# Genetic Algorithms, Tournament Selection, and the Effects of Noise

Brad L. Miller
Dept. of Computer Science
University of Illinois
at Urbana-Champaign

David E. Goldberg
Dept. of General Engineering
University of Illinois
at Urbana-Champaign

IlliGAL Report No. 95006
July 1995

Department of General Engineering
University of Illinois at Urbana-Champaign
117 Transportation Building
104 South Mathews Avenue
Urbana, IL 61801

# Genetic Algorithms, Tournament Selection, and the Effects of Noise

Brad L. Miller
Dept. of Computer Science
University of Illinois
at Urbana-Champaign
bmiller@uiuc.edu

David E. Goldberg
Dept. of General Engineering
University of Illinois
at Urbana-Champaign
deg@uiuc.edu

July 12, 1995

### Abstract

Tournament selection is a useful and robust selection mechanism commonly used by genetic algorithms. The selection pressure of tournament selection directly varies with the tournament size — the more competitors, the higher the resulting selection pressure. This article develops a model, based on order statistics, that can be used to quantitatively predict the resulting selection pressure of a tournament of a given size. This model is used to predict the convergence rates of genetic algorithms utilizing tournament selection.

While tournament selection is often used in conjunction with noisy (imperfect) fitness functions, little is understood about how the noise affects the resulting selection pressure. The model is extended to quantitatively predict the selection pressure for tournament selection utilizing noisy fitness functions. Given the tournament size and noise level of a noisy fitness function, the extended model is used to predict the resulting selection pressure of tournament selection. The accuracy of the model is verified using a simple test domain, the onemax (bit-counting) domain. The model is shown to accurately predict the convergence rate of a genetic algorithm using tournament selection in the onemax domain for a wide range of tournament sizes and noise levels.

The model developed in this paper has a number of immediate practical uses as well as a number of longer term ramifications. Immediately, the model may be used for determining appropriate ranges of control parameters, for estimating stopping times to achieve a specified level of solution quality, and for approximating convergence times in important classes of function evaluations that utilize sampling. Longer term, the approach of this study may be applied to better understand the delaying effects of function noise in other selection schemes or to approximate the convergence delays that result from inherently noisy operators such as selection, crossover, and mutation.

## 1 Introduction

There are many selection schemes for genetic algorithms (GAs), each with different characteristics. An ideal selection scheme would be simple to code, and efficient for both non-parallel and parallel architectures. Furthermore, a selection scheme should be able to adjust its selection pressure so as to tune its performance for different domains. Tournament selection is increasingly being used as a GA selection scheme because it satisfies all of the above criteria. It is simple to code and is efficient for both non-parallel and parallel architectures. Tournament selection can also adjust the selection pressure to adapt to different domain. Tournament selection pressure is increased (decreased) by simply increasing (decreasing) the tournament size. All of these factors have contributed to the increased usage of tournament selection as a selection mechanism for GAs.

Good progress was made some time ago (Goldberg and Deb, 1991) in understanding the convergence rates of various selection schemes, including tournament selection. Recently, building on work by Mühlenbein and Schlierkamp-Voosen (1993), this understanding has been refined to better understand the timing and degree of convergence more accurately (Thierens and Goldberg, 1994). Despite this progress, this detailed timing and degree of convergence analysis has not yet been extended to tournaments other than binary ($s = 2$); nor has the analysis been applied to domains other than deterministic ones. In this paper, we do these two things.

The purpose of this paper is to develop a model for the selection pressure of tournament selection. This model, based on order statistics, quantitatively predicts the selection pressure resulting from both different tournament sizes and noise levels. Given the current population fitness mean and variance, the model can predict the average population fitness of the next generation. The model can also be used iteratively to predict the convergence rate of the GA over time. The predictive model is verified, using the onemax domain, under a range of tournament sizes and noise levels.

Section 2 provides the reader with background information needed to understand this paper, including tournament selection, noise, and order statistics. Sections 3 and 4 develop the predictive model for tournament selection. Section 3 develops a predictive model that handles varying tournament sizes for noiseless environments, and section 4 extends this model for noisy environments. Section 5 assesses the accuracy of the predictive model, using the onemax domain, for a variety of tournament sizes and noise levels. Application of the model for other research issues is described in Section 6. Some general conclusions from this research are presented in section 7.

# 2 Background

This section gives some background information needed to understand this paper. The first subsection describes selection schemes, selection pressure, and tournament section. The second subsection details noise, noisy fitness functions, and approximate fitness functions. Lastly, a brief overview of order statistics is given, focusing on the maximal order statistic for normal distributions.

## 2.1 Tournament Selection

Genetic algorithms use a selection mechanism to select individuals from the population to insert into a mating pool. Individuals from the mating pool are used to generate new offspring, with the resulting offspring forming the basis of the next generation. As the individuals in the mating pool are the ones whose genes are inherited by the next generation, it is desirable that the mating pool be comprised of "good" individuals. A selection mechanism in GAs is simply a process that favors the selection of better individuals in the population for the mating pool. The *selection pressure* is the degree to which the better individuals are favored: the higher the selection pressure, the more the better individuals are favored. This selection pressure drives the GA to improve the population fitness over succeeding generations. The convergence rate of a GA is largely determined by the selection pressure, with higher selection pressures resulting in higher convergence rates. Genetic algorithms are able to to identify optimal or near-optimal solutions under a wide range of selection pressure (Goldberg, Deb, and Thierens, 1993). However, if the selection pressure is too low, the convergence rate will be slow, and the GA will unnecessarily take longer to find the optimal solution. If the selection pressure is too high, there is an increased chance of the GA prematurely converging to an incorrect (sub-optimal) solution.

Tournament selection provides selection pressure by holding a tournament among $s$ competitors, with $s$ being the tournament size. The winner of the tournament is the individual with the highest fitness of the $s$ tournament competitors, and the winner is then inserted into the mating pool. The mating pool, being comprised of tournament winners, has a higher average fitness than the average population fitness. This fitness difference provides the selection pressure, which drives the GA to improve the fitness of each succeeding generation. Increased selection pressure can be provided by simply increasing the tournament size $s$, as the winner from a larger tournament will, on average, have a higher fitness than the winner of a smaller tournament.

## 2.2 Noise and Noisy Fitness Functions

The noise inherent in noisy fitness functions causes the tournament selection process to also be noisy. We assume that a noisy fitness function returns a fitness score for an individual equal to the sum of the real fitness of the individual plus some noise. In this paper, we assume that the noise is normally distributed and unbiased (mean of zero). This assumption is true or approximately true in many noisy domains, and allows the effects of noise to be more easily modeled.

There are many factors that may necessitate the use of noisy fitness functions. In some domains, there may be no known fitness function that can accurately assess an individual's fitness, so an approximate (noisy) fitness function must be used. Noisy information can also negatively affect the fitness evaluation. Noisy information can come from a variety of sources, including noisy sensor input, noisy data, knowledge uncertainty, and human error. To improve run-time performance, some GAs will utilize fast, but noisier, fitness functions instead of more accurate, but slower, fitness functions that may also be available. Sampling fitness functions are a good example of this phenomena, as a fitness function that uses sampling to assess an individual's fitness can use smaller sample sizes to increase run-time speed, at the expense of decreased accuracy of the fitness evaluation.

## 2.3 Order Statistics

This paper uses *order statistics* to further our understanding of tournament selection, and this section briefly reviews them. For a detailed description of order statistics, the reader should see David (1981).

If a random sample of size $n$ is arranged in ascending order of magnitude and then written as

$$x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n},$$

we can let the random variable $X_{i:n}$ represent the distribution of the corresponding $x_{i:n}$ over the space of all possible samples of size $n$. The variable $X_{i:n}$ is called the $i$th order statistic. The field of order statistics deals with the properties and applications of these random variables.

Of particular interest is the maximal order statistic $X_{n:n}$, which represents the distribution of the maximum member of a sample of size $n$. This is directly analogous to tournament selection, where the competitor with the maximum fitness is selected as the tournament winner.

The probability density function $p_{i:n}(x)$ of the $i$th order statistic, $X_{i:n}$, gives the probability that the $i$th highest individual from a sample of size $n$ will have a value of $x$. The value of $p_{i:n}(x)$ is calculated by

$$p_{i:n}(x) = n\binom{n-1}{i-1}P(x)^{i-1}(1-P(x))^{n-i},$$

where $P(x)$ represents the cumulative distribution function of $x$ (the probability that $\{X \leq x\}$). The probability that a single combination will have $i-1$ individuals less than or equal to $x$ and $n-i$ individuals greater than $x$ is given by the product $P(x)^{i-1}(1-P(x))^{n-i}$. However, there are many possible sample combinations that will yield the desired distribution of having $i-1$ individuals less than $x$ and $n-i$ individuals greater or equal to $x$. For $n$ individuals, there are $n$ slots that the $i$th greatest individual could occupy. For each of these slots, there are $\binom{n-1}{i-1}$ different ways of arranging the $i-1$ individuals that are less than or equal to $x$ among the $n-1$ remaining slots.

The expected value (mean) $u_{i:n}$ of an order statistic $X_{i:n}$ can thus be determined by

$$\begin{aligned} u_{i:n} &= \int_{-\infty}^{+\infty} x\ p_{i:n}(x)dx, \\ &= n\binom{n-1}{i-1}\int_{-\infty}^{+\infty} x\ P(x)^{i-1}(1-P(x))^{n-i}dP(x). \end{aligned}$$

For the maximal order statistic ($i=n$), the mean $u_{n:n}$ simplifies to

$$u_{n:n} = n\int_{-\infty}^{+\infty} x\ P(x)^{n-1}dP(x).$$

In this paper we are particularly interested in the normal distribution $N(\mu,\sigma^2)$, where $\mu$ and $\sigma^2$ are the mean and variance, respectively, of the normal distribution. For the standard normal distribution $N(0,1)$, $P(x)$ is the cumulative distribution function for the unit normal $\Phi(x)$, and thus $dP(x)$ is $\phi(x)dx = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$. The expected value (mean) of the maximal order statistic for the standard normal distribution is thus

$$u_{n:n} = n\int_{-\infty}^{+\infty} x\ \Phi(x)^{n-1}\phi(x)dx. \tag{1}$$

For samples of size $n = \{2, 3, 4, 5\}$, Equation 1 for $u_{n:n}$ can be solved exactly in terms of elementary functions. Table 1 gives the values for the mean of the maximal order statistic for $n = \{2, 3, 4, 5\}$ (see Balakrishnan and Cohen (1991) for derivations). For larger values of n, the means of the order statistics for the standard normal distribution have been tabulated extensively (Harter, 1961). The variances and covariances of the standard normal distribution order statistics can also be calculated, and are tabulated for $n \leq 20$ in Teichroew (1956), and for $n \leq 50$ in Tietjen, Kahaner, and Beckman (1977).

| $n$ | $\mu_{n:n}$ | Values of $\mu_{n:n}$ |
|---|---|---|
| 2 | $\frac{1}{\sqrt{\pi}}$ | 0.5642 |
| 3 | $\frac{3}{2\sqrt{\pi}}$ | 0.8463 |
| 4 | $\frac{6}{\pi\sqrt{\pi}} \tan^{-1}(\sqrt{2})$ | 1.0294 |
| 5 | $\frac{5}{4\sqrt{\pi}} + \frac{15}{2\pi\sqrt{\pi}} \sin^{-1}(\frac{1}{3})$ | 1.1630 |

Table 1: Expected Value of Maximal Order Statistic for Standard Normal Distribution.

## 3 Tournament Selection in Deterministic Environments

This section develops a predictive model for the selection pressure resulting from a tournament of size $s$ in a deterministic (noiseless) environment. In a noiseless environment, the fitness function can accurately assess the true fitness of an individual. We show that for a population whose fitness is normally distributed, the resulting tournament selection pressure is proportional to the product of the standard deviation of the population fitness and the maximal order statistic $\mu_{s:s}$.

In a deterministic environment, the fitness function returns the true fitness value of an individual. The population's fitness values, after crossover and mutation, are assumed to be normally distributed over the population. Although tournament selection by itself will generate a skewed (non-normal) distribution, the crossover and mutation operations 'remix' the population, which forces the distribution to become more normal. This normalizing effect of crossover and mutation allows the assumption of normally distributed population fitness to be reasonable for a wide variety of domains.

Let the population fitness in generation $t$ be normally distributed $N(\mu_{F,t}, \sigma^2_{F,t})$. The probability that an individual with fitness $f$ will win a tournament of $s$ individuals randomly picked from the population is given by

$$p(f = max(f_1 \ldots f_s)) = s \; P(F < f)^{s-1} p(f),$$

which represents the probability of an individual with fitness $f$ occurring along with $s-1$ individuals having lower fitness scores. There are $s$ different ways of arranging the $s-1$ "losers" and the "winner." The expected value of the tournament winner $\mu_{F,t+1}$ from a tournament of size $s$ can thus be calculated by

$$\begin{aligned}
\mu_{F,t+1} &= E[f = max(f_1 \ldots f_s)], \\
&= \int_{-\infty}^{+\infty} f \; p(f = max(f_1 \ldots f_s)) df, \\
&= s \int_{-\infty}^{+\infty} f \; P(f)^{s-1} p(f) df.
\end{aligned}$$

However, for a normally distributed population $N(\mu_{F,t}, \sigma^2_{F,t})$, $P(f) = \Phi(\frac{f - \mu_{F,t}}{\sigma_{F,t}})$, and

$$p(f) = \frac{dP(f)}{df} = \frac{1}{\sigma_{F,t}} \phi(\frac{f - \mu_{F,t}}{\sigma_{F,t}}).$$

Thus

$$\mu_{F,t+1} = \frac{s}{\sigma_{F,t}} \int_{-\infty}^{+\infty} f \; \Phi(\frac{f - \mu_{F,t}}{\sigma_{F,t}})^{s-1} \phi(\frac{f - \mu_{F,t}}{\sigma_{F,t}}) df.$$

Substituting $z = \frac{f-\mu_{F,t}}{\sigma_{F,t}}$ gives

$$\begin{aligned}
\mu_{F,t+1} &= s \int_{-\infty}^{+\infty} (\sigma_{F,t} z + \mu_{F,t}) \; \Phi(z)^{s-1} \phi(z) dz, \\
&= \mu_{F,t} \; s \int_{-\infty}^{+\infty} \Phi(z)^{s-1} \phi(z) dz + \sigma_{F,t} \; (s \int_{-\infty}^{+\infty} z \; \Phi(z)^{s-1} \phi(z) dz), \\
&= \mu_{F,t} [\Phi(z)^s]_{-\infty}^{+\infty} + \sigma_{F,t} \mu_{s:s}, \\
&= \mu_{F,t} + \sigma_{F,t} \mu_{s:s}. \qquad (2)
\end{aligned}$$

In Equation 2, $\mu_{s:s}$ is the effective selection pressure for a tournament of size $s$, and can be directly obtained from Table 1. From Equation 2, it is also apparent that the change in fitness between generations is given by:

$$\begin{aligned}
\Delta\mu_{F,t} &= \mu_{F,t+1} - \mu_{F,t}, \\
&= \sigma_{F,t} \mu_{s:s}. \qquad (3)
\end{aligned}$$

For binary tournaments ($s = 2$), this matches the result obtained in (Thierens and Goldberg, 1994), where the expected increase in the average population fitness for tournaments of size $s = 2$ was derived in a different manner using the difference between normal distributions. Their result, using the notation in this paper, was $\mu_{F,t+1} = \mu_{F,t} + \sigma_{F,t} \frac{1}{\sqrt{\pi}}$. This matches the result obtained using Equation 2 with a tournament size of $s = 2$, as $\mu_{2:2} = \frac{1}{\sqrt{\pi}}$ from Table 1. Note that the order statistic model derived in this paper is generalizable to all tournament sizes, and is not limited to $s = 2$.

Equation 3 shows that for tournaments of size $s$, the expected average population fitness increase is directly proportional to $\mu_{s:s}$, the expected value of the maximal order statistic of size $s$. Table 1 gives $\mu_{s:s}$, demonstrating that increasing the tournament size will cause successively smaller increases in the expected average population fitness.

## 4 Tournament Selection in Noisy Environments

This section extends the model developed above to accurately predict the selection pressure in the presence of noise. With noisy fitness functions, there is a chance that the winner of a tournament might not be the individual with the highest true fitness. This section concentrates on quantifying the reduction in tournament selection pressure due to noisy fitness functions.

The model derivation in this section has three major steps. First, the relationship between an individual's noisy fitness and true fitness values is determined, so that the expected true fitness value of an individual can be calculated from the noisy fitness evaluation. Next, the relationship is extended to handle subsets of individuals, so that the true fitness average of a subset of the population can be estimated from the average noisy fitness value of the subset. Lastly, we use the model derived in the previous section to estimate the average noisy fitness value of a particular subset of the population - the subset consisting of noisy tournament winners. This average noisy fitness value is then plugged into the formula found in the second step to estimate the average true fitness of the winners of noisy tournaments. The selection pressure, based on the expected true fitness value of the tournament winners, is thus be determined. The result is a predictive model for tournament selection that can handle varying noise and tournament sizes.

In a noisy environment, the noisy fitness $f'$ of an individual is given by $f' = f + noise$, where $f$ is the real fitness of the individual, and $noise$ is the noise inherent in the fitness function evaluation. As in the last section, the real fitness of the population $F$ is assumed to be normally distributed $N(\mu_{F,t}, \sigma_{F,t}^2)$. This section further assumes that the noise is unbiased and normally distributed $N(0, \sigma_N^2)$. This facilitates modeling the effects of the noise, and is a reasonable assumption for many domains. Using these assumptions, along with the additive property of normal distributions, gives that $F'$ is normally distributed $N(\mu_{F,t}, \sigma_{F,t}^2 + \sigma_N^2)$.

Although the real fitness value for an individual is unknown, the expected value can be determined from the individual's noisy fitness value, which is generated by a noisy fitness function evaluation. As both the true fitness and the noisy fitness are normally distributed, the bivariate normal distribution can be used to obtain the expected true fitness value of $F$ for a given noisy fitness value $f'$ of $F'$. For normal random variables $X$ and $Y$, the bivariate normal distribution states that the expected value of $Y$ for a specific value $x$ of $X$ is

$$E(Y|x) = \mu_Y + \rho_{XY}\frac{\sigma_Y}{\sigma_X}(x - \mu_X),$$

where $\rho_{XY}$ is the correlation coefficient for $X$ and $Y$. The correlation coefficient $\rho_{XY}$ can be calculated by $\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$, where $\sigma_{XY}$ is the covariance of $X$ and $Y$. The covariance between $F$ and $F'$ is simply $\sigma_F^2$, thus

$$\begin{aligned} E(F|f') &= \mu_F + \frac{\sigma_F^2}{\sigma_F \sigma_{F'}}\frac{\sigma_F}{\sigma_{F'}}(f' - \mu_{F'}), \\ &= \mu_F + \frac{\sigma_F^2}{\sigma_{F'}^2}(f' - \mu_{F'}), \\ &= \mu_F + \frac{\sigma_F^2}{\sigma_F^2 + \sigma_N^2}(f' - \mu_{F'}). \end{aligned} \tag{4}$$

As the above formula is linear, the expected value of $F$ for any subset $R$ of the population can be calculated using equation 4, with $f'$ set to the noisy fitness mean $\mu_R$ of the subset. Of course, the subset we are interested in is the noisy tournament winners. The expected mean of the noisy tournament winners of tournament size $s$ can be derived using the same derivation as for the deterministic case:

$$\begin{aligned} \mu_{F',t+1} &= \mu_{F',t} + \sigma_{F',t}\mu_{s:s}, \\ &= \mu_{F',t} + \sqrt{\sigma_{F,t}^2 + \sigma_N^2}\mu_{s:s}. \end{aligned}$$

Setting $f'$ to $\mu_{F',t+1}$ in equation 4 produces the expected true fitness value of the tournament winners:

$$\begin{aligned} E(F_{t+1}|\mu_{F',t+1}) &= \mu_{F,t+1}, \\ &= \mu_{F,t} + \frac{\sigma_{F,t}^2}{\sigma_{F,t}^2 + \sigma_N^2}(\mu_{F',t} + \sqrt{\sigma_{F,t}^2 + \sigma_N^2}\mu_{s:s} - \mu_{F',t}), \\ &= \mu_{F,t} + \frac{\sigma_{F,t}^2}{\sqrt{\sigma_{F,t}^2 + \sigma_N^2}}\mu_{s:s}. \end{aligned} \tag{5}$$

As expected, equation 5 reduces to equation 1, the formula for the deterministic (noiseless) case, when the noise variance $\sigma_N^2$ equals zero. Equation 5 is significant in that it predicts the convergence rate of a genetic algorithm using tournament selection for any tournament size or noise level.

# 5 Validation of Model

This section tests the accuracy of the predictive model, equation 5, using a sample domain. The domain used is the bit-counting, or onemax, domain, which works well for analysis as the variance can be determined from the average population fitness. This section uses equation 5 to predict the performance under a range of tournament sizes and noise levels. Experiments are then run that show that the predictive model is very accurate in determining the tournament selection pressure for different tournament sizes and noise levels.

## 5.1 Onemax Domain

The domain of interest is the onemax, which is also referred to as the bit-counting problem. The real fitness of an individual in this domain is simply the number of one bits in the chromosome. The optimal solution is the chromosome consisting of all one bits. This population fitness in this domain is binomially distributed, and the mean and variance of the the population fitness can therefore be calculated using binomial distribution properties. The population mean fitness at generation $t$ is given by $\overline{f(t)} = \mu_{F,t} = l\ p(t)$, where $l$ is the chromosome length, and $p(t)$ is the percentage of correct alleles in the population. The variance of the population at time $t$ is simply $\sigma_{F,t}^2 = l\ p(t)(1 - p(t))$.

The experiments in this paper all use the following GA configuration parameters. The chromosome length is $l = 100$, crossover is performed using the uniform crossover operator, and no mutation is used so as to better isolate the selection pressure effects. The proportion of correct alleles in the initial random population is assumed to be $p(0) = 0.5$. The population size is adjusted for different noise levels, as described in Goldberg, Deb, and Clark (1992). For the onemax domain, the population sizing equation reduces to $N = 8(\sigma_f^2 + \sigma_n^2)$, with the population variance $\sigma_f^2$ conservatively set to $\sigma_{F,0}^2 = l\ p(0)(1 - p(0)) = 25$. The noise variance $\sigma_n^2$ is user specified for each experiment. For experiments with a non-zero noise variance $\sigma_n^2$, a random number generated from the noisy distribution $N(0, \sigma_n^2)$ is added to the real fitness score for each individual to produce a noisy fitness score. For the noisy experiments, tournament selection is based solely on the noisy fitness values of the individuals.

## 5.2 Predictive Model for the Onemax Domain

This section adapts Equation 5 to determine the convergence rate of the percentage of correct alleles over time for the onemax domain. From equation 5 the fitness increase between two generation is given by:

$$\begin{aligned}\overline{f(t+1)} - \overline{f(t)} &= \mu_{F,t+1} - \mu_{F,t}, \\ &= \frac{\sigma_{F,t}^2}{\sqrt{\sigma_{F,t}^2 + \sigma_N^2}} \mu_{s:s}.\end{aligned}$$

For the onemax domain, $\mu_{F,t} = l\ p(t)$ and $\sigma_{F,t}^2 = l\ p(t)(1 - p(t))$. Thus

$$\begin{aligned}\overline{p(t+1)} - \overline{p(t)} &= \frac{1}{l}(\overline{f(t+1)} - \overline{f(t)}), \\ &= \frac{\mu_{s:s}}{l} \frac{\sigma_{F,t}^2}{\sqrt{\sigma_{F,t}^2 + \sigma_N^2}}, \\ &= \mu_{s:s} \frac{p(t)(1 - p(t))}{\sqrt{l\ p(t)(1 - p(t)) + \sigma_N^2}}.\end{aligned}$$

Approximating the above difference equation with a differential equation yields

$$\frac{dp}{dt} = \mu_{s:s} \frac{p(t)(1 - p(t))}{\sqrt{l\ p(t)(1 - p(t)) + \sigma_N^2}}. \tag{6}$$

Although equation 6 is integrable, it does not reduce to convenient form in the general case; however, it can be easily solved numerically for $p(t)$, and for the noiseless case ($\sigma_N^2 = 0$) $p(t)$ can be determined exactly. Subsection 5.3 will deal with solving equation 6 for $t(p)$. Given the initial percentage of correct alleles is $p(0) = 0.5$, equation 6 can be solved exactly for $p(t)$ in the noiseless case to yield:

$$p(t) = 0.5(1 + \sin(\frac{\mu_{s:s}\ t}{\sqrt{l}})). \tag{7}$$

Equations 6 and 7 together make up the predictive model for the onemax domain. Equation 6 is numerically solved to predict $p(t)$ for noisy domains, while equation 7 is directly used to obtain $p(t)$ for noiseless domains. In both equations, $\mu_{s:s}$ determines the selection pressure for a tournament of size $s$. For noisy domains, the term $\frac{\sigma_{F,t}^2}{\sqrt{\sigma_{F,t}^2 + \sigma_N^2}}$ causes the predicted convergence rate to decrease as the noise is increased. In the next section we assess the accuracy of these equations for a variety of tournament sizes and noise levels.

## 5.3 Convergence Time for the Onemax Domain

While equation 6 is not directly solvable for $p(t)$, it can be solved for $t$ as a function of $p$:

$$t(p) = \frac{1}{\mu_{s:s}} \left[ \sqrt{l}\ \arctan\left( \frac{\sqrt{l}(2p - 1)}{2\sqrt{\sigma_N^2 + l\ p(1 - p)}} \right) + \sigma_N \log\left( \frac{p}{p - 1}\ \frac{-l - 2\sigma_N^2 + l\ p - 2\sigma_N\sqrt{\sigma_N^2 + l\ p(1 - p)}}{2\sigma_N^2 + l\ p + 2\sigma_N\sqrt{\sigma_N^2 + l\ p(1 - p)}} \right) + c \right]. \tag{8}$$

For binary alleles, at time $t = 0$ we can assume that half of the alleles are initially correct $p = 0.5$. Using this to solve for $c$ in equation 8 gives that $c = 0$. For the case where $p = 1$ (convergence), $s = 2$, and $\sigma_N = 0$, equation 8 reduces to $t(1.0) = \sqrt{\pi l}\ \frac{\pi}{2}$, which agrees with convergence time found in Thierens and Goldberg (1994) for binary tournament selection. Of course, equation 8 is more general than the convergence equation in Thierens and Goldberg (1994), as it can handle tournaments of different sizes and noise levels.

We are particularly interested in the time $t_c$ it takes for all alleles to converge ($p = 1$). For the deterministic case, equation 8 reduces to

$$t_c = \frac{\pi\sqrt{l}}{2\mu_{s:s}}. \tag{9}$$

A useful approximation of the convergence time for the noisy cases is

$$t_c = \frac{1}{\mu_{s:s}}\left[\sqrt{l}\ \arctan\left(\frac{\sqrt{l}}{2\sigma_N}\right) + \sigma_N \log\left(\frac{(l-1)4\sigma_N^2}{l+4\sigma_N^2}\right)\right]. \tag{10}$$

This approximation is obtained by setting $p = 1$ in equation 8, except for the $\frac{p}{p-1}$ fraction in the log term. For the $\frac{p}{p-1}$ term, we relax the convergence criterion by setting $p = \frac{l-1}{l}$, indicating that $100(\frac{l-1}{l})$ percent of the bits in the population have correctly converged. Setting $p = \frac{l-1}{l}$ in the $\frac{p}{p-1}$ term yields $(1-l)$. Equation 10 is used to develop approximations for domains characterized by small, medium, and large amounts of noise.

For domains characterized by small levels of noise ($\sigma_N \approx 0$), equation 10 can be approximated by

$$t_c = \frac{1}{\mu_{s:s}}\left[\sqrt{l}\ \arctan\left(\frac{\sqrt{l}}{2\sigma_N}\right) + 2\sigma_N \log(2\sigma_N)\right], \tag{11}$$

as the log term is insignificant for very small levels of noise.

A medium noise level is defined as having the fitness function noise variance $\sigma_N$ approximately equal the initial population fitness variance $\sigma_f$ level, which for the onemax domain is at most $\sqrt{l/4} = \frac{\sqrt{l}}{2}$. Approximating equation 10 using $\sigma_N \approx \frac{\sqrt{l}}{2}$ yields

$$t_c = \frac{1}{\mu_{s:s}}\left[\sqrt{l}\ \arctan\left(\frac{\sqrt{l}}{2\sigma_N}\right) + 2\sigma_N \log(\sqrt{2}\sigma_N)\right]. \tag{12}$$

For large amounts of noise ($\sigma_N \approx \infty$), equation 10 can be approximated by

$$t_c = \frac{1}{\mu_{s:s}}\left[\frac{l}{2\sigma_N} + \sigma_N \log(l-1)\right], \tag{13}$$

as for small angles, $\arctan\theta \approx \theta$.

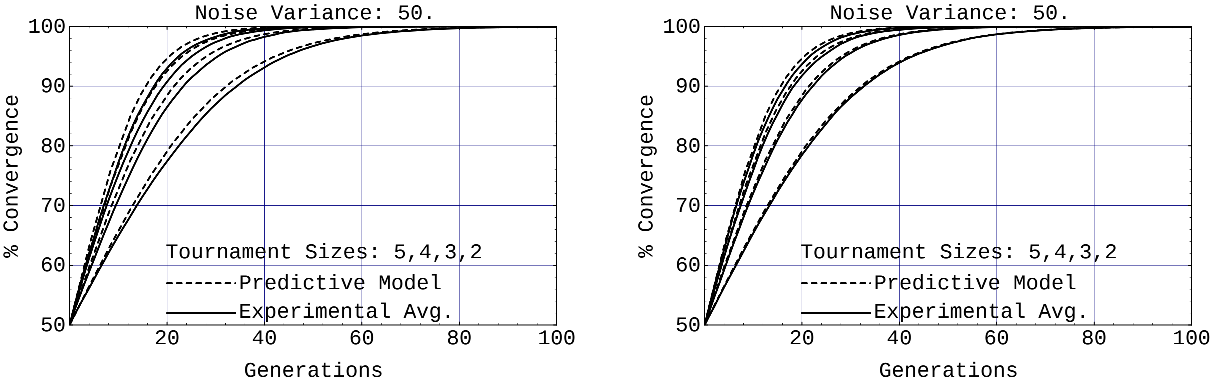
The approximations equations for convergence with small, medium, and large amounts of noise can be used to quickly estimate the convergence time for a GA. These are useful for the GA designer trying to gauge the delaying effects of noise on population convergence.

## 5.4 Experimental Results

In this section we assess the accuracy of our predictive model. We compare the predicted performance versus the actual performance obtained from GA runs for varying noise levels and tournament sizes to validate our predictive model.

To assess the accuracy of the predictive model, GA runs were made at five different noise variance levels $\sigma_n^2 = \{0, \frac{1}{2}\sigma_f^2, \sigma_f^2, 2\sigma_f^2, 4\sigma_f^2\}$. At each noise variance level, GA runs were made with tournaments sizes of $s = \{2, 3, 4, 5\}$. For each combination of noise variance and tournament size, 10 GA runs were made, and the results were averaged. The experimental results were then compared to the results predicted from Equations 6 (noisy) and 7 (deterministic).

A sample plot is shown in figure 1a. The noise variance, given on the top line, is 50. The dashed lines in the plot represent the predicted performance obtained using Equations 6 and 7 for tournament sizes $s = \{2, 3, 4, 5\}$. The solid lines display the GA performance, averaged over 10 runs, for the given noise variance and tournament sizes. The dashed lines, from left to right, correspond to the predicted performance with tournament sizes of five (highest selection pressure), four, three, and two (lowest selection pressure). Similarly, the solid lines correspond to the performance with tournament sizes, from left to right, of five (highest selection pressure), four, three, and two (lowest selection pressure).



(a) One Crossover

(b) Two Crossovers

Figure 1: Effects of Multiple Crossovers

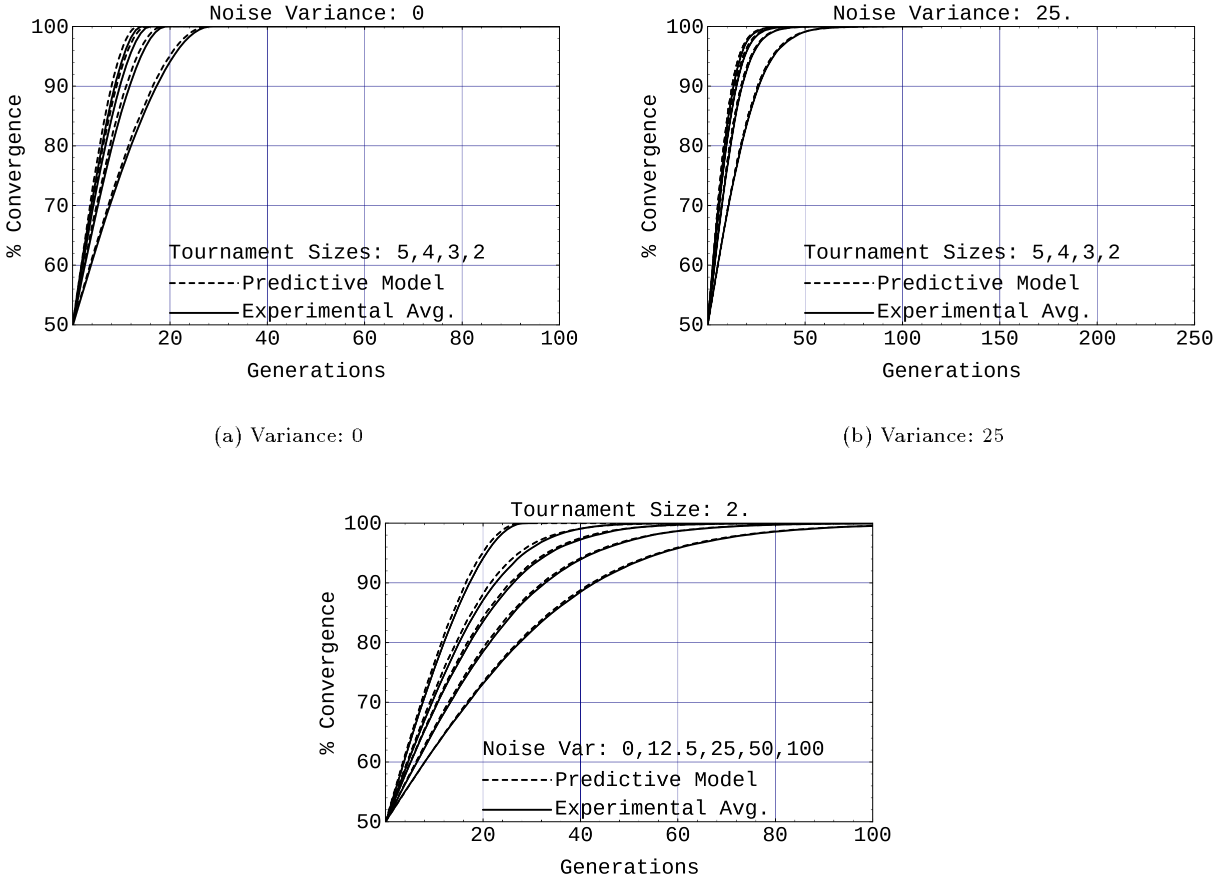
Figure 1 compares the effects of performing one crossover versus two when the noise variance is equal to the fitness variance. While the predictive model slightly overestimates the performance of GAs using one crossover (figure 1a), it accurately estimates the performance of the experiments using two crossovers (figure 1b). This is a result of crossover decreasing the correlation between alleles (Thierens and Goldberg, 1994), and the tendency of crossover to 'normalize' the population fitness distributions, making our model assumption of a normal population distribution more accurate. As done in Thierens and Goldberg (1994), we perform two crossovers per generation in our experiments, so after the usual procedure of tournament selection and recombination, we randomly shuffle the population and again recombine the population using crossover. This has the beneficial effects of reducing the correlation between alleles (Thierens and Goldberg, 1994), and 'normalizing' the population fitness distribution. However, for domains characterized by high allele interaction, this could reduce the overall performance of the GA, as there would be a higher disruption rate for the evolving building blocks.

Figure 2 summarizes our experimental results. Figure 2a plots the deterministic case, where the noise variance is zero, for a variety of tournament sizes. Figure 2b plots the experiments where the noise variance is 25, for a variety of tournament sizes. Figure 2c takes a different view, in that it plots the results of using a fixed tournament size ($s = 2$) for a variety of noise levels. These figures demonstrate that our model is very accurate for predicting GA performance in the onemax domain for a wide range of tournament sizes and noise levels.

## 5.5 Discussion of Results

This subsection discusses the general accuracy of the model, and how selection pressure affects the accuracy of the model. The accuracy of the approximation convergence equations is also discussed.

The model proved to be very accurate over a wide range of noise levels and tournament sizes. On many experiments, the predicted and experimental results were practically identical. However, the model is marginally less accurate in domains characterized by high selection pressure. This is primarily a result of the high selection pressure causing the tournament selection process to generate a non-normal (skewed) distribution, which violates the model assumption of a normally distributed population. For our experiments, high selection pressure was caused by high tournament sizes ($s = 5$). Interestingly, higher levels of noise actually reduces the tournament selection pressure, making the model more accurate. For our experiments, the highest selection pressure was for $s = 5$ and $\sigma_N = 0$. The results of this experiment are shown in figure 2a, in the upper left



(a) Variance: 0

(b) Variance: 25

(c) Tournament Size: 2

Figure 2: Onemax Experiments for Various Tournament Sizes and Noise Levels

two lines (predicted and experimental results) of the plot. This demonstrates that even with high selection pressure, the model is still fairly accurate.

Crossover has a 'normalizing' effect on the population fitness. When the selection pressure is high, the mating pool selected is non-normal (skewed). Performing multiple crossovers per generation has a 'normalizing' effect on the fitness distribution of the new offspring generated through crossover from the mating pool, which in turn makes the model more accurate (see figure 1). The experiments performed in this paper used two crossovers per generation so as facilitate comparison of results with those obtained in Thierens and Goldberg (1994). For very high selection pressures, the model accuracy can be increased by simply increasing the number of crossovers performed per generation. However, this increased accuracy does not come free, for performing multiple crossovers per generation in domains characterized by high allele interaction will retard building block growth. This slows the GA convergence rate, and would thus make the model less accurate. For the onemax domain, which has no allele interaction, multiple crossovers only increases the model accuracy.

These experiments also verified the accuracy of our approximation equations for the convergence time. Table 2 presents the average convergence time of the experiments for a variety of noise levels when the tournament size is two, as well as the convergence times predicted by the exact models (equation 8 and 9), and the small, medium and large noise approximations (equations 11, 12, and 13). For the GA experiments, convergence was defined as the first generation in which the experimental average of the average population fitness was over 99% converged. The exact model for the noisy case (equation 8) also used $p = .99$ convergence for the noisy cases, as it evaluates to infinity if $p = 1.0$, while the deterministic model (equation 9) was used for the noiseless case. The approximation equations all estimate the time until absolute convergence $p = 1.0$.

The table shows that the exact convergence equations (equations 8 and 9) predict the experimental results quite well. The small approximation equation turns out to be fairly accurate at $\sigma_N = \frac{\sqrt{l}}{2} = 12.5$, but it was designed for smaller amounts of noise. For lower noise levels, it should be more accurate than the medium approximation. The medium approximation, being designed around a noise level of $\sigma_N = \frac{\sqrt{l}}{2} = 25$ is fairly accurate for all noise levels up to 100. At the high noise level of 100, the large approximation model is the most accurate approximation. These results indicate that the approximation equations do very well as a quick estimate of the convergence time for GAs utilizing tournament selection.

| Noise | Exper. | Exact $t$ | Approximate $t_c$ | | |
|---|---|---|---|---|---|
| $\sigma_N^2$ | $t_c$ | $(p = .99)$ | Small | Med. | Large |
| 0 | 28.0 | 28.8 | 28.8 | 28.8 | NA |
| 12.5 | 40.0 | 39.9 | 42.8 | 38.3 | 55.6 |
| 25.0 | 49.0 | 49.9 | 56.5 | 50.2 | 60.4 |
| 50.0 | 64.0 | 65.4 | 79.9 | 70.9 | 72.4 |
| 100.0 | 87.0 | 88.4 | 118.2 | 105.5 | 93.3 |

Table 2: Convergence Times for $s = 2$.

# 6 Future Research

This section describes a number of important areas for future research:

- Applying newfound understanding of noise for other selection schemes
- Modeling other GA operators by utilizing the noise component of the model to account for their effect on convergence
- Testing the model in more complex domains
- Using the model to answer basic performance questions for a GA
- Applying the model to help tune GA configuration parameters
- Determining appropriate sample sizes for fitness function employing sampling, so as to maximize performance within a given environment

This research is important in that it has furthered our understanding of noise and its delaying effect on convergence. The model has proved accurate at predicting the convergence rate of a GA utilizing tournament selection for a variety of noise levels and tournament sizes in the onemax domain. The approach taken in section 4, where the deterministic tournament selection model is extended to handle noise, shows promise for adapting other deterministic selection models to handle noise.

Within this paper, the noise component was considered to be produced by noise present in the fitness functions. However, there is no reason why the noise input for the developed model can not include noise introduced from other GA components. The noise input indirectly indicates the degradation of the mating pool fitness as compared to the mating pool selected with no noise present. The noise input can thus be used to account for other degradations of mating pool fitness from other GA mechanisms besides noisy fitness functions. Other GA mechanisms that also introduce noise could be included in the noise component, such as different mutation mechanisms and mutation rates, and tournament selection used with or without replacement. This would increase the model's predictive accuracy for a wider range of GA configurations.

As discussed in Subsection 5.5, the use of order statistics has proved very accurate in predicting tournament selection pressure for the onemax domain. We would like to extend our model to handle other domains that have different characteristics than the onemax domain. The onemax domain is characterized by equal allele weighting (each bit is equally important), and no allele interaction (no higher order building blocks). We would like to extend our model to handle more complex domains, including "domino-like" domains, where the alleles have unequal weighting, and domains characterized by high allele correlation. However, the current model is still of use for these types of domain, in that it can provide an lower bound on the convergence rate. Extending our model to directly handle more complex domains will increase the accuracy of the predicted convergence rates.

One of our model's strengths is that it predicts the distribution of the population fitness over successive generations. Order statistics can be used not only to predict both the increase in fitness between generations, but also to predict the population fitness variance in the next generation. As the population fitness mean and variance can be accurately modeled over time using order statistics, our model can be used to answer questions relating to population fitness distribution. The model could thus be applied to answer PAC-like performance questions like "What is the probability that a solution of quality $Y$ will be present at generation $X$," or "At what generation are we $Z$ percent confident that a solution of at least quality $Y$ will be present in the population." Answers to these questions could be used to determine how long a GA should run before a solution of acceptable quality is likely to be be produced. This would allow for a GA designer to set GA stopping criteria that achieves a desired solution quality.

This model should be very helpful in determining appropriate settings for many GA parameters. While GAs with generic parameter settings are good at finding good solutions in a reasonable amount of time, their performance can be improved by tuning the parameter settings for a specific domain. As discussed above, our model can be used to determining appropriate stopping criteria for the GA. It can also be used to design a GA that has a desired convergence rate for a given domain, by applying the model to determine the appropriate tournament size for achieving a specified convergence rate. It could even be used to design "custom" tournament that achieves a given selection pressure (i.e., a tournament where the best 2 out of 5 competitors are selected for the mating pool).

In some domains, a GA designer is faced with a range of possible fitness functions, all with different noise and run-time performance characteristics. The model can be applied to help select a fitness function that achieves an acceptable solution in an acceptable amount of runtime for a given domain. Some fitness functions are based on sampling, with the sampling fitness function's noise and run-time performance characteristics directly controlled by the sample size. Our model, in conjunction with sampling theory being used to predict the noise from a given sample size, should be able to determine the appropriate sample size needed for the GA to achieve a given convergence rate.

# 7 Conclusions

Tournament selection is an important selection mechanism for GAs. It is simple to code, easy to implement on non-parallel or parallel architectures, robust in the presence of noise, and has adjustable selection pressure. This paper has developed a model that works under a wide range of noise levels and tournament sizes to accurately predict the convergence rate of a GA utilizing tournament selection. The model has been verified using the onemax domain, and shown to be accurate for predicting the convergence rate under a wide range of noise levels and tournament sizes.

The paper has discussed a number of immediate practical uses of the model. It can be used to correctly set various GA control parameters for a given domain, including tournament sizes and mutation rates. The model can determine appropriate stopping criteria for achieving a desired solution quality. The model can be used to answer basic performance questions, such as "What is the probability that a solution of quality $Y$ will be present at generation $X$?" In addition, the model can also be used to determine appropriate sample sizes for the class of fitness functions that employ sampling so as to maximize GA performance.

This research has several long term ramifications. Through the study of one selection scheme, tournament selection, this paper has opened the door for understanding noise and its delaying effect for all selection schemes. The approach used in this study may also be applied to predict the convergence delays resulting from inherently noisy operators such as selection, crossover, and mutation.

# 8 Acknowledgments

This work was supported under NASA Grant No. NGT 9-4. This effort was also sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant numbers F4960-94-1-0103 and

F49620-95-1-0338. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

## References

Balakrishnan, N. & Cohen, A. C. (1991). *Order statistics and inference: estimation methods* (pp. 51–54). Boston: Harcourt Brace Jovanovich.

David, H. A. (1981). *Order statistics* (2nd ed.). New York: John Wiley & Sons, Inc.

Goldberg, D. E. & Deb, K. (1991). A comparative analysis of selection schemes used in genetic algorithms. *Foundations of Genetic Algorithms*, *1*, 69–93. (Also TCGA Report 90007).

Goldberg, D. E., Deb, K., & Clark, J. H. (1992). Genetic algorithms, noise, and the sizing of populations. *Complex Systems*, *6*, 333–362.

Goldberg, D. E., Deb, K., & Thierens, D. (1993). Toward a better understanding of mixing in genetic algorithms. *Journal of the Society of Instrument and Control Engineers*, *32*(1), 10–16.

Harter, H. L. (1961). Expected values of normal order statistics. *Biometrika*, *48*, 151–165.

Mühlenbein, H. & Schlierkamp-Voosen, D. (1993). Predictive models for the breeder genetic algorithm: I. Continuous parameter optimization. *Evolutionary Computation*, *1*(1), 25–49.

Teichroew, D. (1956). Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution. *Annals of Mathematical Statistics*, *27*, 410–426.

Thierens, D. & Goldberg, D. (1994). Convergence models of genetic algorithm selection schemes. In Davidor, Y., Schwefel, H.-P., & Männer, R. (Eds.), *Parallel Problem Solving from Nature- PPSN III* (pp. 119–129). Berlin: Springer-Verlag.

Tietjen, G. L., Kahaner, D. K., & Beckman, R. J. (1977). Variances and covariances of the normal order statistics for sample sizes 2 to 50. *Selected Tables in Mathematical Statistics*, *5*, 1–73.