Chemotaxis Rules

Chemotaxis is the motion of an organism in response to external chemical stimuli. A brief and partial examination of chemotaxis and physiology literature suggests that the organism may sense either the absolute or relative amount of the chemical present, and may also estimate spatial or temporal gradients by sampling the signal in space or time. They respond to these signals by adjusting their heading, speed, or both to either approach or avoid the source of the chemical. We seek to model chemotaxis in a simplified setting to assess the results of different assumptions of organism sensing and response. The organisms are simulated in an agent-based framework to seek the source of a stationary chemical gradient. The agents "sense" either absolute concentration or spatial gradients, or are able to estimate temporal gradients using a single time step memory. They choose their heading, speed, or both at each time step, and their positions are recorded over time. Additionally, the agents within a simulation do not interact – they can superpose in space and do not affect the chemical distribution.

When an organism senses a high level of a positive chemical signal, it may slow down to pinpoint the location of the source. Likewise, in low chemical concentrations the organism may increase its speed to cover more ground. An agent that has location (x, y) at time t and controls its speed, s, uses the following model:

$$s(\phi) = S_{\max} - \frac{|\phi(x, y, t)|}{\gamma}(S_{\max} - S_{\min}), \tag{1}$$

where $\phi(x, y, t)$ is either the magnitude of the spatial gradient, the value of the temporal gradient, or the absolute concentration of the chemical as required. Note that only in the case of the temporal gradient can $\phi(x, y, t)$ be negative. The parameters $S_{\max}$ and $S_{\min}$ are the maximum and minimum values of the speed, and γ is the known or estimated maximum value of ϕ over time and space. For absolute concentration, $\gamma = \gamma_c$ is maximum value of the concentration on the grid (≈ 1), which is fixed for all time. For the spatial gradient, $\gamma = \gamma_g$ is the maximum magnitude of the center-difference approximation to the gradient, also fixed over time. And for the temporal gradient, $\gamma = \gamma_t$ is estimated to be $\gamma_t = S\gamma_g dt$, where $S = S_{\max}$ (unless the speed is fixed in which case $S = 1$) and dt is the length of the time step. This is the maximum change in concentration that an agent will sense after moving at top speed for one time step. Using this rule, agents slow down at high concentrations or gradients and speed up at low concentrations or gradients. This means that at the source, where the concentration is high and the gradients are low, agents that choose speed according to gradients will move around a lot more than those sensing concentration. This may or may not be a biologically reasonable choice.

A similar method is used for calculating the choice of direction. For an agent with location (x, y) at time t, the direction is calculated as follows:

$$p(\phi) = \alpha_{\max} - \frac{|\phi(x, y, t)|}{\gamma}(\alpha_{\max} - \alpha_{\min})$$
$$\theta = \theta_0 + p(\phi)\eta(x, y, t) + \pi H(-\text{sgn}(\phi)), \tag{2}$$

where θ is the agent's new heading. $\phi(x, y, t))$ and γ are the same as in (1), although γ_c is not used because a gradient estimation is needed to choose direction. θ_0 is a reference heading – either the direction of the spatial gradient or the direction from which the agent came during the last time step. $p(\phi)$ defines a "precision" cone around the reference angle θ_0 that models uncertainty in the sensory apparatus or that encodes bias to continue in the same direction. $\alpha_{\max}$ and $\alpha_{\min}$ are the maximal and minimal possible widths of the precision cone (the width to either side of

θ_0). We choose $\alpha_{\max} = \pi$ to ensure random motion in a null sensory environment, although other circumstances may require a smaller $\alpha_{\max}$. $\eta(x, y, t)$ is a uniformly distributed random number in $[-1, 1]$, which has the effect of choosing a particular angle within the precision cone. H is the Heaviside function and depends on the sign of $\phi(x, y, t)$. When $\phi(x, y, t) < 0$, π is added to the final heading θ, turning the agent around when it is headed down-gradient. Note that this term is relevant only when $\phi(x, y, t)$ represents the value of the temporal gradient.

When we say that an agent picks its direction using an estimate of the temporal gradient, we are not referring to a lab-frame chemical gradient. Indeed, in our toy problem, the chemical gradient in time is precisely zero. Instead, we are referring to a temporal gradient in the "agent-frame", a coordinate system that travels along with the agent. The agent at first picks a random direction in which to move. It then compares the current chemical concentration to the concentration that it sensed at the previous time and location. If the difference is high, then the agent is more strongly biased to move in the same direction it has been going (or the opposite direction if the difference is negative and large in magnitude). If the difference is low, then the agent chooses a direction more randomly. In contrast, when using the spatial gradient to choose direction, the agent is biased always to move up-gradient but can sense the direction more precisely if the magnitude of the gradient is high.

In our simple problem, an odorant point source of magnitude is located at $(0.4L, 0.3L)$ in a box that is L units on a side. The source decays exponentially as $C = e^{-40r^2}$, where r is Euclidean distance from the source, and is constant in time. Two hundred agents are randomly distributed in the area $(0.75L, 0.85L)$ by $(0.75L, 0.85L)$ at time $t = 0$. At every time step, the agents make a decision about where to go based on a pre-determined rule set from Table 1. Entries in Table 1 that say "Fixed" mean that the agent's speed is constantly $s = 1$ regardless of sensory input. "Random" entries mean that the direction or speed is chosen randomly from a range: $[0, 2\pi)$ for the angular heading and $[S_{\min}, S_{\max}]$ for the speed. The rest of the entries in the second and third columns indicate the sensory method that the agents use to alter direction or speed as described in the preceding paragraphs. Note that rule set #10 is diffusion; it is included for comparison purposes. The last two columns indicate whether changing direction or speed parameters will change agent behavior, and are included only for clarity.

Table 1: Rule sets for chemotactic motion.

Rule Set	Direction	Speed	$\alpha_{\min}$	$S_{\min}$, $S_{\max}$
1	Spatial Gradient	Spatial Gradient	y	y
2	Spatial Gradient	Absolute Concentration	y	y
3	Spatial Gradient	Fixed	y	n
4	Spatial Gradient	Random	y	y
5	Temporal Gradient	Temporal Gradient	y	y
6	Temporal Gradient	Absolute Concentration	y	y
7	Temporal Gradient	Fixed	y	n
8	Temporal Gradient	Random	y	y
9	Random	Absolute Concentration	n	y
10	Random	Fixed	n	n

We nondimensionalize the problem by taking the length scale to be the length of one side of the square domain, L, so that all of our simulations occur on the unit square. We choose a typical velocity of the modeled organism, denoted s_0, to be a scaling factor for the speed; an agent with a fixed speed in Table 1 is assigned the typical nondimensional speed of 1. These choices lead to

a time scaling of L/s_0, so that a nondimensional time of $\Delta t = 1$ can be thought of as the average time it would take an agent to cross the shortest side of the domain provided that it never changed direction. The agents make navigation decisions at time steps of $\Delta t = 0.01$. As an example of the output of our simulations, Fig. 1 shows the positions of each of the 200 independent agents using rule #2 to navigate with $S_{\min} = 0.5$, $S_{\max} = 1.5$, and $\alpha_{\min} = 2\pi/3$.
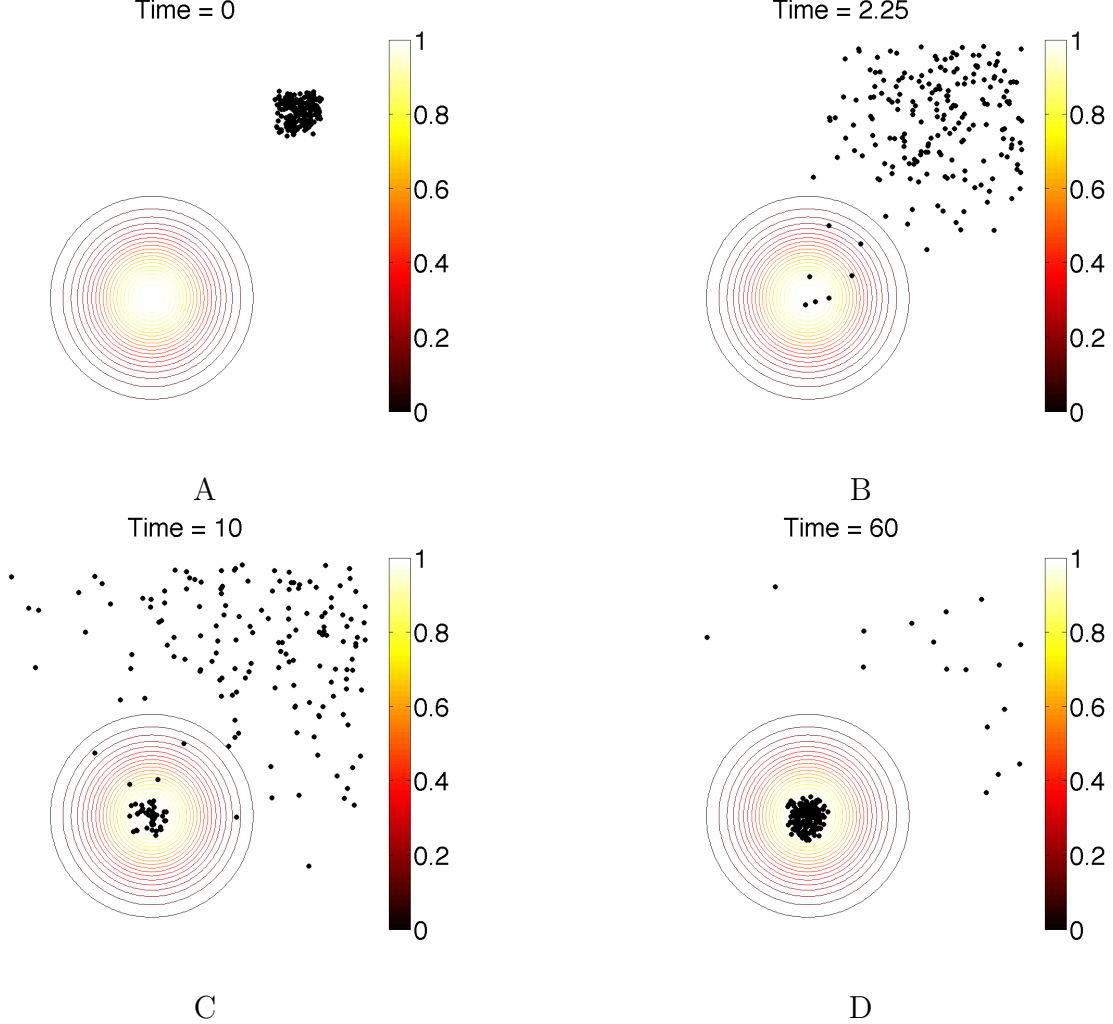

Figure 1: Snapshots at various times showing the locations of each of 200 agents navigating using rule #2 (parameters in the text). The contour plot shows the chemical concentration.

We ultimately want to compare this 2D, time-dependent output between rule sets with varying parameters. Since the "goal" of the agents is to reach the chemical source (or maximal concentration) we consider the distance from each agent to the stationary chemical source at each moment in time, a vector that we denote $\mathbf{r}(t)$. We plot in Fig. 2.A ten time-dependent agent trajectories from the simulation in Fig. 1, chosen to show the large range of agent behaviors. After each agent comes into contact with a sufficiently high chemical gradient, it travels rapidly to the vicinity of the source. Since agent response does not change after reaching the source, the agents irregularly orbit the target, maintaining a relatively close distance. As a comparison, we plot the same set of agents (i.e., the agents with the same initial conditions) using rule #9 in Fig. 2.B, and see that the agents do not seem to home in on the source in the allotted time. These two patterns are clearly different, and the purpose of this section is to rigorously quantify the difference in general between
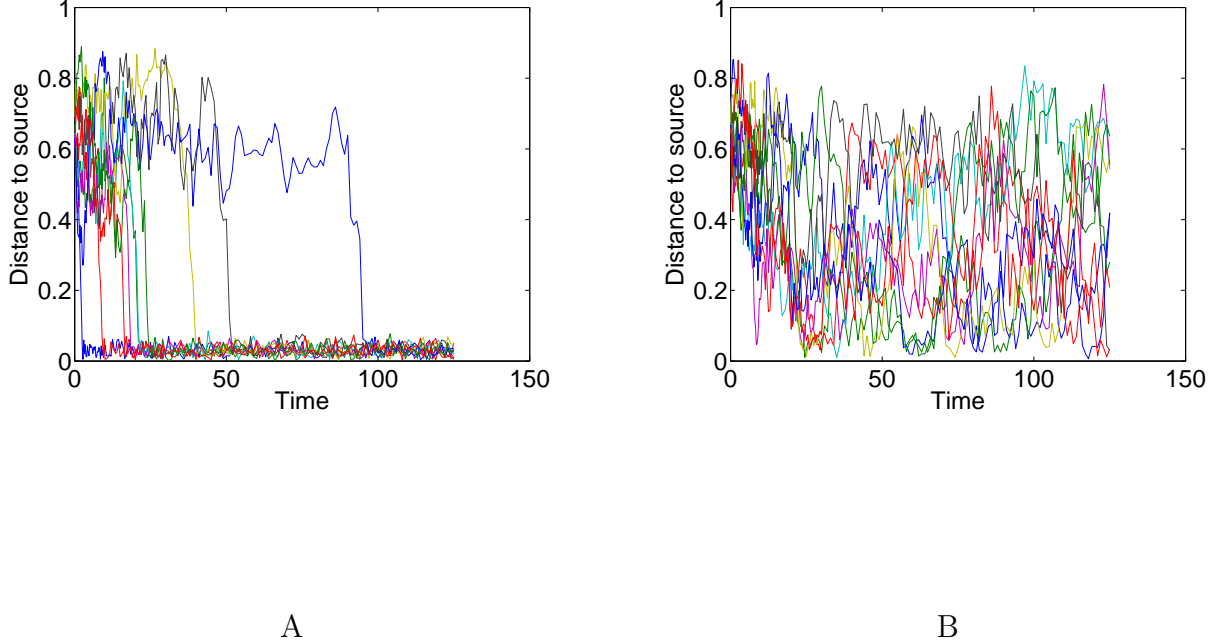
two rule sets.

Figure 2: The distance of each of 10 agents to the chemical source as a function of time for A. the simulation in Fig. 1; and B. rule set #9. The mosquitoes in B. have the same set of initial conditions as those in A.

There are two main differences between rule sets: how the organism responds (speed, heading, or both), and the type of sensory input used to respond (absolute concentration or a gradient). From Eqs. (1) and (2), it can be seen that all the differences in choice of speed and direction are due to the quantity $|\phi(x, y, t)|/\gamma$. γ is intended to be a scaling factor that restricts $|\phi(x, y, t)|/\gamma$ to the range $[0, 1]$, although if information about the environment is uncertain this will not be an exact mapping. If distributions of $\phi(x, y, t)$ are similar across sensory inputs and γ is chosen well, then in principal it *should not matter* what sensory modality the agent uses to make decisions. If, for example, the curves $\phi_c(x, y, t)$ and $\phi_g(x, y, t)$ can be scaled such that they approximate each other, then we would expect the agents' movement choices to be equivalent under these two modalities. However, even if the curves are initially dissimilar, it may be possible to tweak the other parameters ($S_{\max}$ and $S_{\min}$ for speed and $\alpha_{\min}$ for direction) in order to offset the difference. We specifically aim to see if rule sets can be made similar through parameter choices regardless of sensory modality or even response type (speed, heading, or both).

In order to obtain an estimate of the variability within and between rule sets, we performed 25 simulations with $N = 200$ agents each for every one of the 10 rule sets and every one of 120 different parameter sets. We varied $S_{\min}$ from 0.25 to 0.75 at 0.25 intervals; $S_{\max}$ from 1.25 to 2 at 0.25 intervals; and $\alpha_{\min}$ from $\pi/12$ to $10\pi/12$ at $\pi/12$ intervals. Not all of the rule sets changed with every parameter set – for example, rule set #9 is not affected by changes to $\alpha_{\min}$ (see the last two columns of Table 1). The output is a set of vectors $\mathbf{r}(t)$ of length N for each combination of rule
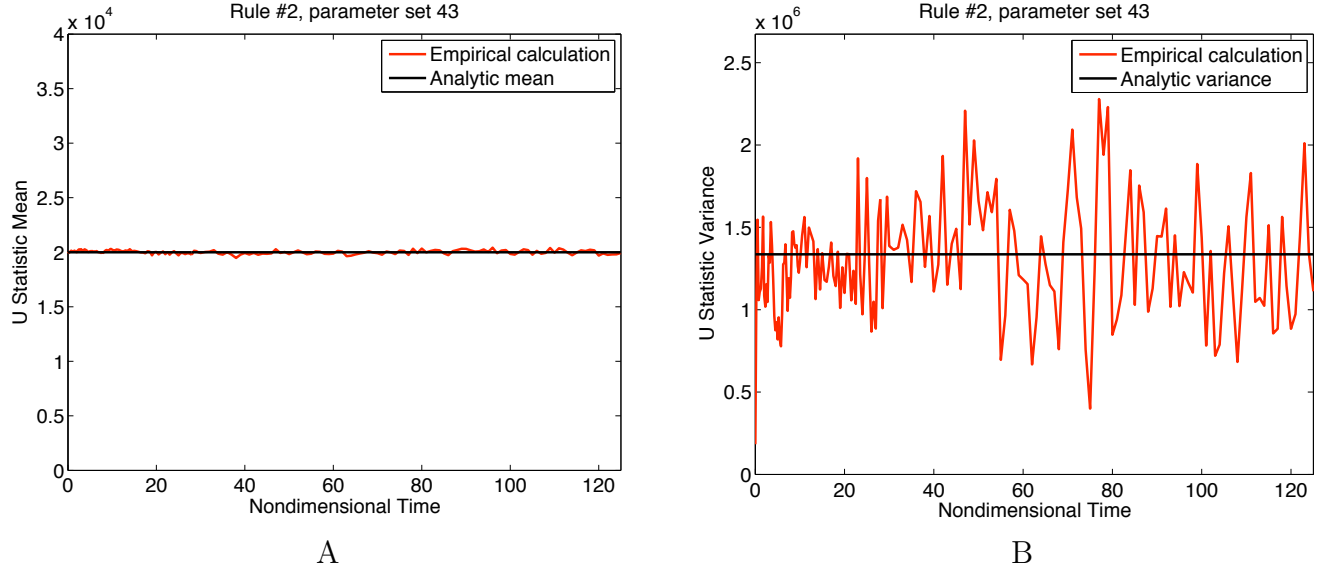
4

set, parameter set, and simulation number.

Figure 3: Comparison of empirical calculations of the mean and variance of the Mann-Whitney U statistic over time to their analytical asymptotic values. 300 U statistics at each time step were used in the calculation.

We use the nonparametric Mann-Whitney U statistic to pairwise compare $\mathbf{r}(t)$ across simulations within each time step. Briefly, we jointly rank two $\mathbf{r}(t)$ vectors at the same time, t, and sum the ranks associated with the first vector. We then subtract the expected value of the sum provided that the two vectors come from the same distribution: $N(N+1)/2$ in this case. This gives the U statistic for the pair based on the first vector (the other U statistic, U_2, will be $N^2 - U_1$).

For a given rule set with a specific parameter set, we would like to establish the baseline variance between simulations. To do this, we calculate the Mann-Whitney U statistic between all pairs of the $\mathbf{r}(t)$ from the 25 simulations mentioned above – a total of 300 independent U statistics at each time step. We find the mean and variance of this collection of U statistics at each moment and plot it versus time. The Mann-Whitney U statistic calculated between samples from the same distribution is known to be normally distributed with a mean of $N^2/2$ and a variance of $N^2(2N-1)/12$ for large N. Our empirically calculated values match the mean well, but the variance poorly (Fig. 3). [maybe perform bootstrapping to smooth things out]

Now that we know what the baseline variation looks like within a rule set, we are in a position to judge the similarity or difference between rule sets. When we calculate the U statistic between rule sets, we get 625 independent U statistics at each moment in time. We again find the mean and variance of the collection of U statistics at each time and plot it on the same graph as the reference distribution in Fig. 3. We find the parameter set for the second rule set that best matches the reference rule set and parameter set, and base our estimation of equivalence on the plots of mean and variance. Figure 4 shows examples of poor matches (A and B) and good matches (C–F). The reference rule set is #2, with $S_{\min} = 0.50$, $S_{\max} = 1.50$, and $\alpha_{\min} = \pi/4$ (parameter set 43). By examining Fig. 4.E and F, we conclude that rule #2 with parameter set 43 is equivalent to rule #1 with any of parameter sets 33, 43, or 53. Interestingly, only the minimum speed varies between these parameter sets, suggesting that it is the least sensitive parameter of the three. The minimum angle shows the most sensitivity – in Fig. 4.A and C the lines group into 5 different clusters, each of which has the same value for $\alpha_{\min}$. Although more difficult to see, some of the same clustering

can also be seen in Fig. 4.B. [maybe add a bit about statistical significance? – might be tough with the crazy jaggedy bits]

Figure 5 shows good matches between rule #2, with $S_{\mathrm{min}} = 0.50$, $S_{\mathrm{max}} = 1.50$, and $\alpha_{\mathrm{min}} = 2\pi/3$ (parameter set 48) and the other three rule sets in which the agent controls both direction and speed. In general, in order to get good matches between the spatial and temporal gradient methods it is necessary to choose a much smaller α_{min}. This figure supports the conjecture that sensory modality (as modeled in this paper) does not strongly affect agent behavior.

Figure 6 compares rule set #2 to the rule sets that control direction, but leave the speed to be determined randomly (4 and 8). It is easy to see that relatively good matches can still be obtained under these circumstances.

[Discuss response classes – heading, speed, or both.]
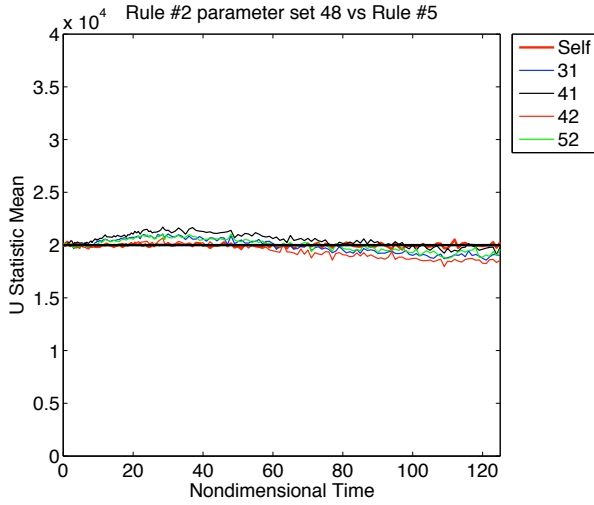
[Show results for diffusion.]
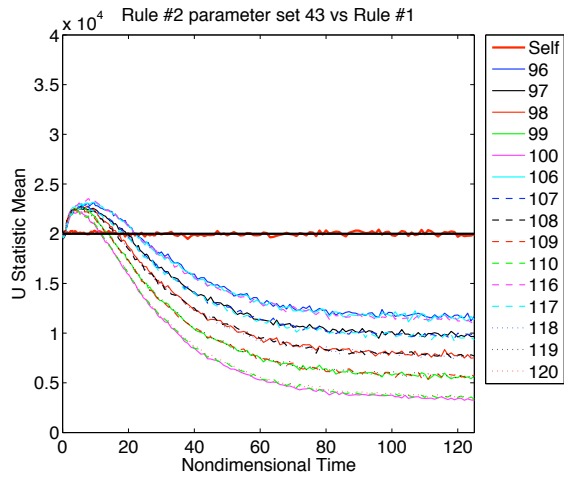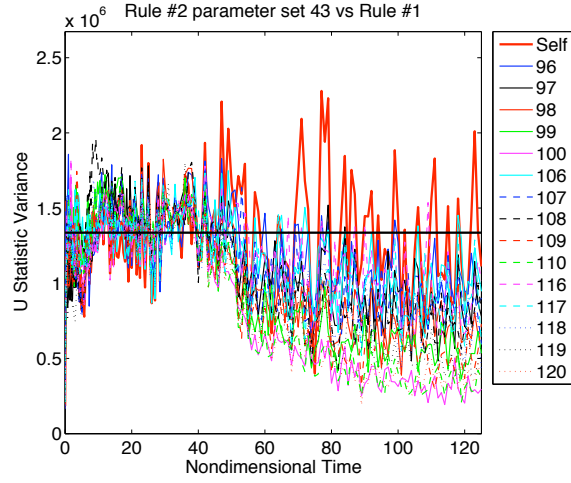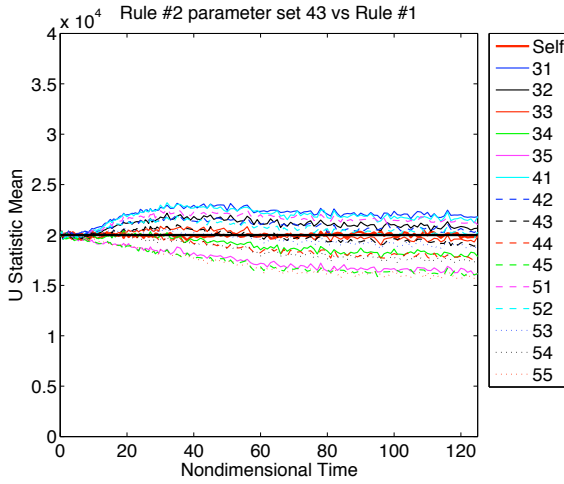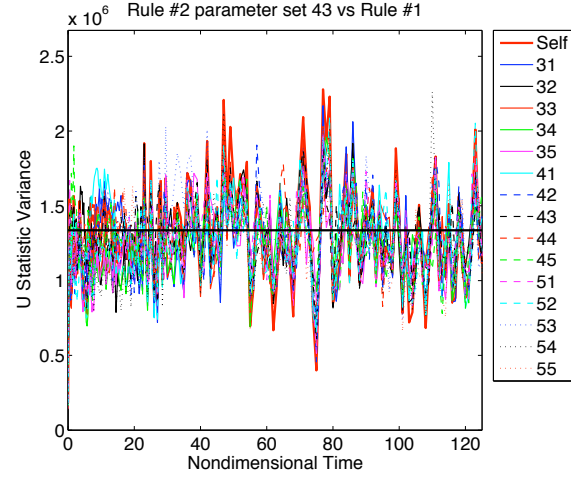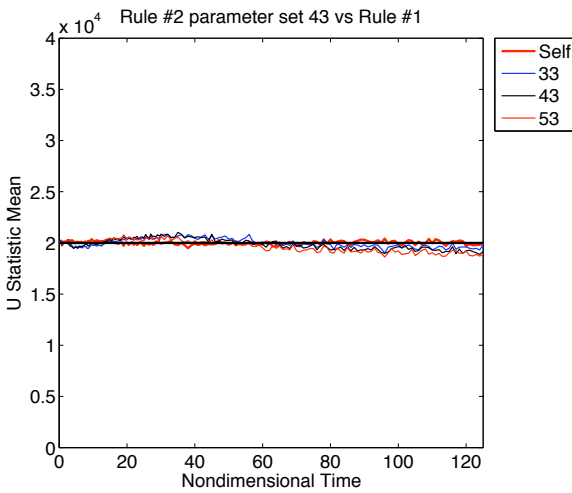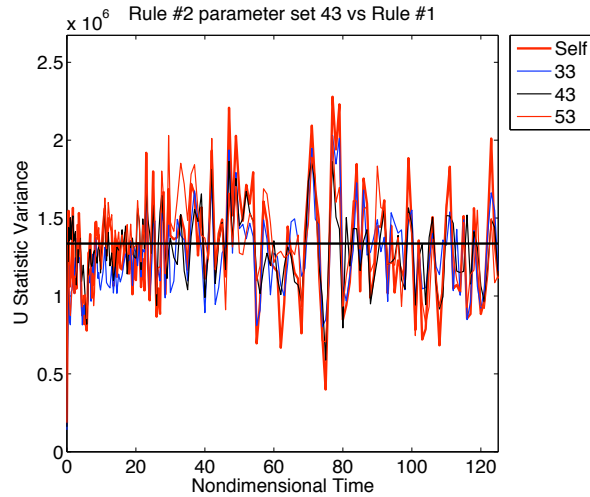
Figure 4: Comparison of the mean and variance of the Mann-Whitney U statistic between rules 1 and 2 over time. 625 U statistics at each time step were used in the calculation. The legend (except for the first entry) denotes the parameter set used with rule #1. The thick black and red lines are as in Fig. 3; if the rule sets are similar then the thin line should match the thick red line. A. & B. show poor matches for the mean and variance; C. & D. show a better set of matches; and E. & F. show the best matches overall.
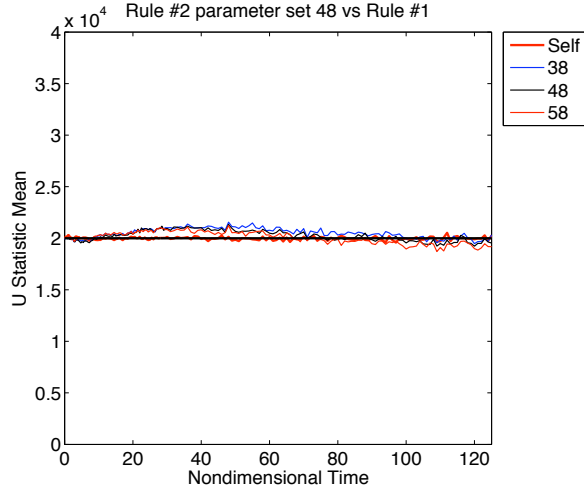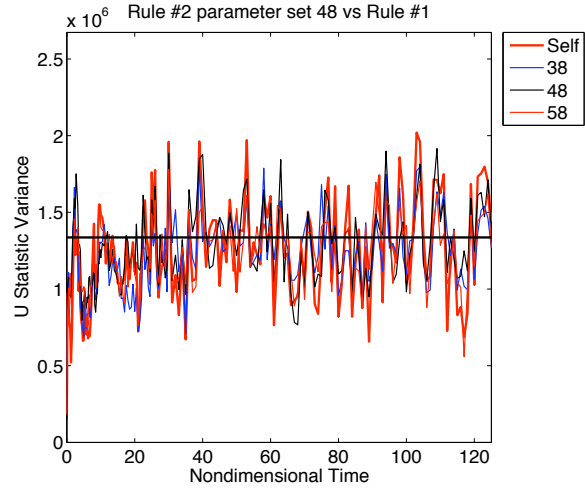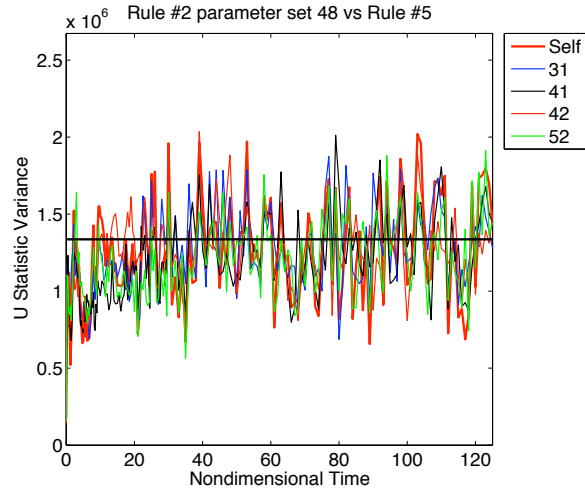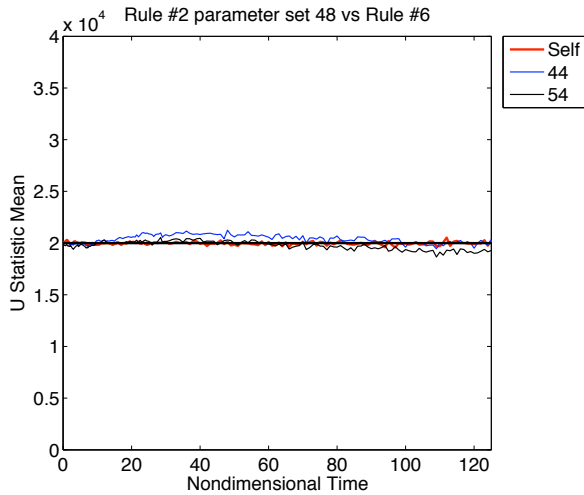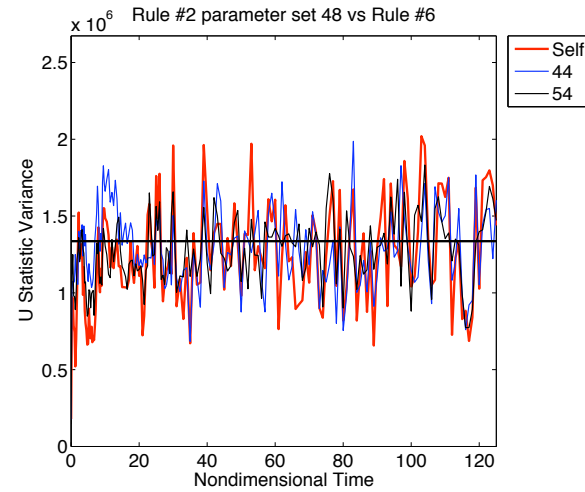
Figure 5: Best matches to the mean and variance of the Mann-Whitney U statistic between rule 2 parameter set 48 and rules 1, 5, and 6 over time. These rule sets all control both speed and direction. 625 U statistics at each time step were used in the calculation. The legend (except for the first entry) denotes the parameter set used with rules #1, 5, and 6 that provided the best match. The thick black is as in Fig. 3; the thick red line is like the one in Fig. 3 except that it is calculated for rule 2 parameter set 48. If the rule sets are similar then the thin line should match the thick red line. A. & B. Rule 2, parameter set 48 versus rule set 1; C. & D. versus rule set 5; and E. & F.
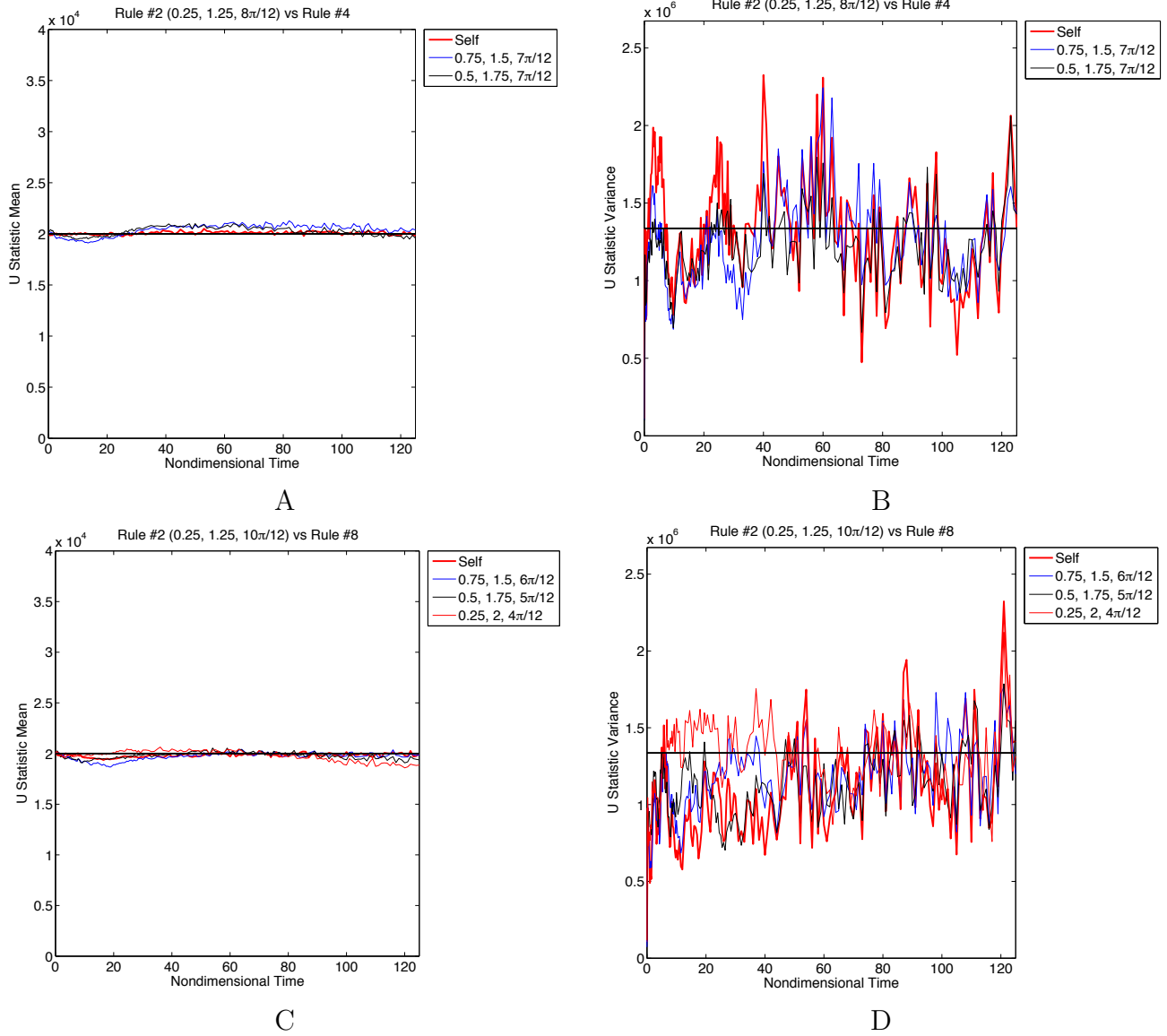
Figure 6: Best matches to the mean and variance of the Mann-Whitney U statistic between rule 2 parameter set 10 and rules 4 and 8 over time. Rule sets 4 and 8 control only direction, with speed determined randomly within $[S_{\min}, S_{\max}]$. 625 U statistics at each time step were used in the calculation. The legend (except for the first entry) denotes the parameter set used with rules #4 and 8 that provided the best match. The thick black is as in Fig. 3; the thick red line is like the one in Fig. 3 except that it is calculated for rule 2 parameter set 10. If the rule sets are similar then the thin line should match the thick red line. A. & B. Rule 2, parameter set 10 versus rule set 4; C. & D. versus rule set 8.