Toward a comprehensive system for constructing compartmental epidemic models

Darren Flynn-Primrose[1], Steve Walker[1], Michael Li[1,2], Benjamin M. Bolker[1,3*], David J. D. Earn[1]

[1*]Department of Mathematics & Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario, L8S 4K1, Canada.
[2]Public Health Risk Science Division, National Microbiology Laboratory, Public Health Agency of Canada, 370 Speedvale Avenue West, Guelph, Ontario, N1H 7M7, Canada.
[3]Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario, L8S 4L8, Canada.

*Corresponding author(s). E-mail(s): bolker@mcmaster.ca;

Keywords: epidemiology, transmission dynamics, compartmental model, graph theory, Cartesian product

MSC Classification: 92D30 , 92-10 , 05C76

1 Introduction

The COVID-19 pandemic has reemphasized the importance of compartmental epidemic models (Abou-Ismail, 2020; Massonis et al, 2021; Adam, 2020; Currie et al, 2020; Lofgren et al, 2014; McBryde et al, 2020; Enserink and Kupferschmidt, 2020) and has resulted in a flood of new compartmental models (e.g., Friston et al (2020); Fields et al (2021); Chang and Liu (2022); Lavielle et al (2020); Balabdaoui and Mohr (2020); Leontitsis et al (2021); Levine and Earn (2022)). This abundance of new model variants is expected given the number of public health modelers trying to use scientific understanding of emerging infectious diseases to contribute to public policy. Modelers must be able to build models rapidly to explore scenarios and generate high quality forecasts; public health recommendations have the biggest impact if they can be acted on promptly. However, the speed at which modelers can develop new models typically trades off with model quality. We therefore need tools that allow modelers to build models more quickly without sacrificing quality.

One approach to this speed-quality trade-off is to build infectious disease models incrementally. Information is scarce early in an epidemic, and so early models should be simple to reflect ignorance. As epidemics progress, we learn more about the characteristics of the pathogen and its transmission; at the same time the public health landscape becomes clearer but more complex. Because policy choices require fast input from scientists, modelers need to add complexity to their models quickly if they are to be relevant to policy.

Proceeding in this way eventually results in complex and fragile models, much of whose complexity is no longer relevant. This sort of complexity also makes it harder to add additional features to the model. Therefore, modelers need tools that make it easier to flexibly add and remove model structure.

Savageau (1988) and Voit (1988, 1990) made an early attempt to create such a toolbox by recasting the underlying differential equations of a model into a canonical

form they call an "S-model". Unfortunately this effort focused on the model's differential equations rather than its graphical structure, thus making it unsuitable for less mathematically inclined modelers. It does not seem to have been widely adopted.

Friston et al (2020) describe how the state space of a complex epidemiological model can be constructed from the product of different latent state dimensions (their Figure 1 shows an example with infection status, clinical status, testing status, and location), but the definition of which compartments are connected, and the rates of flow between them, is left up to the modeller.

A promising recent project to formalize the construction of compartmental models employs the language of category theory (Fong and Spivak, 2018; Libkind et al, 2022, 2021; Baez et al, 2023; Baez and Pollard, 2017). This powerful approach addresses many of the concepts we discuss here; however, at its current stage of development it requires considerable knowledge of advanced mathematics to use effectively. An ongoing project to implement the category theoretic approach in the Julia language can be found at https://github.com/AlgebraicJulia/AlgebraicPetri.jl (Halter et al, 2022).

Worden and Porco (2017) use the relatively simple language of graph theory to describe common methods of "multiplying" a set of compartmental models (***factor models***) into a new ***product model*** that incorporates the structure of all of its factors. The current paper is a result of the our efforts to build software implementations of the products described by Worden and Porco.

Procedure for Model Multiplication

We view model multiplication as a three-step procedure: JD: The simplification part is mentioned twice, but is pretty clearly not one of the three steps. Should simplification be the third step, and resolution be a fourth step?.

1. Generate the vertices of the product model by combining the vertices of the factor models. This typically means taking the Cartesian product of the vertices of factor

models. In many cases, we will want only a subset of the Cartesian product in the final model (e.g., some combinations are physically or biologically impossible).

2. Generate the edges of the product model. Again, we will typically take the Cartesian product of edges in each factor model with the vertices in the other. Some transitions may be disallowed, in which case we would drop those edges from the product or set the flows across them to zero. In other cases, we may want to add edges to the product to allow state changes in multiple strata to occur simultaneously.
3. Resolve ambiguities in how flow functions are generalized to accommodate the presence of additional strata; establish parameters for the product model, if possible by appropriately combining the parameters of the factor models.

The combination of the first two steps is a graph-theoretic Cartesian product. Recognizing the need to adjust flows in the second step above is Worden and Porco's contribution. JD: It is not clear where in the second step flows are adjusted. This requirement can arise in a number of ways; for example, when combining a standard SIR model with some other form of structure, we need to decide whether the susceptible population of any particular stratum can be infected by the infectious populations of other strata. If the stratification is based on age groups then it is reasonable to allow cross-infection, i.e., old people can infect young people and *vice versa*. BB: forward reference to thinking about transmission matrices, if we talk about them? In other cases we might prohibit cross-stratum infection, for example by allowing infection within but not between geographic regions. Our approach follows Worden and Porco's; when computing the magnitude of a flow between compartments we separately compute the contribution from each individual stratum and then sum the resulting quantities to find the total magnitude of the flow (we discuss other possibilities in Section 4.1). One could imagine that a product model for infectious disease is completely defined once we have specified the rules for cross-infection, but in practice a practical and reasonably general framework for combining models must consider many other details. JD:

Still feels a bit rambling; maybe keep things as general as possible here and do the details later.

2 What is a Compartmental Model?

BB: make sure that there is a payoff to concepts we're introducing here, i.e. that they will be further discussed/used in some concrete sense later in the paper (in particular, labeled partitions get mentioned *once* before the conclusion ...) JD: I feel like we don't need the labeled partitions terminology, certainly not here. Am tempted to edit it out. Is anyone taking a first-author-like role here, or are we just a committee for now? Should we try to have some sort of meeting where we make precipitate decisions once we've all been through?

Figure 1 shows two factor models: the "Epi" model represents individuals in a population being infected and then recovering from some infectious disease; the "Age" model represents individuals aging. Our goal is to design an algorithm to generate an appropriate product of these two factors (Figure 2). Each compartment in the product model inherits labels from one compartment in each of the factor models (e.g. SY = "susceptible, young"); the flows also typically have two labels (e.g. old infection flow). This system of "labeled partitions" (so called because each collection of labels will partition the compartments or flows into distinct sets) is useful for constructing

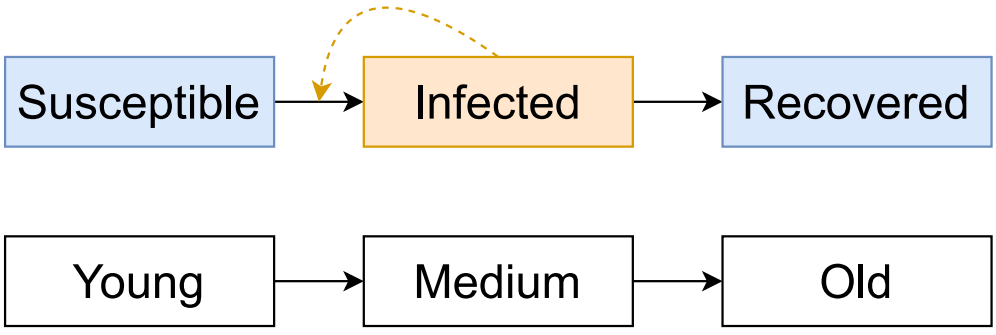

Fig. 1 A standard SIR model and a simple age stratification model. Orange shading denotes the infectious compartment, while blue shading denotes non-infectious compartments.

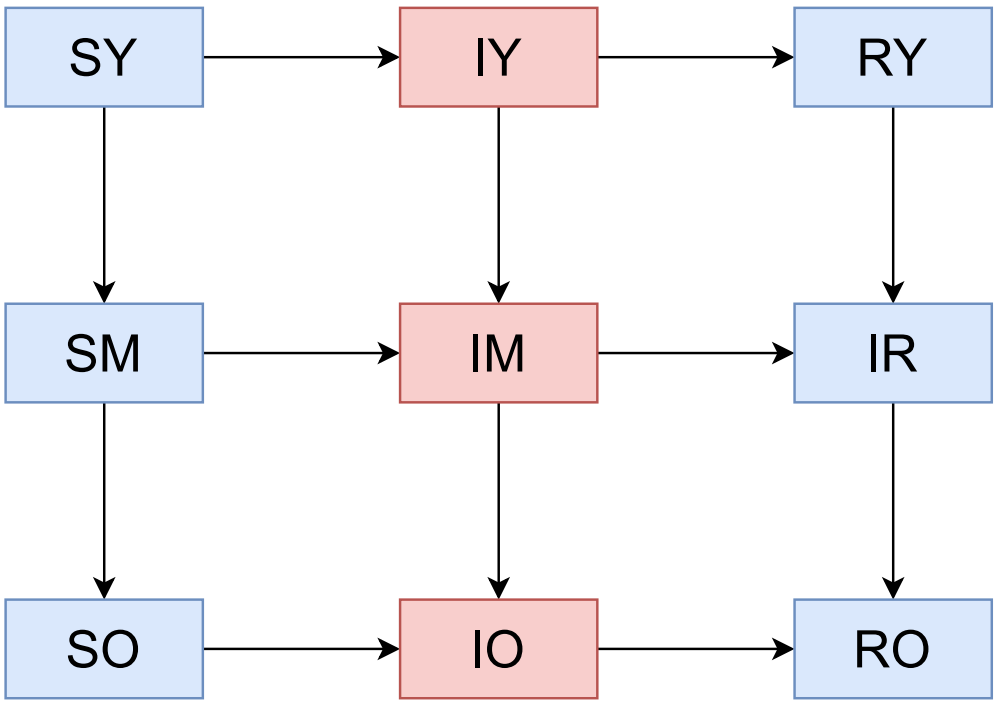

Fig. 2 The product of the two models shown in Figure 1. JD: These two panels should be combined into one Figure for readability, I think.

complex models. For example, instead of specifying each infection flow individually we can simply say that every susceptible compartment should have a flow to the infected compartment that shares the same age-related label.

Infection flows and infectious compartments

All of the flows in the Age model are constant *per capita* flows; that is, the rate of flow between any two compartments is strictly proportional to the number of individuals in the "from" compartment. The flow from "infected" to "recovered" in the Epi model is also a constant *per capita* flow. In contrast, the (total) flow from Susceptible to Infected is typically given by βSI or $(\beta SI)/N$ where β is a transmission parameter and N is the total population size. The *per capita* flow rate depends on occupancy of another compartment, as shown by the dashed arrow in Figure 1. In the specific

context of compartmental models for epidemiological dynamics, *per capita* flows that are state-dependent will generally relate to the infection (transmission) process; we will call these flows "infection" flows, and call the compartments that determine their rates "infectious" compartments.

BB: side comment, interesting connections to next-gen matrices: there we need flows into "infected compartments" and out of that class of compartments, where "infected" means "potentially infectious now or in the future" (from the same initial infection) In the SIR model the infectious compartment is the "to" compartment of the infection flow; however, this relationship does not hold in all models. In the SEIR model, for example, the infection flow goes to the "exposed" compartment, people who have been infected but cannot yet infect others. Models may also have multiple infection flows and infectious compartments — for example separating the infected population on the basis of symptom severity (mild vs. severe), in which case there would be an infection flow to every infected compartment. (Figure 2 does not include dashed arrows connecting infection flows with infectious compartments; these relationships will be discussed in Section 3.)

One-sided flows (sources and sinks)

Consider the long-term outcome of the "Age" model in Figure 1. Since the population is closed and people only flow from younger compartments to older ones, the entire population will eventually accumulate in the "old" compartment. If we want to simulate the system over time scales comparable to the host lifespan, we need to find a way to add and remove people from the model ("vital dynamics"). This can be done with one-sided flows, "sources" (flows with no "from" compartment") and "sinks" (flows with no "to" compartment). We could add a source flow into the "young" compartment and a sink flow out of the "old" compartment, representing the effects of

birth and death. A slightly more realistic model would add a sink for every age compartment, with a very low death rate for the youngest compartment and gradually increasing death rates with age.

Sources are also useful in models of viral presence in wastewater. The rate at which new viral particles enter the wastewater compartment is proportional to the number of infected people in the population. But infected people don't *become* contaminated wastewater; instead, we add a source to the model whose flow rate depends on the infected population.

Combining factor models with sources and sinks into product models may require extra decisions. In most models incorporating vital dynamics, for example, newborns entering the youngest age class are assumed to be susceptible (Earn, 2008), but in cases where vertical transmission or maternal immunity are possible, we might need to consider that they could enter the young/infected or young/recovered compartments instead. In contrast, mortality sinks could occur from any age or infectious class, although mortality rates could be higher in older age classes or infected classes (or highest in the combined old/infected class). Finally, in a model that included strata for age and geographic location, we would want birth flows to always enter the the youngest compartment for each location.

Absolute and* per capita *flows

BB: There is an ambiguity here. Is a *per capita* flow always defined relative to a "from" compartment? If so, then we should mention that sources must always be p.c. flows. The discussion below hints that if we defined birth as $\mu N = \mu(S+I+R)$, rather than a constant b, we might consider it a p.c. flow, but this seems wrong because (1) people aren't actually moving from (say) I to S when they produce offspring, (2) the rate of flow isn't dependent on the concentration of a "from" compartment. I'm not sure this matters in practice, but does birth proportional to N count as a p.c. or an absolute flow? Most of the flows discussed above are *per capita* flows, *i.e.* where the total flow

between compartments is given by a *per capita* rate multiplied by the population in the "from" compartment (for infection flows, the *per capita* rates themselves depend on the states of infectious compartments). In contrast, ***absolute flows*** are specified only in terms of the total flow. Births are often implemented as absolute flows (Earn, 2008). if population-level vaccination rates are recorded in public health data, it may be simplest to implement these as absolute flows (although specifying an absolute flow from a non-source population may sometimes lead to a mathematically ill-posed model (Gharouni et al, 2022).

BB: I commented out a paragraph about state ordering here; I don't understand why it's useful JD: It could be useful, and presumably resulted from some mistake made somewhere along the way. I'm fine without it, though.

Metadata

BB: do we need to know this? more generally, I think we only need to know about the internal machinery of macpan2 when it informs our ideas about model structure more generally

DF: In macpan2 this is taken care of by vectors automatically created when "Compartmental" is called. These vectors hold the index for the relevant compartments in the state vector. So for example "per_capita_from" is a vector with indices locating the "from" compartment of all *per capita* flows in the state vector. There is also a vector called "per_capita_flows" which locates all *per capita* flows in the flow vector. So to compute the magnitude of the *per capita* flows we can write "$state[per\_capita\_flows\_from] * flow[per\_capita\_flows]$. Note first that the multiplication above is element wise not a dot product and second that in macpan2 *per capita* flows that are sources or sinks are treated separately so the above only calculates the flow for *per capita* flows that are NOT sources or sinks.

BB: make sure we are treating vocabulary around nodes/compartments/states consistently In the underlying data structure for a compartmental model, we keep

Name	Notation	Description
from	$\vec{x}^{\text{from}}$	All elements of the state vector that are the "from" compartment for a flow. Note that this may be larger than $\vec{x}$ itself as some compartments may have more than one flow exiting from them.
to	$\vec{x}^{\text{to}}$	All elements of the state vector that are the "to" compartment for a flow. Note that this may be larger than $\vec{x}$ itself as some compartments may have more than one flow entering them.
inflow	$\vec{f}^{\text{in}}$	All flows in the flow vector that have an associated inflow (i.e. no sink flows)
outflow	$\vec{f}^{\text{out}}$	All flows in the flow vector that have an associated outflow (i.e. no source flows)
per capita	$\vec{f}^{\text{pc}}$	All *per capita* flows in the flow vector
absolute	$\vec{f}^{\text{abs}}$	All absolute flows in the flow vector
infectious	$\vec{x}^{\text{inf}}$, $\vec{f}^{\text{inf}}$	All elements of the given vector that are infectious
label	$\vec{x}^{label}$, $\vec{f}^{label}$	All elements of the given vector that share the specified *label* (e.g. "old", "Toronto", "vaccinated")

Table 1 Names, notation, and descriptions of metadata associated with a compartmental model. Note that $\vec{x}^{\text{from}}$ and $\vec{x}^{\text{to}}$ will typically be invoked with respect to a specific type of flow. $\vec{x}^{\text{pc, from}}$ in particular denotes the "from" compartments of all *per capita* flows.

track of compartments (nodes) and flows (edges) separately. Table 1 lists the information we need to track for edges and flows, including (1) which compartments are origins ("from") and destinations ("to") for particular flows; (2) whether a given flow is *per capita* or absolute; (3) whether flows have inflows and/or outflows (i.e., whether they are sources, sinks, or two-sided flows). In addition, we record which states and flows are infectious/infective. Finally, in particular models we may additionally track specific ***labels*** that correspond to useful distinctions for a particular modeling task: young vs. old, hospitalized vs. non-hospitalized, etc.. (While labeling compartments as "infectious" and "non-infectious" is the most important distinction for maintaining model structure, we may also want to consider "infected" vs. "non-infected" compartments; a person in a latent, or exposed, class would be infected but non-infectious.). The metadata vectors represent ordered subsets of the full sets of states or flows in the model; the ordering (which is kept consistent between sets of *per capita* rates and flows) allows us to do addition and elementwise multiplication of vectors. BB: I'm still not sure how we handle multiple outflows from a single compartment (say, flows from I to acute care and ICU); are there repeated elements in these vectors?

The dynamics of compartmental models are typically implemented using ordinary differential equations (ODES), although modelers also use other approaches such as discrete-time models or continuous-time Markov processes. No matter what dynamical framework is used, we need to compute the combined rates of flow among compartments. Depending on the framework, these may be combined into a total rate of change of a compartment (as in ODEs) or remain as rates of different processes (as in continuous-time Markov models).

BB: $\vec{}$ notation can be ugly, e.g. as in $\vec{f}^{\text{inf}}$ (collision between arrow and i-dot). Switch to boldface? (Could redefine 'vec' or the 'xvec'/'fvec'/'rvec' macros in ms_submit.tex)

When we compute overall rates of change $\vec{r}(\vec{x}, \vec{f})$, elements of $\vec{f}$ that correspond to *per capita* flows will be multiplied by the population of their "from" compartment before they are added to $\vec{r}$; absolute flows are added directly to $\vec{r}$. Elements in $\vec{f}$ that correspond to two-sided flows contribute twice to $\vec{r}$, as a decrease in the population of the "from" compartment and an increase in the "to" compartment; one-sided flows (from sources or to sinks) contribute only once.

The rate (or gradient) vector $\vec{r}$ is state and flow dependent, so must be recalculated at every time step in a dynamical model.

In a simple epidemic model the only state-dependent flows (i.e., flows that depend on compartments other than their "from" compartments) are the infection flows. We will typically assume that infection flows are additive; thus we can write

$$\vec{f}^{\text{inf}} = T\vec{x}^{\text{inf}} \tag{1}$$

where T is a transmission matrix with columns corresponding to infectious states and rows corresponding to infection flows. After finding the infected flows we separately calculate the total inflow and the total outflow of each compartment. The total inflow is

$$\vec{r}^{\text{total inflow}} = \vec{x}^{\text{pc, from}} * \vec{f}^{\text{pc, in}} + \vec{f}^{\text{abs, in}}$$

and the total outflow is

$$\vec{r}^{\text{total outflow}} = \vec{x}^{\text{pc, from}} * \vec{f}^{\text{pc, out}} + \vec{f}^{\text{abs, out}}$$

where $*$ denotes elementwise multiplication. BB: is there a standard term for a dynamical system expressed in terms of total changes in compartments rather than in changes due to particular processes? Finally, the total rate of change for every compartment (which we need in an ODE-like system) is

$$\vec{r} = \vec{r}^{\text{total inflow}} - \vec{r}^{\text{total outflow}}.$$

2.1 Parameterizing product models

In addition to understanding how to construct compartments and flows of a product model from the compartments and flows of its factors, we would like to be able to compute parameter values for the product model using the known values from the factor models. Product model parameters are often related to parameters in the original model factors in simple mechanistic ways. However, there is an enormous range of possible relationships between the parameters of the factor models and the parameters of their product. Some parameters, such as those describing intrinsic properties of a pathogen, may be constant across all strata of a product model. Others, such as recovery time, are constant with respect to some strata (e.g., location) but variable with respect to others (e.g., age). In other cases, we may want to simplify relationships due to data constraints — for example, we may know that recovery time varies with age, but choose to treat it as constant for modeling purposes. Ultimately, the question of how factor model parameters should be generalized to the product model depends on the intentions of the modeler. In principle we could develop a framework to deduce,

using only information present in the factor models, how the product model parameters should be related to their factor model equivalents, or at least to define a formal system for specifying which parameters vary across which strata. For now, however, our framework defaults to the most general possible case and leaves it to modelers to construct appropriate simplifications (mapping several stratum-specific parameters in the product model to a single value) for themselves. However we discuss a few common scenarios here for the purpose of illustration.

If a parameter describes processes that are purely pathogen-dependent it will be constant across strata that represent variation among hosts, and conversely for host-dependent parameters across pathogen strata. In other cases, the value of a parameter at each stratum in the product model may be derived from the factor model version of the parameter by a simple scalar. For example, it might be useful to parameterize the recovery time of hosts in different strata as proportional to some baseline value (e.g., a population average or the value in healthy adults). If α is the parameter value in the factor model and $\vec{\beta}$ denotes the values of the derived parameters at each different strata in the product model then $\vec{\beta} = \alpha\vec{w}$ where $\vec{w}$ is a vector of weights.

Relationships among factor model parameters can be constrained in other ways. When multiple flows emanate from a single compartment, we may want to parameterize them as a partition of the total outflow. For example, an exposed host may have multiple possible fates such as asymptomatic infection, mild infection, or severe symptoms. In that case the factor model in question will have three parameters $(\alpha_i \in (0,1), i = 1 \ldots 3)$ that sum to one $(\sum \alpha_i = 1)$. Every stratum of the corresponding product model will have three parameters derived from the original α values, but each stratum may have different partitions. For example people in different age groups may be more or less likely to experience severe, mild, or no symptoms. In this case, we

need to be careful to define the weights so as to make sure the constraints are still satisfied, e.g. with a softmax function ($\alpha_{ij} = \exp(a_{ij}) / \sum \exp(a_{ij})$, with some constraint such as $\alpha_{i1} = 1$ for identifiability).

Different strata of a product model may interact. BB: This was originally framed as being about SI models, but I don't see why we have to introduce a new model; the FOI formulation is general to most directly transmitted disease models ... A standard formulation for the force of infection of a pathogen (i.e., the per-susceptible rate of infection) is $\Lambda = \frac{\beta I}{N}$. Suppose we now stratify this model to represent a scenario where each person lives in one of three different locations but may come in contact with anyone in any location. Our model would then have three infected compartments ($\vec{I} = (I_1, I_2, I_3)$) and three susceptible compartments ($\vec{S} = (S_1, S_2, S_3)$; the force of infection is also a vector of location-specific values ($\vec{\Lambda} = (\lambda_1, \lambda_2, \lambda_3)$). In the most general case, where the force of infection does not take the standard form given above, each λ_i would be expressed as some function of the infected populations as well as a vector of parameters $\vec{\beta_i}$ which gives some information about how people at different locations interact with each other. Thus, we would be left with $\vec{\Lambda} = (f(\vec{\beta_1}, \vec{I}), f(\vec{\beta_2}, \vec{I}), f(\vec{\beta_3}, \vec{I}))$. When the force of infection is a linear equation with respect to the population of infected compartments as in eq. 1, we can be more specific: the factor model parameter β generalizes to a 3×3 matrix of transmission parameters $B = (\beta_{ij})$ so that we can write

$$\vec{\Lambda} = \frac{1}{N} B \vec{I}$$

While it preserves generality, this approach makes the number of required parameters increase quadratically with the number of locations (strata). Epidemiological modelers have devised many ways to add structure to the B parameter, sometimes in terms of a "who acquires infection from whom" (WAIFW) matrix Anderson and May (1985, 1992); Grenfell and Anderson (1985). In practice the likelihood of a person residing in one location coming into contact with a person somewhere else varies according to

the distance between the two locations. We can construct a contact matrix C that makes the rate of contact between people at two locations a (typically monotonically decreasing) function of distance, e.g. $c_{i,j} = e^{-\nu d_{i,j}}$ (or $c_{i,j} = 1/d_{i,j}^2$, in the case of a ***gravity model*** (Xia et al, 2004)). In this way we can write

$$\vec{\Lambda} = \frac{\beta}{N} C \vec{I}$$

which preserves the original meaning of the parameter β and only introduces one new parameter (ν) instead of nine.

There are, of course, other ways to handle this kind of parameter simplification. Most situations will allow for a parameter space mapping of this kind that relates the default parameter space generated by model products to a smaller parameter space dictated by the specific data available to the modeler. However, as we have tried to make clear in this subsection, there are so many potentially useful mappings that for now we will maintain generality by treating all parameters in the product model as independent.

In our general parameterization, the l_i parameters in the product model that come from the i^{th} factor model (of two) can be organized in a $l_i \times k_{2-i} \times k_{2-i}$ order-three tensor, $B^{(i)}$. The modeler will have some set of parameters known to them which we call $\vec{\theta}$ and will be able to compose, from a library of standard relations, a mapping g so that $B^{(i)}_{hij} = g^{(i)}_{hij}(\vec{\theta})$.

3 Cartesian Model Products

Returning to Figure 1, recall that the dashed arrow in the SIR model indicates that the flow from S to I is an infected flow and that its magnitude is partly determined by the number of infected people. When we combined the SIR and age models in Figure 2 we omitted any dashed arrows on the grounds that they could drawn in

several different ways. Worden and Porco (2017) describe two separate products, each with the dashed lines drawn differently. In their ***naive product***, each susceptible age group can only be infected by infectious people in the same age group (Figure 3). In their ***modified product***, each susceptible age group can be infected by infectious people in any age group (see Figure 4). The underlying graphs of these product models are identical; both are the Cartesian product of the factor model graphs. The only difference between them is the functional form of the infection flows in the product model (with respect to the procedure for model multiplication) BB: what???. The graph underlying a product model is not always the Cartesian product of factor model graphs: for example, Worden and Porco's ***strong product*** adds additional edges to the Cartesian product. In the model illustrated in Figure 2, individuals cannot move directly from the young susceptible compartment to the middle infected one; to make that transition they must either become infected and then age or age and then become infected. In both cases it takes a minimum of two time steps to complete the transition. In the strong product the (young, susceptible) compartment would have an edge leading directly to the (infected, middle) compartment so the transition could be done in a single time step. In a scenario where strata were determined by infection with multiple infectious pathogens, we might choose to use the strong product so people can be infected with both pathogens in a single time step. In a model implemented in continuous time, or with short discrete time steps, the chances of being infected by two pathogens simultaneously would be negligible and so the strong product would increase the complexity of the model while having minimal effect on the results. Alternatively there are some cases (Section 4.3) where the digraph underlying the product model is a proper subset of the Cartesian product of the factor models, for example in models with multiple pathogen strains that disallow the possibility of co-infection. For now, however, we will focus on products where the graph of the product model is the Cartesian product of the factor model graphs.

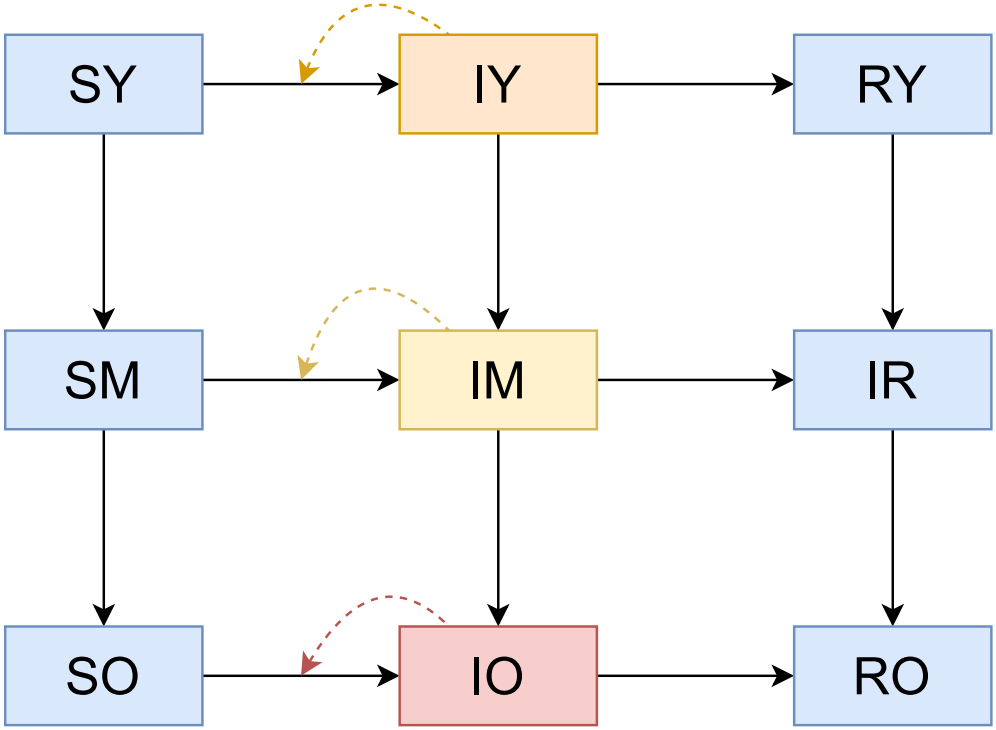

Fig. 3 The naive product of the two models from Figure 1. Blue denotes non-infectious compartments, yellow/orange/red denote infectious compartments. The force of infection is only influenced by the infected population within the same age stratum. In this example, people of different age groups have no contact (or very limited contact) with each other.

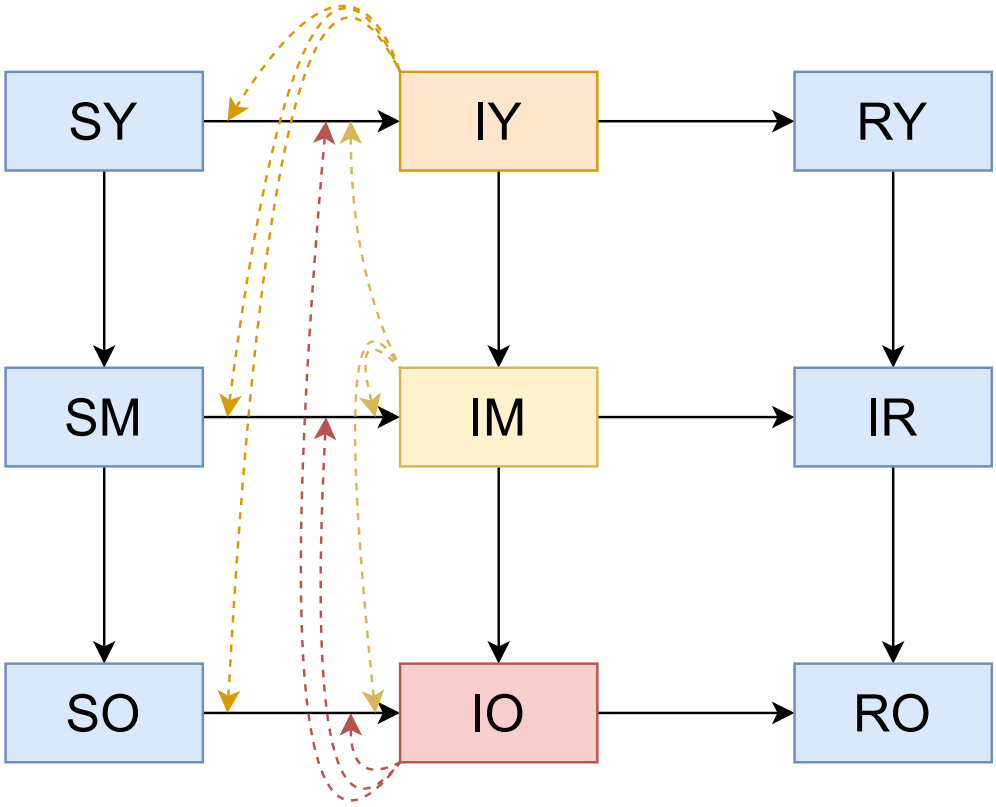

Fig. 4 The modified product of the two models from Figure 1. Unlike in Figure 3, individuals make epidemiological contacts across age strata, so the force of infection for each age stratum is influenced by the infected population in all age strata.

To illustrate the difference between naive and modified products consider an SIR model where the infection flow is given by $\vec{f}^{\text{inf}} = \frac{\beta I}{N}$ where N, the total population of the model, is constant. In this case the transmission matrix T from Equation 1 would be the one-by-one matrix $[\beta/N]$. When moving to the product model, we could either let T be given by the diagonal matrix $\text{Diag}(\beta_i/N)$ which would yield the naive product, or we could let T be given most generally by $[\beta_{ij}/N]$ (with all $\beta_{ij} > 0$), yielding the modified product.

In the age-stratified example the modified product is likely to be preferred because people of all ages commonly interact with each other. In the case of spatial stratification, one might want to use the naive or the modified product, or another alternative, depending on the specifics of the epidemiological system. At first glance the naive product seems appropriate because it encodes the assumption that people in different locations cannot infect one other because they cannot physically interact. This approach works well in product models that do simulate movement explicitly (e.g. Mohammadi et al, 2023), where the flows between different locations are included in the original factor model describing spatial structure. Spatial models that model movement implicitly (Dietz and Sattenspiel, 1995) compute the rates of contact between pairs of people who live in different locations during times when one or both are away from. In this case the modified product would seem the most appropriate.

The naive product restricts interaction to people in the same stratum; the modified product allows people in any strata to interact. We propose a new, generalized product that allows for people in each stratum to interact with people in an arbitrary subset of the other strata, allowing the creation of models where people at a given location can interact at the same location or neighboring locations, but not globally.

BB: A lot of the following section seems to be re-inventing network models. How specific is this to understanding product models? How much detail should we retain?

Below we show three different ways an SI model could be stratified with location. Figures 5 and 6 show the naive and modified products respectively. Figure 7 shows a generalized product where interactions can only occur within a single geographic region or between neighboring regions. For example, an infected person in the Toronto region could infect a susceptible person in Toronto or Ottawa but not one in Montreal, while the neighbourhood of an infected person in Ottawa includes all three regions.

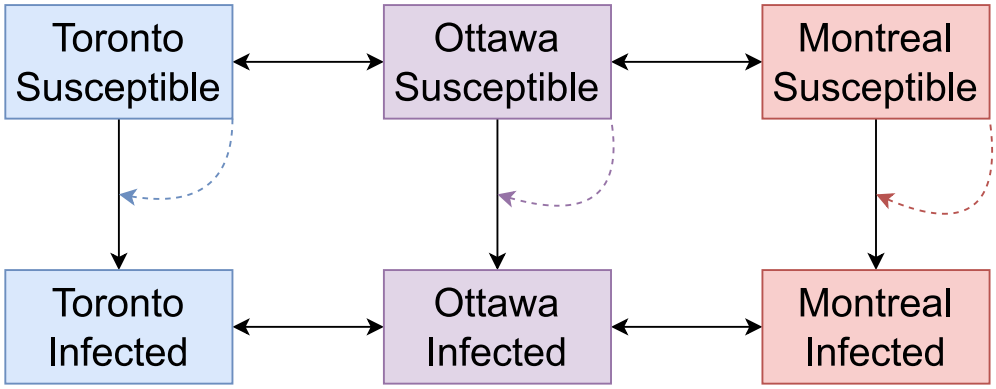

Fig. 5 The naive product of an SI model with location model including Toronto, Ottawa, and Montreal. The force of infection at any given location depends only on the infectious population at that location.

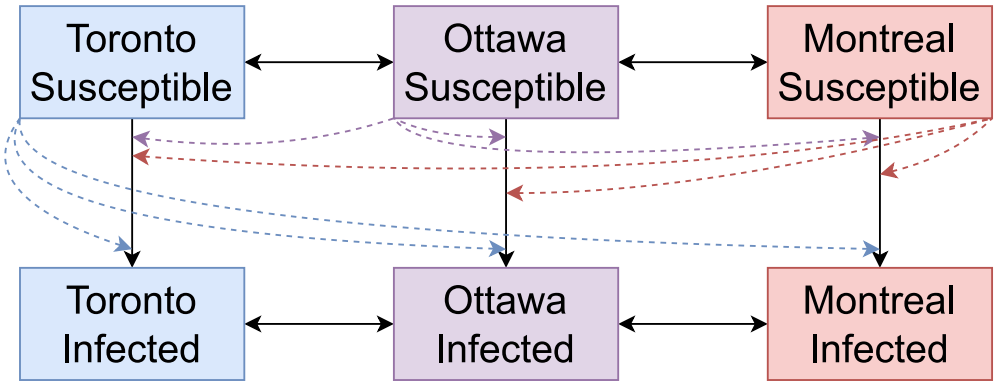

Fig. 6 The modified product of an SI model with location model including Toronto, Ottawa, and Montreal; the forces of infection depend on infectious populations in all three regions.

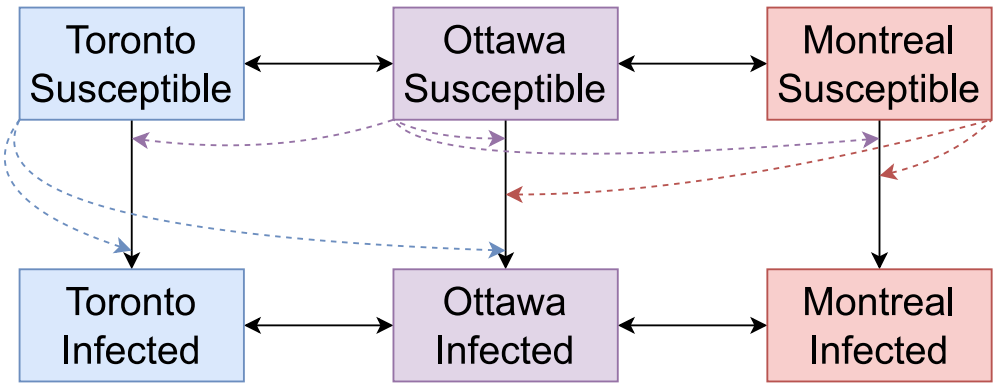

Fig. 7 A generalized product of an SI model with location model including Toronto, Ottawa, and Montreal. The forces of infection depend on the infectious population at the same location and in neighboring locations.

The only change from the contact matrices defined above for the naive and modified products is that in the case of the generalized product we would set $\beta_{31} = \beta_{13} = 0$ to reflect the impossibility of transmission between Montreal and Toronto. BB: I don't really get this. I agree that leaving out a computational term rather than retaining it as a division by zero is slightly better, but this seems fairly trivial/unimportant? Are we losing anything if we delete the following para?

Ultimately we could treat all such special cases by setting the value of β to zero in places where we don't want different strata to interact but in many cases this would result in doing a large amount of unnecessary computational work (e.g. multiplying things by zero) which can significantly increase to time required to simulate a model. For the sake of efficiency it is desirable to think of these products as distinct operations and always use the one that creates the least amount of unnecessary computational baggage.

4 Challenging Examples

While the operations defined above allow us to construct a wide range of compartmental models by taking products of simpler factor models, they cannot account for

every possible combination of models. In this section we discuss a number of examples where the model products we have defined so far are insufficient.

4.1 Models with alternate functional forms

BB: I think this section can be simplified/clarified/better motivated, but not sure how yet (start with examples? STD models with varying mixing rates? Frequency-dependent infection? Are we really thinking about incidence here, or could we equivalently talk about forces of infection?)

Up till now we have assumed that the force of infection on a particular susceptible compartment (and hence the absolute rate of infection of those susceptibles) is a linear function of the numbers of people in the infectious compartments ($\vec{\Lambda} = \frac{\beta}{N} C \vec{I}$). This formulation makes two related assumptions: (1) that the force of infection due to each infectious compartment is strictly proportional to that compartment's occupancy, and does not depend on any other compartments; (2) that the forces of infection can be summed to get the total force of infection.

BB: still trying to figure out how to put this section together sensibly. Some practical examples that occur to me:

- a behavioural model where the FOI decreases with increasing hospital occupancy (or hospitalization rate, or mortality rate, or ...). Now FOI will depend (nonlinearly) on non-infectious compartments. When we create the product model, are those influences local to a stratum or global (e.g., do people pay attention to the population-wide mortality rate, or only to the mortality rate in their age group?)
- incidences of the form $S^{\alpha_S} I^{\alpha_I}$
- frequency-dependent infection (with non-constant population sizes)
- F-D plus activity-weighted mixing, as in STD models: the relevant value is an activity-weighted average, $\sum_j \beta_{ij} c_j I_j / \sum c_j N_j$ (or something like that)

Can we connect what's already here with these ideas? we could write the force of infection as $f(\vec{\beta}, \vec{x})$ where $\vec{\beta}$ is a parameter vector and $\vec{x}$ is the complete state vector. Than suppose we stratified this factor model by age so the state vector of the product model is $\vec{x} = (\vec{x}_1, \ldots, \vec{x}_n)$ where the subscripts denote components that belong to separate age strata. Then the *per capita* rate rate of infection for age stratum i in the product model will be $f(\vec{\beta}_{i1}, \vec{x}_1) + \ldots + f(\vec{\beta}_{in}, \vec{x}_n)$. An alternate approach would be to instead take a weighted average of the compartment populations in each stratum and use this new average as the input to the flow rate function. Using this idea the *per capita* rate of new infections for age strata i would be $f(\vec{\beta}, w_{i1}\vec{x}_1 + \ldots + w_{in}\vec{x}_n)$ where the w's are weights. This approach is particularly useful when incorporating inhibitory influences in a model. For example, during an epidemic, individuals will be more careful if they know hospitals are at capacity than they would be when there are ample medical resources available.

The weighted states approach is equivalent to the summation method provided f is a linear function. One important instance where this will not be the case is if f involves normalizing by the total population of the model and that population is not constant. To see this let $N(\vec{x})$ be a function that sums every component in a vector and let $f(\vec{\beta}, \vec{x}) = \frac{\vec{\beta} \cdot \vec{x}}{N(\vec{x})}$. In this example the summation method would produce

$$\frac{w_1\vec{\beta}_i \cdot \vec{x}_1}{N(\vec{x}_1} + \ldots + \frac{w_n\vec{\beta}_i \cdot \vec{x}_n}{N(\vec{x}_n)} \tag{2}$$

but the weighted states method would produce

$$\frac{\vec{\beta}_i \cdot (w_1\vec{x}_1 + \ldots + w_n\vec{x}_n)}{N(w_1\vec{x}_1 + \ldots + w_n\vec{x}_n)} \tag{3}$$

Notice that in Equation 2 each term in the sum is divided by the population of a single stratum whereas in Equation 4.1 every term in the numerator is divided by the total (weighted) population of the entire model.

Another case where the two approaches may differ is when using non-linear incidence rates. Typically in an SIR model the *per capita* rate of infection is given by $\frac{\beta I}{N}$ however in some cases it might be desirable to use $\frac{\beta S^\kappa I}{N^{\kappa+1}}$. Here again the two approaches will produce different results.

Of course the two approaches are not mutually exclusive, we could find each stratum's contribution to the total number of newly infected people using the sum of weighted states rather than just the state vector for that specific stratum. In fact the summation approach is equivalent to doing that using the weights $w_j = \delta_{ij}$ where δ is the Kronecker delta function.

4.2 Models with Testing

One such example (where simple model products alone seem insufficient) involves modeling the effects of testing for infection, inspired by the dynamics of testing during the COVID-19 pandemic (Gharouni et al, 2022). Consider the epidemiological model in Figure 8 and the testing process depicted in Figure 9. The modified product of these two models includes a compartment for untested individuals at the hospital. However, this product is not what we want (Figure 10). The key difference is that untested individuals entering the hospital are typically tested (i.e., moved from "untested" to "awaiting results"); in our model, we will assume that they are always tested. Therefore, the "untested hospitalized" compartment in product model is always empty; the flow that goes to that compartment should instead be directed to the "hospitalized/awaiting test result" compartment. Constructing the desired model would thus require an extra step to redirect this flow and remove the superfluous compartment.

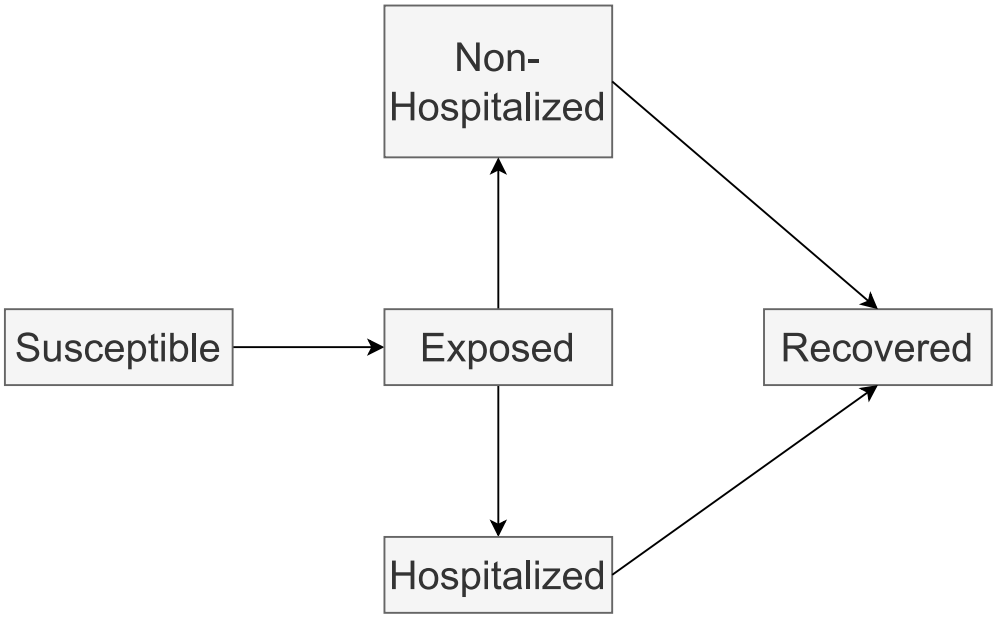

Fig. 8 A simple epidemiological model that we will expand to include testing. In this model, some exposed individuals will develop asymptomatic or mild illness, in which case they stay in the community during their infectious period (and potentially transmit to others); those who instead develop severe illness will be hospitalized. (This model allows neither for within-hospital transmission nor for disease-induced mortality either inside or outside the hospital.)

BB: can we condense/combine Figs 8-10 without making them impossible to read? (Fig 10 is pretty challenging in any case ...)

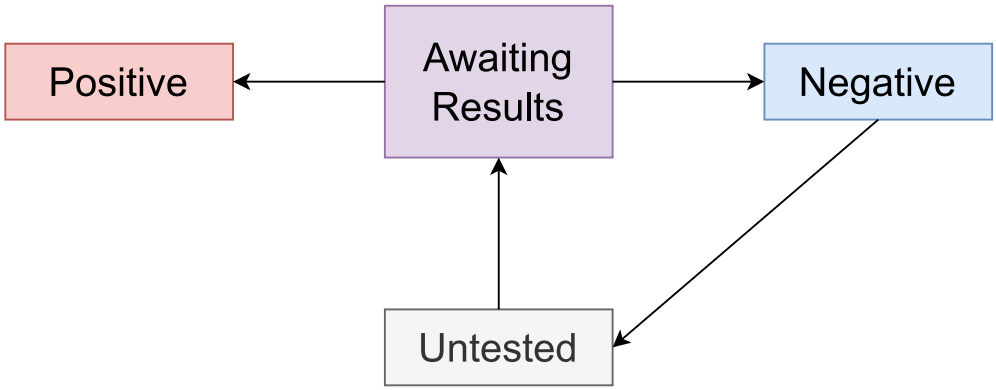

Fig. 9 A simple testing model. Individuals who test negative will, over time, revert back to the "untested" status. This is not the case for those that test positive; at least during the early stages of the COVID-19 pandemic, someone who had tested positive for COVID-19 would assume that they were immune and would not be re-tested even if they developed COVID-like symptoms.

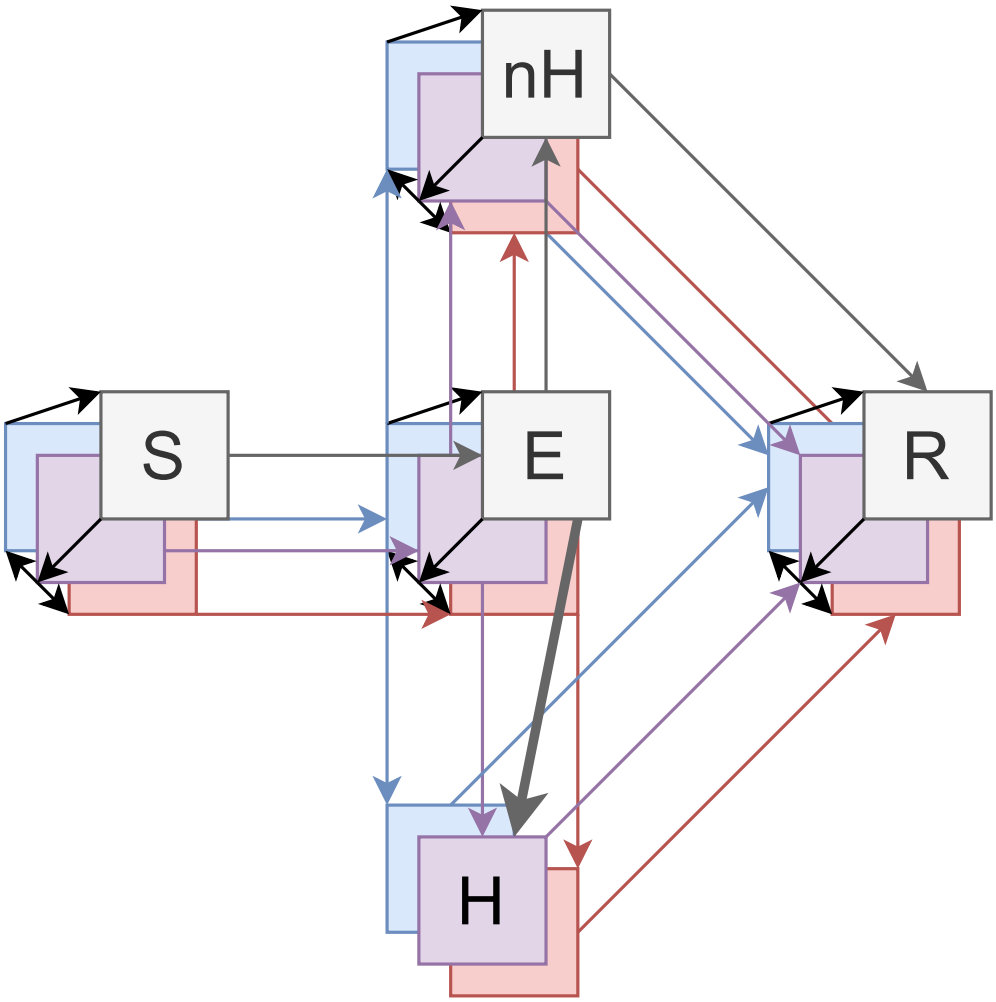

Fig. 10 The desired result of combining Figure 8 with Figure 9. Note the missing grey "untested" box associated with the hospital location; exposed individuals going into the hospital (enlarged, grey downward arrow starting at E) flow into the purple "awaiting results" subcompartment.

4.3 Multistrain Models and a Weak Product

Many epidemics involve multiple co-circulating strains of the same pathogen (Gog and Grenfell, 2002; Williams et al, 2021). In the case of COVID-19 such variants have significant implications for the efficacy of vaccines (Abu-Raddad et al, 2021; Koyama et al, 2020) and diagnostic tests (Vasireddy et al, 2021). In more complex models, including multiple strains rapidly inflates the size of both the state space and the parameter space (Kryazhimskiy et al, 2007). One way to limit the size of these unwieldy models while continuing to include the effects of multiple strains in our model is to disallow the possibility of superinfection (i.e., an individual being infected with multiple strains at the same time). BB: should we worry about distinguishing superinfection (infection A then B) vs coinfection (simultaneous inf with A and B)? It would therefore be useful to define a weak product similar to the operations proposed by Worden and Porco (2017) but which excludes all states corresponding to a superinfected status. One way to do this, which works well for two-strain models, is to use the standard Cartesian product but include only flows that emanate from compartments with no inflow (***sources***) or enter compartments with no outflow (***sinks***). Flows in this category typically represent initial infection (i.e., flows out of a susceptible class that is a source) or final recovery (i.e., flows into a recovered class that is a sink). In making this restriction, we exclude all the flows within the set of compartments that represent infected states, such as from an exposed to an infected compartment). For convenience we denote this operation by $\boxplus$ and call it the ***weak product***.

Figure 11 depicts a two-strain SIR model without superinfection, corresponding to the weak product of two SIR models. Unfortunately, the weak product is not associative! Figures 12 and 13 depict two different results for the weak product of three SIR models with different grouping. If the factor models for the three strains are M_1, M_2, M_3 then Figure 12 depicts $(M_1 \boxplus M_2) \boxplus M_3$ and Figure 13 depicts

$M_1 \boxplus (M_2 \boxplus M_3)$. Figure 14 depicts the desired result for a three-strain SIR model with no super-infection.

BB: I got a little bit confused by the flow here. The three-strain model, and why the weak product doesn't work in this case, is presented (are there other applications of the weak product other than multi-strain models? Seems a little limited if the *only* thing it is good for is two-strain models without superinfection). Then we say "but we can't do that with these products". Then we talk about labeled partitions. Then we present a *different* two-strain model, which I guess represents another problematic case for the weak product (but which can also be solved with labeled partitions)? The only thing we say about this model (in the last few lines of the body of the paper) is that it's problematic [and could be solved by labeled partitions?]

It is possible to create a version of the weak product defined above that will produce the model shown in Figure 14. However, it requires us to distinguish between compartments that are *global* sources or sinks and compartments that are sources or sinks with respect to one of the three strains specifically. That is to say, while a global sink must have no outflows, a weaker condition says that a compartment is a sink with respect to a specific pathogen if every compartment that can be reached via the outflow has the same infection status with respect to that pathogen as the original compartment. Programmatically we achieve this by introducing a concept of 'labeled partitions' which separates the vertices of the model into disjoint sets corresponding to the vertices' status with respect to a specific pathogen. Each dimension of stratification in the model corresponds to a different labeled partition with each stratum corresponding to a different disjoint set. In this way we can define sources and sinks with respect to a specific set of labels rather than globally. For example, we can say a compartment A is a sink with respect to a specific labeled partition if every compartment that can be reached after being in A is in the same set as A. Figure 15 outlines a compartmental model with one source compartment but two sink compartments and

Figure 16 shows the weak product of two such models. An unfortunate aspect of this construction is that several of the compartments can only be reached by individuals after they are already dead (!). If there are relatively few such compartments a modeler may choose simply to leave them in the model and treat them all as a single compartment. But if there are many such "zombie compartments", or if computational efficiency is a pressing concern, they could be removed from the model.

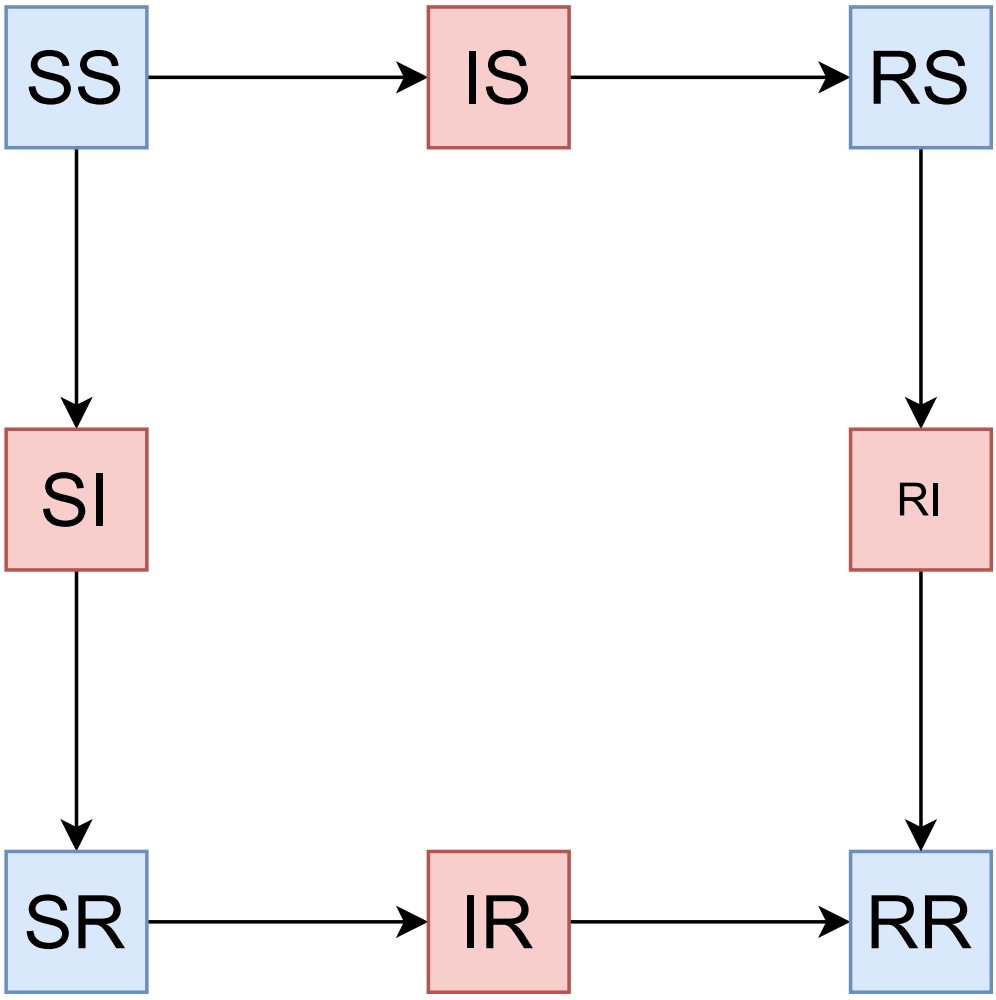

Fig. 11 A two-strain SIR model admitting no superinfection. Red Compartments indicate an infectious population whereas the population in blue compartments are not infectious

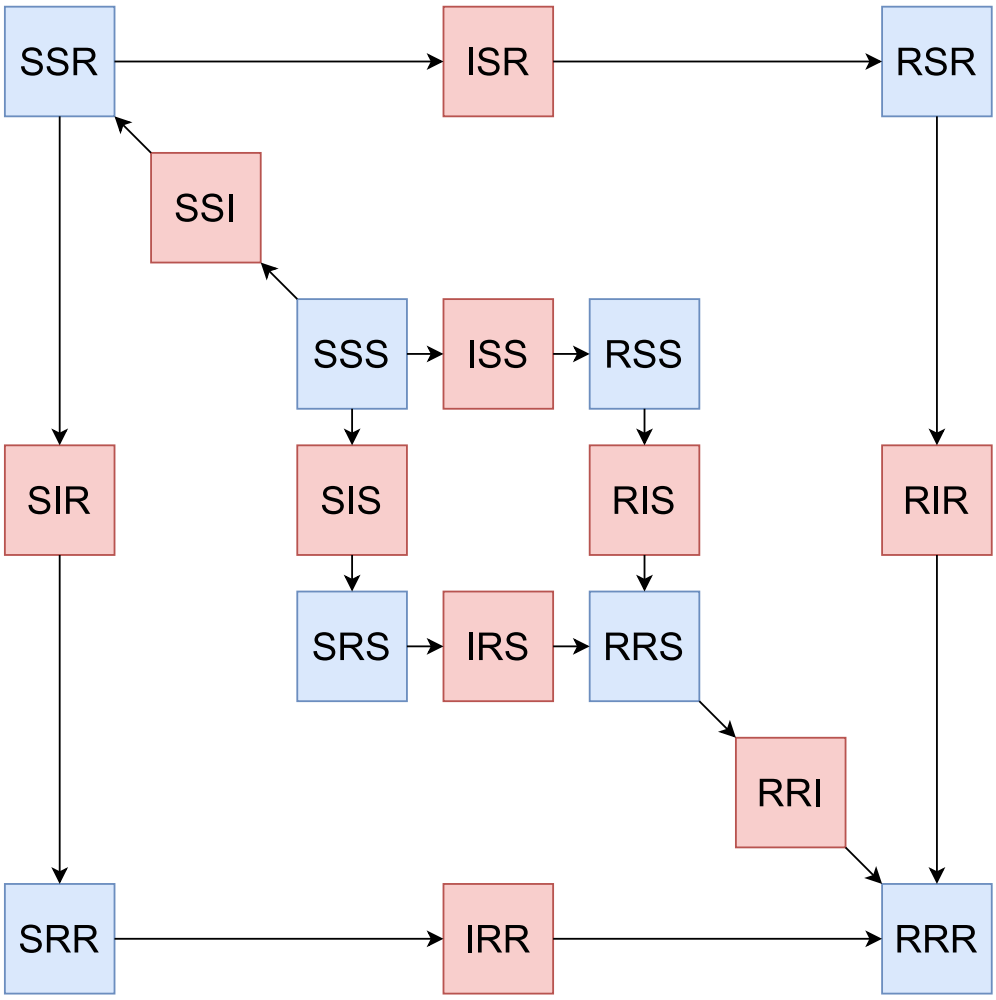

Fig. 12 A model corresponding to the product $(M_1 \boxplus M_2) \boxplus M_3$. The "SRS" and "RSS" compartments are not sinks or sources in $M_1 \boxplus M_2$; hence they have no paths to "SRR" and "RSR" respectively.

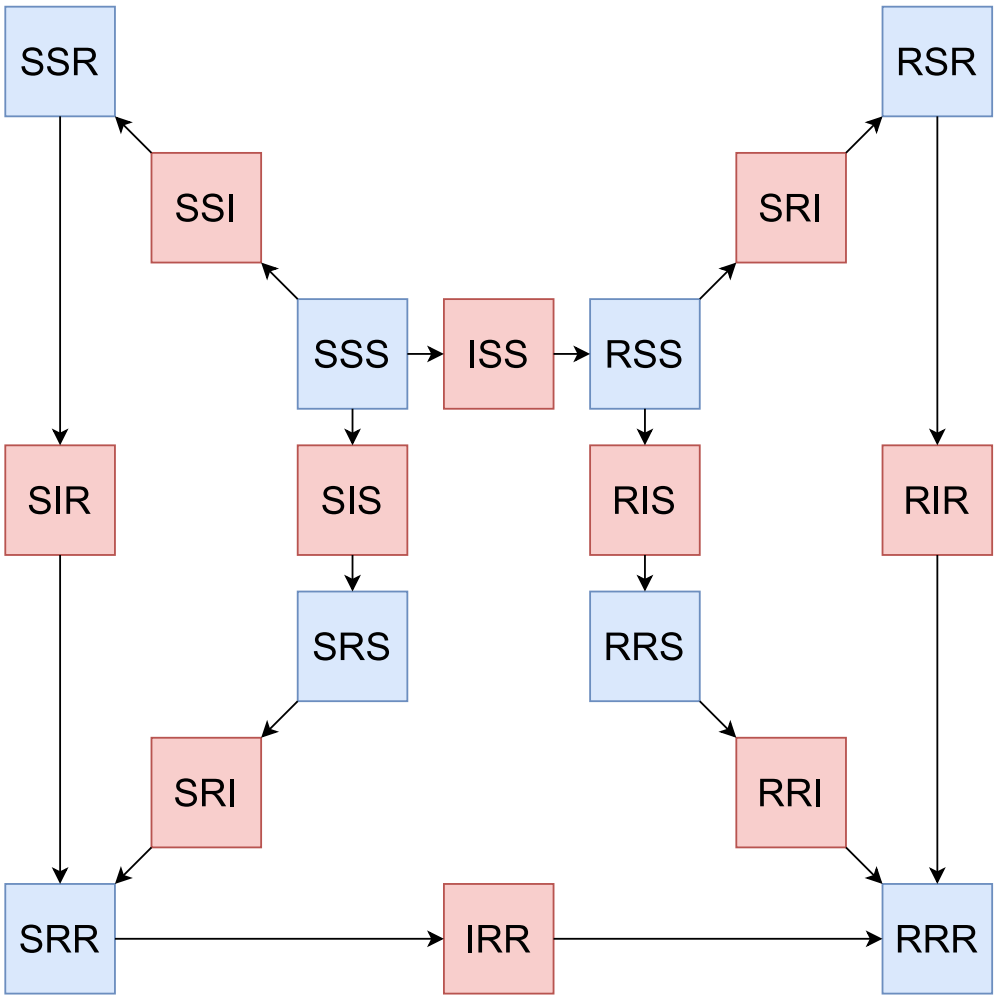

Fig. 13 A model corresponding to the product $M_1 \boxplus (M_2 \boxplus M_3)$. The "SRS" and "SSR" compartments are not sources or sinks in $(M_2 \boxplus M_3)$; hence they have no paths to "RRS" and "RSR" respectively.

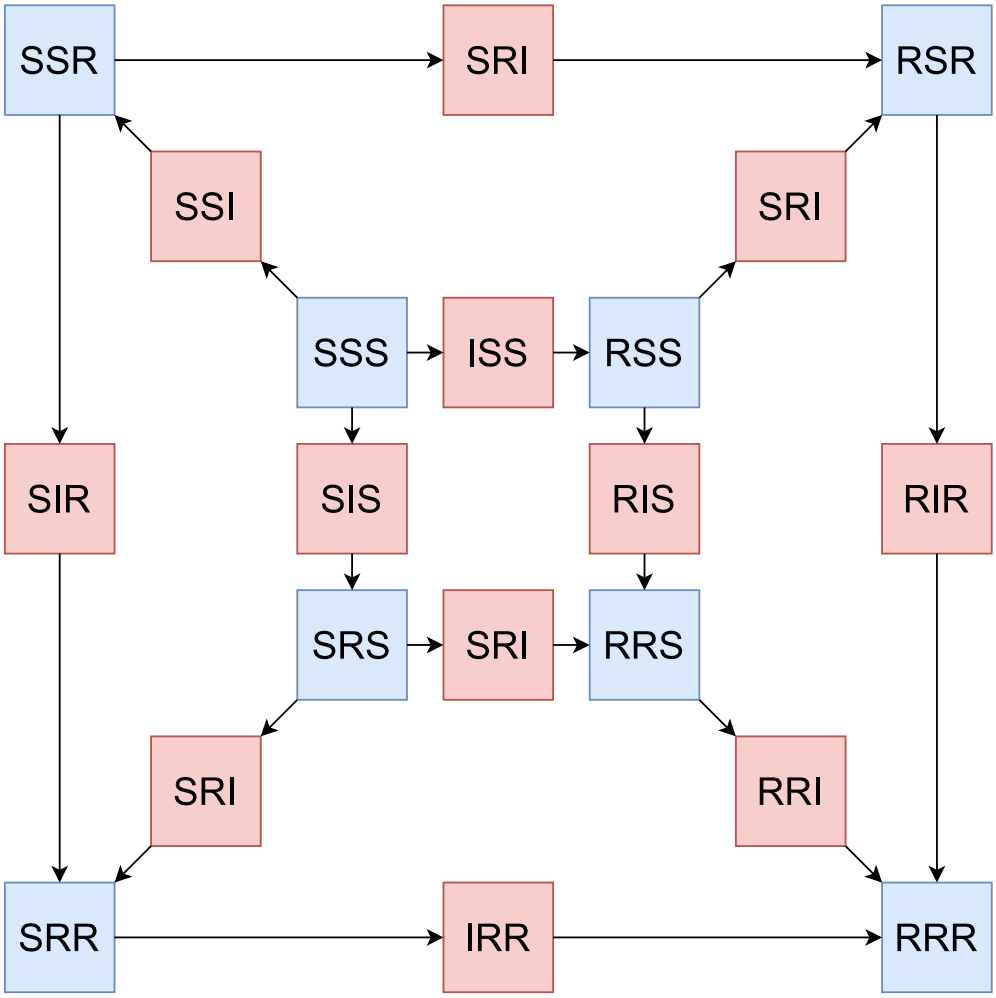

Fig. 14 A three-strain SIR model admitting no superinfection. This model cannot be constructed using only the products defined in this article.

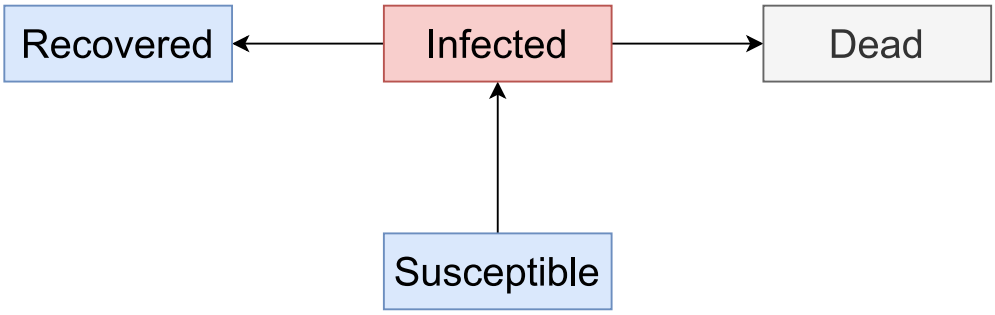

Fig. 15 A single strain model with two sinks and one source.

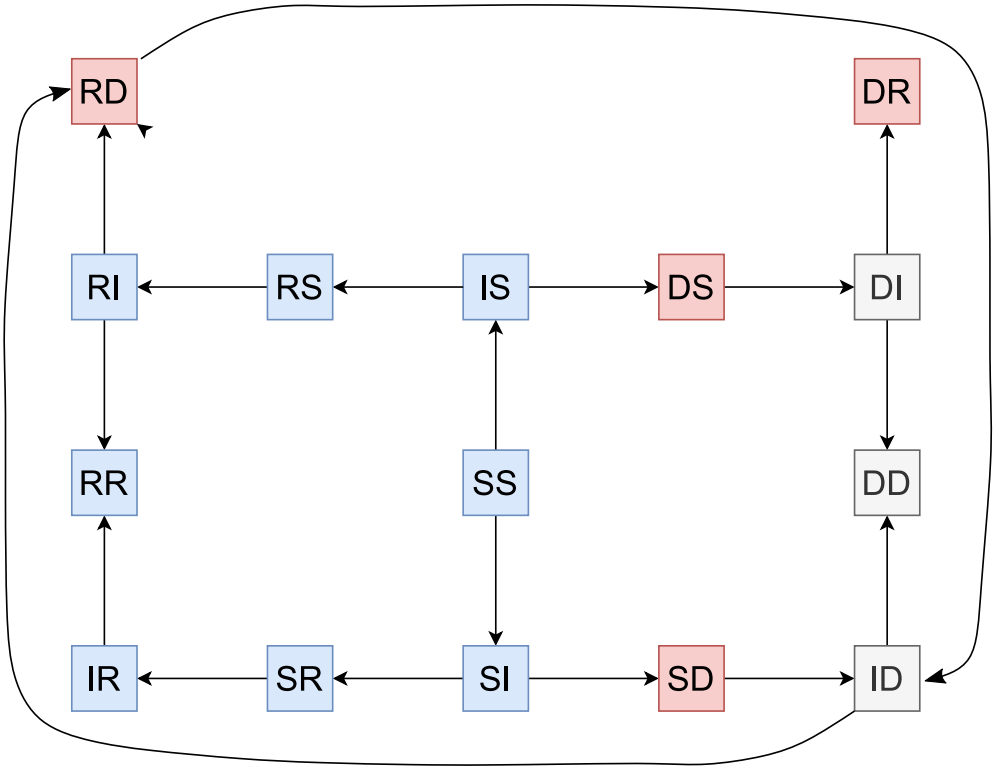

Fig. 16 The weak product of two of the single strain models depicted in Figure 15. The grey compartments are superfluous as they correspond to changes in infection status occurring after death.

5 Conclusion

Adding new strata to simple epidemiological models is closely related to taking the Cartesian product of digraphs. Modellers who want to combine sets of simple models into a single large stratified model would benefit from a toolkit based on well-defined mathematical operations. This toolkit must contain a variety of operations representing a useful subset of the numerous ways that separate strata in a model can interact. Ideally, it should also be specified in such a way that it can act *automatically*, once users have clearly specified which set of rules should be used for constructing the model product. We have developed a mathematical formalism for defining such operations and used it to restate two previously proposed model operations, the naive and modified products, which represent extremes of a spectrum of interactions between strata. The naive product corresponds to the case where different strata never interact, while the modified product corresponds to scenarios where any stratum can interact with any other stratum. We generalize these previously proposed operations to a third operation that allows any level of interaction between model strata, for example to construct geographically stratified models where interactions can occur within a single location and its neighbours but not more distantly.

BB: this needs to be restated in a more positive and useful way! Ultimately, all we have really done here is discuss ways to reduce the size and complexity of compartmental model products by eliminating unnecessary flows and compartments. If time and resources were unlimited we could treat every compartmental model as though it was a complete graph with every compartment connected to every other compartment. Doing this would certainly simplify the process of finding the product of two models, but simulating the result would involve spending a lot of time computing things like the probability of someone spontaneously aging thirty years and developing immunity to a pathogen they never actually encountered. Our goal however is not only to produce mathematically accurate models, but also to use those models to provide policy

makers with useful and timely advice. This means that finding ways to reduce the time needed create relevant models and generate their output is a highly beneficial exercise.

Several challenges remain for anyone wishing to further develop a model construction toolkit. (1) While we have considered the structure of model products in detail, we have been less thorough in building a parallel scheme for the rules used to generalize from factor model parameters to the parameters of the corresponding product model; a set of product types similar to the naive, modified, weak, etc. products described here would be useful. Many models (e.g., models with infection status testing) also have asymmetries in their structure that cannot be reproduced with Cartesian-like products. This suggests the need for additive operations to supplement the multiplicative operations presented here. In fact, such operations already exist in the category-theoretic approach to model operations, one reason why it could be a worthwhile project to unite the category theory and graph theory approaches. This would involve finding "type-graphs" that cause the category theoretic operations known as "pull-backs" and "push-outs" to reproduce the results of graph theory operations (Fong and Spivak, 2018; Libkind et al, 2022, 2021; Baez et al, 2023; Baez and Pollard, 2017).

We are heavily motivated by the desire to develop software to facilitate model construction. One insight of our investigations is the utility of a system of so-called "labeled partitions", which divide the compartments of a model into mutually exclusive groups. Each group in such a division will contain all compartments that are in the same level of some dimension of stratification and the groups can be labeled accordingly. By applying several such divisions to a model, one for each dimension of stratification, it becomes possible to specify important subsets of the model compartments. Using this system of labels and partitions provides an easy way to address issues like the non-commutativity of the weak product and the presence of "zombie compartments" discussed in Section 4.3.

Although theoretical and practical challenges with the application of binary operations on model space remain, our approach forms the basis of a powerful toolkit for the construction of complex, stratified, compartmental models.

BB: beyond further streamlining and clarification, the two things I might like to see in this paper (if they fit in) are (1) something about the decomposition of transmission matrices into (susceptibility $\times$ contact $\times$ infectivity, as in the state-dependent rates vignette ... and (2) more on labeled partitions, which we introduce at the beginning of the paper and only get back to at the very end, a little bit. Can we add stuff on labeled partitions without driving ourselves crazy or making this take much longer ...?

6 Acknowledgements

This project was supported by the Canadian Network for Modelling Infectious Diseases (CANMOD), which is funded through the Emerging Infectious Disease Modelling programme of the Natural Sciences and Engineering Research Council of Canada (NSERC).

7 Declarations

The authors declare that they have no competing interests.

References

Abou-Ismail A (2020) Compartmental models of the COVID-19 pandemic for physicians and physician-scientists. SN Comprehensive Clinical Medicine 2(7):852–858

Abu-Raddad LJ, Chemaitelly H, Butt AA (2021) Effectiveness of the BNT162b2 COVID-19 vaccine against the B.1.1.7 and B.1.351 variants. New England Journal of Medicine 385(2):187–189

Adam D (2020) Special report: The simulations driving the world's response to COVID-19. Nature 580(7802):316–319

Anderson RM, May RM (1985) Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. Journal of Hygiene (Cambridge) 94:365–436

Anderson RM, May RM (1992) Infectious Diseases of Humans: Dynamics and Control. Oxford University Press

Baez JC, Pollard BS (2017) A compositional framework for reaction networks. Reviews in Mathematical Physics 29(09):1750028

612 Baez JC, Li X, Libkind S, et al (2023) A Categorical Framework for Modeling
613 with Stock and Flow Diagrams. In: David J, Wu J (eds) Mathematics of Public
614 Health: Mathematical Modelling from the Next Generation. Springer International
615 Publishing, Cham, p 175–207, https://doi.org/10.1007/978-3-031-40805-2_8, URL
616 https://doi.org/10.1007/978-3-031-40805-2_8

617 Balabdaoui F, Mohr D (2020) Age-stratified discrete compartment model of the
618 COVID-19 epidemic with application to Switzerland. Scientific reports 10(1):1–12

619 Chang YC, Liu CT (2022) A stochastic multi-strain SIR model with two-dose
620 vaccination rate. Mathematics 10(11):1804

621 Currie CS, Fowler JW, Kotiadis K, et al (2020) How simulation modelling can help
622 reduce the impact of COVID-19. Journal of Simulation 14(2):83–97

623 Dietz K, Sattenspiel L (1995) A structured epidemic model incorporating geographic
624 mobility among regions. Mathematical Biosciences 128(1-2):71–91. https://doi.org/
625 10.1016/0025-5564(94)00068-B

626 Earn DJD (2008) A Light Introduction to Modelling Recurrent Epidemics. In: Brauer
627 F, van den Driessche P, Wu J (eds) Mathematical Epidemiology. Springer, Berlin,
628 Heidelberg, p 3–17, https://doi.org/10.1007/978-3-540-78911-6_1, URL https://doi.
629 org/10.1007/978-3-540-78911-6_1

630 Enserink M, Kupferschmidt K (2020) With COVID-19, modeling takes on life and
631 death importance

632 Fields R, Humphrey L, Flynn-Primrose D, et al (2021) Age-stratified transmission
633 model of COVID-19 in Ontario with human mobility during pandemic's first wave.
634 Heliyon 7(9):e07905

635 Fong B, Spivak DI (2018) Seven sketches in compositionality: An invitation to applied
636 category theory. arXiv preprint arXiv:180305316

637 Friston KJ, Parr T, Zeidman P, et al (2020) Dynamic causal modelling of COVID-19.
638 Wellcome open research 5

639 Gharouni A, Abdelmalek FM, Earn DJD, et al (2022) Testing and Isolation Effi-
640 cacy: Insights from a Simple Epidemic Model. Bulletin of Mathematical Biology
641 84(6):66. https://doi.org/10.1007/s11538-022-01018-2, URL https://link.springer.
642 com/10.1007/s11538-022-01018-2

643 Gog JR, Grenfell BT (2002) Dynamics and selection of many-strain pathogens.
644 Proceedings of the National Academy of Sciences 99(26):17209–17214

645 Grenfell BT, Anderson RM (1985) The estimation of age-related rates of infection from
646 case notifications and serological data. Journal of Hygiene (Cambridge) 95:419–436

647 Halter M, Baas A, Patterson E, et al (2022) AlgebraicJulia/AlgebraicPetri.jl:
648 v0.7.2. https://doi.org/10.5281/zenodo.6928476, URL https://doi.org/10.5281/
649 zenodo.6928476

650 Koyama T, Weeraratne D, Snowdon JL, et al (2020) Emergence of drift variants
651 that may affect COVID-19 vaccine development and antibody treatment. Pathogens
652 9(5):324

653 Kryazhimskiy S, Dieckmann U, Levin SA, et al (2007) On state-space reduction in
654 multi-strain pathogen models, with an application to antigenic drift in influenza A.
655 PLoS Computational Biology 3(8):e159

656 Lavielle M, Faron M, Zeitoun JD, et al (2020) Extension of a SIR model for modelling
657 the propagation of COVID-19 in several countries. medRxiv

Leontitsis A, Senok A, Alsheikh-Ali A, et al (2021) SEAHIR: A specialized compartmental model for COVID-19. International journal of environmental research and public health 18(5):2667

Levine Z, Earn DJD (2022) Face masking and COVID-19: Potential effects of variolation on transmission dynamics. Journal of the Royal Society of London, Interface 19:20210781. https://doi.org/10.1098/rsif.2021.0781, URL https://dx.doi.org/10.1098/rsif.2021.0781

Libkind S, Baas A, Patterson E, et al (2021) Operadic modeling of dynamical systems: mathematics and computation. arXiv preprint arXiv:210512282

Libkind S, Baas A, Halter M, et al (2022) An algebraic framework for structured epidemic modeling. arXiv preprint arXiv:220316345 https://doi.org/10.48550/ARXIV.2203.16345, URL https://arxiv.org/abs/2203.16345

Lofgren ET, Halloran ME, Rivers CM, et al (2014) Mathematical models: A key tool for outbreak response. Proceedings of the National Academy of Sciences 111(51):18095–18096

Massonis G, Banga JR, Villaverde AF (2021) Structural identifiability and observability of compartmental models of the COVID-19 pandemic. Annual reviews in control 51:441–459

McBryde ES, Meehan MT, Adegboye OA, et al (2020) Role of modelling in COVID-19 policy development. Paediatric respiratory reviews 35:57–60

Mohammadi Z, Cojocaru M, Arino J, et al (2023) Importation models for travel-related SARS-CoV-2 cases reported in Newfoundland and Labrador during the COVID-19 pandemic. https://doi.org/10.1101/2023.06.08.23291136, URL https://www.medrxiv.org/content/10.1101/2023.06.08.23291136v1, pages:

2023.06.08.23291136

Savageau MA (1988) Introduction to S-systems and the underlying power-law formalism. Mathematical and Computer Modelling 11:546–551

Vasireddy D, Vanaparthy R, Mohan G, et al (2021) Review of COVID-19 variants and COVID-19 vaccine efficacy: what the clinician should know? Journal of Clinical Medicine Research 13(6):317

Voit E (1990) S-system modelling of endemic infections. Computers & Mathematics with Applications 20(4-6):161–173

Voit EO (1988) Recasting nonlinear models as S-systems. Mathematical and Computer Modelling 11:140–145

Williams BJ, St-Onge G, Hébert-Dufresne L (2021) Localization, epidemic transitions, and unpredictability of multistrain epidemics with an underlying genotype network. PLoS Computational Biology 17(2):e1008606

Worden L, Porco TC (2017) Products of compartmental models in epidemiology. Computational and Mathematical Methods in Medicine 2017:e8613878. https://doi.org/10.1155/2017/8613878

Xia Y, Bjørnstad ON, Grenfell BT (2004) Measles Metapopulation Dynamics: A Gravity Model for Epidemiological Coupling and Dynamics. American Naturalist 164(2):267–281. https://doi.org/10.1086/422341