

Dominating occupancy processes by the independent site approximation

Ross McVinish*

Abstract

Occupancy processes are a broad class of discrete time Markov chains on $\{0, 1\}^n$ encompassing models from ecology and epidemiology. This model is compared to a collection of n independent Markov chains on $\{0, 1\}$, which we call the independent site model. We establish conditions under which an occupancy process is smaller in the lower orthant order than the independent site model. An analogous result for spin systems follows by a limiting argument.

Keywords: lower orthant order; multivariate stochastic ordering; propagation of chaos; spin system.

MSC2020 subject classifications: 60J10; 60J27; 60K35.

Submitted to ECP on December 6, 2021, final version accepted on November 12, 2022.

1 Introduction

Occupancy processes [10] are a class of discrete time Markov chains on $\{0, 1\}^n$. This class encompasses models from diverse areas including Hanski's incidence function model [9], which is one of the most important models in metapopulation ecology, contact-based epidemic spreading processes [7] and dynamic random graph models [8]. Furthermore, it was shown in [13] that occupancy processes are natural time discretisation for finite spin systems such as contact process, voter model and Ising model [12].

We define the occupancy process $(X, t \in \mathbb{N})$ as a discrete time Markov chain on $\{0, 1\}^n$ where, conditional on X_t , the $X_{i,t+1}$, $i = 1, \dots, n$, are independent with transition probabilities

$$\mathbb{P}(X_{i,t+1} = 1 \mid X_t) = C_i(X_t)(1 - X_{i,t}) + S_i(X_t)X_{i,t}, \quad (1.1)$$

where the functions $C_i : \{0, 1\}^n \rightarrow [0, 1]$ and $S_i : \{0, 1\}^n \rightarrow [0, 1]$ are called the colonisation and survival functions of site i in reference to metapopulation modelling. We interpret $X_{i,t} = 1$ as site i supporting a population at time t , and $X_{i,t} = 0$ as site i not supporting a population at time t . If site i does not support a population at time t , then the site will be colonised at time $t + 1$ with probability $C_i(X_t)$. Similarly, if site i supports a population at time t , the population will survive to time $t + 1$ with probability $S_i(X_t)$.

Although the occupancy process is a finite state Markov chain, the size of the state space usually renders standard analysis intractable. Instead a variety of approximations are employed to understand the process's behaviour. Provided the colonisation and

*University of Queensland, Australia. E-mail: r.mcvinish@uq.edu.au

survival functions can be extended from $\{0, 1\}^n$ to $[0, 1]^n$, a natural approximation of (1.1) is the deterministic process

$$p_{i,t+1} = C_i(p_t)(1 - p_{i,t}) + S_i(p_t)p_{i,t}, \tag{1.2}$$

where $p_{i,0} = X_{i,0}$. It is known that for suitable sets $\mathcal{H} \subset \mathbb{R}^n$ and assuming all the colonisation and survival functions are influenced by a large number of sites,

$$\sup_{h \in \mathcal{H}} \left| n^{-1} \sum_{i=1}^n h_i(X_{i,t} - p_{i,t}) \right|$$

is small in probability when n is large [3, 10].

Demonstrating the closeness of paths is not the only way to relate stochastic and deterministic models. Allen [1] (see also [16]) showed that the expectation of the stochastic logistic model is overestimated by its deterministic counterpart. A similar result has been demonstrated for the SIR epidemic model [20] and a general non-Markovian network based SIR model [21]. Our first aim is to establish conditions under which the analogous result for model (1.1) holds, namely $\mathbb{E}_0 X_{i,t} \leq p_{i,t}$, where \mathbb{E}_0 denotes expectation conditioned on the initial state X_0 .

We then consider another type of approximation to (1.1) called the independent site approximation. Define $W_t = (W_{1,t}, \dots, W_{n,t})$ where the $W_{i,\cdot}$ are independent Markov chains on $\{0, 1\}$ such that

$$\mathbb{P}(W_{i,t+1} = 1 | W_{i,t}) = C_i(p_t)(1 - W_{i,t}) + S_i(p_t)W_{i,t}, \tag{1.3}$$

$W_{i,0} = X_{i,0}$ and p_t satisfies (1.2). By construction $p_{i,t} = \mathbb{E}_0(W_{i,t})$ for all i and all t . The independent site approximation is motivated by propagation of chaos type results where finite collections of particles in interacting particle systems evolve almost independently of one another under certain conditions [17]. This phenomenon has been demonstrated for a number of population models that exhibit a law of large numbers [2, 6]. If for a fixed i and t the inequality $\mathbb{E}_0 X_{i,t} \leq p_{i,t}$ holds, then $X_{i,t} \leq_{st} W_{i,t}$, where \leq_{st} denotes the usual stochastic ordering. Our second aim is to show that (1.1) is smaller than (1.3) in a form of multivariate stochastic ordering called the lower orthant order. This result will not require the process to display any law of large numbers behaviour for the process.

As occupancy processes are natural time discretisation for finite spin systems, we obtain analogous results for spin systems. The bound on the expectations is obtained using the positive correlations property of spin systems. The stochastic ordering result for spin systems is obtained by applying a limiting argument to the occupancy process.

2 The deterministic system bounds the probability of occupation

In this section we show the deterministic process (1.2) provides a bound on the expected state of the occupancy process (1.1). The main step in the proof is the application of the Harris inequality.

Theorem 2.1. *Assume that for each i the functions C_i and S_i extended to $[0, 1]^n$ are increasing and concave, and the $S_i - C_i$ are decreasing and non-negative. If $p_{i,0} = X_{i,0}$ for all i , then $\mathbb{E}_0 X_{i,t} \leq p_{i,t}$ for all i and all $t \geq 0$.*

Proof. We can express the Markov chain X_t as

$$\begin{aligned} X_{i,t+1} &= (1 - X_{i,t})\mathbb{I}(U_{i,t+1} \leq C_i(X_t)) + X_{i,t}\mathbb{I}(U_{i,t+1} \leq S_i(X_t)) \\ &= \mathbb{I}(U_{i,t+1} \leq C_i(X_t)) + X_{i,t}\mathbb{I}(C_i(X_t) \leq U_{i,t+1} \leq S_i(X_t)) =: \mathcal{X}_i(U_{i,t+1}, X_t), \end{aligned} \tag{2.1}$$

where the $U_{i,t}$ form an array of independent standard uniform random variables. The function $\mathcal{X}_i(u, x)$ is decreasing in u for fixed $x \in \{0, 1\}^n$. Also, for any $x, y \in \{0, 1\}^n$ such that $x \leq y$ in the partial ordering on $\{0, 1\}^n$ (that is, $x \leq y \iff x_i \leq y_i$ for all i), we have $\mathcal{X}_i(u, x) \leq \mathcal{X}_i(u, y)$ for any $u \in [0, 1]$ as $C_i(x)$ and $S_i(x)$ are increasing in x . Hence $X_{i,t}$ is a decreasing function of the array $\{U_{i,t}\}$. As $S_i(x) - C_i(x)$ is decreasing in x , we see $S_i(X_t) - C_i(X_t)$ is an increasing function of the array $\{U_{i,t}\}$. Taking conditional expectations

$$\mathbb{E}(X_{i,t+1} \mid X_t) = C_i(X_t) + (S_i(X_t) - C_i(X_t)) X_{i,t}.$$

Then taking expectations and applying the Harris inequality

$$\begin{aligned} \mathbb{E}_0 X_{i,t+1} &\leq \mathbb{E}_0 C_i(X_t) + \mathbb{E}_0 (S_i(X_t) - C_i(X_t)) \mathbb{E}_0 X_{i,t} \\ &= (1 - \mathbb{E}_0 X_{i,t}) \mathbb{E}_0 C_i(X_t) + \mathbb{E}_0 X_{i,t} \mathbb{E}_0 (S_i(X_t)). \end{aligned}$$

As C_i and S_i are concave, we can apply Jensen's inequality to obtain

$$\mathbb{E}_0 X_{i,t+1} \leq (1 - \mathbb{E}_0 X_{i,t}) C_i(\mathbb{E}_0 X_t) + \mathbb{E}_0 X_{i,t} S_i(\mathbb{E}_0 X_t). \tag{2.2}$$

Write $\pi_{i,t} = \mathbb{E}_0 X_{i,t}$. Suppose $\pi_{i,t} \leq p_{i,t}$ for all i , where p_t satisfies the recursion (1.2) with $p_{i,0} = X_{i,0}$. Then

$$\begin{aligned} p_{i,t+1} &= C_i(p_t) + (S_i(p_t) - C_i(p_t)) p_{i,t} \\ &\geq C_i(p_t) + (S_i(p_t) - C_i(p_t)) \pi_{i,t} = (1 - \pi_{i,t}) C_i(p_t) + S_i(p_t) \pi_{i,t}, \end{aligned}$$

as $S_i - C_i \geq 0$ and $\pi_{i,t} \leq p_{i,t}$. Since C_i and S_i are increasing,

$$p_{i,t+1} \geq (1 - \pi_{i,t}) C_i(\pi_t) + S_i(\pi_t) \pi_{i,t} \geq \pi_{i,t+1}.$$

Hence, $\pi_{i,t} \leq p_{i,t}$ for all i and all $t \geq 0$. □

The deterministic process (1.2) requires the functions C_i and S_i to be extended from $\{0, 1\}^n$ to $[0, 1]^n$. Without imposing additional restrictions, these functions do not have a unique extension, but some extensions will be better than others in terms of how close $p_{i,t}$ is to $\mathbb{E}_0 X_{i,t}$. Let \tilde{p}_t be the solution to (1.2) with the functions C_i and S_i replaced by \tilde{C}_i and \tilde{S}_i satisfying $C_i(p) \leq \tilde{C}_i(p)$ and $S_i(p) \leq \tilde{S}_i(p)$ for all $p \in [0, 1]^n$. If $p_t \leq \tilde{p}_t$ in the partial order on $[0, 1]^n$, then

$$p_{i,t+1} = (1 - p_{i,t}) C_i(p_t) + p_{i,t} S_i(p_t) \leq (1 - \tilde{p}_i) \tilde{C}_i(\tilde{p}_t) + \tilde{p}_i \tilde{S}_i(\tilde{p}_t) = \tilde{p}_{i,t+1}.$$

In light of Theorem 2.1 we prefer smaller extensions of C_i and S_i that are increasing and concave. Methods for constructing the smallest concave extension are discussed in [18], though the gains achieved with these methods are unlikely to repay the computational effort required for their calculation. A relatively simple improvement can be obtained by noting that the occupancy process is not affected by the value assigned to $C_i(x)$ when $x_i = 1$. Suppose C_i is an increasing concave extension and define $\bar{C}_i(p) = C_i(\bar{p})$, where $\bar{p}_j = p_j$ for $j \neq i$ and $\bar{p}_i = 0$. The function \bar{C}_i is an increasing concave function which satisfies $C_i(p) \geq \bar{C}_i(p)$ for all $p \in [0, 1]^n$. This means we should avoid extensions of C_i which result in the deterministic process being 'self-colonising', that is the value of p_i affecting the value of $C_i(p)$. Similar comments apply to the extension of S_i since the process is not affected by the value of $S_i(x)$ when $x_i = 0$.

Since occupancy processes can be viewed as a time discretisation of finite spin systems [13, Algorithms 1 & 2], it is natural consider a version of Theorem 2.1 for those

processes. Any finite spin system $(X, t \in \mathbb{R}_+)$ can be represented as a Markov jump process in the usual transition notation:

$$\begin{aligned} X_i : \quad & 0 \rightarrow 1 \quad \text{at rate } \lambda_i(X) \\ & 1 \rightarrow 0 \quad \text{at rate } \mu_i(X) \end{aligned} \quad \text{for } i = 1, \dots, n, \tag{2.3}$$

where $\lambda_i, \mu_i : \{0, 1\}^n \rightarrow \mathbb{R}_+$. The expectation of $X_{i,t}$ satisfies

$$\mathbb{E}_0 X_{i,t} = X_{i,0} + \int_0^t \mathbb{E}_0 ((1 - X_{i,s})\lambda_i(X_s) - X_{i,s}\mu_i(X_s)) ds. \tag{2.4}$$

Provided the functions λ_i and μ_i can be extended from $\{0, 1\}^n$ to $[0, 1]^n$, this suggests the deterministic approximation for the spin system is the solution to the system of ordinary differential equations

$$p'_{i,t} = (1 - p_{i,t})\lambda_i(p_t) - p_{i,t}\mu_i(p_t). \tag{2.5}$$

Theorem 2.2. *Assume that for each i the functions λ_i extended to $[0, 1]^n$ are increasing and concave, μ_i extended to $[0, 1]^n$ are decreasing and convex, and the $\lambda_i + \mu_i$ are increasing. If $p_{i,0} = X_{i,0}$ for all i , then $\mathbb{E}_0 X_{i,t} \leq p_{i,t}$ for all i and all $t \geq 0$.*

Proof. With the λ_i increasing and the μ_i decreasing, the spin system (2.3) is said to be attractive [12, III Defintion 2.1]. An attractive spin system X with fixed initial condition X_0 has positive correlations at all times $t \geq 0$, that is

$$\mathbb{E}(f(X_t)g(X_t)) \geq \mathbb{E}f(X_t)\mathbb{E}g(X_t)$$

for all continuous functions f and g that are monotone in the sense $f(\eta) \leq f(\zeta)$ whenever $\eta \leq \zeta$ [12, II Theorem 2.14, III Theorem 2.2]. Let $\pi_{i,t} = \mathbb{E}_0 X_{i,t}$. Differentiating (2.4) gives

$$\pi'_{i,t} = \mathbb{E}_0 ((1 - X_{i,t})\lambda_i(X_t) - X_{i,t}\mu_i(X_t)). \tag{2.6}$$

As $\lambda_i(\cdot) + \mu_i(\cdot)$ is increasing, we can apply the positive correlations property to (2.6) to obtain

$$\begin{aligned} \pi'_{i,t} &\leq \mathbb{E}_0 \lambda_i(X_t) - \mathbb{E}_0 X_{i,t} \mathbb{E}_0 (\lambda_i(X_t) + \mu_i(X_t)) \\ &\leq (1 - \mathbb{E}_0 X_{i,t}) \mathbb{E}_0 \lambda_i(X_t) - \mathbb{E}_0 X_{i,t} \mathbb{E}_0 \mu_i(X_t). \end{aligned}$$

As λ_i is concave and μ_i is convex, Jensen's inequality yields

$$\pi'_{i,t} \leq (1 - \mathbb{E}_0 X_{i,t}) \lambda_i(\mathbb{E}_0 X_t) - \mathbb{E}_0 X_{i,t} \mu_i(\mathbb{E}_0 X_t) = (1 - \pi_{i,t}) \lambda_i(\pi_t) - \pi_{i,t} \mu_i(\pi_t) \tag{2.7}$$

Define the functions $\phi_i : [0, 1]^n \rightarrow \mathbb{R}_+$ such that

$$\phi_i(u_1, \dots, u_n) := (1 - u_i) \lambda_i(u) - u_i \mu_i(u).$$

As the λ_i are increasing and the μ_i are decreasing, each function ϕ_i is non-decreasing in each u_j for $j \neq i$. The system of differential inequalities (2.7) satisfies the conditions of a result by Ważewski [19] (see also [14, Theorem 1 of Section 13 in Chapter XI]), which allows us to conclude that if $p_{i,0} = X_{i,0}$ for all i , then $\pi_{i,t} \leq p_{i,t}$ for all i and $t \geq 0$. \square

3 Propagation of chaos and stochastic ordering

An interacting particle system is said to display propagation of chaos if the particles evolve almost independently of one another when the system size is large. Demonstrating this behaviour usually involves showing a law of large numbers holds so that the transition rates of the individual particles are well approximated by some deterministic

process. A propagation of chaos result was established for the occupancy process in [3, 10], where the independent site approximation was coupled to the occupancy process and the two processes shown to be close over finite time intervals.

Instead of attempting to show the occupancy process is close to the independent site approximation, in this section we show that the occupancy process is dominated by the independent site approximation in a certain sense.

A weak notion of multivariate stochastic ordering is the lower orthant order [15, Section 6.G.1]. We say that the random vector \mathbf{Y} is smaller than the random vector \mathbf{Z} in the lower orthant order, denoted $\mathbf{Y} \leq_{lo} \mathbf{Z}$, if

$$\mathbb{P}(Y_1 \leq \zeta_1, \dots, Y_m \leq \zeta_m) \geq \mathbb{P}(Z_1 \leq \zeta_1, \dots, Z_m \leq \zeta_m)$$

for all $(\zeta_1, \dots, \zeta_m) \in \mathbb{R}^m$. For distributions on the hypercube $\{0, 1\}^m$, this condition reduces to

$$\mathbb{P}(Y_i = 0 \text{ for all } i \in A) \geq \mathbb{P}(Z_i = 0 \text{ for all } i \in A),$$

for all subsets $A \subset \{1, 2, \dots, m\}$. Write \mathbb{P}_0 to denote conditioning on the initial state X_0 . The Harris inequality applied to the construction (2.1) shows

$$\mathbb{P}_0(X_{i,t} = 0 \text{ for all } i \in A) = \mathbb{E}_0 \left(\prod_{i \in A} (1 - X_{i,t}) \right) \geq \prod_{i \in A} \mathbb{E}_0(1 - X_{i,t}),$$

for a given t and all subsets $A \subset \{1, 2, \dots, n\}$. Then applying Theorem 2.1 we see

$$\prod_{i \in A} \mathbb{E}_0(1 - X_{i,t}) \geq \prod_{i \in A} \mathbb{E}_0(1 - W_{i,t}) = \mathbb{P}_0(W_{i,t} = 0 \text{ for all } i \in A).$$

This establishes $X_t \leq_{lo} W_t$ for a given time $t \geq 0$. We would like to establish the ordering relation between X and W for all times in the sense that for any subset $A \subseteq \{1, \dots, n\}$, positive integers m_i and times $t_{i,1}, \dots, t_{i,m_i}$

$$\begin{aligned} \mathbb{P}_0(X_{i,t_{i,j}} = 0 \text{ for all } i \in A, j \in \{1, \dots, m_i\}) \\ \geq \mathbb{P}_0(W_{i,t_{i,j}} = 0 \text{ for all } i \in A, j \in \{1, \dots, m_i\}). \end{aligned} \tag{3.1}$$

Note that for each $i \in A$, the set of times $t_{i,1}, \dots, t_{i,m_i}$ may be different.

Theorem 3.1. *Assume the conditions of Theorem 2.1 hold. Assume also that for all i , $S_i - C_i$ is convex. The process $(X, t \in \mathbb{N})$ given by (1.1) is smaller in the lower orthant order than the process $(W, t \in \mathbb{N})$ given by (1.3).*

Proof. Let A be a subset of $\{1, \dots, n\}$, and for each $i \in A$ take a positive integer m_i and times $t_{i,1}, \dots, t_{i,m_i}$. Then by the Harris inequality

$$\begin{aligned} &\mathbb{P}_0(X_{i,t_{i,j}} = 0 \text{ for all } i \in A, j \in \{1, \dots, m_i\}) \\ &= \mathbb{E}_0 \left(\prod_{i \in A} \prod_{j=1}^{m_i} (1 - X_{i,t_{i,j}}) \right) \geq \prod_{i \in A} \mathbb{E}_0 \left(\prod_{j=1}^{m_i} (1 - X_{i,t_{i,j}}) \right) \\ &\geq \prod_{i \in A} \mathbb{P}_0(X_{i,t_{i,j}} = 0 \text{ for all } j \in \{1, \dots, m_i\}). \end{aligned}$$

It remains to show that for each i

$$\mathbb{P}_0(X_{i,t_{i,j}} = 0 \text{ for all } j \in \{1, \dots, m_i\}) \geq \mathbb{P}_0(W_{i,t_{i,j}} = 0 \text{ for all } j \in \{1, \dots, m_i\}).$$

Now for $\omega = (\omega_1, \dots, \omega_m) \in \{0, 1\}^m$ define

$$P_m^X(\omega) := \mathbb{P}_0(X_{i,1} \leq \omega_1, \dots, X_{i,m} \leq \omega_m) = \mathbb{E}_0 \left[\prod_{t=1}^m (1 - X_{i,t})^{1-\omega_t} \right].$$

and define $P_m^W(\omega)$ similarly. We prove by induction that $P_m^X(\omega) \geq P_m^W(\omega)$ for all $\omega \in \{0, 1\}^m$ and all $m \geq 1$.

Assume $\omega \neq \mathbf{1}$ and let $\phi(\omega) = \max\{j : \omega_j = 0\}$. If $\phi(\omega) = 1$, then from Theorem 2.1

$$P_m^X(\omega_1, \dots, \omega_m) = \mathbb{E}_0(1 - X_{i,1}) \geq \mathbb{E}_0(1 - W_{i,1}) = P_m^W(\omega_1, \dots, \omega_m).$$

Suppose now that $\phi(\omega) = \tilde{m} \geq 2$. Then

$$\begin{aligned} P_m^X(\omega) &= \mathbb{E}_0 \left[\prod_{t=1}^{\tilde{m}} (1 - X_{i,t})^{1-\omega_t} \right] = \mathbb{E}_0 \left[\mathbb{E}_0 [(1 - X_{i,\tilde{m}}) | X_{\tilde{m}-1}] \prod_{t=1}^{\tilde{m}-1} (1 - X_{i,t})^{1-\omega_t} \right] \\ &= \mathbb{E}_0 \left[((1 - S_i(X_{\tilde{m}-1})) + (1 - X_{i,\tilde{m}-1})(S_i(X_{\tilde{m}-1}) - C_i(X_{\tilde{m}-1}))) \prod_{t=1}^{\tilde{m}-1} (1 - X_{i,t})^{1-\omega_t} \right]. \end{aligned}$$

The function $S_i(x) - C_i(x)$ is decreasing in x by assumption and $X_{i,t}$ is a decreasing function of the array $\{U_{i,t}\}$ by construction (2.1). As it is composition of two decreasing functions, $S_i(X_t) - C_i(X_t)$ is an increasing function of the array $\{U_{i,t}\}$. Applying the Harris inequality shows

$$\begin{aligned} \mathbb{E}_0 \left[(1 - X_{i,\tilde{m}-1})(S_i(X_{\tilde{m}-1}) - C_i(X_{\tilde{m}-1})) \prod_{t=1}^{\tilde{m}-1} (1 - X_{i,t})^{1-\omega_t} \right] \\ \geq \mathbb{E}_0 [(S_i(X_{\tilde{m}-1}) - C_i(X_{\tilde{m}-1})) P_m^X(\omega_1, \dots, \omega_{\tilde{m}-2}, 0, 1, \dots, 1)]. \end{aligned}$$

Since $S_i - C_i$ is also convex, Jensen's inequality with Theorem 2.1 shows

$$\begin{aligned} \mathbb{E}_0 \left[(1 - X_{i,\tilde{m}-1})(S_i(X_{\tilde{m}-1}) - C_i(X_{\tilde{m}-1})) \prod_{t=1}^{\tilde{m}-1} (1 - X_{i,t})^{1-\omega_t} \right] \\ \geq (S_i(p_{\tilde{m}-1}) - C_i(p_{\tilde{m}-1})) P_m^X(\omega_1, \dots, \omega_{\tilde{m}-2}, 0, 1, \dots, 1). \end{aligned}$$

The same argument shows

$$\mathbb{E}_0 \left[(1 - S_i(X_{\tilde{m}-1})) \prod_{t=1}^{\tilde{m}-1} (1 - X_{i,t})^{1-\omega_t} \right] \geq (1 - S_i(p_{\tilde{m}-1})) P_m^X(\omega_1, \dots, \omega_{\tilde{m}-1}, 1, \dots, 1).$$

Therefore,

$$\begin{aligned} P_m^X(\omega) &\geq (1 - S_i(p_{\tilde{m}-1})) P_m^X(\omega_1, \dots, \omega_{\tilde{m}-1}, 1, \dots, 1) \\ &\quad + (S_i(p_{\tilde{m}-1}) - C_i(p_{\tilde{m}-1})) P_m^X(\omega_1, \dots, \omega_{\tilde{m}-2}, 0, 1, \dots, 1). \end{aligned} \quad (3.2)$$

On the other hand, for the process W given in (1.3)

$$\begin{aligned} P_m^W(\omega) &= \mathbb{E}_0 \left[\prod_{t=1}^{\tilde{m}} (1 - W_{i,t})^{1-\omega_t} \right] = \mathbb{E}_0 \left[\mathbb{E}_0 [(1 - W_{i,\tilde{m}}) | W_{\tilde{m}-1}] \prod_{t=1}^{\tilde{m}-1} (1 - W_{i,t})^{1-\omega_t} \right] \\ &= \mathbb{E}_0 \left[((1 - S_i(p_{\tilde{m}-1})) + (1 - W_{i,\tilde{m}-1})(S_i(p_{\tilde{m}-1}) - C_i(p_{\tilde{m}-1}))) \prod_{t=1}^{\tilde{m}-1} (1 - W_{i,t})^{1-\omega_t} \right]. \end{aligned}$$

Therefore,

$$P_m^W(\omega) = (1 - S_i(p_{\tilde{m}-1}))P_m^W(\omega_1, \dots, \omega_{\tilde{m}-1}, 1, \dots, 1) + (S_i(p_{\tilde{m}-1}) - C_i(p_{\tilde{m}-1}))P_m^W(\omega_1, \dots, \omega_{\tilde{m}-2}, 0, 1, \dots, 1). \quad (3.3)$$

If $P_m^X(\omega) \geq P_m^W(\omega)$ for all $\omega \in \{0, 1\}^m$ such that $\phi(\omega) \leq \tilde{m} - 1$, then comparing (3.2) and (3.3) shows $P_m^X(\omega) \geq P_m^W(\omega)$ for all $\omega \in \{0, 1\}^m$ such that $\phi(\omega) \leq \tilde{m}$. \square

For spin systems the independent site approximation is given by $W_t = (W_{1,t}, \dots, W_{n,t})$ where the $W_{i,\cdot}$ are independent Markov chains on $\{0, 1\}$ such that

$$W_i : \begin{array}{ll} 0 \rightarrow 1 & \text{at rate } \lambda_i(p_t) \\ 1 \rightarrow 0 & \text{at rate } \mu_i(p_t) \end{array} \quad \text{for } i = 1, \dots, n, \quad (3.4)$$

$W_{i,0} = X_{i,0}$ and p_t satisfies (2.5). Since the lower orthant order is closed under convergence in distribution [15, Theorem 6.G.3(d)], a limiting argument can be used to prove the following result.

Theorem 3.2. *Assume the conditions of Theorem 2.2 hold. Assume also that for all i , $\lambda_i + \mu_i$ is concave, and each of the λ_i and μ_i are Lipschitz continuous. The process $(X, t \in \mathbb{R}_+)$ given by (2.3) is smaller in the lower orthant order (3.1) than the process $(W, t \in \mathbb{R}_+)$ given by (3.4).*

Proof. For $\delta > 0$ sufficiently small, let $(X^\delta, t \in \mathbb{N})$ be the occupancy process with

$$C_i(x) = \delta\lambda_i(x), \quad \text{and} \quad S_i(x) = 1 - \delta\mu_i(x).$$

The assumptions of Theorem 3.1 are satisfied by X^δ . Let $(W^\delta, t \in \mathbb{N})$ be the corresponding independent site approximation (1.3) so X^δ is smaller than W^δ in the lower orthant order. Let $(N, t > 0)$ be a unit rate Poisson process independent of X^δ and W^δ . Define the continuous time process $(\tilde{X}^\delta, t \in \mathbb{R}_+)$ by $\tilde{X}_t^\delta := X_{N(\delta^{-1}t)}^\delta$ and $(\tilde{W}^\delta, t \in \mathbb{R}_+)$ by $\tilde{W}_t^\delta := W_{N(\delta^{-1}t)}^\delta$. Then \tilde{X}^δ is smaller than \tilde{W}^δ in the lower orthant order as the lower orthant order is closed under mixtures [15, Theorem 6.G.3 (e)]. It remains to show $\tilde{X}^\delta \xrightarrow{d} X$ and $\tilde{W}^\delta \xrightarrow{d} W$ since the lower orthant order is preserved under convergence in distribution [15, Theorem 6.G.3 (d)].

From the uniformization construction [11, Section 2.1], the process X^δ is a continuous time Markov chain on $\{0, 1\}^n$ with transition rates:

$$q_X^\delta(x, y) = \delta^{-1} \prod_{i=1}^n \left[(\delta\lambda_i(x))^{(1-x_i)(y_i-x_i)_+} (1 - \delta\lambda_i(x))^{(1-x_i)(1-(y_i-x_i)_+)} \times (\delta\mu_i(x))^{x_i(x_i-y_i)_+} (1 - \delta\mu_i(x))^{x_i(1-(x_i-y_i)_+)} \right],$$

for any $x, y \in \{0, 1\}^n$. As $\delta \rightarrow 0$, q_X^δ converges to the transition rates of (2.3) and since the state space is finite, this is sufficient to show $\tilde{X}^\delta \xrightarrow{d} X$.

The process $(N(\delta^{-1}t), \tilde{W}_t^\delta)$ is a continuous time Markov chain on $\mathbb{N}_0 \times \{0, 1\}^n$ with transitions $(m, w) \rightarrow (m + 1, w + u)$ for $u \in \{-1, 0, 1\}^n$ at rate

$$\beta_u(m, w) = \delta^{-1} \prod_{i=1}^n \left[(\delta\lambda_i(p_m^\delta))^{(1-w_i)(u_i)_+} (1 - \delta\lambda_i(p_m^\delta))^{(1-w_i)(1-(u_i)_+)} \times (\delta\mu_i(p_m^\delta))^{w_i(-u_i)_+} (1 - \delta\mu_i(p_m^\delta))^{w_i(1-(-u_i)_+)} \right],$$

where

$$p_{i,m+1}^\delta = \delta \lambda_i(p_m^\delta)(1 - p_{i,m}^\delta) + (1 - \delta \mu_i(p_m^\delta))p_{i,m}^\delta.$$

We can represent $(N(\delta^{-1}t), \widetilde{W}_t^\delta)$ as a random time change of Poisson processes [5, Chapter 6, Section 4]

$$\begin{aligned} \widetilde{W}_t^\delta &= X_0 + \sum_{u \in \{-1,0,1\}^n} u N_u \left(\int_0^t \beta_u(N(\delta^{-1}s), \widetilde{W}_s^\delta) ds \right) \\ N(\delta^{-1}t) &= \sum_{u \in \{-1,0,1\}^n} N_u \left(\int_0^t \beta_u(N(\delta^{-1}s), \widetilde{W}_s^\delta) ds \right), \end{aligned}$$

where the N_u are independent unit rate Poisson processes. The process $(W, t \in \mathbb{R}_+)$ can be constructed on the same probability space as $(N(\delta^{-1}t), \widetilde{W}_t^\delta)$ by representing $(W, t \in \mathbb{R}_+)$ as

$$W_{i,t} = X_{i,0} + N_i^+ \left(\int_0^t (1 - W_{i,s}) \lambda_i(p_s) ds \right) - N_i^- \left(\int_0^t W_{i,s} \mu_i(p_s) ds \right),$$

where p_t is the solution to (2.5) and identifying N_i^+ and N_i^- with the unit rate Poisson processes N_u such that $u_i = \pm 1$ and $u_j = 0$ for all $j \neq i$. We now use Gronwall's inequality to show $\mathbb{E}|W_{i,t} - \widetilde{W}_{i,t}^\delta| \rightarrow 0$ as $\delta \rightarrow 0$ for all t and all i , hence $\widetilde{W}^\delta \xrightarrow{d} W$. By the triangle inequality

$$\begin{aligned} \mathbb{E}|W_{i,t} - \widetilde{W}_{i,t}^\delta| &\leq \mathbb{E} \left| \int_0^t (1 - W_{i,s}) \lambda_i(p_s) ds - \int_0^t (1 - \widetilde{W}_{i,s}^\delta) \lambda_i(p_{N(\delta^{-1}s)}^\delta) ds \right| \\ &\quad + \mathbb{E} \left| \int_0^t W_{i,s} \mu_i(p_s) ds - \int_0^t \widetilde{W}_{i,s}^\delta \mu_i(p_{N(\delta^{-1}s)}^\delta) ds \right| \\ &\quad + \sum_{u: \|u\| \geq 2} \mathbb{E} \int_0^t \beta_u(N(\delta^{-1}s), \widetilde{W}^\delta(s)) ds, \end{aligned} \tag{3.5}$$

where $\|u\| = \sum_{i=1}^n |u_i|$. As the λ_i and μ_i are Lipschitz continuous, there exists constants C_1 and C_2 such that

$$\begin{aligned} &\mathbb{E} \left| \int_0^t (1 - W_{i,s}) \lambda_i(p_s) ds - \int_0^t (1 - \widetilde{W}_{i,s}^\delta) \lambda_i(p_{N(\delta^{-1}s)}^\delta) ds \right| \\ &\quad + \mathbb{E} \left| \int_0^t W_{i,s} \mu_i(p_s) ds - \int_0^t \widetilde{W}_{i,s}^\delta \mu_i(p_{N(\delta^{-1}s)}^\delta) ds \right| \\ &\leq C_1 \int_0^t \mathbb{E}|W_{i,s} - \widetilde{W}_{i,s}^\delta| ds + C_2 \int_0^t \mathbb{E}\|p_{N(\delta^{-1}s)}^\delta - p_s\| ds. \end{aligned} \tag{3.6}$$

The usual argument for proving convergence of Euler's method [4, Theorem 212A] shows that for any $t \geq 0$, there exists a constant C_3 such that

$$\|p_{N(\delta^{-1}t)}^\delta - p_t\| \leq \frac{e^{C_3 t} - 1}{C_3} |\delta N(\delta^{-1}t) - t|$$

so

$$\mathbb{E}\|p_{N(\delta^{-1}t)}^\delta - p_t\| \leq \delta t \frac{e^{C_3 t} - 1}{C_3}. \tag{3.7}$$

For any u such that $\sum_{i=1}^n |u_i| \geq 2$, there exists a constant C_4 such that

$$\beta_u(N(\delta^{-1}t), \widetilde{W}^\delta(t)) \leq C_4 \delta. \tag{3.8}$$

Combining (3.5) – (3.8) and applying Gronwall's inequality, we see that $\mathbb{E}|W_{i,t} - \widetilde{W}_{i,t}^\delta| \rightarrow 0$ as $\delta \rightarrow 0$ for all t and all i . \square

References

- [1] Linda J. S. Allen, *An introduction to stochastic epidemic models*, Mathematical epidemiology, Lecture Notes in Math., vol. 1945, Springer, Berlin, 2008, pp. 81–130. MR2428373
- [2] A. D. Barbour and M. J. Luczak, *Individual and patch behaviour in structured metapopulation models*, J. Math. Biol. **71** (2015), no. 3, 713–733. MR3382730
- [3] A. D. Barbour, R. McVinish, and P. K. Pollett, *Connecting deterministic and stochastic metapopulation models*, J. Math. Biol. **71** (2015), no. 6-7, 1481–1504. MR3419897
- [4] J. C. Butcher, *Numerical methods for ordinary differential equations*, third ed., John Wiley & Sons, Ltd., Chichester, 2016, With a foreword by J. M. Sanz-Serna. MR3559553
- [5] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 2005. MR838085
- [6] Raphaël Forien and Étienne Pardoux, *Household epidemic models and McKean-Vlasov Poisson driven stochastic differential equations*, Ann. Appl. Probab. **32** (2022), no. 2, 1210–1233. MR4414704
- [7] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, *Discrete-time markov chain approach to contact-based disease spreading in complex networks*, Europhys. Lett. EPL **89** (2010), no. 3, 38009.
- [8] Steve Hanneke, Wenjie Fu, and Eric P. Xing, *Discrete temporal models of social networks*, Electron. J. Stat. **4** (2010), 585–605. MR2660534
- [9] Ilkka Hanski, *A practical model of metapopulation dynamics*, J. Anim. Ecol. **63** (1994), no. 1, 151–162.
- [10] Liam Hodgkinson, Ross McVinish, and Philip K. Pollett, *Normal approximations for discrete-time occupancy processes*, Stochastic Process. Appl. **130** (2020), no. 10, 6414–6444. MR4140039
- [11] Julian Keilson, *Markov chain models—rarity and exponentiality*, Applied Mathematical Sciences, vol. 28, Springer-Verlag, New York-Berlin, 1979. MR528293
- [12] Thomas M. Liggett, *Interacting particle systems*, Classics in Mathematics, Springer-Verlag, Berlin, 2005, Reprint of the 1985 original. MR2108619
- [13] Ross McVinish and Liam Hodgkinson, *Fast approximate simulation of finite long-range spin systems*, Ann. Appl. Probab. **31** (2021), no. 3, 1443–1473. MR4278790
- [14] Dragoslav S. Mitrinović, Arlington M. Fink, and Josip E. Pečarić, *Inequalities involving functions and their integrals and derivatives*, Mathematics and its applications (East European series), 53, Kluwer Academic Publishers, Dordrecht; Boston, 1991. MR1190927
- [15] Moshe Shaked and J. George Shanthikumar, *Stochastic orders*, Springer Series in Statistics, Springer, New York, 2007. MR2265633
- [16] Péter L. Simon and Istvan Z. Kiss, *From exact stochastic to mean-field ODE models: a new approach to prove convergence results*, IMA J. Appl. Math. **78** (2013), no. 5, 945–964. MR3116158
- [17] Alain-Sol Sznitman, *Topics in propagation of chaos*, École d’Été de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math., vol. 1464, Springer, Berlin, 1991, pp. 165–251. MR1108185
- [18] Mohit Tawarmalani, Jean-Philippe P. Richard, and Chuanhui Xiong, *Explicit convex and concave envelopes through polyhedral subdivisions*, Math. Program. **138** (2013), no. 1-2, Ser. A, 531–577. MR3034816
- [19] Tadeusz Ważewski, *Systèmes des équations et des inégalités différentielles ordinaires aux deuxièmes membres monotones et leurs applications.*, Ann. Soc. Polon. Math. **23** (1950), 112–166. MR0040506
- [20] Robert R. Wilkinson, Frank G. Ball, and Kieran J. Sharkey, *The deterministic Kermack-McKendrick model bounds the general stochastic epidemic*, J. Appl. Probab. **53** (2016), no. 4, 1031–1040. MR3581239

- [21] Robert R. Wilkinson, Frank G. Ball, and Kieran J. Sharkey, *The relationships between message passing, pairwise, Kermack-McKendrick and stochastic SIR epidemic models*, J. Math. Biol. **75** (2017), no. 6-7, 1563–1590. MR3712322

Acknowledgments. We would like to thank the two reviewers for their helpful comments and suggestions.