Pair Approximations of Takeover Dynamics in Regular Population Structures

Joshua L. Payne   Joshua.Payne@uvm.edu

Department of Computer Science, The University of Vermont, Burlington,
Vermont, USA

Margaret J. Eppstein   Maggie.Eppstein@uvm.edu

Department of Computer Science, The University of Vermont, Burlington,
Vermont, USA

Abstract

In complex adaptive systems, the topological properties of the interaction network are strong governing influences on the rate of flow of information throughout the system. For example, in epidemiological models, the structure of the underlying contact network has a pronounced impact on the rate of spread of infectious disease throughout a population. Similarly, in evolutionary systems, the topology of potential mating interactions (i.e., population structure) affects the rate of flow of genetic information and therefore affects selective pressure. One commonly employed method for quantifying selective pressure in evolutionary algorithms is through the analysis of the dynamics with which a single favorable mutation spreads throughout the population (a.k.a. takeover time analysis). While models of takeover dynamics have been previously derived for several specific regular population structures, these models lack generality. In contrast, so-called pair approximations have been touted as a general technique for rapidly approximating the flow of information in spatially structured populations with a constant (or nearly constant) degree of nodal connectivities, such as in epidemiological and ecological studies. In this work, we reformulate takeover time analysis in terms of the well-known Susceptible-Infectious-Susceptible model of disease spread and adapt the pair approximation for takeover dynamics. Our results show that the pair approximation, as originally formulated, is insufficient for approximating pre-equibilibrium dynamics, since it does not properly account for the interaction between the size and shape of the local neighborhood and the population size. After parameterizing the pair approximation to account for these influences, we demonstrate that the resulting pair approximation can serve as a general and rapid approximator for takeover dynamics on a variety of spatially-explicit regular interaction topologies with varying population sizes and varying uptake and reversion probabilities. Strengths, limitations, and potential applications of the pair approximation to evolutionary computation are discussed.

Keywords

Interaction topologies, pair approximations, saturation dynamics, spatial structure, takeover time analysis.

1 Introduction

In complex adaptive systems, the topological properties of the interaction network are strong governing influences on the rate of flow of information throughout the

system. For example, in epidemiological models, the structure of the underlying contact network has a pronounced impact on the rate of spread of disease throughout a population (Keeling, 1999; Newman, 2002; Pastor-Satorras and Vespignani, 2001). Similarly, in evolutionary systems, the topology of potential mating interactions (i.e., population structure) affects the rate of flow of genetic information and therefore affects selective pressure (Giacobini, Tomassini, and Tettamanzi, 2005; Giacobini, Tomassini, Tettamanzi et al., 2005; Payne and Eppstein, 2007b, 2008; Rudolph, 2000; Sarma and De Jong, 1996). In classical mathematical epidemiology (Anderson and May, 1995), quantitative genetics (Falconer and Mackay, 1996) and canonical evolutionary algorithms (Holland, 1992), inter-individual interactions are typically assumed or allowed to be random. Whether studying the proliferation of infectious disease or the spread of an advantageous genetic mutation, such well-mixed (i.e., panmictic) contact networks facilitate the rapid propagation of information throughout the population. In contrast, when inter-individual interactions are spatially constrained, the rate of dissemination of information is significantly mitigated.

Most natural populations exhibit some form of spatial structure, and the important influence of the spatial scale of inter-individual interactions has thus become increasingly appreciated in many recent modeling efforts (e.g., Eppstein and Molofsky, 2007; Payne and Eppstein, 2007a; Rauch and Bar-Yam, 2006; Sayama et al., 2003; Werfel and Bar-Yam, 2004). For example, human populations have been shown to exhibit complex networks of interactions (e.g., Ebel et al., 2002; Liljeros et al., 2001; Watts and Strogatz, 1998). The assumption of panmixia in a model of disease spread may therefore limit the applicability of its results. Accordingly, several recent epidemiological models have employed more biologically realistic contact networks (Eames and Keeling, 2002; Keeling, 1999; Keeling and Eames, 2005; Meyers et al., 2006; Newman, 2002; Pastor-Satorras and Vespignani, 2001), and the topological characteristics of these networks have been shown to dramatically impact the dynamics of disease spread. Similarly, in evolutionary models, the spatial nature of inter-individual interactions significantly impacts the emergent dynamics. Imposing constraints on the spatial locality of interaction events has been shown to facilitate the emergence of evolutionary phenomena that would be otherwise impossible in well-mixed systems. For example, the maintenance of genetic diversity (Kerr et al., 2002; Sayama et al., 2003), the suppression of evolutionary pathologies (Altenberg, 2005), and the evolution of altruism (Matsuda et al., 1992; Van Baalen and Rand, 1998; Werfel and Bar-Yam, 2004) and cooperation (Hauert and Doebeli, 2004; Ohtsuki et al., 2006; Santos and Pacheco, 2005) have all been shown to be largely influenced by the spatial nature of the underlying contact network. Such spatially-explicit contact networks have also received an increasing amount of attention for use as population structures in evolutionary algorithms (Bryden et al., 2005; Giacobini, Tomassini, and Tettamanzi, 2005; Giacobini, Tomassini, Tettamanzi et al., 2005; Giacobini et al., 2006; Kirley and Stewart, 2007; Payne and Eppstein, 2006, 2007b, 2008; Rudolph, 2000; Sarma and De Jong, 1996; Sprave, 1999; Whitacre et al., 2008). For example, cellular evolutionary algorithms (cEAs), in which populations are structured on low-order regular graphs and mating events are restricted to occur in spatially localized, overlapping neighborhoods, can be used as a means of maintaining population diversity by mitigating selective pressure (Giacobini, Tomassini, Tettamanzi, et al., 2005).

One useful method for quantifying how the flow of information is influenced by a given interaction topology is through the analysis of *takeover time* (Goldberg and Deb, 1991). Takeover time is defined as the expected number of generations until a population consists entirely of copies of the best individual, starting from an initial

population that contains only one copy of the best individual. This analysis removes the confounding effects of variation operators, such as recombination and mutation, with selection acting as the only evolutionary operator. Higher takeover times imply lower selective pressure, and vice versa.

Takeover dynamics have been previously investigated for several specific interaction topologies. Goldberg and Deb (1991) analyzed takeover dynamics in panmictic population structures under a variety of selection mechanisms and showed that takeover is quite rapid in such well-mixed populations. Takeover dynamics in one-dimensional (1D) and two-dimensional (2D) cEAs have also received a large amount of attention. For example, Rudolph (2000) provided exact analytical solutions for ring topologies (1D toroidal lattices) and lower and upper bounds for array topologies (1D non-toroidal lattices). Sarma and De Jong (1996) analyzed 2D toroidal lattices with spatially localized neighborhoods of various shapes and sizes and showed that selection pressure is strongly influenced by the *radius* of the local mating neighborhood (see Equation (2) in Section 2). Under both synchronous and asynchronous updating policies, Giacobini, Tomassini, Tettamanzi, et al. (2005) provided mathematical models of takeover dynamics in 1D and 2D toroidal lattices (with von Neumann neighborhoods). The general result of these studies is that regular lattice population structures with localized interaction neighborhoods reduce selective pressure, relative to panmictic interaction topologies, and can thus enhance the exploratory power of evolutionary search. However, all of these models of takeover dynamics are specific to the particular interaction topologies for which they were designed. The hypergraph model (Sprave, 1999) is somewhat more general, and has been shown to model takeover dynamics in panmictic, metapopulation, and toroidal ring interaction topologies with reasonable accuracy. However, this model relies upon the success probabilities of selection operators that were derived (Chakraborty et al., 1997) under the assumption of panmixia and thus is not directly transferable to spatially structured populations with localized interactions. In addition, the hypergraph model requires computationally expensive calculations of recursive probabilistic formulations using possibly large and dense matrices. It would be useful to formulate a general method for estimating takeover dynamics that is not only broadly applicable, but also quickly computable, so that takeover dynamics could be rapidly predicted for a variety of regular population structures with varying selection and reversion probabilities.

Pair approximations (PAs) were originally derived as a statistical mechanics formulation to approximate equilibrium conditions in spatially structured biological populations, in order to determine conditions for the evolution of altruism (Matsuda et al., 1992). PAs are so named because they use differential equations to estimate the dynamics of states of neighboring *pairs* of individuals. This technique has gained popularity in theoretical biology, ecology, and epidemiology (Eames and Keeling, 2002; Hauert and Doebeli, 2004; Heibeler, 2000; Joo and Lebowitz, 2004; Keeling, 1999; Keeling and Eames, 2005; Ohtsuki et al., 2006; Petermann and De Los Rios, 2004; Satō et al., 1994; Satō and Iwasa, 2000; Van Baalen and Rand, 1998; Van Baalen, 2000) due to its purported generality and the rapidity with which the resulting analytical expressions can be solved. For example, Van Baalen and Rand (1998) used PAs to derive explicit conditions for the invasion of altruistic mutants into non-altruistic populations on a variety of population structures, including triangular and hexagonal lattices and random graphs. Hauert and Doebeli (2004) used PAs to estimate equilibrium proportions of cooperators in the snowdrift game on rectangular (with both von Neumann and Moore neighborhoods), triangular, and hexagonal lattices. Ohtsuki et al. (2006) formulated a PA of the spatially

extended prisoner's dilemma on regular graphs, in order to derive a simple rule for the evolution of cooperation. In most cases, the PA has been used to predict equilibrium frequencies, in accordance with their original intent. An exception to this is the work of Keeling (1999), where the PA was used to estimate the dynamics of disease spread for a single population size in a specific type of random graph, which was designed to have nearly constant nodal degree and spatially localized clustering. Another notable exception is the work of Petermann and De Los Rios (2004), in which the PA was extended and generalized to model higher order interactions between individuals and was used to estimate the saturation dynamics of infectious disease in random graphs and lattices with triangular and von Neumann neighborhoods.

While PAs have thus been applied to a wide variety of spatial topologies in the context of epidemiology, ecology, and evolutionary biology, each of the previous studies has been restricted to a single population size for a given population structure. To date, there has been no analysis of the accuracy of the PA as a function of the interaction between the topological characteristics of the local neighborhood structure and population size. And despite the potential advantages of general applicability and low computational effort, PAs have yet to be applied in the context of takeover dynamics in evolutionary algorithms.

Our intent is to investigate whether the PA can be used as a rapid and general method for approximating takeover dynamics of evolutionary algorithms using spatially explicit local interaction neighborhoods with regular topologies, for various population sizes and selection and reversion probabilities. First, we reformulate takeover time analysis for evolutionary algorithms in terms of the well-known *Susceptible-Infectious-Susceptible* (SIS) model of disease spread (Anderson and May, 1995; Keeling and Eames, 2005; Newman, 2002) and adapt the PA to predict takeover dynamics. We show that the PA, as originally formulated (Matsuda et al., 1992), over-predicts the rate of spread of advantageous alleles, since it does not properly account for the interaction between the size and shape of the local neighborhood and the population size. After parameterizing the PA to account for these influences, we demonstrate that the resulting PA is an efficient approximator of takeover dynamics on a variety of spatially-explicit regular interaction topologies. We discuss the strengths, limitations, and potential improvements to the PA, and suggest how this approach may be useful to practitioners of evolutionary computation.

2 Methods

2.1 Representing Population Structure as a Graph

The population structure of an evolutionary algorithm can be represented as a graph as follows. A graph, $G = (V, E)$, is defined as a nonempty finite set of vertices (V) and a finite set of edges (E) connecting these vertices. Each individual in the population is represented by a vertex $i \in V$, so that $|V| = \mu$, where μ is the population size. An undirected edge $\langle i, j \rangle$ is added to E for each individual j in the mating neighborhood of individual i, for all $i \in V$. Note that many commonly implemented population structures (Giacobini, Tomassini, Tettamanzi et al. 2005; Rudolph, 2000; Sarma and De Jong, 1996), including all of the population structures considered herein, are embedded in Cartesian space. However, it is important to note that adjacency in *Cartesian* space does not imply adjacency in the *interaction graph*, and vice versa. That is, two individuals that share an edge in the graph representation of the population structure are not necessarily

spatially proximal in Cartesian space. In this manuscript, the term *neighbor* is always used to mean adjacency in the interaction graph G.

2.2 Structural Properties of Graphs

When quantifying the structural properties of a graph, there are several metrics of potential interest (e.g., see Newman, 2003). In this section, we briefly define the structural properties considered in the current study. The so-called "clustering" metric (ϕ) of a graph G, stored as an adjacency matrix A, can be computed as the ratio of closed triangles to total triplets (Keeling, 1999), as follows:

$$\phi = \frac{\#triangles}{\#triplets} = \frac{trace(A^3)}{||A^2|| - trace(A^2)} \tag{1}$$

where the superscripts denote matrix exponentiation, $\|M\|$ denotes the sum of all of the elements in a matrix M, and *trace* denotes the sum of the elements on the main diagonal.

The *radius* of an interaction neighborhood N captures the level of dispersion present in that neighborhood (Sarma and De Jong, 1996). Selective pressure has been shown to increase as the ratio of the radius of the local neighborhood ($radius_N$) to the radius of the underlying lattice ($radius_G$) increases (Sarma and De Jong, 1996). For an interaction neighborhood N of size k (i.e., k is the degree of the nodes in the population structure) centered on a vertex located at $\langle x, y \rangle$ in Cartesian space, this metric is formally defined as

$$radius_N = \sqrt{\frac{1}{k+1}\left(\sum_{i=1}^{k+1}(x_i - \bar{x})^2 + \sum_{i=1}^{k+1}(y_i - \bar{y})^2\right)} \tag{2}$$

where

$$\bar{x} = \frac{1}{k+1}\sum_{i=1}^{k+1}x_i, \quad \bar{y} = \frac{1}{k+1}\sum_{i=1}^{k+1}y_i \tag{3}$$

and x_i and y_i are the Cartesian coordinates of the vertex $\langle x, y \rangle$ and the k vertices in its interaction neighborhood. The radius of the entire graph ($radius_G$) can be calculated using Equations (2) and (3) by assuming that the central vertex ($x = \frac{\sqrt{\mu}}{2}$, $y = \frac{\sqrt{\mu}}{2}$) of graph G is connected to every other vertex in the topology (i.e., $k = \mu - 1$). We calculate the ratio ρ of the radius of the interaction neighborhood $radius_N$ to the radius of the underlying population structure $radius_G$:

$$\rho = \frac{radius_N}{radius_G}. \tag{4}$$

Thus, ρ will take on a value of 1 only in the limiting case of a well-mixed system. Note that both the clustering coefficient (ϕ) and radius of an interaction neighborhood ($radius_N$) are independent of population size, while graph radius ($radius_G$), and thus also ρ, are both a function of population size.

Table 1: Naming conventions and schematic diagrams of the (a) rectangular and (b) triangular neighborhood structures considered in this study. (M7 and VN3 are not shown, but can be easily inferred.) Abbreviated names for each population structure are provided, as are the names as they appear in Sarma and De Jong (1996), where applicable (in italics). In (a), the links between vertices are implicit; each vertex in the interaction neighborhood (black circles) centered around a given vertex ($\times$) is connected to this center vertex. In (b), the links between individuals are shown explicitly (solid lines). For clarity, only one representative interaction neighborhood is shown for each type of graph.

(a)

M3	M5	VN1	VN2	MVN
Moore ($d = 3$), C9	Moore ($d = 5$)	von Neumann, L5	L9	C13

(b)

T	AT	ST
Triangular	Alternating Triangular	Semi Triangular
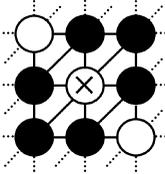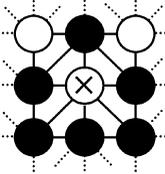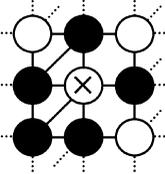
2.3 Population Structures

A regular graph is one in which every vertex has the same degree k. In this study, we investigate takeover dynamics on 10 distinct types of regular population structures, each based on 2D toroidal lattices, but with different local interaction neighborhoods. The naming conventions and corresponding schematic diagrams of the population structures are provided in Table 1, and their relevant structural characteristics are provided in Table 2a. Note that Table 2a presents the population structures in order of increasing radius; all subsequent tables and figures will follow this convention. For each population structure, we considered a total of eight population sizes $\mu \in \{576, 1024, 1600, 2304, 3136, 4096, 5184, 6400\}$, each structured on a ($\sqrt{\mu} \times \sqrt{\mu}$) node toroidal lattice. The corresponding graph radius for each population size is provided in Table 2b.

The local neighborhood structures considered in this study vary in both the number and the spatial locality of the individuals they contain, resulting in differing vertex degrees (k) and clustering (ϕ) characteristics (Table 2a). The topologies considered in this study include the most commonly implemented locally interacting population structures in the literature. For example, the von Neumann (*VN1*, Table 1a) and Moore

Table 2: (a) Structural metrics of the spatial interaction topologies considered in this study, presented in order of increasing radius of the local neighborhoods. (b) Graph radius as a function of population size on square lattices.

(a)

Population Structure	k	ϕ	$radius_N$
VN1	4	0	0.89
ST	5	0.3	0.97
AT	6	0.4	1.03
T	6	0.4	1.07
M3	8	0.43	1.16
MVN	12	0.46	1.47
VN2	8	0.21	1.49
M5	24	0.52	2.00
VN3	12	0.27	2.08
M7	48	0.54	2.83

(b)

Population Size (μ)	$radius_G$
576	9.79
1024	13.06
1600	16.32
2304	19.59
3136	22.86
4096	26.13
5184	29.40
6400	32.66

neighborhoods (*M3*, Table 1a) are the most frequently employed interaction networks in spatially structured evolutionary algorithms (e.g., Giacobini, Tomassini, Tettamanzi, et al., 2005; Sarma and De Jong, 1996). We also investigated common variations of these with larger interaction neighborhoods; specifically, Moore neighborhoods with diameter 5 (*M5*, Table 1a) and 7 (*M7*, not shown) and extended von Neumann neighborhoods (*VN2*, Table 1a, and *VN3*, not shown). The triangular population structures considered herein (*T*, *AT*, and *ST*, Table 1b) are commonly employed in various models of ecological (Van Baalen, 2000), evolutionary (Hauert and Doebeli, 2004; Van Baalen and Rand, 1998), and physical systems (Ong and Cava, 2004).

2.4 Takeover Time

Consider a population with only two levels of fitness; that is, let $\Lambda_i(t)$ be the fitness value of vertex $i \in V$ at time t, where $\Lambda_i(t) \in \{0, 1\}$ and 1 is more fit than 0. In the initial population, $\Lambda_i(0) = 1$ for exactly one $i \in V$ and $\Lambda_j(0) = 0 \forall j \neq i \in V$. Let N_t denote the *proportion* of nodes with value 1 at time t:

$$N_t = \frac{1}{|V|} \sum_{\forall i \in V} \Lambda_i(t) \qquad (5)$$

Following Rudolph (2000), we define the takeover time $T = \min\{t \mid N_t = 1\}$ to be the minimum number of generations such that copies of the most fit individual fully saturate the entire population, starting with only one such individual in the initial population. This definition of takeover time thus assumes that N_t can never decrease.

$E_i[T]$ is defined as the empirical estimate of the expected takeover time given that the initial best individual is located in vertex i. Thus, the overall empirically estimated expected takeover time of a beneficial mutation, averaged over all potential initial conditions, is simply

$$E[T] = \frac{1}{|V|} \sum_{\forall i \in V} E_i[T] \tag{6}$$

assuming that the initial best individual is equally likely to appear in any given node.

2.5 Selection

In this study, we adopt a simple "replace if better" selection mechanism (a.k.a. uniform selection, Gorges-Schleuter, 1999), where nodes are updated synchronously, as follows. For each node $i \in V$, a node j is selected at random with uniform probability from the mating neighborhood of node i, with neighborhood size k. Thus, if there are x nodes containing the fittest value in the mating neighborhood of node i, then the probability of selecting one of them (P_{sel}) is simply

$$P_{sel} = \frac{x}{k}. \tag{7}$$

With uptake probability p_{up}, the value of the selected node j then replaces the value of node i if $\Lambda_j(t) > \Lambda_i(t)$. Therefore, the probability of a high fitness individual replacing a given node (P_{rep}) is given by the product of the uptake probability (p_{up}) and the selection probability (P_{sel}):

$$P_{rep} = p_{up} P_{sel}. \tag{8}$$

Decreasing p_{up} serves as a simple means for decreasing selective pressure in models of takeover dynamics (Rudolph, 2000), although $p_{up} < 1$ is nonstandard in evolutionary algorithms.

2.6 Reversion

While the definition of takeover time assumes that N_t can never decrease, this nonextinction assumption can be relaxed by allowing vertices of value 1 to revert back to 0 with some probability g (Rudolph, 2001). This reversion probability is analogous to mutation in evolutionary systems, where genetic information is occasionally lost, or to recovery from infection in models of disease spread, where infected individuals may recover and either become immune or again become susceptible.

In all experiments performed herein, reversion of high fitness individuals occurred, with probability g, at the end of each generation, after the population had been updated with the selection mechanism. Thus, the probability of a given high fitness individual

reverting to low fitness was independent of the fitness values of the other individuals in its mating neighborhood.

2.7 Reformulating Takeover Time Analysis in Terms of the SIS Model

In the SIS model, a population of μ individuals is compartmentalized into two discrete states: susceptible (S) and infected (I). This model evolves according to the following transition rules

$$S \xrightarrow{p_{up}} I \xrightarrow{g} S \tag{9}$$

where p_{up} governs the rate at which infection occurs and g governs the rate of recovery from infection. Once a node has recovered, it again becomes susceptible. In a spatially structured population, the transmissibility of disease across a connection is (Keeling, 1999)

$$\tau = \frac{p_{up}}{k}. \tag{10}$$

The SIS model and our model of takeover dynamics are thus equivalent. The susceptible state (S) corresponds to low fitness individuals (with fitness 0) and the infected state (I) corresponds to high fitness individuals (with fitness 1). The probability of a given node becoming infected (P_{inf}) is given by the product of the number of infected neighbors (x) in its contact neighborhood and the probability of disease transmission across a connection (τ):

$$P_{inf} = \tau x. \tag{11}$$

By substituting Equation (10) into Equation (11) and Equation (7) into Equation (8), it is clear that the probability of infection in the SIS model (P_{inf}, Equation (11)) is equivalent to the replacement probability in our formulation of takeover time analysis (P_{rep}, Equation (8)). Lastly, the reversion probability (g) corresponds to the probability of mutation back from high to low fitness. Note that when there is no reversion ($g = 0$), the system behaves as an SI model, or "contact process" (Satō and Iwasa, 2000). Despite the clear relationship between these two classes of models, this is the first time, to the best of our knowledge, that this correspondence has been explicitly made.

2.8 Pair Approximations

Instead of estimating the dynamics of the states of *nodes* in a contact network, pair approximations (PAs) estimate the dynamics of states of neighboring *pairs* of nodes. By capturing the correlations between pairs of vertices, some aspects of the structure of the interaction topology can be accounted for.

The PA works as follows. Consider a population of size μ structured on an interaction topology wherein every node has k neighbors. (It is important to note that PAs assume that the underlying contact network is regular, or at least possesses a well defined average degree k.) Following Keeling (1999), let $[X]$ denote the number of nodes in state X, $[XY]$ denote the number of pairs of connected nodes in state XY, and $[XYZ]$ denote the number of connected triplets of nodes in state XYZ, such that XY pairs are always counted once in each direction (i.e., $[XY] = [YX]$) and XX pairs are counted

twice (i.e., $[XX]$ is always even). PAs work by tracking the changes in the numbers of all possible combinations of pairs $[XY]$. Since the interaction topology is regular with constant degree k, the number of singles can always be recovered from the number of pairs, as follows (Keeling, 1999):

$$[X] = \frac{1}{k} \sum_W [XW].$$ (12)

However, the rates of change in the number of pairs depend upon the numbers of configurations larger than pairs, such as triplets, and this information is not maintained by the PA. Even if the number of triplets were maintained, the rates of change in the number of triplets would similarly depend upon the numbers of quadruplets, and so on for larger connected motifs. Thus, in order to estimate the dynamics in terms of the numbers of pairs, the numbers of configurations larger than pairs must be approximated to some degree of accuracy. This is referred to as "closing" the system (Keeling, 1999; Matsuda et al., 1992; Van Baalen, 2000).

The simplest closure strategy (Keeling, 1999; Satō and Iwasa, 2000) is to assume that the nodes at the ends of triplets are not connected to one another (i.e., that triplets are linear, not triangular). Under this assumption, the number of triplets $[XYZ]$ can be approximated as (Keeling, 1999):

$$[XYZ] \approx \frac{(k-1)[XY][YZ]}{\sum_W [YW]} = \frac{(k-1)}{k} \frac{[XY][YZ]}{[Y]}.$$ (13)

While this assumption closes the system at the level of pairs, it can introduce a significant amount of error, since it ignores all of the spatial structure beyond pairwise interactions and therefore neglects possible correlations between the nodes at the ends of triplets. For example, consider the takeover dynamics of a population structured on a square 2D lattice with 3×3 (Moore) interaction neighborhoods (Table 1a, M3). In the early stages of the dynamics, only a few of the most fit individuals (with state 1) are present in the topology, and they are propagating locally into a sea of less fit individuals (with state 0). Under the closure assumption of Equation (13), [101] (i.e., the number of triplets in state 101) would be approximated by $((k-1)/k)([10][01]/[0])$. Both [10] and [01] are computed globally, and so can be expected to be quite small (and [0] quite large) during the early stages of the takeover dynamics, when there are only a few 1's in the topology. Consequently, the approximation of [101] by Equation (13) will also be very small. However, since fit individuals are spreading only locally, the true value of [101] will be larger than the number estimated by Equation (13), because the 1's that are present in the graph are in the same local region. This error is especially pronounced if the interaction topology has a preponderance of triangular paths (e.g., M3), in which case it is not safe to assume that the distant ends of a triplet are not connected.

In order to more accurately estimate the number of triplets $[XYZ]$, one can explicitly take into consideration the proportion of triplets in the interaction topology that form closed triangles (Keeling, 1999). The ratio of closed triangles to total triplets (ϕ, Equation (1), also known as the clustering coefficient) can be incorporated directly into the closure

method, as follows (Keeling, 1999):

$$[XYZ] \approx \frac{(k-1)}{k} \frac{[XY][YZ]}{[Y]} \left((1-\phi) + \frac{\phi\mu}{k} \frac{[XZ]}{[X][Z]} \right). \qquad (14)$$

Thus, the closure method of Equation (14) captures the correlation between nodes at the opposing ends of a triplet in proportion to the ratio of the number of closed triangles to total triplets inherent in the underlying population structure. Note that when there is no clustering ($\phi = 0$), the closure method of Equation (14) reduces to Equation (13). Spatial structure beyond triplets, however, is not considered in Equation (14).

2.9   Estimating Takeover Dynamics with Pair Approximations

In this section, we develop a PA of takeover dynamics by modifying the *Susceptible-Infectious-Recovered* (SIR) PA proposed by Keeling (1999). In the SIR model each vertex has three potential states, and the PA developed by Keeling (1999) thus requires a total of five coupled differential equations. However, in the reformulation of takeover time analysis in terms of the SIS model (Equation (9)) each vertex has only two potential states: 1 for high fitness individuals and 0 for low fitness individuals. With a binary state space, there exist four distinct types of pairs; due to symmetry, the following three differential equations suffice:

$$\frac{d[00]}{dt} = c(-\tau[001] + g[01] + g^2[11])$$

$$\frac{d[01]}{dt} = 0.5c(\tau([001] - [101] - [01]) + g([11] - [01]))$$

$$\frac{d[11]}{dt} = c(\tau([101] + [01]) - g[11] - g^2[11]) \qquad (15)$$

where the additional factor of 0.5 in the second equation accounts for the symmetry between [01] and [10]. In the original derivation of the PA (Matsuda et al., 1992), designed for the prediction of population densities at equilibrium, the value of the coefficient c was proposed to be the constant 2, and this value has since been used in numerous other studies (e.g., Keeling, 1999; Petermann and De Los Rios, 2004; Satō et al., 1994; Van Baalen, 2000). In fact, it is trivial to show that equilibrium conditions are independent of the particular choice of the coefficient c. However, if one is to use the PA to approximate *pre-equilbrium dynamics*, then the value of c becomes important. As will be shown in Section 3.1, our results clearly demonstrate that the optimal choice of c is a function of ρ (Equation (4)), which depends on both the local neighborhood of interactions and the population size.

   In this study, all occurrences of $[XYZ]$ in Equation (15) are computed using Equation (14), τ from Equation (10), and the parameterized coefficient c from Equation (24), which we present in Section 3.1. Thus, the PA employed herein is parameterized by the population size (μ), the ratio (ρ, Equation (4)) of *radius$_N$* to *radius$_G$*, the vertex degree (k), the ratio of closed triplets to total triangles (ϕ, Equation (1)), the reversion probability (g), and the uptake probability (p_{up}). For small reversion probabilities the effect of g^2 in Equation (15) is negligible, so this term is often ignored (e.g., Keeling, 1999). However, since we wanted to test the accuracy of the PA as a function of g, the inclusion of the g^2 term was necessary.

2.10   Experimental Design

For each combination of the 10 population structures (Table 2a) and eight population sizes (Table 2b), takeover dynamics were observed by placing a single copy of the best individual in only one node and then observing the rate with which this advantageous allele spreads through the population. Since each of the 10 population structures considered in this study are regular, the takeover dynamics are unaffected by the placement of the initial copy of the high fitness individual—in sharp contrast to irregular spatial structures (e.g., Giacobini, Tomassini, and Tettamanzi, 2005; Payne and Eppstein, 2007b, 2008). For each of the 80 distinct combinations of population structure and population size, 50 such simulations were performed and averaged, in order to mitigate the stochasticity inherent in the selection and reversion policies. When $g = 0$, the simulations were carried out until saturation occurred; when $g > 0$, simulations were performed for a total of 500 generations, long after the saturation sill had been reached.

Takeover dynamics were also approximated by solving the coupled differential equations of the PA (Equation (15)) via numerical integration, using a Runga-Kutta method with adaptive step size (Matlab's ode45 function). For each combination of population size and population structure, we considered $p_{up} \in \{0.5, 0.75, 1\}$ (which spans the range of feasible selection probabilities in evolutionary algorithms) and $g \in \{0, 0.05, 0.1\}$ (which are common mutation rates in evolutionary algorithms).

2.11   Assessing Error

Since we are interested in using the PA to estimate pre-equilibrium takeover dynamics, we employ several error metrics in order to assess the accuracy of the approximation. All error metrics are represented as percentages, relative to the values observed through direct simulation. To clarify the presentation of these error metrics, we depict representative curves for takeover dynamics predicted by the PA (Figure 1, dashed black line) and observed through direct simulation (Figure 1, solid black line) for the *VN2* population structure with $p_{up} = 1$, $g = 0$, and $\mu = 1024$. While these curves are later described in the results section, they are provided here as a visual aid to help elucidate the definition of the error metrics. Throughout the rest of this manuscript, *Sim(t)* refers to the simulation data for N_t, while *PA(t)* refers to the pair approximation of N_t.

Since the PA is a continuous approximation of discrete system dynamics, N_t only asymptotically approaches 1 (when $g = 0$). Consequently, the expected takeover time of the simulation data (Figure 1, $E[T]_{Sim}$) and the PA (Figure 1, $E[T]_{PA}$) were calculated as the first generation in which N_t was within 1% of the maximum saturation value of 1. When $g > 0$, $E[T]_{Sim}$ and $E[T]_{PA}$ are technically undefined, since complete takeover will never occur. In this case, we use $E[T]_{Sim}$ and $E[T]_{PA}$ to denote the first generation in which N_t is within 1% of its respective saturation sill. The values of the saturation sills were estimated for both the PA and the simulation data by averaging N_t over generations 300 through 500 (long after saturation had occurred).

The first error metric, *area error*, quantifies the area between the predicted and observed takeover curves, as a percentage of the total area under the observed takeover curve. (Area errors were thus normalized by the total area under the simulation curve A_{Sim}, in order to provide a relative error metric that is independent of takeover time.) A portion of the area error is negative (Figure 1, A^-) where the PA is predicting a slower spread of the high fitness individuals than that observed through direct simulation, and a portion is positive (Figure 1, A^+) where the PA is predicting a more rapid spread.
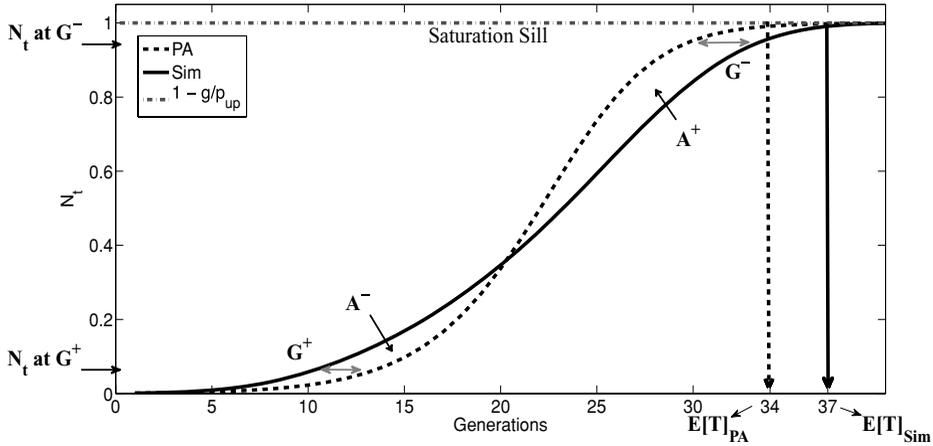
Figure 1: Error metrics used in this study are depicted in a schematic diagram of representative takeover curves approximated by the PA (dashed black lines) and observed as the average of 50 direct simulations (Sim, solid black lines). G represents generation error, A represents area error, and $E[T]$ represents expected takeover time. See text for details.

Specifically, if t_1 denotes the number of generations when the two curves intersect and t_2 denotes the maximum of $E[T]_{Sim}$ and $E[T]_{PA}$, then

$$A_{Sim} = \int_{t=0}^{t_2} Sim(t)\, dt \tag{16}$$

$$A^- = \int_{t=0}^{t_1} PA(t)\, dt - \int_{t=0}^{t_1} Sim(t)\, dt \tag{17}$$

$$A^+ = \int_{t=1}^{t_2} PA(t)\, dt - \int_{t=1}^{t_2} Sim(t)\, dt \tag{18}$$

and the area error metrics are defined as:

$$NegativeAreaError = \frac{A^-}{A_{Sim}} \times 100\% \tag{19}$$

$$PositiveAreaError = \frac{A^+}{A_{Sim}} \times 100\%. \tag{20}$$

After fitting spline curves to both the PA and simulation data, Equations (16–18) were computed using the trapezoidal rule at 1000 uniformly spaced points on the spline curves, between the lower and upper bounds of each integration. The total absolute area error was then computed as:

$$|Area\ Error| = |NegativeAreaError| + |PositiveAreaError|. \tag{21}$$

Generation error quantifies the maximum generation difference (in either direction), for the same degree of saturation N_t, between the takeover dynamics predicted by the PA

and as observed through simulation, and is given as a percentage of the takeover time observed through simulation ($E[T]_{Sim}$). Generation error is more robust for assessing dynamics than simply assessing error in takeover time, since it is assessed over the entire pre-equilibrium range. These discrepancies take on a positive value (Figure 1, G^+) when the PA is predicting a slower spread of the high fitness individuals than that observed in the simulation data and negative value (Figure 1, G^-) when the PA is predicting a more rapid spread. Generation error is defined as the maximum of the absolute values of G^- and G^+ and is normalized by the simulation data, as follows,

$$Generation\ Error = \frac{max\{|G^-|, G^+\}}{E[T]_{Sim}} \times 100\%. \qquad (22)$$

When $g > 0$, we also assess (a) the *percentage of unsuccessful introductions*, out of the 50 trials, where unsuccessful introductions occurred when all high fitness individuals disappeared, due to reversion, in the first few generations, and (b) the *Sill Error*, which is the error in the proportion of high fitness individuals that are predicted by the PA to exist at saturation if the introduction was successful. As before, the values of the saturation sills for the PA (S_{PA}) and the simulation data (S_{Sim}) were calculated as the average value of N_t over generations 300 to 500 of the PA and simulation data, respectively. The sill error was then normalized by the sill of the simulation data, as follows,

$$Sill\ Error = \frac{S_{PA} - S_{Sim}}{S_{Sim}} \times 100\%. \qquad (23)$$

Thus, the sill error is negative when the PA predicts a smaller sill than that observed through simulation, and is positive otherwise.

3   Results

3.1   Accounting for ρ in the PA

In order to determine the values of c needed for Equation (15), we varied the coefficient from 0.1 to 2.0, in increments of 0.01, for each combination of population structure and population size (using $p_{up} = 1$ and $g = 0$), and then selected the value that minimized generation error (Equation (22)) as the empirically determined optimal value for c. In Figure 2, we show these optimal values for c, as a function of ρ, along with the best fit logarithmic curve. The equation for the best fit curve is

$$c \approx 1.82 + 0.9 \log_{10}(\rho) \qquad (24)$$

with $R^2 = 0.88$, and a maximum residual of 0.19 (which occurred with neighborhood structure T, $\mu = 576$). Our intent is to try to create a rapid and general method which can be applied to arbitrary regular population structures of arbitrary size. If one had to perform extensive simulations in order to estimate the optimal coefficient c before the PA could even be applied, this would defeat the purpose. Thus, Equation (24) was used in all subsequent experiments (rather than the optimally determined values of c) to assess the accuracy of the PA (Equation (15)) with the c parameterized by ρ. The value of $c = 2$, used in previous studies, is shown for reference (Figure 2, horizontal dash-dot line); in all cases, the optimal coefficient values for c were well below 2.
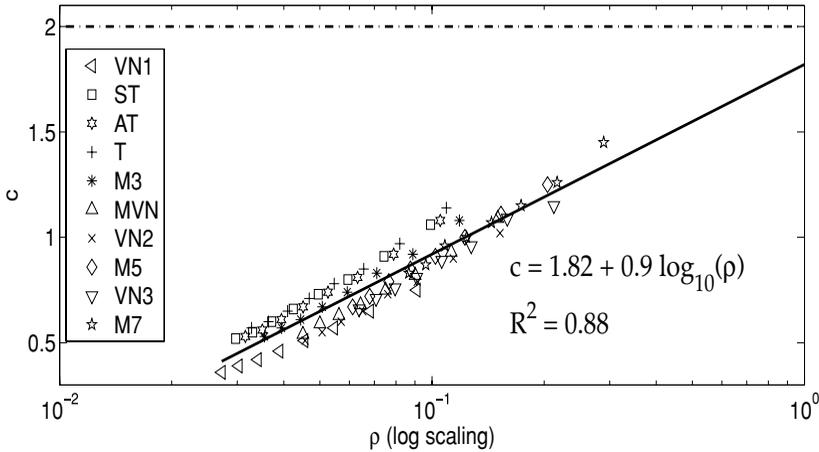
Figure 2: The values of c that minimized generation error as a function of ρ, along with the best logarithmic fit. The dash-dot horizontal line, representing $c = 2$, is provided as reference.

Sarma and De Jong (1996) found that the rate of saturation increased logarithmically with the ratio of local neighborhood radius to grid radius. Thus, it makes sense that the coefficient c must increase logarithmically with ρ, in order to compensate for the decreased takeover times that occur with larger ρ. *A key finding of this work is that, if one is to use the PA to approximate dynamics, the PA must be parameterized to account for the interaction between the size and shape of the local neighborhood and the population size.*

3.2 Non-Extinctive Dynamics

We first consider the nonextinctive case, where $g = 0$. For all combinations of population size and population structure considered, takeover time was found to decrease as the radius of the local neighborhood increased, in concurrence with the results of Sarma and De Jong (1996). Figure 3 shows takeover dynamics predicted by the PA (dashed black lines) and observed by direct simulation (solid black lines) on six representative regular population structures with $p_{up} = 1$ and $\mu = 1024$. When plotting simulation results here and elsewhere in this paper, we depict the proportion of nodes (N_t) containing maximum fitness at generation t, averaged over all 50 independent simulations on that combination of graph type and population size. The scale of the horizontal axis varies for each topology in order to best elucidate the discrepancies between the approximation and simulation data, since this is the relevant measure of accuracy in this case (as opposed to a comparison of takeover dynamics between topologies, in which case the scale of the horizontal axes would be held constant). Error metrics for all graph types and population sizes are quantified in the subsequent sections, but first we offer some general observations.

For the PA, all takeover curves are sigmoidal, exhibiting exponential growth ($R^2 >$ 0.95 for an exponential fit below the inflection point, for all PA curves) followed by saturation. In contrast, takeover curves on 2D lattice topologies with local interaction neighborhoods are known to be polynomial (Gorges-Schleuter, 1999; Giacobini, Tomassini, and Tettamanzi, 2005) below the inflection point (polynomial exponent varied from 1.9 to 2.8, $R^2 >$ 0.96, for all topologies and population sizes considered herein, where the
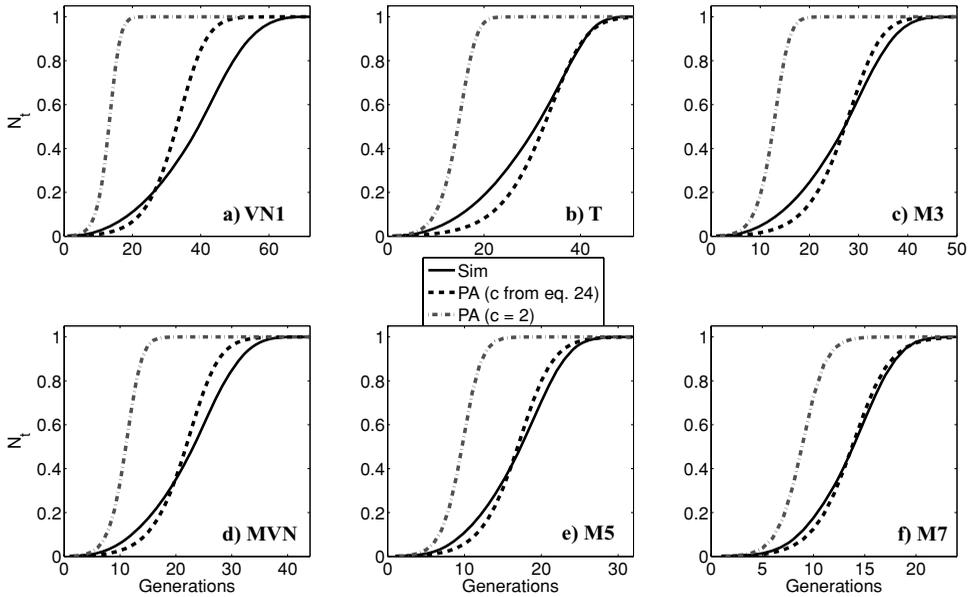
Figure 3: Takeover dynamics as predicted by the PA using Equation (24) (dashed black lines) and as observed through direct simulation (solid black lines) on six regular population structures using $\mu = 1024$, $p_{up} = 1$, and $g = 0$. For reference, we also show the dynamics predicted by the PA using $c = 2$ (dash-dot lines). The legend and vertical axis applies to all panels. Note the change in scale among the horizontal axes of each panel.

exponent increased linearly with increasing $radius_N$, $R^2 > 0.96$ for all population sizes). Thus, the PA is not an actual mechanistic model of the governing dynamics of this system, but is more appropriately characterized as an approximation, and all PA curves were statistically different from the corresponding simulation curves ($p < .001$, χ^2). In general, the slope of the PA tends to increase too slowly early in the growth phase and later, too rapidly, as compared to the simulation data. However, in most cases the resulting curves are in reasonably good agreement with the data observed through simulation (Figure 3c–f). (In contrast, the curves resulting from the PA using $c = 2$, shown in Figure 3 as dash-dot lines, always dramatically overestimate the rate of spread of the advantageous allele.)

If the optimal c were used, the *Sim(t)* and *PA(t)* curves would intersect at the inflection point. However, the use of Equation (24) to estimate c introduces a small amount of error, sometimes causing the curves to intersect a little too early (as for *VN1*, Figure 3a), or a little too late (as for T, Figure 3b), thus resulting in an over- or underestimate of the rate of takeover, respectively. Other errors are due to simplifications in the PA itself, causing the PA to be least accurate on the *VN1* topology, and most accurate for populations with large neighborhood radii and small population sizes. These results are quantified in the next section and discussed in Section 4.

3.2.1 Area Error and Generation Error

Both area error (Figure 4a) and generation error (Figure 4b) were found to decrease approximately exponentially as a function of ρ. The average goodness of fit was $R^2 =$
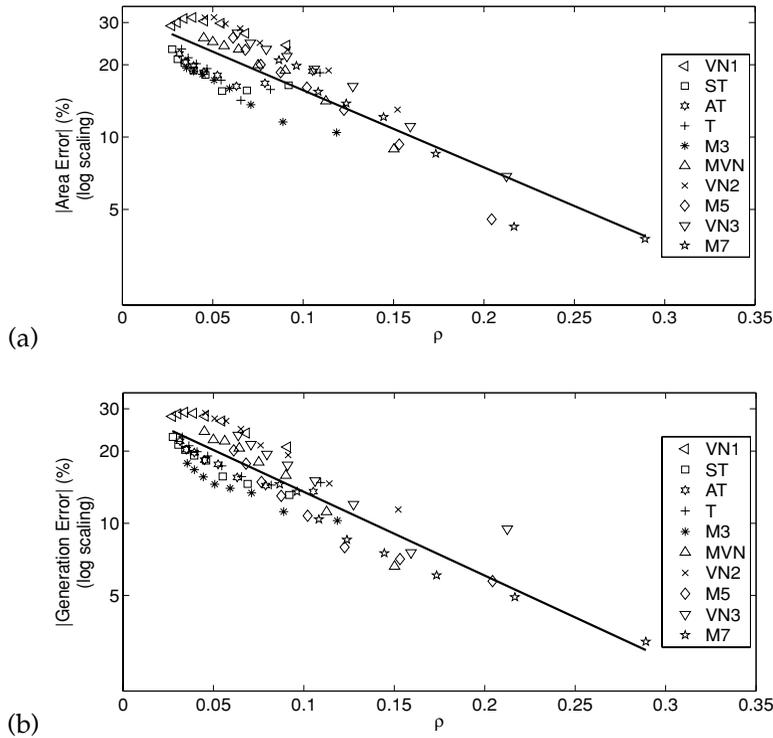
Figure 4: (a) Area error and (b) generation error as a function of ρ for all population sizes and population structures, with $p_{up} = 1$ and $g = 0$. The solid line in each figure depicts the best exponential fit to the data for all population structures, and is provided as a visual aid only. Note the logarithmic scale on the vertical axis.

0.73 between the logarithm of area error and ρ and $R^2 = 0.79$ between the logarithm of generation error and ρ. In both cases, the decay coefficient varied with population structure, as can be seen from the individual symbol types shown in Figure 4. The maximum area error (31%) was observed on the *VN2* population structure with $\mu = 5184$, the minimum (5%) was observed on the *M7* population structure with $\mu = 1024$, and the overall average area error was 19%. The maximum generation error (29%) was observed on the *VN1* population structure with $\mu = 4096$, the minimum generation error (3.2%) was observed on the *M7* population structure with $\mu = 576$, and the average generation error was 17%. Since area error was highly correlated to generation error across all population structures, population sizes, and uptake/reversion probabilities ($R^2 > 0.9$), we present generation error, but not area error, in our subsequent results.

The exponential decrease of generation error as ρ increases has the ironic implication that the PA, even after the formulation has been adjusted to compensate for ρ, becomes more accurate for estimating saturation dynamics as the system approaches the well-mixed case. This finding also has the important implication that for any given local interaction neighborhood, the accuracy of the PA depends heavily upon population size (μ), with smaller population sizes yielding more accurate results. To better elucidate the dependence of the accuracy of the PA on population size, Figure 5 depicts the
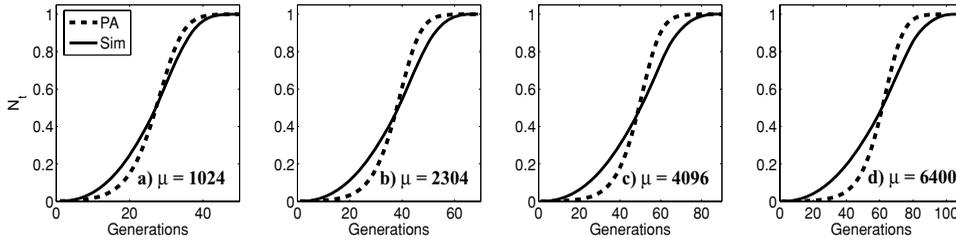
Figure 5: Influence of population size on the accuracy of the PA. Takeover dynamics as estimated by the PA (dashed line) and as observed through direct simulation (solid line) on the *M3* population structure, with a population size of (a) $\mu = 1024$, (b) $\mu = 2304$, (c) $\mu = 4096$, and (d) $\mu = 6400$.

takeover dynamics observed on the *M3* population structure as population size is increased from $\mu = 1024$ to $\mu = 6400$. As population size increases, the discrepancy between the PA and simulation curves becomes more pronounced, with generation error increasing from 10% (Figure 5a) to 18% (Figure 5d). This results from the increase in the number of generations required for complete saturation to occur as population size increases, providing the exponential trend of the PA curve more time to diverge from the polynomial trend of the simulation curve.

3.2.2 Relaxing Selection Pressure, $p_{up} < 1$

For a given combination of population structure and population size, takeover time was found to increase in proportion to p_{up}^{-1}, shifting both the PA and simulation curves to the right as p_{up} decreases (Figure 6). Decreasing p_{up} causes the PA to predict a slower spread of the high fitness allele than that observed through direct simulation (as explicitly shown for two graph types, *VN2* and *M3*, in Figures 6a–c and 6d–f, respectively), shifting the PA curve more to the right, relative to the curve observed through simulation (e.g., compare Figures 6a and 6c). Thus, decreasing p_{up} causes an improvement in the accuracy of the PA for combinations of population structure/population size in which $PA(t)$ intersected $Sim(t)$ before the inflection point when $p_{up} = 1$ (such as *VN2*, $\mu = 1024$, Figure 6a–c); and it causes a degradation in accuracy for populations in which $PA(t)$ intersected $Sim(t)$ at or after the inflection point $p_{up} = 1$ (such as *M3*, $\mu = 1024$, Figure 6d–f). Decreasing p_{up} from 1 to 0.5 resulted in a maximum improvement in generation error of 11%, observed on the *VN2* population structure with $\mu = 6400$, and a maximum degradation in accuracy of 12%, observed on the *M7* population structure with $\mu = 1024$.

3.3 Relaxing the Non-Extinction Assumption, $g > 0$

We now consider the case in which reversion is possible, where high fitness individuals can revert back to low fitness with probability g.

3.3.1 Percentage of Unsuccessful Introductions

Table 3 shows the percentage of all 50 trials with unsuccessful introductions, using $p_{up} \in \{0.5, 0.75, 1\}$ and $g \in \{0.05, 0.1\}$. Since the percentage of unsuccessful introductions was unaffected by population size, the data presented in Table 3 depict the average of the percentages observed on all eight population sizes. Note that data for $g = 0$ are
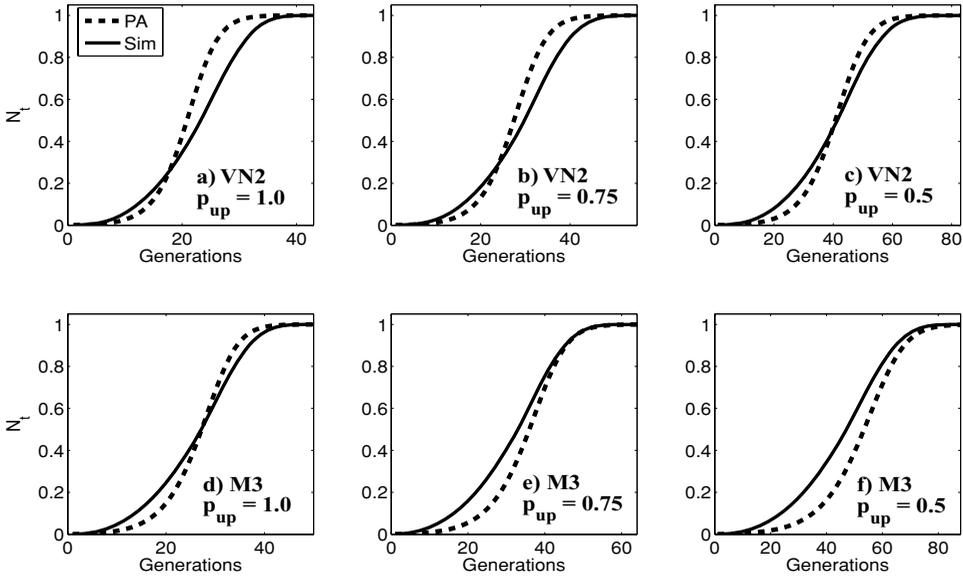
Figure 6: Influence of uptake probabilities p_{up} on the accuracy of the PA. Representative takeover dynamics as estimated by the PA (dashed line) and as observed through direct simulation (solid line) on the (a–c) *VN2* and (d–f) *M3* population structures with $\mu = 1024$, $g = 0$. The three plots in each row correspond to $p_{up} \in \{1, 0.75, 0.5\}$, from left to right. The legend and vertical axis applies to all panels. Note the change in scale among the horizontal axes of each panel.

Table 3: Percentage of unsuccessful introductions as a function of $p_{up} \in \{1, 0.75, 0.5\}$ and $g \in \{0.05, 0.1\}$. The data represent the percentage out of 50 independent trials on each graph type, averaged over all population sizes. Note that the data for $g = 0$ are not displayed, since an unsuccessful introduction is impossible in this case.

Population Structure	Unsuccessful Introductions (%)					
	$p_{up} = 1$		$p_{up} = 0.75$		$p_{up} = 0.5$	
	$g = 0.05$	$g = 0.1$	$g = 0.05$	$g = 0.1$	$g = 0.05$	$g = 0.1$
VN1	2.50	5.00	6.25	15.25	12.00	26.50
ST	2.75	4.25	4.25	11.75	8.75	16.50
AT	4.50	6.25	5.50	12.75	7.00	18.75
T	2.75	10.25	5.25	8.25	7.25	16.50
M3	2.00	7.50	4.00	10.00	8.75	19.75
MVN	2.25	9.00	6.50	8.75	9.00	21.75
VN2	3.00	5.50	4.50	11.75	10.25	21.50
M5	1.25	6.75	4.25	13.00	8.00	18.00
VN3	3.00	6.25	5.50	11.00	6.50	20.75
M7	2.25	7.00	4.75	12.75	6.25	17.00

not displayed, as extinction is not possible in this case. As expected, increasing g or decreasing p_{up} generally increased the number of unsuccessful introductions. However, no obvious relationship was observed between introduction success and neighborhood radius ($radius_N$), degree (k), or ρ.
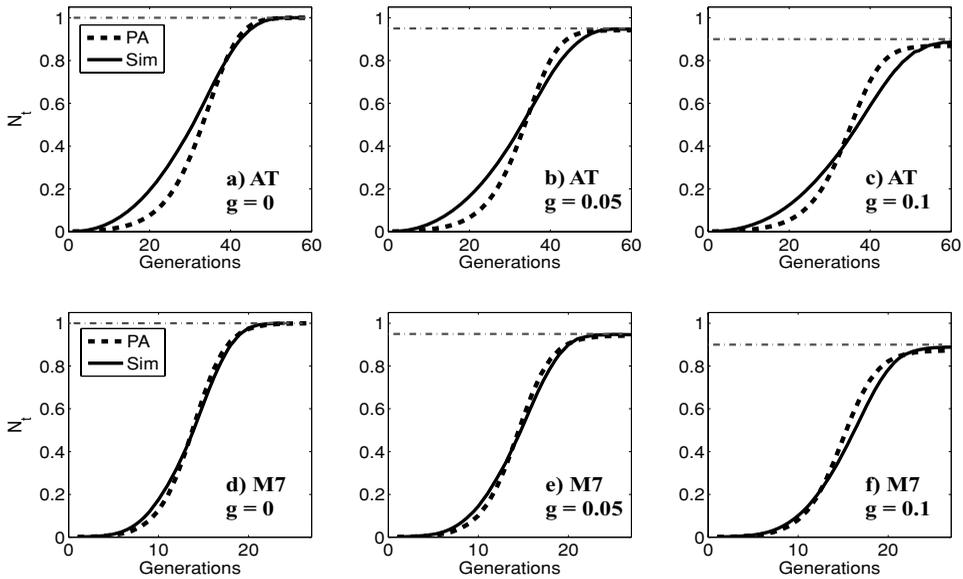
Figure 7: Influence of reversion probabilities g on the accuracy of the PA. Representative takeover dynamics with $p_{up} = 1$ and $\mu = 1024$, as estimated by the PA (dashed black line) and as observed through direct simulation (solid black line), for the *AT* population structure (top row, a–c) and for the *M7* population structure (bottom row, d–f). The three plots in each row correspond to $g \in \{0, 0.05, 0.1\}$, from left to right. The dash-dot horizontal line denotes the saturation sill, as approximated by $1 - g/p_{up}$. The scale of the vertical and horizontal axes are held fixed in each panel.

3.3.2 Sill Error

As g increases, the saturation sill decreases, as shown for representative population structures *AT* and *M7* in Figure 7a–c and Figure 7d–f, respectively. The dash-dot horizontal line represents the expected sill, which can be accurately predicted by $1 - g/p_{up}$ for regular population structures with $k > p_{up}/g$. For the simulation data, we plot the proportion of nodes (N_t) containing maximum fitness at time t, averaged over all of the trials in which extinction did not occur. Note that the proportion of high fitness individuals observed at saturation using the PA (dashed lines) and direct simulation (solid lines) are both in good agreement with the predicted sill ($1 - g/p_{up}$). For each value of p_{up}, the absolute sill error (Equation (23)) was found to increase slightly as g increased from 0.05 to 0.1. For all values of p_{up} and g tested, the absolute sill error remained low, with an average value of 1%. The worst sill error was observed on the *VN1* population structure, with absolute sill errors of 2.3% ($\mu = 6400$ and $p_{up} = 0.5$) and 7.5% ($\mu = 5184$ and $p_{up} = 0.5$), for $g = 0.05$ and $g = 0.1$, respectively.

3.3.3 Generation Error

Figure 7 shows that, as expected, increasing g slightly increases the number of generations required to reach the sill. This shifts both the PA and simulation curves to the right, even though the height of the sill is reduced for $g > 0$. For the *AT* population structure with $\mu = 1024$, Figure 7a–c shows the PA becoming increasingly accurate as g increases, whereas Figure 7d–f shows the PA becoming increasingly less accurate with

increasing g on the $M7$ population structure. Similar to the changes in the PA observed with decreasing p_{up} (Figure 6), these improvements and degradations in accuracy are the result of a shift in the dynamics predicted by the PA as g changes, relative to the shift in the dynamics observed through simulation. Thus, those populations for which the PA was predicting a slower spread of high fitness individuals than that observed through simulation when $g = 0$ become more accurate as g increases, and vice versa.

Generation error was found to be more sensitive to changes in g than to changes in p_{up}. For a given value of p_{up}, increasing g from 0 to 0.1 caused an average decrease in generation error of 3.9% among those populations for which error decreased as g increased and an average increase in error of 6.1% among those populations for which error decreased as g increased. For a given value of g, decreasing p_{up} from 1 to 0.5 caused an average decrease in error of 3.8% and an average increase of 5.6%.

4 Discussion

The aim of this study was to investigate whether the pair approximation (PA) could be used to easily and accurately estimate takeover dynamics for evolutionary algorithms that are spatially-structured on regular graphs. PAs were originally derived as a statistical mechanics formulation to approximate *equilibrium conditions* in spatially structured biological populations, in order to determine conditions for the evolution of altruism (Matsuda et al., 1992). PAs have since been widely applied in various biological applications, as reviewed in Section 1. In this manuscript, we first show that the takeover dynamics of an advantageous allele are equivalent to the dynamics of the well-known SIS model of disease spread. We then modify the epidemiological formulation of the PA of Keeling (1999) to model takeover dynamics in this system.

The original presentation of the PA (Matsuda et al., 1992) incorporated a constant coefficient of $c = 2$, and this value has since been used in numerous other studies. While the value of this coefficient does not affect equilibrium frequencies, we show that it does dramatically impact predictions of pre-equilibrium dynamics. Our results show that, in all 10 population structures tested, a value of $c = 2$ caused the PA to dramatically overestimate the rate of spread of advantages alleles, consistent with the results of Petermann and De Los Rios (2004). Furthermore, we show that, if one is trying to predict pre-equilibrium dynamics, the value of c that minimizes generational errors depends on the interaction between (a) the structure of the local interaction neighborhood and (b) the population size. We combine these two influences into a single parameter ρ (defined as the ratio of the radius of the local interaction neighborhood to the radius of the entire population) and show that the optimal value for the coefficient c can be estimated as a logarithmic function of ρ. Using this formula for c in the PA, we systematically assessed the accuracy of both equilibrium conditions and pre-equilibrium takeover dynamics for 10 regular neighborhood interaction structures, at eight population sizes, and with a variety of uptake and reversion probabilities. Parameterizing the coefficient c by ρ effectively shifts the saturation curve predicted by the PA so that it intersects with the simulation data at (or near) the inflection point, thereby minimizing generational errors in the PA. In general, two primary sources of errors remain: (a) residual errors introduced by using the regression curve to estimate c, and (b) errors introduced by the simplifications in the PA itself.

If the optimal values of c were used for each specific ρ, the PA and simulation takeover curves would intersect at the inflection point, thus balancing out positive and negative area and generational errors and minimizing overall error. However, in

general the optimal c would not be known for a particular neighborhood structure and population size unless extensive simulations were first performed, which would defeat the purpose of using the PA to rapidly and easily predict takeover dynamics without simulations. Thus, the PA was assessed using the values of c predicted by the empirically derived logarithmic fit relating c to ρ. Although this fit was quite good ($R^2 > 0.88$), individual values of c selected according to this formula did contain some small residual error, which caused some predicted curves to intersect the simulation curve a little before the inflection points, thus causing a slight net over-prediction in the rate of spread, or a little after the inflection points, causing a slight net under-prediction in the rate of spread. In general, decreasing the uptake probability or increasing the reversion probability slows takeover time and has the effect of shifting the PA curve slightly more to the right than the simulation curve. Whether this increased or decreased overall error depended on where the two curves intersected at $p_{up} = 1, g = 0$, relative to the inflection point. We suspect that the coefficient c could be further parameterized in terms of both $p_{up} = 1$ and $g = 0$, to reduce this variability.

The PA uses differential equations to model the dynamics of states of neighboring pairs of vertices. Higher order interactions are not explicitly modeled, but must be approximated using some method of closure. We employed the closure method of Keeling (1999), which incorporates the proportion of closed to total triplets that exist in the local interaction neighborhood (a.k.a. the clustering coefficient), thus closing the system at the level of triplets. Interactions higher than triplets were ignored. Consequently, the PA yielded less accurate predictions of dynamics on populations with von Neumann neighborhoods, which have no closed triangles (and thus no clustering) but do have a preponderance of closed quadruplets. To illustrate why this is a problem, consider a closed *wxyz* quadruplet (i.e., a square) in the *VN1* topology. While the state of node y cannot directly affect the state of node w, it can have an affect on z, which in turn may affect w. By ignoring quadruplet correlations and assuming no correlation between distant ends of triplets, the PA treats the *VN1* population structure as if it were a regular random graph with degree k. This explains the more rapid saturation predicted by the PA than that observed through direct simulation. Thus, if one wanted to use the PA to estimate takeover dynamics in populations with von Neumann (or other) neighborhoods with higher order interactions, accuracy could be improved by explicitly accounting for these correlations. For example, Van Baalen (2000) presented a closure method that captures quadruplet correlations, in order to improve the accuracy of the PA in predicting the equilibrium dynamics of a simple birth-death-movement process in lattice based population structures with von Neumann neighborhoods. Satō and Iwasa (2000) introduced a closure method, referred to as variable discounting, which similarly improves the accuracy of the PA in estimating pathogen invasion in lattice based spatial structures with spatially adjacent interactions (i.e., von Neumann and Moore). In contrast to these higher-order closure methods, Petermann and De Los Rios (2004) developed a technique referred to as the "cluster approximation" that explicitly tracks higher order correlations up to a specified degree, improving the accuracy of the approximation in estimating the dynamics of disease spread in lattice structured (Triangular and von Neumann) and random interaction topologies. Although the higher order approximations of Petermann and De Los Rios (2004) are able to more accurately predict equilibrium frequencies, the required number of coupled differential equations grows exponentially in the size of the correlation being tracked and growth rates are still over-predicted, especially in the lattice structured populations. While these higher order methods have yet to be explored on different population sizes, our results

suggest that the interaction between neighborhood size and shape and population size will still need to be accounted for. Such higher order improvements could be used in combination with a parameterized coefficient c as suggested in this paper, although the best fit curve for this would need to be recomputed for a given approximation method.

The PA exhibits exponential growth below the inflection point, while the takeover curves on lattices with local interactions increase only polynomially. Consequently, the accuracy of the PA decreases exponentially with increasing ρ (i.e., with increasing neighborhood interaction radii and/or decreasing population size), in part because larger neighborhood radii have higher polynomial exponents, and in part because as population sizes are decreased, individual neighborhoods cover a larger portion of the population (and thus approach panmixia). Furthermore, as population size increases, the exponential trend of the PA curve has more time to diverge from the polynomial trend of the simulation curve, thereby increasing discrepancies between the two curves. It is ironic that PAs, which were designed to model local interactions, actually work better when interactions are far-ranging. It is also usually assumed that continuous models of discrete systems work better as population sizes approach infinity, but in this case the PA works better for small populations, a counterintuitive finding.

The "replace if better" selection mechanism employed herein is clearly a simplification of the selection operators that are commonly used in evolutionary algorithms. However, PAs can be modified to employ much more complicated selection policies, such as frequency dependence (Van Baalen and Rand, 1998) and game-theoretic payoff matrices (Hauert and Doebeli, 2004; Ohtsuki et al., 2006). The PA could be similarly adjusted to include a more sophisticated reversion mechanism as well (e.g., frequency dependence), and recent work (Ellner, 2001) has demonstrated that the PA can even be adjusted to deal with processes (e.g., selection and reversion) that operate on differing spatial scales.

One definite limitation of the PA is the assumption of topological regularity. For example, while the proposed method for adjusting the PA as a function of ρ allows for nonuniform lattice dimensions, the exponential trend of the PA will become increasingly inaccurate as the dimensions of the lattice become more unequal; in the limiting case of the ring topology, in which the actual saturation dynamics are linear (Rudolph, 2000), the PA would be extremely inaccurate in capturing the saturation dynamics. The PA has been applied in a few population structures with mildly heterogeneous degree distributions, such as random graphs (e.g., Petermann and De Los Rios, 2004). However the extreme heterogeneity of other spatial structures of recent interest, including small-world (Watts and Strogatz, 1998) and scale-free (Barabàsi and Albert, 1999) topologies, prohibits the development of a PA for predicting dynamics on these networks. For example, Ohtsuki et al. (2006) noted that the simple rule for the evolution of cooperation in the graph-based prisoner's dilemma, derived using a PA based on the assumption of a regular graph, was inaccurate for populations structured on highly irregular topologies (e.g., scale-free). While the extreme irregularity of some topologies may preclude the development of such a generalized methodology for estimating takeover *dynamics*, the recent results of Payne and Eppstein (2007b) suggest that it may be possible to rapidly predict expected takeover *times* on arbitrary topologies, using only a few readily computable metrics of the underlying spatial structure, and work is underway to refine this empirical prediction approach (Payne and Eppstein, 2008).

The PA, as formulated herein, implicitly assumes synchronous updating. The choice of synchronous vs. asynchronous updating can have a significant impact on takeover dynamics (Giacobini, Tomassini, and Tettamanzi, 2005; Giacobini, Tomassini, Tettamanzi et al., 2005), and it has been reported that the PA produces more accurate results when

the update policy is synchronous (Hauert and Doebeli, 2004). Thus, this represents a further limitation of the PA.

5   Summary

In summary, takeover time analysis for evolutionary algorithms was reformulated in terms of the well-known SIS model of disease spread. An analytical technique, referred to as the pair approximation (PA), was then adapted to predict takeover dynamics and parameterized by ρ, the ratio of the radius of the local neighborhood and the radius of the graph. The accuracy of our parameterized reformulation of the PA was then assessed on a total of 10 distinct types of regular population structures, each with different configurations of interaction neighborhoods, using eight population sizes, and several combinations of selection and reversion probabilities. The results of this study demonstrate that our parameterized formulation of the PA, using the closure method of Keeling (1999), is a fast and reasonably accurate way to estimate both equilibrium and pre-equilibrium takeover characteristics of synchronously updated populations embedded on a variety of regularly structured graphs. A key result of this study is that the coefficient c in the PA (which is commonly assumed to be $c = 2$) should be parameterized by ρ if one is interested in approximating pre-equilibrium conditions. PAs are not appropriate for heterogeneous graphs, unless the nodal degree is reasonably close to constant. Thus, in biological or epidemiological studies, where interaction topologies are typically dynamic, heterogeneous, and difficult to ascertain, we caution that PAs may produce misleading results. However, in evolutionary computation, population structures are user-defined, frequently regular, and their topological properties are readily computable. We conclude that PAs can be a useful tool for rapidly estimating takeover dynamics in evolutionary algorithms on synchronously updated regular graphs, as long as care is taken to assess the topological characteristics of the graph in advance and the PA is appropriately formulated. Future work will seek to demonstrate if such insights may prove useful for guiding choices of local neighborhood structures in evolving populations, as a potential means of statically or dynamically optimizing selection pressure and convergence.

Acknowledgments

References

Altenberg, L. (2005). Evolvability suppression to stabilize far-sighted adaptations. *Artificial Life*, 11(4):427–443.

Anderson, R. M., and May, R. M. (1995). *Infectious Diseases of Humans*. Oxford, UK: Oxford University Press.

Barabàsi, A. L., and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286:509–512.

Bryden, K. M., Ashlock, D., Corns, S., and Wilson, S. (2005). Graph based evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 10(5):550–567.

Chakraborty, U. K., Deb, L., and Chakraborty, M. (1997). Analysis of selection algorithms: A Markov chain approach. *Evolutionary Computation*, 4(2):133–167.

Eames, K. T. D., and Keeling, M. J. (2002). Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of the Sciences*, 99(20):13330–13335.

Ebel, H., Mielsch, L., and Bornholdt, S. (2002). Scale-free topology of email networks. *Physical Review E*, 66:035103.

Ellner, S. P. (2001). Pair approximations for lattice models with multiple interaction scales. *Journal of Theoretical Biology*, 210(4):435–447.

Eppstein, M. J., and Molofsky, J. (2007). Invasiveness in plant communities with feedbacks. *Ecology Letters*, 10:253–263.

Falconer, D. S., and Mackay, T. F. C. (1996). *Quantitative Genetics*. London: Pearson Education.

Giacobini, M., Preuss, M., and Tomassini, M. (2006). Effects of scale-free and small-world topologies on binary coded self-adaptive cEA. In J. Gottlieb and G. R. Raidl, (Eds.), *Evolutionary Computation and Combinatorial Optimization*, pp. 86–98. Berlin: Springer-Verlag.

Giacobini, M., Tomassini, M., and Tettamanzi, A. (2005). Takeover times curves in random and small-world structured populations. In H. Beyer, (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2005*, pp. 1333–1340. New York: ACM Press.

Giacobini, M., Tomassini, M., Tettamanzi, A., and Alba, E. (2005). Selection intensity in cellular evolutionary algorithms for regular lattices. *IEEE Transactions on Evolutionary Computation*, 9(5):489–505.

Goldberg, D., and Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. In G. Rawlins, (Ed.), *Foundations of Genetic Algorithms*, pp. 69–93. San Mateo, CA: Morgan Kaufmann.

Gorges-Schleuter, M. (1999). An analysis of local selection in evolution strategies. In W. Banzhaf, J. Daida, A. E. Eiben, M. Garzon, V. Honavar, M. Jakiela, and R. Smith, (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-1999*, pp. 847–854. San Mateo, CA: Morgan Kaufmann.

Hauert, C., and Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428:643–646.

Heibeler, D. (2000). Populations on fragmented landscapes with spatially structured heterogeneities: Landscape generation and local dispersal. *Ecology*, 81(6):1629–1641.

Holland, J. (1992). *Adaptation in Natural and Artificial Systems*. Cambridge, MA: MIT Press.

Joo, J., and Lebowitz, J. L. (2004). Pair approximation of the stochastic susceptible-infected-recovered-susceptible model on the hypercubic lattice. *Physical Review E*, 70:036114.

Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London B*, 266:859–867.

Keeling, M. J., and Eames, K. T. D. (2005). Networks and epidemic models. *Journal of the Royal Society Interface*, 2:295–307.

Kerr, B., Riley, M. A., Feldman, M. W., and Bohannan, B. J. M. (2002). Local dispersal promotes biodiversity in a real life game of rock-paper-scissors. *Nature*, 418:171–174.

Kirley, M., and Stewart, R. (2007). An analysis of the effects of population structure on scalable multiobjective optimization problems. In D. Thierens, (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2007*, pp. 845–852. New York: ACM Press.

Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanely, H. E., and Åberg, Y. (2001). The web of human sexual contacts. *Nature*, 411:907–908.

Matsuda, H., Ogita, N., Sasaki, A., and Soto, K. (1992). Statistical mechanics of population: The lattice Lotka-Volterra model. *Progress of Theoretical Physics*, 88:1035–1049.

Meyers, L. A., Newman, M. E. J., and Pourbohloul, B. (2006). Predicting epidemics on directed contact networks. *Journal of Theoretical Biology*, 240(3):400–418.

Newman, M. E. J. (2002). Spread of epidemic disease on networks. *Physical Review E*, 66:016128.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45(2):167–256.

Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441:502–505.

Ong, N. P., and Cava, R. J. (2004). Electronic frustration on a triangular lattice. *Science*, 305:52–53.

Pastor-Satorras, R., and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(4):3200–3203.

Payne, J. L., and Eppstein, M. J. (2006). Emergent mating topologies in spatially structured genetic algorithms. In M. Keijzer, (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2006*, pp. 207–214. New York: ACM Press.

Payne, J. L., and Eppstein, M. J. (2007a). Sensitivity of self-organized speciation to long-distance dispersal. In *Proceedings of the IEEE Symposium on Artificial Life*, pp. 1–7.

Payne, J. L., and Eppstein, M. J. (2007b). Takeover times on scale-free topologies. In D. Thierens, (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2007*, pp. 308–315. New York: ACM Press.

Payne, J. L., and Eppstein, M. J. (2008). The influence of scaling and assortativity on takeover times in scale-free topologies. In M. Keijzer, (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2008*. New York: ACM Press. (to appear.)

Petermann, T., and De Los Rios, P. (2004). Cluster approximations for epidemic processes: A systematic description of correlations beyond the pair level. *Journal of Theoretical Biology*, 229:1–11.

Rauch, E. M., and Bar-Yam, Y. (2006). Long-range interactions and evolutionary stability in predator-prey systems. *Physical Review E*, 73:020903.

Rudolph, G. (2000). On takeover times in spatially structured populations: Array and ring. In K. Lai, O. Katai, and A. B. L. M. Gen, (Eds.), *Proceedings of the Second Asia-Pacific Conference on Genetic Algorithms and Applications, APGA-2000*, pp. 144–151. Hong Kong: Global Link.

Rudolph, G. (2001). Takeover times of noisy non-generational selection rules that undo extinction. In V. Kurkova, (Ed.), *Proceedings of the Fifth International Conference on Artificial Neural Networks and Genetic Algorithms*, pp. 268–271. Berlin: Springer-Verlag.

Santos, F. C., and Pacheco, J. M. (2005). Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters*, 95:098104.

Sarma, J., and De Jong, K. (1996). An analysis of the effect of neighborhood size and shape on local selection algorithms. In H. Voigt, W. Ebeling, I. Rechenberg, and H. Schwefel, (Eds.), *Parallel Problem Solving from Nature*, pp. 236–244. Berlin: Springer-Verlag.

Satō, K., and Iwasa, Y. (2000). Pair approximations for lattice-based ecological models. In U. Dieckmann, R. Law, and J. A. J. Metz, (Eds.), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, pp. 341–358. Cambridge, UK: Cambridge University Press.

Satō, K., Matsuda, H., and Sasaki, A. (1994). Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology*, 32:251–268.

Sayama, H., Kaufman, L., and Bar-Yam, Y. (2003). Spontaneous pattern formation and genetic diversity in habitats with irregular geographic features. *Conservation Biology*, 17:893–900.

Sprave, J. (1999). A unified model of non-panmictic population structures in evolutionary algorithms. In W. Banzhaf, (Ed.), *Proceedings of the Congress of Evolutionary Computation Conference, CEC-1999*, pp. 1384–1391. San Mateo, CA: Morgan Kaufmann.

Van Baalen, M. (2000). Pair approximations for different spatial geometries. In U. Dieckmann, R. Law, and J. A. J. Metz, (Ed.), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, pp. 359–387. Cambridge, UK: Cambridge University Press.

Van Baalen, M., and Rand, D. A. (1998). The unit of selection in viscous populations and the evolution of altruism. *Journal of Theoretical Biology*, 193:631–648.

Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393:440–442.

Werfel, J., and Bar-Yam, Y. (2004). The evolution of reproductive restraint through social communication. *Proceedings of the National Academy of Science*, 101(30):11019–11024.

Whitacre, J. M., Sarker, R. A., and Pham, Q. T. (2008). The self-organization of interaction networks for nature-inspired optimization. *IEEE Transactions on Evolutionary Computation*, 12(2):220–230.