

## Research



**Cite this article:** Harding N, Nigmatullin R, Prokopenko M. 2018 Thermodynamic efficiency of contagions: a statistical mechanical analysis of the SIS epidemic model. *Interface Focus* **8**: 20180036.

<http://dx.doi.org/10.1098/rsfs.2018.0036>

Accepted: 10 September 2018

One contribution of 10 to a theme issue 'Computation by natural systems'.

### Subject Areas:

biophysics, biomathematics

### Keywords:

SIS epidemics, thermodynamic efficiency, maximum entropy principle, Fisher information, criticality

### Author for correspondence:

Nathan Harding

e-mail: [nathan.harding@sydney.edu.au](mailto:nathan.harding@sydney.edu.au)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4238465>.

# Thermodynamic efficiency of contagions: a statistical mechanical analysis of the SIS epidemic model

Nathan Harding<sup>1</sup>, Ramil Nigmatullin<sup>1</sup> and Mikhail Prokopenko<sup>1,2</sup>

<sup>1</sup>Centre for Complex Systems, Faculty of Engineering and IT, University of Sydney, Sydney, New South Wales 2006, Australia

<sup>2</sup>Marie Bashir Institute for Infectious Diseases and Biosecurity, University of Sydney, Westmead, New South Wales 2145, Australia

NH, 0000-0003-1707-9248; RN, 0000-0003-2577-6561; MP, 0000-0002-4215-0344

We present a novel approach to the study of epidemics on networks as thermodynamic phenomena, quantifying the thermodynamic efficiency of contagions, considered as distributed computational processes. Modelling SIS dynamics on a contact network statistical-mechanically, we follow the maximum entropy (MaxEnt) principle to obtain steady-state distributions and derive, under certain assumptions, relevant thermodynamic quantities both analytically and numerically. In particular, we obtain closed-form solutions for some cases, while interpreting key epidemic variables, such as the reproductive ratio  $R_0$  of a SIS model, in a statistical mechanical setting. On the other hand, we consider configuration and free entropy, as well as the Fisher information, in the epidemiological context. This allowed us to identify criticality and distinct phases of epidemic processes. For each of the considered thermodynamic quantities, we compare the analytical solutions informed by the MaxEnt principle with the numerical estimates for SIS epidemics simulated on Watts–Strogatz random graphs.

## 1. Introduction

Various real-world crises and disruptive events, such as epidemics, cascading technological failures, ecological and economic tipping points, can be quantitatively studied as critical phenomena, so that the corresponding critical thresholds can be identified, predicted and used in planning suitable crisis interventions (e.g. vaccinations and quarantine, power-grid safety margins, climate change policies). Modelling critical dynamics typically involves analysis of sensitivities to initial conditions and the overall spatio-temporal behaviour at the system level. Within physics, such behaviour is characterized in terms of the control and order parameters, allowing the modellers to investigate phase transitions. A canonical example is a second-order phase transition in a ferromagnetic system, which separates two qualitatively different phases: a disordered paramagnetic phase characterized by the absence of net magnetization in the high-temperature regime, and the ordered ferromagnetic phase with a net magnetization in the low-temperature regime. Importantly, the change between these phases is sudden and is driven by varying the control parameter (temperature). The resulting magnetization outcome is traced by the order parameter: the net magnetization vector which quantifies the emerged preferred direction in space. Formally, a phase transition manifests itself as 'a sharp change in the properties (state) of a substance (system)' occurring when 'there is a singularity in the free energy or one of its derivatives' [1].

Like many other fields of research, epidemiology has also been drawing on the results obtained within statistical physics in terms of critical thresholds and phase transitions. For example, several studies have successfully modelled epidemic spread as a specific example of percolation in networks [2–6]. Under certain (fairly strong) assumptions, the problem of when an epidemic

takes place becomes equivalent to a standard percolation problem on a graph, whose objective is to compute the fraction of sites that must be occupied before the formation of a ‘giant component’ of connected sites. The size of such giant component scales extensively with the total number of sites [3], demonstrating scale-invariance, a well-known feature observed during critical regimes.

A critical threshold that is often studied in epidemiology is the epidemic threshold defined with respect to the pathogen’s reproductive ratio ( $R_0$ ), that is, the number of secondary infections generated on average, within a susceptible population, by an infected host. The well-known result is that  $R_0$  has to exceed one for an epidemic outbreak to occur. As pointed out in many studies, this prediction strictly holds only in deterministic models with infinite population [7]. The underlying contact network also strongly influences the epidemic threshold and its predictions [8,9].

Furthermore, in finite populations, due to finite-size estimation challenges, an accurate identification of the epidemic threshold is problematic, and instead, an epidemic (critical) interval may be considered [10]. Following [11,12], the study of Erten *et al.* [10] applied an information-theoretic model of distributed computation to a homogeneous network. It identified (i) the lower bound of the interval, on the ordered side of the transition, by the peak of the active information storage, quantifying the ‘memory’ of the computation during a contagion, and (ii) the upper bound of the interval, on the disordered side of criticality, with the maximum of the transfer entropy, quantifying the ‘communication’ aspect of the contagion [10].

Identifying and detecting critical regimes of a contagion remains a subject of a vigorous research. The study of Hartfield & Alizon [7] contrasted the critical community size, defined as the total population size needed to sustain an outbreak once it has appeared, with the outbreak threshold ( $T_0$ ), computed at the onset of an outbreak, and measuring how many infected individuals are ‘needed to ensure that an outbreak is very unlikely to go extinct by drift’.

Under some circumstances (such as pathogen mutations or changes in the host population), even a maladapted pathogen with  $R_0$  below but close to 1 still has a potential for an outbreak, when the changes cause its  $R_0$  to exceed 1. This situation has been considered for new pathogens emerging by crossing the species barrier [13]. The sudden changes in the disease spread are obviously similar to critical dynamics and have been studied in this context, in an attempt to predict, and hopefully prevent, the emergence of criticality [14]. Again, the underlying network topology and its mixing patterns can substantially affect the disease emergence [15].

In summary, some of the present challenges relate to reliably detecting critical thresholds in finite-size systems, within complex network topologies, and dealing with distributed data generated by nonlinear dynamics.

A recent thermodynamics-based framework that has successfully dealt with such challenges in several abstract settings uses Fisher information, a measure that is directly connected to the rate of change of the corresponding order parameters [16–19]. These studies accurately identified phase transitions in the global spatio-temporal behaviour via an estimation of Fisher information for a number of network topologies. Critical thresholds have been pinpointed when the observed variables were most sensitive to the control parameters, resulting in divergence of Fisher information

in infinite systems and its maximization in finite-size systems. Crucially, this method relies on estimation of underlying probability densities and, thus, is applicable even when the corresponding order parameter is unknown or cumbersome to compute.

The approach based on Fisher information has not, to date, been applied in an epidemiological setting, and promises to strengthen the prediction accuracy of epidemic thresholds in complex scenarios, involving heterogeneous network topologies, large-scale distributed data and probable emerging pathogens. It may also enable derivations of closed-form solutions in specific cases. To fully exploit its potential, however, the framework needs to be well grounded in a statistical-mechanical setting and complemented with rigorous methods for the suitable estimation of probability densities.

Typically, in statistical mechanics, the model of the systems is specified in full by providing the microscopic coupling constants between all components of the systems. Once the model is given, the challenge is to infer the emerging macroscopic properties of the ensemble using analytic and computational methods. An inverse problem, that is the determination of the microscopic coupling constants from some known macroscopic constraints is solved using the maximum entropy (MaxEnt) principle. The MaxEnt principle states that the least biased model is obtained by maximizing the entropy of the distribution while at the same time respecting the imposed constraints.

The MaxEnt principle has been applied to analyse activity in various complex networks including ecological networks, networks of neurons [20], biochemical and genetic networks and flocking birds [21]. Typically, the solution of the variational MaxEnt problem relies on detailed computer simulations. Even in these typically hard cases, reasonable approximations can simplify the problem and lead to closed-form solutions. Closed-form solutions reveal strong analogies between macroscopic systems such as flocks of birds and well understood statistically mechanical models such as the Ising model. These analogies elucidate the statistical mechanical origins of non-trivial collective phenomena, such as phase transitions, in complex systems.

Applications of the MaxEnt principle in computational epidemiology has so far been limited. The MaxEnt principle has been applied to stochastic SIS and SIR dynamic models, using real data to fit probability distributions of various epidemic characteristics such as time to infection, number of recovered individuals, number of infected individuals at a specific time interval  $[0, t]$ , etc. [22]. Going beyond this goal, we aim to apply the MaxEnt principle in construction of statistical mechanical models of epidemics enabling a statistical mechanical analysis of epidemic phase transitions.

The state space of the individual nodes on an SIS epidemic network is binary since the nodes can be either infected or susceptible. This is reminiscent of the Ising model, where the state of each node is also binary (either up or down). However, unlike the Ising model, the contact interactions in epidemic networks are directional. An infected individual has the capacity to flip the state of a susceptible neighbour while the susceptible individual has no effect on the infected neighbour. This directionality of interactions poses a novel challenge to the application of the MaxEnt method.

In this paper, we apply the MaxEnt principle to derive a statistical mechanical model of a contact network undergoing

SIS dynamics. We obtain an analytic solution for a simple case, characterizing an initial (seeding) state of an outbreak. In order to arrive at analytic models for more general cases, we propose a simplification that assumes independence between the number of infected individuals and the number of infective links. We assess the impact of this assumption by comparing the results of the MaxEnt model of SIS process on Watts–Strogatz network [23] to the results generated by computer simulations of the underlying dynamics. The derived MaxEnt models are used to evaluate quantities such as entropy, free energy and Fisher information in an epidemiological context. The statistical mechanical setting is used to provide a novel interpretation of epidemic thresholds. We specifically study the thermodynamic efficiency of contagion, considered as a distributed computational process. The thermodynamics of computation have recently been investigated in various contexts [18,24–28], but have not been applied to studies of epidemics.

The paper is structured as follows. In Background, we describe relevant epidemic and network models, while Technical preliminaries outline the MaxEnt principle and an approach to criticality analysis based on Fisher information. We then develop our framework applying the MaxEnt principle to an SIS epidemic model, including a closed-form solution derived in specific cases. This is followed by computational results demonstrating criticality in statistical mechanical terms.

## 2. Background

### 2.1. Models of epidemics

The SIS model of epidemics captures the dynamics of diseases which are transmitted by individual to individual contact. Additionally, it refers to a type of disease in which individuals can be infected multiple times throughout their lives without developing long-lasting immunity. Examples of diseases following SIS dynamics are rotaviruses, sexually transmitted infections and bacterial infections [29]. The SIS model of epidemics refers both to a differential equation model and a discrete time update process [30]. In the case of the differential equation model of the SIS dynamics, the progression of the disease within the population is described by a pair of coupled ordinary differential equations:

$$\frac{dS}{dt} = \gamma I - \beta IS \quad (2.1)$$

and

$$\frac{dI}{dt} = \beta IS - \gamma I, \quad (2.2)$$

where  $I$  is the number of infected individuals,  $S$  is the number of susceptible individuals, the parameter  $\beta$  is the transmission rate and the parameter  $\gamma$  is the recovery rate [29]. This model assumes that all individuals within the host population interact with equal probability [29] and is often referred to as the mass action model of infection.

As the population is generally considered in isolation, it is also common to consider a single differential equation, normalizing the population:

$$\frac{dI}{dt} = \beta I(1 - I) - \gamma I \quad (2.3)$$

and

$$S + I = 1. \quad (2.4)$$

A well-known result of Kermack & McKendrick [31] shows that if the initial fraction of susceptibles is less than  $\gamma/\beta$ ,  $dI/dt < 0$ , the infection dies out. This is referred to as the ‘threshold phenomenon’. This result can also be interpreted as requiring  $\beta/\gamma = R_0$ , commonly known as the basic reproductive ratio, to be large enough that the initial infected population increases with time. In a more general sense, the reproductive ratio  $R_0$  is defined as ‘A measure of the number of infections produced, on average, by an infected individual in the early stages of an epidemic when virtually all contacts are susceptible’ [32].  $R_0$  is frequently used in order to broadly quantify the transmissibility of an epidemic strain; in general, epidemics emerge when  $R_0 > 1$  [33].

### 2.2. Network models

Original approaches have assumed that interactions occur completely at random within the population [31]. It has since been argued that because the ‘structure of a contact network can have a profound effect on the dynamics of infectious disease’ [34] it is imperative to use network models as opposed to the more traditional mass action models [31]. Thus networks have become a standard model for studying the spread of disease, quantifying interactions between individuals or populations of individuals [2,4,8,30,35,36].

There are a number of networks which are commonly investigated within the epidemiological literature. The most commonly studied have been random networks [35], lattices [37], scale-free networks [8] and small-world networks [30].

Unlike the mass action mixing approach, network-based approaches define a neighbourhood for each individual in which they can infect others and be infected by others. Importantly, in some cases, this representation yields a closed-form solution; for example, the study of [4] shows that a large class of the SIR models of epidemic disease can be solved exactly on networks of various kinds using a combination of percolation models and generating function methods. Another key distinguishing result in the study of epidemics on networks shows that in some cases there is no critical threshold [8], with disease propagating regardless of the probability of infection [8]. As such, for network models there is no general result analogous to  $R_0 = 1$  from differential equations models.

Of special importance to studies of contagion processes is the class of small-world networks introduced by Watts & Strogatz [23]. The algorithm constructing a small-world network essentially interpolates between regular and random networks, beginning with a lattice, and rewiring edges with a given probability. The small-world networks are characterized by a high clustering coefficient and small average path length. The clustering coefficient considers each vertex of the graph individually, and compares the number of edges between neighbours to a complete graph as a ratio, while the average path length is the average minimum distance between two vertices. The Watts–Strogatz model [23] produces a graph with a small average path length, and a resulting clustering coefficient which is significantly higher than the corresponding coefficient of an Erdos–Renyi random graph model [38]. Importantly, networks of different

topologies may be considered as small-world networks as long as they are characterized by a relatively high clustering coefficient and a relatively low average path length. These features are of particular relevance to studies of contagion in many real-world scenarios. The degree distribution for a Watts–Strogatz network, interpolating between the ring lattice and a random graph, is similar to the distribution of a random graph but has a pronounced peak centred on the mean degree, decaying exponentially for degrees deviating from the mean [39].

Typically in SIS discrete time update models, the infection parameter  $\nu$  defines a per contact (edge) per time-step probability of transmission. That is, given an individual  $x_i$  in a neighbourhood with  $r$  infected individuals, the per time-step probability of infection  $P(x_i)$  [30] is

$$P(x_i) = 1 - (1 - \nu)^r. \quad (2.5)$$

The parameter  $\nu$  is analogous to the parameter  $\beta$  in the differential equations model; however, there is a subtle difference in interpretation:  $\beta$  is a continuous rate of transmission while  $\nu$  is a discrete probability of transmission per time step. As the update scheme is parallel, individuals all change states at the same time, i.e. recovery events from the current time step cannot affect infection on the same time step and vice versa. Given our objective of studying criticality in a complex distributed setting via the MaxEnt principle, the various network topologies provide a natural constraint on the testable information with respect to interactions within the population. This constraint imposed by heterogeneous networks presents the key challenge in obtaining a closed-form solution, unlike approaches based on mean field approximations.

### 3. Technical preliminaries

#### 3.1. The maximum entropy method

Often, we are faced with the problem of determining the least biased probability distribution, consistent with a set of specific constraints on the average values of measurable quantities. These may, for example, represent relevant conserved quantities in a thermodynamic system, or generic constraints in any probabilistic system of multinomial form (i.e. a system composed of a number of distinguishable entities allocated to equiprobable distinguishable categories) [40]. In the most general setting, this problem can be resolved by extracting the highest amount of (Shannon) information available. As pointed out by Jaynes, ‘in making inferences on the basis of partial information we must use that probability distribution which has maximal entropy subject to whatever is known. This is the only unbiased assignment we can make’ [41].

The Shannon entropy  $S$  of a discrete random variable  $A$  with state space  $\mathcal{A} = \{\sigma_1, \sigma_2, \sigma_3, \dots\}$ , the set of all possible states, is given by [42]

$$S(A) = \sum_{\sigma \in \mathcal{A}} -P(\sigma) \log(P(\sigma)), \quad (3.1)$$

where  $P(\sigma)$  is the probability that the system is in the state  $\sigma$ .

In order to extract the least biased probability distribution, one typically maximizes the Shannon entropy (3.1), subject to the normalization and  $K$  moment constraints on the system, shaped by some functions  $f_k(\sigma)$  with measurable

expectations  $\langle f_k \rangle$ , for  $k = 1, \dots, K$ :

$$\sum_{\sigma \in \mathcal{A}} P(\sigma) = 1 \quad (3.2)$$

and

$$\sum_{\sigma \in \mathcal{A}} P(\sigma) f_k(\sigma) = \langle f_k \rangle. \quad (3.3)$$

In practice, the problem is an optimization problem which can be solved using the method of Lagrange multipliers. Using the method of Lagrange multipliers, the form of the distribution which maximizes the entropy is

$$P(\sigma) = \exp\left(-\lambda_0 - \sum_{k=1}^K \lambda_k f_k(\sigma)\right), \quad (3.4)$$

where  $\lambda \equiv \{\lambda_1, \dots, \lambda_k, \dots, \lambda_K\}$  is the set of Lagrange multipliers corresponding to  $K$  constraints and  $\lambda_0$  is the ‘Massieu function’, the Lagrange multiplier corresponding to the normalization constraint. Introducing the generalized partition function  $Z(\lambda) = \exp \lambda_0$  yields

$$P(\sigma) = Z(\lambda)^{-1} \exp\left(-\sum_{k=1}^K \lambda_k f_k(\sigma)\right). \quad (3.5)$$

#### 3.2. Fisher information

The Fisher information is a measure of the information that an observable random variable  $X$  contains about a set of unknown parameters  $\lambda$ , defined as

$$F_X(\lambda) = E\left[\left(\frac{\partial}{\partial \lambda} \log P(x; \lambda)\right)^2\right], \quad (3.6)$$

which for continuous random variables is

$$F_X(\lambda) = \int \left(\frac{\partial}{\partial \lambda} \log P(x; \lambda)\right)^2 P(x; \lambda) dx, \quad (3.7)$$

where  $P(x; \lambda)$  is the probability density function of  $X$  conditional on the parameters  $\lambda$ .

For a joint random variable, the Fisher information has a chain rule decomposition [43] such that if  $X$  and  $Y$  are jointly distributed random variables,

$$F_{X,Y}(\lambda) = F_X(\lambda) + F_{Y|X}(\lambda). \quad (3.8)$$

If  $X$  and  $Y$  are independent random variables, the distribution of  $Y$  given  $X$  is the same as the distribution of  $Y$ , and therefore,  $F_{Y|X} = F_Y$  implying that

$$F_{X,Y}(\lambda) = F_X(\lambda) + F_Y(\lambda). \quad (3.9)$$

Often, it is important to reparametrize the Fisher information [44]:

$$F_X(\mu) = \left(\frac{d\lambda}{d\mu}\right)^2 F_X(\lambda(\mu)), \quad (3.10)$$

where  $\lambda$  and  $\mu$  are both parameterizations of  $X$ , and  $\lambda$  is a continuously differentiable function of  $\mu$ .

For many distributions the Fisher information is known exactly, and in particular, we shall use the closed form representation for the Fisher information of a Binomial( $n, q$ ) random variable  $\chi$ :

$$F_\chi(q) = \frac{n}{q(1-q)}. \quad (3.11)$$



### 3.3. Thermodynamic efficiency of computation

In a statistical mechanical setting, for thermodynamic variables  $\lambda$ , the solutions obtained according to the MaxEnt principle are characterized by probability densities in the form of the Gibbs measure:

$$P(\sigma|\lambda) = \frac{1}{Z(\lambda)} e^{-\beta H(\sigma, \lambda)} = \frac{1}{Z(\lambda)} e^{-\sum_k \lambda_k f_k(\sigma)}, \quad (3.12)$$

where the state functions  $f_k(\sigma)$  are defined over the configuration space,  $\beta = 1/k_B T$  is the inverse temperature  $T$  ( $k_B$  is the Boltzmann constant), and the Hamiltonian  $H(\sigma, \lambda)$  defines the total energy at state  $\sigma$  [45,46]. In other words, equation (3.12) expresses the state probability in terms of the state energy.

The Gibbs free energy of such system is given by

$$G(T, \lambda_m) = U(S, \phi_m) - TS - \phi_m \lambda_m, \quad (3.13)$$

where  $U$  is the internal energy of the system,  $S$  is the configuration entropy and  $\phi_m$  is an order parameter. In such a setting, the Fisher information quantifies the size of the fluctuations around equilibrium in the collective variables  $f_m$  and  $f_n$ , and is proportional to the curvature of the free entropy  $\psi = \ln Z = -\beta G$  [45–48]:

$$F_{mn}(\lambda) = \langle (X_m(x) - \langle X_m \rangle)(X_n(x) - \langle X_n \rangle) \rangle = \frac{\partial^2 \psi}{\partial \lambda_m \partial \lambda_n}. \quad (3.14)$$

It also identifies phase transitions and the corresponding critical thresholds [16], being proportional to the derivatives of the corresponding order parameters with respect to the thermodynamic variables  $\lambda$  [17]:

$$F_{mn}(\lambda) = \beta \frac{\partial \phi_m}{\partial \lambda_n}. \quad (3.15)$$

Furthermore, under a quasi-static protocol, the Fisher information can be interpreted as the generalized work  $W_{\text{gen}}$  [18]:

$$F(\lambda) = -\frac{d^2 \langle \beta W_{\text{gen}} \rangle}{d\lambda^2}. \quad (3.16)$$

Using equation (3.16), the rate of expended work can be expressed as follows:

$$\frac{d \langle \beta W_{\text{gen}} \rangle}{d\nu} = -\int_{\lambda^*}^{\lambda} F(\lambda') d\lambda', \quad (3.17)$$

where  $\lambda^*$  is the zero-response point for which small changes in the control parameter incur no work:

$$\left. \frac{d \langle \beta W_{\text{gen}} \rangle}{d\lambda} \right|_{\lambda=\lambda^*} = 0. \quad (3.18)$$

Having determined, via Fisher information, the rate of expended work carried out to generate order within the system, one may define the thermodynamic efficiency of computation [18] as ‘the reduction in uncertainty (i.e. the increase in order) from an expenditure of work given a value of the control parameter’:

$$\eta = \frac{-dS/d\lambda}{d \langle \beta W_{\text{gen}} \rangle / d\lambda}. \quad (3.19)$$

In this work, we shall extend the notion of the thermodynamic efficiency of computation to contagion processes.

## 4. Maximum entropy framework for epidemics

### 4.1. A network model of SIS epidemic

We will consider a graph  $G(\mathcal{V}, \mathcal{E})$  with vertex set  $\mathcal{V}$  and edge set  $\mathcal{E}$ . The nodes  $i \in \mathcal{V} = \{1, 2, 3, \dots\}$  represent individuals in the population taking one of two states: susceptible or infected. The state of vertex  $i$  will be denoted  $\sigma_i$  and will take value 0 if the individual is susceptible or 1 if the individual is infected. The edges  $(i, j) \in \mathcal{E}$  represent connections between two individuals  $i$  and  $j$  along which infection can spread. The state of the entire system,  $\sigma$  will be expressed as the vector comprising the states of all individuals. That is, the  $i$ th element of  $\sigma$  is  $\sigma_i$ . We will denote the set of all states  $\sigma$  as  $\mathcal{A}$ .

The epidemic dynamics which we will investigate within the MaxEnt framework are SIS epidemics spreading on a network. The analytical results will be contrasted with the simulation results obtained by stochastic discrete-time parallel-update SIS model introduced by [30]. Similar to the deterministic SIS model, recovery of individuals in this model occurs independently of their neighbours, with a constant probability of recovery, denoted  $\delta$ . This is analogous to the parameter  $\gamma$  from the SIS differential equation model, with the difference being that  $\gamma$  represents a rate, whereas  $\delta$  is a probability per discrete time step.

We aim to investigate the MaxEnt distribution corresponding to a SIS epidemic spreading on a graph  $G(\mathcal{V}, \mathcal{E})$ , constrained by the testable information formed by the averages of two variables in the steady state of the SIS dynamics. These averages, thermodynamically corresponding to conserved quantities, are defined as follows:

$$I(\sigma) = \sum_{\sigma_i \in \sigma} \sigma_i \quad (4.1)$$

and

$$C(\sigma) = \sum_{\sigma_i \in \sigma} \sigma_i \sum_{j \in N_i} 1 - \sigma_j, \quad (4.2)$$

where  $N_i$  denotes the neighbourhood of node  $i$ . The quantity  $I(\sigma)$  is the total number of infected individuals in a configuration  $\sigma$  and is of clear interest in the study of infectious diseases. The quantity  $C(\sigma)$  is the number of neighbouring individuals who have opposite states. In the context of epidemic modelling, this is the number of potentially infective connections in a configuration  $\sigma$ . In analogy to an electromagnetic spin model such as the Ising model, the potentially infective connections ( $C$ ) correspond to the node–node interaction energy, whereas the number of infected individuals ( $I$ ) corresponds to the energy due to the applied external magnetic field.

The SIS discrete-time update dynamics have both an equilibrium state (i.e. the absorbing state of the system) and a metastable state (i.e. a state which takes time exponential in the number of vertices to leave). The equilibrium state corresponds to the trivial case in which  $I = 0$ . At the metastable state, however, the rate at which infected individuals are recovering and the rate at which susceptible individuals are being infected are equal. Between each time step, the rate of recovery is proportional to  $I$ , whereas the instantaneous rate of infection is proportional to  $C$ .

## 4.2. Maximizing entropy

In order to obtain the most probable distribution of infection within the simulated population during a steady state, we use the MaxEnt method with constraints on the average value of  $I$  as given by equation (4.1) and  $C$  as given by equation (4.2). Formally, the MaxEnt principle for this system with SIS discrete time update dynamics forms the optimization problem

$$\max_{P(\boldsymbol{\sigma})} S(A), \quad \text{for } S(A) = - \sum_{\boldsymbol{\sigma} \in \mathcal{A}} P(\boldsymbol{\sigma}) \log(P(\boldsymbol{\sigma})), \quad (4.3)$$

subject to

$$\langle I \rangle = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} P(\boldsymbol{\sigma}) \sum_{\sigma_i \in \boldsymbol{\sigma}} \sigma_i, \quad (4.4)$$

$$\langle C \rangle = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} P(\boldsymbol{\sigma}) \sum_{\sigma_i \in \boldsymbol{\sigma}} \sigma_i \left( \sum_{j \in N_i} 1 - \sigma_j \right) \quad (4.5)$$

and

$$1 = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} P(\boldsymbol{\sigma}). \quad (4.6)$$

The form of this MaxEnt solution for  $K$  constraints is given by equation (3.5) specifying Gibbs distribution. The specific Gibbs distribution consistent with the constraints given by equations (4.4)–(4.6) is

$$P(\boldsymbol{\sigma}) = \frac{e^{\lambda_1 I(\boldsymbol{\sigma}) + \lambda_2 C(\boldsymbol{\sigma})}}{Z}, \quad (4.7)$$

where  $I(\boldsymbol{\sigma})$  and  $C(\boldsymbol{\sigma})$  are defined by (4.1) and (4.2), respectively, and  $\lambda_1, \lambda_2$  and  $Z$  are unknown Lagrange multipliers. Thus, in order to solve for the Lagrange multipliers and obtain the MaxEnt distribution consistent with known information, one must use the the averages of  $I$  and  $C$ .

For clarity, we abbreviate

$$x = e^{\lambda_1} \quad \text{and} \quad y = e^{\lambda_2}. \quad (4.8)$$

Substituting (4.8) into (4.4)–(4.7), we obtain

$$P(\boldsymbol{\sigma}) = \frac{x^{I(\boldsymbol{\sigma})} y^{C(\boldsymbol{\sigma})}}{Z}, \quad (4.9)$$

$$\langle I \rangle = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} I(\boldsymbol{\sigma}) \frac{x^{I(\boldsymbol{\sigma})} y^{C(\boldsymbol{\sigma})}}{Z}, \quad (4.10)$$

$$\langle C \rangle = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} C(\boldsymbol{\sigma}) \frac{x^{I(\boldsymbol{\sigma})} y^{C(\boldsymbol{\sigma})}}{Z} \quad (4.11)$$

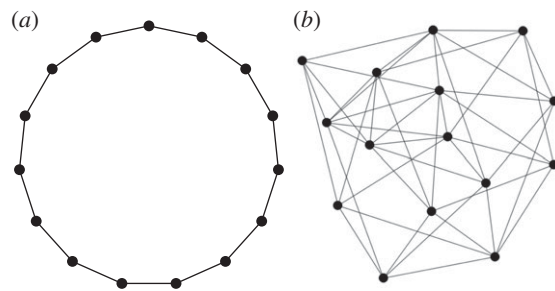
and

$$1 = \sum_{\boldsymbol{\sigma} \in \mathcal{A}} \frac{x^{I(\boldsymbol{\sigma})} y^{C(\boldsymbol{\sigma})}}{Z}. \quad (4.12)$$

Let us define the sets  $\aleph_{I,C}$ , the sets of all configurations  $\boldsymbol{\sigma}$  such that the total number of infected individuals  $I(\boldsymbol{\sigma}) = I$ , and the number of potentially infective connections  $C(\boldsymbol{\sigma}) = C$ ; formally, we have

$$\aleph_{I,C} = \{\boldsymbol{\sigma} : I(\boldsymbol{\sigma}) = I, C(\boldsymbol{\sigma}) = C\}. \quad (4.13)$$

The sets  $\aleph_{I,C}$  form a partition of  $\mathcal{A}$ , the space of all configurations. Hence, each configuration  $\boldsymbol{\sigma}$  belongs to exactly one of the sets  $\aleph_{I,C}$ . As all elements of the set  $\aleph_{I,C}$  have the same value of  $I(\boldsymbol{\sigma})$  and  $C(\boldsymbol{\sigma})$ , we see from equation (4.9) that all states which belong to the same set  $\aleph_{I,C}$  have identical probability.



**Figure 1.** Example networks of 15 nodes for numerically testing the MaxEnt model. (a) A ring network and (b) Watts–Strogatz network with  $p = 1$  and  $k = 6$ .

Specifically, the probability of a state  $\boldsymbol{\sigma} \in \aleph_{I,C}$  is

$$P(\boldsymbol{\sigma} \in \aleph_{I,C}) = \frac{x^I y^C}{Z}. \quad (4.14)$$

As each of the states  $\boldsymbol{\sigma}$  in a set  $\aleph_{I,C}$  has identical probability, this naturally follows the concept of a macrostate  $(I, C)$ . By denoting the cardinality of the set  $\aleph_{I,C}$  by  $N(I, C)$ , we simplify the system of equations (4.10)–(4.12) as

$$\langle I \rangle = \sum_{I,C} N(I, C) I \frac{x^I y^C}{Z}, \quad (4.15)$$

$$\langle C \rangle = \sum_{I,C} N(I, C) C \frac{x^I y^C}{Z} \quad (4.16)$$

and

$$1 = \sum_{I,C} N(I, C) \frac{x^I y^C}{Z}. \quad (4.17)$$

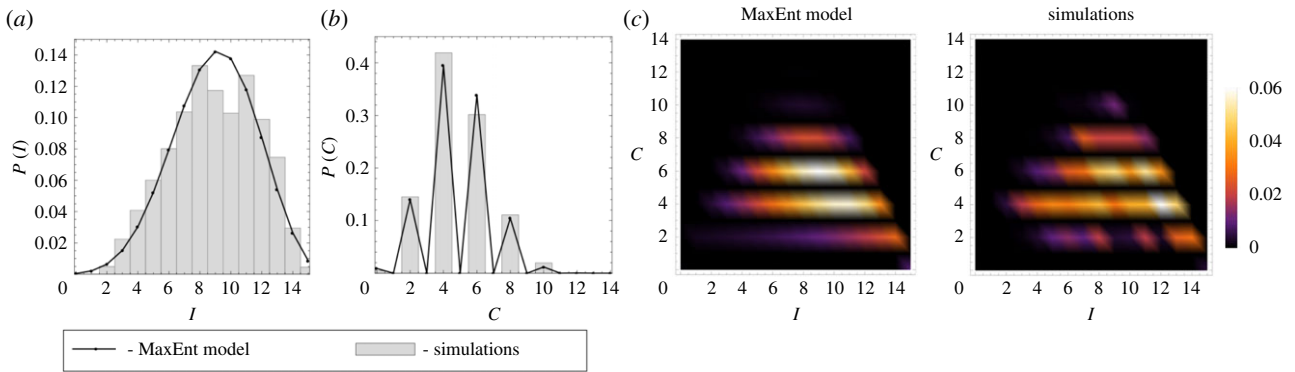
This reformulation opens a way to MaxEnt solutions expressed in terms of the probabilities  $P(I, C)$  defined over the macrostates  $(I, C)$ .

## 4.3. Example: numerical solutions for small graphs

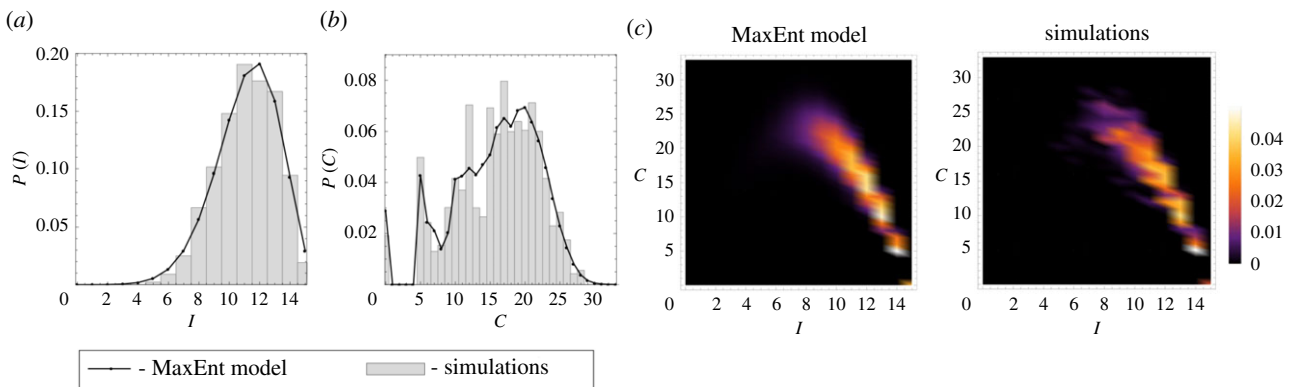
The MaxEnt equations (4.15)–(4.17) is a system of polynomial equations in  $x, y$  and  $Z$ , which in general do not have an analytic solution. The solutions to these equations can be obtained numerically as we will demonstrate in this section.

To exemplify the numerical construction of our MaxEnt model of SIS epidemics on complex networks, we have chosen two graph topologies shown in figure 1. The first graph is a ring and the second graph is a Watts–Strogatz random graph with  $p = 1$  and  $k = 6$ . Both graphs consist of 15 nodes. As these networks are small it is straightforward to compute the functions  $N(I, C)$  by listing all possible state configurations and counting the number of configurations with specific values of  $I$  and  $C$ . Such method of computing  $N(I, C)$  is feasible for small networks, but for larger networks one should make use of more sophisticated combinatorial algorithms.

In practical applications, the values of constraints,  $\langle I \rangle$  and  $\langle C \rangle$ , would be taken from field data. For our purposes, we can obtain the ‘data’ by computationally simulating the dynamics of an SIS epidemic. We simulate the dynamics using SIS parallel update processes stochastic simulation with arbitrarily chosen values of probability of infection transmission  $\nu$  and probability of recovery  $\delta$ . The simulation involves initializing the system at random and then updating the state of the system for 120 000 time steps. The first 20 000 time steps are



**Figure 2.** The observed simulation and MaxEnt probability density functions (a)  $P(I)$ , (b)  $P(C)$  and (c)  $P(I, C)$  for a 15 node ring graph (figure 1a). The SIS simulation parameters were  $\nu = 1.0 \times 10^{-3}$ ,  $\delta = 6.0 \times 10^{-4}$ . The empirical values of the constraints for the MaxEnt equations were  $\langle I \rangle = 8.9323$  and  $\langle C \rangle = 4.8542$ . The computed Lagrange multipliers were  $\lambda_1 = 0.1938$  and  $\lambda_2 = -0.6957$ .



**Figure 3.** The observed simulation and MaxEnt probability density functions (a)  $P(I)$ , (b)  $P(C)$  and (c)  $P(I, C)$  for a 15 node random graph (figure 1b). The SIS simulation parameters were  $\nu = 3.288 \times 10^{-4}$  and  $\delta = 5.0 \times 10^{-4}$ . The empirical values of the constraints for the MaxEnt equations were  $\langle I \rangle = 11.1624$  and  $\langle C \rangle = 15.8917$ . The computed Lagrange multipliers were  $\lambda_1 = 0.6541$  and  $\lambda_2 = -0.1500$ .

used to equilibrate the system and the final 100 000 time steps are used for sampling the equilibrium state. At every time step, the number of infected individuals  $I$  and number of infected connections  $C$  is recorded, and this is used to compute  $\langle I \rangle$  and  $\langle C \rangle$ . The values of  $\langle I \rangle$  and  $\langle C \rangle$  are then used as constraints in equations (4.15)–(4.16). The equations (4.15)–(4.17) are solved numerically for  $x$ ,  $y$  and  $Z$ .

Figures 2 and 3 compare the computed MaxEnt probability distributions with distributions from simulation data for the ring and random networks. There is a good agreement between the empirical and MaxEnt distributions  $P(I)$ ,  $P(C)$  and  $P(I, C)$  for both ring and random network topologies supporting the validity of the MaxEnt method.

The numerical solutions of MaxEnt equations are of general applicability. However, in order to arrive at epidemiological interpretation of the Lagrange multipliers  $\lambda_1$  and  $\lambda_2$ , we would like to express them analytically in terms of  $\langle I \rangle$  and  $\langle C \rangle$ . The analytic solution may only be obtained after invoking certain simplifying assumptions, in particular, independence between  $P(I)$  and  $P(C)$ . The derivation and the analysis of the analytic solutions are presented in the following sections.

#### 4.4. Example: complete graph

We will start our analytic analysis by considering the specific case in which  $G(V, E)$  is a complete graph, with  $V$  vertices and  $E = V(V - 1)/2$  edges. Given  $I$  infected individuals, for

each infected node there are  $V - I$  susceptible neighbours. Therefore, there are  $C = I(V - I)$  potentially infective connections. It is worth noting that although a given number of infected individuals on this topology defines precisely the number of potentially infective connections, the converse is not true in general. That is, for a number of potentially infective connections there is generally more than one corresponding number of infected individuals.

Each  $I$  specifies the value of  $C$ ; however each value of  $C$  often defines exactly two values of  $I$ . As knowing  $I$  uniquely defines  $C$ ,  $\mathfrak{N}_{I,C} = \mathfrak{N}_I = \{\sigma : I(\sigma) = I\}$ . Consequently,  $N(I, C) = N(I)$ . In a complete graph with  $V$  vertices,  $N(I) = \binom{V}{I}$ , analogous to the outcome of placing  $I$  balls into  $V$  buckets, resulting in the binomial coefficient. Therefore, the total number of infected individuals, represented by the constraint (4.15), is a sum from  $I = 0$  to  $V$ , giving

$$\langle I \rangle = \sum_{I=0}^V \binom{V}{I} I \frac{x^I y^{I(V-I)}}{Z}. \quad (4.18)$$

Similarly, the constraint (4.16) yields

$$\langle I(V - I) \rangle = \sum_{I=0}^V \binom{V}{I} I(V - I) \frac{x^I y^{I(V-I)}}{Z}, \quad (4.19)$$

while the normalization constraint (4.17) becomes

$$1 = \sum_{I=0}^V \binom{V}{I} \frac{x^I y^{I(V-I)}}{Z}. \quad (4.20)$$

This is not analytically reducible further, and in the next subsection we consider a simplified system where an exact solution can be found.

Henceforth, for a graph with  $V$  vertices and  $E$  edges, we shall use average quantities  $\langle I^* \rangle = \langle I \rangle / V$  and  $\langle C^* \rangle = \langle C \rangle / E$ .

#### 4.5. Example: an initial seeding state

In order to illustrate how the MaxEnt principle yields an analytical solution, we consider a very simple system with two constraints: on the number of infected individuals  $I$ , and the usual normalization constraint. This abridged case (not limited to the complete graph topology) describes the precursor state of the epidemic with a number of infection sources distributed within the network. This state essentially corresponds to the initial state of an outbreak, before any infective transmissions have taken place, and agent-based simulation studies often focus on such initial 'seeding' state in quantifying effects of different seeding scenarios [49–51]. Nevertheless, this case will reveal an important thermodynamic analogy between key variables of SIS model and the inverse temperature of Gibbs distribution resulting from entropy maximization. Explicitly stated, we wish to find the MaxEnt solution in the form

$$P(\sigma) = \frac{e^{\lambda I(\sigma)}}{Z}, \quad (4.21)$$

subject to the following constraints:

$$\frac{1}{Z} \sum_{I=0}^V I \binom{V}{I} x^I = \langle I \rangle \quad (4.22)$$

and

$$\frac{1}{Z} \sum_{I=0}^V \binom{V}{I} x^I = 1. \quad (4.23)$$

The binomial theorem transforms (4.23) into

$$Z = (x + 1)^V, \quad (4.24)$$

where  $x = e^\lambda$  is a positive real number. Substituting (4.24) into (4.22) yields

$$\langle I \rangle = \frac{\sum_{I=0}^V I \binom{V}{I} x^I}{(x + 1)^V} = \frac{Vx(x + 1)^{V-1}}{(x + 1)^V} = \frac{Vx}{(x + 1)}. \quad (4.25)$$

One may be interested in tracing the proportion of the infected individuals within the total population,  $\langle I^* \rangle = \langle I \rangle / V$ , which is immediately obtained from (4.25):

$$\langle I^* \rangle = \frac{x}{x + 1}. \quad (4.26)$$

Hence, the Lagrange multiplier corresponding to the constraint (4.22),  $\lambda = \log x$ , is

$$\lambda = \log \frac{\langle I^* \rangle}{1 - \langle I^* \rangle}.$$

Finally, substituting this expression into the solution (4.21), we obtain the Gibbs distribution:

$$P(\sigma) = \frac{e^{(\log(\langle I^* \rangle / (1 - \langle I^* \rangle))) I(\sigma)}}{Z}. \quad (4.27)$$

Thus, interpreting  $I(\sigma)$  as the energy of the system, we may derive a thermodynamic analogy of the equilibrium inverse temperature in the SIS epidemic model as negative  $\lambda$ , that is,

$\beta = \log((1 - \langle I^* \rangle) / \langle I^* \rangle)$ , i.e. the log-ratio between non-infected and infected proportions during the initial state of an outbreak.

The partition function, i.e. the Lagrange multiplier corresponding to the normalization constraint (4.23), can be explicitly resolved by using (4.26) and (4.24):

$$Z = \left( \frac{1}{1 - \langle I^* \rangle} \right)^V.$$

Therefore,

$$P(I) = \binom{V}{I} \langle I^* \rangle^I (1 - \langle I^* \rangle)^{V-I}. \quad (4.28)$$

This is the binomial distribution with the 'probability of success'  $\langle I^* \rangle$  during  $V$  trials, characterizing the initial state of an epidemic outbreak.

#### 4.6. Assuming independence

In order to obtain an analytical solution to the more general problem that includes, in addition, the constraint on the number of infective links (4.5), we introduce an assumption about the random variables  $I$  and  $C$ . Specifically, we assume that  $I$  and  $C$  are independent:  $P(I, C) = P(I)P(C)$ , and as a result,  $N(I, C) = N(I)N(C)$ .

Under this assumption, the average of  $I$  may only be used to infer the distribution of  $I$ . Similarly, the average of  $C$  may only be used to infer the distribution of  $C$ . This allows us to reduce the constraints (4.15)–(4.17) as follows:

$$\langle I \rangle = \sum_{I \in V} IN(I) \frac{x^I}{Z_I}, \quad (4.29)$$

$$\langle C \rangle = \sum_{C \in E} CN(C) \frac{y^C}{Z_C}, \quad (4.30)$$

$$1 = \sum_{I \in V} \frac{x^I}{Z_I} \quad (4.31)$$

and

$$1 = \sum_{C \in E} \frac{y^C}{Z_C}, \quad (4.32)$$

where  $Z_C Z_I = Z$ . We then follow the derivations outlined in §§4.4 and 4.5, solving (4.29) and (4.31) for  $P(I) = x^I / Z_I$ , and (4.30) and (4.32) for  $P(C) = y^C / Z_C$  separately. Doing so, we find that  $P(I)$  is a binomial random variable with  $V$  trials and the probability of success  $\langle I^* \rangle$  and that  $P(C)$  is a binomial random variable with  $E$  trials and the probability of success  $\langle C^* \rangle$ . Thus, our next result is that

$$P(I, C) = \binom{V}{I} \binom{E}{C} \langle I^* \rangle^I (1 - \langle I^* \rangle)^{V-I} \langle C^* \rangle^C (1 - \langle C^* \rangle)^{E-C}. \quad (4.33)$$

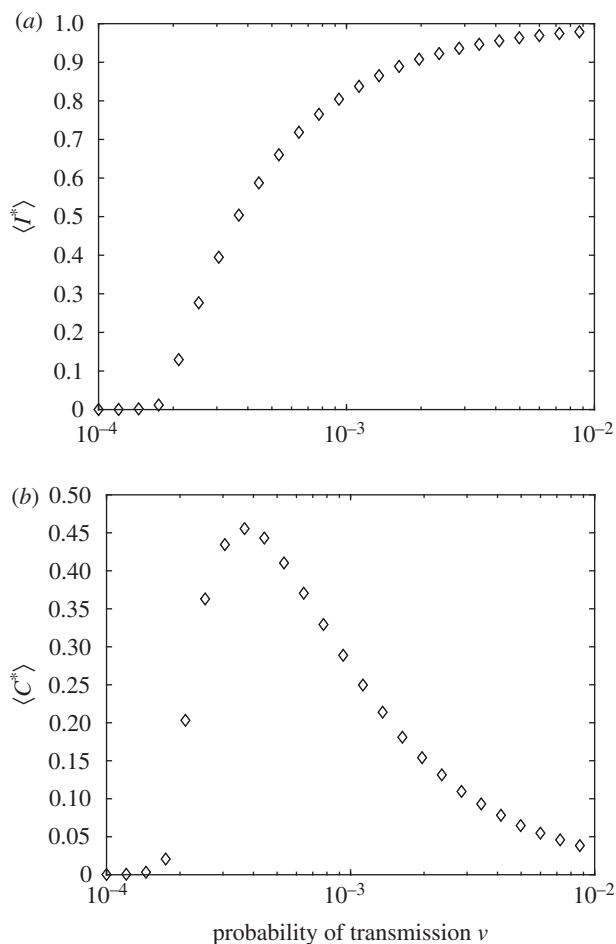
Under the assumption that  $I$  and  $C$  are independent, we may express the entropy of the system in its steady state, that is, the entropy of the joint variable  $(I, C)$ , as the sum of the individual entropies of  $I$  and  $C$ :

$$H(I, C) = H(I) + H(C).$$

Furthermore, given the partition functions of the marginal probability distributions obtained as

$$Z_I = \left( \frac{1}{1 - \langle I^* \rangle} \right)^V \quad \text{and} \quad Z_C = \left( \frac{1}{1 - \langle C^* \rangle} \right)^E,$$





**Figure 4.** The constraints  $\langle I^* \rangle$  (a) and  $\langle C^* \rangle$  (b) as functions of the probability of infection  $\nu$ .

we derive the free entropy of the entire system as

$$\begin{aligned} \log(Z) &= \log(Z_I Z_C) = \log(Z_I) + \log(Z_C) \\ &= -V \log(1 - \langle I^* \rangle) - E \log(1 - \langle C^* \rangle). \end{aligned} \quad (4.34)$$

#### 4.7. Fisher information

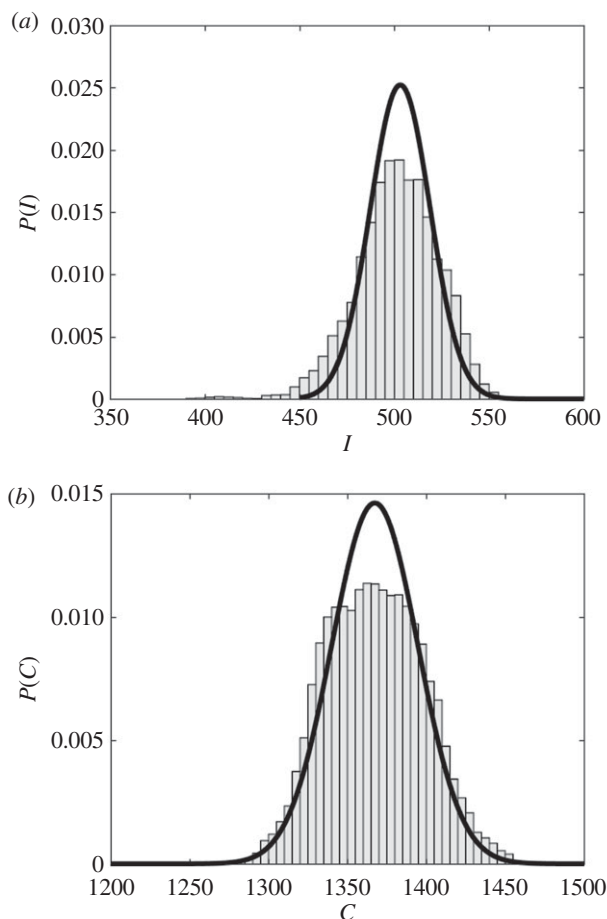
Now we will use the derived Gibbs distributions in expressing the Fisher information of the joint random variable  $(I, C)$  and characterizing critical regimes of SIS epidemics. Under the assumption that  $I$  and  $C$  are independent,  $F_{I,C}(\nu) = F_I(\nu) + F_C(\nu)$  [43].

Our control parameter in this case is  $\nu$ , and the Fisher information of the joint variable  $(I, C)$  describing the macroscopic state of the system will be derived with respect to  $\nu$ , and then with respect to the average constraints  $I^*$  and  $C^*$ .

$$\begin{aligned} F_I(\nu) &= E \left[ \frac{\partial^2}{\partial \nu^2} \log P(I|\nu) \right] \\ &= E \left[ \frac{\partial^2}{\partial \langle I^* \rangle^2} \log P(I|\langle I^* \rangle) \left( \frac{\partial \langle I^* \rangle}{\partial \nu} \right)^2 \right]. \end{aligned} \quad (4.35)$$

The first term within the expectation operator is exactly the Fisher information of  $I$  with respect to  $\langle I^* \rangle$ , i.e.  $F_I(\langle I^* \rangle)$ , yielding

$$F_I(\nu) = F_I(\langle I^* \rangle) \left( \frac{\partial \langle I^* \rangle}{\partial \nu} \right)^2. \quad (4.36)$$



**Figure 5.** The observed simulation and MaxEnt probability density functions of  $I$  (a) and  $C$  (b). The histogram shows the empirically observed probability density function with bin widths of 5, whereas the solid line gives the probability density function of the MaxEnt solution for  $\nu = 3.68 \times 10^{-4}$  and  $\delta = 0.001$ .

Analogously,

$$F_C(\nu) = F_C(\langle C^* \rangle) \left( \frac{\partial \langle C^* \rangle}{\partial \nu} \right)^2. \quad (4.37)$$

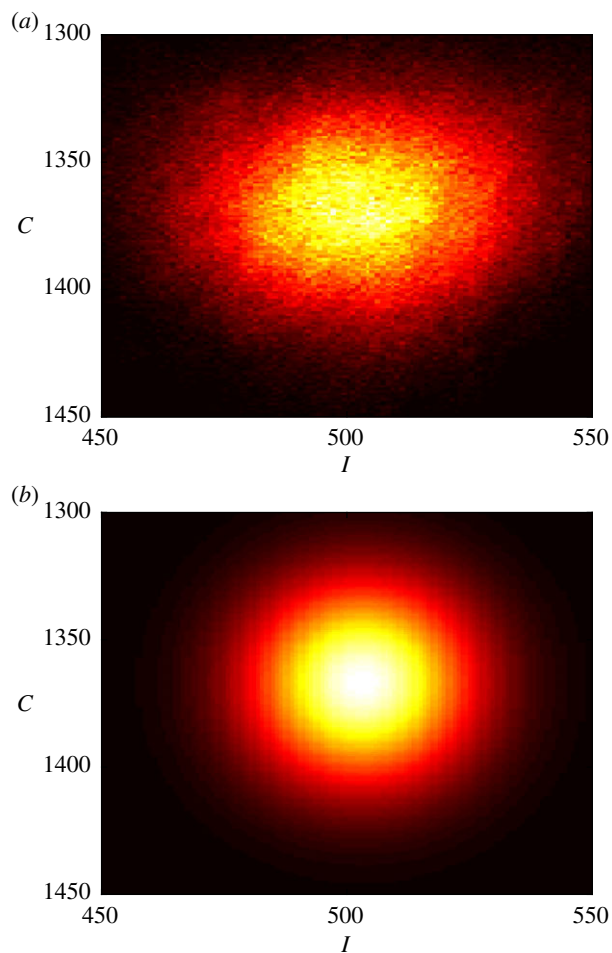
Thus, the Fisher information of the joint variable  $(I, C)$  with respect to the control parameter  $\beta$  can be expressed via Fisher information with respect to the conserved quantities  $I^*$  and  $C^*$ :

$$F_{I,C}(\nu) = F_I(\langle I^* \rangle) \left( \frac{\partial \langle I^* \rangle}{\partial \nu} \right)^2 + F_C(\langle C^* \rangle) \left( \frac{\partial \langle C^* \rangle}{\partial \nu} \right)^2.$$

Using binomial random variables  $(V, \langle I^* \rangle)$  and  $(E, \langle C^* \rangle)$  in the expression for Fisher information (3.11), we can produce our next result, directly expressing the Fisher information of the MaxEnt distribution in terms of the conserved quantities  $I^*$  and  $C^*$ :

$$F_{I,C}(\nu) = \frac{V}{\langle I^* \rangle (1 - \langle I^* \rangle)} \left( \frac{\partial \langle I^* \rangle}{\partial \nu} \right)^2 + \frac{E}{\langle C^* \rangle (1 - \langle C^* \rangle)} \left( \frac{\partial \langle C^* \rangle}{\partial \nu} \right)^2. \quad (4.38)$$

Importantly, the independence assumption, which allowed us to obtain analytic expressions for the Fisher information (4.36)–(4.38), reveals the mechanism for its divergence at criticality. It is known that the divergence of the Fisher information indicates a critical regime and pinpoints the critical threshold. In this instance, it shows that, when  $0 < \langle I^* \rangle < 1$  or  $0 < \langle C^* \rangle < 1$ , the phase transition



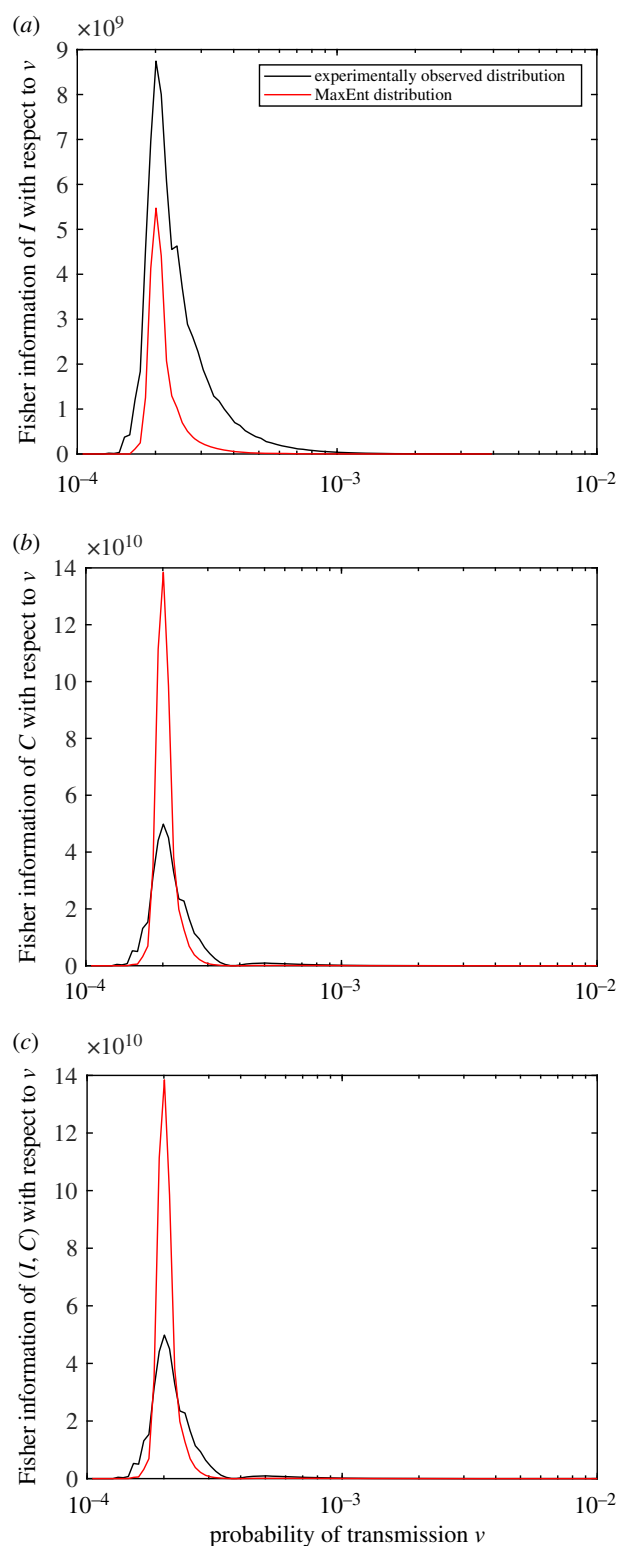
**Figure 6.** The observed (a) and the MaxEnt (b) probability density functions of the joint random variable  $(I, C)$ . The probability distributions are generated from simulations with  $\nu = 3.68 \times 10^{-4}$  and  $\delta = 1 \times 10^{-3}$ .

should occur when the derivative of  $\langle I^* \rangle$  or  $\langle C^* \rangle$  with respect to the control parameter  $\nu$  becomes infinite (i.e. has a vertical tangent). In other words, one of the implications of the independence assumption is that the mechanism behind criticality is explicitly linked to non-differentiability of  $\langle I^* \rangle$  or  $\langle C^* \rangle$ .

In order to explicitly distinguish between the MaxEnt distribution and the distributions observed from simulations, we introduce  $I_M, C_M$ , the random variables distributed according to the MaxEnt solutions given in (4.33); and  $I_O, C_O$ , the random variables associated to the observed probability distributions of our simulations (see electronic supplementary material, S.1 and S.2, for details of numerically calculating the corresponding Fisher information and thermodynamic efficiency of computation, respectively).

## 5. Numerical results

We constrain our analysis to the specific case of Watts–Strogatz random graphs [23] across a broad range of simulated pathogens. Having considered the statistical mechanics of SIS processes under the assumption of independence, we will compare our analytical derivations with the results of numerical simulations of the steady-state probability distributions. To reiterate, these probability distributions are obtained from computational simulations of stochastic discrete-time parallel-update SIS dynamics on Watts–Strogatz random



**Figure 7.** The Fisher information of (a)  $I$ , (b)  $C$  and (c) the joint variable  $(I, C)$  as a function of  $\nu$ , the probability of infection per time step. The value of the Fisher information is an average of values calculated for individual graphs. The peak of each of these plots occurs at  $\nu = 2.01 \times 10^{-4}$ , corresponding to  $R_0 = 1.004$ . (Online version in colour.)

graphs with 1000 nodes, parameters  $k = 6$  and  $p = 1$ , varying the probability of infection  $\nu$  and holding the probability of recovery,  $\delta = 0.001$ , constant. Probability distributions are observed for each of 30 graphs across 30 realizations over 100 000 time steps per run for each value of  $\nu$ . For each run, we perform a logarithmic sweep of the parameter  $\nu$  from  $10^{-4}$  to  $10^{-2}$  with 100 steps in this range. In order to obtain

an appropriate representation of the stationary distributions, we do not record any data until the system has sufficient time to equilibrate (40 000 time steps). In addition to analysis of the steady-state probability distributions, we study important thermodynamic quantities such as the entropy, Fisher information and free entropy of the system.

Firstly, we study the constraints  $\langle I^* \rangle$  and  $\langle C^* \rangle$  as a function of the parameter  $\nu$ , the probability of infection, as shown in figure 4. These constraints are averaged over all simulations described above.  $\langle I^* \rangle$  and  $\langle C^* \rangle$  set the constraints used in order to obtain the MaxEnt probability distribution.

Secondly, we investigate the two types of marginal and joint probability distributions, shown in figures 5 and 6, respectively. Each of the figures shows both the distribution obtained empirically from computer simulations and the distribution obtained from the MaxEnt solution (4.33), given the average quantities. Specifically, the empirical distributions allowed us to estimate the average quantities  $\langle I \rangle$  and  $\langle C \rangle$ , shown earlier in figure 4*a,b*, used in deriving the MaxEnt solutions. These results are obtained from simulations with  $\nu = 3.68 \times 10^{-4}$  and  $\delta = 0.001$ , where  $\nu$  is the per time step probability of infection and  $\delta$  is the per time step probability of recovery. Figure 5 compares the marginal distributions of  $I$  and  $C$  observed from simulations to the MaxEnt distributions obtained only from the observations of the average values  $\langle I \rangle$  and  $\langle C \rangle$  on a particular Watts–Strogatz graph. Secondly, figure 6*a,b* demonstrates the differences between the experimentally observed distribution of  $(I, C)$  and the distributions of  $(I, C)$  according to the MaxEnt principle.

### 5.1. Fisher information

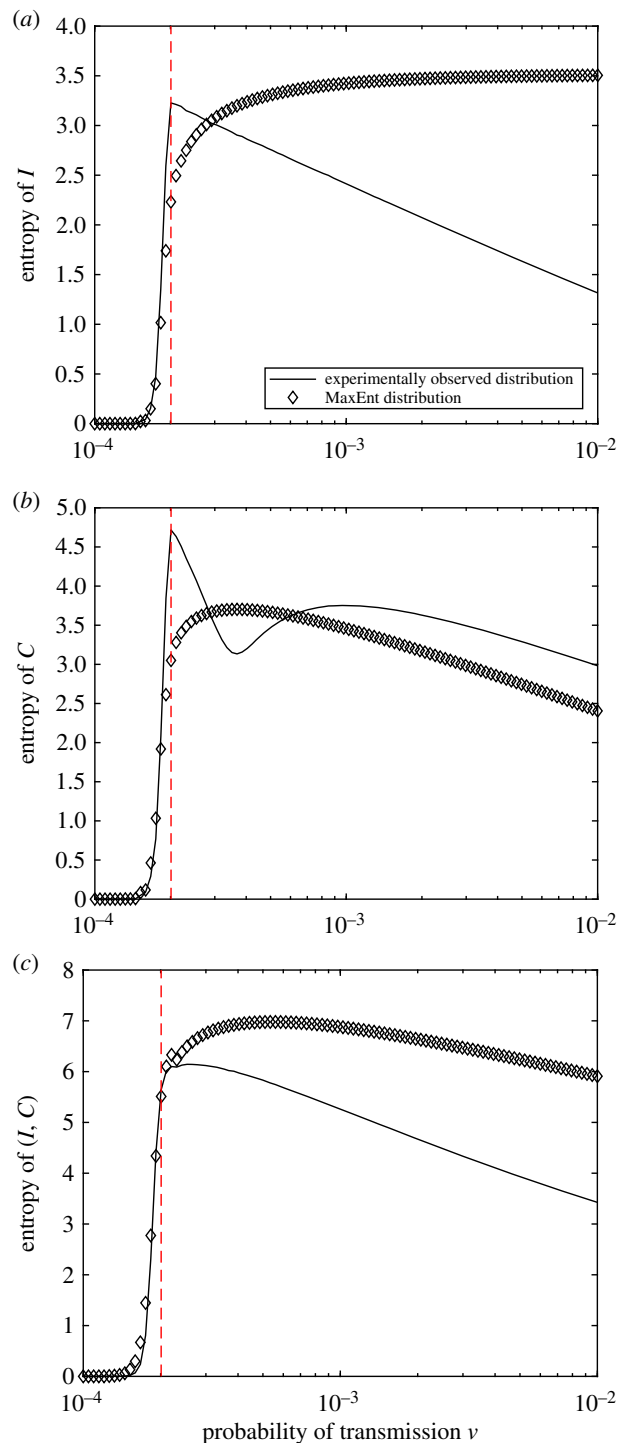
As mentioned earlier, it is well known that the Fisher information diverges at a phase transition [17]. We now compare the Fisher information for both the MaxEnt and observed distributions. This comparison is carried out with the change of variables given by electronic supplementary material, equation (S.1), with  $u_{\min} = -4$  and  $\Delta u = 4.65 \times 10^{-2}$  for 100 values of  $u_i = u_{\min} + i\Delta u$ , using electronic supplementary material, equation (S.7), to calculate the Fisher information of the observed distributions. For each value of  $\nu$ , a single probability distribution is obtained for each graph using multiple runs, and then the Fisher information is calculated, as outlined in §5, by averaging over each of the graphs.

As shown in figure 7, for  $\delta = 1 \times 10^{-3}$ , the Fisher information of each of the random variables peaks around  $2.01 \times 10^{-4}$ . For this system, this corresponds to  $R_0 = 1.004$  (see electronic supplementary material, equation (S.16)). Owing to uncertainty in the location of this maximum,  $\nu \in (1.9179 \times 10^{-4}, 2.1049 \times 10^{-4})$ , there is corresponding uncertainty in the value of  $R_0$ . In this case,  $R_0$  is in the interval (0.9657, 1.0435). Importantly, the approach identifies the critical threshold for the epidemic phase transition.

As pointed out in §3.3, under a quasi-static protocol changing the control parameter  $\nu$ , the Fisher information may also be interpreted as the generalized work,  $W_{\text{gen}}$  (3.16).

### 5.2. Entropy and free entropy

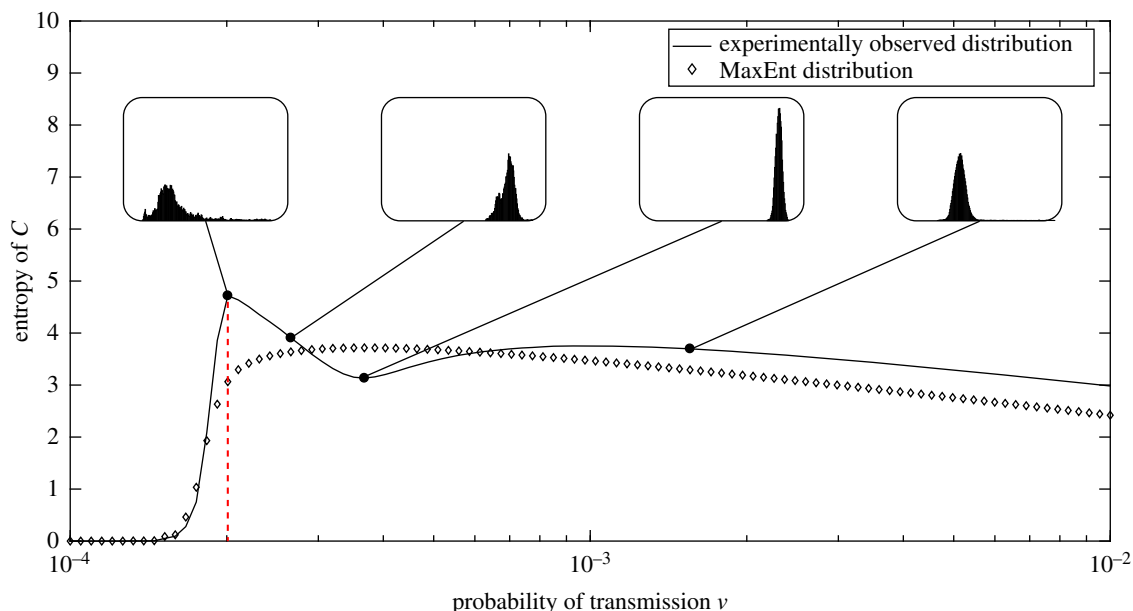
We now turn our attention to other thermodynamic characteristics, beginning with the configuration entropy (3.1), and again comparing the observed and the MaxEnt distributions across a range of values of  $\nu$ . Figure 8*a* shows that, after a



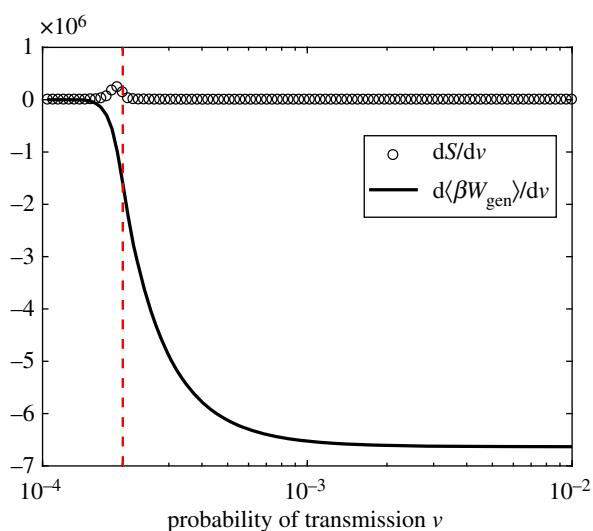
**Figure 8.** The Shannon entropy of the random variables (a)  $I$ , (b)  $C$  and (c)  $(I, C)$  as a function of  $\nu$ , the probability of transmission. The dotted line indicates the peak of the Fisher information. (Online version in colour.)

phase transition, there is decreasing log-linear relationship between the entropy of  $I$  and  $\nu$  for the distributions obtained from the simulations. By contrast, the entropy of the MaxEnt distributions does not show this dependency, resulting in an asymptotic overestimate.

We also note that although the entropy of the MaxEnt solution for  $C$  captures the overall trend of the observed entropy, as shown in figure 8*b*, it does not precisely capture the qualitative behaviour of the entropy of  $C$ , ‘smoothing’ the drop in entropy immediately following the phase transition. Figure 9 shows the distributions of  $C$  inset for multiple values of  $\nu$ . Furthermore, we note a good general agreement for the entropy of the joint variable, illustrated by figure 8*c*.



**Figure 9.** The entropy of the random variable  $C$  as a function of the parameter  $\nu$  for distributions obtained from simulations and the MaxEnt distributions. The dotted line indicates the value of  $\nu$  for which the Fisher information peaks, i.e. the critical value of  $\nu$ . Shown inset for comparison are the observed probability distributions for several values of  $\nu$ . (Online version in colour.)

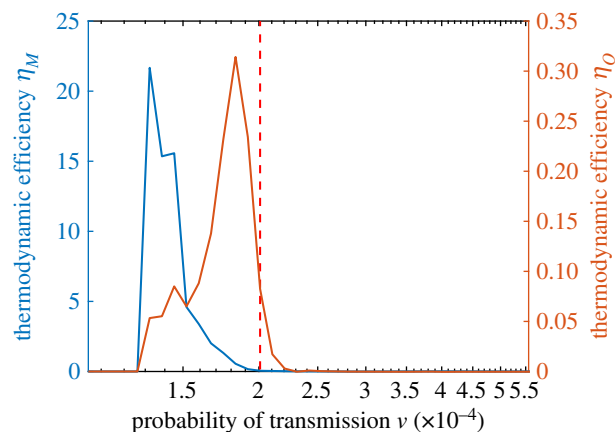


**Figure 10.** Rates of change of work and configuration entropy with respect to  $\nu$ , the probability of transmission, for observed distributions. The dotted line indicates the critical point at which the Fisher information peaks. (Online version in colour.)

These discrepancies and similarities clarify the impact of the independence assumption on the thermodynamics of the epidemics: despite disagreements in a ‘super-critical’ phase, the phase transition itself has been identified equally well by both the MaxEnt distributions and the distributions obtained from the simulations.

### 5.3. Thermodynamic efficiency of computation

Figure 10 illustrates the two key components of the thermodynamic efficiency of computation. We observe three qualitatively distinct regions in the space of  $\nu$ , which can be associated with subcritical, critical and super-critical regions. Most strikingly, we note that in both the subcritical and super-critical regions, the thermodynamic efficiency is very low, while around criticality, the system is most efficient, as

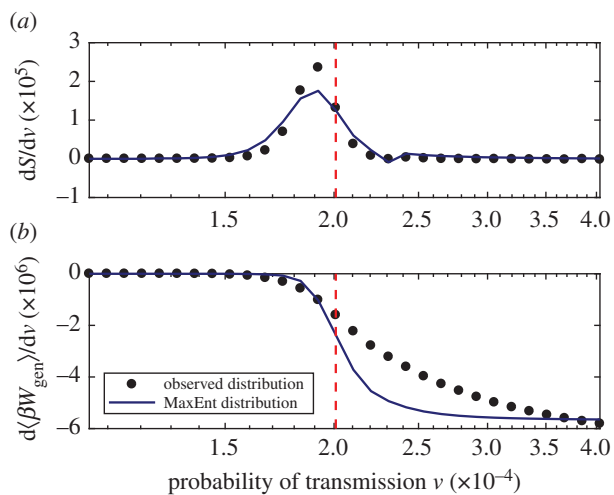


**Figure 11.** Thermodynamic efficiency of computation against the probability of infection  $\nu$ : in red for observed distributions, in blue for the MaxEnt distributions. The dotted line indicates the critical point at which the Fisher information peaks. (Online version in colour.)

shown in figure 11. The peak of the thermodynamic efficiency does not concur precisely with the peak of the Fisher information, due to finite-size effects, and the discrepancy is higher for the estimates based on the MaxEnt method, rather than those based on the observed distributions. The specific reasons for this discrepancy can be seen in figure 12, which contrasts both components in more detail, showing that the reduction in uncertainty (the numerator) estimated by the MaxEnt method starts to diverge earlier, while the rate of work (the denominator) estimated by the MaxEnt method starts to diverge later, than the corresponding estimates based on observed distributions. As a result, the ratio  $\eta_M$  ‘suffers’ from the finite-size effects more than  $\eta_O$ .

In summary, at criticality, where a higher disorder is generated, there is relatively more work extracted out of the system. In an epidemiological context, considering an intervention process that reduces the transmission probability from the super-critical phase towards the subcritical phase, expending the work, the thermodynamic efficiency would





**Figure 12.** Rates of change of the configuration entropy (a) and generalized work (b). The dotted line indicates the critical value of  $\nu$ . (Online version in colour.)

tend to increase as the critical point is passed. On the other hand, we can consider the efficiency of the contagion itself, as a biological phenomenon, i.e. a pathogen emergence process that increases the transmission probability from the subcritical phase towards the super-critical phase. In this case, from the pathogen ‘perspective’, the thermodynamic efficiency would tend to peak on the approach towards the critical point.

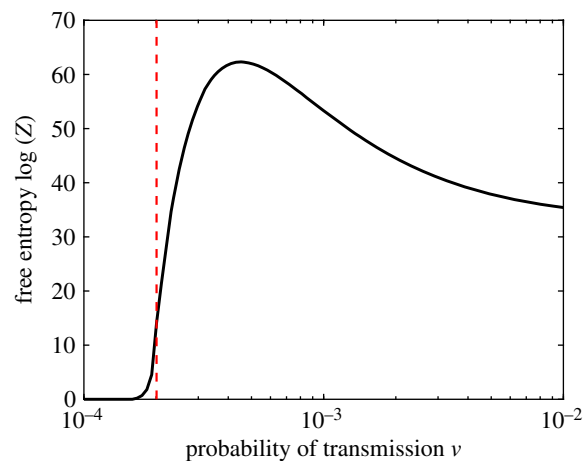
Finally, figure 13 shows the free entropy of the system,  $\log(Z)$ , which is proportional to the free energy. Again, we observe a clear critical regime separating the ‘subcritical’, non-epidemic, phase from the ‘super-critical’, epidemic, phase.

## 6. Discussion and future work

In this paper, we considered epidemics as thermodynamic phenomena, modelling SIS dynamics on a contact network statistical-mechanically. Applying the MaxEnt principle, under certain assumptions, allowed us to derive closed-form solutions for specific cases and interpret key epidemic variables in a statistical mechanical setting. Specifically, the reproductive ratio  $R_0$  of a SIS model was related to the inverse temperature of a Gibbs distribution resulting from entropy maximization, applied to the initial state of an outbreak.

Using the model, we evaluated the Fisher information, configuration and free entropy, as well as the thermodynamic efficiency of contagion (considered as a computational process), in an epidemiological context. This allowed us to identify critical regimes and distinct phases of epidemic processes. Analytical derivations of MaxEnt modelling of SIS process were contrasted with results of the simulated dynamics on Watts–Strogatz random graphs, confirming the critical thresholds, while assessing the impact of our simplifying assumptions. Importantly, the analytical derivations highlighted the mechanism for the emergence of critical regime—via non-differentiability of key variables  $\langle I \rangle$  and  $\langle C \rangle$ —manifesting itself in the divergence of the Fisher information.

The statistical-mechanical perspective taken in this work placed epidemic processes within a broad class of distributed processes. This allowed us to define thermodynamic efficiency of contagions which was shown to peak at criticality. Interestingly, this can be interpreted as (i) the efficiency of



**Figure 13.** The free entropy of the MaxEnt distributions. The dotted line indicates the critical value of  $\nu$ . (Online version in colour.)

an intervention process that expends the work needed to reduce the transmission probability or (ii) the efficiency of the pathogen emergence that extracts the work by increasing the transmission probability. These processes explore the parameter space in opposing directions at the expense/extraction of (thermodynamic) work, but despite the opposite directions the maximal thermodynamic efficiency is attained at the same critical point. Furthermore, the concept of thermodynamic efficiency enables comparative analysis of various interventions and pathogen emergence paths.

We presented both numerical and analytical techniques based on the MaxEnt method. The numerical technique does not need the independence assumption between  $I$  and  $C$ , and attains a high degree of accuracy. However, it requires an explicit combinatorial calculation of the macrostates which is prohibitive for large networks without further simplifying assumptions. The analytical technique allowed us to derive important general thermodynamic properties of the contagion phenomena, but relies on the independence assumption and is less accurate in the super-critical phase. We believe that both presented techniques provide useful arguments for the utility of the MaxEnt method and serve their distinct purposes.

We would like to reiterate that in practical scenarios, when simulation results are not available and the exact topology of the underlying interaction network is not known, one may still rely on the MaxEnt solutions derived under the constraints on  $\langle I \rangle$  and  $\langle C \rangle$  which are obtainable from the real-world observations of the epidemic dynamics. If, on the other hand, the underlying interaction network belongs to a specific type, e.g. Watts–Strogatz random graph, one may go one step further and, under the simplifying independence assumption, obtain analytic solutions and estimate their parameters.

Determining the precise range of applicability of the independence assumption, and the extent of the resultant approximations, remains the subject of future work. In general, this assumption is reasonable in networks with a significant number of random connections, and so would be applicable to many real-world networks in which epidemics take place, e.g. epidemics in urbanized societies [1,52].

We believe that the presented approach can be extended to other contact network topologies and contact processes. In addition, the constraints used during the entropy

maximization can be effectively generated by large-scale simulations of epidemics based on demographic datasets and real-world epidemic dynamics.

**Data accessibility.** Code used to generate the data used in this paper can be found at <https://github.com/NathanHarding/Therm-eff-cont>.

**Author's contributions.** All authors contributed to conception, analysis, interpretation and drafting. N.H. performed computational analysis. R.N. carried out the work described in §4.3. All authors gave final approval for publication.

**Competing interests.** We declare we have no competing interests.

**Funding.** M.P. and N.H. were supported through Australian Research Council grant DP160102742. Additionally, N.H. was supported by an Australian Government Research Training Program (RTP) Scholarship.

**Acknowledgements.** This research was supported by the Sydney Informatics Hub, through the use of HPC services funded by the University of Sydney. The authors thank Emanuele Crosato for numerous helpful discussions on the intricacies of thermodynamic efficiency.

## Reference

- Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. 2015 Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979. (doi:10.1103/RevModPhys.87.925)
- Newman ME, Watts DJ. 1999 Scaling and percolation in the small-world network model. *Phys. Rev. E* **60**, 7332–7342. (doi:10.1103/PhysRevE.60.7332)
- Moore C, Newman ME. 2000 Epidemics and percolation in small-world networks. *Phys. Rev. E* **61**, 5678–5682. (doi:10.1103/PhysRevE.61.5678)
- Newman ME. 2002 Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128. (doi:10.1103/PhysRevE.66.016128)
- Sander L, Warren C, Sokolov I, Simon C, Koopman J. 2002 Percolation on heterogeneous networks as a model for epidemics. *Math. Biosci.* **180**, 293–305. (doi:10.1016/S0025-5564(02)00117-7)
- Meyers L. 2007 Contact network epidemiology: bond percolation applied to infectious disease prediction and control. *Bull. Am. Math. Soc.* **44**, 63–87. (doi:10.1090/S0273-0979-06-01148-7)
- Hartfield M, Alizon S. 2013 Introducing the outbreak threshold in epidemiology. *PLoS Pathog.* **9**, e1003277. (doi:10.1371/journal.ppat.1003277)
- Pastor-Satorras R, Vespignani A. 2001 Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203. (doi:10.1103/PhysRevLett.86.3200)
- Wang W, Liu Q-H, Zhong L-F, Tang M, Gao H, Stanley HE. 2016 Predicting the epidemic threshold of the susceptible-infected-recovered model. *Sci. Rep.* **6**, 24676. (doi:10.1038/srep24676)
- Erten EY, Lizier JT, Piraveenan M, Prokopenko M. 2017 Criticality and information dynamics in epidemiological models. *Entropy* **19**, 194. (doi:10.3390/e19050194)
- Lizier JT, Prokopenko M, Zomaya AY. 2012 Coherent information structure in complex computation. *Theory Biosci.* **131**, 193–203. (doi:10.1007/s12064-011-0145-9)
- Lizier JT, Prokopenko M, Zomaya AY. 2014 A framework for the local information dynamics of distributed computation in complex systems. In *Guided self-organization: inception* (ed. M Prokopenko), pp. 115–158. Berlin, Germany: Springer. (doi:10.1007/978-3-642-53734-9\_5)
- Antia R, Regoes RR, Koella JC, Bergstrom CT. 2003 The role of evolution in the emergence of infectious diseases. *Nature* **426**, 658–661. (doi:10.1038/nature02104)
- O'Regan SM, Drake JM. 2013 Theory of early warning signals of disease emergence and leading indicators of elimination. *Theor. Ecol.* **6**, 333–357. (doi:10.1007/s12080-013-0185-5)
- Leventhal GE, Hill AL, Nowak MA, Bonhoeffer S. 2015 Evolution and emergence of infectious diseases in theoretical and real-world networks. *Nat. Commun.* **6**, 6101. (doi:10.1038/ncomms7101)
- Wang XR, Lizier JT, Prokopenko M. 2011 Fisher information at the edge of chaos in random boolean networks. *Artif. Life* **17**, 315–329. (doi:10.1162/artl\_a\_00041)
- Prokopenko M, Lizier JT, Obst O, Wang XR. 2011 Relating Fisher information to order parameters. *Phys. Rev. E* **84**, 041116. (doi:10.1103/PhysRevE.84.041116)
- Crosato E, Spinney RE, Nigmatullin R, Lizier JT, Prokopenko M. 2018 Thermodynamics and computation during collective motion near criticality. *Phys. Rev. E* **97**, 012120. (doi:10.1103/PhysRevE.97.012120)
- Kalloniatis AC, Zuparic ML, Prokopenko M. 2018 Fisher information and criticality in the Kuramoto model of nonidentical oscillators. *Phys. Rev. E* **98**, 022302. (doi:10.1103/PhysRevE.98.022302)
- Tkačik G, Marre O, Mora T, Amodei D, Berry II MJ, Bialek W. 2013 The simplest maximum entropy model for collective behavior in a neural network. *J. Stat. Mech. Theory Exp.* **2013**, P03011. (doi:10.1088/1742-5468/2013/03/p03011)
- Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M, Walczak AM. 2012 Statistical mechanics for natural flocks of birds. *Proc. Natl Acad. Sci. USA* **109**, 4786–4791. (doi:10.1073/pnas.1118633109)
- Artalejo JR, Lopez-Herrero M. 2011 The SIS and SIR stochastic epidemic models: a maximum entropy approach. *Theor. Popul. Biol.* **80**, 256–264. (doi:10.1016/j.tpb.2011.09.005)
- Watts DJ, Strogatz SH. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442. (doi:10.1038/30918)
- Prokopenko M, Lizier JT. 2014 Transfer entropy and transient limits of computation. *Sci. Rep.* **4**, srep05394. (doi:10.1038/srep05394)
- Spinney RE, Lizier JT, Prokopenko M. 2016 Transfer entropy in physical systems and the arrow of time. *Phys. Rev. E* **94**, 022135. (doi:10.1103/PhysRevE.94.022135)
- Spinney RE, Prokopenko M, Lizier JT. 2017 Transfer entropy in continuous time, with applications to jump and neural spiking processes. *Phys. Rev. E* **95**, 032319. (doi:10.1103/PhysRevE.95.032319)
- Kempes CP, Wolpert D, Cohen Z, Pérez-Mercader J. 2017 The thermodynamic efficiency of computations made in cells across the range of life. *Phil. Trans. R. Soc. A* **375**, 20160343. (doi:10.1098/rsta.2016.0343)
- Spinney RE, Lizier JT, Prokopenko M. In press. Entropy balance and information processing in bipartite and nonbipartite composite systems. *Phys. Rev. E*.
- Keeling M, Rohani P. 2008 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press. See <https://books.google.com.au/books?id=G8enmS23c6YC>.
- Pastor-Satorras R, Vespignani A. 2001 Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E* **63**, 066117. (doi:10.1103/PhysRevE.63.066117)
- Kermack WO, McKendrick AG. 1927 A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* **115**, 700–721. (doi:10.1098/rspa.1927.0118)
- Porta M. 2014 *A dictionary of epidemiology*. Oxford, UK: Oxford University Press.
- Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. 2005 Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359. (doi:10.1038/nature04153)
- Keeling M. 2005 The implications of network structure for epidemic dynamics. *Theor. Popul. Biol.* **67**, 1–8. (doi:10.1016/j.tpb.2004.08.002)
- Andersson H. 1998 Limit theorems for a random graph epidemic model. *Ann. Appl. Probab.* **8**, 1331–1349. (doi:10.1214/aop/1028903384)
- Keeling MJ, Eames KT. 2005 Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307. (doi:10.1098/rsif.2005.0051)
- Rhodes C, Anderson R. 1996 Dynamics in a lattice epidemic model. *Mod. Phys. Lett. A* **210**, 183–188. (doi:10.1016/s0375-9601(96)80007-7)

38. Erdős P, Rényi A. 1959 On random graphs I. *Publicationes Mathematicae (Debrecen)* **6**, 290–297.
39. Barrat A, Weigt M. 2000 On the properties of small-world network models. *Eur. Phys. J. B* **13**, 547–560. (doi:10.1007/s100510050067)
40. Niven R. 2010 Minimization of a free-energy-like potential for non-equilibrium flow systems at steady state. *Phil. Trans. R. Soc. B* **365**, 1323–1331. (doi:10.1098/rstb.2009.0296)
41. Jaynes ET. 1957 Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630. (doi:10.1103/PhysRev.106.620)
42. Shannon CE. 2001 A mathematical theory of communication. *ACM SIGMOBILE Mobile Comput. Commun. Rev.* **5**, 3–55. (doi:10.1145/584091.584093)
43. Zamir R. 1998 A proof of the Fisher information inequality via a data processing argument. *IEEE Trans. Inf. Theory* **44**, 1246–1250. (doi:10.1109/18.669301)
44. Lehmann E, Casella G. 2003 *Theory of point estimation*. Springer Texts in Statistics. New York, NY: Springer.
45. Brody D, Rivier N. 1995 Geometrical aspects of statistical mechanics. *Phys. Rev. E* **51**, 1006–1011. (doi:10.1103/PhysRevE.51.1006)
46. Crooks G. 2007 Measuring thermodynamic length. *Phys. Rev. Lett.* **99**, 100602. (doi:10.1103/PhysRevLett.99.100602)
47. Brody D, Ritz A. 2003 Information geometry of finite Ising models. *J. Geom. Phys.* **47**, 207–220. (doi:10.1016/S0393-0440(02)00190-0)
48. Janke W, Johnston D, Kenna R. 2004 Information geometry and phase transitions. *Phys. A: Stat. Mech. Appl.* **336**, 181–186. (doi:10.1016/j.physa.2004.01.023)
49. Germann TC, Kadau K, Longini IM, Macken CA. 2006 Mitigation strategies for pandemic influenza in the United States. *Proc. Natl Acad. Sci. USA* **103**, 5935–5940. (doi:10.1073/pnas.0601266103)
50. Cliff OM, Harding N, Piraveenan M, Erten EY, Gambhir M, Prokopenko M. 2018 Investigating spatiotemporal dynamics and synchrony of influenza epidemics in Australia: an agent-based modelling approach. *Simul. Model. Pract. Theory* **87**, 412–431. (doi:10.1016/j.simpat.2018.07.005)
51. Zachreson C, Fair KM, Cliff OM, Harding N, Piraveenan M, Prokopenko M. 2018 Vulnerability to pandemics in a rapidly urbanizing society. (<http://arxiv.org/abs/1806.04001>)
52. Neiderud C-J. 2015 How urbanization affects the epidemiology of emerging infectious diseases. *Infect. Ecol. Epidemiol.* **5**, 27060. (doi:10.3402/iee.v5.27060)