SARS-CoV-2 infection dynamics in Denmark, February through October 2020: Nature of the past epidemic and how it may develop in the future

Steen Rasmussen1,2*, Michael Skytte Petersen3* and Niels Hoiby4,5*
1,* Center for Fundamental Living Technology (FLinT) Department for Physics, Chemistry and Pharmacy, University of Southern Denmark
2 Santa Fe Institute, Santa Fe, NM 87501, USA
3 Center for Biosecurity and Biopreparedness (CBB), Statens Serum Institut, Denmark
4 Department of Clinical Microbiology, Rigshospitalet, Copenhagen, Denmark
5 Institute of Immunology and Microbiology, Panum Institute, University of Copenhagen, Denmark

(Dated: November 4, 2020)

Background: There has long been uncertainty about the relative size of the "dark" numbers, the infected population sizes and the actual fatality rate in the COVID-19 pandemic and thus how the pandemic impacts the healthcare system. As a result it was initially predicted that the COVID-19 epidemic in Denmark would overwhelm the healthcare system and thus both the diagnosis and treatment of other hospital patients were compromised for an extended period.

Aim: To develop a robust method for reliable estimation of the epidemic and the healthcare system load in Denmark, both retrospectively and prospectively. To do this a new pandemic simulation had to be developed that accounts for the size and the infection impact of the infectious incubating and asymptomatic infected individuals (dark numbers).

Methods: Our epidemic simulation is based on a SEIRS (Susceptible - Exposed - Infected - Recovered - Susceptible) model, coupled to a simple healthcare model that also includes deaths outside hospital settings. The SEIRS model has separate assessments of asymptomatic and symptomatic cases with different immunological memories. The main data used for parameter estimation in the models are hospital and ICU occupations, death data, serological data of antibody prevalence from the onset through August 2020 together with hospital data and clinical data about the viral infection. Optimal model parameters are in part identified by Monte Carlo based Least Square Error methods while micro-outbreaks are modeled by noise and explored in Monte Carlo simulations. Estimates for the infected population sizes are obtained by using a quasi steady state method.

Results: The age adjusted antibody prevalence in the general population in May 2020 was 1.37%, which yields a relative frequency of symptomatic and asymptomatic cases of 1 to 5.2. Due to the large asymptomatic population found, the actual mortality rate to date is 0.4%. However, with no behavioral and policy restrictions the COVID-19 death toll would have more than doubled the national average yearly deaths within a year. The transmission rate $R_0$ was 5.4 in the initial free epidemic period, 0.4 in the lock-down period and 0.8 -1.0 in the successive re-opening periods through August 2020. The estimated infected population size July 15 to August 15 was 2,100 and 12,200 for October 1 - 20, 2020. The efficiency of the applied daily testing strategy for both periods is estimated to be 40% of the PCR observable infected. Of more theoretical interest we demonstrate how the critical infection parameters for COVID-19 are tightly related in a so-called iso-symptomatic infection diagram.

Conclusions: Our simulation may be useful if a major infection wave occurs in the winter season as it could make robust estimates both for the scale of an ongoing expanding epidemic and for the expected load on the healthcare system. Our simulation may also be useful to assess a future controlled epidemic, e.g. as a basis for evaluating different testing strategies based on estimated infected population sizes. Finally, we believe our simulation can be adjusted and scaled to other regions and countries, which we illustrate with Spain and the US.

Contents

I. Introduction 2

II. Empirical epidemiological data 3
   A. Main data source 3
   B. Asymptomatic versus symptomatic
      SARS-CoV-2 positive persons 3
   C. Immunological memory 4

III. The epidemic model system 4
   A. Infection parameters and relative frequency 6

IV. Methods 7
   A. $R_0$ for the SEIRS model 7
   B. Lockdown and re-opening of the country 8
   C. Monte Carlo optimization of parameters 8
   D. The impact of noise 9
   E. Quasi steady state approximation 9

V. Results 10
   A. Visual inspection for infection parameter fitting 10

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

* steen@sdu.dk, mspe@ssi.dk and hoiby@dadlnet.dk

© 2020 Rasmussen et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
COVID-19 was a new infection that initially hit the Chinese city Wuhan. Therefore it took some time before its severity and pandemic properties were realized by China, WHO, CDC and ECDC. The Chinese outbreak was reported to WHO December 31, 2019 and its Emergency Committee declared a Public Health Emergency of International Concern January 30, 2020 and a global pandemic situation March 11, 2020 [29]. The international spread of the COVID-19 pandemic was facilitated by the Chinese New Year celebration, which took place January 15 - February 11. During this period millions of Chinese people traveled inside China and from abroad e.g. Europe and North America to China to visit their families and then returned after the end of the New Year celebration [20]. Wuhan locked down January 23, 2020 and the rest of China subsequently followed that decision [44]. Many Chinese stay and work in large European cities. One of these cities is Milan, Lombardy, Italy south of the Alps. This was the first region in a western country to be heavily hit by the pandemic officially beginning February 21, 2020 [20] [1]. In February, many schools in Europe closed for one week for winter holiday over which thousands of adults and children traveled from all over Europe to the Alps for skiing.

The first case of COVID-19 reported in Denmark came from Northern Italy February 27, 2020 and subsequently 139 Danes came home after ski-holiday in Northern Italy and Austria, mostly from Ischgl [4], where they had contracted COVID-19 during the school holiday period, which in Denmark ended February 23. The epidemic in Denmark (5.82 mil. inhabitants) developed rapidly and then returned after the end of the New Year celebration [20]. Wuhan locked down January 23, 2020 and the rest of China subsequently followed that decision [44].

Many Chinese stay and work in large European cities. One of these cities is Milan, Lombardy, Italy south of the Alps. This was the first region in a western country to be heavily hit by the pandemic officially beginning February 21, 2020 [20] [1]. In February, many schools in Europe closed for one week for winter holiday over which thousands of adults and children traveled from all over Europe to the Alps for skiing.

The first case of COVID-19 reported in Denmark came from Northern Italy February 27, 2020 and subsequently 139 Danes came home after ski-holiday in Northern Italy and Austria, mostly from Ischgl [4], where they had contracted COVID-19 during the school holiday period, which in Denmark ended February 23. The epidemic in Denmark (5.82 mil. inhabitants) developed rapidly and then returned after the end of the New Year celebration [20]. Wuhan locked down January 23, 2020 and the rest of China subsequently followed that decision [44].

Many Chinese stay and work in large European cities. One of these cities is Milan, Lombardy, Italy south of the Alps. This was the first region in a western country to be heavily hit by the pandemic officially beginning February 21, 2020 [20] [1]. In February, many schools in Europe closed for one week for winter holiday over which thousands of adults and children traveled from all over Europe to the Alps for skiing.

The first case of COVID-19 reported in Denmark came from Northern Italy February 27, 2020 and subsequently 139 Danes came home after ski-holiday in Northern Italy and Austria, mostly from Ischgl [4], where they had contracted COVID-19 during the school holiday period, which in Denmark ended February 23. The epidemic in Denmark (5.82 mil. inhabitants) developed rapidly and then returned after the end of the New Year celebration [20]. Wuhan locked down January 23, 2020 and the rest of China subsequently followed that decision [44].
dramatically. The observed number of hospitalized patients already peaked April 7-8 (555 total hospital patients \(H_{tot} \), of which 146 were patients in ICUs). Thereafter the numbers gradually decreased.

After Easter a gradual re-opening of Denmark began April 20, 2020, and continued with fixed intervals of 2-4 weeks until June 8, 2020. Altogether, by August 31, 2020 there had been a total of 17,084 SARS-CoV-2 positively tested persons, 3,031 of whom were hospitalized (\(\sim 17.7\%)\) and 416 of those patients were treated in ICUs (\(\sim 13.7\%)\), while 625 infected died (\(\sim 3.7\%)\). As of August 31, 2020 there were 20 hospitalized including four in ICUs\[40\].

There were five local outbreaks in Denmark between April 10 and August 31, 2020, with the last two started around August 1, 2020. One outbreak was at a meat processing plant in the city of Ringsted where the infection was introduced by a Polish worker who then mainly infected other Polish workers. Approximately 150 workers from the meat processing plant were infected before the outbreak was eliminated by mid August. The other outbreak was in the city of Århus and was associated with a soccer game July 26th and associated celebrations, a university weekend seminar attended by 80 students, the Muslim Eid festival July 30 - August 3, and a funeral of a Somali rap-musician, which was attended by 500 people. Some of the infected unfortunately included bus drivers that spread the infection to neighboring regions. About 2/3 of the Århus cases were among people originating from Somalia and Palestine and the majority of the infected were young people. The highest incidence per 100,000 inhabitants in Århus during the outbreak was 99 and the outbreak was conquered by the end of August.\[54\].

Since we now have more information about the nature of COVID-19, and since there are still many unanswered questions regarding the future of the Danish epidemic, we have developed a mathematical model and a corresponding simulation of it. We calibrate the simulation using the official retrospective Danish hospital, ICU occupations, as well as non-hospital and total death toll numbers from the onset of the Danish epidemic through August 2020 that are published daily by SSI, Statens Serum Institut (SSI)\[54\], the Danish national center for disease control. The time-series of these data are shown in Fig 1.

We use additional data from SSI. These include the initial influx of newly infected individuals into Denmark in February 2020, recall discussion in previous section. We also use hospital admission data and the daily number of PCR positive individuals from both outside and inside of the hospital system, that are used to estimate the so-called daily transmission contact number, also called "contact number", \(R_t\)\[54\]. The \(R_t\)s measure the current changes in the infection trend and thus allow us to better understand the detailed (micro) properties of the empirical data after the initial infection peak.

It should be noted that SSI has modified minor details in the COVID-19 data multiple times during the period February through August 2020. However, these small changes have not had any measurable impact on our investigations or conclusions. All data are shown and further discussed in Appendix A.

### A. Main data source

To inform, test and validate our simulation we use the available data for the total hospital \(H_{tot}\) and ICU occupations, as well as non-hospital and total death toll numbers from the onset of the Danish epidemic through August 2020 that are published daily by SSI, Statens Serum Institut (SSI)\[54\], the Danish national center for disease control. The time-series of these data are shown in Appendix A.

### B. Asymptomatic versus symptomatic SARS-CoV-2 positive persons

Several publications have investigated the infectious potential of COVID-19 positive individuals and the ratio between asymptomatic persons and symptomatic patients. Of the 567 elderly SARS-CoV-2 positive passengers on the cruise ship Diamond Princess (median age 68 years), 46.5% were asymptomatic\[35\][51]. The young staff (median age 30 years) of the USS Navy carrier Theodore Roosevelt arriving in Guam was infected with COVID-19 and of the 382 surveyed service members 81.5% reported one or more symptoms whereas 18.5%
were asymptomatic [31]. The passengers and staff of a cruise ship departing from Ushuaia, Argentina on a 3 week cruise to Antarctic were infected with COVID-19, of which 128 passengers and crew tested positive for COVID-19. Only 19% were symptomatic whereas 81% were asymptomatic [21]. Based on these three carefully observed groups of patients, the percentage of symptomatic persons ranges from 19 to 81.5%. The percentage of asymptomatic persons ranges from 18.5 to 81% between these 3 groups. A meta-analysis estimated that about 40% of children are asymptomatic [56]. Obviously, there is a grey zone between asymptomatic and symptomatic infected where some authors differentiate asymptomatic and weakly symptomatic while others join these two groups into one class of asymptomatic and only view more severely symptomatic as being symptomatic.

Further, SARS-CoV-2 positive persons in the last part of their incubation time, before they turn symptomatic (thus still being pre-symptomatic), have been shown to harbor the same level of viral RNA in saliva and nasopharyngeal secretions as symptomatic COVID-19 patients and therefore are assumed to be nearly as infectious as symptomatic patients [36][29][13][47][7][25]. Note that this is not the case for incubating persons that later turn asymptomatic.

The mortality of COVID-19 is calculated in two ways: A Case Fatality Rate (CFR) which is the proportion of deaths from COVID-19 compared to the total number of individuals diagnosed with the disease for a particular period. The mortality is also calculated as Infection Fatality Rate (IFR) which is the proportion of deaths among all infected individuals, including symptomatic and asymptomatic and undiagnosed subjects (e.g. based on sero-prevalence studies).

C. Immunological memory

During COVID-19 infection an antibody response develops which involves IgM, IgG and IgA. At least 91 commercial antibody tests have been designed mostly using quantitative ELISA- or qualitative Rapid lateral flow immunoassays [42]. Nearly all patients have developed IgM, IgG and IgA antibodies 3 weeks after onset of the infection [50][55][46][8]. In one study of 70 Chinese patients, virus neutralizing antibodies correlated with antibodies measured by ELISA test [48], but in a study of 62 European patients only 66% developed protective antibodies as determined by virus neutralization tests [22]. The duration of the antibody response to SARS-CoV-2 is not yet known, but the experience from the related SARS-CoV infected patients is, that 94% of the patients still had detectable IgG antibodies after 1 year and 50-74% of the patients after 4 years [23][26]. After mild or asymptomatic MERS CoV infection, antibodies were either limited or rapidly declined within 3 months whereas the antibody response in severe infections would last for 2-3 years [23]. Preliminary results and calculations from China indicate similar dynamics of the SARS-CoV-2 antibody response [45][49].

III. THE EPIDEMIC MODEL SYSTEM

Our basic SEIRS (Susceptible - Exposed - Infectious - Recovered - Susceptible) style model assumes a single population that can be in one of three infectious states: incubation \( I_i \), asymptomatic \( I_a \), and symptomatic \( I_s \). The \( I_i \) state is analogous to the Exposed state in the classical SEIR model although infection is also possible from the \( I_i \) state in our model formulation. The population starts from being susceptible \( S \). After transitioning through incubation \( I_i \) into one of the two infectious states \( I_a \) and \( I_s \), the cases and ends up as either recovered \( R \) or dead \( D \). All three of the infectious states \( I_x, x = i, s, a \) can reduce the susceptible population \( S \) by transitioning members thereof into \( I_i \).

A small fraction of the recovered \( R \) slowly moves back to the susceptible population \( S \) as only a time limited immunity is assumed, which we refer to as the population’s average immunological memory. Since there likely is a significant difference in the immunological memory for the symptomatic and the asymptomatic infected we have disaggregated the recovered population into \( R_a \) and \( R_s \) with short and long immunological memory respectively. The rate of recovered population members once again becoming susceptible is given by \( \xi_y R_y, y = a, s \), which is reflected in the positive terms in the first expression in equation set (1).
During the infection, a fraction of the symptomatic population $I_s$ can become seriously ill and is either transferred to a hospital $H$ or an ICU. From the regular hospital unit $H$ patients can recover, move to an ICU, or die $D$. Patients from ICU can either move to $H$ or die $D$. Seriously ill individuals can also die at home or in an eldercare facility. These non-hospital deaths are indicated by $M$, where terminally ill patients arrive directly from $I_s$.

A flow diagram of the infection and healthcare dynamics is shown in Fig. 2.

The differential equations that define the flow in Fig. 2 are given in Eqs. 1 and 2.

The epidemic dynamics is defined by SEIRS model as follows:

$$\begin{align*}
\frac{dS}{dt} &= - (\beta_i I_i + \beta_a I_a + \beta_s I_s) S / N + \xi_r R_s + \xi_a R_a \\
\frac{dI_i}{dt} &= (\beta_i I_i + \beta_a I_a + \beta_s I_s) S / N - \gamma_i I_i \\
\frac{dI_a}{dt} &= (1 - \rho_s) \gamma_i I_i - \gamma_a I_a \\
\frac{dI_s}{dt} &= \rho_s \gamma_i I_i - \gamma_s I_s \\
\frac{dR_a}{dt} &= \gamma_a I_a - \xi_a R_a, \\
\frac{dR_s}{dt} &= (1 - h_{frac} - m_{frac}) \gamma_s I_s + \rho_{h,r} \gamma_h H - \xi_s R_s,
\end{align*}$$

where the parameters are further discussed in Table 1. Negative terms in the above equations indicate loss of respective population members with time, while positive terms indicate population member growth. The equation for the change in the susceptible population $S$ describes how susceptible are infected from three different populations $I_i, I_a, \text{ and } I_s$ with different rates $\beta_i, \beta_a, \text{ and } \beta_s$ normalized by the population size $N$, while recovered individuals slowly becomes susceptible again as they lose their immunity, where we assume longer immunity time $1/\xi_s$ days for the recovered symptomatic $R_s$ and a shorter immunity time $1/\xi_a$ days for the recovered asymptomatic $R_a$. The newly infected are moved from $S$ to $I_i$, the incubating population where they on average reside for $1/\gamma_i$ days. On average $1 - \rho_s$ of the incubated $I_i$ becomes asymptomatic $I_a$ where they on average reside for $1/\gamma_a$ days. Also, $\rho_s$ of the incubated $I_i$ becomes symptomatic $I_s$ where they on average reside for $1/\gamma_s$ days.

However, a small fraction $h_{frac}$ of the $I_s$ population becomes severely ill and are hospitalized, see Eqs. (2), while another small fraction $m_{frac}$ becomes so ill that they are not moved into the hospital system but stay at home or in a home care facility before dying, see Eqs. (2). All asymptomatic individuals $I_a$ recover and are moved into $R_a$ where we assume they retain some immunity for an average of $1/\xi_a$ days. Finally, most symptomatic infected recover $R_s$ directly from $I_s$, as well as from the hospital system $H$, see Eqs. (2), where we assume they retain some immunity for an average of $1/\xi_s$ days.

The dynamics of the hospital including its ICU and their deaths, as well as the non-hospital death dynamics are defined as follows:

$$\begin{align*}
\frac{dH}{dt} &= \alpha h_{frac} \gamma_s I_s + \rho_{h,cu,h} \gamma_{icu} ICU - \gamma_h H \\
\frac{dICU}{dt} &= (1 - \alpha) h_{frac} \gamma_s I_s + \rho_{h,cu,h} \gamma_{icu} ICU - \gamma_{icu} ICU \\
\frac{dD_{tot}}{dt} &= \rho_{h,d} \gamma_h H + (1 - \rho_{h,cu,h}) \gamma_{icu} ICU + \gamma_m M \\
\frac{dM}{dt} &= m_{frac} \gamma_s I_s - \gamma_m M \\
\frac{dD_m}{dt} &= \gamma_m M
\end{align*}$$
where the parameters are further discussed in Table I. Severely ill symptomatic individuals enter the hospital system at a rate given by $h_{frac}\gamma_{i}I_{s}$ where a fraction $\alpha h_{frac}\gamma_{i}I_{s}$ goes directly to the regular hospital unit while $(1-\alpha)h_{frac}\gamma_{i}I_{s}$ directly enters into an ICU. The size of $h_{frac}$ is critical for the scaling of the pandemic as the hospital data are the central empirical data by which the simulation is adjusted. We use the age aggregated data for hospital admissions from Ferguson et al. 2020 [18] together with the actual age distribution within Denmark to estimate an expected aggregated $h_{frac}$, see Appendix A for details. The regular hospital $H$ also receives a certain fraction $\rho_{icu,h}$ of the improved patients from the ICU as expressed by $\rho_{icu,h}\gamma_{icu}ICU$, while patients from $H$ leave the hospital after an average period of $1/\gamma_{h}$ days. Patients from $H$ either recover into $R_{s}$, or become more ill and move into ICU, or die and move into $D_{tot}$. The ICU population, as mentioned above, receives $(1-\alpha)h_{frac}\gamma_{i}I_{s}$ directly from $I_{s}$ as well as $\rho_{h,icu,h}\gamma_{h}H$ patients from $H$. ICU patients leave the population with the rate $\gamma_{icu}ICU$ and either improve and move to ICU with the already mentioned rate $\rho_{icu,h}\gamma_{icu}ICU$, or die with the rate $(1-\rho_{icu,h})\gamma_{icu}ICU$. The accumulated death toll stems from the hospital rate as $\rho_{h,d}\gamma_{h}$, from the ICU rate as $(1-\rho_{icu,h})\gamma_{icu}ICU$ and from the non-hospital/ICU associated rate as $\gamma_{m,M}$. Finally, the symptomatic population that eventually dies in a non-hospital location $M$ is given by $m_{frac}\gamma_{i}I_{s}$ with a sick period of an average $1/\gamma_{m}$ days where $m_{frac}$ is significantly smaller than $h_{frac}$. The non-hospital death tally are accounted for separately on a weekly basis in $D_{m}$, while they also are counted in the total death toll $D_{tot}$.

### A. Infection parameters and relative frequency of symptomatic and asymptomatic infected

With $\rho_{s}$ we express the relative frequency by which an individual moves from the incubation state to the symptomatic state, while $\rho_{a} = 1 - \rho_{s}$ expresses the frequency of movement to the asymptomatic state. We may now define the relation between the infection parameters $\beta_{i}, \beta_{a}, \beta_{s}$ in the SEIRS model as follows:

\[
\beta_{a} = \zeta \beta_{s} \tag{3}
\]

\[
\beta_{i} = (\rho_{s} \beta_{a} + (1-\rho_{s}) \beta_{a}) \tau \tag{4}
\]

\[
\beta_{i} = \beta_{s}(\rho_{s} + (1-\rho_{s})\zeta) \tau \tag{5}
\]

assuming that $\beta_{i}$ and $\beta_{a}$ can both be defined relative to $\beta_{s}$ and $0 < \zeta \leq 1.0$ since we assume the asymptomatic are always less infectious than the symptomatic infected. The value of $\beta_{i}$ in Eq. (4) is defined as the weighted sum of individuals that eventually become symptomatic and asymptomatic respectively, and where $\tau$ may be viewed as the fraction of the incubation time they are infectious. If we e.g. assume the incubating individuals are infectious the last 1.5 days (0.3) of the average 5 day incubation time [25] we get

\[
\beta_{i} = \beta_{s}(\rho_{s} + (1-\rho_{s})\zeta)0.3. \tag{6}
\]

We may use clinical data to estimate a reasonable $\zeta$ value. Clinical investigations indicate the viral load in saliva from symptomatic and asymptomatic is comparable and that the peak viral load in saliva is found in

<table>
<thead>
<tr>
<th>Parameter List</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
</tr>
<tr>
<td>$\beta_{i}$</td>
</tr>
<tr>
<td>$\beta_{a}$</td>
</tr>
<tr>
<td>$\beta_{s}$</td>
</tr>
<tr>
<td>$\gamma_{i}$</td>
</tr>
<tr>
<td>$\gamma_{a}$</td>
</tr>
<tr>
<td>$\gamma_{s}$</td>
</tr>
<tr>
<td>$\rho_{s}$</td>
</tr>
<tr>
<td>$\xi_{a}$</td>
</tr>
<tr>
<td>$\xi_{s}$</td>
</tr>
<tr>
<td>$h_{frac}$</td>
</tr>
<tr>
<td>$m_{frac}$</td>
</tr>
<tr>
<td>$\alpha$</td>
</tr>
<tr>
<td>$\rho_{h,r}$</td>
</tr>
<tr>
<td>$\rho_{h,icu}$</td>
</tr>
<tr>
<td>$\rho_{icu,h}$</td>
</tr>
<tr>
<td>$\rho_{h,d}$</td>
</tr>
<tr>
<td>$\gamma_{m}$</td>
</tr>
<tr>
<td>$\gamma_{h}$</td>
</tr>
<tr>
<td>$\gamma_{icu}$</td>
</tr>
</tbody>
</table>
last day of the incubation time [36] [29] [13] [47] [25]. Symptomatic children and adults have the same amount of SARS-CoV-2 in the upper respiratory tract [3].

From [28] it is further assumed that the relative infectiousness for the last day of pre-symptomatic (incubated that turns symptomatic) can be set to 1.00, while the average of the severe symptomatic, the weak symptomatic and the asymptomatic can be set to 0.89, 0.44 and 0.11 respectively during their infection time. In our study we do not distinguish between weak symptomatic and asymptomatic so it has the same form as Eq. (6) with the insertion of an extra factor \( \delta \) so that Eq. (9) is identical.

The mathematical models defined in Eqs. (1) and (2) presented in Fig. 2 and Eqs. (1) and (2). Further, by adding noise to the mean field approach defined in Eqs. (1) and (2), we can also interpret the impact of localized micro-outbreaks of the pandemic, as well as other fluctuations found in the empirical data. We investigate the impact of noise and compare the interpretation with the mean field approach without noise.

### IV. METHODS

The mathematical models defined in Eqs. (1) and (2) have a variety of parameters, where some are well defined, e.g. from clinical data, while most are restricted by the necessity for the simulations to reproduce the measured antibody data, the historical hospital and ICU occupation data, as well as the death data both from the hospital system and elsewhere.

The relationship between the infection parameters \( \beta_s, \beta_a \), and \( \beta_i \) is based on clinical studies, recall Eqs. (3) - (6) and Eqs. (7) - (9). The value of \( \rho_s \), the fraction of symptomatic infected, is mainly determined by the observed serological antibody response, which was conducted May 8 - 28, 2020, and published in early June 2020 [54], together with the average antibody decay-times \( 1/\xi_s, 1/\xi_s \) from the asymptomatic and symptomatic infected respectively.

The average antibody decay times \( 1/\xi_s, 1/\xi_s \) are estimated based on existing knowledge about SARS and other corona viruses, recall Section II C. Despite these restrictions a few parameter combinations still have some degrees of freedom, in particular \( h_{JraC} \) and \( m_{JraC} \) that scale the size of \( I_s \) and thus the whole pandemic. Also, it has been difficult to obtain explicit data for the fraction of the seriously ill that are directly admitted to a hospital \( H(\alpha) \) versus ICU (1- \( \alpha \)), as well as \( \rho_{h,icu} \) and \( \rho_{h,ICU} \), the fraction that are moved from a hospital to an ICU and the fraction that dies from a hospital without going to ICU respectively. We will further discuss all parameters in the Results and Discussion sections.

Besides parameter variations, the mathematical models in Eqs. (1) and (2) could be made more realistic by including more details, i.e. more detailed functional relationships and more state variables. Both a geographic disaggregation and an age disaggregation would be useful, but these model extensions are outside the scope of the current study.

Our current study seeks a better understanding of the relationships between \( \beta_s \) (and thus \( R_0 \), \( \xi_s, \xi_s \)) and the \( \xi_s, \xi_s \) pair for the COVID-19 pandemic. It also seeks to provide useful information about "What if?" scenarios, including potential infection wave scenarios assumed to follow the well-known winter influenza patterns. We base our investigations on the simplest model we believe is able to provide such insights. Such a model is presented in Fig. 2 and Eqs. (1) and (2). Further, by adding noise to the mean field approach defined in Eqs. (1) and (2), we can also interpret the impact of localized micro-outbreaks of the pandemic, as well as other fluctuations found in the empirical data. We investigate the impact of noise and compare the interpretation with the mean field approach without noise.

#### A. \( R_0 \) for the SEIRS model

The basic reproduction number \( R_0 \) can be derived from our SEIRS model system, recall Eqs. (1), and is given
the average infection period for each of the infected subpopulations, $I$, Here appropriate, estimate the numerical value for $R_0$ during the different stages of the Danish epidemic.

\[ R_0 = \frac{\beta_i}{\gamma_i} + \frac{(1 - \rho_s)\beta_a}{\gamma_a} + \rho_s\beta_s \gamma_s. \]  

Here $\beta_i$ is the average infection parameter for each of the infected subpopulations, $I$, $\gamma_i$ is one divided with the average infection period for each of the infected subpopulations and $\rho_s$ is the fraction of incubating that becomes symptomatic. In the following sections we shall, when appropriate, estimate the numerical value for $R_0$.

### B. Lockdown and re-opening of the country

The first documented COVID-19 infected arrived in Denmark in late February. On March 11, 2020 the Danish Prime minister announced that the country would be closing down in the following days, which meant that all non-essential professional activities would be shut down, social distancing (2 meters) and enhanced hand sanitation were imposed, only essential shopping was recommended (food and pharmacy), at most 10 individuals were allowed to gather at the same time, movement between different parts of the country was discouraged, international borders were closed, but face masks were not made mandatory. Further, any symptomatic individual with influenza-like symptoms were highly encouraged to self-isolate and seek medical advice and everybody who had been in contact with an infected were also highly encouraged to self-isolate.

We model the closedown process of the country as a sigmoid function for the infection parameters $\beta_s, \beta_a, \beta_i$ over a period of approximately 7 days. For all $\beta_x, x = s, a, i$, as defined in Eqs. (3) - (9) we get

\[ \beta_x(t) = \frac{\beta_x - d_x}{1 + e^{-\omega(t-t_0)}} + d_x \]  

where $\omega$ is the steepness of the sigmoid function, $t_0$ the middle of the approximately 7 days lockdown process, and $d_s, d_a, d_i$ the infection parameters for $\beta_s, \beta_a, \beta_i$ immediately after country shut down. Once it was realized that Denmark had been hit by the COVID-19 pandemic, all individuals with flu-like symptoms were highly encouraged to self-isolate, so in the following we assume $\beta_s$ remains at the low level $\sim 1\%$ of the free pandemic value, both during lockdown and after the country reopens.

We model the reopening of the country in a manner similar to the closedown process, i.e. by increasing $\beta_a$ and $\beta_i$ following the dates where the national reopening policies change. The reopening starts April 20 with the reopening of schools for the youngest kids (K-5) together with a number of small businesses including dentists, hairdressers and other businesses where services are rendered and where close physical proximity between a provider and a costumer is necessary. Later reopening activities were implemented May 7, 20, June 8, and August 14, 2020 [33].

We chose to model the reopening process as a slow (linear) increase of $\beta_a$ and $\beta_i$ over that interval. Thus, at time $date_1$ April 20, 2020, when the country starts reopening activities to $date_2$, June 8, 2020, a slow linear increase in $\beta_y, y = i, a$ occurs as defined below:

\[ \beta_y(t) = d_{y1} + \frac{d_{y2} - d_{y1}}{date_2-date_1} \times (t - date_1). \]  

At time $date_2$, June 8, 2020, this linear increase levels off and becomes flat with a constant $\beta_y^{reopen1}$ value. For the second reopening, August 14 - 31, 2020, a linear increase is again similarly defined leading to a constant $\beta_y^{reopen2}$ using the August 31, 2020 value.

A graphical depiction of the reopening dynamics as expressed in Eqs. (11) and (12) is shown in Fig. 3.

![Average β (infection) profiles over time](image)

**Figure 3:** The middle of the governmental lockdown process of the country occurs approximately March 16, 2020, 20 days after the arrival of the initial group of COVID-19 infected. The lockdown is modeled by a sigmoid function with the infection point at the middle of the process, see Eq. (11) and text for details. Successive governmental interventions that gradually reopen the country are approximated as a linear ramp starting April 20 and leveling off to a constant level June 8, 2020. The last reopening, August 14 - 31, 2020, is handled similarly, see Eqs. (11) and (12) and text for details.

### C. Monte Carlo optimization of parameters

We may adjust the $\beta$ parameters in Eqs. (1) and (2) manually while at the same time ensuring that the scale of the infection fits the data by adjusting $\rho_s$ that determines the relative sizes of the symptomatic and asymptomatic populations. From the national serological test conducted
May 2020, the estimated number of recovered individuals is reported to be about 1.3% of the population (May 28, 2020) although this study only includes individuals of age 18 or above.

The advantage of such an approach is that after some time experimenting with the simulation, the experimenter obtains an intuitive understanding of how the simulation reacts to parameter changes and minor parameter adjustments can be made iteratively.

Alternatively, we may use a Monte Carlo method (MC) to randomly choose the parameter combinations from a uniform distribution in given intervals in a given scenario. This is done until a desired minimum value of the sum of squared differences (LS) between the daily reported data for total Danish hospital occupations and the same data generated in the corresponding simulation is achieved. For more details, see Appendix C.

Using the MC-LS method we initially need to ensure the optimization process takes into account the appropriate scale of the outbreak that can be obtained by adjusting $\rho_s$. Thus we start the formal parameter optimization by initially looping over slightly different $\rho_s$ and different $\beta^{\text{lockdown}}$ and $\beta^{\text{reopen}}_a$ values followed by a two-step MC-LS process:

(i) In the first step we optimize the $\beta^0_i$s at the onset of the outbreak using the infection model given in Eqs. (7) - (9) and optimize for the period February 24 - April 20, 2020. We assume the lockdown $\beta^{\text{lockdown}}_i \simeq 0.01 \times \beta^0_i$ of the outbreak value once the lockdown in a place, as virtually all symptomatic individuals are isolated. We can now optimize for $\beta^{\text{lockdown}}_i$ (as well as $\beta^{\text{lockdown}}_a$ and $\beta^{\text{reopen}}_a$) in the free outbreak and lockdown period from February 24 to April 20 for different fixed $\beta^{\text{lockdown}}_a$ (and $\beta^{\text{lockdown}}_i$) values, where each optimization yields a slightly different $\beta^{\text{lockdown}}_i$. The lower $\beta^{\text{lockdown}}_a$ (and $\beta^{\text{lockdown}}_i$), the higher $\beta^{\text{lockdown}}_i$.

(ii) In the second step we optimize two parameters simultaneously, $\beta^{\text{lockdown}}_a$ and $\beta^{\text{reopen}}_a$ (as well as $\beta^{\text{lockdown}}_i$ and $\beta^{\text{reopen}}_i$) from April 20 to August 14 with the corresponding $\beta^{\text{lockdown}}_i$ obtained in step (i), also using the infection model given in Eqs. (7) - (9).

Going through the simulation with correct seroprevalence numbers for May 28, 2020, the MC solutions with the smallest LS error in the second step is then selected. Then the output from the MC-LS optimization yields the following set of five parameters: $\beta^0_i, \beta^{\text{lockdown}}_i, \beta^{\text{reopen}}_a$ and thus the associated $\beta^{\text{lockdown}}_a$ and $\beta^{\text{reopen}}_i$.

As we shall see in the Results Section, the MC-LS method tends to slightly underestimate $\beta^0_i$ compared to a visual inspection mainly because of the historical details of the data that are discussed in the next subsection.

D. The impact of noise

It is somewhat problematic using our mean field modeling approach to analyze and understand the successive, small and localized epidemic outbreaks over the late spring and summer of 2020. Although these outbreaks are small due to the low national hospital occupation numbers from May through August, most of these later micro-outbreaks are actually reflected in the hospital admittance numbers. Further, as we inspect the hospital admittance numbers for mid April 2020 they reveal a "shoulder" in the data that can be seen as one or more larger localized outbreaks overlaying a steadily decreasing macroscopic trend in the hospital admittance data, which is most clearly seen in Fig. 13(2), Appendix A. Both these earlier (larger) and later (smaller) fluctuations in the hospital data are also picked up in SSI’s daily estimated "contact number" $R_t$ that essentially measures the deviation from the currently observed number of infected compared to an average over that last week [54]. $R_t < 1$ indicates decreasing observed infected, $R_t > 1$ indicates increasing observed infected, while $R_t \sim 1$ means no change in the measured infection level. $R_t$ is estimated both from hospital admittance data and from the observed infected (positively PCR tested) individuals in the population at large.

When minimizing the least square error between the generated total hospital occupation $H + ICU$ in simulation and the reported total hospital occupation, we obscure these fluctuations in the observed data, which is unsatisfactory as we also hope to understand the smaller details in the pandemic, in particular the micro-outbreaks.

This issue can be circumvented by viewing the localized outbreaks as events generated by coincidences that can be implemented by adding a noise term to the equation set (1) that determines the dynamics for $I_i$ as follows:

$$dS/dt = -S(\beta_i I_i + \beta_a I_a + \beta_s I_s)/N + \xi_s R_s + \xi_a R_a - Y(A(t),f)$$

$$dI_i/dt = S(\beta_i I_i + \beta_a I_a + \beta_s I_s)/N - \gamma_i I_i + Y(A(t),f).$$

Here $Y(A(t),f)$ is a Poisson point process with frequency $f$ and amplitude $A(t)$, a local outbreak of size $A$ at time $t$ that on average occurs with frequency $f$. We have included $-Y(A(t),f)$ in Eq. (1a) to maintain population conservation by balancing out the same term in Eq. (1b).

The rest of the Eqs. (1) and (2) remain the same.

In the Results Section we review manual parameter adjustment with and without noise as well as the MC-LS parameter optimization without noise.

E. Quasi steady state approximation

For extended periods the hospital occupation may be approximately constant indicating a steady state situation of the pandemic. For example, we saw that during the period July - August, 2020. If we assume a quasi steady-
state pandemic situation we approximately have:
\[ \frac{dI_i}{dt} = \frac{dI_a}{dt} = \frac{dI_s}{dt} = \frac{dH_{tot}}{dt} = 0. \] (13)

We can then use our model to estimate the sizes of the infected background populations \( I_s, I_a, \) and \( I_i \) as we can calculate backwards from a known approximately constant hospital occupation. In a steady state situation we have \( B_{frac} \gamma_s I_s - (1 - \rho_{h,icu}) \gamma_h H - (1 - \rho_{h,icu,h}) \gamma_{icu} ICU = 0 \) that expresses: what goes into the hospital system equals what leaves the hospital system. With known parameters \( B_{frac}, \rho_{h,icu}, \gamma_h \) and known \( H \) and ICU populations the size of \( I_s \) can be estimated. Then \( I_a = ((1 - \rho_s)/\rho_s) I_s \) and \( I_i = (I_a + I_s)/2 \) as the average incubation time is half the average time of both the symptomatic and asymptomatic.

V. RESULTS

A. Visual inspection for infection parameter fitting

The typical infection dynamics generated by the SEIRS model as defined by the equation system (1) are discussed in Fig. 4 while the corresponding hospital and non-hospital dynamics defined by equation systems (2) are discussed in Fig. 5. For this simulation, the fraction of symptomatic infected is \( \rho_s = 0.16 \) and the infection model is defined by Eqs. (7) - (9), where \( \beta_s^{lockdown} = \beta_s^{reopen1} = 0.01 \times \beta_0^{s}, \beta_s^{lockdown} = 0.1019 \times \beta_0^{s}, \beta_s^{reopen1} = 0.2305 \times \beta_0^{s} \). These parameters are selected by visual inspection of simulation output ensuring an infection scale such that \( \sim 79,700 \) individuals had been infected by May 28, 2020 [54]. Other parameter values used are shown in Table I.

It is noted that the large size of the infected total used above is because we also include the asymptomatic infected, which are estimated to be \( \sim 5.2 \) times larger than the symptomatic infected. Initially only the symptomatic were tested and this is still reflected in the verified positive PCR test numbers that are commonly quoted.

B. Monte Carlo Least Squares optimization of infection parameters

Alternatively, we can utilize a Monte Carlo based Least Squares (MC-LS) optimization method to identify the different \( \beta_x \) where \( x = s, a, i \) as discussed in Section IV C. The result of the simulation is shown in Fig. 6. The MC-LS parameter optimization (8,000 iterations) yields (assuming \( \beta_s^{lockdown} = \beta_s^{reopen1} = \beta_s^{reopen2} = 0.01 \times \beta_0^{s} \), \( \rho_s = 0.1570 \) and \( \beta_0^{s} = 1.0583 \), \( \beta_s^{lockdown} = 0.1366 \times \beta_0^{s}, \beta_s^{reopen1} = 0.1907 \times \beta_0^{s}, \beta_s^{reopen2} = 0.2717 \times \beta_0^{s} \) using the same infection model as above defined by Eqs. (7) - (9). Also the death rates have to be slightly increased to

Figure 4: Infection dynamics; upper panel linear scale, lower panel logarithmic scale. We assume the initial infection \( I_i \) stems from 690 individuals entering Denmark February 24, 2020. Parameter adjustments were done by visual inspection to match the observed hospital and death data in Fig. 5, as well as the number of persons with COVID-19 antibodies \( \sim 1.37\% \) of the population, \( \sim 79,700 \) recovered on May 28, 2020 (day 95 of the simulation). The governmental interventions, a lockdown and a successive reopening of the country, are discussed in Fig. 3. In this simulation \( \rho_s = 0.16 \), which means that a little more than 5 out of 6 infected individuals end up being asymptomatic, and we use the relationships between the \( \beta_x \)’s given in Eqs. (7) - (9), \( \beta_s^{lockdown} = \beta_s^{reopen1} = \beta_s^{reopen2} = 0.01 \times \beta_0^{s}, \beta_s^{lockdown} = 0.1019 \times \beta_0^{s}, \beta_s^{reopen1} = 0.2305 \times \beta_0^{s} \). The immunity loss parameters for asymptomatic infected are \( 1/\xi_s = 60 \) days and for symptomatic infected \( 1/\xi_s = 700 \) days. Health care parameters: \( \alpha = 0.85, \gamma_h = 1/8, \rho_{h,cr} = 0.775, \rho_{h,icu} = 0.135, \rho_{h,ad} = 0.09, \gamma_{icu} = 1/10, \rho_{icu,h} = 0.72, m_{frac} = 0.0081, \gamma_m = 1/8 \). These are the standard parameters shown in Table I. The central relationship between \( \xi, \rho_s, \beta_s \) and \( R_0 \) is discussed in Fig. 8.
Figure 5: Dynamics of total daily hospital occupation $H_{tot}$, ICU occupation, total death $D_{tot}$ as well as weekly non-hospital deaths $D_{nh}$ corresponding to (same simulation as) infection dynamics in Fig. 4; upper panel linear scale; lower panel logarithmic scale. Dots represent the observed hospital and death data, lines represent the simulation data. Parameter adjustments were done by visual inspection to match observed data. Health care parameters used are $\alpha = 0.85$, $\gamma_h = 1/8$, $\rho_{h,r} = 0.775$, $\rho_{h,icu} = 0.139$, $\rho_{h,d} = 0.09$, $\gamma_{icu} = 1/10$, $\rho_{icu,h} = 0.72$, $\mu_{frac} = 0.0081$, $\gamma_m = 1/8$, which are the standard parameters discussed in Table 1. These hospital and death times series data together with the hospital parameters define our main means of verification (or falsification) of our simulations. The central relationship between $\zeta$, $\rho_s$, $\beta_s$ and $R_0$ is discussed in Fig. 8.

Figure 6: Dynamics of total daily hospital occupation $H_{tot}$, ICU occupation and total death $D_{tot}$ dynamics as well as weekly non-hospital deaths $D_{nh}$, where the parameter adjustments are done by 8,000 iterations of MC-LS optimization. Upper panel linear scale and lower panel logarithmic scale. Dots represent the observed hospital and death data, lines represent the simulation data. See text for parameter details and compare with Fig. 5. Note how the MC-LS algorithm optimized for best fit of $H_{tot}$ chooses a slightly lower $\beta_0 = 1.0583$ because it minimizes the sum of the squares of the errors, which means that it misses the full amplitude of initial infection peak. A slightly lower $\beta_0$ implies a slightly lower $\rho_s = 0.157$ (relatively more asymptomatic infected needed to reach the measured seroprevalence level per May 28, 2020). Further, a lower $\beta_0$ implies a higher $\beta_{lockdown}^a$ level (March 17 - April 20, 2020) and a relatively higher $\beta_{lockdown}^a$ level implies a relatively lower $\beta_{reopen}^a$ level (after June 8 and before August 15, 2020). Finally, a lower $\beta_0$ requires a slightly higher death rates to match reported data so $\rho_{h,r} = 0.775 \rightarrow 0.755$; $\rho_{h,icu} = 0.135 \rightarrow 0.15$; $\rho_{h,d} = 0.09 \rightarrow 0.095$; and $\mu_{frac} = 0.0081 \rightarrow 0.0083$. 

fit the data, see discussion in Fig. 6. Other parameter values used are shown in Table I, except the parameters controlling the death rates both from the hospitals and the non-hospital locations.

The hospital occupation data has a "shoulder" after the peak (see Appendix A and Fig. 13(2)). This means there is an extended period in mid April 2020 with close
to constant hospital admissions. Thus the MC-LS algorithm selects a $\beta_s^0$ that underestimates the observed impact on the health care system as the LS error becomes smallest if the simulated hospital curve is close to as many observed data points as possible. Therefore the MC-LS estimate misses the full scale of the hospital occupation peak. As a consequence of the initial underestimation of $\beta_s^0 = 1.0583$, the MC-LS algorithm appropriately compensates by slightly increasing the estimate for $\beta_s^{lockdown} = 0.1366 \times \beta_s^0$, again to minimize the least square error between reported and generated data. Compare Figs. 5 and 6.

C. Impact of noise

Yet another way to interpret the Danish COVID-19 pandemic data is to view the macroscopic infection dynamics as described by the SEIRS model and the microscopic events as superimposed noise. Thus the noise can be interpreted as microscopic, localized infection events, as discussed in Section IV D.

We may assume the impact of noise in the hospital data only becomes visible around April 1, as singular microscopic events would be difficult to distinguish earlier because of the dominating macroscopic growth of the pandemic given by $\beta_s^0$ and the successive macroscopic disruption of the infection dynamics due to the sudden lockdown of the whole country. Inspecting the data from April 1 to August 31 supported by $R_t$ [54] measured from both hospital admissions and observed infected numbers, yields five distinct peaks corresponding to microscopic events, see Fig. 13(5 - 6) Appendix A. Thus we may estimate $f$ as 5 events/150 days = 0.033 events/day.

Now recall Eqs. (1a) and (1b). For April 1 - 20, 2020, we assume $A(t) = 2,500$ based on the size of shoulder in the hospital admittance for mid to late April; recall Fig. 13(5 - 6). $A(t)$ is thus a bit smaller than 10% of the average size of the total infected population $I(t) = I_s + I_o + I_a$ in that time interval according to our simulations. For April 21 - August 31, 2020, we assume $A(t) = 1,000$ as there are no more major "shoulders" in the data.

Thereby $A(t)$ becomes approximately 1/3 of the size of the total infected population from early July through August, 2020 according to our simulations.

As a basis for the noise added simulation we in part use the standard parameters previously used to generate Figs. 4 and 5: $\rho_s = 0.16, \beta_s^0 = 1.09$ and $\beta_s^{lockdown} = \beta_s^{reopen 1} = 0.01 \times \beta_s^0$. However, the reopening parameters have to be lower $\beta_s^{reopen 1} = 0.175 \times \beta_s^0$ and $\beta_s^{reopen 1} = 0.175 \times \beta_s^0$. These $\beta_s$ values are lower so that they, together with the noise added infections, on average generate a similar total number of infections as in the reported data. For simplicity we do not include the last reopening August 14 - 31, 2020, as it does not impact the reported data. Fig. 7 shows a comparison between the observed data and data generated by 100 Monte Carlo simulations each with a different noise sequence realization. Two such noise induced simulations are also highlighted in Fig. 7 to give an impression of the trajectories of individual realizations. Compare Fig. 7 with Figs. 5 and 6.

VI. SYMPTOMATIC VERSUS ASYMPTOMATIC POPULATIONS

In the current study we define the symptomatic infected population as consisting of individuals with serious symptoms, while the asymptomatic population for simplicity includes individuals with no symptoms as well as weak symptoms, e.g. minor symptoms from upper airways (nose, throat) and no fever. Thus, in our model a symptomatic infected individual knows that he or she is sick and will likely seek medical advice as well as self-isolate, while an asymptomatic individual likely would continue his or her daily routines at home, go to school or go to work as well as engage in social activities. In some model studies asymptomatic and weakly symptomatic are treated separately [28].

Because historical total hospital $H_{tot}$ and ICU occupations as well as death data only stem from the symptomatic infected population $I_s$, and we assume a constant fraction of seriously ill individuals are in need of hospitalization [18], the symptomatic population size is at any given time constrained by these real life data. The same is true about the non-hospital deaths. In contrast, the asymptomatic population could in principle vary freely if we didn’t include additional empirical information.

Randomized serological tests for SARS-CoV-2 antibodies in a population yields an estimate of the recovered, both symptomatic and asymptomatic, individuals at any give time. These recovery numbers can then be compared to the recovered population generated in simulation. The size of the recovered populations at any given time depends on both the number of recovered infected individuals and on the decay rate of the average immunological memory of symptomatic $\xi_s$ and asymptomatic $\xi_a$ individuals, recall Eqs. (1). This means that with only one serological measurement at $t_1$ infinitely many combinations of $\rho_s$ and $\xi_s, \xi_a$ fit the observed serological prevalence. In principle we need two serological samples at times $t_1$ and $t_2$, with some months in between, to fix the parameter values for $\rho_s$ and the average of the two $\xi’s$. Thus, measuring the recovered population at two different times $t_1$ and $t_2$ fixes the average $\xi_s$ which then can be expressed in days.

Alternatively, we can estimate the immunological memory based on our knowledge of other corona viruses, recall Section II C, and use these as parameter inputs to our simulations. Assuming a half-life for $1/\xi_s$ to be 700 days and $1/\xi_a$ to be 60 days respectively, we can deter-
mine $\rho_s$ and thus the relative frequency of symptomatic versus asymptomatic individuals from only one randomized serological test at time $t_1$ [49].

The Danish SARS-CoV-2 seroprevalence was found to be 34 out of 2,424 tested $\sim 1.4\%$ of the Danish population during the period May 8 - 28, 2020 [54], where it should be noted that this seroprevalence study only includes adults of more than 18 years of age. However, about 1/3 of the Danish population has been PCR-tested for SARS-CoV-2 virus by the end of August 2020 and the infection prevalence was found to be $\sim 1.1\%$ in adults and $\sim 0.6\%$ in children [54]. Since there are $\sim 1.30$ mil. 0-19 aged and $\sim 4.52$ mil. 20-90+ aged in Denmark we estimate $0.6 \times (1.30/5.82) \times 1.4\% + 1.1 \times (4.52/5.82) \times 1.4\% = 1.3692\% \sim 1.37\%$, where we assume the age distribution of infection prevalence is similar to the age distribution in seroprevalence.

We now need to adjust $\rho_s$ to obtain $\sim 0.0137 \times 5.82 \times 10^6 \sim 79,790$ recovered individuals in our simulation by the end of May, 2020. This yields a $\rho_s \simeq 0.16$, which means that we estimate that approximately 16% of the infected are symptomatic while 84% of the infected are asymptomatic; thus $\sim 5.2$ times more asymptomatic than symptomatic infected individuals. Also, as Denmark only has conducted one prevalence study so far, our estimated scale of the pandemic significantly relies on this one measurement.

### A. Iso-symptomatic infection diagram

The parameters $\zeta$: the relative infectiousness of asymptomatic versus symptomatic; $\rho_s$: the relative fraction of infected (incubated) that becomes symptomatic; $\beta_0$: the infection parameter; and the initial $R_0$, define critical properties of COVID-19 and other communicable infectious diseases. In Fig. 8 shows the relationship that exists between $\zeta$, $\beta_0$, $R_0$, and $\rho_s$, which we may call an iso-symptomatic infection diagram. We assume $\zeta$, $\rho_s$, $\beta_0$: constant over time and throughout a free (non-restricted) pandemic. Thus $\beta_s = \beta_0$ and also the initial $R_0$ are defined for the initial free pandemic period.

The iso-symptomatic infection diagram is constructed from multiple Monte Carlo - Least Square parameter optimized simulations that all match the observed Danish COVID-19 hospital occupation data at the onset of the pandemic. In this construction we use the more generic infection model with variable $\zeta$ as defined in Eqs. (3) - (6).

Quantitatively almost identical iso-symptomatic infection diagrams are found when we optimize the involve parameters not only for the initial free pandemic, but also include data through April 19 (lockdown) or include data through August 14, 2020 (after the first reopening).

Note that in Fig. 8 the "+" curves for $R_0$ are virtually horizontal, which means that $\zeta$ and $\beta_0$ balance each other.

---

Figure 7: 100 Monte Carlo realizations of the epidemic with noise added micro-outbreaks with a frequency of 0.033 events/day and an amplitude of size $A(t) = 2,500$ for April 1 - 20, and 1,000 after April 21, 2020. Upper panel: linear scale dynamics of total daily hospital $H_{tot}$ and ICU occupations together with total daily deaths $D_{tot}$ as well as weekly non-hospital deaths $D_{nh}$. Dots represent the observed hospital and death data, lines represent the simulation data. Middle panel: logarithmic scale dynamics of the same observables. Lower panel: two realizations of the noise induced impact of local outbreaks on $I_i(t)$, $I_a(t)$, $I_s(t)$ (logarithmic scale). See text for details.
The best parameter combination for describing the Danish COVID-19 epidemic given by Eqs. (7) - (9). In another location of the Danish Kingdom, the Faeroe Islands (52, 484 inhabitants), testing was far more extensive from the beginning of the epidemic (109, 233 tests out; as \( \zeta \) increases \( \beta_s^0 \) decreases and vice versa. This can be seen by substituting \( \zeta \) into the expression of \( R_0 \) in Eq. (10) that after reducing the expression yields \( R_0 = 11.5 \times \beta_s^0 (\zeta (1 - \rho_s) + \rho_s) \) where we have also inserted \( \gamma_a = \gamma_s = 0.5 \times \gamma_t = 0.1 \).

Since all simulations in Fig. 8 approximate the observed initial pandemic we need to empirically identify at least two of the three parameters \( \zeta, \beta_s, \rho_s \) to know which parameter combination most likely represents the Danish epidemic. We have previously argued for \( \zeta = 0.309 \) in the infection model given by Eqs. (7) - (9).

If we fix \( \zeta = 0.309 \) for the infection model in Eqs. (3) - (6) and use visual inspection we obtain a slightly lower \( \beta_s^0 \simeq 1.08 \) and a higher \( \rho_s \simeq 0.185 \) to match the initial infection peak in the reported \( H_{\text{tot}} \) data as well as the infection scale measured by May 28, 2020 (population wide antibody prevalence). However, with these parameters the later observed data requires a bit higher \( \beta_s \).

If we instead fix \( \beta_s^0 = 1.09 \) and use visual inspection we obtain \( \rho_s \simeq 0.18 \) and a slightly higher \( \zeta \simeq 0.31 \) that both approximate the observed \( H_{\text{tot}} \) data and the observed May 28, 2020 population wide antibody prevalence.

Finally, if we fix \( \rho_s = 0.16 \) it is becomes more difficult to simultaneously fit the observed \( H_{\text{tot}} \) and antibody prevalence as of May 28, 2020. We either obtain a good seroprevalence fit and too low an initial infection wave or a good fit for the infection wave and too high a seroprevalence fit.

The above investigations with different fixed \( \zeta, \beta_s^0, \rho_s \) parameters underscore that even though it is always possible to identify a best MC-LS fit to \( H_{\text{tot}} \) it might not represent the actual pandemic as it misses the seroprevalence observation - and possibly other observables.

 Iso-symptomatic infection diagrams could be constructed readily for most infectious diseases. With such diagrams available for the medical community, a single diagram would have crucial and quantitative information about critical parameters that define the dynamic characteristics of a given epidemic. Further, it would also at a glance make comparisons between different infectious diseases easier. Similar diagrams where critical thermodynamic properties of materials are depicted have had tremendous practical importance for the development of engineering since the early days of the industrial revolution.

**B. Estimation of mortality rates of the Danish COVID-19 epidemic**

The estimated Infection Fatality Rate (IFR) as of August 31, 2020, based on the simulations that include asymptomatic cases, can be calculated by dividing the total number of deaths with the total number of recovered: 628/173,544 = 0.0036 or \( \sim 0.4\% \).

The official calculations of mortality of COVID-19 in Denmark in the beginning of the pandemic was based on hospitalized patients. The value found to be 3.3% April 3, 2020, just before Easter, and 4.8% June 19, 2020, when the testing strategy had been changed for about 2 months to also include non-hospitalized individuals. However, asymptomatic cases of COVID-19 were not included in these calculations and therefore the numbers were Case Fatality Rates (CFR). Thus, according to our calculations the true mortality (Infection Fatality Rate) of COVID-19 in Denmark has been overestimated by a factor of approximately 10.

In another location of the Danish Kingdom, the Faeroe Islands (52, 484 inhabitants), testing was far more extensive from the beginning of the epidemic (109, 233 tests...
as of September 14, 2020) to detect and isolate all contacts. They have recorded 423 COVID-19 cases (0.8% including asymptomatic cases) and no deaths (both IFR and CFR) (0% - 95% c.i. 0.0 - 0.9%) [54]. Note that our estimated IFR of 0.4% in Denmark is in the middle of their confidence interval.

In Iceland (364,260 inhabitants), a serological study estimated that 0.9% of the population had been infected with SARS-CoV-2 and that the IFR was 0.3%, which is again similar to our estimates [19].

C. Estimation of the $R_0$ for the different phases of the Danish COVID-19 epidemic

$R_0$ can be estimated for the different phases of the epidemic using Eq. (10) if we adjust the $\beta_x(t), x = i, a, s$ values appropriately. For standard simulation parameters and the simulation started on February 24, 2020 with 690 incubated individuals, we estimate $R_0 = 5.361 \pm 5.4$ for $\beta_0^i \approx 1.09$ in the free pandemic until the lockdown March 16, 2020, recall Figs. 4 and 5. Using a bit lower $\beta_0^i \approx 1.06$ instead for the MC-LS optimized parameters yields $R_0 = 5.213 \approx 5.2$. During the lockdown and the successive re-opening of the country we assume symptomatic infected individuals are effectively isolated to about 1% of the initial $\beta_0^i$ infection parameter so the majority of the infection transmission occurs due to asymptomatic and incubating individuals. We can derive $R_{\text{lockdown}}$ during the lockdown March 16 - April 20, 2020, and $R_{\text{reopen}}$ after the reopening June 8 - August 14, 2020, with infection assumptions either as in Eqs. (3) - (6) or (7) - (9). Using the standard parameters for the simulation discussed in Figs 4 and 5 based on the infection model from Eqs. (7) - (9), we have

$$R_{\text{lockdown}}^0 = \frac{0.1019\beta_0^i}{\gamma_i} + \frac{(1 - \rho_s)0.1019\beta_0^a}{\gamma_a} + \frac{\rho_s0.013\beta_0^s}{\gamma_s}$$

$$= 0.386 \approx 0.4 \quad \text{and}$$

$$R_{\text{reopen}}^0 = \frac{0.2305\beta_0^i}{\gamma_i} + \frac{(1 - \rho_s)0.2305\beta_0^a}{\gamma_a} + \frac{\rho_s0.013\beta_0^s}{\gamma_s}$$

$$= 0.8511 \approx 0.9. \quad (14)$$

An additional Danish reopening phase occurred over the period August 14 to August 31, 2020, that most significantly allowed physical attendance at higher educational institutions and longer opening times for pubs and restaurants [53]. This last re-opening, however, was partly rolled back in the first half of September due to a significant infection increase.

An expanding epidemic requires $R_0 > 1$. For this to occur after the reopening of the country, the increased infection pressure must almost solely come from the asymptomatic and incubating populations as the symptomatic population is assumed to be almost completely isolated, i.e. $\beta_s = 0.01 \times \beta_0^s$. Taking this value and setting

$$R_{\text{reopen}}^1 = 1.0$$

in the above equation, we can solve for the factor to multiply $\beta_a$ and $\beta_i$ by for the steady state reopening case and find the multiplier to be 0.2717.

From mid July to mid August 2020 the noise induced simulations generate a better approximation to the low-level steady state background infection that the observed data indicate. Neither of the deterministic simulations with parameters adjusted either by visual inspection nor by Monte Carlo Least Square optimization are able to capture the balance between the frequent micro-outbreaks and the otherwise slightly decreasing pandemic, compare Figs. 5, 6 and 7.

An approximate steady state is indicated in the following observables:

(i) a fluctuating low level of hospitalized ($H_{\text{tot}} \sim 20$) and

(ii) a close to stationary fluctuating number of observed infected (positively PCR tested) individuals in the population at large. The infection numbers fluctuate with an average of 85/day currently based on 20,000 - 40,000 tests/day or 0.3 - 0.6% of the population/day, if we average across the re-occurring micro-outbreaks.

(iii) The two contact numbers $R_0$ fluctuate around 1.0 estimated from the hospital admissions and the national testing program respectively.

We may thus view this situation as a low-level steady state for the background pandemic, with re-occurring, localized, micro-outbreaks, counteracted in part by testing and contact tracing operations. This means that Eq. (16) should modified to

$$R_{\text{reopen}}^0 \approx 1.0 \quad (16a)$$

for July 15 to August 15, 2020 that indicates a quasi steady state epidemic.

VII. ESTIMATION OF INFECTED POPULATION SIZES AND TESTING EFFICIENCIES

A. Quasi steady state approximation for populations sizes

If we assume a steady-state pandemic situation we again have the relations given by Eq. (13). We can then use our model to estimate the sizes of the infected background populations $I_i, I_a,$ and $I_s$ during a quasi steady state period. It should be noted that the above mentioned steady state equation system is neutrally stable, which means that any perturbation to a solution we pick is also a solution, because the equation system is degenerate (determinant = 0). Thus, the system has infinitely many solutions all with an appropriate balance between the different infected populations and thus the resulting hospitalized patients.
Recall the hospitalization fraction $h_{frac} \simeq 8.2\%$ (Appendix B and [18]), which is the percentage of severely symptomatic infected that need hospital care, recall Eqs. (1) and (2). We can equate $I_s \times \gamma_s \times h_{frac}$ with the average in-rate to the hospital system consisting of $H_{tot} = H + ICU$. Recall steady state where $dH/dt + dICU/dt = 0$ that can then be written out is

$$0 = h_{frac}\gamma_s I_s - (1 - \rho_{h,icu})\gamma_h H$$

$$- (1 - \rho_{icu,h})\gamma_{icu} ICU \Leftrightarrow$$

$$I_s = \frac{(1 - \rho_{h,icu})\gamma_h H + (1 - \rho_{icu,h})\gamma_{icu} ICU}{h_{frac}\gamma_s}$$

$$I_s = \frac{(1 - \rho_{h,icu})\gamma_h H_{tot}k + (1 - \rho_{icu,h})\gamma_{icu} H_{tot}(1 - k)}{\gamma_h h_{frac}}$$

where in the last equation we have expressed $H$ and ICU as a function of $H_{tot}$ where $k$ is defined as $H/H_{tot}$. Inserting the parameter values from Table I we now have an estimate for the symptomatic population $I_s$. From this we can also estimate the other infected populations as:

$$I_a = ((1 - \rho_s)/\rho_a)I_s$$

$$I_i = (I_a + I_s)/2,$$

where the last relationship follows from $\gamma_a = \gamma_s = 0.5\gamma_i$. Thus, the size of the total infected population $I_{tot}$ is:

$$I_{tot} = I_i + I_a + I_s.$$  (19)

Currently PCR tests should be able to pick up SARS-CoV-2 RNA starting the last 1.5 days of the ~5 day incubation period and through the ~10 day period where the infected individuals are either symptomatic or asymptomatic.

Still assuming $\gamma_a = \gamma_s = 0.1$/day and no PCR detection after the 10 day infection period associated with $I_a$ and $I_s$, the detectable steady state population is approximately

$$I_{obs} = 0.3 \times I_i + I_a + I_s.$$  (20)

as $I_i$ on average can only be identified the last 1.5 days of the 5 day incubation time; thus the factor 0.3. It should be noted, however, that RNA virus can be excreted and be picked up by PCR measurements significantly longer than the approximately 10 days we have assumed as the duration of the infection both for the symptomatic and asymptomatic cases [6].

From Eqs. (1) for the steady state case we note that

$$\frac{dI_i}{dt} = (\beta_i I_i + \beta_a I_a + \beta_s I_s)S/N - I_i\gamma_i = 0.$$  (21)

In steady state this means that the two terms should be identical; what goes into $I_i$ is equal to what goes out of $I_i$. Thus, we can estimate the daily rate at which newly infected are recruited from the susceptible population as

$$I_{new} = I_i\gamma_i.$$  (22)

Based on hospitalization data and using Eqs. (16) to (22), we can estimate the corresponding infected and observable populations sizes as well as the daily infection rate, none of which are directly observable.

Note the linear proportionality between the numbers; e.g. 10 times higher hospital population yields 10 times higher infected populations and infection rates. Also note that these estimates may be impacted by testing and removal of identified infected individuals so their ability to further spread the infection is reduced. Finally, it should be noted that quasi steady state approximations are often successfully used even when true steady state conditions do not yet apply.

During the quasi steady state period from mid July to mid August, 2020, Denmark on average had 20 hospitalized COVID-19 patients ($H_{tot}$) where the intensive care unit patients are approximately one fifth of every patient, $ICU = 0.2 \times H_{tot} = (1 - k) \times H_{tot}$. Inserting these numbers and default values from Table I into Eqs. (16) to (20) and (22) we obtain the following population infection numbers:

$$I_s \simeq \frac{(1 - 0.135)200}{8}0.8 + (1 - 0.72)200\frac{1}{0.082}$$

$$\simeq 225$$

$$I_a \simeq (0.84/0.16)225 \simeq 1179$$

$$I_i \simeq (225 + 1179)/2 \simeq 702$$

$$I_{tot} \simeq 225 + 1179 + 702 \simeq 2106$$  (23)

$$I_{obs} \simeq 225 + 1179 + 0.3 \times 702 \simeq 1615$$

$$I_{new} \simeq 702 \times 0.2 \simeq 140.$$  

The reopening of the country August 14 - 31, 2020 unfortunately caused a new significant infection rise with an increase in the hospitalizations from early to late September. The infection rise initially had a doubling time of approximately 8 days and by the end of September, 2020 the hospital occupation largely leveled off around 110 due to a number of national initiatives including renewed information campaigns and reinstated restrictions including early closing of bars (10 pm) and at most 50 people at private events. If we assume a quasi steady state for October 1 - 20, 2020 we get from [52] the averages $H_{tot} = 111, k = 0.85$ and $T_d = 407$ that inserted in the Eqs. (16) - (22) yield:

$$I_i \simeq 4066$$

$$I_a \simeq 6830,$$

$$I_s \simeq 1301,$$

$$I_{tot} \simeq 12196,$$

$$I_{obs} \simeq 9351,$$

$$I_{new} \simeq 813.$$  (24)

The key population numbers for different $H_{tot}$ population sizes are tabulated in Table II (all numbers are calculated
on a computer with double precision and then rounded to integer values).

**B. Testing efficiencies**

From mid July to mid August, 2020 the nationwide voluntary testing program together with medically recommended testing and contact tracing on average identify and isolate $T_d \sim 85$ infected/day from the contagious populations [54]. By comparing the number of identified infected $T_d$ with the number observable contagious individuals $I_{obs}$ gives a simple indicator of the testing efficiency. We may define a single day detection efficiency as the detected within a day divided with the estimated number of infected on that same day: $DT_{abs} = T_d / I_{obs} = 85/1615 = 0.0526$ or approximately 5.3%. So in a theoretical situation where every individual in the country could be tested in a single day, all observably infected would be identified in one day with a total efficiency of 100%.

However, from an operational point of view a more relevant efficiency number is $DT_{daily}$ the net daily detected and isolated compared to the daily newly infected defined $I_{new}$. Obviously, the impact of the detection and isolation critically depends on which time point in an infected individual’s disease process it occurs. Since an explicit modeling of the infection and contact tracing processes is outside of the scope of our current study, we will only provide upper and lower bounds for the impact of the infection and contact tracing processes.

The impact would be maximal if the detection and isolation of all individuals occur during the incubation period and before the incubating infected become contagious, i.e. in the first 3.5 days of the incubation period.

Thus, we can calculate an upper bound of the impact of the net daily detection and isolation efficiency as $DT_{dmax_{daily}} = T_d / I_{new} \approx 85/140 = 0.6071$ or approximately 61%.

The impact would be minimal if the detection and isolation of all individuals are asymptomatic infected occurs the very last day of their time as infectiousness. However, we need to recall that throughout this study we have assumed that almost all the symptomatic infected are detected and isolated in Denmark so part of the daily identification and isolation process involves the symptomatic infected. Since $I_s = 225$ in the steady state, we can deduce that the epidemic produces $I_s \gamma_s = 22.5$ new symptomatic infected each day that are assumed detected. So 85 detections/day minus 22.5 detections/day = 62.5 detections/day are available for detection from $I_a$ at their last contagious day. In our model there are 11.5 days of infectiousness for both the symptomatic and asymptomatic populations: 1.5 days of the incubation times plus 10 days of the symptomatic or asymptomatic time. If for simplicity we assume equal contagiousness every day we get that the symptomatic can only infect 1.5 days during the incubation time as we assume they are detected and isolated once they become symptomatic. The asymptomatic can infect 1.5 days during the incubation time and 9 days during the asymptomatic infection period, because they are only detected on the 10th and last day of this period.

Thus, we get a lower bound on the net daily detection efficiency $DT_{min_{daily}} \approx (11.5−0.5 \times 22.5/140 = 0.1786$ or approximately 18.0%.

We now have estimates for a lower and upper bound on the daily testing efficiency for the period July 15 and August 15, 2020 given as $18% < DT_{daily} < 61%$.

As we have no detailed knowledge about the details of the infection and contact tracing processes we may use the average (center point) of the upper and the lower bound as a rough estimate for the daily testing efficiency: $(18 + 61)/2 \approx 39.5$ or approximately 40%. Obviously, more accurate estimates could be calculated from an explicit modeling and simulation of the involved infection and contact tracing activities.

Similarly for the period October 1 - 20, 2020 we obtain a single day detection efficiency $DT_{abs} = T_d / I_{obs} = 407/9351 = 0.0435$ or approximately 4.4%. $DT_{dmax_{daily}} = T_d / I_{new} = 407/813 = 0.501$ or approximately 50% while $DT_{dmin_{daily}} = (11.5−0.5 \times 130 + 11.5−0.5 \times (407−130))/813 \approx 0.3102$ or approximately 31%.

Thus, a rough estimate for the daily testing efficiency is $(50.1+31.0)/2 = 40.55$ or approximately 41% that within rounding errors is the same daily testing efficiency as over the summer.

There are two important lessons to take home from the above estimates:

(i) The testing, contact tracing, and isolation program in Denmark both provides a key indicator for the current geographic infection trends, but also takes part in balancing the background infection increase. This is particularly true when geographically focused efforts are undertaken. Removing and isolating on average 85 infected per day July 15 - August 15, 2020 together with behavioral restrictions had a significant impact in controlling the epidemic. Without the testing and the resulting daily removal of infected individuals the Danish pandemic would likely have been in an expansion phase during this period.

(ii) The Danish testing program, which is among the most ambitious in the world, has a testing efficiency of approximately 40% for both periods investigated, while the single day infection tracing efficiency indicator is found to be 5.3% mid July to mid August 2020 and 4.4% for October 1 - 20, 2020. Many countries have both lower net daily and single day testing efficiencies as we shall see in next subsection.
C. Hospitalized and estimated infected populations in other regions: Spain and the US

If we assume the biology of the COVID-19 pandemic is similar in Denmark, Spain, and the US, and we further assume that the hospital system including ICU’s in these countries function in a similar manner regarding treatment of COVID-19 patients, we may scale the Danish hospitalization numbers and the corresponding infected population estimates to larger populations. Obviously more detailed studies are necessary to verify - or falsify - the validity of such a comparison, but as a simplified approximation we believe such estimates provide appropriate order of magnitude estimates.

In Table II we list estimates for infected and observed population sizes based on the total hospitalized population size and the quasi steady state approximations given in Eqs. (16) to (22). For the estimates we assume the number of intensive care unit patients are approximately one fifth of every patient, thus $ICU = 0.2 \times H_{tot} (k = 0.8$ in Eq. (16)).

Potential healthcare differences across different countries can be adjusted by adjusting the appropriate healthcare system parameters in Eq. (16).

<table>
<thead>
<tr>
<th>Hospital population</th>
<th>Infected population</th>
<th>Daily infected</th>
<th>Observable population</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1053</td>
<td>70</td>
<td>807</td>
</tr>
<tr>
<td>20</td>
<td>2106</td>
<td>140</td>
<td>1615</td>
</tr>
<tr>
<td>50</td>
<td>5265</td>
<td>351</td>
<td>4036</td>
</tr>
<tr>
<td>100</td>
<td>10530</td>
<td>702</td>
<td>8073</td>
</tr>
<tr>
<td>250</td>
<td>26324</td>
<td>1755</td>
<td>20182</td>
</tr>
<tr>
<td>500</td>
<td>52649</td>
<td>3501</td>
<td>40364</td>
</tr>
<tr>
<td>1000</td>
<td>105297</td>
<td>7020</td>
<td>80728</td>
</tr>
<tr>
<td>6000</td>
<td>631784</td>
<td>42119</td>
<td>484367</td>
</tr>
<tr>
<td>15000</td>
<td>1579459</td>
<td>105297</td>
<td>1210918</td>
</tr>
<tr>
<td>30000</td>
<td>3158918</td>
<td>210595</td>
<td>2421837</td>
</tr>
<tr>
<td>60000</td>
<td>6317835</td>
<td>421189</td>
<td>4843674</td>
</tr>
</tbody>
</table>

Table II: Estimated infected population sizes and infection rates as a function of hospital population using Eqs. (16) - (22). Note the linear proportionality between the numbers; e.g., 10 times higher hospital population yields 10 times higher infected populations and infection rates. See text for details.

By the end of August, 2020, Spain had about 6,000 hospitalized with COVID-19 [39], which according to our estimates corresponds to a total infected population of about 630,000 with about 484,000 observable and about 42,000 newly infected per day. In late August Spain detected and isolated a little less than 10,000 positive COVID-19 cases/day [38]. Thus the Spanish daily single day testing efficiency indicator is approximately $10,000/484,000 \approx 0.0207$ or 2.1%.

To estimate the net daily detection efficiency we need to approximate $(DT_{max}^{\text{daily}} + DT_{\min}^{\text{daily}})/2$. $DT_{\max}^{\text{daily}} = T_d/I_{new} \approx 10,000/42,000 = 0.2381$ or approximately 24%. To calculate $DT_{\min}^{\text{daily}}$ we note that $I_s \approx 67,399$ so $I_s^{\gamma_s} \approx 6,739$ new symptomatic infected are produced every day that we assume are detected. Since $T_a \approx 10,000$, $10,000 - 6,739$ detections are thus available for individuals in $I_t$ and $I_a$, recall discussion in the last subsection. $DT_{\min}^{\text{daily}} \approx (11.5 - 10.5 \times 6,739 + 11.5 - 1.5 \times (10,000 - 6,739))/2 = 11.4900$ or approximately 11.5%. This gives an approximate net daily detection efficiency of $(23.81 + 8.12)/2 = 15.965$ or approximately 16% by the end of August 2020.

By the end of September, 2020, the US had two hospitalization peaks, both with about 60,000 patients, one in mid April and one in late July 2020 [43], as well as two valleys, both with about 30,000 patients, one in mid June and one mid September 2020. The corresponding estimates from Table II indicate that 30,000 hospitalizations corresponds to about 3.16 mil. infected while 60,000 hospitalizations corresponds to about 6.32 mil. infected. The daily detection average for the US during September 2020 was about 40,000 [43] while the daily newly infected average was about 210,000. Thus the single day testing efficiency indicator in the US during September was about $40,000/2,420,000 \approx 0.0165$ or 1.7%.

To estimate the daily detection efficiency we need to approximate $(DT_{\max}^{\text{daily}} + DT_{\min}^{\text{daily}})/2$. $DT_{\max}^{\text{daily}} = T_d/I_{new} \approx 40,000/210,595 \approx 0.1899$ or approximately 19%. To calculate $DT_{\min}^{\text{daily}}$, we note that $I_s \approx 336,951$, so $I_s^{\gamma_s} \approx 33,695$ new symptomatic infected are produced every day that we assume are detected. This assumption seems less appropriate with the actual numbers involved, as almost all of the detections would then stem from symptomatic infected. However, to keep consistency in the estimations we keep this assumption. Since $T_a \approx 40,000, 40,000 - 33,695$ detections are thus available for individuals in $I_t$ and $I_a$, recall discussion in last subsection. $DT_{\min}^{\text{daily}} \approx (11.5 - 10.5 \times 33,695 + 11.5 - 1.5 \times (40,000 - 33,695))/210,595 \approx 0.0399$ or approximately 4%. This gives an approximate net daily detection efficiency of $(18.99 + 3.99)/2 = 11.4900$ or approximately 11.5% during September 2020.

The above estimates from Spain and the US should only be viewed as an illustration of the method developed in Eqs. (16) - (22) applied to other countries. Healthcare differences across different countries should be implemented by adjusting the appropriate healthcare system parameters in Eq. (16) as well as our assumption about detection and isolation of symptomatic infected.
D. Impact of the fraction of hospitalized symptomatic

Because $h_{frac}$ is a critical part of these population estimates, we investigate in simulation the impact of both a lower $h_{frac} = 0.061$ and a higher $h_{frac} = 0.139$ using visual inspection for appropriate fits between reported $H, ICU$ occupation data as well as death data and the corresponding values generated in simulation. Recall the original $h_{frac} = 0.082$. For details, see Appendix E.

Obviously a modified value for $h_{frac}$ requires a significant re-adjustment of the main infection parameters. If $h_{frac}$ increases, a larger fraction of the symptomatic infected $I_s$ needs hospitalization. As the reported hospitalization volume at any given time does not change, this implies that the total number of symptomatic infected $I_s$ must also be smaller to ensure correspondence between the reported and simulated data. Thus, increasing $h_{frac}$ means a decreasing $\rho_s$ so a smaller fraction of incubated $I_i$ becomes symptomatic $I_s$ while a larger fraction $(1 - \rho_s)$ becomes asymptomatic $I_a$. As we know from the COVID-19 iso-symptomatic infection diagram investigation, a smaller $\rho_s$ means a larger $\beta^0_s$ and a larger $R_0$ for a constant $\zeta$ value, that in our case is set to 0.309. As $\beta^0_s$ and $R_0$ change for the initial free pandemic so must $\beta^{lockdown}_s, \beta^{reopen1}_s, \beta^{reopen2}_s, x = i, a$ and their corresponding $R_0$ values.

Conversely, a smaller $h_{frac}$ implies a larger $\rho_s$ and a smaller $m_{frac}$ as well as smaller $\beta_s$ and $R_0$.

Nice hospital occupation and death data fits can be obtained both for $h_{frac} = 0.061$ and 0.139, but only for the $h_{frac} = 0.139$ case is it in the same simulation also possible to obtain the correct seroprevalence number of $\approx 79,700$ or $\approx 1.37\%$ of the adult population per May 28, 2020. For $h_{frac} = 0.061$ the simulated seroprevalence is $\approx 96,700$ and thus overshoots with a factor of $96,800/79,700 \approx 1.21$, if we at the same time require appropriate fits for $H, ICU$ and deaths tolls.

For more details, see Appendix E.

VIII. "WHAT IF?" SCENARIOS

A. Best case scenario: Continued micro-outbreak scenarios with $R_0 \simeq 1$.

Since April 2020 Denmark has experienced a number of localized micro-outbreaks that were contained due to isolation, contact tracing, and increasingly more extensive testing. A narrative discussion of these micro-outbreaks was given in the Introduction while in the Results Section we discuss how we model and simulate the impact of such micro-outbreaks by means of added internal noise.

Previously in Section VC we used noise to trigger injections of newly infected into the population to generate local micro-outbreaks. Local fluctuations in the behavior from time to time triggers micro-outbreaks probably in part caused by a combination of events or places where many people are gathered and super-spreaders are present. In these situations the general background dynamics is characterized by $R_0$ slightly smaller than 1.0 so that the background dynamics and the noise induced micro-outbreaks together generate a close to steady state infection level that corresponds to what we may call a macroscopic effective $R_0 \simeq 1$. We interpret the general background dynamics as the sum of the behavior of the general population together with the testing and contact tracing activities that result in a daily removal of infected from the epidemic, which together defines a $R_0$ that is smaller then one.

It should be emphasized that there is a critical difference between injecting additional infected into the general population versus changing the behavior in the general population. If the background infection is decreasing, $R_0 < 1$, an injection of newly infected has the same impact whether they e.g. are returning travellers that were infected elsewhere or they stem from a local micro-outbreak from within the country. In an $R_0 > 1$ situation both effects add equally to an already expanding pandemic.

From August 14 - 31, 2020, Denmark adjusted the reopening of the country. The most significant changes included opening of higher educational and vocational institutions as well as longer opening hours of restaurants and bars. However, the restaurant and bar openings were partly rolled back in the first part of September together with a number of additional restrictions due to observed increase in infection levels and hospitalizations, see [53].

If, as a "best case scenario" we assume no changes in the general human behavior compared to the summer of 2020, as well as no impact by the cooler fall and winter temperatures, we obtain a development as seen in Fig. 9. We still assume recurring localized micro-outbreaks with the same frequency and amplitude as previously had, which is an optimal dynamical regime to operate in until a vaccine becomes available. Since the pandemic dynamics is operating around $R_0 \simeq 1$ a quasi steady state can occur for any hospital occupation level, recall Section VII B. A low hospital occupation is obviously the most desirable. Unfortunately, Denmark did not follow this scenario.

B. Aggressive infection scenario following the influenza pattern

If all behavioral and policy restrictions are lifted without a universally available vaccine, a new aggressive infection wave would immediately emerge. If we further assume
Figure 9: Monte Carlo simulated health care system and death toll dynamics (100 realizations) with standard parameters, recall Table 1 and Fig. 7. The pandemic is assumed to be close to steady state with successive micro-outbreaks ($R_0 \sim 1$), where we assume no changes in policies and human behavior since the Summer of 2020, as well as no impact of colder fall and winter temperatures. Linear upper panel and logarithmic lower panel scales respectively. Note the continued low hospitalization and death toll. Dots represent the observed hospital and death data, lines represent the simulation data.

Figure 10: Vertical axis: number of laboratory verified influenza cases in Denmark shown for the flu seasons 2014/15 to 2019/20 until the onset of the COVID-19 pandemic. Horizontal axis: week number in a calendar year. Note the clear yearly epidemic pattern that starts in the early winter and peaks between early February and mid March. It should also be noted that the closing of Denmark due to COVID-19 March 16, 2020, stopped the influenza epidemic and actually also a smaller whooping cough epidemic, not shown in the above graph [16].

that we still keep the symptomatic infected effectively isolated, i.e. $\beta_s = 0.01\beta_0^a$, we get a scenario where the pandemic is governed by a free infection spread almost solely by the asymptomatic $I_a$ and incubating $I_i$ populations. Such a scenario yields, recall Eqs. (7) - (9),

$$R_0 = \frac{\beta_0^a}{\gamma_1} + \frac{(1 - \rho_s)\beta_0^a}{\gamma_a} + \frac{\rho_s0.01\beta_0^a}{\gamma_s}$$

$$= \frac{\beta_0^a (\rho_s + (1 - \rho_s)\zeta)0.351}{\gamma_a} + \frac{(1 - \rho_s)0.309\beta_0^a}{\gamma_a} + \frac{\rho_s0.01\beta_0^a}{\gamma_s}$$

$$= 3.004 \sim 3.0.$$  

Both the initial growth in the $H_{tot}$ observations as well as the simulations indicate that the initial pandemic February - March 2020 with $R_0 \sim 5.4$ had a doubling time of $\sim 2.4$ days while a pandemic with $R_0 \sim 3.0$ would have a doubling time of $\sim 4.2$ days. Thus, as long as we effectively isolate the symptomatic infected a worst case scenario of such an infection wave cannot be as aggressive as the initial wave we experienced early 2020. This means that future infection waves would be expected to emerge with $R_0$ less than 3.0 and a doubling time of more than 4.2 days.

The most vulnerable period for an aggressive Danish COVID-19 infection wave would likely be at the time the country usually experiences its yearly influenza epidemic, see Fig. 10.

Fig. 11 shows an aggressive scenario with $R_0 = 3.0$ and $\beta_s = 0.01\beta_0^a$ (isolated symptomatic infected) adapted to the yearly epidemic influenza pattern. To explore the full impact of such an infection wave no policy interventions are assumed, which means that incubating and asymptomatic individuals move freely to interact with and potentially infect the susceptible population. In particular, note that in particular due to the assumed limited immunological memory for the asymptotic infected ($1/\xi_a = 60$ days; decay time) the pandemic does not "burn out" as the susceptible population is continuously replenished by the recovered population.
Without intervention the pandemic peaks for \( I_t \) December 30, 2020 and January 7, 2021 for \( I_s \) and \( I_a \), and by February 7, 2021, approximately 4,071,000 people (\( \sim 70\% \) of the population) would have been infected and recovered from the pandemic. The simulated hospital impact peaks around January 17, 2021 with \( \sim 19,680 \) hospitalized and January 20, 2021 with \( \sim 4,786 \) in ICUs which would completely overwhelm the Danish healthcare system. It is estimated that the Danish ICU capacity is about 1,000 beds and the total hospital capacity is about 15,000 beds. A year after the onset of this infection wave (November 1, 2021) about 56,600 people would have died from the SARS-CoV-2 virus, which is about the same as the total number of deaths in Denmark in the year 2019: 53,958.

C. Free pandemic: Worst case scenario

"What would have happened if no policy interventions were imposed at the beginning of the pandemic?" This is a question many people have asked us. In Fig. 12 data are generated in simulation using standard the parameters also used in Figs 4 and 5, but assuming no behavioral modifications or policies imposed at the onset of the pandemic, thus \( R_0 \approx 5.36 \). This is a hypothetical scenario where the pandemic rages freely. The pandemic does not "burn out" in this scenario either as the susceptible population is continuously replenished by the recovered population, recall discussion in Fig. 11. This means that a relatively stable hospital population of about \( H_{tot} \approx 6,500 \) by October 2020 would slowly decrease to \( \sim 4,500 \) in October 2021. A conservative estimate of the expected excess deaths from such a free pandemic is \( \sim 64,960 \) after one year, by February 24, 2021, which is more than all the deaths in Denmark in 2019 (53,958).

IX. DISCUSSION

The goal of this work is to provide a theoretical platform to help understand the details of the Danish COVID-19 epidemic that can hopefully also provide useful insights about the nature of the pandemic in other countries and regions. Our models and simulations can easily be adapted to other regions or countries where hospital, ICU and death data are available.

We have chosen a simple macroscopic description and used a differential equation based simulation of the pandemic and health care system dynamics that seems appropriate for most of the issues we seek to address, although a microscopic agent based simulation scheme would have been more appropriate e.g. to capture the dynamics of the observed micro-outbreaks. Our approach is also limited in two other aspects. We neither have a geographic representation of the pandemic nor do we have an age disaggregated population although reported data...
clearly show features caused both by geography and age differences in the population.

However, by adding noise to our model and conducting Monte Carlo simulations of large ensembles of the system we can to some extent compensate for the shortcomings caused by the lack of geography representations and the occurrence of localized micro-outbreaks. This approach is further supported because Denmark is a small country with a relatively high mobility and population mixing. The missing age distribution does not cause serious problems for understanding the initial epidemic dynamics, but for the later micro-outbreaks we in principle need to know the nature of their age distribution as this significantly impacts the health care system. We have not implemented such parameter changes in our simulations. All parameters are kept constant over time in each of our simulations except the externally imposed lockdown and reopening policies/behaviors. By adjusting the amplitude and frequency of the added noise we can adjust the expected impact on the health care system from such micro-outbreaks.

The micro-outbreaks since late July 2020 began in a slaughterhouse in the city of Ringsted. Then came the outbreaks in the three largest cities in Denmark: first in Århus and then around September 1, 2020 in Odense and Copenhagen. These outbreaks were driven by young people aged 20-29 years and immigrants from Middle East and Somalia who live in rather closed communities. The outbreaks spread to other groups but fewer old and vulnerable persons became infected compared to the epidemic in the spring. The number of patients hospitalized, in particular patients treated in ICUs and the number of deaths, remained very low in the fall period compared to the spring epidemic although it was still the old people who dominated the hospitalizations and deaths. In each of the micro-outbreaks non-pharmacological measures were re-introduced and the $R_0$ of each micro-outbreak went from above 1.5 to approximately 0.8 within a month, which probably is to be expected considering the epidemic dynamics of COVID-19.

The testing capacity during these outbreaks went up to nearly 1% of the whole Danish population each day and by October 7, 43% of the population had been tested for SARS-CoV-2. That means that many persons that contacted SARS-CoV-2 positive persons including asymptomatic persons were detected. Therefore of the 9,623 SARS-CoV-2 positive persons in the fall outbreak only 34 (0.37%) died, whereas 434 (4.8%) of the 8,851 SARS-CoV-2 positive patients who were hospitalized during the spring epidemic died. The reduced mortality is similar to a case-fatality rate of 0.4% in a recent report from Hong Kong [27]. The early high rate mainly reflects that only patients who were hospitalized were tested for SARS-CoV-2 in the spring epidemic. Therefore, the true mortality of the spring epidemic was calculated too high.

This is important to realize, because a restricted testing capacity may mislead the health-care system, the population and the precautions taken to manage the epidemic,

Figure 12: Upper panel: The simulated dynamics of the infected and recovered for a hypothetical free COVID-19 pandemic in Denmark that started February 24, 2020 and runs until November 1, 2021. This hypothetical pandemic would have peaked April 7, 2020 with $I_s \approx 1,451,000$ and peaked April 13 with $I_s \approx 1,729,000$ and $I_s \approx 329,400$ infected individuals. By May 13, 2020, the recovered population would have peaked with more than 4.32 mil. or about 74 % of the population ($R_{tot} \approx 4,320,000$). Lower panel: Simulated and reported health care system occupation as well as death toll data (logarithmic scale). Dots represent the observed hospital and death data, lines represent the simulation data. The hypothetical death toll of 64,960 is clearly underestimated as we use the health care model from Eqs. (2) that in this simulation is assumed to be able to scale with the pandemic, which it clearly cannot in real life. The hospitalization occupation peaks April 23, 2020, with $H_{tot} \approx 22,850$ patients and the intensive care units peaks April 26 with $ICU \approx 5,472$. The Danish healthcare systems would be completely overwhelmed in such a scenario and we would expect an increased mortality rate because of that.
e.g. based on the calculated $R_0$ and $R_t$. The burden on the hospitals and ICUs was lower in the outbreaks in the fall. The spring/fall ratio of the maximum number of normal hospital unit COVID-19 patients was 535/147 while the corresponding ratio for ICU patients was 153/20. The improved treatment of COVID-19 patients (with e.g. Remdesivir, dexamethasone, and prophylaxis and treatment of bleeding disorders, [5][12]) is probably the reason for the proportionally lower number of hospitalized patients who in the fall were treated in ICUs or died.

The estimated percentage of symptomatic infected that needs hospital care $h_{frac}$ is a critical parameter as it scales the connection between the main observables, the time series for $H$, ICU occupations, and the death toll, and the infection dynamics. We have used an age aggregated estimate for $h_{frac} = 8.2\%$ based on the age distribution of health care needs from [18]. With an increase of $h_{frac}$ to 13.9% the simulation still corresponds reasonably well to the observed data we have access to, while a decrease in $h_{frac}$ to 6.1% does not allow the simulation to both satisfy the time series of our main observables and the measured seroprevalence of May 28, 2020.

For our hospital model we have also chosen simplicity over including more details that potentially could help capture the health care dynamics in greater detail. The simple hospital model is nonetheless able to reproduce both the historical hospital and death data quite well, but it has difficulties reproducing some of the reported data details. For example, both the reported $H_{tot}$ and ICU occupation time series have sharper peaks at the onset of the pandemic than the simulated peaks. Also to enable the simulated ICU data to follow the reported time series data in a reasonable manner we need to use an average ICU occupation time $1/\gamma_{icu} = 10$ days, which is shorter than the reported average occupation time, although it should also be noted that a more recent improved treatment of COVID-19 patients means that fewer have a severe course [5] [12].

Requiring our healthcare model to reproduce the historical hospital $H_{tot}$ and ICU data, the simulation underestimates the total accumulated number of hospital patients $H_{tot}$ by 7-14% and it also underestimates the accumulated number of ICU patients. Thus, our simple health care model has difficulties simultaneously matching both the time series data (daily hospital occupation numbers) and the accumulated data (total number of hospitalized patients).

There is an ongoing discussion of the best value for average period of infectiousness both for the symptomatic and asymptomatic populations, where [6] suggests $1/\gamma_s = 1/\gamma_a = 10$ days while [28] suggests 7 days as a better estimate. Lowering the infectious period $1/\gamma_s$ and $1/\gamma_a$ both for the symptomatic and the asymptomatic populations from 10 to 8 days in our simulation makes it difficult to reproduce the total hospital occupation numbers (too fast a decline of simulation numbers after lockdown), while for the simulated death numbers from the hospitals (too fast a rise). However, shorter $1/\gamma_s$ and $1/\gamma_a$ values make it easier to reproduce the initial relatively sharp peak for the ICU occupation number. To compensate for the lower $1/\gamma_a$ and $1/\gamma_s$ values the $\beta$ values must increase a bit. With lower infection times minor adjustments are needed for the parameters associated with hospitals, ICUs, and death tolls.

We can use Monte Carlo (MC) optimization of the parameters where we minimize the Least Square (LS) difference between the reported and the simulated time series of the total hospital occupation ($H + ICU$). Despite the minor mismatches between what the MC-LS parameter optimization yields and what visual parameter inspection yields, the MC-LS capability enables us to do large scale explorations of the parameter impacts.

Using a steady state approximation for the background infection that is indicated by approximately constant low values in our main observables, i.e. $H$, ICU and death data, enables us to make simple analytical estimates of the sizes of the corresponding infected populations $I_s$, $I_a$ and $I_c$. These analytical estimates are robust to most parameter perturbations and they are confirmed by numerical simulations using steady state conditions. This approach and the corresponding infected population estimates should give more accurate predictions of the actual sizes of the infected populations than direct infection testing of the daily scale currently being done.

Another general finding coming out of this study is the tight relationship that exists between $\beta_s$, $R_0$, $\rho_s$ and $\zeta$, which can be represented by an iso-symptomatic-infection diagram, see Section VIA, that is made possible by using MC-LS optimization for each $\beta_s$, $R_0$, $\rho_s$ and $\zeta$ parameter combination. We propose such diagrams could readily be constructed for most infectious diseases, so that one diagram would hold crucial and comparable quantitative information about critical parameters that define the dynamic characteristics of a pandemic.

X. CONCLUSION

Our computational platform explores the connection between the dynamics of the viral epidemic, the national health care system and the imposed policies.

Based on reported seroprevalence data [54] we can estimate the relative frequency between symptomatic and asymptomatic infected to be 16% and 84% respectively. We estimate $R_0 \sim 5.4$ for the initial free pandemic, $R_0 \sim 0.4$ for the lockdown period and $R_0 \sim 0.8 - 1$ for most of the successive reopening periods.

For the immediate future, as the northern hemisphere prepares for the upcoming winter and potential new infection waves, it should be noted that over the summer Denmark has been operating with $R_0 \sim 1$ from June 2020.
through August 2020. The estimated infected population sizes for the quasi steady state period mid July to mid August, 2020 are $I_i \simeq 702$; $I_a \simeq 1,179$; $I_d \simeq 225$; $I_{tot} \simeq 2,106$; and $I_{obs} \simeq 1,615$, while the daily infection rate is $I_{new} \simeq 140$.

The daily net testing efficiency is estimated to be approximately 40% for the periods July 15 - August 15, 2020 and October 1 - 20, 2020. The single day identification efficiency indicator in these period, defined as the number of positively identified infected over the estimated observable infected population the same day, is about 5%.

Since the symptomatic infected are effectively isolated from the population, and if behavioral policies are kept in place and observed, future COVID-19 infection waves fortunately would not be nearly as aggressive as the first wave assuming the viral biology does not change significantly due to mutations. We estimate that a new infection wave would have a lower bound with doubling time of more than 4.2 days and less than $R_0 \sim 3.0$ compared to the initial wave with a doubling time of about 2.4 days and $R_0 \sim 5.4$.

We believe our simulation platform is suitable for exploring forecasting scenarios for the upcoming winter and beyond or until a vaccine hopefully becomes available. Our results indicate that it is possible to more precisely than previously calculate the burden on the hospitals and ICUs of new COVID-19 infection waves or a new pandemic. Because of that, our simulation could hopefully also help minimize the impact on hospital treatment of non-COVID-19 patients.

Importantly we have shown, that the mortality of COVID-19 in Denmark is only 0.4% like in Iceland and the Faeroe Islands. Although a mortality rate of $\sim 0.4\%$ may seem small, our simulation indicates that had Denmark not adopted any behavioral measures to counteract the Danish pandemic the death toll would be $\sim 64,960$ by February 24, 2021 one year into the pandemic. That surpasses the total number of deaths (53,958) in Denmark in the year 2019. This simulation based COVID-19 death estimate is conservative as we have assumed a fully functional healthcare system, which is unrealistic as the free pandemic in such a scenario would have resulted in a hospital $H_{tot}$ and ICU occupation peaks in late April, 2020 that would have completely overwhelmed the Danish health care system with $\sim 22,850$ hospitalized and $\sim 5,472$ in ICUs. Per January 1, 2021 there would still be $\sim 5,860$ hospitalized and $\sim 1,500$ in the ICUs in a free COVID-19 pandemic according to such a simulation scenario.

Clearly Denmark’s adopted behavioral modifications, lockdown, measured reopening, testing, and contact tracing procedures, have been highly successful in mitigating the epidemic from the onset through August 2020.

Acknowledgment

We are grateful for the general guidance of John Erik Hansen regarding all epidemic matters in the early stages of this study as well as a final critical review of the manuscript. We are also grateful for the enlightening conversations with Gitte Kronborg regarding the COVID-19 situation at Danish hospitals in the early stages of the epidemic. Oana Ciofu and Asbjørn Dahl are acknowledged for their critical review of the manuscript and we thank Hans Ziock for his in depth critical reading, corrections, and constructive suggestions throughout the last versions of the manuscript. All remaining errors in the manuscript are the responsibility of the authors.

Appendix A: Statens Serum Institut data

COVID-19 data from Statens Serum Institut is plotted in Fig. 13 (data accessed per August 31, 2020) See caption for details.

It should be noted that SSI has modified minor details in data multiple times during the period February through August 2020. However, these changes have not had any measurable impact on our investigations or conclusions.

Appendix B: Fraction of hospitalized symptomatic infected

The current age distribution of the population in Denmark is given by Danmarks Statistik (Statistics Denmark) [11] and by combining it with the age dependent hospitalization fractions given in Ferguson et al.(2020), p5, table 1 [18] yields:

<table>
<thead>
<tr>
<th>Age distribution of hospitalization</th>
<th>( H_{frac0019} )</th>
<th>( H_{frac4059} )</th>
<th>( H_{frac6079} )</th>
<th>( H_{frac80up} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0019=1.30/5.83</td>
<td>(0.1+0.3)/2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2039=1.47/5.83</td>
<td>(1.2+3.2)/2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4059=1.56/5.83</td>
<td>(4.9+10.2)/2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A6079=1.23/5.83</td>
<td>(16.6+24.3)/2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A80up=0.27/5.83</td>
<td></td>
<td>(27.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From these age weighted hospitalization fractions, $h_{frac}$ is obtained as

\[
h_{frac} = (A0019 \times H_{frac0019} + A2039 \times H_{frac2039} + A4059 \times H_{frac4059} + A6079 \times H_{frac6079} + A80up \times H_{frac80up})/100,
\]

which yields $h_{frac} = 0.0820$.

Obviously, the value of $h_{frac}$ is an important parameter as it scales the epidemic because it determines how
many symptomatic infected $I_s$ enter the hospital system, which is the empirical data we use to adjust our simulation. A smaller $h_{frac}$ implies a larger scale epidemic with higher total number of infected individuals nationwide while a larger $h_{frac}$ implies a smaller scale epidemic. Note that the overall scale of the epidemic is in part restricted by the empirical hospital and death data and in part by the empirical data from the population wide serological antibody count. A detailed discussion of the impact of the size of $h_{frac}$ is given in Appendix E.

Appendix C: Monte Carlo Least Square optimization

The optimization algorithm includes a Monte Carlo method to minimize the sum of squared differences between the Danish hospitalization data and corresponding data generated in simulation, which is the sum of the state variables $H$ and $ICU$ in the model, recall equations (1) and (2). The Least Squares (LS) difference is defined as follows:

$$\text{Least Squares error} = \sum_{i=1}^{n} (f(t_i, \Theta) - y_{t_i})^2 \quad (28)$$

where $y_{t_i} = y_{t_1}, y_{t_2}, ..., y_{t_n}$ is the observed data and $f(t_i, \Theta)$ is the corresponding solution of the simulation with a given parameter set $\Theta = (\theta_1, \theta_2, ..., \theta_m)$. A pseudo description of the algorithm is as follows:

(i) Generate random numbers within a given domain and choose these numbers as values for selected parameters,
(ii) Run a simulation with these parameter settings,
(iii) Calculate the LS difference between reported and simulated data with chosen parameters,
(iv) If LS error is less than the former LS error then save the parameter values,
(v) Repeat a predefined number of times - or until a certain small LS error value is reached.

Figure 14 illustrates the Least Squares error for a standard run with $\rho_s = 0.16$ and $\zeta = 0.309$ only sampling on values of $\beta^0_s$ from a uniform distribution. Note the minimum of the LS error for $\beta^0_s \simeq 1.0583$.

Appendix D: Derivation of $R_0$

The basic reproduction number, $R_0$, can be calculated from an epidemic compartment model using the Next-Generation Matrix Method [10]. Our epidemic compartment model given in Eqs. (1) consists of five compartments $(S, I_t, I_a, I_s, R)$ with three $(I_t, I_a, I_s)$ infected populations. Letting $S = N$ (total population) and only...
including the infected individuals Eqs. (1) reduce to

\[
\frac{dI_i}{dt} = \beta_i I_i + \beta_a I_a + \beta_s I_s - \gamma_i I_i \\
\frac{dI_a}{dt} = (1 - \rho_s) \gamma_i I_i - \gamma_a I_a \\
\frac{dI_s}{dt} = \rho_s \gamma_i I_i - \gamma_s I_s, \tag{29}
\]

where we have ignored the hospital dynamics. We may now write Eqs (29) in matrix form as

\[
\frac{dI}{dt} = (B + \Gamma) I, \tag{30}
\]

where the \(B\) matrix express the epidemic transmissions while the \(\Gamma\) matrix express the transitions between the infected populations each defined as

\[
B = \begin{pmatrix} \beta_i & \beta_a & \beta_s \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{31}
\]

and

\[
\Gamma = \begin{pmatrix} -\gamma_i & 0 & 0 \\ (1 - \rho_s) & 0 & 0 \\ \rho_s & 0 & -\gamma_s \end{pmatrix}. \tag{32}
\]

\(\mathcal{R}_0\) can now be defined as \([10]\]

\[
\mathcal{R}_0 = \rho(B\Gamma^{-1}) = \rho(K) \tag{33}
\]

where \(\rho(K)\) defines the dominating eigenvalue of the resulting \(K\) matrix.

The inverse of \(\Gamma\) is calculated by \(\Gamma^{-1} = \frac{1}{\det(\Gamma)} \text{Adj}(\Gamma)\). The determinant of \(\Gamma\) is

\[
\det(\Gamma) = \gamma_i \gamma_a \gamma_s. \tag{34}
\]

By first transposing \(\Gamma\) we may calculate the adjunct of \(\Gamma\) by first transposing, then finding the matrix of the minor and then calculating the matrix of cofactors.

\[
\Gamma^T = \begin{pmatrix} \gamma_i - (1 - \rho_s) \gamma_i & -\rho_s \gamma_i \\ \gamma_a & 0 & 0 \\ 0 & 0 & \gamma_s \end{pmatrix}. \tag{35}
\]

The matrix of minors is given as

\[
M(\Gamma^T) = \begin{pmatrix} \gamma_a \gamma_s & 0 & 0 \\ -(1 - \rho_s) \gamma_i \gamma_s & \gamma_i \gamma_s & 0 \\ \rho_s \gamma_i \gamma_a & 0 & \gamma_i \gamma_a \end{pmatrix}. \tag{36}
\]

Applying the matrix of cofactors to (36) to obtain the adjunct of \(\Gamma\)

\[
\text{Adj}(\Gamma) = \begin{pmatrix} \gamma_a \gamma_s & 0 & 0 \\ (1 - \rho_s) \gamma_i \gamma_s & \gamma_i \gamma_s & 0 \\ \rho_i \gamma_i \gamma_a & 0 & \gamma_i \gamma_a \end{pmatrix}. \tag{37}
\]

This implies

\[
\Gamma^{-1} = \begin{pmatrix} \frac{1}{\gamma_i} & 0 & 0 \\ \frac{1}{\gamma_a} & 0 & 0 \\ \frac{1}{\gamma_s} & 0 & 0 \end{pmatrix}. \tag{38}
\]

By matrix multiplication of \(B\) and \(\Gamma^{-1}\) we get

\[
K = B\Gamma^{-1} = \begin{pmatrix} \beta_i & \beta_a & \beta_s \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} \frac{1}{\gamma_i} & 0 & 0 \\ \frac{1}{\gamma_a} & 0 & 0 \\ \frac{1}{\gamma_s} & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{\beta_i}{\gamma_i} & \frac{\beta_a (1 - \rho_s)}{\gamma_a} & \frac{\beta_s \rho_s}{\gamma_s} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{39}
\]

Since \(K = B\Gamma^{-1}\) is an upper triangular matrix the (only) eigenvalue can be read off the diagonal and the basic reproduction number is

\[
\mathcal{R}_0 = \frac{\beta_i}{\gamma_i} + \frac{\beta_a (1 - \rho_s)}{\gamma_a} + \frac{\beta_s \rho_s}{\gamma_s}. \tag{39}
\]

**Appendix E: Impact of size of \(h_{frac}\)**

Here we address the impact of changing the size of \(h_{frac}\). If \(h_{frac}\) increases a larger fraction of the symptomatic infected \(I_s\) needs hospitalization. As the reported hospitalization volume at any given time does not change, this
implies that the total number of symptomatic infected $I_s$ must also be smaller so as to ensure correspondence between the reported and simulated data. Thus, increasing $h_{frac}$ means a decreasing $\rho_s$ so a smaller fraction of incubated $I_i$ becomes symptomatic $I_s$ while a larger fraction $(1 - \rho_s)$ becomes asymptomatic $I_a$. As we know from the COVID-19 iso-symptomatic infection diagram investigation, a smaller $\rho_s$ means a larger $\beta_s^0$ and a larger $R_0$ for a constant $\zeta$ value, that in our case is set to 0.309. As $\beta_s^0$ and $R_0$ change for the initial free pandemic so must $\beta_a^0$, $\beta_{lockdown}^0$, $\beta_{reopen1}^0$, $\beta_{reopen2}^0$, $x = i, a$, and their corresponding $R_0$ values.

We test two concrete $h_{frac}$ values away from the standard $h_{frac} = 8.2%$:
(i) \( h_{frac} = 13.9\% \), a \( \sim 70\% \) increase, see Fig. 15.
(ii) \( h_{frac} = 6.1\% \), a \( \sim 25\% \) decrease, see Fig. 16.
Figs. 15 and 16 should be compared with Figs. 4, 5 and 6.

Although both simulation experiments in Figs. 15 and 16 generate nice correspondence with the observed primary data, only \( h_{frac} = 13.9\% \) is simultaneously able to satisfy the reported seroprevalence of May 28, 2020. This means that 0.082 \( \leq h_{frac} \leq 0.139 \) can satisfy all observed data with parameter adjustments, while \( h_{frac} = 6.1\% \) cannot. See more details in the figure captions.

[14] DR Detector (National Broadcasting Company), August 6, 2020
[16] EPI-NYT, Statens Serum Institut
[18] N M Ferguson, et al., Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team, March 16, 2020
[24] Li et al., Emergence of SARS-CoV-2 through recombination and strong purifying selection. Science Advances 01 Jul 2020; Vol. 6, no. 27, eabb9153 DOI: 10.1126/sci-adv.eabb9153
[35] Timothy W Russell, Joel Hellewell, Christopher I Jarvis, Kevin Van Zandvoort, Sam Abbott, Ruwan Ratnayake, CMMID COVID-19 working group, Stefan Flasche,

Sakurai A et al. Natural history of asymptomatic SARS-CoV-infection. NEJM, June 12, 2020 doi 10.1056/NEJMc2013020


Yang H, Chen D, Jiang Q, Yuan Z. High intensities of population movement were associated with high incidence of COVID-19 during the pandemic. Epidemiol Infect 2020 dot 10.1017/S0950268820001703


https://www.politi.dk. Home page for the Politi (Danish Police) with focus on COVID-19 (accessed June - September, 2020)

https://www.ssi.dk. Home page for Statens Serum Institut (Danish Center for Decease Control) with focus on COVID-19 (accessed June - September, 2020)
