Optimality Under Noise: Higher Memory Strategies for the Alternating Prisoner's Dilemma

Daniel B. Neill *†
Duke University

December 2, 2000

*This work is supported by NSF VIGRE grant number DMS-9983320.
†Many thanks to David and Vivian Kraines for their helpful comments and suggestions.

Abstract

The Alternating Prisoner's Dilemma is a variant of the iterated Prisoner's Dilemma in which the players alternate in the roles of actor and recipient. We searched for strategies which are "optimal" in the Alternating Prisoner's Dilemma with noise (a non-zero probability that a player's decision will be transmitted incorrectly). In order to achieve success against a variety of other strategies, a strategy must be "self-cooperating" (able to achieve mutual cooperation with its clone), "C-exploiting" (able to exploit unconditional cooperators), and "D-unexploitable" (able to resist exploitation by defectors). It must also have high evolutionary "dominance", a general measure of evolutionary performance which considers both resistance to invasion and ability to invade other strategies. A strategy which meets these optimality criteria can evolve cooperation by invading a population of defectors and establishing a stable cooperative society.

Most of the strategies commonly discussed in the Alternating Prisoner's Dilemma literature are low-memory strategies such as Tit For Tat, Pavlov, and Firm But Fair, but none of these strategies can simultaneously meet all of the optimality criteria. However, we discovered a class of higher memory "Firm Pavlov" strategies, which not only meet our stringent optimality criteria, but also achieve remarkable success in round-robin tournaments and evolutionary interactions. These higher memory strategies are friendly enough to cooperate with their clone, pragmatic enough to exploit unconditional cooperators, and wary enough to resist exploitation by defectors: they are truly "optimal under noise" in the Alternating Prisoner's Dilemma.

1 Introduction

The Prisoner's Dilemma is a widely used mathematical model of interactions between individuals with partially conflicting goals. In this paradoxical 2x2 non-zero sum game, each player must choose whether to *cooperate* for mutual benefit or to *defect* for individual benefit (harming the other player). The outcomes of the Prisoner's Dilemma game are given in the following payoff table:

		P1 C	P1 D
P2	C	R/R	T/S
	D	S/T	P/P

Table 1: Payoffs to P1/P2

If both players cooperate, each receives the payoff R as a *reward* for mutual cooperation. If both players defect, each receives the payoff P as a *punishment* for defecting. If one player defects while the other cooperates, the defector receives payoff T (as a *temptation* to defect), while the cooperator receives the *sucker* payoff S. The Prisoner's Dilemma is defined by $T > R > P > S$. Additionally, $2R > T + S$, so cooperation achieves more points than alternating T and S payoffs.

What is the optimal strategy for the Prisoner's Dilemma? It is clear that mutual cooperation is preferable to mutual defection, but in the one-shot Prisoner's Dilemma game, no matter what the other player chooses, a player can achieve a higher payoff by defection. Hence the optimal strategy is to defect, and mutual defection is the only rational result of a one-shot Prisoner's Dilemma. An *iterated Prisoner's Dilemma (IPD)* occurs when a Prisoner's Dilemma interaction is repeated by the same players over a number of rounds. Based on past results, reciprocal altruism can develop, enabling mutual cooperation to become a rational option. A *strategy* for the IPD is an algorithm for deciding whether or not to cooperate on a given round, based on the results of previous rounds. It is clear that some strategies are better than others. If a player cooperates too often, the other player can take advantage of him by defecting; if he defects too often, the other player is likely to retaliate, resulting in low scores for both players. It is much less obvious what an "optimal" IPD strategy would be, since a strategy's performance is highly dependent on the other player's strategy as well as the specific parameters of the Prisoner's Dilemma game. In 1980, Robert Axelrod conducted a computer "tournament" in which fourteen strategies designed by leading game theorists were matched against each other. The winning strategy was *Tit for Tat (TFT)*, submitted by Anatol Rapoport. This strategy cooperates initially, then echoes what the other player chose last round (Axelrod & Hamilton, 1981). Tit for Tat is very successful in many variants of the Prisoner's Dilemma game, and its success has sparked theories of the evolution of cooperation based upon reciprocity (Axelrod, 1984). These models have been applied to fields ranging from economics to biology (ex. Maynard Smith, 1982; Axelrod, 1984; Milinski, 1987).

Tit for Tat, however, is clearly not optimal in an iterated Prisoner's Dilemma with "noise", a non-zero probability that a player's decision will be transmitted incorrectly. In an IPD interaction between two TFT strategies, a single accidental defection can lead to an endless sequence of mutual recriminations (Nowak & Sigmund, 1990). Various other strategies have been examined for the the noisy iterated Prisoner's Dilemma, including the *Pavlov* (Win-Stay, Lose-Switch) strategy, which cooperates only after mutual cooperation (R payoff) or mutual defection (P payoff). Pavlov is an "error correcting" strategy: against its clone, it can recover from an accidental defection, resuming mutual cooperation in a relatively small number of rounds (Kraines & Kraines, 1989, 1993).

The *Alternating Prisoner's Dilemma (APD)* is a variant of the iterated Prisoner's Dilemma in which the players alternate in the roles of actor and recipient rather than acting simultaneously (Nowak & Sigmund, 1994; Frean, 1994). In this game, players alternate turns, and on each player's turn he must choose whether to cooperate or defect. This game models many situations in which participants must take turns helping each other, such as gift-giving, or bipartisan cooperation in politics. Examples of alternating reciprocal altruism can be seen in the feeding habits of South American vampire bats, and in the fighting of young male baboons (Wilkinson, 1984; Trivers, 1985).

There are several possible methods for keeping score in the Alternating Prisoner's Dilemma. Consider an APD game between two players, A and B. Player A's choices are denoted in upper case (C or D), and player B's choices are denoted in lower case (c or d). Nowak and Sigmund consider a "unit" consisting of one turn by each player, treating each unit as a normal Prisoner's Dilemma interaction (1994). Thus the unit Cd (a cooperation by A, followed by a defection by B) would result in A receiving a sucker payoff and B receiving a temptation payoff.

With this scoring method, alternation between players may lead to a lack of symmetry in the APD game. Consider a repeated sequence of one cooperation by each player, followed by one defection by each player: CcDdCcDd, etc. If we consider a round to be A's turn followed by B's turn, we get alternating Cc and Dd results, and each player scores $(R+P)/2$. If we consider a round to be B's turn followed by A's turn, we get alternating cD and dC results, and each player scores $(T+S)/2$. Thus, in order to maintain the symmetry of the game, Nowak and Sigmund require that $R+P=T+S$.

However, it is possible to eliminate this restriction on the payoff table through an alternate method for keeping score. We treat each turn after the first one as if it were a complete Prisoner's Dilemma, and thus each player receives two payoffs every round. For example, consider the following APD game:

1. Player A chooses C. Since this is the first turn, neither player receives a payoff.
2. Player B chooses d. Since Player A's last choice was C, A receives a sucker payoff S, and B receives a temptation payoff T.
3. Player A chooses D. Since Player B's last choice was d, both players receive a punishment payoff P.
4. Player B chooses c. Since player A's last choice was D, A receives a temptation payoff T, and B receives a sucker payoff S.

Thus after the sequence CdDc, each player has received a total payoff of $T+P+S$. This process continues indefinitely, with players receiving payoffs based on the last two turns.

Though some have argued that the IPD and the APD are mathematically equivalent (Axelrod, 1984), the two games can result in dramatically different interactions between strategies. For example, in an IPD game between two Tit for Tat players, a single accidental defection leads to a sequence of alternating temptation and sucker payoffs, while in an APD game an error leads to a sequence of punishment payoffs. As a result, strategies may be successful in IPD but not APD interactions, or vice versa. For example, two Pavlov players quickly restore mutual cooperation after an error in the IPD (CD DD CC CC...), but in the APD an error leads to a 6-turn cycle of temptation, punishment, and sucker payoffs (CdDcDdCdDcDd...). Thus the Pavlov strategy is "self-cooperating" in the IPD but not the APD. The *Firm but Fair (FBF)* strategy, which cooperates except after receiving a sucker payoff, is self-coooperating in the APD but not the IPD (Frean, 1994). Two FBF players can restore mutual cooperation after an error in the APD (CdDcCc...), but in the IPD an error leads to alternating temptation and sucker payoffs (CD DC CD DC...). From these examples, it is clear that the IPD and APD are distinct problems, and "optimal" strategies for one game may perform poorly in the other.

2 N-ply strategies for the Alternating Prisoner's Dilemma

We now define some strategies which are commonly discussed in the Alternating Prisoner's Dilemma literature, and introduce the general notion of an "n-ply" strategy. The *n-move history* of a game is a string of C's and D's $H = h_n h_{n-1} \dots h_2 h_1$, where $h_k = C$ if a player cooperated k turns ago, and $h_k = D$ if a player defected k turns ago. Odd indices correspond to the opponent's moves, and even indices correspond to the player's moves. The opponent's moves are generally written in lower case, and the player's moves in upper case. For example, a 3-move history of cCd would mean that the opponent's most recent move was defection, the player's most recent move was cooperation, and the opponent's previous move was cooperation.

We can enumerate the 2^n possible history strings H in lexicographic order, with $H_1 = C \cdots Cc$, $H_2 = C \cdots Cd$, $\dots$, $H_{2^n - 1} = D \cdots Dc$, $H_{2^n} = D \cdots Dd$. An *n-ply strategy* for the APD chooses whether to cooperate or defect based on the n-move history H. The strategy is defined by 2^n numbers $(a_1, \dots, a_{2^n})$, where $0 \leq a_i \leq 1$ for all i, and each a_i corresponds to the probability of cooperation if $H = H_i$.

For example, the two ply strategy (1 0 1 0.2) will cooperate with probability 1 if $H = H_0$ or $H = H_2$, that is, if the two-move history is Cc or Dc. Similarly, it will defect if the two-move history is Cd, and cooperate with probability 0.2 if Dd.

Commonly discussed 1-ply strategies include ALLC (1 1), ALLD (0 0), TFT (1 0), and Random (1/2 1/2). Commonly discussed 2-ply strategies include Pavlov (1 0 0 1), Firm But Fair (1 0 1 1), and the "generous strategies" g1 (1 1/3 1 1/2), g2 (1 1/2 1 1/2), and g3 (1 2/3 1 1/2). We can also form a 2-ply strategy by trivial extension of a 1-ply strategy: the 2-ply strategy (a b a b) behaves equivalently to the one ply strategy (a b). In general, the n-ply strategy S_n is equivalent to the $n+1$-ply strategy formed by concatenating S_n with itself: $S_{n+1} = S_n \oplus S_n$, where $\oplus$ is the concatenation operator.

To fully specify an n-ply strategy, we should also specify its behavior at the start of the game (turns $1..n$). We generally assume that $h_k = C$ for $k \geq i$, that is, "moves" in the history list before the start of the game are assumed to be cooperation. For example, on the first turn of a game, a 3-ply strategy will assume that the history is $H = cCc$. Alternatively, we can specify how the strategy behaves at the start of the game (ex. a Pavlov strategy that defects on the first move), or take the average over all possible initial sequences. In general, the behavior of an n-ply strategy at the start of a game is only relevant for the no-noise case.

3 Optimality criteria

Before defining the optimality criteria, we first attempt to gain a general idea of what it means for a strategy to be "successful" in a Alternating Prisoner's Dilemma interaction. We consider a single APD game, with a finite number N of turns, between two strategies X and Y. The goal of each strategy is to maximize its average payoff per turn,

$$w = \frac{1}{N-1} \sum_{i=2}^{N} p(i)$$

where $p(i)$ is the payoff received by that strategy after turn i. Similarly, in an APD game with an infinite number of turns, each strategy attempts to maximize its average payoff,

$$w = \lim_{N \to \infty} \frac{1}{N-1} \sum_{i=2}^{N} p(i)$$

We denote the expected value of the payoff w to strategy X in an infinitely long game against strategy Y as $w(X|Y)$. In particular, we define $w(X|X)$, the "self-payoff" of strategy X, to be the expected value of its payoff w in an infinitely long game against its clone. In certain cases (ie. interactions between deterministic strategies with no noise), payoffs may vary depending on which strategy goes first. In these cases, we average the two possibilities.

We assume that an APD strategy X will interact with many different strategies Y over time, and thus the strategy's goal is to maximize its average payoff against all strategies it interacts with. Assuming a finite set of opponents $Y = \{Y_1..Y_M\}$:

$$w(X) = \frac{1}{M} \sum_{i=1}^{M} w(X|Y_i)$$

In a population of opponents Y_i with varying frequencies $f(Y_i)$, with $\sum_i f(Y_i) = 1$:

$$w(X) = \sum_i f(Y_i) w(X|Y_i)$$

In a continuous "population space" S of opponents:

$$w(S) = \int_S f(s) w(X|s) \mathrm{d}s$$

The relative success of two strategies X and Y can be evaluated by comparing $w(X)$ and $w(Y)$: strategies with higher w are considered to be more successful. However, it is clear from these measures that a strategy's success in a given APD interaction depends highly on which other strategies it faces. Our goal is to find strategies that perform well in interactions with many other sets of strategies, and hence we attempt to find optimality criteria that are strongly correlated with high average performance.

In general, an APD strategy X should choose to cooperate with a strategy Y if cooperating will make Y significantly more likely to cooperate on future turns. As a simple approximation to this, imagine that strategy Y is the 1-ply strategy $(a\ b)$: that is, it cooperates with probability a after X's cooperation and probability b after X's defection. Then strategy X could score an average payoff of $Ra+S(1-a)$ by cooperating continually, and an average payoff of $Tb+P(1-b)$ by defecting continually. Hence long-term cooperation is reasonable if:

$$Ra + S(1-a) > Tb + P(1-b) \implies a > \left(\frac{T-P}{R-S}\right) b + \left(\frac{P-S}{R-S}\right)$$

Of course, continual cooperation and continual defection are not a strategy's only options: against many strategies the highest payoff can be achieved through a mixture of cooperation and defection. Nevertheless, the general principle holds: cooperate with strategies only if they are likely to reciprocate cooperation with cooperation, and defection with defection. In particular, an "optimal" strategy should be able to exploit unconditional cooperators, as well as preventing other strategies from exploiting it. It should also cooperate with nice, but unexploitable, strategies such as Tit for Tat. Thus one possibility for optimality criteria would be to consider a strategy's scores against ALLC, ALLD, and TFT. However, since TFT performs poorly in games with noise, and choosing any specific variation of TFT adds a subjective bias to our criteria, we choose instead to consider a strategy's score against its clone. Strategies which can achieve cooperation with their clone are likely to cooperate with a wide range of other strategies, while strategies which do not cooperate with their clone are unlikely to achieve cooperation with most other strategies.

Combining these desired characteristics, we find that an "optimal" strategy must possess three essential properties. It must be "self-cooperating" (able to achieve mutual cooperation with its clone), "unexploitable" (able to resist exploitation by ALLD and other strategies), and "exploiting" (able to exploit unconditional cooperators).

These properties are also essential for success in the "evolutionary" APD. Maynard Smith and Price (1973) introduced ideas from evolutionary biology into game theory, dramatically changing the concept of "optimality" by allowing strategies to evolve over time. In the evolutionary APD, a strategy's payoff w is used as a measure of its reproductive fitness. Strategies which receive higher payoffs are able to produce more offspring, while those with poor performance quickly die off. Thus the population of strategies (and hence, the average payoff a strategy receives against this population) changes over time. The "success" of a strategy in the evolutionary APD can be measured by its proportion of the total population.

In the evolutionary APD, an "optimal" strategy should be able to invade and resist invasion by other strategies. According to Maynard Smith's definition (1982), strategy X *invades* strategy Y if $w(X|Y) > w(Y|Y)$, or if $w(X|Y) = w(Y|Y)$ and $w(X|X) > w(Y|X)$. We will discuss the evolutionary APD, defining a more precise measure of evolutionary "dominance", but for now we simply consider a strategy's evolutionary interactions with unconditional cooperators (the $ALLC$ strategy) and unconditional defectors (the $ALLD$ strategy). An optimal strategy should be able to evolve cooperation by invading a population of defectors and establishing a cooperative "society". Once established, this society should be "stable" enough to resist invasion by defectors, as well as resisting *genetic drift*, the infiltration of the population by unconditional cooperators. Strategies which are unexploitable and self-cooperating can establish a stable cooperative society of this sort, and strategies which are exploiting can resist genetic drift.

We now define and examine the three "optimality criteria" for APD games with varying amounts and types of noise.

3.1 Optimality criteria for no noise APD

In the no noise APD, we denote a payoff $w(X|Y)$ by $w_0(X|Y)$ in order to distinguish it from payoffs in APD games with noise.

We first define a "self-cooperating" strategy in terms of its self-payoff $w_0(X|X)$. $w_0(X|X)$ must be between P and R inclusive.

Definition 1 *A strategy X is* self-cooperating *under no noise if* $w_0(X|X) > P$. *A strategy X is* totally self-cooperating *under no noise if* $w_0(X|X) = R$.

Second, we define a "C-exploiting" strategy in terms of its payoff against ALLC. $w_0(X|ALLC)$ must be between R and T inclusive.

Definition 2 *A strategy X is* C-exploiting *under no noise if* $w_0(X|ALLC) > R$.
A strategy X is totally C-exploiting *under no noise if* $w_0(X|ALLC) = T$.

Third, we define a "D-unexploitable" strategy in terms of its payoff against ALLD. $w_0(X|ALLD)$ must be between S and P inclusive.

Definition 3 *A strategy X is* D-unexploitable *under no noise if* $w_0(X|ALLD) > S$.
A strategy X is totally D-unexploitable *under no noise if* $w_0(X|ALLD) = P$.

More generally, we define the "relative performance" of strategy X against strategy Y to be:

$$\sigma_Y(X) = \frac{w(X|Y) - \inf_Z w(Z|Y)}{\sup_Z w(Z|Y) - \inf_Z w(Z|Y)}$$

Thus $\sigma_Y(X) = 1$ if X achieves the maximum possible score against Y, and $\sigma_Y(X) = 0$ if X achieves the minimum possible score against Y. We now consider the "relative performance" of a strategy X against ALLC and ALLD, which we denote by σ_C and σ_D respectively:

$$\sigma_C = \sigma_{ALLC}(X) = \frac{w(X|ALLC) - R}{T - R}$$

$$\sigma_D = \sigma_{ALLD}(X) = \frac{w(X|ALLD) - S}{P - S}$$

Thus a strategy is "C-exploiting" if $\sigma_C > 0$, and "totally C-exploiting" if $\sigma_C = 1$. Similarly, a strategy is "D-unexploitable" if $\sigma_D > 0$, and "totally D-unexploitable" if $\sigma_D = 1$. Note that each of these quantities is the proportion of defections by X against ALLC or ALLD respectively.

We also define the "relative self-performance" of a strategy to be:

$$\sigma_S = \frac{w(X|X) - P}{R - P}$$

Note that this quantity is not necessarily equal to $\sigma_X(X)$. For example, consider the Bully strategy B=(0 1). For this strategy: $\inf_Z w(Z|B) = w(ALLC|B) = S$, $\sup_Z w(Z|B) = w(ALLD|B) = T$, and $w(B|B) = (S + T)/2$. Thus for the (5,3,1,0) payoff table we calculate:

$$\sigma_S = \frac{w(B|B) - P}{R - P} = \frac{2.5 - 1}{3 - 1} = 3/4$$

$$\sigma_B(B) = \frac{w(B|B) - S}{T - S} = \frac{2.5 - 0}{5 - 0} = 1/2$$

Note that a strategy is "self-cooperating" if $\sigma_S > 0$, and "totally self-cooperating" if $\sigma_S = 1$.

Now we consider what σ values a strategy must have to be considered "optimal". A "perfectly optimal" strategy X should attain the maximum possible score against all strategies Y: that is, $\sigma_Y(X) = 1$ for all Y. However, a simple argument shows that this is impossible to achieve:

Proposition 1 *No perfectly optimal strategy exists.*

Proof: Consider the *GRIM* strategy, which cooperates until the opponent defects once, then defects for the rest of the game. No strategy can achieve the maximum possible score against both GRIM and ALLC: if the strategy ever defects after mutual cooperation, it will perform poorly against GRIM ($\sigma_{GRIM}(X) = 0$), and if it never defects after mutual cooperation, it will have a low relative performance against ALLC ($\sigma_C = 0$).

We now consider strategies which are optimal in their interactions with cooperators and defectors. To quantify this, we apply the Maynard Smith criteria. We first note that the following hold for all X, allowing us to simplify our calculations considerably:
$w(X|ALLC) \geq w(ALLC|ALLC)$
$w(ALLC|X) \leq w(ALLC|ALLC)$
$w(X|ALLD) \leq w(ALLD|ALLD)$
$w(ALLD|X) \geq w(ALLD|ALLD)$.

For X to invade ALLC, $w(X|ALLC) > w(ALLC|ALLC)$, which implies $\sigma_C > 0$. For X to invade ALLD, $w(X|ALLD) = w(ALLD|ALLD)$ and $w(X|X) > w(ALLD|X)$, which implies $\sigma_D = 1$ and $\sigma_S > 0$. For X to resist invasion by ALLC, $w(X|X) \geq w(ALLC|X)$. Solving for $w(X|X)$ in terms of σ_S, and $w(ALLC|X)$ in terms of σ_C, we obtain:

$$R - (R-S)\sigma_C \leq P + (R-P)\sigma_S \longrightarrow \sigma_C \geq \frac{R-P}{R-S}(1-\sigma_S)$$

For X to resist invasion by ALLD, $w(X|X) > w(ALLD|X)$. Solving for $w(X|X)$ in terms of σ_S, and $w(ALLD|X)$ in terms of σ_D, we obtain:

$$T - (T-P)\sigma_D < P + (R-P)\sigma_S \longrightarrow \sigma_D > 1 - \frac{R-P}{T-P}\sigma_S$$

We define a "strongly optimal" strategy as one which can invade and take over a population of defectors, establishing total self-cooperation, and resist invasion by ALLD and ALLC. A "weakly optimal" strategy can invade and take over a population of defectors, establishing some degree of self-cooperation, and resist invasion by ALLD and ALLC. This implies:

Definition 4 *A* strongly optimal *strategy is a strategy with $\sigma_D = 1$, $\sigma_S = 1$, and $\sigma_C > 0$.*

Definition 5 *A* weakly optimal *strategy is a strategy with $\sigma_D = 1$, $\sigma_S > 0$, and $\sigma_C > \frac{R-P}{R-S}(1-\sigma_S)$.*

We now consider the strategies which are most commonly discussed in the Prisoner's Dilemma: Tit for Tat (TFT), Generous Tit For Tat (GTFT), ALLC, ALLD, Pavlov (PAV), and Firm But Fair (FBF). The four numbers in parentheses are the probabilities of cooperation after R,S,T,P payoffs, and g is the "generosity" of the GTFT or FBF strategy. We then compute σ_S, σ_C, and σ_D for each strategy in an APD with no noise:

Strategy	σ_S	σ_C	σ_D
TFT (1 0 1 0)	1	0	1
GTFT (1 g 1 g)	1	0	$1-g$
ALLC (1 1 1 1)	1	0	0
ALLD (0 0 0 0)	0	1	1
PAV (1 0 0 1)	1	0	1/2
FBF (1 0 1 g)	1	0	$\frac{1}{1+g}$

Table 2: σ values for no noise APD

As can be seen from the table, none of these strategies are even weakly optimal. TFT is closest to optimal, since it can invade a population of defectors and establish complete cooperation. However, since TFT does not exploit unconditional cooperators, it is vulnerable to infiltration by ALLC, allowing defectors to re-invade. It should be noted that all of these strategies have a "2-ply memory" or less: they react based only on the last two turns (strategy's move followed by opponent's move). In fact, no strategy with memory less than three ply is strongly optimal:

Lemma 1 *For a 2-ply strategy X=(a b c d) with $d \neq 0$, ALLD can exploit the strategy with a probability of at least $\frac{d}{1+d}$.*

Proof: On any given turn, X will either cooperate (earning payoff S) or defect (earning payoff P). X will cooperate with probability d the turn after defecting, or with probability b the turn after cooperating. Since any non-zero b will only increase the probability of being exploited, we may assume that $b = 0$, and X will always defect the turn after cooperating. Thus X spends $1/d$ turns defecting for each turn it spends cooperating, so its probability of being exploited is:

$$\frac{1}{1+1/d} = \frac{d}{1+d}$$

Proposition 2 *No 0, 1, or 2-ply strategy is strongly optimal in the APD with no noise.*

Proof: For any 0-ply strategy S_0=(a), S_0 is equivalent to the 2-ply strategy (a a a a). For any 1-ply strategy S_1=$(a\ b)$, S_1 is equivalent to the 2-ply strategy (a b a b). Now consider the 2-ply strategy (a b c d) with initial cooperation probability e. If $b = 1$, ALLD can exploit the strategy continually. Similarly, if $d \neq 0$, we know that ALLD can exploit the strategy with a probability of at least $\frac{d}{1+d}$, and $\sigma_D \leq \frac{1}{1+d}$. Thus for $\sigma_D = 1$, we know $b < 1$ and $d = 0$. But how can a strategy with $b < 1$ and $d = 0$ be totally self-cooperating? Any defection has a non-zero probability $1 - b$ of leading to an unending rut of mutual defection, so we must ensure that no defections occur: $a = 1$ and $e = 1$. But a strategy with $a = 1$ and $e = 1$ cooperates continually against ALLC, so no 2-ply strategy with $\sigma_D = \sigma_S = 1$ can have $\sigma_C > 0$.

A 2-ply strategy may be weakly optimal: the GRIM strategy (1 0 0 0) with initial cooperation probability $1/3$ has $\sigma_S = 1/9$, $\sigma_C = 2/3$, and $\sigma_D = 1$, and thus meets the criteria for weak optimality. In order to find a strongly optimal strategy, however, we must consider strategies of 3-ply memory or higher.

3.2 Optimality criteria for infinitesimal noise APD

We now consider APD games with noise. Noise is defined as a non-zero probability ϵ that an error occurs on a given turn: a strategy will defect when it meant to cooperate, or cooperate when it meant to defect, with probability ϵ. In an APD game with noise ϵ, we denote the payoff $w(X|Y)$ by $w_\epsilon(X|Y)$.

In an APD game with "infinitesimal noise", it is possible for a player's decision to be transmitted incorrectly, but the probability of this occuring is extremely low. The infinitesimal noise case can be thought of as the limiting case of finite noise, with a noise probability $\epsilon \to 0$, or the results can be applied as an approximation for small $\epsilon > 0$. Thus the payoff $w(X|Y)$ in an APD game with infinitesimal noise is $\lim_{\epsilon\to 0} w_\epsilon(X|Y)$. We extend our definitions of "self-cooperating", "D-unexploitable", and "C-exploiting" strategies to the infinitesimal noise case:

Definition 6 *A strategy X is* self-cooperating *under infinitesimal noise if*
$\lim_{\epsilon\to 0} w_\epsilon(X|X) > P$.
A strategy X is totally self-cooperating *under infinitesimal noise if*
$\lim_{\epsilon\to 0} w_\epsilon(X|X) = R$.

Definition 7 *A strategy X is* C-exploiting *under infinitesimal noise if*
$\lim_{\epsilon\to 0} w_\epsilon(X|ALLC) > R$.
A strategy X is totally C-exploiting *under infinitesimal noise if*
$\lim_{\epsilon\to 0} w_\epsilon(X|ALLC) = T$.

Definition 8 *A strategy X is* D-unexploitable *under infinitesimal noise if*
$\lim_{\epsilon\to 0} w_\epsilon(X|ALLD) > S$.
A strategy X is totally D-unexploitable *under infinitesimal noise if*
$\lim_{\epsilon\to 0} w_\epsilon(X|ALLD) = P$.

We also extend our definitions of σ_S, σ_C, and σ_D for the infinitesimal noise case as follows:

$$\sigma_S = \frac{\lim_{\epsilon\to 0} w_\epsilon(X|X) - P}{R - P}$$

$$\sigma_C = \frac{\lim_{\epsilon\to 0} w_\epsilon(X|ALLC) - R}{T - R}$$

$$\sigma_D = \frac{\lim_{\epsilon\to 0} w_\epsilon(X|ALLD) - S}{P - S}$$

The optimality criteria "strongly optimal" and "weakly optimal" are defined in terms of σ_S, σ_C, and σ_D as in the no-noise case. Recall that a strongly optimal strategy has a $\sigma_D = 1$, $\sigma_S = 1$, and $\sigma_C > 0$. A weakly optimal strategy has $\sigma_D = 1$, $\sigma_S > 0$, and $\sigma_C > \frac{R-P}{R-S}(1 - \sigma_S)$.

It is important to note that $w(X|Y)$ can have a discontinuity at $\epsilon = 0$: in other words, the performance of a strategy can differ significantly between the no noise case ($\epsilon = 0$) and the infinitesimal noise case ($\epsilon \to 0$). For example, consider the behavior of the Tit for Tat strategy, in an APD game with its clone. For the no-noise case, Tit for Tat and its clone will cooperate continually, giving $\sigma_S = 1$. For the infinitesimal noise case, a single error will lead to a rut of mutual defection with payoff P, while a second error will restore mutual cooperation. In an infinitely long game we can expect TFT and its clone to be in the mutual defection rut half the time, giving each an average payoff of $\frac{R+P}{2}$. Thus $\sigma_S = 1/2$ for the TFT strategy. We use similar reasoning to compute σ_S, σ_C, and σ_D for TFT, GTFT, ALLC, ALLD, PAV, and FBF. We assume that $g \gg \epsilon$ for the GTFT and FBF strategies (see Table 3).

Strategy	σ_S	σ_C	σ_D
TFT	1/2	0	1
GTFT	1	0	$1-g$
ALLC	1	0	0
ALLD	0	1	1
PAV	1/2	1/2	1/2
FBF	1	0	$\frac{1}{1+g}$

Table 3: σ values for infinitesimal noise APD

Thus we can see that the lower memory strategies (TFT, GTFT, ALLC, ALLD, PAV, and FBF) are not even weakly optimal. As in the no-noise case, we can prove that no strategy of less than three ply is strongly optimal. But for infinitesimal noise, we can go even further, proving that no lower memory strategy is even weakly optimal:

Proposition 3 *No 0, 1, or 2-ply strategy is optimal in the infinitesimal noise APD.*

Proof: As in the no-noise case, we consider the 2-ply strategy $(a\ b\ c\ d)$. We will show that the strategy cannot have $\sigma_D = 1$, $\sigma_S > 0$, and $\sigma_C > 0$, and hence is not even weakly optimal. For $\sigma_D = 1$ to hold, the strategy's probability of entering the Dd rut against ALLD must be greater than its probability of leaving the Dd rut by $O(\epsilon)$. Thus we must have $1 - b = O(1)$ and $d = O(\epsilon)$, which implies $b < 1$ and $d = 0$. For $\sigma_C > 0$ to hold, the strategy's probability of entering the Dc rut against ALLC must *not* be less than its probability of leaving the Dc rut by $O(\epsilon)$. Thus we must have $1 - a = O(1)$ or $c = O(\epsilon)$, which implies $a < 1$ or $c = 0$. This gives us three possibilities for $\sigma_D = 1$ and $\sigma_C > 0$:

1. $a < 1$, $b < 1$, $c > 0$, $d = 0$. Against its clone, this strategy has probability $O(1)$ of entering a Dd rut and probability $O(\epsilon)$ of leaving the Dd rut, so $\sigma_S = 0$.
2. $a < 1$, $b < 1$, $c = 0$, $d = 0$. Against its clone, this strategy has probability $O(1)$ of entering a Dd rut and probability $O(\epsilon^2)$ of leaving the Dd rut, so $\sigma_S = 0$.
3. $a = 1$, $b < 1$, $c = 0$, $d = 0$. Against its clone, this strategy has probability $O(\epsilon)$ of entering a Dd rut and probability $O(\epsilon^2)$ of leaving the Dd rut, so $\sigma_S = 0$.

Thus no totally D-unexploitable 2-ply strategy can be both self-cooperating and C-exploiting. Therefore, in order to find optimal strategies for the infintesimal noise APD, we must examine strategies with 3-ply or higher memory.

3.3 Optimality criteria for finite noise APD

We now consider APD games with a noise level $0 < \epsilon < 1/2$, examining how a strategy's payoff changes as a function of ϵ. This enables us to analyze games where the approximation $\epsilon \approx 0$ does not hold.

First, it should be noted that the expected average payoff from continued "mutual cooperation" (ie. when both strategies intend to cooperate) is less than R in the presence of noise. If two strategies are intending to cooperate with an error level ϵ, this means that each will cooperate with probability $1 - \epsilon$, and defect with probability ϵ. Thus the expected payoff per turn to each strategy in mutual cooperation is:

$$E_{cc} = (1-\epsilon)^2 R + \epsilon^2 P + \epsilon(1-\epsilon)(T+S)$$

Similarly, the expected average payoff from a run of continued "mutual defection" (ie. when both strategies intend to defect) is greater than P in games with finite noise. If two strategies are defecting with an error level ϵ, each will cooperate with probability ϵ, and defect with probability $1 - \epsilon$. Thus the expected payoff to each strategy in mutual defection is:

$$E_{dd} = (1-\epsilon)^2 P + \epsilon^2 R + \epsilon(1-\epsilon)(T+S)$$

We can also compute E_{dc}, the expected payoff for a strategy which is continually exploiting another strategy, and E_{cd}, the expected payoff for the strategy being exploited:

$$E_{dc} = (1-\epsilon)^2 T + \epsilon^2 S + \epsilon(1-\epsilon)(R+P)$$

$$E_{cd} = (1-\epsilon)^2 S + \epsilon^2 T + \epsilon(1-\epsilon)(R+P)$$

We now compute explicit values of these quantities for two commonly used payoff tables (T,R,P,S). For (5,3,1,0) we know $E_{cc} = 3-\epsilon-\epsilon^2$, $E_{dd} = 1+3\epsilon-\epsilon^2$, $E_{dc} = 5-6\epsilon+\epsilon^2$, $E_{cd} = 4\epsilon+\epsilon^2$. For (4,3,1,0) we know $E_{cc} = 3-2\epsilon$, $E_{dd} = 1+2\epsilon$, $E_{dc} = 4-4\epsilon$, and $E_{cd} = 4\epsilon$.

Next, we extend the definitions of σ_S, σ_C, and σ_D for the finite noise case, replacing R by E_{cc}, P by E_{dd}, T by E_{dc}, and S by E_{cd}.

$$\sigma_S = \frac{w_\epsilon(X|X) - E_{dd}}{E_{cc} - E_{dd}}$$

$$\sigma_C = \frac{w_\epsilon(X|ALLC) - E_{cc}}{E_{dc} - E_{cc}}$$

$$\sigma_D = \frac{w_\epsilon(X|ALLD) - E_{cd}}{E_{dd} - E_{cd}}$$

We note that $\lim_{\epsilon \to 0} \sigma_S(\epsilon)$ is equal to the value of σ_S defined for infinitesimal noise, as expected. This is also true for σ_C and σ_D. These definitions allow us to obtain the expected σ values for ALLC and ALLD. For ALLC, $\sigma_S = 1$ and $\sigma_C = \sigma_D = 0$. For ALLD, $\sigma_S = 0$ and $\sigma_C = \sigma_D$=1. For nearly all strategies, however, the σ values vary as a function of ϵ. In most cases, σ_C increases, and σ_S and σ_D decrease, with increasing ϵ.

For example, we can compute the σ values for FBF (1 0 1 1) as a function of ϵ (see Table 4). Recall that for infinitesimal noise, $\sigma_S = 1$, $\sigma_C = 0$, and $\sigma_D = 0.5$.

ϵ	σ_S	σ_C	σ_D
.001	.999	.001	.499
.010	.990	.007	.495
.050	.953	.034	.475
.100	.913	.064	.450
.200	.843	.111	.396

Table 4: σ values for FBF as a function of ϵ

Using Markov chains, we can compute values of σ_S, σ_C, and σ_D for a strategy as a function of ϵ. For example, for the (5,3,1,0) payoff table, we compute:

$$w(FBF|FBF) = \frac{3 + 3\epsilon}{1 + 2\epsilon}$$

Substituting this value into the expression for σ_S, we know:

$$\sigma_S = \frac{2 + 2\epsilon - \epsilon^2}{2 + 4\epsilon}$$

Now, since the σ values of a strategy vary with ϵ, how do we determine if a strategy is self-cooperating, C-exploiting, or D-unexploitable with finite noise? One possibility is to choose a constant "noise resistance threshold" NRT. Then for any given ϵ, we define:

Definition 9 *A strategy is* self-cooperating *at noise* ϵ *if* $\sigma_S \geq NRT$.

Definition 10 *A strategy is* C-exploiting *at noise* ϵ *if* $\sigma_C \geq NRT$.

Definition 11 *A strategy is* D-unexploitable *at noise* ϵ *if* $\sigma_D \geq NRT$.

Using a noise resistance threshold of $NRT = 0.9$ and the $T = 5$ payoff table, we find that FBF is self-cooperating for $\epsilon \leq 0.117$, but is not D-unexploitable. TFT, on the other hand, is D-unexploitable for $\epsilon \leq .0697$, but is not self-cooperating.

In addition to defining the optimality criteria for specific values of ϵ, we also examine how a given payoff $w_\epsilon(X|Y)$ changes as a function of ϵ. In particular, we define a "noise resistance coefficient" NRC, which gives a lower bound for w_ϵ as a linear function of ϵ. This means that $w_{\epsilon=e} \geq w_{\epsilon\to 0} - (NRC)e$ for all e. We assume that the noise level can vary from 0 to some maximum value ϵ_{max}: for our experiments, we assume ϵ_{max}=0.2. Thus we define:

$$NRC = \sup_{(e=0\ldots\epsilon_{max})} \frac{w_{\epsilon\to 0} - w_{\epsilon=e}}{e}$$

We now consider strategies that are totally self-cooperating under infinitesimal noise. For these strategies, $w(X|X)_{\epsilon\to 0} = R$. We define NRC_S, the noise resistance coefficient for self-cooperation, to be:

$$NRC_S = \sup_{(e=0\ldots\epsilon_{max})} \frac{R - w(X|X)_{\epsilon=e}}{e}$$

Once we have found NRC_S, this gives us a lower bound on the performance of a strategy for all $\epsilon \leq \epsilon_{max}$: we know $w(X|X) \geq R - (NRC_S)\epsilon$. Thus if a strategy has $NRC_S = 5$, and assuming $R = 3$, we know $w(X|X) \geq 3 - 5\epsilon$. It should also be noted that, for strategies which are not totally self-cooperating under infinitesimal noise, $NRC_S = \infty$.

It is clear that a strategy with lower NRC_S is more resistant to noise. How low must NRC_S be for a strategy to be considered "finite noise resistant"? To answer this question, we consider the results of a single error on the total self-payoff of a strategy. For ALLC against its clone, an error results in the pattern CdCcCc: the defector gains $(2T - 2R)$ and the cooperator loses $(2R - 2S)$ as the result of this error, for an average loss of $2R - (T + S)$ per player. For the (5,3,1,0) payoff table, each player loses an average of $2(3) - (5 + 0) = 1$ point per error, and thus $w \approx 3 - \epsilon$, or $NRC_S \approx 1$. For FBF against its clone, an error results in the pattern CdDcCc: each player loses $3R - (T + S + P) = 3$ points per error, and thus $w \approx 3 - 3\epsilon$, or $NRC_S \approx 3$. In general, if a single error leads to N straight defections on average, each player loses $(N + 1)R - (N - 1)P - (T + S)$ points per error. We define a strategy to be "finite noise resistant" if one error leads to an average of two defections or less. This implies:

Definition 12 *A strategy is* finite noise resistant *if* $NRC_S \leq 3R - P - T - S$.

For the (5,3,1,0) payoff table, $NRC_S = 1 + 2(N-1)$, and thus a strategy is finite noise resistant if $NRC_S \leq 3$. For the (4,3,1,0) payoff table, $NRC_S = 2 + 2(N-1)$, and thus a strategy is finite noise resistant if $NRC_S \leq 4$. We now propose a new standard for optimality under finite noise:

Definition 13 *A strategy X is* finite noise optimal *if it is finite noise resistant, totally D-unexploitable under infinitesimal noise, and C-exploiting under infinitesimal noise.*

Thus a finite noise optimal strategy must meet all the criteria for strong optimality under infinitesimal noise, as well as having $NRC_S \leq 3$ (for $T = 5$) or $NRC_S \leq 4$ (for $T = 4$). No strategy discussed so far meets these criteria. We will attempt to find such strategies later, after defining the FBF and FP strategy classes.

3.4 Optimality criteria for arbitrary noise APD

Consider first the 3-ply strategy X=(1 0 1 1, 1 0 1 0), in a game against its clone with finite random noise ϵ. If two errors occur in close succession, it will be stuck in a "Dd rut", unable to restore cooperation until another error occurs: CcCc<u>D</u>d<u>D</u>dDd...Dd<u>C</u>cCc. In general, it enters the Dd rut an average of once per $1/\epsilon^2$ moves, and spends $1/\epsilon$ moves in the Dd rut. Since the proportion of time spent in the Dd rut is approximately ϵ, $\lim_{\epsilon \to 0} w_\epsilon(X|X) = R$, so X is totally self-cooperating under infinitesimal noise. Also, one error results in an average of 2 defections, so X is also finite noise resistant.

In real life examples, however, the probability of an error may vary over the course of the game. Consider again strategy X against its clone. In the worst case scenario, errors occur on the first and third moves of the game, then never again: <u>D</u>d<u>D</u>dDd ...

Cooperation is not restored, and both players receive the punishment payoff indefinitely. A strategy that is truly "noise-proof" (rather than just "noise resistant") should be able to restore self-cooperation after any finite sequence of errors. To formalize this definition, we first define an "error sequence". An *error sequence* ω is a finite-length string of zeros and ones $\omega_1 \omega_2 \dots \omega_n$, where $\omega_i = 1$ if an error occurs on the ith turn. The total number of errors occurring, $\sum_{i=1}^{n} \omega_i$, is the norm Ω of the error sequence. For a given error sequence ω, we define a strategy's error response D(ω) to be the expected total number of defections in a game against itself, given no other errors:

$$D(\omega) = \sum_{k=0}^{\infty} k \mathrm{Prob}(D = k)$$

We then define a strategy's "noise proof coefficient" NPC_S to be the maximum number of defections caused per error:

$$NPC_S = \max_{\forall \omega} \frac{D(\omega)}{\Omega}$$

We define a strategy to be "noise proof" if no error sequence causes more than two defections per error:

Definition 14 *A strategy is* noise proof *if it has* $NPC_S \leq 2$.

It is clear from this definition that a noise proof strategy must also be finite noise resistant, and hence, totally self-cooperating under infinitesimal noise.

We now consider the NPC_S of various strategies. For ALLC, the number of defections $D(\omega)$ caused by an error sequence ω is equal to its number of errors Ω. Thus $NPC_S = 1$ for ALLC, and ALLC is noise proof. For TFT, a single error ($\omega = 1$) results in continuing defection $D(\omega) = \infty$. Thus $NPC_S = \infty$ for TFT, and TFT is not noise proof. For FBF, on the other hand, a single error results in two defections, and no sequence of defections will result in more than two defections per error. Thus $NPC_S = 2$ for FBF, and FBF is noise proof. When we consider FBF strategies with varying generosity, (1 0 1 g) for $0 \leq g \leq 1$, we find that $NPC_S = \max(1 + 1/g, 2)$. Thus FBF(g) is noise proof only for $g = 1$. Finally, we prove:

Proposition 4 *No strategy X with finite memory is noise proof and totally D-unexploitable.*

Proof: Assume that strategy X has n-ply memory for some finite n. Now consider the strategy's generosity g, its probability of cooperating after a sequence of n defections. If $g = 0$, then in a game against its clone, any sequence of n defections will result in an endless rut of mutual defection. In other words, for some error sequence ω with $\Omega \leq n$, we know that D(ω)=∞. Thus the strategy has $NPC_S = \infty$, and is not noise proof. If $g > 0$, then strategy X can be exploited by ALLD at least some fraction of the time. We can compute a lower bound on this fraction by assuming that the strategy will defect against ALLD except after a sequence of n defections. Then each exploitation by ALLD results in $\lfloor n/2 \rfloor - 1$ defections by strategy X (to return to the n defections state), plus an average of $1/g$ defections before X cooperates again. Thus the strategy will defect against ALLD with probability:

$$\sigma_D = \frac{\lfloor \frac{n}{2} \rfloor - 1 + \frac{1}{g}}{\lfloor \frac{n}{2} \rfloor + \frac{1}{g}} = \frac{(\lfloor \frac{n}{2} \rfloor - 1)g + 1}{\lfloor \frac{n}{2} \rfloor g + 1}$$

Thus if $g > 0$, the strategy has $\sigma_D < 1$, and is not totally D-unexploitable.

Only infinite memory strategies can be noise proof and totally D-unexploitable, and thus we define:

Definition 15 *An infinite memory strategy which is noise proof, totally D-unexploitable, and C-exploiting, is* arbitrary noise optimal.

4 A new class of strategies

Thus we are presented with two main tasks: to find a finite memory strategy which is "finite noise optimal", and to find an infinite memory strategy which is "arbitrary noise optimal". Each strategy must be totally D-unexploitable, totally self-cooperating, and (at least weakly) C-exploiting under infinitesimal noise. In other words, each must have $\sigma_S = \sigma_D = 1$ and $\sigma_C > 0$ for $\epsilon \to 0$. They must also be "finite noise resistant", having $NRC_S \leq 3$ for the (5,3,1,0) payoff table and $NRC_S \leq 4$ for the (4,3,1,0) payoff table. The "arbitrary noise optimal" strategy must meet an additional criterion: it must be "noise proof", having $NPC_S \leq 2$.

4.1 First defector strategies

We now define the "first defector strategies", a class of infinite memory strategies. An *n-ply first defector strategy* for the APD chooses to cooperate or defect based on the n-move history H, together with a "memory bit" b. This bit records which player most recently made an unprovoked defection: it is set to 1 whenever the 2-move history is Cd (opponent made an unprovoked defection) and set to 0 whenever the 3-move history is cDc or cDd (player made an unprovoked defection). The strategy is defined by $2^n + 1$ numbers $(a_1, \ldots, a_{2^n+1})$, where $0 \leq a_i \leq 1$ for all i. For $i = 1 \ldots 2^n - 1$, a_i corresponds to the probability of cooperation if $H = H_i$. However, if $H = H_{2^n}$ (the n-move history is all defections) either a_{2^n} or a_{2^n+1} is used: the probability of cooperation is a_{2^n} if $b = 0$, and a_{2^n+1} if $b = 1$. For example, consider the 2-ply first defector strategy (1 0 0 [1 0]). This strategy will defect if the two move history is Cd or Dc, and cooperate if the two move history is Cc. If the history is Dd, it will cooperate if $b = 0$ (the player's unprovoked defection started the sequence of D's) and defect if $b = 1$ (the opponent's unprovoked defection started the sequence of D's). First defector strategies are denoted by FD_∞ to distinguish them from their finite memory approximations FD_n.

4.2 Finite memory approximations to FD strategies

Given a k-ply first defector strategy $FD_\infty = (a_1, \ldots, a_{2^k+1})$, an *n-ply approximation* FD_n to the first defector strategy is defined as follows. For $n = k$, $FD_n = (a_1, \ldots, a_{2^k-1}, g)$, where g is the strategy's "generosity" (probability of cooperation after n defections). Thus a two ply approximation to (1 0 0 [1 0]) would be the 2-ply strategy (1 0 0 g). For $n > k$, we define FD_n recursively by concatenating two copies of FD_{n-1}: $FD_n = FD^*_{n-1} \oplus FD_{n-1}$, where $\oplus$ is the concatenation operator and FD^*_{n-1} is defined as FD_{n-1} with:

$$g = \begin{cases} a_{2^k} & \text{if } n \text{ is odd} \\ a_{2^k+1} & \text{if } n \text{ is even} \end{cases}$$

Thus a 3-ply approximation to (1 0 0 [1 0]) would be (1 0 0 1)$\oplus$(1 0 0 g) = (1 0 0 1, 1 0 0 g). A 4-ply approximation would be (1 0 0 1, 1 0 0 0)$\oplus$(1 0 0 1, 1 0 0 g) = (1 0 0 1, 1 0 0 0, 1 0 0 1, 1 0 0 g). Thus we can approximate a k-ply first defector strategy FD_∞ with an n-ply strategy FD_n for any desired $n \geq k$.

4.3 The Firm But Fair strategy class

The Firm But Fair strategy class consists of the FBF_∞ first defector strategy, as well as its n-ply approximations FBF_n. FBF_∞ is the 2-ply first defector strategy (1 0 1 [1 0]), which behaves equivalently to "contrite TFT" (Wu & Axelrod, 1995). The two ply approximation $FBF_2(g)$ is the standard Firm But Fair strategy (1 0 1 g). The three ply and higher approximations are defined as follows, where $\oplus$ is the concatenation operator:
$FBF_3(g) = FBF_2(g=1) \oplus FBF_2(g) = (1\ 0\ 1\ 1,\ 1\ 0\ 1\ g)$
$FBF_4(g) = FBF_3(g=0) \oplus FBF_3(g) = (1\ 0\ 1\ 1,\ 1\ 0\ 1\ 0,\ 1\ 0\ 1\ 1,\ 1\ 0\ 1\ g)$
$$FBF_n(g) = \begin{cases} FBF_{n-1}(g=1) \oplus FBF_{n-1} & \text{if } n \text{ is odd} \\ FBF_{n-1}(g=0) \oplus FBF_{n-1} & \text{if } n \text{ is even} \end{cases}$$
These strategies cooperate with a cooperating opponent, and defect in response to unprovoked defection by the opponent. In the event of a sequence of defections, they will cooperate if they were the first unprovoked defector (cDd, cDdDd, etc.), defect if the opponent was the first unprovoked defector (Cd, CdDd, etc.), or cooperate with probability g if their memory capacity is exceeded (the first defection occurred at least n turns ago, so they do not know who defected first).

4.4 The Firm Pavlov strategy class

The Firm Pavlov class of strategies is a variation of Firm But Fair which also can exploit an unconditionally cooperating opponent. The Firm Pavlov strategy class consists of the FP_∞ first defector strategy, as well as its n-ply approximations FP_n. FP_∞ is the 3-ply first defector strategy (1 0 0 1, 1 0 1 [1 0]). The three ply approximation $FP_3(g)$ is the strategy (1 0 0 1, 1 0 1 g). The four ply and higher approximations are defined as follows, where $\oplus$ is the concatenation operator:
$FP_4(g) = FP_3(g=0) \oplus FP_3(g) = (1\ 0\ 0\ 1,\ 1\ 0\ 1\ 0,\ 1\ 0\ 0\ 1,\ 1\ 0\ 1\ g)$
$$FP_n(g) = \begin{cases} FP_{n-1}(g=1) \oplus FP_{n-1} & \text{if } n \text{ is odd} \\ FP_{n-1}(g=0) \oplus FP_{n-1} & \text{if } n \text{ is even} \end{cases}$$
These strategies are similar to FBF_n, but defect after they successfully exploit the opponent (3-move history cDc). They do not defect after a 3-move history of dDc, because this would prevent the strategy from breaking out of sequences of mutual defection against itself.

We now examine the performance of various FBF and FP strategies, searching for strategies which meet the criteria for infinitesimal noise optimality, finite noise optimality, and arbitrary noise optimality.

5 Optimality of FBF and FP

5.1 Optimality under no noise and infinitesimal noise

For the no-noise case, $FBF_n(g)$ has $\sigma_S = 1$ and $\sigma_C = 0$, since it will cooperate continually against its clone or ALLC. Its value of σ_D depends on both n and

g. Against ALLD, $FBF_n(g)$ will cooperate with probability g if it has made at least $\lfloor n/2 \rfloor$ straight defections, and defect otherwise. Thus it cooperates once every $\lfloor n/2 \rfloor + 1/g$ turns. This gives us:

$$1 - \sigma_D = \frac{1}{\lfloor \frac{n}{2} \rfloor + \frac{1}{g}} \longrightarrow \sigma_D = \frac{(\lfloor \frac{n}{2} \rfloor - 1)g + 1}{(\lfloor \frac{n}{2} \rfloor)g + 1}$$

Thus $FBF_n(g)$ has $\sigma_D = 1$ for $g = 0$, and the infinite memory FBF_∞ strategy also has $\sigma_D = 1$. $FP_n(g)$ with initial cooperation has identical σ values to $FBF_n(g)$. However, $FP_n(g)$ with initial defection will exploit ALLC continually, giving it $\sigma_C = 1$. Thus for all n, $FP_n(0)$ with initial defection has $\sigma_C = \sigma_D = \sigma_S = 1$, and thus it is strongly optimal for no noise. Similarly, the infinite memory FP_∞ with initial defection has $\sigma_C = \sigma_D = \sigma_S = 1$, and is strongly optimal for the no noise APD.

For infinitesimal noise, we first note that an FBF or FP strategy's value of σ_D will be unchanged from the no noise case. The FBF strategies have $\sigma_C = 0$ as in the no noise case, but the FP strategies have $\sigma_C = 1/3$. This results since, in cooperation, only an error by FP will cause it to start exploiting ALLC, but while exploiting, an error by either FP or ALLC will cause it to stop exploiting. Thus FP is twice as likely to leave the exploiting rut as to enter, and it will spend one-third of the time exploiting ALLC.

We now consider self-cooperation under infinitesimal noise, computing the proportion of time that an FBF or FP strategy will be in mutual cooperation with its clone. We first examine the $FP_3(0)$ strategy (1 0 0 1, 1 0 1 0), demonstrating that $\sigma_S = 1$. But how can $FP_3(0)$ be totally self-cooperating if its generosity (probability of cooperation after dDd) is 0? It is possible for $FP_3(0)$ to fall into a rut of mutual defection with its clone, but it requires two errors to enter the rut and only a single error to restore cooperation. It thus enters the rut once every $1/\epsilon^2$ turns, and leaves the rut every $1/\epsilon$ turns. Hence, in an infinitely long game, we can expect $FP_3(0)$ to be in the defection rut with probability:

$$\frac{1/\epsilon}{1/\epsilon^2 + 1/\epsilon} = \frac{\epsilon}{1+\epsilon} \approx 0$$

Thus $FP_3(0)$ has $\sigma_S = 1$, $\sigma_D = 1$, and $\sigma_C = 1/3$. We have found a strategy that meets the criteria for strong optimality in the infinitesimal noise case!

We now compute the proportion of time that an FBF or FP strategy will be in a rut of mutual defection with its clone in the infinitesimal noise APD. In general, it takes $\lceil n/2 \rceil$ errors to enter the mutual defection rut, and the strategies can escape the rut with probability g ($g \geq \epsilon$). Thus the proportion of time spent in this state is:

$$\frac{\epsilon^{\lceil n/2 \rceil}}{g + \epsilon^{\lceil n/2 \rceil}}$$

For $\epsilon \approx 0$, this quantity is negligible unless $n \leq 2$ and g is $O(\epsilon)$. Thus for FBF and FP strategies of 3-ply or higher memory, $\sigma_S = 1$. Similarly, $\sigma_S = 1$ for

FBF_∞ and FP_∞. For the 2-ply FBF strategy, $\sigma_S = 1/2$ for $g = \epsilon$, and $\sigma_S = 1$ for $g \gg \epsilon$.

A strategy which is strongly optimal for infinitesimal noise must have $\sigma_S = \sigma_D = 1$ and $\sigma_C > 0$. Thus $FP_n(0)$ is strongly optimal for all n, and FP_∞ is also strongly optimal for the infinitesimal noise APD.

5.2 Optimality under finite noise

For the finite noise case, we first calculate the noise resistance coefficient NRC_S for FP and FBF strategies with various values of n and g. Recall that for a strategy X with noise resistance coefficient NRC_S, $w(X|X) \geq R - (NRC_S)\epsilon$ for all $\epsilon \leq 0.2$. As can be seen from Table 5, FBF strategies with $n \geq 5$ or $g \gg \epsilon$ are finite noise resistant. Similarly, FP strategies with $n \geq 5$ or $g \gg \epsilon$ are finite noise resistant for $T = 4$. (For $T = 5$, FP strategies with $n \geq 7$ or $g \gg \epsilon$ are finite noise resistant.) For $T = 4$, the $FP_5(0)$ strategy meets all the criteria for finite noise optimality: it is finite noise resistant, and under infinitesimal noise, it is C-exploiting, totally self-cooperating, and totally D-unexploitable. We have thus accomplished the first part of our goal, to find a finite memory strategy which is finite noise optimal. $FP_n(0)$ is also finite noise optimal for $n > 5$, as is FP_∞.

We next examine the various FP and FBF strategies in more depth. For each, we compute the maximum noise level $\epsilon = e$ for which each strategy is self-cooperating ($\sigma_S \geq 0.9$) and D-unexploitable ($\sigma_D \geq 0.9$). As can be seen from Table 5, increasing generosity allows a strategy to achieve self-cooperation at higher noise levels, but dramatically reduces D-unexploitability. Increasing the memory n also increases self-cooperation, but only slightly reduces D-unexploitability. Finally, note that the $FP_5(0)$ strategy is both self-cooperating and D-unexploitable for noise levels up to approximately .09.

Next we compute the minimum value of g necessary for each finite memory strategy to be finite noise resistant (see Table 6). Again, we note that with the T=4 payoff table, $FBF_5(g)$ and $FP_5(g)$ are finite noise resistant for all values of g.

Next, we examine the strategies' evolutionary interaction with ALLD, calculating the minimum proportion of each necessary to invade a population of defectors for various values of ϵ (see Table 7). It can be seen that only a small proportion of FP or FBF is needed to invade ALLD for low noise levels, but for higher noise levels a large proportion is needed to invade. It should also be noted that higher memory strategies invade better, and that FBF_n invades slightly better than FP_n.

Finally, we calculate the maximum generosity that will allow each strategy to resist an invasion by ALLD. As can be seen from Table 8, all strategies can resist an ALLD invasion for $\epsilon \leq .1$ and $T = 4$. For $T = 5$, or $\epsilon \geq .1$, some very generous versions of these strategies will allow ALLD to invade. Finally, we

Strategy	NRC_S $T=4$	NRC_S $T=5$	$e(\sigma_S=0.9)$ $T=4$	$e(\sigma_S=0.9)$ $T=5$	$e(\sigma_D=0.9)$ $T=4$	$e(\sigma_D=0.9)$ $T=5$
$FBF_2(0)$	∞	∞	0	0	.100	.0697
$FBF_2(.5)$	6	5	.0605	.0585	0	0
$FBF_2(1)$	4	3	.125	.117	0	0
$FBF_3(0)$	6	5	.0605	.0585	.0922	.0668
$FBF_3(.5)$	4	3	.112	.106	0	0
$FBF_3(1)$	4	3	.125	.117	0	0
$FP_3(0)$	6	5	.0556	.0547	.0925	.0669
$FP_3(.5)$	4	3	.0949	.0942	0	0
$FP_3(1)$	4	3	.102	.101	0	0
$FBF_5(0)$	4	3	.104	.0988	.0917	.0666
$FBF_5(.5)$	4	3	.111	.105	0	0
$FBF_5(1)$	4	3	.112	.106	0	0
$FP_5(0)$	4	3.011	.0898	.0890	.0919	.0667
$FP_5(.5)$	4	3	.0943	.0937	0	0
$FP_5(1)$	4	3	.0949	.0942	0	0
FBF_∞	4	3	.111	.105	.0917	.0666
FP_∞	4	3	.0943	.0936	.0918	.0667

Table 5: Performance of strategies under finite noise

Strategy	g $(T=4)$	g $(T=5)$
FBF_2	1	1
FBF_3	.20	.25
FP_3	.33	.40
FBF_5	0	0
FP_5	0	.047

Table 6: Minimum generosity for finite noise resistance

note that FBF_∞ and FP_∞ cannot be invaded by ALLD. Assuming noise levels up to 20%, FBF_n cannot be invaded by ALLD for $n \geq 5$, and FP_n cannot be invaded by ALLD for $n \geq 7$ ($n \geq 5$ for T=4).

5.3 Noise resistance coefficients: a closer look

We now consider the concept of noise resistance coefficient NRC_S more closely, focusing on the 3-ply and 5-ply FP and FBF strategies for the (5,3,1,0) payoff table. Why does the noise resistance of the 3-ply strategies vary between 3 and 5, and why does $FP_5(0)$ have a noise resistance coefficient slightly greater than 3? To answer these questions, we first use Maple to calculate $w(X|X)$ for each

Strategy	$T=4$				$T=5$			
	$\epsilon = .01$	$\epsilon = .05$	$\epsilon = .1$	$\epsilon = .2$	$\epsilon = .01$	$\epsilon = .05$	$\epsilon = .1$	$\epsilon = .2$
FBF_2	.010	.056	.125	.333	.015	.087	.204	.593
FBF_3	.005	.030	.073	.218	.008	.046	.112	.345
FP_3	.005	.031	.075	.233	.008	.046	.113	.362
FBF_5	.005	.029	.069	.203	.008	.044	.105	.316
FP_5	.005	.029	.070	.215	.008	.044	.106	.328
FBF_∞	.005	.029	.069	.200	.008	.044	.104	.310
FP_∞	.005	.029	.070	.212	.008	.044	.105	.322

Table 7: Proportion needed to invade ALLD

Strategy	$T=4$				$T=5$			
	$\epsilon = .01$	$\epsilon = .05$	$\epsilon = .1$	$\epsilon = .2$	$\epsilon = .01$	$\epsilon = .05$	$\epsilon = .1$	$\epsilon = .2$
FBF_2	1	1	1	.814	.968	.846	.708	.456
FBF_3	1	1	1	.811	.968	.855	.734	.530
FP_3	1	1	1	.744	.968	.850	.720	.502
FBF_5	1	1	1	1	1	1	1	1
FP_5	1	1	1	1	1	1	1	.938

Table 8: Maximum generosity to resist ALLD invasion

strategy as a function of its generosity g, and a finite noise level $\epsilon = e$:

$$w(FP_3|FP_3) = \frac{5ge^2 + 3g + e^2 - e^3}{ge^2 + g + ge + e^2 - e^3}$$

$$w(FBF_3|FBF_3) = \frac{-ge^2 + 3g + 3ge + e^2 - e^3}{-ge^2 + g + 2ge + e^2 - e^3}$$

$$w(FP_5|FP_5) = \frac{7ge^2 + 3g - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}{3ge^2 + g + ge - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}$$

$$w(FBF_5|FBF_5) = \frac{ge^2 + 3g + 3ge - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}{ge^2 + g + 2ge - 3ge^3 + ge^4 + e^3 - 2e^4 + e^5}$$

We first consider the case $g \gg e$, and expand each of these functions as a Taylor polynomial in e and g:

$$w(FP_3|FP_3) = 3 - 3e + \left(5 - \frac{2}{g}\right) e^2 + \left(\frac{7}{g} - 2\right) e^3 + O(e^4)$$

$$w(FBF_3|FBF_3) = 3 - 3e + \left(8 - \frac{2}{g}\right) e^2 + \left(\frac{9}{g} - 19\right) e^3 + O(e^4)$$

$$w(FP_5|FP_5) = 3 - 3e + e^2 + \left(14 - \frac{2}{g}\right) e^3 + O(e^4)$$

$$w(FBF_5|FBF_5) = 3 - 3e + 4e^2 + \left(1 - \frac{2}{g}\right) e^3 + O(e^4)$$

In each case, the linearization of $w(X|X)$ around $e = 0$ is $3 - 3e$, and thus:

$$NRC_S = \sup_{(e=0\ldots\epsilon_{max})} \frac{R - w(X|X)_{\epsilon=e}}{e} \approx \lim_{\epsilon\to 0} \frac{R - w(X|X)}{\epsilon} = \frac{3 - (3 - 3\epsilon)}{\epsilon} = 3$$

However, the supremum value does not necessarily occur at $\epsilon \to 0$. To see this, we plot $(R - w)/e$ as a function of e for the FP_3 and FP_5 strategies, with various values of g (see Figure 1, Figure 2). As can be seen from the figures, the supremum value of $(R - w)/e$ may occur at $g = 0$ as for $FP_5(.2)$, at $g = \epsilon$ as for $FP_3(.1)$, or at an intermediate value as for $FP_3(.2)$. Therefore, we must consider higher order terms to find the exact value of NRC_S.

For FBF_5 and FP_5, $w(X|X) > 3 - 3e$, so the strategies have a noise resistance coefficient of 3 for $g \gg \epsilon$. For FP_3, $w(X|X) \approx 3 - 3e + (5 - 2/g)e^2$, so FP_3 has a noise resistance coefficient of 3 for $5 - 2/g > 0$, or $g > 0.4$. For $g < 0.4$, the quadratic term will be negative, so FP_3 will have NRC_S greater than 3. Similarly, for FBF_3, $w(X|X) \approx 3 - 3e + (8 - 2/g)e^2$, so FBF_3 has a noise resistance coefficient of 3 for $8 - 2/g > 0$, or $g > 0.25$.

We compute the following noise resistance coefficients for FP_3 and FBF_3 as a function of g, assuming that the noise ϵ varies from 0 to g (see Table 9).

g	$NRC_S(FP_3)$	$NRC_S(FBF_3)$
.01	4.863	4.816
.02	4.731	4.645
.05	4.364	4.192
.10	3.833	3.594
.20	3.206	3.043
.40	3	3

Table 9: Noise resistance coefficients of $FP_3(g)$ and $FBF_3(g)$

Next we consider the case $g = O(\epsilon)$. In this case, we let $k = g/e$ and find the power series for $w(X|X)$ as a function of k and e:

$$w(FP_3|FP_3) = 3 - \left(3 + \frac{2}{k}\right) e + O(e^2)$$

$$w(FBF_3|FBF_3) = 3 - \left(3 + \frac{2}{k}\right) e + O(e^2)$$

$$w(FP_5|FP_5) = 3 - 3e + \left(1 - \frac{2}{k}\right) e^2 + O(e^3)$$

$$w(FBF_5|FBF_5) = 3 - 3e + \left(4 - \frac{2}{k}\right) e^2 + O(e^3)$$

Thus the noise resistance coefficients of FBF_3 and FP_3 are approximately $3 + 2/k$. For $k = 1$ ($g = e$) we get $NRC_S = 3 + 2 = 5$, and for large k we get $NRC_S = 3$. This explains why the noise resistance coefficients of these strategies vary between 3 and 5. For FP_5, we obtain:

$$\frac{R - w(X|X)}{\epsilon} \approx \frac{3 - (3 - 3e + (1 - 2/k)e^2)}{e} = 3 + \left(\frac{2}{k} - 1\right) e$$

This implies that the noise resistance coefficient of FP_5 is 3 for $(2/k - 1) < 0$, that is, for $k > 2$. For $k < 2$, we know that NRC_S will be slightly greater than 3. For the worst case, $k = 1$, we find that the supremum occurs at a value $e \approx .023$, giving $NRC_S \approx 3.011$ as in Table 5.

For FBF_5, on the other hand:

$$\frac{R - w(X|X)}{\epsilon} \approx \frac{3 - (3 - 3e + (4 - 2/k)e^2)}{e} = 3 + \left(\frac{2}{k} - 4\right) e$$

Even for the worst case, $k = 1$, FBF_5 will have a noise resistance coefficient of 3.

Thus we have discovered several important results from our analysis of the 3-ply and 5-ply FP and FBF strategies. First, the supremum NRC_S does not necessarily occur at the endpoints of the interval, so we must take the supremum of $(R - w)/\epsilon$ over the whole interval in order to obtain a bound. We also found that the self-payoff of each strategy is $w \geq 3 - 3\epsilon$ for sufficiently high generosity. For FP_3 we must have $g > 0.4$, for FBF_3 we must have $g > 0.25$, and for FP_5 we need only $g > 2\epsilon$. FBF_5 has $w \geq 3 - 3\epsilon$ for all $g \geq \epsilon$. The self-payoff of FP_n is reduced by approximately $3e^2$ from the self-payoff of FBF_n, but its ability to exploit unconditional cooperators outweighs this disadvantage.

5.4 Optimality under arbitrary noise

For the arbitrary noise case, we first calculate the noise proof coefficient NPC_S for FP and FBF strategies with various values of n and g. Recall that the noise proof coefficient is the maximum (over all error sequences) of the expected number of defections per error. To calculate the noise proof coefficient, we consider two possible "worst-case" error sequences: the correction of a single error by two defections (CcCc<u>D</u>dCc) and the mutual defection rut resulting from $\lceil n/2 \rceil$ defections. This rut will result in an average of $n + 1/g - 1$ defections, and thus we find:

$$NPC_S = \max\left(\frac{n + \frac{1}{g} - 1}{\lceil \frac{n}{2} \rceil}, 2\right)$$

This implies that the FBF or FP strategy is noise proof for $1/2 < g < 1$ (for n odd) or $g = 1$ (for n even). Also, we find that $NPC_S = 2$ for FBF_∞ and FP_∞. Thus the FP_∞ strategy meets all the criteria for arbitrary noise optimality: it is noise proof, and meets the four criteria for finite noise optimality discussed above. We have thus accomplished the second part of our goal, finding an infinite memory strategy which is arbitrary noise optimal.

6 Performance of FBF and FP

We have shown above that higher-memory strategies of the FP and FBF strategy classes meet many or all of our optimality criteria. In particular, we have found that the $FP_n(0)$ strategies are finite noise optimal for $n > 5$ (for $T = 4$) or $n > 7$ (for $T = 5$). Also, the FP_∞ strategy is arbitrary noise optimal. The corresponding FBF strategies meet almost all of the optimality criteria, but are not C-exploiting.

We must now ask an essential question: do strategies which meet these optimality criteria perform well in (evolutionary and non-evolutionary) APD simulations? To answer this question, we first ran several round-robin tournaments. (In these tournaments, Markov chains were used to calculate each strategy's average score per round against every other strategy in the tournament, then the strategies' total scores were compared.) The first four tournaments consisted of 63 strategies: g1, g2, g3, Pavlov, RAND, ALLD, ALLC, 11 FBF_2 strategies (generosity 0-1), 10 FBF_3 strategies (generosity 0-.9), 11 FP_3 strategies (generosity 0-1), 11 FBF_5 strategies (generosity 0-1), 11 FP_5 strategies (generosity 0-1), FBF_∞, and FP_∞. Tournaments were run at noise levels of .01 and .05, and with payoff tables (4,3,1,0) and (5,3,1,0). See appendix for table of results. The FP_5 strategies won all four tournaments: for the two tournaments with noise .01, $FP_5(.2)$ was the champion, and for the two tournaments with noise .05, $FP_5(1)$ was the champion. In general, FP strategies outperformed FBF strategies: $FP_n(g)$ outscored $FBF_n(g)$ for any given n and g. Also, for the finite memory strategies, performance improved with increasing n. However, the infinite memory strategies FBF_∞ and FP_∞ performed poorly in these tournaments. We hypothesized that this resulted from poor performance against strategies with low memory and low generosity. For example, consider FBF_∞ against the Tit for Tat strategy, $FBF_2(0)$. An error by FBF_∞ would be quickly corrected, but an error by Tit for Tat would result in a rut of mutual defection. To test this hypothesis, a fifth tournament was run with 58 of the 63 strategies, excluding the five non-generous FBF/FP strategies ($FBF_2(0)$, $FBF_3(0)$, $FP_3(0)$, $FBF_5(0)$, and $FP_5(0)$). A .01 noise level, and the (5,3,1,0) payoff table, were used. As expected, FP_∞ was the champion. We also noted that the optimum value of g varied with the memory n (see Table 10).

Next we selected 29 of the FBF and FP strategies: 4 FBF_2 strategies (generosity .2-.8), 5 FBF_3 strategies (generosity 0-.8), 6 FP_3 strategies (generosity 0-1), 6 FBF_5 strategies (generosity 0-1), 6 FP_5 strategies (generosity 0-1),

Strategy	$g(\epsilon = .01)$	g($\epsilon = .05$)
FBF_2	1	1
FBF_3	.2	.7
FP_3	.2	.7
FBF_5	.2	1
FP_5	.2	1

Table 10: Optimum generosity, round-robin

FBF_∞, and FP_∞. Each of the strategies was run individually in a 10-strategy round robin tournament against g1, g2, g3, Pavlov, RAND, ALLD, ALLC, TFT ($FBF_2(0)$), and FBF ($FBF_2(1)$). A .01 noise level and (4,3,1,0) payoff table were used. Of the 29 strategies tested, 28 won their tournament: the only exception was $FBF_3(0)$, which finished second to FBF. The $FP_5(.2)$ and $FP_3(.2)$ strategies won their tournaments by the largest margins, .0713 and .0682 respectively, and FP strategies again outperformed FBF strategies.

From these results, it is clear that both the FBF and FP strategies perform well in round-robin tournaments, with FP performing significantly better than FBF. The ability to exploit unconditional cooperators (without significantly reducing the strategy's ability to cooperate with conditional cooperators) is essential for success in round-robin tournament simulation. The $FP_5(.2)$ strategy was particularly successful in the round-robin tournaments we conducted. It is perfectly self-cooperating, mostly D-unexploitable, and somewhat C-exploiting; the non-zero generosity value also allows it to cooperate with low-memory/low-generosity conditional cooperators such as TFT.

7 Evolutionary optimality criteria

7.1 A brief overview of standard evolutionary models

Theoretical biologists have proposed a number of evolutionary models for population dynamics in the iterated Prisoner's Dilemma. In all of these models, evolution is driven by natural selection: strategies which earn higher average payoffs have higher "evolutionary fitness", and are able to survive and reproduce. The "Genetic Algorithm" model used by Axelrod (1987) and others represents each strategy as a "chromosome", a collection of genes describing the strategy's behavior in any given situation. This model creates genetic variation in a strategy's offspring by both mutation (randomly occurring changes in an individual's genes) and crossover (genetic recombination of features from two different "parent" strategies). This approach enables continued evolution of new strategies, but relies heavily on chance: the evolution of a population is strongly dependent on which random mutations or recombinations occur, as well as the specific parameters of the genetic model.

A second model is the invasion model proposed by Nowak and Sigmund

(1992,1993). This model assumes that a large homogeneous population of a given stategy X is visited periodically by mutant strains. In this model, a mutant strain Y is selected randomly from the hyperspace of all possible n-ply strategies, then John Maynard Smith's invasion criteria (1982) are used to calculate whether Y can invade X. This invasion will occur if $w(Y|X) > w(X|X)$, or $w(Y|X) = w(X|X)$ and $w(Y|Y) > w(X|Y)$. The model assumes that if Y invades X, Y will take over the population; otherwise, the initial population will continue. Invasion models of this sort tend to result in cyclical or complex invasion behavior: for example, ALLD is invaded by Tit for Tat, which is invaded by Generous Tit for Tat, which is invaded by ALLC, which is invaded by ALLD. Though it is a fairly realistic model of population dynamics, this model is very dependent on which mutants are chosen to invade, and its complex time-dependent behavior makes it very difficult to decide which strategy is in some sense "optimal". One solution to this problem is given by the notion of an "evolutionarily stable strategy" (ESS). An ESS is a strategy X which cannot be invaded by any other strategy: that is, for all strategies Y, $w(Y|X) < w(X|X)$, or $w(Y|X) = w(X|X)$ and $w(Y|Y) \leq w(X|Y)$. An ESS is in one sense an "evolutionarily optimal" strategy: once it is established in a population, no other single strategy can invade. However, as shown by Boyd and Lorberbaum (1987), an ESS may be invaded by a combination of two mutant strategies. More importantly, an ESS may have difficulty establishing itself in a population, because it has difficulty invading other strategies. In fact, some evolutionarily stable strategies are "inaccessible" in the sense that they are unable to invade any other homogeneous population (Nowak, 1990).

Thus it is clear that a measure of evolutionary "optimality" should take into account not only a strategy's resistance to invasion (evolutionary stability) but its ability to invade other strategies (evolutionary potency). One model which considers both of these factors is the "Adaptive Dynamics" model used in Nowak & Sigmund (1989) and Hofbauer & Sigmund (1990). This model assumes that an initial homogeneous population of strategy S has numerous offspring X, each slightly different from S. Each of these offspring interacts with all other offspring, and the variety with the highest total payoff survives. Assuming a symmetric distribution of offspring around S, the offspring with the highest total payoff will be the offspring with the highest payoff against the parent. We thus define the fitness function $F(X)$ to be $w(X|S)$. This leads to the adaptive dynamics equation:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = h\nabla F$$

where the gradient ∇F is evaluated at X=S, and h is a constant corresponding to the rate of variation (Kraines & Kraines, 2000). Thus the population evolves in the direction which is most advantageous for the single mutant (Nowak & Sigmund, 1989). One use of this model is to examine an individual strategy by allowing it to evolve according to the adaptive dynamics equation: this process of "self-evolution" gives information about the stability of the strategy, as well as its possible evolutionary path. We will examine the self-evolution of FBF_2,

FBF_3, and FP_3 using this model. Again, the behavior of this dynamical system is time-dependent and dependent on the initial strategy chosen, making it difficult to argue that a given strategy is "optimal". But, as in the Nowak-Sigmund model, we can find strategies which are "optimal" in the sense of resistance to invasion. A "relative evolutionarily stable strategy" (RESS) is a strategy which cannot be invaded (according to the Maynard Smith criterion) by any nearby strategy. In other words, strategy X is a RESS in some space H of strategies if, for some $\delta > 0$, for all strategies Y with $|X - Y| < \delta$, either $w(Y|X) < w(X|X)$, or $w(Y|X) = w(X|X)$ and $w(Y|Y) \leq w(X|Y)$. Once we have found a RESS X, we can also examine its "basin of attraction": the set of strategies in H which evolve into X. A RESS with a large basin of attraction is in some sense "optimal", since it is resistant to (local) invasion and can eventually take over many self-evolving populations. As Kraines and Kraines state, the adaptive dynamics model is "consistent with modern evolutionary theory for a population with limited genetic diversity and frequent minor mutations" (2000). However, it excludes the possibility of invasion by strategies which are not similar to the original strategy, and thus does not allow for major mutations, migrants, or interactions of multiple populations. We will propose another evolutionary model which addresses these shortcomings, but first we examine the self-evolution of various FBF and FP strategies.

7.2 Self-evolution of FBF and FP

To examine the self-evolution of the FBF and FP strategies, we use a discrete approximation to the adaptive dynamics equation as in Kraines & Kraines (2000). This algorithm takes three parameters: the starting strategy $S(0) = [S_1(0), \ldots, S_{2^n}(0)]$, the step size δ, and the rate of variation h. For each time step t, it modifies S using the following equations:

$$\Delta S = S(t+1) - S(t) = h\Delta F$$

$$\Delta F[i] = F(S + \delta\hat{x_i}) - F(S)$$

where $\hat{x_i}$ is the vector with a 1 in position i and zeros elsewhere.

We first examine the $FBF_2(0)$ strategy, (1 0 1 0). We self-evolve this strategy with $T = 4$, $\epsilon = .01$, $\delta = .01$, and $h = .1$. The strategy's probabilities of cooperation after Cd and Dd quickly increase, leveling off after $2 \cdot 10^5$ generations to (1 .2607 1 .6554).

We next examine the $FBF_3(0)$ strategy, (1 0 1 1, 1 0 1 0). We self-evolve this strategy with $T = 4$, $\epsilon = .01$, $\delta = .01$, and $h = .1$. The strategy's probability of cooperation after dDd increases rapidly at first (to .3383 after $2 \cdot 10^5$ generations) then levels off after 10^7 generations. The resulting strategy is (1 0 1 1, 1 0 1 .6566).

We note several interesting facts from the self-evolution of these two strategies. Both strategies increase in generosity, evolving toward increasing self-cooperation and decreasing D-unexploitability. Both generosity parameters seem to level off at approximately 2/3, but the $FBF_2(0)$ strategy also evolves some unconditional cooperation (cooperating over 1/4 of the time after Cd). Thus $FBF_2(0)$ evolves away from the FBF family of strategies, but $FBF_3(0)$ remains an FBF strategy under evolution. The FBF_3 strategy also evolves much more slowly than FBF_2, suggesting a higher degree of evolutionary stability.

Finally, we examine the $FP_3(0)$ strategy, (1 0 0 1, 1 0 1 0). We self-evolve this strategy with $T = 4$, $\epsilon = .01$, $\delta = .01$, and $h = .1$. The strategy's probability of cooperation after dDd increases rapidly at first (to .3392 after $2 \cdot 10^5$ generations) then levels off after 10^7 generations. However, its probability of cooperation after cDc also increases slowly and at a fairly constant rate, to $1-\epsilon$ after $5.1 \cdot 10^6$ generations. The resulting strategy is (1 0 1 1, 1 0 1 .6566).

Thus both the $FP_3(0)$ and $FBF_3(0)$ strategies evolve to the relative evolutionarily stable strategy $FBF_3(.6566)$. We can conclude from this that FBF_3 is much more stable than FP_3 under self-evolution. This result is likely to apply in the general case: there is no evolutionary advantage to exploitation when all strategies are sufficiently similar.

7.3 Evolutionary fitness and the Dominance Criterion

We now consider the question of a general measure of evolutionary performance. Are any of the previously mentioned models suitable for this measure? To answer this question, we consider what properties the measure should have, and apply these to the five models/techniques discussed above: genetic algorithms, Nowak-Sigmund invasion model, ESS theory, self-evolution, and RESS theory. We argue that three properties are essential. First, the measure must take into account both a strategy's evolutionary stability (resistance to invasion) and evolutionary potency (ability to invade other strategies). Second, the measure must be able to evaluate a strategy's performance against a large number of different strategies, not only a carefully chosen few. In particular, it must consider interactions with strategies which are substantially different from the given strategy, not only its evolutionary kin. Third, the measure must not be time-varying: it should not depend on the current composition of the population, but only the set of strategies which is initially present.

ESS theory violates the first property, since it only takes evolutionary stability into account. Self-evolution violates the second property, since it evaluates only the given strategy, not its relation to other strategies. RESS theory also violates the second property, since it considers the interaction of a strategy only with strategies very similar to it. Using the genetic algorithm or Nowak-Sigmund invasion models, the proportion of a strategy tends to vary significantly (and possibly chaotically) over time, so it is difficult to postulate a non-time-varying measure which accurately reflects a strategy's performance. Thus none of the commonly used models of the evolutionary APD lend themselves easily to a general measure of evolutionary performance.

In order to derive our measure, we return to the Maynard-Smith invasion criteria (1982). From Maynard Smith, strategy X invades strategy Y if $w(X|Y) > w(Y|Y)$, or $w(X|Y) = w(Y|Y)$ and $w(X|X) > w(Y|X)$. If strategy X invades strategy Y, we write $X > Y$. Otherwise, we write $X \not> Y$. We must first observe that whether X invades Y and whether Y invades X are independent: thus there are four possibilities:

I. $X > Y$ and $Y \not> X$. In this case, any initial proportion of strategy X can take over, and completely wipe out, strategy Y. We say that X *dominates* Y, and write $X \gg Y$.

II. $X \not> Y$ and $Y > X$, In this case, any intial proportion of strategy Y can take over, and completely wipe out, strategy X. We say that Y *dominates* X, and write $Y \gg X$.

III. $X > Y$ and $Y > X$. In this case, no matter what the initial proportions of strategies X and Y, the two strategies reach a balance in which the population is a mix of X and Y. We say that X and Y are *stable*, and write $X \leftrightarrow Y$. To be more precise, we can calculate the proportion p of strategy X at equilibrium:

$$p = \frac{w(X|Y) - w(Y|Y)}{w(X|Y) - w(Y|Y) + w(Y|X) - w(X|X)}$$

Thus we write $X \overset{p}{\leftrightarrow} Y$.

IV. $X \not> Y$ and $Y \not> X$. In this case, either strategy X or strategy Y will take over the population, driving the other strategy to extinction, depending on the initial proportions of the two strategies. We say that X and Y are *bistable*, and write $X \Leftrightarrow Y$. To be more precise, we can calculate the minimum proportion m of strategy X needed to take over the population:

$$m = \frac{w(Y|Y) - w(X|Y)}{w(X|X) - w(Y|X) + w(Y|Y) - w(X|Y)}$$

Thus we write $X \overset{m}{\Leftrightarrow} Y$. We must also deal with the case $w(X|Y) = w(Y|Y)$ and $w(Y|X) = w(X|X)$, in which case m is defined to be 0.5.

We now define the *dominance measure* $\mathrm{dom}(X|Y)$, which is a measure of the relative evolutionary performance of strategies X and Y.

$$\mathrm{dom}(X|Y) = \begin{cases} 1 & \text{if } X \gg Y \\ 0 & \text{if } Y \gg X \\ p & \text{if } X \overset{p}{\leftrightarrow} Y \\ 1-m & \text{if } X \overset{m}{\Leftrightarrow} Y \end{cases}$$

Thus the measure $\mathrm{dom}(X|Y)$ is between 0 and 1, with higher dominance scores corresponding to better evolutionary performance of X against Y. It should also be noted that $\mathrm{dom}(X|Y) + \mathrm{dom}(Y|X) = 1$, and the dominance score presents a "constant-sum" measure of relative evolutionary performance (ie. if X performs well against Y, Y performs poorly against X).

The measure $\mathrm{dom}(X|Y)$ measures evolutionary fitness in a substantively different manner than other models. Most evolutionary models assume that a homogeneous population is invaded by a small number of mutants: these mutants can be nearly identical to the original population (as in the adaptive dynamics model) or very different (as in the Nowak-Sigmund invasion model). The dominance measure, on the other hand, assumes the mixing of two large homogeneous populations, and the evolution of the resulting population until it reaches a stable state.

Finally, we define the dominance score of a strategy X to be its average dominance score against all strategies it interacts with. Assuming a finite set of opponents $Y = \{Y_1..Y_M\}$:

$$\mathrm{dom}(X) = \frac{1}{M}\sum_{i=1}^{M} \mathrm{dom}(X|Y_i)$$

In a population of opponents Y_i with varying frequencies $f(Y_i)$, with $\sum_i f(Y_i) = 1$:

$$\mathrm{dom}(X) = \sum_i f(Y_i)\mathrm{dom}(X|Y_i)$$

In a continuous "population space" S of opponents:

$$\mathrm{dom}(X) = \int_S f(s)\mathrm{dom}(X|s)\mathrm{d}s$$

Thus a strategy's dominance score is a measure of its average relative evolutionary fitness. It has several advantages over other fitness measures: since it is time-invariant, we can compute a single score (independent of the current proportions of each strategy) for any two strategies. This allows us to measure the evolutionary performance of a strategy in a round-robin tournament format similar to Axelrod's (1984), except that dominance scores (rather than average payoffs) are computed for each pair of strategies. Since the dominance measure is constant-sum, the performance of a strategy is not affected as much by the set of strategies it interacts with: we expect successful strategies to have high dominance scores against most other strategies, while even "successful" strategies will have poor average payoffs against strategies such as ALLD. The disadvantage of the dominance measure is that it does not take into account the more complex interactions that result from the mixing of more than two strategies: in a multi-strategy interaction, a strategy's dominance score is essentially its performance after the first round of evolutionary "battles". But once the system evolves and weaker strategies die off, the strategy may be interacting with a very different set of strategies, dramatically affecting its evolutionary performance. However, there appears to be no way to take this evolutionary complexity into account without sacrificing our main objective: a simple and useful measure of overall evolutionary fitness. While a strategy with a very high dominance score may not succeed in every evolutionary interaction, it is likely to perform well in the great majority of such interactions. Thus we consider evolutionary dominance as another criterion for optimality: an "optimal" strategy must have a high dominance score against a large and varied set of opponents. In particular, we consider an n-ply strategy to satisfy this criterion if it has a high dominance score against the set of all n-ply strategies.

8 Evolutionary performance of FBF and FP

We now examine the evolutionary performance of the FBF and FP strategies through several "round-robin dominance tournaments". In a round-robin dominance tournament with M strategies, we compute:

$$\mathrm{dom}(X_i) = \frac{1}{M}\sum_{j=1}^{M}\mathrm{dom}(X_i|X_j)$$

We then compare each strategy's dominance score; the strategy with the highest dominance wins the tournament.

8.1 Round robin dominance tournaments

We first conducted round robin dominance tournaments among the same 63 strategies as for our initial round robin tournaments. These strategies were g1, g2, g3, Pavlov, RAND, ALLD, ALLC, 11 FBF_2 strategies (generosity 0-1), 10 FBF_3 strategies (generosity 0-0.9), 11 FP_3 strategies (generosity 0-1), 11 FBF_5 strategies (generosity 0-1), 11 FP_5 strategies (generosity 0-1), FBF_∞, and FP_∞. Tournaments were run with the (4,3,1,0) and (5,3,1,0) payoff tables, and a noise level of .01 was used. See appendix for table of results.

Both tournaments were won by the FBF_∞ strategy, with dominance scores of .9567 and .9612 for $T = 5$ and $T = 4$ respectively. This strategy dominated most of the other 62 strategies (43 for $T = 5$, 45 for $T = 4$), and had $\mathrm{dom}(X|Y) > 0.5$ against every other strategy. FBF strategies outperformed FP strategies: $FBF_n(g)$ outscored $FP_n(g)$ for all n and g. Additionally, performance improved with increasing memory: $FBF_m(g)$ outscored $FBF_n(g)$, and $FP_m(g)$ outscored $FP_n(g)$, for $m > n$. Also, strategies with relatively high generosities tended to perform better. The optimum value of g varied with the memory n, and was identical for $T = 4$ and $T = 5$.

Strategy	g
FBF_2	1
FBF_3	.6
FP_3	.6
FBF_5	.7
FP_5	.7

Table 11: Optimum generosity, round-robin dominance

These results are not unexpected: since FP evolves into FBF in the adaptive dynamics model, we would expect that FBF strategies have a higher dominance score than FP in head-to-head competition. That is, $\mathrm{dom}(FBF|FP) > \mathrm{dom}(FP|FBF)$, so $\mathrm{dom}(FBF) > \mathrm{dom}(FP)$ in tournaments consisting mainly of FBF and FP strategies.

8.2 Dominance versus corner and edge strategies

We now ask a more interesting question: how do the FBF and FP strategies perform against the set of all n-ply strategies? It is very difficult to compute the dominance score of a given strategy against all n-ply strategies. An n-ply strategy has 2^n parameters, each of which can vary continuously from 0 to 1, so to compute this score we must integrate $w(X|Y)$ over a continuous 2^n-dimensional subspace of R^{2^n}. Rather than performing the integration, we approximate the strategy's performance by considering only "corner strategies" (n-ply strategies for which each parameter is either 0 or 1), or "edge strateges" (n-ply strategies for which each parameter is either 0, 1/2, or 1). There are 2^{2^n} n-ply corner strategies, and 3^{2^n} n-ply edge strategies.

We considered the performance of each of the 63 strategies in the previous tournament: for each, we calculated an average dominance score against the 256 3-ply corner strategies, and against the 6561 3-ply edge strategies, for both $T = 4$ and $T = 5$. A .01 noise level was used. See appendix for results.

The $FP5(.1)$ strategy was the champion for three of the four tournaments (versus corner strategies with $T = 5$, versus edge strategies with $T = 5$ and $T = 4$). $FP\infty$ won the other tournament (versus corner strategies with $T = 4$). Both strategies performed consistently well, with dominance scores between .88 and .96 for all four tournaments. In general, FP strategies outperformed FBF strategies: all of the top 15 strategies for each tournament were FP strategies. For $T = 5$, the ALLD strategy also performed well (ranking 22nd and 18th versus corner and edge strategies respectively), but it was near the bottom (ranking 56th and 55th) for $T = 4$. Performance generally improved with increasing memory for the finite memory strategies, though in some tournaments the 5-ply strategies (with generosity 0.1 or 0.2) could beat the infinite-memory strategies. We also calculated the optimum value of g for each strategy (results are identical for $T = 5$ and $T = 4$). It should be noted that these generosity values are significantly lower than for the previous round robin dominance tournaments.

Strategy	g (corner)	g (edge)
FBF_2	.5	.4
FBF_3	.1	.1
FP_3	.1	.1
FBF_5	.1	.1
FP_5	.1	.1

Table 12: Optimum generosity vs. 3-ply strategies

It is clear from these results that the FP strategies (particularly those with low but non-zero generosities) perform extremely well on average in evolutionary interactions against 3-ply strategies. Though FBF strategies outperform FP strategies in head-to-head dominance tournaments, FP has a significantly higher dominance score when averaged over all strategies.

8.3 Dominance of 3-ply corner and edge strategies

Next we examine the space of 3-ply strategies in more detail, comparing the dominance scores of all corner and edge strategies rather than focusing on a specific group of strategies. First, round-robin dominance tournaments were run for the 256 3-ply corner strategies. The strategies were numbered from 1 to 256: ALLD (0 0 0 0, 0 0 0 0) was assigned number 1, (0 0 0 0, 0 0 0 1) was assigned number 2, etc. Then each strategy's average dominance score against strategies 1-256 was computed, and the 15 highest-scoring strategies were recorded. Tournaments were run for both $T = 4$ and $T = 5$, and a .01 noise level was used. See appendix for results.

The top corner strategy for both $T = 4$ and $T = 5$ was strategy #90, with a dominance score of .8910 for $T = 5$ and .9380 for $T = 4$. Strategy #90 corresponds to (1 0 0 1, 1 0 1 0): this is the $FP_3(0)$ strategy! Thus $FP_3(0)$ has the highest dominance score of all 3-ply corner strategies: since dominance score is a measure of overall evolutionary fitness, this suggests that $FP_3(0)$ is in some sense evolutionarily optimal among the 3-ply corner strategies. For $T = 4$, the second place strategy was strategy #94, with a dominance score of .8979. Strategy #94 corresponds to (1 0 1 1, 1 0 1 0), the $FBF_3(0)$ strategy. Thus for $T = 4$, $FP_3(0)$ and $FBF_3(0)$ were the top two corner strategies, with FP outperforming FBF by a significant margin. Additionally, $FP_3(1)$ placed tenth with a dominance score of .7992.

Other than FP_3 and (for $T = 4$) FBF_3, most of the top corner strategies were defectors (variants of strategies such as ALLD and GRIM). ALLD finished 11th for $T = 5$ and 14th for $T = 4$, with dominance scores of .8081 and .7753 respectively. Strategy #50, a slightly more cooperative version of GRIM (1 0 0 0, 1 1 0 0) performed particularly well, finishing 2nd for $T = 5$ and 3rd for $T = 4$, with dominance scores of .8570 and .8266 respectively. For $T = 4$, it is clear from this that both $FP_3(0)$ and $FBF_3(0)$ had much higher dominance scores than any other strategy.

Next, round-robin dominance tournaments were run for the 6561 3-ply edge strategies. The strategies were numbered from 1 to 6561: ALLD (0 0 0 0, 0 0 0 0) was assigned number 1, (0 0 0 0, 0 0 0 .5) was assigned number 2, (0 0 0 0, 0 0 0 1) was assigned number 3, etc. Then each strategy's average dominance score against strategies 1-6561 was computed, and the 15 highest-scoring strategies were recorded. Tournaments were run for both $T = 4$ and $T = 5$, and a .01 noise level was used. See appendix for results.

The top edge strategy for $T = 4$ was strategy #1677, with a dominance score of .9514. Strategy #1677 corresponds to (1 0 0 1, 1 0 1 0): thus the $FP_3(0)$ strategy has the highest dominance score among all 3-ply edge strategies, and this suggests that it is also likely to have the highest dominance score over the entire space of 3-ply strategies. The top 15 3-ply edge strategies were all variants of $FP_3(0)$, with no more than three parameters different. One of these strategies was $FBF_3(0)$, which finished 12th with a dominance score of .9211. Thus for $T = 4$, $FP_3(0)$ is the top edge strategy, and $FBF_3(0)$ also performs well.

For $T = 5$, however, the top edge strategy was not $FP_3(0)$ but a slightly more exploiting variant. This was strategy #948, which corresponds to (1 0 0 1, 1 0 .5 0), with a dominance score of .9327. Again the top 15 edge strategies were all variants of $FP_3(0)$, with no more than four parameters different. $FP_3(0)$ finished 8th, with a dominance score of .8796. This suggests that the 3-ply edge strategy with the highest dominance score over the entire space of 3-ply strategies is likely to be of the form (1 0 0 1, 1 0 k 0), where $.5 < k < 1$.

To test this hypothesis, we next compared the dominance scores of 99 strategies (1 0 0 1, 1 0 k 0) against all 3-ply edge strategies. A noise level of .01 was used, and k was varied from .01 to .99. The winner of this tournament was $k = 0.59$, with a dominance score of .9361. This strategy, (1 0 0 1, 1 0 0.59 0), is likely to be extremely successful against the space of all n-ply strategies. Thus $FP_3(0)$ appears to be optimal for $T = 4$, but for $T = 5$, strategies which sometimes defect after dDc perform slightly better.

9 Conclusions

When is a strategy "optimal" in the Alternating Prisoner's Dilemma with noise? Since no strategy can achieve a maximum score against every other strategy, we instead search for *optimality criteria* which are strongly correlated with high average performance in various Prisoner's Dilemma models (including round-robin tournaments and evolutionary invasion models). In order to perform well against a variety of other strategies, an "optimal" strategy must be able to achieve mutual cooperation with its clone, resist exploitation by defectors, and exploit unconditional cooperators. These three criteria (*self-cooperating*, *D-unexploitable*, *C-exploiting*) were defined and examined for Alternating Prisoner's Dilemma games with no noise, infinitesimal noise, finite noise, and arbitrary noise. Most of the strategies commonly discussed in the Alternating Prisoner's Dilemma literature are "2-ply strategies" such as Pavlov and Firm But Fair. These "low memory" strategies make decisions based only on the last two turns (where a "turn" is one player's decision whether to cooperate or defect). However, none of these strategies can simultaneously meet all three of the optimality criteria: a 2-ply strategy which is totally unexploitable by defectors, and can exploit unconditional cooperators, cannot attain any level of self-cooperation in the infinitesimal noise APD.

Higher memory strategies such as *Firm Pavlov*, however, can meet all three optimality criteria: the $FP_3(0)$ strategy (1 0 0 1, 1 0 1 0) is "totally self-cooperating", "totally D-unexploitable", and "C-exploiting", and thus is optimal for the infinitesimal noise APD. Other members of the Firm Pavlov and Firm But Fair strategy classes (each including an infinite memory "first defector" strategy and its n-ply approximations with varying generosity), can meet even more rigorous optimality criteria. For example, a 5-ply Firm Pavlov strategy has a self-payoff of $R - 4\epsilon$ or higher for any finite noise level ϵ, and thus it is "finite noise resistant". Since it also meets the three criteria for infinitesimal noise optimality, it is "finite noise optimal". The infinite memory Firm Pavlov strategy not only meets these criteria, but also can recover quickly from any sequence of errors in a game against its clone, and hence it is "arbitrary noise optimal".

These optimality criteria can be applied to analyze other Prisoner's Dilemma games with or without noise. Strategies which meet the criteria are likely to be successful in a variety of other performance measures. The Firm Pavlov and Firm But Fair strategy classes were shown to perform well in standard round-robin tournament simulations, adaptive dynamics models, and evolutionary invasion models. In addition to considering standard evolutionary and non-evolutionary models, we also proposed the *dominance criterion*, a general measure of evolutionary performance. This measure takes into account both evolutionary stability (ability to resist invasion) and evolutionary potency (ability to invade other strategies), producing a combined measure which can be used to examine relative evolutionary performance for any set of strategies. We conducted a number of "dominance tournaments", a round-robin tournament format similar to Axelrod's, except that dominance scores (instead of average payoffs) are computed for each pair of strategies. In particular, the Firm Pavlov strategies achieved extremely high dominance scores against the set of all 3-ply edge strategies, suggesting that these strategies will be extremely successful in a wide variety of evolutionary interactions. In fact, when a round robin dominance tournament was conducted among all 3-ply edge strategies, the winner was a 3-ply Firm Pavlov strategy.

To succeed in the Prisoner's Dilemma, a strategy must be "friendly" enough to cooperate with its clone, "pragmatic" enough to exploit unconditional cooperators, and "wary" enough to resist exploitation by defectors. The Firm Pavlov strategy $FP_3(0)$ can be thought of as a model for this type of behavior: it responds to cooperation with cooperation, except when it can continue exploiting an unconditional cooperator. Similarly, it responds to defection with defection, except when the opponent's defection is in response to its provocation. This 3-ply strategy combines the advantages of several 2-ply strategies: like Firm But Fair, it is self-cooperating under noise; like Tit for Tat, it is unexploitable; and like Pavlov, it can exploit unconditional cooperators. The Firm Pavlov strategies not only meet our stringent optimality criteria, but also achieve remarkable success in round-robin tournaments and evolutionary interactions: these higher memory strategies are truly "optimal under noise" in the Alternating Prisoner's Dilemma.

10 Bibliography

Axelrod, R. (1984). *The Evolution of Cooperation.* New York: Basic Books.

Axelrod, R. (1987). The evolution of strategies in the iterated prisoner's dilemma, in Davis, D. (ed.), *Genetic Algorithms and Simulated Annealing*, London: Pitman.

Axelrod, R. & Hamilton, W.D. (1981). The evolution of cooperation. *Science*, 211: 1390-1396.

Boyd, R. & Lorberbaum, J. (1987). No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, 327: 58-59.

Frean, M.R. (1994). The prisoner's dilemma without synchrony. *Proc. Roy. Soc. Lond. B*, 257: 75-79.

Hofbauer, J. & Sigmund, K. (1990). Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.*, 4: 75-79.

Kraines, D. & Kraines, V. (1989). Pavlov and the prisoner's dilemma. *Theory and Decision*, 26: 47-79.

Kraines, D. & Kraines, V. (1993). Learning to cooperate with Pavlov. *Theory and Decision*, 35: 107-150.

Kraines, D. & Kraines, V. (2000). Natural selection of memory-one strategies in the iterated prisoner's dilemma. *J. Theor. Biol.*, 203: 335-355.

Maynard Smith, J. (1982). *Evolution and the Theory of Games.* Cambridge: Cambridge University Press.

Maynard Smith, J. & Price, G.R. (1973). The logic of animal conflict. *Nature*, 246: 15-18.

Milinski, M. (1987). Tit for tat in sticklebacks and the evolution of cooperation. *Nature*, 325: 434-435.

Nowak, M. (1990). An evolutionarily stable strategy may be inaccessible. *Theoret. Population Biol.*, 142: 237-241.

Nowak, M. & Sigmund, K. (1989). Game dynamical aspects of the prisoner's dilemma. *J. Appl. Math. Comp.*, 30: 191-213.

Nowak, M. & Sigmund, K. (1990). The evolution of stochastic strategies in the prisoner's dilemma. *Acta Appl. Math.*, 20: 247-265.

Nowak, M. & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, 355: 250-253.

Nowak, M. & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit for tat in the prisoner's dilemma game. *Nature*, 364: 56-58.

Nowak, M. & Sigmund, K. (1994). The alternating prisoner's dilemma. *J. Theor. Biol.*, 168: 219-226.

Trivers, R. (1985). *Social Evolution.* Menlo Park: Benjamin Cummings.

Wilkinson, G.S. (1984). Reciprocal food-sharing in the vampire bat. *Nature*, 308: 181-184.

Wu, H. & Axelrod, R. (1995). How to cope with noise in the iterated prisoner's dilemma. *Journal of Conflict Resolution*, 39: 183-189.

11 Appendix: Round Robin Tournament Results

Figures 1-2 are graphs of $(R-w)/e$ as a function of e and g for the FP_3 and FP_5 strategies respectively.

Table 13 lists the results for the five 63-strategy round-robin tournaments performed in the "Performance of FBF and FP" section. For each of the 63 strategies, a score (average payoff per turn) and ranking (1-63) is given for each tournament. The five tournaments are: $T=5$, $\epsilon=.01$; $T=4$, $\epsilon=.01$; $T=5$, $\epsilon=.05$; $T=4$, $\epsilon=.05$; $T=5$, $\epsilon=.01$, no $g=0$ strategies.

Table 14 lists the results for the 29 mini-tournaments conducted in the "Performance of FBF and FP" section. For each strategy, we give its margin of victory (except for $FBF_3(0)$, which finished 2nd in its tournament).

Table 15 lists the results for the six 63-strategy dominance tournaments performed in the "Evolutionary Performance of FBF and FP" section. For each of the 63 strategies, a dominance score and ranking (1-63) is given for each tournament. The six tournaments are: $T = 5$ round robin; $T = 4$ round robin; $T = 5$ vs. corner; $T = 4$ vs. corner; $T = 5$ vs. edge; $T = 4$ vs. edge.

Tables 16-19 list the top 15 3-ply corner and edge strategies for $T = 5$ and $T = 4$, along with each strategy's dominance score.

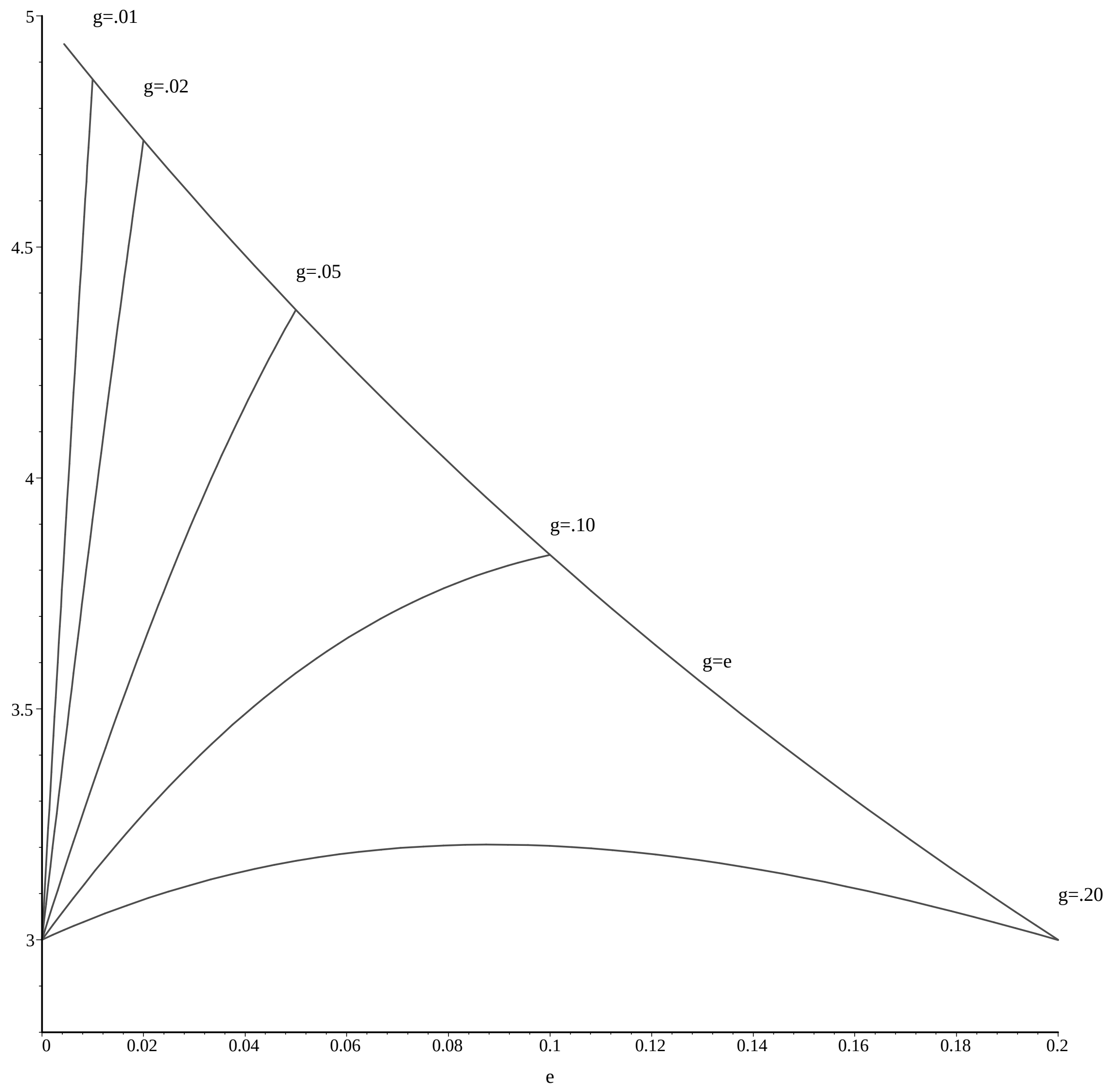

Figure 1: $(R - w)/e$ as a function of e,g for FP_3

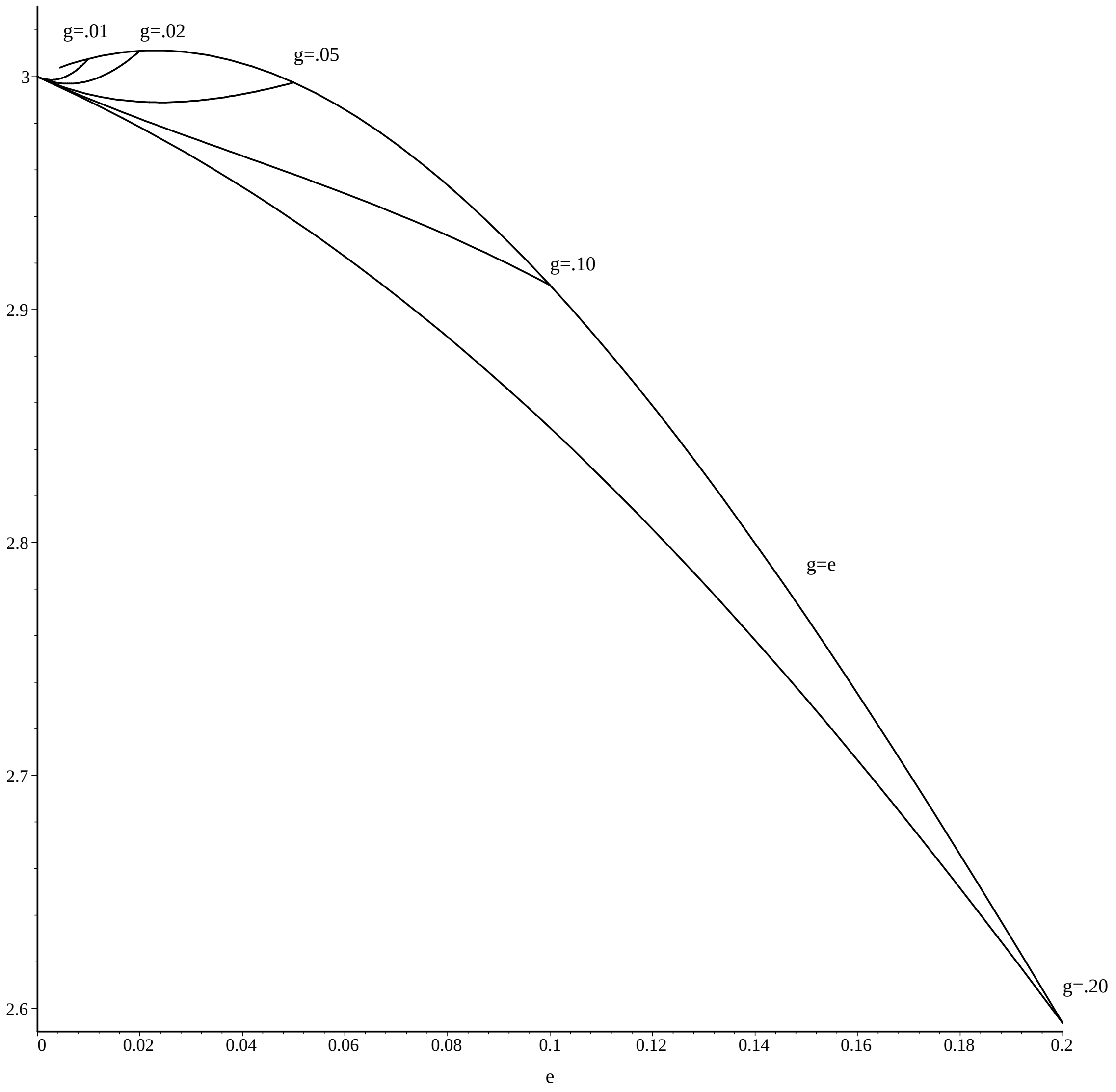

Figure 2: $(R-w)/e$ as a function of e,g for FP_5

Strategy	(.01,5)	(.01,4)	(.05,5)	(.05,4)	(.01,5,$g \neq 0$)
g1	2.8515 (59)	2.8421 (59)	2.7053 (57)	2.6630 (57)	2.8440 (55)
g2	2.8907 (56)	2.8812 (56)	2.7259 (56)	2.6831 (56)	2.8870 (53)
g3	2.8769 (58)	2.8674 (58)	2.6845 (59)	2.6423 (59)	2.8729 (54)
Pavlov	2.9154 (28.5)	2.8817 (55)	2.7869 (49)	2.6931 (55)	2.9118 (30)
RAND	2.5007 (62)	2.2149 (62)	2.4834 (61)	2.2001 (62)	2.5220 (56)
ALLD	2.1482 (63)	1.8608 (63)	2.2205 (63)	1.9149 (63)	2.2412 (58)
ALLC	2.5189 (61)	2.5106 (61)	2.4439 (62)	2.4047 (61)	2.5140 (57)
$FBF_2(0)$	2.8190 (60)	2.8100 (60)	2.6744 (60)	2.6343 (60)	X
$FBF_2(.1)$	2.8841 (57)	2.8747 (57)	2.7029 (58)	2.6620 (58)	2.8916 (52)
$FBF_2(.2)$	2.8969 (55)	2.8875 (54)	2.7355 (55)	2.6936 (54)	2.8982 (51)
$FBF_2(.3)$	2.9024 (54)	2.8929 (53)	2.7540 (54)	2.7115 (53)	2.9013 (50)
$FBF_2(.4)$	2.9054 (53)	2.8959 (52)	2.7663 (53)	2.7233 (52)	2.9031 (49)
$FBF_2(.5)$	2.9073 (51)	2.8977 (50)	2.7751 (52)	2.7318 (51)	2.9042 (48)
$FBF_2(.6)$	2.9086 (48)	2.8990 (47)	2.7819 (51)	2.7384 (50)	2.9050 (47)
$FBF_2(.7)$	2.9095 (47)	2.9000 (46)	2.7873 (48)	2.7436 (47)	2.9056 (46)
$FBF_2(.8)$	2.9102 (46)	2.9007 (45)	2.7919 (42.5)	2.7479 (40)	2.9060 (45)
$FBF_2(.9)$	2.9108 (45)	2.9013 (44)	2.7957 (38)	2.7516 (34)	2.9064 (44)
$FBF_2(1)$	2.9113 (44)	2.9017 (43)	2.7975 (32)	2.7533 (27)	2.9067 (43)
$FBF_3(0)$	2.9068 (52)	2.8973 (51)	2.7827 (50)	2.7393 (49)	X
$FBF_3(.1)$	2.9152 (32)	2.9056 (28.5)	2.7882 (47)	2.7445 (46)	2.9135 (24)
$FBF_3(.2)$	2.9156 (26.5)	2.9061 (23.5)	2.7933 (41)	2.7495 (37)	2.9127 (26.5)
$FBF_3(.3)$	2.9152 (32)	2.9057 (27)	2.7956 (39)	2.7517 (33)	2.9118 (30)
$FBF_3(.4)$	2.9146 (36.5)	2.9051 (32)	2.7968 (36)	2.7528 (29)	2.9108 (34.5)
$FBF_3(.5)$	2.9140 (39)	2.9044 (35)	2.7974 (33.5)	2.7533 (27)	2.9100 (38)
$FBF_3(.6)$	2.9134 (40)	2.9038 (36)	2.7976 (30.5)	2.7535 (22.5)	2.9092 (39)
$FBF_3(.7)$	2.9128 (41)	2.9032 (38)	2.7977 (28.5)	2.7536 (20.5)	2.9085 (40)
$FBF_3(.8)$	2.9122 (42)	2.9027 (40)	2.7977 (28.5)	2.7535 (22.5)	2.9078 (41)
$FBF_3(.9)$	2.9117 (43)	2.9021 (41.5)	2.7976 (30.5)	2.7534 (24.5)	2.9072 (42)
$FP_3(0)$	2.9174 (23)	2.9021 (41.5)	2.7919 (42.5)	2.7400 (48)	X
$FP_3(.1)$	2.9258 (7.5)	2.9104 (8.5)	2.7974 (33.5)	2.7453 (43)	2.9250 (3)
$FP_3(.2)$	2.9262 (3.5)	2.9109 (3.5)	2.8026 (18)	2.7503 (36)	2.9242 (5.5)
$FP_3(.3)$	2.9258 (7.5)	2.9105 (7)	2.8050 (16)	2.7525 (30.5)	2.9233 (8.5)
$FP_3(.4)$	2.9253 (11.5)	2.9099 (12)	2.8061 (14)	2.7536 (20.5)	2.9224 (12.5)
$FP_3(.5)$	2.9246 (15)	2.9093 (15)	2.8067 (12)	2.7542 (16)	2.9215 (16)
$FP_3(.6)$	2.9240 (16)	2.9086 (16)	2.8070 (9.5)	2.7544 (11.5)	2.9207 (17)
$FP_3(.7)$	2.9234 (17)	2.9080 (17)	2.8071 (7.5)	2.7545 (9.5)	2.9200 (18)
$FP_3(.8)$	2.9228 (18)	2.9075 (18)	2.8071 (7.5)	2.7544 (11.5)	2.9193 (19)
$FP_3(.9)$	2.9223 (19)	2.9070 (19)	2.8070 (9.5)	2.7543 (13.5)	2.9187 (20)
$FP_3(1)$	2.9219 (20)	2.9065 (20)	2.8069 (11)	2.7542 (16)	2.9182 (21)
$FBF_5(0)$	2.9079 (49.5)	2.8984 (48.5)	2.7884 (46)	2.7447 (45)	X
$FBF_5(.1)$	2.9153 (30)	2.9058 (25.5)	2.7917 (44)	2.7480 (39)	2.9138 (23)
$FBF_5(.2)$	2.9159 (24.5)	2.9064 (21)	2.7949 (40)	2.7510 (35)	2.9132 (25)
$FBF_5(.3)$	2.9159 (24.5)	2.9063 (22)	2.7964 (37)	2.7525 (30.5)	2.9127 (26.5)

$FBF_5(.4)$	2.9156 (26.5)	2.9061 (23.5)	2.7972 (35)	2.7533 (27)	2.9122 (28)
$FBF_5(.5)$	2.9154 (28.5)	2.9058 (25.5)	2.7978 (26.5)	2.7538 (19)	2.9118 (30)
$FBF_5(.6)$	2.9152 (32)	2.9056 (28.5)	2.7981 (24.5)	2.7541 (18)	2.9114 (32)
$FBF_5(.7)$	2.9149 (34)	2.9054 (30)	2.7983 (23)	2.7543 (13.5)	2.9111 (33)
$FBF_5(.8)$	2.9147 (35)	2.9052 (31)	2.7985 (22)	2.7545 (9.5)	2.9108 (34.5)
$FBF_5(.9)$	2.9146 (36.5)	2.9050 (33)	2.7986 (21)	2.7546 (7.5)	2.9106 (36)
$FBF_5(1)$	2.9144 (38)	2.9048 (34)	2.7987 (20)	2.7546 (7.5)	2.9104 (37)
$FP_5(0)$	2.9185 (21.5)	2.9032 (38)	2.7978 (26.5)	2.7457 (42)	X
$FP_5(.1)$	2.9259 (6)	2.9106 (5.5)	2.8011 (19)	2.7489 (38)	2.9253 (2)
$FP_5(.2)$	2.9265 (1.5)	2.9112 (1)	2.8043 (17)	2.7520 (32)	2.9247 (4)
$FP_5(.3)$	2.9265 (1.5)	2.9111 (2)	2.8058 (15)	2.7534 (24.5)	2.9242 (5.5)
$FP_5(.4)$	2.9262 (3.5)	2.9109 (3.5)	2.8066 (13)	2.7542 (16)	2.9237 (7)
$FP_5(.5)$	2.9260 (5)	2.9106 (5.5)	2.8072 (6)	2.7547 (6)	2.9233 (8.5)
$FP_5(.6)$	2.9258 (9)	2.9104 (8.5)	2.8075 (5)	2.7550 (5)	2.9229 (10)
$FP_5(.7)$	2.9255 (10)	2.9102 (10)	2.8077 (4)	2.7552 (4)	2.9226 (11)
$FP_5(.8)$	2.9253 (11.5)	2.9100 (11)	2.8079 (3)	2.7554 (3)	2.9224 (12.5)
$FP_5(.9)$	2.9252 (13)	2.9098 (13)	2.8080 (2)	2.7555 (1.5)	2.9221 (14)
$FP_5(1)$	2.9250 (14)	2.9097 (14)	2.8081 (1)	2.7555 (1.5)	2.9219 (15)
FBF_∞	2.9079 (49.5)	2.8984 (48.5)	2.7886 (45)	2.7450 (44)	2.9140 (22)
FP_∞	2.9185 (21.5)	2.9032 (38)	2.7981 (24.5)	2.7460 (41)	2.9255 (1)

Table 13: Round robin tournament results

Strategy	Margin of victory	Rank
$FBF_2(.2)$	.0281	18
$FBF_2(.4)$	.0241	21
$FBF_2(.6)$	.0155	24
$FBF_2(.8)$	.0073	25
$FBF_3(0)$	-.0006	29
$FBF_3(.2)$	.0369	11
$FBF_3(.4)$	.0259	20
$FBF_3(.6)$	.0156	23
$FBF_3(.8)$	.0069	26
$FP_3(0)$	.0308	16
$FP_3(.2)$	.0682	2
$FP_3(.4)$	.0572	6
$FP_3(.6)$	.0469	8
$FP_3(.8)$	.0382	10
$FP_3(1)$	.0313	15
$FBF_5(0)$	.0017	27.5
$FBF_5(.2)$	.0399	9
$FBF_5(.4)$	.0345	12
$FBF_5(.6)$	.0297	17
$FBF_5(.8)$	.0260	19
$FBF_5(1)$	.0233	22
$FP_5(0)$	.0330	14
$FP_5(.2)$	.0713	1
$FP_5(.4)$	.0659	3
$FP_5(.6)$	.0611	4
$FP_5(.8)$	.0574	5
$FP_5(1)$	.0546	7
FBF_∞	.0017	27.5
FP_∞	.0331	13

Table 14: Round robin mini-tournament results

Strategy	(5,R)	(4,R)	(5,C)	(4,C)	(5,E)	(4,E)
g1	.2286 (53)	.3037 (49)	.4547 (59)	.7813 (55)	.4202 (59)	.7863 (56)
g2	.2601 (52)	.3358 (45)	.3343 (60)	.6686 (58)	.2885 (60)	.6693 (58)
g3	.2991 (49)	.3210 (47)	.3135 (61)	.4311 (62)	.2386 (61)	.4158 (62)
Pavlov	.0541 (62)	.0433 (62)	.5455 (57)	.5513 (59)	.5335 (57)	.5350 (59)
RAND	.4495 (32)	.1316 (57)	.5149 (58)	.5122 (61)	.5279 (58)	.5039 (60)
ALLD	.3716 (43)	.2648 (52)	.8081 (22)	.7753 (56)	.8285 (18)	.7986 (55)
ALLC	.1164 (59)	.1161 (59)	.2089 (63)	.2247 (63)	.1617 (63)	.2014 (63)
$FBF_2(0)$	.0317 (63)	.0409 (63)	.2241 (62)	.5098 (60)	.2174 (62)	.5029 (61)
$FBF_2(.1)$	.0699 (61)	.0759 (61)	.6485 (51)	.8080 (51)	.6506 (50)	.8180 (51)
$FBF_2(.2)$	.0950 (60)	.1008 (60)	.6772 (48)	.8181 (48)	.6834 (46)	.8510 (45)
$FBF_2(.3)$	.1220 (58)	.1283 (58)	.6841 (46)	.8225 (47)	.6973 (42)	.8676 (40)
$FBF_2(.4)$	.1536 (57)	.1611 (55)	.6846 (45)	.8297 (45)	.7004 (41)	.8689 (39)
$FBF_2(.5)$	.2161 (54)	.2304 (53)	.6947 (43)	.8392 (41)	.6942 (43)	.8646 (41)
$FBF_2(.6)$	.2925 (51)	.3052 (48)	.6899 (44)	.8384 (42)	.6787 (47)	.8540 (44)
$FBF_2(.7)$	.3854 (41)	.4028 (38)	.6736 (49)	.8273 (46)	.6540 (49)	.8383 (48)
$FBF_2(.8)$	.3963 (39)	.4180 (36)	.6464 (53)	.8098 (50)	.6230 (53)	.8196 (50)
$FBF_2(.9)$	.3925 (40)	.4321 (34)	.6121 (55)	.7910 (54)	.5865 (55)	.7993 (54)
$FBF_2(1)$	.4416 (33)	.6112 (22)	.5628 (56)	.7641 (57)	.5433 (56)	.7768 (57)
$FBF_3(0)$	.1913 (56)	.1916 (54)	.7678 (37)	.8979 (34)	.7834 (31)	.9211 (28)
$FBF_3(.1)$	.3328 (46)	.3555 (43)	.7967 (28)	.9186 (26)	.8016 (25)	.9302 (22)
$FBF_3(.2)$	.4123 (36)	.5041 (31)	.7856 (33)	.9089 (31)	.7852 (30)	.9206 (30)
$FBF_3(.3)$	.5818 (27)	.6728 (18)	.7713 (36)	.8982 (33)	.7659 (36)	.9090 (34)
$FBF_3(.4)$	.6816 (16)	.7163 (16)	.7545 (39)	.8866 (35)	.7439 (38)	.8951 (36)
$FBF_3(.5)$	.7289 (13)	.7597 (13)	.7329 (40)	.8726 (37)	.7193 (40)	.8794 (37)
$FBF_3(.6)$	.7472 (12)	.7738 (12)	.7073 (42)	.8527 (39)	.6919 (44)	.8615 (42)
$FBF_3(.7)$	.7035 (15)	.7271 (15)	.6786 (47)	.8327 (43)	.6612 (48)	.8422 (46)
$FBF_3(.8)$	.6611 (17)	.6830 (17)	.6481 (52)	.8136 (49)	.6273 (52)	.8220 (49)
$FBF_3(.9)$	.6269 (21)	.6473 (20)	.6151 (54)	.7937 (53)	.5886 (54)	.8004 (53)
$FP_3(0)$	.1917 (55)	.1560 (56)	.8910 (14)	.9380 (15)	.8796 (8)	.9514 (9)
$FP_3(.1)$	.2970 (50)	.2880 (51)	.9044 (7)	.9477 (7)	.8833 (6)	.9552 (6)
$FP_3(.2)$	.3422 (45)	.3381 (44)	.8908 (15)	.9393 (14)	.8692 (12)	.9483 (12)
$FP_3(.3)$	.3766 (42)	.3603 (41)	.8739 (16)	.9289 (20)	.8515 (16)	.9386 (16)
$FP_3(.4)$	.4024 (38)	.3895 (40)	.8543 (17)	.9167 (29)	.8311 (17)	.9265 (25)
$FP_3(.5)$	.4339 (34)	.4210 (35)	.8328 (18)	.9028 (32)	.8084 (20)	.9126 (33)
$FP_3(.6)$	.4520 (31)	.4366 (33)	.8094 (20)	.8856 (36)	.7833 (32)	.8967 (35)
$FP_3(.7)$	.4087 (37)	.3972 (39)	.7843 (35)	.8680 (38)	.7552 (37)	.8792 (38)
$FP_3(.8)$	.3636 (44)	.3557 (42)	.7575 (38)	.8501 (40)	.7245 (39)	.8607 (43)
$FP_3(.9)$	.3319 (47)	.3270 (46)	.7244 (41)	.8304 (44)	.6888 (45)	.8407 (47)
$FP_3(1)$	.3064 (48)	.3031 (50)	.6725 (50)	.7992 (52)	.6447 (51)	.8174 (52)
$FBF_5(0)$	.6544 (18)	.7357 (14)	.8048 (24)	.9292 (19)	.8071 (23)	.9346 (20)
$FBF_5(.1)$	.8142 (11)	.8186 (11)	.8130 (19)	.9340 (16)	.8119 (19)	.9374 (17)
$FBF_5(.2)$	.8361 (10)	.8407 (10)	.8083 (21)	.9309 (18)	.8077 (21)	.9349 (18)
$FBF_5(.3)$	.8563 (8)	.8612 (8)	.8040 (25)	.9279 (21)	.8031 (24)	.9323 (21)

$FBF_5(.4)$	.8780 (7)	.8832 (7)	.8003 (26)	.9254 (22)	.7988 (26)	.9298 (23)
$FBF_5(.5)$	.9019 (5)	.9074 (5)	.7969 (27)	.9232 (23)	.7945 (27)	.9274 (24)
$FBF_5(.6)$	.9320 (3)	.9380 (3)	.7938 (29)	.9213 (24)	.7905 (28)	.9252 (26)
$FBF_5(.7)$	.9420 (2)	.9476 (2)	.7910 (30)	.9196 (25)	.7867 (29)	.9230 (27)
$FBF_5(.8)$	.9114 (4)	.9172 (4)	.7885 (31)	.9182 (27)	.7832 (33)	.9210 (29)
$FBF_5(.9)$	.8830 (6)	.8891 (6)	.7863 (32)	.9170 (28)	.7798 (34)	.9192 (31)
$FBF_5(1)$	.8524 (9)	.8561 (9)	.7845 (34)	.9161 (30)	.7769 (35)	.9175 (32)
$FP_5(0)$	.4240 (35)	.4106 (37)	.9079 (5)	.9540 (3)	.8884 (4)	.9573 (4)
$FP_5(.1)$	.5469 (30)	.5027 (32)	.9146 (1)	.9555 (2)	.8923 (1)	.9599 (1)
$FP_5(.2)$	.5726 (28)	.5218 (30)	.9122 (3)	.9526 (4)	.8890 (2.5)	.9583 (2)
$FP_5(.3)$	.5899 (26)	.5421 (28)	.9100 (4)	.9501 (5)	.8854 (5)	.9563 (5)
$FP_5(.4)$	.5998 (24)	.5571 (27)	.9071 (6)	.9480 (6)	.8817 (7)	.9543 (7)
$FP_5(.5)$	.6124 (23)	.5788 (25)	.9042 (8)	.9462 (8)	.8781 (9)	.9524 (8)
$FP_5(.6)$	.6417 (20)	.6082 (23)	.9015 (9)	.9446 (9)	.8747 (10)	.9505 (10)
$FP_5(.7)$	.6515 (19)	.6169 (21)	.8991 (10)	.9432 (10)	.8714 (11)	.9487 (11)
$FP_5(.8)$	.6204 (22)	.5860 (24)	.8969 (11)	.9419 (11)	.8683 (13)	.9469 (13)
$FP_5(.9)$	.5916 (25)	.5574 (26)	.8948 (12)	.9408 (12)	.8654 (14)	.9453 (14)
$FP_5(1)$	.5646 (29)	.5301 (29)	.8931 (13)	.9399 (13)	.8628 (15)	.9438 (15)
FBF_∞	.9567 (1)	.9612 (1)	.8068 (23)	.9311 (17)	.8073 (22)	.9348 (19)
FP_∞	.7089 (14)	.6499 (19)	.9136 (2)	.9559 (1)	.8890 (2.5)	.9577 (3)

Table 15: Dominance tournament results

Strategy #	Strategy	Dominance
90	(1 0 0 1, 1 0 1 0)	.8910
50	(1 0 0 0, 1 1 0 0)	.8570
34	(1 0 0 0, 0 1 0 0)	.8468
33	(0 0 0 0, 0 1 0 0)	.8422
49	(0 0 0 0, 1 1 0 0)	.8379
35	(0 1 0 0, 0 1 0 0)	.8357
129	(0 0 0 0, 0 0 0 1)	.8275
161	(0 0 0 0, 0 1 0 1)	.8267
162	(1 0 0 0, 0 1 0 1)	.8226
131	(0 1 0 0, 0 0 0 1)	.8187
1	(0 0 0 0, 0 0 0 0)	.8081
2	(1 0 0 0, 0 0 0 0)	.8057
3	(0 1 0 0, 0 0 0 0)	.8049
130	(1 0 0 0, 0 0 0 1)	.8015
36	(1 1 0 0, 0 1 0 0)	.8010

Table 16: Top 3-ply corner strategies, $T = 5$

Strategy #	Strategy	Dominance
90	(1 0 0 1, 1 0 1 0)	.9380
94	(1 0 1 1, 1 0 1 0)	.8979
50	(1 0 0 0, 1 1 0 0)	.8266
33	(0 0 0 0, 0 1 0 0)	.8202
49	(0 0 0 0, 1 1 0 0)	.8165
34	(1 0 0 0, 0 1 0 0)	.8159
35	(0 1 0 0, 0 1 0 0)	.8123
89	(0 0 0 1, 1 0 1 0)	.8090
210	(1 0 0 0, 1 0 1 1)	.8066
218	(1 0 0 1, 1 0 1 1)	.7992
122	(1 0 0 1, 1 1 1 0)	.7976
74	(1 0 0 1, 0 0 1 0)	.7783
73	(0 0 0 1, 0 0 1 0)	.7765
1	(0 0 0 0, 0 0 0 0)	.7753
161	(0 0 0 0, 0 1 0 1)	.7735

Table 17: Top 3-ply corner strategies, $T = 4$

Strategy #	Strategy	Dominance
948	(1 0 0 1, 1 0 .5 0)	.9327
867	(1 0 0 1, .5 0 .5 0)	.9309
1191	(1 0 0 1, 1 .5 .5 0)	.9045
957	(1 0 .5 1, 1 0 .5 0)	.9035
870	(1 .5 0 1, .5 0 .5 0)	.9033
1110	(1 0 0 1, .5 .5 .5 0)	.9012
1200	(1 0 .5 1, 1 .5 .5 0)	.8941
1677	(1 0 0 1, 1 0 1 0)	.8796
1920	(1 0 0 1, 1 .5 1 0)	.8764
1209	(1 0 1 1, 1 .5 .5 0)	.8700
876	(1 0 1 1, .5 0 .5 0)	.8682
1596	(1 0 0 1, .5 0 1 0)	.8682
1839	(1 0 0 1, .5 .5 1 0)	.8623
1119	(1 0 .5 1, .5 .5 .5 0)	.8614
966	(1 0 1 1, 1 0 .5 0)	.8611

Table 18: Top 3-ply edge strategies, $T = 5$

Strategy #	Strategy	Dominance
1677	(1 0 0 1, 1 0 1 0)	.9514
1920	(1 0 0 1, 1 .5 1 0)	.9500
948	(1 0 0 1, 1 0 .5 0)	.9419
1596	(1 0 0 1, .5 0 1 0)	.9321
1686	(1 0 .5 1, 1 0 1 0)	.9271
1839	(1 0 0 1, .5 .5 1 0)	.9262
1200	(1 0 .5 1, 1 .5 .5 0)	.9261
1191	(1 0 0 1, 1 .5 .5 0)	.9243
1929	(1 0 .5 1, 1 .5 1 0)	.9242
957	(1 0 .5 1, 1 0 .5 0)	.9231
867	(1 0 0 1, .5 0 .5 0)	.9219
1695	(1 0 1 1, 1 0 1 0)	.9211
1209	(1 0 1 1, 1 .5 .5 0)	.9196
1938	(1 0 1 1, 1 .5 1 0)	.9182
870	(1 .5 0 1, .5 0 .5 0)	.9027

Table 19: Top 3-ply edge strategies, $T = 4$