Stability in Graph Dynamical Systems

Joseph A. McNitt

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Mathematics

Henning S. Mortveit, Chair
Jeff T. Borggaard
William J. Floyd

2018
Blacksburg, Virginia

Keywords: Stability, Sensitivity, Graph dynamical systems, Networks, Tuta Absoluta

Stability in Graph Dynamical Systems

Joseph A. McNitt

Academic Abstract

The underlying mathematical model of many simulation models is graph dynamical systems (GDS). This dynamical system, its implementation, and analyses on each will be the focus of this paper. When using a simulation model to answer a research question, it is important to describe this underlying mathematical model in which we are operating for verification and validation. In this paper we discuss analyses commonly used in simulation models. These include sensitivity analyses and uncertainty quantification, which provide motivation for stability and structure-to-function research in GDS. We review various results in these areas, which contribute toward validation and computationally tractable analyses of our simulation model. We then present two new areas of research - stability of transient structure with respect to update order permutations, and an application of GDS in which a time-varying generalized cellular automata is implemented as a simulation model.

Stability in Graph Dynamical Systems

Joseph A. McNitt

General Audience Abstract

There are many systems in our society which are vital, and require quantitative analysis. These include population dynamics, transportation, and energy. To answer research questions about these systems, one may construct a mathematical model of the system and conduct simulations. It is important to define both the mathematical model and the simulation model in order to better understand the source of errors, or to be confident in the validity of the models. One source of error may be in parameters of our simulation model. It can be difficult to gather reliable and precise data, especially in massively interacting systems. Thus we would like to know that there is a range of values which will result in similar outcomes. Stability results can give us this assurance. This paper mainly focuses on stability results in graph dynamical systems (GDS), which is the underlying mathematical model of many simulation models, especially ones with a networked structure.

Acknowledgments

I would like to extend my thanks to all co-authors. Thank you especially to Dr. Mortveit for tremendous advising. Thank you to Dr. Adiga for advising and guiding me in many aspects of my time here. Thank you to Dr. Wu for his help in GENEUS and with the publication. Thank you to Arindam Fadikar for answering various questions throughout. Thank you to Dr. Lewis for guiding me in our project when I first arrived at NDSSL. Thank you to all others at the Biocomplexity Institute, NDSSL, and the Virginia Tech Mathematics Department who have contributed and helped.

Contents

Chapter 1

Background

Complex systems can occur in biology, economics, sociology, transportation, and others. These systems would benefit from analysis but are often too large or vital to expose to exhaustive experimentation. An economy changing, or a disease propagating through a population are two examples in which direct experimentation would be impractical, unethical, or simply impossible. Therefore these real systems must be approximately described and simulated.

Models can be created that provide a mathematical framework which describes a system. We can analyze the model in order to gain general information about the processes of these systems. In order to then simulate these processes, we implement the mathematical model as a computational simulation model. We first must understand the mathematical system in which we will focus, and which is the underpinning of many simulation models (in particular networked ones): graph dynamical systems (GDS).

A GDS contains three elements: a dependency graph, a function, and a state space. In general, our dependency graph G can be directed or undirected. With $n = |V(G)|$ and for some set K (called the *vertex state space*), each vertex, v, has a state, $x_v \in K$. The system state is then the collection of these vertex states, denoted $(x_1, ..., x_n) \in K^n$ where K^n is called the *system state space*. To complete the description, we define the GDS function $F : K^n \to K^n$. To construct F, we first construct *vertex functions* $(f_v)_v$ which are rules defining how each vertex state changes. Our graph G corresponds to the dependent variables in these functions in the following manner. For vertex v, with $N(v)$ as the closed neighborhood, each f_v is a function from $K^{|N(v)|} \to K$.

We now define the input to each f_v as $x[v] \in K^{|N(v)|}$, where each element is the state of the corresponding vertex in the closed neighborhood. This is defined more precisely as follows:

letting $n[v]$ be the sorted sequence of the neighbors of v along with v itself, we have

$$x[v] = (x_{n[v](1)}, x_{n[v](2)}, ..., x_{n[v](d(v)+1)}) \ ,$$

where $d(v)$ is the degree of vertex v. The GDS function, F, is assembled using some update scheme which is discussed in the next definitions. Once we have our GDS function, we denote the graph dynamical system as the triple (F, G, K^n).

Definition 1. *A parallel or synchronous GDS is one in which the GDS function, $F : K^n \to K^n$, is defined as $F(x) = (f_1(x), f_2(x), ..., f_n(x))$.*

Definition 2. *A sequential or asynchronous GDS is one in which a sequence of the vertices is given, π, and the GDS function $F_\pi : K^n \to K^n$ is defined as*

$$F_\pi = F_{\pi(n)} \circ F_{\pi(n-1)} \circ \cdots \circ F_{\pi(1)} \ .$$

Here $(F_v)_v$ is the set of *local functions* where $F_v(x) = (x_1, x_2, ..., f_v(x[v]), ..., x_n)$ and $x = (x_1, ..., x_n) \in K^n$.

Though these update schemes are most common, we can also generalize the update sequence to a block sequential update:

Definition 3. *We let G be a graph, and $\mathcal{B} = \{B_1, B_2, ..., B_m\}$ be some partition of $V(G)$. We define a function $F_{B_k} : K^n \to K^n$ as*

$$(F_{B_k}(x))_v = \begin{cases} f_v(x[v]) & \text{if } v \in B_k \\ x_v & \text{if } v \notin B_k \ . \end{cases}$$

For some $\pi \in S_m$, we have the block sequential map, $F_{\mathcal{B}} : K^n \to K^n$, defined as

$$F_{\mathcal{B}} = F_{B_{\pi(m)}} \circ F_{B_{\pi(m-1)}} \circ \cdots \circ F_{B_{\pi(1)}} \ .$$

Once we have our GDS function F, it is useful to work in the space of state transitions under function applications. Thus we define the phase space:

Definition 4. *Given a GDS function, F, the phase space, $\Gamma(F)$, is the directed graph in which the vertices are the system states $s \in K^n$, and for every state $s \in K^n$, there exists a directed edge $(s, F(s))$.*

We now define three concepts used to describe the dynamics in the phase space.

Definition 5. *For a GDS function F, the orbit of state $x \in K^n$ for a discrete dynamical system F is $\{x, F(x), F^2(x), ...\}$.*

Definition 6. *For a GDS function F, a periodic orbit is an orbit with some finite period T. In other words: for state x, a periodic orbit is such that for some $N > 0$, and some $T > 0$, we have $F^n(x) = F^{n+T}(x)$ for any $n \geq N$.*

Definition 7. *For a GDS function F, a limit cycle of length T is a sorted sequence of system states, $(x, F(x), ..., F^{T-1}(x))$ such that $F^T(x) = x$, T is minimal and $T \geq 1$. Thus it is the sequence of states which repeat in a periodic orbit.*

Definition 8. *For a GDS function F, a fixed point $x \in K^n$ is one in which $F(x) = x$, and the set of fixed points is denoted* $\mathrm{Fix}(F)$.

There are systems closely related to GDS that have been studied:

- **Automata networks**
  This is the same system as a GDS. It is called automata network because it was seen as a generalization of cellular automata (CA), which is a GDS on a regular grid. The motivation for this framework also comes from discrete automata.
  In [11], Goles and Martinez discuss this framework and especially possibilities of using Lyapunov-type functionals. These can give information about the limit cycles and will be discussed further in Section 1.2.2.

- **Polynomial dynamical systems**
  These are GDS where the vertex functions are polynomials of the neighboring vertex states. The GDS is updated synchronously. This approach was taken because of the algebraic structure which results, supporting the use of computational algebra. In [21], Laubenbacher applies this framework to modeling and analyzing a competition between viruses in a network of cells.

- **Boolean networks**
  These are GDS where the state space is $\{0, 1\}$ which can represent genes being on or off. In terms of a mathematical structure, this was studied as the minimum size of the state space. Kauffman [17] and Kauffman and Ribeiro [30] studied this application, especially limit cycle structure, in order to describe various cell types.

- **Interacting particle systems**
  These are stochastic systems with continuous time, constructed in a physics framework. They are collections of Markov processes. There is an underlying countable dependency graph, and thus it is similar to a GDS but there is stochasticity and time continuity. An example of an interacting particle system in physics is an Ising model. There is often a lattice structure that represents the dependencies, so it is similar to a CA. Vertices take states of either -1 or 1 representing spin, and thus we see similarities to a boolean network. The defining characteristics that push it into the category of an IPS is that it is continuous time and the transition function is stochastic [22].

- **Hopfield networks**
  These are similar to GDS in that there is an underlying graph defining the dependencies. The local functions are threshold functions (see Section 1.2.2 for definition) which do not depend on their own state, only states in the open neighborhood. Garimella [10] discusses conditions in which the dynamical system will reach a steady state. In other words there are only fixed points as limit cycles. These conditions were given by Hopfield [15] and Goles [13].

To capture a system with a mathematical model, assumptions and generalizations must typically be made. When modeling, we have a question or questions which need answering. To be sure that the mathematical model aligns with the real system we must validate the model to the extent necessary to answer the research questions. There are different aspects of the real system that can be prioritized. A mechanistic validation would focus on important features of the system and be sure they were represented. A predictive validation would be statistical and run the simulation with the inputs of a historical example and ensure it gives similar results. The process of evaluating whether the simulation model correctly implements the mathematical model is called verification. This verification and validation ensures that we can meaningfully use our simulation model to hypothesize about the underlying system.

In the section below, we will discuss common simulation models, as well as verification methods. This will give motivation for stability theory in the underlying mathematical model. Stability results can contribute toward verification of our simulation model. The section below will also motivate our later results which take some aspects of the structure of a GDS and gain information about the dynamics of that function (structure-to-function results). Structure-to-function results allow us to examine dynamics in a computationally tractable way.

1.1 Motivation

Graph Dynamical Systems can describe many simulation models, including interaction based models. In this section we will describe such simulation models, and reveal sensitivity analyses of quantities of interest with respect to our model parameters. This will be useful information in itself, but will also inform and encourage analyses of the mathematical model.

1.1.1 Description

An interaction-based model is defined at the level of distinct interactions. Each agent in a specified resolution is modeled, and the interactions between agents are distinct. Various

resolutions are possible - examples of an agent could be a person, a household, or a larger region. Each agent will have a set of possible states and rules which govern the adjustment of those states over time. Agents interact with other agents, which is captured by these rules.

This is in contrast to models in continuous systems such as differential equations or difference equations, which describes the average agent and average interactions. One positive of working with an interaction-based approach is that each agent can have its own rules. Huff et al. [16] studies secondary effects of a pandemic - especially the effect of a pandemic on food supply. They apply a system dynamics model to this system and use aggregated information about production, distribution, and consumption. This method is useful to find general sensitivities, but can miss important information about individual interactions. This reveals a possibility in interaction based models which is not present in other methods - heterogeneity among the agents. This could be useful specifying behavior and interventions at the level of the chosen agent as well as monitoring individual agents, but quickly becomes computationally intractable. Another difficulty is that it is more data demanding, as information must be gained at a finer resolution (e.g. per person/entity).

Parameterization and stochasticity

To make interaction based models more tractable, we consider parameterization and stochasticity. Often when implementing a mathematical model as a simulation model, we will include classes of different types of agents. These classes will often have similar update functions, and simply be parameterized. This takes away some heterogeneity as each agent may not have a unique function.

Heterogeneity can then be reintroduced through (random) parameters as in [5]. Barrett et al. created a model of traffic flow with simple acceleration and deceleration rules. Each car followed the same rules and thus it is an unencapsulated agent representation. The interaction based aspect allowed them to track each car's position individually, and thus compare the simulation results to traffic data. The simulation model results did not match the data well. But when a random deceleration was introduced, the simulation model could be calibrated to the data.

Encapsulated vs unencapsulated agents

In many interaction based models, the inputs to the agent rules come from various systems. An example is [6], in which Barrett et al. considered the public response to a detonation of an improvised nuclear device. They modeled the communication system, power network,

transportation, health dynamics, and individual behavior as well as dependencies between these. In an *encapsulated agent*, the system information for all of these would be included in each agent at all times. This can become computationally intractable. Thus to model these multiple systems in the same scenario, they factorize the simulation model by simulating each system for a short amount of time, and use the output to update the next system. Thus each system is not updated at each time step, but takes the output from the previous systems to use as input. The agents in this simulation model are called *unencapsulated agents* and they allow interdependencies to be accounted for in a tractable way.

Benefits of an unencapsulated agent based model include extensibility, clarity, and speed. It is relatively simple to add another system on top of an existing model. A question that is difficult to answer is: does this model come close to the system dynamics?

Multiple spatial and temporal scales

There are multiple aspects in which the scale of the model can vary. These include the temporal and spatial aspects. As our interaction based model has some update method, we must consider time scales. In the reference above [6], Barrett et al. considers time scales and ways in which they can be used to capture the desired model. As implementation of GDS is discrete, defining the time scale is necessary.

Spatial information can often be important to a given research question. One possible method to capture spatiality is to create a grid of the focus region. Each cell would then correspond to a vertex in a cellular automaton. In the spread of organisms or of disease, this can be a useful tool. Guimapi et al. [14] studied the spread of a pest in Africa using this method. Each cell had certain properties related to its area such as vegetation, humidity, and temperature. This could be defined as the parameters to each vertex function. The rules describing the state changes (infested/uninfested) are dependent on neighbor states and require a certain threshold value for each parameter. This framework is similar to a GDS, but the vertex functions are periodic due to the various temperature values across a season.

This application of a cellular automaton is also similar to some differential equations techniques. Through finite element or finite volume methods, one partitions the space in order to calculate.

There can also be multiple levels to our system. There may be a spatial scale of a disease spreading through short distances, and also through air travel. Balcan et al. [4] study the interplay between commuting and long distance travel during an infectious disease epidemic.

This requires not only multiple spatial scales but also multiple time scales, as long distance travel data is daily, whereas commuting data is hourly. To implement multiple spatial scales, Voronoi regions were constructed around airport locations. This allows a collection of the finer resolution commuting space into the air travel network. This can be an adequate way to implement interaction based models which contain multiscale networks.

Role of GDS in modeling

These interaction-based models align well with GDS, and are often implementations of these mathematical models [18]. We know that a GDS is a triple, so to represent interaction based models we must identify these three components. The dependency graph can be induced by each vertex function. The vertex functions will be the rules which are assigned to the updating of an agent's state. With a specified update scheme, these produce a GDS function. The state space will be the allowable states in the simulation model. Since the concept of our interaction based model can thus be described with a GDS, we need verification methods to ensure that the simulation model correctly implements the mathematical model. We can also find stability results in GDS to ensure that there are allowable uncertainties in aspects of our simulation model which will not cause a bifurcation in quantities of interest. Structure to function results in GDS allow for quick comparisons between different simulation models.

1.1.2 Analysis

This section describes possible analyses on the simulation model. This information and suggestions come from the National Academy of Sciences [8]. They discuss verification, uncertainty quantification, sensitivity analysis, and others.

Verification

We define verification as the comparison between the simulation model implementation and the mathematical model describing the real system. There are two aspects of verification, code verification and solution verification. Code verification ensures that our program has no bugs. Solution verification attempts to accurately quantify and limit numerical error. This is especially important with a continuous model such as PDEs.

To conduct verification, we may utilize the hierarchical nature of a computer code, and assess smaller methods first. We can use test problems to verify different aspects of our code. Test problems are a set of inputs to the problem in which the solution is known. Thus we can input these values and ensure that the output is correct.

Uncertainty quantification and sensitivity analysis

When working with simulation models, understanding which aspects have the greatest effect on certain dynamics is generally required for system understanding, for focusing computational burden, analysis, predictions, and validation.

There are two ways in which to study sensitivity - one is a computational/statistical approach in which we study the simulation model and its output, and one is an analytic approach in which we study the mathematical model. In the following, we will consider methods and issues surrounding computational/statistical sensitivity analysis. We will then discuss results in analysis of sensitivities in GDS.

Sensitivity analysis (SA) and uncertainty quantification (UQ) are two analysis classes that quantify variance in simulation model output. Uncertainty quantification is the study of causes of variation in an estimated quantity of interest (QOI) [8]. There are two possible sources of variation in computed QOI - variation in inputs and stochasticity. So an important step is to quantify both input variation and stochastic variation. Some of these come from variation in the underlying system, but some come from lack of knowledge of the underlying system. A direct way to then conduct uncertainty quantification would be to select a range of input variation and simulate to understand how it affects the output variation, but we would like to have a less time-consuming way to do this.

UQ is dealt with in terms of probability, we need to understand what is probabilistic in the nature of the underlying system and what is probabilistic in the model itself. To understand how variation propagates, simulations must be performed. There are different types of statistical designs which can be utilized. A full factorial design is the most expensive but also the most detailed. This design spans all variation at high resolution and runs simulations for all of these inputs. This is expensive and usually infeasible.

One option to limit the number of simulations is sampling. This process allows for a smaller amount of simulations in order to get data points that will fit to another less expensive model. So a function is estimated from these data points and then a full-factorial design can be implemented on this model. One example is a Gaussian process interpolator, as well as other regressions. This will lower computational cost, but many aspects of the behavior of the model are lost.

Another similar method is model reduction. This gives an estimate of the model by reducing the number of dimensions or degrees of freedom [8]. One question that comes up in this area is which variables could be discarded. It can be difficult to accurately reduce non-linear models.

A third method spoken of in [8] is forward propagation of uncertainty. This speaks to limiting a full-factorial design with ideas such as Monte Carlo sampling. One issue with this approach is that rare but consequential parameter values can often be missed. This can be limited by using latin hypercube sampling.

Sensitivity analysis is defined here as uncertainty quantification specific to inputs. Sampling as above is vital to sensitivity analysis. Local sensitivity analysis is specific to certain trial points and global SA tries to better capture the effects of relationships between inputs. Local sensitivity often is described with partial derivatives, in which case it may be useful in differentiable models.

Though these strategies are less than ideal, they can contribute toward answering a given research question. When required to provide an assessment of some problem, one has various constraints, and these types of analyses may be the best possible given time and resources available. Ideally, these statistical analyses will give insight into the real system. But much of this analysis directly compares the real system and the simulation model results. In the section below, we will discuss stability results and structure to function results in the GDS model. Stability results allow us to be confident in small errors in our choice of parameters, while structure to function results allow us to quickly test our simulation model in a computationally tractable way. The following section discusses these ideas in further depth.

1.2 Stability results

A strong understanding of the underlying mathematical model is vital to have confidence in any sensitivity analysis of the simulation model. There are also analyses that can be done on the mathematical model which give results which lend toward verification. Our focus will be stability and structure-to-function results in GDS.

There are different types of stability - state stability, and structural stability. State stability is the first type of stability most people are exposed to - it is whether small state perturbations will result in the same state in the long run. This is defined as follows:
We consider omega limit sets in dynamical systems, developed as follows:

Definition 9. *For some phase space Γ, and state $x \in \Gamma$, an accumulation point of x is a limit of some subsequence of the forward orbit of x.*

Definition 10. *For some function F and a state x, its ω-limit set, $\omega_F(x)$ is the set of accumulation points of x.*

Structural stability is the stability of the dynamics to aspects of the system - this means that perturbing the GDS "slightly" will still result in similar dynamics. More precisely, structural stability as defined in [9] is developed as follows:

We first define a homeomorphism:

Definition 11. *For topological spaces, A and B, a function $h : A \to B$ is a homeomorphism if it is continuously invertible.*

We then define topological conjugation:

Definition 12. *Let $f : A \to A$ and $g : B \to B$ be two maps. f and g are said to be topologically conjugate if there exists a homeomorphism $h : A \to B$ such that, $h \circ f = g \circ h$.*

We can then define structural stability by using a distance metric in the function space, $d(\cdot, \cdot)$:

Definition 13. *Let $f : J \to J$. f is said to be structurally stable on J if there exists $\varepsilon > 0$ such that whenever $d(f, g) < \varepsilon$ for $g : J \to J$, it follows that f is topologically conjugate to g.*

We see that the definition of topologically conjugate speaks to continuity through the homeomorphism. These definitions, therefore, can apply easily to continuous dynamical systems; but with discrete dynamical systems, it may be less directly applicable. To be able to use these definitions in a meaningful way, one can use GDS in which certain aspects are infinite.

Although we will be mostly dealing with GDS where all aspects (elements of the GDS triple) are finite, two aspects can be made infinite - state space and the underlying graph. One benefit of working with sensitivity in an infinite state space is that tools from continuous analysis can be utilized.

Devaney [9] and Lind [23] include a distance metric which reveals how to make states arbitrarily close in an infinite discrete case. First we define a state space which is Boolean at each vertex, $K = \{0, 1\}$. Our graph then is infinite, we call the resulting state space $K^\infty = K \times K \times K \times \cdots$.

Now we define a distance metric between states $x = (x_0, x_1, ...)$ and $y = (y_0, y_1, ...)$:

$$d(x, y) = \sum_{i=0}^{\infty} \frac{|x_i - y_i|}{2^i} \ .$$

Once we have this metric, we can assess stability through our previous definitions. This distance is useful because the distance can be arbitrarily small, and thus a small ε could

then be used in the structural stability definition to discover a topological conjugate.

Lind [23] and Devaney [9] study a shift map, $\sigma((x_0, x_1, ...)) = (x_1, x_2, ...)$. They are able to show that this map is continuous with the distance defined above. Devaney then shows that this map is topologically conjugate to the quadratic map, $F_\mu(x) = \mu x(1-x)$ for $\mu > 2+\sqrt{5}$, to consider the structural stability of this quadratic map.

Lind takes the above distance and considers its application to GDS functions with the graph as $\mathbb{Z}$, i.e. an infinite one-dimensional lattice. He shows that a local function in this setting is continuous, and that it commutes with the shift operator. He shows that f_v is a local function in this setting if and only if it commutes with the shift operator and is continuous over the state space. He is then able to use this to discuss state stability.

Lind also investigates the number of 1's in a state, or other patterns in the limit. With the above distance, he is able to connect sums over the system state with integrals. This ability to study limiting behavior is powerful, but rare for cases in which the state space is finite, because a defined distance in the finite case cannot be arbitrarily close.

In a finite state space, we could define our topology so that everything is an open set to give us continuity, but this wouldn't lead to any interesting results because we have not added any structure to our dynamical system with this definition, and any distance will still have a discrete co-domain.

When possible, usefully employed continuity can be beneficial because continuous systems have a lot of theory in sensitivity, while discrete systems are relatively sparse. One reason for this discrepancy is that a well known sensitivity measure of continuous systems does not easily translate to discrete systems - the derivative. Now, tools such as the derivative were developed for continuous systems in order to conduct the type of analysis that we have seen over the past hundreds of years. Since GDS and similar discrete dynamical systems have only been rigorously studied for around 50 years, it is important to similarly develop these tools - as Newton developed the derivative, which drove Calculus.

To begin our discussion of sensitivity ideas in finite GDS, we will consider a metric that acts similarly to a derivative. To understand this metric we must dig deeper into distance between states in finite GDS.

There are multiple distances which we can define between states in a finite state space. We could define distance as the minimum length of the directed path in the phase space. Hamming distance, which is the number of states which are different, is one type of distance

which is used by Kauffman [17], defined as follows:

Definition 14. *For system states $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$, we define the Hamming distance, $d(\cdot,\cdot) : K^n \times K^n \to$ as $d(x, y) = \sum_{i=1}^{n} \mathbb{I}(x_i \neq y_i)$.*

This can be useful in state stability. Another type of distance that can be useful is a Boolean distance which is 0 if states are the same, and 1 otherwise. A concept of derivative has been conceived with this distance, activity [1].

1.2.1 State stability

In GDS, to assess state stability, we consider measures on the phase space and their properties under perturbation to states. This differs from structural stability because the GDS function is not perturbed - we only consider one phase space. We can consider state stability in the short term by analyzing $||F(x) - x||$ for a GDS (F, G, K^n), a state $x \in K^n$, and some norm: $|| \cdot || : K^n \to \mathbb{Z}$. Here $\mathbb{I}$ is the indicator function

We will first discuss activity, which applies to the short term, and applies to a Boolean state space.

Activity

In this section, our state space will be $K^n = \{0, 1\}^n$. Activity is the probability that a change in a vertex state will cause a change in the image under the GDS function. Formally, it is developed as follows.

For $v \in V(G)$, we define the vth unit vector e_v as:

$$e_v = (0, 0, 0, \ldots, 0, 1, 0, \ldots, 0)$$

so that the state of vertex v is 1 and the state of every other vertex is 0. We then define a function on a specific vertex and system state:

$$\alpha_{F,v}(x) = \mathbb{I}[F(x + e_v) \neq F(x)] \ .$$

Here addition is modulo 2 as we are in the Boolean state space. This function is a partial derivative. As a derivative gives rate of change of the function with respect to a variable, our function also considers the local change of a function with respect to a change in one variable of the state, using the Boolean distance described above:

$$\alpha_{F,v}(x) = \frac{\partial F}{\partial x_v}(x) \ .$$

We then take the expected value over all states, x. It is useful to take a probabilistic approach because of the combinatorial nature of our problem. We would like to consider each state in which the only difference is in x_v. So our sample space is the state space. We set the probability of choosing any state equal, and our random variable is $\alpha_{F,v}(\cdot)$. The *activity of F with respect to v* is this value:

$$\bar{\alpha}_{F,v} = \mathbb{E}[\alpha_{F,v}] \ .$$

This can be viewed as the sensitivity of the GDS function with respect to initial vertex state x_v. Next we look at the overall sensitivity and activity. The sensitivity of F is defined:

$$\bar{\alpha} = \sum_{i=1}^{n} \bar{\alpha}_{F,i}/n \ .$$

The activity of F is defined as the vector:

$$\bar{\alpha}_F = (\bar{\alpha}_{F,1}, \bar{\alpha}_{F,2}, ..., \bar{\alpha}_{F,n}) \ .$$

The sensitivity allows us to rank vertices on "importance" of their state to the function, i.e. how often a change in each vertex will result in a change to the image. These values are computationally intractable to compute. The results in [1] give activity values for specific graphs including lattices, d-regular trees, and random graphs. These activity values are calculated applying various GDS functions.

A downside of activity (as with a derivative) is that it is short term. A longer term activity concept mentioned in [1] is k-step activity. It has the same type of definition as activity, but with the function F replaced by F^k for some $k \in \mathbb{Z}$:

$$\alpha_{F^k,v}(x) = \mathbb{I}[F^k(x + e_v) \neq F^k(x)] \ ,$$

$$\alpha_{F^k,v} = \mathbb{E}[\alpha_{F^k,v}(x)] \ .$$

This considers whether states which differ in the ith spot will stay different after k steps. Another long-term idea is to consider state stability of limit cycles.

Attractor stability

Though we will always eventually have periodic orbits in a GDS with a finite state space, there are a variety of possibilities in the structure and stability of these limits.

Ribeiro and Kauffman [30], as well as Kuhlman and Mortveit [19], consider state stability of limit cycles. This research has biological applications in the classification of cell types. Since genes are being turned on and off rapidly, and cell states are likewise updated rapidly, the focus must be on state stability of limit cycles. To study these ideas, they construct an attractor graph, defined as follows:

Definition 15. *For some function F, an attractor, $A \subset \Gamma(F)$, is a subset of the system state space, K^n, such that $F(A) = A$:*

Definition 16. *For some function F, and state space K^n, the attractor graph of F is a graph $G = (V, E)$ where $V = \{A_i\}_i$ is the set of attractors, and there exists an edge $e = (A_i, A_j) \in E$ if for some $x \in A_i$, we have $\omega_F(x + e_k) = A_j$ for some $1 \leq k \leq n$].*

In other words, this is a graph where the vertices are attractors in the phase space and there exists a directed edge if a state in one attractor can be perturbed at a single vertex to become a state in the basin of another attractor. These graphs may result in strongly connected components. These strongly connected components in the attractor graph are called ergodic sets.

Ribeiro and Kauffman wanted to find GDS in which there were multiple ergodic sets. This could suggest that cell type is driven by a system similar to GDS, and that the different cell types could correspond to different ergodic sets. They found that multiple ergodic sets was indeed possible, but only after constraining the noise to specific vertices.

In the results of Kuhlman and Mortveit [19], they show that there exist GDS which lead to an arbitrarily large number of ergodic sets. One of the main theorems is as follows:

Theorem 17. *There exist permutation block sequential bi-threshold GDS maps F for which strongly connected components of the attractor graph $G(A)$ have binary hyper-cubes Q_2^n as sub-graphs for any integer $n \geq 1$.*

Here Q_2^n is a graph in which vertices are binary tuples, there are n vertices, and an edge exists if the hamming distance between vertices is 1.

This shows that the attractors can be sensitive to vertex changes. Other results in this paper give conditions for which attractor graphs contain only singletons. These results show that there can be multiple ergodic sets, even when the dynamics of switching is not restricted. This account of the range of possible ergodic set structure gives credence to the idea of Ribeiro and Kauffman that ergodic sets could drive cell type.

1.2.2 Structural stability with respect to a measure

In the following, we will apply the idea of structural stability to a measure of our phase space. The definition below speaks to a similarity in phase space between two nearby functions:

Definition 18. *A function f in a family of functions F is called structurally stable with respect to μ if for any $g \in F$, there exists $\varepsilon > 0$ such that $|f - g| < \varepsilon$ then we have $\mu(\Gamma(f)) = \mu(\Gamma(g))$ for some measure μ.*

In the following we may consider structural stability with respect to some measure of a GDS function on a restricted function space. This restriction could correspond to a specific local function form, a specific type of graph, or a specific range of possible update schemes. We first consider a restriction of perturbations on the function space to confine functions of a particular form.

Function perturbations

There are results which discuss stability of limit cycle length under perturbations in a restricted function space. Thus the structure of the function is constant, and perturbations are applied to function parameters. Threshold, bi-threshold, dynamic threshold, block threshold have been studied extensively in this regard, and will be discussed in the following [20,29,32]. To characterize the length of limit cycles, potential functions can be used. We define a potential function as a function of this form, $P : K^n \to \mathbb{R}_{\geq 0}$ where K^n is the state space:

Definition 19. *A function $P : K^n \to \mathbb{R}$ is called a potential function of F if $P(x) \geq 0$ for all $x \in K^n$, and $P(F(x)) - P(x) < 0$ for all $x \in K^n$ such that $x \neq F(x)$.*

It is often used to find cases in which it will continue to decrease and therefore create a contradiction to its positivity. These functions are similar to Lyapunov functions in the continuous setting.

Results on GDS limit cycle structure have been developed using this framework. We first look at generalized threshold functions defined as follows:

Definition 20. *For vertex i and neighborhood $N(i)$ with corresponding states $x[i]$, a generalized threshold function on i is defined:*

$$f_i(x[i]) = \begin{cases} 1 & \sum_{j \in x[i]} a_{ij} x_j \geq k_i \\ 0 & \sum_{j \in x[i]} a_{ij} x_j < k_i \end{cases},$$

where x_i is the state of vertex i, $k_i \in \{0, 1, ..., d(i) + 2\}$, and $a_{ij} \in \mathbb{R}$ is a weight assigned to the edge (i, j).

Barret et al. [7] characterized limit cycle lengths for asynchronous threshold functions. Goles and Olivos (1981) were the first to characterize limit cycle lengths for synchronous threshold functions. They are cited in Kuhlman et al. [20], who then extended the ideas to more general threshold functions. The major result of Goles and Olivos shows that the length of the limit cycles under synchronous threshold functions has an upper bound of 2. We define our GDS function as $F : K^n \to K^n$, updated synchronously, with a Boolean state space, $K^n = \{0, 1\}$. We can now see the theorem:

Theorem 21. *For our GDS function, F, and for all $x = (x_1, ..., x_2) \in K^n$, there exists $s \in \mathbb{N}$ such that $F^{s+2}(x) = F^s(x)$.*

They proved this by showing a contradiction for long limit cycles in their effect on a potential function. This is a very technical argument and will not be discussed here.

Kuhlman et al. [20] extends these results to a bi-threshold case, as well as considering the case of sequential updating. Bi-threshold functions are generalizations of threshold functions as follows:

Definition 22. *A bi-threshold function for a single vertex, i, is:*

$$f_i(x[i]) = \begin{cases} 1 & \text{if } x_i = 0 \text{ and } \sum_{j \in x[i]} a_{ij} x_j \geq k_i^{\uparrow} \\ 0 & \text{if } x_i = 1 \text{ and } \sum_{j \in x[i]} a_{ij} x_j < k_i^{\downarrow} \\ x_i & \text{otherwise} \end{cases},$$

where $k_i^{\downarrow}, k_i^{\uparrow} \in \{0, 1, ..., d(i) + 2\}$, and a_{ij} are the elements of the dependency matrix. Their work reveals a bifurcation in the sensitivity of max limit cycle length. And thus there is a sensitivity of the limit cycle length to the function parameters. One main result in [20] extends Goles and Olivos' proof technique to generalize their result:

Theorem 23. *A synchronous bi-threshold GDS may only have fixed points and 2-cycles as limit sets.*

This gives stability of maximum length of limit cycle with respect to function parameter changes. The second main result in [20] gives a bifurcation of the max limit cycle for the asynchronous case.

Theorem 24. *Let G be a graph, let $\pi \in S_G$, and let $(f_v)_v$ be bi-threshold functions all satisfying $k_v^{\downarrow} - k_v^{\uparrow} \leq 1$. The sequential dynamical system map F_π only has fixed points as limit sets.*

This theorem is again proved using a potential function argument, but with a different function than above. We will call this the potential function, and it takes into consideration threshold values as follows:

For a given vertex v, and system state s, the vertex potential is defined:

$$P(v,s) = \begin{cases} k_v^{\downarrow} & x_v = 1 \\ d(v) + 2 - k_v^{\uparrow} & x_v = 0 \end{cases} .$$

This vertex potential is the minimum number of neighboring states which are equal to x_v that would cause x_v to remain constant - i.e. if $x_v = 1$ and there are $k_v^{\downarrow}$ neighboring states equal to 1, after an update, we will get $x_v = 1$, but if there are $k_v^{\downarrow} - 1$ neighboring states equal to 1, we will get $x_v = 0$.

The edge potential is then defined as follows:

$$P(e = (v, v'), s) = \begin{cases} 1 & x_v \neq x_{v'} \\ 0 & x_v = x_{v'} \end{cases} .$$

Now the system potential, for an underlying graph G and a given system state s, is the sum of the vertex and edge potentials:

$$P(x) = \sum_{v \in V(G)} P(v,x) + \sum_{e \in E(G)} P(e,x) .$$

Note that the system potential will always be positive. Now we will consider the change in this system potential. Our GDS is (F_π, G, K^n) for $\pi \in S_G$. Our sequential function is

$$F_\pi = F_{\pi(n)} \circ F_{\pi(n-1)} \circ \cdots \circ F_{\pi(1)} .$$

We consider change in potential between $F_{\pi(v-1)} \circ F_{\pi(v-2)} \circ \cdots \circ F_{\pi(1)}(x) = y$ for some $x \in K^n$ and $F_{\pi(v)}(y)$. We call this application of $F_{\pi(v)}$ a *vertex update*. Considering the sequential aspect of our GDS, the above composition allows simply adding the changes in potential on a vertex update:

$$\Delta_v(P(x)) = P(F_{\pi(v)}(x)) - P(x) ,$$

to give us our difference of potential when applying F_π: $P(F_\pi(x)) - P(x)$. Thus we have: if each $\Delta_v(P(x)) < 0$, then $P(F_\pi(x)) - P(x) < 0$. It is clear that on a vertex update, if the vertex state remains constant, the system potential remains constant. Thus we consider a change in the vertex state. In [20] the authors derive that on a vertex state change, the potential change is $\Delta_v(P(x)) \leq k_v^{\downarrow} - k_v^{\uparrow} - 2$. Thus if $k_v^{\downarrow} - k_v^{\uparrow} \leq 1$, we have that the change in system potential is negative on a vertex update. And thus any limit cycle of length greater

than 1 would result in a negative potential function, and thus a contradiction. Therefore we have the result.

Having found useful potential functions to use for synchronous and asynchronous threshold functions, attempts were made to extend these ideas to more generalizations of the threshold function. In [32], Wu et al. consider a transition-based dynamic threshold system, in which for a given vertex, the threshold may change on a change in state. It is developed as follows:

Definition 25. *We define $T_v^{\uparrow}$ and $T_v^{\downarrow}$ as the sets of permissible up and down thresholds, respectively. We define these as*

$$T_v^{\uparrow} = \{a_v, a_{v+1}, ..., b_v\} \text{ and } T_v^{\downarrow} = \{a'_v, a'_v + 1, ..., b'_v\} \ ,$$

where $a_v \geq 0$, $a'_v \geq 1$, $b_v \leq d(v) + 1$, and $b'_v \leq d(v) + 2$.

The vertex state space includes the current threshold values, and the current transition index as follows:

Definition 26. *We will define $M_v = \{0, 1\} \times T_v^{\uparrow} \times T_v^{\downarrow} \times T$ and we will also define $M_v^* = \{0, 1\}^{d(v)+1} \times T_v^{\uparrow} \times T_v^{\downarrow} \times T$.*

where T is the space of time/transition indices. We can then define functions that capture the transitions of threshold values and time as: $g_v : M_v^* \to T_v^{\uparrow} \times T_v^{\downarrow}$ and $h_v : M_v^* \to T$, respectively. We can now define our vertex functions:

Definition 27. *A transition based dynamic threshold function has vertex functions:*

$$f_v(x[v]) = \left(\theta_{v,k_v^{\uparrow},k_v^{\downarrow}}(x[v]), g_v(x[v], k_v^{\uparrow}, k_v^{\downarrow}, \tau_v), h_v(x[v], k_v^{\uparrow}, k_v^{\downarrow}, \tau_v)\right) \ .$$

This can apply to certain disease models in which one can contract a disease multiple times, but the probability of contraction changes with each instance. It is a step toward analysis of GDS where the functions directly depend on time. The main result for the synchronous update case is:

Theorem 28. *Let F be a synchronous, transition-based dynamic threshold function. Then the projection of every periodic orbit of F onto its $x = (x_1, ..., x_n)$-component has period at most two. In particular, if $k_v^{\uparrow}$ or $k_v^{\downarrow}$ has period $p > 2$ on a periodic orbit, then the x-component on this orbit is fixed.*

The major result for the asynchronous case is as follows:

Theorem 29. *Consider a transition-based dynamic SDS map. If there exists a τ_{max} such that for all vertices v and $\tau \geq \tau_{max}, \Delta_v(\tau) = k_v^{\downarrow}(\tau) - k_v^{\uparrow}(\tau) < 2$, then, it has only fixed points as limit sets and from any initial configuration, a fixed point is reached in at most $(n\tau_{max} + 1)(3m + 2n + 1)$ time steps where m and n are the number of edges and vertices in the underlying graph X respectively.*

These two theorems are able to apply the same potential functions to this more generalized function to reason about the periodic structure.

We have considered ways of generalizing the threshold function, and considering the case of a parallel update and a sequential update. We can also consider standard threshold functions with a general block update scheme. A block update is defined above - where disjoint subsets of vertices, blocks, are updated in parallel, and there is a sequence of block updates. This type of update mechanism can be powerful because it is a generalization of both the synchronous and asynchronous update schemes - synchronous update corresponds to a block size of $|V(G)|$ for some GDS (F, G, K^n), while asynchronous corresponds to block sizes of 1. The results above show that there is a difference in the periodic structure of a parallel and sequential update of various threshold systems. The following investigates how the limit cycles change as we consider block sizes between synchronous and asynchronous.

Mortveit [29] considers the possible lengths of limit cycles in the block sequential setting. The main result is as follows:

Theorem 30. *Let G be a simple graph and $\mathcal{B}$ a block partition of $V(G)$. If the largest nontrivial block of $\mathcal{B}$ has size at most three, then any block-sequential threshold finite dynamical system with update sequence $\pi \in S_{\mathcal{B}}$ only has fixed points as limit sets. If the largest nontrivial block has size at least four, then a block-sequential threshold finite dynamical system may have limit cycles of length at least two.*

This proof similarly uses a potential function argument. This proof is different since for the block sequential case one cannot simply track the change in a vertex state over the period - because there are some parallel blocks. Thus the potential function changes must be calculated for possible changes to each block of vertices.

In [29], Mortveit conjectured that block size of at least four have limit cycles of length at most two. This was disproved by Goles and Montealegre [12] - where they showed that there exist block-sequential threshold GDS with large limit cycles, suggesting some deeper dependence of limiting behavior on update scheme.

The results above and the method used to find them can be extended to more general cases if graph structure is taken into account. Adiga et al. [3] found a way to extend the potential function argument above to show some limit cycle stability with respect to a block-sequential update structure of any size if there is a connection with the underlying graph:

Theorem 31. *Let G be a simple graph with vertex set $V(G)$ and edge set $E(G)$. Let $\mathcal{B}$ be a block partition of $V(G)$. If every block $B \in \mathcal{B}$ satisfies Condition 32 below, then, any block-sequential standard threshold system induced by $\mathcal{B}$ for any update order on the blocks has only fixed points as limit sets. Also, the transient length is at most $(|E(G)| + |V(G)| + 1)/2$.*

Condition 32. *For any non-empty $B' \subseteq B$ and any assignment x of vertex states for B', $||B'||-2|\Lambda_{B'}(x)| - |B'| < 0$, where $||B'||$ is the number of edges in the subgraph induced by B' and:*

$$\Lambda_{B'}(x) = \{\{u, v\} \in E(X) \mid u, v \in B', \text{ and } x_u = x_v\} .$$

This result gives more information about the structure of limit cycles with ranges of block sequential updates. It could be seen as stability of limit cycle length with respect to correlated graph and update perturbations.

Graph perturbations

Adiga et al. [2] considered stability under graph perturbations. They modeled a disease spread over a network, and the measure on which to assess stability was the number of infections. In the Boolean network, this measure can be defined as mass:

Definition 33. *For state space $K^n = \{0, 1\}^n$, the mass of a state $x = (x_1, x_2, ..., x_n) \in K^n$ is defined as $m(x) = \sum_{i=1}^{n} x_i$.*

Mass can be a useful concept in Boolean networks as it will give a count of the number of vertices affected. Adiga et al. [2] considered the structural stability of the mass of GDS functions under network perturbations. Perturbing the graph can be applicable as it will reflect some graph noise or uncertainty that can arise in simulation models. The result could also be applied to a system in which the number of edges in a graph slowly grows.

Graph perturbations are also function perturbations, but can be more difficult to apply because there must be some mechanism to change the function accordingly. Therefore, this paper considered threshold models as above, a dynamical system in which the function is closely aligned with the graph properties. To analyze random graph perturbations, we must first define the setting in which we will work:

We define a sample space and a weight to this space. The space which we will use is $\Omega_1 = G_{n,p}$ in which Ω_1 is the set of all subgraphs of K_n, and p relates to the probability of each edge appearing. This gives, for a chosen graph $G \in \Omega_1$, the associated weight

$$P_1(G) = p^{|E(G)|}(1-p)^{|E(K^n)|-|E(G)|}.$$

We note that $|E(K^n)| = \binom{n}{2}$. We now consider graph perturbations in a uniform edge addition model. This involves selecting a graph from our sample space, and adding edges:

Definition 34. *A uniform edge addition model is developed as follows: We choose $R_u(\epsilon) \in \Omega_1$, with probability $P_1\big(R_u(\epsilon)\big)$ where $p = \epsilon/n$. We then take a given graph G, and consider the perturbed graph as the edge union: $G' = \big(V(G), E(G) \cup E(R_u(\epsilon)\big)$.*

Adiga et al. [2] considered threshold models for the GDS function. The vertices are updated sequentially, and thus given an initial state, we know from previously stated results that we will reach an equilibrium point. This will be the point at which to measure the system mass. Adiga et al. [2] also consider a random seed set, in which vertices are selected as 1 at random, as this size will have an effect on the stability. Thus, given a graph G, and given seed size, we define a sample space $\Omega_2 = \{x \in K^n \mid m(x) = s\}$ and the weight of the elements of this space is uniform: $P_2(x) = 1/\binom{n}{s}$.

Thus our final sample space is $\Omega = \Omega_1 \times \Omega_2$. Then the weight for a given graph G and initial state x over this space is $P(G, x) = P_1(G) \cdot P_2(x)$. We let the random variable, X, be the mass of the corresponding equilibrium point.

The major theorem for the linear threshold case shows that the number of infections in these threshold GDS models is stable with respect to graph perturbations:

Theorem 35. *Let $G(V, E)$ be a graph with maximum degree Δ and $G+R_u(\varepsilon)$ be the perturbed graph obtained by adding edges uniformly at random with factor ε. For a given threshold GDS, given an initial random seed set of size s, the expected number of infected vertices in the perturbed graph $G' = \big(V(G), E(G) \cup E(R_u(\epsilon))\big)$ is*

$$\mathbb{E}(X_{G'}) = O(s(\Delta + \varepsilon + \log n) \log n) \ ,$$

whereas the expected number of infected vertices in the unperturbed graph is

$$\mathbb{E}(X_G) = O(s\Delta \log n) \ .$$

This theorem quantitatively reflects the stability of mass with respect to graph perturbations, as we consider the expected behavior of a graph G and nearby graphs. Thus when implementing a simulation model which reflects a threshold model, we can be confident that uncertainty in graph structure will not have a disproportionate effect on mass of resulting equilibrium.

Update order

We have discussed various GDS update schemes in the previous sections. In this section we will consider Sequential Dynamical Systems (SDS), GDS with asynchronous update. One main area of study in SDS is stability with respect to perturbations of update sequence.

In the above sections, when considering SDS, each vertex appeared precisely once in the update sequence. When this occurs, the update sequence is a *permutation*. For a given graph G, the set of permutations of G is denoted S_G. This can be generalized to any finite

sequence of vertices, without restrictions on the number of appearances of vertices. We call this general update sequence a *word* and denote the set of words as W_G for a given G. To consider stability, we will consider specific types of update sequence perturbations represented by an update graph, defined as follows:

Definition 36. *For some graph G, the update graph $U(G) = (V, E)$ is one in which $V = S_G$ and there exists an edge $e = \{\pi, \pi'\}$ if for some integer k where $1 \leq k \leq n$, we have*

- $\pi' = \big(\pi(1), \pi(2), ..., \pi(k-1), \pi(k+1), \pi(k), \pi(k+2), \pi(k+3), .., \pi(n)\big)$ *and*
- $\{\pi(k), \pi(k+1)\} \notin E(G)$.

Equivalence relations are useful in stability topics, as one can then assess these equivalences under two perturbed systems. We thus state this definition:

Definition 37. *For a set S, we say that the relation $\sim$ is an equivalence relation if the following hold*

- *For all $a \in S$, $a \sim a$,*
- *If $a, b \in S$ and $a \sim b$, then $b \sim a$,*
- *If $a, b, c \in S$, $a \sim b$, and $b \sim c$, then $a \sim c$.*

The components of our update graph above will define an equivalence relation on S_n as follows.

Given a graph G, for $a, b \in S_n$, we define a relation $\sim_{\alpha'}$ by $a \sim_{\alpha'} b$ if $\{a, b\} \in E\big(U(G)\big)$. We then take the transitive closure and reflexive closure to get an equivalence relation $\sim_\alpha$, and we call this α-equivalence. We use this structure to reason about the function dynamics.

Proposition 38. *Given a graph G, local functions $(F_i)_i$, and $\pi, \pi' \in S_n$, we have*

$$\pi \sim_\alpha \pi' \implies F_\pi = F_{\pi'} \ .$$

Thus we use an equivalence relation on our permutations to assess equality of the corresponding SDS functions. We can develop a simple test of this equivalence class as follows.

We consider the equivalence classes under our relation, $S_G/\sim_\alpha = \{[\pi]_\alpha \mid \pi \in S_G\}$. These can be represented as acyclic orientations as follows. We first define an ordering relation on permutations π:

Definition 39. *For a graph G and a given permutation π, we write $i <_\pi j$ if i appears before j in π.*

For a given G, we define $O : S_G \to \mathrm{Acyc}(G)$ with the rule $O(\pi) = (V(G), E)$ where E is the orientation of G such that $(i, j) \in E$ if $\{i, j\} \in E(G)$ and $i <_\pi j$. We write O_π for convenience.

Once we have defined this O_π, we can state the following theorem:

Theorem 40. *For given graph G and $\pi, \pi' \in S_G$, we have $O_\pi = O_{\pi'} \iff \pi \sim_\alpha \pi'$.*

Thus we can simply direct our graph as prescribed for O_π, and if equal, we have $F_\pi = F_{\pi'}$. Note that the converse of the previous theorem does not necessarily hold- if $O_\pi \neq O_{\pi'}$, this theorem will give no information about the function. Thus we investigate dynamics of the function through more general equivalence relations.

When there exists a graph isomorphism between two finite dynamical system phase spaces, we call them *dynamically equivalent*:

Definition 41. *Two finite dynamical systems with maps $\phi, \theta : K^n \to K^n$ are dynamically equivalent if there exists a bijection $h : K^n \to K^n$ such that $\phi \circ h = h \circ \theta$.*

Clearly functional equivalence implies dynamical equivalence, but there are other conditions between GDS which give dynamical equivalence. To construct these conditions, we first provide definitions. We define $\mathrm{Aut}(G)$ as the set of all graph automorphisms of G. We now define $\mathrm{Aut}(G)$-invariance:

Definition 42. *A sequence of vertex functions, $(f_v)_v$, is $\mathrm{Aut}(G)$-invariant if either of the following conditions hold:*

- *$f_v = f_{\gamma(v)}$ for all $\gamma \in \mathrm{Aut}(G)$;*
- *$\gamma \circ F_v \circ \gamma^{-1} = F_{\gamma(v)}$.*

Definition 43. *For a graph G we let $\alpha(G) = |\mathrm{Acyc}(G)|$.*

With these definitions, we come to the first theorem in [28]:

Theorem 44. *For any sequence $(f_v)_v$ of Aut-invariant vertex functions, the maps F_π and $F_{\gamma\pi}$ are dynamically equivalent.*

We can also bound the number of dynamical equivalence classes, to give the result: $\bar{\alpha}(G)$ is an upper bound for the number of such maps, up to dynamical equivalence, constructed as follows

Definition 45. *For a given graph G and automorphism γ, we define the operation $\langle\gamma\rangle\backslash G$ as the orbit graph of G and the cyclic group $\langle\gamma\rangle$. An orbit graph is a multi graph where the vertices resp. edges are the orbits of γ acting on $V(G)$ resp. $E(G)$.*

With this definition, we can construct an equation which bounds the number of dynamical equivalence classes:

$$\bar{\alpha}(G) = \frac{1}{|\mathrm{Aut}(G)|} \sum_{\gamma \in \mathrm{Aut}(G)} |\mathrm{Fix}(\gamma)| = \frac{1}{|\mathrm{Aut}(G)|} \sum_{\gamma \in \mathrm{Aut}(G)} \alpha(\langle\gamma\rangle\backslash G) \ .$$

We will now consider only certain permutation perturbations and how they relate to the limit cycles of the SDS phase space. We define a *click* to an acyclic orientation as the conversion of a source to a sink. This means that for some source, i, for any edge (i, j), we re-orient this edge (j, i). We consider a relation, $\sim_{\kappa'}$, between acyclic orientations O, O' as $O \sim_{\kappa'} O'$ if O becomes O' under a single click. We take the transitive and reflexive closures of this relation to get κ-equivalence, $\sim_\kappa$. Thus $O \sim_\kappa O'$ if some number of clicks of O gives you O'. This will relate to the SDS map in its cycles. We therefore define a stability measure, cycle equivalence:

Definition 46. *Two finite dynamical systems with maps $\phi : K_1^n \to K_1^n$ and $\theta : K_2^n \to K_2^n$ are cycle equivalent if there exists a bijection $h : Per(\theta) \to Per(\phi)$ such that*

$$\phi|_{Per(\phi)} \circ h = h \circ \theta|_{Per(\theta)} \ .$$

In Macauley and Mortveit [28], they relate the limit cycle structure to this equivalence class:

Theorem 47. *For a graph G, and for permutations π, and π', if $O_\pi \sim_\kappa O_{\pi'}$ then F_π and $F_{\pi'}$ are cycle equivalent.*

This equivalence relates to update sequence by the one-to-one correspondence between acyclic orientations and α-equivalence classes. Thus, given an acyclic orientation, we choose a representative $\pi = \big(\pi(1), \pi(2), ..., \pi(n)\big)$ from the respective α-equivalence class - and a click to the acyclic orientation can be defined in terms of our permutation:

$$\sigma_s(\pi) = \big(\pi(s+1), \pi(s+2), ..., \pi(n), \pi(1), \pi(2), ..., \pi(s)\big) \ .$$

With this, we have related this action on $\mathrm{Acyc}(G)$ to a perturbation of the update sequence, and we can thus state the following:

Theorem 48. *For a graph G, for any $\pi \in S_G$, the SDS maps F_π and $F_{\sigma_s(\pi)}$ are cycle equivalent.*

Macauley [27] generalized this to words:

Theorem 49. *For a graph G, for any $w \in W_G$, the SDS maps F_w and $F_{\sigma_s(w)}$ are cycle equivalent.*

As the previous result could be tested through O_π, this can also be tested. If given a phase space, we can test cycle equivalence through comparing the multi-sets of cycle lengths. But we may not be given a phase space, and constructing a phase space can be computationally expensive/intractable. To construct our more efficient test, we must first provide some definitions.

Definition 50. *A cycle basis $\mathcal{C}$ for a graph, G, is a minimal set of cycles such that:*

$$G - \sum_{C \in \mathcal{C}} C = (V(G), E(G) - \sum_{C \in \mathcal{C}} e) \text{ contains no cycles, where } e \in E(C).$$

We now define Coleman's ν-function:

Definition 51. *For a given cycle basis, $\mathcal{C} = \{C_1, ..., C_k\}$, and a given acyclic orientation, A, of graph G, we define $\nu_{\mathcal{C}} : \mathrm{Acyc}(G)^k \to \mathbb{Z}$ as:*

$$\nu_{\mathcal{C}}(A)_i = |n_+(i)| - |n_-(i)| \ ,$$

where $n_+(i)$ are the edges in the cycle represented by C_i which have the same orientation in A, and $n_-(i)$ are the edges in the cycle represented by C_i which have the opposite orientation in A.

Macauley and Mortveit [28] use this function on the acyclic orientation to give another structure-to-function theorem:

Theorem 52. *For a given graph G, cycle basis $\mathcal{C}$, and permutations $\pi, \pi' \in S_G$, we have*

$$\nu_{\mathcal{C}}(O_\pi) = \nu_{\mathcal{C}}(O_{\pi'}) \iff \pi \sim_\kappa \pi' \ .$$

This theorem allows us to analyze κ-equivalence in a way that is computationally tractable. This type of structure-to-function result can be useful in analyzing large simulation models and their underlying mathematical models.

Another aspect of stability which we can consider is stability of the set of periodic points of an SDS with respect to update order perturbations. We develop these ideas as follows:

Definition 53. *For a given graph G and vertex functions for each $v \in V(G)$, if we have $\mathrm{Per}(F_\pi) = \mathrm{Per}(F_{\pi'})$ for any $\pi, \pi' \in S_G$ resp. $\mathrm{Per}(F_\omega) = \mathrm{Per}(F_{\omega'})$ for any $\omega, \omega' \in W_G$ then we say that for graph G, the set of local functions, $(F_v)_v$, is π-independent resp. ω-independent.*

Definition 54. *Given a graph, G, the local functions $(f_v)_v$ are block invariant if* $\mathrm{Per}(F_{\mathcal{B}}) = \mathrm{Per}(F_{\mathcal{B}'})$ *for any block sequential functions induced by $(f_v)_v$ and block partitions $\mathcal{B}$, and $\mathcal{B}'$ as well as any sequences from $S_{\mathcal{B}}$ and $S_{\mathcal{B}'}$ respectively.*

There are results related to functional and cycle equivalence of a specific type of SDS, ECAs [24–26]. We have defined automata networks above. Cellular automata are Boolean GDS's where the graph is a regular grid. Elementary cellular automata are Boolean GDS in which the graph is an n-circle and the update is parallel. Each local function is also equal. Wolfram was one of the pioneers in these systems [31], and thus the local functions, when discussing cellular automata are often called *Wolfram rules* or simply *rules*. Since the graph is a circle, each local function will have 3 inputs. Thus there are $2^{2^3} = 256$ possible rules.

Asynchronous cellular automata are ECAs in which the update order is sequential. There are results that characterize update sequence stability of the rules in ACA.

- There are 104 rules that are permutation independent, 86 of which are word independent [24, 25].
- There are 41 rules that are block invariant [26].

These results are examples of the differences between word, permutation, and block invariance.

Chapter 2

Contributions

2.1 Phase space invariance for asynchronous Boolean networks under shifts of update sequence

2.1.1 Introduction

In the previous sections we have shown that kappa equivalence implies cycle equivalence, it can also imply stability of portions of transient structure. In the Appendix A, we describe the mechanisms of this, which will be summarized here. A driver of Theorem 48 above was the graph morphism in the phase space $F_{\pi(1)} \circ F_\pi = F_{\sigma_1(\pi)} \circ F_{\pi(1)}$. This is an isomorphism when restricted to periodic points, but can still be useful when the transient paths are considered. We include the following definitions in this paper in order to describe this behavior.

Definition 55. *For a given GDS function, F, a state, x, is a garden of Eden state (*GoE*) state if $F^{-1}(x) = \varnothing$.*

Definition 56. *For a given GDS function F, state space K^n, and* GoE *state x, a transient path $P_F(x)$ is the sequence defined as follows:*

$$P_F(x) = (x, F(x), F^2(x), ..., F^{\tau-1}(x))$$

where τ is the smallest number such that $F^\tau(x) \in \mathrm{Per}(x)$. We denote $\ell(P_F(x)) = \tau$.

2.1.2 Results

The main theorem bounds the range and size of the set of maximum transient lengths along update sequence shifts:

Theorem 57. *Assume that* $x = x^{(0)} \in \mathrm{GoE}(F_\pi)$ *with maximal transient path* $P_0 = P_{F_\pi}(x^{(0)})$ *satisfies* (i) $F_\pi(x) \notin \mathrm{Per}(F_\pi)$ *and* (ii) $F_\pi^{-1}\big(F_\pi(x)\big) \subset \mathrm{GoE}(F_\pi)$. *Then* (a) *the states* $x^{(k)} \in K^n$ *defined by*

$$x^{(k)} = F_{\pi_k} \circ \cdots \circ F_{\pi_1}(x) \text{ in } \Gamma(F_{\sigma_k(\pi)}), \quad \text{with} \quad 0 \le k \le n-1 \ , \tag{2.1}$$

are all transient states of their respective phase spaces. Moreover, (b) *any sequence of maximal transient paths* $(P_k)_k$ *with* P_k *containing* $x^{(k)}$ *satisfies the inequality*

$$|\ell(P_k) - \ell(P_0)| \le 1 \ , \tag{2.2}$$

and (c)

$$\big|\{k \mid \ell(P_k) \neq \ell(P_{k+1})\}\big| \le 2 \ . \tag{2.3}$$

Thus, given a GDS with a specific update sequence, we know that much of the transient structure will remain constant. Especially useful is that the maximum transient length may only vary by 1. This may be applicable in implementation, as it gives assurance that update order clicks will not cause a bifurcation in the length of transient paths.

2.1.3 Contributions

In this paper, I helped to construct the condition and the proof of Theorem 6 (see Appendix A). I contributed to writing the transient path definition. I contributed to writing the condition and proof of the Theorem 6. I contributed to the proof of Proposition 2 (see Appendix A). I also edited and helped prepare for publication. See Appendix A for a copy of the manuscript.

2.2 A multi-pathway modeling approach to assess the threat of Tuta absoluta in southeast Asia

2.2.1 Introduction

Above, we have considered threshold functions, and generalizations which depend on time. In Appendix B below, we consider an implementation of a more general periodic GDS. We extended the work of Guimapi [14] and their investigation into the spread of Tuta absoluta. Tuta absoluta is a pest which infects many plants but especially tomato plants. They considered Africa as a focus region, while we consider Southeast Asia. Tuta absoluta can spread through flying - in [14] they considered a radius of 50-75 km per month. Tuta absoluta can also spread through tomato trade. Thus our research extends [14] through including the

long-distance spread. Future work could include a mathematical treatment of this framework of periodic GDS in studying stability of limit cycle length.

2.2.2 Methods

We consider a stochastic periodic model which takes into account three pathways of spread: short-distance, farm-to-market, and market-to-market. The second pathway resembles the spread of Tuta absoluta due to transport from farm to market and from market to village. The third pathway reflects market to market trading of tomatoes. Thus we consider a generalized cellular automata in which edges are added due to being contained in the same region, as well as trade between regions. The periodicity results from the nature of tomato production in these regions being dependent on rainfall. We use an SEI model, in which a cell changes states from susceptible to exposed depending on suitability and infectivity of neighboring cells (along each of the above pathways). See details in Appendix B. An exposed cell then becomes infected after 1 month. This is reasonable as this is the estimated life cycle of the pest, and thus if conditions are suitable enough to become exposed there will be another generation after a month. An infected cell then stays infected as there is currently no successful attempts to remove Tuta absoluta.

2.2.3 Contributions

In this paper, I contributed by collecting data relevant to our model. I did literature reviews, and implemented the model from Guimapi [14]. I implemented many aspects of our mathematical model in the simulation model, including adding the distance edges, and the functions which output probability of infection. I helped to construct statistical designs. I simulated the resulting model and analyzed output. I contributed to writing section 2.2. See Appendix B for selected sections of the manuscript.

Chapter 3

Conclusion

This thesis has explored stability, structure-to-function, validation, and implementation concepts in GDS. These concepts can be applied to simulation models given a research question. Stability results have been given which will help with verification of the simulation model, as well as validation of the mathematical model. Structure-to-function results have been given which help with gaining information about dynamics of a model in a computationally tractable way. Implementation concepts have been explored which help with understanding the dynamics of a given simulation model, uncertainty quantification, as well as verification and validation techniques. This background creates a base on which the contributions could be built.

This thesis documents the contributions made throughout this Master's program. In GDS, transient length was found to be stable under update order shifts. This bound on the transient length under update order shifts can be important to simulation science as a bound on the time needed to reach a limit cycle. This can help with computational tractability in certain cases. It also gives more structure to GDS, and is the first instance of using the generalized phase space to investigate changes in update order. This is an area which may contribute much in future studies.

An implementation of GDS was considered - modeling pest spread in southeast Asia. This contributes to simulation science by adding to a cellular automata with long-distance edges. Many pests and diseases can be spread locally and through trade and travel. Thus, the methods developed in this project can contribute toward improving these spatially multi-scale networked models.

Bibliography

[1] Abhijin Adiga, Hilton Galyean, Chris J. Kuhlman, Michael Levet, Henning S. Mortveit, and Sichao Wu, *Activity in boolean networks*, Natural Computing **16** (2017), no. 3, 427–439.

[2] Abhijin Adiga, Chris Kuhlman, Henning S. Mortveit, and Anil Kumar S. Vullikanti, *Sensitivity of diffusion dynamics to network uncertainty.*, J. Artif. Intell. Res.(JAIR) **51** (2014), 207–226.

[3] Abhijin Adiga, Chris J. Kuhlman, and Henning S. Mortveit, *Effect of graph structure on the limit sets of threshold dynamical systems*, International Workshop on Cellular Automata and Discrete Complex Systems, Springer, 2015, pp. 59–70.

[4] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J. Ramasco, and Alessandro Vespignani, *Multiscale mobility networks and the spatial spreading of infectious diseases*, Proceedings of the National Academy of Sciences **106** (2009), no. 51, 21484–21489.

[5] C. L. Barrett, M. Wolinsky, and M. W. Olesen, *Emergent local control properties in particle hopping traffic simulations*, Traffic and granular flow (1996), 169–173.

[6] Christopher Barrett, Keith Bisset, Shridhar Chandan, Jiangzhuo Chen, Youngyun Chungbaek, Stephen Eubank, Yaman Evrenosoglu, Bryan Lewis, Kristian Lum, Achla Marathe, Madhav Marathe, Henning Mortveit, Nidhi Parikh, Arun Phadke, Jeffrey Reed, Caitlin Rivers, Sudip Saha, Paula Stretz, Samarth Swarup, James Thorp, Anil Vullikanti, and Dawen Xie, *Planning and response in the aftermath of a large crisis: An agent-based informatics framework*, Proceedings of the 2013 Winter Simulation Conference - Simulation: Making Decisions in a Complex World, WSC 2013, 2013.

[7] Christopher L. Barrett, Harry B. Hunt, Madhav V. Marathe, S. S. Ravi, Daniel J. Rosenkrantz, and Richard E. Stearns, *Complexity of reachability problems for finite discrete dynamical systems*, Journal of Computer and System Sciences **72** (2006), no. 8, 1317–1345.

[8] National Research Council et al., *Assessing the reliability of complex models: mathematical and statistical foundations of verification, validation, and uncertainty quantification*, National Academies Press, 2012.

[9] Robert L. Devaney, *An introduction to chaotic dynamical systems*, Westview press, 1989.

[10] Rama Garimella, Berkay Kicanaoglu, and Moncef Gabbouj, *On the dynamics of a recurrent hopfield network*, Proceedings of the International Joint Conference on Neural Networks (2015), 1–8.

[11] Eric Goles and Servet Martinez, *Neural and automata networks: Dynamical behaviour and applications*, Kluwer Academic Publishers, 1990.

[12] Eric Goles and Pedro Montealegre, *Computational complexity of threshold automata networks under different updating schemes*, Theoretical Computer Science **559** (2014), 3–19.

[13] Eric Goles-Chacc, Françoise Fogelman-Soulie, and Didier Pellegrin, *Decreasing energy functions as a tool for studying threshold networks*, Discrete Applied Mathematics **12** (1985), no. 3, 261–277.

[14] Ritter Y. A. Guimapi, Samira A. Mohamed, George O. Okeyo, Frank T. Ndjomatchoua, Sunday Ekesi, and Henri E. Z. Tonnang, *Modeling the risk of invasion and spread of Tuta absoluta in Africa*, Ecological Complexity (2016), 1–17.

[15] J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities.*, Proceedings of the National Academy of Sciences **79** (1982), no. 8, 2554–2558.

[16] A.G. Huff, W.E. Beyeler, N.S. Kelley, and J.A. McNitt, *How resilient is the United States' food system to pandemics?*, Journal of Environmental Studies and Sciences **5** (2015), no. 3.

[17] S.A. Kauffman, *Metabolic stability and epigenesis in randomly constructed genetic nets*, Journal of Theoretical Biology **22** (1969), no. 3, 437–467.

[18] Chris J. Kuhlman, V. S. Anil Kumar, Madhav V. Marathe, Henning S. Mortveit, Samarth Ravi, S. S. Swarup, Gaurav Tuli, S.S. Ravi, Daniel J. Rosenkrantz, and Daniel J. Rosenkrantz, *A general-purpose graph dynamical system modeling framework*, (2011), 296–308.

[19] Chris J. Kuhlman and Henning S. Mortveit, *Attractor stability in nonuniform Boolean networks*, Theoretical Computer Science **559** (2014), 20–33.

[20] Chris J. Kuhlman, David Murrugarra, Henning S. Mortveit, and V. S. Kumar, *Bifurcations in Boolean networks*, Proceedings of Automata 2011 (2011), 29–46.

[21] Reinhard Laubenbacher, David Murrugarra, and Alan Veliz-Cuba, *Structure and dynamics of polynomial dynamical systems*, arXiv preprint arXiv:1108.0209.

[22] Thomas M. Liggett, *Interacting particle systems-an introduction*, School and conference on probability theory, vol. 17, 2002, pp. 307–340.

[23] D. A. Lind, *Applications of ergodic theory and sofic systems to cellular automata*, Physica D: Nonlinear Phenomena **10** (1984), no. 1-2, 36–44.

[24] M. Macauley and H. S. Mortveit, *An atlas of limit set dynamics for asynchronous elementary cellular automata*, Theoretical Computer Science **504** (2013), 26–37.

[25] Matthew Macauley, Jon McCammond, and Henning S. Mortveit, *Order Independence in Asynchronous Cellular Automata*, (2007), 1–18.

[26] _____, *Block Invariance in Elementary Cellular Automata*, Journal of Cellular Automata **3** (2008), no. 1, 37–56.

[27] Matthew Macauley and Henning S. Mortveit, *Cycle equivalence of graph dynamical systems*, Nonlinearity **22** (2009), no. 2, 421–436.

[28] Matthew Macauley and Henning S. Mortveit, *Cycle equivalence of finite dynamical systems containing symmetries*, (2014), 70–82.

[29] Henning S. Mortveit, *Limit cycle structure for block-sequential threshold systems*, International Conference on Cellular Automata, Springer, 2012, pp. 672–678.

[30] Andre S. Ribeiro and Stuart A. Kauffman, *Noisy attractors and ergodic sets in models of gene regulatory networks*, Journal of Theoretical Biology **247** (2007), no. 4, 743–755.

[31] Stephen Wolfram, *Universality and complexity in cellular automata*, Physica D: Nonlinear Phenomena **10** (1984), no. 1-2, 1–35.

[32] Sichao Wu, Abhijin Adiga, and Henning S. Mortveit, *Limit cycle structure for dynamic bi-threshold systems*, Theoretical Computer Science **559** (2014), 34–41.

Chapter 4

Appendices

The page numbers to be noted in the following are in the upper right. Please disregard other page numbers as they are from the copied manuscripts.

4.1 Appendix A

PHASE SPACE INVARIANCE FOR ASYNCHRONOUS BOOLEAN NETWORKS UNDER SHIFTS OF UPDATE SEQUENCE

RICKY X. F. CHEN, JOSEPH A. MCNITT, HENNING S. MORTVEIT†,
RYAN PEDERSON, AND CHRISTIAN M. REIDYS

Abstract. For an asynchronous Boolean network (ABN), how much of its phase space structure is preserved under cyclic shifts of the update sequence, or more generally, under κ-equivalence of permutation update sequences? In this paper we give an answer to this by extending work of Macauley & Mortveit (2009) who showed that the periodic orbit structures of ABNs are completely preserved under κ-equivalence. In particular, their result implies that attractor structure is invariant under cyclic shift of the update sequence. Here we extend their results to describe the parts of the transient structures of the phase space that are preserved under κ-equivalence. Our main result is a Lipschitz-continuity type theorem for sequences of maximal transient paths, thus demonstrating that these phase space structures are essentially preserved under cyclic shifts of update sequence. Moreover, for fixed vertex functions, we prove that the set of maximal transient lengths of induced SDS maps over a κ-class has cardinality at most 2. In our proofs, we also derive insight into the generalized phase space of asynchronous Boolean networks, a much more general object whose structure can provide information about dynamics of for example ABNs with stochastic update mechanisms.

1. Introduction

Asynchronous Boolean Networks (ABN) and the more general framework of graph dynamical systems (GDS), is a precise and useful dynamical system representation capturing distributed, dynamical phenomena with local interactions. Applications range from disease dynamics on social contact graphs [2], to packet transport in wireless networks,

Key words and phrases. Boolean networks; finite dynamical system; asynchronous; transient structure; update sequence; shifts; Lipschitz.

†Corresponding author: Henning S. Mortveit; Telephone: +1 540-231-5327; Fax: +1 540-231-2606.

traffic systems with individual cars [13], and dynamics of biological systems [6, 15]. We will assume that the reader is already familiar with notation and terminology for asynchronous Boolean networks constructed as composed maps $F\colon K^n \longrightarrow K^n$ of the form

$$F_\pi = F_{\pi_n} \circ F_{\pi_{n-1}} \circ \cdots \circ F_{\pi_1} \ ,$$

where $\pi = (\pi_1, \ldots, \pi_n)$ is a permutation of $\{1, 2, \ldots, n\}$, and where each map $F_v\colon K^n \longrightarrow K^n$ may only change the state of its associated vertex x_v while leaving all other states of the *system state* $x = (x_1, \ldots, x_n)$ fixed. If necessary, please review definitions and details in Section 2. It is assumed that K is finite, typically $K = \{0, 1\}$. We remark that this particular class of finite dynamical systems is often referred to as *sequential dynamical systems (SDS)*, see [7, 12], as *automata networks*, see [3–5], and as *polynomial dynamical systems*, see [1].

Let $\sigma_k(\pi) = (\pi_{k+1}, \ldots, \pi_n, \pi_1, \ldots, \pi_k)$ denote the cyclic, left k-shift of π and observe that

$$F_{\sigma_k(\pi)} = F_{\pi_k} \circ F_{\pi_{k-1}} \circ \cdots \circ F_{\pi_1} \circ F_{\pi_n} \circ \cdots \circ F_{\pi_{k+1}} \ .$$

In [9] it was shown that the restricted map

$$F_{\pi_k}\colon \mathrm{Per}(F_{\sigma_{k-1}(\pi)}) \longrightarrow \mathrm{Per}(F_{\sigma_k(\pi)}) \ , \tag{1}$$

where $\mathrm{Per}(F)$ denotes the periodic points of F, induces an isomorphism from the set of periodic orbits of $F_{\sigma_{k-1}(\pi)}$ to the set of periodic orbits of $F_{\sigma_k(\pi)}$ for $1 \leq k \leq n$. This result implies that the periodic orbit structure of F_π is preserved under cyclic shifts of the update order π. The starting point for their proof is the re-bracketing of the expression

$$(F_{\pi_k} \circ F_{\pi_{k-1}} \circ \cdots \circ F_{\pi_1} \circ F_{\pi_n} \circ \cdots \circ F_{\pi_{k+1}}) \circ F_{\pi_k}$$

which leads to

$$F_{\sigma_k(\pi)} \circ F_{\pi_k} = F_{\pi_k} \circ F_{\sigma_{k-1}(\pi)} \ , \quad 1 \leq k \leq n, \tag{2}$$

which in turn gives graph morphisms

$$\gamma_k\colon \Gamma(F_{\sigma_{k-1}(\pi)}) \longrightarrow \Gamma(F_{\sigma_k(\pi)}) \ , \tag{3}$$

where $\Gamma(F)$ denotes the phase space of F, see Section 2. In this paper we use the sequence $(\gamma_k)_k$ to analyze the extent to which the transient portions of the phase space of F_π are preserved under shifts of the update sequence. Note that a GoE state is a state with no pre-image. Our key results can be summarized as follows:

Proposition: States $x, y \in K^n$ can only be identified under γ_k if they have the same immediate successor in $\Gamma(F_{\sigma_{k-1}(\pi)})$, that is, if $F_{\sigma_{k-1}(\pi)}(x) = F_{\sigma_{k-1}(\pi)}(y)$, see Proposition 3.

Proposition: Each morphism γ_k maps K^n onto the set of states $K^n \setminus \mathrm{GoE}(F_{\sigma_k(\pi)})$, that is, the non-leaf vertices of $\Gamma(F_{\sigma_k(\pi)})$, see Proposition 4. Consequently, we have $K^n \setminus \gamma_k(K^n) \subset \mathrm{GoE}(F_{\sigma_k(\pi)})$.

Theorem: Suppose $x \in \mathrm{GoE}(F_\pi)$, $F_\pi(x) \notin \mathrm{Per}(F_\pi)$, and $F_\pi^{-1}(F_\pi(x)) \subset \mathrm{GoE}(F_\pi)$. Then the following three statements hold, see Theorem 6.

(i) none of the states in the sequence $x^{(k)} = F_{\pi_k} \circ \cdots \circ F_{\pi_1}(x)$ in $\Gamma(F_{\sigma_k(\pi)})$ are periodic.

(ii) the length of any maximal transient path P_k through $x^{(k)}$ in $\Gamma(F_{\sigma_k(\pi)})$ with $0 \leq k \leq n-1$ differs in length from P_0 (the path through x in $\Gamma(F_\pi)$) by at most 1.

(iii) the set $\{\ell(F_{\pi'}) \mid \pi' \in [\pi]_\kappa\}$ of maximal transient lengths of SDS maps over a κ-class $[\pi]_\kappa$ has size at most 2.

The reason for the first condition is to remove the degenerate cases in which transient states are mapped to periodic points under F_{π_k}. The second condition is to avoid the situation with GoE states x such as the one illustrated in the upper left phase space of Figure 1 and the GoE state marked by $(*)$.

Example 1. This example illustrates the main results for the case where the graph G is Circle_4 (vertex set $\{1,2,3,4\}$ and edges all $\{i, i+1\}$ with indices modulo 4) with the added diagonal edge $\{1,3\}$, and bi-threshold functions $\tau_{v,m,k_{01},k_{10}} \colon \{0,1\}^m \longrightarrow \{0,1\}$ defined by

$$\tau_{v,m,k_{01},k_{10}}(x[v]) = \begin{cases} 1, & x_v = 0 \text{ and } \sum_i x[v]_i \geq k_{01}\ , \\ 0, & x_v = 1 \text{ and } \sum_i x[v]_i < k_{10}\ , \\ x_v, & \text{otherwise.} \end{cases} \tag{4}$$

Here the index v is the vertex to which the function is assigned. The phase space for $F_{(2,1,3,4)}$ is shown in the upper left of Figure 1 with the other shifted phase spaces in clockwise order. Here $k_{01} = 1$ and $k_{10} = 3$ in all cases.

The underlined/red vertices shown in the phase space of $F_{(1,3,4,2)}$ in the upper right diagram represent the image of $F_{\pi_1} = F_2$ of $\Gamma(F_{(2,1,3,4)})$ shown in the upper left diagram. Similarly for $F_{\pi_2} = F_1$ from the upper right to the lower right, for $F_{\pi_3} = F_3$ from the lower right to the lower left, and for $F_{\pi_4} = F_4$ from the lower left to the upper left.

Organization of the work. In Section 2 we introduce the basic definitions. Main results are given in Section 3 and the generalization to κ-equivalence is derived in Section 3.1. They are followed by a brief discussion and outline of future work in Section 4.

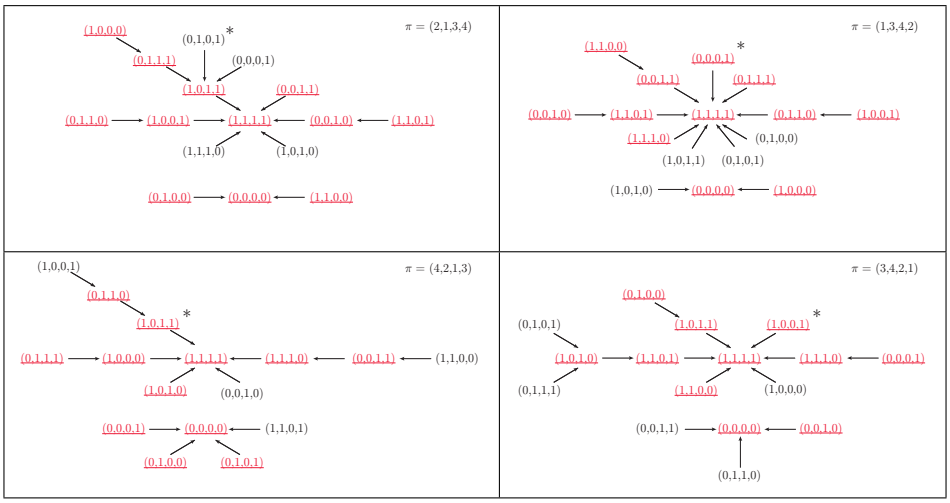


FIGURE 1. Phase spaces for the SDS maps of Example 1. For these $(1,3)$-bithreshold SDS maps, the base update sequence is $\pi = (2,1,3,4)$, shown in the upper left phase space. The remaining phase spaces are for update orders $(1,3,4,2)$ in the upper right, $(3,4,2,1)$ in the lower right, and $(4,2,1,3)$ in the lower left.

2. Background and Terminology

Let G be a finite graph with vertex set $V(G) = \{1, 2, \ldots, n\}$ and undirected edges $E(G)$. For $v \in V(G)$ we let $n[v]$ denote the sorted sequence consisting of v and its neighbors in G. We denote the symmetric group over $V(G)$ by S_G and refer to its elements as *update orders*. To each vertex v we associate a *vertex state* x_v from a finite set K, typically $\{0, 1\}$. The *system state* is $x = (x_1, \ldots, x_n)$. We set $x[v] = (x_{n[v]_i})_i$ and introduce *local dynamics* through a list of *vertex functions* $(f_v)_{v=1}^n$. The role of the function f_v is, essentially, to map the state of vertex v at time t to its state at time $t+1$ based on $x[v]$ at time t. To make this precise, we introduce the functions $F_v \colon K^n \longrightarrow K^n$ by

$$F_v(x_1, \ldots, x_n) = (x_1, \ldots, x_{v-1}, f_v(x[v]), \ldots, x_n) \ .$$

Definition 1. Let $(f_v)_v$ be a sequence of vertex functions over G and let $\pi = (\pi_1, \ldots, \pi_n) \in S_G$. The corresponding *sequential dynamical system map* $F_\pi \colon K^n \longrightarrow K^n$ is the function composition defined by

$$F_\pi = F_{\pi_n} \circ \cdots F_{\pi_1} \ . \tag{5}$$

The *phase space* of a $F \colon K^n \longrightarrow K^n$ is the directed graph $\Gamma(F)$ with $V(\Gamma(F)) = K^n$ and edges $\big\{\big(x, F(x)\big) \mid x \in K^n\big\}$. We let $\mathrm{Per}(F)$

and $\mathrm{Fix}(F)$ denote the set of periodic points and fixed points of F, respectively. A *Garden-of-Eden (*GoE*) state* x of F is a state for which $F^{-1}(x) = \varnothing$. A state $x \in K^n \setminus \mathrm{Per}(F)$ is called a *transient state*, and the first periodic point of the sequence $\{F^t(x)\}_t$ with $t \geq 0$ is called the *periodic point incident to* x, denoted by x_ω. For $x \in \mathrm{GoE}(F)$ we define $P_F(x)$ by

$$P_F(x) = (x, F(x), F^2(x), \ldots, F^{\tau-1}(x)) \ , \tag{6}$$

a path in $\Gamma(F)$, where τ is the maximal integer for which $F^{\tau-1}(x) \notin \mathrm{Per}(F)$. We set $\ell(P_F(x)) = \tau$, which we call the *transient length* of x, and we refer to $P_F(x)$ as a *maximal transient path of* F. We define the maximal transient length $\ell(F)$ of $F\colon K^n \longrightarrow K^n$ by

$$\ell(F) = \max_{x \in \mathrm{GoE}(F)} \ell(P_F(x)) \ . \tag{7}$$

The *generalized phase space* of the function sequence $F = (F_v)_v$ over G is the directed graph $\bar{\Gamma}(F)$ with vertex set K^n and edge set

$$\big\{ \big(x, F_v(x)\big) \mid x \in K^n,\ v \in V(G) \big\} \ , \tag{8}$$

where the edge $(x, F_v(x))$ is labeled by v. A π*-edge* in $\bar{\Gamma}(F)$ is a sequence (or path in $\bar{\Gamma}(F)$) of $n+1$ states $(x(0), x(1), \ldots, x(n))$ where $x(k) = F_{\pi_k}(x(k-1))$ for $1 \leq k \leq n$, and thus $F_\pi(x(0)) = x(n)$. It is clear that the π-edges in $\bar{\Gamma}(F)$ correspond to the edges of $\Gamma(F_\pi)$ and that for each $x \in K^n$ there is a unique π-path in $\bar{\Gamma}(F)$ with origin x. A π*-path* is a concatenation of incident π-edges. Thus π-paths correspond to orbits of F_π. A *partial* π*-edge* is an end portion of a π-edge, that is, a sequence of the form $(x(m), x(m+1), \ldots, x(n))$ where $m > 1$ and with the same conditions as for a π-edge. Part of the generalized phase space of $F = (F_v)_v$ is illustrated in Figure 2. Finally, for $\pi = (\pi_1, \ldots, \pi_n) \in$

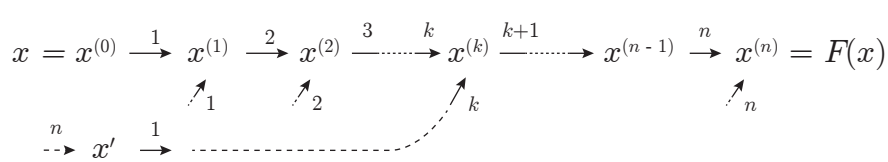


FIGURE 2. A portion of the generalized phase space of $F = (F_v)_v$. For simplicity, we have used edge labels 1, 2, and so on rather than π_1 and π_2. This local structure detailing the transition $x \mapsto F_\pi(x)$, as well as how other transitions $x' \mapsto F_\pi(x)$ may occur, are central to the proof of our Theorem 6.

S_G we write $\sigma_k(\pi) \in S_G$ for the cyclic, left k-shift of π, that is

$$\sigma_k(\pi) = (\pi_{k+1}, \ldots, \pi_n, \pi_1, \ldots, \pi_k) \ .$$

3. Results

We start by recording the following result, parts of which was shown in [9].

Proposition 2. *The map*

$$F_{\pi_k} \colon \Gamma(F_{\sigma_{k-1}(\pi)}) \longrightarrow \Gamma(F_{\sigma_k(\pi)}) \tag{9}$$

is a graph morphism for $1 \leq k \leq n$. *Restricted to* $\mathrm{Per}(F_{\sigma_{k-1}(\pi)})$, F_{π_k} *is an isomorphism.*

Proof. We address the case $k = 1$, the remaining cases being identical. From Equation (2) it is clear that if $F_\pi(x) = z$, then $F_{\sigma_1(\pi)}(F_{\pi_1}(x)) = F_{\pi_1}(z)$, and $(F_{\pi_1}(x), F_{\pi_1}(z))$ is an edge (possibly a loop) in $\Gamma(F_{\sigma_1(\pi)})$, proving the first statement. The second statement is shown in [9]. □

Naturally, if the map F_{π_k} is bijective, then $F_{\sigma_{k-1}(\pi)}$ and $F_{\sigma_k(\pi)}$ have isomorphic phase spaces. Next, we present a criterion guaranteeing that states will not be identified under the graph morphism F_{π_1}.

Proposition 3. *Let* (x, y) *and* (x', y') *be disjoint edges in* $\Gamma(F_\pi)$ *with* $y \neq y'$. *Then* $F_{\pi_1}(x) \neq F_{\pi_1}(x')$.

Proof. Assume that $F_{\pi_1}(x) = F_{\pi_1}(x')$. Then $F_{\pi_n} \circ \cdots \circ F_{\pi_2} \circ F_{\pi_1}(x) = F_\pi(x) = F_\pi(x')$, that is, $y = y'$, a contradiction. □

In other words, for states $x, x' \in \Gamma(F_\pi)$ to be identified under F_{π_1}, a necessary condition is that they have the same image under F_π. We next demonstrate that all non-GoE states of $\Gamma(F_{\sigma_1(\pi)})$ have a non-empty preimage in $\Gamma(F_\pi)$ under F_{π_1}.

Proposition 4. *If* $x' \in \Gamma(F_{\sigma_1(\pi)})$ *is a non-*GoE *state, then there exists* $x \in \Gamma(F_\pi)$ *with* $F_{\pi_1}(x) = x'$.

Proof. Let $x' \in \Gamma(F_{\sigma_1(\pi)})$ be a non-GoE state, and let $z' \in F^{-1}_{\sigma_1(\pi)}(x')$. Then $F_{\sigma_1(\pi)}(z') = F_{\pi_1}(F_{\pi_n} \circ \cdots \circ F_{\pi_2}(z')) = x'$. Clearly, $x = F_{\pi_n} \circ \cdots \circ F_{\pi_2}(z') \in K^n$ satisfies $F_{\pi_1}(x) = x'$. □

Corollary 5. *If* $x \in K^n \setminus F_{\pi_1}(K^n)$ *then* $x \in \mathrm{GoE}(F_{\sigma_1(\pi)})$

Theorem 6. *Assume that $x = x^{(0)} \in \mathrm{GoE}(F_\pi)$ with maximal transient path $P_0 = P_{F_\pi}(x^{(0)})$ satisfies (i) $F_\pi(x) \notin \mathrm{Per}(F_\pi)$ and (ii) $F_\pi^{-1}\big(F_\pi(x)\big) \subset \mathrm{GoE}(F_\pi)$. Then (a) the states $x^{(k)} \in K^n$ defined by*

$$x^{(k)} = F_{\pi_k} \circ \cdots \circ F_{\pi_1}(x) \text{ in } \Gamma(F_{\sigma_k(\pi)}), \quad \text{with} \quad 0 \le k \le n-1 \,, \tag{10}$$

are all transient states of their respective phase spaces. Moreover, (b) any sequence of maximal transient paths $(P_k)_k$ with P_k containing $x^{(k)}$ satisfies the inequality

$$|\ell(P_k) - \ell(P_0)| \le 1 \,, \tag{11}$$

and (c)

$$\big|\{k \mid \ell(P_k) \ne \ell(P_{k+1})\}\big| \le 2 \,. \tag{12}$$

Proof. We first show that $x^{(k)} \notin \mathrm{Per}(F_{\sigma_k(\pi)})$ by constructing a contradiction. Assume that $x^{(k)}$ is periodic. Since $F_{\pi_{k+1}}$ preserves periodic points, see [9], it follows that $x^{(k+1)}$ is periodic. By induction, we conclude that $x^{(n)} = F_\pi(x^{(0)})$ is periodic. This is impossible in light of the first assumption, proving the first part of the theorem. From this we conclude that there is at least one maximal transient path P_k containing $x^{(k)}$ for $0 \le k < n$.

Let $P_0 = \big(x, F_\pi(x), F_\pi^2(x), \dots, F_\pi^{\tau-1}(x)\big)$ where $\tau \ge 2$, and let $h_k = F_{\pi_k} \circ \cdots \circ F_{\pi_1}$. Then, for any maximal transient path P_k containing $x^{(k)} = h_k(x)$ in $\Gamma(F_{\sigma_k(\pi)})$, the path $\big(h_k(x), h_k(F_\pi(x)), \dots, h_k(F_\pi^{\tau-2}(x))\big)$ is a (consecutive) part of P_k. That is, P_k is of the form

$$\big(\dots, x^{(k)}, h_k(F_\pi(x)), \dots, h_k(F_\pi^{\tau-2}(x)), \dots\big) \,.$$

The remaining task is to analyze if in P_k there is a state preceding $x^{(k)}$ and if there is a non-periodic state following $h_k(F_\pi^{\tau-2}(x))$.

For any $z \in F_\pi^{-1}(F_\pi(x))$, there might be many partial π-edges ending at z, see the following illustration where the edge label i indicates π_i:

$$\begin{tikzcd}
\xi_x \arrow[r, "m_x"] & \cdots \arrow[r, "n"] & x \arrow[r, "1"] & x^{(1)} \arrow[r, "2"] & \cdots \arrow[r, "k"] & x^{(k)} \arrow[r, "k+1"] & \cdots \arrow[r, "n"] & F_\pi(x) \arrow[d, dashed] \\
\xi_z \arrow[r, "m_z"] & \cdots \arrow[r, "n"] & z \arrow[r, "1"] & z^{(1)} \arrow[r, "2"] & \cdots \arrow[urr, "k"] & & & \\
 & \xrightarrow{m \ge m_z} \cdots \arrow[ur, "n"] & z' \arrow[r, "1"] & z'^{(1)} \arrow[r, "2"] & \cdots \arrow[r, "k"] & z'^{(k)} \arrow[r, "k+1"] & \cdots \arrow[uurr, "n"] & x_\omega
\end{tikzcd}$$

Let m_z denote the minimum number such that $(\pi_{m_z}, \dots, \pi_n)$ induces a partial π-edge of maximum length among all the partial π-edges ending

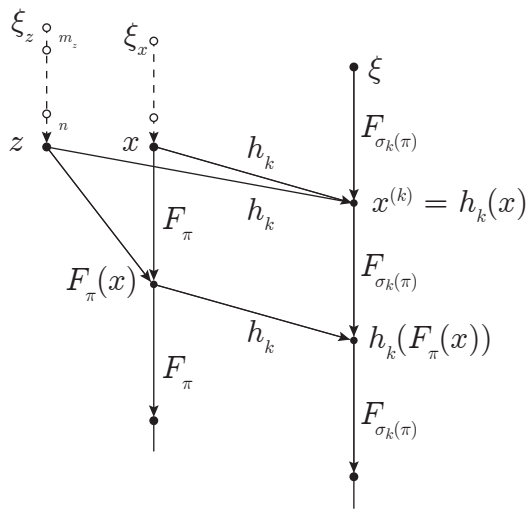


FIGURE 3. The diagram for the proof of Part (b) and Case (1) of Theorem 6. Partial π-edges are dashed. Hollow circles indicate states from the generalized phase space.

at z. Then $m_z \geq 2$ since otherwise we have a full π-edge, and thus z would have at least one preimage in F_π, contradicting the assumption that $z \in F_\pi^{-1}(F_\pi(x)) \subset \mathrm{GoE}(F_\pi)$.

Note that not all states in $F_\pi^{-1}(F_\pi(x))$ necessarily have π-paths in $\bar{\Gamma}(F)$ through $x^{(k)}$. We are only interested in those states z that satisfy $h_k(z) = x^{(k)}$. For a fixed k, let $m'_k = \min\{m_z : z \in F_\pi^{-1}(F_\pi(x)) \wedge h_k(z) = x^{(k)}\}$. There are two cases:

Case (1): $m'_k \leq k+1$. We claim that in any maximal transient path P_k containing $x^{(k)}$, the state $x^{(k)}$ has a preimage which is a GoE state in $F_{\sigma_k(\pi)}$. By the definition of m'_k, if $m'_k \leq k+1$, there exist $z \in F_\pi^{-1}(F_\pi(x))$ and $2 \leq m_z \leq k+1$ such that $h(z) = x^{(k)}$, and $(\pi_{m_z}, \ldots, \pi_n)$ induces a partial π-edge ending at z. Assuming that the starting state of the partial π-edge is ξ_z, we have

$$F_{\pi_n} \circ \cdots \circ F_{\pi_{m_z}}(\xi_z) = z, \quad \text{and} \quad h_k(z) = x^{(k)} \ ,$$

and since $m_z \leq k+1$, we have

$$F_{\pi_k} \circ \cdots \circ F_{\pi_1}[F_{\pi_n} \circ \cdots \circ F_{\pi_{k+1}}(F_{\pi_k} \circ \cdots \circ F_{\pi_{m_z}}(\xi_z))] = x^{(k)} \ .$$

That is, $F_{\sigma_k(\pi)}(F_{\pi_k} \circ \cdots \circ F_{\pi_{m_z}}(\xi_z)) = x^{(k)}$, and the preimage $F_{\pi_k} \circ \cdots \circ F_{\pi_{m_z}}(\xi_z)$ must be a GoE state since otherwise z would not be a GoE state by contradicting $m_z \geq 2$.

Case (2): $m'_k > k+1$. Then $x^{(k)}$ is a GoE state in $F_{\sigma_k(\pi)}$ by reasoning similar to the one of the previous case.

Summarizing Cases (1) and (2), in any maximal transient path P_k, there will be at most one state preceding $x^{(k)}$. Furthermore, by noticing that $m'_k \geq m'_l \geq 2$ due to the easy-to-check observation $\{z \mid h_k(z) = x^{(k)}\} \subseteq \{z \mid h_l(z) = x^{(l)}\}$ for $k < l$, we have shown: if for some k, in any maximal transient path P_k, there is one state preceding $x^{(k)}$, then for any $l > k$, in any maximal transient path P_l, there is one state preceding $x^{(l)}$.

Let x_ω be the periodic point incident to x, and let y be the periodic point such that $F_\pi(y) = x_\omega$. Let $0 < m'' \leq n$ be the minimum integer such that $h_{m''}(y) = h_{m''}(F_\pi^{\tau-1}(x))$. This is illustrated in the following:

$$\begin{tikzcd}
x \arrow[r, dashed, two heads] & F_\pi^{\tau-1}(x) \arrow[r, "1"] & \cdots \arrow[r, "m''"] & \xi \arrow[r, "m''+1"] & \cdots \arrow[r, "n"] & x_\omega \arrow[dddll, dashed, bend left] \\
 & & & \xi \arrow[u, "m''"] & & \\
 & & & \vdots \arrow[u] & & \\
 & & & y \arrow[u, "1"] & &
\end{tikzcd}$$

Then, it is clear that:

(i). For $k < m''$ the state $h_k(F_\pi^{\tau-1}(x))$ is not a periodic point in $\Gamma(F_{\sigma_k(\pi)})$, and $h_k(F_\pi^{\tau-1}(x))$ belongs to any maximal transient path P_k.
(ii). For $k \geq m''$, the state $h_k(F_\pi^{\tau-1}(x))$ is a periodic point in $\Gamma(F_{\sigma_k(\pi)})$ and is therefore not in P_k.

Thus (i) and (ii) imply that in any maximal transient path P_k, there exists at most one state following $h_k(F_\pi^{\tau-2}(x))$.

Note that the cases from the $x^{(k)}$-end and the cases from the $h_k(F_\pi^{\tau-2}(x))$-end are independent. Taking all of the cases we discussed above into account, the maximal transient path P_k can have a possible length contained in the set $\{\tau-1, \tau, \tau+1\}$, completing the proof of Part (b).

For Part (c), let $k_{\min}$ be the minimum integer such that in a maximal transient path $P_{k_{\min}}$, there is a state preceding $x^{(k_{\min})}$. From (i) and (ii) above, we can conclude:

- If $k_{\min} < m''$ we have: for $0 \leq k < k_{\min}$, $\ell(P_k) = \tau$; for $k_{\min} \leq k < m''$, $\ell(P_k) = \tau + 1$; and for $k \geq m''$, $\ell(P_k) = \tau$.
- If $k_{\min} = m''$, we have: for $0 \leq k < n$, $\ell(P_k) = \tau$.
- If $k_{\min} > m''$, we have: for $0 \leq k < m''$, $\ell(P_k) = \tau$; for $m'' \leq k < k_{\min}$, $\ell(P_k) = \tau - 1$; and for $k \geq k_{\min}$, $\ell(P_k) = \tau$.

Hence, Part (c) follows, completing the proof. □

Remark 7. From Part (c) of the previous proof it is clear that there may be SDS maps for which the length of the paths P_k are all the same.

Corollary 8. *Let $(f_i)_i$ be a sequence of vertex functions over G, let $\pi \in S_G$, and define $\lambda_k = \ell(F_{\sigma_k(\pi)})$ for $0 \leq k < n$. There is a positive integer m such that*

$$\left\{\lambda_k \mid 0 \leq k < n\right\} \subset \left\{m-1, m\right\} .$$

Proof. Set $M = \max_k \lambda_k$. If $M = 0$ (true if and only if F_π is invertible), or if $M = 1$, then the statement holds with $m = 1$. If $M \geq 2$ we can apply Theorem 6 choosing a GoE state $x^{(0)}$ whose maximal transient path has length M, possibly using $\pi' = \sigma_k(\pi)$ for a suitable value of k. In this case, the statement holds for $m = M$. □

Example 1. (Continued) In the phase space in the upper left of Figure 1 the marked GoE state $x^{(0)} = (0, 1, 0, 1)$ does not satisfy the conditions of Theorem 6, nor does the point $x^{(2)}$. If we set $z^{(0)} = x^{(2)}$ and ignore the fact that $F_\pi(z^{(0)})$ is periodic, then the fact that condition (ii) of the theorem is not satisfied allows for the possibility that the corresponding maximal transient paths may exceed P_0 in length by more than 1.

3.1. **Generalization to κ-Equivalence.** So far, we have only considered cyclic shifts of update sequences. However, the theory we have developed extends directly to κ-equivalence as shown in this section. To start, we recall necessary definitions; for full details see [9]. First, and without loss of generality, we may assume that the dependency graph G is simple and undirected. For $\pi \in S_G$ we let $O(\pi)$ denote its induced acyclic orientation, which we view as a map, and we write $\mathrm{Acyc}(G)$ for the set of all acyclic orientations of G. Two update sequences $\pi, \pi' \in S_G$ are *α-equivalent* if $O(\pi) = O(\pi')$, and, if this is the case, it follows that $F_\pi = F_{\pi'}$, see [11, 14] where it is also shown that $\pi \sim_\alpha \pi'$ if and only if π and π' differ by a sequence of adjacent transpositions none of which corresponds to an edge in G. Thus, α-equivalence implies functional equivalence of maps F_π.

Continuing, two acyclic orientations $O, O' \in \mathrm{Acyc}(G)$ are *κ-equivalent* if the differ by a sequence of *source-to-sink conversions*. A source-to-sink conversion of an acyclic orientation O is a conversion of one of its sources to a sink. It is done by reversing the orientation of all edges incident to the given source vertex. A *click sequence* mapping O to O' is a sequence $c = (c_1, \ldots, c_m)$ of vertices where each vertex c_i is a source at the time it is applied. We write $O' = c(O)$ for this, and set $c^m = (c_1, \ldots, c_m)$ where $c^0 = \varnothing$ is the empty sequence. Thus, c_{i+1} is a source in $c^i(O)$ for $0 \leq i \leq m-1$. We may always assume that $m < n$, see [10]. At the level of permutation update sequences, a source-to-sink conversion corresponds precisely to (i) a (possibly empty) sequence of adjacent transpositions of non-adjacent vertices followed by (ii) a cyclic, left 1-shift bringing the selected source vertex to the end of the update sequence. This gives rise to the equivalence relation $\sim_\kappa$ on $\mathrm{Acyc}(G)$, and this extends directly to S_G. Since the first type of transformation renders the resulting map F_π invariant (α-equivalence), it follows that the theory we have developed extends from sequences of cyclic shifts of update sequences $\pi \in S_G$ to the more general setting of sequences of source-to-sink conversions of acyclic orientations $O \in \mathrm{Acyc}(G)$. Let $\pi \in S_G$ be a linear extension of $O \in \mathrm{Acyc}(G)$. By α-equivalence, $F_O := F_\pi$ is well-defined.

Theorem 9. *Assume that $x = x^{(0)} \in \mathrm{GoE}(F_\pi)$ satisfies all conditions of Theorem 6, and let $\pi \in S_G$. For any $\pi' \in [\pi]_\kappa$ and click-sequence $c = (c_1, \ldots, c_m)$ for $O(\pi)$, the sequence of states $x^{(k)} \in K^n$ defined by*

$$x^{(k)} = F_{c_k} \circ \cdots \circ F_{c_1}(x) \text{ in } \Gamma(F_{c^k(O(\pi))}), \quad \text{with} \quad 0 \leq k \leq m \ ,$$

are all transient states of their respective phase spaces. Moreover, any sequence of maximal transient paths $(P_k)_k$ with P_k containing $x^{(k)}$ satisfies the inequality

$$|\ell(P_k) - \ell(P_0)| \leq 1 \ .$$

Corollary 10. *Let G be a simple graph. For any sequence $(f_i)_i$ of vertex functions, and for any $\pi \in S_G$, there exists a positive integer m such that*

$$\left\{\ell(F_{\pi'}) \mid \pi \sim_\kappa \pi'\right\} \subset \left\{m-1, m\right\} \ .$$

An immediate consequence is the following: for fixed vertex functions $(f_i)_i$, the set L of all possible maximal transient lengths of maps F_π is bounded in size by

$$|L| \leq 2T_G(1,0) = 2\kappa(G) \ ,$$

where T_G is the Tutte polynomial of G and $\kappa(G)$ is the number of κ-classes for G, [8]. Again, for complete details on κ-equivalence, we refer to [8–10].

4. Summary and Questions

In this paper we have extended the work of [9] and demonstrated that, in addition to the periodic structure, a large portion of the transient structure of phase spaces of SDS (and ABN) is preserved under κ-equivalence. We have shown that the maps $(F_i\colon K^n \longrightarrow K^n)_i$ are morphisms connecting the phase spaces $\big(\Gamma(F_{\sigma_k(\pi)})\big)_k$, and have used this to establish several shared properties of the underlying, shifted SDS maps. In particular, Theorem 6 demonstrates how sequences of transient path structures are essentially preserved across all the shifted phase spaces.

Returning to Figure 1 we see that there are in fact three candidates for the state $x^{(0)}$ of Theorem 6. These are $(1,0,0,0)$, $(1,1,0,1)$ and $(0,1,1,0)$. As a result, Theorem 6 guarantees that there are 3 separate sequences of maximal transient paths. For future work, it may be interesting to enumerate or bound the number of independent sequences of maximal transient paths $(P_k)_k$.

Another aspect of this work is that it offers new insight into the structure of the generalized phase space of $F = (F_i)_i$. This object encodes the phase space of all maps of the form F_π with $\pi \in S_G$, but also information about more complex structures such as sequential dynamical systems where the update sequence is a word or where the update mechanism is part of a stochastic process. An aspect not addressed in this paper is what may loosely be called the dynamics of GoE-states under shifts to the update sequence: for the non-preserved portions of phase space, how do these states "re-attach" themselves to the preserved portions of phase space. After n cyclic shifts, one returns to the original phase space, so it seems reasonable to expect that there is some structure and regularity to this process. Insights on this would possibly offer more information about GoE-states and their enumeration.

Acknowledgements. We thank our external collaborators, members of the Network Dynamics and Simulation Science Laboratory (NDSSL) and the Mathematical Biocomplexity Laboratory (MBL) for their suggestions and comments. In particular, we thank Sichao Wu. This work has been partially supported by grants HDTRA1-17-0118 and HDTRA1-11-D-0016-0001.

References

1. Omar Colón-Reyes, R. Laubenbacher, and B. Pareigis, *Boolean Monomial Dynamical Systems*, Annals of Combinatorics **8** (2004), 425–439.
2. S. Eubank, H. Guclu, V. S. A. Kumar, M. V. Marathe, Srinivasan A., Z. Toroczkai, and N. Wang, *Modelling disease outbreaks in realistic urban social networks*, Nature **429** (2004), no. 6988, 180–184.
3. E. Goles and J. Olivos, *Periodic behavior in generalized threshold functions*, Discrete Mathematics **30** (1980), 187–189.
4. _____, *Comportement periodique des fonctions a seuil binaires et applications*, Discrete Applied Mathematics **3** (1981), 93–105.
5. Eric Goles and Servet Martinez, *Neural and automata networks: Dynamical behaviour and applications*, Kluwer Academic Publishers, 1990.
6. Qijun He, Matthew Macauley, and Robin Davies, *Dynamics of complex boolean networks: canalization, stability, and criticality*, pp. 93–119, Academic Press, 2015.
7. Reinhard Laubenbacher and Bodo Pareigis, *Update schedules of sequential dynamical systems*, Discrete Applied Mathematics **154** (2006), no. 6, 980–994.
8. Matthew Macauley and Henning S. Mortveit, *On enumeration of conjugacy classes of Coxeter elements*, Proceedings of the American Mathematical Society **136** (2008), no. 12, 4157–4165, math.CO/0711.1140.
9. _____, *Cycle equivalence of graph dynamical systems*, Nonlinearity **22** (2009), no. 2, 421–436, math.DS/0709.0291.
10. _____, *Posets from admissible coxeter sequences*, The Electronic Journal of Combinatorics **18** (2011), no. P197, Preprint: math.DS/0910.4376.
11. Henning S. Mortveit and Christian M. Reidys, *Discrete, sequential dynamical systems*, Discrete Mathematics **226** (2001), 281–295.
12. _____, *An introduction to sequential dynamical systems*, Universitext, Springer Verlag, 2007.
13. Kai Nagel and Peter Wagner, *Traffic flow: Approaches to modelling and control*, John Wiley & Sons, 2006.
14. C. M. Reidys, *Acyclic orientations of random graphs*, Advances in Applied Mathematics **21** (1998), no. 2, 181–192.
15. Alan Veliz-Cuba and Brandilyn Stigler, *Boolean models can explain bistability in the lac operon*, Journal of Computational Biology **18** (2011), 783–794.

Mathematical Biocomplexity Laboratory, Biocomplexity Institute, Virginia Tech

E-mail address: cxiaof6@vt.edu

Network Dynamics and Simulation Science Laboratory, Biocomplexity Institute, Virginia Tech

E-mail address: jamcnitt@vt.edu

Department of Mathematics & Network Dynamics and Simulation Science Laboratory, Virginia Tech

E-mail address: henning.mortveit@vt.edu

Network Dynamics and Simulation Science Laboratory, Biocomplexity Institute, Virginia Tech

E-mail address: vgnc@vt.edu

Department of Mathematics & Mathematical Biocomplexity Laboratory, Virginia Tech

E-mail address: duckcr@vt.edu

4.2 Appendix B

A Multi-pathway Modeling Approach to Assess the Threat of *Tuta absoluta* in Southeast Asia

Joseph McNitt[1], Young Yun Chungbaek[1], Henning Mortveit[1], Madhav Marathe[1], Mateus Ribeiro de Campos[2], Nicolas Desneux[2], Thierry Brévault[3], Rangaswamy Muniappan[4], and Abhijin Adiga[1]

[1]*Biocomplexity Institute of Virginia Tech,* [2]*French National Institute for Agricultural Research,* [3]*CIRAD, BIOPASS,* [4]*Feed the Future Integrated Pest Management Innovation Lab*

Abstract

Tuta absoluta is a devastating pest of the tomato crop which has spread in Europe, Africa, and Asia over the last decade. There is strong evidence of multiple pathways, both natural and human-assisted, for its rapid range expansion. We propose a generic data-driven epidemiological modeling approach to study this complex phenomenon accounting for biology, seasonal production, trade and demographic information. We apply this model to study the dynamics of *T. absoluta* spread in South and Southeast Asia – a region at the frontier of its current range. Our objective is to assess the possible routes of introduction, the role of different pathways in the spread and predict its spread pattern in the study region.

Our analysis with respect to incidence reports strongly suggests the role of both natural and human-assisted pathways in the spread of *T. absoluta*. Applying the model to the rest of the study region, we predict that within five years *T. absoluta* will invade all the major vegetable growing areas of mainland Southeast Asia if no steps are taken to mitigate the spread. We also consider alternate scenarios of introduction of the pest to the region through trade and travel. Further, we show that monitoring and effective interventions at the market level can reduce the speed of the spread.

1 Introduction

The world is witnessing a rapid increase in global trade and travel [15]. Due to this increased global connectivity, both international and domestic, no region is spared of the threat from exotic species invasion [21].

climate change and detrimental impact of intensive agriculture on natural resources are likely to further aggravate the problem. As a result, global food security, human health and social welfare will be adversely impacted. The South American Tomato leafminer, or *Tuta absoluta*, is a representative example of biological invasion that has significantly affected tomato production worldwide in the last decade.

Indigenous to South America, *T. absoluta* was accidentally introduced to Spain in 2006 [6, 12]. Since then, it has rapidly spread throughout Europe, Africa, Western Asia, the Indian subcontinent, and parts of Central America [9]. In South Asia, the pest was first reported by India in 2014 [22, 43]. By early 2016, it was discovered in the Kathmandu area of Nepal [4], the northern part of Bangladesh in May 2016 [20], and then in northeastern India [41]. The speed of he spread poses a significant time critical threat to tomato cultivation in Southeast Asia. In order to effectively respond, one needs to identify possible routes of introduction and likely spatio-temporal patterns of spread in this region. With tomato being a commercially important crop, this invastion has had significant global impact. For example, in the Netherlands and Turkey alone, the annual estimated intervention cost are €4 and €167 millions per year, respectively [30, 35]. Due to extensive insecticide treatment in Europe, insecticidal resistance has been recently observed in populations [18]. Overall, lack of effective indigenous predators has made integrated pest management (IPM) a challenging task.

Since tomato is among the top two traded vegetables in the world (http://www.fao.org), it is strongly suspected that trade played a critical role in *T. absoluta*'s rapid spread [9]. Indeed, on multiple occasions it has been discovered in packaging stations ([14] for example). It was observed that the spread pattern in Bulgaria was correlated with prime trade routes [24]. The Animal and Plant Health Inspection Service of the United States Department of Agriculture (USDA-APHIS) has instituted quarantine regulations for imports from regions where the pest is present [47]. Rapid increase in protected cultivation methods such as green houses and tunnel farming have allowed it to overwinter as well as survive the wet season.

Faced with the challenging task of preparing for the invasion of pests and pathogens, and responding effectively to mitigate such incursions should they happen, decision makers are increasingly relying on computational models of pest risk maps and spread to aid decision making [48]. While models that create pest risk maps (e.g. CLIMEX) are useful to identify locations which are suitable for long-term establishment, spread models help specify the spatio-temporal dynamics of how the pest spreads. The latter approach is particularly useful in planning for the possible threat from an invasive species [5, 31, 32, 32, 33] through *in silico* experiments simulating hypothetical invasion scenarios. Together, these tools can be used to address questions such as which locations to monitor, what control measures to take (areawide IPM practices, trade

restrictions, etc.), what is the impact on the economy and health, and so on.

Although there is a general consensus that vegetable and seedling trade is a primary driver of *T. absoluta* spread, previous modeling efforts have exclusively focused on ecological aspects. Two studies [12, 46] provide risk maps using CLIMEX and take additional factors into account. Guimapi et al. [19] used a cellular automata approach to capture the global spread of the pest by factoring in temporal variations and spatial distribution of vegetation, temperature, and tomato production. In recent years, the role of human-mediated dispersal is increasingly accounted for by the modeling community (see for example [11, 15, 28, 39]). In general, modeling the multi-pathway dispersal of pests such as *T. absoluta* is a challenging task due to inadequate understanding of the complex interconnected food system. To add to the problem, most countries ended up not being prepared for the infestation, either due to lack of awareness of the pest or the sheer speed of invasion. Absence of quality incidence records makes callibration and validation hard. Some of these problems were highlighted in the modeling effort by Venkatramanan et al. [49], where a network diffusion model was developed to study the role of trade in the spread of *T. absoluta* in Nepal.

Our work. We describe a multi-pathway propagation model to study the spread of invasive species, which is applied to study the possible spread of *T. absoluta* in the region of South and Southeast Asia comprising of 10 countries: the Association of Southeast Asian Nations (ASEAN) and Bangladesh. We believe this is a timely study, given that *T. absoluta* has already invaded almost all major tomato growing areas in Bangladesh and according to unofficial reports parts of Myanmar. To develop the model, we identified, analyzed, and fused disparate datasets corresponding to natural factors (precipitation, elevation, vegetation, humidity, temperature), biology (host preference, suitability, population growth), seasonal production of host crops, trade dynamics (imports, exports, domestic trade, city locations, city to city distances), and demographic factors (consumption, GDP, population). To our knowledge, this is the first study that explicitly accounts for multiple pathways of introduction and spread of *T. absoluta*. Furthermore, earlier work has not studied the entire Southeast Asia region, particularly in the context of agricultural crops. With the pest having already spread to major tomato producing areas in Bangladesh [20], there is a high chance that it will be introduced to the remaining countries in the near future. From a methodological perspective, this work contributes a generic and modular framework to study invasive species spread accounting natural and human assisted dispersal pathways.

Our analysis suggests the role of multiple pathways in the spread of *T. absoluta*. The model was evaluated

with respect to incidence reports in Bangladesh. In particular, while short distance spread (self-mediated) is important for rapid spread locally, critical long distance jumps (human-mediated) explain the non-radial expansion in the country. Applying the model to the rest of the study region, we predict that within five years *T. absoluta* will invade all the major vegetable growing areas of mainland Southeast Asia if no steps are taken to mitigate the spread. We also consider alternate scenarios of introduction of the pest to the region through trade and travel. Finally, we analyze the effect of monitoring and controlling with respect to different pathways.

2 Methods

We developed a stochastic network propagation model to simulate the multi-pathway spread of *T. absoluta*. See Figure 1a for an illustration. The network is constructed by overlaying the focus region with a grid. The grid cells form the nodes of the network. A directed-weighted edge from one cell to another indicates that there is a pathway by which the pest can be introduced from the former to the latter. There are three pathways in which a cell can become infected: short-distance dispersal, local human-mediated dispersal and long-distance dispersal. Short-distance dispersal captures the spread through natural means, from an infested cell to its adjacent susceptible cells. For human-assisted spread we consider large urban areas in the region which we refer to as *localities*, and model the interactions within and between localities. Local human-mediated dispersal captures the spread due to "local" or farm–city–farm interactions, where the pest can spread from infested cells in the vicinity of a large urban area to other cells in the locality. Long-distance human-mediated dispersal corresponds to spread through trade between cities. Trade flows are estimated using a gravity model approach accounting for total production and population in and around each city. The susceptibility of a cell to pest invasion is determined by the availability of preferred hosts in that cell. If the pest is already established in a cell, then its "infectiousness" depends on the seasonal production volume of these host crops (a surrogate for pest population) and relative preference of the pest to these hosts.

This section is organized as follows. First we provide information on the ecology of *T. absoluta*, setting the background for model structure and some of the key assumptions. Then, we describe the model, followed by the methodology used to determine various parameter values or ranges. The parameter values and major datasets used are summarized in Table 1. The final section is on experiment design describing the scenarios studied.

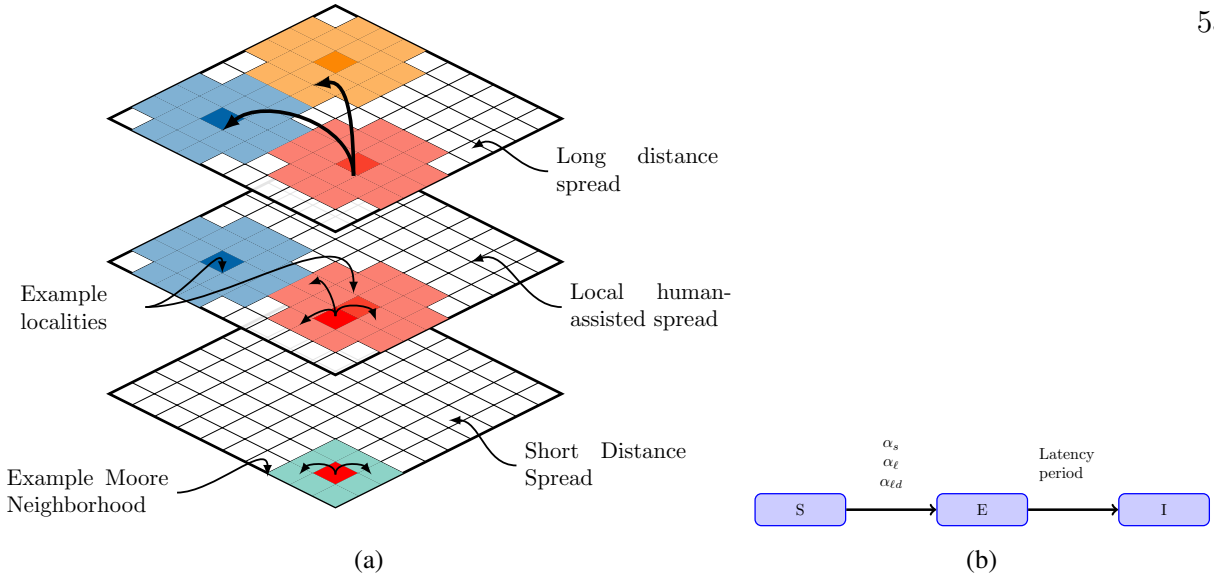


Figure 1: **Schematic of the model.** (a) The multiple pathways of spread are illustrated. (b) The states and factors that influence state transitions.

2.1 *T. absoluta* biology

The tomato leafminer exhibits a short life cycle of about 24–38 days (temperature at $25 \pm 3°$C), from egg to adult, as it is a multivoltine species with overlapping generations in the field [17]. This species causes serious damage to numerous solanaceae crops such as eggplants, potatoes, and especially tomato crops [45]. It penetrates into tomato leaves, stems, or fruits, wherein it feeds and develops by creating conspicuous mines as well as galleries. *T. absoluta* additionally restricts tomato plant growth by feeding on the growing tips. Considering the warm weather throughout the year, particularly in the dry season, the study region presents ideal conditions for rapid development and spread of *T. absoluta*. Pest risk analysis [46] shows that the Ecoclimatic Index for this region is above 50 (highly suitable). Sylla et al. [45] analyzed host preference of *T. absoluta* in France and Senegal. While the highest preference is for tomato, it can survive well on eggplant and potato, which happen to be major vegetable crops in the study region. However, since *T. absoluta* primarily attacks leaves of eggplant and potato, the chance of the pest spreading through trade of these crops seems to be low.

2.2 Model description

We use a discrete-time Susceptible-Exposed-Infected (SEI) epidemic model to simulate multi-pathway pest dispersal. See Figure 1a for an illustration. The focus region is overlayed with a grid of cell size 0.25 arc

degree $\times 0.25$ arc degree, which is approximately $27.8\text{km} \times 27.8\text{km}$ at the equator. These dimensions are comparable to that used in Guimapi et. al. [19] ($25\text{km} \times 25\text{km}$). Each cell can be in three states: susceptible (S) denoting pest free state, exposed (E) denoting that the pest has been introduced but the population has not yet built up to influence other cells, and infectious (I) denoting that the pest has established and the cell can influence its neighbors (in each pathway). The simulation progresses in discrete time steps, where each step corresponds to a month. The likelihood that a cell transitions from state S to E depends on (i) suitability of the cell for *T. absoluta* to establish at that time step, (ii) influence of cells in state I in each pathway, and (iii) level of infestation in each cell with state I. An exposed cell goes to the state I after a latency period of ℓ time steps. This is the time required for the population to build up to infect other cells.

State transitions. The rules for state transitions are shown in Figure 1b. A susceptible (S) cell can be influenced by infectious (state I) cells through the three different pathways. Each such cell infects the susceptible cell with some probability. If *T. absoluta* is succesfully introduced to a cell, it moves from state S to E. The exposed state corresponds to the situation where *T. absoluta* has established, but it is not widespread in the area to influence other cells. It stays in state E for one time step before transitioning to state I. This is a reasonable assumption considering that the conditions are favorable for the pest to complete a life cycle within one month. Once the pest has established in a cell, the cell remains infected forever, a fair assumption considering that, historically, eradication of *T. absoluta* has not been successful[1].

Susceptibility and infectiousness of a cell. We have used monthly production to determine the susceptiblity of a cell in state S and infectiousness of a cell in state I. The suitability of a cell v for pest establishment at time t is denoted by $\epsilon(v, t)$. This is 1 if production at t is non-zero and 0 otherwise. For a cell in state I, the level of infestation in an infected cell v at time t is denoted by $\rho(v, t)$. It is modeled as a linear function of host presence at time t, for which we use the weighted sum of production volume of tomato, eggplant, and potato in that cell at time t. The weights correspond to relative oviposition preference of *T. absoluta* on the three hosts.

Localities. To account for human-assisted dispersal, we considered dense centers of human activity such as big cities or towns. Typically, these locations are the major consumers of vegetables. Since horticultural products are less durable, in developing countries, owing to lack of good storage and transport infrastructure [2],

[1]The only exception is United Kingdom where the pest was detected early [13].

major producing regions are close to cities [8]. In addition, they house big wholesale markets and the trader and distributor infrastructure. Therefore, capturing the spread due to activities within and between cities is critical to model human-mediated spread. For this purpose we constructed *localities*, each representing dense pockets of human activity in the following manner. We identified major urban centers in the study region. A *locality* corresponding to each city consists of all grid cells for which this is the nearest city and the distance is at most the *locality radius* (in kilometers), with the additional constraint that each cell belongs to the same country as this city. We also account for administrative restrictions. Local human-mediated dispersal is modeled as the interaction between cells belonging to a city. Long distance human-mediated dispersal is modeled as trade flows of considered host crops from one city to another. For a city i, $\mathrm{L}(i)$ denotes all cells which are assigned to it.

Now, we describe the three different pathways of spread.

Short distance dispersal. At any time step, a cell's state is influenced by its Moore neighborhood of range r. When range $r = 1$, it corresponds to the adjacent cells (eight at most) and when $r = 2$, it corresponds to $r = 1$ neighbors and cells adjacent to them. The Moore neighborhood of v is denoted by M_v.

$$p_s(v,t) = \alpha_s \epsilon(v,t)\Big(1 - \exp\Big(-\sum_{v' \in \mathrm{M}_v(r)} \rho(v',t)\Big)\Big), \tag{1}$$

where α_s is a tuning parameter for this pathway. The probability that v is influenced by its neighbors is $1 - \prod_{v' \in \mathrm{M}_v(r)} 1 - \rho(v',t)$, which is approximated by equation (1).

Local human-mediated dispersal. Here, the cell's state is influenced by the infected cells in its locality through the marketing chain. In general, it is hard to model the local dynamics as there are several actors in bringing the commodities from farm to market to consumers. Further, these are country and commodity specific. See for example Kethonga et al. [25] for the typical structure of marketing chains and Rebaudo et al. [36] for modeling human interactions in the context of invasive species spread. Here, we use a simple approach. Every cell v is influenced by cells in its locality L based on their infectiousness. The expression is similar to that in (1), but with cells in the locality instead of the Moore neighborhood.

$$p_\ell(v,t) = \alpha_\ell \epsilon(v,t)\Big(1 - \exp\Big(-\sum_{v'\in \mathrm{L}} \rho(v',t)\Big)\Big), \tag{2}$$

where α_ℓ is a tuning parameter for this pathway.

Long-distance human-mediated dispersal. We only consider trade of tomato in our model (see Section 2.1). We model the flow of vegetables among markets based on the following assumptions: (i) the total outflow from a city depends on the amount of produce in its surrounding regions and imports from countries outside the focus region at time t, and (ii) the total inflow depends on total consumption, processing demand, and exports from the city to countries outside the focus region. For each country, the domestic flow is estimated using a doubly constrained gravity model [3, 23]. For a city i, let O_i and I_i denote total outflow and total inflow respectively. The flow F_{ij} from city i to city j is given by $F_{ij}(t) = a_i(t)b_j(t)O_i(t)I_j(t)f(d_{ij})$, where, d_{ij} is the time to travel from i to j, and $f(\cdot)$ is the *distance deterrence function*: $d_{ij}^{-\beta}\exp(-d_{ij}/\kappa)$, where β and κ are tunable parameters. The coefficients a_i and b_j are computed through an iterative process such that the total outflow and total inflow at each node agree with the input values [23]. Overall, we have 12 networks representing flows for each month. The outflows and inflows are calculated as follows:

$$O_i(t) = \mathrm{Prod}(i,t) + \mathrm{Import}(i,t) - \mathrm{Export}(i,t) - \mathrm{Proc}(i,t), \tag{3}$$

$$I_i(t) = \mathrm{Cons}(i)\,. \tag{4}$$

Here, $\mathrm{Prod}(i,t)$ and $\mathrm{Cons}(i)$ are the monthly production and consumption (assumed same every month) at the locality. Export and Import are the monthly total export to, import from or both from outside the country. $\mathrm{Proc}(i,t)$ is the tomato produced for processing. Typically, tomato meant for processing is cultivated locally [51]. In the next section, we will discuss how each of these locality attributes are computed. Given a locality j, let its total infectiousness be denoted by $\rho_{\mathrm{T}}(j,t) = \sum_{v'\in \mathrm{L}(j)} \rho(j,t)$. Suppose cell v belongs to locality i. Then, the probability of cell v transitioning from S to I due to long-distance dispersal is given by:

$$p_{\ell d}(v,t) = \alpha_{\ell d}\epsilon(v,t)\Big(1 - \exp\Big(-\sum_{j\neq i}\sum_{v'\in \mathrm{L}(j)} F_{ji}\rho_{\mathrm{T}}(j,t)\Big)\Big). \tag{5}$$

Here, $\alpha_{\ell d}$ is a tunable parameter for this pathway and i is the locality to which v belongs to. The influence depends on total infectiousness of a locality and the flow from that locality to the locality i.

2.3 Cell, locality, and network attributes

In this section, we will describe how attributes such as production, consumption, and trade flows between localities were computed. Table 1 provides the list of model parameters with their values or ranges and data source as applicable.

Seasonal production. Broadly, in this region, the cropping pattern depends on two factors: seasons–dry and wet, and elevation–highland (upland) and lowland. The aim was to estimate monthly production volume of tomato, eggplant and potato for each cell. This was accomplished in two steps. First, we estimated annual production in each cell. Then, this value was disaggregated to monthly production. The annual production was estimated as follows. From SPAM [53], we obtained annual production estimates for each cell. However, there are several issues with directly using this data. Firstly, these are estimates for the year 2005, and secondly, tomato and eggplant production estimates are not available. Instead, vegetable production volume is available. Also, for countries where data was available, we did not find any correlation between reported tomato (eggplant) production and total SPAM vegetable production for that region. Therefore, for each country, we obtained the most recent production data available (2013 or later) at the highest spatial resolution (region/province/country) (Table S1). The production of a particular vegetable type at a cell was computed as follows:

$$\frac{\text{Total production in the region}}{\text{Total SPAM production for cells in the region}} \times \text{SPAM production in the cell}$$

. For tomato and eggplant, we used vegetable production as the surrogate. There were also cases where no data was available (Cambodia, Myanmar and Laos for example). In such cases, for potato, we used SPAM data as is. For tomato and eggplant, the SPAM value for vegetables was scaled by a scaling factor which was determined as follows. For countries where data was available, we computed the ratio of total tomato production and total SPAM vegetable production for the country. The median value (≈ 0.05) was used as the scaling factor. The same procedure was used for eggplant.

To predict seasonal production, we used regional quarterly tomato and eggplant production data that was available for Philippines [34] (16 regions), precipitation and elevation data at the cell level. For each

region, we obtained the product rate by normalizing quarterly production values with respect to maximum value among these. We used production rate instead of production values since there are several factors that determine a region's production: climate, vegetable preference, demand, etc. Therefore, it may not be meaningful to compare production across regions. We conducted a linear regression with the product rate as a dependent variable and precipitation and elevation as independent variables (SPSS 24.0). Since the dependent variable was highly skewed, we used a log-transformation. To control elevation, we classified the elevations into two groups, high and low, using k-means clustering (SPSS 24.0). Due to the small sample size, we excluded the samples in the high-elevation group and conducted a linear regression analysis for the group of low elevation (< 235). The total 56 samples in the group showed negatively strong correlation between precipitation and logarithm of product rate ($r = -0.734$). More information is provided in Supplementary Information. For most countries, only qualitative information on seasonal production is available making validation near impossible. However, we observed that for the entire region seasonal production was strongly tied to rainfall; during the wet season, the amount of production is considerably less compared to the dry season.

Locality construction. We identified cities with population greater than a certain population threshold in the entire study region to determine localities in the model. We considered a range of threshold values, and chose $250,000$ as the threshold for the model with the main criteria for the choice being coverage of production and population, and knowledge of major wholesale markets (Supplementary Information). The locality radius was chosen to be 100kms since local production in an urban area was within 50–60kms from the city center (Table S2, Supplementary Information). To obtain long distance trade flows, the travel times between pairs of cities were computed using Google API [1]. For the distance deterrence function of the gravity model, we set $\beta = 2$ and $\kappa = 300$ (kms). The former was set based on the analysis in Venkatramanan et al. [49] while for the latter we observed from literature survey that long distance trade is typically between nearby towns.

Production, consumption, imports, exports and processing. Monthly tomato production at a locality was obtained by aggregating production at all cells that belong to it. For consumption, we used country-level production [16], which was available for half of the countries. For Singapore, we estimated it as the difference between total inflow (production and imports) and total outflow (exports) based on FAO data. For Vietnam,

Wijk et al. [51] provides this information. For Myanmar, Cambodia and Laos, we found no information. We used median consumption for the region. We also analyzed consumption with respect to per capita gross domestic product (GDP). However, we did not find any correlation between GDP and consumption both globally as well as restricted to the study region.

For most countries, imports and exports are a small fraction of domestic production. We have ignored these quantities. However, these flows were accounted for when evaluating possibility of pest introduction. The main exceptions were the significant tomato imports from India to Bangladesh and trade between Malaysia and Singapore. We identified major routes of trade from India to Bangladesh [29] and the total imports from India (FAOSTAT) was distributed uniformly between three cities close to the border with India along these routes. Finally, these imports were evenly distributed for the later half of the year since Bangladesh imports mostly during the rainy season. To capture the significant trade between Singapore and Malaysia, we included Singapore in the domestic flow network of Malaysia as there is high interaction between the two countries. The resulting flow from the gravity model from Malaysia to Singapore was obtained by aggregating network flows across months and across edges with Singapore as destination. This flow was comparable to the annual imports from Malaysia to Singapore. With the exception for Thailand [26], there is no information on the amount of tomato production consumed by the processing industry even though there is evidence of processing industry in Vietnam [51], Malaysia and Indonesia. For each locality in Thailand, we scaled the monthly production by the ratio annual production of fresh tomatoes over total annual production (fresh and processed).

2.4 Experiment design

The tunable parameters of the model are the pathway scaling factors α_s, α_ℓ, and $\alpha_{\ell d}$, Moore range r_{M}, latency period ℓ and the scenario , i.e., start point of infestation (time and location). These are listed in Table 1. First, we simulated the spread in Bangladesh where *T. absoluta* is already present and incidence reports are available. We evaluated the model output for each set of parameter values by comparing it with ground truth using a maximum likelihood estimation method for comparison. We used a sparse adaptive design with refinements using Classification and Regression Trees (CART) method to parameterize the model. The selected model was subsequently used to study the possible spread in the rest of the region under various scenarios. Sensitivity analysis was performed to assess uncertainty of model outcomes.

Evaluation of the model is a particularly challenging task as available incidence data is highly inadequate

Table 1: Model parameters and variables, their ranges and references.

Parameter	Description	Range/values	Source†
r_{M}	Range of Moore neighborhood	1,2,3	[19]
ℓ	Latency period to transition from E to I	1,2,3	
β, κ	Gravity model parameters	300–500,2	[49]
–	Locality population threshold and radius	250,000, 100kms	population of cities [42] and reports
ϵ	Suitability threshold	0	
ρ	Infectivity of a cell based on amount of production		Host production (SPAM), host preference [45], precipitation and elevation
Scenarios	Seeding simulations: time and cells to infect	4 cases: Bangladesh (B1 and B2), Malaysia (M1) and Philippines (P1)	*T. absoluta* incidence reports in Bangladesh, FAOSTAT for international trade and migration reports.
Start month		March, April, May	Based on incidence reports
α_s	Short-distance spread scaling factor	0–500	
α_ℓ	Local human-mediated dispersal scaling factor	0–500	
$\alpha_{\ell d}$	Long distance spread scaling factor	0–500	

†The appropriate references are provided in the Supplementary Information.

for calibration and validation. For example, there are only eight locations in Bangladesh for which data is available. Further, it is possible that there were delays in identifying and reporting. Therefore, during the model evaluation phase, to account for both spatial and temporal observation noise, we considered six scenarios: Start times were March, April or May (the month *T. absoluta* was reported) and two locations of initial infections (i) B1: cell corresponding to location of first report (26.19°N, 88.43°E) and (ii) B2: all cells in the Moore neighborhood (range 1) for the location of first report. For each parameter setting the simulation was run with 100 repetitions. We computed the empirical probability that a cell is in state I at time t. The output was compared with ground truth using a maximum likelihood method adapted from [10]. Let C be a reporting cell with t_C denote the month of actual report. To account for uncertainty in reporting, we consider a time window rather than the exact time of report during comparison. Let $U_\tau = [t_C - \tau, t_C]$ be the time window or interval, where τ is the uncertainty parameter. Here, for simplicity, we will assume that t_C is greater than or equal to τ. Supposing $\mathcal{C}_R$ is the set of cells corresponding to ground truth, and $p(C,t)$ is the probability that cell C is infected at time t in the model, then, the likelihood $\mathcal{L}$ is given by,

$$\mathcal{L} = \sum_{C \in \mathcal{C}_R} \Big(\sum_{t \in U_\tau} p(C,t) + \sum_{t \notin U_\tau} \big(1 - p(C,t)\big) \Big)$$

In our case, the uncertainty parameter $\tau = 2$. First, we coarsely sampled the parameter space. Applying CART, we identified subspaces of the parameter space for which likelihood was high (see Figure 3c). We refined our search in the next stage to improve the parameterization. In addition, we implemented the cellular automata model by Guimapi et al. [19] which serves as the baseline model for comparison.

Software and Computational aspects. The model was implemented in Python 2.7. For data management and processing, we used PostGreSQL and SQlite. Statistical analysis was done using Python, R 3.4.3 and SPSS 24.0. The experiments were run using Discovery, a high performance computing cluster with 232 nodes (16-core Sandy Bridge-EP E5-2670 2.60GHz (3.30GHz Turbo) Dual Processor (8 Cores per Processor) nodes with 32 GB of Memory and 500 GB Internal Hard Drive).

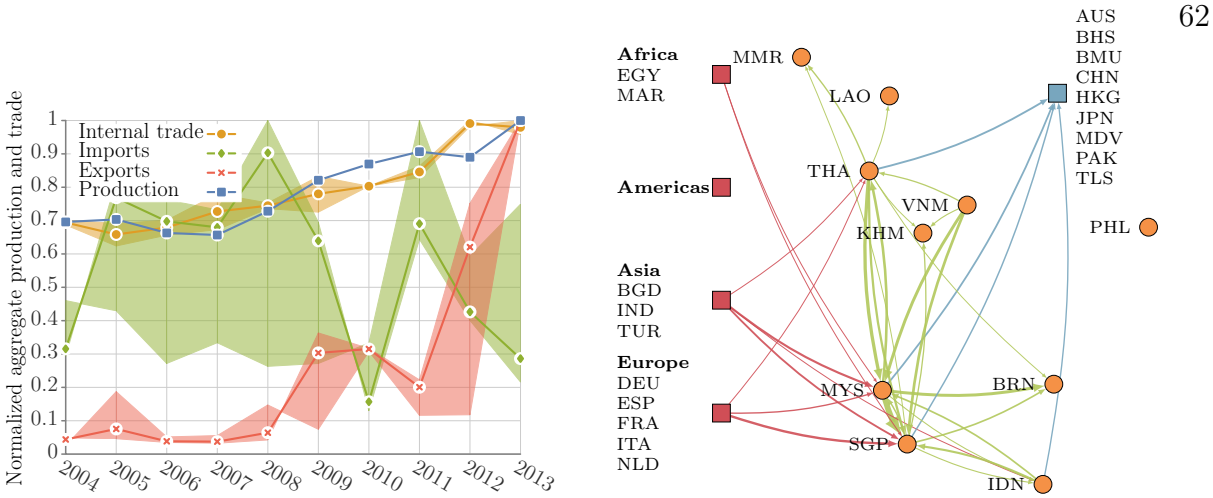
3 Results

3.1 Production, trade and pathways of pest entry in the study region

Using data from FAO (FAOSTAT 2013, Table 1), we analyzed the dynamics of production and trade of various hosts of *T. absoluta*. Here, the objective is to study the dynamics of country-level production and trade flow of various hosts of *T. absoluta*.

General trends. A plot of normalized aggregated tomato production, imports, and exports in the region is shown in Figure 2a (Supplementary Information). There is a steady increase in the production and amount of internal trade, with more or less the same rate of change for both quantities. While the export of tomato outside of the focus region (plot "Exports") has risen steeply in the recent years, the imports generally indicate a downward trend (plot "Imports"). However, there is a lot of discrepancy in reporting import (or export) figures, particularly the commodity flows with countries outside the region. The filled curves shows the difference in these estimates when considered individually. However, the plots clearly indicate that there is a general thrust towards tomato production and trade in the region. Recent efforts to increase production and trade infrastructure in these countries endorse this increasing trend. Under these circumstances, *T. absoluta*'s invasion can have a high negative impact on the economy and livelihood of the people in this region.

Production. Among the countries listed in the FAOSTAT dataset, Indonesia is by far the largest producer ($\approx$ 1M tonnes) followed by Phillipines and Malaysia ($\approx$ 200K tonnes each). FAOSTAT does not have tomato



(a) (b)

Figure 2: **Country-level production and trade of tomato in the focus region using FAOSTAT 2013.** (a) General trends of overall production and trade in the region over a decade. We aggregated tomato production across countries in the region for each year from 2004 to 2013 and normalized by dividing it by the maximum value. We did the same with imports into and exports out of the region as well as internal trade. The trade data presented for each year is the average of total quantity reported by importing countries and exporting countries. (b) The trade of tomato in the focus region. The trade between countries of the focus region is represented by green edges. Also shown are imports from *T. absoluta* affected countries categorized by region. The edge thickness is a function of the trade volume.

production data for Vietnam and Cambodia. However, alternate data sources (personal communication) indicate that Vietnam is the second largest producer ($\approx$ 400K tonnes) (see [50] for 2005 information). The data also indicates that Indonesia is the largest producer of eggplants in the region, followed by Philippines.

Trade. The network capturing tomato trade within the focus region, imports from *T. absoluta* affected countries and exports to countries which have not reported the pest, is shown in Figure 2b. The network was constructed using data from FAOSTAT Trade matrix (Table 1). We note that with the exception of Philippines, there is a lot of trade activity among countries in the focus region. However, the majority of trading happens between neighboring countries. Malaysia and Singapore serve as major hubs for tomato trade with considerable amount of trade among themselves, within the focus region and beyond. Most significant flows are from Malaysia to Singapore ($> 30,000$ tonnes) followed by Thailand to Malaysia ($> 2,000$ tonnes). Philippines does not report any fresh tomato trade with other countries. However, it does report imports of processed tomato. We also analyzed how the network structure evolved across years. While we did not observe much variation in the network structure, there is some change in the countries importing to and

exporting from this region (Supplementary Information for networks of alternate hosts).

Pathways of entry. *T. absoluta* was first reported in northern Bangladesh near Rangpur (Table S3). Since then it has spread to almost all major tomato producing areas of the country. The pest could enter the study region from Bangladesh to neighboring countries such as Myanmar. In fact, our analysis indicates that there is a high chance that the pest is already present in this country. Also, the network structure does indicate a high risk of introduction of *T. absoluta* through trade. In particular, our analysis of how the network has evolved over the years (Supplementary Information) shows that while the basic network structure does not vary much, the imports from *T. absoluta* infested countries is steadily increasing. By virtue of being important hubs, there is a high chance that the pest will enter through the ports of Malaysia and Singapore, and spread to Thailand, Brunei, and Indonesia due to their high interactions with these neighbors. Since Philippines does not share its borders with any country in the region and there is no evidence of tomato trade with rest of the countries, there is a low chance that the pest will be introduced through natural spread or trade. However, human mobility is a possible pathway. The Middle East is the top destination for Filipino workers [40]. Therefore, there is a possibility of introduction through travel. Besides, it is possible that this was the case with India. In 2014, *T. absoluta* was reported in the South Western part much farther away from the country's border with its neighbor Pakisthan which is yet to report presence of the pest. Afghanisthan reported only in 2015. Also, India does not show any imports from *T. absoluta* infested countries.

3.2 Assessing the role of each pathway in the spread of *T. absoluta*

In Bangladesh, following the first report of *T. absoluta* in India in 2014 [22], pheromone traps were deployed in eight places (Table S3). In May 2016, the pest was detected in Panchagarh District in the northern part of Bangladesh [20]. Within 10 months, the pest was reported in almost all major areas of vegetable production. Our analysis strongly indicates that both short distance spread and trade are important pathways in the spread dynamics. For the parameter set with the best fit, the likelihood was close to 6, i.e., on an average, simulation output matched six out of seven locations. If we consider the case where the human-influence was turned off ($\alpha_\ell = 0$ and $\alpha_{\ell d} = 0$), the highest likelihood was 5. Also, we observed an interesting pattern with respect to Moore range (r_{M}) and latency period (ℓ). When only natural spread is considered, the rate of expansion is higher ($r_{\mathrm{M}} = 2$) and latency period (1) is lower than when we account for human-mediated spread ($r_{\mathrm{M}} = 1, \ell = 2$). This is in some sense expected due to long distance jumps facilitated by human-assisted

spread. It also means that *T. absoluta* can achieve the same (or greater) range expansion with considerably less flying capacity and population growth.

While our results highlight the role of human-mediated spread, it is possible that this pathway is more important for the spread than it appears. First of all, vegetable production data for Bangladesh (SPAM) indicates that almost all cells have non-zero production of hosts of *T. absoluta*. Therefore, there is a contiguous landscape of suitable areas for the pest to spread naturally. But in general this may not be the case; the only way two locations can be connected is by trade or travel pathways. One obvious example is two land masses separated by sea. Secondly, and more importantly, it does not explain the gap between the first report and the report in Gaibandha district ($25.15°$N, $89.23°$E, Table **??** in Supplementary Information). The distance between the two locations is only 185kms while this location reported the presence only after nine months of first report suggesting that natural spread might be much slower. According to our best fit model accounting only for natural spread, the corresponding cell gets infected between the third and fifth months.

We simulated the spread using the model developed by Guimapi [19] for Bangladesh. For Moore neighborhoods of 2 and 3 the spread was too rapid. The highest score for the evaluation metric in all cases was 1 for a time window of 2.

4 Discussion

This study is very relevant and timely considering that crops such as tomato, eggplant, and potato are among the top vegetables in most of these countries. Traditionally, these crops have been grown in the winter during the dry season. However, over the past decade, due to rising demand–domestic as well as from neighboring countries–there has been a thrust towards year-round production using protected cultivation methods and resilient varieties [2, 27]. Also the food industry is gradually restructuring: supermarkets replacing traditional food chains. Therefore, invasions from pests such as *T. absoluta* can have a huge negative impact on the socioeconomic fabric of this region. Besides invasive species spread, there are other applications where understanding production and trade dynamics of major crops is useful or even essential. These include studies of natural or human-initiated disasters, climate change, nutrition, etc.

In recent years, there has been a thrust towards integrated modeling approaches to understand invasive species dynamics. Multi-pathway models have been analyzed to study the role of human-mediated dispersal [10, 37, 39]. Robinet et al. [39] show that the distribution and spread pattern of the pinewood nematode in

China is strongly correlated with density of human population and infrastructure such as railways and river ports. A similar approach was applied to the pine processionary moth [37]. Carrasco et al. [10] combine spatially explicit models with a phenology model to incorporate population dynamics of the pest (western corn rootworm). Like the previous works, this work accounts for human population. Carrasco et al. consider two types of long-distance dispersals – domestic and international. The domestic mode is modeled as a flow network between cities using a gravity model approach. The flow between two cities is a function of human population of these cities and the distance between them. Nopsa et al. [28] use a network science approach to studying the role of transport and storage infrastructure in the spread of pests and pathogens of wheat. Sutrave et al. [44] use a time-varying network model to study the spread of Soybean rust. However, in all these cases, there was more information available on the incidence of the pest or pathogen under study enabling validation of the developed models. For emerging pests, this is nearly impossible to obtain such incidence reports, particularly in countries that lack awareness or infrastructure,

Our model is in part motivated by the hybrid approaches used in the study of infectious diseases of humans and livestock (for example [7, 52]). Bradhurst et al. [7] study the spread of foot and mouth disease in livestock by using an aggregate population-level model to capture within-herd spread and an individual-based model for between herd spread. This is much like our approach of capturing local and long-distance human-mediated spread. A similar approach is used by Yang et al. [52] to forecast influenza outbreaks. They use a patch network model where a compartmental model is used to simulate intra-locale spread and a gravity model based approach is used for inter-locality spread of flu in the neighborhoods of New York.

Key challenges. We have used a multitude of datasets– production, consumption, trade dynamics, climate, biology, etc.– to capture, to a reasonable extent, the complexity of the invasion process. A major challenge was data inadequacy. For many countries such as Vietnam, Cambodia, and Laos, production data had to be collected (or even inferred) from several publications and reports. Therefore, data came from disparate sources (multi-type, different countries, etc.), and were misaligned in time (different years) and spatial resolution (grid to country level). Production and trade data from FAOSTAT also had gaps in it. In particular, it was hard to model seasonal production and human assisted spread. Seasonal production is dependent not only on the host, climate, and geography, but also on people's preferences and market demand. For most countries, only qualitative information is available. Another example is the modeling of consumption. It is possible that there is lot of variation in consumption within a country [51] as well as across seasons. Even at

the country level, data is available for only half of the countries. Also, we did not find any correlation between consumption and GDP or tomato production (correlation < 0.01). To determine outflows and inflows for each locality, we had to identify major ports for imports and exports as well as estimate fraction of production which was used for processing. This data was available only for a couple of countries.

Limitations. The fidelity of models such as the one presented here crucially depends on the availability of quality data as well as a good understanding of the processes involved. As pointed out on multiple occasions, the scarcity of data has forced us to greatly simplify some of the processes. However, the developed framework is modular and extensible. Given high-resolution accurate datasets and a better knowledge of the processes, individual modules can be replaced with more sophisticated modeling approaches. One example is the use of volume of production of preferred host crops as a surrogate for population. In the future, complex phenology models can be used instead to model accurately the growth of the pest under specific conditions. This would require years of study of the tritrophic interactions concerning the pest, host, and its indigenous predators (like for example, the work by Carrasco et al. [10]). However, as cautioned by Robinet et al. [38], this would add to the complexity of the model making it near impossible to verify and validate.

The farm–market-consumer interactions (local human-mediated spread) are well understood at a conceptual level. More or less, in all the countries of the focus region, the structure remains the same. It involves various actors such as farmers, wholesalers, retailers, wet markets, supermarkets, etc. Nevertheless, it is nearly impossible to accurately model the dynamics. Rebaudo et al. [36], for example, use a complex agent-based model to study just the interaction between farmers of two villages in the context of the potato moth in Ecuador. Similarly, given data on actual flow of vegetables, the gravity model can be improved or replaced by more sophisticated approaches. The natural spread pathway can be further improved by taking into account wind trajectories.

Applying this generic modeling approach to other study regions or other pests would require taking into account additional factors. For example, seedling trade could be an important pathway, particularly in the European and the Mediterranean region as well as North America. From a production perspective, it is becoming important to factor in protected cultivation methods, which have enabled farmers to extend the growing season. Also, it is important to study account for damage done by the pest.

Acknowledgments

67

This work was supported in part by the United States Agency for International Development under the Cooperative Agreement NO. AID-OAA-L-15-00001 Feed the Future Innovation Lab for Integrated Pest Management, DTRA CNIMS Contract HDTRA1-11-D-0016-0001, NSF BIG DATA Grant IIS-1633028, NSF DIBBS Grant ACI-1443054, NIH Grant 1R01GM109718 and NSF NRT-DESE Grant DGE-154362.

We are grateful to Yousuf Mian, Nguyen Van Hoa, and Kim Hian for their help with obtaining country-specific information on production, trade, and pest incidence.

References

[1] Google API. Retrieved March, 2018.

[2] M. Ali and V. T. B. Hau. Vegetables in Bangladesh: Economic and nutritional impact of new varieties and technologies. *Asian Vegetable Research and Development Centre*, (Technical Bulletin No. 25):1–15, 2001.

[3] James E Anderson. The gravity model. *Annual Review of Economics*, 3(1):133–160, 2011.

[4] Ajaya Shree Ranta Bajracharya, Ram Prasad Mainali, Binu Bhat, Sanjaya Bista, PR Shashank, and NM Meshram. The first record of South American tomato leaf miner, Tuta absoluta (Meyrick 1917)(Lepidoptera: Gelechiidae) in Nepal. *J. Entomol. Zool. Stud*, 4:1359–1363, 2016.

[5] Edward B Barbier, Duncan Knowler, Johnson Gwatipedza, Sarah H Reichard, and Arianne Ransom Hodges. Implementing policies to control invasive plant species. *BioScience*, 63(2):132–138, 2013.

[6] A Biondi, RNC Guedes, and FH Wan. Ecology, Worldwide Spread, and Management of the Invasive South American Tomato Pinworm, Tuta absoluta: Past, Present, and Future. *Annual Review of Entomology*, (63):239–258, 2017.

[7] Richard A Bradhurst, Sharon E Roche, Iain J East, Paul Kwan, and M Graeme Garner. A hybrid modeling approach to simulating foot-and-mouth disease outbreaks in australian livestock. *Frontiers in Environmental Science*, 3:17, 2015.

[8] Amy Buckmaster, Jeffrey Alwang, Everett Peterson, and Mauricio Rivera. Going the distance: How does market access affect demand for ipm packages? *Journal of Integrated Pest Management*, 5(1):B1–B7, 2014.

[9] Mateus R Campos, Antonio Biondi, Abhijin Adiga, Raul NC Guedes, and Nicolas Desneux. From the western palaearctic region to beyond: Tuta absoluta 10 years after invading europe. *Journal of Pest Science*, pages 1–10, 2017.

[10] LR Carrasco, JD Mumford, A MacLeod, T Harwood, Giselher Grabenweger, AW Leach, JD Knight, and RHA Baker. Unveiling human-assisted dispersal mechanisms in invasive alien insects: integration of spatial stochastic simulation and phenology models. *Ecological Modelling*, 221(17):2068–2075, 2010.

[11] Manuel Colunga-Garcia and Robert A Haack. Following the transportation trail to anticipate human-mediated invasions in terrestrial ecosystems. *Pest Risk Modelling and Mapping for Invasive Alien Species. CAB International, Wallingford, UK*, pages 35–48, 2015.

[12] Nicolas Desneux, Eric Wajnberg, Kris AG Wyckhuys, Giovanni Burgio, Salvatore Arpaia, Consuelo A Narváez-Vasquez, Joel González-Cabrera, Diana Catalán Ruescas, Elisabeth Tabone, Jacques Frandon, et al. Biological invasion of European tomato crops by *Tuta absoluta*, ecology, geographic expansion and prospects for biological control. *Journal of Pest Science*, 83(3):197–215, 2010.

[13] EPPO. First report of *Tuta absoluta* from the United Kingdom. https://gd.eppo.int/reporting/article-340, Accessed February 2018.

[14] EPPO article-105. Tuta absoluta caught in a tomato packing station in the Netherlands. https://gd.eppo.int/reporting/article-105, 2009.

[15] Mária Ercsey-Ravasz, Zoltán Toroczkai, Zoltán Lakner, and József Baranyi. Complexity of the international agro-food trade network and its impact on food safety. *PloS one*, 7(5):e37810, 2012.

[16] FAO. Per capita tomato consumption. http://www.helgilibrary.com/indicators/tomato-consumption-per-capita/world/, Accessed in February 2018.

[17] RNC Guedes and MC Picanço. The tomato borer tuta absoluta in south america: pest status, management and insecticide resistance. *EPPO bulletin*, 42(2):211–216, 2012.

[18] RNC Guedes and HAA Siqueira. The tomato borer *Tuta absoluta*: insecticide resistance and control failure. *Plant Sciences Reviews 2012*, page 245, 2013.

[19] Ritter YA Guimapi, Samira A Mohamed, George O Okeyo, Frank T Ndjomatchoua, Sunday Ekesi, and Henri EZ Tonnang. Modeling the risk of invasion and spread of *Tuta absoluta* in Africa. *Ecological Complexity*, 28:77–93, 2016.

[20] MS Hossain, MY Mian, and R Muniappan. First record of tuta absoluta (lepidoptera: Gelechiidae) from bangladesh. *Journal of Agricultural and Urban Entomology*, 32(1):101–105, 2016.

[21] Philip E Hulme. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, 46(1):10–18, 2009.

[22] CM Kalleshwaraswamy, M Shankara Murthy, CA Viraktamath, and NK Krishna Kumar. Occurrence of tuta absoluta (lepidoptera: Gelechiidae) in the malnad and hyderabad-karnataka regions of karnataka, india. *Florida entomologist*, 98(3):970–971, 2015.

[23] Pablo Kaluza, Andrea Kölzsch, Michael T Gastner, and Bernd Blasius. The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48):1093–1103, 2010.

[24] O Karadjova, Z Ilieva, V Krumov, E Petrova, V Ventsislavov, et al. *Tuta absoluta* (Meyrick)(Lepidoptera: Gelechiidae): Potential for entry, establishment and spread in Bulgaria. *Bulgarian Journal of Agricultural Science*, 19(3):563–571, 2013.

[25] Somsack Kethonga, Khamtanh Thadavong, and Paule Moustier. Vegetable Marketing in Vientiane. (November):56, 2004.

[26] Ministry of Information and Communication Technology, Thailand. Agricultural census. http://web.nso.go.th/en/census/agricult/cen_agri03.htm, 2013.

[27] Paule MOUSTIER. Final summary report of SUSPER. *Sustainable Development of Peri-Urban Agriculture in South-East Asia*, (April):152, 2007.

[28] John F Hernandez Nopsa, Gregory J Daglish, David W Hagstrum, John F Leslie, Thomas W Phillips, Caterina Scoglio, Sara Thomas-Sharma, Gimme H Walter, and Karen A Garrett. Ecological networks

in stored grain: Key postharvest nodes for emerging pests, pathogens, and mycotoxins. *BioScience*, 70, page biv122, 2015.

[29] Export-Import Bank of India. Bangladesh: a Study of India's Trade and Investment Potential. 2015.

[30] Sevcan Oztemiz. Tuta absoluta povolny (lepidoptera: Gelechiidae), the exotic pest in turkey. *Romanian Journal of Biology*, 59:47–58, 2014.

[31] Dean R Paini, Andy W Sheppard, David C Cook, Paul J De Barro, Susan P Worner, and Matthew B Thomas. Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences*, 113(27):7575–7579, 2016.

[32] Dean R Paini, Susan P Worner, David C Cook, Paul J De Barro, and Matthew B Thomas. Threat of invasive pests from within national borders. *Nature Communications*, 1:115, 2010.

[33] Charles Perrings, Carlos Castillo-Chavez, Gerardo Chowell, Peter Daszak, Eli P Fenichel, David Finnoff, Richard D Horan, A Marm Kilpatrick, Ann P Kinzig, Nicolai V Kuminoff, et al. Merging economics and epidemiology to improve the prediction and management of infectious disease. *EcoHealth*, 11(4):464–475, 2014.

[34] Philippines Statistics Authority. Major vegetables and rootcrops quarterly bulletin. https://psa.gov.ph/content/major-vegetables-and-rootcrops-quarterly-bulletin, 2017.

[35] RPJ Potting, DJ Van Der Gaag, A Loomans, M Van der Straten, H Anderson, A MacLeod, JMG Castrillón, and GV Cambra. Tuta absoluta, tomato leaf miner moth or south american tomato moth. ministry of agriculture, nature and food quality. *Plant Protection Service of the Netherlands, Utrecht, The Netherlands*, 2013.

[36] Francois Rebaudo, Veronica Crespo-Perez, Jean-Francois Silvain, and Olivier Dangles. Agent-Based Modeling of Human-Induced Spread of Invasive Species in Agricultural Landscapes: Insights from the Potato Moth in Ecuador. *Journal of Artificial Societies and Social Simulation*, 14:1–22, 2011.

[37] Christelle Robinet, Charles-Edouard Imbert, Jérôme Rousselet, Daniel Sauvard, Jacques Garcia, Francis Goussard, and Alain Roques. Human-mediated long-distance jumps of the pine processionary moth in Europe. *Biological invasions*, 14(8):1557–1569, 2012.

[38] Christelle Robinet, Hella Kehlenbeck, Darren J Kriticos, Richard HA Baker, Andrea Battisti, Sarah Brunel, Maxime Dupin, Dominic Eyre, Massimo Faccoli, Zhenya Ilieva, et al. A suite of models to support the quantitative assessment of spread in pest risk analysis. *PLoS One*, 7(10):e43366, 2012.

[39] Christelle Robinet, Alain Roques, Hongyang Pan, Guofei Fang, Jianren Ye, Yanzhuo Zhang, and Jianghua Sun. Role of human-mediated dispersal in the spread of the pinewood nematode in China. *PLoS One*, 4(2):e4646, 2009.

[40] Robyn Magalit Rodriguez. Philippine migrant workers' transnationalism in the middle east. *International Labor and Working-Class History*, 79(1):48–61, 2011.

[41] E. Sankarganesh, D.M. Firake, B. Sharma, V.K. Verma, and G.T. Behere. Invasion of the South American Tomato Pinworm, Tuta absoluta, in northeastern India: a new challenge and biosecurity concerns. *Entomologia Generalis*, 36(4):335–345, 2017.

[42] Simplemaps. Cities and their population. https://simplemaps.com/data/world-cities, Accessed in February 2018.

[43] V Sridhar, AK Chakravarthy, and R Asokan. New record of the invasive South American tomato leaf miner, *Tuta absoluta* (Meyrick)(Lepidoptera: Gelechiidae) in India. *Pest Management in Horticultural Ecosystems*, 20(2):148–154, 2014.

[44] Sweta Sutrave, Caterina Scoglio, Scott A Isard, JM Shawn Hutchinson, and Karen A Garrett. Identifying highly connected counties compensates for resource limitations when evaluating national spread of an invasive pathogen. *PLoS One*, 7(6):e37793, 2012.

[45] Serigne Elhadji Sylla, Thierry Br\'{e}vault, Lucie Monticelli, Karamoko Diarra, and Nicolas Desneux. Geographic variation of host preference by the invasive tomato leafminer Tuta absoluta : implications for host range expansion. Technical report, 2018.

[46] Henri EZ Tonnang, Samira F Mohamed, Fathiya Khamis, and Sunday Ekesi. Identification and risk assessment for worldwide invasion and spread of *Tuta absoluta* with a focus on Sub-Saharan Africa: implications for phytosanitary measures and management. *PloS one*, 10(8):e0135283, 2015.

[47] USDA. New Pest Response Guidelines: Tomato Leafminer (*Tuta absoluta*). *Animal and Plant Health Inspection Service, Plant Protection and Quarantine*, 2012.

[48] Robert C. Venette, Darren J. Kriticos, Roger D. Magarey, Frank H. Koch, Richard H. a. Baker, Susan P. Worner, Nadilia N. Gómez Raboteaux, Daniel W. McKenney, Erhard J. Dobesberger, Denys Yemshanov, Paul J. De Barro, William D. Hutchison, Glenn Fowler, Tom M. Kalaris, and John Pedlar. Pest Risk Maps for Invasive Alien Species: A Roadmap for Improvement. *BioScience*, 60(5):349–362, 2010.

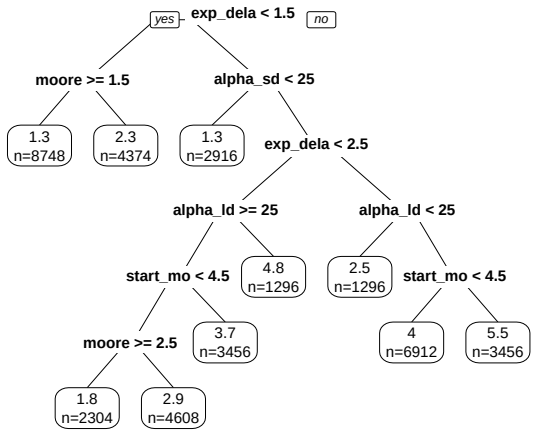
[49] Srinivasan Venkatramanan, Sichao Wu, Bowen Shi, Achla Marathe, Madhav Marathe, Stephen Eubank, Lalit P Sah, AP Giri, Luke A Colavito, KS Nitin, et al. Towards robust models of food flows and their role in invasive species spread. In *Big Data (Big Data), 2017 IEEE International Conference on*, pages 435–444. IEEE, 2017.

[50] Tran Duc Vien. Overview on tomato production and tomato varieties in vietnam. *Collection of Articles in Center for Agricultural Research and Ecological Studies. Hanoi University of Agriculture*, 2006.

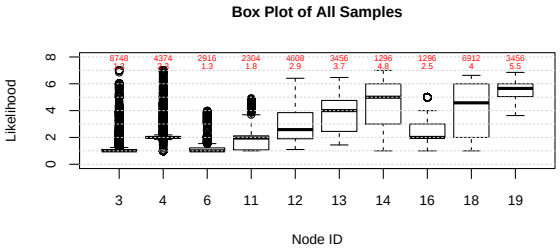
[51] Siebe Van Wijk. The market for vegetables in North Vietnam.

[52] Wan Yang, Donald R. Olson, and Jeffrey Shaman. Forecasting Influenza Outbreaks in Boroughs and Neighborhoods of New York City. *PLoS Computational Biology*, 12(11):1–19, 2016.

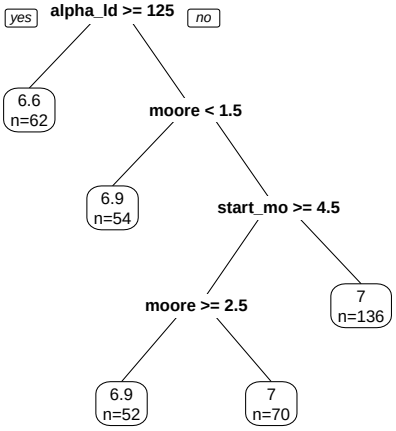
[53] L. You, U. Wood-Sichra, S. Fritz, Z. Guo, L. See, and J. Koo. Spatial production allocation model (spam) 2005 v3.2. http://mapspam.info, 2017.



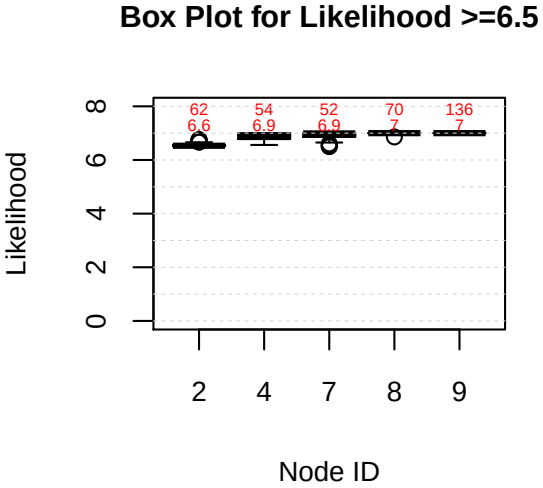
(a)



(b)



(c)



(d)

Figure 3: CART analysis of the parameter space. There are two regimes that can be observed which provide better fit.