval we recover the evolution of the density from the joint action of relaxation*

$$\frac{\partial f_S(x, t)}{\partial t} = \frac{1}{\tau} \bar{Q}_S(f_S)(x, t) \tag{27}$$

11

and SIR interaction

$$\frac{\partial f_S(x,t)}{\partial t} = -\beta x \, f_S(x,t) m_I(t) \tag{28}$$

We recall once again that the typical time to consider is $\tau \ll 1$, which identifies a faster adaptation of individuals to social contacts with respect to the evolution time of the epidemic disease. Since the operator $\bar{Q}_S$ is mass preserving, in the considered time interval the relaxation (27) with a value $\tau \ll 1$ pushes the solution of equation (24) towards the Gamma equilibrium density with the same mass fraction $S(t)$ of the initial datum, and momentum

$$m_S(t + \Delta t) = \bar{x}_S S(t) H(I(t)), \tag{29}$$

as it can be easily verified from the differential expression of the interaction operator $\bar{Q}_S$.

   Indeed, if $\tau \ll 1$ is sufficiently small, one can easily argue from the exponential convergence of the solution to (24) towards the equilibrium [48], that the solution $f_S(x, t + \Delta t)$ is sufficiently close to the Gamma density with mass $S(t)$ and momentum given by (29), and this density can be used into the SIR step (28) to close the splitted system (24)-(28). Analogous procedure can be done with the second equation in system (21), which leads to relaxation towards a Gamma density with mass fraction $I(t)$ and momentum given by

$$m_I(t + \Delta t) = \bar{x}_I I(t) H(I(t)). \tag{30}$$

Consequently, substituting into system (22) we obtain the closed system (26).

   It remains to quantify the action of the social contacts on the evolution of the epidemic. For this reason, let us introduce for a given constant $N \gg 1$ the decreasing function [34]

$$H(r) = \frac{1}{\sqrt{1 + Nr}}, \quad 0 \le r \le 1, \tag{31}$$

which describes a possible way in which, in presence of the spread of the disease, the susceptible and the infected population tend to reduce the mean number of daily social contacts $\bar{x}_J$, $J \in \{S, I\}$. This choice produces a SIR model with global incidence rate

$$D(S, I)(t) = \int_{\mathbb{R}^+} K(f_S, f_I)(x, t) \, dx = \bar{\beta} S(t) \frac{I(t)}{1 + N \, I(t)}, \tag{32}$$

that fulfils all the properties required by the non-linear incidence rates considered in [34]. Indeed, $D(S, 0) = 0$, and the function $D(S, I)$ satisfies

$$\frac{\partial D(S, I)}{\partial I} > 0, \quad \frac{\partial D(S, I)}{\partial S} > 0 \tag{33}$$

for all $S, I > 0$. Moreover $D(S, I)$ is concave with respect to the variable I, i.e.

$$\frac{\partial^2 D(S, I)}{\partial I^2} \le 0, \tag{34}$$

for all $S, I > 0$. In addition to the form (31), we can also consider the following function

$$H(t, r) = \frac{1}{\sqrt{1 + Nr(t) \displaystyle\int_0^t r(s) ds}}, \quad 0 \le r \le 1. \tag{35}$$

This function satisfies the same properties than (31) in terms of the incidence rate requests detailed in [34] and takes into account memory effects on the population's behavior. In fact, people may adapt their life style in terms of possible daily contacts to answer to the actual pandemic situation as
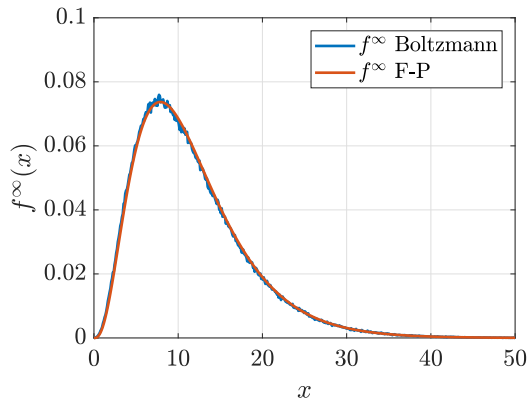
Figure 1: **Test 1**. Initial data and convergence of the Boltzmann dynamics towards the equilibrium state of the corresponding Fokker-Planck equation.

expressed by (31), however further effects based on the global dynamics of the disease can influence the choices of the population.

To conclude, let us introduce the basic reproduction number $R_0$ of this model, i.e. the average number of secondary cases produced by a single infected agent introduced into an entirely susceptible population. This is given by

$$R_0 = \frac{1}{\gamma} \lim_{I \to 0, S \to 1} \frac{\partial D(S,I)}{\partial I} = \frac{\bar{\beta}}{\gamma}, \tag{36}$$

where $\gamma$ is the recovery rate of infected. According to the analysis of [34], an autonomous compartmental epidemiological model with the non-linear incidence rate (32) under the constant population size assumption is stable. Such system has either a unique and stable endemic equilibrium state or no endemic equilibrium state at all. Since the incidence rate $D(S,I)$ satisfies the conditions (33)-(34), if $R_0 > 1$, then the endemic equilibrium state $Q^* = (S^*, I^*)$ of system (26) is asymptotically stable. If $R_0 \leq 1$, then there is no endemic equilibrium state, and the infection-free equilibrium state is asymptotically stable.

## 5 Numerics

In addition to analytic expressions, numerical experiments allow us to visualize and quantify the effects of social contacts on the SIR dynamics used to describe the time evolution of the epidemic. More precisely, starting from a given equilibrium distribution detailing in a probability setting the daily number of contacts of the population, we show how the coupling between social behaviors and number of infected people may modify the epidemic by slowing down the number of encounters leading to infection. In a second part, we discuss how some external forcing, mimicking political choices and acting on restrictions on the mobility, may additionally improve the reduction of the epidemic trend avoiding concentration in time of people affected by the disease and then decreasing the probability of hospitalization peaks. In a third part, we focus on experimental data for the specific case of COVID-19 in different European countries and we extrapolate from them the main features characterizing the contact function $H(\cdot)$.

### 5.1 Test 1: On the effects of the social contacts on the epidemic dynamics

The starting point is represented by a population composed of 99% of susceptibles and 0.01% of infected. The distribution of the number of contact is described by (15) with $\nu = 4$, $\delta = 8$ and
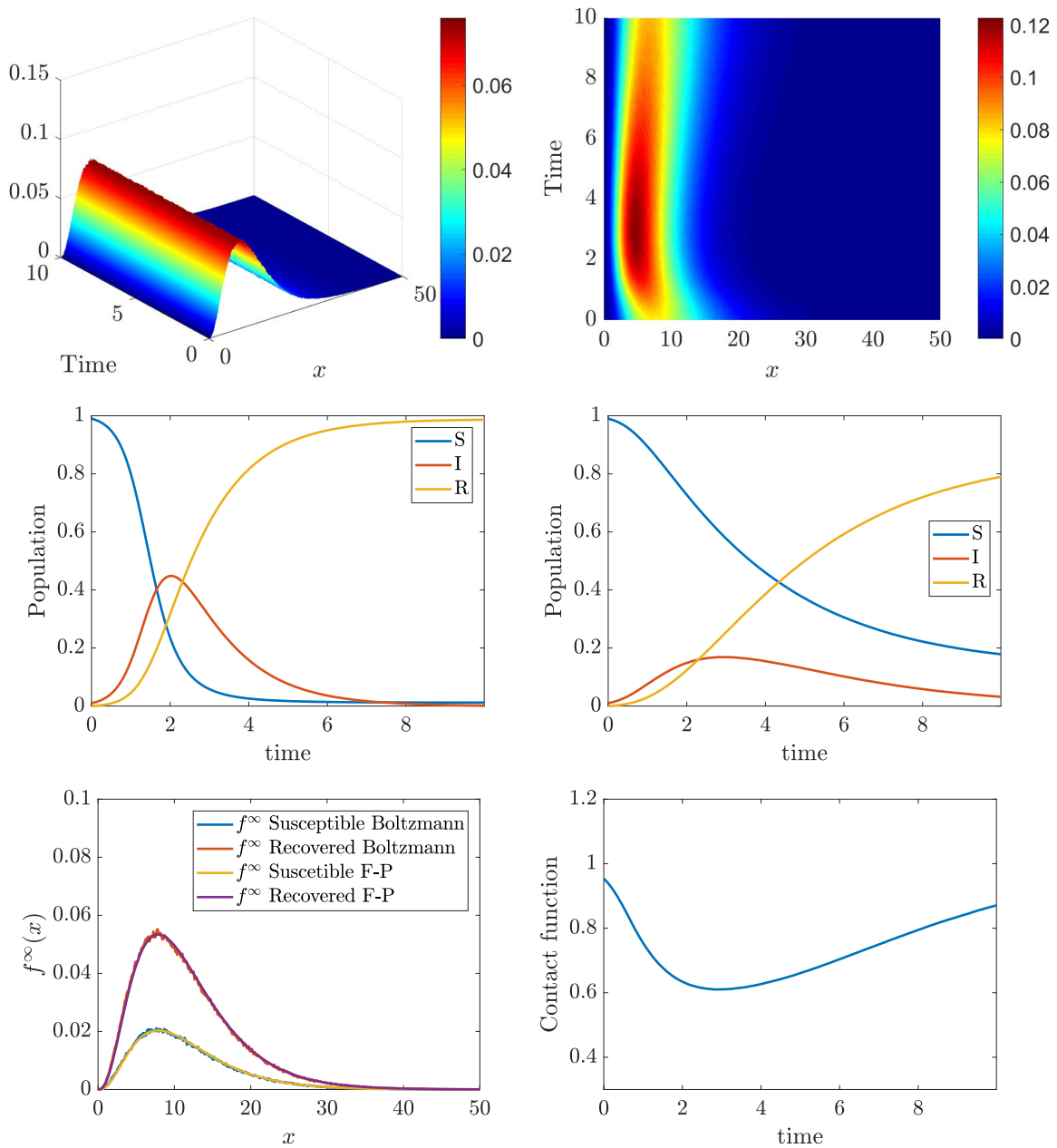
Figure 2: **Test 1**. Top: distribution of the social contacts at equilibrium over time for the two choices of the function $H$. Middle: SIR dynamics corresponding to the different choices of the different mean number of daily contact (left constant case, right as a function of the epidemic). Bottom left: final distribution of the number of daily contact for the susceptible and recovered agents. Bottom right: time evolution of the contact function $H$.
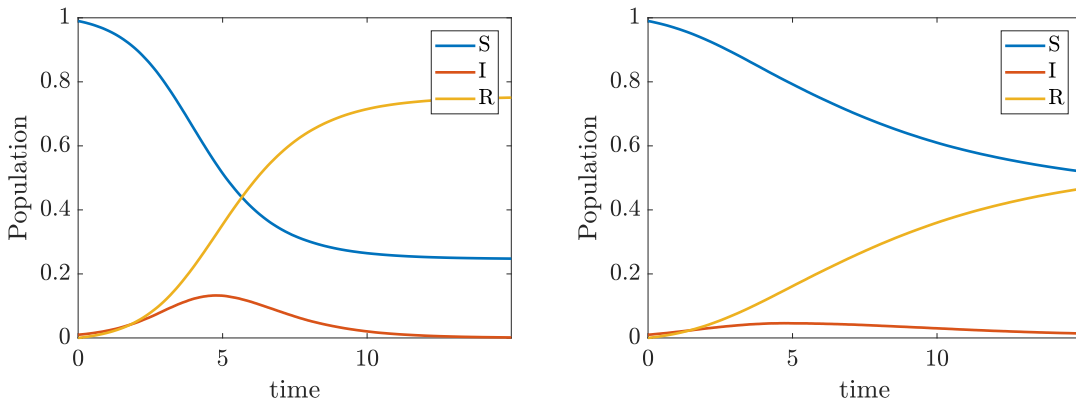
Figure 3: Test 1. SIR dynamic corresponding to different choices for the mean number of daily contacts (left constant case, right as a function of the epidemic) with parameters $k(x) = 0.2/x_J x$ and $\gamma = 1$.

$x_J = 15$ while the epidemic parameters are $k(x) = \frac{0.8}{x_J} x$ and $\gamma = 0.8$. The kinetic model (14) is solved by a splitting strategy together with a Monte Carlo approach where the number of samples used to described the population is fixed to $M = 10^6$. The time step is fixed to $\Delta t = 10^{-2}$ and the scaling parameter is $\tau = 10^{-2}$. These choices are enough to observe the convergence of the Boltzmann dynamics to the Fokker-Planck one as shown in Figure 1 where the analytical equilibrium distribution is plotted together with the results of the Boltzmann dynamics. We considered also uniform initial distribution

$$f_0(x) = \frac{1}{c}\chi(x \in [0,c]), \quad c = 30, \tag{37}$$

being $\chi(\cdot)$ the indicator function. In the introduced setting, we then compare two distinct cases: in the first one we suppose that the social contacts do not affect the solution, meaning $H(I(t)) = 1$, while the second includes the effects of the function $H(I(t))$ given in (31) with $N = 10$. The results are depicted in Figure 2. The top right images show the time evolution of the distribution of the number of contacts for the two distinct cases, while the middle images report the corresponding evolution of the epidemic. For this the second case, the function $H(I(t))$ as well as the distribution of contacts for respectively the susceptible and the recovered are shown at the bottom of the same figure. We clearly observe a reduction of the peak of infected in the case in which the dynamics depends on the number of contacts with $H(I(t))$ given by (31).

We now repeat the same simulation by only changing the considered epidemic parameters. In particular, we consider a lower infection and recovery rate given by $k(x) = 0.2/x_J x$ and $\gamma = 1$, respectively. In Figure 3, we show the evolution of two epidemic profiles in time for the case in which social contacts do not affect the solution and for the case in which $H(I(t))$ is a function of the number of infected as for the previous situation. Results show the same qualitative behavior: peak reduction and spread of the number of infected over time is observed when sociality is taken into account. The total number of infected is also reduced in the second case.

## 5.2 Test 2: Forcing a change in the social attitudes.

Next, we compare the effects on the spread of the disease when the population adapts its habits with a time delay with respect to the onset of the epidemic. This kind of dynamics corresponds to a modeling of a possible lockdown strategy whose effects are to reduce the mobility of the population and, correspondingly, to reduce the number of daily contacts in the population.

The setting is similar to the one introduced in Section 5.1 and we consider a switch between $H = 1$ to $H(I(t)) = 1/\sqrt{1 + NI(t)}$ when the number of infected increases. The social parameters

15

are $\nu = 4$, $\delta = 8$ and $x_J = 15$, as before, while the epidemic parameters are $k(x) = 0.4/x_J x$ and $\gamma = 0.6$, the final time is fixed to $T = 15$. The initial distribution of contacts is also assumed to be of the form (37).

We consider three different settings, in the first one $H = 1$ up to $t < T/4$, in the second one up to $t < T/2$ while in the third one we prescribe a lockdown for a limited amount of time ($1 < t < 2$) and then we relax back to $H = 1$. The results are shown in Figure 4 for both the distribution of daily contacts over time and the SIR evolution. We can identify three scenarios. The first on the top gives a slightly change to the standard epidemic dynamics. Indeed, we can observe around $t = 7.5$ a reduction of the speed of the infection. For the second, we observe an inversion in the trend of the epidemic around $t = 4$, while for the third case we first observe inversion and then the resurgence of the number of infected when the lockdown measures are relaxed.

## 5.3 Test 3: Extrapolation of the contact function from data

In this part, we consider data about the dynamics of COVID-19 in three European countries: France, Italy and Spain. For these three countries, the evolution of the disease, in terms of reported cases, evolved in rather different ways. The estimation of epidemiological parameters of compartmental models is an inverse problem of generally difficult solution for which different approaches can be considered. We mention in this direction a very recent comparison study [37]. It is also worth to mention that often the data are partial and heterogeneous with respect to their assimilation, see for instance discussions in [1, 7, 9, 45]. This makes the fitting problem challenging and the results naturally affected by uncertainty.

The data concerning the actual number of infected, recovered and deaths of COVID-19 are publicly available from the John Hopkins University GitHub repository [17]. For the specific case of Italy, we considered instead the GitHub repository of the Italian Civil Protection Department [42]. In the following, we present the adopted approach which is based on a strategy with two optimisation horizons (pre-lockdown and lockdown time spans) depending on the different strategies enacted by the governments of the considered European countries.

In details, we considered first the time interval $t \in [t_0, t_\ell]$, being $t_\ell$ the day in which lockdown started in each country (Spain, Italy and France) and $t_0$ the day in which the reported cases hit 200 units. The lower bound $t_0$ has been imposed to reduce the effects of fluctuations caused by the way in which data are measured which have a stronger impact when the number of infected is low. Once the time span has been fixed, we then considered a least square problem based on the minimization of a cost functional $\mathcal{J}$ which takes into account the relative $L^2$ norm of the difference between the reported number of infected and the reported total cases $\hat{I}(t)$, $\hat{I}(t) + \hat{R}(t)$ and the evolution of $I(t)$ and $I(t) + R(t)$ prescribed by system (26) with $H \equiv 1$. In practice, we solved the following constrained optimisation problem

$$\min_{\beta, \gamma} \mathcal{J}(\hat{I}, \hat{R}, I, R) \tag{38}$$

where the cost functional $\mathcal{J}$ is a convex combination of the mentioned norms and assumes the following form

$$J(\hat{I}, \hat{R}, I, R) = p \frac{\|\hat{I}(t) - I(t)\|_2}{\|\hat{I}(t)\|_2} + (1 - p) \frac{\|\hat{I}(t) + \hat{R}(t) - I(t) - R(t)\|_2}{\|\hat{I}(t) + \hat{R}(t)\|_2}$$

We choose $p = 0.1$ and we look for a minimum under the constraints

$$0 \leq \beta \leq 0.6, \qquad 0.04 \leq \gamma \leq 0.06.$$

In Table 1 we report the results of the performed parameter estimation together with the resulting attach rate $R_0$ defined in (36).

Once that the contagion parameters have been estimated in the pre-lockdown time span, we successively proceeded with the estimation of the shape of the function $H$ from the data. To estimate
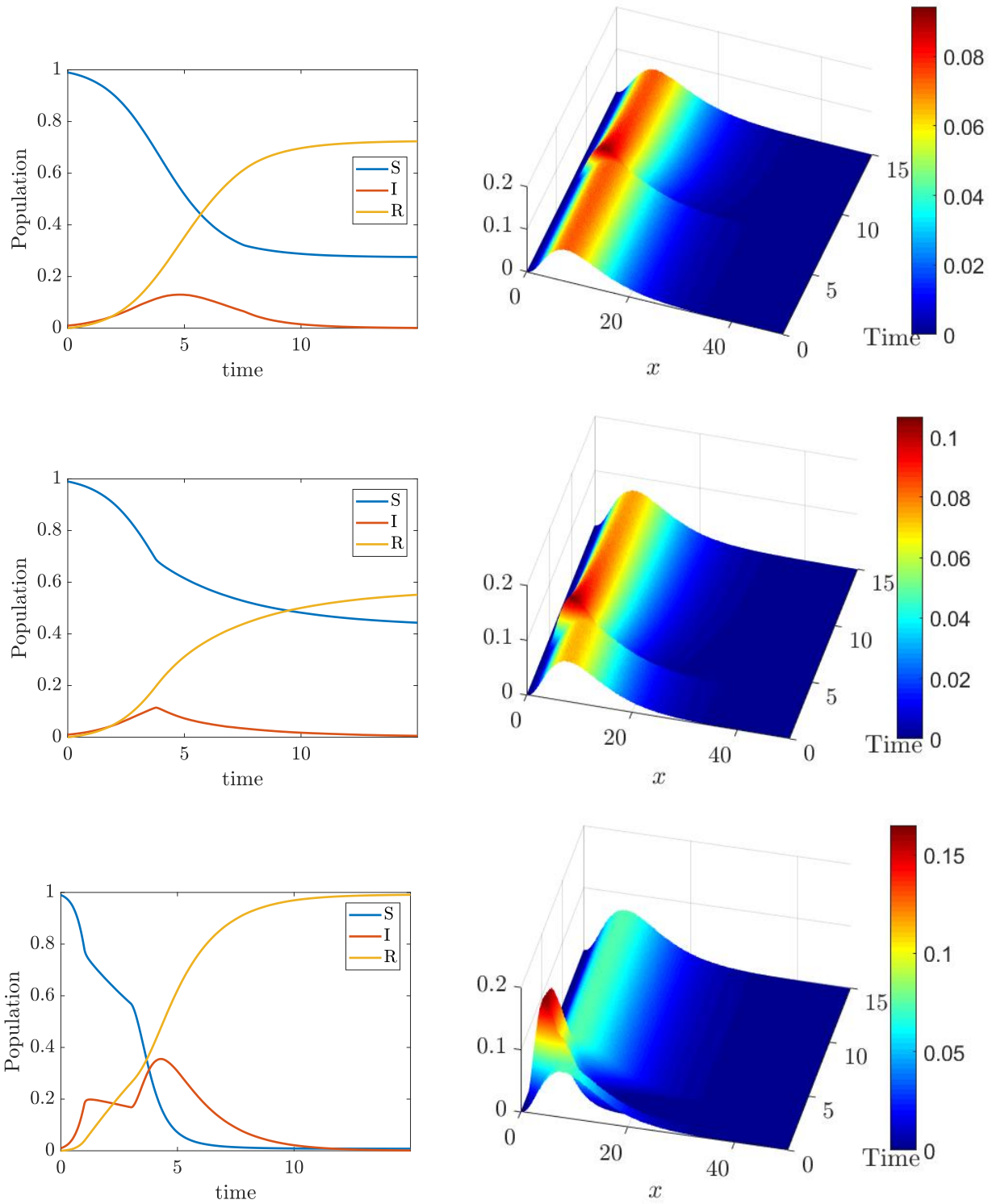
16

Figure 4: **Test 2**. Comparisons of different lockdown behaviors. Top: late lockdown. Middle: early lockdown. Bottom: early lockdown and successive relaxation.
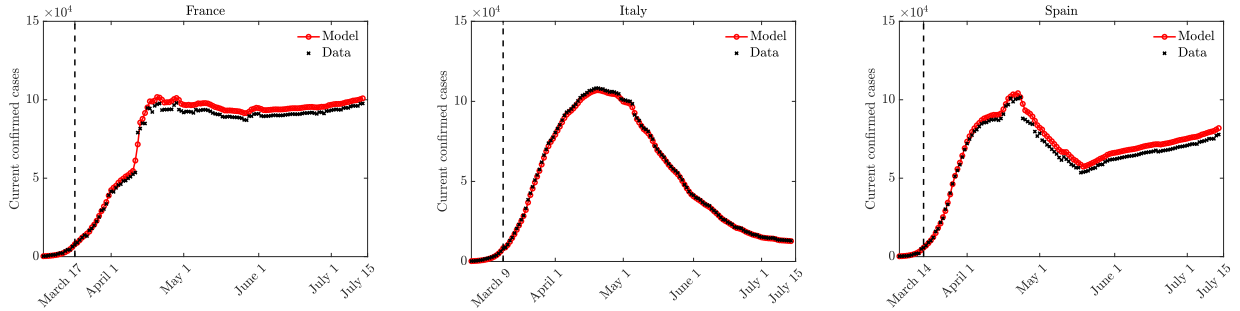
Figure 5: **Test 3**. Fitting of the parameters of model (26) where $\beta, \gamma > 0$ were estimated before the lockdown measures assuming $H \equiv 1$. The parameters characterizing the function $H(\cdot)$ in (26) have been computed during and after lockdown at regular interval of time, up to July 15. The lockdown measures change in each county (dashed line).

|       | France<br>Mar 5-Mar 17 | Italy<br>Feb 24-Mar 9 | Spain<br>Mar 5-Mar 14 |
|-------|-----------------------|----------------------|----------------------|
| $\beta$    | 0.300686 | 0.317627 | 0.370362 |
| $\gamma$   | 0.040000 | 0.058391 | 0.043168 |
| $R_0$ | 7.5172   | 5.4397   | 8.5795   |

Table 1: **Test 3**. Model fitting parameters in estimating attack values for the COVID-19 outbreak before lockdown in various European countries.

this latter quantity, we solved the following optimization problem

$$\min_H \mathcal{J} \tag{39}$$

in terms of $H$ where $\mathcal{J}$ is the same functional of the previous step and where in the evolution of the macroscopic model the values $\beta, \gamma$ have been fixed as a result of the first optimization in the pre-lockdown period. The parameters chosen for (39) are $p = 0.1$ while the constraint is

$$0 \leq H \leq 1.$$

The second optimisation problem has been solved up to last available data for each country with daily time stepping $h = 1$ and over a time window of three days. This has been done with the scope of regularizing possible errors due to late reported infected and smoothing the shape of $H$. Both optimisation problems (38)-(39) have been tackled using the Matlab functions `fmincon` in combination with a RK4 integration method of the system of ODEs.

In Figure 5, we present the result of such fitting procedure between the model (26) and the experimental data. The evolution of the estimated $H(t)$, $t \in [t_\ell, T]$, is instead presented in the left column of Figure 6. From this figure, it can be observed in the case of Italy how, even if the daily number of infected decreases after May 1st, the estimated $H$ remains quite stable after this day. This behavior cannot be reproduced by using a function $H(I(t))$ as the one given in (31). Instead, a function $H(t, I(t))$ of the form (35), which takes into account both the instantaneous number of confirmed infections $NI(t)$ and the total number of infected in the population $N \sum_{s<t} I(s)$ up to time $t$, is able to better better fit the observed results. In fact, it appears that taking into account possible memory effects on the contact function is a good path to follow. In the right column of Figure 5, the two discussed contact functions have been denoted with $H_1 = H_1(NI(t))$ and

18

| $H_1$ | France Mar 18 - Jul 15 | Italy Mar 10 - Jul 15 | Spain Mar 15 - Jul 15 |
|---|---|---|---|
| $a$ | $3 \cdot 10^{-5}$ | $9.038 \cdot 10^{-5}$ | $1.749 \cdot 10^{-5}$ |
| $b$ | 0.7168 | 0.5808 | 1.249 |
| $R^2$ | 0.7096 | 0.7542 | 0.5795 |
| $H_2$ | France Mar 18 - Jul 15 | Italy Mar 10 - Jul 15 | Spain Mar 15 - Jul 15 |
| $a$ | 1.41 | 0.01744 | 0.5201 |
| $b$ | 0.1055 | 0.2566 | 0.1211 |
| $R^2$ | 0.7007 | 0.7959 | 0.9899 |

Table 2: Fitting parameters for the estimate of the contact functions $H_i$, $i = 1, 2$ in different countries based on the evolution of $H(t)$ solution of the optimisation problem (39). The corresponding $R^2$ coefficient is also reported.

$H_2 = H_2(t, NI(t) \sum_{s<t} I(s))$. Both are assumed of the form

$$H_i(r) = \frac{1}{(1 + ar)^b}, \qquad i = 1, 2.$$

A least square minimization approach has been finally developed to determine $a > 0, b \in \mathbb{R}$ that best fit the estimated curve whose results are presented in Table 2. We can easily observe in the right column of Figure 6 how the function $H_2$ is capable to better explain the estimated values of $H$ especially after the epidemic peak. To evaluate the goodness of fit we can finally use the so-called coefficient $R^2$, as reported in Table 2. Results show that the function $H_2$ appears more suitable in terms of the fitting results for all tested situations. This fact may indicate that people are rather fast to apply social distancing, and therefore to reduce their average number of contacts, whereas they tend to restore the pre-pandemic average contact rate more slowly, possibly due to further memory effects.

## 5.4 S-SIR model with fitted contact function

In this last part, we discuss the results of the S-SIR model when the contact function has the shape extrapolated in the previous paragraph. In particular, we devoted it to show whether the obtained extrapolated function $H$, which depends on the product between the current number of infected and the total number of infected, produces qualitative trends that are in agreement with the data.

As discussed in Section 5.3, it appears that the curve of infection may be better explained in the case in which the contact function is both a function of the instantaneous number of infected and of the total number of people that contracted the disease up to time $t$. In order to compare qualitatively the observed curve of infected and the theoretical one, we consider the following setting for the three countries under study: $\nu = 4, \delta = 8$ and $x_J = 15$, $\Delta t = 0.01$, $\tau = 0.01$, $M = 10^5$. Moreover, we suppose $S(t = 0)$ and $I(t = 0)$ to match the relative number of susceptible and infected of each country at the time in which we start our comparison. Finally, we consider

$$H\left(t, NI(t) \sum_{s<t} I(s)\right) = \frac{1}{(1 + aNI(t) \sum_{s<t} I(s))^b}, \tag{40}$$

where we use the parameters of Table 2 that we recall here for the seek of clarity: $(a, b) = (1.41, 0.15)$ in the case of France, and $(a, b) = (0.017, 0.25)$ in the case of Italy. The case of Spain will be discussed later.
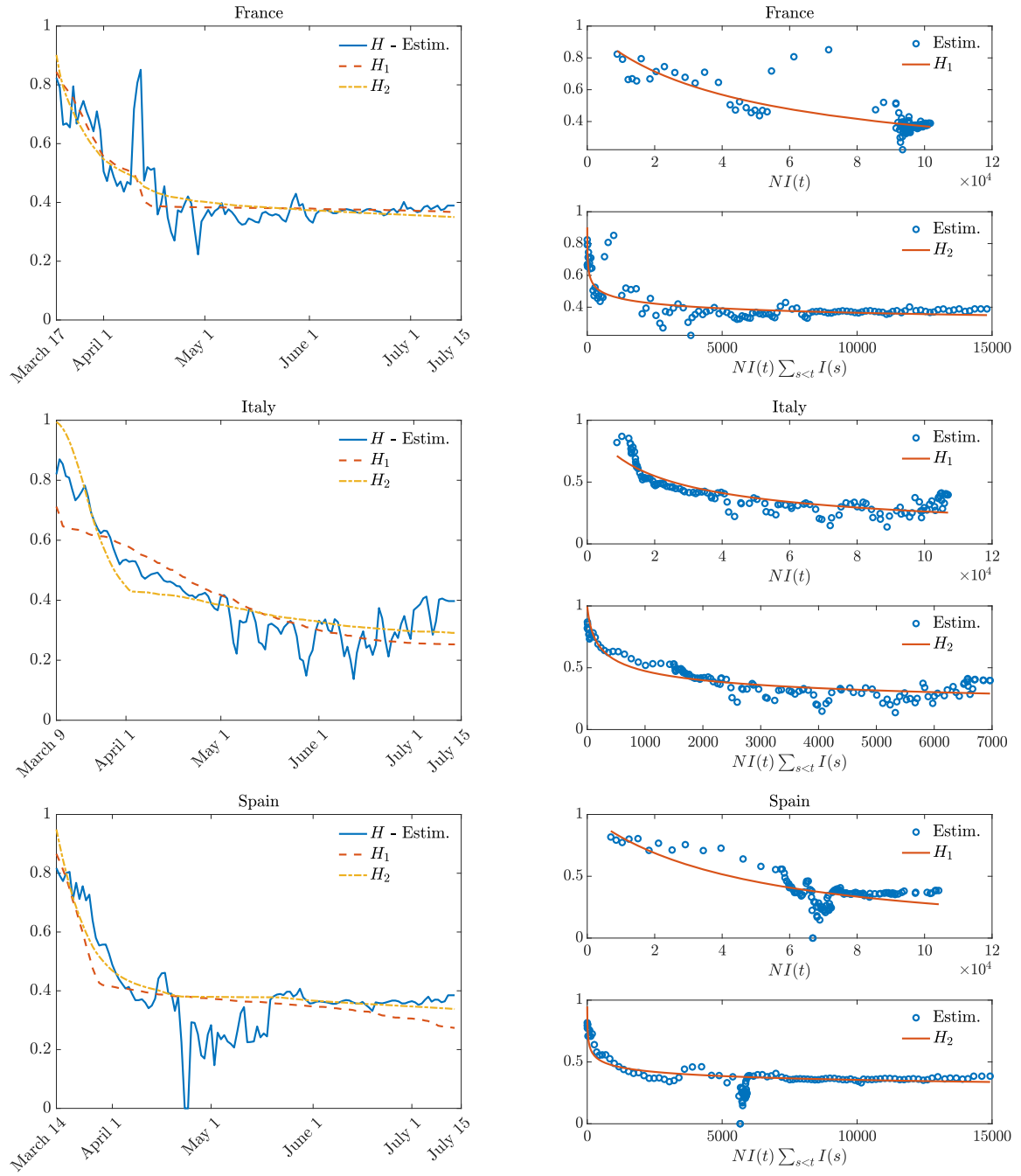
Figure 6: **Test 3**. Estimated shape of the function $H$ in several European countries (left plots) and its dependency on the variables $NI(t)$ and $NI(t)\sum_{s<t} I(s)$ (right plots).
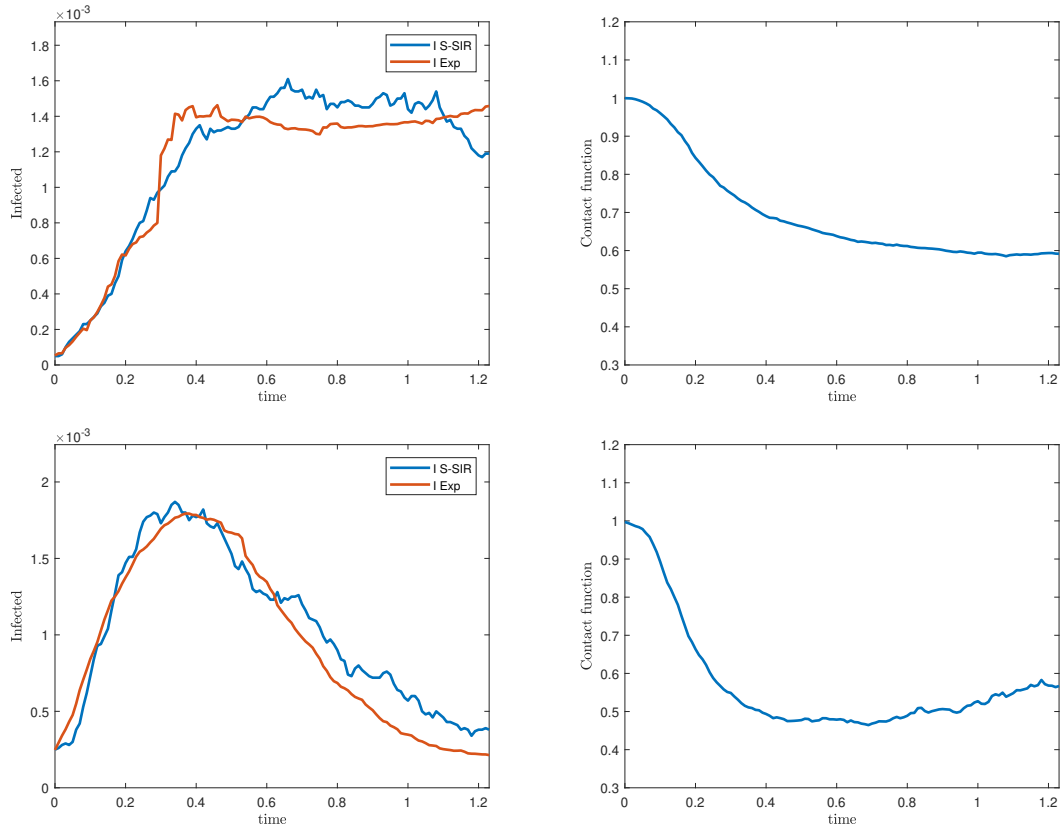
Figure 7: **Test 4**. Left: relative number of infected over time for the S-SIR model when memory effects are taken into account in the contact function. Right: corresponding contact function over time. Top: the French case. Bottom: the Italy case.

In Figure 7 we show the profiles of the infected over time together with the shape of the function $H$ again over time. The results show that with the choices done for the contact function, it is possible to reproduce at least qualitatively the shape of the trend of infected during the pandemic observed in Italy and in France.

It is worth to remark that the considered social parameters have been estimated only the in the case of France, see [3], and we assumed that the initial contact distribution is the same for the Italian case. We now consider the case of Spain. For this country, according to Figure 5, the trend of infected undergoes a deceleration during the lockdown period. This can be also clearly observed in Figure 6 where the extrapolated shape of the contact function $H$ is shown. Let also observe that while the global behavior of this function is captured by the fitting procedure, we however lose the minimum which takes place around end of April. This minimum is responsible of the deceleration in the number of infected and can be brought back to a strong external intervention in the lifestyle of Spain country with the scope of reducing the hospitalizations. This effect can be reproduced by our model by imposing the same behavior in the function $H$. To that aim, the Figure 8 reports finally the profile of the infected over time together with the shape of the function $H$ again over time for this last case. The results show that also in this case, the S-SIR model is capable to qualitatively reproduce the data.
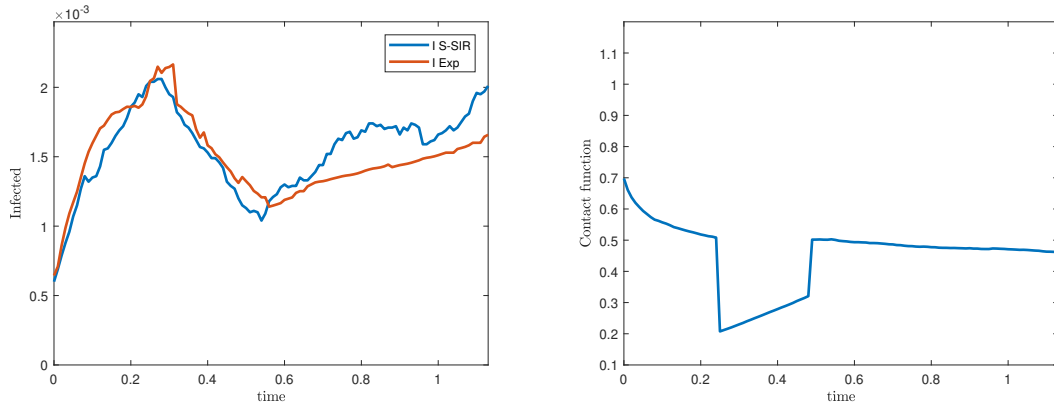
Figure 8: **Test 4**. Left: relative number of infected over time for the S-SIR model when memory effects are taken into account in the contact function. Right: corresponding contact function over time. Spain case.

## Conclusions

The development of strategies for mitigating the spreading of a pandemic is an important public health priority. The recent case of COVID-19 pandemic has seen as main strategy restrictive measures on the social contacts of the population, obtained by household quarantine, school or workplace closure, restrictions on travels, and, ultimately, a total lockdown. Mathematical models represent powerful tools for a better understanding of this complex landscape of intervention strategies and for a precise quantification of the relationships among potential costs and benefits of different options [18]. In this direction, we introduced a system of kinetic equations coupling the distribution of social contacts with the spreading of a pandemic driven by the rules of the SIR model, aiming to explicitly quantify the mitigation of the pandemic in terms of the reduction of the number of social contacts of individuals. The kinetic modeling of the statistical distribution of social contacts has been developed according to the recent results in [3], which present an exhaustive description of contacts in the France population, divided by categories. The numerical experiments then show that the kinetic system is able to capture most of the phenomena related to the effects of partial lockdown strategies, and, eventually to maintain pandemic under control.

## Acknowledgement

## References

[1] Albi, G.; Pareschi, L.; Zanella, M.: Control with uncertain data of socially structured compartmental models. Preprint `arXiv:2004.13067` (2020)

[2] Aitchison, J.; Brown J.A.C.: *The Log-normal Distribution*, Cambridge University Press, Cambridge, UK (1957)

[3] Béraud, G. et al.: The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLoS ONE* **10**(7): e0133203 (2015)

[4] Bobylev, A.: The theory of the nonlinear, spatially uniform Boltzmann equation for Maxwellian molecules, *Sov. Sco. Rev. C Math. Phys.* **7**: 111–233 (1988)

[5] Bonaccorsi, G. et al.: Economic and social consequences of human mobility restrictions under COVID-19, *PNAS*, **117**(27): 15530-15535 (2020)

[6] Brauer, F.; Castillo-Chavez, C.; Feng, Z.: *Mathematical Models in Epidemiology.* With a foreword by Simon Levin. Texts in Applied Mathematics, 69. Springer, New York (2019)

[7] Capaldi, A. et al.: Parameter estimation and uncertainty quantification for an epidemic model. *Math. Biosci. Eng.* **9**(3): 553–576 (2012)

[8] Cercignani, C.: *The Boltzmann Equation and its Applications*, Springer Series in Applied Mathematical Sciences, Vol. **67** Springer–Verlag, New York, NY (1988)

[9] Chowell, G.: Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecast. *Infect. Dis. Model.*, **2**(3): 379–398 (2017)

[10] Cooke, K; Van Den Driessche, P.; Zou, X.: Interaction of maturation delay and nonlinear birth in population and epidemic models. *J. Math. Biol.* **39**: 332–352 (1999)

[11] Cordier, S.; Pareschi, L.; Toscani, G.: On a Kinetic Model for a Simple Market Economy. *J. Stat. Phys.* **120**: 253–277 (2005)

[12] Diekmann, O.; Heesterbeek, J.A.P.; Metz, J.A.J.: On the definition and the computation of the basic reproduction ratio $R_0$ in models for infectious diseases in heterogeneous populations, *J. Math. Biol.* **28**(4): 365–382 (1990)

[13] Diekmann, O.; Heesterbeek, J.A.P.: *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation* John Wiley & Sons, Chichester, UK 2000

[14] Dimarco, G.; Toscani, G.: Kinetic modeling of alcohol consumption. *J. Stat. Phys.* **177**: 1022–1042 (2019)

[15] Dimarco, G.; Pareschi, L.; Toscani, G.; Zanella M.: Wealth distribution under the spread of infectious diseases. *Phys. Rev. E* **102**: 022303 (2020)

[16] Dolbeault, J.; Turinici, G.: Heterogeneous social interactions and the COVID-19 lockdown outcome in a multi-group SEIR model. *Math. Model. Nat. Pheno.*, **15**: Article Number 36 (2020).

[17] Dong, E.; Du, H.; Gardner, L.: An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* (2020). `https://plague.com`

[18] Ferguson, N.M. et al.: Strategies for mitigating an influenza pandemic. *Nature* **442**: 448–452 (2006)

[19] Flaxman et al.: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries, *Report 13. Imperial College COVID-19 Response Team* (2020)

[20] Fumanelli, L. et al.: Inferring the Structure of Social Contacts from Demographic Data in the Analysis of Infectious Diseases Spread. Salathé M., editor. *PLoS Comput Biol.* **8**: e1002673 (2012)

[21] Furioli, G.; Pulvirenti, A.; Terraneo, E.; Toscani, G.: Fokker–Planck equations in the modelling of socio-economic phenomena, *Math. Mod. Meth. Appl. Sci.* **27**: 115–158 (2017)

[22] Furioli, G.; Pulvirenti, A.; Terraneo, E.; Toscani, G.: Fokker–Planck equations in the modelling of socio-economic phenomena. *Math. Mod. Meth. Appl. Sci.* **27**(1): 115–158 (2017)

[23] Furioli, G.; Pulvirenti, A.; Terraneo, E.; Toscani, G.: Non-Maxwellian kinetic equations modeling the evolution of wealth distribution. *Math. Mod. Meth. Appl. Sci.* **30**(4): 685–725 (2020)

[24] Gabetta, E .; Pareschi, L .; Toscani, G.: Relaxation schemes for nonlinear kinetic equations. *SIAM J. Num. Anal.* **34**: 2168–2194 (1997)

[25] Gaeta, G.: A simple SIR model with a large set of asymptomatic infectives. Preprint `arXiv:2003.08720` (2020).

[26] Gatto, M. et al.: Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *PNAS* **117**(19): 10484–10491 (2020)

[27] Giordano, G. et al.: Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine* **26**: 855–860 (2020)

[28] Gualandi, S.; Toscani, G.: Call center service times are lognormal. A Fokker–Planck description. *Math. Mod. Meth. Appl. Sci.*, **28**(08): 1513–1527 (2018)

[29] Gualandi, S.; Toscani, G.: Human behavior and lognormal distribution. A kinetic description. *Math. Mod. Meth. Appl. Sci.* **29**(4): 717–753 (2019)

[30] Hethcote, H.W.: The Mathematics of Infectious Diseases. *SIAM Rev.* **42**(4): 599–653 (2000)

[31] Kahneman, D.; Tversky, A.: Prospect theory: an analysis of decision under risk, *Econometrica* **47**(2): 263–292 (1979)

[32] Kahneman, D.; Tversky, A.: *Choices, values, and frames*, Cambridge University Press, Cambridge, UK 2000

[33] Kehoe, T. et al.: Determining the best population–level alcohol consumption model and its impact on estimates of alcohol–attributable harms. *Population Health Metrics* **10**(6) (2012)

[34] Korobeinikov, A; Maini, P.K.: Non-linear incidence and stability of infectious disease models. *Mathematical Medicine and Biology* **22**: 113–128 (2005)

[35] Lienhard, J.H.; Meyer, P.L.: A physical basis for the generalized Gamma distribution. *Quarterly of Applied Mathematics*, **25**(3): 330–334 (1967)

[36] Limpert, E.; Stahel, W.A.; Abbt, M.: Log-normal distributions across the sciences: keys and clues. *BioScience* **51**(5): 341–352 (2001)

[37] Liu, Y.; Gayle, A. A. ; Wilder-Smith, A.; Rocklöv, J.: The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J. Travel Med.*, **27**(2): 1–4 (2020)

[38] Mossong, J. et al.: Social contacts and mixing pat- terns relevant to the spread of infectious diseases. *PLoS Med.* **5**: e74 (2008)

[39] Novozhilov, A.S.: On the spread of epidemics in a closed heterogeneous population. *Mathematical Biosciences* **215**: 177–185 (2008)

[40] Pareschi L.; Russo G.: Time Relaxed Monte Carlo Methods for the Boltzmann Equation. *SIAM J. Sci. Comput.* **23**: 1253–1273 (2001)

[41] Pareschi, L.; Toscani, G.: *Interacting multiagent systems: kinetic equations and Monte Carlo methods*, Oxford University Press, Oxford 2014

[42] Presidenza del Consiglio dei Ministri, Dipartimento della Protezione Civile. GitHub: COVID-19 Italia - Monitoraggio Situazione, `https:github.com/pcm-dpc/COVID-19`.

[43] Rehm, J. et al.: Statistical modeling of volume of alcohol exposure for epidemiological studies of population health: the US example. *Population Health Metrics* **8**(3) (2010)

[44] Riley, S. et al.: Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. *Science* **300**: 1961–1966 (2003)

[45] Roberts M. G.: Epidemic models with uncertainty in the reproduction. *J. Math. Biol.*, **66**: 1463–1474 (2013)

[46] Stacy, E.W.: A generalization of the Gamma distribution. *Ann. Math. Statist.* **33**: 1187–1192 (1962)

[47] Toscani, G.: Statistical description of human addiction phenomena. In "Trails in Kinetic Theory: foundational aspects and numerical methods", A. Nota, G. Albi, S. Merino-Aceituno, M. Zanella Eds, *SEMA SIMAI Springer Series* (in press) (2020)

[48] Toscani, G.: Entropy-type inequalities for generalized Gamma densities. *Ric. Mat.* (in press) (2020)

[49] Van den Driessche, P.; Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci. 180*: 29–48 (2002)