Supplemental Material:

Comprehensive compartmental model and calibration algorithm for the study of clinical implications of the population-level spread of COVID-19: a study protocol

Brandon Robinson[1], Jodi D. Edwards[2,3], Tetyana Kendzerska[3,4,5], Chris L. Pettit[6], Dominique Poirel[7], John M. Daly[8], Mehdi Ammi[9], Mohammad Khalil[10], Peter J. Taillon[11], Rimple Sandhu[12], Shirley Mills[13], Sunita Mulpuru[4,5], Thomas Walker[1], Valerie Percival[14], Victorita Dolean[15,16], and Abhijit Sarkar *[1]

[1]Department of Civil and Environmental Engineering, Carleton University, Ottawa, ON, Canada K1S 5B6
[2]School of Epidemiology and Public Health, University of Ottawa and University of Ottawa Heart Institute, Ottawa, ON, Canada K1Y 4W7
[3]ICES, Ottawa, ON, Canada K1Y 4E9
[4]The Ottawa Hospital Research Institute, Ottawa, ON, Canada
[5]Department of Medicine, Faculty of Medicine, Division of Respirology, University of Ottawa, Ottawa, ON, Canada K1H 8L6
[6]US Naval Academy, Aerospace Engineering Department, Annapolis, MD 21402, United States
[7]Royal Military College of Canada, Department of Mechanical and Aerospace Engineering, Kingston, ON, Canada K7K 7B4
[8]Independent Control Systems Engineer, Ottawa, ON, Canada K1S 4H6
[9]School of Public Policy and Administration, Carleton University, Ottawa, ON, Canada K1S 5B6
[10]Sandia National Laboratories†, Livermore, CA 94550, United States
[11]Schaffen Research Inc. Ottawa, ON, Canada K1H 7S7
[12]National Renewable Energy Laboratory, Golden, CO 80401, United States
[13]School of Mathematics and Statistics, Carleton University, Ottawa, ON, Canada K1S 5B6
[14]School of International Affairs, Carleton University, Ottawa, ON, Canada K1S 5B6
[15]Department of Mathematics and Statistics, University of Strathclyde, Glasgow, Scotland G1 1XQ
[16]Laboratoire J.A. Dieudonné, CNRS, Université Côte d'Azur, Nice, France 06108

The 22-compartment model described in the paper consists of a number of extensions to the detailed 16-compartment model by Tuite et al. [1], including the addition of multiple compartments, model stratifications, and model discrepancy. This supplemental material contains the equations of motion for the model described in "Study protocol: a comprehensive compartmental model and calibration algorithm for the study of clinical implications of the population-level spread of COVID-19." The mathematical form of the model is described, along with the inclusion of model error to extend the deterministic model to a stochastic compartmental model. Subsequently, the pertinent details of the proposed nonlinear sparse Bayesian learning (NSBL) algorithm are provided. The necessary modifications to the ordinary differential equation model capturing only the temporal variation of COVID-19 case

*Corresponding author; email: abhijit.sarkar@carleton.ca

†Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

counts to include the spatio-temporal variation of disease dynamics through a partial differential equation model are discussed.

A Modelling population-level spread through ordinary differential equations

The model by Tuite et al. [1] is stratified by age into 16 age groups with equal widths of five years, and includes a second stratification indicating whether or not an individual has a pre-existing health condition. In this case, each compartment is indexed by $i = (1,\ldots,16)$ for age group, and $j = (1,2)$ for comorbidity. The proposed model considers the stratification more generally; each compartment in Figure 1 would be indexed by $i = (1,\ldots,N_i)$ for age group, $j = (1,\ldots,N_j)$ for comorbidity ($j = 1$ for no comorbidity, and $j = 2,\ldots,N_j$ for specific pre-existing health conditions). Furthermore, in the equations below, an additional index $k = (1,\ldots,N_k)$ is introduced that has a distinct index for unvaccinated individuals and $N_k - 1$ indices different vaccines (see discussion in Section 2.2.2). The total number of states in the model is $(22 \times N_i \times N_j \times N_k)$. Clearly, as the number of phenomena modelled by stratification increases, the number of states in the model increases exponentially.

Note that the other phenomena described in Sections 2.2.1 and 2.2.2 may be modelled through further stratification (by introducing additional indices) with minimal modification to the system of equations below. The equations are shown for the three above-mentioned stratifications as the model includes cross-terms that allow individuals to move from one index to others in time (i.e., by aging out of their current age index, i, by developing long-term health complications resulting in a change in comorbidity index, j, and by becoming vaccinated, or having the period of vaccine effectiveness elapse resulting in a change in vaccination index, k).

The system of equations (Eqs. (A.1 - A.22), derived by extending the equations in [1]) is a direct mathematical representation of the flowchart in Figure 1, hence, each state (summarized in Table 1) relates to one of the compartments of that flowchart. Likewise, the coefficients (rate parameters or probabilities, summarized in Table 2) are represented by the arrows in the flowchart, as these parameters dictate the flow of individuals between adjacent compartments. Each equation represents the net flow of the population into the respective compartment. The indices of the parameters are omitted from Eqs. (A.1 - A.22) for simplicity. Note the model discrepancy terms in red for each model equation. These consist of a strength parameter, $q_S,\ldots,q_D$, and correlated noise processes, $\xi_S,\ldots,\xi_D$, thereby extending the model equations to a stochastic form. Note that the model errors must be statistically correlated in order to maintain the conservation of the population in the mechanistic model [2]. The model error strength parameters will be estimated in the inference through the NSBL algorithm.

Susceptible:

$$\frac{dS^{ijk}}{dt} = -\lambda^{ijk}S^{ijk} + \sum_{\psi=1}^{N_k}(-\gamma_V^{\psi}S^{ijk} + \gamma_T^{\psi}\Psi^{\psi k}V^{ij\psi}) + \gamma_R(R1^{ijk} + \sum_{\upsilon=1}^{N_j}\Upsilon^{\upsilon j}R2^{i\upsilon k}) - \mu S^{ijk} + \mu S^{(i-1)jk} + q_S\xi_S \quad \text{(A.1)}$$

Vaccinated:

$$\frac{dV^{ijk}}{dt} = -(1-r_V)\lambda^{ijk}V^{ijk} + \gamma_V S^{ijk} - \gamma_T V^{ijk} - \mu V^{ijk} + \mu V^{(i-1)jk} + q_V\xi_V \quad \text{(A.2)}$$

Exposed:

$$\frac{dE^{ijk}}{dt} = (1-\delta_S)\lambda^{ijk}S^{ijk} + (1-\delta_V)(1-r_V)\lambda^{ijk}V^{ijk} - \gamma_E E^{ijk} - \mu E^{ijk} + \mu E^{(i-1)jk} + q_E\xi_E \quad \text{(A.3)}$$

Exposed, isolating:

$$\frac{dQ^{ijk}}{dt} = \delta_S\lambda^{ijk}S^{ijk} + \delta_V(1-r_V)\lambda^{ijk}V^{ijk} - \gamma_E Q^{ijk} - \mu Q^{ijk} + \mu Q^{(i-1)jk} + q_Q\xi_Q \quad \text{(A.4)}$$

Infectious, presymtomatic:

$$\frac{dA^{ijk}}{dt} = \gamma_E E^{ijk} - \gamma_P A^{ijk} - \mu A^{ijk} + \mu A^{(i-1)jk} + q_A \xi_A \tag{A.5}$$

Infectious, pre-symptomatic, isolating:

$$\frac{dW^{ijk}}{dt} = \gamma_E Q^{ijk} - \gamma_P W^{ijk} - \mu W^{ijk} + \mu W^{(i-1)jk} + q_W \xi_W \tag{A.6}$$

Infectous, asymptomatic:

$$\frac{dF^{ijk}}{dt} = \sigma_A \gamma_P A^{ijk} - \gamma_A F^{ijk} - \gamma_{DA} F^{ijk} - \mu F^{ijk} + \mu F^{(i-1)jk} + q_F \xi_F \tag{A.7}$$

Infectious, mild-to-moderate symptoms:

$$\frac{dB^{ijk}}{dt} = (1-\sigma_A)(1-\sigma_S)\gamma_P A^{ijk} - \gamma_M B^{ijk} - \gamma_{DM} B^{ijk} - \mu B^{ijk} + \mu B^{(i-1)jk} + q_B \xi_B \tag{A.8}$$

Infectious, severe symptoms:

$$\frac{dC^{ijk}}{dt} = (1-\sigma_A)\sigma_S \gamma_p A^{ijk} - \gamma_{S1} C^{ijk} - \mu C^{ijk} + \mu C^{(i-1)jk} + q_C \xi_C \tag{A.9}$$

Infectious, asymptomatic, isolating:

$$\frac{dX^{ijk}}{dt} = \sigma_A \gamma_P W^{ijk} - \gamma_A X^{ijk} - \gamma_{DA} X^{ijk} - \mu X^{ijk} + \mu X^{(i-1)jk} + q_X \xi_X \tag{A.10}$$

Infectious, mild-to-moderate symptoms, isolating:

$$\frac{dY^{ijk}}{dt} = (1-\sigma_A)(1-\sigma_S)\gamma_P W^{ijk} - \gamma_M Y^{ijk} - \gamma_{DM} Y^{ijk} - \mu Y^{ijk} + \mu Y^{(i-1)jk} + q_{Y2} \xi_Y \tag{A.11}$$

Infectious, severe symptoms, isolating:

$$\frac{dZ^{ijk}}{dt} = (1-\sigma_A)\sigma_S \gamma_P W^{ijk} - \gamma_{S1} Z^{ijk} - \mu Z^{ijk} + \mu Z^{(i-1)jk} + q_Z \xi_Z \tag{A.12}$$

Infectious, isolating after testing positive:

$$\frac{dG^{ijk}}{dt} = \gamma_{DA}(F^{ijk} + X^{ijk}) + \gamma_{DM}(B^{ijk} + Y^{ijk}) - \gamma_I G^{ijk} - \mu G^{ijk} + \mu G^{(i-1)jk} + q_G \xi_G \tag{A.13}$$

Inadequate access to health care resources:

$$\frac{dN^{ijk}}{dt} = (1-\sigma_H)\gamma_{S1}(C^{ijk} + Z^{ijk}) - \gamma_{S2} N^{ijk} - \mu N^{ijk} + \mu N^{(i-1)jk} + q_N \xi_N \tag{A.14}$$

Hospital:

$$\frac{dH^{ijk}}{dt} = \sigma_H(1-\sigma_C)\gamma_{S1}(C^{ijk} + Z^{ijk}) - \pi_H H^{ijk} - \mu H^{ijk} + \mu H^{(i-1)jk} + q_H \xi_H \tag{A.15}$$

Pre-ICU:

$$\frac{dH1^{ijk}}{dt} = \sigma_H \sigma_C \gamma_{S1}(C^{ijk} + Z^{ijk}) - \pi_A H1^{ijk} - \mu H1^{ijk} + \mu H1^{(i-1)jk} + q_{H1} \xi_{H1} \tag{A.16}$$

ICU:

$$\frac{dI^{ijk}}{dt} = \pi_A H1^{ijk} - \pi_B I^{ijk} - \mu I^{ijk} + \mu I^{(i-1)jk} + q_I \xi_I \tag{A.17}$$

Post-ICU:

$$\frac{dH2^{ijk}}{dt} = (1-\kappa_I)\pi_B I^{ijk} - \pi_C H2^{ijk} - \mu H2^{ijk} + \mu H2^{(i-1)jk} + q_{H2}\xi_{H2} \tag{A.18}$$

Recovered:

$$\begin{aligned}\frac{dR1^{ijk}}{dt} = (1-\phi_C)\Big((1-\phi_M)\left(\gamma_I G^{ijk} + \gamma_A(F^{ijk}+X^{ijk}) + \gamma_M(B^{ijk}+Y^{ijk})\right) \\ + (1-\phi_S)\left((1-\kappa_N)\gamma_{S2}N^{ijk} + (1-\kappa_H)\pi_H H^{ijk} + \pi_C H2^{ijk}\right)\Big) \\ +(1-\phi_P)(1-\kappa_P)\gamma_C P^{ijk} - \gamma_R R1^{ijk} - \mu R1^{ijk} + \mu R1^{(i-1)jk} + q_{R1}\xi_{R1}\end{aligned} \tag{A.19}$$

Recovered with long-term health complications:

$$\begin{aligned}\frac{dR2^{ijk}}{dt} = (1-\phi_C)\Big(\phi_M\left(\gamma_I G^{ijk} + \gamma_A(F^{ijk}+X^{ijk}) + \gamma_M(B^{ijk}+Y^{ijk})\right) \\ + \phi_S\left((1-\kappa_N)\gamma_{S2}N^{ijk} + (1-\kappa_H)\pi_H H^{ijk} + \pi_C H2^{ijk}\right)\Big) \\ +\phi_P(1-\kappa_P)\gamma_C P^{ijk} - \gamma_R R2^{ijk} - \gamma_L R2^{ijk} - \mu R2^{ijk} + \mu R2^{(i-1)jk} + q_{R2}\xi_{R2}\end{aligned} \tag{A.20}$$

Post-acute COVID-19:

$$\begin{aligned}\frac{dP^{ijk}}{dt} = \phi_C\left(\gamma_I G^{ijk} + \gamma_A(F^{ijk}+X^{ijk}) + \gamma_M(B^{ijk}+Y^{ijk}) + (1-\kappa_N)\gamma_{S2}N^{ijk} + (1-\kappa_H)\pi_H H^{ijk} + \pi_C H2^{ijk}\right) \\ -\gamma_C P^{ijk} - \mu P^{ijk} + \mu P^{(i-1)jk} + q_P\xi_P\end{aligned} \tag{A.21}$$

Death:

$$\frac{dD^{ijk}}{dt} = \kappa_H\pi_H H^{ijk} + \kappa_I\pi_B I^{ijk} + \kappa_N\gamma_{S2}N^{ijk} + \gamma_L R2^{ijk} + \kappa_P\gamma_C P^{ijk} + q_D\xi_D \tag{A.22}$$

Eq. (A.1) accounts for the rate at which people flow into and out of the susceptible (S) compartment. Individuals leave the compartment as they become infected as a result of the interaction with infectious individuals through the force of infection term, λ^{ijk}. This term introduces nonlinearity to the model through the multiplicative nonlinear coupling between susceptible and infectious compartments (B, C, F, X, Y, Z), as in Eq. A.23.

$$\begin{aligned}\lambda^{ijk} = \beta\sum_{l=1}^{N_i}\sum_{m=1}^{N_j}\sum_{n=1}^{N_k}\frac{c^{ijklmn}}{N^{lmn}}\Big(rr_C\left(rr_A\left(A^{lmn}+F^{lmn}\right)+B^{lmn}+C^{lmn}\right) \\ + rr_Q\left(G^{lmn}+W^{lmn}+X^{lmn}+Y^{lmn}+Z^{lmn}+N^{lmn}\right) + rr_H\left(H^{lmn}+H1^{lmn}+I^{lmn}+H2^{lmn}\right)\Big)\end{aligned} \tag{A.23}$$

where c^{ijklmn} is a tensor that quantifies the average number of daily interactions between individuals indexed by $\{i,j,k\}$ and those indexed by $\{l,m,n\}$. The terms rr_C and rr_Q are factors accounting for reductions in daily contacts due to social distancing and quarantining policies, respectively, and hence may vary in time as restrictions are imposed and removed as in Tuite et al.[1]. The force of infection term differs from [1], as follows: (i) the force of infection λ may explicitly vary in time to capture the changing reproduction number due to temporal variations in β (for example due to seasonal trends in transmission, human mobility patterns, and the implementation of control measures) whereas in [1], this term was a static value, multiplied by a random process that was meant to induce volatility in the transmission rate in time [3] [4]), (ii) with the inclusion of asymptomatic individuals, a

reduction factor, rr_A, is introduced to allow potential reductions or increases in transmission rates for asymptomatic carriers, and (iii) the above allows for individuals within the hospital track (compartments N, H, $H1$, I, $H2$) to infect susceptible individuals subject to a reduction factor rr_H to account for heightened precautions taken in hospital settings, but allowing for outbreaks to occur. The expression from Tuite et al. [1] can be recovered by setting those factors equal to one and zero, respectively.

The second term in Eq. (A.1) accounts for the individuals who leave the susceptible compartment as they are vaccinated and who return once the effective duration of their vaccine has elapsed. The third term quantifies people who re-enter the susceptible compartment after a period of temporary immunity from having recovered from the disease. The final terms account for individuals who remain susceptible, but are aging in and out of the current age index, i. To ensure that the total population is conserved (i.e., the sum of all 22 compartments across all indices is equal to unity for all time instances), individuals are introduced to the susceptible compartment (S) of age group $i = 1$ at a rate equal to the aging out of members from all compartments of age group $i = N_i$ as in [1]. The interpretation of all other equations follows the same reasoning with the flow in and out of the respective compartments having a one-to-one correspondence with the directed arrows in Figure 1 in the manuscript.

In a discrete state-space form, the model equation can be represented as in Eq. (A.24),

$$\mathbf{u}_{t+1} = \mathbf{g}_t(\mathbf{u}_t, \mathbf{q}_t; \boldsymbol{\phi}), \tag{A.24}$$

where $\mathbf{g}$ is the model operator and $\mathbf{u}$ is the state vector containing all of the above-mentioned $22 \times N_i \times N_j \times N_k$ model states. $\mathbf{q}$ is a vector consisting of the model error associated with each model equation, which corresponds to the red highlighted terms in Eqs.(A.1 - A.22). $\boldsymbol{\phi}$ is a vector of uncertain time-invariant parameters, meaning the subset of parameters from Table 2 that are to be estimated.

Similarly, the model must be accompanied by data; the measurement equation is given in Eq. (A.25),

$$\mathbf{y} = \mathbf{h}_k(\mathbf{u}_{t(k)}, \boldsymbol{\varepsilon}_k; \boldsymbol{\phi}), \tag{A.25}$$

which states that measurements, $\mathbf{y}$, of the state, $\mathbf{u}$, are obtained according to the model operator, $\mathbf{h}$, at time, $t(k)$, and are corrupted by noise, $\boldsymbol{\varepsilon}$. Note that in our case, the dimension of $\mathbf{y}$ will not be equal to that of $\mathbf{u}$, as a number of the states are not directly observable. The above model and measurement equations are necessary for the concurrent estimation of model error, state, and parameters [5, 6, 7, 8].

The initial conditions required for simulations consist of a single snapshot of all model states at a given instant because the model is a system of coupled first-order ordinary differential equations. These initial conditions will be subject to uncertainty, and thus may need to be . As the forcing of the dynamics are driven by the interaction of infectious individuals with susceptible individuals, a nonzero initial condition must be assigned in one of the infectious or exposed compartments (as it is intermediate to susceptible and infectious). This initial condition would model the immigration of a small number of infectious individuals into the region of interest and initiate community transmission through forward iterations of the state-space model.

The increased number of states and the inclusion of time-varying parameters and stochastic source terms has significantly increased the complexity of the model. Thus, it would be beneficial to study the stability of the model itself [9] prior to investigating the inverse problem. Understanding the influence of parametric uncertainty on various quantities of interest by conducting global sensitivity analysis (GSA) [10] would also be beneficial to facilitate the reduction of dimension of the uncertain parameter space.

Table 1: Model states/compartments (indices omitted for brevity)

Symbol	Definition
S	Susceptible
V	Vaccinated
E	Exposed
Q	Exposed, isolating
A	Infectious, pre-symptomatic
W	Infectious, pre-symptomatic, isolating
F	Infectious, asymptomatic
B	Infectious, mild-to-moderate symptomatic (i.e., symptoms not requiring hospitalization)
C	Infectious, severe symptomatic (i.e., symptoms requiring hospitalization)
X	Infectious, asymptomatic, isolating
Y	Infectious, mild-to-moderate symptomatic, isolating
Z	Infectious, severe symptomatic, isolating
G	Infectious, mild-to-moderate symptomatic, isolating but not previously in isolation
N	No access to hospital care
H	Hospitalized, never to be admitted to the intensive care unit (ICU)
$H1$	Hospitalized, to be admitted to the ICU
I	Hospitalized, in the ICU
$H2$	Hospitalized, after being discharged from the ICU
$R1$	Recovered, without long-term health complications
$R2$	Recovered, with long-term health complications
P	Post-acute COVID-19
D	Death

B Nonlinear sparse Bayesian learning algorithm at a glance

The nonlinear sparse Bayesian learning algorithm seeks to obtain parameter posterior distributions in a hierarchical Bayesian setting. The parameter posterior distribution is written as

$$\mathrm{p}(\boldsymbol{\phi}|\mathscr{D},\boldsymbol{\alpha}) = \frac{\mathrm{p}(\mathscr{D}|\boldsymbol{\phi})\mathrm{p}(\boldsymbol{\phi}|\boldsymbol{\alpha})}{\mathrm{p}(\mathscr{D}|\boldsymbol{\alpha})}, \tag{B.1}$$

where $\mathrm{p}(\mathscr{D}|\boldsymbol{\phi})$ is the likelihood function, $\mathrm{p}(\boldsymbol{\phi}|\boldsymbol{\alpha})$ parameter prior distribution, and $\mathrm{p}(\mathscr{D}|\boldsymbol{\alpha})$ is the model evidence. As in Eqs. (A.24 and A.25), $\boldsymbol{\phi}$ is the set of uncertain model parameters, $\mathscr{D}$ represents the data, which is a realization of $\mathbf{y}$ from Eq. (A.25), and $\boldsymbol{\alpha}$ is a hyperparameter, which we will examine next.

This algorithm leverages the concept of automatic relevance determination (ARD) to induce sparsity among the parameters. The set of parameters is decomposed into two distinct subsets, parameters that are known to be relevant to the observed dynamics (denoted $\boldsymbol{\phi}_{-\alpha}$), and parameters whose relevance is questionable (denoted $\boldsymbol{\phi}_{\alpha}$). The parameters that are known to be relevant are assigned prior distributions as in a standard Bayesian approach, whereas questionable parameters are assigned ARD priors which are zero mean normal distributions with precision controlled by the hyperparameter $\boldsymbol{\alpha}$. The ARD priors are independent; hence, the parameter prior can be written as

$$\mathrm{p}(\boldsymbol{\phi}|\boldsymbol{\alpha}) = \mathrm{p}(\boldsymbol{\phi}_{-\alpha})\mathrm{p}(\boldsymbol{\phi}_{\alpha}|\boldsymbol{\alpha}) = \mathrm{p}(\boldsymbol{\phi}_{-\alpha})\mathscr{N}(\boldsymbol{\phi}_{\alpha}|\mathbf{0},\mathrm{diag}(\boldsymbol{\alpha})^{-1}). \tag{B.2}$$

The algorithm optimally selects the values of the elements of $\boldsymbol{\alpha}$, by selecting the values based on the available data, which maximize the model evidence from Eq. (B.1). A parameter whose precision is determined to be large, results in a highly peaked Gaussian distribution centered at zero. As the precision tends to infinity, the prior pdf tends to a Dirac delta function at zero, effectively fixing the parameter to a value of zero and removing it from the analysis. This results in a data-optimal sparse model. The computational efficiency of this algorithm is achieved by substituting the prior pdf from Eq. (B.2) into Eq. (B.1), and creating a Gaussian mixture model of the known

Table 2: Model parameters with definition and reference (indices omitted for brevity)

Symbol	Definition
μ	Rate that individuals pass from age group i to $i+1$ (1/width of age group)
r_V	Vaccine effectiveness (1 indicates 100% immunity, 0 indicates no immunity)
δ_S	Probability that a susceptible individual exposed to the virus will self-isolate (without prior testing)
δ_V	Probability that a vaccinated individual exposed to the virus will self-isolate (without prior testing)
γ_A	1/the average duration of the infectious period for asymptomatic individuals
γ_C	1/the average duration of sub-acute COVID-19
γ_{DA}	Rate of detection among asymptomatic cases
γ_{DM}	Rate of detection among mild-to-moderate cases
γ_E	1/the average incubation period
γ_I	1/the average duration of self-isolation
γ_L	Rate of deaths due to long-term health complications
γ_M	1/the average duration of the infectious period for individuals with mild-to-moderate symptoms
γ_P	1/the average duration of the pre-symptomatic infectious period
γ_R	1/the average effective duration of temporary immunity from having recovered from the virus
γ_{S1}	1/the average duration of severe symptoms before seeking hospitalization
γ_{S2}	1/the average remaining duration of symptomatic period for individuals with severe symptoms
γ_T	1/the average effective duration of temporary immunity from vaccination
γ_V	Rate of vaccination
σ_A	Probability that an infectious individual is asymptomatic
σ_C	Probability that a hospitalized case will be admitted to the ICU
σ_H	Probability that an individual has access to hospital care
σ_S	Probability that a case displaying symptoms will require hospitalization
π_A	1/the average time in hospital prior to ICU
π_B	1/the average time in ICU
π_C	1/the average time in hospital following ICU
π_H	1/the average duration hospitalization (non-ICU track)
ϕ_A	Probability of acute COVID
ϕ_M	Probability of long-term complications for asymptomatic, mild-to-moderate cases
ϕ_S	Probability of long-term complications for severe cases
ϕ_P	Probability of long-term complications for post-acute COVID-19 cases
κ_H	Probability of death among hospital cases
κ_I	Probability of death among ICU cases
κ_N	Probability of death among cases without access to hospital care
κ_P	Probability of death among post-acute COVID-19 cases
Ψ	Parameter controlling the transition of vaccination indices
Υ	Parameter controlling the transition of comorbidity indices

prior $\mathrm{p}(\boldsymbol{\phi}_{-\alpha})$ times the likelihood function $\mathrm{p}(\mathscr{D}|\boldsymbol{\phi})$. This allows for the semianalytical computation of many Bayesian entities and quantities of interest that expedite the optimization of the hyperparameters as summarized in [11]. A detailed derivation of the algorithm is available in [12].

The NSBL algorithm (and more generally, ARD) has some apparent similarities with the aforementioned GSA, particularly as both methods can be used for model reduction. However, they are two distinct concepts with fundamentally different approaches to model reduction which may be complimentary to one another. Some key differentiating features of the two methods are as follows: (i) GSA considers a single output quantity of interest at a time, whereas NSBL may involve assimilating observations of multiple quantities of interest which are functions of the model states, (ii) in GSA, the input parameters' uncertainty is predefined (akin to a prior distribution), whereas NSBL considers both prior knowledge and available data, and (iii) for model reduction, GSA does not consider model complexity, whereas NSBL explicitly handles model complexity through evidence optimization.

C Modelling the spatio-temporal spread through partial differential equations

In order to extend Eqs. (A.1 – A.22) to model the spatio-temporal spread of COVID-19, the model are to be interpreted as population densities. The following nine compartments (indicated by their associated symbol from Table 1) would require the addition of a diffusion term to account for the movement of individuals in those compartments in space: S, V, E, A, F, C, B, P, $R1$, $R2$. This implies that the remaining 12 model compartments do not move in space, which makes use of the fact that compartments Q, W, X, Y, Z, and G on the quarantine track and N, H, $H1$, I, $H2$ on the hospital track will remain fixed in place until they proceed to one of the recovered compartments, or to the death compartment, D, where they will remain fixed permanently.

To illustrate how the form of the ordinary differential equations can be extended to partial differential equations, we take Eq. (A.3), which models infectious, pre-symptomatic individuals as an example. In Eq. (C.1), we see that the form of the equation is largely unchanged, a diffusion term is added (in blue text), and the state variables (uppercase letters) are now population densities in space and time. The model error term, ξ_E, may now also be modelled as a spatio-temporal white or coloured noise process,

$$\frac{\partial E^{ijk}}{\partial t} = (1-\delta_s)\lambda^{ijk}S^{ijk} + (1-\delta_v)(1-r_v)\lambda^{ijk}V^{ijk} - \varepsilon E^{ijk} - \mu E^{ijk} + \mu E^{(i-1)jk} + \nabla\cdot(\nu_E \nabla E) + q_E\xi_E \tag{C.1}$$

where ν_E is a diffusion coefficient for compartment E; one such parameter exists per model state. For more details on partial differential equation-based compartmental models, see [13, 14, 15], which describe spatio-temporal compartmental models for COVID-19. For details on the implementation of similar large-scale problems, see [16, 17].

References

[1] A.R. Tuite, D.N. Fisman, and A.L. Greer. Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada. CMAJ, 192(19):E497–E505, 2020.

[2] L.J.S. Allen. A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. Infectious Disease Modelling, 2(2):128–142, 2017.

[3] A. Camacho, A. Kucharski, Y. Aki-Sawyerr, M.A. White, S. Flasche, M. Baguelin, T. Pollington, J.R. Carney, R. Glover, E. Smout, et al. Temporal changes in Ebola transmission in Sierra Leone and implications for control requirements: a real-time modelling study. PLoS currents, 7, 2015.

[4] J.O. Lloyd-Smith, S.J. Schreiber, P.E. Kopp, and W.M. Getz. Superspreading and the effect of individual variation on disease emergence. Nature, 438(7066):355–359, 2005.

[5] M. Khalil, A. Sarkar, and S. Adhikari. Nonlinear filters for chaotic oscillatory systems. Nonlinear Dynamics, 55(1):113–137, 2009.

[6] P. Bisaillon, R. Sandhu, M. Khalil, C. Pettit, D. Poirel, and A. Sarkar. Bayesian parameter estimation and model selection for strongly nonlinear dynamical systems. Nonlinear Dynamics, 82(3):1061–1080, 2015.

[7] M. Khalil, A. Sarkar, S. Adhikari, and D. Poirel. The estimation of time-invariant parameters of noisy nonlinear oscillatory systems. Journal of Sound and Vibration, 344:81–100, 2015.

[8] R. Sandhu, D. Poirel, C. Pettit, M. Khalil, and A. Sarkar. Bayesian inference of nonlinear unsteady aerodynamics from aeroelastic limit cycle oscillations. Journal of Computational Physics, 316:534–557, 2016.

[9] E. Tornatore, S.M. Buccellato, and P. Vetro. Stability of a stochastic SIR system. Physica A: Statistical Mechanics and its Applications, 354:111–126, 2005.

[10] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. Global sensitivity analysis: the primer. John Wiley & Sons, 2008.

[11] R. Sandhu, M. Khalil, C. Pettit, D. Poirel, and A. Sarkar. Nonlinear sparse bayesian learning for physics-based models. Journal of Computational Physics, 426:109728, 2021.

[12] R. Sandhu. Model Comparison and Sparse Learning of Nonlinear Physics-Based Models Using Bayesian Inference. PhD thesis, Carleton University, 2020.

[13] A. Viguerie, G. Lorenzo, F. Auricchio, D. Baroli, T.J.R. Hughes, A. Patton, A. Reali, T.E. Yankeelov, and A. Veneziani. Simulating the spread of COVID-19 via a spatially-resolved susceptible–exposed–infected–recovered–deceased (SEIRD) model with heterogeneous diffusion. Applied Mathematics Letters, 111:106617, 2021.

[14] A. Viguerie, A. Veneziani, G. Lorenzo, D. Baroli, N. Aretz-Nellesen, A. Patton, T.E. Yankeelov, A. Reali, T.J.R. Hughes, and F. Auricchio. Diffusion–reaction compartmental models formulated in a continuum mechanics framework: application to COVID-19, mathematical analysis, and numerical study. Computational Mechanics, 66(5):1131–1152, 2020.

[15] P.K. Jha, L. Cao, and J.T. Oden. Bayesian-based predictions of COVID-19 evolution in Texas using multispecies mixture-theoretic continuum models. Computational Mechanics, 66(5):1055–1068, 2020.

[16] V. Dolean, P. Jolivet, and F. Nataf. An Introduction to Domain Decomposition Methods: Algorithms, Theory, and Parallel Implementation. SIAM, 2015.

[17] A. Desai, M. Khalil, C. Pettit, D. Poirel, and A. Sarkar. Scalable domain decomposition solvers for stochastic PDEs in high performance computing. Computer Methods in Applied Mechanics and Engineering, 335:194–222, 2018.