Modelling strong control measures for epidemic propagation with networks — A COVID-19 case study

Michael Small[1, 2, 3, a)] and David Cavanagh[3, b)]
[1)]*Complex Systems Group, Department of Mathematics and Statistics, University of Western Australia, Crawley, Perth, Western Australia*
[2)]*Mineral Resources, Commonwealth Scientific and Industrial Research Organisation, Kensington, Perth, Western Australia*
[3)]*Integrated Energy Pty Ltd, Como, Perth, Western Australia*

(Dated: 24 April 2020)

We show that precise knowledge of epidemic transmission parameters is not required to build an informative model of the spread of disease. We propose a detailed model of the topology of the contact network under various control regimes and demonstrate that this is sufficient to capture the salient dynamical characteristics and to inform decisions. Our model of city-level transmission of an infectious agent (SEIR model) characterises spread via a (a) scale-free contact network (no control); (b) a random graph (elimination of mass gatherings; and (c) small world lattice (partial to full lockdown — "social" distancing). This model exhibits good agreement between simulation and data from the 2020 pandemic spread of COVID-19. Also, estimates of the relevant rate parameters of the SEIR model are obtained and we demonstrate the robustness of our model predictions under uncertainty of those estimates. Finally, the social context and utility of this work is identified.

Modelling of disease transmission via compartmental models is well established and generally very effective. However, the differential equations depend on good estimates of underlying rate parameters and provide a continuous solution under the assumption that the population is well-mixed and homogeneous (i.e. all individuals have equal contact with all others). Under these assumptions disease propagation is driven by the infamous parameter R_0 — the ratio of the rate of new infections to the rate of removal of infectious individuals from the transmission pool. Typically, and particularly for contemporary and evolving transmission, these parameters can be somewhat difficult to estimate.

We propose an alternative approach to modelling the dynamic transmission of diseases. The physics literature is rife with work on models of propagation dynamics on networks. We observe that different societal control measures manifest as distinct topological structures and model city-level transmission of an infectious agent. Our approach models changing control strategies by changing the features of the underlying contact network with time. This approach allows us to model the likely time course of a disease and, perhaps surprisingly, we find that these approaches are both quantifiable and robust to uncertainty of the underlying rate parameters.

This report is intended as a guide to computational modelling of reported epidemic infection rates when good estimates of underlying epidemiological rate parameters are not available. The model provides a useful prediction of current control strategies. Nonetheless, we emphasise that the methodology and techniques are not (of themselves) novel, they have been discussed extensively in the references cited herein. What is new is the interpretation of complex network topologies as the principal relevant parameter to characterise control, and it's live application in pandemic response and recovery.

[a)]Electronic mail: michael.small@uwa.edu.au
[b)]Electronic mail: david.cavanagh@integratedenergy.com.au

I. BACKGROUND

It is no exaggeration to say that pandemic spread of infectious agents has very recently attracted wide interest. Mathematical epidemiology is a venerable and well respected field[1]. Propagation of disease in a community is modelled, under the assumption of a well-mixed and homogeneous population via differential equations characterising movement of individuals between disease classes: susceptible (S), exposed (E), infectious (I) or removed (R).

The standard compartmental (i.e. SIR) model dates back to the mathematical *tour de force* of Kermack and McKendrick[2]. The model assumes individuals can be categorised into one of several *compartments*: S or I; S, I or R; or, S, E, I, or R being the most common. Transition between the various compartments is governed by rate parameters and it is the job of the mathematical epidemiologist to estimate those rates — and hence, when $\frac{dI}{dt} < 0$ and the transmission is under control. In the standard SIR formulation the condition $\frac{dI}{dt} < 0$ can be expressed as $\frac{aI(t)}{r} \equiv R_0 < 1$. Efforts to estimate the relevant parameters for the coronavirus pandemic are currently underway and are best summarised (from our local perspective) by the technical reports of Shearer[3], Moss[4] and co-workers.

Conversely, since the renaissance of interest in mathematical graphs under the guise of complex networks[5,6], there has been considerable interest in propagation of infectious agent-like dynamics on such structures. Commonly, the agent is either modelling the spread of information or infection. When one is restricted to the spread of infectious agents on a network (in the context of epidemiology, a contact graph) several interesting features arise. In particular, if that contact graph

is a scale free network (i.e. it has a power-law degree distribution), then the criterion on the key epidemic threshold to ensure control of the outbreak (for the SIS model) becomes $R_0 = 0$.[7] In effect, what happens is that the power-law distribution of the scale-free network ensures that there is finite probability of the epidemic reaching an individual with an arbitrary large number of contacts. The number of secondary infections arising from that individual will be unbounded and transmission is guaranteed to persist. Of course, in the real-world nothing is unbounded and Fu and co-workers[8] showed that a piece-wise linear/constant infectivity was enough to ensure a positive epidemic threshold.

Surprisingly, however, little of the work in the physics literature on epidemic transmission has examined transmission on real-world networks. The first evidence (to the best of our knowledge) that epidemic transmission did really occur on a scale-free contact graph was provided by Small and others[9] for the transmission of avian influenza in migratory bird populations. Curiously, though, the data presented there gave an exponent for the scale-free distribution of approximately 1.2, significantly lower than the "usual" range of $(2,3)$ — that is, the distribution not only had divergent variance, but also divergent mean.

The emergence of an earlier coronavirus, associated with the Severe Acute Respiratory Syndrome (SARS), gave us an opportunity to apply the structures and concepts of complex systems to the modelling of infectious diseases. Small and Tse[10] introduced a complex network based model of propagation and showed good agreement between simulations of that model and available case data. We found that epidemic parameters widely quoted in the literature were only consistent with observed case data when including significant nosocomial transmission.[11] Finally, and most importantly for the current discussion, the scale-free topology of the model we proposed explained super-spreader events through contact rather than requiring pathologically highly infectious individuals.[12]

Unlike what we will propose in this current communication, this model was topologically stationary.[10–12] The model assumed a lattice with long-range (i.e. small-world) connections following a power-law degree distribution. In those papers[10–12] time varying control strategy was reflected only in changes of the rate parameters. The current coronovirus outbreak (that is, COVID-19) poses a different and unique challenge. Since February 2020 (and up to the time of writing) global transport networks and daily movement of individuals have been disrupted on a global scale. Entire cities and countries have engaged in various levels of "lockdown". We argue that it is not appropriate to model this simply by modifying the rate of transmission or rate of removal. In the current work we propose a network switching model through which the topology of the network changes to reflect varying control strategy.

In Sec. II we introduce our model structure. Section III presents our results for the case study of most interest to us, and Sec. IV describes a process for optimising model parameters based on observed caseload data.

II. THE MODEL

We assume nodes in our network can be in one of four states, corresponding to the four states of the standard SEIR model: susceptible S, exposed E, infected I, and removed R. Rates govern the probability of transition between these states, with transition from S to E occurring only through neighbour-to-neighbour contact (on the graph with a node in state I. For comparison, the standard SEIR compartmental differential equation based formulation is given by

$$\begin{pmatrix} \frac{dS}{dt} \\ \frac{dE}{dt} \\ \frac{dI}{dt} \\ \frac{dR}{dt} \end{pmatrix} \equiv \begin{pmatrix} S' \\ E' \\ I' \\ R' \end{pmatrix} = \begin{pmatrix} -pSI \\ +pSI - qE \\ +qE - rI \\ rI \end{pmatrix}, \tag{1}$$

where $p,q,r \in [0,1]$ determine the rate of infection, latency and removal respectively. Clearly, if we desire $\frac{dE}{dt} + \frac{dI}{dt} < 0$ we need $\frac{pS(t)}{r} < 1$. The parameter q determines the average latency period, and hence the ratio of p and r determines the rate of spread.

The assumptions underpinning models of the form (1) are that the population is fully mixed — that is, contacts exist between all members of the population, or, rather, every individual is indistinguishable and contacts occur at a constant rate between individuals. One way to extend this model is to introduce stratified (perhaps by age, vulnerability, comorbidity, or location) transmission models. Doing so for COVID-19 is reasonable and has been extensively covered elsewhere: for the Australian perspective see[13,14] However, this would require estimating distinct values of p, q and r for each strata. We choose an approach which we feel has fewer free parameters.[15]

Let A be an N-by-N binary symmetric matrix, $a_{ij} = 1$ iff individual i and j are in contact. We suppose that all individuals, excluding a small number whom are exposed (E), are initially susceptible (S). Then, at each time step:

$S \to E$ a susceptible node i becomes exposed if there exists a node j that is infectious (I) and $a_{ij} = 1$ with probability p;

$E \to I$ an exposed node becomes infectious with probability q; and,

$I \to R$ an infectious node becomes removed (R) with probability r.

Structural patterns of contact within the community are then modelled by varying the structure of the network A. We consider transition amongst four distinct network structures, described in the following subsections.

A. Scale-free network B

Let B denote an N-by-N unweighted and undirected scale-free network. For simplicity (and rapidity of calculation)

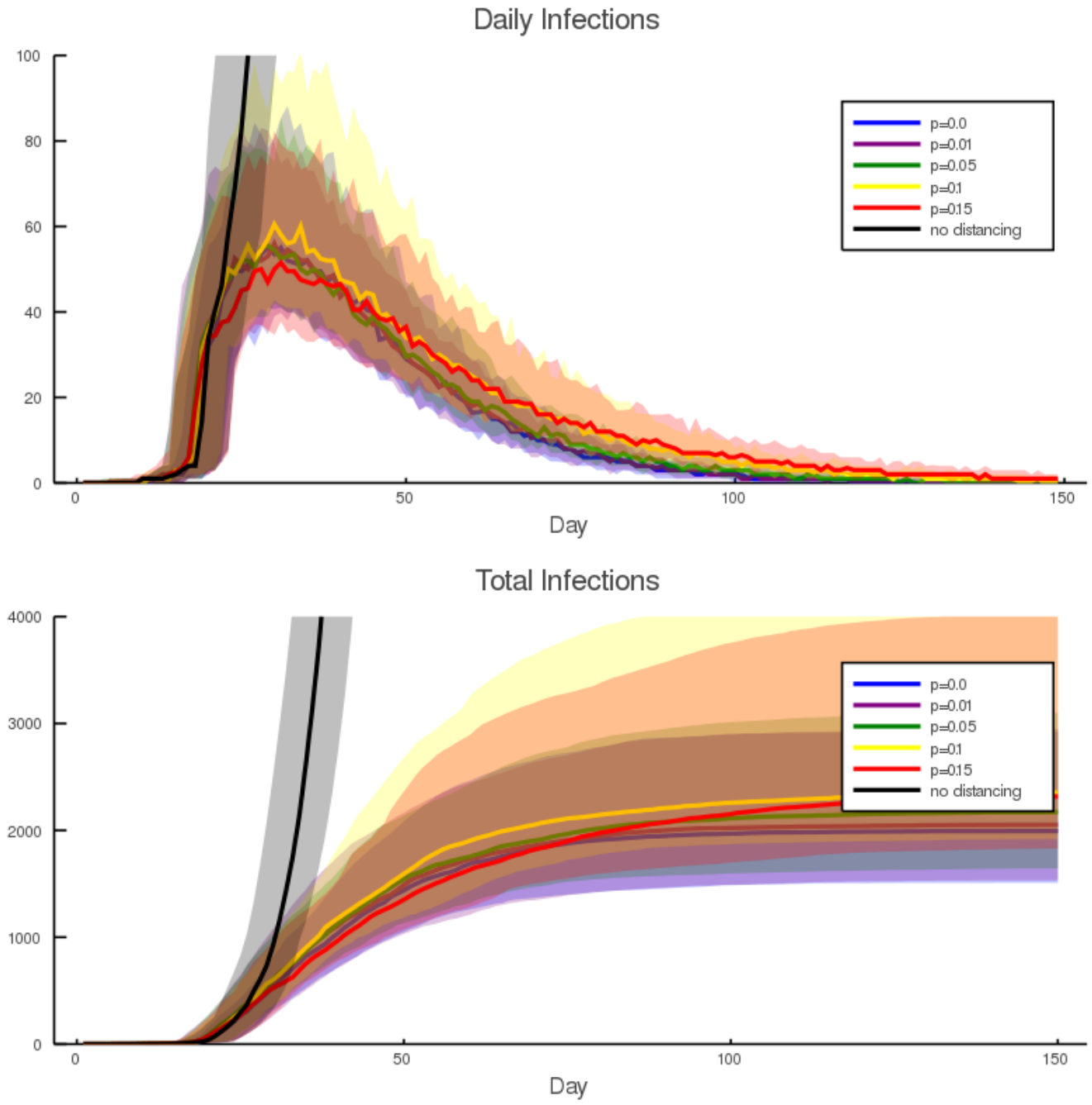

FIG. 1. Predicted epidemic time series. The upper panel is daily new infected individuals (i.e. $-\Delta(S(t)+E(t))$), lower panel is total number infected $(S(0)-S(t))$. For each network configuration, results show mean and distribution of 100 simulations over 240 days. In black $A = B(4)$ for all time. In other simulations $A = B(k)$ for all t until $I(t) > 150$, otherwise $A = L(s)$ with values of s from 0 to 1. The shaded envelopes are 90% confidence intervals. To compute the fraction of population compliant with social isolation measures d we compute $P(\text{no rewired links}) = (1-s)^k$ (here, the number of neighbours $k = 4$).

we generate this network via the preferential attachment algorithm of Barabasi and Albert[5] — there are good reasons for not doing this (notably the rich club will be highly connected[16–18]). Nonetheless, simulations presented here did not depend on the choice of the Barabasi-Albert model over alternatives including the configuration model or likelihood approaches[18]. The network B is parameterised by the number of new edges associated with each new node k and so we represent it as $B(k)$. Here, to ensure comparable number of edges, we choose $k = 2$.

The network $B(k)$ provides a model of random contacts in a community. There is ample evidence that individual con-

tact patterns follow an approximately scale-free distribution. Specifically, in the context of the current pandemic, there is clear evidence in large scale community spread of COVID-19 at sporting events and other mass gatherings which are well modelled via the tail of a scale-free distribution. Due to the random wiring of connections between nodes we expect contact network B to yield an exponential growth of infection.

B. Regular lattice $L(0)$

Let L denote a regular two dimensional lattice with periodic boundary conditions. Each node has four adjacent neighbours. For consistency with what follows we denote this as $L(0)$. Growth of infection on a lattice will be equivalent to diffusion in two dimensions and hence the infected population will grow geometrically – in the case of the configuration discussed here growth is sub-linear.

Lattice configuration is used here as an approximation to *hard isolation*: individuals do not move in geographical space and are therefore only connected to neighbours. Intuitively, one might expect a hard isolation model to consist of small isolated clusters corresponding to individual family units. In addition to being uninteresting – for the very obvious reason that transmission would cease – such a model is overly optimistic. Transmission would still be expected to occur between neighbours (in the ordinary sense of the word). The regular lattice configuration model is able to model such infection between family units, and adjacent dwellings.

C. Random Graph $L(1)$

Let $L(1)$ denote an random graph (ala Erdös-Renyi[6]) with degree of all nodes equal to four. Connections between nodes are chosen uniformly at random and constrained to avoid multiple edges or self-loops. At the opposite extreme to $L(0)$ we denote by $L(1)$ the lattice graph with no lattice structure — here all connections have been rewired and hence correspond to complete random wiring. Unlike B the degree distribution of $L(1)$ is binomial (a link exists between any two random nodes with a fixed probability, independent of all other structure). Hence, while B will be characterised by super-spreader events (spiky outliers in the daily infection count), $L(1)$ is exponential but devoid of extreme events.

The random graph model represents a mixing populace with limitations placed on mass gatherings.

D. Small-World lattice $L(s)$

Finally, let $L(s)$ denote a Watts-Strogatz[19] *two*-dimensional lattice with random rewiring with probability s. That is, the network $L(s)$ is constructed as a regular lattice $L(0)$ each edge emanating from node-i has a probability s of being disconnected from the neighbour node-j and then rewired between node i and random node-k (in doing so, one node will decrease in degree by one, and one will increase by one).

For $s > 0$, the graph $L(s)$ is an imperfect approximation to $L(0)$. That is, individuals are bound in a lattice configuration due to being geographically constrained. However, a fraction of individuals still exhibit long range connections. Effectively, the model $L(s)$ assumes that the populace is practising what is referred to in the popular press as "social distancing" (everyone is fixed at a home location and connected only to others in the same vicinity). However, there is some finite limit to compliance with the enforced isolation. A probability s of a given link switching and therefore connecting random nodes corresponds to a fraction $c = (1-s)^k$ of nodes compliant with these distancing measures since all there k edges are not switched.

In opposition to the standard and rather flawed nomenclature, we will refer to this control strategy as *physical distancing*.

E. Growth rates

We now provide estimates of the characteristic growth rates for propagation on the network structures described above. The standard approach[20] is to replace the compartmental equations Eqn. (1) with distinct equations for nodes of each degree. Let S_k, E_k, I_k and R_k denote the number of nodes of degree k in state S, E, I or R. The system (1) then becomes

$$\begin{pmatrix} S_k' \\ E_k' \\ I_k' \\ R_k' \end{pmatrix}_k = \begin{pmatrix} -pS_k \sum_\ell kP(\ell)I_\ell \\ pS_k \sum_\ell kP(\ell)I_\ell - qE_k \\ qE_k - rI_k \\ rI_k \end{pmatrix}_k , \tag{2}$$

where $P(\ell)$ is the degree distribution of the network A. In general $P(\ell)$ is a little unsatisfactory as we should really compute the sum over $P(\ell|k)$. But even in the SIS case, doing so becomes rather unwieldy. Conversely, for SEIR-type (or (SIR) epidemics the asymptotic state is trivial: $S_k(t) \to S_k^* \in (0, S(0))$, $I_k(t), E_k(t) \to 0$ and $R_k(t) \to R_k^* \in (0, S(0))$. This provides no insight.

Nonetheless, we are interested in growth rate which is determined via decrease in the susceptible population

$$pkS_k \sum_\ell P(\ell) I_\ell .$$

In the scale free case $P(\ell) \propto \ell^{-\gamma}$ and hence growth is super-exponential: high degree nodes have a contact rate proportional to their degree and a non-zero probability of connecting to other high degree nodes.

Conversely, suppose that each node has a fixed degree

$$P(\ell) = \begin{cases} 1 & \ell = L \\ 0 & \text{otherwise} \end{cases} .$$

In our lattice model $L = 4$. The growth rate is then given by $pS_k L I_L$, system (2) immediately reduces to (1), and one is left with the usual exponential growth or decay. Hereafter, we are considering only nodes of degree $k = L$ and will drop the subscript k for convenience. However, for $s < 1$ this reasoning is flawed.

Employing (1), assumes perfect mixing and hence random distribution of infectious and susceptible nodes on the lattice.

Under diffusion the infectious nodes will spread in a single cluster: nodes in that cluster will be in class E, I or R and the remainder of the population will be susceptible. The cluster will be of size $E+I+R$ and the exposed boundary will be of size scaling with $\sqrt{E+I+R}$, nodes on that exterior will be either E or I (we assume that diffusion is fast enough that the removed nodes are interior — this is certainly only an approximation and will depend on relative values of p, q and r), but only the nodes in state I are infectious. Hence, the number of infectious nodes in contact with susceptibles will scale with a quantity bounded by $\frac{I}{E+I+R}\sqrt{E+I+R}$ and $\sqrt{E+I+R}$ — mostly likely around $\frac{I}{E+I}\sqrt{E+I+R}$.[21] On average, only half the links from an infectious node will point to a susceptible (the remainder will point to other nodes in the cluster), hence, the number of susceptible nodes connected to an infected node is approximated by $\frac{I}{2(E+I)}\sqrt{E+I+R}$ and the proportion of susceptible nodes that satisfy this condition will be

$$\frac{I\sqrt{E+I+R}}{2S(E+I)}$$

and the expected number of new infections from a lattice diffusion model is obtained from the product of the rate p and the contact between these exposed infected and susceptible individuals

$$p\times\left(\frac{I\sqrt{E+I+R}}{2(E+I)}\right)\times\left(\frac{I}{E+I}\sqrt{E+I+R}\right) = p\times\left(\frac{I^2(E+I+R)}{2(E+I)^2}\right).$$

We note in passing that typically $S\gg E\propto I>R$ — certainly during initial growth, or in the case of limited penetration. Moreover, the arguments above only hold when $S\gg E,I$.

Finally, in a small-world model there is probability s of a link pointing to a random distant location. With $S\gg E+I+R$ we assume that that link is pointing to a susceptible node and so the expected number of new infections is now

$$(1-s)p\left(\frac{I^2(E+I+R)}{2(E+I)^2}\right)+spSI.$$

Since, E and I are linearly proportionate, the first term scales (very roughly) like $(E+I+R)$ the second like SI. That is, a mixture of the sub-linear growth dictated by the lattice (with proportion $1-s$) and the classical compartmental model (1) with probability s.

III. RESULTS

In this section we first present results of the application of this model. We choose a city of population of approximately 2.1×10^6 (Perth, Western Australia) and perform a simulation with initial exposed seeds and contact network $A=B$ (for $0\leq t<t^*$. The transition time t^* is the time with $I(t)>I_{\rm Th}$ for some threshold infection load $I_{\rm Th}$ for the first time (i.e. $I(t)<I_{\rm Th}$ for all $t<t^*$ and $I(t^*)\geq I_{\rm Th}$. For $t>t^*$ we set $B=L(s)$ for various values of s.

	$t<t^*$	$t>t*$
N	1450^2	
$I_{\rm Th}$	150	
p	0.2	$\frac{1}{12}$
q	$\frac{1}{7}$	$\frac{1}{7}$
r	$\frac{1}{14}$	$\frac{1}{4}$

TABLE I. Epidemic simulation parameters. The simulation size N is chosen to be a square number to make the construction of $L(s)$ simpler. Latency period of $q=\frac{1}{7}$ is comparable to observation, the other parameters are guesses designed to ensure growth in infection for $t<t^*$ but barely endemic otherwise (for $A\neq B$). That is, these parameters are selected in an *ad hoc* manner to match the observed data for our principle region of interest. Moderate[a] change of these parameters does not change the qualitative features, only the scale of the observed simulations.

[a] "Moderate" will be better quantified in the following sub-sections.

The epidemic parameters which we have chosen for this simulation are illustrated in Table I. We do not wish to dwell on the epidemiological appropriateness of these parameters — except to say that the were chosen by hand because they were consistent with our understanding of epidemiology and also gave results that looked good and coincided with the available time series data.

Some brief notes on the effect of parameter selection are in order. First, varying $I_{\rm Th}$ will delay the transition to a "controlled" regime and produce a larger peak. The parameter q is largely determined by the epidemiology of the infection, and for coronavirus COVID-19 fairly well established. It does have an important influence on the time delay of the system, but that is not evident from Fig. 1. Second, the parameters p and r for $t<t^*$ also determine the initial rate of spread — as standard epidemiology would expect. Third, the value of these parameters for $t>t^*$ determine the length of the "tail". In all simulations these parameters are chosen so that a well mixed population would sustain endemic infection. It is the network structure, not fudging of these parameters that causes extinction of the infection — this will be further illustrated in Fig. 2.

Figure 1 depicts one ensemble of simulations. Of note from Fig. 1 is the complete infection of the population without control. Conversely, the random Erdös-Renyi graph $L(1)$ has a sufficiently narrow degree distribution that the infection does (slowly) die away. Various values of $L(s)$ with $s\in(0,1)$ have the expected effect of gradually decreasing the total extent and duration of the outbreak. However, it is important to note that the 90% confidence windows are very wide and overlap almost entirely — while, on average smaller s is better this is very often not evident from individual simulations. This is due to random variation in the initial spread for $t<t^*$: one must be both lucky and clever.

It is very clear from Fig. 1 that the variance between simulations is similar in magnitude to variation in parameters. However, parameters in Fig. 1 correspond to moderate parameters p and r and a wide variation in social isolation. In an effort to understand the parameter sensitivity of this simulation we perform repeated simulations over a wide range of

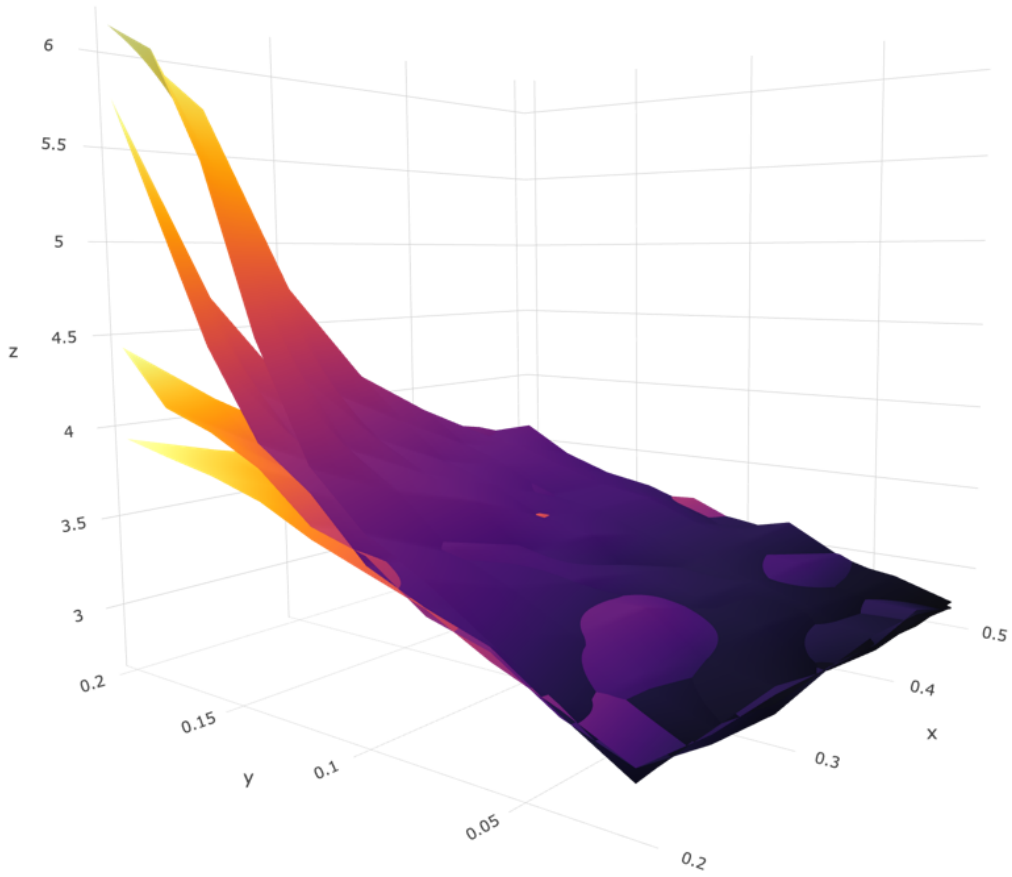

FIG. 2. Parameter sensitivity. The four panels explore the expected total number of infections (population $N = 1450^2$) for various parameter values p and r (for $t > t^*$) and different control strategies (i.e. $L(s)$ for different s). The four surfaces depicted here correspond to (a) $s = 0.0025$; (b) $s = 0.026$; (c) $s = 0.054$; (d) $s = 0.065$. The three coordinates are (x) r; (y) p; and (z) $\log(\max_t(S(0) - S(t)))$ (the logarithm[a] of the total number of infections). In each case we computed 80 simulations of 300 days. Other parameters are as reported in Table I. Surface (a) and (b) exhibit linear scaling with changing parameter values $p(t > t^*)$ and $r(t > t^*)$, while for (c) and (d) that growth is exponential. That is, when compliance with isolation measures drops below 90% there is an explosive growth in the level of infection with $p(t > t^*)$ and $r(t > t^*)$.

[a] Base 10.

$p(t > t^*)$ and $r(t > t^*)$. For all selected values we generate 20 simulations of 300 days each and compute several indicators of infection penetration

- **Mean total infection:** The total number of individuals that become exposed, infected or removed during the duration of the simulation. That is, $\max_t S(0) - S(t) = S(0) - S(300)$.
- **Mean maximum infected:** The maximum daily reported number of infections - that is, the maximum number of new infected individuals: $-\max_t(S(t) + E(t) - S(t-1) - E(t-1)$
- **Half recovered time:** The time in days required for half the simulations to entirely eliminate infection. That is, the median (over simulations) of the minimum (over time) t such that $E(t) + I(t) = 0$

Results for $I_{\mathrm{Th}} = 100$ are reported in Fig. 2, varying I_{Th} simply scales the reported numbers. Depicted in Figs. 2 and 3 are computed values of the mean total infection. The other parameters described above behave in a consistent manner.

Figure 2 and 3 starkly illustrated the importance, for the 2020 COVID-19 pandemic, of implementing and stringently enforcing isolation. Without isolation the epidemic impact is limited for very optimistic values of $p(t > t^*)$ and $r(t > t^*)$. Otherwise, the mean behaviour indicates infection growth by two orders of magnitude within 300 days - almost complete penetration. Our simulations indicate that this first becomes a risk as the containment is less than 90% effective. There is a boundary in our simulations which appears below 90% isolation and grows to include even moderate values of the other epidemic control parameters $p(t > t^*)$ and $r(t > t^*)$.

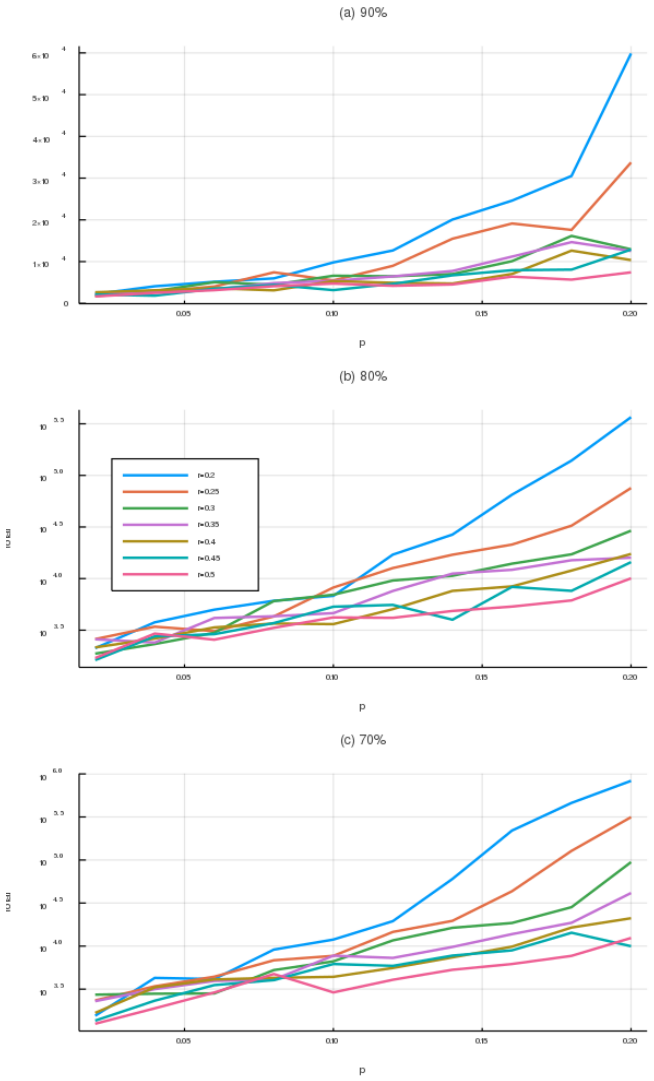

FIG. 3. Parameter sensitivity. The three panels explore the expected total number of infections (population $N = 1450^2$) for various parameter values $p(t > t^*)$ and $r(t > t^*)$ (i.e p and r for $t > t^*$) and different control strategies (i.e. $L(s)$ for different s). The four panels depicted here correspond to (a) $s = 0.013$; (b) $s = 0.026$; (c) $s = 0.054$. The panel headings report the corresponding social isolation compliance (the probability of no non-local connections). In each case we computed 80 simulations of 300 days. Other parameters are as reported in Table I). Note that panel (a) has a linear ordinate, panel (b) and (c) are depicted with a logarithmic scale. As in Fig. 2 we observe explosive growth in impact with lower levels of compliance.

IV. PARAMETER SELECTION

In part, our aim with this communication is to dissuade the application of modelling of times series to predict certain specific futures. That is, we are interested in simulation and inferring structure from the ensemble of such simulations. The random variation reported in Fig. 1 should dissuade all but the most foolhardy from prediction. Nonetheless, it is valid to ask two questions of observed time series data: (1) what parameter values are most likely given this observed trajectory, and (2) which trajectory (or set of trajectories) are most consistent with the current state. The first question we will address via a greedy optimisation procedure, to be described below. The second question is equivalent to asking for an ensemble estimate of the current state of exposed but undetected individuals within the community. A complete study of this second problem is beyond the scope of the present discussion, but some points are worth considering before we return to the issue of parameter estimation in Sec. IV B.

A. State estimation

As noted previously, there is very significant variation between trajectories for the same model parameter settings. While this means that the construction of more complex models – solely from time series data – is inadvisable, it is natural to seek to explain this variability. Simulations conducted above for an SEIR model with nontrivial latency period mean that at any instance in time there is a large number of exposed but undetected individuals within the network. The location of this exposed class within the network (their distribution relative to hubs, for example) explains the variation we observe. This has been demonstrated by simulation from repeated random distributions of exposed individuals. It is easy to estimate the expected number $E(t)$ from the time series $I(t)$ and $R(t)$, however, the distribution of these individuals on the network is not uniform. The question that must be addressed to resolve this issue is what is the expected distribution of $E(t)$ random walkers on a network A? In the interest of clarity and succinctness, we do not address this issue here.

B. Parameter estimation

A separate problem is to determine the maximum likelihood values of the parameters p, q, and r for a given population N and I_{Th} from an observed time series. This can be decomposed to several discrete steps.

1. We suppose that q is fixed and estimable by other means. For COVID-19, for example, q should yield a latency period of 7-14 days. Hence $q \in (\frac{1}{14}, \frac{1}{7})$.

2. Determine the epidemic peak from the time series — this will define the turning point and the time when growth changes from exponential for geometric. This will allow one to determine I_{Th} and the corresponding t^*. In effect we are now seeking a turning point of the total number of infections $(S(0) - S(t))$ and not just $I(t)$ as done in Fig. 1.

3. For $t < t^*$ determine $p(t < t^*)$ and $r(t < t^*)$. The ratio of these two parameters determines the epidemic growth rate via R_0

4. For $t > t^*$ it remains to determine s, $p(t > t^*)$ and $r(t > t^*)$. We note that s controls the extent to which

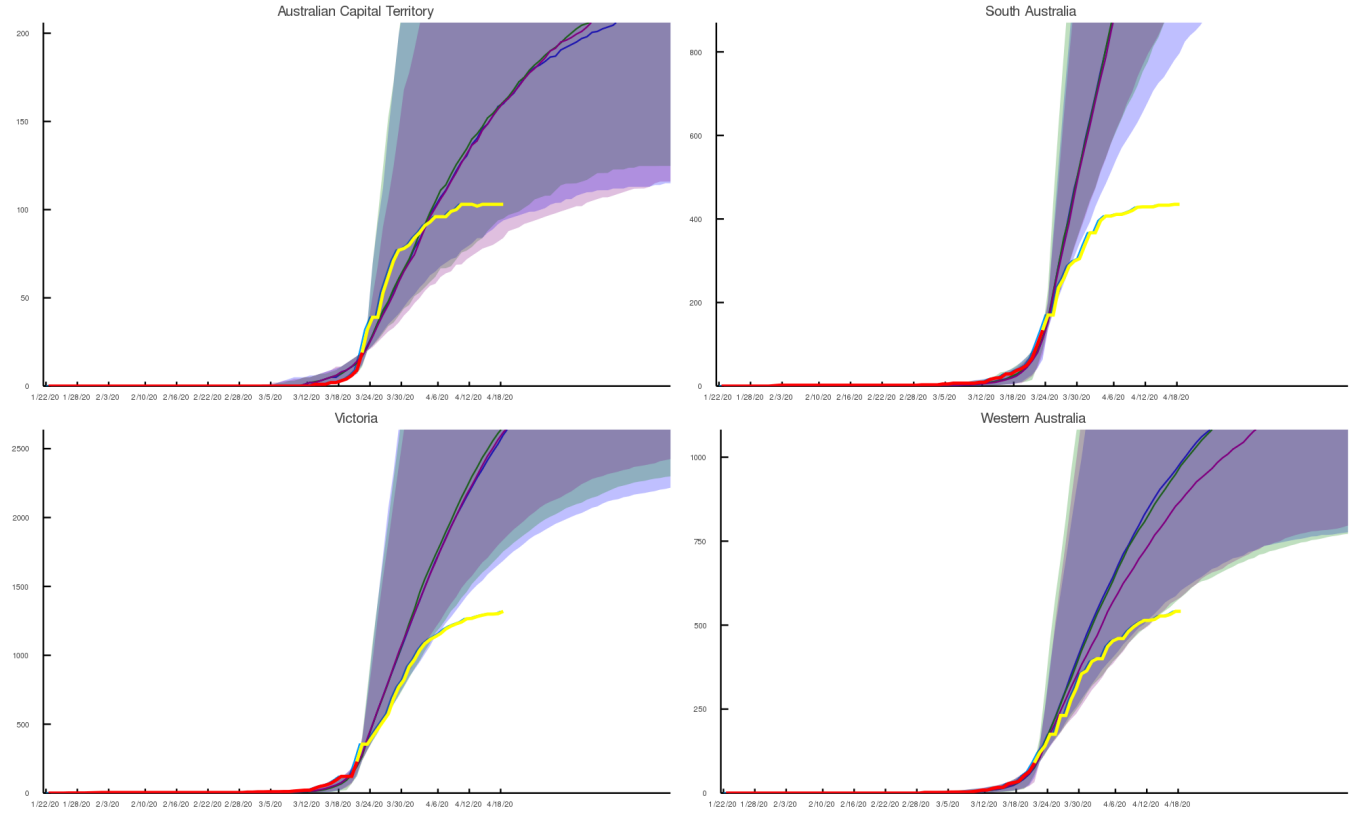

FIG. 4. Control evaluation. We depict the effectiveness of control measures for a variety of representative cities. In each case the epidemic diffusion is fitted to data up to the end of the exponential growth phase (that is, the point of inflexion on curves $S(0) - S(t)$). Simulations up to this time point t^* effectively seed the network and provide a distribution of infectious and exposed individuals within the community. Beyond this point we simulate the application of small-world control network structure $L(s)$ for various values of s. Here we illustrate $s = 0.013$, $s = 0.026$ and $s = 0.054$ corresponding to 95%, 90% and 80% control. Actual observed time series data is also shown and illustrates exception effectiveness of control measures for various Australian cities.

the system is driven by diffusion (geometric) versus exponential growth. But, for now, the best we can do is a greedy likelihood maximisation process.

Note that, in the event that the peak has not yet been reached (i.e. $t < t^*$) it is not even sensible to attempt to estimate the parameters s, $p(t > t^*)$ and $r(t > t^*)$. Nonetheless, in this situation one can estimate instantaneous (or windowed) values for R_0 and attempt to pick the end of the exponential growth phase. The latency introduced by q complicates this process and such procedures have been the subject of considerable discussion in the statistical epidemiological literature — as well as on physics blogs. Figure 4 illustrates the result of such a calculation. Finally, we note (as is indicated in the illustrated exemplars) that we assume a single policy change-point t^* — this is clearly inappropriate for more complex time dependent responses.[22]

V. SOCIAL CONTEXT AND UTILITY

This model has its origins in the severe societal challenge of COVID-19, when the population of Perth was facing the prospect of loss of 30,000 lives, and hospitals being over-run within two or three weeks if the rate of escalation continued. The model was first used in a pandemic response workshop for a city of 100,000 people, where it served to demonstrate the dramatic range of outcomes which were possible, depending on the behaviour of constituents of the city, and degree of social distancing achieved. This proved very effective in enabling appropriate action, both in the workshop and afterwards. Subsequently the results were shared on professional social media, and an online conference, influencing thousands more.

In combination with effective timely coordinated state and federal government polices, and a high level of societal compliance, a very strong result of virus suppression was achieved. The model was further developed to update progress, within two weeks, and at the time of the workshop debrief this was used to show the importance of continuing measures in suppression, the rate at which rapid outbreak could occur, even in the context of strong initial suppression. This allowed the appropriate focus to shift towards a positive recovery. again this was shared local and internationally to provide hope for others and influence behaviour.

Subsequently, the actual case data within the state was plotted against the forecast range, and this was shared with state

FIG. 5. Immersive multiperson visualisation of the model in context using virtual reality. Here illustrated in our implementation of a digital Public Health Emergency Operations Centre, where the model is integrated into wider contextual information such as national trends and geospatial information. The model is used to communicate scenarios allowing stakeholders able to draw conclusions collaboratively in context.

scientific authorities, enabling a constructive discussion about the correlation between application of selected state and national control measures and outcomes. The extension to modelling different approaches to recovery continues in a similar mode.

To gain most value from the model, its results have been interpreted in a variety of environments, including most recently in collaborative virtual reality mode, in a digital Public Health Emergency Operations Centre (PHEOC) (Fig. 5). This has the advantage of rich immersion in the data, while allowing deep multi-party interaction and dialogue to discern appropriate observations, and at the same time allow parties to engage together from anywhere in the world. At the time of writing, the number of new cases of COVID-19 has for the first time reached zero, with only seven fatalities in the State to date, remarkably low compared to world averages.

VI. CONCLUSION

The model we present here has a small – perhaps minimal – number of parameters, and describes the observed dynamics of pandemic disease transmission. When applied to data from the global outbreak of COVID-19 in 2019/2020, the model closely matches observed data across population centres. Nonetheless, identical simulations with new initial conditions yield vastly different outcomes. The variance of our model predictions is large, and in fact much larger than the variance observed between distinct epidemiological parameter values. Our results indicate that particular simulations of models that claim to have predictive power within that prediction envelope may be prone to over-interpretation. Finally, despite modelling a complex systems with complex networks we have demonstrated the sufficiency of a minimal model. Models with large numbers of parameters *which are fitted to time series data* are unnecessary and likely to be unreliable and misrepresent the underlying dynamical process. This does not preclude the construction of more complex models when sound epidemiological reasoning dictates it is necessary and when informed with direct evidence to allow for quantitative estimation of the relevant parameters.

While we are reluctant to make predictions from, or over interpret the application of, this model to the current coronovirus pandemic, our results indicate that strict physical isolation in combination with monitoring and the usual transmission mitigation strategies are required to minimise impact. Below 80% compliance with physical isolation measures risks catastrophic spread of infection (Fig. 3). This data is consistent with the evidence of explosive growth of infection experienced in some localities. Without decisive and potentially severe intervention, similar disasters are likely to occur in regions with weaker health systems.

In the simulations described above, we do not make any attempt to ensure "pseudo-continuity" between time varying manifestations of A. That is, nodes that are connected for one network are not more likely to be connected after switching the network topology. We could see no simple and generic way in which to achieve this. Moreover, we did not detect any excessive mixing that one might expect should this mismatch be an issue.

Also, in Sec. IV A we raise the issue of estimating the expected distribution of unobserved infection sites (i.e. state E) on a network. Should the model described here prove relevant, this will be an issue of immense importance to the proper

quantification of uncertain future behaviour.

Finally, the social context and utility of this modelling is demonstrated by its live use in shaping the planning and implementation of a highly effective response to COVID-19 on a city and state level. Ultimately, one must ask what is the purpose of modelling. Epidemic disease transmission is a fairly simple mathematical problem — exponential growth followed by decay. The difficulty is in reliably estimating parameters. We show that the contact structure provides a direct and effective approach to model control strategies. In addition to the information provided by our simulations, we describe in Sec. V the application of these methods to effectively inform and influence policy makers.

ACKNOWLEDGMENTS

MS is supported by Australian Research Council Discovery Grants DP180100718 and DP200102961. Source code for all calculations described in the manuscript is available on `https://github.com/m-small/epinets`. Data was obtained from `https://github.com/CSSEGISandData/COVID-19`.

[1] J. D. Murray, *Mathematical Biology*, 2nd ed., Biomathematics Texts, Vol. 19 (Springer, 1993).
[2] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," Proc Royal Soc A **115**, 700–721 (1927).
[3] F. M. Shearer, J. Walker, N. Tellioglu, J. M. McCaw, J. McVernon, A. Black, and N. Geard, "Assessing the risk of spread of COVID-19 to the Asia Pacific region," Tech. Rep. (Doherty Institute, 2020) draft, unpublished.
[4] R. Moss, J. Wood, D. B. D, F. Shearer, A. Black, A. Cheng, J. McCaw, and J. McVernon, "Modelling the impact of COVID-19 in Australia to inform transmission reducing measures and health system preparedness," Tech. Rep. (Doherty Institute, 2020) draft, unpublished.
[5] A. Barabási and R. Albert, "Emergence of scaling in random networks," Science **286**, 509–512 (1999).
[6] P. Erdös and A. Rényi, "On random graphs I," Publ. Math. Debrecen **6**, 290–297 (1959).
[7] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," Physical Review Letters **86**, 3200–3203 (2001).
[8] X. Fu, M. Small, D. M. Walker, and H. Zhang, "Epidemic dynamics on scale-free networks with piecewise linear infectivity and immunization," Physical Review E **77**, 036113 (2008).
[9] M. Small, D. M. Walker, and C. K. Tse, "Scale-free distribution of avian influenza outbreaks," Physical Review Letters **99**, 188702 (2007).
[10] M. Small and C. K. Tse, "Small world and scale free model for transmission of SARS," International Journal of Bifurcation and Chaos **15**, 1745–1755 (2005).
[11] M. Small and C. K. Tse, "Clustering model for transmission of the SARS virus: Application to epidmic control and risk assesment," Physica A **351**, 499–511 (2005).
[12] M. Small, C. K. Tse, and D. M. Walker, "Super-spreaders and the rate of transmission of the SARS virus," Physica D **215**, 146–158 (2006).
[13] G. J. Milne and S. Xie, "The effectiveness of social distancing in mitigating COVID-19 spread : a modelling analysis," (2020), medRxiv https://doi.org/10.1101/2020.03.20.20040055 draft, unpublished.
[14] S. L. Chang, N. Harding, C. Zachreson, O. M. Cliff, and M. Prokopenko, "Modelling transmission and control of the COVID-19 pandemic in Australia," (2020), arXiv:2003.10218 [q-bio.PE] draft, unpublished.
[15] Our model is a network. One could argue that a network of N individuals has $N(N-1)/2$ parameters p_{ij} governing contact between individual i and j and (worse) potentially unique r_i and q_i for each individual. We prefer the statistical physics approach of describing the key features of a network with a very small number of parameters.
[16] L. Zhang, M. Small, and K. Judd, "Exactly scale-free scale-free networks," Physica A **433**, 182–197 (2015).
[17] M. Small, Y. Li, T. Stemler, and K. Judd, "Growing optimal scale-free networks via likelihood," Physical Review E **91**, 042801 (2015).
[18] M. Small, Y. Li, K. Judd, and T. Stemler, "Growing optimal scale-free networks by likelihood principles," Physical Review E (2015), arXiv: 1305.7296v1.
[19] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," Nature **393**, 440–442 (1998).
[20] X. Fu, M. Small, and G. Chen, *Propagation dynamics on complex networks: Models, analysis and stability* (John Wiley & Son, 2014).
[21] Throughout, we've assumed a 2-D lattice. Of course, this choice is arbitrary and an N-D lattice would naturally lead to an expression involve the exponent $N-1N$ with the random model prevailing at $n \to \infty$.
[22] Or for regimes with inconsistent, indecisive or ineffectual responses.