Fingerprint Analysis of the Noisy Prisoner’s Dilemma Using a Finite State Representation

Daniel Ashlock  
Mathematics and Statistics  
University of Guelph,  
Guelph, Ontario  
Canada N1G 2W1  
dashlock@uoguelph.ca

Eun-Youn Kim  
National Institute For  
Mathematical Sciences  
385-16, Doryong-dong,  
Yuseong-gu, DaeJeon 305-340,  
South Korea  
eunykim@nims.re.kr

Wendy Ashlock  
Department of Computer Science  
and Engineering  
York University  
Toronto, Ontario  
Canada M3J 1T3  
washlock@cse.yorku.ca

Abstract—Fingerprinting is a technique that permits automatic classification of strategies for playing a game. In this study the evolution of strategies for playing the iterated prisoner’s dilemma at three different noise levels is analyzed using fingerprinting and other techniques including a novel quantity, evolutionary velocity, derived from fingerprinting. The results are at odds with initial expectations and permit the detection of a critical difference in the evolution of agents with and without noise. Noise during fitness evaluation places a larger fraction of an agent’s genome under selective pressure, resulting in substantially more efficient training. In this case efficiency is the production of superior competitive ability at a lower evolutionary velocity. Prisoner’s dilemma playing agents are evolved for 6400 generations, taking samples at eight exponentially-spaced epochs. This permits assessment of the change in populations over long evolutionary time. Agents are evaluated for competitive ability between those evolved using distinct noise levels. The presence of noise during agent training is found to convey a commanding competitive advantage. A novel analysis is done in which a tournament is run with no two agents from the same evolutionary line and one-thousand agents from each noise level studied. This analysis simulates contributed agent tournaments without any genetic relation between agents. It is found that in early epochs the agents evolved without noise have the best average tournament rank, but that in later epochs they have the worst.

I. INTRODUCTION

This study is a substantially lengthened version of [13]. The experiments were re-run to permit the gathering of new types of data and several additional analyses were performed. A new quantity, evolutionary velocity, is introduced in this study and is used to expose an unexpected and critical difference between agents evolved with and without noise using the finite state representation.

The prisoner’s dilemma [18], [17] is a classic model in game theory. Two agents each decide, without communication, whether to cooperate (C) or defect (D). The agents receive individual payoffs depending on the actions taken. The payoffs used in this study are shown in Figure 1. The payoff for mutual cooperation C is the cooperation payoff. The payoff for mutual defection D is the defection payoff. The two asymmetric action payoffs S and T, are the sucker and temptation payoffs, respectively. In order for a two-player simultaneous game to be considered prisoner’s dilemma, it must obey the following pair of inequalities:

\[ S \leq D \leq C \leq T \]  
\[ 2C \leq (S + T). \]

In the iterated prisoner’s dilemma (IPD) the agents play many rounds of the prisoner’s dilemma. IPD is widely used to model emergent cooperative behaviors in populations of selfishly acting agents and is often used to model systems in biology [29], sociology [23], psychology [28], and economics [22].

![Fig. 1.](image-url) (1)The payoff matrix for prisoner’s dilemma used in this study – scores are earned by strategy \( \mathcal{P} \) based on its actions and those of its opponent \( \mathcal{P} \). A payoff matrix of the general two player game \( (C, T, S, D) \) are the scores awarded.

The study in [10], continued in [5], investigates the effect of changing the representation used for a prisoner’s dilemma agent. The representations covered by the two studies are two versions of feed forward neural nets (one biased at the neuron level toward cooperation), Boolean parse trees [19], with and without a one-step time-delay operation, a linear genetic programming representation called an ISAc list [1], lookup tables, a type of Markov chain [27], and both a direct and cellular [5] representation of finite state machines. The change of representation, with other factors held as near to constant as possible, yielded a change from 0% to 95% in the probability that final populations were cooperative. This study uses one of these representations, finite state machines, investigating the probability of cooperative behavior in the
The duration of evolution in [10], [5] was 250 generations with samples taken at 50, 100, 150, 200, and 250 generations. Little effect on the cooperativeness of agents was observed at different epochs. This study demonstrates a different result for noisy strategies evolved for a longer time.

In [11] it was found that evolving agents to play the iterated prisoner’s dilemma for a long time gave them a substantial competitive advantage against agents evolved for less time from different evolutionary lines. This phenomenon, called non-local adaptation, suggests that agents are gaining just not skill playing the agents with whom they are co-evolving but general skill at playing the prisoner’s dilemma; they have a statistically significant advantage against agents they have never been evaluated against before. In [12] it was demonstrated that non-local adaptation takes place in a steady fashion across much of evolution. This study extends measurement of competitive advantage to include the effects of noise as well. Non-local adaptation is also observed in other co-evolutionary contexts. In [21] it was observed in competitive exclusion in a spatial model of plant growth. In [3] it was observed in populations of virtual robots evolving to paint a floor in two competing colors. The effect was observed in predator-prey models in [2] and in populations of agents in a game called divide the dollar in [4].

A prisoner’s dilemma fingerprint is a function in two variables, \(x\) and \(y\) ranging from 0 to 1, whose value is the strategy’s expected score when playing against an infinite test suite of strategies called Joss-Ann strategies. These strategies were chosen to be as representative as possible of the full space of strategies. They are defined using the strategies TFT and Psycho described in Table I. If \(x + y \leq 1\), the Joss-Ann strategy plays \(C\) with a probability of \(x\) and \(D\) with a probability of \(y\) and TFT otherwise. If \(x + y \geq 1\), the Joss-Ann strategy plays \(C\) with a probability of \(1 - y\), \(D\) with a probability of \(1 - x\) and Psycho otherwise.

Note that in either case when \(x + y = 1\), the strategy is just the random strategy which cooperates with probability \(x\) and defects with probability \(y\). For more details about fingerprinting see [24], [7], [5], [1], [6], [15]. This study uses an approximation to the fingerprint function consisting of 25 numbers which are the values of the fingerprint function on a 5 by 5 evenly-spaced grid from 1/6 to 5/6. Given that fingerprints are functions they are a subset of a potentially infinitely dimensioned space, making the use of a 25-point sampling grid potentially suspect. In [15] it is proved (Corollary 2) that the space of fingerprint functions is at most 6-dimensional, making a 25 point sample a generous one.

Fingerprints are developed in detail in three theses [30], [24], [15]. These results are summarized and extended in [6]. A portion of the theory of fingerprints and an initial application to the visualization of evolved agents appears in [7]. Additional applications as well as a marriage of fingerprints with a new technique called multi-clustering appear in [8]. One application of fingerprints is to place a metric-space structure on the space of prisoner’s dilemma strategies. In this study, the metric space structure induced by fingerprints is used to compute a new quantity: evolutionary velocity. The evolutionary velocity of a population is the distance that its mean fingerprint moves in one generation.

Fingerprinting was used in [16], with a finite state representation, to demonstrate that the strategies that arise have different distributions for different population sizes and in different epochs. The latter result, that strategies rare or absent at the beginning of evolution become common after thousands of generations of evolution, was surprising. In [10] fingerprints were used to demonstrate that the rate of appearance of several well-known strategies varied between a direct finite state representation for prisoner’s dilemma playing agents, a cellular representation for finite state agents, and a new type of representation called a function stack, a modified form of Cartesian Genetic Programming [25].

The remainder of the study is structured as follows. The details of experiments performed are given in Section II as are details of the analysis techniques used. Results and discussion are presented in Section III. Conclusions and possible next steps for both the improvement of fingerprinting techniques and applications are given in Section IV.

II. Experimental Design

The agent representation used in this study is 8-state finite state machines with actions associated with transitions between states (Mealy machines). State transitions are driven by the opponent’s last action. Access to state information permits the machine to condition its play on several of its opponent’s previous moves. The machines are stored as linear chromosomes listing the states. The initial state and action for a given machine are stored with and undergo crossover with the first state in this linear chromosome.

Two variation operators are employed, a binary variation operator and a unary variation operator. The binary variation operator used is two-point crossover on the list of states. Crossover treats states as atomic objects. The mutation operator changes a single state transition 40% of the time, the initial state used by the machine 5% of the time, the initial action 5% of the time, or an action associated with a transition 50% of the time. The unary variation operator replaces the current value of whatever it is changing with a valid value selected uniformly at random.

The evolutionary algorithm used in this study operates on a population of 36 agents, a number chosen for compatibility with previous studies [12], [9], [10]. Agent fitness is assessed by a round-robin tournament in which each pair of players engage in 150 rounds of the iterated prisoner’s dilemma. Three collections of runs were performed with different versions of this round-robin tournament. The first, serving as a baseline, had perfect communication between the agents. The second and third had an average of 1% and 5%, respectively, of actions misunderstood by the opponent. This is one of two forms of possible noise; the other (not used in this study) has the actual nature of the action reversed rather than the opponent’s perception of it.
Reproduction is elitist with an elite of the 24 highest scoring strategies, another choice that maintains consistency with past studies. When constructing the elite, ties are broken uniformly at random. Twelve pairs of parents are picked by fitness-proportional selection with replacement on the elite. Parents are copied, and the copies are subjected to crossover and mutation.

In each simulation the evolutionary algorithm was run for 6400 generations with 100 replicates (experiments with distinct random number seeds). The elite portion of the population in generations 50, 100, 200, 400, 800, 1600, 3200, and 6400 was saved for analysis. This yields 100 sets of 24 machines at each of eight epochs for each of the three noise levels. A number of descriptive statistics are saved for each generation of each replicate. These include the mean fitness, a 95% confidence interval on the fitness, the maximum fitness, and the evolutionary velocity of the population, defined in Section II-C.

Many of the parameters used in this study are either selected in imitation of previous studies or are essentially arbitrarily. A careful parameter-variation study would require millions of collections of evolutionary runs because of combinatorial explosion of the parameter space. The current study represents a small slice of this diverse space of possible experiments. We now carefully describe the ways in which the evolved machines were analyzed.

A. Probability of Cooperativeness

For each epoch and noise level, the fitness (average prisoner’s dilemma score) of the saved elite population in a noise-free tournament was computed. Following [10], [5] a population was judged to be cooperative if its average score was at least 2.8 out of a maximum possible of 3.0. The fraction of populations that were cooperative was modeled as the parameter of a binomial distribution; Figure 2 summarizes the results. The choice to evaluate all populations without noise permits an apples-to-apples comparison of the populations evolved at different noise levels.

B. Competitiveness and Nonlocal-Adaptation

While both competitiveness and non-local adaptation were analyzed in [13] a different form of these analyses are performed in this study. The results are in substantial agreement, but the method of analysis has been revised and the presentation of the information has been substantially modified in a manner that enhances clarity.

Competitiveness is measured between sets of agents evolved using different fitness evaluation processes. It is intended to assess the impact on the agent’s competitive ability of those distinct fitness evaluation processes. In this study there are three different fitness evaluation processes, all based on round robin tournaments, but using different levels of noise. We evaluate competitiveness in the zero-noise environment – a practice that would logically give the agents evolved in that environment an advantage. The zero noise environment was chosen because it yields deterministic results on any given pair of agents. As we shall see, agents evolved without noise did not benefit from its lack when evaluated against agents evolved in the presence of noise. Non-local adaptation is measured between agents from distinct epochs but evolved using the same fitness evaluation process. It measures the impact of additional evolution.

Save for the identity of the sets of agents being compared both competitiveness and non-local adaptation are measured in the same fashion. All agents for a given noise level and epoch are loaded into agent-pools. All pairs of distinct agent-pools are then compared. The comparison is performed by sampling 10,000 pairs of agents, each pair consisting of one agent from each of the agent-pools being compared. Each pair plays 150 rounds of iterated prisoner’s dilemma. Each time an agent from the first pool out scores an agent from the second pool a success is recorded. Each time an agent from the second pool out scores an agent from the first a failure is recorded. The successes and failures are modeled as Bernoulli experiments using a normal approximation to the binomial distribution. The Bernoulli parameter being estimated is the probability an agent from the first pool will beat an agent from the second. A 95% confidence interval is constructed for the Bernoulli parameter for each pair of agent-pools. Examples of the presentation of these confidence intervals for competitiveness appear in Figure 5 while those for non-local adaptation appear in Figure 6.

C. Evolutionary Velocity

A velocity is the change of a quantity over time. The natural measure of time in the evolutionary algorithms used in this study is generations. The fingerprints of prisoner’s dilemma playing agents used in this study map the strategies to points in 25-dimensional Euclidean space. A population may then be said to have a mean position in fingerprint space in each generation. The evolutionary velocity of an evolving population in a given generation is defined to be the change in the position of the mean fingerprint of the population between the last generation and the current one using the standard Euclidean distance.

The small, well-mixed populations used in this study ensure that a population will move to and stay near a state of relatively low diversity. When a population is relatively non-diverse, the fingerprints will be close together. A common event in an evolving population of game playing agents is a succession event. This happens when a new agent type appears that can exploit a flaw in the current dominant agent type. The population diversity briefly increases as the new agent type takes over. There should be some succession events visible as simultaneous changes in the population average score and spikes in the evolutionary velocity.

Based on earlier work [13] we know that being exposed to noise yields a competitive advantage to prisoner’s dilemma playing agents, an observation confirmed via a different evaluation of competitive ability in Section III-A of this study. Because of this we conjectured that evolutionary velocity would be lowest in the noise-free environment and highest in the environment where agents were evolved with the most
noise. In Section III-B we test and discuss this incorrect conjecture.

D. Two Forms of Competitiveness

Non-local adaptation and competitiveness both measure the ability of two agent types to out-score one another in a pairwise competition. Axelrod’s famous tournaments [18], [17] measured a completely different sort of competitive ability. This is the ability to have the highest score in a tournament with a diverse collection of other players. The strategy TFT cannot beat another player in a pairwise contest. The only way to get ahead of another player is to defect against them when they cooperate. TFT never defects first and so always defects against its opponent’s cooperation on a move directly after the opponent does the same. In spite of this, TFT won both of Axelrod’s tournaments. This is because, even though it was out-scored by all its opponents, it elicited cooperation in many of them. Its opponents, on the other hand, sometimes got into very low-scoring sequences of defections with one another. At the end of the tournaments TFT had the highest total score.

We create a contest similar to a contributed agent contest in the following fashion. A contest consists of a round robin tournament with 36 agents, the number used in fitness evaluation during evolution, using 150 rounds of iterated prisoner’s dilemma. In order to prevent genetically-based collaboration, no two agents in a given tournament were selected from the same replicate. Twelve agents were selected from each of the three available noise levels and all agents in a given tournament were selected from the same epoch. The agents were selected from the saved elite population uniformly at random and the replicates used in a tournament were also selected uniformly at random, without replacement. A set of 10,000 tournaments were performed. For each epoch 95% confidence intervals for the mean population rank of the agents from each noise level were computed.

E. Fingerprint Analysis: Voronoi Tiles

In order to analyze prisoner’s dilemma strategies one can try to find some way to define and name each individual strategy. Table I lists twelve strategies with names. However, this naming method quickly becomes intractable as the number of strategies increases. Also, it gives no objective way to compare the strategies. A prisoner’s dilemma fingerprint is a way of assigning a real-valued function to each strategy, or as with the simplified version used in this study, a point in 25-space which consists of 25 values of that function. In order to analyze the evolved strategies, twelve reference strategies, given in Table I, were selected. These strategies represent all unique fingerprints that appear in the collection of lookup tables for prisoner’s dilemma based on the opponent’s last two moves[15]. They are all equivalent to one-state or two-state finite state machines. A depiction of their relative positions in fingerprint space is shown in Figure 3.

The evolved strategies in this study are analyzed in terms of which of these strategies they are closest to. This technique is called a Voronoi tiling or a Dirichlet tessellation. Explanations of the technique can be found in many places; one good one is [20]. If a strategy is closer to Strategy S than it is to any other reference strategy, then it is said to be in the Strategy S tile. The assumption is that strategies with fingerprints that are close together behave similarly, so we can group them together and analyze them in terms of the simple strategies that we understand well and in terms of where they fall in fingerprint space.

The space of fingerprint functions might be as much as 6-dimensional. However, the error is small when strategies are projected onto two dimensions using non-linear projection, a technique defined in [8] and applied in [14] to RNA folding. The technique uses an evolutionary algorithm to project multi-dimensional points onto a 2-dimensional picture preserving the distance relationships between the points as closely as possible. Thus, it is reasonable to conceptualize the space in two dimensions. The greatest possible fingerprint distance between any two strategies is that between ALLD and ALLC (proved in [15] Theorem 3); the next greatest distance is between TFT and Psycho. One can visualize fingerprint space as the interior of the diamond formed by these four strategies as shown in Figure 3.

III. RESULTS AND DISCUSSION

In an early study of the impact of noise on the iterated prisoner’s dilemma [26] it was found that higher levels of noise yielded lower levels of cooperation. This study evolved 16-state finite state machines using the Moore encoding for 50 generations with the same noise levels used in this study. The top panel of Figure 2 demonstrates that allowing additional evolutionary time changes this result. In the final epoch samples, and only in that epoch, the significant differences between the probabilities that a population will be cooperative at different noise levels disappears. The noise free and 1% noise populations never exhibit a significant difference, though the 1% noise populations are less cooperative in 7 of 8 epochs. The 5% noise populations, however, show a steady and statistically significant increase in their ability to cooperate. It appears that they are learning strategies which allow cooperation in the presence of noise.

The assessment of cooperativeness was performed in a noise-free environment to permit apples-to-apples comparison of the probability of cooperativeness. The bottom panel of Figure 2 displays the average fitness of populations, averaged over all 100 replicates. The experiments with 1% noise appear to pay a small fitness penalty for noise events, while the 5% noise experiments pay a larger one. The 5% noise level experiments have a substantial upward trend in fitness over the course of evolution. This is evidence that the strategies evolving under 5% noise are adapting nontrivially to the presence of noise.

A. Competitiveness and Non-local Adaptation

Results for competitiveness are presented in Figure 5. Sampling pairs of players as a win/loss Bernoulli variable yields entirely comparable results to placing whole populations into
TABLE I

<table>
<thead>
<tr>
<th>Abbrev.</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFT</td>
<td>Fit-for-tat</td>
<td>This strategy does whatever its opponent did last time.</td>
</tr>
<tr>
<td>TF2T</td>
<td>Fit-for-two-tats</td>
<td>This strategy cooperates except after a sequence of two defections.</td>
</tr>
<tr>
<td>UC</td>
<td>Usually cooperate</td>
<td>This strategy cooperates except after a C following a D.</td>
</tr>
<tr>
<td>ALLC</td>
<td>Always cooperate</td>
<td>This strategy always cooperates.</td>
</tr>
<tr>
<td>PSY-TFT</td>
<td>Psycho-tit-for-tat</td>
<td>This strategy plays like PSYCHO until its opponent defects; then it plays like TFT until its opponent cooperates.</td>
</tr>
<tr>
<td>TFT-PSY</td>
<td>Tit-for-tat-psycho</td>
<td>This strategy plays like TFT until its opponent defects; then it plays like PSYCHO until its opponent cooperates.</td>
</tr>
<tr>
<td>PSY-TFT</td>
<td>Psycho-tit-for-tat</td>
<td>This strategy plays like PSYCHO until its opponent defects; then it plays like TFT until its opponent cooperates.</td>
</tr>
<tr>
<td>TF2T-UC</td>
<td>Inverse tit-for-two-tats</td>
<td>This strategy does the opposite of what 2TFT would do.</td>
</tr>
</tbody>
</table>

The main result is that evolution in the presence of noise yields a competitive advantage. This is in spite of the fact that the arena for evaluating competitiveness is the noise-free environment. The use of two noise level in the experimentation also demonstrates that 5% noise yields a more commanding advantage.

Figure 5 displays 64 comparisons representing each possible pair of saved epochs for each of three noise level. Each comparison comprises 10,000 pairwise tests of agents and is made in the form of a confidence interval on the probability that the agent type indexing the row will beat the agent type indexing the column. Agents evolved with a lower noise level index the columns. There are three categories of outcome: row victory, column victory, or a failure to find a significant competitive advantage. These have been gray-scale coded with light gray representing column victory, dark gray representing row victory, and medium gray representing a result in which the confidence interval for probability of row victory includes 0.5.

The grey-scale coding permits the viewer to immediately see that all epochs of agents evolved at 5% noise beat all epochs of agents evolved at 0% noise, in spite of the fact that additional evolution grants competitive advantage when the noise level is not varied. The time factor becomes important in the current experiments in the 0:1 and 1:5 comparisons where agents from a substantially later epoch achieve victory over agents evolved at a higher noise level. One is tempted to conjecture that all three populations are becoming more competitive, and that the agents evolving in the presence of noise are simply doing so faster. The evidence presented in Section III-B suggests that this is, at best, an oversimplification of the true state of affairs.

The results for non-local adaptation, shown in Figure 6 use the same grey-scale encoding scheme as the competitiveness results. These matrices have a form of skew-symmetry, with entries opposite across the main diagonal being mirror images of one another. This yields 28 distinct comparisons, in the form of pairs of symmetric entries off of the diagonal. However, they are less easy to interpret. The original observation of non-local adaptation [11] compared two epochs, 100 and 10,000 generations, of a noise free prisoner’s dilemma structured as a spatial evolutionary algorithm. The current experiments have a much simpler design but report and compare eight epochs rather than two. The hypothesis tested in the initial non-local adaptation study was that additional evolution grants a competitive advantage even against agent types not encountered before. The initial study supported this hypothesis; the results in [13] and here are consistent, but the new results reveal a more complex situation.

A comparison of two epochs is consistent with the non-local adaptation hypothesis (NLA hypothesis) if either the more-evolved population has a significant advantage or if neither population has a significant advantage. In the zero-noise experiments, summarized in the top panel of Figure 6, 26 of the 28 comparisons are consistent with the NLA hypothesis. Both the inconsistent comparisons are on the sub-diagonal representing adjacent epochs. In the experiments with 1% noise, 25 of 28 comparisons are consistent with the NLA hypothesis. The three comparisons that are inconsistent are the comparisons of the three most evolved populations. These three populations also have the greatest temporal separation. The situation grows substantially worse for the NLA hypothesis in the experiments with 5% noise, where only 18 of the 28 comparisons are consistent with the hypothesis. The 5% noise agents sampled at epoch 6400 are beaten by epochs 200-3200. This is clear evidence of a substantial decline in competitive ability for the agents evolved with 5% noise at deep time against younger agents evolved at 5% noise.

An explanation for this retrograde non-local adaptation lies in Figure 2. The strategies evolved at 5% noise begin a sharp increase in cooperativeness in the 5th-8th epoch. This corresponds exactly with the sudden decline in competitive ability in the third column, 5th-8th rows of the bottom panel of Figure 6. There is an apparent tradeoff between competitive ability and the ability to cooperate in the presence of noise. As they gain the ability to cooperate in a noise environment, the agents evolved in a high-noise environment become more tolerant of (possibly-noise induced) defections. The agents from earlier epochs, presumably less tolerant of defection, exploit this
tolerance. This highlights the dilemma of cooperation in a noisy environment: how do we tell a true defection from the noise-induced appearance of defection?

Note that the additional tolerance of noise events in the more highly evolved agents from the 5% noise environment does not mean they are becoming pushovers. The between-noise-level comparisons show that the agents are still very strong competitors; the loss of competitive ability in later epochs acts in a purely intramural manner.

B. Evolutionary Velocity

Fingerprint based evolutionary velocity is introduced in this study. When small, well-mixed populations of Prisoner’s dilemma playing agents are trained with an evolutionary algorithm the populations rapidly become non-diverse. This means that, most of the time, a single agent type with minor variation dominates the population. An important event is the succession in which a new agent type, able to out-compete the current dominant type, appears and takes over the population. Succession events are sometimes accompanied by sudden changes in the population mean fitness. In other cases the fitness difference is relatively small and the succession is difficult to detect. One potential application of fingerprints is to detect such events.

This type of detection is prototyped in this study by displaying fitness and evolutionary velocity on the same set of axes. One representative population from each noise level are displayed in this fashion in Figure 7. For all three noise levels we note that abrupt changes in average score are often accompanied by pronounced spikes in evolutionary velocity. The size of the changes in fitness is not well correlated with the size of the evolutionary velocity spikes. It is intuitive that these sizes need not be correlated in magnitude. A relatively modest shift in strategy, and hence fingerprint, can result in a large change in score and vice versa.

During the analysis of these results a possible improvement in the experimental design was noted. The evolutionary velocity was computed from the mean fingerprint of the entire population. In retrospect, computing the mean position of the elite would likely have generated a cleaner and more useful signal. All members of the elite have survived at least one fitness evaluation. It is possible for a mutant with very low probability of survival to also exhibit a large shift in its fingerprint. This large fingerprint deviation would be reflected as spurious evolutionary velocity. The fingerprints of mutants that will not survive into the next generation generates a preventable source of noise in evolutionary velocity.

Our hypothesis was that evolutionary velocity would increase with noise level. That simple hypothesis explains the patterns in both Figure 4 and Figure 6. Using both these
Fig. 7. This figure displays overlaid plots of fitness and evolutionary velocity on the same axes. Each of the three panels are results for one replicate. The top panel is from a population evolved with 0% noise, the middle a 1% noise population, and the bottom 5% noise. These plots are illustrative rather than representative.
Fig. 4. The relative size of the contents of the different Voronoi tiles in fingerprint space across the sampling epochs.

Fig. 5. Competitiveness results in the form of 95% confidence intervals for the probability that the agents indexing the row will beat the agents indexing the column. The top panel gives results for no noise versus 1% noise, the middle for no noise versus 5% noise, and the bottom for 1% noise versus 5%. Rows are indexed by the lower-noise agent pools with numerical labels for both rows and columns giving the epoch of the agent pool. Dark gray cells indicate significant superiority of the agent-pool indexing the row, light gray the same for the pool indexing the column. Middle gray indicates no significant difference.
Evolutionary velocity is a measure of change of position in the strategy space. If there is a reason the higher noise agents will tend to a more direct path to good positions in the adaptive landscape, then all the results reported in this study remain consistent. Such an explanation follows with some necessary preliminaries.

Given the substantial competitive advantage that additional noise grants, this suggests that agents evolved in the presence of noise are moving through strategy space more efficiently. Evolutionary velocity is a measure of change of position in strategy space. If there is a reason the higher noise agents will tend to take a more direct path to good positions in the adaptive landscape, then all the results reported in this study remain consistent. Such an explanation follows with some necessary preliminaries.

One feature of fingerprints is that they measure only the asymptotic behavior of agents. Examine the finite state machine in Figure 8. It implements the strategy *vengeful* which cooperates until its opponent’s first defection and defects thereafter. In a noise-free environment this is a fairly competitive

### Table II

<table>
<thead>
<tr>
<th>Noise</th>
<th>Mean evolutionary velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>(0.164,0.168)</td>
</tr>
<tr>
<td>1%</td>
<td>(0.134,0.138)</td>
</tr>
<tr>
<td>5%</td>
<td>(0.125,0.129)</td>
</tr>
</tbody>
</table>

Fig. 8. The finite state diagram for a two-state implementation of the strategy *vengeful.*
strategy, it cooperates perfectly with itself and defends itself aggressively against anything that does not embrace a no-first-use policy toward defection. The noise used in the computation of fingerprints ensures that the asymptotic behavior and hence the fingerprint of vengeful is identical to that of always defect. The ability of vengeful to cooperate with copies of itself are strongly dependent on perfect information: a lack of noise. Its cooperation is embedded in transient states.

The implementation of vengeful given in Figure 8 is a simple example of the dependence of success on transient states – states that the machine will leave and never return to if the opponent makes a particular move. Populations evolved in the absence of noise are likely to exploit such transient states to recognize copies or near-copies of themselves. This recognition strategy is called *handshaking* and grants agents the ability to preferentially cooperate with strategies that share their genes. Noise makes handshaking less practical, because the handshake must be able to tolerate noise. The practicality of handshaking also drops as the level of noise increases both because the imperative to correct for noise grows and because the potential for more complex patterns of noise grows. A strategy that can deal with one noise event and still handshake cannot deal as easily with two. Given that the agents used in this study have eight states, there may not be room in the genome to handle multiple-noise errors in a handshake.

Consider now a population of agents evolved without noise. There is evolutionary pressure to make the portion of the state-diagram used to compete with copies of yourself small. This reduces the probability it will be modified by mutation and so makes it more heritable. This secondary evolutionary pressure to keep the active genome small is a form of *parsimony*. It is also advantageous to have this portion of the state diagram immediately accessible from the starting state of the transition diagram. This evolutionary pressure thus favors the use of “early” states in the transition to manage the interaction of an agent with its close relatives. Such states are often transient and thus not measured by fingerprinting. “Later” states might be rarely used in fitness evaluation. Mutation, thus, would make them more or less random. We deduce that agents evolved without noise will have evolutionary velocity considerably closer to the fundamental random-walk velocity that would result from applying mutations to a population with random selection than would a population evolved with noise. We offer this as a tentative explanation for the higher evolutionary velocity of the zero-noise populations. We turn now to the question of more efficient travel through strategy space by agents evolved at higher noise levels.

We have already discussed why an agent evolved in the presence of noise will find it harder to depend on transient states. In fact, any state accessible from the starting state has a positive probability of being reached when noise is present. This means that the number of states used during a fitness evaluation will be higher when noise is present and, to a degree, the evenness of access to states will increase with noise. This means that agents evolved in the presence of noise have a higher fraction of their genetic material subjected to selection pressure. It is intuitive that this increase in the generality of selection pressure across the genome will grow with noise, at least for “low” noise levels.

At this point a thought experiment may be valuable. A population of agents evolved in the absence of noise inherit many features in their transition diagrams that are not subjected to selection pressure. A mutation will, with positive probability, place these untested features into the portion of the transition diagram that is relevant to selection. Likewise, mutation can detach parts of the transition diagram from those portions undergoing selection and, once they are detached, re-
randomize them. This means the rate at which a population develops, and later expresses, non-adaptive features is much higher when the fraction of the genome undergoing selection is lower. We have already demonstrated that the fraction of the genome undergoing selection is higher in the presence of noise. This means that evolutionary velocity generated by the exploitation of and elimination of non-adaptive features resulting from the long-term action of mutation will be higher as noise levels drop and will be highest at zero noise.

The agents evolved with noise are, in essence, making far better use of their fitness evaluations. Since a greater fraction of their genomes are under test, they are more likely to retain adaptive behaviors once those behaviors are acquired. They are less likely to gain non-adaptive features that, once placed under selection pressure by a mutation, must be eliminated from the gene-pool. We conclude that agents evolved in the presence of noise move toward good locations in the adaptive landscape more efficiently than agents evolved without noise. To be clear: no-noise agents have a higher fingerprint-based evolutionary velocity that generates a lower rate of acquisition of adaptive behavior. This results from the lower proportion of their genome that is under selection in any given generation.

To test this reasoning, we analyzed the state use of the agents in the experiments. The results of this analysis are shown in Figure 9. This figure shows that the noiseless populations use significantly more transient states. It also shows that, in the early epochs (the first five for 1% noise, the first three for 5% noise), the agents evolved with noise use more of their states than those evolved without noise. In later epochs the number of states used by agents evolved with noise decreases. We conjecture that this is driven by the increase in evolutionary efficiency.

All the evolving agents are subject to the same parsimony pressure to have smaller genomes. If the agents evolved with noise are, in fact, evolving more efficiently, wasting less time on weeding out mutation-induced non-adaptive features, then the parsimony pressure acts on them more strongly. The results of parsimony pressure may be set aside during a succession event. The zero noise agents, which we have demonstrated run around the adaptive landscape at a higher rate, are likely to have a higher rate of succession events and so set aside parsimony gains more often. This tentative explanation for the smaller number of states used by agents evolved with noise decreases. We conjecture that this is driven by the increase in evolutionary efficiency.

C. Contest Winners and Evolution Dominators

The results presented in Sections III-A and III-B suggest that it might be worth testing agents sampled from the different noise levels and different populations in a contest setting. The confidence intervals for the mean contest rank of each agent type are displayed, for all epochs, in Figure 10. In these tournaments rank 1 is the best rank, and so the better an agent population is the lower it appears in the figure. The tournament ranks are in the range 1-36 while the confidence intervals have a radius of 0.05-0.065 and so are difficult to see.

In the first three epochs, the agents evolved without noise have a significantly better tournament ranking that those evolved with 1% noise which in turn are superior to those with 5% noise. Between the third and fourth epoch agents evolved with 1% noise obtain significantly superior mean contest rank to those evolved with no noise. Between the fourth and fifth epochs the agents evolved with 5% noise also pass the agents evolved without noise. The 1% noise agents retain their superiority to the 5% noise agents throughout all epochs. In [16] it was found that strategies not observed early in evolution became common at deep time (6400 generations). The results on mean contest ranking echo these results – there is a significant difference in the character of prisoner’s dilemma agents between early and later evolution.

The ability to win a contest is different from the ability to survive the process of evolution. Winning a tournament involves getting high scores against a diversity of opponents. Tournament winners do not consistently beat their opponents; they just get high scores while playing them. Evolutionary algorithms select for strategies which get high scores when playing copies of themselves and which beat mutants. We call agents with these two different sorts of competitive ability contest winners and evolution dominators. A contest winner will receive a high tournament ranking when playing round-robin against a slate of diverse opponents. An evolution dominator will tend to cooperate with agents very similar to itself (its relatives) and beat agents different from itself. These definitions reflect the different objective functions used for scoring contests and for driving the evolution of a small well-mixed population.

It is intuitive that an evolution dominator will do poorly in a contest unless the contest is packed with strategies similar to itself with which it will cooperate. The evolutionary algorithm used in this study selects strongly for evolution dominators. It does so less efficiently when noise is present, as the retrograde non-local adaptation shown in Figure 6 demonstrates. The populations evolved with noise are, however, better contest winners. The 1% noise populations are superior to the 5% noise populations. This, in turn, suggests that there is a difference in the contest-winner ability of agents evolved at different noise levels and hence potentially an optimum noise level for such evolution. This level is probably below 5%. These results also suggest that a well-mixed, noiseless, small-population evolutionary algorithm is not an optimal tool for designing entries for a prisoner’s dilemma contest.

D. Fingerprint Analysis

The outcome of the Voronoi tile fingerprint analysis is given in Figure 4. These figures show the distribution of the strategies in the tiles of the combined populations evolved at each noise level. The difference between the noiseless populations and the two sets of populations evolved with noise is substantial. The populations evolved with no noise change comparatively little in the proportions of strategies in each tile, while both sets of populations evolved with noise show a trend with time towards strategies in the ALLD and (TF2T) tiles.
All three sets of populations contain no significant number of strategies in the ALLC, (2TFT), and PSY tiles (less than 1.0% in the noiseless populations and less than 0.2% in the populations evolved with noise). The strategies in these tiles are likely just bad strategies. Of the other nine tiles, the noiseless populations favor the TFT tile with 45% of strategies falling in it. There does not seem to be much change in the distribution over time, except that strategies begin to appear in significant numbers in the (TF2T) tile starting after 400 generations (epoch 4). In epoch 1 they account for less than 1% of the strategies; in epoch 4 they account for 2%, and by epoch 8 5% of the strategies are in that tile.

In the populations evolved with 1% noise, there is a significant trend into the ALLD and (TF2T) tiles. The ALLD tile contains 1% of the strategies in epoch 1 and 20% in epoch 8. Likewise, the (TF2T) tile contains 1% of the strategies in epoch 1 and 60% in epoch 8. The trend is even more pronounced in the populations evolved with 5% noise. 1% of the strategies in the first epoch are in the (TF2T) tile and 60% in the last epoch. The move to the ALLD tile happens sooner in the 5% noise populations with 12% of the strategies there in the first epoch and 32% in the last epoch.

The Voronoi tiles do not give information about specific strategies found; they only give information about what other strategies have nearby fingerprints. We also analyzed the occurrences of two specific strategies: ALLD and TFT. In the noiseless populations ALLD never amounts to more than 2% of the total strategies; TFT ranges from 10% to 22% with a peak in epoch 2. In the 1% noise populations neither strategy is common, they amount to less than 1% of the total, except in epoch 4 in which 3% of the strategies are ALLD. In the 5% noise populations, TFT is also rare (less than 1%), but ALLD exists in significant numbers, ranging from 8-32% of the strategies, peaking in epoch 4.

Another strategy worth analyzing in particular is the Fortress strategy. It was first discovered and analyzed in [16] when it appeared in populations that had evolved for more than 4096 generations. The three state version of it is pictured in Figure 11. Fortress \( n \) has \( n \) states in its minimal form. It will cooperate indefinitely with another strategy which cooperates with it, but only after a series of \( n - 1 \) defections, and, if the opponent ever defects, it must play the series of defections again before Fortress will cooperate with it. Table III shows the fraction of strategies with fingerprints close to Fortress 3 for all the populations saved. “Close to” is defined by three different distances: \( R = 0.1 \) means the strategies are virtually identical to Fortress 3; \( R = 1.0 \) means they are as similar as strategies in the Voronoi tiles are to their reference strategies, and \( R = 0.32 \) is in between. In the noiseless populations, these strategies are rare, but become more common in the later epochs, peaking at 10% in epoch 8 with \( R = 1.0 \). The story is radically different in the populations evolved with noise. They
also become more common with time, but, instead of being rare, they come to dominate the population with 83% of the strategies within $R = 1.0$ of Fortress 3 in epoch 8 of the 1% noise populations and 86% in the 5% noise populations.

IV. CONCLUSIONS AND NEXT STEPS

This study introduces the idea of fingerprint-based evolutionary velocity and integrates it with previous results on competitiveness, non-local adaptation, and probability of cooperative behavior to reach the following conclusions. First, employing noise in a finite state representation increases the fraction of states that is under selection. Second, having a higher fraction of the genome under selection enhances the efficiency of evolutionary training.

This study also highlights two distinct forms of competitive ability, the ability to dominate an evolving population and the ability to win a contest against a diverse selection of opponents. These two abilities, while not opposites, are demonstrated in this study to have some degree of trade-off. In particular, we conclude that a well-mixed, small-population evolutionary algorithm will favor the production of evolution dominators over contest winners. We also note that evolving with noise enhances the algorithm’s ability to produce contest winners and that adding 1% noise does this to a greater degree than adding 5% noise. This suggests that the noise level can be optimized for the production of contest winners, a possible next step for this research.

Results on the number of transient and unused states in the three different agent types suggest that the more efficient evolution enjoyed by agents evolved with noise may have permitted evolution to exert a greater parsimony pressure on the size of the active genome in those agents. The genome size, eight states or 68 bits, used in this study leaves only modest room for parsimony pressure to act. A follow-up study with more states could be used better understand the interaction of noise and evolution to affect state usage.

This study shows that the non-local adaptation hypothesis from [11] is too simple. While additional evolution often conveys a competitive advantage, it doesn’t do so invariably. This study opens a number of doors for understanding why this is so. The delineation of distinct types of competitiveness (evolution dominator and contest winner) also permits a more nuanced phrasing of the non-local adaptation hypothesis.

In [15] it was found that relaxing both the small-population and well-mixed hypotheses in evolution without noise yields a diverse population that is cooperative and remains stable against invasion. Re-examining this evolutionary design with evolutionary velocity could be used to document the stability. The type of evolutionary algorithm used there might be much better for evolving contest winners than the sort used here.

The hypothesis that the fraction of the genome under selec-

<table>
<thead>
<tr>
<th>Epoch</th>
<th>R=0.1</th>
<th>R=0.32</th>
<th>R=1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% Noise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.0000</td>
<td>0.0004</td>
<td>0.0471</td>
</tr>
<tr>
<td>100</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0271</td>
</tr>
<tr>
<td>200</td>
<td>0.0000</td>
<td>0.0025</td>
<td>0.0354</td>
</tr>
<tr>
<td>400</td>
<td>0.0000</td>
<td>0.0108</td>
<td>0.0387</td>
</tr>
<tr>
<td>800</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0046</td>
</tr>
<tr>
<td>1600</td>
<td>0.0308</td>
<td>0.0308</td>
<td>0.0446</td>
</tr>
<tr>
<td>3200</td>
<td>0.0196</td>
<td>0.0283</td>
<td>0.0867</td>
</tr>
<tr>
<td>6400</td>
<td>0.0413</td>
<td>0.0437</td>
<td>0.1008</td>
</tr>
<tr>
<td>1% Noise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.0000</td>
<td>0.0004</td>
<td>0.0171</td>
</tr>
<tr>
<td>100</td>
<td>0.0175</td>
<td>0.0238</td>
<td>0.0425</td>
</tr>
<tr>
<td>200</td>
<td>0.0646</td>
<td>0.0671</td>
<td>0.1083</td>
</tr>
<tr>
<td>400</td>
<td>0.1179</td>
<td>0.1471</td>
<td>0.2525</td>
</tr>
<tr>
<td>800</td>
<td>0.1988</td>
<td>0.2046</td>
<td>0.3588</td>
</tr>
<tr>
<td>1600</td>
<td>0.3125</td>
<td>0.3350</td>
<td>0.5483</td>
</tr>
<tr>
<td>3200</td>
<td>0.4646</td>
<td>0.4846</td>
<td>0.7396</td>
</tr>
<tr>
<td>6400</td>
<td>0.5729</td>
<td>0.5742</td>
<td>0.8329</td>
</tr>
<tr>
<td>5% Noise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0342</td>
</tr>
<tr>
<td>100</td>
<td>0.0088</td>
<td>0.0092</td>
<td>0.0721</td>
</tr>
<tr>
<td>200</td>
<td>0.0683</td>
<td>0.0742</td>
<td>0.1371</td>
</tr>
<tr>
<td>400</td>
<td>0.1988</td>
<td>0.2112</td>
<td>0.3162</td>
</tr>
<tr>
<td>800</td>
<td>0.2658</td>
<td>0.2804</td>
<td>0.3646</td>
</tr>
<tr>
<td>1600</td>
<td>0.3917</td>
<td>0.4238</td>
<td>0.5529</td>
</tr>
<tr>
<td>3200</td>
<td>0.5575</td>
<td>0.5642</td>
<td>0.7183</td>
</tr>
<tr>
<td>6400</td>
<td>0.6775</td>
<td>0.6775</td>
<td>0.8625</td>
</tr>
</tbody>
</table>
tion is an important variable deserves additional study. This study uses eight-state Mealy automata as its representation. This representation encodes a space of $2^{68} = 2.95E10$ distinct automata (many of which encode the same strategy), but each of these represents only 68 bits of information. It seems likely that the character of evolution in the presence of noise may change significantly if the genome size of agents is increased or permitted to evolve. In addition, the finite state representation is adept at failing to use portions of its genome. The representation used in this study, directly encoded finite state machines, was one of twelve used in [10], [5]. The genome-fraction question could be addressed in the context of any of these representations and many others.

The questions treated in this study of genome size can be asked in many contexts well beyond the prisoner’s dilemma. Fingerprinting may be employed for any simultaneous two-player game with a finite number of moves [24]. Many of the results supporting the results in this study are restricted to games with two moves (the mathematics becomes more complex when three or more moves are available).

V. ACKNOWLEDGMENTS

The first author thanks the University of Guelph and the National Science and Engineering Research Council of Canada (NSERC) for supporting this work. The second author thanks Korean National Institute For Mathematical Sciences for its support of this research. The third author thanks York University and the Ontario Council of Graduate Studies for their support of this research.

REFERENCES