

10-16-2009

Interaction Topologies and Information Flow

Joshua Payne
University of Vermont

Follow this and additional works at: <http://scholarworks.uvm.edu/graddis>

Recommended Citation

Payne, Joshua, "Interaction Topologies and Information Flow" (2009). *Graduate College Dissertations and Theses*. Paper 177.

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.

INTERACTION TOPOLOGIES AND INFORMATION FLOW

A Dissertation Presented

by

Joshua L. Payne

to

The Faculty of the Graduate College

of

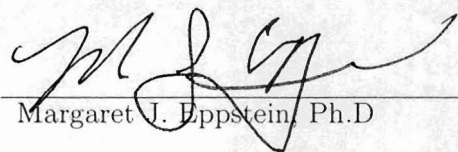
The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Computer Science

October, 2009

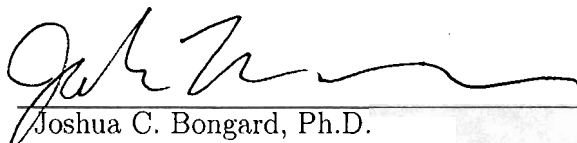
Accepted by the Faculty of the Graduate College, The University of Vermont, in partial fulfillment of the requirements for the degree of Doctor of Philosophy, specializing in Computer Science.

Dissertation Examination Committee:



Margaret J. Eppstein, Ph.D.

Advisor



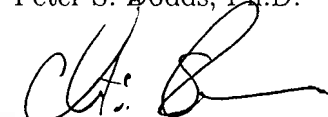
Joshua C. Bongard, Ph.D.



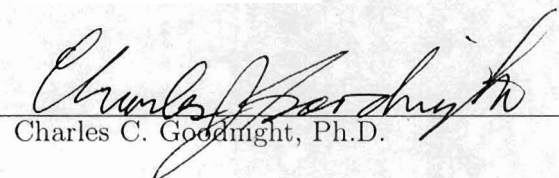
Christopher M. Danforth, Ph.D.



Peter S. Dodds, Ph.D.

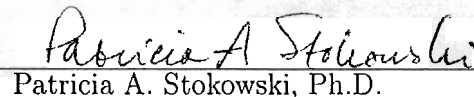


Christian Skalka, Ph.D.



Charles C. Goodnight, Ph.D.

Chairperson



Patricia A. Stokowski, Ph.D.

Interim Dean,
Graduate College

Date: May 6, 2009

Abstract

Networks are ubiquitous, underlying systems as diverse as the Internet, food webs, societal interactions, the cell, and the brain. Of crucial importance is the coupling of network structure with system dynamics, and much recent attention has focused on how information, such as pathogens, mutations, or ideas, flow through networks. In this dissertation, we advance the understanding of how network structure affects information flow in two important classes of models. The first is an independent interaction model, which is used to investigate the propagation of advantageous alleles in evolutionary algorithms. The second is a threshold model, which is used to study the dissemination of ideas, fads, and innovations throughout populations.

This journal-format dissertation comprises three interrelated studies, in which we investigate the influence of network structure on the dynamical properties of information flow. In the first study, we develop an analytical technique to approximate system dynamics in arbitrarily structured regular interaction topologies. In the second study, we investigate the flow of advantageous alleles in degree-correlated scale-free population structures, and provide a simple topological metric for assessing the selective pressures induced by these networks. In the third study, we characterize the conditions in which global information cascades occur in threshold models of binary decisions with externalities, structured on degree-correlated Poisson-distributed random networks.

Citations

Material from this dissertation has been accepted for publication in *Evolutionary Computation* on May 13, 2008 in the following form:

Payne, J.L., & Eppstein, M.J. (2009). Pair Approximations of Takeover Dynamics in Regular Population Structures. *Evolutionary Computation*.

Material from this dissertation has been accepted for publication in the *IEEE Transactions on Evolutionary Computation* on January 19, 2009 in the following form:

Payne, J.L., & Eppstein, M.J. (2009). Evolutionary Dynamics on Scale-Free Interaction Networks. *IEEE Transactions on Evolutionary Computation*.

Material from this dissertation has been submitted for publication to *Physical Review E* on March 13, 2009 in the following form:

Payne, J.L., Dodds, P.S., & Eppstein, M.J. Information Cascades on Degree-Correlated Random Networks. *Physical Review E*.

To my mother, for everything.

Acknowledgements

First thanks go to my family. To my mother for having the courage to bring me into this world and the patience to put up with me for all these years. This dissertation is as much your accomplishment as it is my own. I am forever indebted to you, and always grateful. To my grandparents, for being a constant source of inspiration and support. You are the most amazing couple I've ever met and I love you with all my heart. To my brother and sister, for simply being good people. I've learned more from the two of you than you could possibly know. Many thanks to all of my aunts and uncles for providing a great diversity of interesting examples of how a good life is lived. Lastly, a huge thanks to all of my cousins for keeping things youthful and fun.

The next round goes to my friends here in Burlington, who I've shared some wonderful moments with over the last five years. Summer days on sailboats, powder days on snowboards, mountain bike trips, miniramps, and endless dinner parties, we've really done an extraordinary job at having fun. Thank you for all the memories. To my friends elsewhere in the world, thanks for always being by my side, even though you're far away. Special thanks to Ryan, who is currently off on his own academic adventure. We've had some good times my friend, with many more sure to come. Also, a big thank you to Davnah. I haven't stopped smiling since I met you.

On to academics. Thanks go to my dissertation committee: Josh Bongard, Peter Dodds, Chris Danforth, Chris Skalka, and Charles Goodnight. You've always been supportive of my research endeavors and for that I'm very grateful. Thanks to Dave Bahr and Jim

Seibert for their support and friendship during my undergraduate education. Thanks to those organizations that have provided funding^{1,2} and computational resources³ over the years. Thanks to the Santa Fe Institute for inviting me to their summer school and thanks to the National Academy of Sciences for funding my adventures in Austria.

Finally, huge thanks go to my advisor, Margaret Eppstein. Its been a real pleasure getting to know you over the last five years. From summer school in Boston to winning a best paper award in Honolulu, we've shared some wonderful moments. You've been an amazing mentor, and I'm both a better scientist and a better person for knowing you.

¹Vermont EPSCoR NSF EPS 0701410

²Vermont EPSCoR DOE FG02 00ER45828

³Vermont Advanced Computing Center: NASA NNX 08A096G

Table of Contents

| | |
|---|-----------|
| Citations | ii |
| Dedication..... | iii |
| Acknowledgments | iv |
| List of Tables..... | viii |
| List of Figures..... | ix |
| Chapter 1 Introduction..... | 1 |
| 1.1 Information Flow | 3 |
| 1.1.1 The Flow of Alleles | 4 |
| 1.1.2 The Flow of Ideas | 4 |
| 1.2 Interaction Topologies | 6 |
| 1.2.1 Topological Metrics | 7 |
| 1.2.2 Regular Interaction Topologies | 8 |
| 1.2.3 Irregular Interaction Topologies | 12 |
| 1.3 Motivation and Contributions | 15 |
| Chapter 2 Pair Approximations of Takeover Dynamics in Regular Popula- tion Structures..... | 18 |
| Abstract | 18 |
| 2.1 Introduction | 19 |
| 2.2 Methods | 24 |
| 2.2.1 Representing population structure as a graph | 24 |
| 2.2.2 Structural properties of graphs | 24 |
| 2.2.3 Population structures | 26 |
| 2.2.4 Takeover time | 29 |
| 2.2.5 Selection | 29 |
| 2.2.6 Reversion | 30 |
| 2.2.7 Reformulating takeover time analysis in terms of the SIS model . . . | 30 |
| 2.2.8 Pair approximations | 31 |
| 2.2.9 Estimating takeover dynamics with pair approximations | 34 |
| 2.2.10 Experimental design | 35 |

| | | |
|---|---|------------|
| 2.2.11 | Assessing error | 36 |
| 2.3 | Results | 39 |
| 2.3.1 | Accounting for ρ in the PA | 39 |
| 2.3.2 | Non-extinctive dynamics | 40 |
| 2.3.3 | Relaxing the non-extinction assumption, $g > 0$ | 47 |
| 2.4 | Discussion | 50 |
| 2.5 | Summary | 55 |
| | References | 62 |
| Chapter 3 Evolutionary Dynamics on Scale-Free Interaction Networks . . . | | 63 |
| | Abstract | 63 |
| 3.1 | Introduction | 64 |
| 3.2 | Methods | 71 |
| 3.2.1 | Representing Population Structure as a Graph | 71 |
| 3.2.2 | Structural Properties of Graphs | 72 |
| 3.2.3 | Takeover Time Analysis | 74 |
| 3.2.4 | Generating Scale-Free Topologies with Tunable Scaling and Assortativity | 76 |
| 3.2.5 | Experimental Design | 82 |
| 3.3 | Results | 83 |
| 3.3.1 | Static Properties | 85 |
| 3.3.2 | Dynamic Properties | 86 |
| 3.4 | Discussion | 95 |
| 3.5 | Glossary of Variables | 103 |
| | References | 111 |
| Chapter 4 Information Cascades on Degree-Correlated Random Networks | | 112 |
| | Abstract | 112 |
| 4.1 | Introduction | 113 |
| 4.2 | Model | 115 |
| 4.3 | Methods | 117 |
| 4.4 | Experimental Design | 119 |
| 4.5 | Results | 119 |
| 4.6 | Discussion | 128 |
| | References | 133 |
| Chapter 5 Concluding Remarks | | 134 |
| Comprehensive Bibliography | | 137 |

List of Tables

| | | |
|-----|--|-----|
| 2.1 | 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. | 27 |
| 2.2 | (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. | 28 |
| 2.3 | 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. | 47 |
| 4.1 | Ranges of assortativity considered in this study, as a function of average degree z | 119 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | Illustration of the mechanisms of information transmission considered in this dissertation. The focal individual is colored gray, the black vertices denote high fitness (active) individuals and the white vertices denote low fitness (inactive) individuals. In our formulation of takeover time analysis, the probability of the gray vertex activating is 1/4 in (a) and 1/2 in (b). In our model of binary decisions with a threshold of 1/3, the probability of the gray vertex activating is 0 in (a) and 1 in (b). | 5 |
| 1.2 | Schematic diagram of (a-c) regular and (d-f) irregular interaction topologies. The regular networks depicted are (a) a complete graph, (b) a ring, and (c) a lattice. The irregular networks depicted are (d) an ER random graph, (e) a small-world graph, and (f) an AB scale-free graph. | 10 |
| 1.3 | Fractional size of the giant connected component (dark gray diamonds) and the frequency of global information cascades (light gray circles). The giant connected component was calculated as the average of those observed in 10 random ER graphs. The frequency of global cascades was calculated via simulation on these graph instances, using a vertex threshold of 0.18, replicating the results of Watts (2002) | 13 |
| 2.1 | 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. | 36 |
| 2.2 | 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. | 39 |
| 2.3 | Takeover dynamics as predicted by the PA using equation (2.24) (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. | 41 |

| | | |
|-----|--|----|
| 2.4 | (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. | 43 |
| 2.5 | 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) $\mu = 6400$ | 45 |
| 2.6 | 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. | 46 |
| 2.7 | 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. | 49 |
| 3.1 | Data points show representative complementary cumulative distribution functions (CCDF) of vertex connectivity in randomized scale-free interaction networks created with the GN algorithm (Section 3.2.4), with the smallest and largest scaling exponents used in this study $\gamma \in \{2.8, 4.0\}$ and $\langle k \rangle = 4$. Lines are drawn using the scaling exponents measured using the algorithm described in Section 3.2.4 with a k_{min} of 4, and are deliberately offset in the vertical direction so as not to obscure the data. Data shown pertain to $N = 10,000$. Note the double logarithmic scale. Insets denote the corresponding probability distribution functions (PDF), showing how the observed frequency of occurrence is lower-bounded by $1/N$ in these finite-sized graphs. | 73 |

| | | | |
|-----|-----|--|----|
| 3.2 | (a) | Upper bound of positive assortativities (r_+) and lower bound of negative assortativities (r_-) that we were able to obtain using the reshuffling algorithm described in Section 3.2.4, as a function of the scaling parameter (γ). Data points are the averages of the maximum (or minimum) assortativities that were achieved in 10 million assortative edge swaps on each of 10 graph instances at each of the 5 population sizes with $\langle k \rangle = 4$. Error bars indicate the minima and maxima observed for each combination of parameters. The shaded box indicates the range of experimental parameters examined in this study. (b) Visualization of one representative positively assortative graph ($\gamma = 2.8, r = 0.2, N = 500, \langle k \rangle = 4$), and (c) visualization of one representative negatively assortative graph ($\gamma = 2.8, r = -0.2, N = 500, \langle k \rangle = 4$). In (b) and (c) the degree distribution of the graphs is identical. For visual clarity, the graphs shown in (b) and (c) are deliberately smaller than any used in the experiments; visualizations were made using Pajek (http://vlado.fmf.uni-lj.si/pub/networks/pajek). | 80 |
| 3.3 | (a) | Characteristic path length (L) and (b, c) expected takeover time ($E[T]$), each shown as a surface function of assortativity (r) and scaling (γ) on a representative population size and average degree ($N = 3600, \langle k \rangle = 4$). The symbols represent the data points for each of the 10 graph instances at each parameter combination, with lines connecting the means. Panel (b) corresponds to local uniform selection and (c) to binary tournament selection (\bullet symbols, dashed line) and linear ranking selection (\circ symbols, solid line). In (a, b), the dash-dot lines at $\gamma \sim 2.8$ correspond to the data presented in Figs. 3.5 and 3.6, and the dashed lines at $r \sim 0.2$ correspond to the data presented in Figs. 3.4 and 3.7. | 84 |
| 3.4 | | Distribution of individual path lengths (L_i) in representative networks with $N = 3600, \langle k \rangle = 4$, and $r = 0.2$ for (a) $\gamma \sim 4$, (b) $\gamma \sim 3.6$, (c) $\gamma \sim 3.2$, and (d) $\gamma \sim 2.8$. The * symbols are placed atop each bin as a visual aid. The insets magnify the domain immediately surrounding the average characteristic path lengths L of the ten graph instances (denoted by the thick dashed vertical lines) in order to illustrate the nonlinear relationship between L and γ . Note that the data in this Figure corresponds to the data along the dashed line in Figure 3.3a. | 87 |
| 3.5 | | Distribution of individual path lengths (L_i) in representative networks with $N = 3600, \langle k \rangle = 4$, and $\gamma \sim 2.8$ for (a) $r \sim -0.2$, (b) $r \sim -0.1$, (c) $r \sim 0$, (d) $r \sim 0.1$, and (e) $r \sim 0.2$. The * symbols are placed atop each bin as a visual aid. The insets magnify the domain immediately surrounding the average characteristic path lengths L of the ten graph instances (denoted by the thick dash-dot vertical line) in order to illustrate the nonlinear relationship between L and r . Note that the data in this Figure corresponds to the data along the dash-dot line in Figure 3.3a. | 88 |

| | | |
|-----|---|-----|
| 3.6 | Takeover dynamics in representative networks with $N = 3600$, $\langle k \rangle = 4$, and $\gamma \sim 2.8$ for (a) $r \sim -0.2$, (c) $r \sim 0$, and (e) $r \sim 0.2$ (corresponding to the dash-dot lines in Fig. 3.3b). The gray curves denote takeover dynamics for a single experiment and the dash-dot curves denote their averages. In (b), (d), and (f) we show frequency histograms of observed takeover times (T) for the data shown in (a), (c), and (e), respectively. The * symbols are placed atop each bin as a visual aid. | 90 |
| 3.7 | Takeover dynamics in representative networks with $N = 3600$, $\langle k \rangle = 4$, and $r \sim 0.2$ for (a) $\gamma \sim 4.0$, (c) $\gamma \sim 3.6$, (e) $\gamma \sim 3.2$, and (g) $\gamma \sim 2.8$ (corresponding to the dotted lines in Fig. 3.3b). The gray curves denote takeover dynamics for a single experiment and the dashed curves denote their averages. In (b), (d), (f), and (h) we show frequency histograms of observed takeover times (T) for the data shown in (a), (c), (e), and (h), respectively. The * symbols are placed atop each bin as a visual aid. | 91 |
| 3.8 | (a) Takeover time ($E[T]$) as a function of the average inverse degree (μ_{-1}) and population size (N), for all scaling exponents (γ), assortativities (r), and average degrees ($\langle k \rangle$) considered in this study. The data presented in (a) is also shown in (b) and (c), but from different viewing angles, in order to better elucidate the strength of the relationship between (b) $E[T]$ and μ_{-1} and (c) $E[T]$ and N . Note that in (a) and (b), $\langle k \rangle = 4$ corresponds to $m = 2$ and $\langle k \rangle = 8$ to $m = 4$. In (a)-(c), the plane represents the best fit to the data (see text); open circles denote data points above this plane and filled squares denote data below this plane. Note the logarithmic scaling of takeover time ($E[T]$) and population size (N) in each panel. | 94 |
| 3.9 | (a) Standard deviation of expected takeover times ($\sigma_{E[T]}$) due to differences between the means on the 10 graph instances at each parameter combination, and (b) Standard deviations of individual takeover times (σ_T) on the same graph instances, due to the N initial locations of the high fitness individual. In both (a) and (b), standard deviations are shown as a surface function of assortativity (r) and scaling (γ) at $N = 10,000$, $\langle k \rangle = 4$. Note that the scale of the vertical axis in (a) is an order of magnitude smaller than in (b). | 95 |
| 4.1 | Observed frequency of global cascades, as a function of threshold ϕ and average degree z , for (a) disassortative ($r = -0.5$), (b) uncorrelated ($r = 0$), and (c) assortative ($r = +0.5$) random networks. The color bar corresponds to all panels. The asterisks denote the numerical solutions to Eq. (4.3). The dashed vertical lines at $\phi = 0.18$ correspond to the three curves in Fig. 4.2. | 118 |
| 4.2 | Observed statistics of global cascades as a function of average degree z for $\phi = 0.18$, on assortative ($r = +0.5$), uncorrelated ($r = 0$), and disassortative ($r = -0.5$) random networks. Each data point is the frequency of global cascades observed on 10 network instances. The lines are the corresponding average sizes of the extended vulnerable components (S_e). The x-axis is scaled logarithmically. | 121 |

| | | |
|-----|---|-----|
| 4.3 | Frequency (F_{gc} , left column) and size (S_{gc} , right column) of global cascades, as a function of average degree z and assortativity r , for three threshold values: $\phi = 0.2$ (top row), $\phi = 0.25$ (middle row), and $\phi = 0.33$ (bottom row). The color bar corresponds to all panels. The dashed horizontal lines in panels (a)-(d) are at $r = 0$, and the dashed vertical lines in panels (e) and (f) correspond to the data presented in Fig. 4. We were unable to obtain data in the hatched region (see Table 4.1). | 122 |
| 4.4 | (a) Frequency (asterisks) and size (circles) of global cascades, as a function of assortativity r , for $z = 2.0$ and $\phi = 0.33$. The bars correspond to the sizes of the vulnerable (S_v , black) and peripheral (S_p , white) subcomponents of S_e . The dashed line is provided as a guide for the eye. | 123 |
| 4.5 | Visualizations of random network properties with $z = 2$ and disassortative ($r = -0.8$, top row), uncorrelated ($r = 0$, middle row), and assortative ($r = +0.8$, bottom row) degree-degree correlations. In (a,d,g), we visualize averages of the E matrices for all 10 network instances ($N = 10,000$) used at each parameter combination. Shading (ranging from white to black) of matrix entries in E correspond to the probabilities (ranging from 0 to 0.17) with which a randomly chosen edge connects vertices of degree j and k ; entries corresponding to vertices with $k > 6$ occur with very low frequency and so are not shown. In (b,e,h), we depict the extended vulnerable component (S_e) for $\phi = 0.33$ for individual random networks with $N = 200$ (for visual clarity). Peripheral vertices ($k > 3$) are denoted as white squares; vertices in the vulnerable component are denoted as circles, and are colored according to degree: white ($k = 1$), gray ($k = 2$), and black ($k = 3$). The corresponding pie charts in (c,f,i) denote the proportions of degrees found in these extended vulnerable components (b,e,h). | 124 |
| 4.6 | Illustration of a representative cyclical triggering motif (for $\phi = 0.33$) in the external component. Diamonds correspond to vertices in the external component (Ω_x), the square corresponds to a vertex in the peripheral component (Ω_p), and the circle to a vertex in the vulnerable component (Ω_v). The initial seed is placed in the black diamond. Spreading into Ω_v is indicated by the black arrows. | 126 |
| 4.7 | Average cascade size with $\phi = 0.18$ as a function of average degree z , given that the initial seed was placed in high-degree vertices (in the top 10% of the degree distribution, squares) or in the remaining vertices in the bottom 90% of the degree distribution (asterisks), for (a) disassortative ($r = -0.8$), (b) uncorrelated ($r = 0$), and (c) assortative ($r = +0.8$) random networks. Note that in (a), we were not able to obtain data for $z < 2$ (Table 4.1). Lines are provided as a guide for the eye. | 127 |

Chapter 1

Introduction

The last decade has witnessed an unparalleled advance in our understanding of the interaction patterns of real-world systems, fueled by a dramatic increase in computational power and the availability of large-scale datasets. Protein-protein interactions in the cell (Jeong et al., 2001), router connectivity in the Internet (Faloutsos et al., 1999), and even some human relationships (Ebel et al., 2002; Liljeros et al., 2001) have been mapped, and their structures are both more complex and, in some ways, more similar than most had previously imagined. These findings have spurred an intense research effort to model the structural evolution of large-scale networked systems (Newman et al., 2006) and to investigate the influence of network structure on dynamical processes (Barrat et al., 2008), uncovering several important relationships between the topological properties of interaction networks and system dynamics. For example, the evolution of altruism (Werfel and Bar-Yam, 2004) and cooperation (Hauert and Doebeli, 2004; Nowak and May, 1992; Santos and Pacheco, 2005), the spreading of disease (Keeling, 1999; Kiss et al., 2008; Newman, 2002b; Pastor-Satorras and Vespignani, 2001) and ideas (Watts, 2002), and the propagation of failures in critical technological infrastructures (Kinney et al., 2005; Sachtjen et al., 2000) have all been shown to be strongly affected by the structural characteristics of the underlying interaction network. In addition to understanding how interaction topologies affect dynamics in real-world

systems, there has also been an increasing interest in exploiting explicit population structures to improve the search performance of nature-inspired optimization algorithms (Alba and Dorronsoro, 2005; Alba and Tomassini, 2002; Bryden et al., 2005; Rudolph, 2001; Sarma and De Jong, 1996; Whitacre et al., 2008), as a means to improve diversity maintenance and increase system evolvability.

In order to understand how interaction topologies affect system dynamics in both evolutionary algorithms and real-world dynamical processes, models are often constructed by whittling the system down to something much simpler, abstracting away many of the confounding features of the real system. The states of system components, whether these components are individuals, routers, or web pages, are often represented as binary values (e.g., an individual is either sick or healthy, a router is either broken or functioning properly, etc.). Interaction patterns are commonly represented as static, undirected, unweighted networks, where vertices denote system components and edges (connections between vertices) denote their potential ability to interact. For example, routers may be represented as vertices with edges denoting their physical connection; metabolites may be represented as vertices with edges denoting enzyme-catalyzed reactions that change one metabolite to another. Lastly, an appropriate mechanism for updating vertex states is chosen. For example, in an epidemiological model, the probability of an individual changing from a healthy to an infected state may be dependent on the number of independent contacts with infected individuals. Alternatively, in a model of the power grid, the probability of a transmission component changing from a functioning to a failed state may be dependent on the cumulative energy load transferred from adjacent power generation components.

While such simplified models are idealized abstractions of real systems, their analysis can provide useful insights regarding the influence of the structural characteristics of interaction networks on the dynamical properties of both evolutionary algorithms and real-world systems. For example, simple models of the flow of advantageous alleles in structured populations (Sarma and De Jong, 1996; Giacobini et al., 2005b) have demonstrated that

the performance improvements obtained with lattice-based population structures in evolutionary computing result from a slowed spread of high fitness solutions. In the context of real-world systems, simple lattice-based models of individual competition have demonstrated that increased spatial locality of interaction events facilitate the maintenance of genetic diversity (Kerr et al., 2002; Sayama et al., 2003) and can potentially lead to speciation (Eppstein et al., 2009; Payne et al., 2007); and in social systems, simple models of contagion have demonstrated that the ability of a single piece of information, be it a song or fashion trend, to spread throughout a population results from the presence of an interconnected community of “early adopters” (Watts, 2002; Watts and Dodds, 2007).

Despite these recent findings, many questions regarding the intricate and complex relationships between interaction topologies and information flow still remain unanswered. In this dissertation, we advance the understanding of how network structure affects information flow in both evolutionary computation and real-world systems, through an investigation of two important classes of models on a variety of interaction topologies.

1.1 Information Flow

We use the term *information flow* to denote the spreading dynamics of entities (e.g., pathogens, mutations, ideas, etc.) throughout a system. This term is therefore quite general, and could be as easily applied to the dissemination of ideas between cultures as to the propagation of failures in technological networks.

Models of information flow employ a great variety of information transmission mechanisms. At one extreme lies the independent interaction model, where each contact with a piece of information (referred to as an exposure) yields an independent probability of information transmission (Dodds and Watts, 2004). At another extreme is the threshold model, in which the probability of information transmission abruptly jumps from low to high as a critical number of exposures is exceeded (Dodds and Watts, 2004). In this Section, we describe two instantiations of these models, which will be the focus of investigation in this

dissertation.

1.1.1 The Flow of Alleles

The first model is used to study the flow of advantageous alleles in evolutionary algorithms. Individuals are represented as vertices, and conceptualized as either possessing an advantageous mutation or not, resulting in a binary state space of high and low fitness individuals, respectively. The state of a vertex is updated by randomly selecting an adjacent vertex and then adopting the state of the adjacent vertex if it is of higher fitness (Fig 1.1). Thus, each contact provides an independent probability of transmitting information, making this an independent interaction model, akin to models of disease spread (Dodds and Watts, 2004). Since it is assumed that the population is connected (i.e., a path exists between any two vertices), the quantity of interest in this model is the amount of time required for an advantageous allele to fully saturate the population.

Monitoring the saturation dynamics of advantageous alleles is referred to as takeover time analysis (Goldberg and Deb, 1991), and is used to assess the impact of various tunable parameters on selective pressure in evolutionary algorithms. If the takeover time is rapid, then advantageous mutations flow quickly throughout the population, and selective pressure is high. If takeover time is slow, then the flow of advantageous mutations is more gradual, inducing a lower selective pressure on the population. Lower selective pressures are known to enhance the exploration phase of evolutionary optimization (Goldberg and Deb, 1991; Giacobini et al., 2005b), and aid in curbing premature convergence to a local suboptima.

1.1.2 The Flow of Ideas

The second class of models considered in this dissertation are referred to as binary decisions with externalities, and were originally developed in the social sciences (Granovetter, 1978; Schelling, 1973) to study the dissemination of fads, ideas, and innovations. The term binary decision refers to an either-or situation, such as choosing to ratify a treaty, or subscribe to

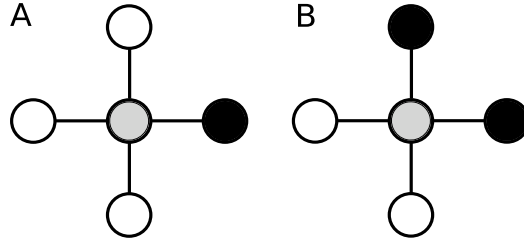


Figure 1.1: Illustration of the mechanisms of information transmission considered in this dissertation. The focal individual is colored gray, the black vertices denote high fitness (active) individuals and the white vertices denote low fitness (inactive) individuals. In our formulation of takeover time analysis, the probability of the gray vertex activating is $1/4$ in (a) and $1/2$ in (b). In our model of binary decisions with a threshold of $1/3$, the probability of the gray vertex activating is 0 in (a) and 1 in (b).

a new telephone service. The state space thus consists of individuals who have, and have not, adopted the piece of information under investigation (referred to as active and inactive individuals, respectively). Externalities refer to the choices of others, which affect your own. For example, the utility of a particular cell phone plan may depend on the proportion of your contacts that have also adopted that plan. A nation’s decision to sign a treaty may depend on the fraction of its allies that have also signed that treaty. This model of information propagation can be similarly applied to technological systems as well (Kinney et al., 2005; Sachtjen et al., 2000). The probability of failure for a transmission component in a power grid may be contingent on its incoming power load; when a specific proportion of adjacent components have failed, this critical load is no longer met, and the focal component fails as well, thus resulting in cascading failures. In these examples, the “decision” of a system component to adopt a piece of information is dependent on the proportion of its contacts that have also adopted this information. As such, binary decisions with externalities are threshold models (Dodds and Watts, 2004).

There are many possible variations in the mechanisms of information transmission in these models. We consider a basic Heaviside threshold mechanism, where vertex activation is dependent on whether or not the proportion of adjacent vertices that are active exceeds the threshold of the focal individual. Every individual is initialized with an identical activation threshold. The dynamics of the model begin by placing a single active individual into a

population of inactive individuals, and then monitoring the spread of this information. In each update, every individual calculates the proportion of their contacts that are active. If this proportion exceeds the individual's threshold, then it becomes active itself (Fig. 1.1). Once an individual has activated, it is assumed to remain active and the spreading dynamics continue until a steady state is reached.

Aside from the mechanism of information transmission, this model differs from our model of the flow of alleles in two primary ways. First, this system is purely deterministic. Second, the propagation of information does not always saturate an entire population, even if the population is connected. The spread of information often fizzles out very quickly, frustrated by the presence of highly connected individuals that are difficult to influence. However, situations also arise in which information spreads to a significantly large proportion of the population, giving rise to what is referred to as a global information cascade (Watts, 2002). The quantities of interest in this model are therefore the frequency with which global cascades occur and their final size.

1.2 Interaction Topologies

The focus of this dissertation is on how interaction patterns influence the flow of information in these two models. These interaction patterns are represented as networks, which provide a useful and general abstraction of system structure. As an object of study, graphs have been a topic of interest in mathematics for centuries (Euler, 1741). However, the sheer volume of data involved in recent mappings of real-world networks, such as the proteome (Jeong et al., 2001) and the World Wide Web (Albert et al., 1999), have necessitated the development of a new suite of topological metrics that can succinctly characterize such complex, large-scale interaction patterns.

1.2.1 Topological Metrics

The number and diversity of topological metrics that have been developed to characterize real-world interaction networks is vast. Here, we outline the metrics that are considered in this dissertation and provide an example of their relevance to real systems. (These metrics will be formally defined in the main body of this dissertation.) Perhaps the most basic is the *degree distribution*, which captures the probability with which a vertex has a given number of connections (referred to as the *degree* of the vertex). Recent analyses of systems as diverse as the Internet (Faloutsos et al., 1999), genetic regulatory networks (Ravasz et al., 2002), and societal interactions (Ebel et al., 2002; Liljeros et al., 2001) have revealed that these seemingly disparate interaction networks possess degree distributions that follow a similar mathematical form. Specifically, their probability distributions follow a power-law, meaning that the majority of vertices possess only a few connections, but a few vertices possess an extremely large number of connections. Such networks are referred to as “scale-free” (Barabási and Albert, 1999), because the degree distribution does not possess a characteristic scale. That is, the network’s average degree is not representative of the overall connectivity distribution, in sharp contrast to Erdős-Rényi (ER) random networks (Erdős and Rényi, 1959), which possess Poisson degree distributions. The extreme heterogeneity in the degree distribution of scale-free networks has important implications for system dynamics. For instance, power-law degree distributions render systems robust to random failure, but vulnerable to targeted attack (Albert et al., 2000).

In an analysis of social, technological, and biological networks, Watts and Strogatz (1998) discovered two common trends: the average distance between any two vertices is quite short and the neighbors of a given vertex are likely to be neighbors themselves. These networks are referred to as “small-world” and these two distinctive topological properties are quantified as the *characteristic path length* and *clustering coefficient*, respectively. Characteristic path length, which is the average all-pairs shortest distance between vertices, has been shown to largely govern the rate of pathogen invasion in a simple model of disease

spreading (Watts and Strogatz, 1998). The clustering coefficient, which is the frequency of triangular subgraphs, has also been shown to impact the dynamics of pathogen invasion. Keeling demonstrated that increased clustering in the contact network decreased the probability of large-scale epidemics (Keeling, 2005).

Subsequent analyses of social, technological, and biological networks revealed a more subtle structure. Newman found that the degrees of adjacent vertices are often correlated, with social networks generally possessing positive degree-degree correlations and biological and technological networks generally possessing negative degree-degree correlations (Newman, 2002a, 2003a). These correlations are quantified in a metric referred to as *assortativity*, which is the Pearson correlation coefficient of the degrees of adjacent vertices. The introduction of degree-degree correlations has been shown to affect the evolution of cooperation in models of social dilemmas (Rong et al., 2007) and the rate of pathogen invasion in simple models of disease spread (Kiss et al., 2008).

Some prior work exists regarding the influence of interaction topologies on our two exemplary models of information flow and we now briefly review these studies on regular (where all vertices have the same degree) and irregular interaction topologies.

1.2.2 Regular Interaction Topologies

The Well-Mixed Case

As a starting point, models of information flow typically treat system interaction patterns as “well-mixed,” in which any individual is equally likely to interact with any other individual in the population. This interaction pattern can be thought of as a complete graph, wherein all vertices are interconnected (Fig. 1.2a). Such an interaction topology often allows for the development of an analytically tractable model and provides a useful frame of reference for any deviation from this assumption.

Goldberg and Deb (1991) introduced the concept of takeover time analysis to the evolutionary computation community, as a means of quantifying selective pressure. Their

focus was on the impact of several commonly implemented selection schemes on the flow of advantageous alleles in well-mixed populations. In a head-to-head comparison of fitness proportionate, linear ranking, and tournament selection, they demonstrated that some selection policies lead to faster population convergence than others. Specifically, larger tournament sizes were shown to lead to faster takeover times than smaller tournaments, and linear ranking and binary tournament selection were shown to induce qualitatively similar selection pressures under certain conditions.

Models of binary decisions with externalities in well-mixed interaction topologies were investigated by Granovetter (1978). His work focused on how the distribution of individual thresholds affected the aggregate behavior of the population, and demonstrated that small perturbations in the threshold distribution could lead to large changes in collective behavior. For example, a slight change in the standard deviation of a Gaussian threshold distribution could tip collective behavior from information spreading through only a few individuals to spreading throughout the entire population. Using riots as an illustrative example of collective behavior, Granovetter employed this result to argue that the population-level composition of rioting preferences (i.e., how many individuals you must see rioting before you decide to riot) need not be substantially different in a rioting population than in a non-rioting population. That is, individual preferences are not easily inferred from aggregate behavior.

The Ring and the Lattice

The assumption of complete mixing is a convenient, but crude approximation to the interaction patterns of most real-world systems. For example, many terrestrial ecological communities are spatially-embedded on the two-dimensional surface of the earth, with interaction events constrained by geographical distance. As such, ecological models developed under the assumption of complete mixing are often unable to accurately characterize system dynamics. Motivated by this realization, modeling efforts began to relax the assumption

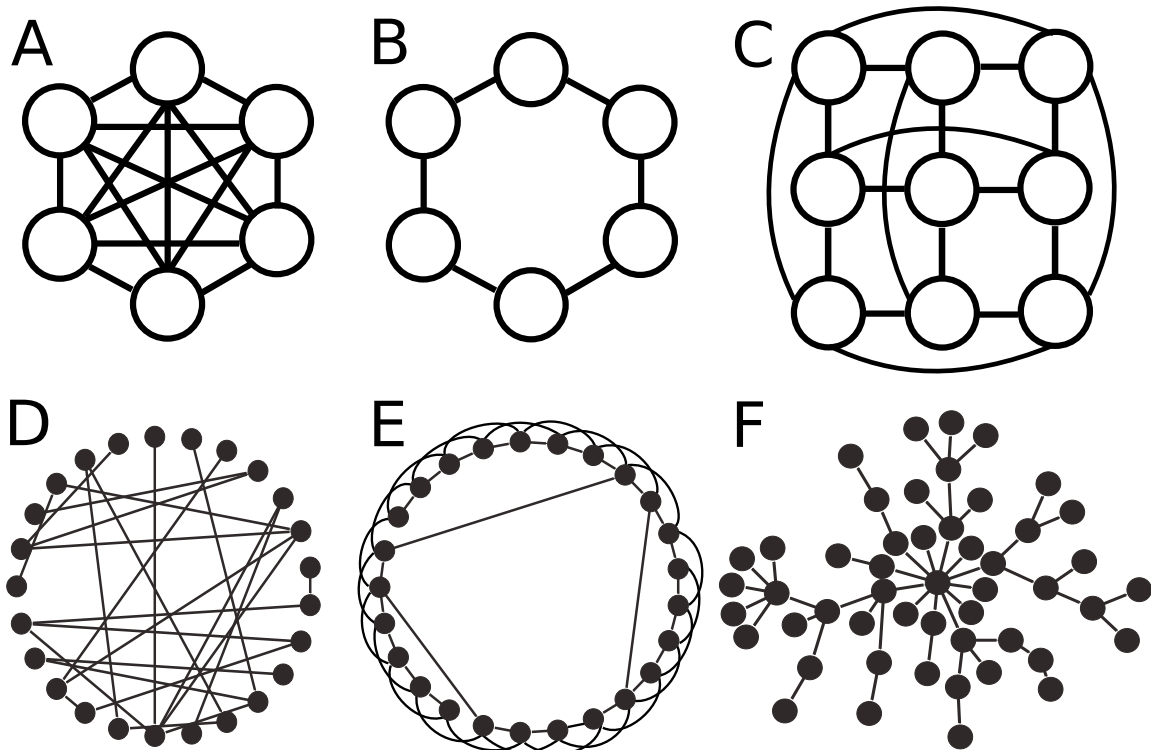


Figure 1.2: Schematic diagram of (a-c) regular and (d-f) irregular interaction topologies. The regular networks depicted are (a) a complete graph, (b) a ring, and (c) a lattice. The irregular networks depicted are (d) an ER random graph, (e) a small-world graph, and (f) an AB scale-free graph.

of a complete interaction topology by structuring populations on regular networks, such as a one (Fig. 1.2b) or two-dimensional (Fig. 1.2c) lattice. In the context of evolutionary computation, the utilization of explicit population structures was initially motivated not by biological realism, but by algorithm parallelization (see Alba and Tomassini (2002) for a review of this topic). Early studies focused on coarse-grained parallelization, where independent, well-mixed subpopulations are placed on separate processing units and coupled by occasional migration events (Grefenstette, 1981). Fine-grained parallelism, where each individual is assigned to a separate processing unit and individuals interact according to a specific communication topology, was investigated later. Specifically, Manderick and Spiessens (1989) considered a two-dimensional lattice topology with localized neighborhoods and found that this population structure not only facilitated parallelism, but also improved the effectiveness of evolutionary search, with smaller neighborhoods yielding higher quality solutions. This observation led to the development of cellular evolutionary algorithms, where individual populations are structured on regular lattices (whether serial or parallel) and mating events are restricted to occur in spatially-localized neighborhoods (Alba and Dorronsoro, 2008).

The first takeover time analyses of regular interaction topologies were performed by Sarma and De Jong (1996). Their study focused on the influence of local neighborhood size and shape on takeover times in two-dimensional toroidal lattices. Their results demonstrated that a simple metric of neighborhood dispersion, referred to as a neighborhood's *radius*, largely governs takeover times in regular lattices. Specifically, they demonstrated that the growth coefficient in a logistic model of takeover dynamics increased logarithmically with the ratio of the local neighborhood radius to the global lattice radius, and asymptotically approached the growth coefficient of a well-mixed system. Thus, increasing the ratio of these two radii allows one to easily tune the selective pressures induced by a regular interaction network from low (nearest neighbor interactions) to high (well-mixed interactions). These empirical observations were soon followed by mathematical models of

takeover dynamics in one and two-dimensional lattices (Rudolph, 2000; Giacobini et al., 2005b).

In an analysis of segregation processes, Schelling (1971) investigated a model of binary decisions with externalities on both one and two-dimensional lattices. As his study was focused on racial segregation, he employed a method of information transmission that is different from the one treated in this dissertation. Regardless, his analyses provided some useful insights regarding information flow on regular interaction topologies. Specifically, Schelling (1971) showed that mild preferences for one type over another within local neighborhoods could lead to global population segregation, again casting doubt on the possibility of determining micromotives from collective behavior.

1.2.3 Irregular Interaction Topologies

The Random Graph

The properties of random graphs have been rigorously studied since the seminal work of Erdős and Rényi (1959). One classic result of random graph theory is that a giant connected component exists in these Poisson-distributed networks so long as its average degree is greater than one, and as the average degree continues to increase, the size of this component approaches that of the entire graph (Fig. 1.3, dark gray diamonds). An average degree of one is thus referred to as the percolation threshold, since this is the point where the giant connected component first occupies a finite fraction of an infinite graph.

In the context of takeover time analysis, ER random interaction topologies (Fig. 1.2d) were first considered by Giacobini et al. (2005a). They provided empirical and analytical results for takeover dynamics on connected ER random graphs well above the percolation threshold (with an average degree ranging from 10 to 50), demonstrating that takeover times decrease as the average degree of the network increases. One important result of their research is that the selective pressures induced by ER random networks with high average degree are quantitatively similar to those induced in the well-mixed (complete graph) case.

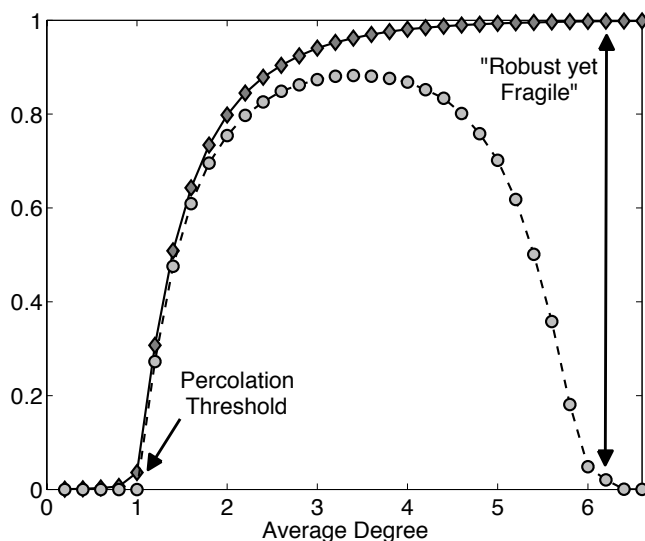


Figure 1.3: Fractional size of the giant connected component (dark gray diamonds) and the frequency of global information cascades (light gray circles). The giant connected component was calculated as the average of those observed in 10 random ER graphs. The frequency of global cascades was calculated via simulation on these graph instances, using a vertex threshold of 0.18, replicating the results of Watts (2002)

However, it remains unclear how their results scale to ER random networks with an average degree closer to the percolation threshold.

Binary decisions with externalities were investigated on ER random networks by Watts (2002). He demonstrated that global information cascades can only occur if there exists a percolating subnetwork of individuals that each activate in the presence of a single active neighbor. This connected component of vulnerable individuals (a.k.a. “early adopters”) acts as the network’s critical mass; once they activate, the rest of the network is easily tipped. For ER random graphs with low average degree, the frequency (Fig 1.3, light gray circles) and size of a cascade was shown to be limited by the connectivity of the network; that is, the spread of information was limited to the giant connected component. In contrast, for random graphs with high average degree, where the size of the giant connected component approaches that of the entire network, the frequency of global cascades was shown to be vanishingly small, but when global cascades did occur, they would topple the entire network (Fig 1.3, vertical arrow). Information cascades on sufficiently dense networks

thus possess a “robust yet fragile” quality, where information spreading is typically limited by the presence of well-connected individuals that act as barriers, but under the right conditions, information can saturate the entire network. Watts (2002) also demonstrated that increased heterogeneity in the threshold distribution expanded the parameter range in which global cascades occur.

The Small-World and the Scale-Free

As previously mentioned, recent analyses have clearly demonstrated that the interaction patterns of numerous systems are highly heterogeneous, often possessing heavy-tailed degree distributions, short characteristic path lengths, high clustering, and degree-degree correlations. These observations have prompted the development of myriad models of dynamical processes on irregular networks (Barrat et al., 2008), and have also inspired research into the utilization of irregular population structures for evolutionary optimization. Giacobini et al. (2005a) were the first to investigate takeover times in both small-world (Fig. 1.2e) and scale-free (Fig. 1.2f) interaction topologies. For small-world networks, they developed mathematical approximations of takeover dynamics and showed that takeover times decreased as the interaction network was rewired to possess a shorter characteristic path length. Takeover dynamics on scale-free interaction topologies were investigated empirically, using the graph generating algorithm provided by Barabási and Albert (1999) (AB). These results demonstrated that takeover dynamics are at least as fast in AB scale-free interaction topologies as they are in random networks, and that takeover times could be decreased significantly if the initial high fitness information is strategically placed in a highly connected vertex.

Centola et al. (2007) investigated a model of binary decisions with externalities on small-world interaction topologies with fixed average degree. Their results demonstrated that edge rewiring dramatically decreased the parameter range in which global information cascades occur, since these random links weaken local structure and thus reduce the

probability of receiving multiple signals from an active cluster. Information cascades have also been investigated on scale-free interaction topologies. Watts (2002) demonstrated that the introduction of power-law degree distributions severely limited the parameter range in which global information cascades occur, resulting from the presence of highly connected vertices that are difficult to tip, and thus frustrate information flow.

1.3 Motivation and Contributions

Regular interaction topologies have been shown to improve performance in many forms of evolutionary algorithms, including genetic programming (Folino et al., 2003b; Spector and Klein, 2005) algorithms (Alba and Dorronsoro, 2004, 2005), and particle swarm optimization (Kennedy and Mendes, 2002). These performance improvements are often attributed to a slowed spread of high fitness solutions throughout the population, which reduces selective pressure and facilitates diversity maintenance. The models of takeover dynamics that have been previously developed provide a useful reference for understanding how a chosen interaction network will affect selective pressure. However, these existing models are network-specific, and for the common case of two-dimensional networks, models only exist for nearest-neighbor interactions (Giacobini et al., 2005b). One goal of this dissertation is to develop a general method for estimating takeover dynamics in arbitrarily structured regular interaction topologies. To this end, we present in Chapter 2 a reformulation of takeover time analysis in terms of the classic Susceptible-Infectious-Susceptible model of disease spread. We then adapt a general technique, referred to as the pair approximation, for estimating takeover dynamics in regular interaction topologies. Our results demonstrate that the pair approximation, as originally formulated, is insufficient for capturing pre-equilibrium dynamics of information flow because it does not properly account for the interaction between neighborhood structure and population size. We then show that the pair approximation can be parameterized by the ratio of the local neighborhood radius to the global lattice radius to yield more accurate results. We empirically analyze this method for estimating takeover

dynamics on a variety of regular interaction topologies with several population sizes.

The use of irregular interaction topologies has also recently received increasing interest in the evolutionary computation community. In particular, scale-free population structures have been investigated in genetic algorithms used for a variety of single (Giacobini et al., 2006) and multiobjective (Kirley and Stewart, 2007a,b) optimization problems, and in a mobile robotics application (Gasparri et al., 2007). However, these studies have produced variable results, with scale-free population structures enhancing search performance in some cases and hindering it in others. In order to properly assess the potential value of scale-free interaction networks for evolutionary algorithms, a deeper understanding of their impact on selective pressure is required. While empirical results of takeover dynamics exist for AB scale-free networks (Giacobini et al., 2005a), these results pertain only to uncorrelated networks possessing a degree distribution with a fixed scaling exponent and a single combination of average degree and population size. In Chapter 3, we investigate takeover dynamics on scale-free interaction topologies with various population sizes, average degrees, scaling exponents, and assortativities. Our results demonstrate that average takeover times increase with both decreased scaling exponents and increased assortativity, with nonlinear interactions between these two topological properties. We show that several topological metrics of interest in related studies of information flow are not able to explain the takeover times observed on the broader class of scale-free interaction networks that we consider. Rather, we demonstrate that 97% of the variability in the logarithmically transformed average takeover times can be simply described using a planar function of two simple topological metrics, providing an easily computable estimation of the selective pressures induced by scale-free interaction topologies. However, this simple relationship does not capture the effects of assortativity, which are shown to interact nonlinearly with scaling to influence the variability in takeover times both within and between graph instances, in distinctly different ways.

In the context of threshold models of binary decisions with externalities, there have

been several recent studies focused on the relationship between the structural properties of interaction topologies and information cascades. These results have demonstrated how modularity (Galstyan and Cohen, 2007; Watts and Dodds, 2007), edge directionality (Gleeson, 2008b), the presence of long distance interactions vs. local clustering (Centola et al., 2007), and degree heterogeneity (Watts, 2002; Watts and Dodds, 2007) affect both the size and frequency of global information cascades. However, relatively little attention has been paid to the influence of degree-degree correlations. Gleeson has recently provided an analytic approximation for the size of global information cascades on degree-correlated random networks (Gleeson, 2008a) and we have complemented these results by providing an analytic solution for global cascade frequencies (Dodds and Payne, 2009). Despite these advances, much remains unclear regarding the influence of degree-degree correlations on information cascades in this model. In Chapter 4, we experimentally investigate a model of binary decisions with externalities on degree-correlated random interaction topologies. We demonstrate that increasing degree-degree correlations generally expands the parameter range in which global information cascades occur. However under certain conditions, global information cascades are found to occur for sufficiently positive or negative degree-degree correlations, but not when correlations are relatively small. We demonstrate that the ability of high degree vertices to trigger a global information cascade is strongly affected by degree-degree correlations; in negatively assortative networks, high degree vertices trigger the largest average information cascades, but in positively assortative networks, the opposite result is found. Lastly, we show that cyclical triggering motifs are present in these large finite graphs that trigger cascades on denser networks than predicted analytically.

Chapters 2-4 thus represent the body of this journal format dissertation. In Chapter 5, we provide some concluding remarks and discuss directions for future research.

Chapter 2

Pair Approximations of Takeover Dynamics in Regular Population Structures

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 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-equilibrium 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.

2.1 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 et al., 2005a; Giacobini et al., 2005b; Payne and Eppstein, 2007, 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, 1975), 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 et al. (2007); 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 et al., 2005a; Giacobini et al., 2005b; Giacobini et al., 2006; Kirley and

Stewart, 2007; Payne and Eppstein, 2006, 2007, 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 et al., 2005b).

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 (2.2) in Section 2.2). Under both synchronous and asynchronous updating policies, Giacobini et al. (2005b) 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 and Iwasa, 2000; Satō et al., 1994; Van Baalen, 2000; Van Baalen and Rand, 1998) 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 neighbor-

hood 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.

2.2 Methods

2.2.1 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 et al., 2005b; 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 .

2.2.2 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:

$$\phi = \frac{\#triangles}{\#triplets} = \frac{\text{trace}(A^3)}{\|A^2\| - \text{trace}(A^2)} \quad (2.1)$$

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,

$$radius_N = \sqrt{\frac{1}{k+1} \left(\sum_{i=1}^{k+1} (x_i - \bar{x})^2 + \sum_{i=1}^{k+1} (y_i - \bar{y})^2 \right)} \quad (2.2)$$

where

$$\bar{x} = \frac{1}{k+1} \sum_{i=1}^{k+1} x_i, \quad \bar{y} = \frac{1}{k+1} \sum_{i=1}^{k+1} y_i \quad (2.3)$$

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 (2.2) and (2.3) 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 $radius_G$:

$$\rho = \frac{radius_N}{radius_G} \quad (2.4)$$

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.

2.2.3 Population structures

A regular graph is one in which every vertex has the same degree k . In this study, we investigate takeover dynamics on ten 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 2.1, and their relevant structural characteristics are provided in Table 2.2a. Note that Table 2.2a 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 2.2b.

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 2.2a). 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 2.1a) and Moore neighborhoods ($M3$, Table 2.1a) are the most frequently employed interaction networks in spatially structured evolutionary algorithms (*e.g.*, (Giacobini et al., 2005b; Sarma and De Jong, 1996)). We also investigated common variations of these with larger interaction neighborhoods; specifically, Moore neighborhoods with diameter 5 ($M5$, Table 2.1a) and 7 ($M7$, not shown) and extended Von Neumann neighborhoods ($VN2$, Table 2.1a, and $VN3$, not shown). The triangular population structures considered herein (T , AT , and ST , Table 2.1b) 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).

Table 2.1: 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.

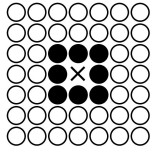
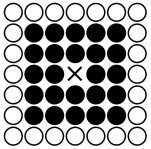
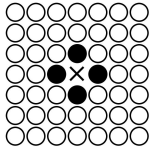
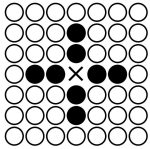
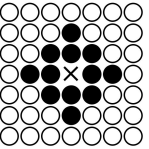
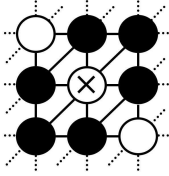
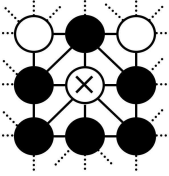
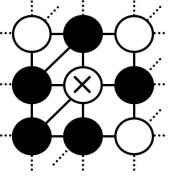
| (a) | | | | |
|---|---|--|--|--|
| <i>M3</i> | <i>M5</i> | <i>VN1</i> | <i>VN2</i> | <i>MVN</i> |
| Moore ($d = 3$), <i>C9</i> | Moore ($d = 5$) | Von Neumann, <i>L5</i> | <i>L9</i> | <i>C13</i> |
|  |  |  |  |  |
| (b) | | | | |
| <i>T</i> | <i>AT</i> | <i>ST</i> | | |
| Triangular | Alternating Triangular | Semi Triangular | | |
|  |  |  | | |

Table 2.2: (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 | $radius_N$ |
|----------------------|-----|--------|------------|
| <i>VN1</i> | 4 | 0 | 0.89 |
| <i>ST</i> | 5 | 0.3 | 0.97 |
| <i>AT</i> | 6 | 0.4 | 1.03 |
| <i>T</i> | 6 | 0.4 | 1.07 |
| <i>M3</i> | 8 | 0.43 | 1.16 |
| <i>MVN</i> | 12 | 0.46 | 1.47 |
| <i>VN2</i> | 8 | 0.21 | 1.49 |
| <i>M5</i> | 24 | 0.52 | 2.00 |
| <i>VN3</i> | 12 | 0.27 | 2.08 |
| <i>M7</i> | 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 |

2.2.4 Takeover time

Consider a population with only two levels of fitness; *i.e.*, 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 :

$$N_t = \frac{1}{|V|} \sum_{\forall i \in V} \Lambda_i(t) \quad (2.5)$$

Following Rudolph (2000), we define the takeover time $T = \min\{t | N_t = 1\}$ 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

$$E[T] = \frac{1}{|V|} \sum_{\forall i \in V} E_i[T] \quad (2.6)$$

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

2.2.5 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

$$P_{sel} = \frac{x}{k} \quad (2.7)$$

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}):

$$P_{rep} = p_{up}P_{sel} \tag{2.8}$$

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

2.2.6 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.

2.2.7 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

$$S \xrightarrow{p_{up}} I \xrightarrow{g} S \tag{2.9}$$

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)

$$\tau = \frac{p_{up}}{k} \quad (2.10)$$

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):

$$P_{inf} = \tau x \quad (2.11)$$

By substituting equation (2.10) into equation (2.11) and equation (2.7) into equation (2.8), it is clear that the probability of infection in the SIS model (P_{inf} , equation (2.11)) is equivalent to the replacement probability in our formulation of takeover time analysis (P_{rep} , equation (2.8)). 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.

2.2.8 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):

$$[X] = \frac{1}{k} \sum_w [XW] \quad (2.12)$$

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):

$$[XYZ] \approx \frac{(k-1)[XY][YZ]}{\sum_w [YW]} = \frac{(k-1)}{k} \frac{[XY][YZ]}{[Y]} \quad (2.13)$$

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 pair-wise 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 2.1a, $M\beta$). 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 (2.13), $[101]$ (*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, 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 (2.13) 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 (2.13), 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.* $M\beta$), 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 (2.1), also known as the clustering coefficient) can be incorporated directly into the closure method, as follows (Keeling, 1999):

$$[XYZ] \approx \frac{(k-1)}{k} \frac{[XY][YZ]}{[Y]} \left((1-\phi) + \frac{\phi\mu}{k} \frac{[XZ]}{[X][Z]} \right) \quad (2.14)$$

Thus, the closure method of equation (2.14) 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 (2.14) reduces to equation (2.13). Spatial structure beyond triplets, however, is not considered in equation (2.14).

2.2.9 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 3 potential states, and the PA developed by Keeling (1999) thus requires a total of 5 coupled differential equations. However, in the reformulation of takeover time analysis in terms of the SIS model (equation (2.9)) each vertex has only 2 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:

$$\begin{aligned}
 \frac{d[00]}{dt} &= c(-\tau[001] + g[01] + g^2[11]) \\
 \frac{d[01]}{dt} &= 0.5c\left(\tau([001] - [101] - [01]) + g([11] - [01])\right) \\
 \frac{d[11]}{dt} &= c(\tau([101] + [01]) - g[11] - g^2[11])
 \end{aligned}
 \tag{2.15}$$

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-equilibrium dynamics*, then the value of c becomes important. As will be shown in Section 2.3.1, our results clearly demonstrate that the optimal choice of c is a function of ρ (equation (2.4)), which depends on both the local neighborhood of interactions and the population size.

In this study, all occurrences of $[XYZ]$ in equation (2.15) are computed using equation (2.14), τ from equation (2.10), and the parameterized coefficient c from equation (2.24), which we present in Section 2.3.1. Thus, the PA employed herein is parameterized by the

population size (μ), the ratio (ρ , equation(2.4)) of $radius_N$ to $radius_G$, the vertex degree (k), the ratio of closed triplets to total triangles (ϕ , equation (2.1)), the reversion probability (g), and the uptake probability (p_{up}). For small reversion probabilities the effect of g^2 in equation (2.15) 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.

2.2.10 Experimental design

For each combination of the ten population structures (Table 2.2a) and eight population sizes (Table 2.2b), 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 et al., 2005b; Payne and Eppstein, 2007, 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 (2.15)) *via* numerical integration, using a Runge-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).

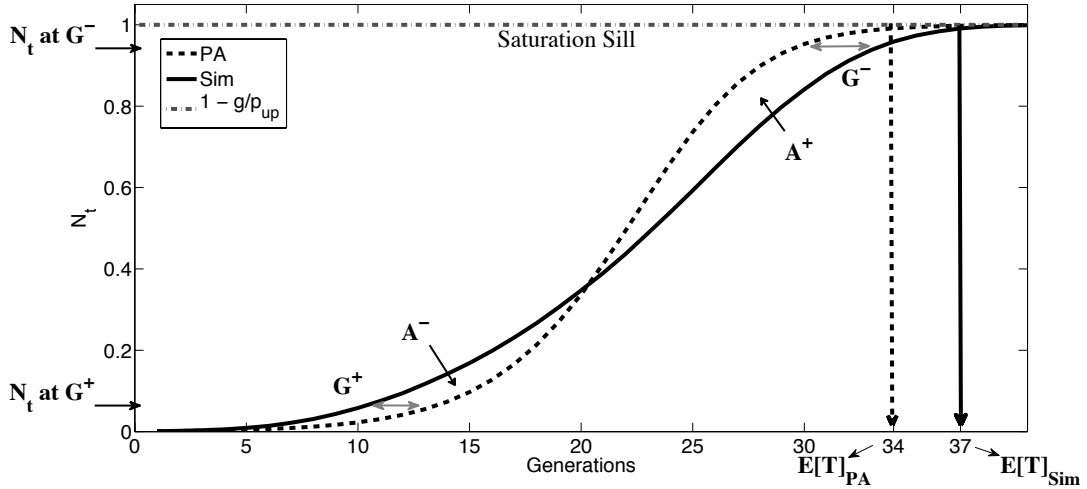


Figure 2.1: Error metrics used in this study are depicted in a schematic diagram of representative takeover curves approximated by the PA (dashed black lines) and observed as the average of 50 direct simulations (Sim, solid black lines). G represents generation error, A represents area error, and $E[T]$ represents expected takeover time. See text for details.

2.2.11 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 2.1, dashed black line) and observed through direct simulation (Figure 2.1, 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 to discrete system dynamics, N_t only asymptotically approaches 1 (when $g = 0$). Consequently, the expected takeover time of the simulation data (Figure 2.1,

$E[T]_{Sim}$) and the PA (Figure 2.1, $E[T]_{PA}$) 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 2.1, 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 2.1, A^+) where the PA is predicting a more rapid spread. 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

$$A_{Sim} = \int_{t=0}^{t_2} Sim(t)dt \quad (2.16)$$

$$A^- = \int_{t=0}^{t_1} PA(t)dt - \int_{t=0}^{t_1} Sim(t)dt \quad (2.17)$$

$$A^+ = \int_{t=1}^{t_2} PA(t)dt - \int_{t=1}^{t_2} Sim(t)dt \quad (2.18)$$

and the area error metrics are defined as:

$$NegativeAreaError = \frac{A^-}{A_{Sim}} \times 100\% \quad (2.19)$$

$$PositiveAreaError = \frac{A^+}{A_{Sim}} \times 100\% \quad (2.20)$$

After fitting spline curves to both the PA and simulation data, equations (2.16-2.18) 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:

$$|AreaError| = |NegativeAreaError| + |PositiveAreaError| \quad (2.21)$$

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 2.1, 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 2.1, 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,

$$GenerationError = \frac{\max\{|G^-|, G^+\}}{E[T]_{Sim}} \times 100\% \quad (2.22)$$

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,

$$SillError = \frac{S_{PA} - S_{Sim}}{S_{Sim}} \times 100\% \quad (2.23)$$

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

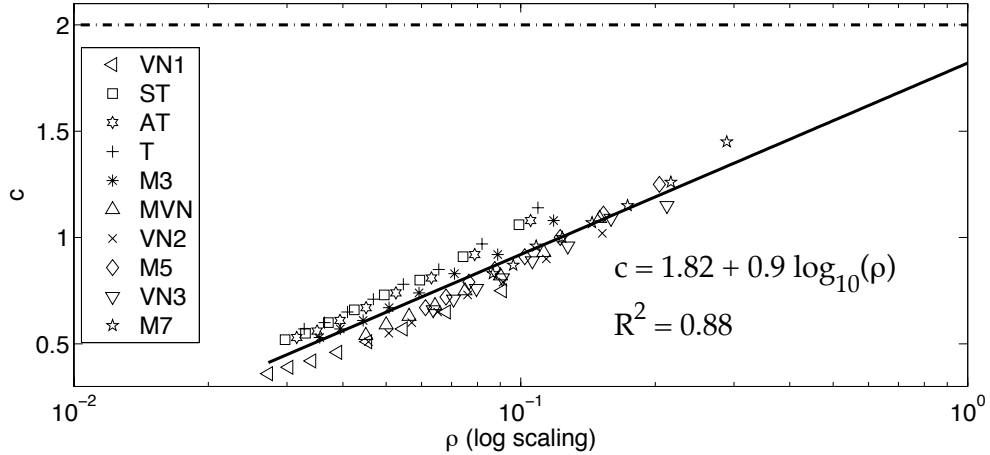


Figure 2.2: The values of c that minimized generation error as a function of ρ , along with the best logarithmic fit. The dash-dot horizontal line, representing $c = 2$, is provided as reference.

2.3 Results

2.3.1 Accounting for ρ in the PA

In order to determine the values of c needed for equation (2.15), 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 (2.22)) as the empirically determined optimal value for c . In Figure 2.2, 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

$$c \approx 1.82 + 0.9 \log_{10}(\rho) \quad (2.24)$$

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 (2.24) was used in

all subsequent experiments (rather than the optimally determined values of c) to assess the accuracy of the PA (equation (2.15)) with the c parameterized by ρ . The value of $c = 2$, used in previous studies, is shown for reference (Figure 2.2, horizontal dash-dot line); in all cases, the optimal coefficient values for c were well below 2.

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.*

2.3.2 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 2.3 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 ($R^2 > 0.95$

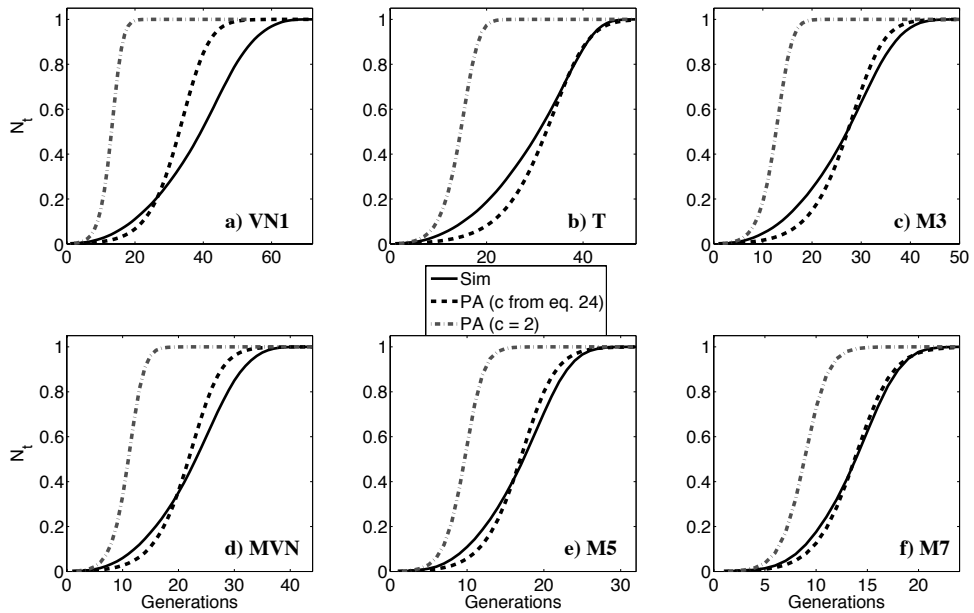


Figure 2.3: Takeover dynamics as predicted by the PA using equation (2.24) (dashed black lines) and as observed through direct simulation (solid black lines) on six regular population structures using $\mu = 1024$, $p_{up} = 1$, and $g = 0$. For reference, we also show the dynamics predicted by the PA 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.

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 (Giacobini et al., 2005b; Gorges-Schleuter, 1999) 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 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 ($p < 0.001, \chi^2$). 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 2.3c-f). (In contrast, the curves resulting from the PA using $c = 2$, shown in Figure 2.3 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 (2.24) to estimate c introduces a small amount of error, sometimes causing the curves to intersect a little too early (as for $VN1$, Figure 2.3a), or a little too late (as for T , Figure 2.3b), thus resulting in an over- or under-estimate 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 2.4.

Area error and generation error

Both area error (Figure 2.4a) and generation error (Figure 2.4b) were found to decrease approximately exponentially as a function of ρ . The average goodness of fit was $R^2 = 0.73$ between the logarithm of area error and ρ and $R^2 = 0.79$ between the logarithm of generation

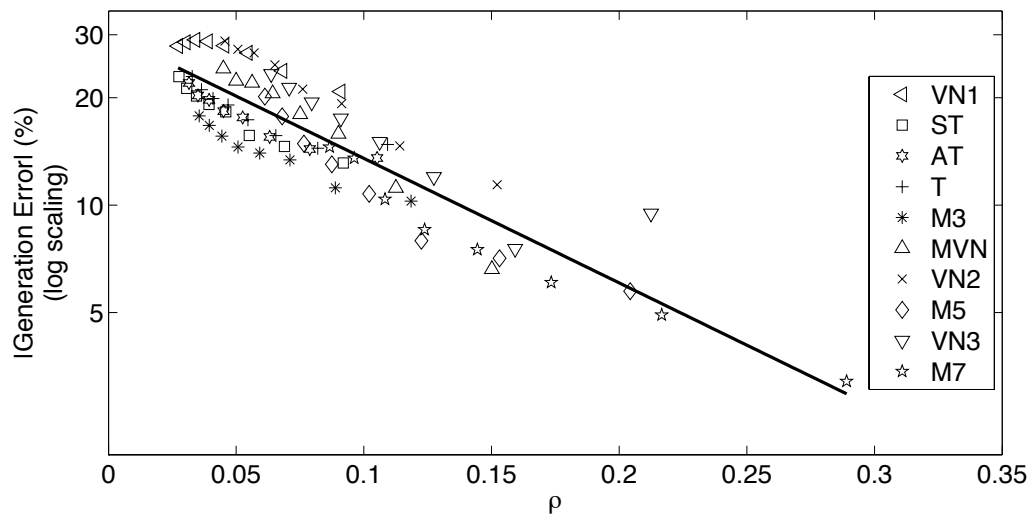
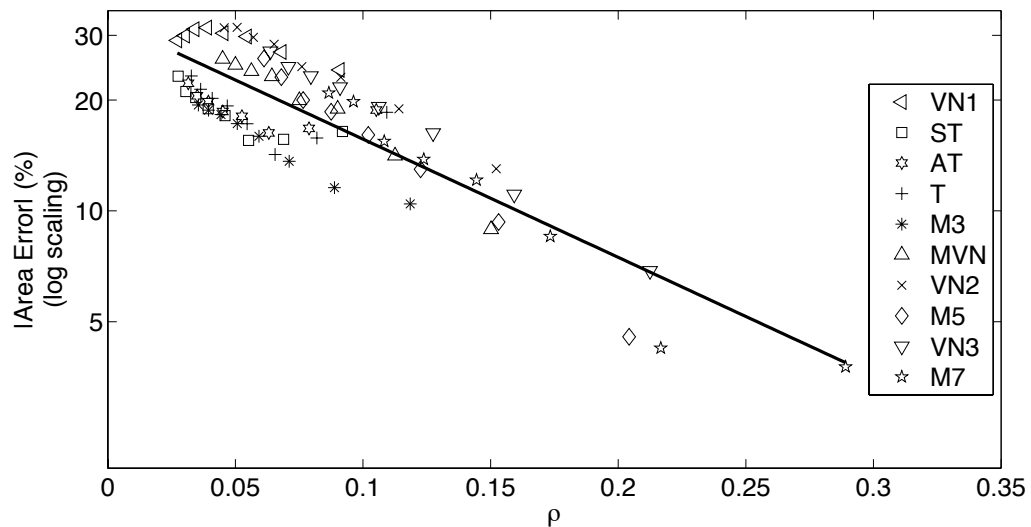


Figure 2.4: (a) Area error and (b) generation error as a function of (ρ) , for all population sizes and population structures, with $p_{up} = 1$ and $g = 0$. The solid line in each figure depicts the best exponential fit to the data for all population structures, and is provided as a visual aid only. Note the logarithmic scale on the vertical axis.

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 2.4. 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 2.5 depicts the 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 2.5a) to 18% (Figure 2.5d). 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.

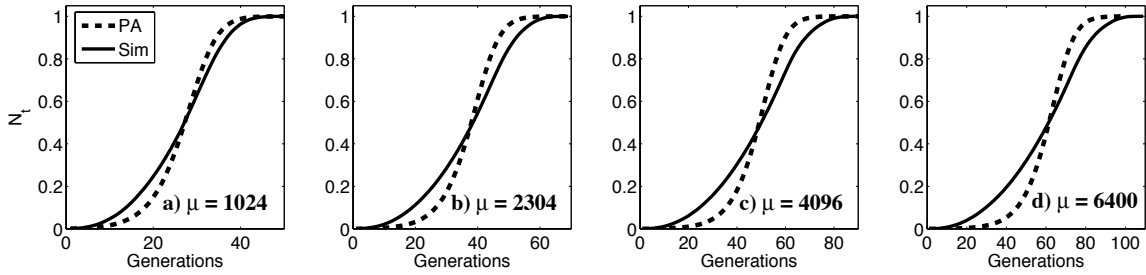


Figure 2.5: Influence of population size on the accuracy of the PA. Takeover dynamics as estimated by the PA (dashed line) and as observed through direct simulation (solid line) on the $M3$ population structure, with a population size of (a) $\mu = 1024$, (b) $\mu = 2304$, (c) $\mu = 4096$, and (d) $\mu = 6400$

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 2.6). 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 2.6a-c and 2.6d-f, respectively), shifting the PA curve more to the right, relative to the curve observed through simulation (*e.g.*, compare Figures 2.6a and 2.6c). 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 2.6a-c) and 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 2.6d-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 $\mu = 1024$.

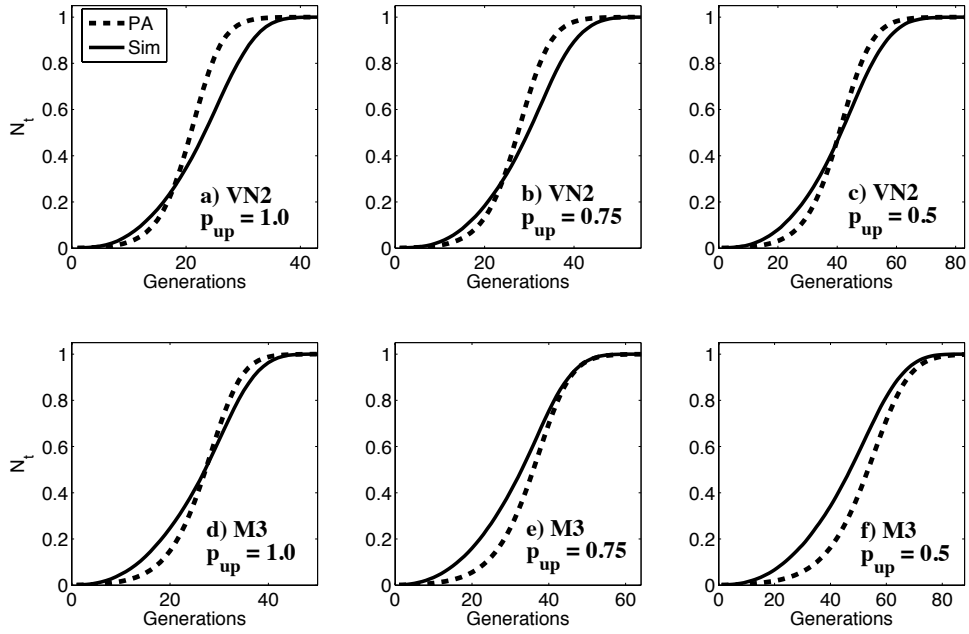


Figure 2.6: Influence of uptake probabilities p_{up} on the accuracy of the PA. Representative takeover dynamics as estimated by the PA (dashed line) and as observed through direct simulation (solid line) on the (a-c) *VN2* and (d-f) *M3* population structures with $\mu = 1024$, $g = 0$. The three plots 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 2.3: 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 (%) | | | | | |
|----------------------|--------------------------------|-----------|-----------------|-----------|----------------|-----------|
| | $p_{up} = 1$ | | $p_{up} = 0.75$ | | $p_{up} = 0.5$ | |
| | $g = 0.05$ | $g = 0.1$ | $g = 0.05$ | $g = 0.1$ | $g = 0.05$ | $g = 0.1$ |
| <i>VN1</i> | 2.50 | 5.00 | 6.25 | 15.25 | 12.00 | 26.50 |
| <i>ST</i> | 2.75 | 4.25 | 4.25 | 11.75 | 8.75 | 16.50 |
| <i>AT</i> | 4.50 | 6.25 | 5.50 | 12.75 | 7.00 | 18.75 |
| <i>T</i> | 2.75 | 10.25 | 5.25 | 8.25 | 7.25 | 16.50 |
| <i>M3</i> | 2.00 | 7.50 | 4.00 | 10.00 | 8.75 | 19.75 |
| <i>MVN</i> | 2.25 | 9.00 | 6.50 | 8.75 | 9.00 | 21.75 |
| <i>VN2</i> | 3.00 | 5.50 | 4.50 | 11.75 | 10.25 | 21.50 |
| <i>M5</i> | 1.25 | 6.75 | 4.25 | 13.00 | 8.00 | 18.00 |
| <i>VN3</i> | 3.00 | 6.25 | 5.50 | 11.00 | 6.50 | 20.75 |
| <i>M7</i> | 2.25 | 7.00 | 4.75 | 12.75 | 6.25 | 17.00 |

2.3.3 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 .

Percentage of unsuccessful introductions

Table 2.3 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 2.3 depict the average of the percentages observed on all eight population sizes. Note that data for $g = 0$ are 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 .

Sill error

As g increases, the saturation sill decreases, as shown for representative population structures AT and $M7$ in Figure 2.7a-c and Figure 2.7d-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 (2.23)) 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.

Generation error

Figure 2.7 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 2.7a-c shows the PA becoming increasingly accurate as g increases, whereas Figure 2.7d-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 2.6), 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*.

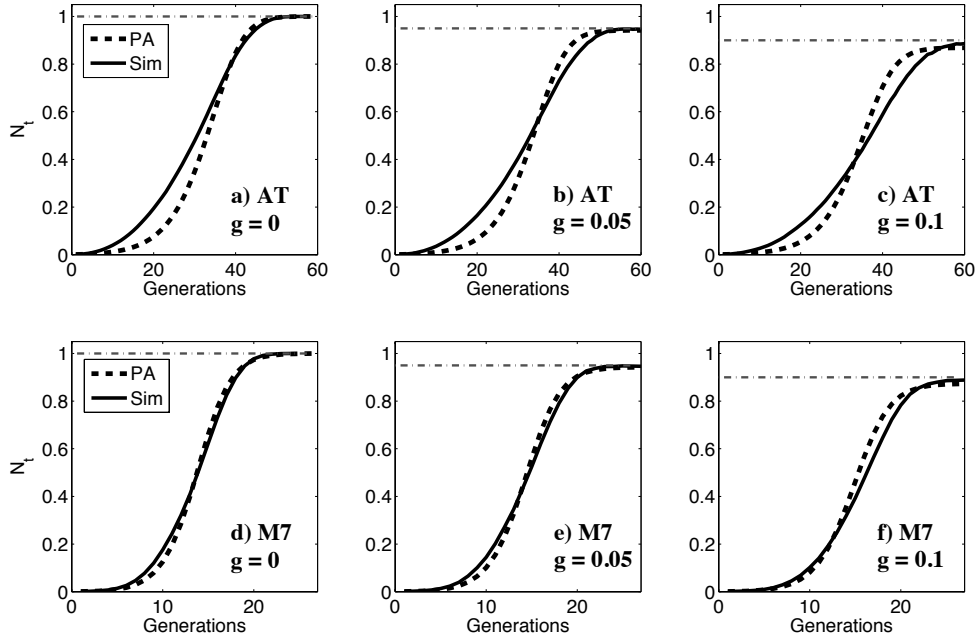


Figure 2.7: Influence of reversion probabilities g on the accuracy of the PA. Representative takeover dynamics with $p_{up} = 1$ and $\mu = 1024$, as estimated by the PA (dashed black line) and as observed through direct simulation (solid black line), for the *AT* population structure (top row, a-c) 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.

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%.

2.4 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 *Suceptible-Infectious-Susceptible* (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 8 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 had 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 manuscript, 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 counter-intuitive 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 (Barabási 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 Prisoners 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 (2007) 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 et al., 2005a; Giacobini et al., 2005b), 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.

2.5 Summary

In summary, takeover time analysis for evolutionary algorithms was reformulated in terms of the well-known *Susceptible-Infectious-Susceptible* (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 ten 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 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.

References

- Altenberg, L. (2005). Evolvability suppression to stabilize far-sighted adaptations. *Artificial Life* 11(4), 427–443.
- Anderson, R. M. and R. M. May (1995). *Infectious Diseases of Humans*. Oxford University Press, Oxford.
- Barabási, A. L. and R. Albert (1999). Emergence of scaling in random networks. *Science* 286, 509–512.
- Bryden, K. M., D. Ashlock, S. Corns, and S. Wilson (2005). Graph based evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 10(5), 550–567.
- Chakraborty, U. K., L. Deb, and M. Chakraborty (1997). Analysis of selection algorithms: a markov chain approach. *Evolutionary Computation* 4(2), 133–167.
- Eames, K. T. D. and M. J. Keeling (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., L. Mielsch, and S. Bornholdt (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 J. Molofsky (2007). Invasiveness in plant communities with feedbacks. *Ecology Letters* 10, 253–263.
- Falconer, D. S. and T. F. C. Mackay (1996). *Quantitative Genetics*. Pearson Education Limited, London.
- Giacobini, M., M. Preuss, and M. Tomassini (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. Springer-Verlag, Heidelberg.
- Giacobini, M., M. Tomassini, and A. Tettamanzi (2005a). Takeover times curves in random and small-world structured populations. In H. G. Beyer (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2005*, pp. 1333–1340. ACM Press, New York.
- Giacobini, M., M. Tomassini, A. Tettamanzi, and E. Alba (2005b). Selection intensity in cellular evolutionary algorithms for regular lattices. *IEEE Transactions on Evolutionary Computation* 9(5), 489–505.
- Goldberg, D. and K. Deb (1991). A comparative analysis of selection schemes used in genetic algorithms. In G. Rawlins (Ed.), *Foundations of Genetic Algorithms*, pp. 69–93. Morgan Kaufmann, San Francisco.
- 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. Morgan Kaufmann, San Francisco.
- Hauert, C. and M. Doebeli (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. H. (1975). *Adaptation in Natural and Artificial Systems*. Ph. D. thesis, University of Michigan.
- Joo, J. and J. L. Lebowitz (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 K. T. D. Eames (2005). Networks and epidemic models. *Journal of the Royal Society Interface* 2, 295–307.
- Kerr, B., M. A. Riley, M. W. Feldman, and B. J. M. Bohannan (2002). Local dispersal promotes biodiversity in a real life game of rock-paper-scissors. *Nature* 418, 171–174.
- Kirley, M. and R. Stewart (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. ACM Press, New York.
- Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg (2001). The web of human sexual contacts. *Nature* 411, 907–908.
- Matsuda, H., N. Ogita, A. Sasaki, and K. Soto (1992). Statistical mechanics of population: the lattice lotka-volterra model. *Progress of Theoretical Physics* 88, 1035–1049.
- Meyers, L. A., M. E. J. Newman, and B. Pourbohloul (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., C. Hauert, E. Lieberman, and M. A. Nowak (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502–505.
- Ong, N. P. and R. J. Cava (2004). Electronic frustration on a triangular lattice. *Science* 305, 52–53.
- Pastor-Satorras, R. and A. Vespignani (2001). Epidemic spreading in scale-free networks. *Physical Review Letters* 86(4), 3200–3203.
- Payne, J. L. and M. J. Eppstein (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. ACM Press, New York.
- Payne, J. L. and M. J. Eppstein (2007). Takeover times on scale-free topologies. In D. Thierens (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2007*, pp. 308–315. ACM Press, New York.
- Payne, J. L. and M. J. Eppstein (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*, pp. 241–248. ACM Press, New York.
- Payne, J. L., M. J. Eppstein, and C. J. Goodnight (2007). Sensitivity of self-organized speciation to long-distance dispersal. In *Proceedings of the IEEE Symposium on Artificial Life*, pp. 1–7.
- Petermann, T. and P. De Los Rios (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 Y. Bar-Yam (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. K. Lai, O. Katai, M. Gen, and B. Lin (Eds.), *Proceedings of the Second Asia-Pacific Conference on Genetic Algorithms and Applications, APGA-2000*, pp. 144–151. Global Link Publishing Company, Hong Kong.
- 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. Springer-Verlag, Heidelberg.
- Santos, F. C. and J. M. Pacheco (2005). Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters* 95, 098104.
- Sarma, J. and K. De Jong (1996). An analysis of the effect of neighborhood size and shape on local selection algorithms. In H. M. Voigt, W. Ebeling, I. Rechenberg, and H. P. Schwefel (Eds.), *Parallel Problem Solving from Nature*, pp. 236–244. Springer-Verlag, Heidelberg.
- Satō, K. and Y. Iwasa (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 University Press, Cambridge.
- Satō, K., H. Matsuda, and A. Sasaki (1994). Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology* 32, 251–268.
- Sayama, H., L. Kaufman, and Y. Bar-Yam (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. Morgan Kaufmann, San Francisco.
- Van Baalen, M. (2000). Pair approximations for different spatial geometries. In U. Dieckmann, R. Law, and J. A. J. Metz (Eds.), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, pp. 359–387. Cambridge University Press, Cambridge.
- Van Baalen, M. and D. A. Rand (1998). The unit of selection in viscous populations and the evolution of altruism. *Journal of Theoretical Biology* 193, 631–648.
- Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of small-world networks. *Nature* 393, 440–442.
- Werfel, J. and Y. Bar-Yam (2004). The evolution of reproductive restraint through social communication. *Proceedings of the National Academy of Science* 101(30), 11019–11024.
- Whitacre, J. M., R. A. Sarker, and Q. T. Pham (2008). The self-organization of interaction networks for nature-inspired optimization. *IEEE Transactions on Evolutionary Computation* 12(2), 220–230.

Chapter 3

Evolutionary Dynamics on Scale-Free Interaction Networks

Abstract

There has been a recent surge of interest in studying dynamical processes, including evolutionary optimization, on scale-free topologies. While various scaling parameters and assortativities have been observed in natural scale-free interaction networks, most previous studies of dynamics on scale-free graphs have employed a graph generating algorithm that yields a topology with an uncorrelated degree distribution and a fixed scaling parameter. In this contribution, we advance the understanding of selective pressure in scale-free networks by systematically investigating takeover times under local uniform selection in scale free topologies with varying scaling exponents, assortativities, average degrees, and numbers of vertices. We demonstrate why the so-called “characteristic path length” of a graph is a nonlinear function of both scaling and assortativity. Neither the eigenvalues of the adjacency matrix nor the effective population size were sufficient to account for the variance in takeover times over the parameter space that was explored. Rather, we show that 97% of the variance of logarithmically transformed average takeover times, on all scale-free graphs

tested, could be accounted for by a planar function of (i) the average inverse degree (which captures the effects of scaling) and (ii) the logarithm of the population size. Additionally, we show that at low scaling exponents, increasingly *positive* assortativities increased the variability between experiments on different random graph instances, while increasingly *negative* assortativities increased the variability between takeover times from different initial conditions on the same graph instances. We explore the mechanisms behind our sometimes counter-intuitive findings, and discuss potential implications for evolutionary computation and other relevant disciplines.

3.1 Introduction

The behavior of a complex adaptive system is governed by the collective dynamics of its interacting system components. Consequently, the topological characteristics of the interaction network that specifies which components can interact with one another have a pronounced influence on the rate of information flow throughout the system, and thus play a critical role in determining emergent system-wide dynamics. For example, the evolution of altruism (Werfel and Bar-Yam, 2004) and cooperation (Hauert and Doebeli, 2004; Nowak and May, 1992; Ohtsuki et al., 2006; Santos and Pacheco, 2005), species invasiveness (Eppstein and Molofsky, 2007), disease propagation (Keeling, 1999; Newman, 2002; Pastor-Satorras and Vespignani, 2001), energy transfer in food webs (Borrett et al., 2007), predator-prey dynamics (Rauch and Bar-Yam, 2006), the maintenance of genetic diversity (Kerr et al., 2002; Reichenbach et al., 2007; Sayama et al., 2003), the suppression of evolutionary pathologies (Altenberg, 2005), and the self-organization of barriers to gene flow (Payne et al., 2007) have all been shown to be highly dependent on the topological properties of the underlying interaction network.

In the context of evolutionary computation, interaction networks are often utilized as population structures. This was initiated in the design of parallel evolutionary algorithms (for a review of this topic, please see Alba and Tomassini (2002)), where populations are

typically structured into subpopulations that are distributed onto separate processing units. Mating interactions are usually panmictic within the subpopulations, with periodic migration occurring between subpopulations according to an explicit subpopulation interaction topology (*e.g.*, all-to-all, grid-based, etc.) (Cantú-Paz, 2000; Folino et al., 2001, 003a; Herrera and Lozano, 2000; Lienig, 1997; Van Veldhuizen et al., 2003). While population structures can thus be exploited for computational efficiency of parallel implementations of evolutionary algorithms, there has also been an increasing interest in exploiting different population structures to improve the effectiveness of evolutionary search. In cellular evolutionary algorithms, populations are structured on low-order regular graphs, such as a one-dimensional (1D) or two-dimensional (2D) lattice, and mating events are restricted to occur within spatially-localized, overlapping neighborhoods. In these regular population structures, the imposition of spatial constraints on recombination events has been shown to improve the maintenance of genetic diversity, in part by retarding the flow of advantageous alleles and consequently reducing the selective pressure (Giacobini et al., 2005b). While 1D and 2D lattice population structures are the most commonly used interaction networks for evolutionary optimization, the utilization of other regular graph structures, such as the generalized Peterson graph and the complete bipartite graph, have recently been investigated as well (Bryden et al., 2005), wherein the performance improvements obtained on each graph structure were shown to be problem dependent.

In contrast to these regular population structures, the interaction networks of numerous natural populations have been found to be heterogeneous (*i.e.*, different nodes in the interaction network have different numbers of connections). Of particular interest are scale-free networks (Barabási and Albert, 1999), a class of highly heterogeneous graphs, which have been shown to be quite ubiquitous in natural systems, ranging from networks of social interactions, such as email networks (Ebel et al., 2002) and sexual contacts (Liljeros et al., 2001), to cellular systems, such as metabolic networks (Jeong et al., 2000; Ravasz et al., 2002) and protein-protein interactions (Jeong et al., 2001). Accordingly, there has been a recent surge

of interest in studying dynamical processes on scale-free topologies; examples include the saturation dynamics of infectious disease in epidemiological models (Pastor-Satorras and Vespignani, 2001), information cascades in binary decision models (Watts, 2002), the emergence of cooperative behavior in evolutionary games (Poncela et al., 2007; Rong et al., 2007; Santos and Pacheco, 2005), and evolutionary optimization (Gasparri et al., 2007; Giacobini et al., 2006; Kirley and Stewart, 2007a,b). In particular, scale-free population structures have been analyzed in the context of genetic algorithms with self-adaptive mutation (Giacobini et al., 2006) and multi-objective optimization (Kirley and Stewart, 2007a,b), and have been applied to the localization problem in robotics (Gasparri et al., 2007), with each study demonstrating varying degrees of success. For example, while Giacobini et al. (2006) found that populations evolving on scale-free topologies were unable to outperform panmictic populations on a variety of benchmark optimization problems, Gasparri et al. (2007) found that scale-free population structures enhanced the genetic algorithm’s ability to solve both the localization and kidnap problems in a mobile robotics application. Such mixed results were also reported by Kirley and Stewart (2007a,b). While populations evolving on scale-free topologies were outperformed by populations evolving on random topologies on a two-objective problem (in terms of convergence speed and spread of solutions across the Pareto-front) (Kirley and Stewart, 2007b), populations evolving on scale-free population structures were shown to outperform populations evolving on random, small-world (Watts and Strogatz, 1998), and regular lattice population structures as the number of objectives increased, on specific multi-objective problems (Kirley and Stewart, 2007a).

Scale-free topologies exhibit a power-law distribution of vertex connectivity, such that the probability $p(k)$ of having a vertex of degree k is of the form $p(k) \propto k^{-\gamma}$, where γ is referred to as the scaling parameter. Various scaling exponents have been measured in natural systems, *e.g.*, $\gamma \sim 2.4$ for the number of species per genus of mammals and $\gamma \sim 3.1$ for the protein-protein interaction network of *S. cerevisiae* (Clauset et al., 2009). The propensity with which vertices of similar degree (*i.e.*, number of connections) are

connected to one another is referred to as the assortativity (r) of the network (Newman, 2002). Some scale-free interaction networks (*e.g.*, societal interactions) exhibit “positive assortativity” (Newman, 2002), where vertices of high degree are more likely to be connected to one another than to vertices of low degree. Other scale-free networks (*e.g.*, protein-protein interactions), exhibit “negative assortativity” (Newman, 2002), where nodes of high degree are more likely to connect to nodes of low degree. Networks in which there is no relationship between the degree of adjacent vertices, such as those produced by the preferential attachment algorithm provided by Albert and Barabási (AB) (Barabási and Albert, 1999), are referred to as “uncorrelated.”

While various scaling parameters and assortativities have been observed in natural systems, most previous studies of dynamics on scale-free graphs (including the evolutionary optimization studies cited above) were generated using the AB algorithm, which yields uncorrelated topologies ($r \sim 0$) (Newman, 2002) with a scaling parameter that approaches $\gamma \sim 3$ (Barabási and Albert, 1999), as the number of vertices approaches infinity. However, a few recent studies have shown that both scaling and assortativity exert important influences on graph-based dynamical processes. For example, the equilibrium proportion of cooperators in the prisoner’s dilemma played on scale-free networks has been shown to decrease as the scaling exponent γ increases (Poncela et al., 2007), or as the assortativity deviates from an uncorrelated mixing pattern ($r \neq 0$) (Rong et al., 2007). In the unbiased Voter Model, recent results (Sood et al., 2008) have demonstrated that the expected consensus time on scale-free graphs increases as the scaling exponent increases, and that assortativity has a smaller secondary effect on the amount of time required to reach consensus.

In order to better understand the potential for evolutionary optimization on scale-free graphs, it is important to understand how both scaling and assortativity affect fundamental system dynamics. Such an understanding may also provide insight into other types of dynamic processes on scale-free networks, including spread of infectious disease and dissemination of fads and ideas.

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”) (Goldberg and Deb, 1991). Higher takeover times imply lower selective pressure, and *vice versa*. Takeover times have been previously investigated and modeled in several regular population structures. Goldberg and Deb (1991) investigated saturation dynamics in panmictic population structures (*i.e.*, complete graphs) under a variety of selection mechanisms and showed that takeover is quite rapid in such well-mixed systems. Rudolph (2000) provided exact analytical solutions for expected saturation times on ring structures (*i.e.*, 1D toroidal lattice) and lower and upper bounds for array structures (*i.e.*, 1D non-toroidal lattice). Sarma and De Jong (1996) investigated takeover times in 2D toroidal lattices with neighborhoods of various shapes and sizes and showed that selective pressure is largely governed by the radius of the local mating neighborhood and Giacobini et al. (2005b) provided mathematical models of takeover dynamics in 1D and 2D (with Von Neumann neighborhoods) toroidal lattices, using synchronous and asynchronous updating policies.

Analysis of takeover times in irregular population structures has received considerably less attention. Giacobini et al. (2005a) provided analytical approximations of takeover dynamics in random graphs and empirical results for small-world topologies (Watts and Strogatz, 1998). Their results demonstrated that random graphs induce a selective pressure qualitatively similar to panmixia and that the selective pressures induced by small-world topologies approach that of a random graph as the probability of creating long-distance interactions increases. In the same study (Giacobini et al., 2005a), they showed that the average selective pressures induced by the AB scale-free population structures they employed were at least as strong as those induced by random graphs and that, when the initial copy of the high fitness individual was strategically placed in a highly connected vertex of an AB graph, takeover was even faster than with panmixia.

In a preliminary study (Payne and Eppstein, 2007), we measured takeover times using

a variety of scale-free graph generating algorithms (Barabási and Albert, 1999; Holme and Kim, 2002; Ravasz et al., 2002) with various degrees of clustering, modularity, and hierarchical organization (which resulted in a variety of scaling exponents and assortativities, although these parameters were neither systematically varied nor explicitly reported there). We showed that selective pressures on different types of scale free graphs vary from very high levels, comparable to those induced by random mixing, to very low levels that are even weaker than those induced by nearest neighbor interactions (Moore neighborhoods). In a follow-up study, we investigated the independent and combined influence of scaling (γ) and assortativity (r) on takeover times in scale-free topologies at a single population size, with fixed average degree (Payne and Eppstein, 2008). In that work, we found that takeover times exhibited a nonlinear negative correlation with the scaling exponent and a nonlinear positive correlation with assortativity, with additional nonlinear interactions between these two topological properties. Further, we showed that with low scaling and high positive assortativity, takeover times were much less dependent on the degree of initial placement than on uncorrelated (Giacobini et al., 2005a; Payne and Eppstein, 2008) and negatively assortative (Payne and Eppstein, 2008) scale-free graphs.

Ideally, expected takeover times in arbitrary population structures (regular or irregular) could be simply predicted using readily pre-computable metrics of the underlying interaction network. In related application domains, such as statistical physics and ecology, simple functional relationships have been determined between the structural properties of the interaction networks and the dynamic properties of spreading behavior. For example, consensus time in the unbiased Voter Model has recently been shown to be a function of the population size and the first and second moments of the degree distribution (Sood et al., 2008). Investigations in heterogeneous ecological networks have shown that *indirect* pathways are an important governing influence in graph-based dynamical processes (Borrett et al., 2007; Caswell, 2001). Specifically, the leading eigenvalue of the network adjacency matrix has been shown to be a good measure of the rate with which the number of paths

between vertices grows as a function of path length (Caswell, 2001) and this has been shown to have a pronounced influence on the rate of flow of matter and energy throughout ecosystems (Borrett et al., 2007). In the context of evolutionary computation, expected takeover time on regular graphs is known to be a positively correlated linear function of the characteristic path length (Payne and Eppstein, 2007; Rudolph, 2000), a metric which quantifies the mean shortest distance between all pairs of vertices. In addition, we have recently shown that differential equation based analytical methods, based on pair approximations, can be parameterized by the ratio of the local neighborhood radius to the global lattice radius in order to rapidly estimate pre-equilibrium takeover dynamics in grid-based regular population structures (Payne and Eppstein, 2009). The problem has proved more elusive in the case of scale-free topologies with stochastic, degree-dependent update policies. In (Payne and Eppstein, 2007), we suggested average takeover times on scale-free topologies were positively linearly correlated with a combination of the maximum and variance of the all-pairs shortest path lengths. However, our subsequent work (Payne and Eppstein, 2008) showed that takeover times were actually a logarithmic function of the metric presented in (Payne and Eppstein, 2007) in uncorrelated and positively assortative scale-free graphs, with the slopes varying as a function of assortativity, while takeover times in negatively assortative scale-free graphs were completely independent of this metric. Conversely, while average takeover time was shown to be an ambiguous multifunction of the characteristic path length in uncorrelated and positively assortative scale-free graphs, characteristic path length was found to have a *negative* nonlinear correlation with takeover times in negatively assortative scale-free graphs (a counter-intuitive result, given that this correlation is *positive* in regular graphs).

The results of our previous studies thus indicate that metrics based on all-pairs shortest direct paths are not sufficient to explain the rate of information flow in scale-free networks with stochastic, degree-dependent update policies. In this research, we seek to understand why this observation is true and to determine if there are other static network properties that

govern takeover times on scale-free topologies with varying scaling exponents, assortativities, average degrees, and population sizes (*i.e.*, number of vertices). We significantly extend our previous studies (Payne and Eppstein, 2007, 2008) on takeover times under local uniform selection in scale-free interaction networks. In Section 3.2, we explain how we generated and employed scale-free graphs with systematically varying scaling exponents, assortativities, average degrees, and population sizes. In Section 3.3 we show *why* characteristic path length is a nonlinear non-monotonic function of both scaling and assortativity, and thus does not govern takeover time. We demonstrate that the influence of scaling and assortativity on average takeover times is qualitatively similar under three different stochastic selection mechanisms: uniform selection, binary tournament selection, and linear ranking selection. We then show that the logarithm of average takeover times in scale-free topologies can be described as a planar function of the average inverse degree and the logarithm of population size, and that assortativity exerts a strong influence on the variability in takeover times at low scaling exponents. In Section 3.4, we discuss our findings and their relevance to other disciplines, and suggest ways in which the structural characteristics of scale-free interaction networks can potentially be exploited in evolutionary optimization.

3.2 Methods

3.2.1 Representing Population Structure as a Graph

The population structure of an evolutionary algorithm can be represented as a graph $G = (V, E)$, 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 $|V| = N$, where N is the population size. The graph is undirected, with an edge $\langle i, j \rangle \in E$ for every individual j in the mating neighborhood of individual i , for all $i \in V$. The number of neighbors in the mating neighborhood of individual i , also referred to as the *degree* of vertex i , is denoted by k_i . Throughout this manuscript, the terms

population structure, topology, network, and graph are used synonymously.

3.2.2 Structural Properties of Graphs

When quantifying the structural properties of graphs, there are several metrics of potential interest (*e.g.*, see Newman (2003)). In this section, we present the topological properties considered in this study (for the readers' convenience, we also present a glossary of variables in Appendix 3.5). The distribution of vertex connectivity $p(k)$ is a probability distribution function (PDF) that depicts the frequency with which a node has degree k . The scale-free topologies considered in this study possess a distribution of vertex connectivity of the form:

$$p(k) = Pr(K = k) \propto k^{-\gamma} \quad (3.1)$$

The complementary cumulative distribution function (CCDF), commonly used to visualize power-law distributions (Clauset et al., 2009), depicts the frequency with which nodes have degree greater than or equal to k :

$$P(k) = Pr(K \geq k) \propto k^{-\gamma+1} \quad (3.2)$$

In equations (3.1) and (3.2), the scaling parameter γ affects the shape of the power-law distribution, such that distributions with smaller γ possess heavier tails (Fig. 3.1).

The i^{th} moment (μ_i) of the degree distribution is given by:

$$\mu_i = \sum_k k^i p(k) \quad (3.3)$$

such that μ_1 is the average degree ($\langle k \rangle$) and μ_{-1} is the average inverse degree. It is important to point out that while the average degree (μ_1) is unaffected by changes in the scaling exponent (γ), the average inverse degree (μ_{-1}) changes as a function of both γ and $\langle k \rangle$. Since the graphs considered are connected, μ_{-1} is always well-defined.

A path is defined as a sequence of vertex-edge pairs, beginning at one vertex i and ending at another vertex j ; the direct distance $dist(i, j)$ between any two nodes i and j is defined as the length of the shortest path between i and j . The average individual path

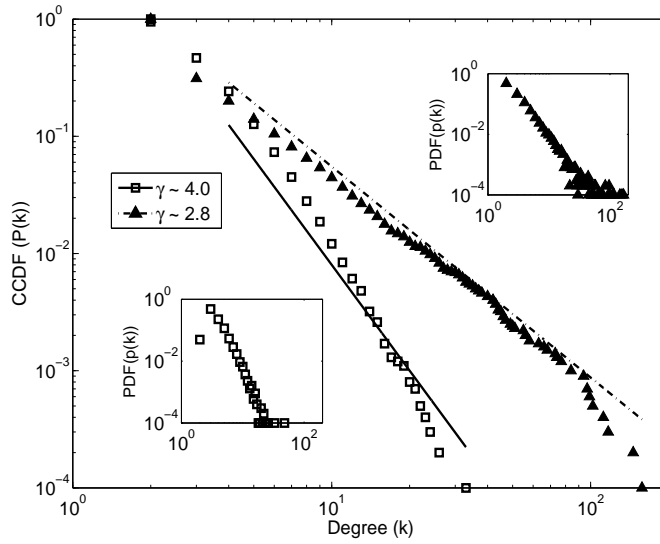


Figure 3.1: Data points show representative complementary cumulative distribution functions (CCDF) of vertex connectivity in randomized scale-free interaction networks created with the GN algorithm (Section 3.2.4), with the smallest and largest scaling exponents used in this study $\gamma \in \{2.8, 4.0\}$ and $\langle k \rangle = 4$. Lines are drawn using the scaling exponents measured using the algorithm described in Section 3.2.4 with a k_{min} of 4, and are deliberately offset in the vertical direction so as not to obscure the data. Data shown pertain to $N = 10,000$. Note the double logarithmic scale. Insets denote the corresponding probability distribution functions (PDF), showing how the observed frequency of occurrence is lower-bounded by $1/N$ in these finite-sized graphs.

length L_i of a vertex i is defined as the mean of the shortest paths between i and all other vertices in the graph:

$$L_i = \frac{1}{|V| - 1} \sum_{\forall j \neq i \in V} dist(i, j) \quad (3.4)$$

The characteristic path length L of a graph G is defined as the mean of the individual path lengths:

$$L = \frac{1}{|V|} \sum_{\forall i \in V} L_i \quad (3.5)$$

The assortativity (r) of a graph G measures the propensity with which vertices of similar degree are connected to one another. Assortativity is formally defined as (Newman, 2002),

$$r = \frac{|E|^{-1} \sum_{\forall \langle i, j \rangle \in E} k_i k_j - \left[|E|^{-1} \sum_{\forall \langle i, j \rangle \in E} \frac{1}{2} (k_i + k_j) \right]^2}{|E|^{-1} \sum_{\forall \langle i, j \rangle \in E} \frac{1}{2} (k_i^2 + k_j^2) - \left[|E|^{-1} \sum_{\forall \langle i, j \rangle \in E} \frac{1}{2} (k_i + k_j) \right]^2} \quad (3.6)$$

which is equivalent to the Pearson correlation coefficient of the degrees of vertices at the opposing ends of an edge (Newman, 2002). A graph is said to be *positively assortative* if $r > 0$, *uncorrelated* if $r = 0$, and *negatively assortative* if $r < 0$.

It is worth noting that the radius (Sarma and De Jong, 1996) of a local interaction neighborhood, which exhibits a strong effect on saturation times in regular population structures (Payne and Eppstein, 2009; Sarma and De Jong, 1996), is not a meaningful metric in the case of scale-free topologies, since these graphs are not embedded in Cartesian space.

3.2.3 Takeover Time Analysis

In order to most directly infer the influence of the structural properties of scale-free interaction networks on the saturation times of advantageous alleles, we minimized (a) the number of different alleles, and (b) the complexity of the selection operator. Specifically, we consider a population with only two levels of fitness (as in (Giacobini et al., 2005b; Payne

and Eppstein, 2007, 2008, 2009; Rudolph, 2000)); *i.e.*, 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 :

$$N_t = \frac{1}{|V|} \sum_{\forall i \in V} \Lambda_i(t) \quad (3.7)$$

Following Rudolph (2000), we define the takeover time $T = \min\{t | N_t = 1\}$ of an experiment to be the first generation in which the most fit genotype fully saturates the population, starting from a single copy of this genotype. 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

$$E[T] = \frac{1}{|V|} \sum_{\forall i \in V} E_i[T] \quad (3.8)$$

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

We tested three stochastic selection mechanisms that have been used in previous takeover time analyses: local uniform selection (Gorges-Schleuter, 1999), binary tournament selection (Giacobini et al., 2005b), and linear ranking selection (Giacobini et al., 2005b). Note that in heterogeneous graphs with a minimum node degree of two, tournaments of larger sizes are not possible for all nodes, and in takeover studies with only two levels of fitness, ranking is arbitrary within nodes of the same fitness level. Our tests verified that the influence of scaling and assortativity on takeover dynamics is qualitatively similar under these three selection mechanisms, (see Section 3.3.2). Thus, consistent with previous studies of takeover dynamics on heterogeneous networks (Giacobini et al., 2005a; Payne and Eppstein, 2007, 2008), the remainder of our experiments utilized local uniform selection (Gorges-Schleuter, 1999), with a simple “replace if better” survivor selection mechanism, as described below.

Each node is 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_i . 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

$$P_{sel} = \frac{x}{k_i} \quad (3.9)$$

The value of node i is then replaced by the value of node j if j has higher fitness (*i.e.* $\Lambda_j(t) > \Lambda_i(t)$).

3.2.4 Generating Scale-Free Topologies with Tunable Scaling and Assortativity

Several methods have been proposed for generating topologies with power-law distributions of vertex connectivity (*e.g.*, Caldarelli et al. (2002); Park et al. (2005); Servedio et al. (2004)). However, most of these graph-generating algorithms produce topologies that consist of a single giant connected component and several small clusters of vertices that are isolated from the rest of the graph, effectively resulting in an unnecessary reduction of the population size. Further, most scale-free graph generating algorithms produce topologies that asymptotically approach fixed structural properties (*e.g.*, the AB algorithm produces graphs with γ approaching 3 and r approaching 0, as N approaches ∞). In this study, we wanted to generate connected scale-free topologies that possess well-controlled scaling exponents and assortativities, but whose inter-connections are otherwise random. In the following subsections we describe the algorithms used to create these graphs.

Tuning the Scaling Exponent

In order to generate scale-free topologies with tunable scaling exponents, we utilized the growing network (GN) model of Krapivsky et al. (2000), which is a generalization of the AB algorithm. Each graph was initialized as a fully connected clique of m_0 nodes (whereas in original GN model (Krapivsky et al., 2000) the graph is initialized with only one node). In each time step t , a single node is added to the graph and is connected to m existing nodes,

such that the probability of connecting to an existing node of degree k is proportional to the linear connection kernel A_k :

$$A_k = \begin{cases} m_0 & \text{if } k = m_0 \\ \alpha k & \text{if } k > m_0 \end{cases} \quad (3.10)$$

The probability Π that a newly introduced vertex attaches to an existing vertex i of degree k_i is then given by

$$\Pi(k_i) = \frac{A_{k_i}}{\sum_{\forall j \in V} A_{k_j}} \quad (3.11)$$

After t time steps, the graph consists of $m_0 + t$ vertices and $\binom{m_0}{2} + mt$ edges, where $t = N - m_0$. So long as $m = m_0$, graph connectivity is ensured, since the first incoming node is forced to connect to all m_0 nodes in the initial clique. Note that if $\alpha = 1$, the AB algorithm is recovered.

By tuning the parameter α in equation (3.10), the scaling exponent can be theoretically tuned anywhere in the range $2 < \gamma < \infty$, for infinitely sized graphs. (Graphs with $\gamma \leq 2$ possess degree distributions with infinite mean and variance.) While an analytical formulation of γ as a function of α is provided in (Krapivsky et al., 2000), this formulation assumes that the graph is infinitely large, and can only be used as a rough approximation for the finite graphs considered herein. Thus, the values of α used in this study to generate graphs with specific scaling exponents, at various population sizes and average degree, were determined empirically. Due to finite size effects and the stochastic nature of the GN algorithm, the mapping of α to γ is not one-to-one, and a range of scaling exponents (γ) will be observed for any given α . In order to generate a graph with a specific desired scaling exponent γ , a graph was generated using the empirically pre-determined value of α , and the observed γ was then calculated for the graph (as described in the subsequent section). Only if the graph had the desired γ (to within 0.01) was it retained and included in the study.

Measuring the Scaling Exponent

Accurately estimating the scaling parameter (γ) of data that are thought to be drawn from a power-law degree distribution is an area of current research (Clauset et al., 2009). One common method for estimating γ is to use the slope of the best linear fit between $\log_{10}(p(k))$ and $\log_{10}(k)$. While this method is straightforward, it has been shown to introduce significant bias in common cases (Clauset et al., 2009).

In order to avoid such bias, we estimated γ using the method provided by Clauset et al. (2009). This method systematically varies γ over a specified range and iteratively applies the Kolmogorov-Smirnov (KS) test to the observed data and fitted model. The γ that minimizes the KS statistic is then chosen as the hypothesized power-law model. The goodness of fit of this model is then calculated using a Monte-Carlo procedure (Clauset et al., 2009) in order to verify that the degree distribution is, in fact, drawn from the hypothesized power-law model. While the minimum degree (k_{min}) of the hypothesized power-law model can be simultaneously estimated (Clauset et al., 2009), in this study it was sufficient to simply use $k_{min} = \langle k \rangle$.

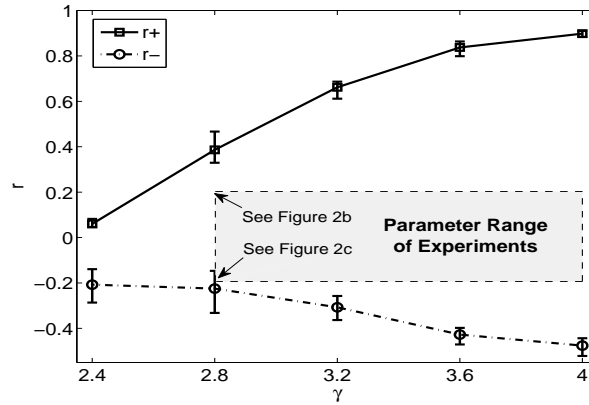
Shuffling the Edge Set

In order to eliminate any structural motifs that may have been inadvertently introduced into the topology as an artifact of the GN algorithm, we utilized the method provided by Maslov and Sneppen (2002) to randomize the edge set of every topology used in this study. The Maslov-Sneppen algorithm is an iterative method that, in each iteration, randomly chooses two edges, $\langle a, b \rangle$ and $\langle x, y \rangle$, from the edge set E (where a, b, x, y are all distinct). These edges are then shuffled to create two new edges, $\langle a, x \rangle$ and $\langle b, y \rangle$, which replace the original edges, so long as they are not already present in the graph. Since the degree of each vertex remains unchanged by a shuffling event, this method exactly preserves the underlying degree distribution.

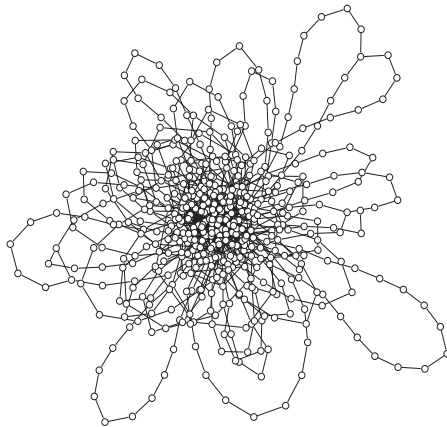
Tuning the Assortativity

In order to investigate the relationship between various assortativities and saturation times in scale-free topologies, we devised a simple iterative method that allows for the direct specification of assortativity within some error tolerance, as follows. In each iteration of our algorithm, the assortativity of the graph is measured using equation (3.6) and compared to the desired assortativity. Two edges, $\langle a, b \rangle$ and $\langle x, y \rangle$, are randomly selected from the edge set E (where a, b, x, y are all distinct) with uniform probability. If the observed assortativity is less than the desired assortativity, the edges are swapped such that the two nodes with the larger degree are connected to one another and the two nodes with the smaller degree are connected together (a “positive” assortative swap). If the observed assortativity is greater than the desired assortativity, the reverse swap is done (a “negative” assortative swap). If either of the new edges is already present in the graph, then the swap is aborted. The algorithm iteratively continues swapping edges in this fashion until the observed assortativity is within 0.01 of the desired assortativity or until a specified maximum number of swaps is reached (in this study, we set the maximum number of swaps to 1,000,000). Like the Maslov-Sneppen algorithm, this method exactly maintains the underlying degree distribution, since the degree of each node remains unchanged after a swapping event. Since neither our assortative shuffling algorithm nor the Maslov-Sneppen algorithm guarantee that the graph remains connected, we discarded graphs that became disconnected, as discussed in Section 3.2.4, step 4.

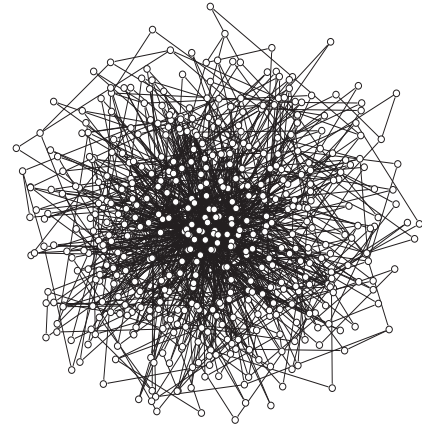
Our algorithm is similar to the single-parameter shuffling algorithm proposed by Xulvi-Brunet and Sokolov (2004), which probabilistically alternates between Maslov-Sneppen edge swaps and edge swaps that alter assortativity. However, the non-determinism in the selection of the edge-swapping algorithm results in a range of assortativities for a given specified probability, whereas our method enables us to more closely achieve the desired degree of assortativity (to within some small error tolerance), thus facilitating the generation of topologies with well-controlled assortativities.



(a)



(b)



(c)

Figure 3.2: (a) Upper bound of positive assortativities (r_+) and lower bound of negative assortativities (r_-) that we were able to obtain using the reshuffling algorithm described in Section 3.2.4, as a function of the scaling parameter (γ). Data points are the averages of the maximum (or minimum) assortativities that were achieved in 10 million assortative edge swaps on each of 10 graph instances at each of the 5 population sizes with $\langle k \rangle = 4$. Error bars indicate the minima and maxima observed for each combination of parameters. The shaded box indicates the range of experimental parameters examined in this study. (b) Visualization of one representative positively assortative graph ($\gamma = 2.8, r = 0.2, N = 500, \langle k \rangle = 4$), and (c) visualization of one representative negatively assortative graph ($\gamma = 2.8, r = -0.2, N = 500, \langle k \rangle = 4$). In (b) and (c) the degree distribution of the graphs is identical. For visual clarity, the graphs shown in (b) and (c) are deliberately smaller than any used in the experiments; visualizations were made using Pajek (<http://vlado.fmf.uni-lj.si/pub/networks/pajek>).

Although assortativity can theoretically range from $-1 \leq r \leq 1$, in reality the actual range of assortativities in scale-free graphs that we were able to obtain, using the algorithms described above, was much more constrained. The positive (negative) assortativity limits were empirically assessed by performing 10 million positive (negative) assortative edge swaps on 10 graph instances, for population sizes $N \in \{1000, 1600, 3600, 6400, 10000\}$, and $\gamma \in \{2.4, 2.8, 3.2, 3.6, 4.0\}$, with $\langle k \rangle = 4$ (Fig. 3.2a) and $\langle k \rangle = 8$ (the data for $\langle k \rangle = 8$ were very similar to those for $\langle k \rangle = 4$, so for clarity are not shown). Population size and number of edges had no detectable effects on the achievable ranges of assortativity, however we observed greater constraints on achievable assortativities at lower scaling exponents (γ). A symmetric range of assortativities for each scaling exponent was implicitly limited by these constraints, and further experiments were thus run using assortativities from $-0.2 \leq r \leq 0.2$, for $2.8 \leq \gamma \leq 4$, as shown by the shaded box in Fig. 3.2a. This range spans to either side of $\gamma = 3$, below which the variance of the degree distribution is theoretically infinite (this is not possible for finite graphs, resulting in a degradation in the scale free behavior at high k , as shown in Fig. 3.1 for $\gamma \sim 2.8$). Parameters within this range have been observed in natural systems (*e.g.*, a recent analysis (Clauset et al., 2009) of empirical power law distributions arising in a variety of systems provides statistical support for scaling exponents ranging from 1.7 to 4.3). We use a relatively small ($N = 500$) scale-free interaction network with $\gamma = 2.8$ and $\langle k \rangle = 4$ to graphically illustrate how the topology qualitatively differs at two extremes of the parameter range shown in Fig. 3.2a, when the assortativity of the graph is reshuffled to have positive assortativity $r = 0.2$ (Fig. 3.2b) or negative assortativity $r = -0.2$ (Fig. 3.2c). Note that the degree distributions of these two graphs are identical.

Summary of Scale-Free Graph Generation

For each desired scaling exponent γ , assortativity r , population size N , and average degree $\langle k \rangle$, we performed the following steps:

1. A graph was created using the GN algorithm, as presented in Section 3.2.4, with

$$m_0 = m.$$

2. The scaling exponent (γ) of the resulting graph was measured using the method presented in Section 3.2.4, with the minimum degree (k_{min}) at which the scaling behavior was assessed set to $\langle k \rangle$ (Fig. 3.1). If the absolute difference between the observed and desired γ was greater than 0.01, or if the goodness of fit of the hypothesized power law model was not satisfactory ($p < 0.1$, as calculated using the Monte Carlo method presented in (Clauset et al., 2009)), then the graph was discarded and step (1) was repeated.
3. Once a graph was created with an acceptable scaling exponent, it was randomized by performing 100,000 Maslov-Sneppen edge swaps, as described in Section 3.2.4. If the graph became disconnected due to a swapping event, the shuffled graph was discarded and step (3) was repeated.
4. The graph was then further shuffled to obtain the desired degree of assortativity (r), using the method presented in Section 3.2.4, until the absolute difference between the observed and desired assortativity was less than 0.01. If the graph became disconnected due to a swapping event, the shuffled graph was discarded and step (4) was repeated. If the graph could not be shuffled to have the desired assortativity within 1,000,000 swapping events, the graph was discarded and step (1) was repeated.

Self loops were implicitly or explicitly precluded by all steps in graph generation.

3.2.5 Experimental Design

In this study, population sizes of $N \in \{1000, 1600, 3600, 6400, 10000\}$ were considered with average degree $\langle k \rangle \in \{4, 8\}$ ($m = 2$ and $m = 4$, respectively), where the scaling exponent was varied between $2.8 \leq \gamma \leq 4.0$ in increments of 0.4 and assortativity was varied between $-0.2 \leq r \leq 0.2$ in increments of 0.1. For each combination of N , $\langle k \rangle$, γ , and r , ten

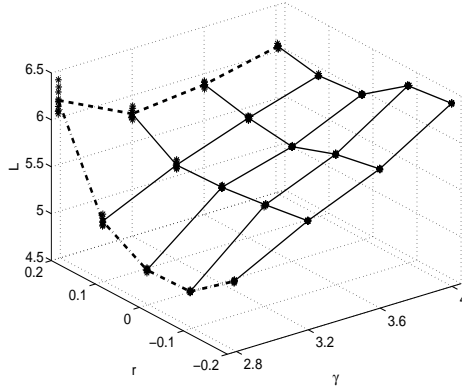
independent graph instances were generated by performing the steps detailed in Section 3.2.4.

In contrast to regular population structures, the expected takeover times in scale-free interaction networks are known to be affected by the placement of the initial high fitness individual (Giacobini et al., 2005a; Payne and Eppstein, 2008). Therefore, for each graph instance we systematically placed the high fitness individual of the initial population in *each* vertex of the population structure, one at a time, and subsequently performed 10 independent takeover time simulations for each individual placement, in order to account for the stochasticity inherent in the selection mechanism. Thus, $100 \times N$ independent simulations were performed for each combination of N , $\langle k \rangle$, γ , and r , resulting in a total of over 140 million independent takeover time simulations.

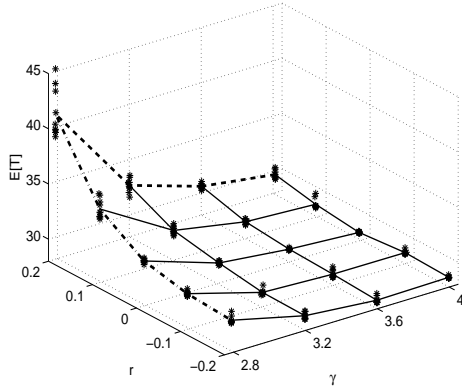
To facilitate this extensive experimental design, simulations and graph generation were performed on a cluster of 128 dual-processor, dual-core (Opteron 2220) IBM x3455s, each with 6GB of memory. All of the graph generating algorithms, edge shuffling routines, and takeover simulations were written in the C programming language for speed. Data analysis was performed using Matlab, as were the methods employed to measure scaling exponents, using the code provided with (Clauset et al., 2009).

3.3 Results

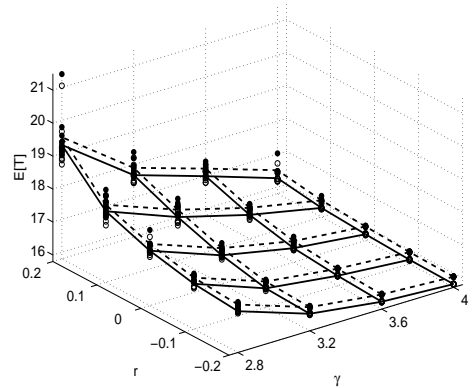
Results were qualitatively similar for all population sizes (N) and average degree ($\langle k \rangle$), so we initially present data for the representative population size of $N = 3600$ and $\langle k \rangle = 4$, and then show how the results scale as a function of population size and average degree. The first subsection pertains to the influence of scaling and assortativity on the static structural properties of the networks considered and the second subsection pertains to the dynamical properties of the takeover times observed on these networks.



(a)



(b)



(c)

Figure 3.3: (a) Characteristic path length (L) and (b, c) expected takeover time ($E[T]$), each shown as a surface function of assortativity (r) and scaling (γ) on a representative population size and average degree ($N = 3600$, $\langle k \rangle = 4$). The symbols represent the data points for each of the 10 graph instances at each parameter combination, with lines connecting the means. Panel (b) corresponds to local uniform selection and (c) to binary tournament selection (\bullet symbols, dashed line) and linear ranking selection (\circ symbols, solid line). In (a, b), the dash-dot lines at $\gamma \sim 2.8$ correspond to the data presented in Figs. 3.5 and 3.6, and the dashed lines at $r \sim 0.2$ correspond to the data presented in Figs. 3.4 and 3.7.

3.3.1 Static Properties

Characteristic path length (L) was found to vary non-monotonically as a function of both the scaling exponent γ and assortativity r (Fig. 3.3a), consistent with the preliminary results presented in (Payne and Eppstein, 2008). For $r \sim 0.2$ (Fig. 3.3a, dashed line), L first decreases from $\gamma \sim 4.0$ to $\gamma \sim 3.2$ ($p \ll 0.01$, ANOVA) and subsequently increases from $\gamma \sim 3.2$ to $\gamma \sim 2.8$ ($p \ll 0.01$, ANOVA). Similarly, for $\gamma \sim 2.8$ (Fig. 3.3a, dash-dot line), L is lowest at $r \sim 0$ and increases as the magnitude of assortativity (either positive or negative) increases. These nonlinear relationships between γ and L , and between r and L , result from changes in the distribution of underlying path lengths L_i , as described below.

Fig. 3.4 depicts the effect of decreasing the scaling exponent (γ) on the distribution of individual path lengths (L_i), and consequently on the characteristic path length (L), for a single representative graph instance with $N = 3600$, $\langle k \rangle = 4$, and $r \sim 0.2$ (this data corresponds to the dashed lines in Fig. 3.3a-b). For $\gamma \sim 4.0$, the distribution of individual path lengths is approximately normally distributed ($p > 0.01$, χ^2 test) with a mean of $L = 6.05$ and standard deviation of 0.6 (Fig. 3.4a). However, as γ decreases, the distribution of path lengths deviates from normality ($p \ll 0.01$, χ^2 test) and begins to flatten and become increasingly skewed, with longer individual path lengths becoming more common as γ continues to decrease (Fig. 3.4b-d). For $\gamma \sim 3.6$ and $\gamma \sim 3.2$, the flattening of the distribution has the effect of decreasing L , but for $\gamma \sim 2.8$ the increase in the frequency of longer individual path lengths begins to play a dominating role, and L consequently increases (dashed vertical lines, $p \ll 0.01$, ANOVA).

Increasing the assortativity (r) has a similar impact on the underlying distribution of path lengths, and on the corresponding characteristic path length (L), as shown in Fig. 3.5, for a representative scale-free interaction network with $N = 3600$, $\langle k \rangle = 4$, and $\gamma \sim 2.8$. The path length distributions deviated from normality for all $\gamma \sim 2.8$ ($p \ll 0.01$, χ^2 test), though for $r \sim -0.2$, the path length distribution was found to be well centered around $L = 5.26$, with very few long paths (Fig. 3.5a). However, for $r > 0.2$, the path length distribution

was found to become increasingly flattened and skewed. Longer path lengths were found to occur more frequently as r increases, and their overall magnitude was found to increase (*e.g.*, compare the horizontal spread of $*$ symbols in Fig. 3.5a with those in 3.5e). Similar to the effect of decreasing γ , increasing r initially has the effect of decreasing L (dash-dot vertical lines for $-0.2 \leq r \leq 0$, $p \ll 0.01$, ANOVA, Fig. 3.5a-c), but as the frequency of longer path lengths and their overall magnitude increases, L subsequently begins to increase (dash-dot vertical lines for $0 < r \leq 0.2$, $p \ll 0.01$, ANOVA, Fig. 3.5d-e). These changes in topology as a function of assortativity are graphically illustrated by comparing Fig. 3.2b (for positive assortativity), where there is a preponderance of long paths through vertices of degree $k = 2$, and Fig. 3.2c (for negative assortativity), where individual paths tend to alternate between high and low degree nodes, and are consequently more homogeneous in length.

3.3.2 Dynamic Properties

While characteristic path length (L) varied non-monotonically with both the scaling exponent (γ) and assortativity (r), (Fig. 3.3a), expected takeover times increased monotonically (*i*) as assortativity was increased from $r \sim -0.2$ to $r \sim 0.2$ and (*ii*) as the scaling exponent was decreased from $\gamma \sim 4.0$ to $\gamma \sim 2.8$ (Fig. 3.3b-c), for all three selection mechanisms considered. The variability in the average takeover times, for each of the 10 graph instances at each parameter combination, also increased with increasing r and decreasing γ (in Fig. 3.3b-c, each $*$, \bullet , and \circ symbol represent data points for the 10 graph instances at each combination of r and γ).

The takeover times observed using binary tournament selection (Fig. 3.3c, \bullet symbols, dashed line) and linear ranking selection (Fig. 3.3c, \circ symbols, solid line) were shorter than those observed under local uniform selection (Fig. 3.3b), because both of these selection mechanisms increase the probability of selecting high fitness nodes from the local neighborhood, relative to uniform selection. However, the qualitative influence of scaling and

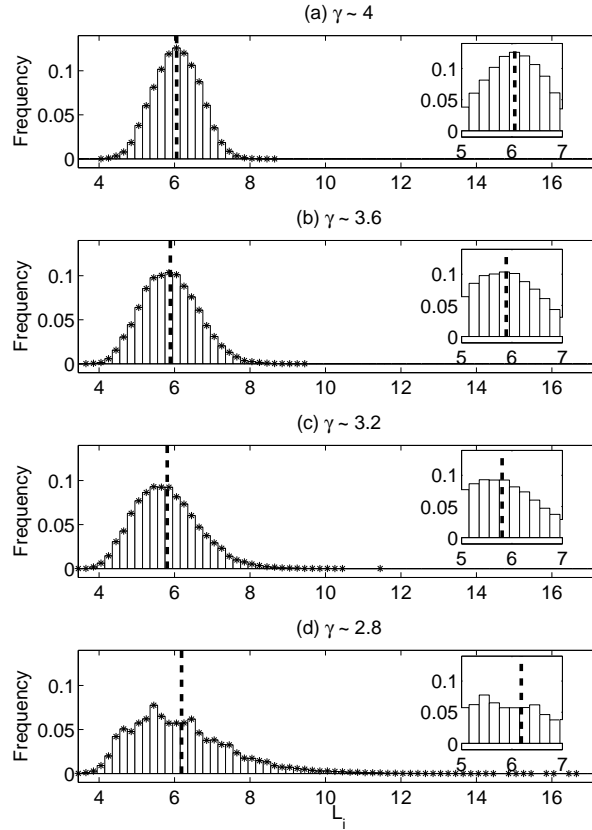


Figure 3.4: Distribution of individual path lengths (L_i) in representative networks with $N = 3600$, $\langle k \rangle = 4$, and $r = 0.2$ for (a) $\gamma \sim 4$, (b) $\gamma \sim 3.6$, (c) $\gamma \sim 3.2$, and (d) $\gamma \sim 2.8$. The * symbols are placed atop each bin as a visual aid. The insets magnify the domain immediately surrounding the average characteristic path lengths L of the ten graph instances (denoted by the thick dashed vertical lines) in order to illustrate the nonlinear relationship between L and γ . Note that the data in this Figure corresponds to the data along the dashed line in Figure 3.3a.

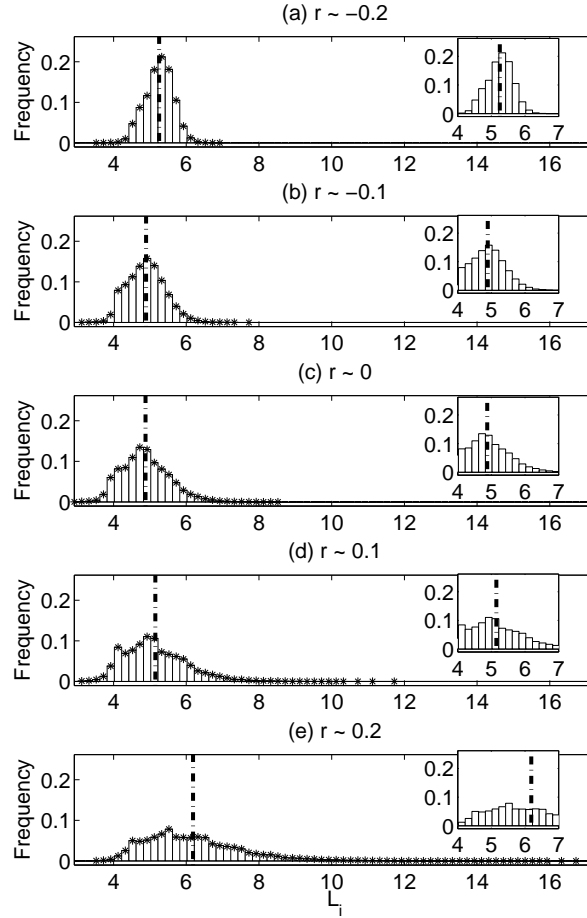


Figure 3.5: Distribution of individual path lengths (L_i) in representative networks with $N = 3600$, $\langle k \rangle = 4$, and $\gamma \sim 2.8$ for (a) $r \sim -0.2$, (b) $r \sim -0.1$, (c) $r \sim 0$, (d) $r \sim 0.1$, and (e) $r \sim 0.2$. The * symbols are placed atop each bin as a visual aid. The insets magnify the domain immediately surrounding the average characteristic path lengths L of the ten graph instances (denoted by the thick dash-dot vertical line) in order to illustrate the nonlinear relationship between L and r . Note that the data in this Figure corresponds to the data along the dash-dot line in Figure 3.3a.

assortativity on average takeover times were similar under all three selection mechanisms (compare Fig. 3b to 3c). The remainder of our results were achieved using local uniform selection.

In Fig. 3.6, we depict the actual takeover dynamics (Fig. 3.6a,c,e), and corresponding distributions of takeover times (Fig. 3.6b,d,f), observed on a single representative scale-free interaction network with $N = 3600$, $\langle k \rangle = 4$, and $\gamma \sim 2.8$ (corresponding to the data around the dash-dot lines in Fig. 3.3a-b). Fig. 3.3b indicates that both expected takeover times, and the variability in the expected takeover times *between graph instances*, increases with *increasing* assortativity. In contrast, Fig. 3.6 shows that the variability in takeover times *on a single graph instance*, due to different initial placements of the high fitness individual, actually *increases* with *decreasing* assortativity, while the average takeover time simultaneously *decreases* (e.g., at $\gamma \sim 2.8$ and $r \sim 0.2$, $E[T] = 43.2$, but this falls to $E[T] = 31.3$ at $r \sim -0.2$, Fig. 3.3b). This occurs because, in negatively assortative graphs, some initial placements result in an extended initial period in which the high fitness genotype is unable to spread (compare Fig 3.6a,c,e, gray lines), but a larger proportion of takeover times are comparatively more rapid than at high assortativity (compare the heights and locations of the peaks of the distributions in Fig. 3.6b,d,f). None of the distributions in Fig. 3.6b,d,f, are normal ($p \ll 0.01$, χ^2 test).

In contrast, decreasing the scaling exponent (γ) of a scale-free interaction network increases *both* the range and average of takeover times observed on a single graph instance, as shown for a representative topology with $N = 3600$, $\langle k \rangle = 4$, and $r \sim 0.2$ in Fig. 3.7 (corresponding to the data around the dashed lines in Fig. 3.3a-b). In this case, the distribution of observed takeover times is shifted to the right and broadened as the scaling exponent decreases, with the expected takeover time increasing from $E[T] = 29.5$ (for $\gamma \sim 4.0$) to $E[T] = 43.2$ (for $\gamma \sim 2.8$), as shown in Figs. 3.7b,d,f,h. Once again, none of the distributions of takeover times are normal ($p \ll 0.01$, χ^2 test).

We now shift our attention to how our results scale with population size (N) and average

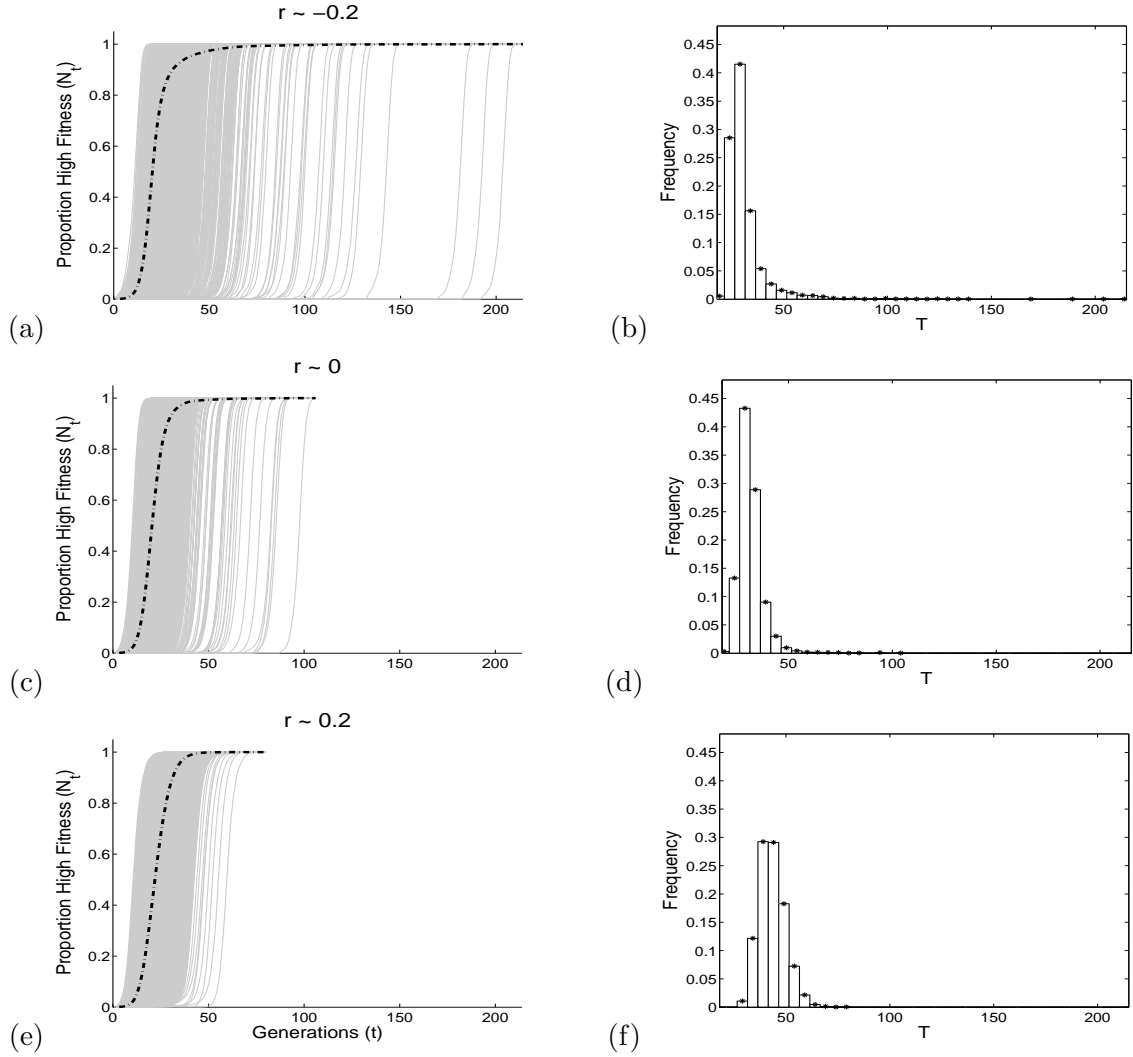


Figure 3.6: Takeover dynamics in representative networks with $N = 3600$, $\langle k \rangle = 4$, and $\gamma \sim 2.8$ for (a) $r \sim -0.2$, (c) $r \sim 0$, and (e) $r \sim 0.2$ (corresponding to the dash-dot lines in Fig. 3.3b). The gray curves denote takeover dynamics for a single experiment and the dash-dot curves denote their averages. In (b), (d), and (f) we show frequency histograms of observed takeover times (T) for the data shown in (a), (c), and (e), respectively. The * symbols are placed atop each bin as a visual aid.

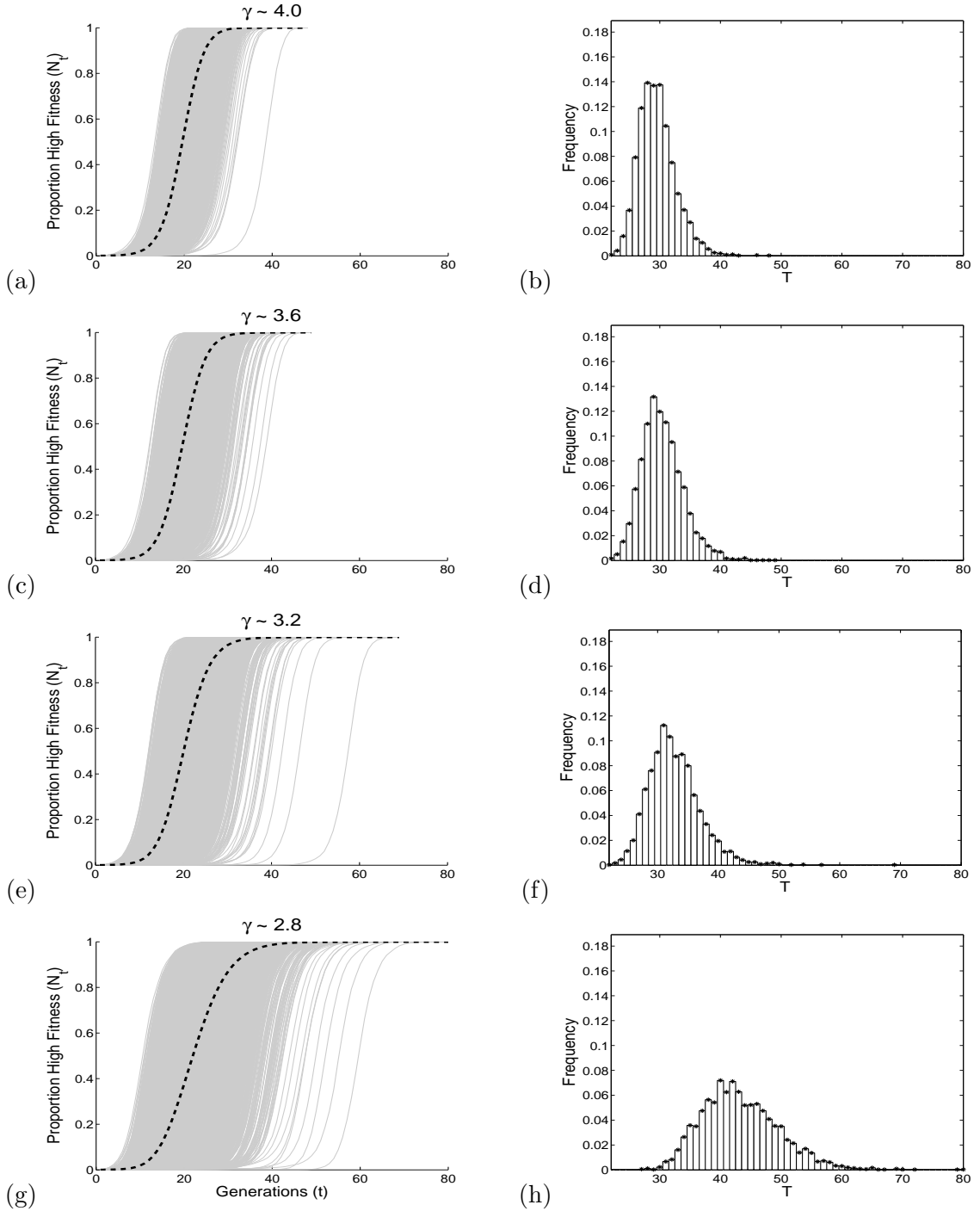


Figure 3.7: Takeover dynamics in representative networks with $N = 3600$, $\langle k \rangle = 4$, and $r \sim 0.2$ for (a) $\gamma \sim 4.0$, (c) $\gamma \sim 3.6$, (e) $\gamma \sim 3.2$, and (g) $\gamma \sim 2.8$ (corresponding to the dotted lines in Fig. 3.3b). The gray curves denote takeover dynamics for a single experiment and the dashed curves denote their averages. In (b), (d), (f), and (h) we show frequency histograms of observed takeover times (T) for the data shown in (a), (c), (e), and (g), respectively. The * symbols are placed atop each bin as a visual aid.

degree ($\langle k \rangle$). Since expected takeover time is an ambiguous multifunction of the characteristic path length, L cannot be used to describe takeover times (or selective pressure) in populations evolving on scale-free topologies. We also found that the leading eigenvalue of the network adjacency matrix is not a good indicator of takeover times in the general case. For the scale-free graphs considered herein with $\langle k \rangle = 4$, the expected takeover time $E[T]$ was found to grow linearly in the leading eigenvalue of the adjacency matrix, for a given population size ($R^2 > 0.86$). However, this relationship was found to deteriorate as the average degree was increased to $\langle k \rangle = 8$ ($R^2 > 0.51$). Further, both the slope and intercept of the linear relationship between leading eigenvalue and expected takeover time were found to differ between $\langle k \rangle = 4$ and $\langle k \rangle = 8$, with a combined R^2 of only 0.02. In contrast, for each population size N , we found that $E[T]$ increased exponentially in the average inverse degree (μ_{-1}) (with the goodness of fit $R^2 > 0.95$ for each N considered in this study); note that (μ_{-1}) varies with *both* $\langle k \rangle$ and scaling exponent γ . Further, for both values of $\langle k \rangle$ considered, we found that the logarithm of the expected takeover time ($\log_{10}(E[T])$) scaled linearly with the logarithm of the population size ($\log_{10}(N)$) (the goodness of fit $R^2 = 0.70$ for $\langle k \rangle = 4$ and $R^2 = 0.87$ for $\langle k \rangle = 8$), with coefficients on $\log_{10}(N)$ less than one, implying that $E[T]$ grows sub-linearly in N . Consequently, we constructed a 2-dimensional best-fit function (depicted in Fig. 3.8a), which accurately describes expected takeover time according to the following relationship:

$$\log_{10}(E[T]) = 0.80 + 0.11 \log_{10}(N) + 0.88\mu_{-1} \quad (3.12)$$

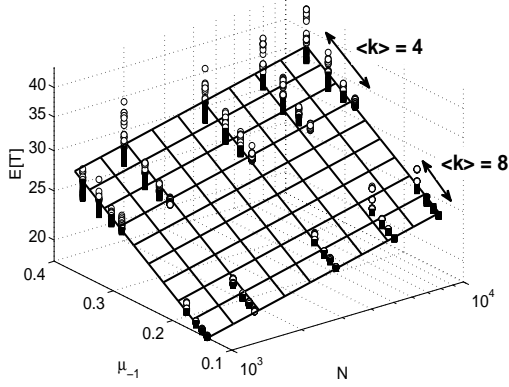
with $R^2 = 0.97$ and a root mean squared error of only 1.32 between this plane and all of the data points (*i.e.*, for all scaling exponents γ , assortativities r , average degrees $\langle k \rangle$, and population sizes N of the scale-free graphs considered herein). In Fig. 3.8b, the viewing angle has been rotated to show the quality of the linear fit as a function of the average inverse degree μ_{-1} ; note that the differences in expected takeover time at each given average degree $\langle k \rangle$ and population size (N) are due to the effects of different scaling exponents (γ), which are captured by the average inverse degree (μ_{-1}). In Fig. 3.8c the viewing angle has been

rotated to show the quality of the linear fit as a function of $\log_{10}(N)$.

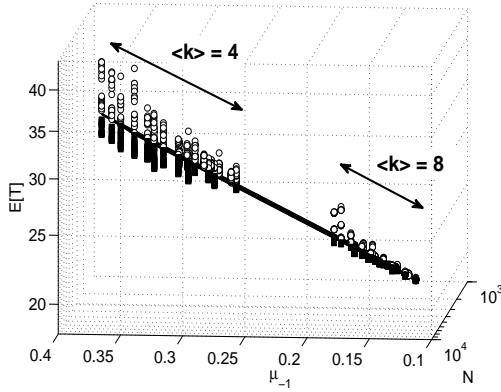
Much of the heteroskedasticity in $E[T]$ as a function of μ_{-1} , as evident in Fig. 3.8b, is attributable to the effect of assortativity (r) on the variability in takeover times. Assortativity has a relatively small effect on the *average* takeover time $E[T]$, that is not modeled by equation (3.12). However, assortativity does have a strong influence on the *variability* of $E[T]$, particularly for scale-free graphs with low scaling exponents (*e.g.*, compare the spread of the data points in Fig. 3.3b at $r \sim 0.2$ for $\gamma \sim 2.8$ and $\gamma \sim 4.0$). In Fig. 3.9a, we explicitly plot the standard deviations of average takeover times ($\sigma_{E[T]}$) as a surface function of r and γ , on the largest population size ($N = 10,000$) with the smallest average degree ($\langle k \rangle = 4$), where this influence is most pronounced; *i.e.*, these indicate the variabilities in the mean takeover times *between the 10 graph instances at each parameter combination* (already averaged over each of the N initial placements, which are already averaged over each of the 10 repetitions at each initial placement). The $\sigma_{E[T]}$ increases nonlinearly as assortativity increases and scaling decreases and contributes to the variable spread of the data points around the plane in Fig 3.8.

As previously noted, assortativity has the opposite effect on the variability in individual takeover times due to *different initial locations of the high fitness value in a single graph instance* (*e.g.*, Fig. 3.6). This is explicitly shown in Fig. 3.9b, where we plot σ_T (*i.e.*, the standard deviations in individual takeover times T due to the N initial placements, which are already averaged over the 10 repetitions at each initial placement, on each given graph instance) as a surface function of r and γ , again on the largest population size ($N = 10,000$) with the smallest average degree ($\langle k \rangle = 4$), where this influence is most pronounced.

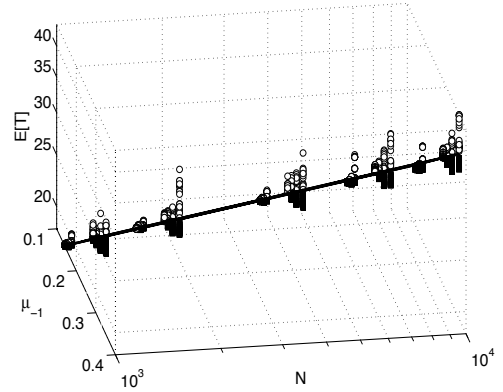
The same general relationships shown in Figs. 3.9a and 3.9b occur at lower population sizes and higher average degree, but are less pronounced because the magnitude of the variability in takeover times decreases with both of these parameters (Fig. 3.8). Note that the scale of the y-axis is an order of magnitude lower for the variability in the mean takeover times observed on the 10 different graph instances (Fig. 3.9a) than for the variability caused



(a)



(b)



(c)

Figure 3.8: (a) Takeover time ($E[T]$) as a function of the average inverse degree (μ_{-1}) and population size (N), for all scaling exponents (γ), assortativities (r), and average degrees ($\langle k \rangle$) considered in this study. The data presented in (a) is also shown in (b) and (c), but from different viewing angles, in order to better elucidate the strength of the relationship between (b) $E[T]$ and μ_{-1} and (c) $E[T]$ and N . Note that in (a) and (b), $\langle k \rangle = 4$ corresponds to $m = 2$ and $\langle k \rangle = 8$ to $m = 4$. In (a)-(c), the plane represents the best fit to the data (see text); open circles denote data points above this plane and filled squares denote data below this plane. Note the logarithmic scaling of takeover time ($E[T]$) and population size (N) in each panel.

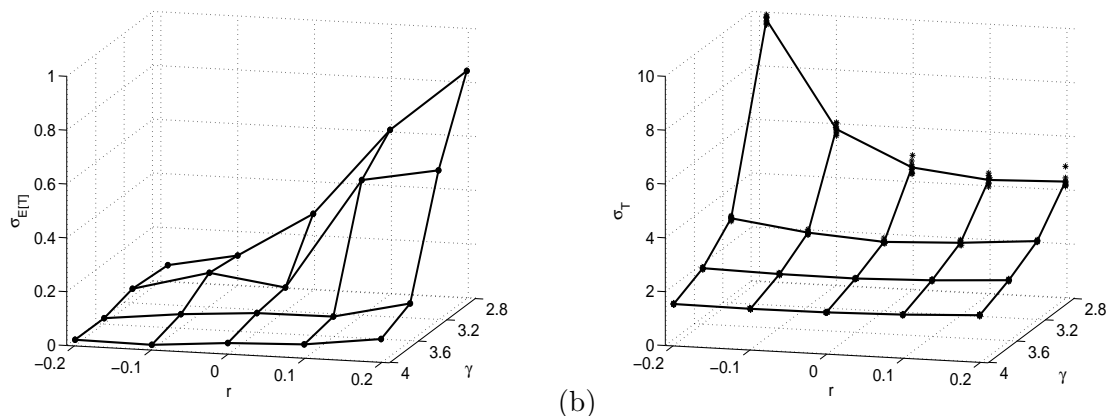


Figure 3.9: (a) Standard deviation of expected takeover times ($\sigma_{E[T]}$) due to differences between the means on the 10 graph instances at each parameter combination, and (b) Standard deviations of individual takeover times (σ_T) on the same graph instances, due to the N initial locations of the high fitness individual. In both (a) and (b), standard deviations are shown as a surface function of assortativity (r) and scaling (γ) at $N = 10,000$, $\langle k \rangle = 4$. Note that the scale of the vertical axis in (a) is an order of magnitude smaller than in (b).

by the N different initial placements (Fig. 3.9b) at each parameter setting.

3.4 Discussion

The results of this study demonstrate that in scale-free interaction networks with randomized edge sets, takeover time ($E[T]$) under local uniform selection increases as either (i) the assortativity r increases or (ii) the scaling exponent γ decreases, with these two topological properties interacting nonlinearly in their effect on takeover time (Fig. 3.3b). In negatively assortative graphs ($r < 0$), nodes with high degree (*a.k.a.* “hubs”) are attached to nodes with low degree, and *vice versa* (Fig. 3.2c). As such, placing the initial high fitness genotype in a hub results in rapid saturation, because the neighbors of a hub have few connections themselves and thus have a high probability of adopting the hub’s genetic information quickly. In contrast, placing the initial high fitness genotype in a low degree vertex that neighbors a hub can result in a severe retardation of the propagation of this genetic information (Fig. 3.6a), because the hub has more neighbors to choose from. For example, if the initial high fitness genotype is placed in a low degree neighbor of a hub with

degree k , then it will take on average k generations for this hub to adopt this high fitness information, under the local uniform selection method outlined in Section 3.2.3. While this has the effect of increasing the variability in the takeover times observed on a single graph instance (Fig. 3.6a), the extremely rapid saturation that occurs when the initial high fitness genotype is placed in a highly connected node outweighs such occasional slow takeover time events, and causes an overall decrease in $E[T]$ as r decreases (Fig. 3.3b). In positively assortative graphs ($r > 0$), hubs are connected to hubs, and nodes of low degree are connected to one another. For example, when $\langle k \rangle = 4$, this causes the formation of long sections of linear chains of nodes (Fig. 3.2b). In such topologies, placing the initial high fitness genotype in a hub does not result in rapid saturation because its neighboring hubs are reluctant to adopt the high fitness genetic information, as they too have many neighbors to choose from. Further, no matter where the initial high fitness genotype is placed, this genetic information must still travel through the long chains of low degree vertices, resulting in the lack of correlation between the degree of the vertex of initial placement and takeover time that was noted in (Payne and Eppstein, 2008).

Decreasing γ causes the interaction network to possess more highly connected vertices (e.g., for $N = 10,000$, $k_{max} = 27$ for $\gamma \sim 4.0$ and $k_{max} = 140$ for $\gamma \sim 2.8$, Fig. 3.1). As previously discussed, in negatively assortative and uncorrelated topologies, such hubs communicate advantageous genetic information rapidly once they have it (Giacobini et al., 2005a; Payne and Eppstein, 2008), but their high connectivity makes them more resistant to it in the first place. This causes an increase in both the average takeover time and the variability in the takeover times observed on a single graph instance (Fig. 3.7). In combination, high assortativity (r) and low scaling exponents (γ) interact nonlinearly to increase takeover time (Fig. 3.3b).

Metrics based on pairwise shortest paths, such as characteristic path length (L), have a strong governing influence on the rate of information flow in many topologies. For example, expected takeover time ($E[T]$) is positively correlated with L in regular population

structures (Payne and Eppstein, 2007). However, the degree to which shortest path metrics govern information flow is dependent upon both network heterogeneity (Payne and Eppstein, 2008) and whether or not the local update mechanism is a function of vertex degree. Even if the update mechanism were independent of vertex degree, as in a *Susceptible-Infectious-Susceptible* model with 100% infection rate, L would still not be sufficient to describe the rate of spread in scale-free graphs, because the so-called “characteristic” path length is not actually characteristic of static structure in these graphs (as shown in Figs. 3.4 and 3.5). An alternative shortest path metric, the average maximum shortest distance from each vertex to any other vertex (*i.e.*, the average eccentricity), will determine average saturation times when the update mechanism is independent of vertex degree, whether the graph is regular or not. However, when the updating mechanism is degree dependent and the graph is heterogeneous, shortest path metrics are not sufficient to describe dynamics.

Our results show that it *is* possible to describe the logarithm of the expected takeover time $E[T]$ using a planar function of (*i*) the inverse average degree (μ_{-1}) and (*ii*) the logarithm of population size ($\log_{10}(N)$) (equation (3.12)). Since the probability with which a vertex of degree k adopts the high fitness allele of a neighboring vertex is inversely proportional to k (equation (3.9)), it makes intuitive sense that the expected takeover time is a function of μ_{-1} , which weights the inverse degrees of the vertices in a graph in proportion to their occurrence (equation (3.3)) and thus captures the effects of both the average degree and the scaling exponent. Further, our observation that the logarithm of takeover time ($\log_{10}(E[T])$) is linearly dependent on the logarithm of the population size ($\log_{10}(N)$) has been noted in other studies of dynamical processes on scale-free graphs, including consensus time (*i.e.*, the number of steps required to reach a state of all 0’s or all 1’s) in the unbiased Voter Model (Sood et al., 2008).

In (Sood et al., 2008), it was also shown that consensus times in the unbiased Voter Model on Molloy-Reed scale-free topologies (Molloy and Reed, 1995) with a fixed average degree were largely governed by the effective population size, defined as $N_{eff} = N\mu_1^2/\mu_2$.

However, this metric did not account for much of the variability in expected takeover times in our study. For the $\langle k \rangle = 4$ data, $R^2 = 0.64$ and for the $\langle k \rangle = 8$ data, $R^2 = 0.72$; however, the slopes and intercepts for these rather weak associations were quite different from each other, indicating a strong sensitivity of N_{eff} to average degree, so the overall R^2 was only 0.10. Although the two state Voter model (VM) used in (Sood et al., 2008) is similar in many ways to the takeover time study presented here, there are several significant differences in these models that can account for the differences in behavior. In the biased VM, state 0 is prescribed a fitness $f = 1$ and state 1 a fitness of $f = r$, with $r > 1$; in the unbiased case, there is no preference for either state, *i.e.*, there is no selection pressure. In the VM, nodes are updated asynchronously, such that in each step a node i is chosen with probability inversely proportional to its fitness (in contrast to our implementation using locally uniform selection with synchronous updates). In the VM, node i then imports the state of a neighboring node j , irrespective of the fitness of node j (again, in contrast to our model which uses a replace-if-better update policy). Thus, the results of this study complement those presented in (Sood et al., 2008), and highlight the sensitivity of information flow to the specific selection and update mechanisms. Future work will explore system responses to additional update mechanisms, such as the heterogeneous threshold-based update policies commonly used in models of binary decisions (Watts, 2002), in which the update rule of any given individual is contingent upon both the individual's particular response function and the states of its neighboring individuals.

Although the effects of assortativity were not explicitly studied in (Sood et al., 2008), the authors do comment that assortativity had a smaller secondary effect (relative to the scaling exponent) on consensus times. In this study, we too found assortativity to have a secondary effect on mean takeover times. For example, equation (3.12), which does not account for the effects of assortativity, explains 97% of the variance in the logarithmically transformed average takeover times. However, a closer examination of the data shows that, while assortativity has little effect on either mean or variability of average takeover times

at the higher scaling exponents, at low scaling exponents assortativity has a pronounced effect on the variability in takeover times. Specifically, increasingly *positive* assortativities increased the variability between experiments on different random graph instances (after averaging over all initial placements on the same graph instance) (Fig. 3.9a), while increasingly *negative* assortativities increased the variability between takeover times from different initial placements on the same graph instances (Fig. 3.9b). These seemingly contradictory findings can be explained as follows. At low scaling exponents (*i.e.*, large hubs), the average system dynamics are sensitive to the specific connectivity patterns found in different graph instances with the same degree of positive assortativity, while the large number of very long individual path lengths (L_i) in these graphs means that they are relatively insensitive to initial placement of the high fitness individual, since these long paths must always be traversed. Conversely, at low scaling exponents with negative assortativity, all graph instances tend to have alternating high and low degree nodes and in this sense the connectivity patterns are relatively similar to each other, yet the takeover time is very sensitive to whether the initial high fitness individual is placed in a high or low degree node, which respectively cause very short or long delays before the onset of the sigmoidal saturation process (Fig. 3.6). Thus, assortativity does have an important role in influencing systems dynamics on scale-free networks, especially if individual trajectories in specific systems are more relevant than average expected behaviors over an entire class of systems.

One of our goals has been to discover if there are simple, and easily computable, metrics of graph topology which govern the rate of spread of information in arbitrary networks. Previously, we have focused on metrics based on all-pairs shortest direct paths, such as characteristic path length, or a combination of the maximum and variance of the all-pairs shortest path lengths (Payne and Eppstein, 2007). However, none of these have proven to govern system dynamics in the general case (Payne and Eppstein, 2008). In this study, we have also shown that neither the leading eigenvalue of the adjacency matrix nor the “effective population size” (N_{eff}) metric are sufficient to predict takeover times in these heterogeneous

graphs. However, we did find that most of the variability in the logarithm of average takeover times in scale-free graphs with a wide range of scaling exponents, assortativities, number of nodes, and average node degree, can be explained by a function of the average inverse degree and the logarithm of the population size. Unfortunately, this relationship will not generalize to the case of regular graphs, in which the average inverse degree is constant for a given topology. For example, consider the contrasting cases of a lattice with Von Neumann neighborhoods ($\langle k \rangle = 4$) and a regular random graph of the same average degree. According to equation (3.12), the takeover times on these two topologies would be identical, while in reality the saturation times on a regular random graph will be much more rapid than those observed on the corresponding lattice with Von Neumann neighborhoods.

The results of this study demonstrate how the topological properties of randomized scale-free interaction networks influence the flow of advantageous alleles, and thus impact the selective pressures induced by such population structures. As previously mentioned, the few studies (Gasparri et al., 2007; Giacobini et al., 2006; Kirley and Stewart, 2007a,b) that have utilized scale-free population structures for computational optimization have reported mixed results. However, to date all of the studies utilizing scale-free population structures for evolutionary optimization have employed graphs generated using the AB (Barabási and Albert, 1999) algorithm, which exhibit no assortativity ($r \sim 0$) and have a scaling exponent that asymptotically approaches $\gamma = 3$. Thus, it remains unclear whether or not, and to what degree, the structural properties of the larger class of scale-free interaction networks (with variable scaling exponents and assortativity) can be exploited to improve the search performance of evolutionary algorithms, and how these topological characteristics relate to various aspects of problem difficulty, such as multimodality, epistasis, and deception. Since our results show that takeover times of individual experiments are highly sensitive to initial placement of the high fitness individual, especially at low scaling exponents in disassortative networks (Fig. 3.9b), we suspect that evolutionary algorithms on such networks would require multiple restarts to assure reliable results. However, the question as to whether

the heterogeneity of these networks can sometimes lead to higher fitness solutions than on regular graphs remains to be seen. We have recently commenced further studies to try to address these questions.

The population structures utilized in evolutionary computation are typically static. That is, the interaction network is grown prior to the evolution of the population and it remains fixed throughout the evolutionary process. However, recent work has demonstrated that dynamic interaction network structures, both regular (Alba and Dorronsoro, 2005) and irregular (Whitacre et al., 2008), can enhance the search capabilities of population-based optimization algorithms. Specifically, Alba and Dorronsoro (2005) showed that the solution quality of genetic algorithms can be improved by dynamically altering the dimensions of rectangular lattice population structures, and Whitacre et al. (2008) demonstrated that self-organizing complex interaction networks improve diversity maintenance in steady-state, asexual, mutation-limited populations. The results presented in this study demonstrate that the selective pressures induced by scale-free interaction networks can be altered by changing the scaling exponent and/or assortativity, both of which can be achieved through edge swaps. Thus, it may be possible to dynamically alter the selective pressure in scale-free interaction networks, as an on-line means for controlling the exploration/exploitation tradeoff in evolutionary optimization.

The selection operators commonly used in evolutionary optimization may differ from the local uniform selection policy with replace-if-better updates employed in the bulk of this study. However, experimentation with selection operators that are commonly employed in evolutionary optimization on regular population structures, such as binary tournament selection and linear ranking selection, produced qualitatively similar results to those observed using local uniform selection.

For each graph instance considered in this study, care was taken to ensure that the edge sets were thoroughly randomized. We expect that scale-free interaction networks with more rigidly defined edge sets, such as the hierarchical and modular scale-free graphs considered

in (Barabási et al., 2001; Ravasz et al., 2002), may induce qualitatively different selective pressures than those reported on herein.

All of the experiments performed in this study were nonextinctive, such that high fitness individuals could not revert back to low fitness. Thus, the formulation of takeover time analysis considered herein implicitly assumed that once an advantageous mutation arose, it spread relentlessly. In evolutionary optimization, and other relevant contexts such as epidemiology, this is clearly a simplifying assumption; individuals with advantageous mutations are often lost before they are chosen for reproduction and infectious individuals often recover before they spread disease. Thus, population structure influences the spread of genetic information by affecting (*i*) the probability with which an advantageous mutation becomes established and, once established, (*ii*) the rate with which it spreads. While this study investigated the latter influence of population structure, work is currently underway to investigate how the structural characteristics of scale-free interaction networks impact the former. As certain population structures are known to be promoters of the spread of advantageous alleles and others are known to be suppressors (Lieberman et al., 2001; Whigham and Dick, 2008), investigating the influence of the topological properties of scale-free interaction networks on establishment probabilities will complement the results presented in this study.

In summary, in this contribution we demonstrated why the so-called “characteristic path length” of a graph is not, in fact, characteristic of system dynamics on scale-free graphs with stochastic degree-dependent update mechanisms, because this metric does not account for variability in node degree. We also showed that the eigenvalues of the connectivity matrix and the effective population size (N_{eff}) are unable to account for observed variability in takeover times. Rather, we found that the logarithms of average takeover times on scale-free graphs, with a variety of scaling exponents, assortativities, and average degrees *were* well-accounted for by a planar function of (*i*) the average inverse degree (which incorporates the influence of scaling) and (*ii*) the logarithm of the population size, and that there is a

strong nonlinear effect of assortativity on the variability of takeover times at low scaling exponents. As the formulation of takeover time analysis considered herein serves as a simplified measure of information flow in general, we believe that the results of this study may provide insights into a variety of dynamical processes which may occur on scale-free networks, such as epidemiological invasion, the dissemination of fads, ideas, and innovations, and evolutionary optimization.

3.5 Glossary of Variables

- A_k : *Attachment kernel* - The attachment weight of a node of degree k , which is used to determine the probability ($\Pi(k)$) that an incoming node will attach to a node of degree k .
- $E[T]$: *Expected takeover time* - Takeover time, averaged over all initial placements, on a single graph instance.
- γ : *Scaling parameter* - The scaling exponent of a power-law degree distribution, $p(k) \propto k^{-\gamma}$
- G : *Graph* - A graph $G = (V, E)$, comprised of a vertex set (V) and an edge set (E). In this study, graph is used synonymously with *population structure*, *topology*, and *network*.
- k : *Vertex degree* - The number of neighbors of a given node.
- $\langle k \rangle$: *Average degree* - The average number of neighbors per node, which is equivalent to the first moment of the degree distribution (μ_1).
- k_{min} : *Minimum degree* - The minimum number of neighbors possessed by a node in a graph G .
- k_{max} : *Maximum degree* - The maximum number of neighbors possessed by a node in a graph G .
- $\Lambda_i(t)$: *Vertex fitness* - The fitness of a vertex i at time t
- L_i : *Average individual path length* - Average shortest path between vertex i and all other vertices in a graph G .
- L : *Characteristic path length* - Average all-pairs shortest distance between vertices.
- μ_1 : *Average degree* - The first moment of the degree distribution, which is equivalent to the average degree ($\langle k \rangle$).
- μ_{-1} : *Average inverse degree* - The first inverse moment of the degree distribution.

- N : *Population size* - The number of vertices in a graph G .
- N_{eff} : *Effective population size* - Defined in (Sood et al., 2008) as $\frac{N\mu_1^2}{\mu^2}$
- N_t : The proportion of high fitness nodes at time t .
- $\Pi(k_i)$: The probability that a newly introduced vertex attaches to an existing vertex i of degree k_i .
- $p(k)$: *Probability distribution function* - The probability of finding a node of degree k .
- $P(k)$: *Complementary cumulative distribution function* - The probability of finding a node of degree greater than or equal to k .
- P_{sel} : *Selection probability* - Probability of selecting a high fitness individual from a mating neighborhood.
- r : *Assortativity* - Propensity of vertices of similar degree to attach to one another.
- σ_T : Standard deviation of the takeover times observed in each individual graph instance (*i.e.*, for each of the N initial placements of the high fitness individual, which are already averaged over the 10 replicates of each initial placement).
- $\sigma_{E[T]}$: Standard deviation of the expected takeover times across graph instances, for a given combination of γ , r , and N .
- T : *Takeover time* - The takeover time observed in a single experiment.

References

- Alba, E. and B. Dorronsoro (2005). The exploration/exploitation tradeoff in dynamic cellular genetic algorithms. *IEEE Transactions on Evolutionary Computation* 9(2), 126–142.
- Alba, E. and M. Tomassini (2002). Parallelism and evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 6(5), 443–462.
- Altenberg, L. (2005). Evolvability suppression to stabilize far-sighted adaptations. *Artificial Life* 11(4), 427–443.
- Barabási, A. L. and R. Albert (1999). Emergence of scaling in random networks. *Science* 286, 509–512.
- Barabási, A. L., E. Ravasz, and T. Vicsek (2001). Deterministic scale-free networks. *Physica A* 299, 559–564.
- Borrett, S. R., B. D. Fath, and B. C. Patten (2007). Functional integration of ecological networks through pathway proliferation. *Journal of Theoretical Biology* 245, 98–111.
- Bryden, K. M., D. Ashlock, S. Corns, and S. Wilson (2005). Graph based evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 10(5), 550–567.
- Caldarelli, G., A. Capocci, P. De Los Rios, and M. A. Munoz (2002). Scale-free networks from varying vertex intrinsic fitness. *Physical Review Letters* 89(25), 258702.
- Cantú-Paz, E. (2000). Markov chain models of parallel genetic algorithms. *IEEE Transactions on Evolutionary Computation* 4(3), 216–226.

- Caswell, H. (2001). *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer Associated, Sunderland.
- Clauset, A., C. R. Shalizi, and M. E. J. Newman (2009). Power-law distributions in empirical data. *arXiv:0706.1062v1*.
- Ebel, H., L. Mielsch, and S. Bornholdt (2002). Scale-free topology of email networks. *Physical Review E* 66, 035103.
- Eppstein, M. J. and J. Molofsky (2007). Invasiveness in plant communities with feedbacks. *Ecology Letters* 10, 253–263.
- Folino, G., C. Pizzuti, and G. Spezzano (2001). Parallel hybrid method for sat that couples genetic algorithms and local search. *IEEE Transactions on Evolutionary Computation* 5(4), 323–334.
- Folino, G., C. Pizzuti, and G. Spezzano (2003a). A scalable cellular implementation of parallel genetic programming. *IEEE Transactions on Evolutionary Computation* 7(1), 37–53.
- Gasparri, A., S. Panziera, F. Pascucci, and G. Ulivi (2007). A spatially structured genetic algorithm over complex networks for mobile robot localisation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 4277–4282. IEEE Press.
- Giacobini, M., M. Preuss, and M. Tomassini (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. Springer-Verlag, Heidelberg.
- Giacobini, M., M. Tomassini, and A. Tettamanzi (2005a). Takeover times curves in random and small-world structured populations. In H. G. Beyer (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2005*, pp. 1333–1340. ACM Press, New York.

- Giacobini, M., M. Tomassini, A. Tettamanzi, and E. Alba (2005b). Selection intensity in cellular evolutionary algorithms for regular lattices. *IEEE Transactions on Evolutionary Computation* 9(5), 489–505.
- Goldberg, D. and K. Deb (1991). A comparative analysis of selection schemes used in genetic algorithms. In G. Rawlins (Ed.), *Foundations of Genetic Algorithms*, pp. 69–93. Morgan Kaufmann, San Francisco.
- 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. Morgan Kaufmann, San Francisco.
- Hauert, C. and M. Doebeli (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* 428, 643–646.
- Herrera, F. and M. Lozano (2000). Gradual distributed real-coded genetic algorithms. *IEEE Transactions on Evolutionary Computation* 4(1), 43–63.
- Holme, P. and B. J. Kim (2002). Growing scale-free networks with tunable clustering. *Physical Review E* 65, 026107.
- Jeong, H., S. P. Mason, A. L. Barabási, and Z. N. Oltvai (2001). Lethality and centrality in protein networks. *Nature* 411, 41.
- Jeong, H., B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási (2000). The large-scale organization of metabolic networks. *Nature* 407, 651–654.
- Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London B* 266, 859–867.
- Kerr, B., M. A. Riley, M. W. Feldman, and B. J. M. Bohannan (2002). Local dispersal promotes biodiversity in a real life game of rock-paper-scissors. *Nature* 418, 171–174.

- Kirley, M. and R. Stewart (2007a). 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. ACM Press, New York.
- Kirley, M. and R. Stewart (2007b). Multiobjective optimization on complex networks. In *Fourth International Conference on Evolutionary Multi-Criterion Optimization*, pp. 81–95. Lecture Notes Computer Science.
- Krapivsky, P. L., S. Redner, and F. Leyvraz (2000). Connectivity in growing random networks. *Physical Review Letters* 85(21), 4629.
- Lieberman, E., C. Hauert, and M. A. Nowak (2001). Evolutionary dynamics on graphs. *Nature* 433, 312–316.
- Lienig, J. (1997). A parallel genetic algorithm for performance-driven VLSI routing. *IEEE Transactions on Evolutionary Computation* 1(1), 29–39.
- Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg (2001). The web of human sexual contacts. *Nature* 411, 907–908.
- Maslov, S. and K. Sneppen (2002). Specificity and stability in topology of protein networks. *Science* 296, 910–913.
- Molloy, M. and B. Reed (1995). A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms* 6, 161–180.
- Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters* 89, 208701.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review* 45(2), 167–256.

- Nowak, M. A. and R. M. May (1992). Evolutionary games and spatial chaos. *Nature* 359, 826–829.
- Ohtsuki, H., C. Hauert, E. Lieberman, and M. A. Nowak (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502–505.
- Park, K., Y. C. Lai, and N. Ye (2005). Self-organized scale-free networks. *Physical Review E* 72, 026131.
- Pastor-Satorras, R. and A. Vespignani (2001). Epidemic spreading in scale-free networks. *Physical Review Letters* 86(4), 3200–3203.
- Payne, J. L. and M. J. Eppstein (2007). Takeover times on scale-free topologies. In D. Thierens (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2007*, pp. 308–315. ACM Press, New York.
- Payne, J. L. and M. J. Eppstein (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*, pp. 241–248. ACM Press, New York.
- Payne, J. L. and M. J. Eppstein (2009). Pair approximations of takeover dynamics in regular population structures. *Evolutionary Computation*. To appear.
- Payne, J. L., M. J. Eppstein, and C. J. Goodnight (2007). Sensitivity of self-organized speciation to long-distance dispersal. In *Proceedings of the IEEE Symposium on Artificial Life*, pp. 1–7.
- Poncela, J., J. Gómez-Gardenes, L. M. Florià, and Y. Moreno (2007). Robustness of cooperation in the evolutionary prisoner’s dilemma on complex networks. *New Journal of Physics* 9, 184.

- Rauch, E. M. and Y. Bar-Yam (2006). Long-range interactions and evolutionary stability in predator-prey systems. *Physical Review E* 73, 020903.
- Ravasz, E., A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabási (2002). Hierarchical organization of modularity in metabolic networks. *Science* 297, 1551–1555.
- Reichenbach, T., M. Mobilia, and E. Frey (2007). Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* 448, 1046–1049.
- Rong, Z., X. Li, and X. Wang (2007). Roles of mixing patterns in cooperation on a scale-free networked game. *Physical Review E* 76, 027101.
- Rudolph, G. (2000). On takeover times in spatially structured populations: array and ring. In K. K. Lai, O. Katai, M. Gen, and B. Lin (Eds.), *Proceedings of the Second Asia-Pacific Conference on Genetic Algorithms and Applications, APGA-2000*, pp. 144–151. Global Link Publishing Company, Hong Kong.
- Santos, F. C. and J. M. Pacheco (2005). Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters* 95, 098104.
- Sarma, J. and K. De Jong (1996). An analysis of the effect of neighborhood size and shape on local selection algorithms. In H. M. Voigt, W. Ebeling, I. Rechenberg, and H. P. Schwefel (Eds.), *Parallel Problem Solving from Nature*, pp. 236–244. Springer-Verlag, Heidelberg.
- Sayama, H., L. Kaufman, and Y. Bar-Yam (2003). Spontaneous pattern formation and genetic diversity in habitats with irregular geographic features. *Conservation Biology* 17, 893–900.
- Servedio, V. D. P., G. Calarelli, and P. Butta (2004). Vertex intrinsic fitness: how to produce arbitrary scale-free networks. *Physical Review E* 70, 056126.
- Sood, V., T. Antal, and S. Redner (2008). Voter models on heterogeneous networks. *Physical Review E* 77, 041121.

- Van Veldhuizen, D. A., J. B. Zydallis, and G. B. Lamont (2003). Considerations in engineering parallel multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 7(2), 144–173.
- Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Science* 99, 5766.
- Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of small-world networks. *Nature* 393, 440–442.
- Werfel, J. and Y. Bar-Yam (2004). The evolution of reproductive restraint through social communication. *Proceedings of the National Academy of Science* 101(30), 11019–11024.
- Whigham, P. A. and G. Dick (2008). Evolutionary dynamics for the spatial moran process. *Genetic Programming and Evolvable Machines* 12(2), 220–230.
- Whitacre, J. M., R. A. Sarker, and Q. T. Pham (2008). The self-organization of interaction networks for nature-inspired optimization. *IEEE Transactions on Evolutionary Computation* 12(2), 220–230.
- Xulvi-Brunet, R. and I. M. Sokolov (2004). Reshuffling scale-free networks: from random to assortative. *Physical Review E* 70, 066102.

Chapter 4

Information Cascades on Degree-Correlated Random Networks

Abstract

We experimentally investigate a threshold model of contagion on degree-correlated random networks. We show that the parameter range in which global information cascades occur generally expands as degree-degree correlations become increasingly positive. However, under certain conditions, large-scale information cascades can paradoxically occur when degree-degree correlations are sufficiently positive or negative, but not when correlations are relatively small. We also show that the ability of high degree vertices to trigger large cascades is strongly affected by degree-degree correlations. We find that in networks with negative degree-degree correlations, high degree vertices trigger larger information cascades than low degree vertices, on average, but in networks with positive degree-degree correlations, the opposite result obtains. Lastly, we show that cyclical triggering motifs are present in large finite random networks that enable cascades on denser networks than predicted an-

alytically. Since real-world networks are both finite and exhibit degree-degree correlations, these results may have relevance in understanding contagion in a variety of application domains.

4.1 Introduction

In the past decade, analytical and computational models have been used extensively to study how the structural characteristics of interaction topologies affect network-based dynamical processes (Barrat et al., 2008). For example, power-law degree distributions have been shown to lead to a vanishing epidemic threshold in the Susceptible-Infectious-Susceptible (SIS) model of disease spread (Pastor-Satorras and Vespignani, 2001), decrease consensus times in the unbiased Voter Model (Sood et al., 2008), and increase saturation times in models of takeover dynamics (Payne and Eppstein, 2008). The update mechanisms in such network-based models of information flow generally come in two basic flavors (Dodds and Watts, 2004); in typical models of disease spread (Murray, 2002), the probability of information spreading along a given edge is independent of the degree of the vertex to which it leads, whereas in threshold models of binary decisions (Watts, 2002), a vertex will change to a new state if a specified fraction of its neighboring vertices are in that state. In both kinds of models, a small number or fraction of vertices are typically “seeded” with a novel piece of information; when this information spreads throughout a significant portion of the network, it is referred to as an information cascade, akin to processes in real-world systems, such as propagating failures in power grids (Kinney et al., 2005; Sachtjen et al., 2000) or the adoption of new ideas and fads in social networks (Schelling, 1973). The size and frequency of such information cascades have been shown to be affected by various factors, including the number of seed vertices (Gleeson and Cahalane, 2007); the degree of the vertices in which these seeds are placed (Watts and Dodds, 2007); the details of the threshold function (Watts, 2002; Watts and Dodds, 2007); edge directionality (Gleeson, 2008b); the presence of long-range connections *vs.* local clustering (Centola et al., 2007); and other structural

characteristics of the underlying interaction network, such as modularity (Galstyan and Cohen, 2007; Watts and Dodds, 2007) and heterogeneity in the degree distribution (Watts, 2002; Watts and Dodds, 2007).

In the most commonly used network generating algorithms, such as those that produce generalized random networks (Milo et al., 2003) or Albert-Barabási scale-free networks (Barabási and Albert, 1999), there are no correlations between the degrees of neighboring vertices (Newman, 2002). However, in an analysis of several real-world networks, Newman found that social interaction networks are typically *assortative* (*i.e.*, degrees of adjacent vertices are positively correlated), while technological and biological networks are commonly *disassortative* (*i.e.*, degrees of adjacent vertices are negatively correlated) (Newman, 2002, 2003). Degree-degree correlations can have a significant impact on network-based dynamical processes. In random networks with exponentially truncated power-law degree distributions, a giant component percolates at lower average degree (thus facilitating the spread of information) in assortative networks than in uncorrelated or disassortative networks, whereas the maximum size of the giant component is larger in disassortative networks than in uncorrelated and assortative networks (Newman, 2002). Further, with these same degree distributions, the giant component of an assortative network is more robust to the random removal of vertices than the giant component of uncorrelated or disassortative networks (Newman, 2003). Positive degree-degree correlations have also been shown to significantly alter the distribution of avalanche sizes in the Bak-Tang-Wiesenfeld sandpile model (Yin et al., 2007), and when the prisoner’s dilemma is played on scale-free networks, the proportion of cooperators at equilibrium is reduced in both assortative and disassortative networks, relative to uncorrelated networks (Rong et al., 2007). In the context of the SIS model, the introduction of degree-degree correlations has no effect on the null epidemic threshold observed in infinitely large scale-free networks with infinite variance (Boguñá et al., 2003), though investigations of large, but finite, scale-free networks have shown that epidemic outbreaks spread more rapidly in assortative networks than in uncorrelated or disassortative

networks (Kiss et al., 2008). Investigations of takeover times on connected scale-free networks have shown that the introduction of positive degree-degree correlations increases the mean saturation time and decreases the variability between initial seed placements (Payne and Eppstein, 2009). Despite these many advances, how degree-degree correlations affect dynamics on complex networks is not yet completely understood.

Here, we focus on a threshold model of binary decisions, originally developed by Watts for the uncorrelated generalized random network case (Watts, 2002). A key aspect of this threshold model is that a single vertex may trigger a global cascade. In the context of technological networks, such as power grids, these ‘triggers’ represent the system’s Achilles’ heel, whereas in social systems, they, at least theoretically, represent the prime targets for marketing strategists to effectively advertise their product (Katz and Lazarsfeld, 1955; Watts and Dodds, 2007). Using this model, we perform extensive numerical simulations of information cascades starting from a single initial seed on large ($N = 10^4$) degree-correlated random networks. Our investigations reveal several results regarding the relationships between degree-degree correlations and the frequency and size of large-scale information cascades, the underlying structure of a network’s triggering component, and the ability of high degree vertices to trigger cascades. Motivated by this study, we have elsewhere developed analytic results concerning the frequency of global cascades on degree-correlated random networks (Dodds and Payne, 2009), complementing the recent results of Gleeson (2008a), who provided analytic results for cascade sizes. While our experimental findings are largely in excellent agreement with analytic work, we show observed discrepancies to be due to the presence of cyclical topological motifs.

4.2 Model

Vertices can be in one of two states, *active* (infected) or *inactive* (susceptible), and once a vertex activates, it cannot deactivate (Watts, 2002). Every vertex is given an identical threshold-based response function, where the probability (Π) that a vertex of degree k

changes its state from inactive to active is a function of the fraction of its k neighbors that are active. Specifically, if x denotes the number of active neighbors of a vertex of degree k , and ϕ denotes the threshold of its response function, then

$$\Pi(x, k) = \begin{cases} 1, & \text{if } \frac{x}{k} \geq \phi \\ 0, & \text{otherwise} \end{cases} \quad (4.1)$$

Vertices are updated synchronously, though for the case of monotonically increasing response functions, as considered herein, asynchronous updating yields equivalent results.

Vertices that activate in the presence of a single active neighbor are referred to as *vulnerable*, and all other vertices are referred to as *stable* (Watts, 2002). Infinite random networks have a zero probability of containing cycles; large-scale information cascades that are triggered by a single vertex can therefore only occur in these strictly branching structures if there exists a sufficiently large connected component of vulnerable vertices (Watts, 2002). This is referred to as the *vulnerable component*, its size is denoted by S_v , given as a fraction of the system size N , and the set of vertices it contains is denoted by Ω_v . The frequency of information cascades is almost completely dictated by S_v , since any initial placement in Ω_v will lead to a cascade of size at least S_v . However, stable vertices that are directly adjacent to the vulnerable component can also trigger large-scale cascades, since an initial seed placed in one of these vertices will ignite the vulnerable component. We will refer to these vertices as *peripheral* and denote the set of peripheral vertices as Ω_p , with size S_p . The union of the vulnerable and peripheral components, $\Omega_e = \Omega_v \cup \Omega_p$, makes up what is referred to as the *extended vulnerable component* (Watts, 2002), the size of which we denote as S_e . In strictly branching networks, S_e is exactly the number of vertices that can trigger a large-scale information cascade, so this quantity is equivalent to the frequency with which large-scale cascades occur if the initial seed is randomly chosen. Vertices outside of Ω_e are referred to as *external*, the set of which is denoted by Ω_x .

We will define a *global information cascade* as any cascade that infects 1% or more of the entire network. This arbitrary distinction (Watts, 2002) is reasonable because the

distribution of cascade sizes is generally bimodal with sizes either well above or well below 1% (except near the onset of the percolating vulnerable component, where the cascade size distribution obeys a power law (Watts, 2002)). Thus, the terms *average global cascade size* (S_{gc}) and *global cascade frequency* (F_{gc}) will refer to global cascades only, omitting the lower mode of the distribution. In contrast, *average cascade size* (S_{avg}) will be used to mean the average of all information cascades, combining both modes of the cascade size distribution. All sizes are given as a proportion of the number of vertices N .

4.3 Methods

We consider undirected random networks of N vertices, with average degree z , and a Poisson degree distribution $p_k = z^k e^{-z}/k!$. Random networks were generated by randomly placing $M = Nz/2$ edges between pairs of vertices selected with uniform probability. Duplicate edges were prohibited, resulting in an exact average degree z .

Following Newman (2002), we use e_{jk} to denote the probability that a randomly chosen edge connects vertices with degree $j + 1$ and $k + 1$. The quantity $q_k = \sum_j e_{jk}$ is then the probability that a randomly chosen edge followed in a random direction leads to a vertex of degree $k + 1$. Degree-degree correlations are quantified by their assortativity r , which is formally defined (Newman, 2002) as

$$r = \frac{1}{\sigma_q^2} \sum_j \sum_k jk(e_{jk} - q_j q_k) \quad (4.2)$$

where $\sigma_q^2 = \sum_k k^2 q_k - [\sum_k k q_k]^2$ is the variance of the distribution q_k . A network is said to be disassortative if $r < 0$, uncorrelated if $r = 0$, and assortative if $r > 0$.

We generate networks with a given assortativity, within some error tolerance, using the following iterative shuffling routine. At each step, we first measure the assortativity of the network using Eq. (4.2). Two edges $\langle a, b \rangle$ and $\langle x, y \rangle$ are then chosen at random with uniform probability (such that a, b, x, y are all distinct). If the observed assortativity is less than the desired assortativity, then we swap the edges such that the two vertices with the

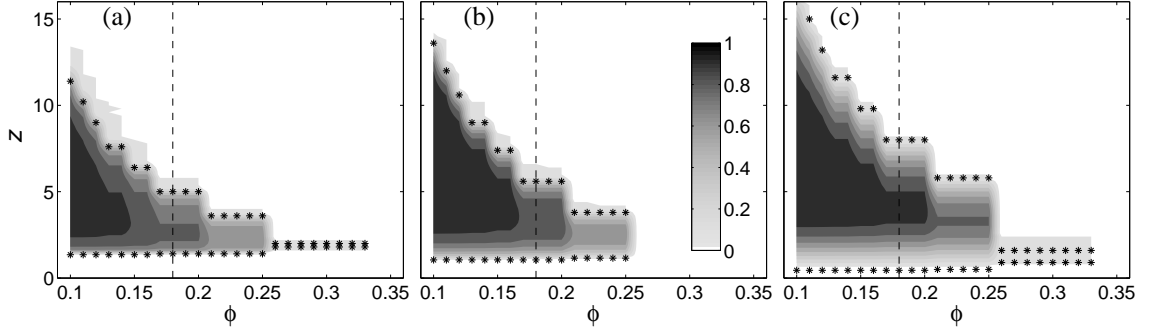


Figure 4.1: Observed frequency of global cascades, as a function of threshold ϕ and average degree z , for (a) disassortative ($r = -0.5$), (b) uncorrelated ($r = 0$), and (c) assortative ($r = +0.5$) random networks. The color bar corresponds to all panels. The asterisks denote the numerical solutions to Eq. (4.3). The dashed vertical lines at $\phi = 0.18$ correspond to the three curves in Fig. 4.2.

larger degree are connected to one another and the two vertices with the smaller degree are connected together. To instead incrementally decrease assortativity, we connect the two vertices with the largest and smallest degree, and the two vertices of intermediate degree. (If either of the new edges were already present, then we leave the network unchanged.) Shuffling continues until the observed assortativity is within 0.01 of the desired assortativity. This method exactly preserves the underlying degree distribution p_k , since the degree of each vertex remains unchanged after a swapping event.

We have recently shown (Dodds and Payne, 2009) that the condition for the onset of global spreading, triggered by a single initial seed, is:

$$|\mathbf{A}| = 0 \text{ where } [\mathbf{A}]_{j+1,k+1} = \delta_{jk}q_k - kb_{k+1}e_{jk}, \quad (4.3)$$

and b_k is the probability that a degree k vertex is vulnerable. Note that round-off errors in the numerical solution of Eq. (4.3) can lead to multiple locations in which this equation is satisfied; in the results shown herein, we used the extremal values of z where the determinant of \mathbf{A} was numerically equal to zero.

Table 4.1: Ranges of assortativity considered in this study, as a function of average degree z .

| Average degree (z) | Range of assortativity (r) |
|------------------------|--------------------------------|
| $z < 1.0$ | $-0.50 \leq r \leq +0.50$ |
| $1.0 \leq z < 2.0$ | $-0.70 \leq r \leq +0.95$ |
| $2.0 \leq z < 3.0$ | $-0.85 \leq r \leq +0.95$ |
| $3.0 \leq z < 5.0$ | $-0.90 \leq r \leq +0.95$ |
| $z \geq 5.0$ | $-0.95 \leq r \leq +0.95$ |

4.4 Experimental Design

In all experiments, the number of vertices was held constant at $N = 10^4$. We considered thresholds, ϕ , in the range $[0.1, 0.35]$ in increments of 0.01. Unless otherwise specified, the average degree z was varied from 1.0 to 20.0 in increments of 0.2. Assortativities ranged from $r = -0.95$ to $r = 0.95$, depending on z , since the shuffling method presented in Section 4.3 suffers from some constraints. In particular, it is considerably more difficult to obtain extremely negative assortativities than it is to obtain extremely positive assortativities, and this effect is especially pronounced for low z (see Dodds and Payne (2009) for details). In Table 4.1, we present the lower and upper bounds of the assortativities of the networks generated for this study, as a function of z .

For each combination of z and r , 10 network instances were generated. For each combination of network instance and threshold ϕ , the initial seed was placed in each vertex, one at a time, and an independent simulation was performed for each of the $N = 10^4$ initial placements, resulting in almost 10^{13} simulations.

4.5 Results

We first investigate the influence of degree-degree correlations on the so-called ‘cascade window’ (Watts, 2002), which delineates the region in which global cascades can occur, as a function of the vertex threshold ϕ and average degree z . In order to concretely ascertain the lower boundary of average degree z for which global cascades occur, we augmented our experimental design with an additional set of experiments on random networks with

$0.2 \leq z \leq 1.0$ (in increments of 0.05). In Fig. 4.1, we depict the frequency of global cascades as a function of vertex threshold ϕ and average degree z in disassortative (Fig. 4.1a), uncorrelated (Fig. 4.1b), and assortative (Fig. 4.1c) random networks. Frequencies were recorded as the proportion of all simulations that resulted in a global cascade, across all initial placements on each of the 10 network instances, for each combination of ϕ and z . The results of our simulations (Fig. 4.1, shaded contours) are in excellent agreement with Eq. (4.3) (Fig. 4.1, asterisks), although this analytic solution occasionally under-predicts the upper z -boundary of the cascade window; we find these discrepancies to be due to the presence of cyclical triggering motifs that are not considered in the analytic treatment, and we discuss this issue later in the present section.

As shown in Fig. 4.1, the cascade window is clearly influenced by degree-degree correlations. In general, increasing assortativity r expands the parameter range in which global cascades are observed (compare the shaded regions in Fig. 4.1, as the panels are read from left to right). Specifically, for $\phi \leq 0.25$, increasing r consistently decreases the minimum, and increases the maximum, average degree z for which global cascades occur (compare the heights of the shaded regions in Fig. 4.1, as the panels are read from left to right). In uncorrelated networks, global cascades were never observed with $\phi > 0.25$ (Fig. 4.1b). However, at low average degree z , global cascades were observed for thresholds ranging from $0.25 < \phi \leq 0.33$, for both disassortative networks (Fig. 4.1a) and assortative networks (Fig. 4.1c).

In Fig. 4.2, we present the percolation phase transition in more detail, depicting both the observed frequencies of global cascades and the corresponding sizes of the extended vulnerable component (S_e), as a function of z . To be consistent with previous work (Watts, 2002; Watts and Dodds, 2007), these cascade simulations pertain to a threshold of $\phi = 0.18$ (corresponding to the vertical dashed lines in Fig. 4.1). For each value of r at $\phi = 0.18$, we find excellent agreement between S_e (Fig. 4.2, lines) and global cascade frequency (Fig. 4.2, symbols), as expected. Increasing r decreases the minimum z for which global cascades

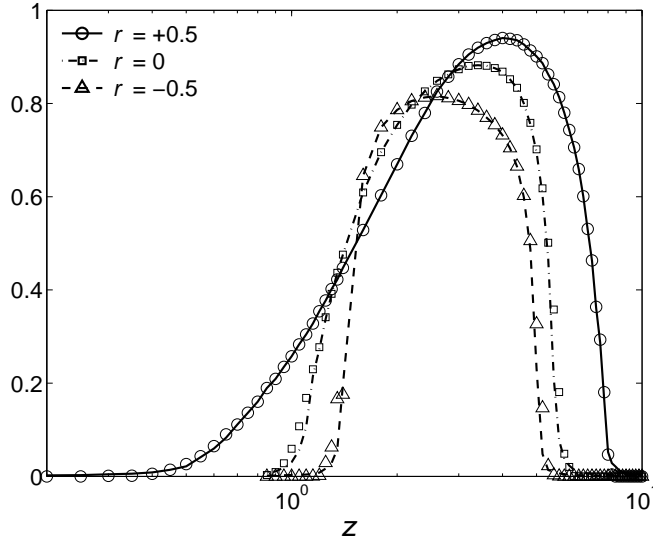


Figure 4.2: Observed statistics of global cascades as a function of average degree z for $\phi = 0.18$, on assortative ($r = +0.5$), uncorrelated ($r = 0$), and disassortative ($r = -0.5$) random networks. Each data point is the frequency of global cascades observed on 10 network instances. The lines are the corresponding average sizes of the extended vulnerable components (S_e). The x-axis is scaled logarithmically.

are observed and increases the maximum observed frequency of global cascades (compare the peaks of the three curves in Fig. 4.2). The range of z for which global cascades occur also increases with increasing r , as seen previously in Fig. 4.1. Further, increasing r alters the skew of the functional relationship between global cascade frequency and z (compare the shape of the curve as assortativity increases from $r = -0.5$ to $r = 0$ to $r = +0.5$ in Fig. 4.2).

In Fig. 4.3, we depict the effects of average degree z and assortativity r on global cascade frequency (F_{gc} , Fig. 4.3, left column) and global cascade size (S_{gc} , Fig. 4.3, right column), for three vertex thresholds, $\phi = 0.20$ (top row), $\phi = 0.25$ (middle row), and $\phi = 0.33$ (bottom row). For $\phi \leq 0.25$ and $r < 0$, the maximum z for which any global cascades occur is relatively insensitive to changes in r (note how the right-most z -boundary of the shaded contours below the horizontal dashed lines in Fig. 4.3a,c are nearly independent of r). However, for $r > 0$, the maximum average degree z for which global cascades occur increases dramatically as r becomes increasingly positive (Fig. 4.3a,c,

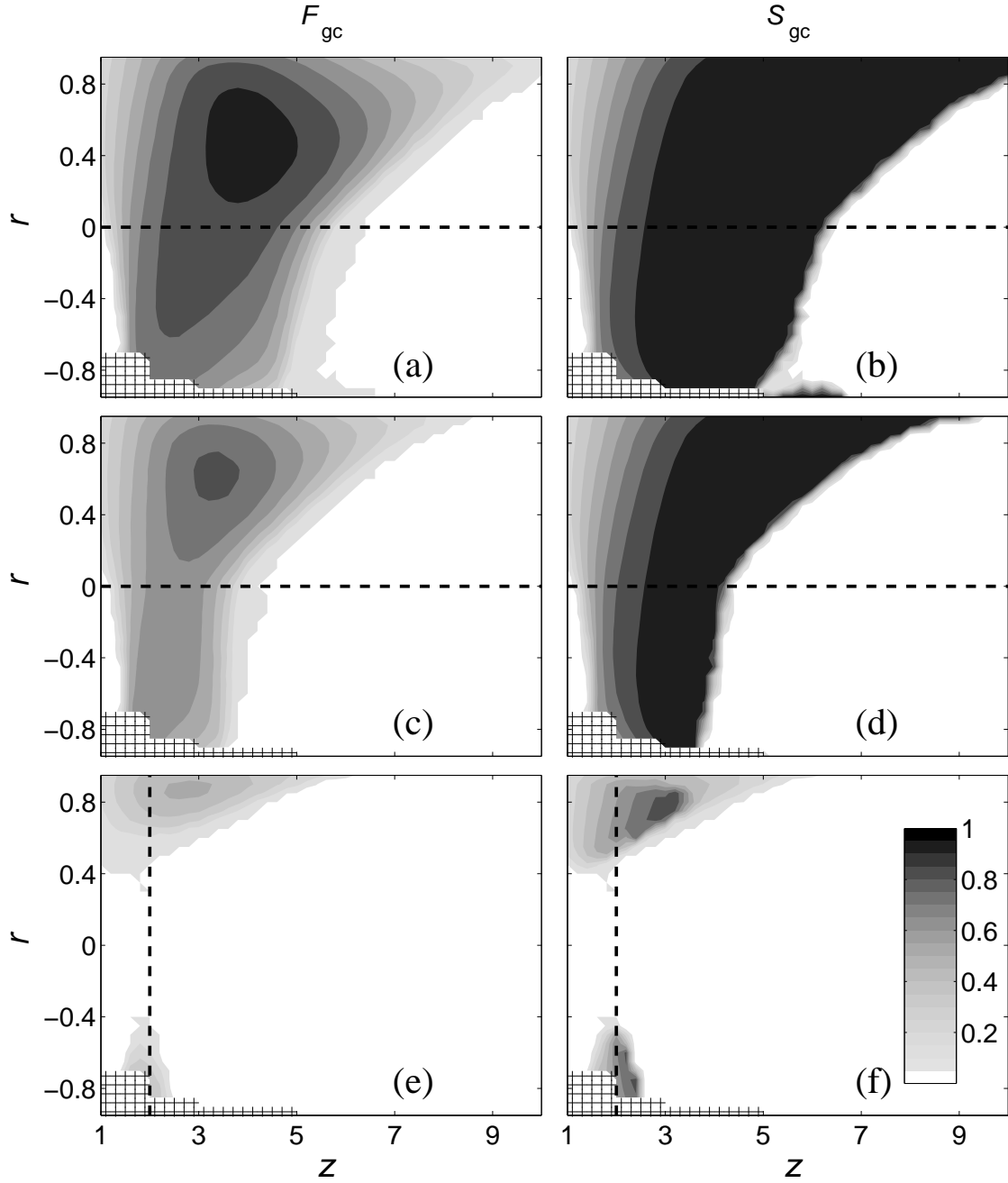


Figure 4.3: Frequency (F_{gc} , left column) and size (S_{gc} , right column) of global cascades, as a function of average degree z and assortativity r , for three threshold values: $\phi = 0.2$ (top row), $\phi = 0.25$ (middle row), and $\phi = 0.33$ (bottom row). The color bar corresponds to all panels. The dashed horizontal lines in panels (a)-(d) are at $r = 0$, and the dashed vertical lines in panels (e) and (f) correspond to the data presented in Fig. 4. We were unable to obtain data in the hatched region (see Table 4.1).

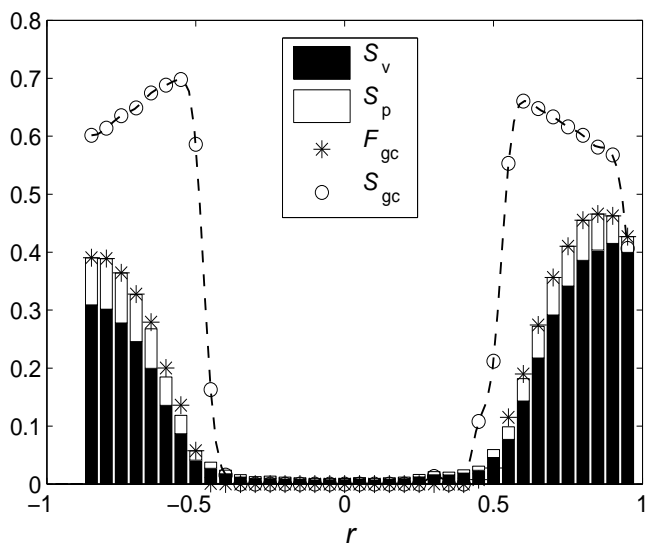


Figure 4.4: (a) Frequency (asterisks) and size (circles) of global cascades, as a function of assortativity r , for $z = 2.0$ and $\phi = 0.33$. The bars correspond to the sizes of the vulnerable (S_v , black) and peripheral (S_p , white) subcomponents of S_e . The dashed line is provided as a guide for the eye.

data above horizontal dashed line, where the right-most z -boundary of the shaded contours increase rapidly with r). Although the frequencies of global cascades near this upper- z boundary are low, when global cascades do occur they consistently spread throughout the entire network (as indicated by the black contours in Fig. 4.3b,d; *e.g.*, for $z = 9$ and $r = +0.8$, compare the low frequency in Fig. 4.3a to the large cascade size in Fig. 4.3b).

At high thresholds ($\phi > 0.25$, Fig. 4.3, bottom row) and low z , the frequency and size of global cascades becomes a bimodal function of r (Fig. 4.3e,f). In addition, the sizes of global cascades near the upper- z boundary are significantly reduced relative to lower thresholds (as shown by the lower grayscale contours in Fig. 4.3f as compared to Fig. 4.3b,d). Further details of the bimodal response to r at $z = 2$ and $\phi = 0.33$ (vertical lines in Figs. 4.3e,f) are depicted in Fig. 4.4, where we also show the sizes of the vulnerable (S_v) and peripheral (S_p) components. Note the sharp transition in global cascade size at $|r| \sim 0.5$, and how the global cascade size decreases as the frequency of global cascades increases for $|r| \rightarrow 1$.

Although the two modes shown in Fig. 4.4 look similar, they are caused by distinctly

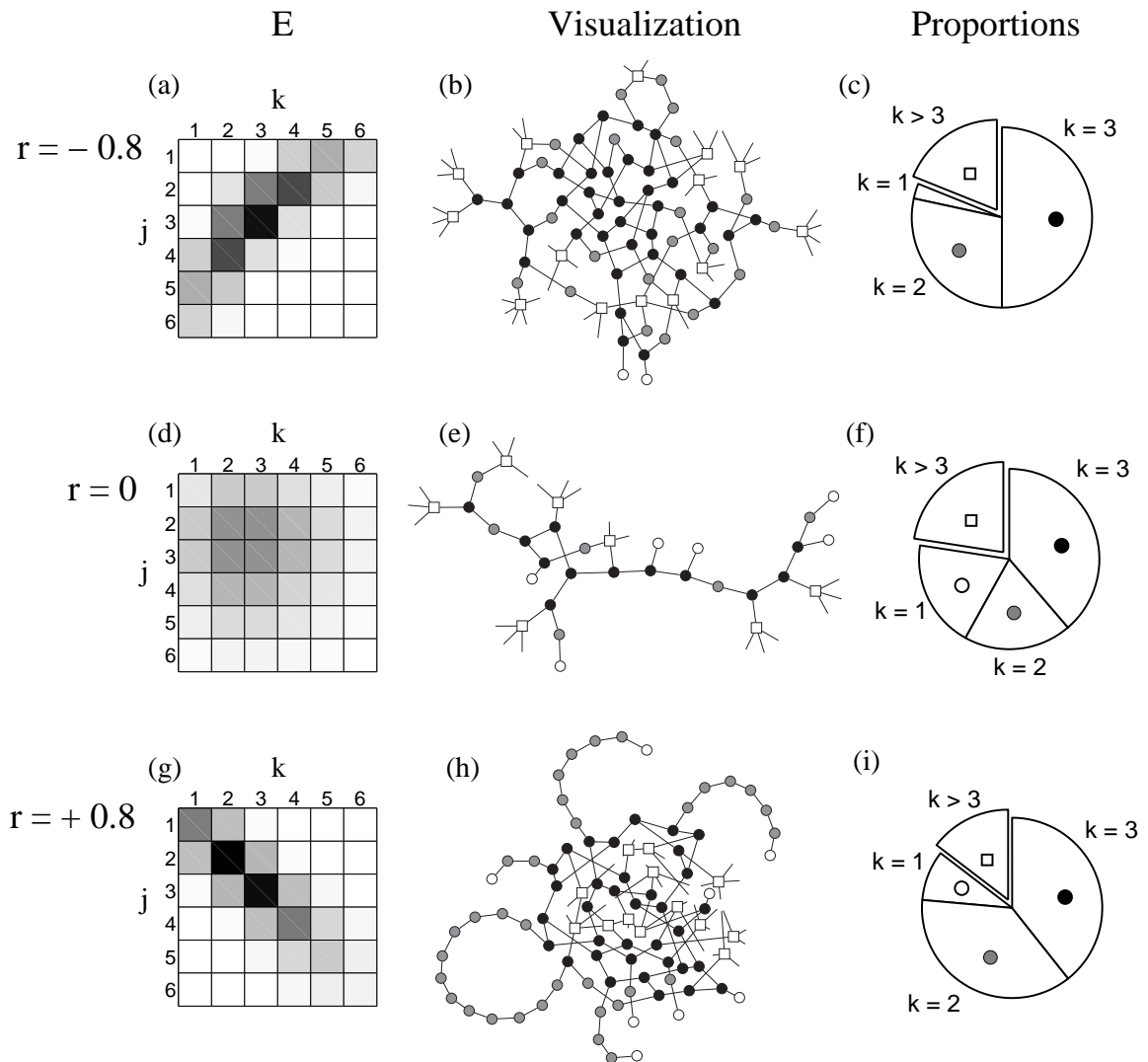


Figure 4.5: Visualizations of random network properties with $z = 2$ and disassortative ($r = -0.8$, top row), uncorrelated ($r = 0$, middle row), and assortative ($r = +0.8$, bottom row) degree-degree correlations. In (a,d,g), we visualize averages of the E matrices for all 10 network instances ($N = 10,000$) used at each parameter combination. Shading (ranging from white to black) of matrix entries in E correspond to the probabilities (ranging from 0 to 0.17) with which a randomly chosen edge connects vertices of degree j and k ; entries corresponding to vertices with $k > 6$ occur with very low frequency and so are not shown. In (b,e,h), we depict the extended vulnerable component (S_e) for $\phi = 0.33$ for individual random networks with $N = 200$ (for visual clarity). Peripheral vertices ($k > 3$) are denoted as white squares; vertices in the vulnerable component are denoted as circles, and are colored according to degree: white ($k = 1$), gray ($k = 2$), and black ($k = 3$). The corresponding pie charts in (c,f,i) denote the proportions of degrees found in these extended vulnerable components (b,e,h).

different underlying topological properties. In disassortative networks (*e.g.*, Fig. 4.5a,b,c), the vulnerable component comprises edges between vertices of alternating high and low degree. For example, the vulnerable component shown in Fig. 4.5b for $z = 2$ and $\phi = 0.33$ comprises alternating vertices of degree $k = 2$ and $k = 3$, with only a few degree $k = 1$ vertices attached. This results from the inherent negative degree correlation, where $k = 1$ vertices frequently connect to stable vertices ($k > 3$, Fig. 4.5a), excluding them from the vulnerable component. In contrast, for assortative networks (*e.g.*, Fig. 4.5g,h,i), vertices of similar degree are frequently connected in the vulnerable component. This is shown for networks with $z = 2$ and $\phi = 0.33$ in Fig. 4.5h, where there are clusters of vertices of $k = 3$ and chains of vertices with $k = 2$, most of which are terminated by $k = 1$ vertices. The probability of stable ($k > 3$) vertices connecting to the core of $k = 2$ and $k = 3$ vertices is higher in disassortative networks (Fig. 4.5a) than in assortative networks (Fig. 4.5g). Consequently, the peripheral component of $k > 3$ vertices is larger in the disassortative networks (Fig. 4.5c) than in the assortative networks (Fig. 4.5i). In uncorrelated networks, vulnerable and stable vertices are likely to connect to one another (Fig. 4.5d), frustrating the formation of a sizable extended vulnerable component (Fig. 4.5e), hence prohibiting global cascades (Fig. 4.4).

In general, there is excellent agreement between the size of the extended vulnerable component ($S_e = S_v + S_p$) and the observed frequency of global cascades (compare the height of the bars with the asterisks in Fig. 4.4). However, the observed frequency of global cascades is slightly higher than can be explained by the size of the extended vulnerable component S_e in the region $0.5 \lesssim |r| \lesssim 0.7$ (in Fig. 4.4, note how the asterisks are slightly above the bars in these regions). Closer examination of individual experiments enabled us to ascertain that these discrepancies were caused by triggering vertices which are not themselves part of Ω_e , but are ultimately connected to vertices in Ω_e by certain cyclical structural motifs, such as the motif illustrated in Fig. 4.6, for $\phi = 0.33$ (such triggering motifs were first noted by D. E. Whitney, private communication). In this illustration, the

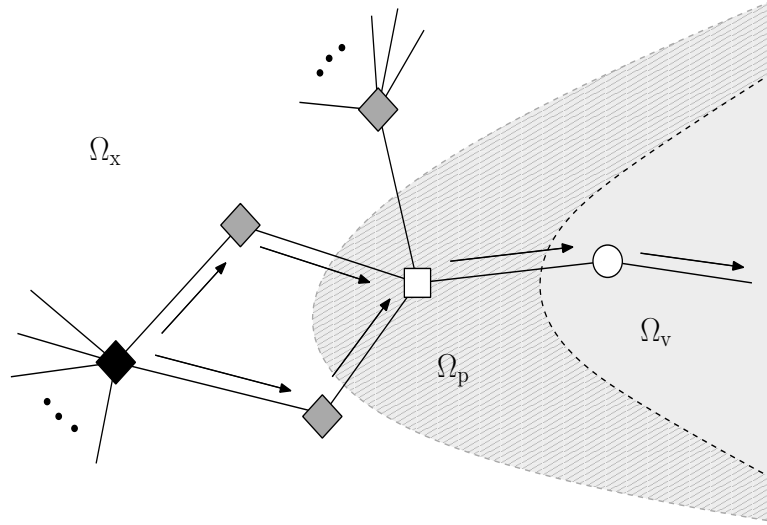


Figure 4.6: Illustration of a representative cyclical triggering motif (for $\phi = 0.33$) in the external component. Diamonds correspond to vertices in the external component (Ω_x), the square corresponds to a vertex in the peripheral component (Ω_p), and the circle to a vertex in the vulnerable component (Ω_v). The initial seed is placed in the black diamond. Spreading into Ω_v is indicated by the black arrows.

vertex in which the initial seed was placed is colored black, and is not part of the extended vulnerable component; neither are its two immediate neighbors (gray diamonds). However, both of these neighbors are vulnerable ($k = 2$), and will therefore become immediately infected. While the adjacent peripheral vertex ($k = 4$, white square) is stable, its threshold will be exceeded as it now has two infected neighbors, and it will subsequently infect its adjacent vulnerable vertex ($k = 2$, white circle), igniting the entire vulnerable component, and possibly the entire network, giving rise to a global cascade that is triggered from outside the extended vulnerable component Ω_e . Such triggering components were found with the highest frequency near the upper- z boundary of the cascade window, explaining the observed discrepancy between our numerical and analytic results (Fig. 4.1).

As shown in Fig. 4.5, the probability with which a vertex of degree k is attached to the vulnerable component (and consequently, the frequency with which it triggers a global cascade) is highly influenced by degree-degree correlations. For an average degree z sufficiently above the percolation threshold, global cascades typically topple the entire

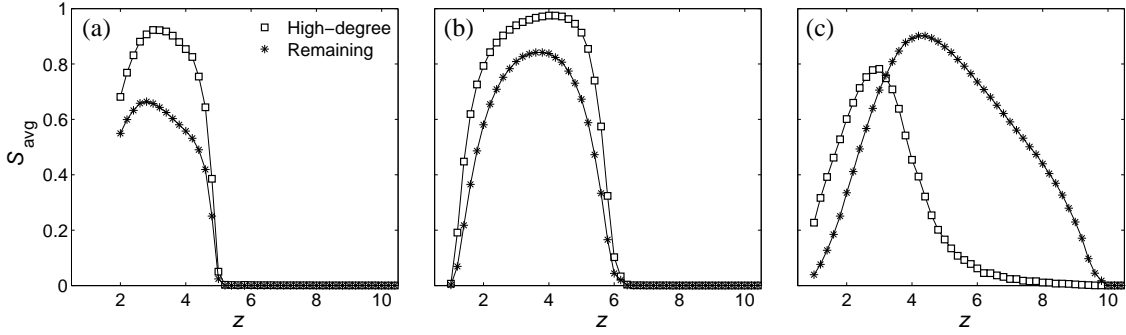


Figure 4.7: Average cascade size with $\phi = 0.18$ as a function of average degree z , given that the initial seed was placed in high-degree vertices (in the top 10% of the degree distribution, squares) or in the remaining vertices in the bottom 90% of the degree distribution (asterisks), for (a) disassortative ($r = -0.8$), (b) uncorrelated ($r = 0$), and (c) assortative ($r = +0.8$) random networks. Note that in (a), we were not able to obtain data for $z < 2$ (Table 4.1). Lines are provided as a guide for the eye.

network (Watts and Dodds, 2007) (Fig. 4.3). Taken together, these two observations imply that the relationship between the degree of the vertex in which the initial seed is placed and the average size of the cascade (S_{avg}) it triggers is also heavily influenced by degree-degree correlations. To illustrate this effect, we present in Fig. 4.7 the average cascade size for $\phi = 0.18$ as a function of average degree z for disassortative (Fig. 4.7a), uncorrelated (Fig. 4.7b), and assortative (Fig. 4.7c) random networks. We make the arbitrary distinction between high-degree vertices (those in the top 10% of the degree distribution p_k), and the remaining vertices (those in the bottom 90% of the degree distribution p_k) (Watts and Dodds, 2007). In disassortative networks (Fig. 4.7a), the high-degree vertices consistently trigger larger average cascades than do the remaining vertices. This observation stems from the negative degree correlation in disassortative networks, where highly connected vertices have a high probability of connecting to vulnerable vertices in the extended vulnerable component (Ω_e). In contrast, vertices with only a few connections are likely to connect to stable vertices in the external component (Ω_x). In uncorrelated random networks (Fig. 4.7b), high-degree vertices offer only a marginal increase in average cascade size over those triggered by other vertices, as discussed in (Watts and Dodds, 2007). However, in assortative networks (Fig. 4.7c), the ability of high-degree vertices to trigger large cascades decreases

with increasing z , and in networks with $z > 3$ the vertices in the bottom 90% of the degree distribution become dominant in triggering global cascades. This occurs because, in networks with high z and strong positive degree-degree correlations, highly connected stable vertices are likely to connect only with one another, forming stable clusters in Ω_x . In contrast, vulnerable vertices possess fewer connections and most often connect with one another, forming a dense core of vulnerable vertices in Ω_v , from which large cascades easily arise.

4.6 Discussion

Information cascades on random networks are often characterized as ‘robust yet fragile’ (Watts, 2002). When the network is sufficiently dense (high average degree z), information cascades occur relatively infrequently, but when they do occur, they often spread throughout the entire network. We have shown that increasing assortativity in a random network can exacerbate this effect, especially in networks whose degree-degree correlation is already positive (Figs. 4.1, 4.3a-d). Thus, not only do cascades occur more frequently in assortative networks, but they also typically cover a greater proportion of the network. While this may be perceived as a blessing to marketing strategists, it is cause for concern regarding other forms of social contagion, such as misinformation and disease. Conversely, the disassortative degree-degree correlations of technological and biological networks (Newman, 2002) may render these systems more resilient to small perturbations.

Increasing assortativity led to an earlier onset (*i.e.*, at lower z) of a percolating extended vulnerable component and an overall change in the shape of the functional relationship between global cascade frequency and average degree (Fig. 4.2). This result is in line with observations concerning the giant connected component in networks with heavy-tailed degree distributions (Newman, 2002). However, in those networks, the maximum size of the giant connected component was found to be larger in disassortative networks than in uncorrelated or assortative networks (Newman, 2002). In contrast, our experiments

on random networks with Poisson degree distributions demonstrate that the size of the extended vulnerable component is largest when degree-degree correlations are assortative (Fig. 4.2).

In uncorrelated random networks with Poisson degree distributions, it is known (Watts, 2002) that global cascades starting from a single initial seed do not occur when $\phi > 0.25$ (Fig. 4.1b). However, for low average degree z , we observed global cascades for $\phi = 0.33$ in both disassortative (Fig. 4.1a) and assortative (Fig. 4.1c) networks. This resulted in a bimodal response of cascade frequency to r (Figs. 4.3e and 4.4), as opposed to the unimodal response observed with $\phi \leq 0.25$ (Fig. 4.3a,c). Another deviation from the behavior observed with $\phi \leq 0.25$ is the resulting size of a global cascade (Fig. 4.3f), which was most often far smaller than the size of the entire system. We note that these interesting phenomena in both global cascade frequency and size are not immediately apparent from the analytical treatments of this model (Dodds and Payne, 2009; Gleeson, 2008a), or from other models of information spreading on degree-correlated networks (e.g., (Kiss et al., 2008)). While this counterintuitive response of cascade frequency and size is interesting in its own right, it is important to point out that bimodality only occurs for low average degree, which is a special case for disassortative networks. For example, when $z = 2$ (Fig. 4.4), the observed maximum degree in any network was only 10. This presents some anomalous constraints on the negative degree correlations found in highly disassortative networks, as the vulnerable degree 2 and 3 vertices have a high probability of connecting to one another (Fig. 4.5a), whereas with higher z these vertices would be more likely to connect to higher degree stable vertices. This can be clearly seen in Fig. 4.3e, where the peak in global cascade frequency decays much more rapidly as z increases in the $r < 0$ peak than in the $r > 0$ peak.

Identifying which vertices will trigger large information cascades has long been of interest, both in applications where the desire is to promote such cascades (as in marketing (Katz and Lazarsfeld, 1955)) as well as in applications where the desire is to prevent such cascades (as on the electrical power grid (Kinney et al., 2005) or in epidemiology (Kiss

et al., 2008)). A common assumption among the social scientists and marketing strategists is that the most connected individuals (so-called ‘influentials’) are the best candidates for triggering a sizable information cascade (Katz and Lazarsfeld, 1955). Recent results in uncorrelated random networks (Watts and Dodds, 2007) have demonstrated that cascades resulting from seeding high-degree vertices (those in the top 10% of the degree distribution) were actually not much larger than those emanating from vertices in the bottom 90% of the degree distribution (this result is recreated in Fig. 4.7b). Here, we found that cascades are possible in much denser networks that are positively assortative, and in these networks the lower degree vertices are responsible for triggering the largest cascades (Fig. 4.7c), resulting from an amalgamation of low-degree vulnerable vertices (a similar result was obtained in Watts and Dodds (2007) for group-based networks). This result may be of direct relevance to the social science and marketing communities, and may speak to the relative success of viral marketing campaigns that leverage online social networking (Freeman and Chapman, 2008), such as the Obama campaign (Stelter, 2008; Tumulty, 2007). In contrast, in disassortative networks the cascade window is restricted to sparser networks, and in these networks we found that cascades resulting from seeding high-degree vertices does result in much larger cascades (Fig. 4.7a). These results may indicate that targeted infection of high degree vertices may be an effective strategy for disseminating information in networks with negative degree-degree correlations, such as the internet.

References

- Barabási, A. L. and R. Albert (1999). Emergence of scaling in random networks. *Science* 286, 509–512.
- Barrat, A., M. Barthélemy, and A. Vespignani (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press.
- Boguñá, M., R. Pastor-Satorras, and A. Vespignani (2003). Absence of epidemic threshold in scale-free networks with degree correlations. *Physical Review Letters* 90, 028701.
- Centola, D., V. M. Eguíluz, and M. W. Macy (2007). Cascade dynamics of complex propagation. *Physica A* 374, 449–456.
- Dodds, P. S. and J. L. Payne (2009). Analysis of a threshold-model of social contagion on degree-correlated networks. *Physical Review E*. Submitted, preprint available at arXiv:0903.0597v1.
- Dodds, P. S. and D. J. Watts (2004). Universal behavior of a generalized model of contagion. *Physical Review Letters* 92, 218701.
- Freeman, B. and S. Chapman (2008). Gone viral? Heard the Buzz? A guide for public health practitioners and researchers on how Web 2.0 can subvert advertising restrictions and spread health information. *Journal of Epidemiology and Community Health* 62, 778–782.

- Galstyan, A. and P. Cohen (2007). Cascading dynamics in modular networks. *Physical Review E* 75, 036109.
- Gleeson, J. P. (2008a). Cascades on correlated and modular random networks. *Physical Review E* 77, 046117.
- Gleeson, J. P. (2008b). Mean size of avalanches on directed random networks with arbitrary degree distributions. *Physical Review E* 77, 057101.
- Gleeson, J. P. and D. J. Cahalane (2007). Seed size strongly affects cascades on random networks. *Physical Review E* 75, 056103.
- Katz, E. and P. F. Lazarsfeld (1955). *Personal Influence: the Part Played by People in the Flow of Mass Communications*. Free Press, Glencoe, IL.
- Kinney, R., P. Crucitti, R. Albert, and V. Latora (2005). Modeling cascading failures in the north american power grid. *European Physical Journal B* 46, 101–107.
- Kiss, I. Z., D. M. Green, and R. R. Kao (2008). The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *Journal of the Royal Society Interface* 5, 791.
- Milo, R., N. Kashtan, S. Itzkovitz, M. E. J. Newman, and U. Alon (2003). On the uniform generation of random graphs with prescribed degree sequences.
- Murray, J. D. (2002). *Mathematical Biology* (2nd ed.). Springer, New York.
- Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters* 89, 208701.
- Newman, M. E. J. (2003). Mixing patterns in networks. *Physical Review E* 67, 026126.
- Pastor-Satorras, R. and A. Vespignani (2001). Epidemic spreading in scale-free networks. *Physical Review Letters* 86(4), 3200–3203.

- Payne, J. L. and M. J. Eppstein (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*, pp. 241–248. ACM Press, New York.
- Payne, J. L. and M. J. Eppstein (2009). Evolutionary dynamics on scale-free interaction networks. *IEEE Transactions on Evolutionary Computation*. To appear.
- Rong, Z., X. Li, and X. Wang (2007). Roles of mixing patterns in cooperation on a scale-free networked game. *Physical Review E* 76, 027101.
- Sachtjen, M. L., B. A. Carreras, and V. E. Lynch (2000). Disturbances in power transmission system. *Physical Review E* 61, 4877.
- Schelling, T. C. (1973). Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities. *Journal of Conflict Resolution* 17, 381.
- Sood, V., T. Antal, and S. Redner (2008). Voter models on heterogeneous networks. *Physical Review E* 77, 041121.
- Stelter, B. (2008, July 7). The Facebooker Who Friendened Obama. The New York Times.
- Tumulty, K. (2007, July 5). Obama’s Viral Marketing Campaign. Time Magazine.
- Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Science* 99, 5766.
- Watts, D. J. and P. S. Dodds (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research* 34, 441.
- Yin, Y., D. Zhang, G. Pan, M. He, and J. Tan (2007). Sandpile on scale-free networks with assortative mixing. *Physical Scripta* 76, 606–612.

Chapter 5

Concluding Remarks

The results of this dissertation have advanced our understanding of how interaction topologies affect information flow in two simple classes of models. We have provided a general method for approximating takeover dynamics in regular interaction topologies; quantified the selective pressures induced by scale-free, degree-correlated interaction topologies; and characterized the conditions in which global information cascades occur in degree-correlated Poisson-distributed random interaction topologies.

This dissertation has laid the foundation for numerous directions of future research. In the context of evolutionary computation, the degree to which scale-free population structures can be exploited to improve search performance remains an open question. However, our results suggest that a systematic investigation of the influence of the structural properties of scale-free interaction topologies on the performance of evolutionary optimization algorithms is warranted, and we have recently commenced research in this direction. Specifically, we have investigated the performance of genetic algorithms on two benchmark optimization problems: massively multimodal deceptive problems (Goldberg et al., 1992) and NK landscapes (Kauffman, 1993). These preliminary experiments provide tentative evidence that the search performance obtained using scale-free population structures improves as the scaling exponent of the power-law degree distribution increases. Thus, decreased

degree heterogeneity appears to improve evolutionary search. Assortativity was found to have a smaller, secondary effect on solution quality, with performance generally improving with increasing assortativity. In both of the benchmark problems considered, higher quality solutions were found, on average, using a simple two-dimensional lattice population structure than with any scale-free interaction topology. However, the variability of the best solutions found between independent experiments was dramatically higher with scale-free population structures than with lattice-based population structures. For highly multimodal or epistatic problems, which are known to be difficult for evolutionary algorithms, it may be possible to capitalize on this variability by performing an ensemble of trials on numerous graph instances and then taking the best solution from these independent runs. Further experimentation is required to test this hypothesis.

Our results demonstrate that takeover times in scale-free interaction networks can be simply described using a planar function of two topological properties. However, a mechanistic explanation of this result has yet to be developed. This, in addition to a rigorous analytical model of takeover dynamics, would be a useful complement to our experimental results. Recent analytical treatments of the Voter Model (Sood et al., 2008) may provide a useful starting point.

In the context of threshold models of binary decisions with externalities, there are several promising avenues of research. The first is to calculate the frequency with which cyclical triggering motifs appear in finite networks, as a function of network size, connectivity, assortativity, and vertex threshold. As analytical treatments of threshold-based spreading models assume infinitely large, perfectly branching networks, such triggering motifs are not considered. However, real-world networks are by definition finite and are often degree-correlated. Understanding the conditions in which triggering motifs appear and determining their frequency may therefore aid in the proper application of threshold-models to a variety of real-world domains.

Several other extensions involve simple changes to the model formulation. For example,

information could possess an inherent quality; the size of the state space could be increased; vertices could possess an incubation or refractory period; vertices could be forgetful, or information may incrementally degrade in each transferral, analogous to the well-known “telephone game.” Another possible direction of future research is to investigate the influence of degree-degree correlations on information cascades in scale-free networks. While Watts (2002) found that the introduction of power-law degree distributions limited the conditions in which global information cascades occurred, it may be possible that the introduction of degree-degree correlations changes this picture, as was the case for Poisson-distributed random networks.

The interaction networks employed in models of real-world systems and evolutionary algorithms are typically static. While this has provided a wealth of useful insight regarding the relationship between network structure and dynamical processes, an obvious direction of future research is to investigate the incorporation of dynamically changing interaction networks. Recent results in this area are encouraging, with dynamic linking in evolutionary models providing novel insights into the emergence of cooperation (Pacheco et al., 2008), and adaptively changing population structures in evolutionary algorithms improving search performance and diversity maintenance (Whitacre et al., 2008). As dynamic interaction networks are a closer approximation to many real-world systems, their incorporation may allow for the development of more realistic models of dynamical processes and for the design of population-based optimization algorithms that more closely resemble the systems from which they draw their inspiration.

Comprehensive Bibliography

- Alba, E. and B. Dorronsoro (2004). Solving the vehicle routing problem by using cellular genetic algorithms. In J. Gottlieb and G. R. Raidl (Eds.), *Evolutionary Computation in Combinatorial Optimization*, pp. 11–20. Springer-Verlag, Berlin.
- Alba, E. and B. Dorronsoro (2005). The exploration/exploitation tradeoff in dynamic cellular genetic algorithms. *IEEE Transactions on Evolutionary Computation* 9(2), 126–142.
- Alba, E. and B. Dorronsoro (2008). *Cellular Genetic Algorithms*. Springer-Verlag.
- Alba, E. and M. Tomassini (2002). Parallelism and evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 6(5), 443–462.
- Albert, R., H. Jeong, and A. L. Barabási (1999). Diameter of the World-Wide Web. *Nature* 401, 130–131.
- Albert, R., H. Jeong, and A. L. Barabási (2000). Error and attack tolerance of complex networks. *Nature* 406, 378–382.
- Altenberg, L. (2005). Evolvability suppression to stabilize far-sighted adaptations. *Artificial Life* 11(4), 427–443.
- Anderson, R. M. and R. M. May (1995). *Infectious Diseases of Humans*. Oxford University Press, Oxford.
- Barabási, A. L. and R. Albert (1999). Emergence of scaling in random networks. *Science* 286, 509–512.

- Barabási, A. L., E. Ravasz, and T. Vicsek (2001). Deterministic scale-free networks. *Physica A* 299, 559–564.
- Barrat, A., M. Barthélemy, and A. Vespignani (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press.
- Boguñá, M., R. Pastor-Satorras, and A. Vespignani (2003). Absence of epidemic threshold in scale-free networks with degree correlations. *Physical Review Letters* 90, 028701.
- Borrett, S. R., B. D. Fath, and B. C. Patten (2007). Functional integration of ecological networks through pathway proliferation. *Journal of Theoretical Biology* 245, 98–111.
- Bryden, K. M., D. Ashlock, S. Corns, and S. Wilson (2005). Graph based evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 10(5), 550–567.
- Caldarelli, G., A. Capocci, P. De Los Rios, and M. A. Munoz (2002). Scale-free networks from varying vertex intrinsic fitness. *Physical Review Letters* 89(25), 258702.
- Cantú-Paz, E. (2000). Markov chain models of parallel genetic algorithms. *IEEE Transactions on Evolutionary Computation* 4(3), 216–226.
- Caswell, H. (2001). *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer Associated, Sunderland.
- Centola, D., V. M. Eguíluz, and M. W. Macy (2007). Cascade dynamics of complex propagation. *Physica A* 374, 449–456.
- Chakraborty, U. K., L. Deb, and M. Chakraborty (1997). Analysis of selection algorithms: a markov chain approach. *Evolutionary Computation* 4(2), 133–167.
- Clauset, A., C. R. Shalizi, and M. E. J. Newman (2009). Power-law distributions in empirical data. *arXiv:0706.1062v1*.

- Dodds, P. S. and J. L. Payne (2009). Analysis of a threshold-model of social contagion on degree-correlated networks. *Physical Review E*. Submitted, preprint available at arXiv:0903.0597v1.
- Dodds, P. S. and D. J. Watts (2004). Universal behavior of a generalized model of contagion. *Physical Review Letters* 92, 218701.
- Eames, K. T. D. and M. J. Keeling (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., L. Mielsch, and S. Bornholdt (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 J. Molofsky (2007). Invasiveness in plant communities with feedbacks. *Ecology Letters* 10, 253–263.
- Eppstein, M. J., J. L. Payne, and C. J. Goodnight (2009). Underdominance, multiscale interactions, and self-organizing barriers to gene flow. *Artificial Evolution and Applications*. Submitted.
- Erdős, P. and A. Rényi (1959). On random graphs. *Publicationes Mathematicae* 6, 290–297.
- Euler, L. (1741). Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 8, 128–140.
- Falconer, D. S. and T. F. C. Mackay (1996). *Quantitative Genetics*. Pearson Education Limited, London.
- Faloutsos, M., P. Faloutsos, and C. Faloutsos (1999). On power-law relationships of the internet topology. *Computer Communication Review* 29, 251–262.

- Folino, G., C. Pizzuti, and G. Spezzano (2001). Parallel hybrid method for sat that couples genetic algorithms and local search. *IEEE Transactions on Evolutionary Computation* 5(4), 323–334.
- Folino, G., C. Pizzuti, and G. Spezzano (2003a). A scalable cellular implementation of parallel genetic programming. *IEEE Transactions on Evolutionary Computation* 7(1), 37–53.
- Folino, G., C. Pizzuti, G. Spezzano, L. Vanneschi, and M. Tomassini (2003b). Diversity analysis in cellular and multipopulation genetic programming. In R. Sarker (Ed.), *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 305–311. IEEE Press.
- Freeman, B. and S. Chapman (2008). Gone viral? Heard the Buzz? A guide for public health practitioners and researchers on how Web 2.0 can subvert advertising restrictions and spread health information. *Journal of Epidemiology and Community Health* 62, 778–782.
- Galstyan, A. and P. Cohen (2007). Cascading dynamics in modular networks. *Physical Review E* 75, 036109.
- Gasparri, A., S. Panziera, F. Pascucci, and G. Ulivi (2007). A spatially structured genetic algorithm over complex networks for mobile robot localisation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 4277–4282. IEEE Press.
- Giacobini, M., M. Preuss, and M. Tomassini (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. Springer-Verlag, Heidelberg.
- Giacobini, M., M. Tomassini, and A. Tettamanzi (2005a). Takeover times curves in random and small-world structured populations. In H. G. Beyer (Ed.), *Proceedings of the Genetic*

- and *Evolutionary Computation Conference, GECCO-2005*, pp. 1333–1340. ACM Press, New York.
- Giacobini, M., M. Tomassini, A. Tettamanzi, and E. Alba (2005b). Selection intensity in cellular evolutionary algorithms for regular lattices. *IEEE Transactions on Evolutionary Computation* 9(5), 489–505.
- Gleeson, J. P. (2008a). Cascades on correlated and modular random networks. *Physical Review E* 77, 046117.
- Gleeson, J. P. (2008b). Mean size of avalanches on directed random networks with arbitrary degree distributions. *Physical Review E* 77, 057101.
- Gleeson, J. P. and D. J. Cahalane (2007). Seed size strongly affects cascades on random networks. *Physical Review E* 75, 056103.
- Goldberg, D. and K. Deb (1991). A comparative analysis of selection schemes used in genetic algorithms. In G. Rawlins (Ed.), *Foundations of Genetic Algorithms*, pp. 69–93. Morgan Kaufmann, San Francisco.
- Goldberg, D., K. Deb, and J. Horn (1992). Massive multimodality, deception, and genetic algorithms. In R. Männer and B. Manderick (Eds.), *Parallel Problem Solving From Nature*, pp. 37–46. North-Holland, Amsterdam.
- 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. Morgan Kaufmann, San Francisco.
- Granovetter, M. (1978). Threshold models of collective behavior. *The American Journal of Sociology* 83(6), 1420–1443.

- Grefenstette, J. J. (1981). Parallel adaptive algorithms for function optimization. Technical Report CS-81-19, Vanderbilt University, Computer Science Department.
- Hauert, C. and M. Doebeli (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.
- Herrera, F. and M. Lozano (2000). Gradual distributed real-coded genetic algorithms. *IEEE Transactions on Evolutionary Computation* *4*(1), 43–63.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Ph. D. thesis, University of Michigan.
- Holme, P. and B. J. Kim (2002). Growing scale-free networks with tunable clustering. *Physical Review E* *65*, 026107.
- Jeong, H., S. P. Mason, A. L. Barabási, and Z. N. Oltvai (2001). Lethality and centrality in protein networks. *Nature* *411*, 41.
- Jeong, H., B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási (2000). The large-scale organization of metabolic networks. *Nature* *407*, 651–654.
- Joo, J. and J. L. Lebowitz (2004). Pair approximation of the stochastic susceptible-infected-recovered-susceptible model on the hypercubic lattice. *Physical Review E* *70*, 036114.
- Katz, E. and P. F. Lazarsfeld (1955). *Personal Influence: the Part Played by People in the Flow of Mass Communications*. Free Press, Glencoe, IL.
- Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.
- 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. (2005). The implications of network structure for epidemic dynamics. *Theoretical Population Biology* 67, 1–8.
- Keeling, M. J. and K. T. D. Eames (2005). Networks and epidemic models. *Journal of the Royal Society Interface* 2, 295–307.
- Kennedy, J. and R. Mendes (2002). Population structure and particle swarm performance. In *Proceedings of the Congress of Evolutionary Computation Conference*, Washington, DC, USA, pp. 1671–1676. IEEE Press.
- Kerr, B., M. A. Riley, M. W. Feldman, and B. J. M. Bohannan (2002). Local dispersal promotes biodiversity in a real life game of rock-paper-scissors. *Nature* 418, 171–174.
- Kinney, R., P. Crucitti, R. Albert, and V. Latora (2005). Modeling cascading failures in the north american power grid. *European Physical Journal B* 46, 101–107.
- Kirley, M. and R. Stewart (2007a). 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. ACM Press, New York.
- Kirley, M. and R. Stewart (2007b). Multiobjective optimization on complex networks. In *Fourth International Conference on Evolutionary Multi-Criterion Optimization*, pp. 81–95. Lecture Notes Computer Science.
- Kiss, I. Z., D. M. Green, and R. R. Kao (2008). The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. *Journal of the Royal Society Interface* 5, 791.
- Krapivsky, P. L., S. Redner, and F. Leyvraz (2000). Connectivity in growing random networks. *Physical Review Letters* 85(21), 4629.

- Lieberman, E., C. Hauert, and M. A. Nowak (2001). Evolutionary dynamics on graphs. *Nature* 433, 312–316.
- Lienig, J. (1997). A parallel genetic algorithm for performance-driven VLSI routing. *IEEE Transactions on Evolutionary Computation* 1(1), 29–39.
- Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg (2001). The web of human sexual contacts. *Nature* 411, 907–908.
- Manderick, B. and P. Spiessens (1989). Fine-grained parallel genetic algorithms. In J. D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms*, pp. 428–433. Morgan Kaufmann, San Mateo.
- Maslov, S. and K. Sneppen (2002). Specificity and stability in topology of protein networks. *Science* 296, 910–913.
- Matsuda, H., N. Ogita, A. Sasaki, and K. Soto (1992). Statistical mechanics of population: the lattice lotka-volterra model. *Progress of Theoretical Physics* 88, 1035–1049.
- Meyers, L. A., M. E. J. Newman, and B. Pourbohloul (2006). Predicting epidemics on directed contact networks. *Journal of Theoretical Biology* 240(3), 400–418.
- Milo, R., N. Kashtan, S. Itzkovitz, M. E. J. Newman, and U. Alon (2003). On the uniform generation of random graphs with prescribed degree sequences.
- Molloy, M. and B. Reed (1995). A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms* 6, 161–180.
- Murray, J. D. (2002). *Mathematical Biology* (2nd ed.). Springer, New York.
- Newman, M. E. J. (2002a). Assortative mixing in networks. *Physical Review Letters* 89, 208701.
- Newman, M. E. J. (2002b). Spread of epidemic disease on networks. *Physical Review E* 66, 016128.

- Newman, M. E. J. (2003a). Mixing patterns in networks. *Physical Review E* 67, 026126.
- Newman, M. E. J. (2003b). The structure and function of complex networks. *SIAM Review* 45(2), 167–256.
- Newman, M. E. J., A. L. Barabási, and D. J. Watts (Eds.) (2006). *The Structure and Dynamics of Networks*. Princeton University Press.
- Nowak, M. A. and R. M. May (1992). Evolutionary games and spatial chaos. *Nature* 359, 826–829.
- Ohtsuki, H., C. Hauert, E. Lieberman, and M. A. Nowak (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502–505.
- Ong, N. P. and R. J. Cava (2004). Electronic frustration on a triangular lattice. *Science* 305, 52–53.
- Pacheco, J. M., A. Traulsen, H. Ohtsuki, and M. A. Nowak (2008). Repeated games and direct reciprocity under active linking. *Journal of Theoretical Biology* 250, 723–731.
- Park, K., Y. C. Lai, and N. Ye (2005). Self-organized scale-free networks. *Physical Review E* 72, 026131.
- Pastor-Satorras, R. and A. Vespignani (2001). Epidemic spreading in scale-free networks. *Physical Review Letters* 86(4), 3200–3203.
- Payne, J. L. and M. J. Eppstein (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. ACM Press, New York.
- Payne, J. L. and M. J. Eppstein (2007). Takeover times on scale-free topologies. In D. Thierens (Ed.), *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2007*, pp. 308–315. ACM Press, New York.

- Payne, J. L. and M. J. Eppstein (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*, pp. 241–248. ACM Press, New York.
- Payne, J. L. and M. J. Eppstein (2009a). Evolutionary dynamics on scale-free interaction networks. *IEEE Transactions on Evolutionary Computation*. To appear.
- Payne, J. L. and M. J. Eppstein (2009b). Pair approximations of takeover dynamics in regular population structures. *Evolutionary Computation*. To appear.
- Payne, J. L., M. J. Eppstein, and C. J. Goodnight (2007). Sensitivity of self-organized speciation to long-distance dispersal. In *Proceedings of the IEEE Symposium on Artificial Life*, pp. 1–7.
- Petermann, T. and P. De Los Rios (2004). Cluster approximations for epidemic processes: a systematic description of correlations beyond the pair level. *Journal of Theoretical Biology* 229, 1–11.
- Poncela, J., J. Gómez-Gardenes, L. M. Florià, and Y. Moreno (2007). Robustness of cooperation in the evolutionary prisoner’s dilemma on complex networks. *New Journal of Physics* 9, 184.
- Rauch, E. M. and Y. Bar-Yam (2006). Long-range interactions and evolutionary stability in predator-prey systems. *Physical Review E* 73, 020903.
- Ravasz, E., A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabási (2002). Hierarchical organization of modularity in metabolic networks. *Science* 297, 1551–1555.
- Reichenbach, T., M. Mobilia, and E. Frey (2007). Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* 448, 1046–1049.

- Rong, Z., X. Li, and X. Wang (2007). Roles of mixing patterns in cooperation on a scale-free networked game. *Physical Review E* 76, 027101.
- Rudolph, G. (2000). On takeover times in spatially structured populations: array and ring. In K. K. Lai, O. Katai, M. Gen, and B. Lin (Eds.), *Proceedings of the Second Asia-Pacific Conference on Genetic Algorithms and Applications, APGA-2000*, pp. 144–151. Global Link Publishing Company, Hong Kong.
- 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. Springer-Verlag, Heidelberg.
- Sachtjen, M. L., B. A. Carreras, and V. E. Lynch (2000). Disturbances in power transmission system. *Physical Review E* 61, 4877.
- Santos, F. C. and J. M. Pacheco (2005). Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters* 95, 098104.
- Sarma, J. and K. De Jong (1996). An analysis of the effect of neighborhood size and shape on local selection algorithms. In H. M. Voigt, W. Ebeling, I. Rechenberg, and H. P. Schwefel (Eds.), *Parallel Problem Solving from Nature*, pp. 236–244. Springer-Verlag, Heidelberg.
- Satō, K. and Y. Iwasa (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 University Press, Cambridge.
- Satō, K., H. Matsuda, and A. Sasaki (1994). Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology* 32, 251–268.
- Sayama, H., L. Kaufman, and Y. Bar-Yam (2003). Spontaneous pattern formation and genetic diversity in habitats with irregular geographic features. *Conservation Biology* 17, 893–900.

- Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology* 1, 143–186.
- Schelling, T. C. (1973). Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities. *Journal of Conflict Resolution* 17, 381.
- Servedio, V. D. P., G. Calarelli, and P. Butta (2004). Vertex intrinsic fitness: how to produce arbitrary scale-free networks. *Physical Review E* 70, 056126.
- Sood, V., T. Antal, and S. Redner (2008). Voter models on heterogeneous networks. *Physical Review E* 77, 041121.
- Spector, L. and J. Klein (2005). Trivial geography in genetic programming. In T. Yu, R. L. Riolo, and B. Worzel (Eds.), *Genetic Programming Theory and Practice III*, pp. 109–124. Kluwer Academic Publishers.
- 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. Morgan Kaufmann, San Francisco.
- Stelter, B. (2008, July 7). The Facebooker Who Friendened Obama. *The New York Times*.
- Tumulty, K. (2007, July 5). Obama’s Viral Marketing Campaign. *Time Magazine*.
- Van Baalen, M. (2000). Pair approximations for different spatial geometries. In U. Dieckmann, R. Law, and J. A. J. Metz (Eds.), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, pp. 359–387. Cambridge University Press, Cambridge.
- Van Baalen, M. and D. A. Rand (1998). The unit of selection in viscous populations and the evolution of altruism. *Journal of Theoretical Biology* 193, 631–648.
- Van Veldhuizen, D. A., J. B. Zydallis, and G. B. Lamont (2003). Considerations in engineering parallel multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation* 7(2), 144–173.

- Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Science* 99, 5766.
- Watts, D. J. and P. S. Dodds (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research* 34, 441.
- Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of small-world networks. *Nature* 393, 440–442.
- Werfel, J. and Y. Bar-Yam (2004). The evolution of reproductive restraint through social communication. *Proceedings of the National Academy of Science* 101(30), 11019–11024.
- Whigham, P. A. and G. Dick (2008). Evolutionary dynamics for the spatial moran process. *Genetic Programming and Evolvable Machines* 12(2), 220–230.
- Whitacre, J. M., R. A. Sarker, and Q. T. Pham (2008). The self-organization of interaction networks for nature-inspired optimization. *IEEE Transactions on Evolutionary Computation* 12(2), 220–230.
- Xulvi-Brunet, R. and I. M. Sokolov (2004). Reshuffling scale-free networks: from random to assortative. *Physical Review E* 70, 066102.
- Yin, Y., D. Zhang, G. Pan, M. He, and J. Tan (2007). Sandpile on scale-free networks with assortative mixing. *Physical Scripta* 76, 606–612.