Segregation Landscape: A New View on the Schelling Segregation Space

Philippe Collard
Univ. Nice Sophia Antipolis, CNRS, I3S Laboratory, UMR 7271
UNS: Parc Valrose, 06108 Nice, France
philippe.collard@unice.fr

Teodor Ghetiu
Wind River Systems
teodor.ghetiu@windriver.com

Thomas C. Schelling showed that global aggregation may occur, even if it does not correspond to agent preferences; thus, to some extent his model supported the view that segregation is unavoidable, whatever the tolerance is. The segregation landscape approach proposed in this paper is seriously weakening this hypothesis; here, we radically change the perspective and propose using the landscape metaphor to represent emergent segregated communities. A segregation landscape is a mapping from situated individuals into an extra dimension that represents the degree of segregation of everyone. This paper uncovers how to interpret hills and valleys, and whether these interpretations are congruent with the intuitive notion of frontier. Such a representation allows us to describe both the static properties of a segregation space and their impact on how information propagates between segregated communities. In order to assess the explanatory power of the landscape metaphor, we devise agent-based simulations. First, we establish the link between the micro-level quantified by individual tolerance and the macro-structure represented by the landscape, then we show how "geographic" properties impact the dynamical behavior on such a population landscape.

1. Introduction

The general context of this paper is that of modeling spatial segregation in order to understand, as Schelling said, how "people who differ conspicuously in binary groups—e.g. blacks and whites, males and females […]—get separated spatially, in residence, in dining halls, at public events" [1]. Tolerance determines the local level of individual behavior, while aggregation characterizes the resulting global level of the entire population. Individuality and aggregation constitute important topics in complex systems research, as they represent facets of emergence, a core concept deemed a "central and constructive player in our understanding of the natural world" [2]. While much research

https://doi.org/10.25088/ComplexSystems.25.3.169

has looked at social phenomena through a two-dimensional perspective, this paper makes use of the landscape metaphor to represent the degree of segregation undergone by each individual.

This metaphor has been essayed in different forms during recent years [3–5]. The motivation to use it is that three-dimensional landscapes provide a richer mapping between the static properties of segregated populations and geographical concepts that everyone intuitively understands. In addition, although it is assumed that individuals are equivalent, the metaphor also provides insights into the different roles that each one has over the emerging social aggregates. In essence, we address the following scientific question, "when aggregates emerge from local segregationist behaviors, do individuals really play an equal role?"

As people get together, a concept of particular interest is that of the frontier. In order to recognize the unity of an aggregate of individuals, one must be able to distinguish between its interior and its exterior; this is where the concept of frontier becomes relevant. The classical scenario sees frontiers as geographical boundaries between two contiguous territorial systems (e.g., river, mountain range). The primary role of a frontier is to separate (e.g., a defense system is aimed at keeping enemies away); however, absolute separation is an ideal and, in reality, this is complemented by exchange (e.g., in the form of information flow). Frontiers can be described in terms of their varying separation–exchange tradeoff.

The segregation landscape metaphor we propose allows us to describe both the static properties of the system and their influence on how information propagates between segregated communities. One fundamental issue is to link these two points of view by determining how "geographic" properties rely on original properties and how they influence the dynamical process of information flow.

The paper is structured as follows. In Section 2, we present a simple model of spatial segregation that will serve as a framework for our study. Section 3 defines the concepts of *segregation landscape*, *frontier*, and *segregation walk*; in addition, we identify the basic properties of such a landscape. Section 4 provides experimental results; we use a multiagent environment to study the influence of the level of tolerance on landscape and frontier features. In Section 5 we offer our conclusion and hints for future research.

2. A Generic Model of Segregation

The generic model of segregation we use was first proposed in [6]. We consider a set of *agents* immersed in a world composed of spatial *locations*. Apart from its own location, an agent is characterized by its *type* attribute, which remains constant over time. All agents with the same type are part of the same community.

2.1 Proximity Network versus Agent Network

The perception of an agent spans only the set of locations adjacent to it, which represent its local neighborhood. A location is either vacant or occupied by one agent. Let V be the set of vacant locations and A the set of agents; as there is up to one agent per location, the *density* of agents is the ratio

$$\delta = \frac{|A|}{|A| + |V|}.$$

The proximity network defines the arrangement or interconnectivity of locations. Two locations are connected by an *edge* if they are neighbors. The degree of a location node, that is, the number of edges that connect it to other locations, is named its proximity degree.

Overlying the proximity network is the agent network. Its nodes are the agents, and two agents are in direct contact if they are located in adjacent locations; a proximity network edge can be seen as a shared communication channel. While the proximity network is static, as agents can move, the agent network evolves over time. We call *agent degree* (denoted as $\mathrm{ad}_a(t)$) the number of connections the agent network a has to other agent networks at time t; it is a measure of the local influence of the agent within the agent network.

2.2 Tolerance and Individual's Satisfaction

We assume that the level of *tolerance* is shared by all agents and is represented by a constant floating number τ in the range $[0..1]$. The state of an agent, satisfied or unsatisfied, is time dependent; it depends on the tolerance and on its own type and the type of its neighbors. For each agent a at time t, the Boolean indicator *satisfied* is defined as $\mathrm{satisfied}_a(t) = \left(\left(1 - s_a(t)\right) \leq \tau\right)$, where $s_a(t)$ is the ratio of the number of neighbors with similar type to the agent degree. (Let us note that obviously an isolated agent is satisfied.) So the tolerance τ denotes the threshold under which an agent is satisfied. A tolerance of 0.5 means that each agent accepts at most half of its neighbors to be different from itself. The agents are said to be intolerant if $\tau \ll 0.5$ and rather tolerant if $\tau \gg 0.5$.

2.3 Micromotive versus Macrobehavior

Agent behavior is oriented toward achieving and maintaining satisfaction: an unsatisfied agent is motivated to move toward another location, whereas a satisfied one has no incentive to move. We assume that an unsatisfied agent uses the eulogy to fleeing rule [7] to find a new place: *a location is randomly chosen from the world and the agent moves into it if and only if the location is vacant*. If all the agents are satisfied at the same time, the eulogy to fleeing rule has no effect, and the system has reached a fixed-point configuration. In this paper, we do not discuss the conditions that guarantee that the system converges toward equilibrium; we select system conditions in which equilibrium will be reached. The parameters of the segregation model are summarized in Table 2 (columns 1 and 2).

2.4 Schelling-like Model

Schelling's checkerboard model of residential segregation has become one of the most cited and studied models in many domains, such as economics and sociology [8–12]. It is also one of the predecessors of agent-based computer models [13].

In Schelling's initial work [14], the starting point is that the individual's satisfaction depends on tolerance and on the size and the social composition of the neighborhood. Global aggregation is measured as the average over the entire population of local information [15–17] where, for each individual a_i, the information is the measure of *similarity* s_i regarding its neighbors; let us note that such a metric is closely related to the *dissimilarity index* used in the demographic literature [18]. Schelling showed that global aggregation may occur even if it does not correspond to agent preferences, that is, even if tolerance is high; thus, to some extent his model supported the assumption that segregation is unavoidable whatever the tolerance is.

In [6], the authors show that the Schelling model can be viewed as an instance of the generic model of segregation. One only needs to assume that the proximity network is a two-dimensional regular grid and the neighborhood of an agent is composed of the eight nearest cells surrounding it (for each location, the proximity degree is 8). In addition, the eulogy to fleeing rule has already been used within Schelling's models, leading toward equilibrium states where correlation between tolerance and global aggregation underlined by Schelling is confirmed [7].

3. Segregation Landscape

In evolutionary biology and combinatorial optimization, a *fitness landscape* is a way of visualizing a problem [3, 5, 19]. In such a landscape, the altitude represents the fitness; there are peaks, and the highest peak is the best solution regarding fitness adaptation or optimiza-

https://doi.org/10.25088/ComplexSystems.25.3.169

tion. Depending on the number of peaks, fitness landscapes can be described as rugged or smooth.

In sociology or geography, the usual way of visualizing the formation of communities is to look at the two-dimensional space in which the agents move. In this paper, we propose to supplement this view by the more informative segregation landscape. The segregation landscape provides a third dimension that expands the two-dimensional human and physical geography; such a dimension represents for each agent its *distance* to the opposite community, meaning that to an extent, the altitude represents the degree of segregation. This approach differs from the classical definition of segregation based on information located in the vicinity of each individual regarding its type and the type of its neighbors [15–17].

In the following, in order to specify the metaphor of segregation landscape, we introduce the concepts of segregation index, segregation shape, and frontier.

3.1 Segregation Index

For each agent a_i in one community C, the *segregation index* σ is its distance to the opposite community $\overline{C}$:

$$\sigma(a_i) = \min_j \left\{ d\left(a_i, a_j\right) \,\middle|\, a_j \in \overline{C} \right\}, \tag{1}$$

where d is the Euclidean distance between two agents. Intuitively, peaks will correspond to high σ values, while coasts will correspond to low values.

In formal terms, the segregation landscape consists of three ingredients: (*i*) the set of agents A; (*ii*) a notion of neighborhood represented by the grid proximity network; and (*iii*) the segregation function σ:

$$\mathrm{SL} = \left\{A, \mathrm{gPN}, \sigma\right\}. \tag{2}$$

Let us note that whenever an agent moves, this affects not only the neighborhood it leaves, but also the one it arrives in, and consequently, the segregation index of all agents with which the agent "interacts" may change. This implies that as long as there are unsatisfied individuals, the segregation landscape may change over time. In the following, we will look at the landscape when all the agents have become satisfied.

3.2 Segregation Shape

Maintaining the landscape metaphor, we match recognizable emergent shapes to concepts such as peaks, valleys, contour lines, coasts, and watersheds (Table 1).

https://doi.org/10.25088/ComplexSystems.25.3.169

Geography	Segregation Landscape
peak	most segregated agents
valley	agents for which all the neighbors have a higher segregation index
contour lines	agents with almost equal segregation index
coast	agents closest to the opposite community
watershed	agents for which the coast of the neighbors is far away from their own coast

Table 1. Mapping between geography and segregation landscape.

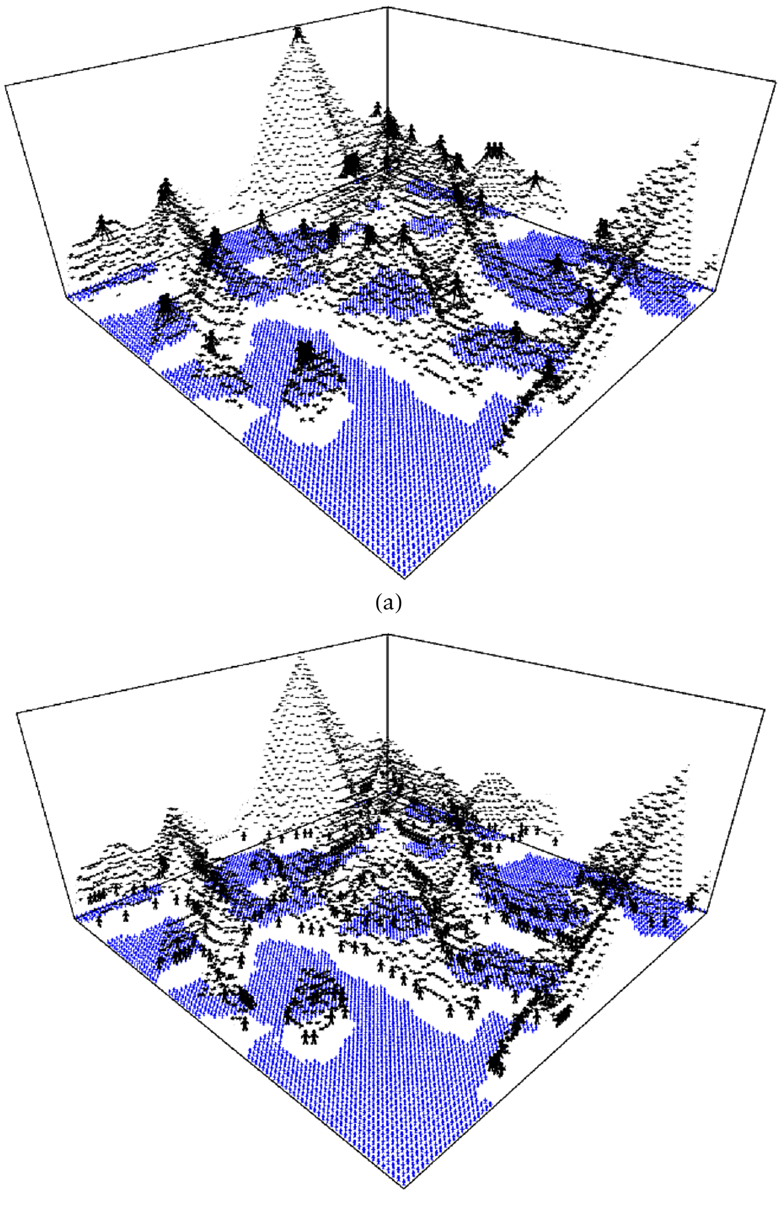
This way, they gain the following semantic:

1. *Peaks* are the agents most distant from the opposite community; that is, the agents for which all the neighbors have a lower segregation index.
2. *Valleys* are the agents for which all the neighbors have a higher segregation index; from an optimization point of view, peaks and valleys are local optima.
3. *Contour lines* are sets of agents with equal or almost equal segregation index.
4. *Coasts* are the agents closest to the opposite community and so the less segregated; to some extent, they foreshadow the proximity of the frontier between the two communities.
5. *Watershed* contains agents for which the coast of the neighbors may be far away from their own coast.

Let us note that with the classical two-dimensional view, there is no structured organization inside a community, and everyone is playing the same role, so, in a way, from a landscape perspective each community stands on a large plateau. As a consequence, two-dimensional models do not match the reality very well; for instance, two individuals that are located next to each other, on both sides of a watershed, are equivalent, whereas, if we have to take into account the way information propagates to the opposite community, their behaviors differ in a significant way.

Figures 1 and 2 illustrate the geographic features on one particular segregation landscape generated using an agent-based simulation. To facilitate understanding, the third dimension allows us to show only one community (the other agents remain on the ground). Each figure illustrates only one feature, and the corresponding agents are plotted as "person."

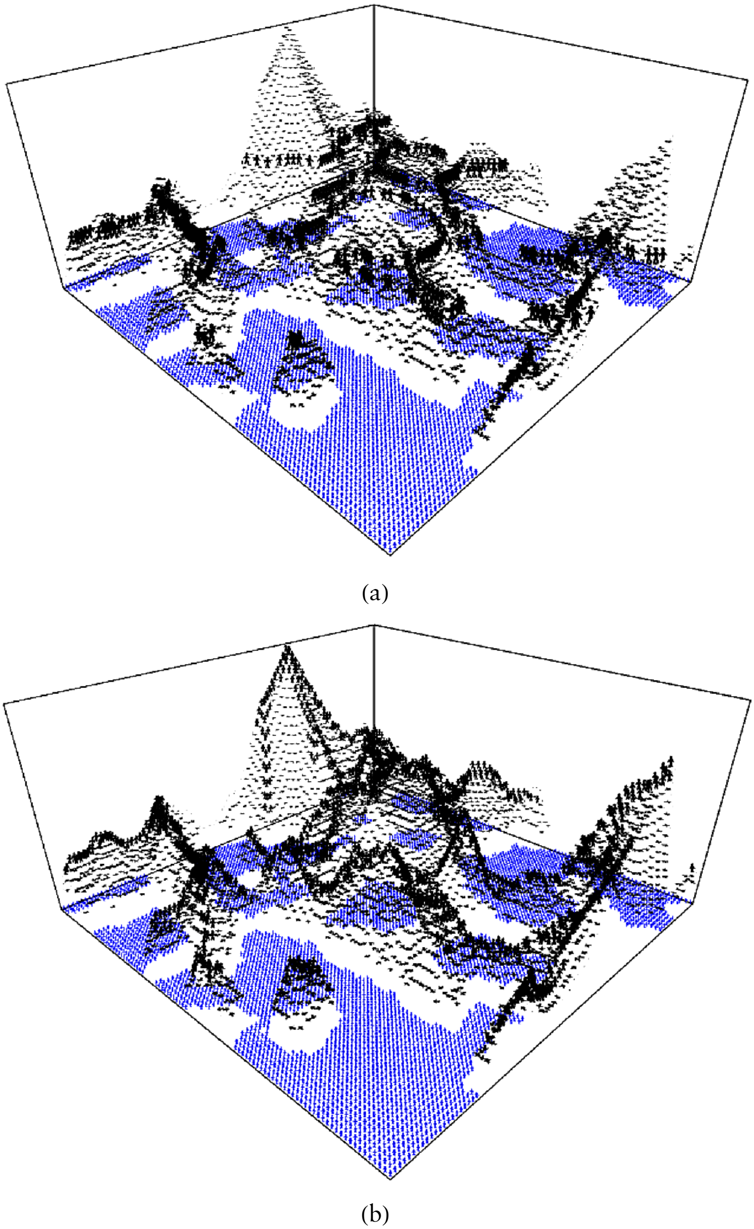
https://doi.org/10.25088/ComplexSystems.25.3.169



(a)

(b)

Figure 1. Geography vs. segregation landscape: (a) peaks and (b) valleys.



(a)

(b)

Figure 2. Geography vs. segregation landscape: (a) contour lines and (b) watershed.

3.3 Frontier

A frontier is a generic concept that has different instantiations, depending on the context in which it is considered. A common class of frontiers is found in the geographical domain where they appear as fronts.

3.3.1 Specifics of This Paper Regarding the Concept of Frontier

Before presenting our approach, it is necessary to clarify its position beside our previous works. In [6], we revisited the concept of frontier in the context of the Schelling model, now seen as the set of locations where contact occurs between two agents of opposite types. We considered two kinds of contact: direct contact refers to agents being directly linked in the agent network, whereas indirect contact is mediated through a vacant location. Two types of proximity networks were considered: grid and scale free. The frontier represents structure that both determines the borderland between two aggregates of opposite types and allows communication between them.

The present study differs from this previous effort in the following aspects: (*i*) the proximity network is a grid only; (*ii*) we consider direct contacts only; (*iii*) the frontier is defined according to the segregation index; (*iv*) we consider the likelihood that a signal emitting from one agent crosses the frontier and reaches the other community; and (*v*) we underline the fact that agents on the frontier are not all equivalent.

3.3.2 Definitions

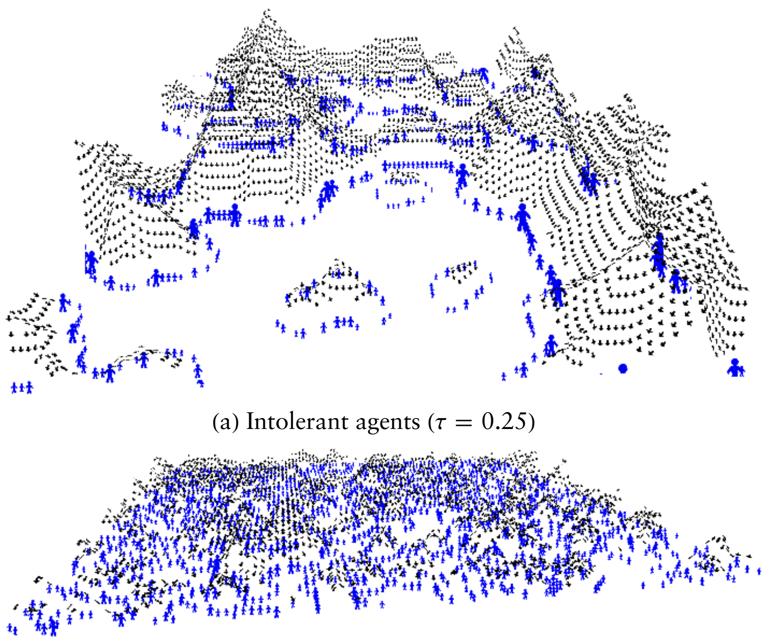
The above points lead to the following definitions.

- *Referrer*. One agent r is a referrer if and only if there is at least one other agent a in the opposite community, such that $d(a, r) = \sigma(a)$. We will denote as $R_C(t)$ the subset of $\overline{C}$ of all the referrers for community C at time t. Let us note that one agent can have more than one referrer, so the set R_a of referrers of an agent a in community C is $\operatorname{argmin}_{r \in \overline{C}} d(a, r)$.
- *Basin of attraction*. The basin of attraction of a referrer r in community C is the set of agents a in the opposite community $\overline{C}$ for which r is a referrer: $B(r) = \left\{a \in \overline{C} \,\middle|\, r \in R_a\right\}$.
- *Scope*. The scope of a referrer is the size of its basin of attraction: $\text{scope}(r) = |B(r)|$.
- *Gate*. A referrer r is said to be a gate if there is at least one other agent a in the opposite community such that $d(a, r) = 1$; in such a case, there is a direct contact between r and one agent in the other community; under these circumstances, the agent a is positioned at a global minimum in the segregation landscape. We will refer to the set of all the gates for community C at time t as $G_C(t)$; let us note that $G_C \subset R_C \subset \overline{C}$.

- *Frontier*. Assuming that all the individuals have become satisfied, we define the frontier F_C of a community C as the set of referrers in the opposite community: $F_C = R_C(\infty)$.

3.4 Overview

Figure 3 provides qualitative views on two extreme segregation landscapes where, in both cases, all the agents have become satisfied. In order to facilitate comprehension and to avoid confusion between the two communities, we chose to represent all the agents in only one community (see small "arrows" in Figure 3), whereas from the other community we depicted only the referrers—the size of a referrer is proportional to its scope (see "persons" in Figure 3). A key point is that the segregation landscape metaphor shows that agents on the frontier are not all equivalent: there are agents with a high scope, while others have only small basins of attraction. As we will detail in Section 4, the former play a central role for disseminating information between communities.



(a) Intolerant agents ($\tau = 0.25$)

(b) Tolerant agents ($\tau = 0.63$)

Figure 3. Segregation landscape: one community ("arrow") and its frontier ("person"); $\delta = 90\%$.

https://doi.org/10.25088/ComplexSystems.25.3.169

3.5 How to Characterize a Segregation Landscape

To enable consistent comparisons between various kinds of segregation landscapes, we must take into account both static and dynamic properties. In the following, we define what we mean by global segregation index, roughness, frontier size and thickness, scope distribution, and penetrability.

3.5.1 Global Segregation Index

First of all, we define the global segregation index $\overline{\sigma}$ as the local segregation index averaged over all the population.

3.5.2 Landscape Ruggedness

To synthesize landscapes with tunable ruggedness, Kauffman [20] proposed the *NK* models, where the ruggedness increases with parameter *K* from a single-peaked landscape to a multi-peaked landscape: walks to local optima become shorter as *K* increases [21]. In the following, we will show that in our model tolerance plays a role similar to that of the *K* parameter. Although ruggedness of a three-dimensional landscape is an intuitive notion, we use several quantitative measures to probe alternative facets of this characteristic.

- *Local roughness*. We define the local roughness of a segregation landscape as the root mean squared difference between the segregation index of an agent and that of its neighbors, averaged over all the population [22]:

$$\mathrm{rg(SL)} = \frac{1}{|A|}\sum_i \sqrt{\frac{1}{\mathrm{ad}_i}\left(\sum_{j=1}^{\mathrm{ad}_i} \left(\sigma(a_i) - \sigma\left(a_j\right)\right)^2\right)}, \tag{3}$$

  where a_j is a neighbor of a_i. As defined, the roughness of a smooth landscape will be small, while that of a rugged landscape will be high.

- *Minima ratio*. This ratio corresponds to the fraction of minima in the landscape. According to the segregation index σ, we will denote as $C_\sigma^{\min}$ the set of local minima in the community C.

- *Peak ratio*. The ratio of peaks is the number of agents with no fitter neighbors divided by the total number of agents.

3.5.3 Frontier

The frontier can be characterized according to the following aspects:

- *Relative size*. We define the relative size of the frontier F as the number of referrers divided by the total number of agents:

$$\mathrm{rs}(F) = \frac{2|F|}{|A|}. \tag{4}$$

https://doi.org/10.25088/ComplexSystems.25.3.169

- *Thickness*. The thickness of the frontier is the mean distance to the opposite community over all the referrers:

$$\mathrm{th}(F) = \frac{\sum_i^{|F|} \sigma(r_i)}{|F|}, \tag{5}$$

where $r_i \in F$. Let us note that the lowest value for thickness is 1 and would be reached only in the situation in which each referrer is a gate.

- *Scope distribution*. As they differ by their scope, the referrers are not all equivalent: while some have a big basin of attraction, others refer only to a few agents in the opposite community. So the scope distribution can give some insight into such a differentiation.

3.5.4 Penetrability

The notion of penetrability has been used in psychology, physics, and geography. It is the quality of being penetrable by people, light, information, and so on [23–25]. One of the most interesting and still not completely understood phenomena happening in social networks is their ability to spread units of information (e.g., rumors or cultural contents); in the context of segregation, the question is whether a signal emitting from one community can reach the other one. We define penetrability as the likelihood that a signal emitting from one agent reaches the edge of the frontier, then crosses the frontier and finally reaches the other community.

In order to pass through the frontier, we assume that information must follow a segregation walk in the basin of attraction of a gate [4]. We name segregation walk any process by which a piece of information passes from agent to agent, depending on segregation. Such a process looks like water from rain descending the line of minimum gradient in a geographical landscape. The information crosses the landscape, each step being assumed to lead to an improvement of the system against segregation. Assuming that all the agents are satisfied, their behavior is based on the act of passing information closer and closer to the frontier; by continuing this process over many iterations, the information will eventually end up on a local minima from a segregation index perspective. Let us note that we assume that each agent knows which of its neighbors, if any, are the less segregated.

We define the theoretical penetrability p of the segregation landscape as the ratio between the sum of all gate scopes and the number of agents:

$$p(\mathrm{SL}) = \frac{\sum_G \mathrm{scope}(g)}{|A|}, \tag{6}$$

where g is a gate and scope(g) is its scope. Obviously, this measure depends both on the number of gates and on the scope distribution for the gates.

4. Simulation and Experimental Results

Experiments are performed via the *NetLogo* multi-agent programmable environment [26]. The pseudocode for simulating the model is defined in Algorithm 1. All the qualitative results we present are averaged over 100 runs. Simulations are performed on an $L \times L$ lattice of locations, with L set to 100. The grid is a toroid where the top and bottom edges, as well the left and right edges, are connected to each other. We set the density of agents to 90%, which is a standard value for which the Schelling model converges. Agents are positioned in a random initial configuration, such that the vacant locations and the two types of agents are well mixed.

Name	Parameter	Value
N	world size	10000
δ	agent density	]0;1]
τ	tolerance	[0;1]
$\|type\|$	type number	2
PN	proximity network	grid
AN	agent network	-

Table 2. Simulation: global parameters.

Name	Parameter
type	the type
$(xcor, ycor)$	the spatial position
satisfied?	the satisfaction
referrer	the referrer
σ	the segregation index

Table 3. Simulation: agent parameters.

```
procedure main
    setup
    while not (all the agents are satisfied) do
        for each agent a do
            if not satisfied?(a) then
                choose a node-location at random
                if the location is vacant then
                    a moves to this location
                end if
            end if
        end for
        for each agent a do
```

```
            computeSatisfaction(a)
        end for
    end while
end procedure
Ensure: All the agents are satisfied

procedure setup
    density ← δ, tolerance ← τ
    create N×δ agents
    fix the type: N/2 with type 1 and N/2 with type 2
    position at random the agents on the grid
    for each agent a do
        computeSatisfaction(a)
    end for
end procedure

procedure computeSatisfaction(a)
    cN ← count neighbors of a
    cNd ← count neighbors with [type != type(a)]
    if cN != 0 then
        satisfied?(a) ← (cNd/cN ≤ tolerance)
    else
        satisfied?(a) ← true
    end if
end procedure
```

Algorithm 1. Simulation of the simple model of segregation.

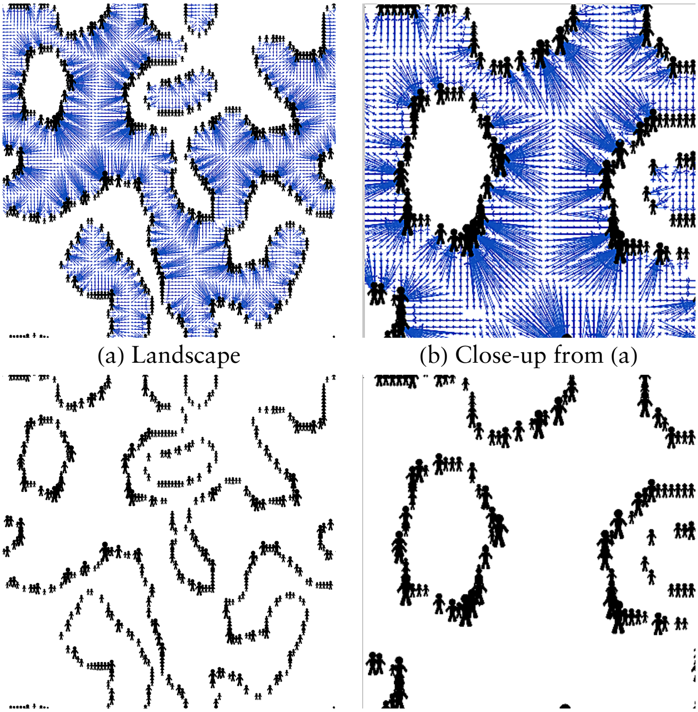
4.1 Geography versus Tolerance

4.1.1 Observation

Overhead views. To provide an overall view of the landscape, we consider two representative runs with two extreme cases for the tolerance. Obviously, we examine the landscape at the end of the process, when all the agents are satisfied. (To ensure convergence, values for the tolerance are above 0.20.) In Figure 4, the tolerance is set to 0.25, while it is 0.63 in Figure 5: in both cases, the top represents the landscape, while the bottom represents the frontier only. Representing links between each agent and its referrer allows us to see the variety of referrers' basins of attraction.

For intolerant agents, the dynamics lead to the emergence of spatial homogeneous patterns (Figure 4(a, b)) isolated by a no-man's-land of vacant nodes (Figure 4(c, d)), but as tolerance increases, communities fragment further (Figure 5(a, b)), and we observe that the smooth shape becomes more complex, as in a real landscape when roughness dictates many meanders to the edge of a lake (Figure 5(c, d)).

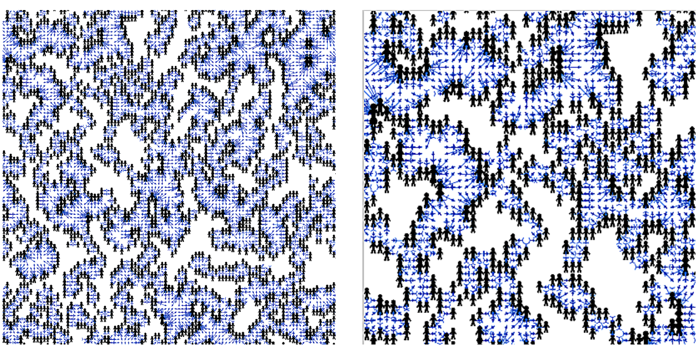
https://doi.org/10.25088/ComplexSystems.25.3.169



(a) Landscape (b) Close-up from (a)

(c) Frontier (d) Close-up from (c)

Figure 4. Segregation landscape: view from above. Intolerant agents ($\tau = 0.25$).



(a) Landscape (b) Close-up from (a)

Figure 5. (*continues*).

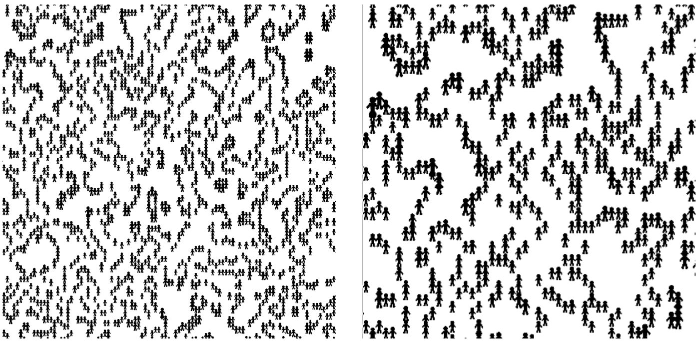


Figure 5. Segregation landscape: view from above. Tolerant agents ($\tau = 0.63$).

Global segregation. Figure 6 shows the value reached at convergence by the global segregation index, according to the level of tolerance. The graph represents a decreasing "S"-shaped curve, whose inflection point roughly coincides with neutral tolerance. As tolerance increases toward its maximum, the global segregation index decreases toward its minimum value 1, which means that the two communities are fully mixed. In addition, let us note that the variance decreases with tolerance; this confirms that intolerant micromotives lead to the emergence of large patterns of individuals belonging to the same community (Figure 4(a)).

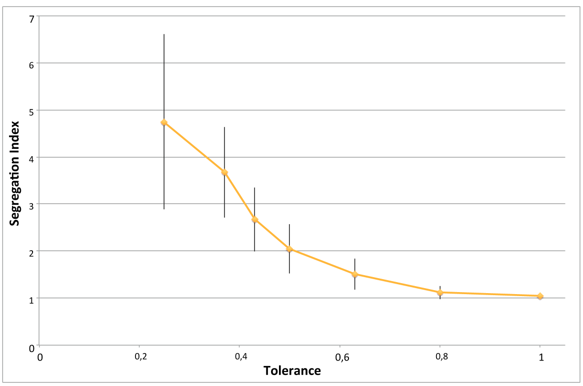


Figure 6. Global segregation vs. tolerance.

Roughness. Figure 7 shows that: (*i*) roughness (equation (3)) increases with tolerance—with a particularly rapid rate of growth between 0.04 and 0.198 when τ increases from 0.25 to 0.5: roughness reaches its high values as soon as tolerance is above neutrality; and

https://doi.org/10.25088/ComplexSystems.25.3.169

(*ii*) the number of extrema in the landscape actually increases with tolerance—it follows an S-curve between two extreme values: for intolerant agents, there are less than 10% of minima and 1% of peaks, while with tolerant agents, almost 90% of them are local minima and 70% are peaks.

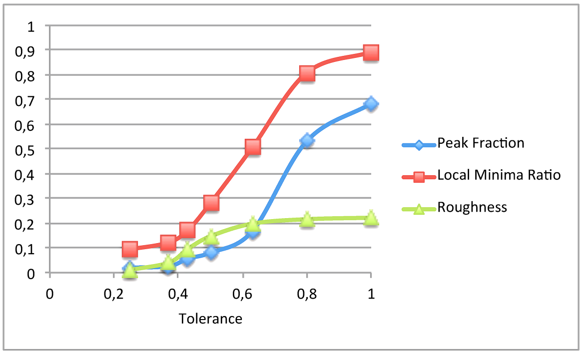


Figure 7. Roughness vs. tolerance.

4.1.2 Discussion

Qualitative observations of the segregation indexes for the same community show that individuals are not all equal with regard to segregation. As tolerance increases, the global segregation index decreases, whereas the shape of the landscape becomes more rough and the frontier more complex. Knowledge of this "geography" is important as long as the population is subject to segregation optimization: the form determines the likelihood of reaching other agents more or less segregated. So tolerance is a key parameter, and its variation determines the roughness; tunable ruggedness captures here the intuition that the number of local "hills and valleys" can be adjusted via changes in tolerance; the geography depends on what people's tolerance is: few "hills" with large "valleys" for intolerance and strong ruggedness for tolerance. Nevertheless, quantitative measures show that tolerance significantly impacts roughness for values below 0.5 only.

4.2 Frontier versus Tolerance

4.2.1 Observation

Frontier size and thickness. Figure 8 gives quantitative results on the shape of the frontier: as tolerance increases, we observe that the size increases, while the thickness decreases. In addition, the size reaches a plateau (~62% of the total population) as soon as the tolerance is above 0.8, whereas the thickness reaches a plateau $\left(t \simeq 1\right)$ as soon as the tolerance is above 0.5.

https://doi.org/10.25088/ComplexSystems.25.3.169

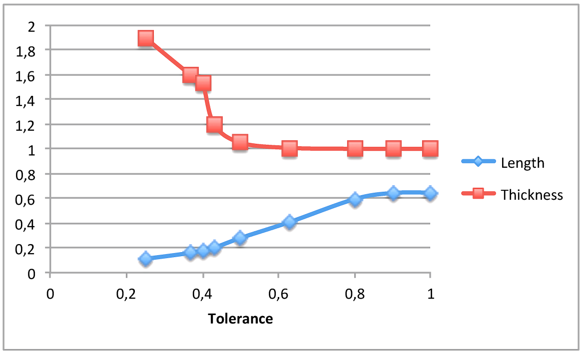
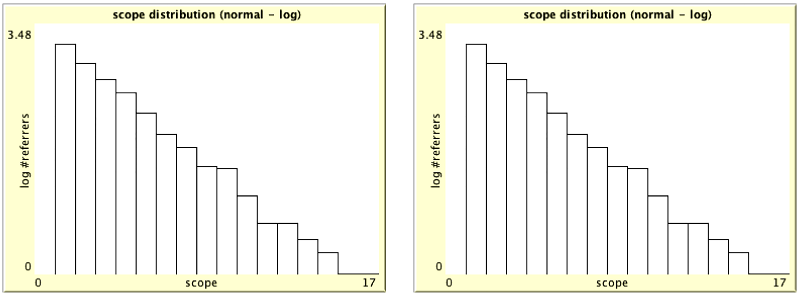


Figure 8. Frontier in a segregation landscape: size and thickness vs. tolerance.

Frontier scope distribution. Figure 9 shows the frontier scope distribution for intolerant and tolerant individuals. For intolerant agents, the log–log plot (Figure 9) shows a reasonably consistent linear curve, which reveals that the distribution approximately follows a power law with a long tail. For tolerant agents, plotting the log of the number of referrers with a given basin size vs. the basin size yields a straight line (Figure 9), so the distribution decrease is exponential.



(a) log–log axis ($\tau = 0.25$) (b) normal–log axis ($\tau = 0.63$)

Figure 9. Scope distribution power law decrease vs. exponential decrease.

4.2.2 Discussion

Qualitative observations show that the complexity of contours increases as agents become more and more tolerant. Intolerant agents lead to a short no-man's-land frontier; such an impervious structure with a high thickness and low size limits the possibility of direct communication between opposite agents (Figure 4). In the opposite case, if agents are tolerant, thickness reaches its minima and size is maximum; the frontier looks like a Peano line (Figure 5).

https://doi.org/10.25088/ComplexSystems.25.3.169

In [27], Goles et al. proposed a criterion called *energy*, by analogy with spin-glass systems, that represents the full states of all the agents:

$$E = -\frac{1}{2} \sum_{a_i \in A} e_i \times \sum_{a_j \in N_i} e_j, \tag{7}$$

where N_i is the neighborhood of agent a_i, and $\forall \left(a_i, a_j\right) \in A^2$, $e_i e_j = +1$ if a_i and a_j are in the same community and -1 elsewhere. The authors show that as the system converges, the energy decreases until reaching a level that is linearly correlated to the perimeter of the interface between the two communities. So it is interesting to know whether or not this is still the case with our definition of frontier. To test the statistical significance of the association between energy (equation (7)) and size (equation (4)), we conduct a regression analysis. (Experiments' values are averaged over 100 runs.) We establish the following regression equation for size on energy: $\mathrm{rs}_\tau \simeq \mathrm{rs}_1 + 0.1414 \times E_\tau$, where the τ subscript refers to tolerance, comes with a correlation coefficient close to 0.995. (If all the agents are maximally tolerant, i.e., $\tau = 1$, the energy E_1 is close to 0.) This result appears to confirm that the landscape metaphor is well suited to determine whether or not an agent stands on the interface between the two communities.

Experimental results show that the (relative) size is not in itself enough to characterize the frontier well; to be clearly understood, such a "complex object" requires deeper analysis. Indeed, we have to take into account: (*i*) the relative importance of the referrers; and (*ii*) its thickness.

Each referrer distinguishes itself by its basin of attraction. The distribution of the size of the basin of attraction is highly reliant on the level of tolerance: for intolerant agents the number of referrers having a certain scope is found to decrease as a power of the size of the basin; it would look similar to an exponential decay, but the tail does not decay as quickly, leaving very large basins still being possible: there are large numbers of referrers with low scope and nevertheless some with high scope. As a consequence, there is not a characteristic value for the distribution. For tolerant agents, the distribution decrease is exponential: there are many referrers with a small basin of attraction and very few with a large basin. Finally, thickness must also be taken into consideration, and we observe that tolerance significantly impacts this measure only for values below 0.5.

4.3 Penetrability versus Tolerance

Let us remember that we have defined penetrability as the likelihood for information spreading from one agent to reach the edge of the frontier and then to pass through the frontier. With the aim of predicting what should happen, equation (6) proposes a theoretical definition of penetrability based on the size of the basins of attraction for the gates. On the other hand, experimental probability will be a result of trials that test predictions, and we must do a sufficient number of dynamic penetrability tests to see how often information crosses the frontier.

Theoretical penetrability. Figure 10 shows that both the theoretical penetrability (equation (6)) and the number of gates increase with tolerance. As tolerance increases, the number of gates roughly follows a steady growth from 0.02 to 0.64. As tolerance increases from 0.25 to 0.5, penetrability quickly increases from 0.39 to 1, then a high plateau with the maximum value of 1 is reached. The difference in the increase can be explained by the scope distribution, which follows, according to the tolerance, either a power law or an exponential decay. All this indicates that the number of gates itself is not sufficient to explain penetrability well, and we really have to also take into account the scope of the gates. In other words, as tolerance increases, there is not only a quantitative, but also a fundamentally qualitative change for the gates.

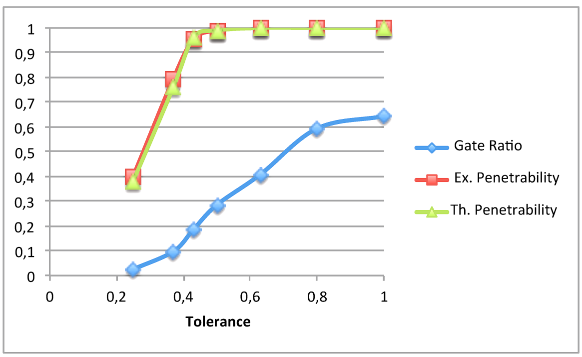


Figure 10. Penetrability and gate ratio vs. tolerance.

Experimental penetrability. As the previous results are based only on a theoretical formula, to complement validation of the landscape metaphor, we conduct penetrability tests (Algorithm 2). Each agent can have two possible states: *informed* or *susceptible*. (We use the term susceptible in reference to the SI epidemic model [28], meaning "susceptible to become infected"—a susceptible agent is not infected, but has the potential of being so.) Whatever the nature of information, virus, rumor, or opinion is, only an informed agent can transmit the signal. Assuming that the goal is to reach the frontier, the movement of propagation toward the opposite community can be seen as

https://doi.org/10.25088/ComplexSystems.25.3.169

moving downhill, since, at any time, a less segregated condition is preferable to a more segregated one: the lower the segregation index, the nearer to the opposite community the agent is.

The penetrability test (Algorithm 2) is executed as follows. First, the segregation walk is initialized by "informing" a randomly chosen agent a_0. The information is then passed from the agent to one of its nearest neighbors that has a lower segregation index. Iteratively, the walk terminates as soon as the signal reaches an agent a_k that is a local minimum. Finally, the test succeeds if there is a gate g in the opposite community such that $d(a_k, g) = 1$. As the objective is to approach the likelihood for the information to reach an agent of the opposite group, we applied the penetrability test a large number of times (10 000) in order to get the ratio of success. The root mean square deviation between the theoretical penetrability given by equation (6) and the ratio of success of penetrability tests is very low, in the order of 0.01 (Figure 10). This result appears to confirm that the segregation landscape is a complementary approach to determine the penetrability of the frontier. The main reason is that the landscape metaphor allows us to distinguish on the frontier between hubs and agents barely connected.

```
Require: all the agents are susceptible
    inform a randomly selected agent a_0 ∈ C
    k ← 0
    while not (a_k ∈ C_σ^min) do
        a_k propagates the information to one neighbor a_{k+1}
such that a_{k+1} ∈ C and σ(a_{k+1}) < σ(a_k)
        k ← k + 1
    end while
    return d(a_k, F_C) = 1
Ensure: a_k ∈ C_σ^min
```

Algorithm 2. Penetrability test.

5. Conclusion and Future Work

Our understanding of the segregated world has been shifted toward a landscape-oriented three-dimensional world. This allows us to highlight the underlying community structure induced by segregating behaviors. This new point of view allows us to refine the concept of frontier using that of referrer-agents that can or cannot act as gates through which information may pass from one community to the other.

https://doi.org/10.25088/ComplexSystems.25.3.169

The benefits of this approach are multiple: while the two-dimensional view already represents a metaphor for social segregation, the three-dimensional view expands our conceptual capabilities by facilitating the mapping between emerging segregation shapes and recognizable features such as hills, valleys, watersheds, and so on. Furthermore, additional metrics can be applied, such as landscape roughness, frontier scope distribution, and others.

One of the more significant findings to emerge from this study is that in fact, agents differ greatly from one place to another. While previous works on Schelling-like models consider that all the agents within a community are equivalent, from a landscape perspective some of them play a critical role; for instance, two individuals that are located next to each other and on both sides of a watershed are not equivalent, provided that they are acting to propagate information to the opposite community.

In relation with the concept of frontier, we identified the role of referrer, differentiating agents inside segregated groups from those situated on the frontier. According to its neighborhood in the opposite community, a referrer can be a gate through which information passes in or out of the segregated group to which it belongs. The power of the metaphor lies in such storytelling: the image of information going through hilly landscapes, from one community to the other, helps to efficiently communicate the complex mechanics of social segregation.

Schelling showed that global aggregation may occur even if it does not correspond to agent preferences; thus, in some way his model supports the assumption that global segregation is unavoidable whatever the tolerance is. Moving from the classical averaged local dissimilarity index to the segregation index metric, the segregation landscape approach changes the point of view and is seriously weakening this hypothesis: (*i*) studying correlation between tolerance and landscape roughness, we show that tolerance seriously impacts the emergent shape only when tolerance is low (Figure 7); (*ii*) studying the frontier properties, we show that thickness is high only if tolerance is below 0.5 (Figure 8); and (*iii*) considering the flow of information through the landscape, the notion of gate, and the size of the basins of attraction, we show that the scope distribution differs considerably between low and high tolerance (Figure 9) and penetrability is low only when tolerance is below 0.5 (Figure 10).

In the future, we plan to further the present work in the following ways: (*i*) study the impact of various strategies (random or targeted) in order to control or prevent penetrability of the frontier; and (*ii*) use the three-dimensional metaphor to study the dynamic that leads to the emergence of a fixed-point landscape that will be satisfactory for everyone.

https://doi.org/10.25088/ComplexSystems.25.3.169

References

[1] Nobelprize.org. "Thomas C. Schelling: Biographical." (Jun 20, 2016) http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2005/schelling-bio.html.

[2] M. A. Bedau, "Weak Emergence," *Philosophical Perspectives*, **11**, 1997 pp. 375–399. http://www.jstor.org/stable/2216138.

[3] S. I. Fabrikant, D. R. Montello, and D. M. Mark, "The Natural Landscape Metaphor in Information Visualization: The Role of Commonsense Geomorphology," *Journal of the American Society for Information Science and Technology*, **61**(2), 2010 pp. 253–270. doi:10.1002/asi.21227.

[4] C. R. Reeves, "Landscapes, Operators and Heuristic Search," *Annals of Operations Research*, **86**, 1997 pp. 473–490. doi:10.1023/A:1018983524911.

[5] P. F. Stadler, "Fitness Landscapes," in *Biological Evolution and Statistical Physics* (M. Lässig and A. Valleriani, eds.), Berlin, Heidelberg: Springer, 2002 pp. 183–204. doi:10.1007/3-540-45692-9_ 10.

[6] P. Collard, S. Mesmoudi, T. Ghetiu, and F. Polack, "Emergence of Frontiers in Networked Schelling Segregationist Models," *Complex Systems*, **22**(1), 2013 pp. 35–39. http://www.complex-systems.com/pdf/22-1-2.pdf.

[7] P. Collard and S. Mesmoudi, "How to Prevent Intolerant Agents from High Segregation?," in *Advances in Artificial Life (ECAL 2011)*, Paris, France, 2011 (T. Lenaerets et al., eds.), Cambridge, MA: The MIT Press, 2011 pp. 168–175. https://mitpress.mit.edu/index.php?q=books/advances-artificial-life-ecal-2011.

[8] R. Pancs and N. J. Vriend, "Schelling's Spatial Proximity Model of Segregation Revisited," *Journal of Public Economics*, **91**(1–2), 2007 pp. 1–24. doi:10.1016/j.jpubeco.2006.03.008.

[9] J. Zhang, "A Dynamic Model of Residential Segregation," *Journal of Mathematical Sociology*, **28**(3), 2004 pp. 147–170. doi:10.1080/00222500490480202.

[10] S. Gerhold, L. Glebsky, C. Schneider, H. Weiss, and B. Zimmermann, "Computing the Complexity for Schelling Segregation Models," *Communications in Nonlinear Science and Numerical Simulation*, **13**(10), 2008 pp. 2236–2245. doi:10.1016/j.cnsns.2007.04.023.

[11] L. Gauvin, J. Vannimenus, and J.-P. Nadal, "Phase Diagram of a Schelling Segregation Model," *The European Physical Journal B*, **70**(2), 2009 pp. 293–304. doi:10.1140/epjb/e2009-00234-0.

[12] A. Banos, "Exploring Network Effects in Schelling's Segregation Model," presentation given at *S4-Modus Workshop: Multi-scale Interactions between Urban Forms and Processes*, Besançon, France, 2009.

[13] J. B. Rosser, "On the Complexities of Complex Economics Dynamics," *Journal of Economic Perspectives*, **13**(4), 1999 pp. 169–192. doi:10.1257/jep.13.4.169.

[14] T. C. Schelling, "Models of Segregation," *The American Economic Review*, **59**(2), 1969 pp. 488–493. http://www.jstor.org/stable/1823701.

[15] T. C. Schelling, "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, **1**(2), 1971 pp. 143–186. doi:10.1080/0022250X .1971.9989794.

[16] W. J. Carrington and K. R. Troske, "On Measuring Segregation in Samples with Small Units," *Journal of Business & Economic Statistics*, **15**(4), 1997 pp. 402–409. doi:10.1080/07350015.1997.10524718.

[17] S. Grauwin, F. Goffette-Nagot, and P. Jensen, "Dynamic Models of Residential Segregation: Brief Review—Analytical Resolution and Study of the Introduction of Coordination," Social Science Research Network (July 1, 2009). doi:10.2139/ssrn.1436978.

[18] A. J. Laurie and N. K. Jaggi, "Role of 'Vision' in Neighbourhood Racial Segregation: A Variant of the Schelling Segregation Model," *Urban Studies*, **40**(13), 2003 pp. 2687–2704. doi:10.1080/0042098032000146849.

[19] S. Wright, "The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution," in *Proceedings of the Sixth International Congress on Genetics*, **1**(6), 1932 pp. 356–366.

[20] S. A. Kauffman, "Adaptation on Rugged Fitness Landscapes," in *Lectures in the Sciences of Complexity: The Proceedings of the 1988 Complex Systems Summer School, held June–July 1988 in Santa Fe, New Mexico* (D. L. Stein, ed.), Redwood City, CA: Addison-Wesley Publishing Company, 1989 pp. 527–618.

[21] S. A. Kauffman and E. D. Weinberger, "The *NK* Model of Rugged Fitness Landscapes and Its Application to Maturation of the Immune Response," *Journal of Theoretical Biology*, **141**(2), 1989 pp. 211–245. doi:10.1016/S0022-5193(89)80019-0.

[22] A. E. Lobkovsky, Y. I. Wolf, and E. V. Koonin, "Predictability of Evolutionary Trajectories in Fitness Landscapes," *PLOS Computational Biology*, 7(12), 2011 1002302. doi:10.1371/journal.pcbi.1002302.

[23] I. Altman and D. A. Taylor, *Social Penetration: The Development of Interpersonal Relationships*, New York: Holt, Rinehart and Winston, 1973.

[24] P. N. Kugler, M. T. Turvey, and R. Shaw, "Is the 'Cognitive Penetrability' Criterion Invalidated by Contemporary Physics?," *Behavioral and Brain Sciences*, **5**(2), 1982 pp. 303–306 doi:10.1017/S0140525X00012139.

[25] J. J. E. van der Meer and G. F. Roach, "The Influence of a Geographical Barrier on the Balance of Selection and Migration," *Mathematical Methods in the Applied Sciences*, **8**(1), 1986 pp. 269–283. doi:10.1002/mma.1670080118.

[26] U. Wilensky. "NetLogo Home Page" (Jun 21, 2016). http://ccl.northwestern.edu/netlogo.

[27] N. G. Domic, E. Goles, and S. Rica, "Dynamics and Complexity of the Schelling Segregation Model," *Physical Review E*, **83**(5), 2011 056111. doi:10.1103/PhysRevE.83.056111.

[28] M. E. J. Newman, *Networks: An Introduction*, New York: Oxford University Press, 2010.