

Satisfiability and Evolution

Adi Livnat

*Department of Biological Sciences
Virginia Tech.
adi@vt.edu*

Christos Papadimitriou

*Computer Science Division
University of California at Berkeley
christos@berkeley.edu*

Aviad Rubinfeld

*Computer Science Division
University of California at Berkeley
aviad@eecs.berkeley.edu*

Gregory Valiant

*Computer Science Department
Stanford University
valiant@stanford.edu*

Andrew Wan

*Simons Institute
University of California at Berkeley
atw12@columbia.edu*

Abstract—We show that, if truth assignments on n variables reproduce through recombination so that satisfaction of a particular Boolean function confers a small evolutionary advantage, then a polynomially large population over polynomially many generations (polynomial in n and the inverse of the initial satisfaction probability) will end up almost surely consisting exclusively of satisfying truth assignments. We argue that this theorem sheds light on the problem of the evolution of complex adaptations.

Keywords—evolution; algorithms; Boolean functions

I. INTRODUCTION

The TCS community has a long history of applying its perspectives and tools to better understand the processes around us; from learning to multi-agent systems, game theory and mechanism design. By and large, the efforts to understand these areas from a rigorous and algorithmic perspective have been very successful, leading to both rich theories and practical contributions. Evolution is, perhaps, one of the most blatantly algorithmic processes, yet our computational understanding of it is still in its infancy (see [16] for a pioneering study), and we currently lack a computational theory explaining its apparent success. Algorithmically, how plausible are the origins of evolution and the emergence of self-replication? Is evolution surprisingly efficient or surprisingly inefficient? What are the necessary criteria for evolution-like algorithms to yield rich, interesting, and diverse ecosystems? Why is recombination (i.e., sexual reproduction) more successful than asexual reproduction? Given the reshuffling of genomes that occurs through recombination, how can complex traits that depend on many different genes arise and spread in a population?

In this work, we begin to tackle this last question of why complex traits that may depend on many different

genes are able to *efficiently* arise in polynomial populations with recombination. In the standard view of evolution, a variant of a particular gene is more likely to spread across a population if it makes its own contribution to the overall fitness, independent of the contributions of variants of other genes. How can complex, multi-gene traits spread in a population? This may seem to be especially problematic for multi-gene traits whose contribution to fitness does not decompose into small additive components associated with each gene variant —traits with the property that even if one gene variant is different from the one that is needed for the right combination, there is no benefit, or even a net negative benefit. Here, we provide one rigorous argument for how such complex traits can efficiently spread throughout a population. While we consider this question in a model that makes considerable (but justifiable) simplifications, this model makes a theoretically rigorous contribution to the fundamental problem of how evolution can produce complexity.

Motivating example: Waddington’s experiment.: In 1953 the great experimentalist Conrad Waddington exposed the pupae of a population of *Drosophila melanogaster* to a heat shock, and noticed that in some of the adults that developed, the appearance of the wings had changed (they lacked a complete posterior crossvein) [17]. He then maintained a population of flies where only those with altered wings were allowed to reproduce. By repeating the procedure of heat shock and selection over the generations, the percentage of flies with altered wings increased over time to values close to one. Even more interestingly, beginning at generation fourteen, some flies exhibited the new trait even without having been treated with heat shock.

At first sight, this surprising phenomenon — known as genetic assimilation — recalls Lamarck’s now discredited belief that acquired traits can be inherited. However, Boolean functions provide a purely genetic explanation, which extends the idea originally offered informally by Stern [15] (see also [2], [5]): Suppose that the phenotype “altered wings” is a Boolean function of n genes x_1, \dots, x_n with

A.L. would like to acknowledge start-up funds from Virginia Tech. C.P. and A.R. are supported by NSF grants CCF0964033 and CCF1408635, and by Templeton Foundation grant 3966. A.W. conducted part of this research while at Harvard University and IIS, Tsinghua University and supported by NSF grant CCF-0964401.

two alleles (variants) each, thought as $\{-1, 1\}$ variables, and of another $\{-1, 1\}$ variable h (standing for “high temperature”).¹

$$x_1 + x_2 + \dots + x_n + \frac{(1+h)}{2} \cdot k \geq n,$$

for some integer k (think of $n \approx 10$, and $k \approx n/3$).

To see how the percentage of flies with altered wings increases in the population over time, we track the allele frequencies from generation to generation. Let μ_i^t be the average value of x_i in the population at time t , and assume the genotype frequencies at time t are distributed according to a distribution μ^t (the reason for denoting the distribution this way will become clear). If mating occurs at random with free recombination² then, in expectation, the average value of each x_i in the next generation is given by:

$$\mu_i^{t+1} = \frac{\mathbf{E}_{\mu^t}[f(x) \cdot x_i]}{\mathbf{E}_{\mu^t}[f]}, \quad (1)$$

where $f(x) = 1$ exactly when a fly having genotype x will develop altered wings (i.e., the above inequality is satisfied) and $f(x) = 0$ otherwise. We then assume that the next generation will be distributed according to a product distribution μ^{t+1} , where each x_i has expectation μ_i^{t+1} . By approximating the genotype frequencies of the population for each generation in this way, it can be shown by calculation that a trait with this genotypic specification (a) is very rare in the population under normal temperature $h = 1$; (b) it becomes much more common under high temperature $h = 1$; (c) jumps to just above 50% after the first breeding under $h = 1$; (d) after successive breedings with $h = 1$ it is nearly fixed; and (e) if after this h becomes -1 , the trait is still quite common.

Note: Our interpretation of Waddington’s experiment is a simplification. First, we consider only the distribution of genotypes in each generation to determine the distribution of the next; instead, we could first take a finite sample according to the present distribution and use that sample to calculate the distribution of the next generation. Such an approximation can only become exact when the population size is infinite, but it is a standard and useful one in population genetics (and we shall eventually consider finite populations for our main result). We also assumed that each individual of the new generation is produced by sampling each gene independently of the other genes, and with probability equal to the frequencies of the two alleles of this gene in the parent population (the adults of the previous generation with altered wings). This assumption turns out to be justified in the settings that we will consider, as will be discussed in the following section.

¹Here we assume for simplicity haploid organisms, that is, each individual has only one copy of each gene.

²See the next section for any unfamiliar terms and concepts from evolution.

Populations of truth assignments

This way of looking at Waddington’s experiment brings about a very natural question: Is this amplification of satisfying truth assignments (outcomes (c) and (d) of experiment described above) a property of threshold functions, or is it more general? Does it hold for all monotone functions, for example? *For all Boolean functions?*

Consider any satisfiable Boolean function $f : \{-1, 1\}^n \rightarrow \{0, 1\}$ of n binary genes (in the absence of the environmental variable h which was crucial in Waddington’s phenomenon). What if genotypes satisfying this Boolean function had a slight advantage under natural selection? (In Waddington’s experiment, they had an absolute advantage because of the experimental design.) For example, imagine that genotypes satisfying f survive to adulthood more than the others, in expectation, by a factor of $(1 + \epsilon)$, for some small $\epsilon > 0$. Would this trait (that is, satisfaction of the Boolean function f) be eventually fixed in the population? And, if so, could this be a subtle mechanism for introducing complex adaptation in a population?

To reflect our assumption that satisfaction of f confers only an ϵ -advantage, we may take a function $f : \{-1, 1\}^n \rightarrow \{1, 1 + \epsilon\}$, where we regard the value $1 + \epsilon$ as “satisfied”, and the value 1 as “unsatisfied”. We track the allele frequencies from generation to generation as in Waddington’s experiment: Equation (1) gives us the average value μ_i^{t+1} of each x_i in the next generation, and we describe the next generation by the product distribution μ^{t+1} .

Suppose that we continue this process, starting from distribution μ^0 , and defining $\{\mu_i^1\}, \{\mu_i^2\}, \dots, \{\mu_i^t\} \dots$ as above. Consider the average fitness of the population at time t , defined as $\mu^t(f) = \Pr_{\mu^t}[f(x) = 1 + \epsilon]$. The question is, when does $\mu^t(f)$ approach one? Our first result states that, for monotone functions, it does after $O\left(\frac{n}{\epsilon\mu^0(f)}\right)$ steps:

Theorem 1. *If f is monotone, then $\mu^t(f) \geq 1 - \frac{n(1+\epsilon)}{\epsilon t \mu^0(f)}$.*

Note: This nontrivial result also serves to illustrate one point: The work is *not* about satisfiability heuristics (monotone functions are not an impressive benchmark in this regard...). Heuristics are about finding good *individuals* in a population. In contrast, evolution is about creating good *populations*. This is our focus here.

Our ambition is to prove the same result for all Boolean functions. Immediately we see that this is impossible if we insist on an infinite population: Consider the function $f = x_1 \oplus x_2$: starting with the uniform distribution at time $t = 0$, the above dynamics would leave the distribution unchanged, for all time, and hence $\mu^t(f) = 1/2$ for all t . The parity function is not the only Boolean function with this property: for example the function “ $\sum_{i=1}^n x_i = k$ ”, if started at $\mu_i = \frac{k}{n}$, will stay at that spot forever, and will

always have $\mu^t(f) = O(1/\sqrt{k})$. However, experimentation shows that these “spurious fixpoints” are not absorbing, and evolution pulls the distribution away from them and towards satisfaction. That is, this disappointing phenomenon is an artifact of the infinite population simplification. Indeed, random genetic drift due to sampling effects has been considered to be a significant component of evolution at the molecular level (it is possible for an allele to become fixed in the population even in the absence of selection). Thus, we need to make the model more realistic.

We adopt a model, consisting of the following process: At each generation t we create a large population of N individuals (we call this the “sampling” step) by sampling N times from the product distribution μ^t to obtain $y^{(1)}, \dots, y^{(N)}$ (N is assumed to remain constant from one generation to the next, which is a standard assumption in population genetics [6]). The empirical allele frequencies of the sample are given by a vector ν^t , where for each i we have:

$$\nu_i^t = \frac{1}{N} \sum_{j=1}^N y_i^{(j)}.$$

We write $\nu^t \sim B(\mu^t)$ to denote a draw from this distribution and use ν^t to denote the implied Bernoulli distribution.

We then enforce the assumed selection advantage of satisfaction to obtain the “in-expectation” frequencies of the subsequent generation:

$$\mu_i^{t+1} = \frac{\mathbf{E}_{\nu^t}[f(x) \cdot x_i]}{\mathbf{E}_{\nu^t}[f]}. \quad (2)$$

We show that when selection is weak, any satisfiable Boolean function will almost surely be always satisfied after polynomially many time steps.

Theorem 2. (informal statement) *For any satisfiable Boolean function f of n variables and any sufficiently small $\epsilon > 0$, after T generations of N individuals $\mu^T(f) = 1$ with probability arbitrarily close to one, where T and N are polynomial functions of n , $\frac{1}{\epsilon}$, and $\frac{1}{\mu^0(f)}$.*

The proof of Theorem 2 shows why the population does not become stuck at the previously discussed “spurious fixpoints;” sampling effects ensure movement over sufficiently many generations, and selection ensures movement is made towards satisfaction.

Outline of the paper

In the next section we introduce some basic concepts from population genetics, we define and justify our simplified model, and we present a result due to Nagylaki [12] implying that, if selection is weak, then one can assume that the genotype distribution is a product distribution. In Section III, we show Theorem 1 on monotone functions. Our main result is given in Section IV, and its proof outlined; the full proof is detailed in the Appendix. In Section V, we conclude with a discussion of our result, and a number of open problems.

II. EVOLUTION BACKGROUND AND OUR MODEL

The genetic makeup of an organism is its *genotype*, which specifies one *allele* (gene variant) for each genetic site, or “locus,” in the haploid case. We shall be focusing on n specific genes of interest (say, a few dozen out of the many thousands of genes of the species). At each locus, we assume that there are two alleles segregating in the population (hence the relevance of Boolean functions). Thus, a genotype will be a vector in $\{\pm 1\}^n$. We assume the species reproduces sexually (this is crucial, see the discussion in the last section). In a sexual species reproduction proceeds through *recombination*, that is, the formation of a new genotype by choosing alleles from two parental genotypes in the previous generation. To produce each generation, the individuals mate at random (we also assume no bipartition into sexes) and there is no generation overlap (that is, the new generation is produced en masse just before death of the previous one). We assume that the population size is constant at some large number N (expressed as a function of n , the number of genes of interest, which is the basic parameter). Each genotype $g \in \{\pm 1\}^n$ is assumed to have a *fitness value* equal to the expected number of offsprings this genotype will produce. We also assume that the genes *recombine freely*, that is, for any two genes i, j of an offspring, the probability that the alleles come from the same parent is exactly half (and not larger, as is the case if the two genes are linked).

These assumptions are simplifications of the standard model of population genetics used broadly in the literature, and generally trusted to preserve the essence of selection in sexual populations. The Boolean assumption is of course meant to bring into play mathematical insights from that field, but we believe that it is not restrictive (for example, allele -1 could stand for “any allele other than allele 1 ”). In this paper we shall make two more assumptions. The first additional assumption is that the fitness values of our genotypes are either 1 or $1 + \epsilon$, where $\epsilon > 0$ is very small: the organism will reproduce slightly more in expectation if an underlying Boolean function is satisfied. We discuss this restriction in Section V.

The final assumption is more problematic in general, but justified in the current context: We assume that generating an individual of the next generation is tantamount to selecting, independently, an allele for each of the n genes, with probability equal to the probability of occurrence of that allele in the parent generation. That is, we assume that the distribution of the genotypes in a generation is a *product distribution*. This situation is called in the population genetics literature *linkage equilibrium*, or the *Wright manifold* [19], [20]. In general, genotype frequencies are known to be correlated, and this correlation — the distance from the product distribution — is called *linkage disequilibrium* [7] and is of importance and interest in the study of evolution. However, in the absence of selection, a standard argument

shows that the distribution of a population quickly reaches linkage equilibrium (arguments exist both for finite and infinite populations). Our previous assumption places our experiment in a regime known as *weak selection*. Weak selection means that the fitness values are in a small interval $[1 - \epsilon, 1 + \epsilon]$, where ϵ is called the *selection strength*. An elegant and powerful result due to Thomas Nagylaki [12] states that, under weak selection, evolution proceeds to a point very close to linkage equilibrium. In particular, assume that a population evolves as we described above in a regime of weak selection of strength ϵ , and let m be the total number of alleles (this is $2n$ in our case; actually, Nagylaki's Theorem also holds under diploid and partial recombination). By linkage disequilibrium we mean formally the L_∞ distance between the genotype distribution and the product distribution:

Theorem 3. (Nagylaki's Theorem, see [12]) *Under weak selection, and after $O(\log m \cdot \log 1/\epsilon)$ generations, linkage disequilibrium is $O(\epsilon)$.*

In our setting ϵ is minuscule, so Nagylaki's Theorem motivates our assumption that populations are formed "by independent sampling of the genetic soup." We strongly believe that our theorem is true for large ϵ as well, but this remains open, as discussed in the last section.

III. WEAK SELECTION ON MONOTONE FUNCTIONS

In this section we give a self-contained proof of Theorem 1. The proof is simple, once a connection is made to discrete Fourier analysis. In what follows, we assume familiarity with Fourier analysis over the Boolean cube for product distributions. We briefly review some basic facts and describe the notation used in our proofs.

For $\mu = (\mu_1, \dots, \mu_n) \in [-1, 1]^n$ and a function $f : \{-1, 1\}_\mu^n \rightarrow \mathbb{R}$, where $\{-1, 1\}_\mu^n$ denotes the Boolean cube endowed with the product distribution given by $\mu_i = \mathbf{E}[x_i]$, we consider the μ -biased Fourier decomposition of f . Let $\sigma_i^2 = 1 - \mu_i^2$ be the variance of each bit. We denote the μ -biased Fourier coefficients by $\hat{f}(S; \mu) = \mathbf{E}_\mu[f \cdot \phi_S^\mu]$, where $\phi_S^\mu = \prod_{i \in S} \frac{x_i - \mu_i}{\sigma_i}$. Let $D_i^{(\mu)} f = \frac{\sigma_i}{2}(f_{i=1} - f_{i=-1})$ be the difference operator for Boolean functions over $\{-1, 1\}_\mu^n$. We have that

$$D_i^{(\mu)} f = \sum_{S \ni i} \hat{f}(S; \mu) \phi_{S \setminus \{i\}}^\mu, \quad (3)$$

and in particular, $\mathbf{E}_\mu[D_i^{(\mu)} f] = \hat{f}(i; \mu)$, which we will use repeatedly throughout our proofs.

Our first step will be to observe that the change in allele frequencies from one generation to the next may be expressed in terms of f 's linear Fourier coefficients. Let μ be the vector which specifies the allele frequency of the population at time t . Then, letting μ' be the allele frequency vector at time $t + 1$ and using the selection specified by

Equation (1), we have that

$$\mu'_i - \mu_i = \sigma_i \frac{\hat{f}(i; \mu)}{\mathbf{E}_\mu[f]}. \quad (4)$$

This follows immediately from the definitions:

$$\begin{aligned} \sigma_i \cdot \hat{f}(i; \mu) &= \sigma_i \cdot \mathbf{E}_\mu[f \cdot \phi_i^\mu] \\ &= \mathbf{E}_\mu[f \cdot x_i] - \mathbf{E}_\mu[f] \cdot \mu_i \\ &= \mathbf{E}_\mu[f] \cdot \mu'_i - \mathbf{E}_\mu[f] \cdot \mu_i. \end{aligned}$$

Our proof uses the following well-known facts, which are easily derived from the basic notions (see Chapter 2.3, [13]). First, we have that the influences of a monotone function are given by its linear coefficients. (For a function $f : \{-1, 1\}_\mu^n \rightarrow \mathbb{R}$, we denote its influence in direction i by $\sum_{S \ni i} \hat{f}(S; \mu)^2$.) Next, the inequality of Poincaré lower bounds the total influence of a function by its variance. The versions below have been scaled to our setting and can be obtained by applying the original facts to a Boolean function $g : \{-1, 1\}^n \rightarrow \{-1, 1\}$ and setting $f(x) = 1 + \frac{\epsilon}{2}(1 + g)$.

Proposition 4. *Let $f : \{-1, 1\}_\mu^n \rightarrow \{1, 1 + \epsilon\}$ be monotone. Then for all $i \in [n]$:*

$$\sum_{S \ni i} \hat{f}(S; \mu)^2 = \frac{\epsilon \sigma_i}{2} \cdot \hat{f}(i; \mu).$$

Proposition 5. *Let $f : \{-1, 1\}_\mu^n \rightarrow \{1, 1 + \epsilon\}$ and $\text{Var}[f] = \mathbf{E}_\mu[f^2] - \mathbf{E}_\mu[f]^2$. Then*

$$\sum_{i \in [n]} \sum_{S \ni i} \hat{f}(S; \mu)^2 = \sum_{S \subseteq [n]} |S| \hat{f}(S; \mu)^2 \geq \text{Var}[f].$$

Equation (4) tells us that the bias of each bit i increases according to the corresponding coefficient $\hat{f}(i)$. Proposition 4 tells us that for monotone f , the linear coefficients correspond to the influences of f . Finally, the inequality of Poincaré tells us that the linear coefficients must be large.

We may now prove Theorem 1.

Theorem 1. *Let $f : \{-1, 1\}_\mu^n \rightarrow \{1, 1 + \epsilon\}$ be monotone. Then $\mu^t(f) \geq 1 - \frac{n(1+\epsilon)}{\epsilon^t \mu^0(f)}$.*

Proof: Combining Equation (4) with Propositions 4 and 5 tells us that the sum of the biases increases at each step:

$$\begin{aligned} \sum_{i \in [n]} (\mu'_i - \mu_i) &= \frac{2}{\epsilon \cdot \mathbf{E}_\mu[f]} \sum_{i \in [n]} \sum_{S \ni i} \hat{f}(S; \mu)^2 \\ &\geq \frac{2}{\epsilon \cdot \mathbf{E}_\mu[f]} \text{Var}[f] \\ &= \frac{2}{\epsilon \cdot \mathbf{E}_\mu[f]} \epsilon^2 \mu(f) (1 - \mu(f)) \end{aligned}$$

Let $\mu^t(f) = 1 - \delta$. Then for all $t' \leq t$, the sum of the biases increases at each step:

$$\begin{aligned} \sum_{i=1}^n \mu_i^{t'+1} - \sum_{i=1}^n \mu_i^{t'} &\geq \frac{2\epsilon \mu^{t'}(f)(1 - \mu^{t'}(f))}{\mathbf{E}_\mu[f]} \\ &\geq \frac{2\epsilon \mu^{t'}(f)\delta}{1 + \epsilon} \geq \frac{2\delta \epsilon \mu^0(f)}{1 + \epsilon}. \end{aligned}$$

On the other hand, we know that $-n \leq \sum_{i=1}^n \mu_i^{t'} \leq n$ for all t' , so $t \leq \frac{n(1+\epsilon)}{\delta \epsilon \mu^0(f)}$. ■

We remark that Theorem 1 (with worse parameters) can also be proven using a generalization of the Russo-Margulis lemma to product distributions, which states that the gradient of $\mathbf{E}_\mu[f]$ (as a function of μ) corresponds to the influences of f (see appendix of the full version).

IV. THE MAIN RESULT

For a function $f : \{-1, 1\}^n \rightarrow \{1, 1 + \epsilon\}$, consider the multilinear extension $\tilde{f} : [-1, 1]^n \rightarrow [1, 1 + \epsilon]$ defined by $\tilde{f}(\mu) = \mathbf{E}_{x \sim \mu}[f(x)]$. Our goal is to understand when $\tilde{f}(\mu) = 1 + \epsilon$. We start with the precise statement of the main result (compare with Theorem 2):

Theorem 6. *Let $\beta = \sqrt{\frac{\epsilon}{N(1-n\epsilon)}}$. If*

$$\tilde{f}(\mu^0) > 1 + \sqrt{2\beta \ln \frac{2}{\beta}}$$

then there is some constant C such that for any $T \geq C \cdot \frac{\epsilon n^8 \cdot N^4}{1-n\epsilon}$:

$$\Pr \left[\tilde{f}(\mu^{(T)}) = 1 + \epsilon \right] \geq 1 - 2\beta - 2/n.$$

Note that the conditions in Theorem 6 imply restrictions on the initial probability of satisfaction and the strength of selection. In particular, the selection coefficient must be in the range $1/N^{1/3} < \epsilon < 1/n$ (we discuss this restriction in the next section), and the initial probability of satisfaction must be at least $N^{-1/4}$. The full proof of the theorem is given in the appendix of the full version; in this section we sketch its salient points.

One first difficulty in the proof is this: The convergence proof gauges the improvement in average population fitness obtained during the second of the two steps per generation (the fitness step). However, the first of the two steps (the sampling step) introduces variance, and we must establish that this variance is insignificant in comparison with the increase in fitness. Our first step is to establish that the difference between the average fitness of the sample and the average fitness, squared (that is to say, the variance introduced), is bounded from above by the increase in average fitness obtained in the fitness step:

$$\begin{aligned} \mathbf{E}_{\nu \sim B}[(\tilde{f}(\nu) - \tilde{f}(\mu))^2] &\leq \\ \mathbf{E}_{\nu \sim B}[\tilde{f}(\mu') - \tilde{f}(\nu)] / [(N-1) \cdot (1-n\epsilon)]. \end{aligned} \quad (5)$$

Here we focus on one generation, so μ denotes the product distribution from which the sampling is made, ν the empirical product distribution of the sample (note that $\tilde{f}(\nu)$ is a random variable with expectation $\tilde{f}(\mu)$), and μ' the product distribution resulting from the selection (or fitness) step. Thus, μ' is the initial product distribution in the next generation.

To establish inequality (5), we first show that the right-hand side is lower bounded by the total mass of the singleton Fourier coefficients of the biased transform:

$$\tilde{f}(\mu') - \tilde{f}(\nu) \geq (1-n\epsilon) \sum_{i=1}^n \hat{f}(i; \nu)^2. \quad (6)$$

The intuition in the proof of (6) is that the fitness step is very close to an ϵ -long step of the gradient ascent of the average fitness function (this intuition is very accurate away from the boundary of the hypercube). Gradient ascent in each coordinate is captured by the corresponding singleton coefficient squared. But then there is an analytical complication of approximating the overall ascent by the sum of sequential coordinate-wise ascents; the difficulty is, of course, that the partial derivatives change after each small ascent, and the change must be bounded.

This establishes that the fitness increase in the selection step is larger than the linear Fourier mass, and hence nonnegative when ϵ is small. However, the linear Fourier mass may be zero, as is the case for the exclusive-or function under the uniform distribution (recall the discussion a few lines after Theorem 1). Here, sampling effects will ensure that progress is made in expectation. We show that, on average, the linear Fourier mass is much larger than the variance:

$$\begin{aligned} \mathbf{E}_{\nu \sim B} \left[(\tilde{f}(\mu) - \tilde{f}(\nu))^2 \right] &\leq \\ \frac{1}{N-1} \mathbf{E}_{\nu \sim B} \left[\sum_{i=1}^n \hat{f}(i; \nu)^2 \right] \end{aligned} \quad (7)$$

The rather involved proof of (7) takes place entirely within the biased Fourier domain. Now notice that (7), combined with (6), completes the proof of inequality (5).

Note that the upper bound on the variance in (5) includes in the denominator a factor of $(1-n\epsilon) \cdot N$. This immediately tells us that our technique is sharpest when the population N is large and the selection strength ϵ is small — in particular, it *must* be smaller than $\frac{1}{n}$. This latter point is a rather puzzling limitation of our result: Why does a theorem about the effectiveness of natural selection become harder to prove when selection is stronger? One intuitive explanation is that in this case selection works very much like gradient ascent, and it is well known that the convergence of gradient ascent is harder to establish when the ascent step is large, as a large step can “skip over” the stationary point sought. Is this upper limit on ϵ necessary? This is an intriguing open question discussed in the last section.

Next, we establish that the total effect of the sampling steps is small: For any $\alpha > \sqrt{2\beta \ln 2\beta^{-1}}$,

$$\Pr \left[\left| \sum_{t=1}^T \tilde{f}(\nu^t) - \tilde{f}(\mu^{t-1}) \right| \geq \alpha \right] \leq 2\beta,$$

where $\beta = \left(\frac{\epsilon}{N(1-n\epsilon)} \right)^{1/2}$.

It is not hard to see that the sum $\sum_{t=1}^T \tilde{f}(\nu^t) - \tilde{f}(\mu^{t-1})$ constitutes a martingale, albeit one with no obvious upper bound on each step. We bound the total effect of the sampling step by resorting to a rather exotic martingale inequality derived from a generalization of Bernstein’s inequality to martingales with unbounded jumps and proved in [4] (we provide a statement of the inequality in the full version).

Incidentally, notice that this is the place where it is proved, quite indirectly, that the sampling step succeeds in getting the process unstuck from spurious fixpoints such as $(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2})$ for the exclusive-or function: Since the total effect of sampling is limited, the increase in average fitness must eventually prevail.

Finally, when the process is near a vertex of the hypercube, fitness increases are too small to help finish the argument, but here we rely on the fact that the process is very likely to drift so close to a vertex that it will eventually get stuck there, completing the proof of the main result.

V. DISCUSSION

We proved a novel and highly nontrivial aspect of Boolean satisfiability: By randomly crossing assignments and favoring satisfaction slightly, one can breed a population of pure satisfying truth assignments. We argued that this rather curious property seems important in understanding one intriguing aspect of evolution: how complex traits controlled by many genes can emerge.

There are many roads of mathematical inquiry opened by this theorem. First, can the limitations/restrictions of our model be relaxed so that it better reflects the realities of life? Some of the assumptions in our model are arguably unrealistic (haploidy, fixed population size, random mating, partly in-expectation fitness calculation), but these follow widely accepted practices in population genetics needed for mathematical simplification. We also make the assumption of weak selection, but this is also a very defensible one for unlinked loci.

There are, however, a few further restrictions of our model that call for discussion:

- *Two alleles per gene.* The motivation is, of course, that this assumption ushers in the powerful analytical toolbox of Boolean functions. We have no doubt that similar results hold for more alleles, but would require a great number of technical adjustments.
- *Fitness landscape.* We assumed a very specialized fitness landscape with values 1 and $1 + \epsilon$ only. This is a natural simplification that facilitates the connection to Boolean functions, but we do not believe it is an essential one. We believe that this result can be extended to much richer landscapes with a small gap, for example to situations in which fitness values are in $[1 - \delta, 1] \cup [1 + \epsilon, 1 + \epsilon + \delta]$ for some small $\delta > 0$.

A harder question is, what happens if the fitness gap ϵ is larger? As we have mentioned, this is an analytical challenge with roots in the difficulty of the analysis of gradient descent. Of course, a constant gap would bring us outside the realm of weak selection, and render our approximation by product distribution baseless. There are two ways we can proceed: One is to prove that the exact recurrence equations of genotype frequencies yield eventual satisfaction. This seems possible but challenging.

Another avenue, which we have followed for some time, is to work with product distributions anyway. In particular, what if the fitness landscape has values $\{0, 1\}$ — that is to say, non-satisfying truth assignments are removed from the population, as in Waddington’s experiment? This is a realistic approximation if, for example, this selection does not happen in every generation but every $O(\log n)$ generations (because breeding without selection is known to take you close to the Wright manifold). In such a setting, our quadratic bound for the in-expectation process of monotone functions no longer requires any dependence on the initial probability of satisfaction $\mu^0(f)$. For the process with sampling, we have the following conjecture.

Conjecture: If the fitness landscape has values $\{0, 1\}$, then the process reaches near universal satisfaction with probability approaching 1 as the population size goes to infinity.

We now want to point out an obvious and yet surprising aspect of our work: In the traditional framework of adaptive evolution, each allele spreads in the population mainly either due to an additive contribution to fitness that it makes in and of itself (let us call this “traditional propagation”) or due to random genetic drift [1], [19], [20], [18]. In our model, however, alleles at different genes are spreading in the population as governed by the complex interactions between them that are continually subject to selection. Thus, a population can change dramatically through a novel process involving subtle changes in genetic statistics and simultaneous gradual emergence in the whole population [8], [11], and not by traditional propagation.

Furthermore, notice that since recombination is a crucial ingredient of our analysis, our results inform the question of the role of sex in evolution. In this regard they add to recent works that have begun to examine the role of sex while giving full weight to the importance of genetic interactions [3], [9].

Finally, can our bounds be improved? For the monotone case, it is easy to see that the TRIBES function with appropriate fan-in provides a matching lower bound. As for the general case, we feel that the very generous bounds of the main result can be improved substantially. For example, the assumed time bound is only necessary in order to finish the last part of the argument (convergence to a vertex) once the vast majority of the population is already satisfying; more

analysis is needed to investigate this subtle phenomenon.

Our proof that the population converges to a single satisfying truth assignment may seem a troubling aspect of our result. Two remarks: First, the loss of genetic diversity should not be surprising in itself. With drift alone, for each locus, one allele will become fixed eventually (where the probability that a particular allele will be the fixed allele is proportional to its current frequency in the population). Second, in our process many satisfying truth assignments are likely to survive for a very long time before the random walk clears the picture. This fact may be more relevant than the characteristics of equilibrium; after all, evolution happens in the transient.

ACKNOWLEDGMENTS

We are grateful to Yu Liu of Tsinghua University for some very interesting conversations in the beginning of this research.

REFERENCES

- [1] *The Genetical Theory of Natural Selection*. Oxford: The Clarendon Press, 1930.
- [2] K G Bateman. The genetic assimilation of four venation phenocopies. *J Genet*, 56:443–447, 1959.
- [3] E Chastain, A Livnat, C Papadimitriou, and U Vazirani. Algorithms, games, and evolution. *Proceedings of the National Academy of Sciences*, 111(29):10620–10623, 2014.
- [4] K. Dzhaparidze and J.H. van Zanten. On Bernstein-type inequalities for martingales. *Stochastic Processes and their Applications*, 93(1):109 – 117, 2001.
- [5] D S Falconer. *Introduction to Quantitative Genetics*. Oliver and Boyd, Edinburgh, 1960.
- [6] J H Gillespie. *Population Genetics: A Concise Guide*. JHU Press, 2004.
- [7] R.C. Lewontin. The interaction of selection and linkage. I. General considerations: heterotic models. *Genetics*, 49:49–67, 1964.
- [8] A Livnat. Interactions-based evolution: how natural selection and nonrandom mutation work together. *Biology Direct*, 8:24, 2013.
- [9] A Livnat, C Papadimitriou, J Dushoff, and M W Feldman. A mixability theory for the role of sex in evolution. *Proceedings of the National Academy of Sciences*, 105(50):19803–19808, 2008.
- [10] G. Margulis. Probabilistic characteristics of graphs with large connectivity. *Prob. Peredachi Inform.*, 10:101–108, 1974.
- [11] E Mayr. *Animal Species and Evolution*. Belknap Press, 1963.
- [12] T. Nagylaki. The evolution of multilocus systems under weak selection. *Genetics*, 134(2):627–647, 1993.
- [13] Ryan O’Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.
- [14] L. Russo. On the critical percolation probabilities. *Z. Wahrsch. verw. Gebiete*, 43:39–48, 1978.
- [15] C. Stern. Selection for sub-threshold differences and the origin of the pseudoexogenous adaptations. *The American Naturalist*, 92, 313-316 1958.
- [16] Leslie G Valiant. Evolvability. *Journal of the ACM (JACM)*, 56(1):3, 2009.
- [17] C. H. Waddington. Genetic assimilation of an acquired character. *Evolution*, 7(2):118–126, 1953.
- [18] M J Wade and C J Goodnight. Perspective: the theories of fisher and wright in the context of metapopulations: when nature does many small experiments. *Evolution*, 52(6):1537–1553, 1998.
- [19] S. Wright. Evolution in Mendelian populations. *Genetics*, 16:97–159, 1931.
- [20] S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proc. 6th International Congress of Genetics*, volume 1, pages 356–366, 1932.