


OPEN

# Hamiltonian dynamics of the SIS epidemic model with stochastic fluctuations

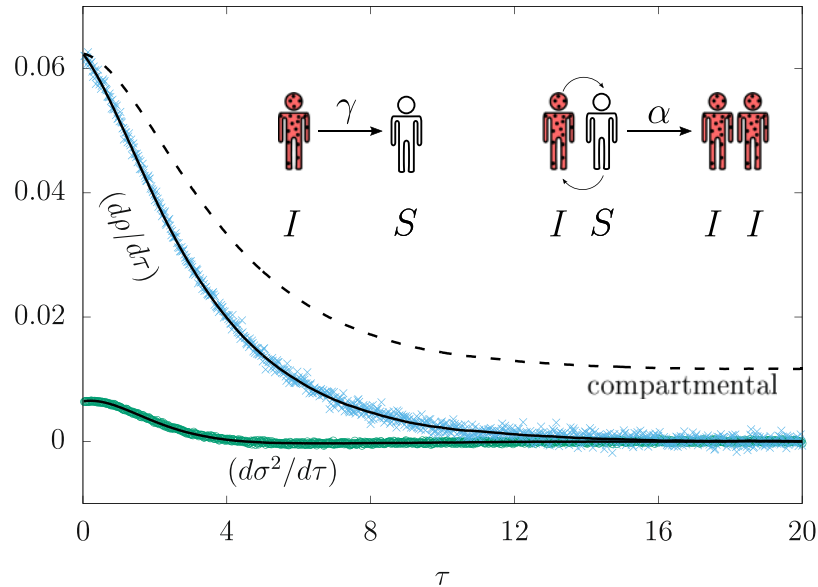
Gilberto M. Nakamura<sup>1,2,3\*</sup> & Alexandre S. Martinez<sup>1,2</sup> 

Empirical records of epidemics reveal that fluctuations are important factors for the spread and prevalence of infectious diseases. The exact manner in which fluctuations affect spreading dynamics remains poorly known. Recent analytical and numerical studies have demonstrated that improved differential equations for mean and variance of infected individuals reproduce certain regimes of the SIS epidemic model. Here, we show they form a dynamical system that follows Hamilton's equations, which allow us to understand the role of fluctuations and their effects on epidemics. Our findings show the Hamiltonian is a constant of motion for large population sizes. For small populations, finite size effects break the temporal symmetry and induce a power-law decay of the Hamiltonian near the outbreak onset, with a parameter-free exponent. Away from the onset, the Hamiltonian decays exponentially according to a constant relaxation time, which we propose as a metric when fluctuations cannot be neglected.

Models of disease transmission, or epidemic models for short, have been an integral part of the epidemiological toolkit, dating back from pioneer models of Kermack and McKendrick<sup>1</sup>. The main goal of epidemic models can be summarized as the ability to accurately predict spreading patterns of a given communicable disease afflicting a specific population. These models allow decision makers to assess the various intervention strategies available to them and to plan accordingly. Several approaches have been developed to model disease outbreaks<sup>2</sup>, namely, compartmental equations, stochastic equations, agent-based simulations, etc. Each approach suits a particular aspect of the outbreak being studied, built upon hypotheses compatible with empirical records or based on a phenomenological context. They include, but are not restricted to, biological content of the disease, mechanisms behind pathogen transmission, social interactions among the target population and its spatial structure<sup>3</sup>. By the same token, different models for the same disease and population may produce inconsistent results, possibly due to conflicting underlying hypotheses. For instance, the random-mixing hypothesis—i.e. the population is assumed to be homogeneous and its elements mix at random—seems reasonable to model pathogen transmission for airborne disease like influenza, but it seems equivocated for sexually transmitted diseases<sup>4</sup>. Recent studies using household data shows that the random-mixing hypothesis can produce reliable predictions for households, even for heterogeneous contact networks<sup>5</sup>. In general, available empiric data and numerical simulations provide evidence that the disease spreading is largely affected by the heterogeneity of contact network of the population<sup>6–12</sup>.

Despite the significant advances obtained in the past few decades, several challenges remain open. One issue concerns the failure to account for effects unrelated to diseases themselves, such as vaccination skepticism, which ultimately reduces children immunization rate. Outbreaks of treatable communicable diseases, like measles, are on the rise<sup>13</sup>. Another issue deals with understanding the complex dynamics and processes behind infections in both small and large scales<sup>14–16</sup>. To put it simply, there are too many variables and their effects are not entirely known due to the non-linear nature of the problem. As a consequence, the full extent of variable changes or their fluctuations remains poorly understood, which may produce sub-optimal intervention strategies. As an example, detailed field data from the recent Ebola epidemic have shown that smaller outbreaks from different localities are asynchronous<sup>17</sup>. The lack of synchronization between different populations reduces the likelihood of pathogen eradication on a global scale, as long as migration is allowed in some form<sup>18,19</sup>. The effects of migration and

<sup>1</sup>Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP), Universidade de São Paulo, Avenida Bandeirantes 3900, 14040-901, Ribeirão Preto, Brazil. <sup>2</sup>Instituto Nacional de Ciência e Tecnologia - Sistemas Complexos (INCT-SC), 22460-320, Rio de Janeiro, Brazil. <sup>3</sup>Present address: Laboratoire d'Imagerie et Modélisation en Neurobiologie et Cancérologie (IMNC), Centre National de la Recherche Scientifique (CNRS), UMR 8165, Universités Paris 11 and Paris 7, Campus d'Orsay, 91405, Orsay, France. \*email: [gmnakamura@usp.br](mailto:gmnakamura@usp.br)



**Figure 1.** Numerical simulations of the SIS model. (inset) Infected hosts (I) recover to susceptible state (S) with rate  $\gamma$  (left). The adequate interaction between an infected host with a susceptible one may trigger a new infection, with rate  $\alpha$  (right). Stochastic effects are far more relevant for small population sizes ( $N = 50$ ,  $\gamma/\alpha = 1/2$ ), reducing the accuracy of compartmental equations. The forward-derivative  $d\rho/d\tau$  from data (cross) agrees with Eq. (5a) (solid line), while the compartmental equation Eq. (1) fails to reproduce the data (dashed line). The forward-derivative  $d\sigma^2/d\tau$  from data (circles) also agrees with the formula in Eq. (5b) (line). All the lines are drawn using the simulated data for  $\langle\rho(\tau)\rangle$ ,  $\sigma^2(\tau)$ , and  $\Delta_3(\tau)$ .

spatial structures in epidemic models and pathogen variability have been under investigation for some time<sup>20,21</sup>, and they have been linked to chaotic dynamics in local population<sup>22</sup>. Experiments on the effects of migration between metapopulations, i.e. similar populations but spatially separated, subjected to temporal fluctuations have shown that pathogen prevalence is greatly influenced by the nature of the fluctuation<sup>23</sup>, highlighting the interplay between synchronization and pathogen prevalence in epidemics<sup>24</sup>.

Traditionally, the detailed examination of fluctuations—either temporal or spatial—and their effects on system dynamics have been largely described by correlation functions<sup>23,25</sup>. More recently, autocorrelation functions have been used to reveal the nature and general aspects of fluctuations in a simple agent-based epidemic model for a population of size  $N$ , in which temporal fluctuations are divided into two broad classes: gaussian and non-gaussian<sup>26</sup>. In the gaussian regime, the prevalence of the disease is well described by its instantaneous average, finite variance, and higher cumulants can be neglected. This is remarkable as it allows one to derive the exact contributions of fluctuations to disease outbreaks in the asymptotic limit  $N \gg 1$ . Here, we show that the dynamical equations form a Hamiltonian dynamical system, and the way external noise can be incorporated to model disease outbreaks. This approach allows us to discuss quantitatively the relevant scales of the problem, and interpret the resulting Lagrangian and canonical transformations.

## Model

We begin our discussion using the susceptible-infectious-susceptible (SIS) epidemic model. The SIS model describes the dissemination of a single communicable disease in a susceptible population of size  $N$ . The transmission of the pathogen occurs when infectious hosts transmit the disease pathogen to healthy susceptible individuals. The infectious period extends throughout the whole course of the disease until the recovery of the patient, warranting a two-stage model: either infected or susceptible. The essence of the model is summarized by inset in Fig. 1.

The traditional formulation of the problem assumes the random-mixing hypothesis (see Introduction) holds for a large population size  $N \gg 1$ , comprised of statistically equivalent individuals. Under these circumstances, the only relevant variable is the instantaneous density of infected elements  $\rho(t)$ , which means that fluctuations can be safely neglected. Furthermore,  $\rho(t)$  decreases with rate  $\gamma\rho$ , where  $\gamma$  is the recovery rate. New infections per unit of time (disease incidence) are proportional to  $\alpha\rho(1 - \rho)$ , i.e., they depend on the chance that infected elements interact with susceptible ones, with intensity given by the transmission rate  $\alpha$ . This picture provides an interpretation where  $\rho(t)$  is continuously exchanged between two compartments, leading a simple description called compartmental equation:  $d\rho(t)/dt = \alpha\rho(1 - \rho) - \gamma\rho$ . For the sake of convenience, redefine the timescale as  $\tau \equiv \alpha t$  and  $\rho_0 \equiv 1 - \gamma/\alpha$ , so that

$$\frac{d}{d\tau}\rho(\tau) = \rho(\rho_0 - \rho). \quad (1)$$

Clearly, the equilibrium density can either be  $\rho_{eq} = 0$  or  $\rho_{eq} = \rho_0$ . Also,  $\rho_0$  is related to the basic reproduction number  $R_0 = N(\alpha/\gamma)$  which provides an estimate on the number of new infections per generation<sup>27</sup>.

In light of its long age, compartmental equations have met considerable success in predicting the time evolution of disease outbreaks, providing valuable insights for intervention strategies and funding allocation<sup>28</sup>. However, outbreaks that fail to meet the underlying hypotheses (random mixing and large population of statistically equivalent elements) can contradict compartmental equations. These inconsistencies are largely attributed to stochastic effects and their inherent fluctuations<sup>2</sup>.

## Improved Compartmental Equations

Stochastic variables are known to cause the emergence of critical phenomena in computer simulations of epidemic models, under certain parameter ranges<sup>29,30</sup>. One key ingredient common to critical phenomena is the scale invariance of fluctuations. This special symmetry remains the foundation of cooperative phenomena and critical phase transitions, whose contributions span over a broad set of research fields such as condensed-matter, quantum field theories, and neuroscience to name a few<sup>31–36</sup>. In these systems, fluctuations occur in all sizes and dictate the general behavior of the problem, which prompts for in-depth studies of their effects in epidemic models.

Stochastic epidemic models include non-deterministic events that intrinsically occurs during the course of the disease spreading process. Examples of these events include the transmission of the pathogen, and the elapsed time required for the complete recovery of patients. The inclusion of these effects brings the models closer to more realistic expectations, which in turn can deviate from predictions using compartmental equations<sup>37–39</sup>. One prime example is the effects of the absorbing state in the SIS model, which has been examined in detail by Näsell and others<sup>40–42</sup>. New experiments on this subject provide evidence that temporal fluctuations can drastically alter the prevalence of pathogens<sup>23</sup>. Spatial heterogeneity also introduces an extra layer of complexity as it may trap or delay the pathogen transmission<sup>20</sup>. As a result, the requirements of statistical equivalence may not hold for all scales. To deal with this issue, stochastic formulations and numerical simulations have been the default tools to investigate fluctuations in disease outbreaks. In what follows, we outline the general ideas behind the agent-based approach to describe the SIS model. A more detailed account of the results listed here are explained in detail in refs<sup>26,43,44</sup>.

Our discussion assumes the disease spreading follows a Markov chain in discrete time  $\delta t$ . Moreover,  $\delta t$  is such that at most a single recovery or transmission event is likely to occur during the course of its duration. Under these requirements, the master equation of the SIS model in discrete time reads

$$\frac{dP_\mu(t)}{dt} = - \sum_{\nu=0}^{2^N-1} H_{\mu\nu} P_\nu(t). \quad (2)$$

Here,  $P_\mu(t)$  refers to the instantaneous probability to observe the system in the  $\mu$ -th configuration. Configuration labels follow the binary ruling  $\mu = n_0 2^0 + n_1 2^1 + \dots + n_{N-1} 2^{N-1}$ , where  $n_k = 1$  if the  $k$ -th agent is infected, or  $n_k = 0$  otherwise, with  $k = 0, 1, \dots, N-1$ . For instance, for  $N = 3$ , the configuration  $|\mu = 3\rangle = |110\rangle$  states that only the agent with label  $k = 2$  is susceptible. The matrix elements  $H_{\mu\nu}$  express the transition rates from configuration  $\nu$  to configuration  $\mu$ . By virtue of probability conservation, in each time step the transition rules satisfy  $\sum_\mu H_{\mu\nu} = 0$ . The matrix elements  $H_{\mu\nu} = \langle \mu | \hat{H} | \nu \rangle$  are computed from projections on the time step operator

$$\hat{H} = \frac{\alpha}{N} \sum_{k,\ell=0}^{N-1} A_{k\ell} (1 - \hat{n}_k - \hat{\sigma}_k^+) \hat{n}_\ell + \gamma \sum_{k=0}^{N-1} (\hat{n}_k - \hat{\sigma}_k^-), \quad (3)$$

where  $A_{k\ell}$  is the adjacency matrix,  $\hat{n}_k$  represents the  $k$ -th occupation operator (with eigenvalues  $n_k = 1$  if infected, 0 otherwise), and  $\hat{\sigma}_k^\pm$  are the localized spin-1/2 ladder operators that produce the transition  $S \rightarrow I$ . Clearly,  $\hat{\sigma}_k^-$  produce the opposite transitions,  $I \rightarrow S$  in relation the  $k$ -th agent. As notation, the hat symbol always accompanies operators to quickly distinguish them from numbers.

The master equation Eq. (2) provides the means to evaluate the time evolution of relevant statistical moments of  $\rho(t)$ . Notice that the average density of infected agents in the system reads

$$\langle \rho(t) \rangle = \frac{1}{N} \sum_{\mu=0}^{2^N-1} \sum_{k=0}^{N-1} \langle \mu | \hat{n}_k | \mu \rangle P_\mu(t). \quad (4)$$

Applying the time derivative, and using Eq. (2), one arrives at the equation of motion for  $\langle \rho(t) \rangle$ . Useful expressions are known only for a few types of adjacency matrix  $A$ . The simplest one is the complete graph  $A_{k\ell} = 1 - \delta_{k\ell}$ , which recovers the random mixing hypothesis. In those particular instances, the complete time evolution of the system comprehends a set of hierarchical equations that involves the statistical moments of  $\rho(t)$ , as shown in refs. <sup>43,45</sup>. More explicitly<sup>45</sup>, the first two equations for instantaneous mean  $\langle \rho \rangle$  and variance  $\sigma^2 = \langle \rho^2 \rangle - \langle \rho \rangle^2$  are

$$\frac{d\langle \rho \rangle}{d\tau} = \langle \rho \rangle [\rho_0 - \langle \rho \rangle] - \sigma^2(\tau), \quad (5a)$$

$$\frac{d\sigma^2}{d\tau} = 2\sigma^2[\rho_0 + \langle \rho \rangle] - 2\Delta_3(\tau) + \frac{1}{N} \langle \rho(1 - \rho) \rangle + \frac{\gamma}{N\alpha} \langle \rho \rangle, \quad (5b)$$

where  $\Delta_3(\tau) = \langle \rho^3(\tau) \rangle - \langle \rho(\tau) \rangle^3$ . These results find excellent agreement with simulated data using an ensemble with  $10^6$  replicas starting from the same initial condition (see Fig. 1).

Comparing Eqs (1) and (5), the case that considers temporal fluctuations decays faster than the compartmental equation by  $\sigma^2(\tau)$ , even in the regime  $N \gg 1$ . Both equations are equivalent whenever  $\sigma(\tau)$  becomes irrelevant compared to  $\langle \rho \rangle$ . Therefore, a generalization of compartmental equations for the SIS model is readily available by retaining both mean and variance, neglecting higher statistical moments. Thus, the dynamical system describes a gaussian variable evolving along time. The skewness coefficient vanishes as a direct consequence of this assumption, so that  $\Delta_3(\tau) \approx 3\langle \rho(\tau) \rangle \sigma^2(\tau)$ . For  $N \gg 1$ , the resulting equations are

$$\frac{d}{d\tau} \ln \langle \rho \rangle = \rho_0 - \langle \rho \rangle - \frac{\sigma^2}{\langle \rho \rangle}, \tag{6a}$$

$$\frac{1}{2} \frac{d}{d\tau} \ln \sigma^2 = \rho_0 - 2\langle \rho \rangle. \tag{6b}$$

We emphasize that the variance in Eq. (6a) slows down the growth rate of  $\langle \rho(\tau) \rangle$ , recalling the Allee effect<sup>28,46</sup>.

The dynamical system represented by Eq. (6) can also be obtained from a stochastic expansion following the guidelines in ref.<sup>47</sup>. In this case, assume the density  $\rho = \langle \rho \rangle + \eta$  is well described by its instantaneous average plus some noise function  $\eta(\tau)$ . We assume  $\langle \eta \rangle = 0$  and  $\langle \eta^2 \rangle = \sigma^2(\tau)$  for the sake of consistency. The additional requirement  $|\eta| \ll \langle \rho \rangle$  ensures stochastic effects act as perturbations. Expanding Eq. (1) as a Taylor series around  $\rho = \langle \rho \rangle$ , and then taking the ensemble average, produces

$$\frac{d\langle \rho \rangle}{d\tau} = \left\langle g(\langle \rho \rangle) + \eta \frac{dg}{d\rho} \Big|_{\langle \rho \rangle} + \frac{\eta^2}{2} \frac{d^2g}{d\rho^2} \Big|_{\langle \rho \rangle} + o(\eta^3) \right\rangle, \tag{7}$$

where  $g(\rho) = \rho(\rho_0 - \rho)$ . Since  $\langle \eta \rangle$  vanishes, the expansion recovers Eq. (6a) up to  $o(\eta^4)$ . The task to obtain the equation of motion of  $\sigma^2$  is greatly simplified by noting that

$$\eta \frac{d\eta}{dt} = \eta \left[ \frac{d\rho}{dt} - \frac{d\langle \rho \rangle}{dt} \right] = \eta^2 [\rho_0 - 2\langle \rho \rangle] - \eta^3 + \eta \langle \eta^2 \rangle. \tag{8}$$

Taking the ensemble average, one obtains  $(1/2)(d/dt) \ln \sigma^2 = [\rho_0 - 2\langle \rho \rangle] - \langle \eta^3 \rangle / \sigma^2$ . Now, since  $\eta$  mimics a gaussian variable,  $\langle \eta^3 \rangle = 0$  and Eq. (6b) is recovered as well.

Equation (6) can be further combined into a single second-order differential equation<sup>26</sup>, with solution

$$\langle \rho(\tau) \rangle = \frac{\rho_0(1 + c_1 e^{-\rho_0 \tau})}{1 + 2c_1 e^{-\rho_0 \tau} + c_2 e^{-2\rho_0 \tau}}, \tag{9a}$$

$$\sigma^2(\tau) = \frac{\langle \rho(\tau) \rangle^2 (c_1^2 - c_2) e^{-2\rho_0 \tau}}{(1 + c_1 e^{-\rho_0 \tau})^2}. \tag{9b}$$

The constants  $c_1$  and  $c_2$  depend solely on the initial conditions. The special case  $c_2 = c_1^2$  recovers the usual solution of Eq. (1). We assumed that fluctuations behave as gaussian fluctuations. While reasonable for various situations, the assumption does not hold for  $\gamma/\alpha$  around unity or small population sizes, according to numerical simulations<sup>26</sup>, in which Eq. (5b) should be used instead of Eq. (6b). For the sake of completeness, there is another solution in which  $\sigma^2(\tau) = \langle \rho(\tau) \rangle^2$ , with  $\langle \rho(\tau) \rangle = \rho_0 / (2 + c_1 e^{-\rho_0 \tau})$ .

### Hamilton Equations

The fact that the dynamical system Eq. (6) can be combined into a single second-order differential equation suggests an interpretation of the epidemic model in terms of Hamilton equations<sup>48</sup>. Hamiltonian systems are ubiquitous in Physics, serving as basis to describe and explain countless physical phenomena. The hallmark of systems are the Hamilton equations:

$$\frac{dq}{d\tau} = \frac{\partial \mathcal{H}}{\partial p}, \tag{10a}$$

$$\frac{dp}{d\tau} = - \frac{\partial \mathcal{H}}{\partial q}, \tag{10b}$$

where  $q(t)$  and  $p(t)$  are conjugated variables, and the Hamiltonian function  $\mathcal{H}$  encodes some information about the problem—usually associated with energy for conservative systems but not restricted to them. Besides classical mechanics and related areas, quantum field theories and statistical mechanics are deeply intertwined with Hamilton's principle and Liouville theorem. Despite its usefulness in Physics, Hamilton formulation and surrounding principles are rarely used in population dynamics, ecological problems, or epidemic models, where first-order differential equations are dominant. The lack of second-order differential equations in these areas, although not prohibitive, raises questions about the description of the dynamics, as discussed extensively in ref.<sup>49</sup>. In part,

because it means some interactions and stochastic effects acting on the system remain unaccounted. Significant advances in the Hamiltonian formulation of stochastic epidemic models have been obtained using the eikonal approximation, with emphasis on the disease extinction and vaccination<sup>50,51</sup>. Generalizations for heterogeneous networks have also been investigated, providing improved control strategies<sup>52,53</sup>. Even more, the aforementioned Hamiltonian formulation explains other effects, such as the emergence of noise-induced metastable states<sup>54</sup>. The crucial ingredient of these studies concerns the eikonal approximation, in which either the generating function or the density of infected is written as the exponential of the classical action of the system<sup>55</sup>. For populations with fixed size, one variable describes  $\langle \rho(t) \rangle$ , while its conjugate variable encompasses fluctuations<sup>52</sup>. However, the description of epidemics using this conjugate pair remains a complex task because the conjugate variable is not directly related to familiar statistics such as the standard deviation or variance. In what follows, instead of using the eikonal approach, we argue that Hamilton dynamics for the stochastic SIS model can be constructed from  $\langle \rho(\tau) \rangle$  and  $\sigma(\tau)$ .

In view of the inherent stochasticity behind disease spreading, it seems necessary to determine whether Eq. (6) form a Hamiltonian system or not. A brief inspection shows the pair  $(\langle \rho \rangle, \sigma^2)$  does not satisfy the usual Hamilton equations. The solution to this issue is obtained by assuming, instead, that the correct conjugated pair is  $(\langle \rho \rangle, h(\sigma^2))$ , where  $h(x)$  is some analytical function. Inspiration from common pairs of conjugate variables can be used to refine the choice of  $h(x)$ . For instance, the product  $\langle \rho \rangle \times h(\sigma^2)$  should be dimensionless, in close analogy the scalar product between position and wave vectors. One possible candidate is  $h(x) = x^{-1/2}$ , which entails  $1/\sigma$  as the conjugated variable to  $\langle \rho \rangle$ .

Define the dynamical variables  $q(\tau) = \langle \rho(\tau) \rangle$  and  $p(\tau) = 1/\sigma(\tau)$  to describe the SIS model. In addition, consider the following Hamiltonian

$$\mathcal{H} = q(\tau)p(\tau)[\rho_0 - q(\tau)] + \frac{1}{p(\tau)}. \quad (11)$$

Plugging these expressions in Eq. (10), one obtains the equations of motion:

$$\frac{dq}{d\tau} = q(\rho_0 - q) - \frac{1}{p^2} \equiv \langle \rho \rangle [\rho_0 - \langle \rho \rangle] - \sigma^2, \quad (12a)$$

$$\frac{dp}{d\tau} = -p(\rho_0 - 2q) \equiv -\frac{1}{\sigma} [\rho_0 - 2\langle \rho \rangle]. \quad (12b)$$

Thus, at first glance  $\mathcal{H}$  appears to be a valid candidate to describe the SIS model. Even more, replacing  $(q, p)$  by Eq. (9) in Eq. (11) shows that the Hamiltonian is a constant of motion  $\mathcal{H}^\infty = \rho_0 c_1 (c_1^2 - c_2)^{-1/2}$ . The upper index in  $\mathcal{H}^\infty$  is a reminder that calculations take place in the absence of finite size corrections.

However, taking finite size corrections into account changes drastically the notion of  $\mathcal{H}$  as a constant of motion. In fact, as Fig. 2 depicts, numerical simulations for finite populations reveal  $\mathcal{H}$  changes continuously along time until equilibrium sets in, akin to a non-conservative system. The precise meaning of  $\mathcal{H}$  in the epidemiological context is still murky, at best. A detailed analysis of correlations between changes in  $\mathcal{H}$  and the spreading pattern of real outbreaks is mandatory to understand the action-reaction analogy. In the meantime, it is instructive to study  $\mathcal{H}$  for  $\tau \ll 1$  and  $\tau \gg 1$  (see Fig. 2). For  $\tau \ll 1$ , where incidentally fluctuations vary the most (see Fig. 1), a remarkable feature appears via the relation  $\mathcal{H} \sim \tau^{-\lambda}$  with  $\lambda = 1/2$ . In particular, the exponent  $\lambda$  seems insensitive to changes in the epidemiological parameter  $\gamma$ . This parameter-free behavior is not observed for the remaining statistics,  $\langle \rho(\tau) \rangle$  and  $\sigma(\tau)$ . Power-laws are crucial to identify scaling relations and the emergence of universal features, and they are usually related to the symmetry of the problem rather than microscopic details. Here, evidence of universal behavior is captured by the data collapse  $\mathcal{H}/\rho_0^2$  (Fig. 2a inset). From these observations, we can infer fluctuations play a larger role in the early disease spreading, being largely independent of exact values of epidemiological parameters.

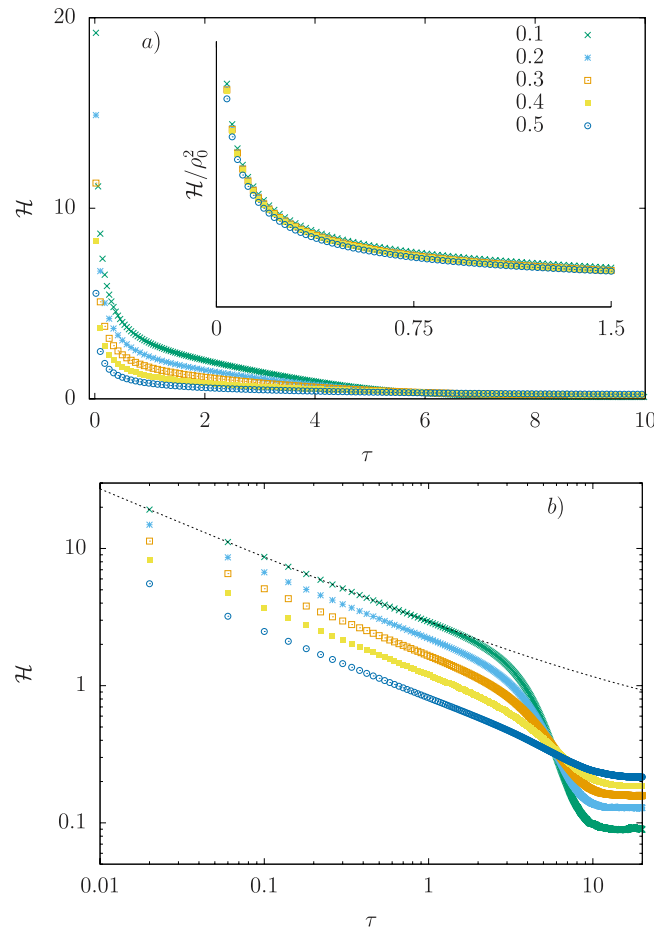
An effective decay  $e^{-\tau/\tau_{\text{eff}}}$  describes the general behavior of  $\mathcal{H}$  away from the outbreak onset. The relaxation time  $\tau_{\text{eff}}$  depends on  $N$  and the ratio  $\gamma/\alpha$ , and it can be estimated from data by fitting  $\mathcal{H}$  to an exponential function plus a constant. Alternatively, it can be evaluated as

$$\tau_{\text{eff}} = \frac{1}{\mathcal{H}(0)} \int_0^\infty d\tau [\mathcal{H}(\tau) - \mathcal{H}(\infty)]. \quad (13)$$

From a formal point of view, the evaluation of  $\tau_{\text{eff}}$  requires the solutions of Eqs (5a) and (5b) in Eq. (11), followed by an integration. Surely, the procedure is arguably more demanding than estimating  $R_0$ . However, as others have reasoned before,  $R_0$  provides a naive estimation on secondary infections because the growth rate of the outbreak changes continuously along time<sup>56</sup>. In contrast,  $\tau_{\text{eff}}$  mimics a constant of motion.

## Lagrangian and Canonical Transformations

Another insight from  $\tau_{\text{eff}}$  links the temporal integral of  $\mathcal{H}$  with the mechanical action  $S$ . A formal connection with  $S$  is desirable because it brings a large machinery revolving around variational principles and conservation laws. However, the action  $S = \int d\tau \mathcal{L}(q, \dot{q}; \tau)$  is a functional of the Lagrangian  $\mathcal{L}$ . It turns out that  $\mathcal{L}$  can be obtained from  $\mathcal{H}$  by inspection. From Eqs (15) and (7),  $\mathcal{H}$  takes the following form:  $\mathcal{H} = p[q(\rho_0 - q) + p^{-2}] = p(dq/d\tau) + 2/p$ . Recalling the formal expression  $\mathcal{H} = p\dot{q} - \mathcal{L}$ , it becomes clear that



**Figure 2.** Finite size effects on the Hamiltonian. **(a)** Simulated data with  $N = 50$  and  $10^6$  Monte Carlo runs for various ratios  $\gamma/\alpha$ . (inset) Data collapse using the scaling factor  $\rho_0^2$ , suggesting an universal behavior at the beginning of the outbreak. **(b)** Initial decay of  $\mathcal{H}$  compatible with power-law,  $\mathcal{H} \sim \tau^{-\lambda}$ . The exponent  $\lambda = 1/2$  remains constant for different ratios  $\gamma/\alpha$ , suggesting an universal behavior.

$$\mathcal{L} = -\frac{2}{p} = -2\sigma(\tau) = -2\sqrt{q(\rho_0 - q) - \frac{dq}{d\tau}}, \tag{14}$$

where we have used Eq. (6a) and considered only the positive root. Thus,  $\mathcal{L}$  is proportional to the standard deviation while the action entails the accumulated deviation over the course of the outbreak. To check our result for large populations  $N \gg 1$ , the minimal action recovers Eq. (1) as expected for a noise-free system. In general, the equation of motion reads

$$\frac{d^2q}{d\tau^2} = 3(\rho_0 - 2q)\left[\frac{dq}{d\tau} - \frac{2}{3}q(\rho_0 - q)\right]. \tag{15}$$

The fact that  $\mathcal{L}$  contains solely the standard deviation allows us to understand how to add uncorrelated fluctuations into the model. By virtue of  $\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y]$  for uncorrelated random variables  $x$  and  $y$ , the perturbed Lagrangian can be obtained by adding a  $\sigma_{ext}^2(\tau)$  to the variance of the system  $\sigma^2(\tau)$ :

$$\mathcal{L}' = -2\sqrt{q(\rho_0 - q) - \frac{dq}{d\tau} + \sigma_{ext}^2(\tau)}. \tag{16}$$

This picture is consistent with addition of a noise function  $\sigma_{ext}^2(\tau)$  to Eq. (6a). The perturbed Lagrangian  $\mathcal{L}'$  describes, ultimately, the time evolution of the disease prevalence in environments with noise. Note that this description differs from the usual derivation of Langevin equations, in which the noise function (force)  $r(\tau)$  couples linearly with  $q$ , i.e.,  $\mathcal{L}' = \mathcal{L} - r(\tau)q(\tau)$ . By the same token, the addition of correlated signals  $\eta(\tau)$  to the Lagrangian entails corrections from the covariance matrix: since  $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$ , then  $\mathcal{L}' = -2\sqrt{\sigma_\rho^2 + \sigma_\eta^2 + 2\text{Cov}[\rho, \eta]}$ . The covariance matrix can be estimated or modeled directly from data, pro-

moting further understanding on the spreading of co-existing diseases, where facilitation or competition processes are in place.

With both Hamiltonian and Lagrangian formalisms secured, canonical transformations become available. These transformations are particularly useful to highlight properties of the dynamical systems and to solve them. They change the old variables  $(q, p)$  into new variables  $(Q, P)$ , while preserving Hamilton's equations. There are a large number of transformations available: it would render impossible to cover all of them here. Instead, we show that at least one canonical transformation exists, and that it promotes the interpretation of the stochastic spreading process as effective mechanical systems. Consider:  $P_1(t) = 2p^{1/2}q$  and  $Q_1(t) = -p^{1/2}$ . The Poisson bracket  $\{Q_1, P_1\}_{qp} = (\partial Q_1/\partial q)(\partial P_1/\partial p) - (\partial Q_1/\partial p)(\partial P_1/\partial q) = 1$  shows the transformation is canonical. Setting  $m = 2$ , the Hamiltonian in terms of the canonical variables  $(Q_1, P_1)$  becomes

$$-\mathcal{H}_1 = \frac{1}{2m}(P_1 + \rho_0 Q_1)^2 - \frac{\rho_0^2 Q_1^2}{2m} - \frac{1}{Q_1^2}. \quad (17)$$

One may interpret  $-\mathcal{H}_1$  as the Hamiltonian of an effective mechanical problem in one-dimension, in which the particle has mechanical momentum  $P_1(\tau)$ , with generalized coordinate  $Q_1(\tau)$ , subjected to a velocity dependent potential.

## Conclusion

The description of several real-world problems often includes stochastic fluctuations. The SIS epidemic model includes them due to uncertainties associated with pathogen transmission. For small fluctuation amplitudes,  $\langle \rho(\tau) \rangle$  and  $\sigma^2(\tau)$  are adequate descriptors. Our findings demonstrate  $\langle \rho(\tau) \rangle$  and  $1/\sigma(\tau)$  are conjugated variables, and they satisfy Hamilton's equation. These results link the stochastic SIS epidemic model with a pure dynamical system, which can be solved and manipulated using standard analytical tools. We find the Hamiltonian is a constant of motion for  $N \gg 1$ . However, finite size effects break the temporal symmetry of the system:  $\mathcal{H} \sim \tau^{-1/2}$  follows a power-law around the outbreak onset. A clear explanation for this scaling is still lacking. The relaxation time  $\tau_{eff}$  portrays the decay of  $\mathcal{H}$  until equilibrium sets in, meaning that it can also be used to characterize the SIS epidemic. Unlike popular estimates of epidemic growth rate,  $\tau_{eff}$  remains constant along time and can be extracted from data values of  $\mathcal{H}$ . Finally, our results also suggest a way to incorporate interactions into the SIS model via the Lagrangian function. This finding has intriguing implications for our understanding of facilitation-competition mechanisms between co-occurring diseases since it does not replicate the canonical procedure to obtain Langevin equations.

## Data availability

Numerical codes, simulation data and data descriptors are available at <https://doi.org/10.17605/OSF.IO/WFCEP>.

Received: 4 March 2019; Accepted: 11 October 2019;

Published online: 01 November 2019

## References

- Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. Royal Soc. A* **115**, 700–721, <https://doi.org/10.1098/rspa.1927.0118> (1927).
- Heesterbeek, H. *et al.* Modeling infectious disease dynamics in the complex landscape of global health. *Sci.* **347**, <https://doi.org/10.1126/science.aaa4339> (2015).
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979, <https://doi.org/10.1103/RevModPhys.87.925> (2015).
- Keeling, M. & Eames, K. Networks and epidemic models. *J. Royal Soc. Interface* **2**, 295–307, <https://doi.org/10.1098/rsif.2005.0051> (2005).
- Goeyvaerts, N. *et al.* Household members do not contact each other at random: implications for infectious disease modelling. *Proc. Royal Soc. B* **285**, 20182201, <https://doi.org/10.1098/rspb.2018.2201> (2018).
- Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203, <https://doi.org/10.1103/PhysRevLett.86.3200> (2001).
- Eames, K. T. D. & Keeling, M. J. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc. Natl. Acad. Sci. USA* **99**, 13330–13335, <https://doi.org/10.1073/pnas.202244299> (2002).
- Keeling, M. The implications of network structure for epidemic dynamics. *Theor. Popul. Biol.* **67**, 1–8, <https://doi.org/10.1016/j.tpb.2004.08.002> (2005).
- Newman, M. E. J. The structure and function of complex networks. *SIAM Rev.* **45**, 167–256, <https://doi.org/10.1137/S003614450342480> (2003).
- Bansal, S., Grenfell, B. T. & Meyers, L. A. When individual behaviour matters: homogeneous and network models in epidemiology. *J. Royal Soc. Interface* **4**, 879–891 (2007).
- Durrett, R. Some features of the spread of epidemics and information on a random graph. *Proc. Natl. Acad. Sci. USA* **107**, 4491–4498, <https://doi.org/10.1073/pnas.0914402107> (2010).
- Wang, H. *et al.* Effect of the interconnected network structure on the epidemic threshold. *Phys. Rev. E* **88**, 022801, <https://doi.org/10.1103/PhysRevE.88.022801> (2013).
- Horne, Z., Powell, D., Hummel, J. E. & Holyoak, K. J. Countering antivaccination attitudes. *Proc. Natl. Acad. Sci. USA* **112**, 10321–10324, <https://doi.org/10.1073/pnas.1504019112> (2015).
- Roberts, M., Andreasen, V., Lloyd, A. & Pellis, L. Nine challenges for deterministic epidemic models. *Epidemics* **10**, 49–53, Challenges in Modelling Infectious Disease Dynamics., <https://doi.org/10.1016/j.epidem.2014.09.006> (2015).
- Britton, T. *et al.* Five challenges for stochastic epidemic models involving global transmission. *Epidemics* **10**, 54–57, Challenges in Modelling Infectious Disease Dynamics., <https://doi.org/10.1016/j.epidem.2014.05.002> (2015).
- Pellis, L. *et al.* Eight challenges for network epidemic models. *Epidemics* **10**, 58–62, Challenges in Modelling Infectious Disease Dynamics., <https://doi.org/10.1016/j.epidem.2014.07.003> (2015).

17. Chowell, G., Sattenspiel, L., Bansal, S. & Viboud, C. Mathematical models to characterize early epidemic growth: A review. *Phys. Life Rev.* **18**, 66–97, <https://doi.org/10.1016/j.plrev.2016.07.005> (2016).
18. Ruxton, G. D. Low levels of immigration between chaotic populations can reduce system extinctions by inducing asynchronous regular cycles. *Proc. Royal Soc. B* **256**, 189–193, <https://doi.org/10.1098/rspb.1994.0069> (1994).
19. Earn, D. J. D., Rohani, P. & Grenfell, B. T. Persistence, chaos and synchrony in ecology and epidemiology. *Proc. Royal Soc. B* **265**, 7–10, <https://doi.org/10.1098/rspb.1998.0256> (1998).
20. Real, L. A. & Biek, R. Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes. *J. Royal Soc. Interface* **4**, 935–948, <https://doi.org/10.1098/rsif.2007.1041> (2007).
21. Hastings, A. Timescales, dynamics, and ecological understanding. *Ecol.* **91**, 3471–3480, <https://doi.org/10.1890/10-0776.1> (2010).
22. Allen, J. C., Schaffer, W. M. & Rosko, D. Chaos reduces species extinction by amplifying local population noise. *Nat.* **364**, 229, <https://doi.org/10.1038/364229a0> (1993).
23. Duncan, A. B., Gonzalez, A. & Kaltz, O. Stochastic environmental fluctuations drive epidemiology in experimental host-parasite metapopulations. *Proc. Royal Soc. B* **280**, 20131747, <https://doi.org/10.1098/rspb.2013.1747> (2013).
24. Marguta, R. & Parisi, A. Periodicity, synchronization and persistence in pre-vaccination measles. *J. Royal Soc. Interface* **13**, 20160258, <https://doi.org/10.1098/rsif.2016.0258> (2016).
25. Cardy, J. L. *Phase Transition and Critical Phenomena*, vol. 11 (Academic Press, New York, 1987).
26. Nakamura, G., Cardoso, G. & Martinez, A. Robust parameter determination in epidemic models with analytical descriptions of uncertainties. *arXiv e-prints* arXiv:1807.05301 Preprint at, <https://arxiv.org/abs/1807.05301v1> (2018).
27. Alves, D., Haas, V. & Caliri, A. The predictive power of  $r_0$  in an epidemic probabilistic system. *J. Biol. Phys.* **29**, 63–75, <https://doi.org/10.1023/A:1022567418081> (2003).
28. Murray, J. D. *Mathematical Biology: an Introduction*, vol. 1, 3<sup>rd</sup> edition edn (Springer, New York, 2002).
29. Rhodes, C. & Anderson, R. Epidemic thresholds and vaccination in a lattice model of disease spread. *Theor. Popul. Biol.* **52**, 101–118, <https://doi.org/10.1006/tpbi.1997.1323> (1997).
30. Rhodes, C. J., Jensen, H. J. & Anderson, R. M. On the critical behaviour of simple epidemics. *Proc. Royal Soc. B* **264**, 1639–1646, <https://doi.org/10.1098/rspb.1997.0228> (1997).
31. Wilson, K. G. Confinement of quarks. *Phys. Rev. D* **10**, <https://doi.org/10.1103/PhysRevD.10.2445> (1974).
32. Kogut, J. B. An introduction to lattice gauge theory and spin systems. *Rev. Mod. Phys.* **51**, 659–713, <https://doi.org/10.1103/RevModPhys.51.659> (1979).
33. Eguíluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M. & Apkarian, A. V. Scale-free brain functional networks. *Phys. Rev. Lett.* **94**, 018102, <https://doi.org/10.1103/PhysRevLett.94.018102> (2005).
34. Schneidman, E., Berry, M. J. II, Segev, R. & Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nat.* **440**, 1007, <https://doi.org/10.1038/nature04701> (2006).
35. Alcaraz, F. C. & Nakamura, G. M. Phase diagram and spectral properties of a new exactly integrable spin-1 quantum chain. *J. Phys. A: Math. Gen.* **43**, 155002, <https://doi.org/10.1088/1751-8113/43/15/155002> (2010).
36. Meloni, F., Nakamura, G. M., Granzotti, C. R. & Martinez, A. S. Vegetation cover reveals the phase diagram of patch patterns in drylands. *Phys. A* **534**, 122048, <https://doi.org/10.1016/j.physa.2019.122048> (2019).
37. Keeling, M. J. Metapopulation moments: coupling, stochasticity and persistence. *J. Animal Ecol.* **69**, 725–736, <https://doi.org/10.1046/j.1365-2656.2000.00430.x> (2000).
38. Van Mieghem, P. & Cator, E. Epidemics in networks with nodal self-infection and the epidemic threshold. *Phys. Rev. E* **86**, 016116, <https://doi.org/10.1103/PhysRevE.86.016116> (2012).
39. Cator, E. & Van Mieghem, P. Second-order mean-field susceptible-infected-susceptible epidemic threshold. *Phys. Rev. E* **85**, 056111, <https://doi.org/10.1103/PhysRevE.85.056111> (2012).
40. Näsell, I. The quasi-stationary distribution of the closed endemic sis model. *Adv. Appl. Prob.* **28**, 895–932, <https://doi.org/10.2307/1428186> (1996).
41. Näsell, I. On the time to extinction in recurrent epidemics. *J. Royal Stat. Soc. B* **61**, 309–330, <https://doi.org/10.1111/1467-9868.00178> (1999).
42. Näsell, I. Extinction and quasi-stationarity in the verhulst logistic model. *J. Theor. Biol.* **211**, 11–27, <https://doi.org/10.1006/jtbi.2001.2328> (2001).
43. Nakamura, G. M., Monteiro, A. C. P., Cardoso, G. C. & Martinez, A. S. Efficient method for comprehensive computation of agent-level epidemic dissemination in networks. *Sci. Rep.* **7**, 40885, <https://doi.org/10.1038/srep40885> (2017).
44. Nakamura, G. M., Monteiro, A. C. P., Cardoso, G. C. & Martinez, A. S. Finite symmetries in agent-based epidemic models. *Math. Comput. Appl.* **24**, <https://doi.org/10.3390/mca24020044> (2019).
45. Kiss, I. Z. & Simon, P. L. New moment closures based on a priori distributions with applications to epidemic dynamics. *Bull. Math. Biol.* **74**, 1501–1515, <https://doi.org/10.1007/s11538-012-9723-3> (2012).
46. dos Santos, R. V., Ribeiro, F. L. & Martinez, A. S. Models for allee effect based on physical principles. *J. Theor. Biol.* **385**, 143–152, <https://doi.org/10.1016/j.jtbi.2015.08.018> (2015).
47. Vilar, J. M. G. & Rubi, J. M. Determinants of population responses to environmental fluctuations. *Sci. Rep.* **8**, 887, <https://doi.org/10.1038/s41598-017-18976-6> (2018).
48. Goldstein, H. *Classical mechanics* (Addison-Wesley, 1950).
49. Chester, M. A law of nature? *Open J. Ecol.* **1**, <https://doi.org/10.4236/oje.2011.13011> (2011).
50. Kamenev, A. & Meerson, B. Extinction of an infectious disease: A large fluctuation in a nonequilibrium system. *Phys. Rev. E* **77**, 061107, <https://doi.org/10.1103/PhysRevE.77.061107> (2008).
51. Dykman, M. I., Schwartz, I. B. & Landsman, A. S. Disease extinction in the presence of random vaccination. *Phys. Rev. Lett.* **101**, 078101, <https://doi.org/10.1103/PhysRevLett.101.078101> (2008).
52. Hinder, J. & Schwartz, I. B. Epidemic extinction and control in heterogeneous networks. *Phys. Rev. Lett.* **117**, 028302, <https://doi.org/10.1103/PhysRevLett.117.028302> (2016).
53. Hinder, J. & Schwartz, I. B. Epidemic extinction paths in complex networks. *Phys. Rev. E* **95**, 052317, <https://doi.org/10.1103/PhysRevE.95.052317> (2017).
54. Parker, M., Kamenev, A. & Meerson, B. Noise-induced stabilization in population dynamics. *Phys. Rev. Lett.* **107**, 180603, <https://doi.org/10.1103/PhysRevLett.107.180603> (2011).
55. Schwartz, I. B., Billings, L., Dykman, M. & Landsman, A. Predicting extinction rates in stochastic epidemic models. *J. Stat. Mech.: Theory Exp.* **2009**, P01005, <https://doi.org/10.1088/1742-5468/2009/01/p01005> (2009).
56. Heffernan, J., Smith, R. & Wahl, L. Perspectives on the basic reproductive ratio. *J. Royal Soc. Interface* **2**, 281–293, <https://doi.org/10.1098/rsif.2005.0042> (2005).

## Acknowledgements

The authors acknowledge funding from CNPq 309851/2018-1 and Capes 88882.317630/2019-01.



### Author contributions

G.M.N. wrote the paper, performed research, and executed numerical simulations; A.S.M. edited the paper and designed the research. All authors reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to G.M.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019