

SCIENTIFIC REPORTS



OPEN

The effects of heterogeneity on stochastic cycles in epidemics

Francisco Herreras-Azcué  & Tobias Galla 

Models of biological processes are often subject to different sources of noise. Developing an understanding of the combined effects of different types of uncertainty is an open challenge. In this paper, we study a variant of the susceptible-infective-recovered model of epidemic spread, which combines both agent-to-agent heterogeneity and intrinsic noise. We focus on epidemic cycles, driven by the stochasticity of infection and recovery events, and study in detail how heterogeneity in susceptibilities and propensities to pass on the disease affects these quasi-cycles. While the system can only be described by a large hierarchical set of equations in the transient regime, we derive a reduced closed set of equations for population-level quantities in the stationary regime. We analytically obtain the spectra of quasi-cycles in the linear-noise approximation. We find that the characteristic frequency of these cycles is typically determined by population averages of susceptibilities and infectivities, but that their amplitude depends on higher-order moments of the heterogeneity. We also investigate the synchronisation properties and phase lag between different groups of susceptible and infected individuals.

It is now widely recognised that noise and uncertainty play an important role in modelling biological systems. Traditional approaches to modelling phenomena in biology¹ are often based on deterministic ordinary or partial differential equations, and do not aim to describe stochasticity. In order to capture epistemic uncertainty, static or dynamic noise variables are introduced in more modern mathematical biology. This randomness reflects the lack of detailed knowledge about phenomena at finer scales than described by the model at hand; any modelling approach necessarily operates at a set scale (e.g. cell, individual, or population), and does not capture in detail the processes at smaller scales. These are ‘emulated’ through effective randomness. Different types of such noise are frequently found in models of biological phenomena, including intrinsic demographic noise, extrinsic stochasticity, parameter uncertainty or heterogeneity between different types of interacting entities^{2,3}. Some of these random variables are static and do not evolve in time, others are described by dynamic time-dependent noise. Intrinsic noise, due to the stochastic dynamics of a system, has lately been the focus of many studies (see for example^{4–6}). Extrinsic or parametric noise, due to variations, heterogeneity or uncertainties in the parameters or the environment surrounding the process, has received similar attention (e.g.^{7–9}). To be able to adequately describe biological systems, however, it may be necessary to account for both these uncertainties which contribute to the noisy dynamics.

In the modelling of epidemics this is of particular importance. The infection process, driven by serendipitous contacts, is inherently stochastic, and heterogeneity in susceptibility to a disease or infectiousness of different individuals are known to exist and play a role in viral spread. Genetic differences that result in heterogeneous susceptibilities to a disease have been suggested to play an important role^{10,11}, and variation in viral reproduction from host to host have been observed in ref.¹². Behavioural, structural or contact differences between individuals are inevitable, but we focus our study on the former type of heterogeneity. However, the better part of the existing work focusing on heterogeneity of this type, does not explicitly seek to capture demographic noise. Instead one often assumes infinite populations and deterministic dynamics. This approach is often taken outside epidemics as well. Much existing work studies *individual* sources of uncertainty, heterogeneity and noise in isolation, but not their interacting together. A notable exception is the modelling of gene regulatory networks, in which the interaction of intrinsic and extrinsic noise is actively studied, see e.g. refs^{13–15}.

The effects of intrinsic noise have been recognised in recent years. In models with demographic processes, for example, intrinsic stochasticity has been seen to lead to sustained quasi-cycles^{16–19} in parameter regimes in which a deterministic model would converge to a stable fixed point. These quasi-cycles have been identified not only

Theoretical Physics, School of Physics and Astronomy, The University of Manchester, Manchester, M13 9PL, United Kingdom. Correspondence and requests for materials should be addressed to F.H.-A. (email: francisco.herreriasazcue@postgrad.manchester.ac.uk) or T.G. (email: tobias.galla@manchester.ac.uk)

in models of epidemic spread, but also in other instances of population dynamics, including in genetic circuits, evolutionary systems and in game theory^{20–23}. Heterogeneity has been and is being considered in epidemics as well. Age structure is studied for example in refs^{24,25}, seasonally changing infection rates in refs^{26,27}, variation in infectivity and/or susceptibility are addressed in refs^{28–32}, spatial structure has been approached in refs^{33–36}, and epidemics on static and dynamic networks are studied in refs^{7,37–41}. Heterogeneity has been found to generate outbreaks that propagate hierarchically^{38,42}, grow faster than in homogeneous populations³⁹, and have a lower total number of infected individuals^{43,44}.

Much of this work, whether describing a well-mixed population, a compartmented or structured one, is based on variants of the celebrated susceptible-infective-recovered (SIR) model. They can be described either by deterministic differential equations, or as a stochastic process involving a population of discrete individuals. In the former approach the population is effectively assumed to be infinite, so that the timing of stochastic infection, recovery or birth-death events ‘averages’ out, and smooth laws for the time evolution of the population are obtained. The latter approach explicitly captures the intrinsic randomness of infection, recovery and demographics. The population is taken to be finite, and its state discrete. The model evolves through discrete events (e.g. infections). In the simplest case this defines a Markovian random process, which often can be analysed further mathematically, at least to a good approximation. Starting from the master equation in a well-mixed population a set of stochastic differential equations can be derived in the limit of large, but finite populations⁴⁵. These can then be studied further within the ‘linear-noise approximation’ (LNA)⁴⁶. The mathematics are tractable and the corresponding theory is now well established. While remarkably powerful, this approach so far has mostly been used for well-mixed populations. The linear-noise approximation has also been applied to networked systems with contact heterogeneity (see e.g refs^{19,47}), but progress is then much harder and often relies on further moment-closure approximations.

The aim of our work is to introduce agent-to-agent heterogeneity into the SIR dynamics in a finite well-mixed population. This provides a middle ground between homogeneous well-mixed models and an explicitly networked population. At the same time, we maintain tractability and are able to characterise stochastic effects in finite populations via the linear-noise approximation. This allows us to systematically investigate the combination of parameter heterogeneity and demographic noise. We divide the population of agents into K different groups of susceptible individuals, where members of different groups have different susceptibilities. Similarly, in our model there are M classes of infective individuals, with each class representing a different propensity to pass on the disease. This follows the lines of ref.³², but we explicitly focus on the combination of heterogeneity and intrinsic noise. Intrinsic stochasticity had not been included in ref.³².

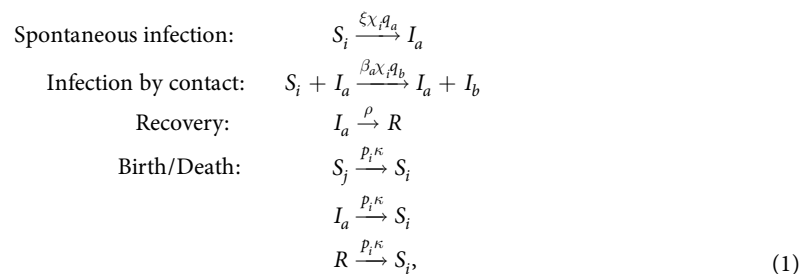
Our paper is organised as follows: In the following section we describe our model in detail. As a baseline we then construct the deterministic rate equations. They describe the deterministic dynamics in the limit of infinite populations, and are required to carry out the LNA. The most natural deterministic description will generally involve $K + M$ coupled non-linear equations (one for each subclass in the population). We discuss when and how these can be reduced to a smaller set of equations for aggregate quantities. In the next section we perform then the linear-noise approximation and use this approximation to characterise the fluctuations about deterministic fixed points. In particular we set up the theory to obtain the spectra of noise-driven quasi-cycles. Using this theory we then present our main results in the section titled “Consequences of heterogeneity”. We investigate in detail how the heterogeneity in the population affects the properties of stochastic outbreaks of the disease. Finally, in the last section we summarize our findings.

Model

We use an extension of the standard SIR model⁴⁸, in a population of fixed size N . Broadly, each individual can be of one of three types: susceptible (S), infective (I) or recovered (R). The spreading of the disease is described by infection events. These occur either through contact of a susceptible with an infective individual, as described below, or through spontaneous infection. Individuals recover at rate ρ , and they die at rate κ . The death rate is assumed to be independent of the disease status of an individual. To keep the number of individuals in the population constant, any death event is immediately followed by a birth of a new susceptible individual. This is of course an assumption, valid for large enough populations, so that fluctuations in the overall size can be neglected. The assumption is mainly made for simplicity and is not uncommon (see e.g refs^{16,49,50}).

We introduce heterogeneity by dividing the groups of susceptibles and infectives into subclasses. We will write S_i and I_a for these, with $i = 1, \dots, K$ and $a = 1, \dots, M$. Individuals in subgroup S_i have susceptibility χ_i to the disease, and infectives in class I_a have infectiousness β_a , which describes the propensity of the infective to pass on the disease to susceptible individuals. We write n_i for the number of individuals of type S_i , and m_a for the number of individuals in class I_a .

The dynamics are illustrated in Fig. 1, and can be summarised in the following reaction scheme:



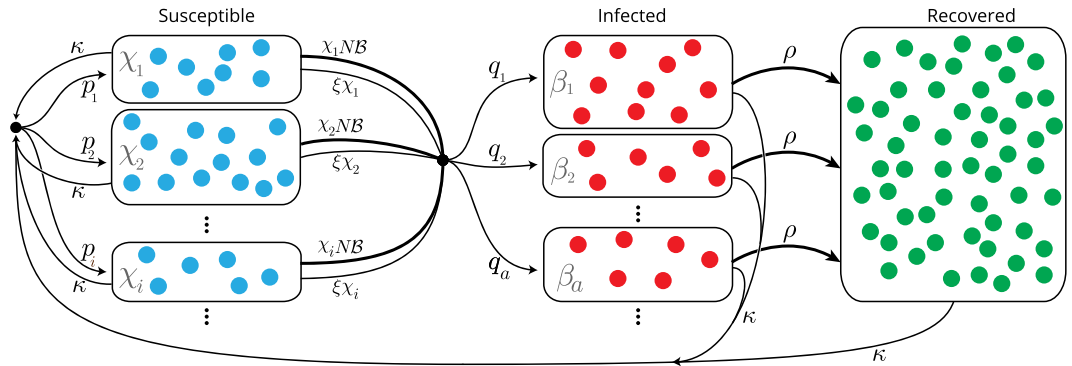


Figure 1. SIR model with heterogeneous susceptibility and infectivity. The diagram illustrates the different processes described by the model. New (susceptible) individuals are born at a rate κ , and they are assigned a susceptibility of χ_i with probability p_i . Susceptible individuals transition to an infected state either by spontaneous infection or by contact with any of the infected classes. The former process occurs with rate $\xi\chi_i$, if the susceptible is of type S_i . Contact infection occurs at a rate χ_iNB , where NB is the total infective power of the population (see Eq. (3)). Once infected, the individual is assigned an infectiousness β_a with probability q_a . All infected individuals recover at the same rate ρ . At any stage, individuals die with a rate κ . To keep the total population N constant, deceased individuals are immediately replaced by a new susceptible individual.

where $\{p_i\}$ and $\{q_a\}$ represent the probabilities of being assigned a susceptibility χ_i or infectiousness β_a at birth or upon infection, respectively. The first of these reactions describes spontaneous infection, converting an individual in class S_i into an individual of type I_a . The per-capita rate of events of this type is $\xi\chi_iq_a$, where ξ is an overall inverse time scale for spontaneous infection, χ_i is the susceptibility of S_i to the disease, and q_a is the probability that the newly infected individual is in class I_a . Similarly, the second reaction describes infection of an individual of type S_i upon contact with an individual of type I_a . The newly infected individual is in class I_b . Events of this particular type occur with a rate proportional to β_a (the propensity of I_a to spread the disease), to χ_i (the susceptibility of S_i) and to q_b . The third reaction describes recovery, and the final three reactions are birth/death events. The newly born individual is assumed to be randomly placed into one of the classes S_i ($i = 1, \dots, K$), occurring with respective probability p_i . We note that our model does not describe potential correlations between the susceptibility of an individual and its infectivity after they become infected; our focus is on heterogeneity of susceptibility due to physiological factors, and not primarily due to contact patterns. Extensions to include correlations can however be constructed along similar lines.

The model defines a continuous-time Markov process, and can be simulated straightforwardly using for example the celebrated Gillespie algorithm⁵¹. The starting point for the analytical study of the model is the master equation. Our analysis below will be based on approximating the solution to this master equation by performing a system-size expansion⁴⁶ and linear-noise approximation, leading to a stochastic differential equation describing the dynamics in the limit of large, but finite population size. In order to do this, it is useful to first introduce

$$\bar{\chi} = \sum_i p_i \chi_i, \quad \text{and} \quad \bar{\mathfrak{X}} = \frac{1}{N} \sum_i \chi_i n_i. \tag{2}$$

The quantity $\bar{\chi}$ is the mean susceptibility of a newly born individual, whereas $N\bar{\mathfrak{X}}$ describes the aggregate susceptibility of the population. Similarly, we define

$$\bar{\beta} = \sum_a q_a \beta_a \quad \text{and} \quad \mathcal{B} = \frac{1}{N} \sum_a \beta_a m_a, \tag{3}$$

where $\bar{\beta}$ represents the mean infectivity of a newly infected individual, and $N\mathcal{B}$ the total ‘infective power’ in the population. We note that $\bar{\chi}$ and $\bar{\beta}$ are fixed in time, and are properties of the distributions $\{p_i, \chi_i\}$ and $\{q_a, \beta_i\}$. The quantities $\bar{\mathfrak{X}}$ and \mathcal{B} , on the other hand, are time-dependent and evolve as the composition of the population changes.

Deterministic analysis

Dynamics. In the limit of an infinite population, the dynamics can be described by deterministic equations for the quantities $x_i = \lim_{N \rightarrow \infty} n_i/N$ and $y_a = \lim_{N \rightarrow \infty} m_a/N$. They are given by

$$\begin{aligned} \dot{x}_i &= \kappa p_i - \kappa x_i - \xi \chi_i x_i - \chi_i x_i \mathcal{B}, \\ \dot{y}_a &= \xi q_a \bar{\mathfrak{X}} + q_a \bar{\mathfrak{X}} \mathcal{B} - \rho y_a - \kappa y_a. \end{aligned} \tag{4}$$

These ordinary differential equations can be derived either by using direct mass-action kinetics, or from the lowest-order expressions in an expansion of the master equation in the inverse system size⁴⁶.

Ultimately we will mostly be interested in aggregate quantities, i.e. the total density of susceptibles or infectives in the population, irrespective of what subclass they belong to. We therefore introduce

$$S = \sum_i x_i \quad \text{and} \quad I = \sum_a y_a. \tag{5}$$

From Eqs (4) we find

$$\begin{aligned} \dot{S} &= \kappa - \kappa S - \xi \mathfrak{X} - \mathcal{B} \mathfrak{X}, \\ \dot{I} &= \xi \mathfrak{X} + \mathfrak{X} \mathcal{B} - \rho I - \kappa I. \end{aligned} \tag{6}$$

This system is not closed due to the presence of \mathfrak{X} and \mathcal{B} on the right-hand side. These quantities in turn evolve in time according to

$$\begin{aligned} \dot{\mathfrak{X}} &= \kappa \bar{\chi} - \kappa \mathfrak{X} - (\xi + \mathcal{B}) \sum_i \chi_i^2 x_i, \\ \dot{\mathcal{B}} &= \xi \mathfrak{X} \bar{\beta} + \bar{\beta} \mathfrak{X} \mathcal{B} - (\rho + \kappa) \mathcal{B}, \end{aligned} \tag{7}$$

which again does not close the set of equations, due to the presence of the term $\mathfrak{X}_2(t) \equiv \sum_i \chi_i^2 x_i(t)$. Modulo normalisation, and recalling that the $\{x_i\}$ are time-dependent, this object is recognised as the second moment of the distribution of susceptibilities among the group of susceptibles at time t . It cannot be determined from Eqs (6) and (7) alone. Instead we find

$$\dot{\mathfrak{X}}_n = \kappa \bar{\chi}^n - \kappa \mathfrak{X}_n - (\xi + \mathcal{B}) \mathfrak{X}_{n+1}, \tag{8}$$

where we have introduced $\bar{\chi}^n = \sum_i p_i \chi_i^n$ and $\mathfrak{X}_n = \sum_i x_i \chi_i^n$. This indicates that the deterministic dynamics at the aggregate level is described by an infinite hierarchy of equations. This set of equations does not close in the transient regime. However, as we will see next, closure can be achieved assuming the system settles down to a fixed point in the long run.

Fixed point. We proceed by a brief analysis of the fixed points of the deterministic dynamics. We will label these by a star. They can be obtained by setting $\dot{x}_i = 0$ and $\dot{y}_a = 0$ in Eq. (4), leading to

$$\begin{aligned} x_i^* &= \frac{\kappa p_i}{\kappa + (\xi + \mathcal{B}^*) \chi_i}, \\ y_a^* &= \frac{(\xi + \mathcal{B}^*) \mathfrak{X}^* q_a}{\rho + \kappa}. \end{aligned} \tag{9}$$

Similarly, we find the fixed points of the aggregate quantities S, I, \mathfrak{X} and \mathcal{B} from Eqs (6,7). After re-arranging and using Eq. (9) we arrive at

$$\begin{aligned} S^* &= 1 - \frac{(\rho + \kappa) \mathcal{B}^*}{\kappa \bar{\beta}}, \\ I^* &= \frac{\mathcal{B}^*}{\beta}, \\ \mathfrak{X}^* &= \frac{(\rho + \kappa) \mathcal{B}^*}{(\xi + \mathcal{B}^*) \bar{\beta}}, \\ \mathcal{B}^* &= \frac{\bar{\beta} \kappa}{(\rho + \kappa)} \sum_i \left(\frac{\chi_i p_i}{\xi + \mathcal{B}^* + \chi_i} \right), \end{aligned} \tag{10}$$

which is a closed set of equations, for a given set of model parameters $\{p_i, \chi_i, q_a, \beta_a\}$.

We highlight that while the transient dynamics of the system described in terms of the four macroscopic variables S, I, \mathfrak{X} and \mathcal{B} generates an infinite hierarchy of equations, potential fixed points can be uniquely described by a closed set of equations, assuming that the distribution of susceptibilities at birth and of the propensity of newly infected individuals to pass on the disease are known. In other words, the fixed point can be obtained in terms of the model parameters $\{q_a, \beta_a\}$ and $\{p_i, \chi_i\}$. While we cannot provide an analytical proof that the deterministic system will always converge to a fixed point, we note that, for the range of parameter used, we have not detected a single case in which numerically integrating Eq. (4) did not lead to a fixed point. In this context it is useful to point out that, in a homogeneous model, any combination of susceptibility and infectivity within the range of parameters used here would lead to a basic reproductive number above unity. For such models it is known that stable fixed points are eventually reached⁵².

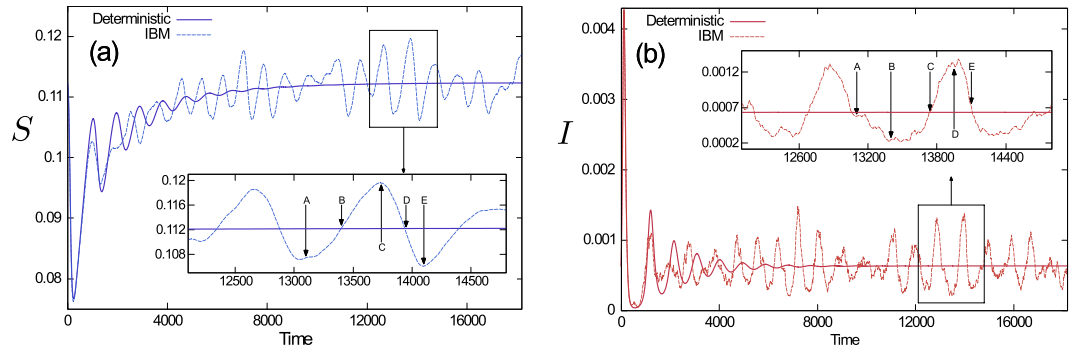


Figure 2. Population dynamics. Time series of the population density of total susceptible (panel (a)) and total infected individuals (panel (b)). Noise-sustained oscillations are clearly seen. The insets show a zoom in on the cycles. Labels A, B, ..., E are for later purposes (see below).

Linear-noise approximation

We now proceed to analyse the effects of stochasticity in the model, with a particular focus on the interaction between heterogeneity of individuals in the population and the noise induced by the demographics of the finite system.

We illustrate these effects in Fig. 2, and show an example of both the deterministic time-evolution of the system (thick continuous lines) and a realization of an individual-based simulation (thin dashed lines); the latter illustrates the intrinsic stochasticity of the process. Even after the deterministic model has reached a fixed point, the individual-based model shows sustained oscillations around it. These oscillations arise from a combination of complex eigenvalues of the underlying deterministic dynamics and the presence of intrinsic noise coming from the Poissonian jump process of the master equation. We will focus our attention on these stochasticity-driven periodic outbreaks in the remainder of this article, and build on the mathematical analysis via the linear-noise approximation²⁰. In particular we will study how the heterogeneity in the population affects the properties of these cycles.

Stochastic Dynamics. In order to carry out an analysis of the stochastic dynamics, we write $n_i/N = x_i + \tilde{x}_i/\sqrt{N}$, and $m_a/N = y_a + \tilde{y}_a/\sqrt{N}$, where $x_i(t)$ and $y_a(t)$ are the solutions of the deterministic equations (4) and the quantities with a tilde describe the stochastic fluctuations about the deterministic trajectory. The above ansatz reflects the anticipation that these fluctuations will have a relative magnitude of order $N^{-1/2}$. We then carry out an expansion in the inverse system size up to and including sub-leading order⁴⁶. In the fixed point regime of the deterministic dynamics we then arrive at

$$\begin{aligned} \dot{\tilde{x}}_i &= -\kappa\tilde{x}_i - (\xi + \mathcal{B}^*)\chi_i\tilde{x}_i - \chi_i x_i^* \tilde{\mathcal{B}} + \eta_i, \\ \dot{\tilde{y}}_a &= q_a(\xi\tilde{\mathcal{X}} + \tilde{\mathcal{X}}\mathcal{B}^* + \mathcal{X}^*\tilde{\mathcal{B}}) - (\rho + \kappa)\tilde{y}_a + \nu_a. \end{aligned} \tag{11}$$

The linear-noise approximation also applies during transients. All objects on the right-hand side of Eq. (11) then become time dependent. Since we ultimately focus on the oscillations about deterministic fixed point, we have not made this more explicit. The $\{\eta_i\}$ and $\{\nu_a\}$ are Gaussian white noise variables, with variance and co-variance (across components) as described in more detail in the Supplement (see Sec. S1). Writing $\tilde{\mathcal{S}} = \sum_i \tilde{x}_i$ and $\tilde{\mathcal{I}} = \sum_a \tilde{y}_a$ we find the following dynamics of fluctuations at the aggregate level,

$$\begin{aligned} \dot{\tilde{\mathcal{S}}} &= -\kappa\tilde{\mathcal{S}} - (\xi + \mathcal{B}^*)\tilde{\mathcal{X}} - \mathcal{X}^*\tilde{\mathcal{B}} + \sum_i \eta_i, \\ \dot{\tilde{\mathcal{I}}} &= (\xi + \mathcal{B}^*)\tilde{\mathcal{X}} + \mathcal{X}^*\tilde{\mathcal{B}} - (\rho + \kappa)\tilde{\mathcal{I}} + \sum_a \nu_a, \\ \dot{\tilde{\mathcal{X}}} &= -\kappa\tilde{\mathcal{X}} - \mathcal{X}^*\tilde{\mathcal{B}} - (\xi + \mathcal{B}^*)\sum_i \chi_i^2 \tilde{x}_i + \sum_i \chi_i \eta_i, \\ \dot{\tilde{\mathcal{B}}} &= (\xi + \mathcal{B}^*)\tilde{\beta}\tilde{\mathcal{X}} + \tilde{\beta}\mathcal{X}^*\tilde{\mathcal{B}} - (\rho + \kappa)\tilde{\mathcal{B}} + \sum_a \beta_a \nu_a. \end{aligned} \tag{12}$$

As in the deterministic analysis, this set of equations is not closed. It describes the dynamics of fluctuations about the deterministic fixed point, but makes no assumption of stationarity of the fluctuations (for example, correlation functions need not be time translation invariant). The lack of closure is due to the term $\sum_i \chi_i^2 \tilde{x}_i$ in the equation for $\tilde{\mathcal{X}}$. However, we will show below that a closed set of equations in the stationary state (of fluctuations) can be derived.

Fluctuation around the deterministic fixed point. We here show that although Eqs (12) are not closed, we can explore noise-induced oscillations around the deterministic fixed point. To this end, we introduce the Fourier transforms (with respect to time) of the variables \tilde{x}_i and \tilde{y}_a . We will denote these by \hat{x}_i and \hat{y}_a . From the Langevin equations (11) we find, after re-arranging,

$$\begin{aligned} \hat{x}_i &= \frac{-\chi_i x_i^* \hat{\mathcal{B}} + \hat{\eta}_i}{i\omega + \kappa + (\xi + \mathcal{B}^*)\chi_i}, \\ \hat{y}_a &= \frac{[(\xi + \mathcal{B}^*)\hat{\mathcal{X}} + \mathcal{X}^* \hat{\mathcal{B}}]q_a + \hat{\nu}_a}{i\omega + \rho + \kappa}. \end{aligned} \tag{13}$$

The noise variables $\{\eta_i\}$ and $\{\nu_a\}$ are uncorrelated in time, and their variance and correlation across components can be expressed in terms of known quantities (see Eq. (S3) in the Supplement). The variable ω is the conjugate of time under Fourier transform. Similarly, we find the following for the relevant aggregate quantities,

$$\begin{aligned} \hat{S} &= \frac{1}{i\omega + \kappa} \left[-\frac{i\omega + D}{\beta} \hat{\mathcal{B}} + \frac{1}{\beta} \sum_a \beta_a \hat{y}_a + \sum_i \hat{\eta}_i \right], \\ \hat{I} &= \frac{1}{i\omega + D} \left[\frac{i\omega + D}{\beta} \hat{\mathcal{B}} - \frac{1}{\beta} \sum_a \beta_a \hat{y}_a + \sum_a \hat{\nu}_a \right], \\ \hat{\mathcal{X}} &= \frac{1}{\beta C} \left[(i\omega + E) \hat{\mathcal{B}} - \sum_a \beta_a \hat{y}_a \right], \\ \hat{\mathcal{B}} &= \frac{\beta C \sum_i \frac{\chi_i \hat{\eta}_i}{i\omega + A_i} + \sum_a \beta_a \hat{y}_a}{i\omega + E + \beta C \kappa \sum_i \frac{\chi_i^2 p_i}{A_i(i\omega + A_i)}}, \end{aligned} \tag{14}$$

where, for simplicity, we have introduced the notation

$$\begin{aligned} A_i &= \kappa + (\xi + \mathcal{B}^*)\chi_i, \\ C &= \xi + \mathcal{B}^*, \\ D &= \rho + \kappa, \\ E &= \rho + \kappa - \beta \bar{\mathcal{X}}^*. \end{aligned} \tag{15}$$

Equations (14) constitute a closed set of equations for the Fourier transforms of the aggregate fluctuations \tilde{S} , \tilde{I} , $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{B}}$ in the stationary state. We thus make an observation similar to that in the deterministic analysis: although we cannot describe the evolution of fluctuations in the transient regime, we can derive a closed description of the statistics of fluctuations about deterministic fixed points within the linear-noise approximation.

Power Spectral Density. Equation (14) can be used to describe the periodic cycles shown in Fig. 2; we will now proceed to analyse these in more detail. Specifically, we will use the above results to compute the power spectral density (PSD) of fluctuations. This allows us to identify the characteristic frequency of noise-driven epidemic cycles, and to infer information about their amplitude.

The (average) power spectral density of a time series, $z(t)$, generated from the stochastic individual-based model, is given by $\mathcal{P}_z(\omega) = \langle |\hat{z}(\omega)|^2 \rangle$, where $\langle \dots \rangle$ stands for an average over realizations of the stochastic dynamics. The PSD can be computed analytically for all individual signals x_i, y_a , and for the aggregate variables S, I, \mathcal{X} and \mathcal{B} . The resulting expressions are lengthy; for completeness we provide them in the Supplement (see Sec. S2). As an illustration we here show the PSD of \mathcal{B} ,

$$\mathcal{P}_{\mathcal{B}}(\omega) = \frac{2\bar{\mathcal{X}}^* C}{|g|^2} \left(\beta^2 - \frac{\beta^2 C \kappa}{D} \sum_i \frac{\chi_i p_i A_i}{\omega^2 + A_i^2} \right) - \frac{(\beta C \kappa)^2}{|g|^2} \left[\sum_{ij} \frac{p_i p_j \chi_i \chi_j (A_i + A_j)(\omega^2 + A_i A_j)}{A_i A_j (\omega^2 + A_i^2)(\omega^2 + A_j^2)} \right], \tag{16}$$

with

$$|g|^2 = \left[E + \beta C \kappa \sum_i \frac{\chi_i^2 p_i}{\omega^2 + A_i^2} \right]^2 + \omega^2 \left[1 - \beta C \kappa \sum_i \frac{\chi_i^2 p_i}{A_i (\omega^2 + A_i^2)} \right]^2. \tag{17}$$

As detailed in the Supplement (see Sec. S2) the power spectra of S, I and \mathcal{X} can be expressed in terms of that of \mathcal{B} ; many of the characteristics of the spectra of S, I and \mathcal{X} are shared with those of \mathcal{B} , or directly related to it. We note that the RHS of Eq. (16) is proportional to $1/|g|^2$, and the same is the case for the spectral densities of \mathcal{X}, S and I (see Eq. (S10) in the Supplement); as a result, some of the key properties of the power spectra are determined by the behaviour of $|g|^2$, as discussed in more detail below.

Test Against Simulations. To illustrate the model and test our analytical results, we sampled possible heterogeneous populations. Specifically, the simulations shown in Fig. 3 are for populations with five susceptible and three infected subclasses. For each example, the probabilities $\{p_i\}$ and $\{q_a\}$ were drawn at random from a flat distribution over the simplexes $\sum_i p_i = 1$ and $\sum_a q_a = 1$. Susceptibilities and infectivities were assigned randomly in the intervals $0.5 \leq \chi_i \leq 2.5$ and $0.3 \leq \beta_a \leq 1.3$. Simulations are for $N = 10^6$, and the rates for recovery, birth/death and immigration were set at $\rho = 0.07, \kappa = 5.5 \times 10^{-5}$ and $\xi = 5 \times 10^{-6}$ respectively. The rates β_a, ρ, κ and

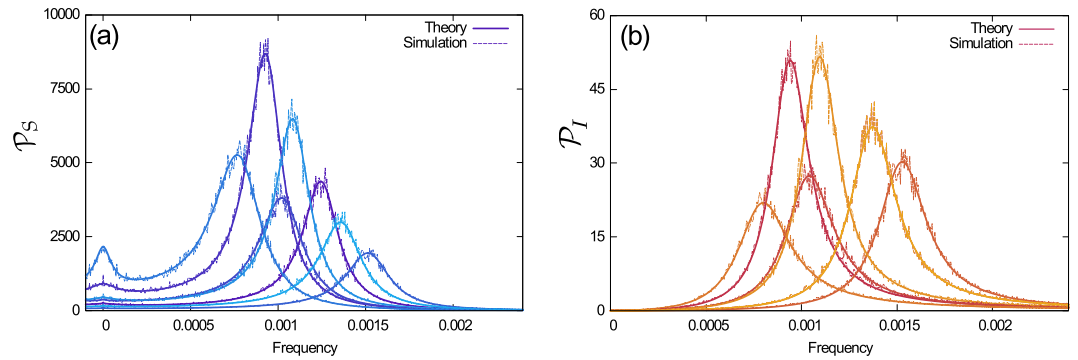


Figure 3. Power spectral densities of the fluctuations of (a) Susceptible and (b) Infected population for seven different examples of the model, generated as explained in more detail in the text. In all cases theory and simulations agree.

ξ have units of days^{-1} , whereas χ_i is dimensionless. The chosen rates are representative of childhood diseases such as whooping cough, measles, rubella or chickenpox⁵³.

The resulting PSDs are shown in Fig. 3. The continuous thick lines show the analytical result, and dashed lines are obtained from simulations, as an average over realizations of the individual-based model. As can be seen from the figure, the predictions of Eq. (S10) precisely match the results from simulations. In all figures, axes labelled ‘frequency’ show $f = \omega/2\pi$, and have units of days^{-1} .

It is interesting to note that the power spectral density can remain non-zero at zero frequency. A more detailed analysis reveals that its value is finite (i.e., not diverging), there is no evidence of e.g. a delta-peak at $\omega = 0$. This indicates that the area under the overall correlation function of fluctuations is non-zero, but finite, and there is no discernible shift of the overall stationary equilibrium (such a shift would result in a diverging contribution to the power spectrum at $\omega = 0$).

Consequences of Heterogeneity

Having established an analytical description of quasi-cycles, we now use this theory to identify which properties of the distribution of p_i , χ_i , q_a and β_a are most relevant for the characteristics of stochastic quasi-cycles in heterogeneous populations. Specifically, we study how heterogeneity in the population affects the dominant frequency of quasi-cycles, their amplitude and the sharpness of the spectra. We will then also discuss if and how the different subgroups synchronise during the epidemic cycles.

Dominant Cycle Frequency. Numerical inspection of the different terms in the analytical solution of the PSDs suggests that the dominating element is the factor $1/|g|^2$, as briefly indicated above. The frequency for which $|g|^2$ reaches its minimum roughly corresponds to the dominant cycle frequency, ω_d , in the PSDs. The minimum of $|g|^2$ can be found by differentiation of the expression in Eq. (17). Assuming that $C\chi_i \gg \kappa$, we further approximate the location of this minimum. This assumption is valid if infection processes occur on a time scale which is much shorter than the life expectancy of an individual. Further, we assume that $\omega \gg A_i$, i.e., that a susceptible individual typically lives through several epidemic events before it becomes infected. Both approximations are intuitively plausible for childhood diseases, known to show periodic outbreaks⁵³. Making these assumptions we find that the frequency for which $|g|^2$ is minimal can be approximated as

$$\omega_d \approx \sqrt{\kappa \bar{\chi} \bar{\beta}}. \quad (18)$$

This implies that the characteristic frequency is determined (mostly) by the mean susceptibility at birth and the mean infectivity at infection ($\bar{\chi}$ and $\bar{\beta}$) and the capacity of replenishment of the susceptible pool (κ).

The validity of our approach is confirmed in Fig. 4(a), where we test the approximation against simulations for a wide set of parameters. A perhaps more intuitive representation of our result can be found in Fig. 4(b), where we show the power spectra of several sample populations, each with different distributions of $\{p_i, \chi_i, q_a, \beta_a\}$, but all with the same first moments $\bar{\chi}$ and $\bar{\beta}$. As seen in the figure, this produces spectra of different amplitudes but with the same characteristic frequency. For comparison we include the homogeneous case $K = M = 1$.

Amplitude of Stochastic Cycles. While we have found above that the dominant frequency of stochastic cycles is largely determined by the first moments $\bar{\chi}$ and $\bar{\beta}$, the results shown in Fig. 4(b) demonstrate that this is not the case for the amplitude of the spectra at the dominant frequency. To investigate this further, we evaluate the analytic expressions for the PSDs in Eq. (S10) at the approximation of ω_d in Eq. (18). Making the same assumptions as in the previous section, we find that the height of the peak in the power spectra can be approximated as

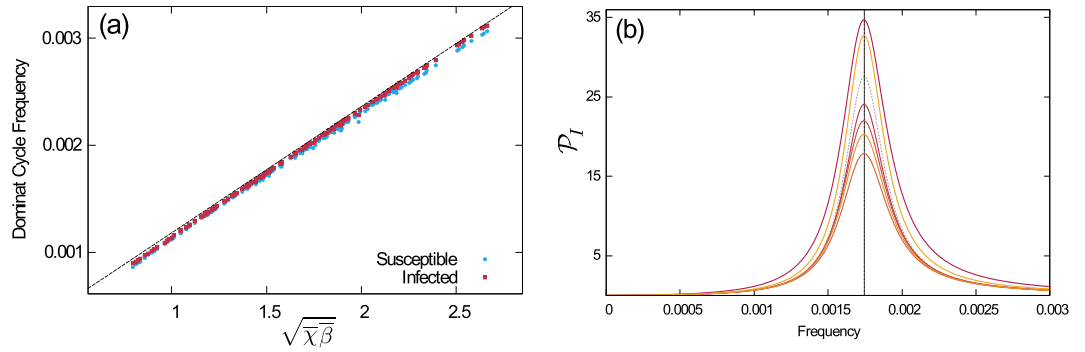


Figure 4. Verification of approximation (18) for the dominating frequency of cycles. **(a)** Frequency $f = \omega/2\pi$ at the maximum of the PSD, determined from Eq. (S10) as a function of $\sqrt{\bar{\chi}\bar{\beta}}$, for fixed κ . The black dashed line corresponds to Eq. (18). Markers are from 200 different populations, each with 5 susceptible and 3 infected subgroups, and with random choices of $\{p_i, \chi_i, q_a, \beta_a\}$. The values of χ_i and β_a were chosen from the interval 1.7 ± 1.69999995 ; q_a and p_i from a flat distribution. This resulted in values of $\bar{\chi}$ and $\bar{\beta}$ in the range 0.3 to 3.3, and for χ^2 and β^2 in the range 0.1 to 10. **(b)** PSD of the total infected population of different random distributions of $\{p_i, \chi_i, q_a, \beta_a\}$, with equal values for $\bar{\chi}$ and $\bar{\beta}$, but different values of χ^2 and β^2 . As a consequence of Eqs (18) and (19), the characteristic frequency is the same for all such samples, but the height of the peak in the PSD varies considerably (the amplitude of the oscillations changes with the square root of the amplitude of the power spectra). The dashed grey line corresponds to the homogeneous model, i.e. $K = M = 1$. The vertical dotted line is a visual aid.

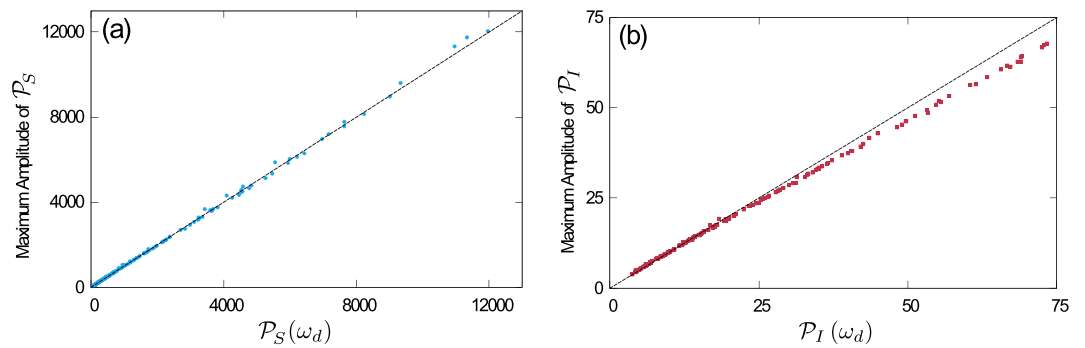


Figure 5. Verification of approximation (19) for the peak-height of the spectral densities. Horizontal axes show the prediction of Eq. (19) for susceptibles **(a)**, and infectives **(b)**. On the vertical axis we show the height at the peak of the spectra, as determined numerically from Eq. (S10) in the Supplement. Black dashed lines are the diagonal ($y = x$), and markers represent the populations described in Fig. 4.

$$\begin{aligned}
 P_I(\omega_d) &\approx \frac{2(\rho + \kappa)}{[\frac{(\rho + \kappa)\xi}{B^*} + \frac{B^*\chi^2}{\bar{\chi}}]^2} \frac{\bar{\beta}^2}{\bar{\beta}^3}, \\
 P_S(\omega_d) &\approx \frac{(\rho + \kappa)^2}{\kappa\bar{\chi}\bar{\beta}} P_I(\omega_d).
 \end{aligned}
 \tag{19}$$

We note the presence of the second moments $\bar{\chi}^2$ and $\bar{\beta}^2$, unlike in Eq. (18). This indicates that the spread of susceptibilities and infectivities is relevant to the size of the fluctuations about the endemic equilibrium. We note that the case $K = M = 1$ in Fig. 4(b) is special, as it leads to zero variance of the disorder by construction. We have experimented with the number of groups, K and M , and to a good approximation we find that the number of groups only affects the height and location of the peak in the spectrum through the mean and variance of the distributions of β and χ .

In Fig. 5 we plot results from the approximation in Eq. (19) against the maximum amplitude of spectra obtained numerically from the full expression (within the LNA), see Eq. (S10) in the Supplement. The data confirms that the approximation is valid for a wide range of parameters. While we find slight deviations at large amplitudes in the case of the infectives, the approximation is very robust for the susceptible population.

Sharpness of the Spectra. We now turn to the sharpness of the peak in the PSDs. The sharper the peak, the closer the stochastic outbreaks are to perfect cyclic behaviour. Conversely, cyclic behaviour is less distinct if

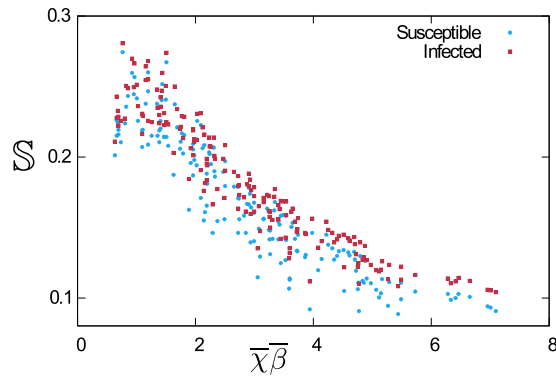


Figure 6. Sharpness of the power spectra as a function of the product of the mean susceptibilities and infectivities at birth/infection. Data is for the populations described in Fig. 4.

the peak in the spectrum is shallow. This has been described before as the ‘coherence’ of the spectra¹⁶. As we will investigate a different notion of coherence in the following section and in order to avoid confusion, we will refer to the concentration of power near the peak of the spectrum as ‘sharpness’.

Following¹⁶, we define the sharpness as the relative spectral power accumulated in an interval around the peak,

$$\mathbb{S} = \frac{\int_{\omega_d - \Delta\omega}^{\omega_d + \Delta\omega} \mathcal{P}(\omega) d\omega}{\int_{-\infty}^{+\infty} \mathcal{P}(\omega) d\omega}. \quad (20)$$

We compute the sharpness numerically, using the expressions in Eq. (S10). In order to evaluate the denominator in Eq. (20) we integrate up to an upper cutoff of $\omega_{max} = \pi/100 \text{ days}^{-1}$. In the numerator we use $\Delta\omega = 0.05\omega_{max}$. The choice of $\Delta\omega$ can be illustrated using Fig. 4(b), where the sharpness \mathbb{S} of the peak roughly corresponds to the fraction of total power concentrated in the interval between frequencies of 0.0015 and 0.002 days⁻¹.

In Fig. 6 we show the sharpness of spectra for 200 random populations (as described in Fig. 4). It is clear from the figure that there is a trend of increasing sharpness as the product of the mean susceptibility and infectivity at birth approaches unity (in the dimensions used here). The spread of the markers on the vertical axis indicates that there are significant effects of heterogeneity. It proves difficult, though, to find a functional dependence on higher moments of the distributions of susceptibilities and/or infectivities which would further collapse the data. While we do not show this data here in detail, we have also experimented with heterogeneity drawn from several distributions (e.g. flat, normal, Gamma). Results suggest that—to a good approximation—the functional shape of the spectra is determined by $\bar{\beta}$, $\bar{\chi}$, $\bar{\beta}^2$ and $\bar{\chi}^2$, i.e. by the first two moments of the heterogeneity. Higher-order features do not seem to play an important role. We have also tested the stronger property of full collapse upon re-scaling by peak height and location of peak, i.e. whether there is a scaling property of the type $P(\omega) = P_{max} \times f(\omega/\omega_d)$. This appears not to be the case.

Synchronization between Subgroups. We have established so far that introducing heterogeneity leads to significant changes in the quasi-cycles of the aggregate numbers of susceptible and infective individuals. However, we have not yet said much about the dynamics of the individual subgroups. In Fig. 7 we show the same example of sustained oscillations as in the inset of Fig. 2, but instead of the total susceptible and infected population we now highlight the time evolution of each of the subgroups.

In the upper two panels, (a) and (b), we show time series of the number of individuals in each subgroup normalised by the total population size. More specifically, we show susceptible subclasses (n_i/N) in panel (a), and infective subclasses (m_a/N) in panel (b). For each of these, stochastic oscillations can be observed. These cycles are pronounced for the case of the infective subgroups, panel (b), and more shallow for the susceptibles, panel (a). This is to be expected, given that the total number of susceptibles is more than an order of magnitude larger than those of the infectives (see also Fig. 2). From Fig. 7(a), (b) it is clear that all subgroups undergo cycling of roughly the same frequency. This is confirmed by the power spectra in Fig. 8.

We note that these statements rely on expressing number of individuals in each class as a fraction of the total population, and not relative to the time-dependent total number of susceptibles or infectives, respectively. We contrast the above with a representation in which we express the occupancy in each infective subgroup as a fraction of the infectives only, and similarly for the susceptibles. To this end we replot the simulation run shown in Fig. 7(a), (b), but now in terms of $n_i/(NS)$ and $m_a/(NI)$, respectively. The quantities $NS = \sum_j n_j$ and $NI = \sum_b m_b$ are the total number susceptible and infective individuals, respectively, and they are time-dependent themselves. Results are shown in Fig. 7(c), (d). Although the overall number of infectives, NI , undergoes the noise-driven cycles shown in Fig. 2, we find no discernible structure within the group of infectives; the time series $m_a/(NI)$ in Fig. 7(d) are essentially flat noisy lines. This is what one would expect, since the allocation to each subgroup, I_a , of infectives is random when an individual is newly infected, and the recovery rate is the same for all infective subgroups.

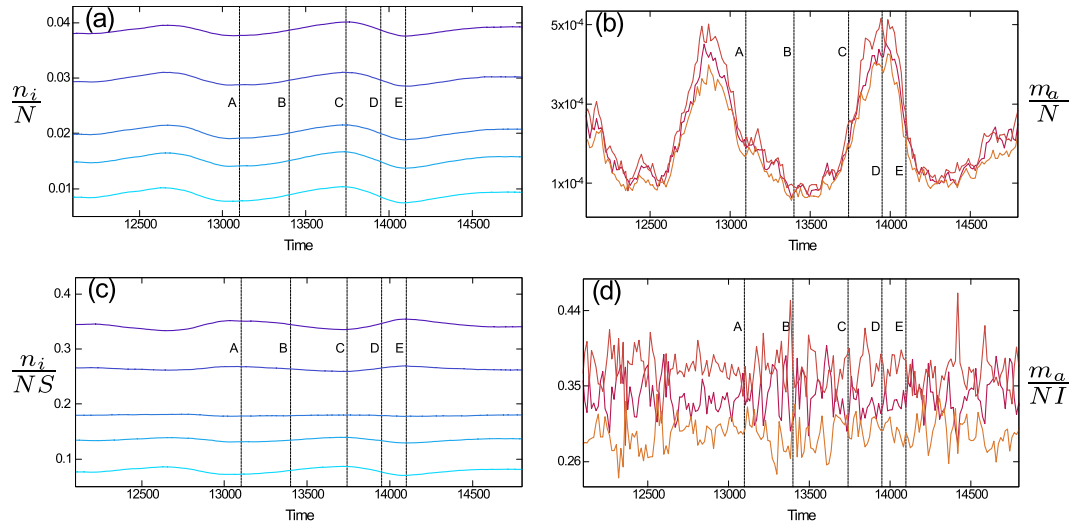


Figure 7. Stochastic cycles in subgroups of susceptibles and infectives. We show the same simulation run as in Fig. 2, but now split up into the different subgroups. Panels (a) and (b) show the number of individuals in each susceptible and infective subgroup normalised by the total population (N). In panels (c) and (d), we show the number of individuals in each subgroup divided by the total number of susceptible or infected individuals, respectively (NS and NI). Lines labelled A to E refer to points in the cycles of the aggregate variables S, I shown in Fig. 2.

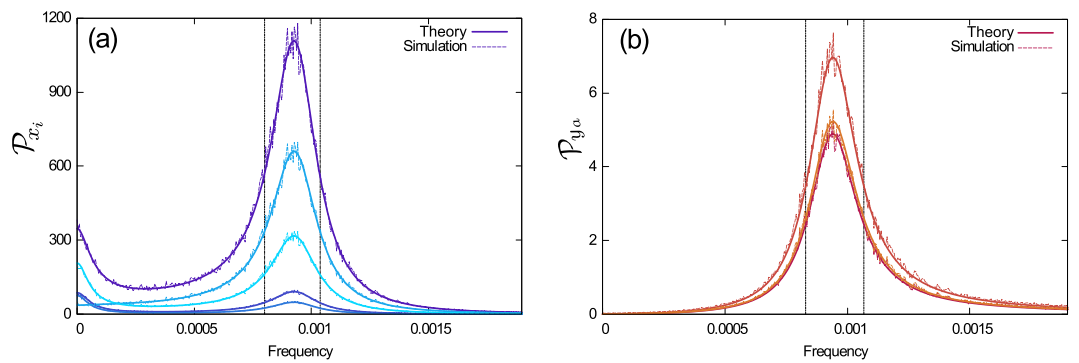


Figure 8. Power spectra of fluctuations for different subclasses of susceptibles and infectives. We use the same sample of the model parameters $\{\chi_i, p_i, \beta_a, q_a\}$ as in Fig. 3. Simulations are averaged over multiple realizations of the stochastic dynamics, at fixed model parameters. The vertical dotted lines are for later purposes and mark the locations at which the power spectra take values approximately equal to half the maximum amplitude.

A more complex behaviour can be seen within the group of susceptibles. This group as a whole undergoes stochastic cycles (see Fig. 2), but an interesting structure is observed within the group of susceptibles as well. The time series $n_i/(NS)$ in Fig. 7(c) show cyclic behaviour, and—to a good approximation—any pair of these time series is either in phase with each other, or they have a phase difference of $\pm\pi$.

To explore the phase lag between the different time series we use the so-called complex coherence function⁵⁴. This technique relies on computing the cross-spectrum $\langle \hat{x}_i(\omega)\hat{x}_j^*(\omega) \rangle$ between time series $x_i(t)$ and $x_j(t)$. The phase lag is then obtained as

$$\mathcal{L}_{x_i x_j}(\omega) = \tan^{-1} \frac{\text{Im} \langle \hat{x}_i(\omega)\hat{x}_j^*(\omega) \rangle}{\text{Re} \langle \hat{x}_i(\omega)\hat{x}_j^*(\omega) \rangle}. \tag{21}$$

We stress that the subscript * denotes complex conjugation, and is not to be confused with *, used earlier to indicate fixed points of the deterministic dynamics. Eq. (21) returns a phase lag for each spectral component, ω . Details can be found in the Supplement (see Sec. S3).

The phase lag between the different groups of susceptible individuals is shown in Fig. 9. The data in panel (a) corresponds to Fig. 7(a). More precisely, in Fig. 9(a) we pick the time series n_1/N as a reference, and show the phase lag of all subgroups n_i/N with respect to this reference time series. We find that the phase lag for frequencies around the dominant frequency in the power spectra is small, consistent with Fig. 7(a); all time series n_i/N

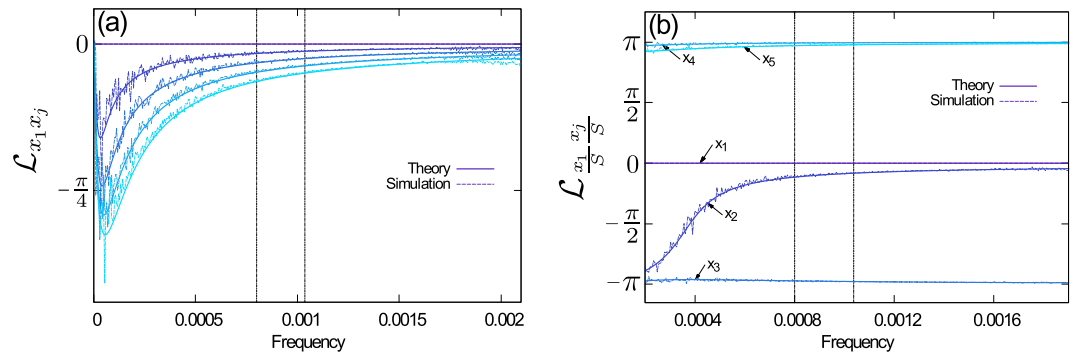


Figure 9. Phase-lag of time series between different subgroups of susceptibles. Data is for the same setup as in Fig. 7. We show the phase-lag between subgroups i and reference subgroup 1. Panel (a) depicts the case in which time series are normalized with respect to the total population, N ; in panel (b) input time series are normalized with respect to the total number of susceptibles NS . As in Fig. 8, the vertical dotted lines mark the half-width of the peaks in the corresponding power spectra.

oscillate (roughly) in phase with each other. In Fig. 9(b) we repeat this procedure, but now taking the time series $n_i/(NS)$ as an input, corresponding to Fig. 7(c). One then finds a rather different picture; the phase lag around the dominant frequency takes values either near zero, or close to $\pm\pi$. This indicates that the different classes of susceptible individuals fall into two groups. The time series in either group are in phase with each other, and in anti-phase with those in the respective other group. A closer inspection shows that these two groups are formed by the time series i with $x_i^* < S^*/K$ and with $x_i^* > S^*/K$ respectively. This behaviour in turn can be understood intuitively by revisiting Eq. (9). Assuming $\kappa \ll (\xi + \mathcal{B}^*)\chi_i$ for all i (a valid approximation for the cases analysed here), we find $x_i^* \propto 1/\chi_i$, indicating that the more susceptible classes are less populated at the deterministic fixed point than the less susceptible ones. During the increasing leg of a stochastic cycle, we expect the number of newly infected individuals among class i to be proportional to $x_i^*\chi_i$, suggesting that all susceptible classes are depleted in equal absolute numbers. This in turn means that subclasses with $x_i^* > S^*/K$ will represent an even larger fraction of the susceptible population as the total susceptible population decreases, while the subclasses with $x_i^* < S^*/K$ will represent a smaller fraction. This is what is observed in Fig. 7(c).

Conclusions

In summary, we have explored the SIR model in finite populations, including demographic processes and allowed for agent-to-agent heterogeneity in both the susceptibility to a disease and the capacity to spread the disease. This system combines the effects of intrinsic demographic stochasticity (due to random infection, recovery and birth-death events), with quenched heterogeneity. The focus of our paper is to characterise the interplay between these two types of stochasticity, and to investigate how the heterogeneity between individuals affects quasi-cycles driven by intrinsic noise. Our analysis relies on the system-size expansion, which allows us to compute the properties of these cycles analytically in the linear-noise approximation.

Our principal results can be summarised as follows: (i) In the deterministic limit of infinite populations, no closed set of equations for macroscopic quantities can be found in the transient regime. Fixed points for aggregate quantities of this deterministic dynamics can however be fully determined from a set of closed equations for the total susceptible (S^*) and infected (I^*) population, and weighted averages of the susceptibility ($\bar{\chi}$) and infectivity ($\bar{\mathcal{B}}$). (ii) Similarly, the Langevin equations in the linear-noise approximation do not close easily at the aggregate level, but a closed set of equations for the spectra of fluctuations in S , I , $\bar{\chi}$ and $\bar{\mathcal{B}}$ about the deterministic fixed point can be found in the stationary state. These can be used to analytically describe the stochastic oscillations about the fixed point. (iii) Within reasonable assumptions, the characteristic frequency of the noise-driven oscillations is determined mostly by the mean susceptibility and infectivity at birth or infection ($\bar{\chi}$ and $\bar{\mathcal{B}}$). However, the amplitude of the oscillations and the sharpness of peaks in the power spectra will generally depend on the higher moments of the distribution of susceptibilities and infectivities, in particular also on the agent-to-agent heterogeneity. (iv) Finally, the number of individuals in the different subclasses of infectives and susceptibles undergo stochastic cycles as well. If expressed in relation to the total population, these time series are synchronised and in phase. Normalized against the time-dependent total number of infectives, however, the different infective subclasses show no discernible oscillatory behaviour. Using a similar normalization within the susceptible population, we find that different subclasses are synchronised and either in phase with each other or have a phase difference of $\pm\pi$. These results are confirmed analytically. Regardless of the normalization, we find that the periodic outbreaks do not follow a hierarchical infection process, and all subgroups have similar absolute depletion/increase in absolute numbers. This is in contrast to what has been reported in single outbreak studies^{38,42}. However, it is important to note that in this existing work the outbreak is tracked in an initial transient period. Our results are valid after this period, at a deterministic fixed point, where the susceptible population is distributed in inverse proportion to their susceptibility (as explained above); this is a scenario different to the one studied in refs^{38,42}.

We think our results can be relevant for future work in several ways. First, our work contributes to the ongoing discussion about when and how a model with heterogeneity can be replaced or approximated by a homogeneous model. In previous studies, heterogeneous models were compared to homogeneous models with susceptibility equivalent to the arithmetic⁵⁵ or harmonic mean⁴⁴ of the susceptibilities in the different groups. More recently, the focus has been placed on equivalent basic reproduction numbers (R_0)⁵⁶. In the heterogeneous model this requires estimating R_0 based on, for example, the outbreak size, and therefore the comparison is not straightforward. Here we have shown that all models within the class we have looked at and with equal values of $\bar{\chi}\bar{\beta}$ generate periodic outbreaks with the same dominating frequency. This characteristic frequency can be used to define a unique homogeneous model to which models of varying degrees of heterogeneity can be compared. Furthermore, the dependence of the spectra of oscillations on both the first and higher moments of the distribution of heterogeneity might provide an avenue towards estimating how heterogeneous a population is from the observation of epidemic cycles. Finally, the formalism we have developed is versatile and can be applied to study quasi-cycles in other areas in which heterogeneity might be relevant, for example in predator-prey dynamics or evolution^{20,22,57–60}. Our findings indicate that the frequency of quasi-cycles can, to a good approximation, be obtained from the first moment of the distribution of heterogeneous agent properties, but that their amplitude depends on higher moments of the disorder. We expect similar behaviour in other heterogeneous systems with noise-driven cycles.

References

- Murray, J. D. *Mathematical biology*, vol. 17 of *Interdisciplinary Applied Mathematics* (Springer-Verlag, Berlin Heidelberg, 2002), 3rd edn.
- Wilkinson, D. J. *Stochastic Modelling for Systems Biology* (CRC Press, Boca Raton, 2011), 2nd edn.
- Goel, N. S. & Richter-Dyn, N. *Stochastic Models in Biology* (Academic Press, New York, NY, 1974).
- Andersson, H. & Britton, T. *Stochastic epidemic models and their statistical analysis*, vol. 151 of *Lecture Notes in Statistics* (Springer New York, New York, NY, 2000).
- Elowitz, M. B. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186, <https://doi.org/10.1126/science.1070919> (2002).
- Paulsson, J. Summing up the noise in gene networks. *Nature* **427**, 415–8, <https://doi.org/10.1038/nature02257> (2004).
- Moreno, Y., Pastor-Satorras, R. & Vespignani, A. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B* **26**, 521–529, <https://doi.org/10.1007/s10051-002-8996-y> (2002).
- Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* **135**, 216–226, <https://doi.org/10.1016/j.cell.2008.09.050> (2008).
- Lafuerza, L. F. & Toral, R. On the effect of heterogeneity in stochastic interacting-particle systems. *Scientific Reports* **3**, 1189, <https://doi.org/10.1038/srep01189> (2013).
- Bauch, C. T., Lloyd-Smith, J. O., Coffee, M. P. & Galvani, A. P. Dynamically Modeling SARS and Other Newly Emerging Respiratory Illnesses. *Epidemiology* **16**, 791–801, <https://doi.org/10.1097/01.ede.0000181633.80269.4c> (2005).
- Smith, D. L., Dushoff, J., Snow, R. W. & Hay, S. I. The entomological inoculation rate and Plasmodium falciparum infection in African children. *Nature* **438**, 492–495, <https://doi.org/10.1038/nature04024> (2005).
- Heldt, F. S. *et al.* Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nature Communications* **6**, 8938, <https://doi.org/10.1038/ncomms9938> (2015).
- Scott, M., Ingalls, B. & Kaern, M. Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **16**, 026107, <https://doi.org/10.1063/1.2211787> (2006).
- Swain, P. S., Elowitz, M. B. & Siggia, E. D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences* **99**, 12795–12800, <https://doi.org/10.1073/pnas.162041399> (2002).
- Hilfinger, A. & Paulsson, J. Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proceedings of the National Academy of Sciences* **108**, 12167–12172, <https://doi.org/10.1073/pnas.1018832108> (2011).
- Alonso, D., McKane, A. J. & Pascual, M. Stochastic amplification in epidemics. *Journal of The Royal Society Interface* **4**, 575–582, <https://doi.org/10.1098/rsif.2006.0192> (2007).
- Olsen, L. & Schaffer, W. Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. *Science* **249**, 499–504, <https://doi.org/10.1126/science.2382131> (1990).
- Black, A. J., McKane, A. J., Nunes, A. & Parisi, A. Stochastic fluctuations in the susceptible-infective-recovered model with distributed infectious periods. *Physical Review E* **80**, 021922, <https://doi.org/10.1103/PhysRevE.80.021922> (2009).
- Rozhnova, G. & Nunes, A. Fluctuations and oscillations in a simple epidemic model. *Physical Review E* **79**, 041922, <https://doi.org/10.1103/PhysRevE.79.041922> (2009).
- McKane, A. J. & Newman, T. J. Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical Review Letters* **94**, 1–4, <https://doi.org/10.1103/PhysRevLett.94.218102> (2005).
- Björnstad, O. N. & Grenfell, B. T. Noisy clockwork: time series analysis of population fluctuations in animals. *Science* **293**, 638–643, <https://doi.org/10.1126/science.1062226> (2001).
- Bladon, A. J., Galla, T. & McKane, A. J. Evolutionary dynamics, intrinsic noise, and cycles of cooperation. *Physical Review E* **81**, <https://doi.org/10.1103/PhysRevE.81.066122> (2010).
- Samoilov, M., Plyasunov, S. & Arkin, A. P. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences* **102**, 2310–2315, <https://doi.org/10.1073/pnas.0406841102> (2005).
- Bolker, B. M. & Grenfell, B. T. Chaos and biological complexity in measles dynamics. *Proceedings of the Royal Society B: Biological Sciences* **251**, 75–81, <https://doi.org/10.1098/rspb.1993.0011> (1993).
- Schenzle, D. An age-structured model of pre- and post-vaccination measles transmission. *Mathematical Medicine and Biology* **1**, 169–191, <https://doi.org/10.1093/imammb/1.2.169> (1984).
- Earn, D. J. D., Rohani, P., Bolker, B. M. & Grenfell, B. T. A simple model for complex dynamical transitions in epidemics. *Science* **287**, 667–670, <https://doi.org/10.1126/science.287.5453.667> (2000).
- Stone, L., Olinky, R. & Huppert, A. Seasonal dynamics of recurrent epidemics. *Nature* **446**, 533–536, <https://doi.org/10.1038/nature05638> (2007).
- Diekmann, O., Heesterbeek, J. & Metz, J. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* **28**, 365–382, <https://doi.org/10.1007/BF00178324> (1990).
- Hethcote, H. W. & Van Ark, J. W. Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs. *Mathematical Biosciences* **84**, 85–118, [https://doi.org/10.1016/0025-5564\(87\)90044-7](https://doi.org/10.1016/0025-5564(87)90044-7) (1987).

30. Nold, A. Heterogeneity in disease-transmission modeling. *Mathematical Biosciences* **52**, 227–240, [https://doi.org/10.1016/0025-5564\(80\)90069-3](https://doi.org/10.1016/0025-5564(80)90069-3) (1980).
31. Hickson, R. I. & Roberts, M. G. How population heterogeneity in susceptibility and infectivity influences epidemic dynamics. *Journal of Theoretical Biology* **350**, 70–80, <https://doi.org/10.1016/j.jtbi.2014.01.014> (2014).
32. Novozhilov, A. Epidemiological models with parametric heterogeneity: Deterministic theory for closed populations. *Mathematical Modelling of Natural Phenomena* **7**, 147–167, <https://doi.org/10.1051/mmnp/20127310> (2012).
33. Keeling, M. J. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society B: Biological Sciences* **266**, 859–867, <https://doi.org/10.1098/rspb.1999.0716> (1999).
34. Rohani, P. Opposite patterns of synchrony in sympatric disease metapopulations. *Science* **286**, 968–971, <https://doi.org/10.1126/science.286.5441.968> (1999).
35. Hagenaars, T. J., Donnelly, C. A. & Ferguson, N. M. Spatial heterogeneity and the persistence of infectious diseases. *Journal of Theoretical Biology* **229**, 349–359, <https://doi.org/10.1016/j.jtbi.2004.04.002> (2004).
36. Yu, J., Jiang, D. & Shi, N. SIR model with random perturbation. *Journal of Mathematical Analysis and Applications* **360**, 235–244, <https://doi.org/10.1016/j.jmaa.2009.06.050> (2009). Global stability of two-group.
37. Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences* **103**, 2015–2020, <https://doi.org/10.1073/pnas.0510525103> (2006).
38. Barthélemy, M., Barrat, A., Pastor-Satorras, R. & Vespignani, A. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Physical Review Letters* **92**, 178701–1, <https://doi.org/10.1103/PhysRevLett.92.178701> (2004).
39. Keeling, M. J. The implications of network structure for epidemic dynamics. *Theoretical Population Biology* **67**, 1–8, <https://doi.org/10.1016/j.tpb.2004.08.002> (2005).
40. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences* **101**, 15124–15129, <https://doi.org/10.1073/pnas.0308344101> (2004).
41. Holme, P. Information content of contact-pattern representations and predictability of epidemic outbreaks. *Scientific Reports* **5**, 14462, <https://doi.org/10.1038/srep14462> (2015).
42. Barthélemy, M., Barrat, A., Pastor-Satorras, R. & Vespignani, A. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology* **235**, 275–288, <https://doi.org/10.1016/j.jtbi.2005.01.011> (2005).
43. Boylan, R. D. A note on epidemics in heterogeneous populations. *Mathematical Biosciences* **105**, 133–137, [https://doi.org/10.1016/0025-5564\(91\)90052-K](https://doi.org/10.1016/0025-5564(91)90052-K) (1991).
44. Andersson, H. & Britton, T. Heterogeneity in epidemic models and its effect on the spread of infection. *Journal of Applied Probability* **35**, 651–661, <https://doi.org/10.1239/jap/1032265213> (1998).
45. Gardiner, C. W. *Handbook of stochastic methods* (Springer-Verlag, Berlin Heidelberg, 2003), 3rd edn.
46. van Kampen, N. G. *Stochastic processes in physics and chemistry*. 3rd edn, (Elsevier, Amsterdam, 1992).
47. Rozhnova, G. & Nunes, A. Cluster approximations for infection dynamics on random networks. *Physical Review E* **80**, 051915, <https://doi.org/10.1103/PhysRevE.80.051915> (2009).
48. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **115**, 700–721, <https://doi.org/10.1098/rspa.1927.0118> (1927).
49. Britton, T. & O'Neill, P. D. Bayesian Inference for Stochastic Epidemics in Populations with Random Social Structure. *Scandinavian Journal of Statistics* **29**, 375–390, <https://doi.org/10.1111/1467-9469.00296> (2002).
50. Shulgin, B., Stone, L. & Agur, Z. Pulse vaccination strategy in the SIR epidemic model. *Bulletin of mathematical biology* **60**, 1123–1148, [https://doi.org/10.1016/S0092-8240\(98\)90005-2](https://doi.org/10.1016/S0092-8240(98)90005-2) (1998).
51. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81**, 2340–2361, <https://doi.org/10.1021/j100540a008> (1977).
52. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, Princeton, NJ, 2008).
53. Anderson, R. M. & May, R. M. *Infectious diseases of humans: Dynamics and control* (Oxford University Press, Oxford, 1992).
54. Stoica, P. & Moses, R. *Spectral analysis of signals* (Pearson Prentice Hall, Upper Saddle River, 2004).
55. Ball, F. Deterministic and stochastic epidemics with several kinds of susceptibles. *Advances in Applied Probability* **17**, 1, <https://doi.org/10.2307/1427049> (1985).
56. Yates, A., Antia, R. & Regoes, R. R. How do pathogen evolution and host heterogeneity interact in disease emergence? *Proceedings of the Royal Society B: Biological Sciences* **273**, 3075–3083, <https://doi.org/10.1098/rspb.2006.3681> (2006).
57. Butler, T. & Goldenfeld, N. Robust ecological pattern formation induced by demographic noise. *Physical Review E* **80**, 030902, <https://doi.org/10.1103/PhysRevE.80.030902> (2009).
58. Black, A. J. & McKane, A. J. Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution* **27**, 337–345, <https://doi.org/10.1016/j.tree.2012.01.014> (2012).
59. Cremer, J., Reichenbach, T. & Frey, E. Anomalous finite-size effects in the Battle of the Sexes. *The European Physical Journal B* **63**, 373–380, <https://doi.org/10.1140/epjb/e2008-00036-x> (2008).
60. Mobilia, M. Oscillatory dynamics in rock “paper” scissors games with mutations. *Journal of Theoretical Biology* **264**, 1–10, <https://doi.org/10.1016/j.jtbi.2010.01.008> (2010).

Acknowledgements

FHA thanks Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexico) for support. TG acknowledges funding by the Engineering and Physical Sciences Research Council (EPSRC, UK) under grant number EP/K037145/1. TG would like to thank the Group of Nonlinear Physics, University of Santiago de Compostela, Spain for hospitality.

Author Contributions

F.H.A. and T.G. conceived and designed the study. F.H.A. carried out the analytical calculations and computer simulations. F.H.A. and T.G. interpreted results. F.H.A. and T.G. wrote and reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-12606-x>.

Competing Interests: The authors declare that they have no competing interests.

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) 2017